Mechanism Design for Multi-dimensional Facility Location Problems: A Computational and Informational Perspective

by

Xin Sui

A thesis submitted in conformity with the requirements
for the degree of Doctor of Philosophy
Graduate Department of Computer Science
University of Toronto

Abstract

Mechanism Design for Multi-dimensional Facility Location Problems: A Computational and Informational Perspective

Xin Sui

Doctor of Philosophy

Graduate Department of Computer Science

University of Toronto

2014

Mechanism design deals with the design of protocols to elicit individual preferences while achieving some social objective (e.g., maximizing social welfare). An important property of mechanisms is *strategy-proofness*, which requires that no agent can gain (or induce a more preferred outcome) by misreporting her preferences to the mechanism. Previous results have shown that when monetary transfer (e.g., in the form of payments) is allowed between agents and the mechanism, the famous *VCG mechanism* is both strategy-proof and efficient (for social welfare maximization).

Despite these positive results, there are still many settings where monetary transfer is difficult to implement, or even prohibited. For instance, political policies are generally determined without payment; monetary compensation between parties that are involved in organ donation is illegal in most countries, etc. It is natural to ask that whether it is possible to design strategy-proof mechanisms when monetary transfer is not allowed. This line of research, referred to as "*mechanism design without payment*", has received much attention from people in economics, political science, and more recently, computer science.

In this thesis, we study a classical embodiment of mechanism design without money called the *facility location problem*: suppose the municipal government plans to build several homogeneous facilities according to the reported preferred locations of the residents. Each resident would prefer one of the facilities built near his home/office (or *ideal* location), and the facilities

should be built to minimize the social cost (i.e., sum of costs over all agents) or other social objectives. We study the facility location from three perspectives: mechanism design, single-peaked preferences and preference elicitation. We first propose a family of strategy-proof mechanisms for multi-dimensional, multi-facility location problem, called *quantile mechanisms*, by extending the classical *generalized median mechanisms*. We also show that the quantile mechanism are approximately *group* strategy-proof for constrained/unconstrained facility location problems, and study the computational complexity of finding an optimal group manipulation. Next, we study the common assumption of *single-peakedness* used in classical mechanism design for facility location problems, and show that agent preferences are far from being single-peaked in one-dimension, but *approximately* single-peaked in two-dimensions. Finally, we study preference elicitation in facility location problems (along with the second price auction), and propose a framework for analyzing the tradeoff between efficiency and privacy.

Acknowledgements

This acknowledgement marks the end of an educational history throughout the last 25 years. Towards the end of my Ph.D. study, I have a mixed feeling and too many people to thank. However, I may not be able to mention every name that has played an important role in this process, so please forgive me if I miss yours.

First and foremost I would like to thank my supervisor Craig Boutilier, without whom this thesis would never be possible. Everyone who knows him should know how great he is, and it is really an honor for me to work with him. He is a wonderful researcher, a great friend, and most important, an excellent supervisor. It is not only the way of doing high quality research that I learn from him, but also lots of other things, including how to manage time as a senior professor, how to communicate with other people, and even how to speak English. These contribute not only to the completion of this thesis, but also every aspect in my future career.

I am also grateful to three other professors in my supervisory committee, Allan Borodin, Toniann Pitassi and Fahiem Bacchus. Thank you all for your advises, feedbacks and supports in my research. I would also like to thank my supervisor while I was in the Chinese University of Hong Kong, Ho-fung Leung, who introduced me to the field of game theory and mechanism design. I would also like to thank all my co-authors and colleagues that I have worked with or have contributed to the completion of this thesis, including Tuomas Sandholm, Alex Nienaber, Jérôme Lang, Yiling Chen, David Parkes, Vincent Conitzer, Ariel Procaccia, Darius Braziunas, Kevin Regan, Laurent Charlin, Tyler Lu, Joel Oren, Amirali Salehi-Abari, Joanna Drummond, Andrew Perrault, etc. I also will thank my fellow friends that I meet in Toronto who have made my life here wonderful, including Joel Yuan, Vanassa Feng, Terry Xu, Joyce Dong, Carlisle Ma, Alice Shan, Tiger Xu, Zishu Liu, Rose Sun, Chao Zhuang, etc.

Finally, I would like to give special thanks to my families. Thank my dearest wife, Dan Chen, for loving, encouraging and accompanying me for the past 7 years. Thank our lovely daughter, Sophia Sui, for offering additional motivation of finishing this thesis. Also thank my parents, Yongcai Sui and Yuxia Zhao for providing selfless supports over more than 30 years.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

There are many settings where multiple self-interested agents have to make a joint decision, e.g., choosing an outcome from a set, in which each agent has his own individual preference. For example, a group of people may have to decide which restaurant to choose for lunch, although each individual has his own preference for meals; the citizens of a country have to elect a president from a set of candidates, where each voter may have a ranking over them; the governments of different nations may have to come to an agreement on how to interact with others, where each may hope to act for its own benefit. All these settings have something in common, i.e., a joint decision has to be made to maximize group satisfaction, based on the individual preferences that may conflict with each other. Such *preference aggregation* problems have existed for millennia, and occur everywhere in our daily lives.

The above preference aggregation problem is a challenging task and requires knowledge from multiple disciplines, including economics, political science, philosophy, mathematics, computer science, etc. In this chapter, we will describe several aspects that are involved in the process of preference aggregation, show what their limitations are, and briefly discuss how they can be improved. The objective of this chapter is to give a high-level picture of the problems studied in this thesis and outline the research contributions of this thesis.

1.1 Preference Aggregation: An Economic Solution

In this section, we describe three aspects that are involved in preference aggregation: social choice, game theory and preference elicitation.

1.1.1 Social Choice

Social choice, which is a theoretical framework for "analysis of combining individual opinions, preferences, interests, or welfares to reach a collective decision or social welfare in some sense" [Sen, 1987], plays an important role in preference aggregation. Generally speaking, social choice blends elements of *voting theory* and *welfare economics*. While the former usually assumes *ordinal preferences*, the latter uses *utility functions* to describe the *degree* or *strength* of preferences.

The use of social choice methods for preference aggregation can be traced back to ancient times. For instance, the *majority rule* in which the joint decision is made based on the opinion of the majorities, was used as a means of not allowing a minority to undo the will of the people respecting the agreed-upon procedures in ancient Greek (around 9th-4th B.C.). However, social choice theory did not become a social scientific discipline with sound mathematical foundations until 1950 with the seminal paper of Kenneth J. Arrow [1950], who introduced the axiomatic method to the study of social choice. In particular, the famous *Arrow's Impossibility Theorem* shows that any social choice method that satisfies a list of seemingly basic requirements must be *dictatorial*. Since then, much work in social choice theory has focused on the possibility/impossibility results of preference aggregation methods that satisfy certain desirable properties, including *Pareto-efficiency*, *monotonicity*, *non-dictatorship*, *non-manipulability*, etc. Some landmark results include the *Muller-Satterthwaite Impossibility Theorem* [Muller and Satterthwaite, 1977] and *Gibbard-Satterthwaite Impossibility Theorem* [Gibbard, 1973, Satterthwaite, 1975].

While these impossibility theorems serve as negative results, they can be avoided by re-

laxing the *unrestricted preference* assumption. In fact, it is natural to make certain restrictions on the possible preferences of agents in many applications. For instance, a simple but elegant domain restriction is *single-peakedness*. Roughly speaking, each individual has a single, most-preferred point in the outcome space and outcomes become less preferred as one moves away from that point. A typical example where such a preference restriction holds is the *facility location problem*, in which some agency intends to build several homogeneous public facilities (e.g., warehouses, libraries, etc.) and each user reports the ideal location at which she would like the facility to be built. Many other social choice problems fit within this class. Voting is one example: political candidates can be ordered along several dimensions (e.g., stance on environment, health care, fiscal policy). Voters have preferences over points in this space, and one must elect several candidates to a legislative body. Other embodiments include product design, customer segmentation, etc. It has been shown that when individual preferences are single-peaked, the *median mechanism and its generalization* [Black, 1948, Moulin, 1980, Barberà et al., 1993] admits *strategy-proof* mechanisms, thus avoiding the Gibbard-Satterthwaite theorem.

1.1.2 Game-Theoretic Paradigm

Social choice theory deals with the aggregation of preferences, assuming such information is given truthfully by each individual. However, when intelligent agents interact with each other, it is a different story. This is because in such settings, the consequences of one agent's action does not only depend on his own preference, but also the actions of other agents. In other words, agents are *strategic*. We will refer to the setting in which intelligent agent interact with each other as a *game*.

Fortunately, *game theory* enables us to analyse a game in a convenient way. Game theory is "the study of mathematical models of conflict and cooperation between intelligent rational decision-makers" [Osborne and Rubinstein, 1994]. In a game where intelligent agents interact with each other, the concept of *equilibrium* is used to look for a stable state from which no

individual will unilaterally deviate. This is because any agent that deviates from the equilibrium while others do not will not be better off. So it is possible to predict the action taken by each agent and the corresponding outcome induced by these joint actions.

Game theory provides a mathematical tool to analyse how intelligent agents interact. However, we still need some interaction protocol to implement the social choice function in equilibrium, and such an interaction protocol is usually called a *mechanism*. A mechanism is a set of rules defining what will happen given the selected actions of individuals. For example, in an auction, an auction mechanism decides who will get the item, and at what price, depending on the bids received; in an election, a voting mechanism determines who will be elected, based on the votes of the voters; and in a matching problem, a matching mechanism assigns medical residents to hospitals, given the declared preferences of both sides.

Mechanism design, a sub-field of game theory and microeconomics, deals with design of protocols to elicit the preferences of self-interested agents so as to achieve a certain social objective [Mas-Colell et al., 1995]. An important property in mechanism design is *strategy-proofness*, which requires that agents have no incentive to misreport their preferences to the mechanism. Previous work has focused on settings where agents can transfer utilities amongst themselves in the form of *payments*. A typical example is the *second price auction* [Vickrey, 1961, Clarke, 1971, Groves, 1973], in which an item is sold to the highest bidder at the price of the second highest bidder. It can be shown that this mechanism maximizes *social welfare* and satisfies strategy-proofness, when agents have quasi-linear utilities.

However, there are also many settings where money cannot be used as a medium of compensation, due to ethical and/or institutional considerations [Schummer and Vohra, 2007]. For example, political decisions should be made without monetary transfers; monetary compensation between parties that are involved in organ donation is illegal in many countries, etc. A natural question is whether it is possible to design strategy-proof mechanisms when monetary transfer is not allowed. This line of work, referred to as "*mechanism design without money*", has received much attention in economics, political science, and recently computer

science in the past several decades [Black, 1948, Moulin, 1980, Procaccia and Tennenholtz, 2009, Schummer and Vohra, 2007].

1.1.3 Preference Elicitation

Preference elicitation is a process of assessing the individual preferences of agents, based on which a joint group decision is made. In some settings the elicitation of individual preferences can be seemingly trivial, e.g., in a single-item auction, the decision maker only has to know the valuation of each bidder for the item. However, there are many other settings where the amount of elicitation necessary to make a joint decision can be impractically large. Consider the example of combinatorial auctions in which an agent is allowed to bid on *bundles* of items. Preference elicitation in such a case may ask each agent to report her valuations for all possible bundles (which is exponential in the number of items being sold), and may require complicated techniques.

Methods for preference elicitation can be classified into three different categories. The first is *complete preference elicitation*, which tries to elicit the full preferences of the agents, and learn them in an efficient way [Zinkevich et al., 2003]. Complete preference elicitation is usually unnecessary and useful in settings where decisions have to be made repeatedly. *Adaptive preference elicitation* focuses on gathering enough information about agents' preferences to make the group decision [Chajewska et al., 2000, Boutilier, 2002], and has been studied under different settings [Parkes and Ungar, 2000, Conen and Sandholm, 2002], although worst case results show that nearly complete information is needed in many settings [Nisan and Segal, 2006]. The third approach is *decision-theoretic preference elicitation*, which allows the trade-offs between the quality of the decision made and the cost of elicitation [Blumrosen and Nisan, 2002, Hyafil and Boutilier, 2006a, 2007]. For example, at a certain stage of the elicitation process, if the cost of extra refinement of preferences exceeds the expected improvement in the quality of the decision, it may be better to stop elicitation, even though an optimal decision may not be reached.

1.2 Algorithmic Approaches and Our Contributions

With the rapid development of computer technology, computers are changing human society dramatically. While computers help people solve problems more efficiently, they also bring new challenges as will be introduced. In this section, we show how computer science is involved in different aspects of preference aggregation and how it can be used to tackle these challenges in this thesis.

1.2.1 Approximate Mechanism Design and Analysis

Classic work on mechanism design aims at implementing some social choice function in equilibrium exactly. However, when it is difficult or impossible to do so, a reasonable solution is to implement the social choice function approximately, under an appropriate approximation measure. Here, the resort to *approximations* is driven by two considerations. The first is the computational difficulties that arise when optimizing the social objective, which is usually ignored in classical mechanism design. For example, traditional approaches to mechanism design with money rely on finding the optimal allocation that maximizes social welfare. While this is computationally trivial in a single-item auction, the problem becomes NP-complete in combinatorial auctions [Rothkopf et al., 1998] where agents are allowed to bid on bundles of items. Moreover, if one uses a sub-optimal allocation instead, the mechanism is in general no longer strategy-proof when Groves payments are used. So an interesting question to ask is whether it is possible to design computationally feasible allocation and payment schemes, such that strategy-proofness can be guaranteed. There has been considerable amount of work on "*algorithmic mechanism design*" in the past decade [Nisan and Ronen, 1999, 2000, Lehman et al., 2002, Archer and Tardos, 2001, Dobzinski et al., 2006, Parkes, 2008], much of it focusing on combinatorial domains, such as combinatorial auctions.

A second complication arises due to the incompatibility of *social efficiency* and strategy-proofness. Consider the k-facility location problem for $k \geq 2$: while the optimal solution is

to cluster the agents into k clusters and choose the "median" within each cluster, such a social choice function cannot be implemented in equilibrium by any mechanism. In this case, we have to give up efficiency to maintain strategy-proofness. Work along these lines includes [Procaccia and Tennenholtz, 2009, Lu et al., 2009, 2010, Fotakis and Tzamos, 2010, Escoffier et al., 2011], which tries to provide worst-case guarantees on the performance of "approximately efficient" mechanisms, and analyse the degree to which efficiency has to be sacrificed to maintain strategy-proofness.

In this thesis, we will focus on the second form of approximation. Our first contribution in this area is to propose a class of *quantile mechanisms*, a type of *generalized median mechanism* [Moulin, 1980, Barberà et al., 1993, Barberà, 2010] for the multi-dimensional, multi-facility location problem (Chapter 3). We derive several worst-case approximation ratios for *social cost* and *maximum load* for the L_1 and L_2 cost models. While the performance guarantees of such mechanisms under worst-case assumptions are quite discouraging, we also develop a *sample-based empirical framework* for optimizing quantile mechanisms relative to a known preference distribution. We use profiles sampled from this distribution to optimize quantiles while maintaining strategy-proofness of our mechanisms. Our empirical results demonstrate that, by exploiting probabilistic domain knowledge, we obtain strategy-proof mechanisms that outperform mechanisms designed to guard against worst-case profiles, and give solutions extremely close to the optimum attainable with exact knowledge of agent preferences.

While quantile mechanisms are individual strategy-proof, they fail to guarantee *group strategy-proofness* in multi-dimensional spaces. Intuitively, a mechanism is group strategy-proof if any group of agents form a coalition and make a joint misreport, then there must be some agent who is not strictly better-off. In fact, the characterization results of Barberà et al. [1993] suggest that there is no (*anonymous*) *non-dictatorial*, *group strategy-proof* mechanisms in such a setting. However, the quantile mechanism may work "reasonably well" in practice. Our second contribution in this area is to bound the maximum incentive for a group of agents to misreport their preferences, assuming some form of cost function (Chapter 4). We

provide several possibility/impossibility results with respect to individual and group strategy-proofness in both unconstrained facility location—in which facilities can be placed at any point in some (metric) space—and constrained facility location—in which agent preferred outcomes may not be feasible (i.e., the feasible outcome space is constrained). We also complement our results with empirical analysis of data from the 2002 Irish General Election and U.S. capital city locations, showing that while the probability of manipulation may remain high, the gains and impact on social welfare are extremely small in practice (much less than worst-case theoretical bounds).

1.2.2 Computational Social Choice

While traditional research on social choice focuses on the axiomatic properties for different rules, their computational issues are rarely considered. For instance, while the Gibbard-Satterthwaite Impossibility Theorem shows that it is impossible to devise a non-manipulable rule satisfying certain desirable properties, computer science may provide tools for making such manipulative behaviors computationally difficult (e.g., *NP-hardness*) to implement. This idea was first explored in the work of Bartholdi et al. [1989b], which, together with that of Bartholdi et al. [1989a, 1991], is broadly considered to be the starting point of *Computational Social Choice*. After nearly 15 years, this groundbreaking result was followed by a number of results on the computational complexity of manipulation in various settings [Conitzer and Sandholm, 2003, Conitzer et al., 2007, Faliszewski et al., 2009a, Faliszewski and Procaccia, 2010].

Much like previous work, our contribution in this area is the analysis of the computational complexity of the group manipulation problem in quantile mechanisms (Chapter 5). Specifically, focusing on the unconstrained facility location problem, we show that for single facility location, the optimal group manipulation problem—in which the objective is to minimize the cost over all manipulators—can be formulated as a *linear program (LP)* or *second-order cone program (SOCP)*, under the L_1- and L_2-norms, respectively, and hence can be solved in polyno-

mial time; and for multi-facility location, the optimal group manipulation problem is NP-Hard (reduction from p-median), but can be formulated as a *mixed integer linear program (MILP)* or *mixed integer second-order cone program (MISOCP)*, under the L_1- and L_2-norms, respectively. We also show that our formulations work extremely well in practice and are scalable to reasonably large problem size, despite the hardness result we provide.

1.2.3 Computational Aspects of Single-peakedness

Single-peakedness allows protocols such as the *median mechanism* to reduce the communication burden on participants in the mechanism, ensure truthful reporting, and can often ease computational demands. While conceptually attractive, single-peakedness is a very strong assumption and unlikely to hold in realistic settings, e.g., elections with thousands of voters and more than a handful of candidates. Recent research has begun to investigate computational methods to test single-peakednes [Escoffier et al., 2008], and to study various forms of approximation (e.g., deleting voters, deleting candidates, clustering candidates, or adding additional axes, etc.) [Escoffier et al., 2008, Faliszewski et al., 2011, Erdélyi et al., 2012, Galand et al., 2012]. However, the extent to which these proposals for approximate single-peakedness can further help explain actual voter preferences is unclear.

Our contributions in this area are as follows. First, we test single-peaked consistency, and several forms of approximation used in the literature (in isolation and in combination) on two election data sets to see if these approximations have any empirical explanatory power. We developed a *branch-and-bound* algorithm to find the best single-dimensional axis given a preference profile, i.e., the ordering of candidates for which the greatest number of voters are single-peaked. The algorithm is easily extended to support various forms of approximation, including *voter deletion*, *local candidate deletion* and *adding new axes*. While the best-axis problem is computationally difficult, our method works well in practice. We show that voter preferences in these elections cannot be explained by any form of approximation that are recently proposed in the literature, and are far from being single-peaked. Next, we extend our

algorithm to find the best *multi-dimensional ordering* to explain a preference profile. We show that voter preferences in our data sets are approximately single-peaked in a two-dimensional space, which suggest that a focus on multi-dimensional rather than single-dimensional models can greatly enhance the applicability of single-peaked models in practice.

The rich literature on *spatial models* also bears a strong relationship to single-peaked preferences [Hotelling, 1929, Hinich, 1978, Poole and Rosenthal, 1985]. Spatial models explain voter choice by inferring the distances between voters and candidates, and typically using some form of probabilistic choice model based on these distance [Bradley and Terry, 1952, Luce, 1959, Shepard, 1959]. While the model is more restrictive than single-peakedness in some sense, stochastic choice allows for accommodation of "misorderings", much like approximations in single-peaked models. We develop an alternating optimization algorithm for fitting both voters and candidates into a latent space (combined with Plackett-Luce as the choice model), when the voter preferences are given. Our findings show similar results to those derived from the single-peaked model, namely that the two-dimensional fit is much better than then one-dimensional fit. They also suggest that party policies plays an important role in the electorates view of candidates.

1.2.4 Cost of Elicitation: Computation, Communication and Privacy

Much work in mechanism design assumes *direct-revelation*, in which agents reveal their full preferences to the mechanism. While direct-revelation seems to be a natural way to elicit preferences, it often elicits more information than needed to make the optimal decision, leading to both communication and computational difficulties [Conitzer and Sandholm, 2004]. For example, in a general *combinatorial auction*, an agent can report her valuations for all possible *bundles* of items (which is exponential in the number of items being sold), most of which she will not receive/win. The line of work on *indirect-revelation* mechanisms, commonly studied in the settings of auctions, showed that each of these difficulties can be alleviated in some settings [Parkes, 1999, Zinkevich et al., 2003], though not in the worst-case [Nisan and Segal,

2006]. Direct revelation also requires a sacrifice of privacy: revealing her full preferences may be undesirable for an agent, especially when some of that information is provably unnecessary for computing the optimal outcome.

Previous work has considered the trade-off between communication and efficiency [Blumrosen and Nisan, 2002, Hyafil and Boutilier, 2007], and the trade-off between privacy and communication [Sandholm and Brandt, 2008, Feigenbaum et al., 2010]. In our thesis, we consider a third trade-off, that between social efficiency and privacy, and provide a general framework for analyzing this trade-off (Chapter 7). Specifically, we consider approximately efficient mechanisms that find ε-optimal solutions to the choice problems, and show how agents' privacy improves as one increases the degree of approximation ε. By extending the *privacy approximation ratios* introduced by Feigenbaum et al. [2010], we analyze the efficiency-privacy trade-off in both *second-price auctions* and *facility location problems* (introducing new incremental mechanisms for facility location along the way). We show that, both theoretically and empirically, small sacrifices in efficiency can provide significant gains in privacy, in both the average and worst case.

1.3 Outline of This Thesis

Generally speaking, this thesis focuses on the facility location problem, which is a typical embodiment of mechanism design without money. We emphasize that facility location problems embody a much richer class of mechanism design problems than suggested by the name. Other problems that fit into this class of problems include voting, product configuration, political decision making, etc. Please refer to Section 2.3 for a detailed discussion of these settings.

The thesis can be divided into three topics: *mechanism design for facility location* (Chapter 3-5), *single-peakedness and approximation* (Chapter 6) and *preference elicitation* (Chapter 7). The overall structure of this thesis is illustrated in Figure 1.1.

We start by reviewing some necessary background in social choice, game theory and mech-

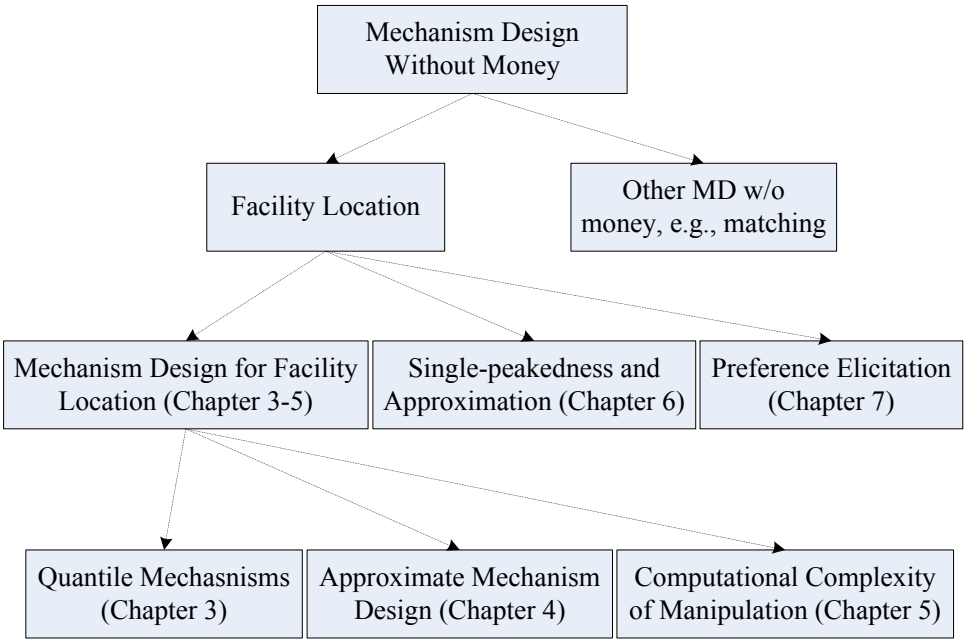


Figure 1.1: The structure of this thesis.

anism design, single-peaked preferences, and introduce some definitions and notation used throughout this thesis in Chapter 2. Then in the first part, we study the facility location problem from the perspective of mechanism design. Specifically,

- In Chapter 3, we propose the *quantile mechanisms*, a family of *strategy-proof* mechanisms for multi-dimensional, multi-facility location problems, and derive worst-case approximation ratios for social cost and maximum load, respectively;

- In Chapter 4, we consider the constrained and unconstrained facility location problem, providing several possibility/impossibility results with respect to individual and group strategy-proofness. We also bound the incentive for manipulation in quantile mechanisms where group strategy-proofness is not possible in general; and

- In Chapter 5, we analyze the computational complexity of the group manipulation problem in quantile mechanisms, showing that it can be solved in polynomial time in certain cases, but that is NP-hard in others. We also provide compact linear and quadratic programming formulations for the optimal group manipulation and analyze their solution

empirically.

In the second part (Chapter 6), we focus on the commonly used assumption of single-peakedness in the literature of facility location and social choice. Moreover,

- We first propose a branch-and-bound algorithm to find the best single-dimensional axis, and extend the algorithm to support several forms of recently proposed approximations. We show that agent preferences are far from single-peakedness in 1D, but are approximately single-peaked in 2D;
- We also study the spatial model and propose an alternating optimization algorithm to estimate both agent and candidate positions in a same multi-dimensional space.

In the last part (Chapter 7), we address the preference elicitation problem in the facility location (also the second price auction), and propose two incremental elicitation protocols (English and bisection). Specifically,

- We propose a framework for analyzing the interesting trade-off between efficiency and privacy, and provide several upper and lower bounds on the privacy approximation ratios for different protocols.

While each of these chapters is independent, they all fit together and complement each other. For instance, strategy-proofness of quantile mechanisms can be achieved if agent preferences are single-peaked (see Chapter 3 for details). However, as we shall see in Chapter 6, agent preferences are often not single-peaked, but may be approximately so in practice. An interesting question is whether we can design approximately strategy-proof mechanism when agent preferences are approximately single-peaked? In the final Chapter 8, we conclude this thesis, talk about the connections between different chapters, and highlight some possible future directions.

Chapter 2

Background

In this chapter, we present some necessary background to understand the the results in the remaining chapters, and introduce some notation we use throughout in this thesis. We start with social choice theory in Section 2.1. Then we introduce mechanism design in Section 2.2. In Section 2.3, we move on to the facility location problem—a classical embodiment of mechanism design without money—and single-peaked preferences—an important preference domain that admits *strategy-proof* mechanisms for facility location and other applications. We also present some work on computational aspects of single-peakedness, and recent proposals for approximate single-peakedness in Section 2.4. We also introduce the spatial theory of voting in that section.

2.1 Social Choice Theory

Social choice deals with the aggregation of individual preferences. In this section, we start by introducing preference relations and utility functions. Then we define social choice functions, introducing some related properties and two impossibility theorems. Finally, we overview some recent work on computational social choice.

2.1.1 Preferences and Utility

We start by defining preference relation and utility function, and defer our definition of social choice problem to the next section.

Suppose that a set $N = \{1, 2, \ldots, n\}$ of agents wants to make a joint decision (or select an outcome) $o \in O$, where O is set of all possible outcomes. Each agent $i \in N$ has a *preference* over the outcome set O:

Definition 2.1 (Preference relation) *A preference relation is a binary relation $\succeq$ over the set of possible outcomes O.*

For two outcomes $o_1, o_2 \in O$, $o_1 \succeq o_2$ is interpreted as "outcome o_1 is at least as preferred as outcome o_2". For example, when the set O consists of different cuisines, *Italian* $\succeq$ *French* means that Italian food is at least as preferred as French food. We use $\succ$ to denote the *strict* preference relation in which $o_1 \succ o_2$ means that $o_1 \succeq o_2$ but $o_2 \not\succeq o_1$, and use $\sim$ to denote *indifference* in which $o_1 \sim o_2$ means $o_1 \succeq o_2$ and $o_2 \succeq o_1$.

It is reasonable to assume that each agent is *rational*:

Definition 2.2 (Rational) *A preference relation $\succeq$ is rational if it is complete, reflexive and transitive, where $\succeq$ on a set O is complete if $o_1 \succeq o_2$ or $o_2 \succeq o_1$ for every two outcomes $o_1, o_2 \in O$, reflexive if $o \succeq o$ for every outcome $o \in O$, and transitive if $o_1 \succeq o_3$ whenever $o_1 \succeq o_2$ and $o_2 \succeq o_3$ for any $o_1, o_2, o_3 \in O$.*

In other words, a rational preference relation is a weak order over the set of possible outcomes. Such a preference is referred to as "ordinal", which is in contrast to "cardinal" in which the preference of the agent is specified by a utility function $u_i : O \to \mathbb{R}$. In other words, the cardinal properties are the numerical values associated with the outcomes, and hence the magnitude of any differences in the utility measure between outcomes. Formally, we define a utility function that is *consistent* with an ordinal preference $\succeq$ as follows:

Definition 2.3 (Consistent Utility Function) *A utility function $u : O \to \mathbb{R}$ is consistent with a preference relation $\succeq$, when $u(o_1) \geq u(o_2)$ if and only if $o_1 \succeq o_2$ for any two outcomes $o_1, o_2 \in O$.*

Note that for any given preference relation $\succeq$, the utility function that can represent $\succeq$ is not unique. However, under very weak condition, there always exists some utility function that represents $\succeq$. In the remainder of this thesis, we will use preference relation $\succeq$ or the utility function $u(\cdot)$, depending on the problem being addressed.

2.1.2 Social Choice Functions and Impossibility Theorems

As mentioned above, social choice theory deals with "the combination of individual preferences to reach a collective decision", which is usually accomplished through a *social choice function*. Formally, let $\succeq_i$ be the preference relation of agent i, and $\succeq = \{\succeq_1, \ldots, \succeq_n\}$ be the *preference profile* of n agents. Also let $\mathcal{R}_i$ be set of all possible preference relations available to agent i, and $\mathcal{R} = \prod_i \mathcal{R}_i$, then we have:

Definition 2.4 (Deterministic social choice function) *A deterministic social choice function f is a mapping $f : \mathcal{R} \to O$.*

In other words, a deterministic social choice function selects an outcome from the set of possible outcomes given a preference profile. A social choice function can also be randomized $f : \mathcal{R} \to \Delta(O)$, which defines a probability distribution over the outcome set O. Note that give a preference profile, there are many social choice functions that can be used to determine an outcome. We give two examples of commonly used social choice functions.

Example 2.1 (Plurality rule) *The* plurality rule *is a voting rule in which the winner is the candidate with the most first-place votes (assuming some tie breaking rule). Consider the following voting problem. Four voters have to choose the president from three candidates*

{Clinton, Obama, McCain}, and their preferences are shown as follows:

$$\textit{Voter 1} : \textit{Clinton} \succ_1 \textit{Obama} \succ_1 \textit{McCain}$$

$$\textit{Voter 2} : \textit{Obama} \succ_2 \textit{McCain} \succ_2 \textit{Clinton}$$

$$\textit{Voter 3} : \textit{McCain} \succ_3 \textit{Clinton} \succ_3 \textit{Obama}$$

$$\textit{Voter 4} : \textit{Obama} \succ_4 \textit{Clinton} \succ_4 \textit{McCain}$$

If the plurality rule is used to select the president, i.e., the candidate who is ranked at top the most time wins, and ties are broken alphabetically, then Obama will be the winner.

Example 2.2 (Borda rule) *Consider the above example, but where the* Borda rule *is used. In the Borda rule, the candidate that is ranked at the s^{th} position in any vote receives a Borda score of $m - s$ from that vote, (where m is the total number of candidates), and the candidate with the highest total score over all votes wins. With the above preferences, the Borda score for Clinton, Obama and McCain are 4, 5 and 3, respectively. The Borda rule will generate a social preference relation of Obama $\succ$ Clinton $\succ$ McCain, and $Obama$ will be elected as the president.*

Traditional research on social choice deals with the axiomatic properties of the social choice function, and provides a number of important impossibility results. An impossibility result is one that shows certain properties cannot be simultaneously satisfied by any single social choice function. In the remaining of this section, we will describe two important results: *Muller-Satterthwaite* and *Gibbard-Satterthwaite* Impossibility Theorem. We first define some related properties.

The first is *Pareto efficiency*, which requires that a social choice function cannot select a *dominated* outcome. An outcome o is dominated by another outcome o' if every agent prefers o' to o. Formally:

Definition 2.5 (Pareto efficient) *A social choice function f is Pareto efficient, if for any pref-*

erence profile $\succeq$ with $o = f(\succeq)$, $\nexists o'$ such that $o' \succeq_i o$ for all i and $o' \succ_{i^*} o$ for some i^*.

Monotonicity says that an outcome must remain the winner if the support for it is increased relative to a preference profile under which it was already winning:

Definition 2.6 (Monotonicity) *A social choice function f is monotonic if, for any preference profile $\succeq$ with winning alternative $o = f(\succeq)$, and any profile $\succeq'$ that is identical to $\succeq$ except that o has a higher rank in the vote $\succeq'_i$ of some voter i than it does in $\succeq_i$, we have $o = f(\succeq')$.*

We also have non-dictatorship, meaning the outcome is not dictated by a single agent:

Definition 2.7 (Non-dictatorship) *Let $\tau(\succeq)$ be the most preferred outcome in $\succeq$. A social choice function f is non-dictatorial if, $\nexists i$ such that $\tau(\succeq_i) \in f(\succeq)$ for all $\succeq$.*

The Muller-Satterthwaite Theorem says that when agents preferences are *unrestricted*, if a social choice function is Pareto efficient and monotonic, then it must be dictatorial. Agents preference are said to be unrestricted if any rational preference can be held by an agent. This result can be formally described as follows:

Theorem 2.1 (Muller-Satterthwaite Impossibility Theorem) *When $|O| \geq 3$ and agent preferences are* unrestricted*, there is no social choice function that is Pareto efficient, monotonic and non-dictatorial.*

Note that the pre-criteria $|O| \geq 3$ is necessary, otherwise there are social choice functions that satisfy all these three properties (e.g., the majority rule, which selects alternatives among two candidates which have a majority). The unrestricted preferences assumption is also critical, otherwise the impossibility result can be avoided. We will discuss this point in Section 2.3.

Another important impossibility result in social choice theory is the *Gibbard-Satterthwaite* Theorem, which says that when there are three or more outcomes and agent preferences are unrestricted, then a *unanimous* social choice function is *strategy-proof* if and only if it is dictatorial. A social choice function is said to be unanimous if all agents prefer the same outcome, then that outcome should be selected as the winner. Strategy-proofness requires that if each

agent prefers the outcome selected given its truthful report, no matter what the reports of other agents' are. Formally, we have:

Definition 2.8 (Unanimous) *A social choice function f is unanimous if, whenever $o \succeq_i o'$ for all $o' \in O$ and all i, then we have $o \in f(\succeq)$.*

Definition 2.9 (Strategy-proof) *Let $\succeq_{-i}$ be the preference profile of all agents but i. A social choice function f is strategy-proof if:*

$$f(\succeq_i, \succeq_{-i}) \succeq_i f(\succeq_i', \succeq_{-i}), \qquad \forall i, \forall \succeq_i', \forall \succeq_{-i}$$

Theorem 2.2 (Gibbard-Satterthwaite Theorem) *Let $|O| \geq 3$ and agent preferences are unrestricted, then a unanimous social choice function f is strategy-proof if and only if it is dictatorial.*

Similarly, the conditions that $|O| \geq 3$ and unrestricted preferences cannot be relaxed, otherwise non-dictatorial and strategy-proof mechanisms exist (e.g., *median mechanism and its generalization* when agents have *single-peaked* preferences). We will discuss this in Section 2.3 in detail.

2.1.3 Computational Aspects of Manipulation

Traditional research on social choice focuses on the axiomatic properties of different social choice functions, however, the computational issues associated with these methods have only been addressed recently. Initiated by Bartholdi et al. [1989b, 1989a, 1991], research began to address problems in computational aspects of preference aggregation, and a burgeoning area, *Computational Social Choice*, has been developed over the past decade. Focusing on the computational aspects of strategy-proofness, we present some recent related work in this section. Note that there are many other research topics in the area of computational social choice, including computationally hard aggregation rules [Bartholdi III et al., 1989a, Conitzer,

2006], communication requirements in social choice [Conitzer and Sandholm, 2002a, 2005, Conitzer, 2009], distributed resource allocation [Lipton et al., 2004, Chen et al., 2013], etc. However, as most of them are not the focus of this thesis,[1] we omit the discussion but refer readers to a survey [Brandt et al., 2015].

There has been extensive study of the manipulation problem in other social choice, especially in the context of voting. While the Gibbard-Satterthwaite impossibility theorem shows that social choice functions that are immune to manipulation do not exist in general, Bartholdi et al. [1989b] demonstrated that manipulation of certain voting rules can be computationally difficult. Specifically, they show manipulating the *single-transferable vote (STV)* is NP-hard. This spawned an important line of research into the complexity of various voting rules, which collectively can be viewed as proposing the use of computational complexity as a barrier to practical manipulation. Many existing voting rules have been proved NP-hard to manipulate, including Copeland, ranked pairs, maximin, etc (see, for example, [Conitzer and Sandholm, 2003, Conitzer et al., 2007, Faliszewski et al., 2009a], and [Faliszewski and Procaccia, 2010] for an excellent survey).[2] Most of this work focuses on a discrete and atomic outcome space, and the objective is to select a single winner.

Exploiting computational complexity to prevent (or reduce the odds of) manipulation is somewhat problematic in that it focuses on worst-case scenarios, and usually assumes full knowledge of agent preferences. However, there may still be an efficient algorithm that can solve most "practical" instances of the manipulation problem, and if so, the computational hardness only provides limited protection against manipulation. It would be much better if one can show that manipulation is "usually" hard. Recent work has shown that when preferences

[1] An exception is the communication requirement, where people try to analyze the number of bits that have to be transmitted to compute the outcome of a social choice function. For example, Sandholm and Brandt [2008] showed that perfect privacy can be achieved using English protocol in the the second-price auction (see Example 2.3) at the expense of exponential communication. Feigenbaum et al. [2010] proposed a general framework to analyze the *trade-off between privacy and communication*, defining several forms of *privacy approximation ratio*. We address the this problem by showing that when approximation is allowed on the social choice function, communication requirement can be further decreased. Moreover, there is a complicated four-way tradeoff between efficiency, privacy, communication and incentives. We will talk about this more in Chapter 7.

[2] As NP-hardness of manipulation is not the focus in this thesis, we will omit the description of these rules.

are single-peaked, the *constructive manipulation problem*—in which a set of manipulators try to find a set of preference rankings (reports) that would make a specific candidate win—is polynomial time solvable for many voting rules [Faliszewski et al., 2009b]. Recent work has also studied average case manipulability (i.e., the probability that a preference profile is "easily" manipulable, assuming some distribution over preferences or preference profiles), and shows that manipulation is often feasible both theoretically and empirically [Friedgut et al., 2008, Isaksson et al., 2012, Conitzer and Sandholm, 2006, Procaccia and Rosenschein, 2007, Walsh, 2009, Xia and Conitzer, 2008]. The complete information assumption has also been challenged, and manipulation given probabilistic knowledge of other agent's preferences has been studied in equilibrium [Majumdar and Sen, 2004, Ángel Ballester and Rey-Biel, 2009] and from an optimization perspective [Lu et al., 2012].

2.2 Game Theory and Mechanism Design

A commonly used assumption in social choice theory is that agents report their preferences truthfully, which does not necessarily hold in practice. Consider the Example 2.1, in which *Obama* is selected as the winner by the plurality rule if all agents report their preferences truthfully. If voter 3 changes his vote to *Clinton* $\succ_3$ *McCain* $\succ_3$ *Obama*, then both *Obama* and *Clinton* will be ranked first twice and *Clinton* will be selected the winner (assuming ties are broken alphabetically). So voter 3 has an *incentive* to misreport her preference and can induce a (personally) preferred outcome by such a misreport.

The above argument does not only hold for voter 3, but also every voter in the election. Such *strategic* behavior makes it hard to predict the outcome. Fortunately, *game theory* provides a mathematical framework for analyzing *games* in which a set of self-interested agents interact with each other. In this section, we provide a brief introduction to game theory and various solution concepts, and then introduce the mechanism design problem, which deals with the design of communication protocols to implement specific social choice functions. We also

introduce a famous family of VCG mechanisms when agents have quasi-linear utility, and discuss other means to circumvent the Gibbard-Satterthwaite Theorem.

2.2.1 Games and Solution Concepts

A *strategic game* is a model used to study how self-interested agents interact with each other. Let N be a finite set of agents, and O be a set of possible outcomes. For each agent $i \in N$, there is a set of possible actions A_i that can be taken by agent i, and the joint action $\mathbf{a} = \prod_{i \in N} a_i \in A = \prod_i A_i$ determines an outcome (where $a_i \in A_i$). An agent's utility function maps joint actions to real numbers, i.e., $u_i : A \to \mathbb{R}$. A game of *complete information* specifies the utilities that each agent receives for all possible joint actions. In this setting, the utilities of all agents are assumed to be common knowledge, and such a game is usually referred to as a *normal-form game*.

However, we are more interested in settings where an agent may be uncertain about the preferences of others. Such games are usually called games with *incomplete information* and modelled as *Bayesian games*. A Bayesian game is similar to a normal-form game except that instead of knowing the utilities of the others, each agent processes a (common) *prior distribution* $Pr(T)$ from which the *types* of agents are drawn. Formally, we define a Bayesian game as follows:

Definition 2.10 (Bayesian Game) *A Bayesian game consists of:*

- *a finite set of agents N, and a set of outcomes O*
- *a set of types T_i for each agent i and a set of joint types $T = \prod_i T_i$*
- *a set of actions A_i available to each agent $i \in N$ and a set of joint actions $A = \prod_i A_i$*
- *an outcome function $h : A \to O$*
- *a joint prior distribution $Pr(T)$ over agent types*

- *a utility function* $u_i : O \times T_i \to \mathbb{R}$ *for each agent* $i \in N$

The outcome function can also be randomized $h : A \to \Delta(O)$, where Δ is a distribution over O. The marginal prior distributions $Pr(T_i)$ can be different for each agent, as long as they are commonly known to all of them.

In a Bayesian game, it is possible to analyze the action taken by each agent and predict the outcome for the joint action, using the notions of *solution concept* in game theory. There are three commonly used solution concepts for Bayesian games: *dominant strategy equilibrium*, *ex-post equilibrium*, and *Bayesian-Nash equilibrium*. All these notions capture a *steady state* of a Bayesian game from which no agent will unilaterally deviate if the others remain unchanged. To define these solution concepts, we first define the *strategy* of an agent.

Definition 2.11 (Strategy) *A strategy of agent* i *is a contingent plan that defines the action for each possible type, i.e.,* $s_i : T_i \to A_i$.

Let $s_i \in \Sigma_i$ be the strategy of agent i, where Σ_i is the set of all possible strategies for agent i. Each strategy defined above maps a single action to each possible type, and is often referred to as a *pure* strategy. Alternatively, an agent can use a *mixed* strategy, denoted as $s_i \in Pr(\Sigma_i)$, which defines a probability distribution over pure strategies. We use s_i to denote both pure and mixed strategies, and use u_i to denote the expected utility for playing a pure/mixed strategy. Let $\mathbf{s} = (s_1, \ldots, s_n)$ be a strategy profile, and $\mathbf{s}_{-i} = (s_1, \ldots, s_{i-1}, s_{i+1}, \ldots, s_n)$ be the joint strategy of all agents except i. Also let $\mathbf{t}_{-i}$ be the joint type of all agents except i. We can define *dominant strategy equilibrium* as follows:

Definition 2.12 (Dominant strategy equilibrium) *A strategy* s_i *is dominant for agent* i *if:*

$$u_i(h(s_i(t_i), \mathbf{s}_{-i}(\mathbf{t}_{-i})), t_i) \geq u_i(h(s_i'(t_i), \mathbf{s}_{-i}(\mathbf{t}_{-i})), t_i), \qquad \forall s_i' \in \Sigma_i, \forall s_{-i} \in \Sigma_{-i}, \mathbf{t}_{-i} \in T_{-i}$$

A strategy profile $\mathbf{s} = (s_1, \ldots, s_n)$ *is in dominant strategy equilibrium (DSE) if* s_i *is a dominant strategy for each agent* i.

In other words, a strategy s_i is said to be a dominant strategy for agent i, if it maximize the (expected) utility of agent i, no matter what strategies the other agents use or the realization of their types. A strategy profile $\mathbf{s}$ is in dominant strategy equilibrium if everyone uses a dominant strategy. Dominant strategy equilibrium is a very strong solution concept because each agent in a game can commit to a dominant strategies without considering the strategies of other agents. The following example shows that truthful bidding is a dominant strategy in the *second price auction*:

Example 2.3 (Second price auction) *Consider a second price auction (a.k.a. Vickrey auction) in which a single item is sold to the highest bidder at the second highest bid (assuming ties are broken alphabetically). Let v_i be the valuation of agent i on the item, $b_i(v_i)$ be the bid of agent i, and b' be the highest bid from any other agent, then the utility of agent i is:*

$$u_i(b_i, b', v_i) = \begin{cases} v_i - b' & \text{if } b_i > b'; \\ 0 & \text{otherwise}. \end{cases}$$

Remark 2.1 (DSE in second price auction) *It is well known that the strategy profile in which each agent bids her true valuation $b_i(v_i) = v_i$ is in DSE. In other words, it is optimal for each agent i to bid her true valuation, no matter what the bids of other agents are. As b' is the highest bid from any other agent, then by case analysis, when $b' < v_i$ it is optimal to bid $b_i > b'$, and when $b' \geq v_i$ then it is optimal to bid $b_i < b'$. Bidding $b_i = v_i$ solves both cases, and is a dominant strategy for agent i.*

One should note that a DSE does not always exists in a Bayesian game, and one may use a *weaker* solution concept, e.g., *ex-post equilibrium (EPE)* or *Bayesian-Nash equilibrium (BNE)*. In an ex-post equilibrium, each agent's action is a best response to the actions taken by others that are dictated by their strategy, for any of their types; In a Bayesian-Nash equilibrium, each agent chooses a strategy that maximizes the *expected* utility against the strategies of other agents, given her belief about the others' types. However, we focus on DSEs in this thesis, as

they exist in most of the games we study here. In situations where a DSE does not exist, we look for equilibria that are approximately dominant, i.e., the utility of each agent for committing the corresponding strategy in the equilibrium is at most ε worse than that of committing any other strategy, no matter what strategies the other agents uses or the realization of their types. We will discuss this in detail in later chapters.

2.2.2 Mechanism Design and The Revelation Principle

While game theory provides a mathematical framework for analyzing the actions taken by individual agents, and the corresponding outcomes that result in a strategic game, mechanism design deals with the reverse problem of designing outcome rules (e.g., the mapping from joint strategies to outcomes) such that an "optimal" outcome can be implemented under certain solution concepts (e.g., Bayesian-Nash, ex-post or DSE). We start by defining a mechanism:

Definition 2.13 ((Deterministic) Mechanism) *A (deterministic) mechanism $\mathcal{M} = (\Sigma_1, \ldots, \Sigma_n, g)$ consists of a set of strategies Σ_i for each agent i, and an outcome rule $g : \prod_i \Sigma_i \to O$ that selects an outcome based on the joint strategy of the agents.*

Note that the outcome rule g defined above maps each strategy profile to a single outcome, i.e., the mechanism is *deterministic*. The outcome rule can also be randomized $g : \prod_i \Sigma_i \to \Delta(O)$, where Δ is the set of probability distributions over O, and such a mechanism is called a *randomized mechanism*. Randomized mechanisms can can offer more flexibility as we will see below in section 2.3.3.

A mechanism offers each agent a set of strategies, and specifies how the outcome is chosen for each selected joint strategy. In this sense, a mechanism can be viewed as a Bayesian game, if associated with a utility function for each agent and a prior over types. The objective of the mechanism designer is to choose the rule g to *implement* some social choice function with respect to some solution concept (e.g., DSE, EPE or BNE). A commonly used social choice function is the *social welfare maximizer*, i.e., $f(\mathbf{t}) \in \arg\max_o \sum_i u_i(o, t_i)$, which provides a

utilitarian view on the society of agents. An alternative social choice function is the *minimum welfare maximizer*, i.e., $f(\mathbf{t}) \in \arg\max_o \min_i u_i(o, t_i)$. Formally, we define:

Definition 2.14 (Implementation) *A mechanism $\mathcal{M}$ is said to implement a social choice function f in DSE (respectively, EPE, BNE) if there exists a DSE (respectively, EPE, BNE) $s = (s_1, \ldots, s_n)$, such that $g(s_1(t_1), \ldots, s_n(t_n)) = f(\mathbf{t})$ for any $\mathbf{t}$. In this case, we also say that f is implementable in DSE (respectively, EPE, BNE) by $\mathcal{M}$.*

In other words, a mechanism is said to implement some social choice function in DSE (respectively, EPE, BNE), if: 1) there exists a DSE (respectively, EPE, BNE); and 2) the outcome chosen by the mechanism under the above equilibrium coincides with the one selected by the social choice function.

Recall that the objective of mechanism design is to choose the outcome rule to implement a social choice function in DSE (respectively, EPE, BNE). However, the outcome rule must be designed carefully to deal with every possible strategy profile of the agents, which makes the search space huge for a mechanism. The *revelation principle*, one of the most important theoretical results in mechanism design, indicates that when looking for mechanisms to implement a social choice function, we can focus only on those in which each agent reveal her type *truthfully* without loss of generality (or *incentive-compatible* ones). In other words, if some mechanism implements a social choice function under a given solution concept, then there exists another mechanism that implements the same social choice function under the same solution concept truthfully. Before describing the revelation principle formally, we provide some definitions.

Definition 2.15 (Direct-revelation) *A mechanism $\mathcal{M}$ is* direct-revelation *if $\Sigma_i = T_i$ for all i.*

In other words, in a direct-revelation mechanism each agent reveals a type (not necessarily her true type), and the mechanism chooses an outcome based on the revealed type profile.

An agent's strategy is *truthful* if she always reports her true type, and a mechanism is *incentive compatible* if the strategy profile in which all agents are truthful is a dominant strategy

(respectively, ex-post, Bayesian-Nash) equilibrium. Formally, we have:

Definition 2.16 (Truthful) *In a direct-revelation mechanism, an agent's strategy s_i is said to be truthful if $s_i(t_i) = t_i$.*

Definition 2.17 (Incentive compatibility) *A direct-revelation mechanism $\mathcal{M}$ is dominant strategy (respectively, ex-post, Bayesian-Nash) incentive compatible if the strategy profile in which all agents report their types truthfully is a DSE (respectively, EPE, BNE) in the game induced by the mechanism $\mathcal{M}$.*

Dominant strategy incentive compatibility is usually referred to as *strategy-proofness* (Definition 2.9), and is a very important property in mechanism design. If a mechanism is strategy-proof, then no agent has an incentive to misreport and can reveal her true type to the mechanism without considering the strategies of others. In addition, the induced strategy profile is a DSE. In fact, one can check that the second price auction defined in Example 2.3 is strategy-proof. We will return to this example in section 2.2.3.

Now we are ready to describe the revelation principle. The result was first formulated for dominant strategy equilibria by Gibbard [1973], and later extended to Bayesian-Nash equilibria by Green and Laffont [1977] and Myerson [1979, 1981]. However, we will describe them in a single theorem:

Theorem 2.3 (Revelation principle) *If f is a social choice function that is implementable in DSE (respectively, EPE, BNE) by a mechanism $\mathcal{M}$, then f is also truthfully implementable in DSE (respectively, EPE, BNE) by some direct mechanism $\mathcal{M}'$.*

2.2.3 The Vickrey-Clarke-Groves Mechanisms

The revelation principle allows one to focus on mechanisms in which agents report their preference truthfully, however, the Gibbard-Satterthwaite Impossibility Theorem indicates that when there are three or more possible outcomes and agent preferences are unrestricted, a mechanism

is strategy-proof if and only if the social choice function it implements is dictatorial. While any dictatorial social choice function is truthfully implementable in DSE, it fails to aggregate preferences in a reasonable way.

The Gibbard-Satterthwaite impossibility theorem serves as a negative result. However, it also makes the strong assumption of unrestricted preferences, and can be avoided if such an assumption is relaxed. In this section, we present an important line of work on mechanism design, namely the family of *Vickrey-Clarke-Groves (VCG)* mechanisms, which are non-dictatorial and strategy-proof when agents have *quasi-linear* preferences. VCG mechanisms allow monetary transfer among agents, and can be used to circumvent the Gibbard-Satterthwaite Theorem in settings like auctions. Another important domain restriction on agent preferences is *single-peakedness*, which also admits non-dictatorial strategy-proof mechanisms, as will be discussed in detail in Section 2.3.

Consider a setting in which monetary transfer is allowed among agents themselves, e.g., an auction. In such a setting, an outcome is composed of two parts: a "non-payment" part determines how the outcome is chosen (e.g., the allocation of the items) and a payment part that determines how much each agent has to pay as a function of the reported types. The utility of each agent is *quasi-linear*:

Definition 2.18 (Quasi-linear utility) *The utility function of agent i is quasi-linear if:*

$$u_i(o, t_i) = v_i(o, t_i) - p_i$$

where $p_i : \prod_i T_i \to \mathbb{R}$ is the payment of agent i, and $v_i(o, t_i)$ is the valuation of agent i on outcome o.

Now, we are ready to describe the VCG mechanisms. We start from defining a more general class of mechanisms called the *Groves* mechanisms:

Definition 2.19 (Groves mechanisms) *A Groves mechanism is a direct-revelation mechanism*

$\mathcal{M} = (T_1, \ldots, T_n, g, p_1, \ldots, p_n)$ *in which:*

- T_i *is the set of types available to agent* i *and*
- g *is an outcome rule that* $g(\mathbf{t}) \in \arg\max_o \sum_i v_i(o, t_i)$*, i.e., an efficient outcome and*
- p_i *is a payment rule that* $p_i(\mathbf{t}) = h_i(\mathbf{t}_{-i}) - \sum_{i' \neq i} v_{i'}(g(\mathbf{t}), t_{i'})$*, where* h_i *is an arbitrary function on the types of all agents but* i*.*

The term $h_i(\mathbf{t}_{-i})$ has no strategic implications and can be viewed as a constant term for agent i since it does not depend on her type t_i, and the term $\sum_{i' \neq i} v_{i'}(g(\mathbf{t}), t_{i'})$ is the sum of valuations over all agents but i in the efficient outcome $g(\mathbf{t})$. The selection of h_i has a significant impact on the money paid by the agents, and some additional properties of the mechanism (see below).

It is well known that Groves mechanisms are efficient and strategy-proof when agents have quasi-linear utilities [Nisan et al., 2007]. The proof is straightforward. For each agent i, the term $\sum_{i' \neq i} v_{i'}(g(\mathbf{t}), t_{i'})$ in the payment, when added with the own valuation $v_i(g(\mathbf{t}), t_i)$, is the totally social welfare of $g(\mathbf{t})$. This means that the mechanism aligns all agents' incentives with the social objective of maximizing social welfare, which is exactly achieved by revealing their types truthfully.

Among the family of Groves mechanisms, one called *Clarke (pivotal) mechanism* is of special interest:

Definition 2.20 (Clarke mechanism) *A Clarke mechanism is a Groves mechanism in which* $h_i(\mathbf{t}_{-i}) = \max_o \sum_{i' \neq i} v_{i'}(o, t_{i'})$.

In other words, in a Clarke mechanism, the function $h_i(\mathbf{t}_{-i})$ is set to be the social welfare of the efficient outcome when agent i is removed. Such a choice of the function $h_i(\mathbf{t}_{-i})$ offers the Clarke mechanism *individual rationality* and *no negative externalities* holds. A Groves mechanism is individual rational if agents always prefer to participate in the mechanism. Besides

individual rational, the Clarke mechanism is also *weakly budget-balanced* that it does not have to be subsidized when another property called *no single-agent effect* condition is satisfied.

2.2.4 Other Possibility Results and Computational Mechanism Design

In this section, we survey some other possibility results in mechanism design (without presenting in full detail). As in the above section, we only focus on the setting of quasi-linear preferences here.

The family of Groves mechanisms characterize all mechanisms that are efficient and strategy-proof, which follows from two results. The first one is *Robert's Theorem*, which says that for quasi-linear preferences, a social choice function is implementable in DSE if and only if it is an *affine maximizer*.[3] It is also easy to see that the social welfare maximizer is a special case of an affine maximizer in which $\gamma_o = 0$ and $\omega_i = 1$ for all i. The second result is by Green and Laffont [1977], who showed that to implement an affine maximizer, it is necessary to use Groves mechanisms.

Robert's theorem also makes a strong assumption that agent valuations (i.e., v_i) are unrestricted, which is unrealistic as additional structure may be imposed in some specific domains. An interesting question is whether we can put structure on agent valuations, and derive a more general class of social choice functions that are truthfully implementable in DSE. The answer to this question is "yes". Lavi et at. [2003] showed that *weak monotonicity*[4] is an exact characterization of the truthful social choice functions in *order-based domains*. A valuation domain is said to be order-based if the domain of agent types can be characterized by ordinal constraints on agent valuations. The order-base domain includes, for example, the domain of combinatorial auctions. Saks and Yu [2005] further extended this characterization, and showed that weak monotonicity actually characterizes truthful social choice functions in the convex valuation do-

[3] A social choice function f is said to be an affine maximizer if $f(\mathbf{t}) \in \arg\max_o (\gamma_o + \sum_i \omega_i v_i(o, t_i))$, where $\gamma_o \in \mathbb{R}$ is an arbitrary constant and $\omega_i \in \mathbb{R}^+$ is the weight of agent i.

[4] A social choice function f is said to satisfy weak monotonicity, if whenever $f(t_i, \mathbf{t}_{-i}) = o_1$ and $f(t'_i, \mathbf{t}_{-i}) = o_2$, then $v_i(o_2, t'_i) - v_i(o_1, t'_i) \geq v_i(o_2, t_i) - v_i(o_1, t_i)$.

main.[5] Convex domains subsume order-based domains and includes many practical economic environments [Saks and Yu, 2005].

The computational issues arising from different aspects of mechanism design with payment are also considered. A striking example is the combinatorial auctions: To apply the VCG mechanism, one must compute the optimal allocation with all agents included, as well as, for each agent, the optimal allocation if she is removed. However, the corresponding optimization problem for each allocation and payment problem is NP-Hard [Rothkopf et al., 1998], and replacing the optimal solution with an approximate one fails to guarantee strategy-proofness of the resulting mechanism. The challenge is to design strategy-proof mechanisms that are computationally feasible, at the same time achieve reasonably good outcomes compared to the optimal one. This line of work, which has been referred to as "computational mechanism design" [Nisan and Ronen, 1999], has also produced several interesting results [Lehman et al., 2002, Nisan and Ronen, 2000, Archer and Tardos, 2001, Archer et al., 2003, Dobzinski et al., 2006].

2.3 Facility Location and Single-peaked Preferences

The Gibbard-Satterthwaite impossibility theorem makes strong assumptions of unrestricted preferences. However, in many real-world applications (e.g., auctions), there may be structure in agent preferences. We have shown in Section 2.2.3 that, when payments are allowed and agent preferences are quasi-linear, the Gibbard-Satterthwaite theorem can be escaped and VCG mechanisms are non-dictatorial and strategy-proof.

However, there are many other settings where money cannot be used as a medium of compensation. This can arise from ethical/institutional considerations: political decisions must be made without monetary transfers; organ donations can be arranged involving multiple needy

[5] An agent i's valuation domain T_i is said to be convex if for two types $t_i, t_i' \in T_i$, and the corresponding utilities $v_i(o, t_i)$ and $v_i(o, t_i')$ for outcome $o \in O$, then we have $t_i'' \in T_i$ such that $v_i(o, t_i'') = \lambda v_i(o, t_i) + (1 - \lambda) v_i(o, t_i'), \forall \lambda \in [0, 1]$. Note that this restriction is on the domain of valuations, not on the utility functions.

patients and their relatives, yet monetary compensation is illegal [Schummer and Vohra, 2007]. It is natural to ask that whether it is possible to design strategy-proof mechanisms without payments.

In this section, we address this question and describe mechanism design without money in detail. More specifically, we illustrate with the facility location problem, a classical problem with *single-peaked preferences*, and show that strategy-proof mechanisms exist in such a setting. Note that we simply use "facility" as a suggestive terminology, and the results can be generalized to many other settings, such like voting, product design, market segmentation, etc.

2.3.1 The Model

In this section, we will introduce the model of facility location problem. We start with the simple case of single-dimensional, single-facility location, and then generalize it to the multi-dimensional, multi-facility case.

Suppose the government wants to build a public library along a street for the use of nearby residents. Each resident has an single, ideal location at which he/she would like the library to be built, and her cost is the distance between her ideal and the selected location of the library. The government asks each resident to report her ideal location, and decides where to build the library based on the received reports. Note that, depending on the rule used to make the joint decision, the residents may have incentives to misreport their ideals and manipulate the outcome.

Let us define this problem more formally: Suppose we have to choose a location x to build a single facility in some one-dimensional space $O = \mathbb{R}$ (or some bounded subspace thereof). We also have a set of agents N, each with a type $t_i \in T_i$ determining her personal cost $c_i : \mathbb{R} \times T_i \to \mathbb{R}$ associated with each possible location. The objective in facility location is to select a location $x \in \mathbb{R}$ to minimize some social objective. Two of the most commonly studied are *social cost (SC)* and *maximum cost (MC)*.[6] Social cost is the sum of costs over all

[6]Alternative terminologies corresponding to social cost and maximum cost would be utilitarianism and egali-

agents and maximum cost is the cost of the worst agent. Formally, we have:

$$SC(x, \mathbf{t}) = \sum_i c_i(x, t_i) \quad \text{and} \quad MC(x, \mathbf{t}) = \max_i c_i(x, t_i)$$

It is natural to assume agent preferences are *single-peaked* in the facility location problem. Intuitively, this means the agent has a single "ideal" location, and its cost for any chosen location increases as it "moves away from" this ideal. Formally, we need only a strict ordering (or *axis*) on outcomes, rather than a distance metric, to define betweenness. Following other work in the literature, we will use the preference relation $\succeq_i$, instead of the type and utility function, to denote the preference of an agent i:

Definition 2.21 (One-dimensional axis) *An one-dimensional axis A on O is any strict ordering $<_A$ of the outcomes in O.*

Definition 2.22 (One-dimensional single-peaked preference) *[Black, 1948] Let O be a set of possible outcomes, and A be an one-dimensional axis on O. An agent i's preference $\succeq_i$ is one-dimensional single-peaked with respect to A if:*

- *There is a single, most-preferred outcome $\tau(\succeq_i) \in O$ (his ideal location or peak), satisfying $\tau(\succeq_i) \succ_i o$, $\forall o \neq \tau(\succeq_i)$*
- *For any two outcomes $\alpha, \beta \in O$, $\alpha \succeq_i \beta$ whenever we have $\beta <_A \alpha <_A \tau(\succeq_i)$ or $\tau(\succeq_i) <_A \alpha <_A \beta$*

A preference profile is one-dimensional single-peaked if there exist an one-dimensional axis A such that every agent is one-dimensional single-peaked with respect to A.

The intuition is that if outcome α is "closer" to $\tau(\succeq_i)$ than β, then α should be more preferred than β. Note that the most preferred outcome is selected based on $\succeq_i$, in the sequel, we will denote $t_i = \tau(\succeq_i)$ for convenience. While single-peaked preferences apply naturally to

tarianism.

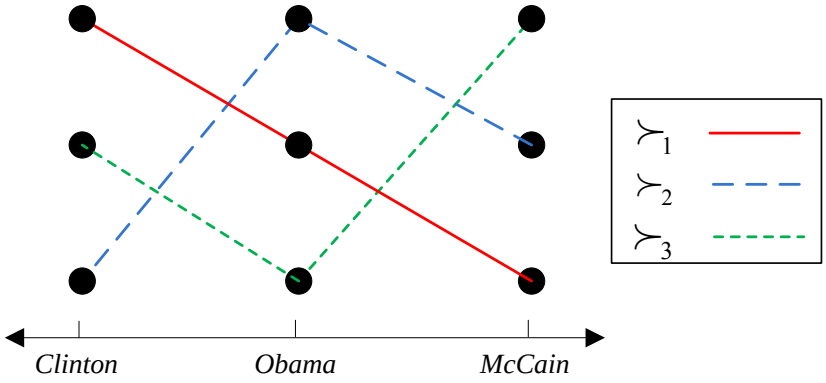


Figure 2.1: An axis and a set of agent preferences in a one-dimensional space. The preference profile $\succ = \{\succ_1, \succ_2\}$ is single-peaked with respect to the axis, however, the preference profile $\succ' = \{\succ_1, \succ_2, \succ_3\}$ is not.

problems that involve geometric/geographic distributions of physical objects like facility location, the concept has much broader application like voting, product design, customer segmentation, etc. The following example shows the use of one-dimensional single-peaked preferences in a voting:

Example 2.4 (Single-peaked preferences in voting) *Consider the voting example in Example 2.1, in which* $O = \{Obama, Clinton, McCain\}$*, and* $A = Clinton <_A Obama <_A McCain$ *be an axis (as shown in Figure 2.1). The preference profile* $\succ = \{\succ_1, \succ_2\}$*, where* $\succ_1 =$ *Clinton* $\succ_1$ *Obama* $\succ_1$ *McCain and* $\succ_2 =$ *Obama* $\succ_2$ *McCain* $\succ_2$ *Clinton is single-peaked. However, the preference profile* $\succ' = \{\succ_1, \succ_2, \succ_3\}$ *is not single-peaked (with respect to* A*). This is because when McCain is the most preferred outcome, single-peakedness (in Definition 2.22) requires that Obama should be at least as preferred as Clinton, which is violated in* $\succ_3$*. In fact, one can check that there is no axis that the preference profile* $\succ'$ *is single-peaked with respect to.*

Note that when determining whether a preference profile is single-peaked or not, one should test single-peakedness for all agents on every possible axis. This could be computationally difficult since there are many axes to consider. However, the structure in preferences imposed by single-peakedness may be used to identify the axis (or axes) efficiently. We will come back to this point later in Section 2.4.

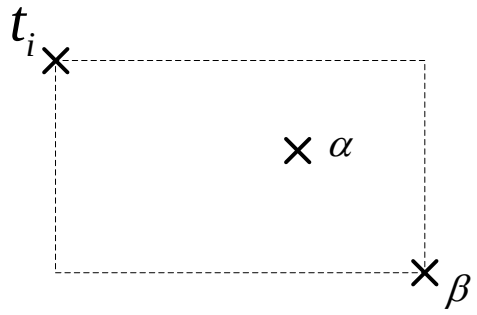


Figure 2.2: Single-peaked preference in a two-dimensional space, where t_i is agent i's peak, and outcome α is at least as preferred as outcome β.

The facility location problem can be generalized to involve multiple facilities in multiple dimensions. We start by defining single-peakedness in multi-dimensional space:

Definition 2.23 (Multi-dimensional axis) *A multi-dimensional axis $A^m = \langle A_1, \ldots, A_m \rangle$ on O is a collection of m distinct axes, each being an one-dimensional axis on O.*

Definition 2.24 (Multi-dimensional single-peaked preference) *[Barberà et al., 1993] Let O be a set of possible outcomes, and A^m be a m-dimensional axis on O. An agent i's preference $\succeq_i$ is m-dimensional single-peaked with respect to A^m if:*

- *There is a single most-preferred outcome $t_i \in O$*
- *For any two outcomes $\alpha, \beta \in O$, $\alpha \succeq_i \beta$ whenever $\beta <_{A_k} \alpha <_{A_k} t_i$ or $t_i <_{A_k} \alpha <_{A_k} \beta$ for all axes $A_k, k \leq m$*

A preference profile is multi-dimensional single-peaked if there exists a multi-dimensional axis A^m such that every agent is multi-dimensional single-peaked with respect to A^m.

In other words, if an outcome α lies within the "bounding box" of t_i and β, then α is at least as preferred as β. It is similar to the one-dimensional case in that if we move further away from i's peak t_i, then we can reach α via some path before we reach β (see Figure 2.2). Note that this requirement does not restrict i's relative preference for α and β if neither lies within the other's bounding box with respect to t_i).

In multiple facility location problems, the objective is to select a location vector $\mathbf{x} = \{x_1, x_2, \ldots, x_q\}$ of q facilities, instead of a single facility, in an m-dimensional space $\mathbb{R}^m$ (or some bounded subspace thereof). Given a location vector $\mathbf{x}$, each agent i uses the location with least cost, i.e., $c_i(\mathbf{x}, t_i) = \min_{j \leq q} c_i(x_j, t_i)$, and the objective is to select a location vector $\mathbf{x}$ to minimize the social cost $SC(\mathbf{x}, \mathbf{t}) = \sum_i c_i(\mathbf{x}, t_i)$ or to minimize maximum cost $MC(\mathbf{x}, \mathbf{t}) = \max_i c_i(\mathbf{x}, t_i)$.

Facility location can be interpreted literally, and naturally models the placement of several facilities (e.g., warehouses, libraries, etc.) in some geographic space to meet the requirement of its users. However, many other social choice problems fit within this class. Here, we list several typical problems:

- **Voting/Representative Selection.** Consider a setting in which a vote needs to be held to select one or more representatives to a committee. A typical example is fully proportional representation [Chamberlin and Courant, 1983], where voters have preferences (generally not single-peaked, but sometimes yes, see [Betzler et al., 2011, Galand et al., 2012, Lan and Elkind, 2013]) over candidates, and one looks for a set of candidates that maximizes social welfare, where the utility of a voter depends on his/her most preferred candidate in the selected set. Each candidate represents a point in a multi-dimensional space, with each axis representing a particular issue, e.g., fiscal policy, health care return, etc., and voter can also be a point in the same space, representing the stance of her ideal (perhaps hypothetical) candidate. The closer a candidate is to a voter, the more preferred that candidate is (relative to others). Modelling this as a facility location problem, voters are agents and candidates are facilities, and the voting method should be designed to respect voters' desires to the greatest extent possible.

- **Product Configuration/Customer Segmentation.** Suppose a vendor want to launch a family of several new, related products, each described by a multi-dimensional feature vector. Each customer has an ideal configuration of the product, and will prefer the product that is closest to her ideal product in the feature space. In such a setting, customers are

agents and products are facilities, and the vendor should configure products to maximize overall customer satisfaction.

- **Political Decision Making.** Consider a scenario where one or several policy decisions have to be agreed among several parties. Each policy consists of several aspects, e.g., stance on the environment, health care, human right, and each party want the policy to be made to maximize its own benefit and be close to its own stance as much as possible. Parties are agents and policy decisions are facilities, and the agreement should be achieved to minimize the objections from all parties.

2.3.2 Characterization of Strategy-proof Mechanisms for Facility Location

In this section, we introduce several results on the characterization of strategy-proof mechanisms, when agents have general/specific single-peaked preferences. While general single-peaked preferences refer to those defined in Definition 2.22 and 2.24, specific single-peaked preferences are more restricted than the general ones and assume some specific form of cost function. All of them focus on the single facility case and deterministic mechanisms.

Black [1948] first studied single-peaked preferences, and showed that when agents' preferences are single-peaked, the mechanism of choosing the median among all peaks strategy-proof. Such a mechanism is referred to as a *median mechanism*. Formally,

Definition 2.25 (Median mechanism) *Let agents have peaks of $t_1, t_2, \ldots, t_n$, and assume w.l.o.g., that $t_1 \leq \ldots \leq t_n$. The median mechanism that selects the median among all peaks. Specifically, if $n = 2t + 1$ for some integer t, then the median mechanism chooses the $(t + 1)$th peak; if $n = 2t$, then the median mechanism chooses the tth peak.*

The median mechanism is strategy-proof. To see why, consider the options of one agent, say agent 1. Let h be the index of the agent whose reported peak is the median position, then consider the following two cases: (i) If she misreports some peak $t'_1 \leq t_h$, then according to

the median mechanism, the outcome is still t_h and she is not better off; (ii) If she misreports some peak $t_1' > t_h$, then the median mechanism will choose a new outcome $o > t_h$. However, recall that agent 1's preference is single-peaked with a true peak of t_1, so we have $t_h \succeq_1 o$ and she is worse off by misreporting.

One can easily check that a mechanism that chooses any order statistic is strategy-proof, using the same reasoning above. In addition, if a mechanism always chooses a fixed location, no matter what the reports of agents are, then such a mechanism is also strategy-proof. In fact, all of these mechanisms mentioned above can be formulated as a *generalized median mechanism*, in which the outcome is computed as the median position among the actual agents and a set of phantom agents.

Moulin [1980] provided an important characterization result showing that the generalized median mechanisms composes the class of all strategy-proof mechanisms, when additional property called *anonymity*[7] is required. Formally, we have:

Theorem 2.4 (One-dimensional generalized median mechanism) *A mechanism $\mathcal{M}$ (with the outcome rule g) for single-peaked preferences in a one-dimensional space is strategy-proof and anonymous if and only if there exist $n+1$ constants $b_1, \ldots, b_{n+1} \in \mathbb{R} \cup \{-\infty, +\infty\}$ such that:*

$$g(\succeq) = \text{med}(t_1, \ldots, t_n, b_1, \ldots, b_{n+1}) \tag{2.1}$$

where med *is the median function.*

The "if" part in the theorem says that a generalized median mechanism is strategy-proof and anonymous, which can be easily checked; the "only if" part shows that there is no other mechanism, beyond the generalized median, that satisfies strategy-proofness and anonymity simultaneously. This serves as a strong theoretical result, and indicates that we can focus on generalized median mechanisms when strategy-proofness and anonymity is required in single-peaked domains. Moreover, generalized median mechanisms only require that each agent re-

[7] A social choice function f is anonymous, if for any permutation $\succeq'$ of $\succeq$, we have $f(\succeq') = f(\succeq)$ for all $\succeq$.

veals her most preferred outcome, i.e., her peak, instead of her full ranking over outcomes, which is both communicationally and computationally efficient. Here are two examples of the generalized median mechanisms:

Example 2.5 (Leftmost mechanism) *Consider the mechanism that chooses the leftmost reported peak. This mechanism is strategy-proof and anonymous, and it can be interpreted as a generalized median mechanism by setting* $b_1 = \ldots = b_n = -\infty$ *and* $b_{n+1} = +\infty$*, where the median in Equation (2.1) is* t_1*, i.e., the leftmost reported peak.*

Example 2.6 (Fixed location mechanism) *Consider the fixed location mechanism that locates the facility at position* 0 *no matter what the reports of agents are. This mechanism is strategy-proof and anonymous, and it can be interpreted as a generalized median mechanism by setting* $b_1 = \ldots = b_{n+1} = 0$*, where the median in Equation (2.1) is 0.*

Moulin [1980] also characterized the class of strategy-proof mechanisms when anonymity is not required. A non-anonymous mechanism can choose the outcome based on the identity of the agents, and is much less interesting for the social choice theory. In the sequel, we will focus only on anonymous mechanisms.

Barberà et al. [1993] generalize the result of Moulin to the multi-dimensional case, using multi-dimensional single-peaked preferences in Definition 2.24. They provided a characterization result showing a mechanism is strategy-proof and anonymous in a multi-dimensional space if and only if it is a *multi-dimensional generalized median mechanism*. A m-dimensional generalized median mechanism can be decomposed into m independent one-dimensional generalized median mechanisms, with the kth mechanism determining the coordinate of the facility on the kth dimension, for all $k \leq m$. Their results can be stated in the following theorem:

Theorem 2.5 (Multi-dimensional generalized median mechanism) *A mechanism for multi-dimensional single-peaked preferences in a multi-dimensional space is strategy-proof and anonymous if and only if it is an* m*-dimensional generalized median mechanism.*

The above theorem says that when generalized to the multi-dimensional space, the multi-dimensional generalized median mechanisms are the only mechanisms that guarantee strategy-proofness and anonymity. Moreover, the mechanism must satisfy strong separability assumption, that the kth coordinate of the facility must be only determined by the peaks on the kth dimension, for all $k \leq m$.

Besides these characterization results for general single-peaked preferences, there are some other results for more restricted preferences and domain assumptions. Massó and Moreno de Barreda [2011] considered the case where agents have *symmetric single-peaked preferences*[8] in one-dimensional spaces, and showed that a mechanism is strategy-proof and anonymous for symmetric single-peaked preferences if and only if it is a *disturbed generalized median mechanism*. The class of disturbed generalized median mechanisms is broader than the generalized median mechanisms by allowing discontinuity in some pre-defined intervals. More specifically, if the median position (with the phantom peaks) does not fall into any interval, then the disturbed generalized median mechanism selects that location; if the median is in the left half of an interval, then the disturbed generalized median mechanism selects the beginning point of the interval; if the median is in the right half of an interval, then the disturbed generalized median mechanism selects the ending point of the interval; otherwise if the median is exactly the midpoint, then a tie-breaking rule is used to determine the outcome.

Border and Jordan [1983] consider multi-dimensional settings where agents' preferences are "separable quadratic".[9] They provide a characterization result (in terms of decomposability) similar to Theorem 2.5 that a multi-dimensional mechanism for separable quadratic pref-

[8] An agent i's preference is symmetric single-peaked if $\forall o_1, o_2 \in O$, $o_1 \succeq_i o_2$ if and only if $|t_i - o_1| \leq |t_i - o_2|$. The symmetric single-peaked preferences are more restricted than general single-peaked preferences, as outcomes on different sides of an agent's peak are comparable to each other now. However, it subsume many forms of cost functions as special cases, e.g., Euclidean distance, Manhattan distance, quadratic distance, etc.

[9] The separable quadric preferences on a m-dimensional space is define as follows: Let $(o^k, o^{-k}) \in \mathbb{R}^m$ be a location in the m-dimensional space, in which o^k is the coordinate of the location on the kth dimension, and o^{-k} is the joint coordinate of the location on other dimensions. An agent i's preference is separable if her cost function satisfies $c_i((o^k, \widetilde{o^{-k}}), \succeq_i) \leq c_i((o^{k'}, \widetilde{o^{-k}}), \succeq_i) \Leftrightarrow c_i((o^k, \overline{o^{-k}}), \succeq_i) \leq c_i((o^{k'}, \overline{o^{-k}}), \succeq_i)$, for all $k \leq m$, all $o^k, o^k \in \mathbb{R}$ and all $\widetilde{o^{-k}}, \overline{o^{-k}} \in \mathbb{R}^{m-1}$. An agent i's preference is quadratic if $c_i(o, \succeq_i) = \sum_{k,k'=1}^{m} a_{kk'}(t_i - o^k)(t_i - o^{k'})$. So an agent i's cost is separable quadratic if and only if $a_{kk'} = 0$ for $k \neq k'$, and the cost function will be $c_i(o, \succeq_i) = \sum_k a_k((t_i)_k - o_k)^2$.

erences is strategy-proof, anonymous and unanimous if and only if it can be decomposed into m independent mechanisms, each being strategy-proof, anonymous and unanimous in one-dimension. In addition, a mechanism for separable quadratic preferences is strategy-proof, anonymous and unanimous if and only if it is a generalized median mechanism with two phantom peaks of $-\infty$ and $+\infty$. Note that as the set of separable quadratic preferences is a strict subset of the set of single-peaked preferences, their characterization result indicates that if we enlarge the preference domain to separable but not quadratic, we still get the same class of mechanisms that are strategy-proof, anonymous and unanimous.[10] They also show that if we allow for non-separable but quadratic preferences, then any strategy-proof mechanism must be dictatorial.

2.3.3 Approximate Mechanism Design Without Money

In the above section, we have introduced previous work on mechanism design for facility location from an economic view. However, this work focuses on the case of a single facility and the characterization of strategy-proof mechanisms. In this section, we present some work from the perspective of computer science, and show that strategy-proofness and efficiency are not compatible. We also describe several mechanisms, showing the degree to which efficiency can be approximated, when strategy-proofness is required. The focus of this section is multi-facility location, and randomized mechanisms are also allowed.

To define efficiency (and approximation), we must adopt some specific form of cost function. Note that for general single-peaked preferences, each agent i's peak t_i does not fully determine her preferences. However, if a specific form of cost function is adopted, then her type and cost function are fully determined by her ideal location. For this reason, we equate an agent i's type with her ideal point t_i.

[10] For single-peaked preferences, Moulin's characterization results indicates that a mechanism is strategy-proof, anonymous and unanimous if and only if it is a generalized median mechanism with two phantom peaks of $-\infty$ and $+\infty$. As separable preferences is between single-peaked and separable quadratic preferences, this conclusion comes immediately.

Given a location vector $\mathbf{x} = (x_1, \ldots, x_q)$, let the cost function of agent i be $c_i(\mathbf{x}, t_i) = \min_{j \leq q} ||t_i - x_j||_2$., i.e., the Euclidean distance between her peak t_i and the closest facility in $\mathbf{x}$. Recall that the objective in facility location is to minimize the social cost $SC = \sum_i c_i(\mathbf{x}, t_i)$ or maximum cost $MC = \max_i c_i(\mathbf{x}, t_i)$, we can define efficiency for them as follows:

Definition 2.26 (Efficient) *A mechanism $\mathcal{M}$ (with outcome rule g) for the facility location problem is efficient for social cost and maximum cost minimization, if $g(\mathbf{t}) \in \arg\min_{\mathbf{x}} \sum_i c_i(\mathbf{x}, t_i)$ and $g(\mathbf{t}) \in \arg\min_{\mathbf{x}} \max_i c_i(\mathbf{x}, t_i)$, respectively.*

Procaccia and Tennenholtz [2009] first studied the problem of approximate mechanism design without money. They observe that minimizing social cost for single facility location can be easily implemented by Black's median mechanism (in Definition 2.25), however, for two or more facilities, there is no way to satisfy strategy-proofness and efficiency simultaneously. They propose to achieve strategy-proofness by considering *approximately efficient* mechanisms:

Definition 2.27 (Approximation ratio) *A mechanism $\mathcal{M}$ (with outcome rule g) has an approximation ratio of ε relative to social cost and maximum cost minimization if for all $\mathbf{t}$, we have:*

$$\sum_i c_i(g(\mathbf{t}), t_i) \leq \varepsilon \cdot \min_{\mathbf{x}} \sum_i c_i(\mathbf{x}, t_i) \quad \textit{and} \quad \max_i c_i(g(\mathbf{t}), t_i) \leq \varepsilon \cdot \min_{\mathbf{x}} \max_i c_i(\mathbf{x}, t_i)$$

In other words, a mechanism has an approximation ratio of ε if the social cost (respectively, maximum cost) it achieves in an equilibrium outcome is at most ε times the minimum social cost (respectively, maximum cost) achieved by any location vector, for all type profiles $\mathbf{t}$. Note that we have $\varepsilon \geq 1$ for a minimization problem, and say a mechanism is *approximately efficient* if $\varepsilon > 1$.

They show that for two-facility location in a one-dimensional space, the *left-right* mechanism is *group strategy-proof*. A mechanism is group strategy-proof, if for any coalition of

agents, there is no joint misreport that makes everyone strictly better off:

Definition 2.28 (Group Strategy-proof) *A mechanism $\mathcal{M}$ (with outcome rule g) is group strategy-proof if, for any $S \subseteq N$, there exists an agent $i \in S$ such that:*

$$c_i(g(t_S, \mathbf{t}_{-S}), t_i) \leq c_i(g(t'_S, \mathbf{t}_{-S}), t_i), \qquad \forall t'_S, \forall \mathbf{t}_{-S}$$

Theorem 2.6 (Left-right mechanism) *Let $\mathbf{t} = (t_1, t_2, \ldots, t_n)$ be the ideal location profile. In addition, denote the leftmost location in $\mathbf{t}$ by $lt(\mathbf{t}) = \min_i t_i$, and the rightmost location by $rt(\mathbf{t}) = \max_i t_i$. The left-right mechanism, which locates the two facilities at $lt(\mathbf{t})$ and $rt(\mathbf{t})$ is group strategy-proof and has an approximation ratio of $(n-2)$ for social cost, and 2 for maximum cost for one-dimensional, two-facility location problem.*

In other words, the mechanism of locating the facilities at the leftmost and rightmost peaks is group strategy-proof and has a bounded approximation ratio. They also provide some lower bounds: for social cost minimization, there is no deterministic strategy-proof mechanism with an approximation ratio less than $3/2$ (this lower bound is improved by Lu et al. [2010] to $(n-1)/2$); for maximum cost minimization, there is no deterministic strategy-proof mechanism can give an approximation ratio smaller than 2.

An interesting observation is that by using randomization, one can allow for a broader class of mechanisms, and achieve better performances. Procaccia and Tennenholtz [2009] provided a randomized mechanism with an approximation ratio of $5/3$, breaking the lower bound of 2 for deterministic mechanisms. As maximum cost is not the focus in this thesis, we will omit further discussion.

Lu et al. [2010] also studied the two facility location problem in multi-dimensional spaces. They first proved a lower bound of $(n-1)/2$ for social cost, which confirms the conjecture of Procaccia and Tennenholtz [2009]. This linear lower bound (in the number of agents) says that there is no deterministic mechanism that guarantees an approximation ratio less than $(n-1)/2$, implying that the left-right mechanism is asymptotically optimal. Another contribution of their

paper is the randomized *proportional mechanism* with a constant approximation ratio for two-facility location problem. The mechanism locates the first facility at the position of some agent uniformly, and locates the second facility at the position of each agent with a probability proportional to the distance between that agents and the first facility. They showed that the proportional mechanism is strategy-proof, and has a constant approximation ratio of 4.

When moving to the case of more than two facilities, things are not easy. As far as we know, no previous work has proposed any strategy-proof mechanism with bounded approximation ratio for general multi-facility location problems in multi-dimensional space. Here, we discuss two results that consider the problem in restricted settings.

Fotakis and Tzamos [2010] considered *winner-imposing mechanisms* for multi-facility location in multi-dimensional spaces. Winner-imposing mechanisms only locate facilities among the reported locations from agents, and the agents whose reports are selected for placing facilities (or the winners) are only allowed to use the facilities located at their reports. Formally,

Definition 2.29 (Winner-imposing) *Given an ideal location profile* $\mathbf{t}$*, let* $P(\mathbf{t})$ *be the set of locations of all facilities, and* $P_i(\mathbf{t})$ *be the set of locations of facilities available to agent* i*. A mechanism is non-imposing if* $P(\mathbf{t}) = P_i(\mathbf{t})$*,* $\forall i$*. A mechanism is winner-imposing if* $t_i \in P(\mathbf{t})$ *implies* $P_i = \{t_i\}$*.*

Winner-imposing mechanisms seem reasonable in the sense that agents should be "responsible" for their reports, especially when positive results for non-imposing mechanisms are unknown. They showed that the winner-imposing version of the proportional mechanism is strategy-proof and has an approximation ratio of $4q$ for social cost, where q is the number of facilities to be located.

Escoffier et al. [2011] studied non-imposing mechanisms in multi-dimensional spaces, but in a very restricted setting where the number of agents is exactly one more than the number of facilities, i.e., $q = n-1$. They proposed the *inverse proportional mechanism*, which locates the

facilities at the position of all agents but i with probability proportional to the inverse distance between agent i and the closest facility located so far, and provided an approximation ratio of $n/2$ and n for social cost and maximum cost, respectively. They also provide several lower bounds: for social cost, no deterministic mechanism has an approximation ratio less than 3 and no randomized mechanism has an approximation ratio less than 1.055; for maximum cost, no deterministic mechanism has an approximation ratio less than 2.

2.4 Computational Aspects of Single-peakedness

In above sections, we have presented some work on mechanism design, assuming agent preferences are single-peaked. In addition, the axis on which the candidates (or outcomes) are positioned is also assumed to be known in advance. However, in many scenarios, whether agent preferences are single-peaked, and if so, the axis on which the candidates are positioned, is partially or fully unknown, and these questions must be answered before aggregating individual preferences. Recent research has begun to investigate computational methods to test the single-peakedness of a profile [Bartholdi and Trick, 1986, Escoffier et al., 2008], and various forms of approximation (e.g., by deleting outlier candidates, clustering candidates, deleting voters, or adding additional axes) [Escoffier et al., 2008, Faliszewski et al., 2011, Erdélyi et al., 2012, Galand et al., 2012], i.e., determining whether, given the agent preferences on a set of candidates, these preferences are (approximately) single-peaked with respect to some axis (which is usually referred to as *single-peaked consistency*), and if so, how one of the possible axes can be identified.

In this section, we discuss some recent approaches to this problem. We first present some recent work on determining single-peaked consistency, and then introduce several recently proposed approximation of single-peakedness, and the corresponding consistency problem.

2.4.1 Single-peaked Consistency

Given a preference profile, the *single-peaked consistency* problem is to determine whether there exists an axis with respect to which the preference profile is single-peaked. Following other work in the literature, we assume that the outcome (candidate) set is a finite set of $C = \{1, 2, \ldots, c\}$, and each preference relation $\succeq_i$ is an (strict or non-strict) ordering on C.

Definition 2.30 (Single-peaked consistency) *A preference profile $\succeq = (\succeq_1, \succeq_2, \ldots, \succeq_n)$ on a finite set C is single-peaked consistent if there exists an axis A (or a strict ordering $<_A$) such that $\succeq_i$ is single-peaked with respect to A for all $i \leq n$.*

The single-peaked consistency problem for a strict preference profile was first considered by Bartholdi et al. [1986] (as well as the problem of determining whether a profile is single peaked with respect to a tree [Trick, 1989], which is weaker than single-peakedness with respect to an axis). A preference profile $\succ = \{\succ_1, \succ_2, \ldots, \succ_n\}$ is strict if each preference relation $\succ_i$ is a strict preference relation, i.e., a total ordering without ties. They give an algorithm whose running time is $O(nc^2)$, where n and c are the number of agents (voters) and candidates, respectively. Escoffier *et al.* [2008] considered the same problem, and developed a more efficient algorithm with a running time of $O(nc)$.[11] Their algorithm exploits the fact that candidates ranked last in any voter's ranking must lie at the extreme points of the axis, and build an axis that is compatible with the preference profile in an "outside-in" fashion. More specifically, the algorithm proceeds by placing one or two last-ranked candidates that have not yet been placed at each step, and is repeated until all candidates have been placed (in which case we have an axis A) or a contradiction with the preference profile has been found.

Recently, Lackner [2014] studied the single-peaked consistency problem for a non-strict preference profile. A preference profile $\succeq = \{\succeq_1, \succeq_2, \ldots, \succeq_n\}$ is non-strict if each preference relation $\succeq_i$ is a weak ordering over the set C, i.e., a total ordering with ties. He showed that when the preference profile contains at least one strict preference relation, then single-peaked

[11] Thanks to Jérôme Lang, who pointed out that a similar algorithm (up to minor details) was already published in a paper by Doignon and Falmagne [1994].

consistency problem for a non-strict preference profile can be solved in $O(nc)$ time (where n and c is the number of agents and candidates, respectively), although the general problem remains open. He also provides a polynomial algorithm with a running time of $O(nc^2)$ for an arbitrary *top order* preference profile, where a preference relation is said to be top-ordered if a voter ranks only her t most-preferred candidates, where $1 \leq t \leq c-2$.

The consistency problem has also been studied for several recent notions of approximate single-peakedness, which we will discuss below.

2.4.2 Approximate Single-peakedness

While single-peakedness is a powerful concept, preference profiles are unlikely to be single-peaked in practice, especially as the number of voters or candidates becomes large.[12] Several forms of *approximate single-peakedness* have been proposed recently that allow limited violations of the constraints imposed by single-peakedness. We now outline some of these. Following this work, we assume that the preference profile $\succ = \{\succ_1, \succ_2, \ldots, \succ_n\}$ is strict.

Several approximation methods attempt to find some minimal change to the profile that would render it single-peaked. Faliszewski *et al.* [2011] consider the removal of *maverick* voters to render a profile single-peaked (e.g., perhaps certain voters are "irrational" in their declared preferences). The aim is to delete as few mavericks as possible, which measures the quality of the approximation.

Definition 2.31 (k-Maverick) *A strict preference profile $\succ$ is* k-maverick single-peaked *if a profile $\succ'$ obtained by removing at most k voters from $\succ$ is single-peaked.*

Erdélyi *et al.* [2012] consider *local candidate deletion (LCD)*, allowing the deletion of misordered candidates from each voter's preference—the notion is "local" since *different* candidates can be deleted from each $\succ_i$. The goal is to minimize the (local) number of candidates

[12] In problems defined on metric spaces, such as facility location [Procaccia and Tennenholtz, 2009, Lu et al., 2010, Escoffier et al., 2011, Dokow et al., 2012] single-peakedness is more likely to hold, but even then may be compromised by considerations apart from distance.

deleted.

Definition 2.32 (k-Local Candidate Deletion) *A strict preference profile $\succ$ is* k-local candidate deletion (k-LCD) single-peaked *if a profile $\succ'$ obtained by removing at most k candidates from each $\succ_i$ is single-peaked.*

One can also approximate single-peakedness by allowing multiple axes, where each voter must be single-peaked with respect to at least one of these axes:

Definition 2.33 (k-Additional Axis) *A strict preference profile $\succ$ is* k-additional axis (k-AA) single-peaked *if there are $k+1$ axes $A_1, \ldots, A_{k+1}$ such $\succ_i$ is single-peaked w.r.t. at least one axis, $\forall i \in N$.*

It is important to note that k-AA single-peakedness, while it implies $k+1$-dimensional single-peakedness, is not equivalent to it. It imposes the stringent requirement that each voter be single-peaked with respect to *one* of the axes, something not needed in true multi-dimensional models.

Several other notions of approximate single-peaked have also been proposed, but these are somewhat weaker than those above, so we do not investigate them. Among these are *k-Dodgson* [Faliszewski et al., 2011], which allows performing at most k swaps of adjacent candidates in each voter's ranking. k-LCD is at least as powerful, since deleting a candidate is at least effective as swapping two candidates [Erdélyi et al., 2012]. Another is *clustered single-peakedness* [Galand et al., 2012], which allows groups of candidates to be clustered and requires single-peakedness with respect to such clusters, with the aim of minimizing maximum cluster size. k-LCD (indeed "global" candidate deletion) can simulate its effects (though their quality measures are somewhat different).

The consistency problem for these notions has also been studied. Formally, we define:

Definition 2.34 (Approximately single-peaked consistency) *For $X \in \{$Maverick, Local Candidate Deletion, Additional Axis$\}$, the decision problem for approximately single-peaked con-*

sistency can be defined as follows: given a strict preference profile $\succ = \{\succ_1, \succ_2, \ldots, \succ_n\}$ *on a finite set of* C *and a positive integer* k*, is* $\succ$ k*-X single-peaked consistent?*

Erdélyi et al. [2012] showed that these decision problems are NP-complete. We will state the theorem formally as follows without further explanation:

Theorem 2.7 *The single-peaked consistency problem for Maverick, Local candidate deletion and additional axis are all NP-complete.*

2.4.3 Spatial Theory of Voting

The main criticism of single-peaked preferences is that the constraints it imposes on agent preferences are too loose, since outcomes are comparable to each other only if the betweenness relationship (with respect to the peak) holds. For instance, in Example 2.4 where the axis is *Clinton* $<_A$ *Obama* $<_A$ *McCain*, knowing *Obama* is the peak will not determine a voter's preference on *Clinton* and *McCain* even it is single-peaked. Such restrictions are even looser in multi-dimensional settings.

A more restricted model is the *spatial model*, in which both voters and candidates are represented by points in some single- or multi-dimensional space, and voter costs are computed as a function of some distance measure between themselves and the candidates. Each dimension can be interpreted as a specific issue (e.g., environment, health care or fiscal policy), and each voter's and candidate's location can be interpreted as her stance on these issues. The spatial model assumes that each voter recognizes her own stance, evaluates all possible candidates, and casts her vote according to such evaluations. A typical example is where for each agent, the most preferred candidate is the one closest to her own stance, the second most preferred candidate is the one second closest, and so on. Formally, let $N = \{1, \ldots, n\}$ be a set of voters and $C = \{1, \ldots, c\}$ be a set of candidate. A spatial model assumes that each agent i has a position $t_i \in \mathbb{R}^m$ and each candidate j has a position $c_j \in \mathbb{R}^m$, and let $d(t_i, c_j)$ be the distance between agent i and candidate j under some distance measure. Possible distance measures

can be L_1, L_2 or squared Euclidean. For each agent i, the preference $\succ_i$ over C is computed according to the distances between himself and the candidates, and a preference profile $\succ$ is a preference vector that $\succ = \{\succ_1, \ldots, \succ_n\}$.

The above model is deterministic in which the rank position of each candidate in a vote is computed deterministically according to the distance between the voter and that candidate. An alternative is a stochastic model in which the rank position of each candidate is a random variable such that the closer a candidate is to a voter, the higher the probability that candidate is more preferred. Popular stochastic choice model includes the Bradley-Terry model [Bradley and Terry, 1952] and the Plackett-Luce model [Plackett, 1975, Luce, 1959].

Plackett-Luce Model In this thesis, we focus on the Plackett-Luce model [Plackett, 1975, Luce, 1959]. Plackett-Luce is a popular stochastic choice model for comparisons involving more than two candidates, and has been widely used in horse racing [Plackett, 1975], document ranking [Cao et al., 2007], electorates modeling [Gormley and Murphy, 2007], etc. This model is known as a multi-stage model, in which the voter keeps choosing the next most preferred candidate from the set of available candidates every time until all candidates have been selected. Formally, the model is parametrized by a vector b_i for each agent i:

$$b_i = (b_{i1}, b_{i2}, \ldots, b_{ic})$$

where $\sum_{j=1}^{c} b_{ij} = 1$, $\forall i \in N$. An intuitive explanation of b_{ij} is the probability that agent i chooses candidate j as the most preferred one in her ranking. However, as we will see shortly that the Plackett-Luce is a multi-stage model and b_{ij} can be interpreted as the relative probability of choosing candidate j over any other candidates in the whole ranking subject to the normalization constraints.

It is reasonable to assume that the probability of a voter choosing a candidate in the first position is a decreasing function of the distance between that voter and candidate in the space. A popular form is some exponentially decreasing function of the distance, in which the proba-

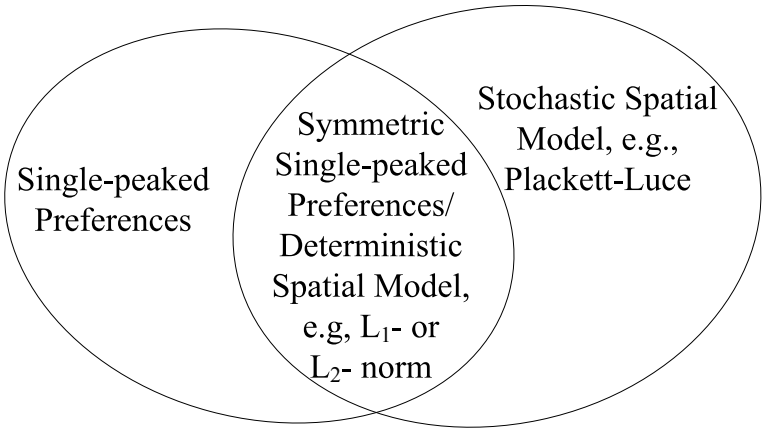


Figure 2.3: The relationship between single-peaked preferences and the spatial model.

bility of b_{ij} is computed as:

$$b_{ij} = \frac{\mathbf{exp}\{-d(t_i, c_j)\}}{\sum_{j'=1}^{c} \mathbf{exp}\{-d(t_i, c_{j'})\}}$$

where $\mathbf{exp}\{\}$ is the exponential function with the base of constant e, and $d(t_i, c_j) = \sum_{k=1}^{m}(t_i^k - c_j^k)^2$ is the squared Euclidean distance between agent i and candidate j.

At each stage, the values of b_{ij} are normalized subject to the constraints that the candidates already chosen are excluded from the vector b_i (i.e., set $b_{ij} = 0$ for those js who have been ranked). Assuming voters cast their vote independently, we can compute the probability that a preference profile $\succ$ is correct under the Plackett-Luce model with the parameter vector $\mathbf{b} = (b_1, b_2, \ldots, b_n)$. Formally, we have:

$$Pr(\succ \mid \mathbf{b}) = \prod_{i=1}^{n} \prod_{j=1}^{c} \frac{b_{ij}}{\sum_{j \succeq_i j'} b_{ij'}} \tag{2.2}$$

where the operator $\succeq_i$ means as least as preferred to. The denominator is the sum of probabilities over all candidates that are at least as preferred to j by agent i, which is used as a normalization factor.

We will end this section by discussing the relationship between single-peaked preferences

and spatial model (as shown in in Figure 2.3). The single-peaked preferences do not have to be defined over a metric space, while the spatial model is distance-based; single-peaked preferences require that the relative preference over two candidates are "deterministic" if the betweenness relationship occurs, while the rank position of each candidate in the spatial model is a random variable. In this sense, the spatial model (with stochastic choice model) can be viewed as an approximation of single-peakedness, and if deterministic choice model is used, it becomes a special case of single-peaked preference (e.g., the L_1- or L_2-cost).

Chapter 3

Quantile Mechanisms

3.1 Introduction

When agent preferences are single-peaked, previous work has shown that choosing a single alternative (e.g., a single facility) can be accomplished in a strategy-proof fashion using the well-known *median mechanism* [Black, 1948] and its generalizations [Moulin, 1980, Barberà, 2010]. Such models are used frequently for political choice, facility location, product design, customer segmentation, and related tasks, as we discussed in section 2.3.1.

Unfortunately, such mechanisms are efficient (e.g., with respect to social cost) only in very limited circumstances. Furthermore, extending these mechanisms to allow the choice of multiple alternatives (e.g., multiple facilities) generally causes even these limited guarantees to evaporate. In response, research has begun to address the question of *approximate mechanism design without money* [Procaccia and Tennenholtz, 2009], which focuses on the design of strategy-proof mechanisms for problems such as multi-facility location that are approximately efficient (i.e., have good approximation ratios) [Lu et al., 2010, Fotakis and Tzamos, 2010]. This work provides some positive results, but is generally restricted to settings involving two facilities (or adopts other restrictions) and L_2 (Euclidean) preferences.

In this chapter, we propose *quantile mechanisms*, a type of *generalized median mechanism*

(GMM) [Barberà et al., 1993, Barberà, 2010]. However, we address a more general class of problems than those tackled by GMMs. Specifically: (a) we consider problems involving selection of *multiple* alternatives (e.g., multi-facilities) in a multi-dimensional outcome space; (b) we address both social cost and maximum load as performance metrics; and (c) we analyze our mechanisms relative to L_1 (Manhattan) and L_2 (Euclidean) preferences.

Our first contribution is the analysis of the approximation ratios of quantile mechanisms under various assumptions. The performance guarantees of such mechanisms under worst-case assumptions are quite discouraging (e.g., the approximation ratio is unbounded for social cost).[1] Indeed, designing mechanisms that have worst-case guarantees may lead to poor performance in practice. Our second contribution is the development of a sample-based *empirical framework for optimizing quantile mechanisms* relative to a known preference distribution. In most realistic applications, such as facility location, product design, and many others, the designer will have *some* knowledge of the preferences of participating agents. Assuming this takes the form of a probability distribution, we use profiles sampled from this distribution to optimize quantiles while maintaining strategy-proofness. Our empirical results demonstrate that, by exploiting probabilistic domain knowledge, we obtain strategy-proof mechanisms that outperform mechanisms designed to guard against worst-case profiles. Our framework can be viewed as a form of *automated mechanism design (AMD)*, which advocates the use of preference (or type) distributions to optimize mechanisms [Conitzer and Sandholm, 2002b, Sandholm, 2003].

3.2 One-dimensional Quantile Mechanisms

We begin with one-dimensional facility location problems to develop intuitions. Following the notation from Section 2.3, we use n and q to denote the number of agents and facilities,

[1] A later characterization result by Fotakis and Tzamos [2012] shows that for multi-facility location problem, there is no deterministic and strategy-proof mechanism with a bounded approximation ratio. In the sequel, we use the term "bounded" if the approximation ratio is bounded by any function of the number of agents; otherwise, we say the approximation ratio is unbounded.

respectively. Our objective is to select q homogeneous facilities in a line, which is represented by a location vector $\mathbf{x} = (x_1, x_2, \ldots, x_q)$, where $x_j \in \mathbb{R}$. We consider two specific forms of cost functions: L_1 (Manhattan) and L_2 (Euclidean) distances. More specifically, given a location vector $\mathbf{x}$, we define the *distance-based cost function* for agent i as follows:

$$c_i(\mathbf{x}, t_i) = \min_{j \leq q} ||t_i - x_j||_p$$

where t_i is the peak of agent i, and $p \in \{1, 2\}$ reflects either L_1 or L_2 distance from agent i's nearest facility. In other words, $c_i(\mathbf{x}, t_i)$ reflects the cost to agent i of using her "closest" facility in $\mathbf{x}$ under the relevant norm (L_1 or L_2). As in Section 2.3, we use $t_i = \tau(\succeq_i)$ to denote the peak of agent i.

We use $x^p[i, \mathbf{x}]$ to denote the closest facility of agent i in the location vector $\mathbf{x}$ under the L_p-norm (where $p \in \{1, 2\}$), and define the *load* of facility j given location vector $\mathbf{x}$ and type profile $\mathbf{t}$ as $l_j(\mathbf{x}, \mathbf{t}) = \#\{i : x^p[i, \mathbf{x}] = j\}$, i.e., the number of agents using facility j. The objective of the facility location problem is to choose a location vector $\mathbf{x}$ to minimize *social cost (SC)* or *maximum load (ML)*:

$$SC(\mathbf{x}, \mathbf{t}) = \sum_i c_i(\mathbf{x}, t_i) \quad \text{or} \quad ML(\mathbf{x}, \mathbf{t}) = \max_j l_j(\mathbf{x}, \mathbf{t})$$

Social cost is a natural objective as it reflects the social welfare over all agents. From an algorithmic perspective, minimizing social cost is equivalent to the geometric p-median problems [Megiddo and Supowit, 1984] (we will talk about the p-median problem in Section 5.3.1), which has been extensively studied in the literature [Kuhn, 1973, Vardi and Zhang, 2000, Arora et al., 1998, Lin and Vitter, 1992]. Maximum load also makes sense, for instance, when a product designer launches a family of q new products, consumers purchase the product closest to their ideal product, but costs are minimized by balancing production; or when facility management costs increase super-linearly with load. Many other fundamental social objectives, such as fairness (e.g., maximum agent distance), and combinations thereof can be adopted

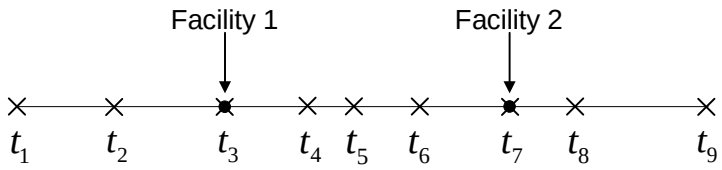


Figure 3.1: The $(0.25, 0.75)$-quantile mechanism for a two-facility location problem when $n = 9$.

depending on one's design goals.

Without loss of generality, we rename the agents so their ideal locations are ordered: $t_1 \leq t_2 \leq \ldots \leq t_n$. Then we can define the *quantile mechanisms* as follows:

Definition 3.1 (Quantile mechanism) *Let* $\mathbf{p} = (p_1, p_2, \ldots, p_q)$ *be a* q*-vector in the unit cube, where* $0 \leq p_1 \leq p_2 \ldots \leq p_q \leq 1$*. A* $\mathbf{p}$*-quantile mechanism locates the* j*th facility at the* p_j*th quantile of the reported ideal location.*[2] *Formally, we have:*

$$x_j = t_{i_j}, \text{ where } i_j = \lfloor (n-1) \cdot p_j \rfloor + 1, \qquad \forall j \leq q$$

In other words, the quantile mechanism locate each facility at a pre-specified quantile among the reported peaks independently. Recall that any quantile (or order statistic) can be implemented by arranging the phantom peaks in the generalized median mechanism, so our quantile mechanism can be decomposed into q independent GMMs. The following example shows a $(0.25, 0.75)$-quantile mechanism for a two-facility location problem:

Example 3.1 *We illustrate the* $(0.25, 0.75)$*-quantile mechanism for a two-facility problem with* $n = 9$ *agents in Figure 3.1. Ordering the nine agents' reported locations so that* $t_1 \leq \ldots \leq t_9$*, the mechanism locates the first facility at* $x_1 = t_3$ *(since* $\lfloor 8 \cdot 0.25 \rfloor + 1 = 3$*) and the second at* $x_2 = t_7$*.*

The following theorem shows that the quantile mechanism is group strategy-proof:

[2] We could equivalently use order statistics; but the quantile formulation removes dependence on the number of the agents in the mechanism's specification.

Theorem 3.1 (Strategy-proofness of quantile mechanism) *The* $\mathbf{p}$*-quantile mechanism is group strategy-proof for any quantile vector* $\mathbf{p}$*.*

Proof: We prove group strategy-proofness for the case of $q = 2$, and describe how it can be extended to $q > 2$ below.

Let $S \subseteq N$ be a coalition of agents, $\mathbf{x} = (x_1, x_2)$ be the location vector if all agents truthfully report their ideals, and $\mathbf{x}' = (x_1', x_2')$ be the location vector if agents in S jointly deviate from their peaks (assuming reports from agents in $N \backslash S$ remain fixed). In addition, let $\Delta_1 = x_1 - x_1'$ and $\Delta_2 = x_2' - x_2$. We show that if either Δ_1 or Δ_2 is strictly greater or strictly less than 0, some agent in S is worse off in the outcome $\mathbf{x}'$ than she is in $\mathbf{x}$, which is sufficient to establish (group) strategy-proofness. There are four cases to consider:

I. $\Delta_1 \geq 0$ and $\Delta_2 \geq 0$. We can ignore the case where both Δ_1 and Δ_2 are 0, since no agent in S gains by misreporting if neither facility moves. Assume, w.l.o.g., that $\Delta_1 > 0$ and $\Delta_2 \geq 0$. Recall that x_1 is the p_1th quantile among all reported peaks. Hence $\Delta_1 > 0$ implies that some agent $i \in S$, with $t_i \geq x_1$, reports a new ideal to the left of x_1. Agent i's cost is now:

$$c_i(\mathbf{x}', t_i) = \min\{t_i - x_1', x_2' - t_i\} \geq \min\{t_i - x_1, x_2 - t_i\} = c_i(\mathbf{x}, t_i)$$

So agent i has a greater or equal cost in $\mathbf{x}'$ than in $\mathbf{x}$.

II. $\Delta_1 \geq 0$ and $\Delta_2 < 0$. In this case, there must be an $i \in S$, with $t_i \geq x_2$, that reports a new ideal to the left of x_2; it's cost is:

$$c_i(\mathbf{x}', t_i) = t_i - x_2' > t_i - x_2 = c_i(\mathbf{x}, t_i)$$

Hence agent i has a greater cost by misreporting.

III. $\Delta_1 < 0$ and $\Delta_2 \geq 0$. This case is completely symmetric to case II.

IV. $\Delta_1 < 0$ and $\Delta_2 < 0$. The case is similar to case II: There must be an $i \in S$ whose ideal is to the right of x_2 but misreports to the left of x_2, increasing its cost.

This establishes group strategy-proofness for $q = 2$.

For the case of $q > 2$, we can define Δ_t for each $1 \leq t \leq q$. By using case analysis on all possible combinations of Δ_t, we can always find an agent in S who is not strictly better off, which completes our proof. ■

Since any quantile mechanism is group strategy-proof for any class of single-peaked preferences, it prevents strategic manipulation even when applied to specific cost/preference functions such as L_1 and L_2 cost. Unfortunately, quantile mechanisms can give rise to poor approximation ratios when we consider specific cost functions, specifically, L_1 or L_2 costs.

Theorem 3.2 *Let agents have L_1 or L_2 preferences. Let $\mathbf{p} = (p_1, p_2, \ldots, p_q)$ define a quantile mechanism $\mathcal{M}$. If $q \geq 3$, the approximation ratio of $\mathcal{M}$ with respect to social cost is unbounded. The approximation ratio with respect to maximum load is $q \cdot z$, where $z = \max_{1 \leq j \leq q}(p_{j+1} - p_{j-1})$ (where $p_0 = 0$ and $p_{q+1} = 1$).*

Proof: We first show the approximation ratio is unbounded when the objective is social cost minimization. The intuition is that for any quantile vector $\mathbf{p}$, there is a type profile for which optimal social cost is arbitrarily small, while the mechanism-induced social cost is constant. We prove this for the case of $q = 3$ and describe how it can be extended to $q > 3$.

Let $q = 3$, and assume that each agent's type is one of only four possible ideal locations, 0, δ, 2 and 3, where $0 < \delta < 1$ (as shown in Figure 3.2). For any quantile vector $\mathbf{p} = (p_1, p_2, p_3)$, consider a type profile $\mathbf{t} = (t_1, t_2, \ldots, t_n)$ where $t_1 = \ldots = t_{i_1} = 0$ and $t_{i_1+1} = \ldots = t_{i_2} = \delta$, with $\lfloor (n-1) \cdot p_1 \rfloor + 1 \leq i_1 < \lfloor (n-1) \cdot p_2 \rfloor + 1 \leq i_2$. Given these reports, the $\mathbf{p}$-quantile mechanism locates the first two facilities at locations 0 and δ. In addition, let n_1, n_2, n_3 and n_4 be the number of agents whose ideal locations are at 0, δ, 2 and 3, respectively. When δ is small enough, the mechanism incurs a social cost of n_3 (if the third facility is located at 3) or n_4 (if the third facility is located at 2). However, the optimal location of the three facilities for

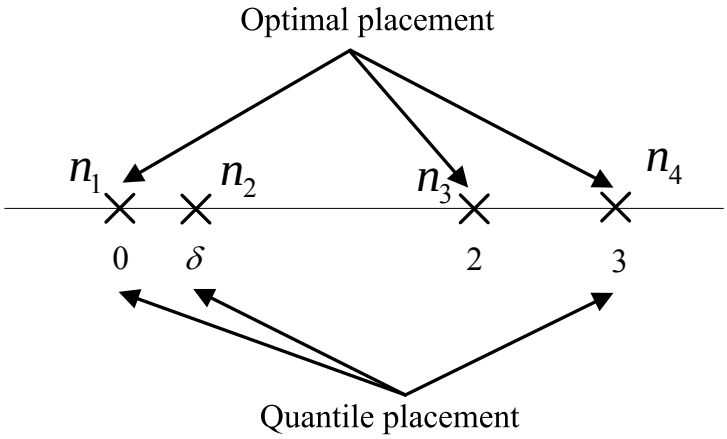


Figure 3.2: Unbounded approximation ratio of quantile mechanism with respect to social cost.

this profile is 0, 2 and 3, which has optimal social cost of $n_2 \cdot \delta$ (assuming $n_1 \geq n_2$). Thus the approximation ratio is $n_3/(n_2 \cdot \delta)$ or $n_4/(n_2 \cdot \delta)$, which is unbounded as $\delta \to 0$.

To extend to $q > 3$, we can construct a similar ideal location profile such that two out of q facilities are located arbitrarily close at 0 and δ. As $\delta \to 0$, the optimal social cost is arbitrarily small and the social cost induced by the quantile mechanism is not, leading to an unbounded approximation ratio as well.

For maximum load, assume a quantile vector $\mathbf{p} = (p_1, p_2, \ldots, p_q)$, and the induced location vector $\mathbf{x} = (x_1, \ldots, x_q)$. For each facility $1 \leq j \leq q$, the number of agents using facility x_j is at most $l_j(f(\mathbf{t}), \mathbf{t}) = n \cdot (p_{j+1} - p_{j-1})$; this occurs when each agent with a peak in (x_{j-1}, x_{j+1}) is closest to x_j, in which case maximum load is $ML(f(\mathbf{t}), \mathbf{t}) = n \cdot z$, where $z = \max_{1 \leq j \leq q}(p_{j+1} - p_{j-1})$ (if we let $p_0 = 0$ and $p_{q+1} = 1$). However, optimal maximum load, which is $\lceil n/q \rceil$, occurs when all the facilities are in the same location and agents are "assigned" to each facility evenly. So the approximation ratio is $\frac{n \cdot z}{\lceil n/q \rceil} \approx q \cdot z$. ∎

Notice that the theorem does not hold for social cost with $q = 2$ facilities: the *left-right mechanism*, or $(0, 1)$-quantile mechanism in our terminology, has a bounded approximation ratio of $n - 1$ for social cost [Procaccia and Tennenholtz, 2009]. Indeed, the $(0, 1)$-quantile mechanism is the *only* mechanism within the quantile family that has a bounded approximation ratio, and the only anonymous, deterministic mechanism with a bounded approximation

ratio for single-dimensional, two-facility location problem (see the characterization results of Fotakis and Tzamos [2012] for details). For $q \geq 3$, not only does no quantile mechanism have bounded approximation ratio, it has recently been shown that no deterministic and strategy-proof mechanism has bounded approximation ratio [Fotakis and Tzamos, 2012]. This gives further motivation for the use of probabilistic prior distributions to optimize quantiles for average-case performance rather than worst-case performance (see Section 3.4).

With respect to maximum load, it is natural to ask which quantile vector $\mathbf{p}$ minimizes z in Theorem 3.2. We can show that the quantile mechanism that "evenly distributes" facilities is approximately optimal, and that it has the smallest approximation ratio within the family.

Proposition 3.1 *Let agents have L_1 or L_2 preferences. If q is odd, then the quantile mechanism with $p_j = \frac{j}{q+1}$, $\forall 1 \leq j \leq q$, is $\frac{2q}{q+1}$-optimal w.r.t. maximum load. If q is even, then the quantile mechanism with $p_j = p_{j+1} = \frac{j+1}{q+2}$, $\forall j = 2j' - 1, 1 \leq j' \leq q/2$, is $\frac{2q}{q+2}$-optimal w.r.t. maximum load. In each case, the mechanism has the smallest approximation ratio within the quantile family.*

Proof: We prove the proposition for cases where $q = 2l + 1$ is odd for some integer l. The case for even q is similar.

We first show the approximation ratio of the mechanism in which $p_j = j/(q+1)$. According to Theorem 3.2, the maximum load is at most $2n/(q+1)$. However, the optimal placements of facilities can induce a maximum load of $\lceil n/q \rceil$, so the approximation ratio is $2n/\left((q+1) \cdot \lceil n/q \rceil\right) = 2q/(q+1)$.

Next, we show that this mechanism achieves the smallest approximation ratio within the family. Suppose by contradiction that there is a mechanism $\mathcal{M}'$ with the quantile vector $\mathbf{p}' = (p'_1, p'_2, \ldots, p'_q)$, who has a smaller approximation ratio, then it must be the case that $z' =$

$\max_{1\leq j\leq q}\left(p'_{j+1}-p'_{j-1}\right) < 2/(q+1)$. Again, defining $p'_0 = 0$ and $p'_{q+1} = 1$, we have:

$$\begin{aligned} p'_{q+1} - p'_0 &= (p'_{q+1} - p'_{q-1}) + \ldots + (p'_2 - p'_0) \\ &< \underbrace{\frac{2}{q+1} + \ldots + \frac{2}{q+1}}_{l+1 \text{ times}} \\ &= \frac{2(l+1)}{2l+2} = 1 \end{aligned}$$

This contradicts the fact that $p'_0 = 0$ and $p'_{q+1} = 1$, so such mechanism $\mathcal{M}'$ does not exists. To summary, the quantile mechanism in which $p_j = j/(q+1)$ achieves the smallest approximation ratio of $2q/(q+1)$ with respect to maximum load within the quantile family. ■

Note that for even q, the mechanism is partially imposing. We locate two facilities at each selected location, and balance the agents choosing any location; they are indifferent to the "imposed" assignment, so it isn't truly imposing mechanisms (we don't remove choice from the agents [Fotakis and Tzamos, 2010]). We use this for convenience; there are strictly non-imposing mechanisms with the same ratio.

3.3 Multi-dimensional Quantile Mechanisms

We have shown that when locating multiple facilities in a single-dimensional space, the quantile mechanisms are (group) strategy-proof. However, as we have discussed in Section 2.3.1, many social choice problems can be interpreted as "facility location" problems when viewed as making choices in a higher dimensional space, such as selection of political/committee representatives, product design, and the like. For example, in a voting where multiple candidates will be selected, each candidate might be represented by his/her stand on economics, medical and social insurance, human rights and foreign policy, assuming these are the most important characteristics that distinguish the candidates in the minds of voters. Thus each candidate can be represented as a point in this 4-dimensional space, with each dimension describing his/her position on one of these issues. In this section, we generalize quantile mechanisms to multi-

dimensional spaces, and prove some properties as in the single-dimensional case.

We assume that agents have multi-dimensional single-peaked preferences (see Definition 2.24). Similarly, let n be the number of agents, q be the number of facilities and m be the number of dimensions. Each agent i has an ideal location $t_i \in \mathbb{R}^m$, and the objective is to select a location vector $\mathbf{x} = (x_1, x_2, \ldots, x_q)$ (where $x_j \in \mathbb{R}^m$) to minimize the social cost $SC(\mathbf{x}, \mathbf{t}) = \sum_i c_i(\mathbf{x}, t_i)$ or maximum load $ML(\mathbf{x}, \mathbf{t}) = \max_j l_j(\mathbf{x}, \mathbf{t})$, assuming each agent uses the facility with least distance under the L_1 or L_2 norm.

For any type profile $\mathbf{t}$, we use $t_1^k \leq t_2^k \leq \ldots \leq t_n^k$ to denote the *ordered projection* of $\mathbf{t}$ in the kth dimension for all $k \leq m$. In other words, we simply order the reported coordinates in each dimension independently. An *m-dimensional quantile mechanism* is defined as follows:

Definition 3.2 (Multi-dimensional quantile mechanism) *Let $\mathbf{P}$ be a $q \times m$ quantile matrix $\mathbf{p} = (\mathbf{p}_1; \mathbf{p}_2; \ldots; \mathbf{p}_q)$, where each $\mathbf{p}_j \in [0,1]^m$ is an m-vector in the unit cube, with $\mathbf{p}_j = (p_j^1, p_j^2, \ldots, p_j^m)$. Given a reported profile $\mathbf{t}$, the $\mathbf{P}$-quantile mechanism locates the jth facility by selecting, for each dimension $k \leq m$, the p_j^kth quantile of the ordered projection of $\mathbf{t}$ in the kth dimension as the coordinate of facility j in that dimension. Formally:*

$$x_j = (t^1_{\lfloor (n-1)\cdot p_j^1 \rfloor + 1}, t^2_{\lfloor (n-1)\cdot p_j^2 \rfloor + 1}, \ldots, t^m_{\lfloor (n-1)\cdot p_j^m \rfloor + 1}), \qquad \forall j \leq q, \forall k \leq m$$

The following example shows a two-dimensional quantile mechanism for two-facility location.

Example 3.2 *Consider the example shown in Figure 3.3, in which two facilities are to be located in a two-dimensional space for $n = 11$ agents. With $\mathbf{P} = (0.2, 0.7; 0.8, 0.3)$, the $\mathbf{P}$-quantile mechanism locates the first facility at the x-coordinate of t_3 (since $\lfloor 10 \cdot 0.2 \rfloor + 1 = 3$) and at the y-coordinate of t_8; and the second facility is placed at the x-coordinate of t_9 and the y-coordinate of t_4. Notice facilities need not be located at the ideal point of any particular agent.*

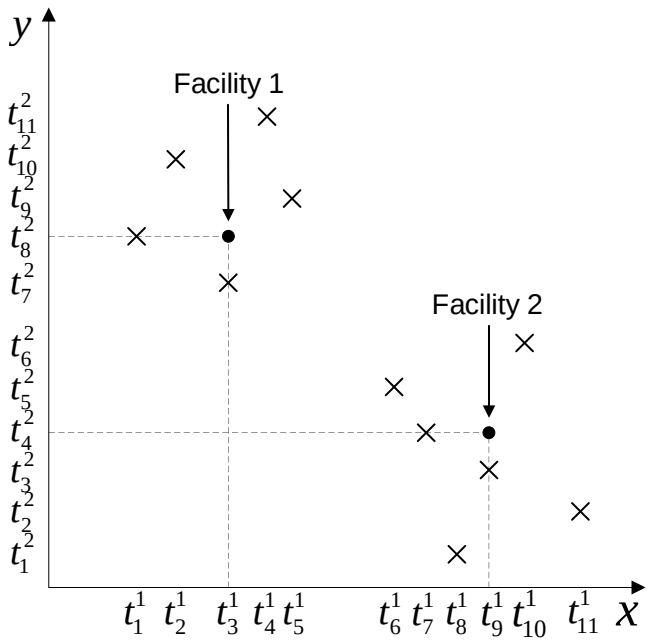


Figure 3.3: A quantile mechanism for a two-dimensional, two-facility location problem with $n = 11$ agents.

The following theorem says that the m-dimensional quantile mechanism is strategy-proof. In fact, as each facility is located using a generalized median mechanism independently, strategy-proofness follows directly from that of GMM.

Theorem 3.3 *The m-dimensional $\mathbf{P}$-quantile mechanism is strategy-proof for any quantile matrix $\mathbf{P}$.*

Unlike the single-dimensional case, the mechanism is not *group strategy-proof*: Consider the two-dimensional case, where two agents i and j can collude to misreport their preferences such that i's misreport benefits j in one dimension, and j's misreport benefits i in the other, making both better off (see Example 4.1). We will talk about the incentive and complexity of group manipulation later Chapter 4 and 5.

The following results generalize the corresponding one-dimensional results above.

Theorem 3.4 *Let agents have L_1 or L_2 preferences, and $\mathbf{P}$ define a quantile mechanism $\mathcal{M}$ for an m-dimensional, q-facility location problem with $m > 1$. The approximation ratio of $\mathcal{M}$*

is unbounded with respect to social cost for $q \geq 2$. The approximation ratio of $\mathcal{M}$ is $q \cdot z$ with respect to maximum load, where $z = \max_{k \leq m} \max_{j \leq q}(p^k_{j+1} - p^k_{j-1})$ (let $p^k_0 = 1$ and $p^k_{q+1} = 1$ for all $k \leq q$).

Proof: The unbounded approximation ratio with respect to social cost for $q \geq 3$ follows directly from Theorem 3.2, so we only have to prove for the case of $q = 2$. Consider the following two-dimensional quantile mechanism in which $\mathbf{P} = (p^1_1, p^2_1; p^1_2, p^2_2)$, and without loss of generality, let us assume $p^1_1 \leq p^1_2$. Then there are two cases:

- $p^2_1 \leq p^2_2$. Consider the type profile $\mathbf{t} = (\overbrace{((0,a),\ldots,(0,a))}^{t \text{ copies}}, \overbrace{(a,0),\ldots,(a,0)}^{n-t \text{ copies}})$, where $a > 0$ is a positive real number. There are five possible outcomes for a quantile mechanism, in which the two facilities at located at $(0,0)$ and $(a,0)$, or $(0,a)$ and (a,a), or $(0,0)$ and (a,a) or both at $(a,0)$ or both at $(0,a)$, all leading to a positive social cost. However, the optimal social cost is to locate two facilities at $(0,a)$ and $(a,0)$, which has a social cost of 0, inducing an unbounded approximation ratio.

- $p^2_1 > p^2_2$. Similarly, if we consider the profile $\mathbf{t} = (\overbrace{((0,0),\ldots,(0,0))}^{t \text{ copies}}, \overbrace{(a,a),\ldots,(a,a)}^{n-t \text{ copies}})$, then the quantile mechanism will have a strictly positive social cost, while the minimum social cost is 0, inducing an unbounded approximation ratio.

  The approximation ratio with respect to maximum load is computed similarly as in Theorem 3.2. For each facility j and each dimension k, the number of agents whose peaks are in (x^k_{j-1}, x^k_{j+1}) is at most $n \cdot (p^k_{j+1} - p^k_{j-1})$, and the maximum load achieved by maximizing over all facilities and dimensions, i.e., $ML(f(\mathbf{t}), \mathbf{t}) = n \cdot z$, where $z = \max_{k \leq m} \max_{j \leq q}(p^k_{j+1} - p^k_{j-1})$. However, the optimal maximum load is $\lceil n/q \rceil$, and the approximation ratio is $n \cdot z / \lceil n/q \rceil \approx q \cdot z$.

■

As in the single-dimensional case, we can optimize the quantiles for maximum load, when $q = \tilde{q}^m$ for some integer $\tilde{q}$ by exploiting Proposition 3.1 in each dimension:

Proposition 3.2 *Let $q = \tilde{q}^m$. If $\tilde{q}$ is odd, the mechanism that locates one facility at each $\frac{1}{\tilde{q}+1}$th quantile in each dimension is $\left(\frac{2\tilde{q}}{\tilde{q}+1}\right)^m$-optimal w.r.t. maximum load. If $\tilde{q}$ is even, the mechanism that locates two facilities at each $\frac{2}{\tilde{q}+2}$th quantile in each dimension is $\left(\frac{2\tilde{q}}{\tilde{q}+2}\right)^m$-optimal w.r.t. maximum load. Moreover, these are the smallest approximation ratios within the family of quantile mechanisms.*

Proof: The proof is similar to that of Proposition 3.1. We only show for cases for $\tilde{q} = 2l + 1$ being odd for some integer l. The case for even $\tilde{q}$ is similar.

We first show the approximation ratio of the given mechanism. In such a case, the maximum load is at most $n \cdot (\prod_{k \le m} \max_{j \le q}(p^k_{j+1} - p^k_{j-1})) = n \cdot (2/(q+1))^m$. Note that this is smaller than the upper bound we achieve in Theorem 3.4 as the facilities are arranged in grids. On the other hand, the optimal placement of facilities can induce a maximum load of $\lceil n/\tilde{q}^m \rceil$, so the approximation ratio is $(2\tilde{q}/(\tilde{q}+1))^m$.

We also show that this mechanism achieves the smallest approximation ratio within the family. Suppose by contradiction that there is a mechanism $\mathcal{M}'$ with the quantile matrix $\mathbf{P}' = ((p^{1'}_1, \ldots, p^{m'}_1); \ldots; (p^{1'}_q, \ldots, p^{m'}_q))$, who has a smaller approximation ratio, then it must be the case that $z' = \max_{k \le q} \max_{j \le q} \left(p^{k'}_{j+1} - p^{k'}_{j-1}\right) < (2/(\tilde{q}+1))^m$. Again, defining $p^{k'}_0 = 0$ and $p^{k'}_{q+1} = 1$ for all $k \le m$, we have:

$$\begin{aligned} p^{k'}_{q+1} - p^{k'}_0 &= (p^{k'}_{q+1} - p^{k'}_{q-1}) + \ldots + (p^{k'}_2 - p^{k'}_0) \\ &< \underbrace{\left(\frac{2}{\tilde{q}+1}\right)^m + \ldots + \left(\frac{2}{\tilde{q}+1}\right)^m}_{l+1 \text{ times}} \\ &= \left(\frac{1}{l+1}\right)^{m-1} < 1 \end{aligned}$$

This contradicts the fact that $p^{k'}_0 = 0$ and $p^{k'}_{q+1} = 1$. Hence there is no mechanism with smaller approximation ratio. ■

3.4 A Sample-based Optimization Framework

We have seen that quantile mechanisms are strategy-proof for general m-dimensional, q-facility location problems, and can offer bounded approximation ratios for L_1 and L_2 preferences (though only under certain conditions for social cost). Unfortunately, these guarantees require optimizing the choice of quantiles with respect to worst-case profiles, which can lead to poor performance in practice. For example, in a single-dimensional, two-facility location problem, decent approximation guarantees for social cost require using the $(0, 1)$-quantile mechanism (or the left-right mechanism [Procaccia and Tennenholtz, 2009]) ; but if agent preferences are uniformly distributed in one dimension, this mechanism will perform quite poorly. Intuitively, the $(0.25, 0.75)$-quantile mechanism should have lower expected social cost due to its "probabilistically suitable" placement of two facilities, each for use by half of the agents.

We consider a framework for empirical optimization of quantiles within the family of quantile mechanisms that admits much better performance in practice. As in *automated mechanism design (AMD)* [Conitzer and Sandholm, 2002b, Sandholm, 2003], we assume a prior distribution D over agent preference profiles. One will often assume a prior model D (e.g., learned from observation) that renders individual agent preferences independent *given* that model, but this is *not* a requirement for our method. In many settings, such as facility location or product design, such distributional information will readily be available. We sample preference profiles from this distribution, and use them to optimize quantiles to ensure the best expected performance with respect to our social objective.

Unlike classic AMD, we restrict ourselves to the specific family of quantile mechanisms. While this limits the space of mechanisms, we do this for several reasons. First, it provides a much more compact mechanism parameterization over which to optimize than in typical AMD settings.[3] Second, since the resulting mechanism is "automatically" strategy-proof, no matter which quantiles are chosen, the optimization need not account for incentive constraints. Third,

[3]Automated mechanism design has been explored in parameterized mechanisms, e.g., in combinatorial auctions [Likhodedov and Sandholm, 2004, 2005].

our optimized quantile mechanisms are *responsive* to specific preferences of agents, such that the locations of facilities vary with different preference profiles. This stands in contrast to what we term *Bayesian optimization*. In Bayesian optimization, the placement of facilities relative to the prior is chosen by sampling profiles from the distribution *without eliciting the actual preferences of agents*, and placing facilities that minimize average social cost relative to the samples. In our model, samples are used to determine the mechanism's *quantiles*; the actual placement is made using these once preferences are elicited. (We empirically compare our sample-optimized quantile mechanisms to direct Bayesian optimization below.)

Let agent type profiles $\mathbf{t} = (t_1, t_2, \ldots, t_n)$ be drawn from distribution D. Given a $\mathbf{P}$-quantile mechanism, let $f_{\mathbf{P}}(\mathbf{t})$ denote the chosen locations when the agent type profile is $\mathbf{t}$. The goal is to select $\mathbf{P}$ to minimize the expected social cost or maximum load:

$$\min_{\mathbf{P}} \mathbb{E}_D \left[SC(f_{\mathbf{P}}(\mathbf{t}), \mathbf{t}) \right]; \quad \text{or} \quad \min_{\mathbf{P}} \mathbb{E}_D \left[ML(f_{\mathbf{P}}(\mathbf{t}), \mathbf{t}) \right]$$

Naturally, other objectives can be modelled in this way too.

Given W sampled preference profiles, we optimize quantile selection relative to the W sampled profiles. For small problems, we use simple exhaustive optimization for this purpose. Specifically, we consider all possible values for the percentile matrix $\mathbf{P}$. For each, we compute the average social cost (maximum load) over W sample profiles, and select the one with minimum objective value. This is feasible for problems of small size we consider.

For large problems, one can formulate the minimization problem as a mixed integer linear program (MILP) for both L_1 and L_2 cost, and use standard optimization tools, e.g., CPLEX, to solve the problem (for social cost minimization only). Relaxed formulations require $O(n^m)$ variables however, rendering them intractable for problems with large numbers of agents. We also we experimented with gradient and coordinate descent algorithms from random starting points (i.e., $\mathbf{P}$-matrices) on all of the problems described below. These worked extremely well: no run of either algorithm on the problems below converged to a solution more than 2%

from optimal (on avg. within 0.5% of optimal); and with 100 random restarts, both methods found the optimal solution in every instance (and did so quickly, in times ranging from 0.88–1.97 sec.). Details on the formulations of the MILP, and the gradient and coordinate descent algorithms are described in the appendix of this chapter for references.

3.5 Empirical Evaluation

In this section, we present an empirical evaluation of the practical performance of our quantile mechanisms. Specifically, we consider problems with $n = 101$ agents in the following experiments, with agent preferences drawn independently from three classes of distributions: uniform D_u, Gaussian D_g and mixtures of Gaussians D_{gm} with 3 components.[4] Each distribution reflects rather different assumptions about agent preferences: that they are spread evenly (D_u); that they are biased toward one specific location (D_g); or that they partitioned are into 2 or 3 loose clusters (D_{gm}). In all cases, $W = 500$ sampled profiles are used for optimization. We examine results for both social cost and maximum load.

3.5.1 One-dimensional mechanisms

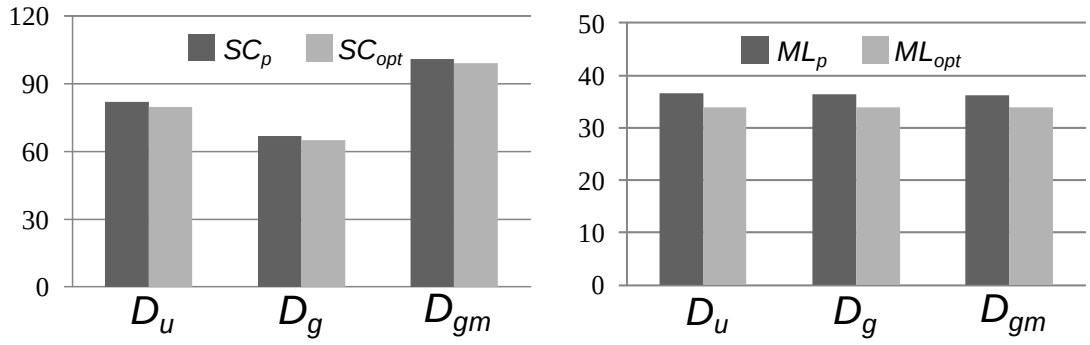
We begin with simple one-dimensional problems with $q = 2$, 3 or 4. Table 3.1 shows the quantiles resulting from our optimization for both SC and ML under each of the three distributions. For example, when agent ideal locations are uniformly distributed, the $(0.25, 0.75)$-quantile mechanism is optimal in terms of minimizing the expected social cost for two facilities. This is expected, since the uniform (and Gaussian) distribution partitions agents into two groups of roughly equal size, and facilities should be located at the median positions of each group.

The performance of the optimized quantile mechanisms are extremely good. Figure 3.4 compares the expected social cost and maximum load of our mechanisms with those given by *optimal placement* of facilities for the case of $q = 3$. Recognize however that optimal

[4] D_u is uniform on $[0, 10]$. D_g is Gaussian $\mathcal{N}(0, 2)$. D_{gm} is a Gaussian mixture with 3 components: $\mathcal{N}(-4, 4)$ (weight 0.4), $\mathcal{N}(0, 1)$ (weight 0.45), and $\mathcal{N}(5, 2)$ (weight 0.15).

Distribution		$q = 2$	$q = 3$	$q = 4$
D_u	SC	$(0.25, 0.75)$	$(0.16, 0.5, 0.84)$	$(0.12, 0.37, 0.63, 0.88)$
	ML	$(0.49, 0.50)$	$(0.33, 0.35, 0.98)$	$(0.25, 0.26, 0.74, 0.75)$
D_g	SC	$(0.25, 0.75)$	$(0.15, 0.5, 0.85)$	$(0.1, 0.35, 0.65, 0.9)$
	ML	$(0.49, 0.50)$	$(0.33, 0.35, 0.9)$	$(0.25, 0.26, 0.74, 0.75)$
D_{gm}	SC	$(0.17, 0.68)$	$(0.16, 0.59, 0.93)$	$(0.12, 0.37, 0.68, 0.94)$
	ML	$(0.49, 0.50)$	$(0.14, 0.65, 0.66)$	$(0.17, 0.34, 0.73, 0.74)$

Table 3.1: Optimal quantiles for different distributions, objectives, and numbers of facilities.



(a) Social Cost (b) Maximum Load

Figure 3.4: Comparison of optimized quantile mechanism and optimal value ($q = 3$).

placement is not realizable with any strategy-proof mechanism. Despite this, the optimized quantile mechanisms perform nearly as well, in expectation, as optimal placement in all three cases. Contrast this with the performance of the mechanisms with provable approximation ratios. When $q = 2$, the $(0, 1)$-quantile mechanism has an average social cost of 242.4, 340.9 and 523.2 for D_u, D_g and D_{gm}, respectively; but the social cost of our mechanisms are only 123.7, 76.5, and 165.1, respectively. When $q = 3$, the $(0.25, 0.5, 0.75)$-quantile mechanism has the best approximation ratio for ML (see Proposition 3.1). Its average maximum loads are 39.5, 38.7 and 38.3, which are close to (but not as good as) the loads of the optimized quantile mechanisms (36.5, 36.5, and 36.2).

We also compare the performance (with respect to social cost) of our optimized quantile mechanism with Bayesian optimization (see column **1D** in Table 3.2). Bayesian optimization performs almost as well as the optimal quantile mechanism when the number of agents is large.

Distr.	**1D**						**2D**		**4D**	
	$n = 101$			$n = 21$			$n = 101$	21	$n = 101$	21
	$q = 2$	3	4	$q = 2$	3	4	$q = 3$		$q = 2$	
D_u	2.2	3.0	3.8	9.7	18.5	24.6	1.4	7.4	1.0	6.2
D_g	1.4	2.3	3.1	11.6	19.7	27.9	1.5	5.4	0.9	2.9
D_{gm}	2.2	1.7	3.8	8.2	11.9	21	1.2	6.2	0.9	3.6

Table 3.2: Percentage improvement in social cost of optimized quantile mechanism vs. Bayesian optimization.

However, for the smaller population, eliciting ideal locations using the quantile mechanism gives much better results than the Bayesian approach. For example, when $q = 4$, the optimized quantile mechanism has an expected cost that is 3.1% better than the Bayesian model with $n = 101$ agents; but the performance gaps grows to 27.9% with $n = 21$ agents. In addition, we see that the agent-facility ratio also matters (i.e., when there are more facilities, the quantile mechanism tends to exhibit a greater performance gap).

These results are not surprising in this i.i.d. setting: indeed simple law-of-large-numbers arguments suggest that no elicitation of ideal points is needed at all for optimal placement given a sufficiently large population.[5] However, our framework does not require this i.i.d. assumption—preferences can be arbitrarily correlated. In such a case, Bayesian optimization can work extremely poorly. For example, consider a 1-D, 2-facility problem in which a latent variable V correlates preferences: if V is true, ideal points are drawn from a Gaussian $\mathcal{N}(\mu_1, \sigma)$; otherwise, they are drawn from $\mathcal{N}(\mu_2, \sigma)$. If each realization of V is equally likely, optimal Bayesian placement selects facilities at each of μ_1 and μ_2. By contrast, the optimal quantile mechanism is a simple function of σ, and will place facilities around the mean of the single "true" Gaussian, greatly improving social cost.

[5]Thanks to Lirong Xia for this observation.

3.5.2 Multi-dimensional mechanisms

We also experimented with two additional problems. **2D** is a two-dimensional, three-facility location problem where agents have L_2 preferences, capturing, say, the placement of three public projects like libraries, or warehouses. **4D** is a four-dimensional, two-facility location problem with L_1 preferences, which might model the selection of 2 products for launch, each with four attributes that predict consumer demand.[6]

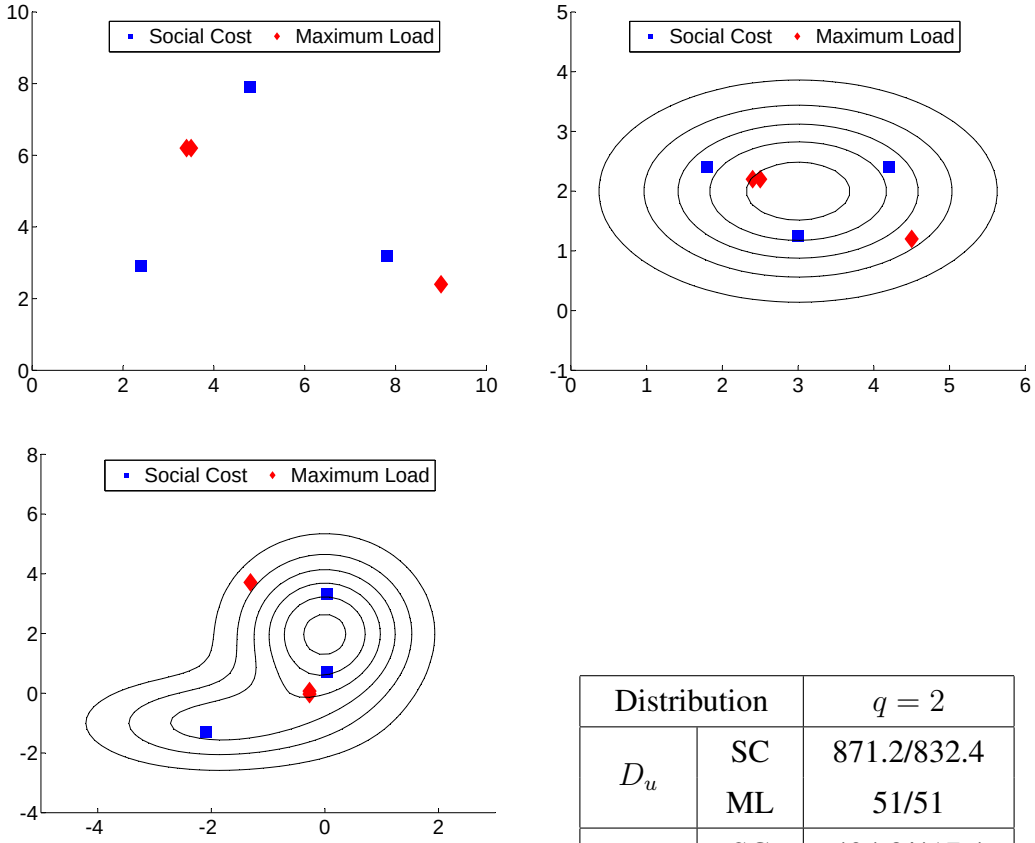
For the problem **2D** we show the *expected placement* of facilities given the selected quantiles in Figure 3.5(a)-(c), for both SC and ML, for each of the three distributions. (*Actual* facility placement will shift to match the reported type profile in each instance.) Placement for SC tends to be distributed appropriately, while ML places two facilities adjacent to one another. For **4D**, we measure performance rather than visualizing locations. Figure 3.5(d) compares expected SC and ML of our optimized quantile mechanisms to those using true optimal facility placements: the quantile mechanisms are always optimal for ML;[7] and for SC, placements using our optimized strategy-proof mechanisms are only 1.77%-4.66% worse than the corresponding non-strategy-proof optimal placements. This strongly suggests that quantile mechanisms, optimized using priors over preferences, are well-suited to multi-dimensional, single-peaked domains. The improvement of optimized quantile mechanisms over Bayesian optimization (see columns **2D** and **4D** in Table 3.2) exhibits trends similar to those in the **1D** case.

3.6 Conclusion

In this chapter, we introduced a family of mechanisms, namely quantile mechanisms, for general multi-dimensional, multi-facility location problems. We showed that the quantile mecha-

[6]For **2D**, D_u is uniform over $[0, 10]$ in each dimension. D_g is normal with mean $\boldsymbol{\mu} = [3, 2]$ and covariance $\boldsymbol{\Sigma} = [2, 1]\mathbf{I}$. D_{gm} is a 2 component mixture: $\mathcal{N}([-2, -1], [2, 1]\mathbf{I})$ (weight 0.3) and $\mathcal{N}([0, 2], [1, 3]\mathbf{I})$ (weight 0.7). For **4D**, D_u is uniform over $[0, 10]$ in each dimension. D_g is $\mathcal{N}([3, 2, 1, 2], [2, 3, 4, 1]\mathbf{I})$. D_{gm} is a 2 component mixture: $\mathcal{N}([2, 1, 0, 1], [4, 6, 8, 5]\mathbf{I})$ (weight 0.4) and $\mathcal{N}([1, 2, 1, 0], [7, 4, 5, 8]\mathbf{I})$ (weight 0.6).

[7]This is because the mechanism locates two facilities at almost the same position, and achieves optimal maximum load. However, this is not always possible for three or more facilities.



Distribution		$q = 2$
D_u	SC	871.2/832.4
	ML	51/51
D_g	SC	424.8/417.4
	ML	51/51
D_{gm}	SC	696.5/681.5
	ML	51/51

Figure 3.5: Optimized quantiles for (a) **2D**: Uniform, (b) **2D**: Gaussian, (c) **2D**: Gaussian mixture, and (d) **4D**.

nisms are strategy-proof for any quantile matrix, and provided several bounds on the approximation ratio for both social cost and maximum cost. We also developed a sample-based framework (and corresponding algorithms) for optimizing the selection of quantiles, when some prior distribution over agent preferences are known. Empirical results show that while the worst-case approximation ratios appear discouraging, the optimized quantile mechanisms work very well, and performs extremely close the optimum attainable even with precise knowledge of agent preferences.

The quantile mechanisms are just a starting point for the design of optimized mechanisms

for single-peaked domains, and can be extended in several ways. First, further development of optimization methods for quantile mechanisms (e.g., our MIP or MIQCP formulations) are needed to make our approach more scalable. Sample complexity results—theoretical bounds on the number of sampled profiles needed by our technique to ensure near-optimal results with high probability—are also of interest. Finally, incremental (or multi-stage) mechanisms that trade off social cost, communication costs, and agent privacy (as will be discussed in Chapter 7) would be extremely valuable. We will discuss potential future research directions in detail in Chapter 8.

Appendix of Chapter 3

MIP Formulation of Quantile Optimization (for Social Cost Only)

In this section, we describe our formulation of mixed integer linear program for social cost minimization in the quantile mechanism. Given W sampled preference profiles, our objective is to select the quantiles to minimize the average social cost relative to the W sampled profiles.

Let n be the number of agents, and m be the number of dimensions. Recall that the quantile mechanism locates the facilities facility on each dimension independently, so for each sampled profile, the total number of possible locations for a facility is n^m. If we use j to index all these possible locations, and define a_j as an indicator variable whose value is 1 iff a facility is placed

in the jth location, then we can formulate the optimization problem as follows:

$$\min_{a_j} \sum_{w=1}^{W} \sum_{i=1}^{n} \left(d_{ij}^{w} \cdot Y_{ij}^{w} \right) \tag{3.1}$$

$$s.t. \quad Y_{ij}^{w} \leq a_j, \qquad \forall i \leq n, \forall j \leq n^m, \forall w \leq W \tag{3.2}$$

$$\sum_{j=1}^{n^m} Y_{ij}^{w} = 1, \qquad \forall i \leq n, \forall w \leq W \tag{3.3}$$

$$\sum_{j=1}^{n^m} a_j = q \tag{3.4}$$

$$Y_{ij}^{w} \in \{0, 1\}, \qquad \forall i \leq n, \forall j \leq n^m, \forall w \leq W \tag{3.5}$$

$$a_j \in \{0, 1\}, \qquad \forall j \leq n^m \tag{3.6}$$

The d_{ij}^{w} is the pre-computed distance between agent i and facility j in sample w, Y_{ij}^{w} is an indicator variable whose value is 1 iff agent i is "assigned" to use facility j in sample w, and the objective function minimizes the sum of social cost over all W sampled profiles. Constraints (2) says that no agent can be assigned to facility j unless it is selected, and constraints (3) says that each agent can only be assigned to exactly 1 facility. Finally, constraints (4) requires that a total number of q facilities have to be selected. Note that the indicator variables Y_{ij}^{w} is binary technically, however, they can be relaxed due to the following reason: when the objective is a minimization problem, each agent will be "assigned" to the facility with least cost automatically, and allowing for fractional assignment will not change the objective value.

In the relaxed formulation, the number of continuous variables is $O(Wn^{m+1})$ and the number of binary variables is $O(n^m)$. This preliminary formulation can be used to solve small problems, but fails to scale very well. Please see the next section for two heuristic algorithms for solving large problems.

Gradient and Coordinate Descent Algorithms

In this section, we propose two heuristic algorithms for the sample-based optimization: gradient and coordinate descent. Both algorithms applies to social cost and maximum load minimization, and performs extremely well compared with the simple exhaustive method.

Recall that the objective is to choose a quantile matrix $\mathbf{P}$ to locate q facilities in some m-dimensional spaces, so the total number of possible quantile matrices is $(n^m)^q$. Let $\mathbf{t}^w$ be the type profile in sample w, and $\mathbf{x}^w = f_{\mathbf{P}}(\mathbf{t}^w)$ be the chosen location vector under the $\mathbf{P}$-quantile mechanism in sample w, we first define the *neighbourhood* as follows:

Definition 3.3 (Neighbourhood) *A quantile matrix $\mathbf{P}'$ is said to be a neighbour of $\mathbf{P}$ if there exist a single facility j^* and a single dimension k^* such that $\lfloor (n-1)p_j^k \rfloor - \lfloor (n-1)p_j'^k \rfloor \in \{-1, 1\}$ if $j = j^*$ and $k = k^*$, and $p_j^k = p_j'^k$ otherwise.*

In other words, if we consider the induced location vector by a quantile matrix, we can only move the location of a single facility on a single dimension to the (ordered) previous or next peak. Also note that we are optimizing the quantile matrix over all W sampled profiles, those moves over all W profiles have to be for the same facility, on the same dimension, and alone the same dirction.

Given a quantile matrix $\mathbf{P}$, we use $N(\mathbf{P})$ to denote the set of neighbours of $\mathbf{P}$. We also use $\mathbf{t}^w$ to denote the type profile in sample w. Then our gradient descent algorithm can be described as follows. The algorithm is presented for minimizing social cost, but can be easily generalized for maximum load minimization with minor changes.

The algorithm starts from a random quantile matrix, and compute the corresponding social cost over all W sampled profiles. Among the neighbours of the current quantile matrix, it chooses the one that improves (decreases) the social cost most, and update the quantile matrix. This process is repeated until a local minimum has been reached. We also use the random restart strategy, i.e., repeat Algorithm 1 for 1000 times, each with a randomly chosen quantile matrix, and keep the one that induces smallest social cost.

Algorithm 1 The Gradient Descent Algorithm for Social Cost Minimization

$\mathbf{P} \leftarrow$ A random quantile matrix, and $SC = \sum_{w=1}^{W} SC(f_{\mathbf{P}}(\mathbf{t}^w), \mathbf{t}^w)$
While True **do**
  $\mathbf{P}^* = \arg\min_{\mathbf{P}' \in N(\mathbf{P})} \sum_w SC(f_{\mathbf{P}'}(\mathbf{t}^w), \mathbf{t}^w)$ and $SC^* = \sum_w SC(f_{\mathbf{P}^*}(\mathbf{t}^w), \mathbf{t}^w)$
  if $SC^* \geq SC$ **then**
    break
  else
    $\mathbf{P} \leftarrow \mathbf{P}^*$, and $SC = SC^*$
return $\mathbf{P}$

Algorithm 2 The Coordinate Descent Algorithm for Social Cost Minimization

$\mathbf{P} \leftarrow$ A random quantile matrix, $SC = \sum_{w=1}^{W} SC(f_{\mathbf{P}}(\mathbf{t}^w), \mathbf{t}^w)$
$j^* = 0$ and $k^* = 0$
While True **do**
  for j from 1 to q **do**
    for k from 1 to m **do**
      if $j = j^*$ and $k = k^*$ **do**
        break
      $p_j^{'k} = \arg\min_{0 \leq p \leq 1} \sum_w SC(f_{\mathbf{P}'}(\mathbf{t}^w), \mathbf{t}^w)$, and $SC' = \sum_w SC(f_{\mathbf{P}'}(\mathbf{t}^w), \mathbf{t}^w)$, where
      $\mathbf{P}'$ is the quantile matrix achieved by replacing p_j^k in $\mathbf{P}$ with $p_j^{'k}$
      if $SC' \geq SC$ **then**
        $j^* = j$ and $k^* = k$
      else
        $\mathbf{P} \leftarrow \mathbf{P}'$, and $SC = SC'$
return $\mathbf{P}$

Note that at each step, the gradient descent algorithm has to compute the best neighbour among all possible neighbour. For any quantile matrix $\mathbf{P}$, the total number of neighbours is $O(q2^m)$, which may make the algorithm intractable when the number of dimension is large. We also propose a coordinate descent algorithm. The intuition of the algorithm is that at each step, we fix all but one quantile, and find the optimal value for that quantile. The process is repeated until no improvement for any quantile can be made. The algorithm is given in Algorithm 2. Similarly, we also use random restart techniques, repeating the algorithm for 1000 time and returning the quantile matrix with the minimum social cost over all W sampled profiles.

Chapter 4

Group Manipulation: Incentives

4.1 Introduction

In the previous chapter, we introduced the family of quantile mechanism. While quantile mechanisms are individual strategy-proof, they fail to guarantee group strategy-proofness in multi-dimensional spaces.

In this chapter, we address mechanism design in the multi-dimensional case when multiple facilities can be chosen, addressing both *unconstrained FLPs*—in which facilities can be placed at any point in some (metric) space—and *constrained FLPs*— in which some outcomes in the preference space are not feasible (i.e., the outcome space is constrained). In particular, we consider cases in which strategy-proofness cannot be achieved, and analyze *approximately strategy-proof mechanisms*. If one can bound the potential gain an agent (or group) can obtain by misreporting their preferences, the cost of determining an optimal misreport may outweigh the benefits, rendering such mechanisms "practically strategy-proof" [Hyafil and Boutilier, 2007, Lu et al., 2012].

In unconstrained problems, individual strategy-proofness can be achieved using *generalized median mechanisms, or GMMs* [Moulin, 1980, Barberà et al., 1993] for single-FLPs, and *quantile mechanisms (QMs)* for multi-FLPs (as we have shown in Chapter 3), but group

strategy-proofness is unachievable in general. Our first contribution is to provide an impossibility result showing that the incentive for any group of agents to misreport is unbounded for arbitrary preference profiles. Then we give a profile-specific bound on the incentive to misreport. Second, we analyze constrained FLPs, defining a new family of *closest candidate mechanisms (CCMs)*. CCMs use QMs to determine *tentative* locations, then project these to the nearest feasible locations using some distance function. While CCMs are not strategy-proof in general, we are able to bound the incentive for individuals and groups to misreport. Finally, we empirically evaluate the performance of our mechanisms using real-world preference data in electoral and geographic facility domains. We evaluate the probability of agents (or groups) successfully manipulating outcomes, and more importantly show that their expected gain and impact on social welfare is quite small in practice. This suggests that the mechanisms analyzed here, namely, GMMs, QMs and CCMs, may be "sufficiently strategy-proof" for practical purposes.

4.2 Unconstrained Facility Location

We follow the notation introduced in Section 2.3. Let n be number of agents, q be the number of facilities to be located, and m be the number of dimensions. The objective is to select q homogeneous facilities in some m-dimensional space $\mathbb{R}^m$. Such an outcome is represented by a location vector $\mathbf{x} = (x_1, x_2, \ldots, x_q)$, where $x_j \in \mathbb{R}^m$. Each agent has a type $t_i \in T_i$ determining her cost associated with any location vector $\mathbf{x}$, i.e., $c_i(\mathbf{x}, t_i) = \min_{j \leq q} c_i(x_j, t_i)$, in which each agent uses the facility with least cost.

As discussed in Chapter 3, when agents have single-peaked preferences (e.g., c_i equals the L_1 or L_2 cost), the median mechanism and its generalization [Black, 1948, Moulin, 1980] guarantee individual strategy-proofness for unconstrained facility location problems. These mechanisms have also been generalized to multi-dimensional spaces [Barberà et al., 1993], i.e., the multi-dimensional generalized median, and for multi-facility location, i.e., the quantile

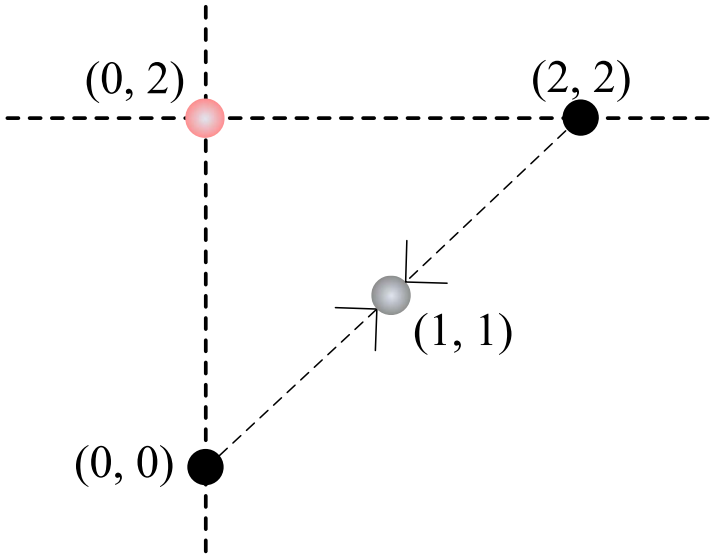


Figure 4.1: A two-dimensional counter example showing that quantile mechanisms are not group strategy-proof.

mechanisms we introduced in Chapter 3.

While the GMMs and the QMs are group strategy-proof in single-dimensional spaces, they fail to guarantee group strategy-proofness in multiple dimensions. Consider a two-dimensional, single facility location problem as shown in the Example 4.1 below. These two agents can collude to misreport their preferences such that agent 1 can benefit from the misreport of agent 2 in one dimension, while agent 2 can benefit from the misreport of agent 1 in the other dimension, making both better off.

Example 4.1 *Consider a two-dimensional, single facility location problem for 2 agents. The quantile mechanism* $\{0; 1\}$ *is used, i.e., locate the facility at the intersection position of the left-most coordinate on the first dimension and the rightmost coordinate on the second dimension. If the agent peaks are* $(0, 0)$ *and* $(2, 2)$*, then the mechanism will locate the facility at* $(0, 2)$ *and the cost of both agents is 2 under the* L_2*-norm. However, if both agents misreport* $(1, 1)$*, then the facility will be located at* $(1, 1)$*, in which case both agents have a cost of* $\sqrt{2} \approx 1.414 < 2$*. This is shown in Figure 4.1.*

In fact, the characterization results of Barberà et al. [1993] indicate that there is no (anony-

mous) group strategy-proof mechanism in such settings.

Remark 4.1 (Non-existence of group strategy-proof mechanism) *[Barberà et al., 1993] There is no (anonymous) group strategy-proof mechanisms for multi-dimensional, unconstrained facility location problems.*[1]

Alternatively, one can try to bound the incentive for any group of agents to misreport, showing GMMs and QMs to be *approximately group strategy-proof*. Here, approximately group strategy-proofness means that a mechanism is not group strategy-proof, but the incentive for any group of agents to misreport is bounded.

Definition 4.1 (Approximately group strategy-proof) *A mechanism f is* ε-group strategy-proof *if, for any $S \subseteq N$, any group misreport $\mathbf{t}'_S$ and any type profile of other agents $\mathbf{t}_{-S}$, there is some $i \in S$ such that:*

$$c_i(f(\mathbf{t}_S, \mathbf{t}_{-S}), t_i) \geq c_i(f(\mathbf{t}'_S, \mathbf{t}_{-S}), t_i) + \varepsilon$$

where $\mathbf{t}_S$ and $\mathbf{t}_{-S}$ are the type profile of all agents in S and $N \setminus S$, respectively.

In other words, if a group of agents form a coalition and misreport their peaks, there must by one of them whose gain is less than or equal to ε. If the value of ε is small enough, then considering the cost of finding a good lie (e.g., information cost, computational cost, etc), such a mechanism is "practically group strategy-proof" (note that if $\varepsilon = 0$, this definition reduces to group strategy-proofness in Definition 2.28). This definition requires that *each* agent in a manipulating coalition S has some gain by participating, which is sensible in settings with non-transferable utility (as is the case in many social choice problems).

Let f be any mechanism, S be a coalition with (fixed) true type profile $\mathbf{t}_S$, and $\mathbf{t}_{-S}$ be the (fixed) reports of the other agents. We define the *gain* of $i \in S$ for a (coalitional) misreport $\mathbf{t}'_S$ to be $G(i, S, \mathbf{t}'_S) = c_i(f(\mathbf{t}_S, \mathbf{t}_{-S}), t_i) - c_i(f(\mathbf{t}'_S, \mathbf{t}_{-S}), t_i)$; the *maximum gain* of i to be

[1] Anonymity is critical, as dictatorial mechanisms belong to the class of GMMs and offers group strategy-proofness.

$G(i,S) = \max_{\mathbf{t}'_S} G(i,S,\mathbf{t}'_S)$; and *incentive for S to misreport* to be $G(S) = \max_{i \in S} G(i,S)$. We say a misreport $\mathbf{t}'_S$ is *viable* iff $G(i,S,\mathbf{t}'_S) \geq 0$ for each $i \in S$ and $G(i,S,\mathbf{t}'_S) > 0$ for some $i \in S$.

To quantify the incentive for a group of agents to misreport, one must make assumptions about agent cost functions. Here we assume that cost is equal to the L_2 distance from the ideal point, i.e., $c_i(x_j,t_i) = ||x_j - t_i||_2$.[2] We first give an impossibility result, showing that the incentive for group manipulation can be arbitrarily large.

Theorem 4.1 (Unbounded group strategy-proofness) *GMMs and QMs are not ε-group strategy-proof for any fixed $\varepsilon > 0$ under the L_2-norm.*

Proof: We give a counter-example for two-dimensional, two-facility location under QMs. The result applies directly to GMMs since QMs are a specific instance of GMMs.

Let $\mathbf{p} = (p^1, p^2)$ be a two-dimensional, quantile matrix used for a QM, in which the facility is located at the coordinate of the p^1th peak in the first dimension, and at the coordinate of the p^2th peak in the second dimension. Consider the following two cases:

I. $p^1 + p^2 \leq 1$.

Consider the following profile $\mathbf{t} = (\underbrace{(a,0),\ldots,(a,0)}_{np^2 \text{ copies}}, \underbrace{(0,a),\ldots,(0,a)}_{np^1 \text{ copies}}, \underbrace{(a,a),\ldots,(a,a)}_{n(1-p^1-p^2) \text{ copies}})$, where $a > 0$ is a positive real number (as shown in Figure 4.2). The QM will locate the facility at position $(0,0)$, and the costs of the agents are: a, for those at $(a,0)$ and $(0,a)$; and $\sqrt{2}a$, for those at (a,a). However, if all n agents are manipulators, then there exists a viable misreport in which all agents report $(a/2, a/2)$, which will then be selected. The cost under this misreport is $\sqrt{2}/2a$ for each agents, and the gains are: $a - \sqrt{2}/2a \approx 0.293a$, for those at $(a,0)$ and $(0,a)$; and $\sqrt{2}/2 \approx 0.707a$, for those at (a,a). As a can be arbitrarily large, so are the gains due to manipulation.

[2] Barberà et al.'s [1993] characterizations do not preclude the existence of group strategy-proof mechanisms when specific cost functions are used (e.g., L_2-norm). However, it is still meaningful to study the group manipulation of GMMs and QMs due to their simplicity and intuitive nature, their (individual) strateg-yproofness, and flexibility (e.g., the fact that they can be optimized or tuned for specific prior distributions over preferences). Similar remarks apply to the negative results of Barberà et al. [1997] for constrained FLPs.

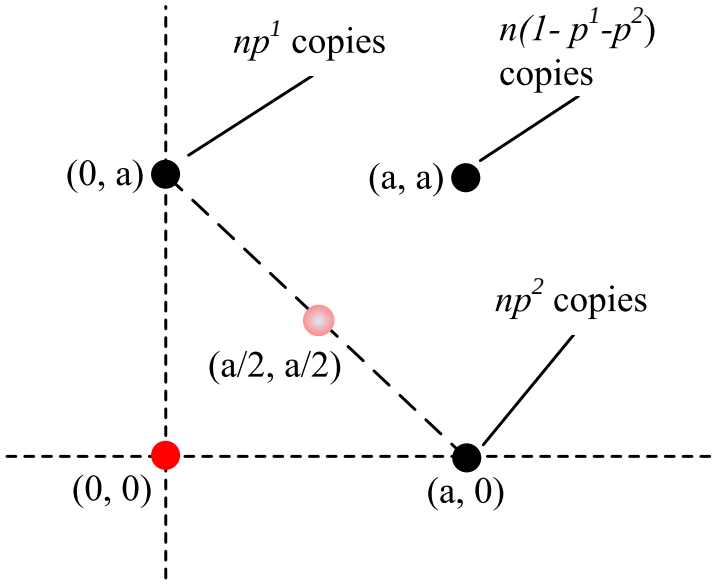


Figure 4.2: A two-dimensional counter example showing the incentive for a group of agents to misreport can be unbounded.

II. $p^1 + p^2 > 1$.

Consider the following profile $\mathbf{t} = (\underbrace{(a,0),\ldots,(a,0)}_{n(1-p^1)\ copies}, \underbrace{(0,a),\ldots,(0,a)}_{n(1-p^2)\ copies}, \underbrace{(0,0),\ldots,(0,0)}_{n(p^1+p^2-1)\ copies})$.

The QM will locate the facility at (a, a), and the agents costs are: a, for those at $(a, 0)$ and $(0, a)$; and $\sqrt{2}a$, for those at $(0, 0)$. As above, a viable manipulation exists in which each manipulator misreports $(a/2, a/2)$, and again the gain of the manipulators is arbitrarily large as $a \to \infty$.

This demonstrates that the (additive) incentive for manipulation is unbounded for GMMs and QMs. ■

We note that while the gain is unbounded in an additive sense, the relative gain is also unbounded (if all the $n(1-p^1-p^2)$ agents are at location $(a/2, a/2)$ in Figure 4.2). While this observation serves as a negative result, it is an *a priori* worst-case analysis, allowing arbitrary preference profiles. In practice, the incentive for a group of agents to misreport depends on the actual ideal points of the sincere agents and the manipulators, i.e., $\mathbf{t}_S$ and $\mathbf{t}_{-S}$. In the remainder of this section, we assume that the true peaks of both sincere agents and manipulators are known, and provide a profile-specific, *a posteriori* bound that relies on this knowledge. Such

a bound represents the maximum gain that the set of manipulators can realize by conducting a joint misreport. In cases where only some but not all of the preferences (peaks) are known, bounds can be developed using similar analysis. Moreover, if some prior distribution of agent peaks is known in advance, incentives for manipulation can be evaluated empirically using the sample-based optimization framework in Chapter 3.

We begin with single-facility case, and providing an upper bound on the incentive for a group of agents to misreport.

Definition 4.2 (Pareto optimal misreport) *Let $S \subseteq N$ be a set of manipulators. A misreport $\mathbf{t}'_S$ is* Pareto optimal *if there is no other misreport $\mathbf{t}''_S$ such that $c_i(f(\mathbf{t}''_S, \mathbf{t}_{-S})) \leq c_i(f(\mathbf{t}'_S, \mathbf{t}_{-S}))$ for all $i \in S$ and for some $i^* \in S$, we have $c_{i^*}(f(\mathbf{t}''_S, \mathbf{t}_{-S})) < c_{i^*}(f(\mathbf{t}'_S, \mathbf{t}_{-S}))$.*

Intuitively, a misreport $\mathbf{t}'_S$ is Pareto optimal if there is no other misreport in which no manipulator is worse off than they were in $\mathbf{t}'_S$ and at least one is strictly better off than she is in $\mathbf{t}'_S$.

When bounding the incentive for a group of agents to misreport, we can focus on Pareto optimal misreports without loss of generality (since making a Pareto improvement to some non-Pareto optimal misreport will improve the lot of the manipulators and can only increase the upper bound on this incentive). The following lemma provides a necessary condition for a misreport to be Pareto optimal.

Lemma 4.1 *Let $S \subseteq N$ be a set of manipulators, and $x = f(\mathbf{t}_S, \mathbf{t}_{-S})$ be the chosen location under truthful reports. We use superscript k to index dimensions, and define $I^k = [\min_{i \in S} t_i^k, \max_{i \in S} t_i^k]$ to be the tightest the bounding interval containing all manipulator peaks in the kth dimension. Also let $\mathbf{t}'_S$ be a Pareto optimal misreport and $x' = f(\mathbf{t}'_S, \mathbf{t}_{-S})$ be the location chosen under $\mathbf{t}'_S$. Then we have $x'^k \in I^k$ if $x^k \in I^k$ and $x'^k = x^k$ otherwise.*

Proof: Suppose the lemma does not hold. Then for each dimension k, one of the following two situations must arise:

I. $x'^k \notin I^k$ and $x^k \in I^k$. Note $x'^k \notin I^k$ means either $x'^k < \min_{i \in S} t_i^k$ or $x'^k > \max_{i \in S} t_i^k$, and w.l.o.g., we assume it is the former case. Recall that we have $f^k(\mathbf{t}'_s, \mathbf{t}_{-S}) = x'^k < x^k = f^k(\mathbf{t}_S, \mathbf{t}_{-S})$, which means there must be some manipulator whose misreport lies to the left of (is less than) $\min_{i \in S} t_i^k$ in the kth dimension. We can construct another misreport $\mathbf{t}''_S$ such that $f^k(\mathbf{t}''_S, \mathbf{t}_{-S}) = \min_{i \in S} t_i^k$ and $f^{\tilde{k}}(\mathbf{t}''_S, \mathbf{t}_{-S}) = x^{\tilde{k}}, \forall \tilde{k} \neq k$, and each manipulator i strictly gains in the kth dimension without losing in any other dimension. This means $c_i(f(\mathbf{t}''_S, \mathbf{t}_{-S})) < c_i(f(\mathbf{t}'_S, \mathbf{t}_{-S}))$, which contradicts our assumption that $\mathbf{t}'_S$ is a Pareto optimal misreport.

II. $x'^k \neq x^k$ and $x^k \notin I^k$. Similarly $x^k \notin I^k$ means either $x^k < \min_{i \in S} t_i^k$ or $x^k > \max_i t_i^k$, and w.l.o.g., we assume it is the former case. Since QMs locate the facility at a specified quantile, we must have $x'^k < x^k$. We can construct another misreport $\mathbf{t}''_S$ such that $f^k(\mathbf{t}''_S, \mathbf{t}_{-S}) = x^k$ and $f^{\tilde{k}}(\mathbf{t}''_S, \mathbf{t}_{-S}) = x^{\tilde{k}}, \forall \tilde{k} \neq k$, and each manipulator i strictly gains in the kth dimension without losing in any other dimension. This too contradicts the Pareto optimality of $\mathbf{t}'_S$.

■

Lemma 4.1 shows that, when bounding the incentive to misreport, we can focus our attention on those dimensions in which the coordinate of the facility selected under truthful reporting lies within the corresponding bounding interval of the manipulator peaks—for those dimensions where this is not true, the manipulators can safely leave their reports on those dimensions unchanged (i.e., report sincerely).

Before describing our bound, we first introduce some notation. Let $S \subseteq N$ be a set of manipulators and x be the chosen location under truthful reporting. We define $C(i, x) = \{\bar{x} \in \mathbb{R}^m : ||\bar{x} - t_i||_2 \leq ||t_i - x||_2\}$ to be the circle centered at t_i with radius $||t_i - x||_2$. Let $C(S) = \cap_{i \in S} C(i, x)$ denote the intersection of these circles. Let I^k be the bounding interval as defined in Lemma 4.1, and $C^{\perp}(S) = \{\bar{x} \in \mathbb{R}^m : \bar{x}^k \in C^k(S)$ *if* $x^k \in I^k$ *and* $\bar{x}^k = x^k$ *otherwise*$\}$ be the projection of $C(S)$ onto the subspace of $\mathbb{R}^m$ in which we fix the coordinates of x in those dimensions k not contained in the bounding intervals to x^k. We have the following theorem:

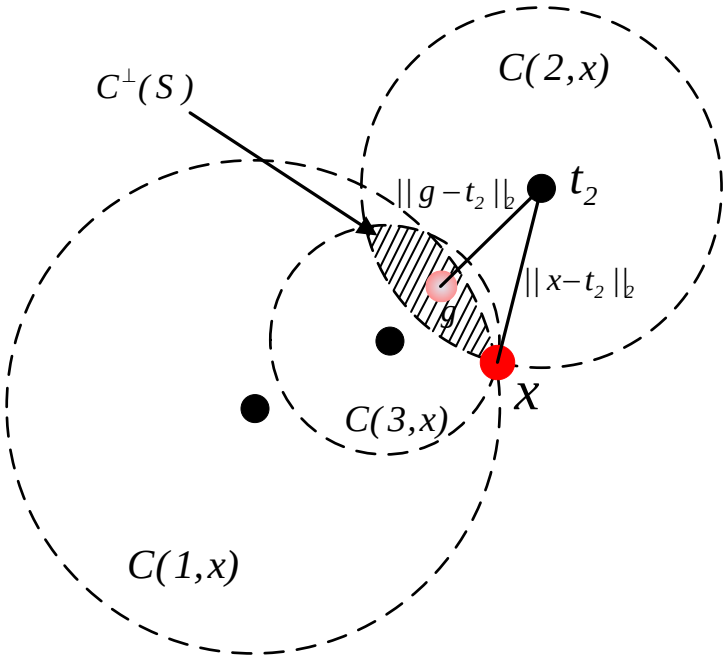


Figure 4.3: An two-dimensional example showing that a viable misreport must induce a location contained in $C^{\perp}(S)$ (the shaded area).

Theorem 4.2 (Incentive under single-FLPs) *Let $S \subseteq N$ be a set of manipulators, and $x = f(\mathbf{t}_S, \mathbf{t}_{-S})$ be the chosen location under truthful reports. For single-facility location problems under GMMs/QMs, the incentive for manipulators in S to misreport is at most:*

$$\varepsilon_S = \max_{g \in C^{\perp}(S)} \left[\max_{i \in S} \left(||x - t_i||_2 - ||g - t_i||_2 \right) \right].$$

Proof: Let $\mathbf{t}'_S$ be any group misreport and $g = f(\mathbf{t}'_S, \mathbf{t}_{-S})$ be the induced location of the facility. The first thing to note is that for the misreport $\mathbf{t}'_S$ to be viable, the induced location g must be contained in $C(S)$, otherwise there will be some manipulator who is strictly worse-off (as shown in Figure 4.3).

By Lemma 4.1, we need only consider the projection of $C(S)$ onto the subspace $C^{\perp}(S)$ (as defined above). For each location g, the gain of manipulator i is $||x - t_i||_2 - ||g - t_i||_2$. If we take the maximum over all manipulator, and all possible locations g, we obtain the stated bound. ∎

In the multi-facility case, we provide an upper bound on the incentive to misreport by

considering each facility independently. Formally, let $S \subseteq N$ be a set of manipulators, and $\mathbf{x} = f(\mathbf{t}_S, \mathbf{t}_{-S})$ be the chosen location vector under truthful reporting. For each facility j with location $x_j \in \mathbf{x}$, we define $S_j = \{i \in S : j = \arg\min_{j' \leq q} ||x_{j'} - t_i||_2\}$ as the set of manipulators whose closest facility is j under truthful report. We also define $C(i, x_j)$ and $C(S_j)$ similarly as in the single-facility case, and $D_j = \{i \in S_j : \exists j' \textit{ s.t. } C(i, x_j) \cap C(S_{j'}) \neq \emptyset\}$ as the set of manipulators in S_j whose circles intersect with $C(S_{j'})$ for some other facility j'. Intuitively, D_j denotes the set of manipulators in S_j who may deviate from using facility j to use other facilities. Then we have:

Theorem 4.3 (Incentive under multi-FLPs) *Let $S \subseteq N$ be a set of manipulators, and $x = f(\mathbf{t}_S, \mathbf{t}_{-S})$ be the chosen location under truthful reports. For multi-facility location problems under GMMs/QMs, the incentive for a set of manipulators S to misreport is at most $\varepsilon_S = \max_j \varepsilon_{S_j}$, where:*

$$\varepsilon_{S_j} = \max_{g \in C^{\perp}(S_j \backslash D_j)} \left[\max_i \left(c_i(\mathbf{x}, t_i) - ||g - t_i||_2 \right) \right].$$

Proof: Let $\mathbf{t}'_S$ be any group misreport and $g = f_j(\mathbf{t}'_S, \mathbf{t}_{-S})$ be the induced location of facility j. Among the manipulators whose closest facility is j, we have to preclude those who may deviate from using j to other facilities, which we denote by D_j.

By Lemma 4.1, we can focus our attention on the projection of $C(S_j \backslash D_j)$ onto subspace in which the coordinates of x_j are not contained in the bounding intervals. If we take the maximum over all manipulators, over all possible locations g for each facility, and over all facilities, we obtain the stated bound. ■

4.3 Constrained Facility Location

We now turn our attention to constrained facility location problems, in which facilities can only be placed at a restricted finite set of *feasible location* $C = \{c_1, \ldots, c_l\}$ (where $l > q$ is

the number of feasible locations). For instance, in political settings, agent i's ideal point may corresponds to a "fictitious" candidate who agrees with i on every issue, while selection is limited to those "actual" candidates who have agreed to stand for election. Similar restrictions often apply in voting (a finite set of candidates under consideration), product design (selecting from an existing assortment), and other forms of facility location problems.

Barberà et al. [1997] studied constrained FLPs and provided an important characterization result that a mechanism is strategy-proof in a constrained setting iff: a) it is a generalized median mechanism; and b) it satisfies the *intersection property*, a condition requiring that the decision rules operating on different dimensions must be coordinated to guarantee a feasible location. The intersection property can be satisfied when the set of feasible locations has some special shapes (e.g., rectangles), however, it will, as they say, "anticipate impossibility theorems in most applications" (for arbitrary shape of the set of feasible locations as we consider in this chapter).

So as above, we turn our attention to *approximately* strategy-proof mechanisms. The definition of approximately strategy-proofness can be reduced from Definition 4.1 in which $|S| = 1$. In other words, a mechanism is ε-strategy-proof if the gain for any agent, any misreport and any other reports is at most ε.

Focusing on QMs (since they apply to single- and multi-FLPs), we deal with constraints by defining *closest candidate mechanisms (CCMs)*, and assume L_2-distance for costs and "projection" to the nearest feasible location:

Mechanism 4.1 (Closest candidate mechanism (CCM)) *Let $C = \{c_1, \ldots, c_l\}$ be a set of feasible locations, and f' a (multi-dimensional) QM. A* closest candidate mechanism (CCM) *f, based on QM f', selects a location vector, given reports $\mathbf{t}$, as follows: (i) let $f'(\mathbf{t}) = \widetilde{\mathbf{x}} = \{\widetilde{x_1}, \ldots, \widetilde{x_q}\}$; (ii) return location vector $\mathbf{x} = \{x_1, \ldots, x_q\}$, where $x_j = \arg\min_{c \in C} ||c - \widetilde{x_j}||_2$.*

In other words, the mechanism runs a QM on the reported peaks and replaces any infeasible location $x'_j \notin C$ with the nearest feasible location in C. The following is an example of how

CCM works in 1D:

Example 4.2 *Let the feasible set of facilities be $C = \{2, 3, 7\}$, and agents' peak profile be $t = \{0, 1, 4, 7, 10\}$. Then a CCM with quantile vector $(0.25, 0.75)$ will locate the first facility at $x_1 = 2$ (as $2 = \arg\min_{c \in C} ||c - 1||_2$) and the second at $x_2 = 7$ (as 7 is in the feasible set).*

While not strategy-proof in general, CCMs are in fact (group) strategy-proof in 1D:

Theorem 4.4 (Group strategy-proofness of one-dimensional CCM) *CCMs are group strategy-proof in 1D for one-dimensional FLPs under L_2 cost.*

Proof: We first describe the proof assuming $q = 2$, and then show how the analysis can be generalized when $q > 2$.

Let $S \subseteq N$, and $\widetilde{\mathbf{x}} = \{\widetilde{x_1}, \widetilde{x_2}\}$ be the location vector chosen by the QM if all agents report truthfully, and $\mathbf{x} = \{x_1, x_2\}$ be the projected location vector into C. Let $\widetilde{\mathbf{x}}' = \{\widetilde{x_1}', \widetilde{x_2}'\}$ be the vector chosen by the QM if agents in S jointly misreport, and $\mathbf{x}' = \{x_1', x_2'\}$ be its projection. W.l.o.g., assume $x_1 < x_2$ and $x_1' < x_2'$. Consider four cases:

I. $x_1 \geq x_1'$ and $x_2 > x_2'$: Both x_2 and x_2' are feasible, so $\widetilde{x_2}' \leq (x_2' + x_2)/2 \leq \widetilde{x_2}$. Since QM chooses each location using quantiles, suppose some i, with peak $t_i > \widetilde{x_2}$, misreports to the left of $\widetilde{x_2}$. Then $i \in S$, and i's cost now is $c_i(\mathbf{x}', t_i) = t_i - x_2' > t_i - x_2 = c_i(\mathbf{x}, t_i)$, and is strictly worse off.

II. $x_1 < x_1'$ and $x_2 > x_2'$: As above, there must be some $i \in S$, with peak $t_i > \widetilde{x_2}$, who misreports to the left of $\widetilde{x_2}$. So i's cost now is $c_i(\mathbf{x}', t_i) = t_i - x_2' > t_i - x_2 = c_i(\mathbf{x}, t_i)$, and is strictly worse off.

III. $x_1 < x_1'$ and $x_2 \leq x_2'$: Symmetric to cases I and II.

IV. $x_1 \geq x_1'$ and $x_2 \leq x_2'$: There must some $i \in S$, with type $\widetilde{x_1} < t_i < \widetilde{x_2}$, who misreports to the left of $\widetilde{x_1}$ or to the right of $\widetilde{x_2}$. W.l.o.g., assume a misreport to the left of $\widetilde{x_1}$. Then i's cost is $c_i(\mathbf{x}', t_i) = \min\{t_i - x_1', x_2' - t_i\} \geq \{t_i - x_1, x_2 - t_2\} = c_i(\mathbf{x}, t_i)$, and is worse off.

This establishes group strategy-proofness for $q = 2$.

For the case of $q > 2$, we can define $\widetilde{\mathbf{x}}$, $\mathbf{x}$, $\widetilde{\mathbf{x}}'$ and $\mathbf{x}'$ similarly. Then by using case analysis as in the case of $q = 2$, we can always find an agent in S who is not strictly better off, which completes our proof. ■

One can show that CCMs in the multi-facility case are a straightforward extension of the family of *disturbed GMMs* [Massó and Moreno de Barreda, 2011] in the 1D setting, which characterize all strategy-proof mechanisms when agents have symmetric single-peaked preferences (of which L_1- and L_2-preferences are a special case). CCMs also satisfy the *intersection property* in 1D, a sufficient condition for a mechanism to be strategy-proof with constraints, hence it is consistent with Barberà et al.'s characterization result.

Evaluating incentives to misreport in multi-dimensional spaces is more involved. Our main results, Theorems 4.5 and 4.6 below, require two preliminary lemmas. The first addresses single-agent misreports. We begin with some some notation.

Definition 4.3 *For each feasible $c \in C$, we define its* electoral zone *to be $\mathbb{Z}_c = \{x \in \mathbb{R}^m, c = \arg\min_{c' \in C} ||c' - x||_2\}$.*

Definition 4.4 *Let $\mathbb{C}_c^k$ be the* potential deviation area *of feasible candidate c if a single manipulator changes her report in all but dimension k, i.e., $\mathbb{C}_c^k = \{x \in \mathbb{R}^m, x^k = x'^k$ for some k where $x' \in \mathbb{Z}_c\}$, and $\mathbb{C}_c = \cup_k \mathbb{C}_c^k$ be the union of potential deviation areas over all dimensions.*

Then we have the following result:

Lemma 4.2 *For any two feasible locations $c_1, c_2 \in C$, an agent i can gain from a misreport that changes the location of a facility from c_1 to c_2 only if $\mathbb{C}_1 \cap \mathbb{Z}_2 \neq \emptyset$.*

Proof: We provide a proof for the case of 2D first, and then show how it can be generalized to any number of dimensions. Consider two feasible locations c_1 and c_2 (see Figure 4.4). Let g be one of the chosen locations under a QM f', and c_1 be its projected feasible location under CCM f (note we must have $g \in \mathbb{Z}_1$, otherwise f will not project g to c_1). Suppose there exists

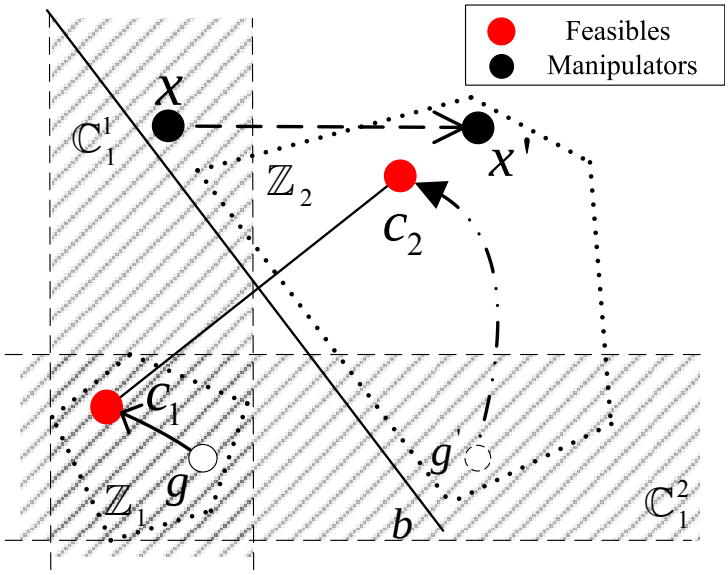


Figure 4.4: An example where a manipulator can benefit by changing the outcome from c_1 to c_2.

a location profile $\mathbf{t}$ in which an agent i (with true peak x) will use facility c_1, but has a positive incentive to change it to c_2. Then we can construct another profile $\mathbf{t}'$ such that if i misreports x', the selected location for g under the QM f' will be $g' \in \mathbb{C}_1$ (more specifically $\mathbb{C}_1^2$ in the figure). Since f projects g' to the closest feasible location, which is c_2 instead of c_1, agent i gains by misreporting. However, f will project g' to c_2 only if there is no other feasible location closer to g', i.e., only if g' is in the electoral zone of c_2. This implies $\mathbb{C}_1 \cap \mathbb{Z}_2 \neq \emptyset$.

For the case of $m > 2$, $\mathbb{C}_1 \cap \mathbb{Z}_2 \neq \emptyset$ implies that there is at least on dimension k such that $\mathbb{C}_1^k \cap \mathbb{Z}_2 \neq \emptyset$. Then we can construct a location profile if some agent misreports her ideal location in all but dimension k and move one of the chosen locations under the quantile mechanism to some point in the electoral zone of c_2. The CCM will project to the feasible location of c_2 instead of c_1 now, completing our proof. ■

This lemma ensures an agent can profitably change a facility only if she can move the corresponding quantile-location into the electoral zone of another feasible outcome. The next lemma bounds the gain an agent can realize by changing one of the CCM's outcomes from one feasible location to another. For each pair of feasible locations $c_1, c_2 \in C$, we define $\mathbf{K}_{1,2} = \{k : \mathbb{Z}_2 \cap \mathbb{C}_1^k \neq \emptyset\}$ as the set of dimensions that $\mathbb{Z}_2$ intersects with $\mathbb{C}_1$. For any two

points $x, y \in \mathbb{R}^m$, let $\mathbf{B}(x, y)$ be the minimum bounding box containing x and y. Then we have:

Lemma 4.3 *Let $c_1, c_2 \in C$. The maximum gain any agent can realize by replacing c_1 with c_2 in a CCM is:*

$$G(c_1, c_2) = \begin{cases} ||c_2 - c_1||_2 & \text{if } \exists x \in \mathbb{C}_1 \cap \mathbb{Z}_2 \text{ s.t. } \mathbf{B}(c_2, x) \cap \mathbb{Z}_1 \neq \emptyset \\ \max_{k' \in \mathbf{K}_{1,2}} \sqrt{\sum_{k \neq k'} |c_1^k - c_2^k|^2} & \text{otherwise.} \end{cases}$$

Proof: We prove the lemma for 2D case first, and show how the analysis can be generalized to higher dimensions.

For the feasible pair of outcomes $c_1, c_2 \in C$, we consider the following two cases:

I. $c_2 \in \mathbb{C}_1^k$ for some k, $\exists x \in \mathbb{C}_1 \cap \mathbb{Z}_2$ and $g \in \mathbf{B}(c_2, x) \cap \mathbb{Z}_1$ (as shown in Figure4.5 left). Consider the situation in which a manipulator's true peak coincides with c_2, which provides the maximum gain for a manipulation that induces location c_2. We can construct a location profile such that g is one of the quantile-location under truthful report before projection. As we have $g \in \mathbb{Z}_1$, the CCM will project it to c_1, and the manipulator cost is at most $||c_2 - c_1||_2$ (equality if c_1 is the closest facility under truthful report). However, the manipulator can misreport and change the quantile-location for g to x (as $g \in \mathbf{B}(c_2, x)$), inducing a projection to c_2 (as $x \in \mathbb{Z}_2$) and a cost of 0, so her gain is at most $||c_2 - c_1||$.

II. $c_2 \notin \mathbb{C}_1^k$ for any k. If $\mathbf{K}_{1,2} = \emptyset$, then by Lemma 4.2 we have $G(c_1, c_2) = 0$, otherwise the upper bound is demonstrated using the properties of a hyperbola. Given two focal points, the difference of the distances to these two foci from any point on a hyperbola is constant. Let a and b be the semi-major and semi-minor axes, and c the half distance between two foci satisfying $c^2 = a^2 + b^2$.

Let c_1 and c_2 be two focal points of a hyperbola (see Figure 4.5 right). Let the angle between line c_1c_2 and the horizontal axis be α, and the angle between the asymptotes and the semi-major axis be θ. Our goal is to bound the maximum value of $2a$, which is the

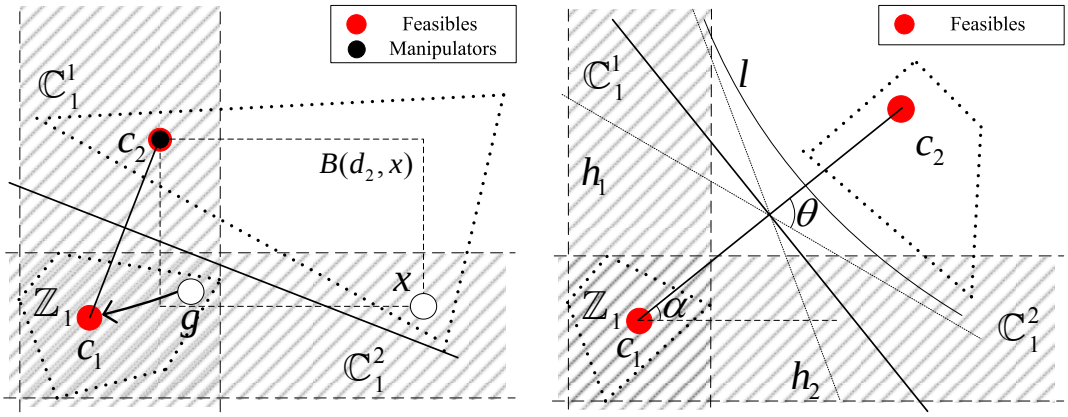


Figure 4.5: The incentive is bounded if some manipulator can benefit from changing the outcome from c_1 to c_2.

difference of distances to the two foci on a hyperbola, s.t. the constraint that hyperbola l intersects the horizontal or vertical axis (otherwise no agent can benefit this much). Suppose w.l.o.g., we have $\mathbb{Z}_2 \cap \mathbb{C}_1^2 \neq \emptyset$. The maximum gain is achieved when the angle $\theta > 90° - \alpha$ and the hyperbola intersects the shaded area $\mathbb{C}_1^1$. Recall that for an asymptote, we have $\tan(\theta) = b/a$, so we can formulate this as a maximization:

$$\begin{aligned} &\max 2a \\ &s.t.\ \left(c_2^1 - c_1^1\right)^2 + \left(c_2^2 - c_1^2\right)^2 = 4(a^2 + b^2) \\ &\quad \frac{b}{a} > \frac{|c_2^1 - c_1^1|}{|c_2^2 - c_1^2|} \end{aligned}$$

Solving the above maximization, we have $2a = |c_1^2 - c_2^2|$. And if we consider every dimension $k \in \mathbf{K}_{1,2}$, we can get the above bound.

When generalizing to higher dimensions, the analysis in case I still applies. For case II, we use two sheeted hyperboloid instead of hyperbola and conical surface instead of asymptotes,

where the above optimization becomes:

$$\max 2a$$
$$s.t. \sum_k (c_2^k - c_1^k)^2 = a^2 + b^2$$
$$\frac{b}{a} > \max_k \{ \frac{|c_2^k - c_1^k|}{\sqrt{\sum_{k' \neq k} |c_2^{k'} - c_1^{k'}|^2}}, \frac{\sqrt{\sum_{k' \neq k} |c_2^{k'} - c_1^{k'}|^2}}{|c_2^k - c_1^k|} \}$$

Solving the above maximization problem, we have $2a = \max_{k'} \left(\sum_{k \neq k'} |c_2^k - c_1^k|^2 \right)^{1/2}$, completing our proof. ■

We now describe the main results of this section, and provide upper bounds on the incentives for individuals and groups misreport in CCMs under the L_2-norm. Unlike the unconstrained case, the bound here applies for any group of manipulators with any preference profile, and is a function of the feasible locations only. The first result is for a single manipulator:

Theorem 4.5 (Approximate strategy-proofness of multi-dimensional CCM) *CCMs are ε-strategy-proof in multi-dimensional FLPs under the L_2-norm, where*

$$\varepsilon = \max_{(c_r, c_s) \in C} G(c_r, c_s)$$

Proof: For each feasible pair of outcomes $c_r, c_s \in C$, the gain of any agent when changing the outcome from c_r to c_s is at most $G(c_r, c_s)$ by Lemma 4.3. Maximizing over all feasible pairs completes the proof. ■

For group misreports, we provide a loose bound:

Theorem 4.6 (Approximate group strategy-proofness of multi-dimensional CCM) *CCMs are ε-group strategy-proof in multi-dimensional FLPs under the L_2-norm, where*

$$\varepsilon = \max_{c_r, c_s \in C} ||c_r - c_s||_2$$

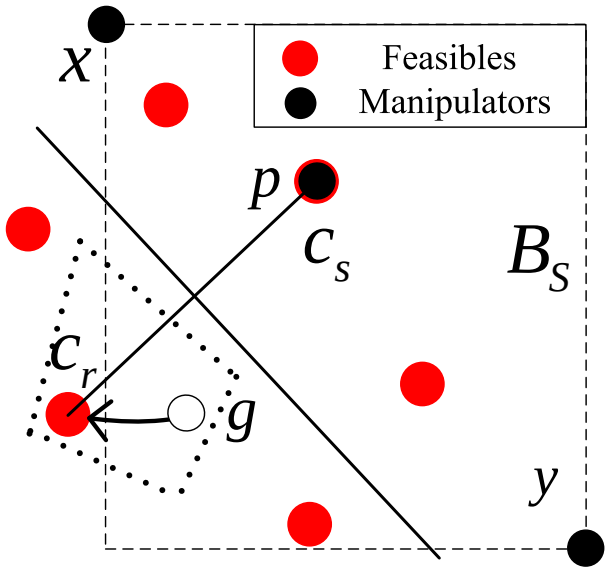


Figure 4.6: An example where a manipulator can benefit from changing the outcome from c_1 to c_2.

Proof: Consider any feasible pair of outcomes $c_r, c_s \in C$. Let g be one of the chosen locations under a QM f', which is projected to c_r under the CCM f. We can construct a location profile and a manipulator set $S = \{x, y, p\}$ such that: (i) all the manipulators are closer to c_s than to c_r; and (ii) one of the manipulators p coincides with c_s. In addition, we can also ensure that x and y are "far enough away" so that the bounding box containing x, y and p intersects with $\mathbb{Z}_{c_r}$ (as shown in Figure 4.6).

A viable group manipulation exists if all three manipulators misreport c_s, and move the selected quantile-location from g to d_s, in which the gain of the each manipulators is at most $||c_r - c_s||_2$ (the bound is tight if $q = 1$). Maximizing over all feasible pairs completes the proof. ■

Recall that this upper bound is a function of the feasible locations, so for the unconstrained facility location problem (where $C = \mathbb{R}^m$), this result reduces to Theorem 4.1. Also note that this bound can be viewed as a negative result, as it naturally holds in any mechanism for constrained FLPs (for any mechanism that maps agent profiles to feasible locations, this bound is the most that any agent can gain by misreporting). However, the proof of the worst-case bound makes strong assumptions about the locations of the peaks of both the sincere agents

and the manipulators. Such worst-case bounds are unlikely to arise in practice, as we explore empirically in the next section.

4.4 Empirical Analysis

The theoretical bounds derived above offer some insight into the performance of GMMs, QMs and CCMs w.r.t. incentive for manipulation. But the tightness of these bounds in practice depends on the distribution of agent preferences (i.e., their peaks in the underlying space). We evaluate these incentives empirically using two real-world data sets.

The first uses voting data from the Dublin West constituency in the 2002 Irish General Election.[3] It consists of 29,989 votes over nine candidates, with each vote a ranking of a subset the candidates. We use the 3800 votes that rank all nine candidates. For this data set, it includes only voter rankings of candidates and not ideal points (which may not correspond to any candidate). Furthermore, the (latent) space in which candidates and voter peaks lie is not given, and voter preferences may not be single-peaked. Fortunately, recent analysis has suggested not only that this data is approximately single-peaked in two dimensions (see Section 6.3 later), but also that a spatial model [Poole and Rosenthal, 1985] using L_2 distance provides a reasonable explanation of voter preferences ([Gormley and Murphy, 2007]). We fit this data to a 2D-spatial model by estimating both voter peaks (ideal points) and candidate positions (i.e., the feasible set), using an alternating optimization algorithm. The details will be introduced in Section 6.4 later. We use the estimated voter peaks in tests of unconstrained QMs (ignoring the candidates) and constrained CCMs (limiting selection to the nine candidates).

The second data set comprises geographic data for facility location [Daskin, 2011], with latitude and longitudes of 88 cities in the continental United States (the 48 state capitals unioned with the 50 largest cities). Following [Snyder and Daskin, 2005], we treat these locations as both the ideal points of 88 agents and the feasible locations in constrained FLPs. In other

[3] Available from www.dublincountyreturningofficer.com.

words, the agents reveal their locations (which we assume to be private, but in fact linked to a specific site) and then place a small set of facilities among themselves (the setup is similar to a voting for representatives from *within a group* [Alon et al., 2011]). This data is used to test CCMs.

To generate unconstrained FLPs from the voting data, we assume $s \in \{2, 5, 8\}$ manipulators and $n \in \{0, 2, 5, 10, 20, 50,80, 100\}$ sincere voters. For each setting, we randomly sample voter peaks from the 3800 estimated (spatial) positions to generate 1000 type profiles. For each profile, we either enumerate all manipulating coalitions of the required size or randomly sample $t \in \{10, 20, 50, 100, 200\}$ sets of s manipulators (depending on problem size). For each of the 1000 profiles, if *any* of the coalitions has a viable manipulation, we say the profile is manipulable and report the average gain of the coalition members in the coalition that has maximal gain.[4] We report the following in our results:

- The *probability of manipulation*, i.e., the proportion of the 1000 profiles that admit a beneficial manipulation for *some* coalition;
- The *(normalized) gain* for the coalition with maximal gain, averaged over the 1000 profiles; and
- The average *loss in social welfare* realized, relative to truthful reporting.

To test CCMs on constrained FLPs, we use a smaller number of manipulators 1, 2 and 4, but otherwise use the same settings as in the unconstrained case.

Figure 4.7 shows results on unconstrained problems for a single facility (winning candidate) using the median mechanism (quantile 0.5). Interestingly, the probability of manipulation increases with the number of sincere agents and converges to 1.0 (see the middle figure of Figure 4.7). This occurs because we simply measure whether *some* coalition among the set

[4]This set up assumes, somewhat unrealistically, that the members of this worst-case coalition can "discover" each other, and that they generate their misreport with *full knowledge* the reports of the sincere agents, as is common in analysis of manipulation in voting. For an analysis of manipulation in voting under more realistic knowledge assumptions, see [Lu et al., 2012].

of agents can successfully manipulate. This suggests that there is almost always some group whose peaks "contain" the median position. However, the left figure in Figure 4.7 shows that the average normalized gain decreases significantly with the number of sincere agents (e.g., with 2 manipulators, manipulation probability increases from 9.7% to 100%, but normalized gain reduces from 6.2% to 0.33%. Manipulative power is limited by the nearby peaks of sincere voters, and diminishes with more sincere voters. Impact on social welfare is also limited and is very small beyond 10 sincere agents, suggesting that QMs (including the median mechanism) are robust to manipulation in practice (note that manipulation may both increase or decrease total social cost).

We next evaluate CCMs in constrained two-facility FLPs, using the QM $\mathbf{q} = \{0.2, 0.3; 0.8, 0.7\}$ to make the initial selections (which are then projected using CCM). Figure 4.8 and 4.9 show the results on both the voting data set and the geographic data set, respectively. The results for the voting data in constrained FLPs is similar to those for the unconstrained FLPs, except that the probability of manipulation initially increases as the number of sincere agents grows, and then decreases. The initial increase occurs for the same reason as in the unconstrained case, and subsequently decreases because the number of feasible locations is fixed and small, which limits the probability of manipulation as the number of sincere agents increases. For the geographic data set, the probability of manipulation remains high, suggesting that there is always some group that can profitably manipulate a QM. Compared with the results on the voting data set, this occurs, in part, because the number of feasible outcomes increases as the number of agents increases, making it more probable for the manipulators to probe new possibilities. Average normalized gain and loss in social welfare is much higher than in the voting data set (e.g., with 2 of each agent type, average gain in constrained FLPs is 52%, compared to 4.7% in the voting data set). This is largely due to the fact that the agents' ideal locations and the feasible locations are much more tightly clustered in the geographic data set (since ideal points coincide with feasible locations) than in the voting data set. Despite this, both average gain and impact on social cost drop quickly with the number of sincere agents.

4.5 Conclusion

In this chapter, we have studied the mechanism design problem for both unconstrained and constrained FLPs, investigating the degree to which individual and group strategy-proofness can be achieved, and providing bounds on the incentive for individuals and groups to misreport in generalized median, quantile, and our newly proposed closet candidate mechanisms. Empirical analysis of Irish electoral data shows that these mechanisms may perform extremely well in practice, limiting the odds of manipulation and especially the potential gains and impact on social welfare.

There are several interesting future directions that extend the results in this chapter. Exploring the approximate incentive properties of additional mechanisms (beyond GMMs, QMs, CCMs) and cost functions (beyond L_2) is of interest. The exploration of incremental (or multi-stage) mechanisms that trade off social cost, incentives, privacy and communication would be extremely valuable (as will be discussed in Chapter 7). In addition, preferences are often not fully single-peaked in realistic domains, but are often approximately so (as we will see in Chapter 6). Extending the theoretical analysis to this setting would be of value. Finally, we are interested in examining the optimization problem facing manipulators when they have only probabilistic knowledge of the potential reports of the sincere agents, as well as the impact of this limited knowledge on the probability of manipulation, average gain/incentive, and loss in social welfare Lu et al. [2012]. This would provide a more realistic assessment of the robustness/resistance of GMMs and QMs to group manipulation. We will discuss more future directions in Chapter 8.

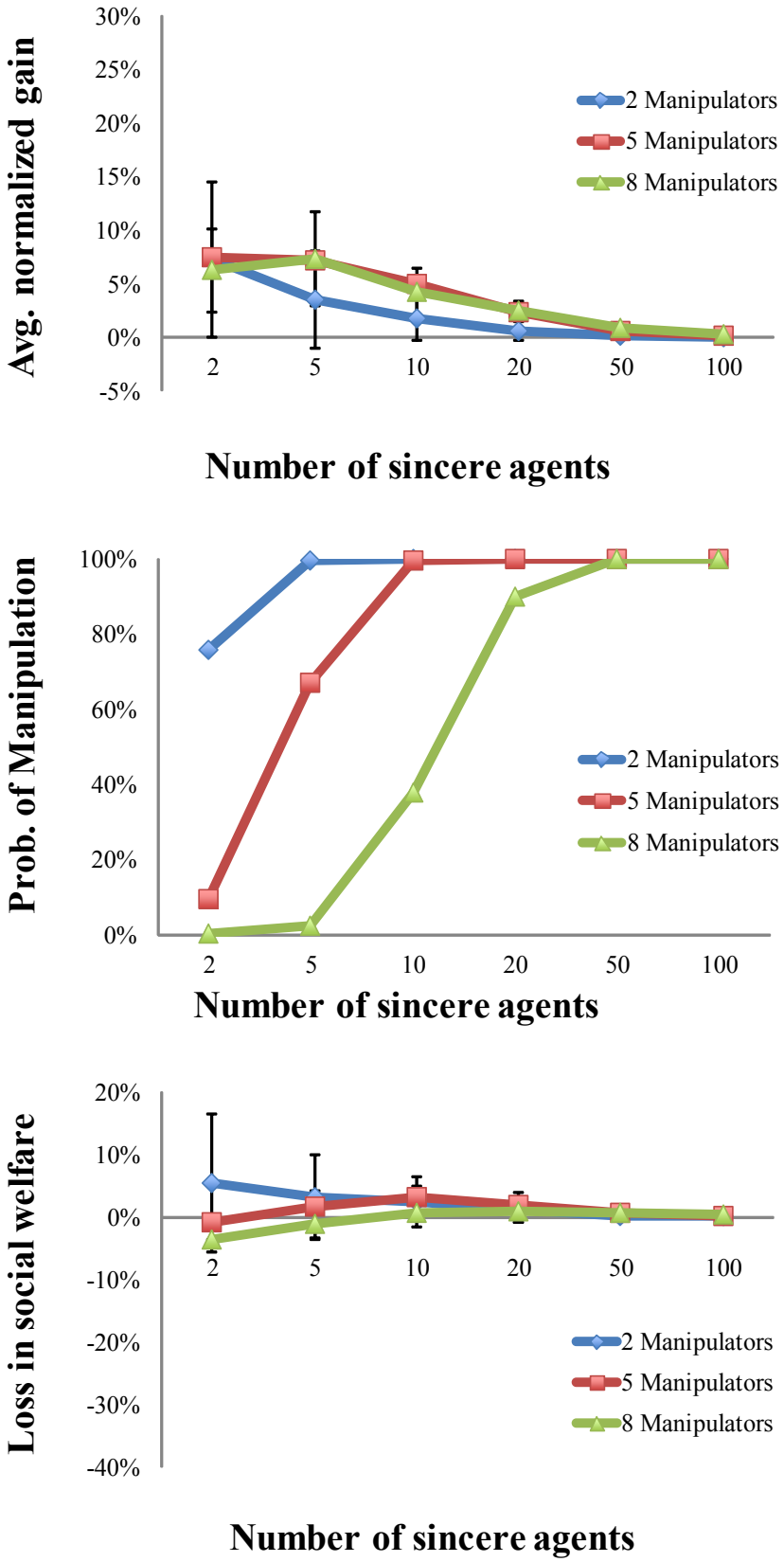


Figure 4.7: Unconstrained, single-FLPs and GMMs: normalized gain (top), prob. of manipulation (middle), and loss in social welfare (bottom). The error bars show the standard deviation for each point.

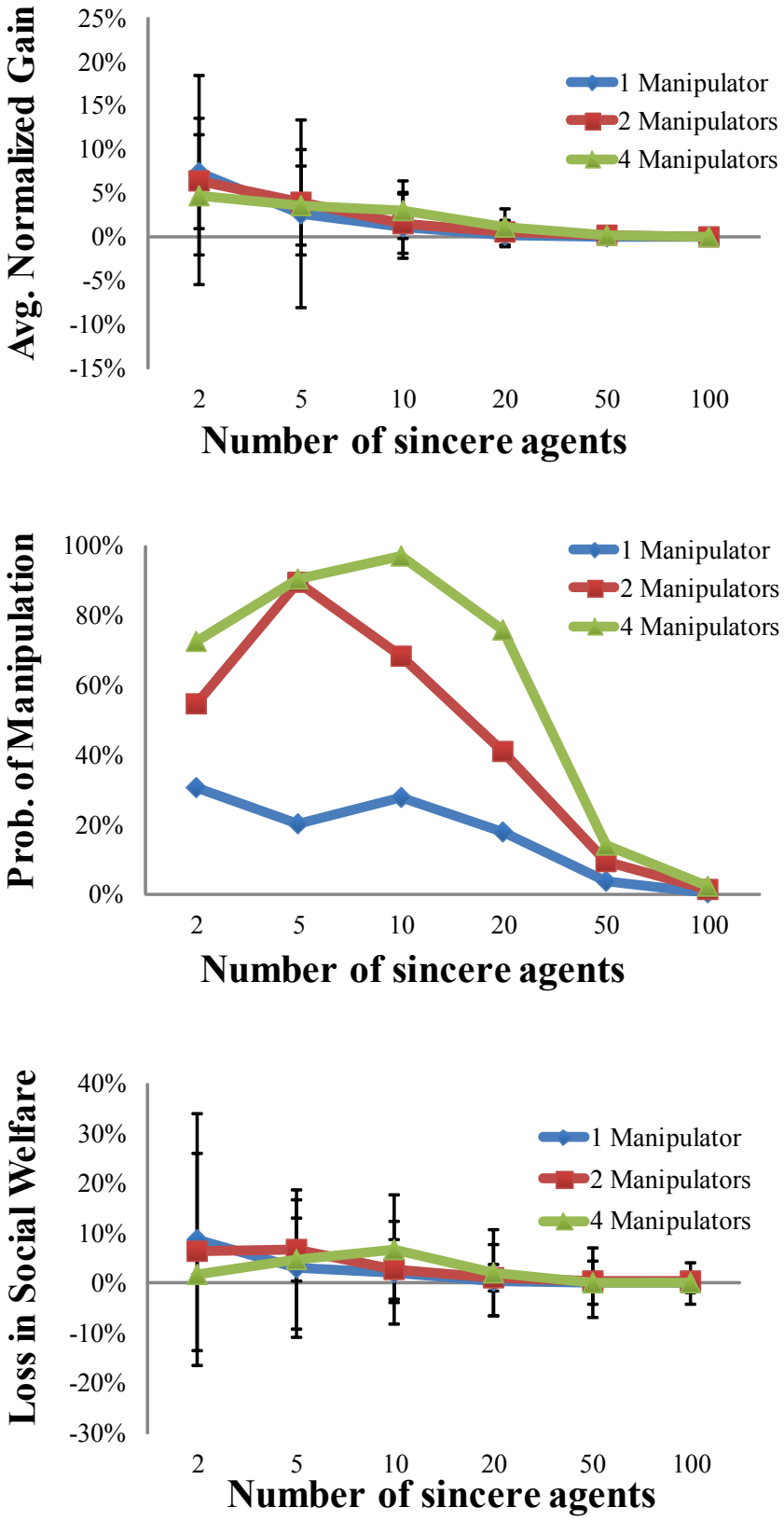
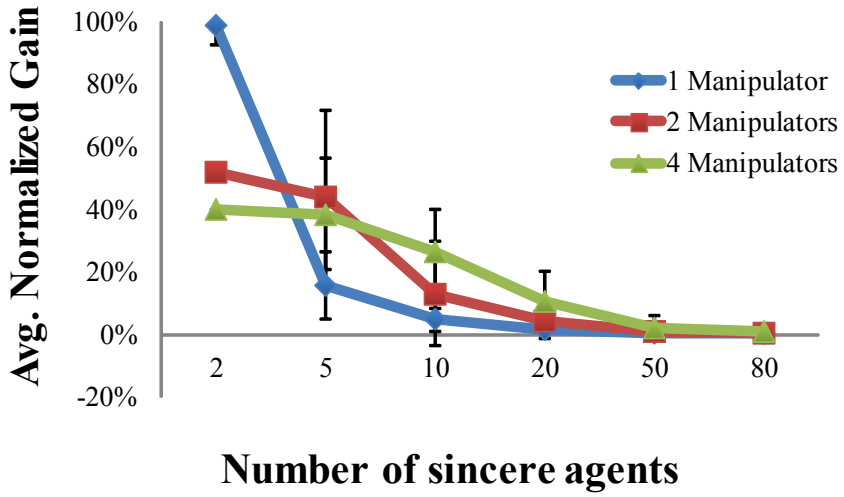
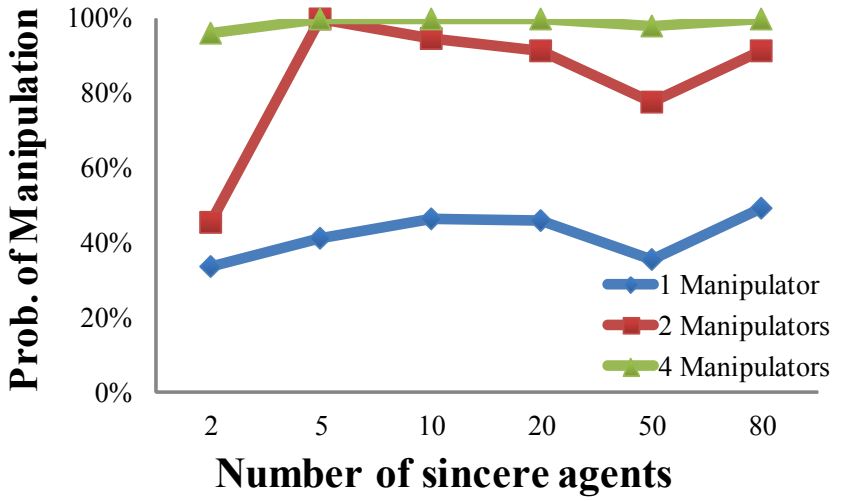


Figure 4.8: Constrained single-FLPs and CCMs for the voting data: normalized gain (top), prob. of manipulation (middle), and impact on social cost (bottom). The error bars show the standard deviation for each point.





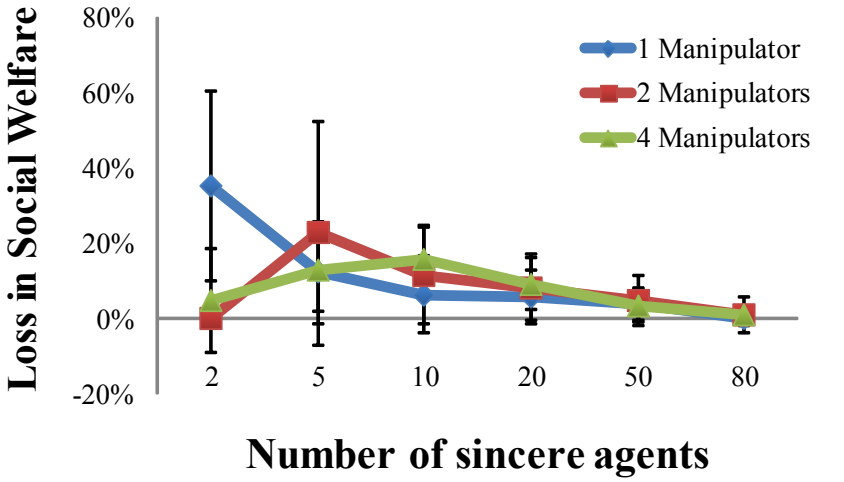


Figure 4.9: Constrained single-FLPs and CCMs for the geographic data: normalized gain (top), prob. of manipulation (middle), and impact on social cost (bottom). The error bars show the standard deviation for each point.

Chapter 5

Group Manipulation: Optimization and Complexity

5.1 Introduction

Quantile mechanisms, as defined in Chapter 3, are strategy-proof, but as shown in Chapter 4, are not group strategy-proof, and provide several possibility/impossibility results on the incentives for a group of agents to misreport. While group strategy-proofness cannot be guaranteed, this does not mean that finding a viable or optimal group manipulation is computationally feasible for a group of agents. In this chapter, we study the computational complexity for group manipulation in unconstrained facility location. Specifically, focusing on quantile mechanisms (QMs) (and to some extent *generalized median mechanisms (GMMs)*), we consider the formulation of the *optimal group manipulation problem* as mathematical programs of various types; the computational complexity of this problem; and how much manipulators might gain given optimal manipulations, under different cost functions, when GMMs/QMs are used.

Our primary contribution is to formulate the group manipulation problem—for both single- and multi-FLPs under both the L_1- and L_2-norms (where these metrics measure distance/cost between ideal points and facilities)—as convex optimization problems, and study their compu-

tational complexity. We show that single-FLPs with L_1 and L_2 costs can be specified as linear programs (LPs) and second-order cone programs (SOCPs), respectively. This means both can be solved in polynomial time (using interior point methods [Boyd and Vandenberghe, 2004]). By contrast, we show that multi-FLPs are NP-hard by reduction from the geometric p-median problem [Megiddo and Supowit, 1984] under both norms. Despite this, we provide formulations of these problems as mixed integer linear (MILPs) and mixed integer SOCPs (MISOCPs) for L_1 and L_2 costs, respectively. We also test these formulations empirically, with results that suggest commercial solvers can compute optimal group manipulations (or prove that none exists) for multi-FLPs of reasonable size rather effectively, despite the theoretical NP-hardness of the problem.

5.2 Group Manipulation for Single-Facility Location Problems

In this section, we address the problem of group manipulation for single-facility location problems, first describing its general form, then describing a linear programming formulation under the L_1-norm (or distance metric), and finally describing a second-order cone programming formulation under the L_2-norm.

Following the notation in Section 2.3, we let n be the number of agents, q be the number of facilities, and m be the number of dimensions. Also let $f_{\mathbf{P}}$ be a quantile mechanism with quantile matrix $\mathbf{P}$, which select q homogeneous facilities in the m-dimensional space $\mathbb{R}^m$. Such an outcome is represented by a location vector $\mathbf{x} = (x_1, x_2, \ldots, x_q)$, where $x_j \in \mathbb{R}^m$. Each agent has a type $t_i \in T_i$ determining her cost associated with any location vector $\mathbf{x}$, i.e., $c_i(\mathbf{x}, t_i) = \min_{j \leq q} c_i(x_j, t_i)$, in which each agent uses the facility with least cost.

Informally, the *optimal group manipulation problem* is that of finding a joint misreport for a group of manipulators such that the outcome induced by this misreport is such that: (a) the sum of costs of the manipulators is minimized; and (b) relative to the outcome that would have

been induced by truthful reporting, no manipulator is worse off (has a higher cost) and at least one is strictly better-off (has lower cost). We formalize this as follows:

Definition 5.1 (Optimal group manipulation) *Let $N = S \cup M$, where S is a set of sincere agents and M is a set of manipulators with type vectors $\mathbf{t}_S$ and $\mathbf{t}_M$, respectively. Let $f_{\mathbf{P}}$ be a QM with quantile matrix $\mathbf{P}$. Let $x_{\mathbf{P}} = f_{\mathbf{P}}(\mathbf{t}_M, \mathbf{t}_S)$ be the location chosen by $f_{\mathbf{P}}$ if all agents report their peaks truthfully, and $x'_{\mathbf{P}} = f_{\mathbf{P}}(\mathbf{t}'_M, \mathbf{t}_S)$ be the location chosen given some misreport $\mathbf{t}'_M$ by the manipulators M. The* optimal group manipulation problem *is to find a joint misreport $\mathbf{t}'_M$ for the agents in M satisfying:*

$$t'_M = \arg\min_{\mathbf{t}'_M} \sum_{i \in M} c_i\left(x'_{\mathbf{P}}, t_i\right) \tag{5.1}$$

$$\text{s.t. } c_i\left(x'_{\mathbf{P}}, t_i\right) \leq c_i\left(x_{\mathbf{P}}, t_i\right), \quad \forall i \in M \tag{5.2}$$

$$c_i\left(x'_{\mathbf{P}}, t_i\right) < c_i\left(x_{\mathbf{P}}, t_i\right), \quad \textit{for some } i \in M \tag{5.3}$$

Notice that we assume agents have non-transferable utilities, otherwise we can optimize objective (5.1) without the constraints (this is because for the case of transferable utilities, all agents can be made better off by transferring some positive fraction of the gain to each other). One may also argue whether the objective of minimizing social cost among the manipulators makes sense. Definitely, one can use other objective functions (e.g., maximum cost minimization), however, we choose this particular one because: 1) this is a natural objective that maximizes the social welfare of the manipulators, and (2) it subsumes the "existence of a misreport", i.e., whether there exists a joint misreport that no one is worse-off and the total gain is greater than zero.

Given a group of manipulators M, we generally refer to the remaining agents $S = N \setminus M$ as "sincere," though we need not presume that their reports are truthful in general, only that M knows (or can anticipate) their reports.

5.2.1 Group Manipulation Specification

Recall from Defn. 5.1 that a group manipulation is a set of misreports by the manipulating coalition M such that no manipulator is worse off and at least one is better off. The optimization formulation of this problem in Eq. (5.16) requires that one find the misreport that provides the greatest total benefit to the coalition. This explicit, straightforward formulation considers all possible misreports (i.e., the vector of purported "preferred" locations of each manipulator), which in principle induces a very large high-dimensional search space, from which the optimal misreport must be selected.

Fortunately, we can decrease the search space dramatically by considering only *viable* locations for manipulator misreports. We first define *viability*:

Definition 5.2 (Viability) *Let $f_{\mathbf{P}}$ be a QM with quantile matrix $\mathbf{P}$, and $\mathbf{t}_S$ be the reported types of the sincere agents in $S = N \setminus M$. A location $x \in \mathbb{R}^m$ is* viable *for a manipulating coalition M if there exists a joint misreport $\mathbf{t}'_M$ s.t. $x = f_{\mathbf{P}}(\mathbf{t}'_M, \mathbf{t}_S)$ We say $\mathbf{t}'_M$* implements *x in this case.*

The following proposition shows that, in single-FLPs, if a mechanism $f_{\mathbf{P}}$ selects a location $x'_{\mathbf{P}} = f_{\mathbf{P}}(\mathbf{t}_S, \mathbf{t}'_M)$ under a group manipulation $\mathbf{t}'_M$, then it also selects $x'_{\mathbf{P}}$ if each manipulator misreports $x'_{\mathbf{P}}$ as her peak.

Proposition 5.1 *For single-FLPs, let $\mathbf{t}'_M$ be a group manipulation and $x'_{\mathbf{P}}$ be a viable location implemented by $\mathbf{t}'_M$ under mechanism $f_{\mathbf{P}}$. Then $x'_{\mathbf{P}}$ is also implemented by the group manipulation $\mathbf{t}^*_M = \{x'_{\mathbf{P}}, \ldots, x'_{\mathbf{P}}\}$.*

Proof: We first provide a proof for $m = 2$ first, and then show how the analysis can be generalized to the case of $m > 2$. Consider an arbitrary group manipulation $\mathbf{t}'_M$, which implements location $x'_{\mathbf{P}} = f_{\mathbf{P}}(\mathbf{t}'_M, \mathbf{t}_S) \in \mathbb{R}^2$ (as shown in Figure 5.1). Let us denote the misreport of each manipulator by $t'_i = (t'^1_i, t'^2_i), \forall i \in M$ and the location by $x'_{\mathbf{P}} = (x'^1_{\mathbf{P}}, x'^2_{\mathbf{P}})$.

Pick an arbitrary manipulator $i \in M$, and assume w.l.o.g. that $t'^1_i \leq x'^1_q$ and $t'^2_i \geq x'^2_q$. We

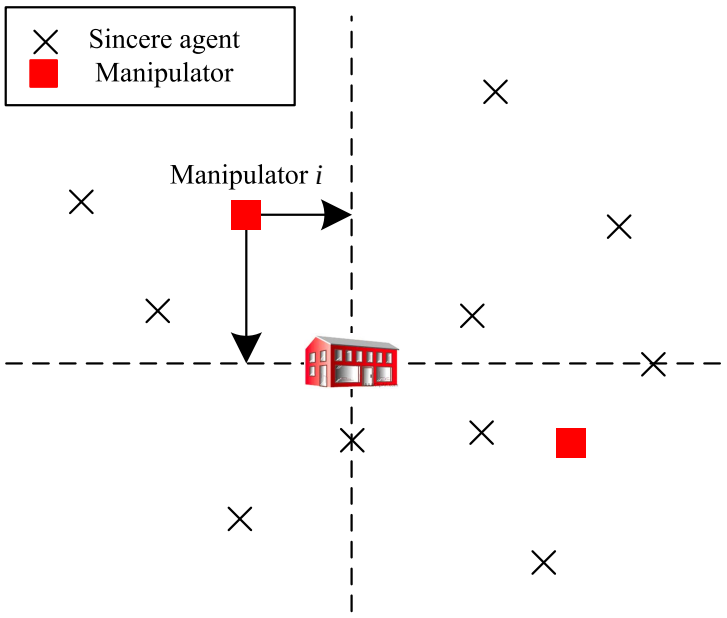


Figure 5.1: Each manipulator can move her misreport to $x'_{\mathbf{P}}$ without changing the outcome.

construct another group manipulation $\mathbf{t}''_M$ by changing the misreport of manipulator i to $x'_{\mathbf{P}}$. Recall that the mechanism $f_{\mathbf{P}}$ locates the facility at a specified quantile, so we have:

$$\begin{aligned} f_{\mathbf{P}}(\mathbf{t}'_M, \mathbf{t}_S) &= f_{\mathbf{P}}((t'_i, \mathbf{t}'_{M\setminus i}), \mathbf{t}_S) \\ &= f_{\mathbf{P}}(((t_i'^1, x_{\mathbf{P}}'^2), \mathbf{t}'_{M\setminus i}), \mathbf{t}_S) \\ &= f_{\mathbf{P}}(((x_{\mathbf{P}}'^1, x_{\mathbf{P}}'^2), \mathbf{t}'_{M\setminus i}), \mathbf{t}_S) \\ &= f_{\mathbf{P}}((x'_{\mathbf{P}}, \mathbf{t}'_{M\setminus i}), \mathbf{t}_S) = f_{\mathbf{P}}(\mathbf{t}''_M, \mathbf{t}_S) \end{aligned}$$

Repeating this procedure over all manipulators completes our proof for $m = 2$.

For the case of $m > 2$, we can also change the misreport from each manipulator to the coordinate of the implemented location on each dimension independently, without changing the location of the facility. This completes our proof. ■

Proposition 5.1 demonstrates that we can limit our attention to the "unanimous" reporting of viable locations when searching for optimal misreports, without considering misreports that reveal locations that cannot be implemented or realized by the manipulators. Since Proposition 5.1 gives us license to consider only viable locations as potential misreports, we can

reformulate the optimal group manipulation problem (Defn. 5.1) as follows:

Definition 5.3 (Optimal group manipulation) *Let $f_{\mathbf{P}}$ be a QM with quantile matrix $\mathbf{P}$. Let $x_{\mathbf{P}} = f_{\mathbf{P}}(\mathbf{t}_M, \mathbf{t}_S)$ be the location chosen by $f_{\mathbf{P}}$ if all agents report their peaks truthfully, and $x'_{\mathbf{P}} = f_{\mathbf{P}}(\mathbf{t}'_M, \mathbf{t}_S)$ be the location chosen given some misreport $\mathbf{t}'_M$ by the manipulators M. The optimal group manipulation problem can be reformulated as:*

$$\min_{x \in \mathbb{R}^m} \sum_{i \in M} c_i\left(x, t_i\right) \tag{5.4}$$

$$\text{s.t.} \quad c_i\left(x, t_i\right) \leq c_i\left(x_{\mathbf{P}}, t_i\right), \quad \forall i \in M \tag{5.5}$$

$$c_i\left(x, t_i\right) < c_i\left(x_{\mathbf{P}}, t_i\right), \quad \textit{for some } i \in M \tag{5.6}$$

$$x \text{ is a viable location under } f_{\mathbf{P}} \text{ and } \mathbf{t}_S \tag{5.7}$$

In the sequel, our specific formulations of the problem will rely on Defn. 5.3. We can also safely omit the constraints embodied in Eq. 5.6, as they can easily be checked after the fact given the optimized location vector—if no manipulator is strictly better off under the optimal misreport, then a group manipulation obviously cannot exist.

5.2.2 LP Formulation under the L_1-norm

We now consider the formulation of optimal manipulation when the L_1-norm is used as the cost function, i.e., $c_i(x, t_i) = \sum_{k \leq m} |x^k - t_i^k|$ for any location $x \in \mathbb{R}^m$. Let $x = (x^1, \ldots, x^m)$ represent the location to be optimized (i.e., the location induced by the manipulation) in single-FLPs, where each x^k is a continuous variable. Let c_i be a continuous variable denoting the cost of manipulator i given outcome x. We can formulate the objective function Eq. (5.4), and the

constraints Eq. (5.5), as follows:

$$\min_{x} \sum_{i \in M} c_i \tag{5.8}$$

$$\textit{s.t.} \quad c_i = \sum_{k \leq d} |x^k - t_i^k|, \quad \forall i \in M \tag{5.9}$$

$$0 \leq c_i \leq u_i, \qquad \forall i \in M \tag{5.10}$$

where $u_i = c_i(x_{\mathbf{P}}, t_i)$ is the cost of manipulator i under a truthful report $\mathbf{t}_M$ by the manipulators.

This formulation contains absolute values in the nonlinear constraints (5.9). We introduce an additional set of variables to linearize these constraints. Letting D_i^k be an upper bound on the distance between t_i and x in the kth dimension, we linearize the constraints (5.9) as follows:

$$-D_i^k \leq t_i^k - x^k \leq D_i^k, \quad \forall i \in M, \forall k \leq m \tag{5.11}$$

$$D_i^k \geq 0, \qquad \forall i \in M, \forall k \leq m \tag{5.12}$$

$$c_i = \sum_{k \leq m} D_i^k, \quad \forall i \in M \tag{5.13}$$

Finally, we need constraints that guarantee that new location x is viable. Recall that a QM locates the facility at a specified quantile of the reported peaks in each dimension independently, and by Proposition 5.1 we can assume w.l.o.g. that all manipulators use the same misreport. This implies that a viable location for the facility is bounded by the reported coordinates of two sincere agents in each dimension. Formally, let $\mathbf{P} = (p^1, \ldots, p^m)$ be the quantile vector for mechanism $f_{\mathbf{P}}$ (for single-FLPs, we have a single vector rather than a full matrix), and let

$$\perp^k = \min\{z \in \mathbb{Z}^+ : z + |M| \geq p^k \cdot n\} \textit{ and}$$

$$\top^k = \max\{z \in \mathbb{Z}^+ : |S| + |M| - z \geq (1 - p^k) \cdot n\}.$$

(For convenience, we assume w.l.o.g. that $p^k \cdot n$ is an integer.) If we let $\tilde{x}_S^k = \{\tilde{x}_1^k, \ldots, \tilde{x}_{|S|}^k\}$ denote the ordered coordinates of the reported peaks of the sincere agents S in the kth dimension, then we have:

Lemma 5.1 *For single-FLPs, a location $x = (x^1, \ldots, x^m) \in \mathbb{R}^m$ is viable if and only if $\tilde{x}_{\perp^k}^k \leq x^k \leq \tilde{x}_{\top^k}^k, \forall k \leq m$.*

Proof: We first show x is viable if $\tilde{x}_{\perp^k}^k \leq x^k \leq \tilde{x}_{\top^k}^k, \forall k \leq m$. By Lemma 5.1, we can assume w.l.o.g. that all manipulators misreport x^k on the kth dimension. Also by the definition of $\perp^k$ and $\top^k$ there are $(p^k \cdot n - |M|)$ and $((1-p^k) \cdot n - |M|)$ sincere agents on the left of (including equal to) $\tilde{x}_{\perp^k}^k$ and on the right of (including equal to) $\tilde{x}_{\top^k}^k$ on the $k$the dimension, respectively[1]. Now consider any x^k satisfying $\tilde{x}_{\perp^k}^k \leq x^k \leq \tilde{x}_{\top^k}^k$. As all manipulators misreport x^k on the kth dimension, then the total number of agent (including sincere agents and manipulators) on the left and right of or equal to x^k is at least $(p^k \cdot n)$ and $((1-p^k) \cdot n)$, respectively. Recall that the QM locates the facility at the p^kth quantile on each dimension we know $x^k = f_{\mathbf{P}}^k$, i.e., x is a viable location.

We prove the converse by contradiction. Suppose x is a viable location where, say w.l.o.g., $x^k < \tilde{x}_{\perp^k}^k$. By Lemma 5.1, we can assume w.l.o.g. that all manipulators misreport x^k on dimension k. By the definition $\perp^k$, the total number of agents (including sincere agents and manipulators) on the left of or equal to $\tilde{x}_{\perp^k}^k$ is at most $(p^k \cdot n - |M| + |M|) = (p^k \cdot n)$. This suggests that $f_{\mathbf{P}}^k = \tilde{x}_{\perp^k}^k > x^k$, contradicts the fact that x is a viable location. ∎

This lemma ensures that we can use the following boundary constraints as to enforce viability (see Eq. (5.7)):

$$\tilde{x}_{\perp^k}^k \leq x^k \leq \tilde{x}_{\top^k}^k, \quad \forall k \leq m \tag{5.14}$$

To summarize, we can formulate the optimal group manipulation under the L_1-norm as an LP, which is stated in the following theorem:

[1]For convenience, we assume w.l.o.g. that $q^k \cdot n$ is an integer.

$$
\begin{aligned}
&\min_{x} \sum_{i \in M} c_i \\
&\textit{s.t.} \quad 0 \leq c_i \leq u_i, \qquad \forall i \in M \\
&-D_i^k \leq t_i^k - x^k \leq D_i^k, \quad \forall i \in M, \forall k \leq m \\
&D_i^k \geq 0, \qquad \forall i \in M, \forall k \leq m \\
&c_i = \sum_{k \leq m} D_i^k, \quad \forall i \in M \\
&\tilde{x}^k_{\perp^k} \leq x^k \leq \tilde{x}^k_{\top^k}, \quad \forall k \leq m
\end{aligned}
$$

Figure 5.2: The complete linear program of optimal group manipulation for single facility location problem under the L_1-norm.

Theorem 5.1 (LP for optimal group manipulation in single-FLPs) *The optimal group manipulation problem for single facility location under the L_1-norm can be formulated as a linear program (LP), with objective function (5.8) and constraints (5.10)-(5.14).*

Proof: Figure 5.2 provides a snapshot of the whole LP. The objective function (5.8) minimizes the sum of costs over all manipulators. Constraints (5.10)-(5.13) guarantee that no manipulators is worse off, and constraints (5.14) ensure that the optimized location induced by the misreport is viable. As both the objective and the constraints are linear, and all variables are continuous, this constitutes a linear program. The total number of variables is $(m+1)|M|+m$ or $O(m|M|)$ (where $|M|$ comes from the manipulator cost c_i, $m|M|$ comes from D_i^ks, and the rest m comes from x). ■

As such, the optimal manipulation problem can be solved in polynomial time.

5.2.3 SOCP Formulation under the L_2-norm

The optimization formulation for the L_1-norm above can be easily modified to account for L_2 costs. Specifically, we need only a minor modification of the constraints (5.13) to incorporate

$$
\begin{aligned}
&\min_{x} \sum_{i \in M} c_i \\
&\textit{s.t.} \quad 0 \leq c_i \leq u_i, \qquad \forall i \in M \\
&-D_i^k \leq t_i^k - x^k \leq D_i^k, \quad \forall i \in M, \forall k \leq m \\
&D_i^k \geq 0, \qquad \forall i \in M, \forall k \leq m \\
&\left(c_i\right)^2 \geq \sum_{k \leq m} \left(D_i^k\right)^2, \quad \forall i \in M \\
&\tilde{x}_{\perp^k}^k \leq x^k \leq \tilde{x}_{\top^k}^k, \quad \forall k \leq m
\end{aligned}
$$

Figure 5.3: The complete second-order cone program of optimal group manipulation for single facility location problem under the L_2-norm.

Euclidean distances as follows:

$$\left(c_i\right)^2 \geq \sum_{k \leq m} \left(D_i^k\right)^2, \quad \forall i \in M \tag{5.15}$$

We also have the following theorem:

Theorem 5.2 (SOCP for optimal group manipulation in single-FLPs) *The optimal group manipulation problem for the single facility location under the L_2-norm can be formulated as a second-order cone program (SOCP), with objective function (5.8) and constraints (5.10)-(5.12) and (5.14)-(5.15).*

Proof: Figure 5.3 provides a snapshot of the whole SOCP. The objective function (5.8) minimizes the sum of costs over all manipulators. Constraints (5.10)-(5.12) and (5.14)-(5.15) guarantee that no manipulator is worst-off and the new location is viable. This constitutes a second-order cone program (SOCP) under the L_2-norm. The total number of variables is also $O(m|M|)$. ■

Since SOCPs can be solved in polynomial time, we have the following:

Remark 5.1 *The optimal group manipulation problem for single-facility location under both the L_1- and L_2-norms can be solved in polynomial time.*

5.3 Group Manipulation for Multi-Facility Location Problems

In this section, we extend our analysis of group manipulation to multi-facility location problems. Unlike single-FLPs, we show that problem in computationally intractable for multi-FLPs, under both the L_1- and L_2-norms. However, we provide mathematical programming models that are often quite efficient in practice.

Following the notation in Section 2.3, we let n be the number of agents, q be the number of facilities, and m be the number of dimensions. Also let $f_{\mathbf{P}}$ be a quantile mechanism with quantile matrix $\mathbf{P}$, which select q homogeneous facilities in the m-dimensional space $\mathbb{R}^m$. Such an outcome is represented by a location vector $\mathbf{x} = (x_1, x_2, \ldots, x_q)$, where $x_j \in \mathbb{R}^m$. Each agent has a type $t_i \in T_i$ determining her cost associated with any location vector $\mathbf{x}$, i.e., $c_i(\mathbf{x}, t_i) = \min_{j \leq q} c_i(x_j, t_i)$, in which each agent uses the facility with least cost. Then the optimal group manipulation problem is to find a joint misreport $\mathbf{t}'_M$ for the agents in M satisfying:

$$t'_M = \arg\min \sum_{i \in M} c_i \left(\mathbf{x}'_{\mathbf{P}}, t_i\right) \tag{5.16}$$

$$\textit{s.t.} \;\; c_i \left(\mathbf{x}'_{\mathbf{P}}, t_i\right) \leq c_i \left(\mathbf{x}_{\mathbf{P}}, t_i\right), \quad \forall i \in M \tag{5.17}$$

$$c_i \left(\mathbf{x}'_{\mathbf{P}}, t_i\right) < c_i \left(\mathbf{x}_{\mathbf{P}}, t_i\right), \quad \text{for some } i \in M \tag{5.18}$$

where $\mathbf{x}_{\mathbf{P}} = f_{\mathbf{P}}(\mathbf{t}_M, \mathbf{t}_S)$ is the location vector chosen by $f_{\mathbf{P}}$ if all agents report their peaks truthfully, and $\mathbf{x}'_{\mathbf{P}} = f_{\mathbf{P}}(\mathbf{t}'_M, \mathbf{t}_S)$ is the location vector chosen given some misreport $\mathbf{t}'_M$ by the manipulators M.

5.3.1 The Complexity of Group Manipulation

We first show that group manipulation is NP-hard for multi-FLPs.

Theorem 5.3 (NP-Hardness of optimal group manipulation in multi-FLPs) *The optimal group manipulation problem for multi-facility location under either the L_1- or L_2-norms is NP-hard.*

Proof: We prove the hardness result using a reduction from p-median. W.l.o.g., we can focus on the two-dimensional version of this problem: Given a set $R = \{(x_1, y_1), \ldots, (x_t, y_t)\}$ of points in the plane, we want to find a set $Q = \{(z_1, s_1), \ldots, (z_p, s_p)\}$ of p points so as to minimize:

$$\sum_{i=1}^{t} \min_{1 \leq j \leq p} \{|x_i - z_j| + |y_i - s_j|\} \tag{5.19}$$

or

$$\sum_{i=1}^{t} \min_{1 \leq j \leq p} \{\sqrt{(x_i - z_j)^2 + (y_i - s_j)^2}\} \tag{5.20}$$

where formula (5.19) and (5.20) are referred to as the Rectilinear and Euclidean p-median problem, respectively.

Consider the following 3 steps.

I. Take an arbitrary instance of the p-median problem. Specifically, let $R = \{(x_1, y_1), \ldots, (x_t, y_t)\}$ be a set of t points in the plane, and $Q^* = \{(z_1^*, s_1^*), \ldots, (z_p^*, s_p^*)\}$ be a p-median of R. We assume all the points in R are in the unit box, and so are the medians (as shown in Figure 5.4) without loss of generality, as for any fixed problem, we can bring all points into this normalized from by rescaling.

II. Let $\mathbf{P}$ be a quantile matrix that is *consistent* with Q^*. We say a quantile matrix is consistent with a set of points Q if there exists a type profile $\mathbf{t}$ such that $f_{\mathbf{P}}(\mathbf{t}) = Q$. It is not hard to see that such a quantile matrix always exists (if may not be unique).[2] Recall that $\mathbf{P}$ is a $p \times 2$ matrix with the form of $\mathbf{P} = \{(p_1^1, p_1^2); \ldots; (p_p^1, p_p^2)\}$. We can create another

[2]Given a set of points Q, we can construct a consistent quantile matrix $\mathbf{P}$ as follows: for any two points $(q_j^1, q_j^2), (q_{j'}^1, q_{j'}^2) \in Q$, we have to guarantee that the corresponding quantile values satisfy $p_j^1 \leq p_{j'}^1$ iff $q_j^1 \leq q_{j'}^1$ and $p_j^2 \leq p_{j'}^2$ iff $q_j^2 \leq q_{j'}^2$. If we do this for every pair of points in Q, we can get a consistent quantile matrix.

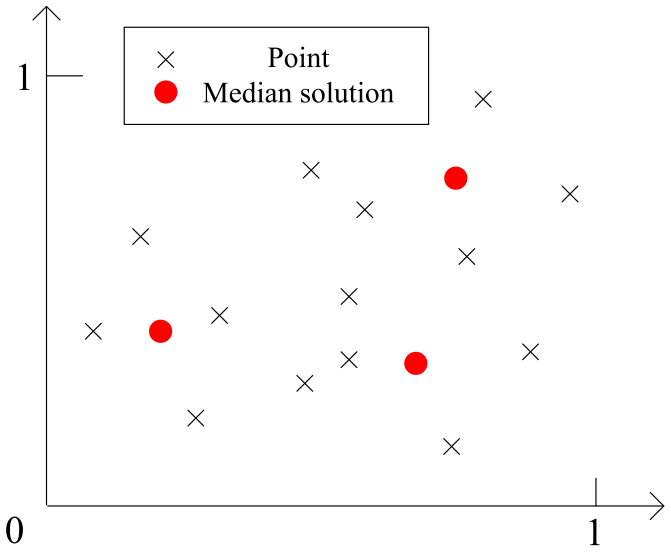


Figure 5.4: The p-medians of a set of points, where each $\times$ represents a given point, and each $\bullet$ represents one solution point of the p-median problem.

matrix $\tilde{\mathbf{P}}$ by adding $(0,0)$ as the first row to $\mathbf{P}$, i.e., $\tilde{\mathbf{P}} = \{(0,0); (p_1^1, p_1^2); \ldots; (p_p^1, p_p^2)\}$.

Let $\top = \max\{p_1^1, \ldots, p_p^1, p_1^2, \ldots, p_p^2\}$, i.e., the maximum quantile value in $\mathbf{P}$. Then we create another set $\tilde{R}$ of $t+U$ points by adding U copies of $(-b,-b)$ into R, i.e., $\tilde{R} = \{\overbrace{(-b,-b), \ldots, (-b,-b)}^{U copies}\} \cup R$, where b is a positive number, and U is a large positive integer satisfying $(U+t) \cdot \top \leq U$. The reason that we add these points is to guarantee that the quantile mechanism $f_{\tilde{P}}$ will locate all facilities at $(-b,-b)$ if if we treat the points in $\tilde{R}$ as peaks from the agents.

Now consider the optimal group manipulation problem for $(p+1)$ facilities under the quantile mechanism $f_{\tilde{\mathbf{P}}}$, where the ideals of manipulators are $\tilde{R}$ and there is no sincere agent. If the U we choose satisfy the above inequality and all manipulators report their peaks truthfully, then all the $(p+1)$ facilities will be located at $(-b,-b)$ (as shown in Figure 5.5), i.e., $\mathbf{x}_{\tilde{\mathbf{P}}} = \{\overbrace{(-b,-b), \ldots, (-b,-b)}^{(p+1) copies}\}$. Under this location vector $\mathbf{x}_{\tilde{\mathbf{P}}}$, every manipulator with the peak of $(-b,-b)$ has a cost of 0, and everyone else has a cost of at least $2b$ under the L_1-norm or $\sqrt{2}b$ under the L_2-norm. However, there exists a group

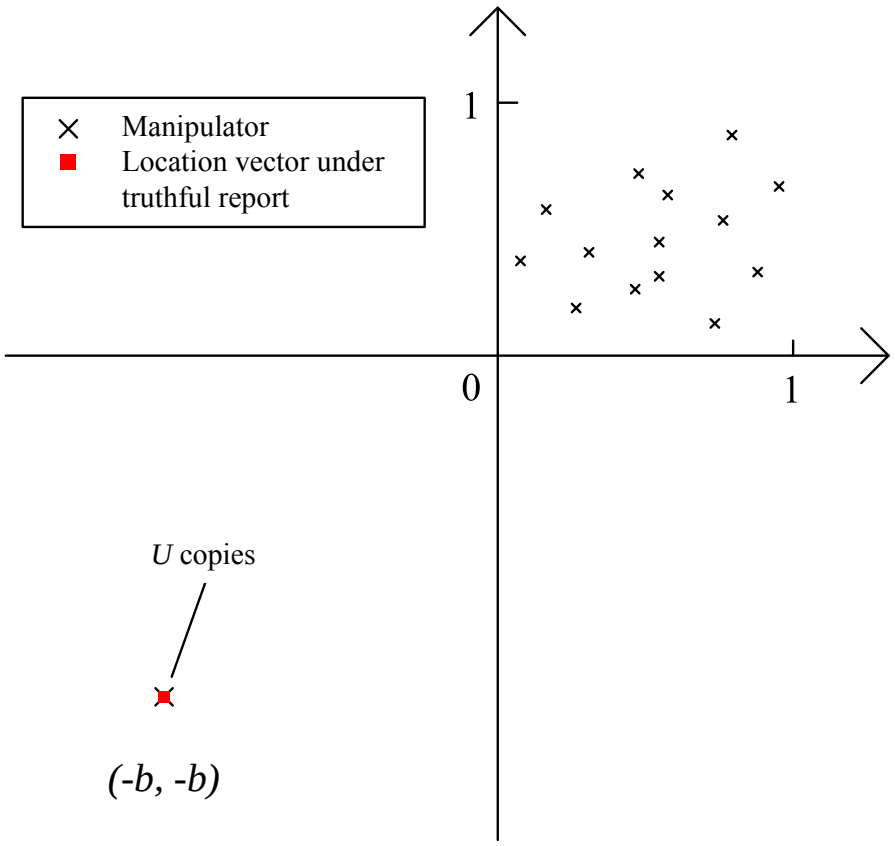


Figure 5.5: The outcome of mechanism $f_{\mathbf{P}}$ if all manipulators report truthfully, where each $\times$ represents the true peak of a manipulator, and each $\square$ represents the location of one facility under $f_{\mathbf{P}}$.

manipulation: if some manipulators at $(-b, -b)$ misreport their peaks, claiming they lie within the unit box, then this will "push" some facilities into the unit box, benefiting the other t manipulators.

III. So far, we have encoded the p-median problem as a group manipulation problem. The last step is to show that if we have a solution to the optimal manipulation problem defined in step II, we also have a solution to the p-median problem defined in step I. This will indicate that the optimal group manipulation is at least as hard as the p-median problem, which is known to be NP-Hard.

Let $\mathbf{x}^*_{\mathbf{P}}$ be the optimal solution of the optimal group manipulation problem defined in step

II, our first claim is that $(-b,-b) \in \mathbf{x}^*_{\tilde{\mathbf{P}}}$. This is trivial because a group manipulation requires that no manipulator is worse off (i.e., constraints (5.5) in Definition 5.3), and there must be one facility located at $(-b,-b)$ for every optimal manipulation.

Our next claim is that $\mathbf{x}^*_{\tilde{\mathbf{P}}} \backslash (-b,-b) \in \arg\min_Q \left\{ \sum_{i=1}^{t} \min_{1 \le j \le p} \{ |x_i - z_j| + |y_i - s_j| \} \right\}$ (or $\mathbf{x}^*_{\tilde{\mathbf{P}}} \backslash (-b,-b) \in \arg\min_Q \left\{ \sum_{i=1}^{t} \min_{1 \le j \le p} \{ \sqrt{(x_i - z_j)^2 + (y_i - s_j)^2} \} \right\}$). In other words, if we remove $(-b,-b)$ from the solution vector $\mathbf{x}^*_{\tilde{\mathbf{P}}}$ we get from step II, then the remaining locations constitutes an optimal solution to the p-median problem defined in step I. Suppose by contradiction that this is not true. If we use $D(Q^*)$ to denote the value of the p-median problem induced by a set of points Q^*, this means we have $\sum_{i \in R} c_i(\mathbf{x}^*_{\tilde{\mathbf{P}}} \backslash (-b,-b), t_i) > D(Q^*)$. However, if we consider the set $Q' = \{(-b,-b)\} \cup Q^*$, we would like to show that Q' is a solution to the group manipulation problem with lower cost for the manipulators if b is large enough, contradicting the optimality of $\mathbf{x}^*_{\tilde{\mathbf{P}}}$.

We first show that Q' actually induces a smaller objective value:

$$
\begin{aligned}
\sum_{i \in \tilde{R}} c_i(Q', t_i) &= \sum_{i \in R} c_i(Q', t_i) + \sum_{i \in \tilde{R} \backslash R} c_i(Q', t_i) \\
&= \sum_{i \in R} c_i(Q', t_i) + 0 \\
&= \sum_{i \in R} c_i(Q^*, t_i) \\
&= D(Q^*) < \sum_{i \in R} c_i(\mathbf{x}^*_{\tilde{\mathbf{P}}} \backslash (-b,-b), t_i) \\
&= \sum_{i \in \tilde{R}} c_i(\mathbf{x}^*_{\tilde{\mathbf{P}}}, t_i)
\end{aligned}
$$

where the third equality comes from the fact that if b is large enough, then no manipulator in R will use the facility located at $(-b,-b)$, and as we will see in Lemma 5.2, there is a closer facility in the unit box where everyone is better-off.

We next show that both constraints (5.5) are satisfied. We first give the following lemma:

Lemma 5.2 *There exists a solution point $x \in \mathbf{x}^*_{\tilde{\mathbf{P}}}$, such that x is in the unit box.*

Proof: Suppose by contradiction that there is no solution point in the unit box, then if we randomly pick one location $x = (x^1, x^2)$ from $\mathbf{x}^*_{\tilde{\mathbf{P}}}$. W.l.o.g., we can consider the following 2 cases: (1) x is in the unit box in only one dimension, e.g., $0 \leq x^1 \leq 1$ and $x^2 \geq 1$, and (2) x is not in the unit box in either dimension, e.g., $x^1 \geq 1$ and $x^2 \geq 1$. For the fist case, we can have a group manipulation where all reports between 1 and x^2 on the second dimension are changed to 1, and the induced location vector (which contains $(x^1, 1)$ instead of (x^1, x^2)) is a better group manipulation. We can do the similar thing for the latter case, in which (x^1, x^2) will be replace by $(1, 1)$ and a better group manipulation is found. This contradicts with our assumption that $\mathbf{x}^*_{\tilde{\mathbf{P}}}$ is optimal, completing our proof.

■

By Lemma 5.2, we know that for every manipulation in R, their cost is at most 2 under the L_1-norm or $\sqrt{2}$ under the L_2-norm. So if $b \gg 1$, then each manipulator located at $(-b, -b)$ has a cost of 0, and everyone else has a cost of at most 2 under the L_1-norm or $\sqrt{2}$ under the L_2-norm, satisfying constraints (5.5).

The last step is to show that Q' is a viable location set under the quantile mechanism $f_{\tilde{\mathbf{P}}}$. Recall that $Q' = \{(-b, -b)\} \cup Q^*$ and $\tilde{\mathbf{P}} = \{(0, 0); q\}$. As $\mathbf{P}$ is consistent with Q^*, so if we add a bottom-left point in Q and a corresponding quantile vector in $\mathbf{P}$, then the new quantile matrix $\tilde{\mathbf{P}}$ is also consistent with Q'.

So we have reduced the p-median problem to our optimal group manipulation problem. As the p-median problem is NP-Hard, so is our group manipulation.

■

From the reduction, we can see that the problem remains NP-hard even for two dimensions. Moreover, the NP-hardness result only holds when the number of facilities p is a variable. If p is fixed, the problem is solvable in polynomial time [Boyd and Vandenberghe, 2004].

Finally, we would like to show that the following decision problem which is derived from the optimal group manipulation problem is NP-Complete.

Problem 5.1 *Given a set of sincere agents and manipulators, a quantile mechanism $f_{\mathbf{P}}$, a location vector $\mathbf{x}_{\mathbf{P}}$ chosen under $f_{\mathbf{P}}$, and a positive value h. Is there a group manipulation $\mathbf{t}'_M$, such that $\sum_{i \in M} c_i(f_{\mathbf{P}}(\mathbf{t}'_M, \mathbf{t}_S), t_i) \leq h$, and no manipulator is worse off?*

This problem belongs to NP, because given a joint misreport, we can check in polynomial time if the objective function has a smaller value and if no manipulator is worse off. In addition, the above theorem implies that its optimization counterpart is an NP-Hard. Therefore, this decision problem is NP-Complete.

As with any computational hardness result, while this implies worst-case instances may be difficult to solve, it does not mean that many (or even most) of the instances that occur in practice can't be solved effectively. In the remainder of this section, we describe formulations of the optimal group manipulation problem for multi-FLPs as integer programs (linear and SOCP) that may be practically solvable. Our formulations are quite compact, and combined with the empirical evaluation provided in Section 5.4, suggest that optimal group manipulations can be found reasonably quickly, the NP-hardness of the problem notwithstanding.

5.3.2 MILP Formulation under the L_1-norm

In this section, we describe our mixed integer linear programming (MILP) formulation of optimal group manipulation in multi-FLPs under the L_1-norm.

The following proposition is the analogous to Proposition 5.1 for single-FLPs.

Proposition 5.2 *Let $\mathbf{t}'_M$ be a group manipulation and $\mathbf{x} = \{(x_1^1, \ldots, x_1^m); \ldots; (x_q^1, \ldots, x_q^m)\}$ be a viable location vector implemented by $\mathbf{t}'_M$ (assuming reports from agents in $S = N \backslash M$ remain fixed to be $\mathbf{t}_S$). Let $\mathbf{X}^k = \{x_1^k, \ldots, x_q^k\}$ denote the set of coordinates of these facilities in the kth dimension. Then there exists a group manipulation $\mathbf{t}^*_M$ that implements $\mathbf{x}$, where $t_i^* \in \prod_{k \leq m} \mathbf{X}^k, \forall i \in M$.*

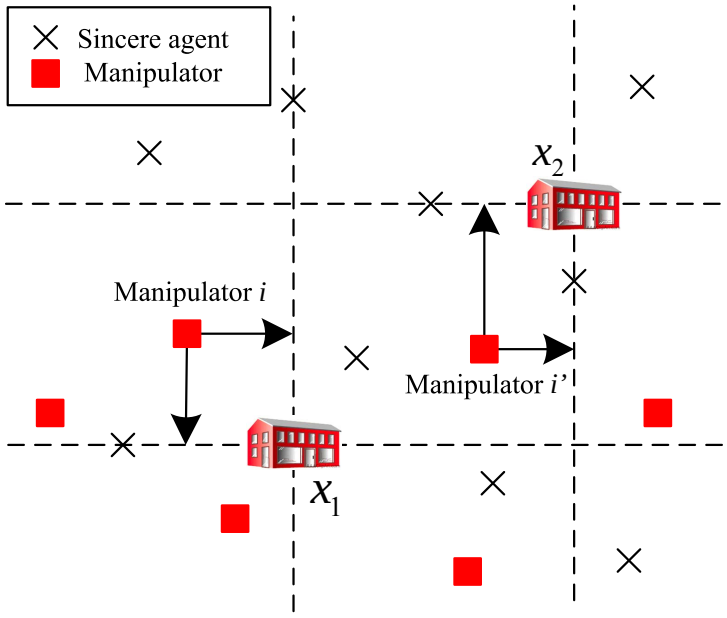


Figure 5.6: The probability that a random drawn point falls into each region.

Proof: We first prove the proposition for a two-dimensional, two-facility location problem, and then show how the analysis can be generalized to any number of dimension and facility.

Consider the group manipulation shown in Figure 5.6, where the optimized new location vector is $\mathbf{x} = \{x_1, x_2\}$. While we can think that $\mathbf{x}$ has partitioned the space into 9 small grids (assuming $x_1^1 \neq x_2^1$ and $x_1^2 \neq x_2^2$), which are separated by (x_1^1, x_1^2), (x_1^1, x_2^2), (x_2^1, x_1^2) and (x_2^1, x_2^2). For each manipulator $i \in M$, we can move her misreports towards the closest of these four locations without passing through it, creating another group manipulation GM' from the original group manipulation GM. Similar to the single-facility case, the way that quantile mechanism works will guarantee that the new location vector $\mathbf{x}$ remains unchanged.

For the case of multi-facility location in multi-dimensional space, we can view the intersection positions induced by different facilities on different dimensions as partitioning the space into small boxes. For each manipulator, we can change her misreport to one corner of the box which her misreport lies in without changing the implemented viable location vector. As each corner corresponds an element in $\prod_{k \leq m} \mathbf{X}^k$, this completes our proof. ■

In other words, we can assume w.l.o.g. that manipulators misreports are drawn from the "intersection positions" in different dimensions induced by the different facilities contained

within some viable location vector. Of course, the precise misreports at these intersection positions must be coordinated to guarantee that the resulting location vector $\mathbf{x}$ is itself viable (a point we return to below).

Let $\mathbf{x} = \{(x_1^1, \ldots, x_1^m), \ldots, (x_q^1, \ldots, x_q^m)\}$ represent the location vector to be optimized (i.e., induced by the manipulation). Let c_i be a continuous variable denoting the cost of manipulator i given outcome $\mathbf{x}$. Finally, let c_{ij} be the cost of manipulator i w.r.t. facility j, and I_{ij} be a 0-1 variable whose value is 1 iff the closest facility for manipulator i is j. We can formulate the objective function Eq. (5.4), and the constraints Eq. (5.5), as follows:[3]

$$\min_{\mathbf{x}\in(\mathbb{R}^m)^q} \sum_{i\in M} c_i \tag{5.21}$$

$$\textit{s.t.} \qquad c_i = \sum_{j\leq q} I_{ij} \cdot c_{ij}, \qquad \forall i \in M \tag{5.22}$$

$$\sum_{j\leq q} I_{ij} = 1, \qquad \forall i \in M \tag{5.23}$$

$$I_{ij} \in \{0, 1\}, \qquad \forall i \in M, \forall j \leq q \tag{5.24}$$

$$0 \leq c_i \leq u_i, \qquad \forall i \in M, \forall j \leq q \tag{5.25}$$

$$c_{ij} \geq 0, \qquad \forall i \in M, \forall j \leq q \tag{5.26}$$

Here $u_i = c_i(\mathbf{x_P}, t_i)$ is the cost of manipulator i under a truthful report $\mathbf{t}_M$ by the manipulators.

Constraints (5.22)-(5.24) require that each manipulator be closest to (or more precisely, decide to use) only one of the facilities, and ensure her cost for that facility is minimized over all facilities. Constraint (5.25) ensures that no manipulator is worse off w.r.t. truthful reporting. Since both I_{ij} and c_{ij} are variables in constraint (5.22), we must linearize these quadratic terms by introducing additional variables. Let O_{ij} represent the product of I_{ij} and c_{ij}. We can then

[3]For convenience, we refer to the formulations for single-FLPs, which can be generalized to multi-FLPs easily.

replace the constraint (5.22) by

$$c_i = \sum_{j \leq q} O_{ij}, \qquad \forall i \in M \tag{5.27}$$

$$O_{ij} \geq c_{ij} + (I_{ij} - 1)U, \quad \forall i \in M, \forall j \leq q \tag{5.28}$$

$$O_{ij} \geq 0, \qquad \forall i \in M, \forall j \leq q \tag{5.29}$$

where U is any upper bound on manipulator cost.

These constraints guarantee that when $I_{ij} = 1$, O_{ij} is equal to c_{ij} (by constraints (5.29)), and when $I_{ij} = 0$, O_{ij} is equal to 0 (by constraints (5.28) and (5.29) together). Note that as the objective function is a minimization problem, we only need lower bounds on O_{ij} here.

We also need constraints similar to constraints (5.11)-(5.13) to bound each manipulator's cost with respect to each facility c_{ij}. Let D_{ij}^k be an upper bound on the distance between manipulator i and facility j in the kth dimension. We have:

$$-D_{ij}^k \leq t_i^k - x_j^k \leq D_{ij}^k, \quad \forall i \in M, \forall j \leq q, \forall k \leq m \tag{5.30}$$

$$D_{ij}^k \geq 0, \qquad \forall i \in M, \forall j \leq q, \forall k \leq m \tag{5.31}$$

$$c_{ij} = \sum_{k \leq m} D_{ij}^k, \qquad \forall i \in M, \forall j \leq q \tag{5.32}$$

Finally, we must ensure that location vector $\mathbf{x}$ is viable. Let

$$\perp_j^k = \min\{z \in \mathbb{Z}^+ : z + |M| \geq p_j^k \cdot n\}, \text{ and}$$
$$\top_j^k = \max\{z \in \mathbb{Z}^+ : |S| + |M| - z \geq (1 - p_j^k) \cdot n\},$$

and $\tilde{x}_S^k = \{\tilde{x}_1, \ldots, \tilde{x}_{|S|}\}$ be the ordered coordinates of the reports of the sincere agents in S in the kth dimension. We break $[\tilde{x}^k_{\perp_j^k}, \tilde{x}^k_{\top_j^k}]$ into several (ordered) close and open intervals: $[\tilde{x}^k_{\perp_j^k}, \tilde{x}^k_{\perp_j^k}], (\tilde{x}^k_{\perp_j^k}, \tilde{x}^k_{\perp_j^k+1}), \ldots, (\tilde{x}^k_{\top_j^k-1}, \tilde{x}^k_{\top_j^k}), [\tilde{x}^k_{\top_j^k}, \tilde{x}^k_{\top_j^k}]$ (see Figure 5.7 for an illustration). Let Δ_j^k index these intervals ($0 \leq \Delta_j^k < 2|M| + 1$), and let $I_{\Delta_j^k}$ be an indicator variable whose

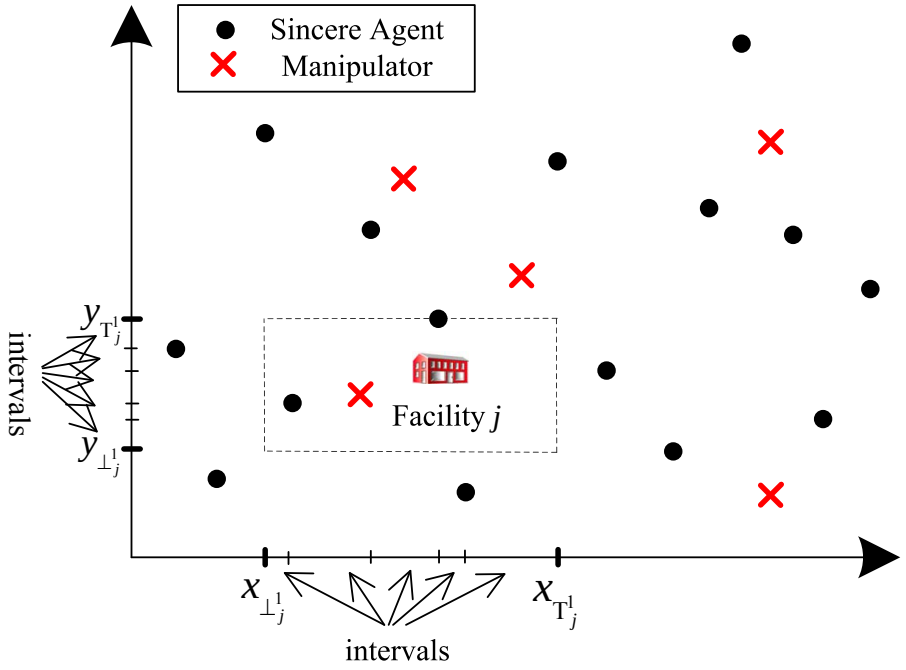


Figure 5.7: For each facility in each dimension, the boundaries are further split into small intervals, each bounded by one/two sincere agents.

value is 1 iff the coordinate of facility j is contained in the Δ_j^kth interval in the kth dimension. We then have:

$$\sum_{\Delta_j^k} I_{\Delta_j^k} = 1, \qquad \forall j \leq q, \forall k \leq m \tag{5.33}$$

$$\sum_{\Delta_j^k} I_{\Delta_j^k} \tilde{x}_l^k \leq x_j^k \leq \sum_{\Delta_j^k} I_{\Delta_j^k} \tilde{x}_r^k, \quad \forall j \leq q, \forall k \leq m \tag{5.34}$$

$$I_{\Delta_j^k} \in \{0, 1\}, \qquad \forall j \leq q, \forall k \leq m \tag{5.35}$$

where

$$l = \perp_j^k + \lfloor \Delta_j^k / 2 \rfloor \tilde{x}_s^k,$$

$$r = \perp_j^k + \lfloor (\Delta_j^k + 1)/2 \rfloor.$$

Constraints (5.33) and (5.35) ensure that each facility is located within only one interval in

each dimension, while constraint (5.34) defines the upper and lower bounds for that interval. For each interval, we can pre-compute the number of sincere agents that lie to the left of (below) and right of (above) it (including equality) in each dimension k, which we denote by $L_{\Delta_j^k}$ and $R_{\Delta_j^k}$, respectively.[4] We also introduce another indicator variable T_{ij}^k whose value is 1 iff manipulator i misreports the location of facility j in the kth dimension (this binary variable can be relaxed, since all terms in (5.36) and (5.37) are integral). Given a quantile matrix $\mathbf{P} = \{(p_1^1, \ldots, p_1^m); \ldots; (p_q^1, \ldots, p_q^m)\}$, the location vector $\mathbf{x}$ to be optimized is viable under $f_{\mathbf{P}}$ if the following constraints are satisfied:

$$\sum_{\Delta_j^k} I_{\Delta_j^k} L_{\Delta_j^k} + \sum_{j' \leq_{\mathbf{P}} j} \sum_{i \in M} T_{ij'}^k \geq n q_j^k, \qquad \forall j \leq q, \forall k \leq m \tag{5.36}$$

$$\sum_{\Delta_j^k} I_{\Delta_j^k} R_{\Delta_j^k} + \sum_{j' \geq_{\mathbf{P}} j} \sum_{i \in M} T_{ij'}^k \geq n(1 - q_j^k), \forall j \leq q, \forall k \leq m \tag{5.37}$$

$$\sum_{j \leq q} T_{ij}^k = 1, \qquad \forall i \in M, \forall k \leq m \tag{5.38}$$

$$T_{ij}^k \in [0, 1], \qquad \forall i \in M, \forall j \leq q, \forall k \leq m \tag{5.39}$$

The LHS of constraint (5.36) indicates the total number of sincere agents (the first term) and manipulators (the second term) to the left of (or at) facility j in the kth dimension, where $j' \leq_{\mathbf{P}} j$ denotes the facility j' to the left of j in the kth dimension (i.e., $p_{j'}^k \leq p_j^k$). According to $f_{\mathbf{P}}$, this number should be greater than or equal to nq_j^k. Constraints (5.37) are similar, but used to count from the right. Constraints (5.38) and (5.39) ensure that each manipulator reports the location of one facility on each dimension.

To summarize, we have the following result:

Theorem 5.4 (MILP for optimal group manipulation in multi-FLPs) *The optimal group ma-*

[4]The reason that we have closed and open intervals is that we have to compute $L_{\Delta_j^k}$ and $R_{\Delta_j^k}$ differently. For a closed interval, the sincere agent whose coordinate coincides with the interval should be counted twice when computing both $L_{\Delta_j^k}$ and $R_{\Delta_j^k}$, so we should have $L_{\Delta_j^k} + R_{\Delta_j^k} \geq |S| + 1$. However, while for an open interval, no sincere agent is double counted, so $L_{\Delta_j^k} + R_{\Delta_j^k} = |S|$.

nipulation for multi-facility location under L_1-norm can be formulated as a mixed integer linear program with objective function (5.21) and constraints (5.23)-(5.39).

Proof: Figure 5.8 provides a snapshot of the whole MILP. The objective function (5.21) minimizes the sum of the costs of all manipulators. Constraints (5.23)-(5.32) guarantee that the new location vector is mutually beneficial and each manipulator uses her closest facility. Finally, constraints (5.33)-(5.39) ensure that the optimized location vector is viable.

Next, we analyse the number of variables. As we have mentioned, there are at most $2|M|+1$ intervals for each facility j on each dimension k, so the total number of variables for $I_{\Delta_j^k}$ is at most $(2|M|+1)qm$. We also need another $q|M|$ variables to denote the closest facility for each manipulator, so the total number of binary variables is $2(|M|+1)qm$ or $O(qm|M|)$, where m is the number of dimensions, q is the number of facilities and $|M|$ is number of manipulators. The number of continuous variables is $(1+2q+2qm)|M|+qm$ or $O(qm|M|)$ (where $|M|$ comes from c_is, $2q|M|$ comes from c_{ij}s and O_{ij}s, $2qm|M|$ comes from D_{ij}^ks and T_{ij}^ks, and the rest qm comes from $\mathbf{x}$). ■

The final step is to construct a misreport profile $\mathbf{t}'_M$ that implements the location vector optimized above. By Proposition 5.2, we can assume each manipulator reports one of the intersection positions of the target outcome vector. So we arbitrarily choose a set of manipulators of size exactly $\sum_i T_{ij}^k$ for each target facility j in each dimension k. This completes the construction of the misreport that implements the target location vector.

5.3.3 MISOCP Formulation under the L_2-norm

When optimizing misreports for multi-FLPs under the L_2-norm, we can use an approach similar to that used in the single-facility case, and formulate the optimal manipulation as an mixed-integer SOCP (MISOCP). We need only modify constraints (5.32) as follows:

$$(c_{ij})^2 \geq \sum_{k\leq m} \left(D_{ij}^k\right)^2, \quad \forall i \in M, \forall j \leq q \tag{5.40}$$

$$
\begin{aligned}
&\min_{\mathbf{x}\in(\mathbb{R}^m)^q} \sum_{i\in M} c_i \\
\textit{s.t.} \quad & c_i = \sum_{j\le q} O_{ij}, && \forall i \in M \\
& O_{ij} \ge c_{ij} + (I_{ij} - 1)U, && \forall i \in M, \forall j \le q \\
& O_{ij} \ge 0, && \forall i \in M, \forall j \le q \\
& \sum_{j\le q} I_{ij} = 1, && \forall i \in M \\
& I_{ij} \in \{0,1\}, && \forall i \in M, \forall j \le q \\
& 0 \le c_i \le u_i, && \forall i \in M, \forall j \le q \\
& c_{ij} \ge 0, && \forall i \in M, \forall j \le q \\
& -D^k_{ij} \le t^k_i - x^k_j \le D^k_{ij}, && \forall i \in M, \forall j \le q, \forall k \le m \\
& D^k_{ij} \ge 0, && \forall i \in M, \forall j \le q, \forall k \le m \\
& c_{ij} = \sum_{k\le m} D^k_{ij}, && \forall i \in M, \forall j \le q \\
& \sum_{\Delta^k_j} I_{\Delta^k_j} = 1, && \forall j \le q, \forall k \le m \\
& \sum_{\Delta^k_j} I_{\Delta^k_j} \tilde{x}^k_l \le x^k_j \le \sum_{\Delta^k_j} I_{\Delta^k_j} \tilde{x}^k_r, && \forall j \le q, \forall k \le m \\
& I_{\Delta^k_j} \in \{0,1\}, && \forall j \le q, \forall k \le m \\
& \sum_{\Delta^k_j} I_{\Delta^k_j} L_{\Delta^k_j} + \sum_{j' \le_{\mathbf{P}} j} \sum_{i\in M} T^k_{ij'} \ge n q^k_j, && \forall j \le q, \forall k \le m \\
& \sum_{\Delta^k_j} I_{\Delta^k_j} R_{\Delta^k_j} + \sum_{j' \ge_{\mathbf{P}} j} \sum_{i\in M} T^k_{ij'} \ge n(1 - q^k_j), && \forall j \le q, \forall k \le m \\
& \sum_{j\le q} T^k_{ij} = 1, && \forall i \in M, \forall k \le m \\
& T^k_{ij} \in [0,1], && \forall i \in M, \forall j \le q, \forall k \le m
\end{aligned}
$$

Figure 5.8: The complete second-order cone program of optimal group manipulation for single facility location problem under the L_1-norm.

Using this we obtain the following result:

Theorem 5.5 (MISOCP for optimal group manipulation in multi-FLPs) *The optimal group manipulation problem for multi-FLPs under the L_2-norm can be formulated as a mixed integer second-order cone program, with objective function (5.21), and constraints (5.23)-(5.31) and (5.33)-(5.40).*

Proof: Figure 5.9 provides a snapshot of the whole MISOCP. The objective function (5.21) minimizes the sum of the costs of all manipulators. Constraints (5.23)-(5.31) and (5.40) guarantee that the new location vector is mutually beneficial and each manipulator uses her closest facility. Finally, constraints (5.33)-(5.39) ensure that the optimized location vector is viable. The number of binary and continuous variables are both $O(qm|M|)$. ■

5.4 Empirical Evaluation

In this section, we evaluate the efficiency of the formulations in outlined above. Since the optimal manipulation problem for single-FLPs is computationally tractable (polynomial in the input size), we provide empirical results only for multi-facility problems here, testing the efficiency of the MILP and MISOCP formulations described in Section 5.3.

We test two problems. The first is a two-dimensional, two-facility location problem under the L_2-norm, where the quantile matrix used is $\mathbf{P} = \{0.3, 0.4; 0.8, 0.7\}$. The second is a four-dimensional, three-facility location problem under the L_1-norm, where the quantile matrix used is $\mathbf{P} = \{0.1, 0.6, 0.4, 0.9; 0.4, 0.2, 0.8, 0.6; 0.7, 0.8, 0.3, 0.4\}$. For both problems, we vary the number of sincere agents $|S| \in \{100, 200, 500\}$, and the number of manipulators $|M| \in \{5, 10, 20, 50, 100, 200\}$. We randomly generated 100 problems instances for each parameter setting in which the peaks of both the sincere agents and the manipulators are randomly drawn from the same data set (data sets are explained in detail below). We compute the average execution time of our MILP/MISOCP models, and the probability of manipulation (i.e, the proportion of the 100 instances in which a viable manipulation exists for the randomly chosen

$$
\begin{aligned}
&\min_{\mathbf{x}\in(\mathbb{R}^m)^q} \sum_{i\in M} c_i \\
\textit{s.t.} \quad & c_i = \sum_{j\le q} O_{ij}, && \forall i \in M \\
& O_{ij} \ge c_{ij} + (I_{ij} - 1)U, && \forall i \in M, \forall j \le q \\
& O_{ij} \ge 0, && \forall i \in M, \forall j \le q \\
& \sum_{j\le q} I_{ij} = 1, && \forall i \in M \\
& I_{ij} \in \{0,1\}, && \forall i \in M, \forall j \le q \\
& 0 \le c_i \le u_i, && \forall i \in M, \forall j \le q \\
& c_{ij} \ge 0, && \forall i \in M, \forall j \le q \\
& -D^k_{ij} \le t^k_i - x^k_j \le D^k_{ij}, && \forall i \in M, \forall j \le q, \forall k \le m \\
& D^k_{ij} \ge 0, && \forall i \in M, \forall j \le q, \forall k \le m \\
& c^2_{ij} \ge \sum_{k\le m} \left(D^k_{ij}\right)^2, && \forall i \in M, \forall j \le q \\
& \sum_{\Delta^k_j} I_{\Delta^k_j} = 1, && \forall j \le q, \forall k \le m \\
& \sum_{\Delta^k_j} I_{\Delta^k_j} \tilde{x}^k_l \le x^k_j \le \sum_{\Delta^k_j} I_{\Delta^k_j} \tilde{x}^k_r, && \forall j \le q, \forall k \le m \\
& I_{\Delta^k_j} \in \{0,1\}, && \forall j \le q, \forall k \le m \\
& \sum_{\Delta^k_j} I_{\Delta^k_j} L_{\Delta^k_j} + \sum_{j'\le_{\mathbf{P}} j} \sum_{i\in M} T^k_{ij'} \ge n q^k_j, && \forall j \le q, \forall k \le m \\
& \sum_{\Delta^k_j} I_{\Delta^k_j} R_{\Delta^k_j} + \sum_{j'\ge_{\mathbf{P}} j} \sum_{i\in M} T^k_{ij'} \ge n(1 - q^k_j), && \forall j \le q, \forall k \le m \\
& \sum_{j\le q} T^k_{ij} = 1, && \forall i \in M, \forall k \le m \\
& T^k_{ij} \in [0,1], && \forall i \in M, \forall j \le q, \forall k \le m
\end{aligned}
$$

Figure 5.9: The complete second-order cone program of optimal group manipulation for single facility location problem under the L_2-norm.

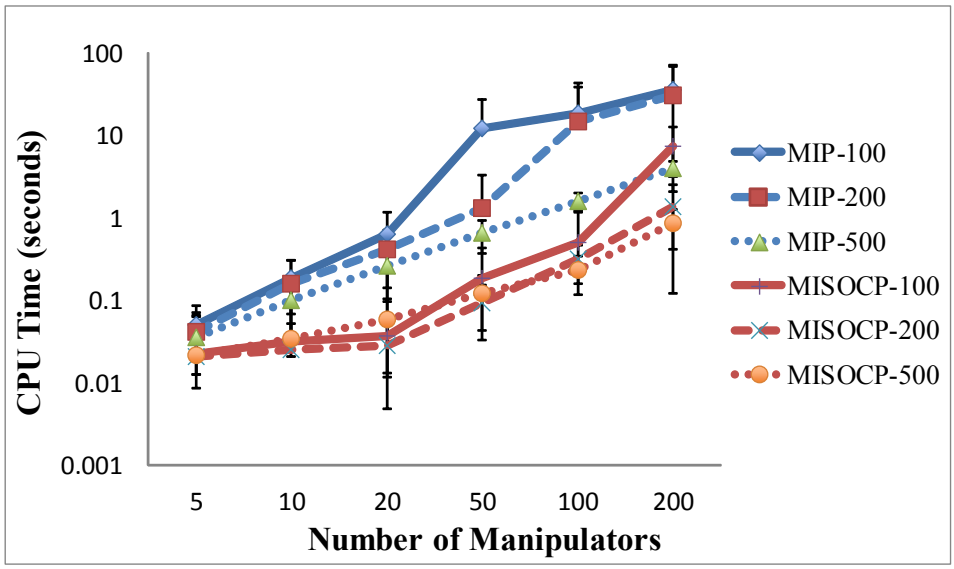


Figure 5.10: Time to solve for an optimal manipulation (both axes are log-scale). The error bars show the standard deviations.

manipulators).

For the two-dimensional problem, we use preference data from the Dublin west constituency in the 2002 Irish General Election. As in Section 4.4, we use the results from fitting the election data to a two-dimensional spatial model and estimating the voter and candidate positions in the underlying latent space.[5] For the four-dimensional problem, we use a synthetic data set in which the peaks of both the sincere agents and the manipulators are randomly generated from a uniform distribution on the unit cube.

For each instance, the MILP/MISOCP is solved using CPLEX (version 12.51), on a machine with a quad-core CPU (2.9Gz/core) and 8GB memory. Figure 5.10 shows the average computation time required to find the optimal group manipulation (or showing that no group manipulation exists) for both formulations. We see that our formulations admit very effective solution—for small problems, the optimal group manipulation can be found in less than 1 second; even for reasonably large problems, such as the four-dimensional, three-facility problem with 100 sincere agents and 200 manipulators, the optimal manipulation is found in 35.47 seconds (on average). The standard deviation also indicates that the performance of our formu-

[5]Please see Section 6.4 for details.

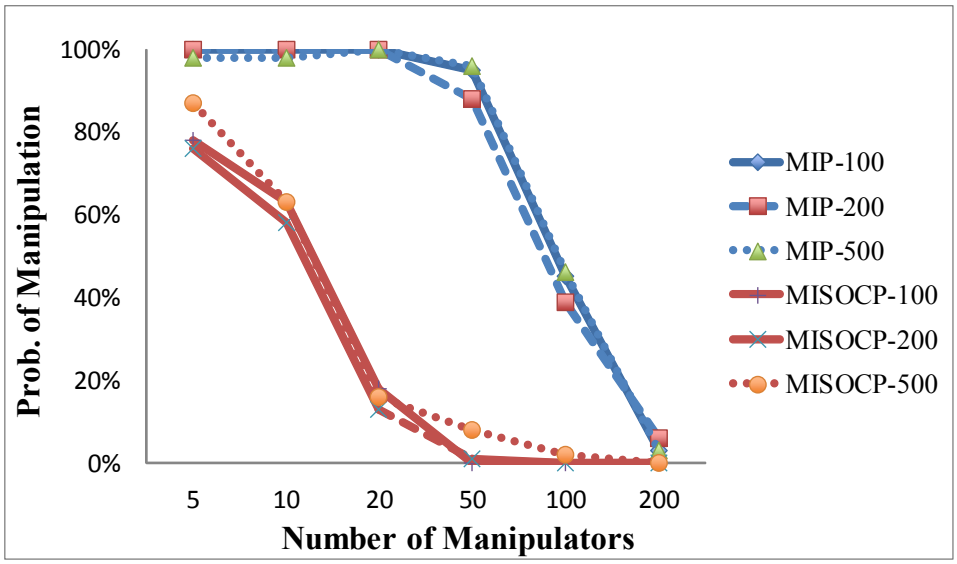


Figure 5.11: Probability to find an optimal manipulation (both axes are log-scale).

lations is very stable—the coefficient of variation (ratio of the standard deviation to the mean) ranges from 0.21 to 1.58 for the 2D problem, and from 0.45 to 1.92 for the 4D problem.

We also show the probability of manipulation for both problems in Figure 5.11. For 2D problems, the probability of manipulation decreases from around 80% (for 5 manipulators) to 20% (for 20 manipulators) and finally to 0 (for 200 manipulators), indicating that it is very hard for a randomly selected set of manipulators to find a viable manipulation; for 4D problems, the probability remains high (close to 1) even with 20 manipulators then decreases with larger sets of manipulators. This is not surprising since, as the number of manipulators get larger, it is harder for them to find a mutually beneficial misreport. The higher probability for 4D problems is due to the fact that we are placing three facilities rather than two, increasing the potential of viable manipulations.

5.5 Conclusion

In this chapter, we have addressed the optimal group manipulation problem in multi-dimensional, multi-facility location problems. Specifically, we analyzed the computational problems of ma-

nipulating quantile mechanisms. We showed that optimal manipulation for single-facility problems can be formulated as an LP or SOCP, under the L_1- and L_2-norm, respectively, and thus can be solved in polynomial time. By contrast, the optimal manipulation problem for multi-facility problems is NP-hard, but can be formulated as an ILP or MISOCP under the L_1- and L_2-norm, respectively. Our empirical evaluation demonstrates that our MILPs/MISOCPs formulation for multi-FLPs scales well, despite the NP-hardness of the problem.

The results in this chapter suggests a number of interesting future directions. First, more empirical results would be helpful in understanding the practical ease or difficulty of group manipulation, as well as the probability of manipulation, the potential gain of manipulators, and the impact of manipulation on social welfare. Second, other objectives for the manipulating coalition (e.g., minimizing the maximum cost), and mechanisms with other cost functions are also of interest. Finally, some research has shown that agent preferences are often not exactly single-peaked, but may be approximately so under some forms of approximation (as we will see in Chapter 6). The theoretical and empirical evaluation of group manipulation in such settings would be extremely valuable. We will discuss more future directions in Chapter 8.

Chapter 6

Multi-dimensional Single-peakedness and its Approximation

6.1 Introduction

In the preceding chapters, we have focused on models in which agent preferences are assumed to be single-peaked. In such a setting, mechanisms like the median and generalized median mechanisms provide strategy-proof guarantees. In addition, they also ease communication and computational demands, since agents need only reveal their peak preferences.

While conceptually attractive, single-peakedness is a very strong assumption, one unlikely to fully hold in many realistic settings [Conitzer, 2009, Escoffier et al., 2008]. Consider an elections with thousands of voters and more than a handful of candidates. While each voter has her own opinion about how candidates can be ordered according to some particular issue (e.g., left-right spectrum, foreign policy, economic policy, etc), it is very unlikely that all of them agree upon a same ordering.

However, one might hope that preferences are *approximately single-peaked*, and thereby retain some of the advantages mentioned above. To this end, recent research has begun to investigate computational methods to test single-peakedness [Escoffier et al., 2008], and various

forms of approximation (e.g., by deleting outlier candidates, clustering candidates, deleting voters, or adding additional axes) [Escoffier et al., 2008, Faliszewski et al., 2011, Erdélyi et al., 2012, Galand et al., 2012]. These techniques, however, have focused on *one-dimensional (1D)* preferences, and have not been tested empirically to determine if these approximations *can explain observed preferences* in, say, real-world elections.

We address this issue in two ways. First, we test single-peaked consistency, and several forms of approximation (in isolation and in combination) on two election data sets to see if these approximations have any empirical explanatory power. To do so, we develop a *branch-and-bound (BB)* algorithm to find the *best (1D) axis* given a preference profile, i.e., the ordering of candidates for which the greatest number of voters are single-peaked (as will be defined formally later). The algorithm is easily extended to support various forms of approximation. While this *best-axis problem* is computationally difficult, our method works well in practice. We show that *none* of the forms of approximation proposed in the literature come close to explaining voter preferences in these election data sets: the best axis explains under 2.9% of voter preferences in one case and under 0.4% in the other; and even aggressive approximation improves this to only 50% and 25%, respectively. To address this problem, we extend our algorithm to support *multi-dimensional single-peaked consistency*. Focusing on the *two-dimensional (2D)* case, we show that exact 2D-single-peakedness explains the datasets much better[1]—without approximation, the best axis set explains over 38% and 26% of voter preferences respectively; and with a very small degree of approximation, the 2D model explains almost all voters. Apart from our algorithmic developments, our findings suggest a focus on multi-dimensional rather than 1D models can greatly enhance the applicability of single-peaked models in practice.

The rich literature on *spatial models* for voter or consumer choice bears a strong relationship with single-peaked preferences as well [Hotelling, 1929, Hinich, 1978, Poole and

[1]An interesting question is whether one can get a better fitting result with higher number of dimensions. The answer is yes, however, we also have to consider the arising computational issues due to the increase of the number of dimensions and the over fitting problem. It turns out that a two-dimensional model is enough to explain voter preferences. The same arguments apply for the spatial model discussed later.

Rosenthal, 1985]. Spatial models explain voter choice by estimating (from data) distances between voters and candidates, and typically using some form of probabilistic choice based on these distances [Bradley and Terry, 1952, Luce, 1959, Shepard, 1959] (see Section 2.4.3). While the model is more restrictive than multi-dimensional single-peakedness in some senses, stochastic choice allows for accommodation of "misorderings," much like approximations in single-peaked models. Spatial models are typically used to explain choice data rather than full preference rankings (see [Gormley and Murphy, 2007] for an exception). We study a spatial model for rank data. In particular, we consider a stochastic choice models, namely Plackett-Luce, and propose an alternating optimization algorithm that optimizes voter and candidate positions alternatively before converge. We also conduct an empirical study on two data sets from the Irish General Election 2002. The results show that, the two- dimensional fit is much better than then single-dimensional fit. It is also suggested that party policies plays an important role in the electorates view of candidates.

6.2 A One-Dimensional Branch and Bound Algorithm

We first define the *best axis problem*, and present our branch and bound algorithm for solving this problem in a one-dimensional space. Following the notations from Section 2.3 and 2.4, let $N = \{1, \ldots, n\}$ be a set of agents (or *voters*) and $C = \{1, \ldots, c\}$ be a set of c options (or *candidates*). Each voter $i \in N$ has a preference (or strict total order) $\succ_i$ over C, with $c_1 \succ_i c_2$ meaning i prefers candidate c_1 to c_2. A *preference profile* $\succ = \{\succ_1, \ldots, \succ_n\}$ reflects the joint preferences of all voters.

a one-dimensional axis A is any strict ordering $<_A$ of the candidates in C. Intuitively, it represents an ordering of candidates relative to some salient quality, e.g., on the left-right political spectrum. We define the *best axis* problem as follows:

Definition 6.1 (Best axis problem) *Given a preference profile $\succ$ and a one-dimensional axis, we define $s(A)$ as the score of A, i.e., the number of voters whose preferences $\succ_i$ are single-*

peaked with respect to A. The best axis problem is to find a single axis $A^* \in \arg\max_A s(A)$.

In other words, the best axis problems finds a single axis that explains the preferences of the greatest number of voters (i.e., renders $\succ$ single-peaked). Note that this problem is harder than the problem of determining single-peaked consistency (as in Definition 2.30) introduced in Section 2.4. For example, the algorithm by Escoffier et al. [2011] finds an axis A that renders the profile (perfectly) single-peaked if one exists; but if no such A exists, it does not find a best axis that fits the greatest number of voters. On the other hand, if we have an algorithm that can solve the best axis problem and return an axis A, then we use this to determine if the preference profile is single-peaked consistent with respect to A if $s(A) = A$. Also the best axis problem is just the optimization variant of the k-maverick problem (see Definition 2.31), which is NP-complete [Erdélyi et al., 2012]. We develop a branch-and-bound algorithm for this problem, and use this as a building block for generating additional axes and for supporting approximations like k-LCD and k-AA (see Definition 2.32 and 2.33).

6.2.1 The Algorithm

Our branch-and-bound algorithm, *1D-SPBB*, is specified in Algorithm 3 and adopts ideas from the single-peaked consistency method of Escoffier *et al.* [2008]. Each node in the search tree is labelled by a *partial axis* in which a subset of the candidates are ordered.

Definition 6.2 (Partial axis) *A partial axis $A_{p,q}$ is an ordering of a subset of candidates of the form $A_{p,q} = \{(c_1, \ldots, c_p), \ldots, (c_q, \ldots, c_z)\}$, where p candidates $c_1 <_{A_{p,q}} \ldots <_{A_{p,q}} c_p$ are ordered on the left of the axis and $z - q + 1$ candidates $c_q <_{A_{p,q}} \ldots <_{A_{p,q}} c_z$ are ordered on the right.*

We also use $C_r = C \backslash \{c_1, \ldots, c_p, c_q, \ldots, c_z\}$ to denote the remaining candidates that are not yet ordered on the axis. If $C_r = \emptyset$, we say that partial axis is complete.

An *extension* of partial axis $A_{p,q}$ is any complete axis A that retains the two sub-orderings and completes the ordering by placing the remaining candidates between the two in some

Algorithm 3 1D-SPBB($A_{0,z+1} \leftarrow \emptyset$, $C_r \leftarrow C$, $lb^* \leftarrow 0$)

```
1:  while There exists an unchecked axis do
2:      Pick two candidates c' and c'' from C_r
3:      Build the axis A_{p+1,q-1} from A_{p,q} by locating c' at
        the position of p + 1 and c'' at q − 1
4:      if C_r \ {c', c''} = ∅ then
5:          Mark the complete axis A_{p+1,q-1} as checked
6:          Compute the score s(A_{p+1,q-1})
7:          if s(A_{p+1,q-1}) > lb* then
8:              lb* ← s(A_{p+1,q-1}) and A* ← A_{p+1,q-1}
9:      else
10:         Compute the upper bound ub for A_{p+1,q-1}
11:         if ub > lb* then
12:             1D-SPBB(A_{p+1,q-1}, C_r \ {c', c''}, lb*)
13:         else
14:             Mark the whole branch as checked
15:             Return
```

fashion. Let $E(A_{p,q})$ be the set of extensions of $A_{p,q}$. We say a voter i is *consistent* with $A_{p,q}$ if $\succ_i$ is single-peaked with respect to some $A \in E(A_{p,q})$. 1D-SPBB also maintains, at each node, the list of voters who are consistent with that node's partial axis.

The algorithm starts with an empty axis and extends it from the "outside in." At each step, 1D-SPBB branches by placing two candidates in C_r at positions $p+1$ and $q-1$ of a partial axis $A_{p,q}$ to form a more complete axis $A_{p+1,q-1}$. It then computes the corresponding consistent voters in the preference profile. The number of consistent voters provides a *upper bound* on the score $s(A)$ of any $A \in E(A_{p+1,q-1})$. If the axis is complete, this gives us the exact score $s(A)$ of this axis, and a *lower bound* on $s(A^*)$. In typical fashion, 1D-SPBB maintains a global lower bound lb^*, corresponding to the score of the best complete axis A^* found so far. It cuts the search for extensions of a partial axis $A_{p,q}$ when the upper bound on $A_{p,q}$ falls below lb^*; and when it terminates, the best axis A^* is the optimal axis.

We now consider several key steps in the algorithm that ensure its practicality despite the theoretical hardness of the problem. First, note that axes of the form $A_{1,z}$ at the first level of the search tree fix only the two extreme points of the axis. Symmetry means that we need not consider any axis with c' at the leftmost position and c at the rightmost, if we have already

Algorithm 4 Compute Score or Upper Bound of (Partial) Axis A

```
1: V ← {} %Set of consistent voters%
2: for agent i ∈ N do
3:    consistent ← TRUE
4:    l ← 1, r ← n %Left and right pointers%
5:    for t from z to 1 do
6:       if A[l] = ≻_{i,t} or unplaced candidate then
7:          l ← l + 1
8:       else if A[r] = ≻_{i,t} or unplaced candidate then
9:          r ← r − 1
10:      else
11:         consistent ← FALSE
12:         break
13:   if consistent = TRUE then
14:      V ← V ∪ {i}
```

expanded the partial axis with c leftmost and c' rightmost. This reduces the search tree size by a factor of two, improving efficiency.

A critical component of 1D-SPBB is the identification of consistent voters given a partial (or complete) axis. Given A, Algorithm 4 computes an upper bound on the score of any $\tilde{A} \in E(A)$; and if A is complete, it will return $s(A)$. We let $A[j]$ denote the candidate at the jth position of A and $\succ_{i,t}$ the candidate ranked tth in voter i's preference $\succ_i$. The algorithm is based on that of Escoffier *et al.* [2008] for testing single-peaked consistency, exploiting the fact that candidates ranked last in any $\succ_i$ must lie at the extreme ends of the axis.

Since Algorithm 4 will be called frequently by 1D-SPBB (see Algorithm 3), its running time should be slight. Fortunately, it is easy to see that this is the case:

Theorem 6.1 *Algorithm 4 has a running time of* $O(nc)$.[2]

Proof: The algorithm checks, for each agent i, that whether the currently ranked last candidate in $\succ_i$ is at the extreme location on the axis (or an unplaced candidate). As there are a total number of n agents and c candidates, the total running time is $O(nc)$. ■

[2]Note that the running time of $O(nc)$ is based on the fact that the computation is done on a single processor. However, it can be accelerated by sharding the voters/agents votes across multiple processors, which has a running time of $O(c)$.

Good heuristics for selecting branches (i.e., partial axes to expand) can have a dramatic impact on any branch-and-bound algorithm: the ability to increase our lower bound quickly can significantly impact the degree of pruning. Our current heuristic simply expands nodes in descending order of their upper bounds, in the hope that a partial axis with a large upper bound will have some completion with a high score, thereby improving our global lower bound. If additional domain-specific information is available, other heuristics may be used. For instance, if a probabilistic prior distribution over voter preferences is known, then the expected degree of consistency can be used to heuristically score nodes for expansion. Other possibilities include expanding nodes that are "least similar" or most likely to be "correct" given the nodes that have already been expanded.

6.2.2 Approximation

We use the best axis algorithm 1D-SPBB as the core of more general algorithms to find optimal axes under various forms of approximation, and to estimate the degree to which a preference profile is approximately single-peaked. We propose several extensions of 1D-SPBB for three different notions of approximately single-peaked consistency (see Definition 2.34). In some cases, the algorithms do not guarantee discovery of the optimal approximation (i.e., the minimum k), but they provide both upper and lower bounds on the degree of approximation.

***k*-maverick consistency.** Computing the minimum k for which a profile is k-maverick consistent is precisely what 1D-SPBB does. The best axis A^* found by the algorithm explains $s(A^*)$ consistent voters (and this is the maximum number of voters explainable by any axis). Hence, the remaining $n - s(A^*)$ voters form the maverick set of minimum size. Hence, Algorithm 3 can be applied without any change.

***k*-additional axis consistency.** The 1D-SPBB algorithm can also be used to compute k-AA consistency. We investigate a simple greedy algorithm to approximate the minimum k for which a profile $\succ$ is k-AA single-peaked consistent. The algorithm, 1D-SPBB-AA, is shown

Algorithm 5 1D-SPBB-AA($N_r \leftarrow N$)

1: $k \leftarrow 0$
2: **while** $N_r \neq \emptyset$ **do**
3: Run Algorithm 3 on N_r, get the best axis A^* and the corresponding set of consistent voters V^*
4: $N_r \leftarrow N_r \backslash \{V^*\}$
5: $k \leftarrow k + 1$
6: Return k

in Algorithm 5 and works as follows: starting with the full profile, we find the best axis A_1 using 1D-SPBB. We then remove all n_1 voters consistent with A_1 from the profile (we use N_r to denote the set of voters that have not been removed, whose initial value is N) and rerun 1D-SPBB on the profile of the $n - n_1$ remaining voters. We repeat until the profile is empty. If it terminates after $k + 1$ iterations, 1D-SPBB-AA verifies k-AA consistency.

The value k determined by 1D-SPBB-AA is only an upper bound on the minimum k required for k-AA consistency because of its greedy nature. Note that deciding if a profile is k-AA consistent is NP-complete [Erdélyi et al., 2012], so our greedy algorithm cannot ensure an optimal k in general. However, The *first* iteration of the algorithm also determines a *lower bound* on k: if the first axis returned explains n_1 voters, then $k \geq \lceil \frac{n}{n_1} \rceil$ is needed to ensure k-AA consistency. We exploit this fact below in analyzing our data sets.

***k*-local candidate deletion consistency.** We can readily adapt 1D-SPBB to work with k-LCD consistency. Specifically, given a fixed value of k, we modify the algorithm to compute the best axis, i.e., the axis that renders the greatest number of voters single-peaked when we allow up to k candidates to be deleted from any voter's ranking. This is useful if we wish to see how single-peaked a profile is when voters, say, make "mistakes" in their ballots, or fail to distinguish certain candidates from one another. We can also combine this with k-AA consistency, finding the number of additional axes needed when each axis is allowed to explain voter preferences using k-LCD.

1D-SPBB can be used directly for this purpose, and requires only a modification in Algorithm 4, when computing the upper bound of a partial axis (or score of a complete axis) for

Algorithm 6 Compute Score or Upper Bound of (Partial) Axis A for k-LCD

$V \leftarrow \{\}$ %Set of consistent voters%
for agent $i \in N$ **do**
  consistent $\leftarrow$ TRUE, $v_i \leftarrow 0$
  $l \leftarrow 1, r \leftarrow n$ %Left and right pointers%
  for t from z to 1 **do**
    if $A[l] = \succ_{i,t}$ or unplaced candidate **then**
      $l \leftarrow l + 1$
    else if $A[r] = \succ_{i,t}$ or unplaced candidate **then**
      $r \leftarrow r - 1$
    else
      $v_i \leftarrow v_i + 1$
      if $v_i \leq k$ **then**
        Delete $\succ_{i,t}$ from A if it has been placed or a place holder otherwise
      else
        consistent $\leftarrow$ FALSE
        break
  if consistent = TRUE **then**
    $V \leftarrow V \cup \{i\}$

k-LCD. Instead of reporting a violation of single-peakedness when $\succ_i$ is inconsistent with the (partial) axis, it records, for each voter i, the number of inconsistencies detected so far (which we use v_i to denote, whose initial value is 0 for all voters)—each inconsistency can be managed by a local deletion. If $k+1$ violations occur, then i is reported as inconsistent with the (partial) axis. The algorithm is presented in Algorithm 6.

We use this method to find the best axis for fixed values of k in experiments below. The algorithm, 1D-SPBB-LCD, is presented in Algorithm 7. Since the number of consistent voters is non-decreasing in k, we can use binary search to find the minimum value of k that ensures k-LCD single-peaked consistency w.r.t. the best axis found by Algorithm 3. Since we can always make any profile single-peaked by removing $z-2$ candidates from each voter's ranking in the worst case, binary search will take at most $\log_2(z-1)$ iterations. Like k-AA, this problem is NP-complete [Erdélyi et al., 2012] and the algorithm may not find the *minimum number* of local deletions required: this is due to the fact that when a violation occurs, we simply remove the lower-ranked candidate in $\succ_i$, whereas a deletion of the higher-ranked candidate may have led to a fewer future deletions for voter i. Thus our method returns only an upper bound of the

Algorithm 7 1D-SPBB-LCD(0, $c-2$)

```
k ← c − 2
t = ⌊(c − 2)/2⌋
Run Algorithm 3 (combined with Algorithm 6 to compute lower (upper) bound for k-LCD), get the best axis A* and the set of consistent voters V*
if V* = N then
    k ← t
    1D-SPBB-LCD(0, t)
else
    1D-SPBB-LCD(t, c − 2)
Return k
```

optimal solution.

We also adapt the greedy algorithm, 1D-SPBB-AA, to find (approximate) the minimal number of additional axes needed when allowing k-LCD: we call this method 1D-SPBB-AA-k-LCD. The algorithm is a combination of Algorithm 3, Algorithm 5 and Algorithm 6, and used to test whether combinational of approximations can explain voter preferences. We will not the present the algorithm here, but refer the readers to the next section for experimental results.

6.2.3 Results from 2002 Irish General Election

We applied our algorithms to two data sets taken from the 2002 Irish general election.[3] The Dublin West election has 9 candidates and 29,989 votes of the top-t form (for varying values of t), of which 3,800 are complete preference rankings. In Dublin North, there are 12 candidates and 43,942 votes, of which 3,662 are complete.[4] Our primary experiments are run on the subset of votes comprising all complete rankings. We first ran 1D-SPBB (Algorithm 3), combined with Algorithm 4, to compute the best single axis for the two data sets. Figure 6.1 shows that the *best* axis explains, assuming single-peaked preferences, a tiny fraction of voter preferences, only 109 of 3,800 (2.87%) and 14 of 3,662 (0.38%) in Dublin West and Dublin

[3] Data sets obtained from: www.dublincountyreturningofficer.com.

[4] A ranking has top-t form if a voter ranks only her t most-preferred candidates. If $t < c-1$, the ranking is incomplete.

	Dublin West	Dublin North
# of consistent voters	109 (2.9%)	14 (0.4%)
Complete axis visited	2	3
Branch out due to bound	9,375	509,202
Running time (in sec.)	0.64	2.92

Figure 6.1: 1-D branch-and-bound results (best single axis).

North, respectively. Clearly voter preferences are far from being single-peaked.

Our methods can easily be adopted to partial rankings in the obvious fashion. Preliminary results on the full voting data sets, including the truly partial rankings, show that 6% (respectively, 6.5%) of voters are single-peaked with respect to the best axis in the West (respectively, North) data sets. This larger fraction is not unexpected, since single-peakedness cannot be violated by unranked candidates (so voters with top-t preferences for smaller values of t are consistent with far more axes). Despite this, voter preferences remain far from being single-peaked.

We also see that 1D-SPBB is quite efficient. While the total number of axes are $9!/2 \approx 181K$ and $12!/2 \approx 240M$, respectively, the algorithm needs only 0.64s. (respectively, 2.92s) to find the best axis, visiting only two (respectively, three) complete axes, and branching out 9,375 (respectively, 509,202) times, indicating good pruning due to strong lower bounds.

We also investigate the various approximations described above. The (single) best axis results immediately tell us that k-maverick consistency requires deletion of 97.13% (respectively, 99.62%) of voters to ensure the preference profile is single-peaked. We also immediately obtain a lower bound on k-AA consistency: Dublin West needs a minimum of $\lceil \frac{3800}{109} \rceil = 35$ additional axes, while North needs 262 additional axes. We also ran the greedy k-LCD approximation algorithm, 1D-SPBB-AA-k-LCD, for different values of k (when $k = 0$, this is just 1D-SPBB-AA). Figure 6.2(a) and (b) show the percentage of voters explained with each additional axis added for different values of k (note the log-scale on the x-axis). Without k-LCD approximation (i.e., when $k = 0$), 447 (respectively, 1,452) axes are needed to explain

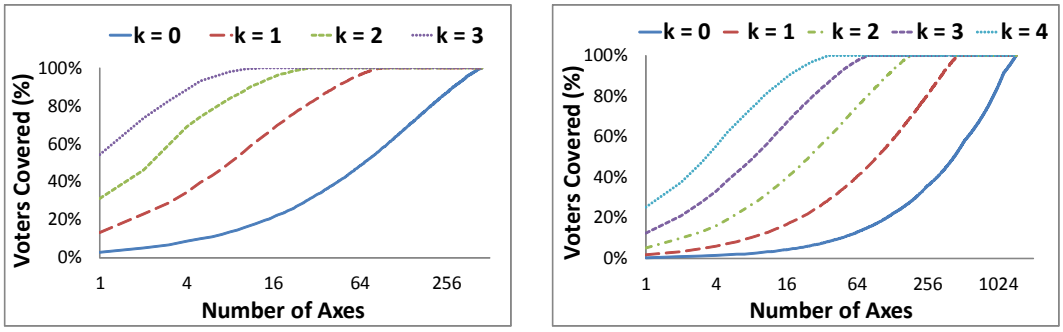


Figure 6.2: 1-D branch-and-bound results, with LCD-approximation: Dublin West (top); Dublin North (bottom).

all voter preferences (this is an upper bound on k-AA consistency). k-LCD without multiple axes requires deletion of $k = 7$ (respectively, $k = 10$) candidates, the maximum possible for each data set. Even with k-LCD for reasonable values of k, many additional axes are needed to explain the data: for instance, 31 axes are needed to explain Dublin West when $k = 2$, while Dublin North, with an aggressive $k = 4$, needs 39 axes. The linear nature of the plots (recall the log-scale) also shows that deletion of maverick voters will not help. This suggests that, even allowing for *combinations of approximations* proposed in the literature, preferences in these data sets are very far from being single-peaked in 1D. This motivates the use of higher-dimensional models, to which we now turn.

6.3 A Two-dimensional Branch and Bound Algorithm

Since voter preferences in the data sets above are not single-peaked in the one-dimensional sense—even when aggressive approximation is considered—the explanatory power of these proposed approximations in 1D is rather limited. We now extend these techniques to two-dimensional (2D) models (see Definition 2.24). Our extensions generalize beyond two dimensions, but we focus on 2D models for ease of presentation, and also because, as we see below, two dimensions suffice for our data sets

6.3.1 The Algorithm

Extending our branch-and-bound algorithm to the 2D case presents several challenges. First, the search space explodes, as we must potentially consider all $O((z!)^2)$ combinations of first and second axes. Second, candidates ranked last in some $\succ_i$ need no longer lie at the extreme point of an axis (as we will see later). Third, in two-dimensions, some axes are *dominated* by others—these should be pruned for computational efficiency to the greatest extent possible. We now outline a 2D extension of the 1D-SPBB called *2D-SPBB*, and explain how to tackle each of the issues above.

To address the combinatorial explosion of possible pairs of axes, instead of considering all candidate permutations as possible first axes, we admit only a relatively small set of potential initial axes, or a limited *sample* of possible axes. For each such (potential) first axis, we fix it as our first dimension, then apply our 1D algorithm 1D-SPBB to compute the second dimension. Our implementation uses 1D-SPBB-AA to find a collection of 1D axes that fully explains the given profile $\succ$—we use this as our set of potential first dimension axes. This guarantees that each $\succ_i$ is single-peaked consistent w.r.t. at least one of the axes.[5] This way of structuring 2D-SPBB means any axis searched/expanded in the first dimension is always complete, never partial. Of course, this is simply a heuristic, and may limit our ability to find a good 2D model.

Computing scores and upper bounds.

To address the second problem, we develop a new algorithm to compute the upper bound for a pair of partial axes in a 2D space (i.e., maximum number of voters that are consistent with some extension of the partial pair), or the score of the pair of axes if they are complete. This includes variants that incorporate the same forms of approximation as above. One key difference between 2D and 1D lies in the computation of consistency. In a 2D space, the inconsistency of $\succ_i$ with single-peakedness only occurs with the violation of some *bounding box constraint*.

[5] If a voter is single-peaked w.r.t. one axis A, then she is also single-peaked w.r.t. any 2D-space using A as one of its axes.

Definition 6.3 (Bounding box constraint) *A bounding box constraint* $\mathbf{b} = \langle b_1, b_2, b_3 \rangle$ *is tuple of three candidates* $b_1, b_2, b_3 \in C$ *such that* $||b_1 - b_3||_1 = ||b_1 - b_2||_1 + ||b_2 - b_3||_1$.

In other words, a tuple $\mathbf{b} = \langle b_1, b_2, b_3 \rangle$ is a bounding box constraint if b_2 is contained in the minimum bounding box of b_1 and b_3. Recall that the definition of multi-dimensional single-peakedness (Definition 2.24), each bounding box constraint imposes two restrictions for single-peakedness on agent preferences: first., if b_1 is the peak of an agent, then b_2 must be ranked before b_3 in her preference; and second if b_3 is the peak, then b_2 must be ranked before b_1. The following example shows a bounding box constraint imposed by a multi-dimensional axis.

Example 6.1 (Bounding box constraint in 2D) *The following 2D example with five candidates* $c_1, \ldots, c_5$ *illustrates the concept of bonding box constraint. Assume two axes,* $A_1 = c_4 <_{A_1} c_3 <_{A_1} c_1 <_{A_1} c_2 <_{A_1} c_5$ *and* $A_2 = c_1 <_{A_2} c_2 <_{A_2} c_4 <_{A_2} c_5 <_{A_2} c_3$ *(as shown in Figure 6.3). The only bounding box constraint is* $\langle c_1, c_2, c_5 \rangle$*, which induces two restrictions on agent preferences: (a) if* $\tau_i = c_1$ *for some* i*, then we must have* $c_2 \succ_i c_5$*; and (b) if* $\tau_i = c_5$ *for some* i*, then we must have* $c_2 \succ_i c_1$*. So the preference profile* $\succ = \{\succ_1, \succ_2\}$ *where* $\succ_1 = c_4 \succ_1 c_3 \succ_1 c_1 \succ_1 c_5 \succ_1 c_2$ *and* $\succ_2 = c_5 \succ_1 c_1 \succ_1 c_2 \succ_1 c_3 \succ_i c_4$ *is not single-peaked, as the condition (b) is violated by agent 2.*

Note that in a multi-dimensional space, candidates ranked last in some $\succ_i$ need no longer lie at the extreme point of an axis, so Algorithm 4 does not apply here. For instance, in the above example, $c_4 \succ_i c_3 \succ_i c_1 \succ_i c_5 \succ_i c_2$ is a valid single-peaked preference, but the least preferred candidate c_2 is not extreme on either axis. So we need a new algorithm as the counterpart of Algorithm 4 to compute the lower and upper bound for a multi-dimensional axis. The algorithm is given in Algorithm 8, which checks for violations of bounding box constraints in agent preferences. Specifically, for each $\succ_i$, we compute the set of bounding box constraints $\mathbf{B}$ induced by the partial axes. Recall that each constraint $\mathbf{b} = \langle b_1, b_2, b_3 \rangle$, involves three candidates: if $\tau_i = b_1$ (respectively, $\tau_i = b_3$), then we must have $b_2 \succ_i b_3$ (respectively,

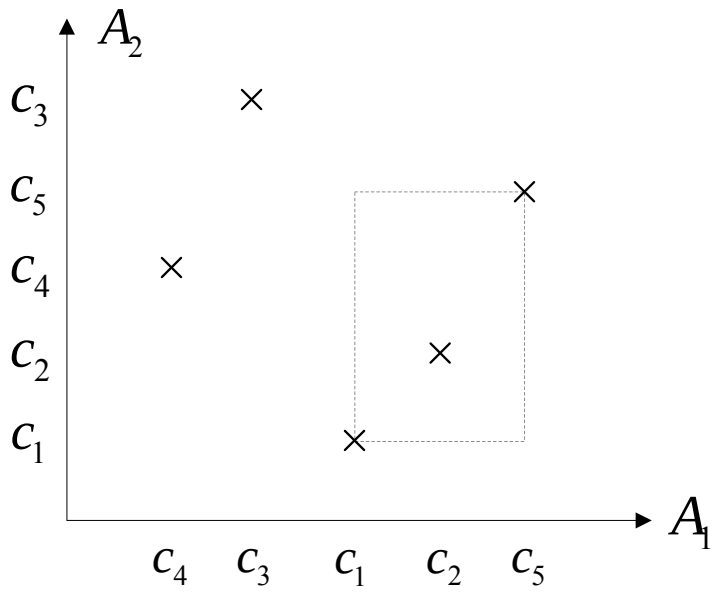


Figure 6.3: Bounding box constraints imposed by an axis.

$b_2 \succ_i b_1$) to ensure single-peakedness of $\succ_i$. If no constraints are violated, i joins the set of consistent voters.

As in the 1D case, consistency testing must be fast to ensure that nodes in the branch-and-bound tree are processed quickly. Consistency testing is polynomial time:

Theorem 6.2 *Given a preference profile $\succ$, the number of voters consistent with a pair of partial axes A_1, A_2 can be computed in $O(nc^4)$ time.*

Proof: The set of bounding box constraints $\mathbf{B}$ can be computed in $O(c^3)$ time, since each constraint involves candidate triples (of which there are at most $\binom{c}{3}$). Testing a ranking $\succ_i$ against each such constraint (as described above) can be accomplished in $O(c)$ time, and must be done at most once for each of n voters (generally, substantially fewer at deeper nodes in the tree).[6] Thus total running time is $O(c^3) + O(nc^4) = O(nc^4)$. ■

We mention two important details regarding the computation of the set of bounding box constraints $\mathbf{B}$. First, it can be done incrementally by inheriting bounding box constraints from nodes higher in the search tree, then adding only the new constraints induced by placing two

[6]Other efficiencies, e.g., caching consistency tests across voters with identical preference orderings, are also possible.

Algorithm 8 Compute Lower (Upper) Bound for A (2D) Axis $A = (A_1, A_2)$

```
1: B ← The set of bounding box constraints induced by A
2: V ← {} %Set of consistent voters%
3: for agent i ∈ N do
4:    consistent ← TRUE
5:    for bounding box constraint b ∈ B do
6:       if ≻_{i,1}= b_1(b_3) then
7:          if b_2 is ranked lower than b_3 (b_1) then
8:             consistent ← FALSE
9:             break
10:   if consistent = TRUE then
11:      V ← V ∪ {i}
```

more candidates on the second axis. Second, for any incomplete axis, apart from "explicit" constraints involving candidates on the axis, we can also compute "implicit" constraints. For example, suppose A_1 is fixed, with $c_1 <_{A_1} c_2 <_{A_1} c_3 <_{A_1} c_4 <_{A_1} c_5 <_{A_1} c_6$, while A_2 is partial, with $A_2 = c_1 <_{A_2} c_6 <_{A_2} \ldots <_{A_2} c_5 <_{A_2} c_2$. The only explicit constraint is $\langle c_6, c_5, c_2\rangle$; but four implicit constraints can be added: $\langle c_1, c_3, c_5\rangle$, $\langle c_1, c_4, c_5\rangle$, $\langle c_6, c_4, c_2\rangle$ and $\langle c_6, c_3, c_2\rangle$. This allows more precise upper bound computation and more aggressive pruning.

Removing dominated axes.

The fact that pairs of axes in 2D give rise to bounding box constraints leads to a form of "domination" that can be exploited to further reduce the combinatorial overhead of searching.

Definition 6.4 *A pair of (partial) axes $A = \langle A_1, A_2\rangle$ is* dominated *by $A' = \langle A'_1, A'_2\rangle$ if the set of bounding box constraints induced by A' is a strict subset of that induced by A.*

Consider $A = \langle A_1, A_2\rangle$, with complete axis $A_1 = c_1 <_{A_1} c_2 <_{A_1} c_3 <_{A_1} c_4 <_{A_1} c_5$ and partial axis $A_2 = c_1 <_{A_2} \ldots <_{A_2} c_5$. $A = \langle A_1, A_2\rangle$ is dominated by a different pair $A = \langle A_1, A'_2\rangle$: we obtain strictly fewer bounding box constraints by swapping c_1 with whichever candidate happens to be placed in the second position of A_2, and c_5 with whichever candidate is placed in the fourth position. As such, assuming A_1 is fixed (as we would in a specific branch of 2D-SPBB), a different axis A'_2 offers strictly more flexibility than A_2.

We exploit this fact by using an algorithm for removing (some, but not all) dominated axes

	Dublin West	Dublin North
$k = 0$	2,498/3800 (65.73%)	1,732/3,662 (47.30%)
$k = 1$	3,553/3800 (93.5%)	2,948/3,662 (80.50%)
$k = 2$	3,788/3800 (99.68%)	3,436/3,662 (93.83%)
$k = 3$	3,800/3800 (100%)	3,645/3,662 (99.54%)

Figure 6.4: 2-D branch-and-bound: number of consistent voters with single best 2D axis using k-LCD approximation.

during 2D-SPBB: detecting this can allow pruning of a large part of the branch-and-bound tree. We test domination by checking whether a swap of two adjacent candidates on any axis can induce a strict subset of original constraints (as in the above example): if yes, the (partial) axis is pruned. This simple test is sound; and while it does not ensure pruning of *all* dominated axes, it improves run-time dramatically.

Approximation.

As in the 1D case, 2D-SPBB automatically generates the minimal k required for k-maverick consistency. Of course, if we use sampling to limit the axes that will be considered for the first dimension, we will obtain only an upper bound on k. It can also be applied repeatedly to greedily approximate the minimal set of additional 2D "axis pairs" needed to explain a profile; and we can easily incorporate k-LCD approximation into 2D-SPBB using similar modifications to those described in Sec. 6.2.2. We focus on k-LCD below.

6.3.2 Results from the 2002 Irish General Election

We use the Dublin West and North data sets to test the effectiveness of 2D-SPBB and specifically the ability of k-LCD approximation to fit the Irish voting data. Figure 6.4 shows the fraction of voters that are explained by the *best axis pair* generated using 2D-SPBB, both without approximation ($k = 0$), and allowing k-LCD approximation for $k \leq 3$. The contrast with the 1D fit is notable: even without approximation, the best 2D-axis pair explains 65.7% (respectively, 47.3%) of all voters. Allowing 2 out of 9 (respectively, 3 out of 12) local candidate

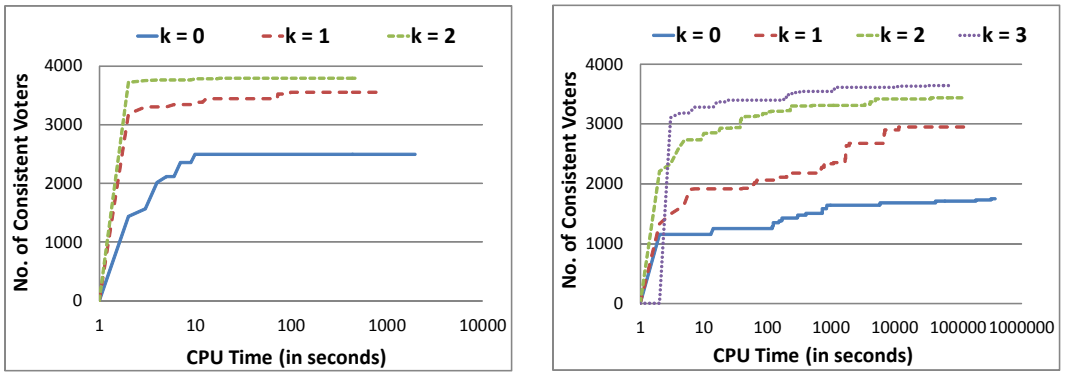


Figure 6.5: 2-D branch-and-bound results, anytime performance: (a) Dublin West; (b) Dublin North.

deletion provides a near-perfect fit, covering 99.68% (respectively, 99.54%) of voters. This strongly suggests that the 2D model carries far more explanatory power for this Irish voting data.

The 2D algorithms are more computationally intensive than their 1D counterparts (though restricting attention to sampled axes in the first dimension helps tremendously). It is instructive to examine the anytime performance of 2D-SPBB to see how quickly it converges to the best 2D model, and how quickly model quality improves for various levels of k-LCD. Figs. 6.5 (a) and (b) show that, while convergence to the best 2D model can take a considerable amount of time, the anytime performance is very good, allowing the discovery of models that capture most of the (explainable) voters extremely quickly (note the log-scale on the x-axis).

6.4 Spatial Model for Rank Data

As mentioned above, single-peaked preference is a "loose" model in the sense that it admits a large number of consistent preference relations. A more restricted model the spatial model introduced in Section 2.4.3, in which both voters and candidates are represented by points in some single- or multi-dimensional space and the closer a candidate is to a voter, the more preferred that candidate is.

Recall from Section 2.4.3 that, if the agent and candidate positions are given in some single-

or multi-dimensional space, then the preference of each agent $\succ_i$ can be computed either deterministically or stochastically. In this section, we consider the reverse problem of given a preference profile $\succ = \{\succ_1, \ldots, \succ_n\}$, can we fit both agents and candidates into some single- or multi-dimensional space such that the "error" observed in $\succ$ is minimized. In particular, we consider the Plackett-Luce model introduced in Section 2.4.3, and propose an alternating optimization algorithm to maximize the log-likelihood value of the observed preferences. We present some preliminary results on fitting spatial models to two data sets from real-world elections. These fitting results have been used for generating single-peaked preferences in Section 5.4 and 4.4 for both unconstrained-FLPs (if only considering agent positions) and constrained-FLPs (if considering both agent and candidate positions).

We use the same notation as in Section 2.4. Let $N = \{1, \ldots, n\}$ be the set of voters and $C = \{1, \ldots, c\}$ be the set of candidates. Each voter $i \in N$ has a strict preference $\succ_i$ over C, and a preference profile $\succ = \{\succ_1, \ldots, \succ_n\}$ is the joint preferences of all voters. Our objective is to estimate the voter positions t_i and candidate positions c_j in the latent feature space, given a preference profile $\succ$.

6.4.1 Log-likelihood Maximization

Recall from Section 2.4.3 that the Plackett-Luce model is parametrized by a vector $b_i = (b_{i1}, \ldots, b_{in})$ for each agent i, where b_{ij} is the probability that agent i will choose candidate j as as her most preferred one. A popular form of this probability is the exponential decreasing function of the squared distances:

$$b_{ij} = \frac{\mathbf{exp}\{-d(t_i, c_j)\}}{\sum_{j'=1}^{c} \mathbf{exp}\{-d(t_i, c_{j'})\}}$$

where $d(t_i, c_j) = ||t_i^k - c_j^k||_2^2$ is the squared distance between t_i and c_j.

The model is a multi-stage model such that each agent keeps choosing the next most preferred candidate from the set of available candidates until all candidates have been selected

in her ranking. At each stage, the probability b_{ij} is normalized subject to the constraints that candidates who have been selected are excluded from the vector b_i. If we use t_i to denote the position of agent i and c_j to denote the location of facility j, then the log-likelihood of the observed preference profile (assuming each agent cast her vote independently) is:

$$\ln(Pr(\succ \mid \mathbf{b})) = \ln \prod_{i=1}^{n} \prod_{j=1}^{c} \frac{b_{ij}}{\sum_{j \succeq_i j'} b_{ij'}} = -\sum_{i=1}^{n} \sum_{j=1}^{c} d(t_i, c_j) - \sum_{i=1}^{n} \sum_{j=1}^{c} \ln \omega_{ij} \tag{6.1}$$

where $\omega_{ij} = \sum_{j \succeq_i j'} \mathbf{exp}\{-d(t_i, c'_j)\}$ is a normalization factor, which is the sum of probabilities over all candidates who are not ranked higher than candidate j in voter i's ranking.

Note that the only unknown variables in Equation (6.1) are t_i and c_j, so our objective is to choose t_i and c_j to maximize the log-likelihood value in Equation (6.1). Similar problem has also been considered by Gormley and Murphy [2007]. They focus on learning the distribution of agent and candidate positions using a Metropolis-Hastings algorithm, while we use a different optimization algorithm to optimize agent and candidate positions alternatively. The details of the algorithm will be introduced in the next section.

6.4.2 An Alternating Optimization Algorithm

In this section, we propose an alternating optimization algorithm that optimizes voters' and candidates' positions alternatively to maximize the value in Equation (6.1). Our algorithm is a natural extension of that proposed by Poole and Rosenthal [1985] in the sense that it can be applied to model preference rankings rather than binary choices.

The algorithm is specified in Algorithm 9, and is composed of two stages. In the first stage, it optimizes each voter's position (t_i) to maximize the log-likelihood with respect to that agent independently (line 2-4), assuming the positions of candidates (c_js) are fixed. This is because each voter cast her vote independently, and her ranking only depends on the distance between herself and the candidates, whose positions are fixed in this stage. In the second stage, it jointly optimize candidates' positions (c_js) to maximize the log-likelihood value in Equation(6.1) (line

Algorithm 9 The Alternating Algorithm for Maximizing Log-likelihood in (6.1)

1: Randomly select initial positions of t_i and c_j.
2: **while** not converged **do**
3: **for** agent $i \in N$ **do**
4: Fix the positions of candidates, and optimize voter i's position t_i to maximize her log-likelihood given her ranking $\succ_i$.
5: **end for**
6: Fix the positions of voters optimized in the first stage, jointly optimize candidates' positions c_js to maximize the log-likelihood in (6.1).
7: **end while**

5), assuming agent positions (t_is) are fixed. Note that the candidate positions cannot be optimized independently because the probability that a voter makes her first choice depends not only on how close she is to that candidate, but also on how close she is from the others. Fortunately, the number of candidates is usually small, and the number of variables that have to be optimized simultaneously is $c \cdot m$ (where c is the number of candidates and m is the number of dimensions). The alternating optimization is repeated until the algorithm has converged.

It is hard to optimize Equation (6.1) analytically because of the second term ($\ln \omega_{ij}$ is the log value of a sum), so we use numerical optimization here. In each stage of the algorithm, we use the Broyden-Fletcher-Goldfarb-Shanno (BFGS) method to maximize the log-likelihood values. The BFGS algorithm is a Quasi-Newton method for solving nonlinear optimization problems [Broyden, 1970, Fletcher, 1970, Goldfarb, 1970, Shanno and Kettler, 1970, Shanno, 1970]. Compared with Newton's method in which evaluating the Hessian matrix is computationally intensive, the BFGS method only approximates the Hessian matrix by (approximate) gradient evaluations. It should be noted that the BFGS method is not guaranteed to converge in general, however, when combined with random restart techniques, it is shown to be very efficient in practice.

6.4.3 An Empirical Study on Irish General Election 2002

We conduct an empirical study on two data sets from the Irish general election 2002, and evaluate the accuracy of the spatial model in terms of the likelihood values. Specifically, we use

the alternating optimization algorithm to maximize the log-likelihood value in Equation (6.1) and estimate both agent and candidates positions in the latent 1D or 2D space. Note that theoretically one can use a space with any number of dimensions, however, the computational effort required and improvement on accuracy should also be factored when moving to higher dimensions. Given the estimated agent and candidate positions, we also compute the inferred rankings of agents using the deterministic choice model introduced in Section 2.4.3 (in particular, L_2 cost), and compare them with the actual observed ranking using two different measures: *Kendall Tau distance* and *Spearman's Footrule distance*. The former measures the number of misranked pairs, and the latter one computes the sum of differences in positions.

Definition 6.5 (Kendall Tau distance) *Let α and β be two rankings. The Kendall Tau (KT) distance of α and β is defined as:*

$$K(\alpha, \beta) = \{(i, j) | i < j, (\alpha(i) - \alpha(j)) \cdot (\beta(i) - \beta(j)) < 0\}$$

where $\alpha(t)$ and $\beta(t)$ are the ranked position of candidate t in α and β, respectively.

Definition 6.6 (Spearman's Footrule distance) *Let α and β be two rankings. The Spearman's Footrule (SF) distance of α and β is defined as:*

$$S(\alpha, \beta) = \sum_{i=1}^{c} |\alpha(i) - \beta(i)|$$

where $\alpha(t)$ and $\beta(t)$ are the ranked position of candidate t in α and β, respectively.

Data Source The Irish general election of 2002 was the first time that electronic voting machines are used in an Irish election, in three constituencies: Dublin-west, Dublin-north and Meath. We will use the two data sets of Dublin-west and Dublin-north here.[7]

There are three and four seats allocated to the constituency of Dublin-west and Dublin-

[7]The data sets are available from http://www.dublincountyreturningofficer.com

Candidate	Party
Brain Lenihan, Jnr (BLJ)	Fianna Fáil (FF)
Joe Higgins (JH)	Socialist Party (SP)
Joan Burton (JB)	Labour Party (LP)
Sheila Terry (ST)	Fine Gael (FG)
Deirdre Doherty Ryan (DDR)	Fianna Fáil (FF)
Tom Morrissey (TM)	Progressive Democrats (PD)
Mary Lou McDonald (MLM)	Sinn Féin (SF)
Robert Bonnie (RB)	Green Party (GP)
John Smyth (JS)	Christian Solidarity (CS)

Table 6.1: Candidates and their belonging parties for the constituency of Dublin-west

north, respectively. The voting mechanism used is called single transferable vote form of proportional representation (PR-STV). In PR-STV, each voter provides a (incomplete) ranking of the candidates, and her vote is initially allocated to the most preferred candidate. During the counting process, if a candidate is elected or eliminated, then votes from voters who place that candidate as her first choice will be transferred to other candidates according to her ranking. This procedure is repeated until all seats are allocated or a sufficient number of candidates are left.

For Dublin-west, there are 9 candidates and 29,989 votes of the top-t form, of which 3,800 are complete; For Dublin-north, there are 12 candidates and 43,942 votes of the top-t form, of which 3,662 are complete. The candidates and their belong parties are listed in Table 6.1 and 6.2 for details (Candidates in bold text are winners).

Results of Learning a Spatial Model Figure 6.6 shows the estimated agent and candidate positions using one- and two-dimensional model, respectively, when the best fitting result (i.e., maximum log-likelihood value in Equation (6.1)) is selected among 100 runs of Algorithm 9. An interesting results is that party politics plays an important role in the electorates view of the candidates: in Dublin-west, there are only two candidates who are from the same party (BLJ and DDR are from FF), and in both our one and two-dimensional fitting results, these two candidates are clustered together; in Dublin-north, there are two parties with more than one candidate (FF has JG, GVW and MK, FG has NO and CB), and in both results, candidates

Candidate	Party
Trevor Sargent (TS)	Green Party (GP)
Seán Ryan (SR)	Labour Party (LP)
Jim Glennon (JG)	Fianna Fáil (FF)
G. V. Wright (GVW)	Fianna Fáil (FF)
Clare Daly (CD)	Socialist Party (SP)
Michael Kennedy (MK)	Fianna Fáil (FF)
Nora Owen (NO)	Fine Gael (FG)
Mick Davis (MD)	Sinn Fèin (SF)
Cathal Boland (CB)	Fine Gael (FG)
Ciaràn Goulding (CG)	Ind. Health Alliance (IHA)
Eamon Quinn (EQ)	Independent (IND)
David Walshe (DW)	Christian Solidarity (CS)

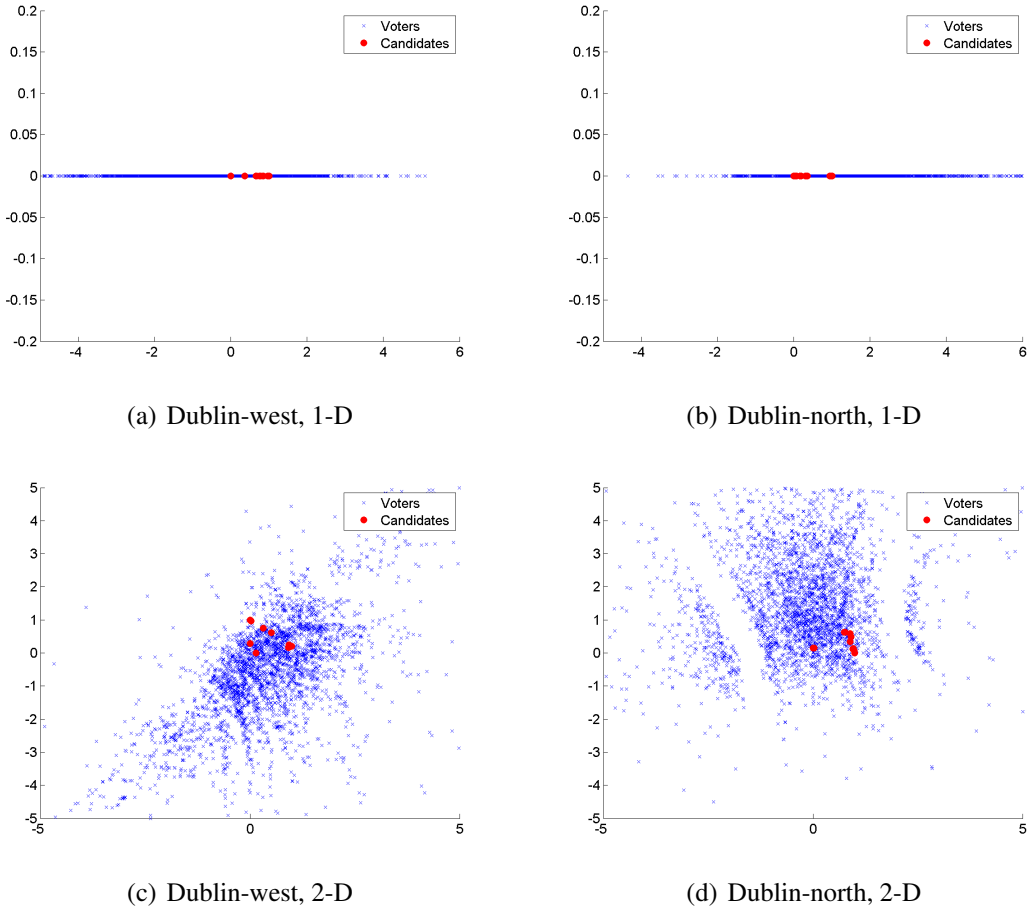
Table 6.2: Candidates and their belonging parties for the constituency of Dublin-north

belonging to the same parties are clustered together.

We also compare the average likelihood, average Kendall Tau distance, and average Spearman's Footrule distance respectively in Table 6.3. We can see that the two-dimensional model beats the one-dimensional model in all three measures: for the average likelihood, the two-dimensional model is 37.72 and 12.25 times better than the one-dimensional model for Dublin-west and Dublin-north, respectively; for the average KT and SF distances, the two-dimensional model is also better than the one-dimensional better, although not that significantly. The results are also consistent with the findings we have in single-peaked model (see earlier results presented in this chapter) and those reported by Gormley and Murphy [2007] such that the two-dimensional model can explain more voter preferences than the single-dimensional one.

	Dublin-west		Dublin-north	
	1-D	2-D	1-D	2-D
Avg. Likelihood	8.35E-05	3.15E-03	8.16E-06	1.00E-04
Avg. KT Dis.	10.56	8.22	19.94	16.86
Avg. SF Dis.	17.05	13.71	30.55	26.43

Table 6.3: Comparison of the one-dimensional and two-dimensional Fittings



(a) Dublin-west, 1-D

(b) Dublin-north, 1-D

(c) Dublin-west, 2-D

(d) Dublin-north, 2-D

Figure 6.6: Results of the one-dimensional fitting: a) One-dimensional Fitting for Dublin-west, b) One-dimensional Fitting for Dublin-north, c) Two-dimensional Fitting for Dublin-west, d) Two-dimensional Fitting for Dublin-north

6.5 Conclusion

In this chapter, we have developed a branch-and-bound algorithm designed to determine the degree to which a preference profile can be viewed as single-peaked in both the single- and multi-dimensional senses. It uses, and combines, various forms of approximation. Experiments on two election data sets demonstrate clearly that one-dimensional models, for any reasonable degree of approximation, cannot explain voter preferences in the two data sets we have explored. By contrast, a two-dimensional model provides an excellent fit, using very low degrees of local candidate deletion (as the only form of approximation) to explain the preferences of over 99% of the voters in each data set. Our algorithms are very effective in practice in

1D spaces, and feasible in 2D with strong anytime performance, despite the NP-completeness of these problems. While these findings are preliminary, and need to be corroborated on further election and other preference data sets, they suggest that the extension to two (or additional) dimensions may render the use of single-peaked modeling, or its approximations, more applicable in practice.

We also studied a spatial model for ranking. Combined with the Plackett-Luce model as the choice model, we fit both voters and candidates in a same m-dimensional space by maximizing the log-likelihood that the observed rankings are correct. We propose an alternating algorithm to optimize the voters' and candidates' position by fixing one or the other. An empirical study on two data sets from the Irish General Election 2002 shows applicability of our method, and suggests that two-dimensional model is better in terms of fitting accuracy.

The results in this chapter can be extended in several ways. One is to further develop the theoretical foundations of single-peaked consistency for different forms of approximations in higher dimensions. Second, multi-dimensional single-peakedness is a much weaker assumptions than its 1D counterpart; so while it may fit preference data better, its predictive power is lessened. Developing a deeper understanding of these tradeoffs is vital. An interesting question is, for instance, minimum conditions on profiles that prevent the fit of *any* m-dimensional model (c.f 1D, where single-peakedness can be violated with only two voters and four candidates, or three voters and three candidates Ballester and Haeringer [2011]. Finally, while much attention has been paid to mechanisms that exploit single-peakedness (e.g., generalized median mechanisms or quantile mechanisms as we have introduced), little work has addressed the impact of approximate single-peakedness on these mechanisms, or the design and analysis of mechanisms specifically for approximate single-peakedness. Having a sense of which forms of approximation best fit real-world preferences can help focus mechanism design efforts on those most likely to have a practical impact. We will address more future work in detail in Chapter 8.

Chapter 7

The Trade-off Between Efficiency and Privacy

7.1 Introduction

We have described the median mechanism and its generalization [Black, 1948, Moulin, 1980, Barberà et al., 1993] in Chapter 2. We also showed that how these mechanism can be generalized to multi-dimensional, multi-facility location problems (e.g., through quantile mechanisms) in Chapter 3. All of these mechanisms assume *direct revelation*, in which agents reveal their full utility functions to the mechanism. However, direct mechanisms often elicit more information than required to make optimal choices, leading to communication and computational difficulties [Conitzer and Sandholm, 2004]. For example, consider the simple median mechanism (in Definition 2.25) which selects the median position among all reported peaks. Revealing full preferences is often unnecessary in terms of communication: while the only information needed by the mechanism is the peaks (most preferred outcomes) of the agents, direct revelation requires agents to evaluate their preferences over all outcomes and report such preferences to the mechanism.

Direct revelation also requires a sacrifice of *privacy*: revealing its full utility function may

be undesirable for an agent, especially when some of that information is *provably unnecessary* for computing the optimal outcome. Again, in the median mechanism where peak positions are the only require information, agents may prefer not to reveal their full preferences over other candidates. Recent work—using techniques similar to those used in the analysis of communication complexity [Kushilevitz and Nisan, 1996]—has analyzed privacy preservation of specific mechanisms in this sense; that is, where the degree to which an agent reveals more than is strictly needed to compute the outcome of the mechanism is the degree to which privacy has been lost.[1] For instance, Sandholm and Brandt [2008] showed that, for the second-price auctions (see Example 2.3), the English auction (defined formally later) preserves complete privacy—no agent reveals any more than is strictly necessary to determine the outcome—but that this comes at the cost of exponential communication complexity. More recently, Feigenbaum et al. [2010] proposed a general framework to analyze the *trade-off between privacy and communication*, defining several forms of *privacy approximation ratio*. They also showed how different mechanisms for second-price auctions (and several other problems) improve privacy at the expense of communication, and vice versa.

Previous work has addressed both the trade-off between communication and efficiency, and the trade-off between privacy and communication. In this chapter, we address a third trade-off, that between efficiency and privacy, and provide a general framework for analyzing this trade-off. Specifically, we consider *approximate mechanisms* that find ε-optimal solutions to a social choice problem, and show how agents' privacy improves as one increases the degree of approximation ε. Our contributions are as follows: In Section 7.2 we define a general framework for analyzing these trade-offs, extending *privacy approximation ratios*, introduced by Feigenbaum et al. [2010], to the case of approximate mechanisms. In Section 7.3, we analyze the efficiency-privacy trade-off in approximate versions of several mechanisms for second-price

[1]Note that the notion of *privacy* used here is quite different from *differential privacy*, which deals with the potential "leakage" of a user's private information associated with a particular set of queries to a database [Dwork, 2006]. Though some connections between differential privacy and mechanism design have been developed [McSherry and Talwar, 2007], these have focused largely on how to exploit differential privacy to design approximately efficient and truthful mechanisms, and do not attempt to limit information revelation in the sense we pursue here.

auctions, including the English, sealed-bid, and bisection protocols (these will be defined formally later) in both the worst and average cases, and compare our ε-privacy approximation ratios with the exact ratios derived by Feigenbaum et al. [2010]. We also generalize their analysis from 2-agent to n-agent auctions. In Section 7.4, we develop incremental protocols for *facility location problems* that implement the classic median mechanism (in Definition 2.25). We analyze the exact privacy approximation ratio for these new protocols, and again derive results demonstrating the efficiency-privacy trade-off induced by approximate versions of these protocols.

Approximate mechanisms will not just improve (increase) privacy, but also generally improve (reduce) communication complexity. Furthermore, sacrificing efficiency usually breaks the incentive properties of standard mechanisms. We note that all of our mechanisms are exactly truthful or ε-incentive compatible,[2] and demonstrate this below.

7.2 Efficiency-Privacy Trade-off

Much of mechanism design deals with *direct revelation mechanisms*, in which each agent reveals its entire type to the mechanism. For simple outcome spaces (e.g., single-item auctions), the precision required by direct revelation is often unnecessary; in complex settings (e.g., combinatorial auctions [Cramton et al., 2005] in which each agent can bid on a combination of items), the outcome set O has exponential size, imposing significant burdens on communication. Incremental mechanisms have been proposed (e.g., ascending auctions [Parkes, 1999] and adaptive elicitation [Zinkevich et al., 2003] for combinatorial auctions) which, by eliciting only information that is "needed," can reduce this burden in practice. In general, however, such methods cannot reduce information requirements in the worst case [Nisan and Segal, 2006].

In a different vein, one can use *informational approximation*, eliciting information about agent valuations that admits only an approximately optimal choice. For example, *priority*

[2]Indeed, when one factors in incentives, there is a more complex four-way trade-off between efficiency, privacy, communication complexity and incentives.

games [Blumrosen and Nisan, 2002] model single-item auctions in which agents express their valuations with limited precision, and provide allocations (and prices) that sacrifice efficiency (since true types are unknown) for communication savings; they are also strategy-proof. *Partial revelation VCG mechanisms* [Hyafil and Boutilier, 2007] apply in any setting where VCG can be used (namely, to maximize social welfare under quasi-linear utility), but again limit revelation and sacrifice efficiency. Without efficient outcome selection, such mechanisms are not strategy-proof; but with approximate variants of VCG pricing, ε-efficiency induces ε-incentive compatibility in dominant strategies.

The *communication complexity* model [Kushilevitz and Nisan, 1996] provides a useful framework for analyzing the communication or informational costs of specific protocols. They can also be adapted to quantify the degree of privacy revelation in mechanisms. Following our notation in Section 2.2, if we let n be the number of agents, $\mathbf{t} = (t_1, \ldots, t_n)$ be a type profile, and f is a social choice function, then one can think of the social choice function f as a n-dimensional matrix (tensor) M^f whose entry at position (a type profile) $\mathbf{t}$ is $f(\mathbf{t}) = f(t_1, \cdots, t_n)$. Then we can define the *ideal monochromatic region* and the *ideal monochromatic partition* as follows:

Definition 7.1 (Ideal monochromatic region and partition) *Let f be an arbitrary social choice function. The* ideal monochromatic region *for a type profile $\mathbf{t} \in T$ with respect to f is $R_f^I(\mathbf{t}) = \{\mathbf{t}' | f(\mathbf{t}') = f(\mathbf{t})\}$. The* ideal monochromatic partition *of f is the set of (disjoint) ideal monochromatic regions with respect to T.*

Intuitively, $R_f^I(\mathbf{t})$ describes the set of type profiles $\mathbf{t}'$ that are indistinguishable from $\mathbf{t}$ relative to the social choice function f: each such $\mathbf{t}'$ leads to the same choice $o = f(\mathbf{t})$. Thus the the ideal monochromatic region in which the true profile $\mathbf{t}$ is contained can be thought of as minimum information required to compute the outcome $f(\mathbf{t})$, and is both necessary and sufficient to determine that outcome.

A *(deterministic) communication protocol p* specifies the rules by which agents with private

information share that information (with a third party or one another) to compute the outcome of a function [Kushilevitz and Nisan, 1996]. If the outcome selected under a communication protocol (or the *run*) is the same as that selected under a social choice function for any input, then we say the protocol *implements* the social choice function. As such, a mechanism is simply a protocol. Formally, we have:

Definition 7.2 (Run of a protocol) *Let p be a communication protocol, and $\mathbf{t}$ be a type profile. We define the* run *of a protocol on a type profile, denoted as $p(\mathbf{t})$, as the outcome selected under the execution of p when agents have preferences $\mathbf{t}$.*

Definition 7.3 (Implementation) *Let p be a communication protocol, and $p(\mathbf{t})$ be run of p on $\mathbf{t}$. We say p implements a social choice function f if $p(\mathbf{t}) = f(\mathbf{t}), \forall \mathbf{t} \in T$.*

A communication protocol induces *rectangles*, corresponding to the information reveal by that protocol. We define a *rectangle* of M^f to be a submatrix of M^f. Formally:

Definition 7.4 (Protocol induced rectangle) *Let p_f be a communication protocol that implements some social choice function f. The p_f-*induced rectangle for a type profile $\mathbf{t} \in T$, *denoted by $R_f^p(\mathbf{t})$, is the maximal submatrix S of M^f containing $\mathbf{t}$ such that the run of p_f on $\mathbf{t}$ is indistinguishable for any $\mathbf{t}' \in S$,[3] i.e., $p_f(\mathbf{t}) = p_f(\mathbf{t}')$.*

Recall from Definition 7.1 that the ideal monochromatic region represents the minimum information required to compute a social choice function f. This means that if a protocol p_f implements f, then the information elicited by p_f is at least as much as the minimum, i.e., $R_f^p(t) \subseteq R_f^I(t)$. Feigenbaum et al. [2010] use the ratio of the sizes of the ideal (maximal) regions of f and the regions (rectangles) induced by p_f to characterize the degree to which p_f discloses *extraneous* private information.[4] We present the definitions using two agents with type vector $\mathbf{t} = (t_1, t_2)$ (as in Feigenbaum et al. [2010]), though they extend to n agents in the obvious way (see below):

[3]The fact that indistinguishable regions of p_f must be rectangles is a consequence of the communication model [Kushilevitz and Nisan, 1996] (e.g., see later in Figure 7.1).

[4]These notions corresponds to the *objective privacy approximation ratios* defined in [Feigenbaum et al., 2010].

Definition 7.5 (Privacy approximation ratio) *The* worst case privacy approximation ratio (WPAR) *of a protocol p_f for a social choice function f is:*

$$\mathsf{wpar}(p_f) = \max_{(t_1,t_2)\in T} \frac{|R_f^I((t_1,t_2))|}{|R_f^p((t_1,t_2))|}.$$

Let D be a distribution over T. The average privacy approximation ratio (APAR) *of a protocol p_f for a social choice function f w.r.t. D is:*

$$\mathsf{apar}(p_f) = E_{(t_1,t_2)\sim D}\left[\frac{|R_f^I((t_1,t_2))|}{|R_f^p((t_1,t_2))|}\right].$$

We can think of perfect privacy as revealing *only enough information* about the type profile of the agents to compute a social choice function f (i.e., reveal only the ideal region). These ratios (PARs) then measures how much *additional* information a protocol p_f reveals about the type vector (in the worst case, or on average given some distribution over types). A smaller PAR indicates that p offers a greater degree of privacy, with the smallest PAR value of 1 meaning that p offers *perfect privacy*. A PAR value of $k > 1$ means that (either in the worst or average case) the protocol learns that the joint type lies in a region that is k times smaller than required to compute f.

Sandholm and Brandt [2008] show that for SPAs, the English protocol is the only perfect privacy preserving protocol for two bidders, though it bears exponential communication cost; furthermore, perfect privacy is not possible for $n > 2$ bidders. Feigenbaum et al. [2010] demonstrate an interesting trade-off between privacy and communication complexity in two-bidder SPAs by analyzing English, sealed-bid and bisection protocols. We discuss these results below when defining approximate versions of these protocols.

The work described above studies the trade-off between efficiency and communication, and the trade-off between privacy and communication. A third natural trade-off suggests itself, namely, that between efficiency and privacy. We exploit the notion of approximate solution [Blumrosen and Nisan, 2002, Hyafil and Boutilier, 2007] and show how it can be used to

improve the privacy approximation ratios of Feigenbaum et al. [2010]: that is, how much additional privacy can be preserved if we allow an ε sacrifice in efficiency. We first define ε-*approximation* and ε-*implementation*:

Definition 7.6 (ε-approximation and ε-implementation) *A social choice function $\tilde{f}$ is said to* ε-approximate *another social choice function f if $|\sum_i v_i(f(\mathbf{t}), t_i) - \sum_i v_i(\tilde{f}(\mathbf{t}), t_i)| \leq \varepsilon$, $\forall \mathbf{t} \in T$. If a protocol $p_{\tilde{f}}$ implements such an $\tilde{f}$, then we say $p_{\tilde{f}}$* ε-implements f.

In other words, $\tilde{f}$ (and any corresponding protocol $p_{\tilde{f}}$) approximates f if the difference in the social welfare between the two is no more than ε for any type profile, where the social welfare is defined as the sum of agent utilities.

We can now introduce privacy approximation ratios relative to approximate implementations of a social choice function f.

Definition 7.7 (Approximate privacy approximation ratio) *Let $p_{\tilde{f}}$ be a communication protocol that ε-implements f with social choice function $\tilde{f}$. The* ε-worst case privacy approximation ratio *of $p_{\tilde{f}}$ is:*

$$\varepsilon\text{-}\mathsf{wpar}(p_{\tilde{f}}) = \max_{\mathbf{t} \in T} \frac{|R_f^I(\mathbf{t})|}{|R_{\tilde{f}}^p(\mathbf{t})|}.$$

Let D be a distribution over T. The ε-average case privacy approximation ratio *of $p_{\tilde{f}}$ is:*

$$\varepsilon\text{-}\mathsf{apar}(p_{\tilde{f}}) = E_D \left[\frac{|R_f^I(\mathbf{t})|}{|R_{\tilde{f}}^p(\mathbf{t})|} \right].$$

These definitions are similar to those in Definition 7.5 except that we compare the ideal monochromatic regions of a social choice function f to the regions (or rectangles) induced by a protocol for its ε-*approximation* $\tilde{f}$. Our definitions in fact reduce to Definition 7.5 when $\tilde{f} = f$ and thus $\varepsilon = 0$. As above, smaller values of ε-par indicate a greater degree of privacy preservation. Unlike exact par which has a minimum value of 1 (perfect privacy), ε-par *can be less than 1*, indicating that strictly less information than required for computing f is revealed.

Indeed, this is only possible because of approximation. While both measures are interesting, we believe the average case measure ε-apar (using appropriate distributions in specific applications) may be more useful in practice.

We note that these definitions can be recast to minimize ε-par over *all* ε-implementations of f, measuring the trade-off *inherent* in f; but we focus here on the analysis of specific families of protocols. Mechanisms for specific problems, e.g., SPAs, can be parameterized by the degree of approximation ε they offer, especially by limiting the precision with which agents reveal their valuations, hence improving ε-par by sacrificing efficiency. We now explore this trade-off.

7.3 Trade-offs in Second Price Auctions

We illustrate the usefulness of our framework by analyzing the efficiency-privacy trade-off for approximate versions of three mechanisms used in *second price auctions*, the English auction, the sealed-bid auction, and the bisection auction. Our contributions are two-fold: first, we generalize the two-agent analysis of Feigenbaum et al. [2010] by providing privacy approximation ratios (or bounds) for n-agent SPAs (whose analysis is somewhat more involved). Second, we demonstrate the additional privacy savings obtained by admitting approximate efficiency.

Consider a setting with n agents, and each agent $i \leq n$ has a valuation v_i for some item. Let $v[h]$ be the h-th highest valuation in (multiset) $V = \{v_1, \cdots, v_n\}$ and $a[h]$ the agent with valuation $v[h]$ (ties broken lexicographically). The SPA allocates the item to $a[1]$ for price $v[2]$. We consider the following three protocols that implement the second price auction:

Definition 7.8 (Sealed-bid auction (for SPAs)) *The* sealed-bid mechanism *is a one-shot protocol for SPAs: each agent submits its valuation to the mechanism, which awards the item to the agent with the highest bid at the price of the second highest bid. Ties (i.e., when more than one agents have the same highest price) are broken lexicographically.*

Definition 7.9 (English auction (for SPAs)) *The* English auction *is incremental: a (small) bid increment δ is chosen, and the asking price p is raised by δ at each round. Each agent i can drop out in any round (however, an agent will drop out when $p > v_i$ strategically). The item is awarded to the last remaining agent, at the current asking price. Ties (i.e., when more than one agent drops out at the last round) are broken lexicographically (at the prior price, which all final agents "accepted").*

Definition 7.10 (Bisection auction (for SPAs)) *The* bisection auction *[Grigorieva et al., 2007] uses a binary search (asking each i whether her value v_i is above specific values) to determine a value $b \in (v[1], v[2]])$ (and $b = v[1]$ if $v[1] = v[2]$). Once b is identified, binary search on the interval containing $v[2]$ is used to identify $v[2]$ to a desired precision σ. Note once $v_i < v[2]$ is proven for some agent i, no further queries are asked.*

Following Feigenbaum et al. [2010], we treat the valuation space as discrete, representable with k bits, allowing $\nu = 2^k$ distinct valuations. We assume, w.l.o.g., that $v_i \in \mathbf{V}^k = \{0, \ldots, 2^k - 1\}$. In the following, we will analyze the privacy approximation ratio (or bounds) for these three protocols, when approximate implementation of the social choice function is used.

English Protocol The English protocol with an "exact" bid increment $\delta = 1$ has exponential communication complexity $O(2^k)$ [Sandholm and Brandt, 2008]: simply consider the case of $v[2] = 2^k - 1$. But this high cost allows for very strong privacy: for two agents, par in both the worst case and average case is 1, i.e., it is perfectly privacy preserving. For $n > 2$ agents, perfect privacy is not possible [Sandholm and Brandt, 2008]. The thin line in Figure 7.1 illustrates the ideal monochromatic partition for a two-agent SPA.

We can approximate the English auction by simply increasing the bid increment, setting $\delta = 1 + \varepsilon = 2^d$ for some precision integer $d > 0$.[5] Clearly this *ε-English protocol*, denoted by p_E^{ε}, ε-approximates the exact SPA, with suboptimal allocation happening only when multiple

[5] We use powers of 2 for convenience only.

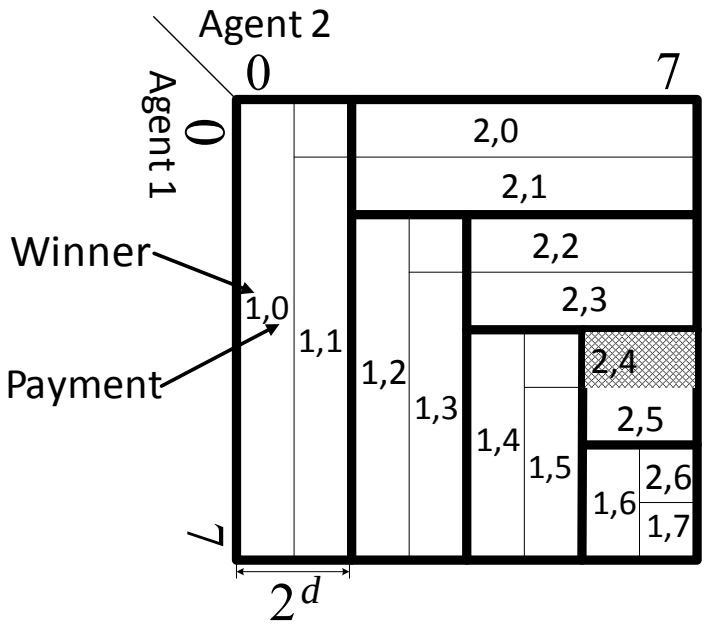


Figure 7.1: Partitions induced by the English auction for 2-bidder SPAs when $\delta = 1$ ($\varepsilon = 0$, thin line) and $\delta = 2$ ($\varepsilon = 1$, thick line). When $\delta = 1$, this is also the ideal monochromatic partition. The shaded region indicates the inputs from which ε-wpar is derived. The numbers indicate the outcome for each ideal rectangle (e.g., in the leftmost rectangle, the item is allocated to agent 1 for a price of 0).

agents drop out at the last round; but all such agents have values within an interval of size $(1 + \varepsilon)$, guaranteeing ε-efficiency. The price paid is also within ε of that dictated by the exact SPA, and p_E^ε is ε-incentive compatible (in fact, ε-strategy-proof). The thick line in Figure 7.1 illustrates the protocol-induced partition for the ε-English auction when $\varepsilon = 1$. Notice that for some type profiles, the outcome is different from that in the exact protocol (e.g., with profile $(2, 3)$, p_E^ε allocates the item to agent 1 for a price of 2, while the exact protocol allocates efficiently to agent 2 for a price of 2). It is easy to verify that, for any type profile $\mathbf{t}$, the protocol induced rectangle for p_E^ε is at least as large as that induced by the exact English protocol, indicating privacy savings. The shaded area denotes the profiles from which we derive ε-wpar: the ideal monochromatic region has size 3 while the protocol-induced rectangle has size 4. Note that ε-wpar $= \frac{3}{4} < 1$, indicating *better than perfect privacy*.

These intuitions can be generalized to n-agent SPAs, where we have:

Theorem 7.1 (ε-worst case privacy approximation ratio of ε-English protocol (for SPAs)) *For n-agent SPAs, the ε-worst case privacy approximation ratio of ε-English protocol (for large enough n) is:*

$$\varepsilon\text{-}\mathsf{wpar}_p(p_E^\varepsilon) = \frac{(2^k)^{n-1} - (2^k - 1)^{n-1}}{(1+\varepsilon)^n}$$

Proof: See appendix of this chapter. ■

The worst case occurs when the valuations of all agents are $2^k - 1$, such that the size of the ideal monochromatic region is maximized and the size of the protocol induced rectangle is minimized. Relative to the exact protocol in which $\varepsilon = 0$, worst-case privacy savings of p_E^ε are $(1+\varepsilon)^n$, as one would expect ($1+\varepsilon$ per agent). It is important to note that privacy loss (or savings) exponential in n is an artifact of the definitions: if each of n agents gives up a bounded amount of privacy κ relative to the ideal region, the "product" of their losses is κ^n. Hence this should not be viewed as especially problematic here or in the sequel.

Next, we consider the average case. Suppose we have a uniform distribution D over type profiles (all average-case analysis in the sequel uses the uniform distribution D). We can bound the ε-average case privacy approximation ratio ε-apar of the ε-English protocol:

Theorem 7.2 (ε-average case privacy approximation ratio of ε-English protocol (for SPAs)) *For n-agent SPAs, the ε-average case privacy approximation ratio of ε-English protocol is:*

$$2\frac{(2^{k-1})^{n-3}}{\binom{n-3}{\lfloor \frac{n}{2} \rfloor}(1+\varepsilon)^n} \leq \varepsilon\text{-}\mathsf{apar}(p_E^\varepsilon) \leq 2\binom{n}{2}\frac{(2^k)^{n-2}}{(1+\varepsilon)^{n-1}}$$

Proof: See appendix of this chapter. ■

In the ε-English protocol, the valuations of at least $n-1$ agents are identified with precision $1+\varepsilon$, so privacy savings are at between $(1+\varepsilon)^{n-1}$ and $(1+\varepsilon)^n$ relative to exact implementation. Exact average case savings are computed numerically below.

Bisection Protocol A natural way to approximate the bisection protocol is to use early termination, stopping when we identify $v[2]$ with some desired precision σ (i.e., when the bisection interval containing $v[2]$ is no larger than σ). We then allocate the item to $a[1]$ using the price at the low end of $v[2]$'s interval (with ties broken lexicographically). To ensure ε-efficiency, thus defining the *ε-bisection protocol* p_B^ε, we must have $\sigma \leq \varepsilon + 1$ (otherwise the mechanism is not ε-efficient). This mechanism is also ε-incentive compatible (an agent's gain by misreporting is at most ε), and the ε-worst case privacy approximation is as follows:

Theorem 7.3 (ε-worst case privacy approximation ratio of ε-bisection protocol (for SPAs)) *For n-agent SPAs, the ε-worst case privacy approximation ratio of ε-bisection protocol is:*

$$\varepsilon\text{-}\mathsf{wpar}(p_B^\varepsilon) = \frac{(2^k)^{n-1} - (2^k - 1)^{n-1}}{(1+\varepsilon)^n}$$

Proof: See appendix of this chapter. ■

The worst case privacy approximation ratio under the ε-bisection protocol is exactly the same as that under the ε-English protocol, which occurs when all agents have values clustered at $v[2] = 2^k - 1$: each reports its valuation with the maximum precision, so ε-wpar is exponential in both k and n. The privacy savings of p_B^ε relative to exact implementation are precisely $(\varepsilon + 1)^n$.

For the average case, we have the following bounds:

Theorem 7.4 (ε-average case privacy approximation ratio of ε-bisection protocol (for SPAs)) *For n-agent SPAs,*

$$\frac{nk}{(1+\varepsilon)^n} \leq \varepsilon\text{-}\mathsf{apar}(p_B^\varepsilon) \leq n(n-1)\frac{(k)^{n-1}}{1+\varepsilon}$$

Proof: See appendix of this chapter. ■

We see that apar for (exact and approximate) bisection is polynomial in k (and exponential in n), which compares favorably to the English protocol (exponential in both k and n). De-

pending on the number of agents whose values fall in the bisection interval containing $v[2]$ at termination, the privacy savings for p_B^ε range from $\varepsilon+1$ to $(\varepsilon+1)^n$. Exact average case savings are computed numerically below, and compared with those for the ε-English protocol p_E^ε.

Sealed-Bid Protocol The *ε-sealed-bid protocol* p_S^ε approximates the exact sealed-bid protocol by simply "coarsening" the valuation space, asking for reports v_i with limited precision σ. The bound $\sigma \leq 1+\varepsilon$ also holds for p_S^ε, requiring termination only when $v[2]$ is known to lie within an interval of length $1+\varepsilon$. The ε-worst case approximation ratio ε-wpar of p_S^ε in n-agent SPAs is identical to that for ε-English and ε-bisection, since it induces the same size of rectangle in the worst case.

Theorem 7.5 (ε-worst case privacy approximation ratio of ε-sealed-bid protocol (for SPAs)) *For n-agent SPAs, the ε-worst case privacy approximation ratio of ε-sealed-bid protocol is:*

$$\varepsilon\text{-}\mathit{wpar}(p_S^\varepsilon) = \frac{(2^k)^{n-1} - (2^k-1)^{n-1}}{(1+\varepsilon)^n}$$

Proof: See appendix of this chapter. ■

Again, the worst case occurs when all agents have valuations 2^k-1, and the size of the ideal monochromatic region is $(2^k-1)^{n-1} - (2^k)^{n-1}$, and the size of the protocol induced rectangle is exactly $(1+\varepsilon)^n$. The privacy savings compared to the exact sealed-bid protocol are also $(1+\varepsilon)^n$ ($1+\varepsilon$ per agent).

Despite the same worst case behavior, ε-sealed-bid protocol is much worse on average than ε-bisection protocol:

Theorem 7.6 (ε-average case privacy approximation ratio of ε-sealed-bid protocol (for SPAs)) *For n-agent SPAs, the ε-average case privacy approximation ratio of ε-sealed-bid protocol is:*

$$2\frac{(2^{k-1})^{n-3}}{\binom{n-3}{\lfloor \frac{n}{2} \rfloor}(1+\varepsilon)^n} \leq \varepsilon\text{-}\mathit{apar}(p_S^\varepsilon) \leq \frac{(2^k)^{n-1} - (2^k-1)^{n-1}}{(1+\varepsilon)^n}$$

ε	Second Price Auctions		
	$n=3$	$n=4$	$n=5$
$\varepsilon=0$	32 / 15 / 410	1225 / 72.4 / 11254	46563 / 350 / 333760
$\varepsilon=1$	8.1 / 5.11 / 51.25	156 / 20.7 / 703.4	2942 / 86 / 10430
$\varepsilon=3$	2.05 / 1.56 / 6.4	19 / 5 / 44	173 / 17.3 / 325.9
$\varepsilon=7$	0.48 / 0.4 / 0.8	1.95 / 0.96 / 2.75	8.0 / 2.6 / 10.2

Table 7.1: ε-apar for SPAs with different n and ε when $k=5$ bits. The three values in each cell indicate ε-apar for the ε-English, ε-bisection and ε-sealed-bid protocols, respectively.

Proof: See appendix of this chapter. ■

Our current lower bound for ε-apar(p_S^ε) is quite loose; but we can use the lower bound for ε-apar(p_E^ε) instead: for each profile, the size of the protocol induced rectangle for p_E^ε is at least as large as that for p_S^ε (because the valuation of the highest bid is only revealed in some extent), so we must have $R^p_{p_S^\varepsilon}(\mathbf{t}) \subseteq R^p_{p_E^\varepsilon}(\mathbf{t}), \forall \mathbf{t}$. According to our definition of ε-apar, this means ε-apar$(p_S^\varepsilon) \geq \varepsilon$-apar$(p_E^\varepsilon)$. Hence, ε-apar(p_S^ε) is exponential in both k and n. In addition, since the size of all induced rectangles is $(\varepsilon+1)^n$, in both the worst and average case, p_S^ε offers privacy savings of $(\varepsilon+1)^n$ over exact sealed-bid.

Summary The average case ε-privacy approximation ratios for SPAs of different sizes, computed numerically, are shown in Table 7.1.[6] Recalling that smaller ε-apar indicates better privacy, we see that our ε-approximate protocols offer significant privacy savings relative to their exact counterparts. For instance, when $n=3$ and $\varepsilon=1$, the ε-English protocol reveals $\frac{8.1}{32} \approx \frac{1}{4}$ of the information revealed by the exact protocol, while $\varepsilon=3$ requires only $\frac{1}{16}$ of information. We also see that ε-bisection has the smallest ε-apar, preserving much more privacy than either ε-English or ε-sealed-bid; e.g., when $\varepsilon=3$ and $n=4$, ε-bisection requires revelation of only $\frac{5}{19}$ and $\frac{5}{44}$ of the information required by ε-English and ε-sealed-bid, respectively. We also notice that ε-apar, and the privacy savings of the approximate protocols over their exact counterparts, grow exponentially with n. This is consistent with our theoretical re-

[6]For each type profile $\mathbf{t}$, we first derive the size of the ideal monochromatic region (see Claim 7.2 in the appendix of this chapter), run simulation of a particular protocol to get the size of the protocol induced rectangle, and compute their ratios. Then the average PAR is the weighted sum over all possible type profiles.

sults. Moreover, though our current lower bound for ε-bisection is linear in ε, these numerical results suggest that the actual savings are much greater. We conjecture that the true savings area $O((1+\varepsilon)^{\frac{n}{c}})$, for some constant $c > 1$.

To summarize, we have derived privacy approximation ratios for the n-agent versions of three key protocols for SPAs. We have also shown that approximate variants of these protocols allow for savings in privacy over their exact counterparts that is exponential in the number of agents n and polynomial in the degree of approximation ε in almost all cases (both worst and average case).

7.4 Tradeoffs in Facility Location

We now consider another classic domain in mechanism design, one-dimensional, single facility location problems as defined in Section 2.3. Following notation in Section 2.3, we assume there are n agents, each with an ideal location t_i in a finite set of possible locations $\mathbf{L} = \{0, \ldots, 2^k - 1\}$. For ease of exposition, we assume an odd number of agents $n = 2m - 1$; we also assume (w.l.o.g.) that agents are sorted by preferred location: $t_1 \leq \ldots \leq t_n$. The objective is to select an optimal location of the facility that maximizes social welfare by minimizing the social choice function $f(x) = \sum_i c(x, t_i) = |x - t_i|$, i.e., the sum of distance between the ideal location of each agent and the facility.

The median mechanism (as in Definition 2.25), which asks each agent i to report her most preferred location t_i and locates the facility at the median t^M of the reported values, is a strategy-proof mechanism that selects the optimal location. Generally direct elicitation of the ideal locations (i.e., sealed-bid) is used to implement such a mechanism; however, incremental elicitation can be accomplished using mechanisms much like those for SPAs. We define the following two incremental mechanisms:

Definition 7.11 (English protocol (for FLPs)) *We define an* English protocol for FLPs *as follows: beginning with a* current location $p = 0$*, we increment* p *by* $\delta = 1$*, asking* i *if* $t_i \geq p$*,*

stopping when at least m agents have dropped out, thereby identifying the median.

Definition 7.12 (Bisection protocol (for FLPs)) *The* bisection protocol for FLPs *simply conducts a binary search to find the median t^M. Starting from an asking position $t^p = 2^k - 1$, we ask for each agent if $t_i \geq t^p$ or $t_i \leq t^p$, and update t^p to be the middle of the interval containing the median t^M. At any stage, if we know $t_i \geq t^M$ or $t_i \leq t^M$, agent i is asked no further queries. The process is repeated until t^M has been identified.*

Approximate versions of both protocols (as well as "sealed-bid") are defined analogously to the case of SPAs.

Before describing our results regarding privacy approximation ratios, we first provide a general negative result showing that there is no perfect privacy preserving protocol for the median mechanism. The intuition is that any protocol requires the revelation of the identity of an agent with the median value in at least *some* instances. We state the results formally in the following theorem:

Theorem 7.7 (Non-existence of perfect privacy preserving protocols for FLPs) *There is no perfect privacy preserving protocol for the median mechanism for n-party FLPs, for any $n \geq 2$.*

Proof: We show the proof for the case of $n = 2$, but the analysis can be generalized to any number of agents. Consider a two-agent facility location problem, where we use the leftmost mechanism (as in Example 2.5) to locate the facility. The leftmost mechanism can be viewed as a median mechanism (as in Definition 2.25) for the two-agent case, and is both strategy-proof and efficient for the two-agent FLPs.

Figure 7.4 shows the ideal monochromatic partition for the two-agent FLPs. Clearly, except the bottom-right one, no other ideal monochromatic regions are rectangles. However, the communication model requires that the indistinguishable regions of any communication protocol must be rectangles [Kushilevitz and Nisan, 1996], indicting that there is no perfect privacy preserving protocol. ■

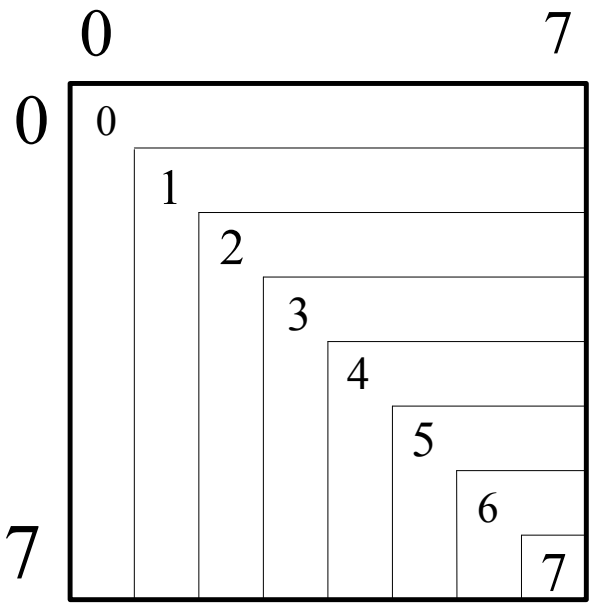


Figure 7.2: Ideal monochromatic partition for 2-agent FLPs.

Before computing the privacy approximation ratio, we first analyze the size of the ideal monochromatic region for a location profile. We have the following:

Claim 7.1 *Let* $\mathbf{t}$ *be a type profile with median* t^M*. Then the size of the ideal monochromatic region is:*

$$R_f^I(\mathbf{t}) = \sum_{r=0}^{m-2} \binom{n}{r+1} [\sum_{s=0}^{r} \binom{n-1-r}{m-1-r+s} \cdot (t^M)^{m-1-r+s} \cdot (2^k - 1 - t^M)^{m-1-s}] + \sum_{r=m}^{n} \binom{n}{r} (2^k - 1)^{n-r}$$

The first term reflects the case when fewer than m agents have an ideal location that coincides with the median t^M, and the second when at least m agents have the ideal location t^M.

It should be noted that the size of the ideal monochromatic region for FLPs is a function $Z(t^M)$ of t^M, not the entire profile $\mathbf{t}$. If we denote:

$$Z(r) = O(r^{m-1}(2^k - 1 - r)^{m-1}),$$

then one can check that the maximum is attained when $r = 2^{k-1}$ or $r = 2^{k-1} - 1$ and minimum is attained when $r = 0$ or $r = 2^k - 1$. This property will be used when proving both worst case and average case PARs for all three protocols.

Next, we provide the ε-worse and ε-average case privacy approximation ratios (or bounds) for n-agent FLPs for each of the three protocols: English, bisection and sealed-bid.

English Protocol We first analyze the worst case privacy approximation ratio of the exact English protocol for FLPs (i.e., where $\delta = 1$): Each agent is asked by the mechanism that whether their valuation is higher or lower than the current asking location p. If $n_1 \leq m - 1$ agents drop out at location $p - 1$ and $n_2 \geq m$ agents drop out at location p, then the median value is exactly p. The result is shown in the following theorem:

Theorem 7.8 (Worst case privacy approximation ratio of English protocol (for FLPs)) *Let p_{EF} be the exact English protocol for n-agent FLPs. Then the worst case privacy approximation ratio is:*

$$Z(0) \leq \mathsf{wpar}(p_{EF}) \leq Z(2^{k-1})$$

Proof: See appendix of this chapter. ■

While we are unable to get an exact wpar value for the English protocol, the lower bound $Z(0)$ means that wpar is exponential in both k and n, indicating weak privacy guarantees.

The ε-English protocol p^{ε}_{EF} uses a bid increment $\delta > 1$, identifying the median with precision δ when the protocol stops, and randomly selecting a location within this δ-interval. To ensure ε-approximation, δ cannot be too large:

Lemma 7.1 *Let p^{ε}_{EF} be the ε-English protocol for n-agent FLPs. Then p^{ε}_{EF} ε-implements the median mechanism only if the bid increment δ satisfies $\delta \leq 1 + \frac{\varepsilon}{n}$.*

Proof: We prove this lemma by contradiction. Let p^{ε}_{EF} be an ε-English protocol that ε-implements the median mechanism, with a bid increment of $\delta > 1 + \frac{\varepsilon}{n}$. Consider the type

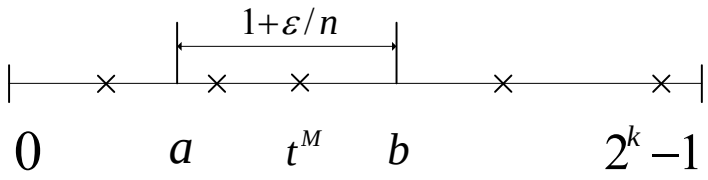


Figure 7.3: English protocol for facility location problem

profile shown in Figure 7.4, in which the median is t^M. Let a be the location that $n_1 \leq m-1$ agents drop out and $b = a + \delta$ be the location that $n_2 \geq m$ agents drop out. Then the ε-English protocol will locate the facility at location a, and the induced social cost is:

$$SC(a) = \sum_{i=1}^{n_1}(t_i - a) + \sum_{i=n_1+1}^{n}(t_i - a) \tag{7.1}$$

However, if the facility is located at the median position t^M, then the social cost is:

$$SC(t^M) = \sum_{i=1}^{m-1}(t^M - t_i) + \sum_{i=m+1}^{n}(t_i - t^M) \tag{7.2}$$

Subtracting Equation 7.2 from Equation 7.1, we have:

$$SC(a) - SC(t^M) = \sum_{i=n_1+1}^{m-1}(2t_i - 2a) + (t^M - a) \geq n(\delta - 1) > \varepsilon$$

which contradicts our assumption that p^{ε}_{EF} ε-implements the median mechanism, completing our proof. ■

The distinction with SPAs, which allow increments of $(1+\varepsilon)$, is due to the fact that an ε-misplacement of the facility can impact all n agents (not just the winner as in SPAs). The mechanism is $\frac{\varepsilon}{n}$-incentive compatible.

By Theorem 7.8 and Lemma 7.1, we have:

Corollary 7.1 (ε-worst case privacy approximation ratio of ε-English protocol (for FLPs))

For n-agent FLPs, the ε-worst case privacy approximation ratio of ε-English protocol satisfies:

$$\lim_{n\to\infty} \varepsilon\text{-}\mathsf{wpar}(p^{\varepsilon}_{EF}) \leq \lim_{n\to\infty} \frac{Z(2^{k-1})}{(1+\frac{\varepsilon}{n})^m} = \frac{Z(2^{k-1})}{e^{\varepsilon/2}}$$

and

$$\lim_{n\to\infty} \varepsilon\text{-}\mathsf{wpar}(p^{\varepsilon}_{EF}) \geq \lim_{n\to\infty} \frac{Z(0)}{(1+\frac{\varepsilon}{n})^n} = \frac{Z(0)}{e^{\varepsilon}}$$

For the ε-English protocol, at least m and at most n agents' locations are identified with a precision of $1+\frac{\varepsilon}{n}$, so the privacy savings are between $(1+\frac{\varepsilon}{n})^m$ and $(1+\frac{\varepsilon}{n})^n$, which converges to $e^{\varepsilon/2}$ and e^{ε} as $n \to \infty$. Similarly, the $Z(0)$ term in the lower bound indicates that ε-wpar is exponential in both k and n.

For the average case, we begin our analysis with the exact protocol, providing upper and lower bounds:

Theorem 7.9 (Average case privacy approximation ratio of English protocol (for FLPs)) *Let p_{EF} be the exact English protocol for n-agent FLPs. Then the worst case privacy approximation ratio is:*

$$m\binom{m-1}{m/2}(2^{k-1})^{m-2} \leq \mathsf{apar}(p_{EF}) \leq m\binom{n}{m-1}(2^k)^{m-1}$$

Proof: See appendix of this chapter. ■

This result allows us to show that the average case privacy savings of p^{ε}_{EF} relative to exact p_{EF} are at most $(1+\frac{\varepsilon}{n})^m$. However, in the ε-English protocol p^{ε}_{EF}, we "coarsen" the revealed locations of at least m and at most n agents, which means the saving is between $(1+\frac{\varepsilon}{n})^m$ and $(1+\frac{\varepsilon}{n})^n$, and converges to $e^{\varepsilon/2}$ and e^{ε} as $n \to \infty$. These values can be multiplied by the terms in the bounds of Theorem 7.9 to derive bounds on ε-$\mathsf{apar}(p^{\varepsilon}_{EF})$.

Bisection Protocol We now consider the bisection protocol p_{BF} for FLPs and analyze its privacy approximation ratios before considering its ε-approximate implementation. We first consider the worst case PAR for p_{BF}:

Theorem 7.10 (Worst-case privacy approximation ratio of bisection protocol (for FLPs)) *Let p_{BF} be the exact bisection protocol for n-agent FLPs. Then the worst case privacy approximation ratio is:*

$$\mathsf{wpar}(p_{BF}) = Z(2^{k-1})$$

Proof: See the appendix of this chapter. ■

This worst case occurs with a type profile in which $m-1$ agents have ideal location $2^{k-1}+1$ and m agents prefer location 2^{k-1}: the p_{BF}-induced rectangle has size 1 while the ideal monochromatic region has size $Z(2^{k-1})$. Similarly, the $Z(2^{k-1})$ term indicates that wpar is exponential in both k and n.

The approximate *ε-bisection protocol* p^{ε}_{BF} for FLPs identifies the median only to some desired precision, but uses a *dynamic precision* parameter to determine termination. Specifically, we terminate when the median is proven to lie in some interval $[t^M_-, t^M_+)$, and a random point in that interval is selected for the facility. The mechanism is $\frac{\varepsilon}{n}$-incentive compatible. To ensure ε-efficiency, we require:

Lemma 7.2 *Let l and r be the number of agents in $[t^M_-, t^M_+)$ whose desired location is left of (less than) and right of (greater than) of t^M, respectively. Then p^{ε}_{BF} ε-implements the median mechanism for FLPs iff $(t^M_+ - t^M_- - 1)(2\max\{l,r\}+1) \leq \varepsilon$.*

Proof: Let t^M_- and t^M_+ be the lower and upper bound of the interval containing t^M, and suppose by contradiction that the condition does not hold, i.e., $(t^M_+ - t^M_- - 1)(2\max\{l,r\}+1) >$

ε. Then if t_-^M is selected as the median, we have:

$$\begin{aligned} SC(t_-^M) - SC(t^M) = & \sum_{i=1}^{m-1-l}(t_-^M - t_i) + \sum_{i=m-l}^{n}(t_i - t_-^M) \\ & - \left(\sum_{i=1}^{m-1}(t^M - t_i) + \sum_{i=m+1}^{n}(t_i - t^M)\right) \\ \geq & (2l+1)(t_+^M - t_-^M - 1) \\ > & \varepsilon \end{aligned}$$

which contradicts with our assumption that p_{BF}^{ε} ε-implements the median mechanism, so we must have $(2l+1)(t_+^M - t_-^M - 1) \leq \varepsilon$. On the other hand, if t_+^M is selected as the median, we can derive $(2r+1)(t_+^M - t_-^M - 1) \leq \varepsilon$, using similar analysis. This completes our proof. ■

This means the "precision" of the final interval $[t_-^M, t_+^M)$ is determined by p_{BF}^{ε} dynamically: if, when the median value interval is identified, no other agents' locations lie within $[t_-^M, t_+^M)$, the protocol can stop when the interval is narrowed to $t_+^M - t_-^M \leq 1 + \varepsilon$; but if $m-1$ agents remain in the interval, and are left of t^M, then the protocol stops only when $t_+^M - t_-^M \leq 1 + \frac{\varepsilon}{n}$. This mechanism is also $\frac{\varepsilon}{n}$-incentive compatible. By Theorem 7.10 and Lemma 7.2, we have the following corollary for the ε-bisection protocol:

Corollary 7.2 (ε-worst case privacy approximation ratio of ε-bisection protocol (for FLPs)) *For n-agent FLPs, the ε-worst case privacy approximation of ε-bisection protocol satisfies:*

$$\lim_{n\to\infty} \varepsilon\text{-}\mathit{wpar}(p_{BF}^{\varepsilon}) = \lim_{n\to\infty} \frac{Z(2^{k-1})}{(1+\frac{\varepsilon}{n})^n} = \frac{Z(2^{k-1})}{e^{\varepsilon}}$$

For average case analysis, we again begin with exact bisection protocol, providing upper and lower bounds in the following theorem:

Theorem 7.11 (Average-case privacy approximation ratio of bisection protocol (for FLPs)) *Let p_{BF} be the ε-bisection protocol for n-agent FLPs. Then the ε-average case privacy approx-*

imation ratio is:

$$\binom{n}{m}k^{m-1} \leq \textsf{apar}(p_{BF}) \leq m\binom{n}{m}k^m$$

Proof: See appendix of this chapter. ■

As with SPAs, apar for bisection in FLPs is polynomial in k, offering significant privacy savings relative to the English protocol. With ε-approximation, we can show that the privacy savings range from $\varepsilon + 1$ to $(\varepsilon + 1)^n$, depending on the number of agents whose locations fall into the bisection interval as t^M. We compare average case savings numerically across these different protocols below.

Sealed-Bid Protocol The sealed-bid protocol p_{SF} for FLPs has each agent reveal her preferred location and returns the median. So the size of the protocol induced rectangle is always 1 regardless of the type profile.

We first give the worst-case privacy approximation ratio in the following theorem:

Theorem 7.12 (Worst-case privacy approximation ratio of sealed-bid protocol (for FLPs)) *Let p_{SF} be the exact sealed-bid protocol for n-agents FLPs. Then the worst case privacy approximation ratio is:*

$$\textsf{wpar}(p_{SF}) = Z(2^{k-1})$$

Proof: See appendix of this chapter. ■

The worst also occurs when $m - 1$ agents have ideal location $2^{k-1} + 1$ and m agents have ideal location 2^{k-1}, as in the exact bisection protocol.

The ε-sealed-bid protocol p^{ε}_{SF} asks for locations with limited precision σ. In the worst case, when all locations lie in the interval of t^M, Lemma 7.2 needs precision $\sigma \leq 1 + \frac{\varepsilon}{n}$, and ε-wpar(p^{ε}_{SF}) is identical to that for p^{ε}_{BF}.

ε	Facility Location Problems	
	$n=3$	$n=5$
$\varepsilon=0$	96 / 42 / 1228	8776 / 1514 / 1.50E+06
$\varepsilon=10$	6.1 / 3.6 / 19.2	374 / 154 / 46766
$\varepsilon=15$	3.0 / 3.6 / 19.2	152 / 61 / 1461
$\varepsilon=22$	1.34 / 0.97 / 2.4	86 / 36 / 1461

Table 7.2: ε-apar for FLPs with different n and ε when $k=5$ bits. The three values in each cell indicate ε-apar for the ε-English, ε-bisection and ε-sealed-bid protocols, respectively.

Corollary 7.3 (ε-worst case privacy approximation ratio of ε-sealed-bid protocol (for FLPs))

For n-agents FLPs, the ε-worst case privacy approximation ratio of ε-sealed-bid protocol is:

$$\lim_{n\to\infty} \varepsilon\text{-}wpar(p^{\varepsilon}_{SF}) = \lim_{n\to\infty} \frac{Z(2^{k-1})}{(1+\frac{\varepsilon}{n})^n} = \frac{Z(2^{k-1})}{e^{\varepsilon}}$$

For the average case, we have upper and lower bounds on apar for exact sealed-bid protocol p_{SF}:

Theorem 7.13 (Average case privacy approximation ratio of sealed-bid protocol (for FLPs))

Let p_{SF} be the exact sealed-bid protocol for n-agents FLPs. Then the average case privacy approximation ratio is:

$$Z(0) \le apar(p_{SF}) \le Z(2^{k-1}-1)$$

Proof: See appendix of this chapter. ■

In the ε-sealed-bid protocol p^{ε}_{SF}, each rectangle has size $\frac{\varepsilon}{n}$ (compared to size 1 for exact sealed-bid protocol p_{SF}), so average privacy saving are $(1+\frac{\varepsilon}{n})^n$, converging to e^{ε} as $n\to\infty$. However, ε-apar(p^{ε}_{SF}) is still exponential in both k and n.

Summary As with SPAs above, Table 7.2 shows average case ε-privacy approximation ratios for FLPs of different sizes computed numerically. Results are similar to those for SPAs that approximation provides significant savings in privacy, and the bisection protocol offers the

greatest privacy preservation.

To summarize, we have proposed two incremental mechanisms for FLPs: the English and the bisection protocols. Together with the sealed-bid protocol, we have provided upper and lower bounds on worst- and average-case par, showing, as with SPAs, that the bisection protocol offers relatively strong privacy guarantees compared to the other two (polynomial in k and exponential in n). With ε-approximation, strong privacy savings are possible (exponential in ε as $n \to \infty$).

7.5 Conclusion

In this chapter, we have presented a framework for analyzing the natural tradeoff between efficiency and privacy in mechanism design. Within this model, we have analyzed second-price auctions and facility location problems, and for each investigated the extent to which privacy is preserved for a variety of different protocols. We have shown that the bisection protocol offers significant privacy advantages over other protocols, and also demonstrated the degree to which additional privacy preservation can be gained through ε-approximation of these protocols over their exact implementations, using both worst and average case analyses.

Our framework can be generalized in several ways. While we have presented our work in the context of mechanism design, it can be applied to any form of distributed function computation. One might also consider other forms of approximate privacy that account for, say, different sensitivity to the reports of different agents, or from different regions of type space. Our analysis can also be extended in several ways, including deriving average case results for more realistic distributions of valuations; and broadening the class of mechanisms and social choice functions. Finally, this work suggests a complicated four-way tradeoff between communication, efficiency, incentives and privacy in the design of mechanisms. Developing optimization models that explicitly trade off these criteria against one another will be important in the automated design of privacy-preserving mechanisms. Incremental mechanisms such

as those discussed here should have even greater potential to offer practical—if not (worst-case) theoretical—privacy and communications savings. We will discuss more future work in Chapter 8.

Appendix of Chapter 7

Proof of Theorems

Theorem 7.1 (ε-worst case privacy approximation ratio of ε-English protocol (for SPAs)) *For n-agent SPAs, the ε-worst case privacy approximation ratio of ε-English protocol (for large enough n) is:*

$$\varepsilon\text{-}\mathsf{wpar}_p(p_E^\varepsilon) = \frac{(2^k)^{n-1} - (2^k - 1)^{n-1}}{(1+\varepsilon)^n}$$

Proof: According to Definition 7.7, the ε-worst case PAR is the ratio between the ideal monochromatic region $R_f^I(\mathbf{t})$ and the protocol induced rectangle $R_{\tilde{f}}^p(\mathbf{t})$ for some type profile $\mathbf{t}$. We prove this theorem by computing the maximum for the numerator and the minimum for the denominator independently, and then showing they can be achieved by a same type profile.

Recall that the second price auction is a Groves mechanism with payment (as in Definition 2.19). We use $\langle i, p\rangle$ to denote an outcome, where $i \in \{1, \ldots, n\}$ is the identity of the winning agent, and p is the payment. Let $R_f^I(i,p)$ be the size of the ideal monochromatic region when the outcome is $\langle i, p\rangle$, then we have:

Claim 7.2 *For SPAs, the size of the ideal monochromatic region for an outcome $\langle i, p\rangle$ is:*

$$R_f^I(i,p) = (2^k - p - 1)\sum_{j=1}^{n-1}\binom{n-1}{j}p^{n-1-j} + p^{i-1}\sum_{j=1}^{n-i}\binom{n-i}{j}p^{n-i-j}$$

The first term corresponds to the case that the winner's valuation is strictly greater than p, and the second term corresponds to the case that the winner's valuation equals to p. Note that

as ties are broken alphabetically, so if agent i wins the item at the price p, then the valuation of all agents whose identities are less than i can only be in $[0, p-1]$.

Next we show the size of the ε-English protocol induced rectangle:

Claim 7.3 *For the ε-English protocol and the corresponding social choice function $\tilde{f}$, the size of the protocol induced rectangle is $R^p_{\tilde{f}}(i,p) = (2^k - p + \varepsilon)(1+\varepsilon)^{n-1}$ for an outcome $\langle i, p\rangle$.*

When the auction stops, the ε-English protocol has identified the valuations all agents but the winner within a precision of $(1+\varepsilon)$, and the valuation of the winner in the interval $[p - \varepsilon, 2^k - 1]$, so the protocol induced rectangle has a size of $(2^k - p + \varepsilon)(1+\varepsilon)^{n-1}$.

Now let us consider the ratio between this two. It is not hard to see that $R^I_f(i,p) \leq R^I_f(1,p)$, i.e., the size of the monochromatic region is maximized when the first agent gets the item (this is because of our tie breaking strategy), and we have $R^I_f(1,p) = (2^k - p)\sum_{j=1}^{n-1}\binom{n-1}{j}p^{n-1-j} = (2^k-p)[(p+1)^{n-1} - p^{n-1}]$. On the other hand we have $R^p_{\tilde{f}}(i,p) = (2^k - p + \varepsilon)(1+\varepsilon)^{n-1}$, so the ε-worst case privacy approximation ratio of ε-English protocol is:

$$\begin{aligned}
\varepsilon\text{-wpar}(p^{\varepsilon}_E) &= \max_{\mathbf{t}} \frac{R^I_f(\mathbf{t})}{R^p_{\tilde{f}}(\mathbf{t})} \\
&= \max_{\langle i,p\rangle} \frac{R^I_f(i,p)}{R^p_{\tilde{f}}(i,p)} \\
&\leq \frac{R^I_f(1,p)}{R^p_{\tilde{f}}(i,p)} \\
&= \frac{(2^k-p)\sum_{j=1}^{n-1}\binom{n-1}{j}p^{n-1-j}}{(2^k-p+\varepsilon)(1+\varepsilon)^{n-1}} \\
&= \frac{(2^k-p)\left[(p+1)^{n-1} - p^{n-1}\right]}{(2^k-p+\varepsilon)(1+\varepsilon)^{n-1}}
\end{aligned}$$

If we view this as a function of p, the maximum is achieved with $p = 2^k - 1$ for large enough n, and we have:

$$\varepsilon\text{-wpar}(p^{\varepsilon}_E) \leq \frac{(2^k)^{n-1} - (2^k-1)^{n-1}}{(1+\varepsilon)^n}$$

This worst case happens when the valuation of all agents are $2^k - 1$, and all the equalities above holds, completing our proof. ■

Theorem 7.2 (ε-average case privacy approximation ratio of ε-English protocol (for SPAs)) *For n-agent SPAs, the ε-average case privacy approximation ratio of ε-English protocol is:*

$$2\frac{(2^{k-1})^{n-3}}{\binom{n-3}{\lfloor \frac{n}{2} \rfloor}(1+\varepsilon)^n} \leq \varepsilon\text{-}\mathit{apar}(p_E^\varepsilon) \leq 2\binom{n}{2}\frac{(2^k)^{n-2}}{(1+\varepsilon)^{n-1}}$$

Proof: The upper and lower bounds on average case PAR for the exact protocol are computed as a special case ($m = n - 1$) of those in Theorem 7.9.

For the approximation case, the valuation of at least $n - 1$ agents are identified with a precision of $\varepsilon + 1$, so the saving is between $(1+\varepsilon)^{n-1}$ and $(1+\varepsilon)^n$, completing our proof. ■

Theorem 7.3 (ε-worst case privacy approximation ratio of ε-bisection protocol (for SPAs)) *For n-agent SPAs, the ε-worst case privacy approximation ratio of ε-bisection protocol is:*

$$\varepsilon\text{-}\mathit{wpar}(p_B^\varepsilon) = \frac{(2^k)^{n-1} - (2^k - 1)^{n-1}}{(1+\varepsilon)^n}$$

Proof: Our proof is similar to that of Theorem 7.1. We have shown the size of the ideal monochromatic region for any outcome $\langle i, p \rangle$ in Claim 7.2, and we consider the size of the ε-bisection protocol induced rectangle:

Claim 7.4 *For the ε-bisection protocol and the corresponding social choice function $\tilde{f}$, the size of the protocol induced rectangle is $R_{\tilde{f}}^p(i,p) \geq (1+\varepsilon)^n$ for an outcome $\langle i, p \rangle$.*

The worst case occurs when the valuations of all agents fall into a same interval of length

$\varepsilon + 1$, so the ε-worst case privacy approximation ratio of ε-bisection protocol is:

$$\begin{aligned}
\varepsilon\text{-}\mathsf{wpar}(p_B^\varepsilon) &= \max_{\mathbf{t}} \frac{R_f^I(\mathbf{t})}{R_{\tilde{f}}^p(\mathbf{t})} \\
&\leq \frac{R_f^I(1,p)}{R_{\tilde{f}}^p(i,p)} \\
&\leq \frac{(2^k - p)\left[(p+1)^{n-1} - p^{n-1}\right]}{(1+\varepsilon)^n} \\
&\leq \frac{(2^k)^{n-1} - (2^k - 1)^{n-1}}{(1+\varepsilon)^n}
\end{aligned}$$

The equalities hold when all agents have values $2^k - 1$, completing our proof. ■

Theorem 7.4 (ε-average case privacy approximation ratio of ε-bisection protocol (for SPAs)) *For n-agent SPAs,*

$$\frac{nk}{(1+\varepsilon)^n} \leq \varepsilon\text{-}\mathit{apar}(p_B^\varepsilon) \leq n(n-1)\frac{(k)^{n-1}}{1+\varepsilon}$$

Proof: The upper and lower bounds on average case PAR for the exact protocol are computed as a special case ($m = n - 1$) of those in Theorem 7.11.

For the approximation case, the valuation of at least 1 agents are identified with a precision of $\varepsilon + 1$, so the saving is between $(1 + \varepsilon)$ and $(1 + \varepsilon)^n$. ■

Theorem 7.5 (ε-worst case privacy approximation ratio of ε-sealed-bid protocol (for SPAs)) *For n-agent SPAs, the ε-worst case privacy approximation ratio of ε-sealed-bid protocol is:*

$$\varepsilon\text{-}\mathit{wpar}(p_S^\varepsilon) = \frac{(2^k)^{n-1} - (2^k - 1)^{n-1}}{(1+\varepsilon)^n}$$

Proof: Our proof is similar to that of Theorem 7.1 and 7.3 .We have shown the size of the ideal monochromatic region for any outcome $\langle i, p\rangle$ in Claim 7.2. For the size of the ε-sealed-bid protocol induced rectangle, we have the following claim:

Claim 7.5 *For the ε-sealed-bid protocol and the corresponding social choice function $\tilde{f}$, the size of the protocol induced rectangle is $R^p_{\tilde{f}}(i,p) = (1+\varepsilon)^n$ for an outcome $\langle i, p\rangle$.*

For any outcome $\langle i, p\rangle$, the size of the ε-sealed-bid protocol induced rectangle is exactly $(1+\varepsilon)^n$ (as the valuation of each agent is identified with a precision of $1+\varepsilon$). So the ε-worst case privacy approximation ratio is:

$$\begin{aligned}
\varepsilon\text{-}\mathsf{wpar}(p^{\varepsilon}_S) &= \max_{\mathbf{t}} \frac{R^I_f(\mathbf{t})}{R^p_{\tilde{f}}(\mathbf{t})} \\
&\leq \frac{R^I_f(1,p)}{R^p_{\tilde{f}}(i,p)} \\
&\leq \frac{(2^k - p)\left[(p+1)^{n-1} - p^{n-1}\right]}{(1+\varepsilon)^n} \\
&\leq \frac{(2^k)^{n-1} - (2^k-1)^{n-1}}{(1+\varepsilon)^n}
\end{aligned}$$

Similarly, the equalities hold when all agents have valuations of $2^k - 1$, completing our proof. ■

Theorem 7.6 (ε-average case privacy approximation ratio of ε-sealed-bid protocol (for SPAs)) *For n-agent SPAs, the ε-average case privacy approximation ratio of ε-sealed-bid protocol is:*

$$2\frac{(2^{k-1})^{n-3}}{\binom{n-3}{\lfloor \frac{n}{2} \rfloor}(1+\varepsilon)^n} \leq \varepsilon\text{-}\mathit{apar}(p^{\varepsilon}_S) \leq \frac{(2^k)^{n-1} - (2^k-1)^{n-1}}{(1+\varepsilon)^n}$$

Proof: We first consider the average case PAR for the exact sealed-bid protocol. Recall that the size of the protocol induced rectangle is always 1, then by Definition 7.5 (or Claim 7.6

that will be described soon) and the property of $R_f^I(i,p)$, we have:

$$\begin{aligned}
\mathsf{apar}(p_f) &= \frac{1}{2^{kn}} \sum_{\mathbf{t} \in T} R_f^I(i,p) \\
&= \frac{1}{2^{kn}} \sum_{\langle i,p \rangle} R_f^p(i,p) \cdot R_I^p(i,p) \\
&\geq R_f^p(n,0) \cdot \frac{1}{2^{kn}} \sum_{i=0}^{2^k-1} R_f^I(i,p) \\
&= R_f^p(n,0) = 2^k - 1
\end{aligned}$$

and

$$\begin{aligned}
\mathsf{apar}(p_f) &= \frac{1}{2^{kn}} \sum_{\langle i,p \rangle} R_f^p(i,p) \cdot R_f^I(i,p) \\
&\leq R_f^p(1, 2^k - 1) \cdot \frac{1}{2^{kn}} \sum_{i=0}^{2^k-1} R_f^I(i,p) \\
&\leq R_f^p(1, 2^k - 1) = (2^k + 1)^{n-1} - (2^k)^{n-1}
\end{aligned}$$

where the equation $(\sum_{i=0}^{2^k-1} R_f^I(i,p))/2^{kn} = 1$ comes from the fact that the sum over all possible ideal monochromatic regions equals the whole type space (see Figure 7.1).

Note that the lower bound we derived here is quite lose—it is exponential in k but not n. However, we can use that of ε-English protocol in its place. This is because for each type profile, the size of the ε-English induced rectangle is at least as large as that induced by the ε-sealed-bid protocol. This will give us a tighter lower bound.

For the approximate case, each valuations of all agents are identified with a precision of $\varepsilon + 1$, so the saving is exactly $(1 + \varepsilon)^n$. ■

Theorem 7.8 (Worst case privacy approximation ratio of English protocol (for FLPs)) *Let p_{EF} be the exact English protocol for n-agent FLPs. Then the worst case privacy approxima-*

tion ratio is:

$$Z(0) \leq \mathsf{wpar}(p_{EF}) \leq Z(2^{k-1})$$

Proof: We first prove the upper bound. By the definition of wpar, we have:

$$\mathsf{wpar}(p_{EF}) = \max_{\mathbf{t}} \frac{R_f^I(\mathbf{t})}{R_{\tilde{f}}^p(\mathbf{t})} \leq \max_{0 \leq r \leq 2^k - 1} Z(r) = Z(2^{k-1})$$

where the inequality comes from the fact that the size of the English protocol induced rectangle is always greater than 1, and the second equality comes from the property of function $Z(r)$.

For the lower bound, we consider the location profile $\mathbf{t} = \{2^k-1, 2^k-2, \ldots, 2^k-2, 2^k-1\}$, in which m agents have peak $2^k - 1$, $m - 1$ agents have peak $2^k - 2$ and the median position is $2^k - 1$. The English protocol induced rectangle has a size of 1 (the median will not be determined until the last agent is queried), and we have:

$$\mathsf{wpar}(p_{EF}) = \max_{\mathbf{t}} \frac{R_f^I(\mathbf{t})}{R_{\tilde{f}}^p(\mathbf{t})} \geq Z(2^k - 1) = Z(0)$$

This completes our proof. ■

Theorem 7.9 (Average case privacy approximation ratio of English protocol (for FLPs)) *Let p_{EF} be the exact English protocol for n-agent FLPs. Then the worst case privacy approximation ratio is:*

$$m\binom{m-1}{m/2}(2^{k-1})^{m-2} \leq \mathsf{apar}(p_{EF}) \leq m\binom{n}{m-1}(2^k)^{m-1}$$

Proof: Recall from Definition 7.5, the average case PAR is defined as:

$$\mathsf{apar}(p_f) = E_D\left[\frac{|R_f^I((t_1,t_2))|}{|R_f^p((t_1,t_2))|}\right]$$

Assuming uniform distribution, we can rewrite the above equation as shown in the following claim:

Claim 7.6 *Let p_f be some protocol that implements a social choice function f. Then the average case privacy approximation ratio satisfies:*

$$\begin{aligned}\mathit{apar}(p_f) &= \frac{1}{(2^k)^n}\sum_{\mathbf{t}}\frac{|R_f^I(\mathbf{t})|}{|R_p^f(\mathbf{t})|}\\ &= \frac{1}{2^{kn}}\sum_{R(p_f)}|R(p_f)|\frac{|R_f^I(R(p_f))|}{|R(p_f)|}\\ &= \frac{1}{2^{kn}}\sum_{R(p_f)}|R_f^I(R(p_f))|\end{aligned}$$

where the sum is over protocol induced rectangles $R(p_f)$, and $R_f^I(R(p_f))$ denotes the ideal region that contains the protocol induced rectangle $R(p_f)$.

In other words, instead of considering each type profile independently, we can consider the protocol induced rectangles. Using this idea, we first show the lower bound of average case PAR. Let $S^n_{[0,2^k-1]}(m) = \sum_{R(p_f)}|R_f^I(R(p_f))|$ be the average case PAR (before divided by 2^{kn}) of finding the median (or mth smallest number) among n agents, where each agent has a valuation in $[0, 2^k-1]$. Then by the English protocol, we have:

$$\begin{aligned}S^n_{[0,2^k-1]}(m) &\geq (m-1)S^{n-1}_{[0,2^k-1]}(m-1) + S^n_{[1,2^k-1]}(m)\\ S^n_{[1,2^k-1]}(m) &\geq (m-1)S^{n-1}_{[1,2^k-1]}(m-1) + S^n_{[2,2^k-1]}(m)\\ &\cdots\cdots\\ S^n_{[2^k-3,2^k-1]}(m) &\geq (m-1)S^{n-1}_{[2^k-3,2^k-1]}(m-1) + S^n_{[2^k-2,2^k-1]}(m)\\ S^n_{[2^k-2,2^k-1]}(m) &\geq (m-1)S^{n-1}_{[2^k-2,2^k-1]}(m-1) + S^n_{[2^k-1,2^k-1]}(m)\end{aligned}$$

The intuition behind this is that the average case PAR (before division) can be computed recursively. More specifically, the problem of finding the mth smallest number among n agents

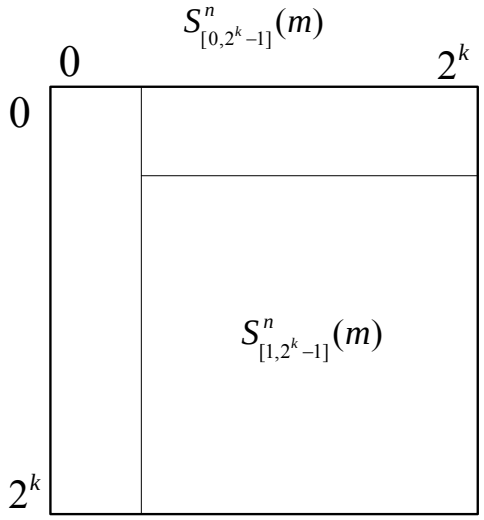


Figure 7.4: The English protocol induced rectangles after the first round for $n = 2$.

in which each agent's valuation is in $[0, 2^k-1]$ can be decomposed to the problem of finding the $(m-1)$th smallest number among $(n-1)$ agents in which each agent's valuation is in $[0, 2^k-1]$ (conditional on knowing one agent has a valuation of 0), and the problem of finding the mth smallest number among n agents in which each agent's valuation is in $[1, 2^k - 1]$ (conditional on knowing none of them has valuation of 0). The multiplier of $(m - 1)$ comes from the fact that we can choose at most $(m - 1)$ non-overlapping valuation spaces for the remaining $n - 1$ agents. The case for $n = 2$ is shown in Figure 7.4.

Summing over all these inequalities, we have:

$$S^n_{[0,2^k-1]}(m) \geq (m-1)\left[S^{n-1}_{[0,2^k-1]}(m-1) + ... + S^{n-1}_{[2^k-2,2^k-1]}(m-1)\right] + Z(2^k-1)$$

By reorganizing the S terms and using the above inequalities iteratively, we have:

$$\begin{aligned}
S^n_{[0,2^k-1]}(m) \geq & (m-1)([S^{n-1}_{[0,2^k-1]}(m-1) - S^{n-1}_{[1,2^k-1]}(m-1)] + ... \\
& + (2^k-2)[S^{n-1}_{[2^k-3,2^k-1]}(m-1) - S^{n-1}_{[2^k-2,2^k-1]}(m-1)] \\
& + (2^k-1)S^{n-1}_{[2^k-2,2^k-1]}(m-1)) + Z(2^k-1) \\
\geq & (m-1)^2(S^{n-2}_{[0,2^k-1]}(m-2) + 2S^{n-2}_{[1,2^k-1]}(m-2) + ... \\
& + (2^k-1)S^{n-2}_{[2^k-2,2^k-1]}(m-2)) \\
& + \left[1 + (2^k-1)(m-1)\right] Z(2^k-1) \\
\geq & ..... \\
\geq & \prod_{t=1}^{(m-1)/2} (m-t)^2 \left[\binom{m-2}{m-2} S^{n-m-1}_{[0,2^k-1]}(1) + ... + \binom{2^k+m-4}{m-2} S^{n-m-1}_{[2^k-2,2^k-1]}(1) \right] \\
& + \left[\sum_{i=0}^{m-2} \left(\binom{2^k-2+i}{i} \prod_{j=0}^{i-1} m-1-\lfloor \frac{j}{2} \rfloor \right) \right] Z(2^k-1)
\end{aligned}$$

Note that the terms $S^m_{[i,2^k-1]}(1)$ are the PARs (before division) for the problem of finding the smallest number among m agents, which is given in the following claim:

Claim 7.7 *Let $S^{n-m-1}_{[i,2^k-1]}(1)$ be the PAR (before division) of finding the smallest number among $n-m-1$ agents, then we have:*

$$S^{n-m-1}_{[i,2^k-1]}(1) = (n-m-1)\sum_{j=i}^{2^k-2} Z(j) + Z(2^k-1)$$

Combining the previous inequality and Claim 7.7, we have:

$$\begin{aligned}
S^n_{[0,2^k-1]}(m) = &\prod_{t=1}^{(m-1)/2}(m-t)^2 \sum_{i=0}^{2^k-2} m \sum_{j=i}^{2^k-2} \binom{m-2+j}{m-2} Z(j) \\
&+ \left[m(2^k-1) + \sum_{i=0}^{m-2} \left(\binom{2^k-2+i}{i} \prod_{j=0}^{i-1} m-1-\lfloor \frac{j}{2} \rfloor \right) \right] Z(2^k-1) \\
\geq &m \prod_{t=1}^{(m-1)/2}(m-t)^2 \binom{2^{k-1}+m-3}{m-2} \sum_{j=0}^{2^k-2} Z(j) \\
&+ \left[m(2^k-1) + \sum_{i=0}^{m-2} \left(\binom{2^k-2+i}{i} \prod_{j=0}^{i-1} m-1-\lfloor \frac{j}{2} \rfloor \right) \right] Z(2^k-1) \\
\geq &m \prod_{t=1}^{(m-1)/2}(m-t)^2 \binom{2^{k-1}+m-3}{m-2} \sum_{j=0}^{2^k-1} Z(j) \\
\geq &m \binom{m-1}{(m-1)/2} (2^{k-1})^{m-2} \sum_{j=0}^{2^k-1} Z(j)
\end{aligned}$$

where the second inequality comes from Claim 7.1 such that $Z(j) = Z(2^k - 1 - j)$ for all integer $j \in [0, 2^k - 1]$.

Note that the term $\sum_{j=0}^{2^k-1} Z(j)$ is the sum of the size of ideal monochromatic region over all possible outputs, so we have $\sum_{j=0}^{2^k-1} Z(j) = 2^{kn}$, which cancels with the denominator 2^{kn}, and we have:

$$\mathsf{apar}(p_f) \geq \frac{1}{2^{kn}} \left[m \binom{m-1}{(m-1)/2} (2^{k-1})^{m-2} \sum_{j=0}^{2^k-1} Z(j) \right] = m \binom{m-1}{(m-1)/2} (2^{k-1})^{m-2}$$

For the upper bound, we use the following property, whose intuition can also be explained in Figure 7.4:

$$S^n_{[i,2^k-1]}(m) \leq n S^{n-1}_{[i+1,2^k-1]}(m) + S^n_{[1,2^k-1]}(m), \quad \forall i$$

Using similar techniques as for the lower bound, we have:

$$\begin{aligned}
S^n_{[0,2^k-1]}(m) \leq & n\left[S^{n-1}_{[0,2^k-1]}(m-1)+...+S^{n-1}_{[2^k-2,2^k-1]}(m-1)\right]+Z(2^k-1)\\
\leq & n(n-1)\left[S^{n-2}_{[0,2^k-1]}(m-2)+...+(2^k-1)S^{n-2}_{[2^k-2,2^k-1]}(m-2)\right]\\
& +\left[1+(2^k-1)n\right]Z(2^k-1)\\
\leq & .....\\
\leq & \prod_{t=0}^{m-2}(n-t)\sum_{i=0}^{2^k-2}m\sum_{j=i}^{2^k-2}\binom{m-2+j}{m-2}Z(j)\\
& +\left[\sum_{i=0}^{m-2}\left(\binom{2^k-2+i}{i}\prod_{j=0}^{i-1}(n-j)\right)\right]Z(2^k-1)\\
\geq & m\binom{n}{m-1}(2^k)^{m-1}\sum_{j=0}^{2^k-1}Z(j)
\end{aligned}$$

Similarly, the term $\sum_{j=0}^{2^k-1}Z(j)=2^{kn}$ and cancels with the denominator, so we have the upper bound as shown above, completing our proof. ■

Theorem 7.10 (Worst-case privacy approximation ratio of bisection protocol (for FLPs)) *Let p_{BF} be the exact bisection protocol for n-agent FLPs. Then the worst case privacy approximation ratio is:*

$$\mathit{wpar}(p_{BF}) = Z(2^{k-1})$$

Proof: Consider the type profile $\mathbf{t} = \{2^{k-1}, 2^{k-1}+1, \ldots, 2^{k-1}+1, 2^{k-1}\}$, in which m agents have peak 2^{k-1}, $m-1$ agents have peak $2^{k-1}+1$ and the median position is 2^{k-1}. If we use the bisection protocol, then the valuations of all agents are identified with a precision of 1 (as the median cannot be determined until every agent is queried), and the size of the protocol

induced rectangle is also 1. Combined with Claim 7.1, we have:

$$\mathsf{wpar}(p_{BF}) = \max_{\mathbf{t}} \frac{R_f^I(\mathbf{t})}{R_{\tilde{f}}^p(\mathbf{t})} \leq \max_{0 \leq r \leq 2^k - 1} Z(r) = Z(2^{k-1})$$

in which equalities holds with the above type profile, completing our proof. ■

Theorem 7.11 (Average-case privacy approximation ratio of bisection protocol (for FLPs)) *Let p_{BF} be the ε-bisection protocol for n-agent FLPs. Then the ε-average case privacy approximation ratio is:*

$$\binom{n}{m} k^{m-1} \leq \mathit{apar}(p_{BF}) \leq m \binom{n}{m} k^m$$

Proof: By Claim 7.6, we have:

$$\mathsf{apar}(p_f) = \frac{1}{2^{kn}} \sum_{R(p_f)} |R_f^I(R(p_f))| = \frac{1}{2^{kn}} \sum_{i=0}^{2^k - 1} n_i \cdot Zi$$

where n_i is the number of p_f induced rectangles where i is the median.

The next claim shows the maximum and minimum value of n_i:

Claim 7.8 *Let n_i be the number of rectangles induced by a protocol in which i is the median, then we have* $\min_j n_j = n_0 = \sum_{i=m}^{n} \binom{n}{i} k^{n-i}$ *and* $\max_j n_j = n_{2^{k-1}} \leq m \binom{n}{m} k^m$.

Using the above claim, we have:

$$\begin{aligned}
\mathsf{apar}(p_f) &= \frac{1}{2^{kn}} \sum_{i=0}^{2^k-1} n_i \cdot Z(i) \\
&\geq \frac{1}{2^{kn}} \sum_{i=0}^{2^k-1} n_0 \cdot Z(i) \\
&\geq n_0 \frac{1}{2^{kn}} \sum_{i=0}^{2^k-1} \cdot Z(i) \\
&\geq \binom{n}{m} k^{m-1}
\end{aligned}$$

and

$$\begin{aligned}
\mathsf{apar}(p_f) &= \frac{1}{2^{kn}} \sum_{i=0}^{2^k-1} n_i \cdot Z(i) \\
&\leq \frac{1}{2^{kn}} \sum_{i=0}^{2^k-1} n_{2^{k-1}} \cdot Z(i) \\
&\leq n_{2^{k-1}} \frac{1}{2^{kn}} \sum_{i=0}^{2^k-1} \cdot Z(i) \\
&\leq m \binom{n}{m} k^m
\end{aligned}$$

This completes our proof of the theorem. ■

Theorem 7.12 (Worst-case privacy approximation ratio of sealed-bid protocol (for FLPs)) *Let p_{SF} be the exact sealed-bid protocol for n-agents FLPs. Then the worst case privacy approximation ratio is:*

$$\mathsf{wpar}(p_{SF}) = Z(2^{k-1})$$

Proof: The proof is similar as that of Theorem 7.10. Note that in a sealed-bid protocol,

the size of the protocol induced rectangle is always 1, so we have:

$$\mathsf{wpar}(p_{SF}) = \max_{\mathbf{t}} \frac{R_f^I(\mathbf{t})}{R_{\tilde{f}}^p(\mathbf{t})} \leq Z(2^{k-1})$$

in which equalities holds with the type profile of $\mathbf{t} = \{2^{k-1}, 2^{k-1}+1, \ldots, 2^{k-1}+1, 2^{k-1}\}$, completing our proof. ■

Theorem 7.13 (Average case privacy approximation ratio of sealed-bid protocol (for FLPs)) *Let p_{SF} be the exact sealed-bid protocol for n-agents FLPs. Then the average case privacy approximation ratio is:*

$$Z(0) \leq \mathit{apar}(p_{SF}) \leq Z(2^{k-1}-1)$$

Proof: Note that in the sealed-bid protocol, all the protocol induced rectangles have size 1. According to Claim 7.6, the average case privacy approximation ratio satisfies:

$$\begin{aligned}
\mathsf{apar}(p_f) &= \frac{1}{2^{kn}} \sum_{i=0}^{2^k-1} Z(i) \cdot Z(i) \\
&\geq \frac{1}{2^{kn}} \sum_{i=0}^{2^k-1} Z(0) \cdot Z(i) \\
&\geq Z(0) \frac{1}{2^{kn}} \sum_{i=0}^{2^k-1} \cdot Z(i) \\
&\geq Z(0)
\end{aligned}$$

and

$$\begin{aligned}
\mathsf{apar}(p_f) &= \frac{1}{2^{kn}} \sum_{i=0}^{2^k-1} Z(i) \cdot Z(i) \\
&\leq \frac{1}{2^{kn}} \sum_{i=0}^{2^k-1} Z(2^{k-1}-1) \cdot Z(i) \\
&\leq Z(2^{k-1}-1) \frac{1}{2^{kn}} \sum_{i=0}^{2^k-1} \cdot Z(i) \\
&\leq Z(2^{k-1}-1)
\end{aligned}$$

This completes our proof of the theorem. ■

Chapter 8

Conclusion and Future Work

Facility location models the placement of facilities (e.g., warehouses, public facilities, etc.) in some geographic space where agents use the least cost (or "closest") facility. Moreover, such a problem represents a general class of social choice problems (e.g., voting, product configuration, customer segmentation, etc.), and receives much attention in economics, political science, and recently computer science. In this thesis, we have studied the facility location problem from three perspectives: mechanism design, single-peaked consistency (and approximation), and preference elicitation. In this chapter, we first summarize the results presented in this thesis, and then highlight some possible future research directions.

8.1 Summary of Results

We provide a brief summary of the results in this thesis.

Mechanism Design for Facility Location Problems

Despite the extensive work on mechanism design for facility location problems, most of them focuses on either single facility or single dimension. In Chapter 3, we proposed a class of quantile mechanisms, a form of generalized median mechanisms for multi-dimensional multi-

facility location problem, that are strategy-proof, and derived worst-case approximation ratios for social cost and maximum load for L_1- and L_2-cost models. More importantly, we proposed a sample-based framework for optimizing the choice of quantiles relative to any prior distribution over preferences, while maintaining strategy-proofness. Our empirical investigations, using social cost and maximum load as objectives, demonstrated the viability of this approach and the value of such optimized mechanisms vis-à-vis mechanisms derived through worst-case analysis.

While quantile mechanisms are strategy-proof for unconstrained facility location problems, they are not group strategy-proof in multi-dimensional spaces. Moreover, guarantees on individual strategy-proofness evaporate in settings allowing constraints on the feasible placement of facilities (i.e., constrained facility location problems). In Chapter 4, we addressed these more general problems, providing several possibility/impossibility results with respect to individual and group strategy-proofness in both constrained and unconstrained problems. We also bounded the incentive for manipulation in median-like mechanisms in settings where individual/group strategy-proofness is not possible. We complemented our results with empirical analysis of both electoral and geographic facility data, showing that the odds of successful manipulation, and more importantly, the gains and impact on social welfare, are small in practice (much less than worst-case theoretical bounds). These results showed that the quantile mechanisms are "practically" strategy-proof considering the cost of find a good lie (e.g., information cost, communication cost, computational cost, etc.).

While the generalized median/quantile mechanisms are not group strategy-proof, this does not mean finding a viable group manipulation is computationally easy for a group of manipulators. In Chapter 5, we addressed optimal group manipulation in unconstrained, multi-dimensional, multi-facility location problems. We focused on two families of mechanisms, generalized median and quantile mechanisms, evaluating how hard it is for a group of agents to manipulate these mechanisms. We showed that, in the case of single-facility problems, optimal group manipulation can be formulated as a linear or second-order cone program, under the L_1-

and L_2-norms, respectively, and hence can be solved in polynomial time. For multiple facilities, we showed that optimal manipulation is NP-hard, but can be formulated as a mixed integer linear or mixed integer second-order cone program, under the L_1- and L_2-norms, respectively. Despite this hardness result, empirical evaluations showed that multi-facility manipulation can be computed in reasonable time with our formulations.

Single-peaked Consistency and Its Approximations

Single-peakedness is one of the most commonly used domain restrictions in social choice. However, whether agent preferences are single-peaked in practice, and the extent to which recent proposals for approximate single-peakedness can further help explain voter preferences, is unclear. In Chapter 6, we assessed the ability of both single-dimensional and multi-dimensional approximations to explain preference profiles drawn from several real-world elections. We developed a simple branch-and-bound algorithm that finds multi-dimensional, single-peaked axes that best fit a given profile, and which works with several forms of approximation. Empirical results on two election data sets showed that preferences in these elections are far from single-peaked in any one-dimensional space, but are nearly single-peaked in two dimensions. Our algorithms are reasonably efficient in practice, and also show excellent any-time performance.

The Trade-off Between Efficiency and Privacy

A key problem in mechanism design is the construction of protocols that reach socially efficient decisions with minimal information revelation. This can reduce agent communication, and further, potentially increase privacy in the sense that agents reveal no more private information than is needed to determine an optimal outcome. This is not always possible: previous work has explored the tradeoff between communication cost and efficiency, and more recently, communication and privacy. We explored a third dimension: the tradeoff between privacy and efficiency. By sacrificing efficiency, we can improve the privacy of a variety of existing mechanisms. We analyzed these tradeoffs in both second-price auctions and facility location

problems (introducing new incremental mechanisms for facility location along the way). Our results showed that sacrifices in efficiency can provide gains in privacy (and communication), in both the average and worst case.

8.2 Future Work

We highlight some possible future directions that would extend the results in this thesis.

Facility Location with Other Objectives

Most work on facility location focuses on the objective of minimizing social cost or minimizing maximum cost. In Chapter 3, we considered another objective, that of minimizing the maximum load. However, there are many other objectives that can be factored in. For example, a natural extension of the facility location problem is to assume a (uniform) opening cost for each facility, with the social cost defined as the total distance of agent peaks to the closest facility plus the opening costs. Here, the number of facilities may vary according to agent preferences, e.g., new facilities decrease social cost, but their expense must be factored in [Lu and Boutilier, 2011a]. How to design strategy-proof mechanisms for such settings, and bound the incentive for misreporting if the social choice functions are not strategy-proof, are interesting future directions.

Mechanism Design With Approximately Single-peaked Preferences

Most previous work on mechanism design for facility location assumes exact single-peakedness of agent preferences. However, as we have shown in Chapter 6 that, preference profiles drawn from real-world applications are only approximately single-peaked under some notions of approximation defined in Section 2.4. A natural question is to extend the previous mechanisms (e.g., median, generalized median and quantile mechanisms) to settings where agent preferences are approximately single-peaked (e.g., as those defined in Section 2.4) and analyse the

incentive for agents to misreport. We conjecture that the incentive is bounded for some forms of approximation, while unbounded for others.

Characterization of Strategy-proof Mechanisms for Multi-FLPs

There has been a fair amount of work on characterization of strategy-proof mechanisms for single facility location problem: Moulin [1980] showed that a (anonymous) mechanism is strategy-proof if and only if it is a generalized median mechanism. Barberà et al. [1993] further generalized this result to the multi-dimensional space, showing that a multi-dimensional mechanism is strategy-proof if and only if it is a multi-dimensional generalized median mechanism that locates the facility by choosing its coordinates in each dimension independently. Border and Jordan [1983] provide a similar characterization results for multi-dimensional separable star-shaped (including quadratic) preferences. Massó and Moreno de Barreda [2011] showed that the disturbed generalized median mechanisms are the only strategy-proof mechanisms for symmetric single-peaked preferences (of which L_1 and L_2 are instances). However, all of these characterization results focus on the single facility location problem. An interesting question is to characterize the class of strategy-proof mechanisms for multi-facility location problems with some specific form of single-peaked preferences (e.g., L_1 or L_2). Our conjecture is that the class of mechanisms should be close to the disturbed generalized median mechanisms for symmetric single-peaked preferences, but with some additional constraints.

Theoretical Foundations of Approximate Single-peaked Consistency

The single-peaked consistency problem has been studied for different notions of approximation. For example, Erdélyi et al. [2012] showed that the consistency (decision) problem for k-maverick, k-LCD and k-AA are NP-complete. However, most previous work focuses on the single-peaked consistency problem in single-dimensional space, and the corresponding consistency problem in multi–dimensional spaces, is unclear. We conjecture that the consistency (decision) problems remain NP-complete in multi-dimensional spaces. In addition, the ex-

planatory power of the multi-dimensional single-peaked model, i.e., the minimum size of a profile that cannot by explained by any multi-dimensional axes, and the maximum size of a profile that can be explained by some multi-dimensional axes, is another interesting future direction.

Deterministic Choice Models for Spatial Theory of Voting

In Section 6.4, we fit both voters and candidates into some latent feature spaces under the spatial model, in which the stochastic choice model of Plackett-Luce is used. Another interesting direction is to use a deterministic (e.g., L_1- or L_2-distance) rather than stochastic choice models, and see how well such models fit these data sets. Knoblauch [2010] showed that the corresponding consistency problem in 1D, i.e., given a preference profile, whether there exists a one-dimensional metric space in which both agents and candidates belong to such that the ranking for each agent is dictated by the distance between her peak and the candidates under the L_1- or L_2-norm, can be solved in polynomial time. However, how the analysis can be generalized to higher dimensions, and how it can be adapted to allow for approximations, is unclear. For example, one can optimize the candidate and voter positions in a way such that the number of candidates that have to be locally deleted from each vote to render the profile consistent with the observed rankings under L_1- or L_2-norm, is minimized (i.e., k-LCD). It is possible to formulate this as a mixed integer linear program (MILP) under L_1-cost, or a mixed integer quadratically constrained program (MIQCP) under the L_2-costs, and use standard optimization tools, e.g., CPLEX, to solve the optimization problem.

Other Elicitation Protocols for FLPs

In Chapter 7, we have proposed two incremental elicitation protocols for single-FLPs, i.e., English and bisection. However, there are many other potential elicitation strategies for eliciting the agent peaks. For example, Conitzer [2009] proposes to use pairwise comparisons of candidates to elicit preferences and identify the one-dimensional axes with respect to which agents

are single-peaked; Lu and Boutilier [2011b] use the both comparison and top-t queries to get agent rankings (although not for single-peaked preferences) for robust decision making, etc. In addition, the one-shot rather than incremental elicitation [Hyafil and Boutilier, 2006b], in which one partitions the type space off-line and ask agents to reveal the partitions which their true types belong to, may also be an interesting direction.

Bibliography

Noga Alon, Felix Fischer, Ariel Procaccia, and Moshe Tennenholtz. Sum of us: Strategyproof selection from the selectors. In *Proceedings of the 13th Conference on Theoretical Aspects of Rationality and Knowledge (TARK-13)*, pages 101–110, Chennai, India, 2011.

Miguel Ángel Ballester and Pedro Rey-Biel. Does uncertainty lead to sincerity? Simple and complex voting mechanisms. *Social Choice and Welfare*, 33(3):477–494, 2009.

Aaron Archer and Éva Tardos. Truthful mechanisms for one-parameter agents. In *Proceedings of the Forty-second IEEE Symposium on Foundations of Computer Science (FOCS-01)*, pages 482–491. IEEE, 2001.

Aaron Archer, Christos Papadimitriou, Kunal Talwar, and Eva Tardos. An approximate truthful mechanism for combinatorial auctions with single parameter agents. In *SODA-03*, pages 205–214, Baltimore, MD, 2003.

Sanjeev Arora, Prabhakar Raghavan, and Satish Rao. Approximation schemes for euclidean k-medians and related problems. In *Proceedings of the Thirtieth Annual ACM Symposium on Theory of Computing*, pages 106–113. ACM, 1998.

Kenneth J. Arrow. A difficulty in the concept of social welfare. *Journal of Political Economy*, 58:328–346, August 1950.

Miguel Angel Ballester and Guillaume Haeringer. A characterization of the single-peaked domain. *Social Choice and Welfare*, 36(2):305–322, 2011.

Salvador Barberà. Strategy-proof social choice. In K. J. Arrow, A. K. Sen, and K. Suzumura, editors, *Handbook of Social Choice and Welfare*, volume 2, pages 731–832. North-Holland, Amsterdam, 2010.

Salvador Barberà, Faruk Gul, and Ennio Stacchetti. Generalized median voter schemes and committees. *Journal of Economic Theory*, 61(2):262–289, 1993.

Salvador Barberà, Jordi Massó, and Alejandro Neme. Voting under constraints. *Journal of Economic Theory*, 76:298–321, 1997.

John Bartholdi and Michael A. Trick. Stable matching with preferences derived from a psychological model. *Operations Research Letters*, 5(4):165–169, 1986.

John Bartholdi III, Craig Tovey, and Michael Trick. Voting schemes for which it can be difficult to tell who won the election. *Social Choice and Welfare*, 6(2):157–165, 1989a.

John Bartholdi III, Craig Tovey, and Michael Trick. The computational difficulty of manipulating an election. *Social Choice and Welfare*, 6(3):227–241, 1989b.

John J. Bartholdi III and James B. Orlin. Single transferable vote resists strategic voting. *Social Choice and Welfare*, 8(4):341–354, 1991.

Nadja Betzler, Arkadii Slinko, and Johannes Uhlmann. On the computation of fully proportional representation. SSRN: 1952497, 2011.

Duncan Black. On the rationale of group decision-making. *Journal of Political Economy*, 56 (1):23–34, 1948.

Liad Blumrosen and Noam Nisan. Auctions with severely bounded communication. In *Proceedings of the 43rd Annual Symposium on Foundations of Computer Science (FOCS-02)*, pages 406–416, Vancouver, 2002.

Kim C. Border and J. S. Jordan. Straightforward elections, unanimity and phantom voters. *The Review of Economic Studies*, 50(1):153–170, 1983.

Craig Boutilier. A POMDP formulation of preference elicitation problems. In *Proceedings of the Eighteenth National Conference on Artificial Intelligence (AAAI-02)*, pages 239–246, Edmonton, 2002.

Stephen Boyd and Lieven Vandenberghe. *Convex Optimization*. Cambridge University Press, Cambridge, 2004.

Ralph Allen Bradley and Milton E. Terry. Rank analysis of incomplete block designs. *Biometrika*, 39:324–345, 1952.

Felix Brandt, Vincent Conitzer, Ulle Endriss, Jérôme Lang, and Ariel Procaccia, editors. *Handbook of Computational Social Choice*. Cambridge University Press, Cambridge, 2015. To appear.

Charles George Broyden. The convergence of a class of double-rank minimization algorithms. *IMA Journal of Applied Mathematics*, 6(1):76–90, 1970.

Zhe Cao, Tao Qin, Tie-Yan Liu, Ming-Feng Tsai, and Hang Li. Learning to rank: From pairwise approach to listwise approach. In *Proceedings of the Twenty-fourth International Conference on Machine Learning (ICML-07)*, pages 129–136, Corvallis, OR, 2007. ACM.

Urszula Chajewska, Daphne Koller, and Ronald Parr. Making rational decisions using adaptive utility elicitation. In *Proceedings of the Seventeenth National Conference on Artificial Intelligence (AAAI-00)*, pages 363–369, Austin, TX, 2000.

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

Yiling Chen, John K. Lai, David C. Parkes, and Ariel D. Procaccia. Truth, justice, and cake cutting. *Games and Economic Behavior*, 77(1):284–297, 2013.

Edward H. Clarke. Multipart pricing of public goods. *Public Choice*, 11(1):17–33, 1971.

Wolfram Conen and Tuomas Sandholm. Partial-revelation VCG mechanisms for combinatorial auctions. In *Proceedings of the Eighteenth National Conference on Artificial Intelligence (AAAI-02)*, pages 367–372, Edmonton, 2002.

Vincent Conitzer. Computing slater rankings using similarities among candidates. In *Proceedings of the Twenty-first National Conference on Artificial Intelligence (AAAI-06)*, pages 613–619, Boston, 2006.

Vincent Conitzer. Eliciting single-peaked preferences using comparison queries. *Journal of Artificial Intelligence Research*, 35:161–191, 2009.

Vincent Conitzer and Tuomas Sandholm. Vote elicitation: Complexity and strategy-proofness. In *Proceedings of the Eighteenth National Conference on Artificial Intelligence (AAAI-02)*, pages 392–397, Edmonton, 2002a.

Vincent Conitzer and Tuomas Sandholm. Complexity of mechanism design. In *Proceedings of the Eighteenth Conference on Uncertainty in Artificial Intelligence (UAI-02)*, pages 103–110, Edmonton, 2002b.

Vincent Conitzer and Tuomas Sandholm. Universal voting protocol tweaks to make manipulation hard. In *Proceedings of the Eighteenth International Joint Conference on Artificial Intelligence (IJCAI-03)*, pages 781–788, Acapulco, 2003.

Vincent Conitzer and Tuomas Sandholm. Computational criticisms of the revelation principle. Unpublished manuscript, 2004.

Vincent Conitzer and Tuomas Sandholm. Communication complexity of common voting rules. In *Proceedings of the Sixth ACM Conference on Electronic Commerce (EC'05)*, pages 78–87, Vancouver, 2005.

Vincent Conitzer and Tuomas Sandholm. Nonexistence of voting rules that are usually hard to

manipulate. In *Proceedings of the Twenty-first National Conference on Artificial Intelligence (AAAI-06)*, pages 627–634, Boston, 2006.

Vincent Conitzer, Tuomas Sandholm, and Jérôme Lang. When are elections with few candidates hard to manipulate? *Journal of the ACM*, 54(3):1–33, 2007.

Peter Cramton, Yoav Shoham, and Richard Steinberg, editors. *Combinatorial Auctions*. MIT Press, Cambridge, 2005.

Mark S. Daskin. *Network and Discrete Location: Models, Algorithms, and Applications*. John Wiley & Sons, Hoboken, NJ, 2011.

Shahar Dobzinski, Noam Nisan, and Michael Schapira. Truthful randomized mechanisms for combinatorial auctions. In *Proceedings of the Thirty-eighth Annual ACM Symposium on Theory of Computing (STOC-06)*, pages 644–652. ACM, 2006.

Jean-Paul Doignon and Jean-Claude Falmagne. A polynomial time algorithm for unidimensional unfolding representations. *Journal of Algorithms*, 16(2):218–233, 1994.

Elad Dokow, Michal Feldman, Reshef Meir, and Ilan Nehama. Mechanism design on discrete lines and cycles. In *Proceedings of the Thirteenth ACM Conference on Electronic Commerce (EC'12)*, pages 423–440, Valencia, Spain, 2012.

Cynthia Dwork. Differential privacy. In *Proceedings of the 33rd International Colloquium on Automata, Languages and Programming*, pages 1–12, S. Servolo, Italy, 2006.

Gábor Erdélyi, Martin Lackner, and Andreas Pfandler. The complexity of nearly single-peaked consistency. In *Proceedings of the Fourth International Workshop on Computational Social Choice (COMSOC-2012)*, Kraków, Poland, 2012.

Bruno Escoffier, Jérôme Lang, and Meltem Öztürk. Single-peaked consistency and its complexity. In *Proceedings of the Eighteenth European Conference on Artificial Intelligence (ECAI-08)*, pages 366–370, Patras, Greece, 2008.

Bruno Escoffier, Laurent Gourvès, Thang Nguyen Kim, Fanny Pascual, and Olivier Spanjaard. Strategy-proof mechanisms for facility location games with many facilities. In *Proceedings of the Second International Conference on Algorithmic Decision Theory (ADT-11)*, pages 67–81, Piscataway, NJ, 2011.

Piotr Faliszewski and Ariel D. Procaccia. AI's war on manipulation: Are we winning? *AI Magazine*, 31(4):53–64, 2010.

Piotr Faliszewski, Edith Hemaspaandra, Lane A. Hemaspaandra, and Jörg Rothe. A richer understanding of the complexity of election systems. In *Fundamental Problems in Computing*, pages 375–406. Springer, 2009a.

Piotr Faliszewski, Edith Hemaspaandra, Lane A. Hemaspaandra, and Jörg Rothe. The shield that never was: Societies with single-peaked preferences are more open to manipulation and control. In *Proceedings of the 12th Conference on Theoretical Aspects of Rationality and Knowledge*, pages 118–127. ACM, 2009b.

Piotr Faliszewski, Edith Hemaspaandra, and Lane A. Hemaspaandra. The complexity of manipulative attacks in nearly single-peaked electorates. In *Proceedings of the Thirteenth Conference on Theoretical Aspects of Rationality and Knowledge (TARK-11)*, pages 228–237, Groningen, The Netherlands, 2011.

Joan Feigenbaum, Aaron D. Jaggard, and Michael Schapira. Approximate privacy: Foundations and quantification (extended abstract). In *Proceedings of the Eleventh ACM Conference on Electronic Commerce (EC'10)*, pages 167–178, Cambridge, MA, 2010. doi: http://doi.acm.org/10.1145/1807342.1807369.

Roger Fletcher. A new approach to variable metric algorithms. *The Computer Journal*, 13(3): 317–322, 1970.

Dimitris Fotakis and Christos Tzamos. Winner-imposing strategyproof mechanisms for multi-

ple facility location games. In *Proceedings of the Sixth International Workshop on Internet and Network Economics (WINE-10)*, pages 234–245, Stanford, CA, 2010.

Dimitris Fotakis and Christos Tzamos. On the power of deterministic mechanisms for facility location games. arXiv: 1207.0935, 2012.

Ehud Friedgut, Gil Kalai, and Noam Nisan. Elections can be manipulated often. In *Proceedings of the 49th Annual IEEE Symposium on the Foundations of Computer Science (FOCS'08)*, pages 243–249, Philadelphia, 2008.

Lucie Galand, Denis Cornaz, and Olivier Spanjaard. Bounded single-peaked width and proportional representation. In *Proceedings of the Fourth International Workshop on Computational Social Choice (COMSOC-2012)*, Kraków, Poland, 2012.

Allan Gibbard. Manipulation of voting schemes: A general result. *Econometrica*, 41(4):587–601, 1973.

Donald Goldfarb. A family of variable-metric methods derived by variational means. *Mathematics of Computation*, 24(109):23–26, 1970.

Isobel Claire Gormley and Thomas Brendan Murphy. A latent space model for rank data. In Edoardo M. Airoldi, David M. Blei, Stephen E. Fienberg, Anna Goldenberg, Eric P. Xing, and Alice X. Zheng, editors, *Statistical Network Analysis: Models, Issues, and New Directions*, pages 90–102. Springer, 2007.

Jerry Green and Jean-Jacques Laffont. Characterization of satisfactory mechanisms for the revelation of preferences for public goods. *Econometrica*, 45:427–438, 1977.

Elana Grigorieva, P. Jean-Jacques Herings, Rudolf Müller, and Dries Vermeulen. The private value single item bisection auction. *Economic Theory*, 30(1):107–118, 2007.

Theodore Groves. Incentives in teams. *Econometrica*, 41:617–631, 1973.

Melvin J. Hinich. Some evidence on non-voting models in the spatial theory of electoral competition. *Public Choice*, 33(2):83–102, 1978.

Harold Hotelling. Stability in competition. *The Economic Journal*, 39(153):41–57, 1929.

Nathanaël Hyafil and Craig Boutilier. Regret-based incremental partial revelation mechanisms. In *Proceedings of the Twenty-first National Conference on Artificial Intelligence (AAAI-06)*, pages 672–678, Boston, 2006a.

Nathanaël Hyafil and Craig Boutilier. Regret-based incremental partial revelation mechanisms. In *Proceedings of the Twenty-second National Conference on Artificial Intelligence (AAAI-07)*, pages 72–78, Vancouver, 2006b.

Nathanaël Hyafil and Craig Boutilier. Mechanism design with partial revelation. In *Proceedings of the Twentieth International Joint Conference on Artificial Intelligence (IJCAI-07)*, pages 1333–1340, Hyderabad, India, 2007.

Marcus Isaksson, Guy Kindler, and Elchanan Mossel. The geometry of manipulationa quantitative proof of the gibbard-satterthwaite theorem. *Combinatorica*, 32(2):221–250, 2012.

Vicki Knoblauch. Recognizing one-dimensional euclidean preference profiles. *Journal of Mathematical Economics*, 46(1):1–5, 2010.

Harold W Kuhn. A note on fermats problem. *Mathematical programming*, 4(1):98–107, 1973.

Eyal Kushilevitz and Noam Nisan. *Communication Complexity*. Cambridge University Press, Cambridge, 1996.

Martin Lackner. Incomplete preferences in single-peaked electorates. In *Proceedings of the 28th AAAI Conference on Artificial Intelligence (AAAI 2014)*. AAAI Press, 2014.

Lan Lan and Hau Chanand Edith Elkind. Multiwinner elections under preferences that are single-peaked on a tree. In *Proceedings of the Twenty-Third international joint conference on Artificial Intelligence*, pages 425–431. AAAI Press, 2013.

Ron Lavi, Ahuva Mu'alem, and Noam Nisan. Towards a characterization of truthful combinatorial auctions. In *Proceedings of the 44th Annual IEEE Symposium on Foundations of Computer Science (FOCS-03)*, pages 574–583, Cambridge, MA, 2003.

Daniel Lehman, Liaden Ita O'Callaghan, and Yoav Shoham. Truth revelation in approximately efficient combinatorial auctions. *Journal of the ACM*, 49:577–602, 2002.

Anton Likhodedov and Tuomas Sandholm. Methods for boosting revenue in combinatorial auctions. In *Proceedings of the Nineteenth National Conference on Artificial Intelligence (AAAI-04)*, pages 232–237, Pittsburgh, PA, 2004.

Anton Likhodedov and Tuomas Sandholm. Approximating revenue-maximizing combinatorial auctions. In *Proceedings of the Twentieth National Conference on Artificial Intelligence (AAAI-05)*, pages 267–274, Pittsburgh, PA, 2005.

Jyh-Han Lin and Jeffrey Scott Vitter. Approximation algorithms for geometric median problems. *Information Processing Letters*, 44(5):245–249, 1992.

Richard J Lipton, Evangelos Markakis, Elchanan Mossel, and Amin Saberi. On approximately fair allocations of indivisible goods. In *Proceedings of the 5th ACM conference on Electronic commerce*, pages 125–131. ACM, 2004.

Pinyan Lu, Yajun Wang, and Yuan Zhou. Tighter bounds for facility games. *Internet and Network Economics*, pages 137–148, 2009.

Pinyan Lu, Xiaorui Sun, Yajun Wang, and Zeyuan A. Zhu. Asymptotically optimal strategy-proof mechanisms for two-facility games. In *Proceedings of the Eleventh ACM Conference on Electronic Commerce (EC'10)*, pages 315–324, Cambridge, MA, 2010.

Tyler Lu and Craig Boutilier. Budgeted social choice: From consensus to personalized decision making. In *Proceedings of the Twenty-second International Joint Conference on Artificial Intelligence (IJCAI-11)*, pages 280–286, Barcelona, 2011a.

Tyler Lu and Craig Boutilier. Robust approximation and incremental elicitation in voting protocols. In *Proceedings of the Twenty-second International Joint Conference on Artificial Intelligence (IJCAI-11)*, pages 287–293, Barcelona, 2011b.

Tyler Lu, Pingzhong Tang, Ariel D. Procaccia, and Craig Boutilier. Bayesian vote manipulation: Optimal strategies and impact on welfare. In *Proceedings of the Twenty-eighth Conference on Uncertainty in Artificial Intelligence (UAI-12)*, Catalina, CA, 2012. 543–553.

R. Duncan Luce. *Individual Choice Behavior: A Theoretical Analysis*. Wiley, 1959.

Dipjyoti Majumdar and Arunava Sen. Ordinally Bayesian incentive compatible voting rules. *Econometrica*, 72(2):523–540, 2004.

Andreu Mas-Colell, Micheal D. Whinston, and Jerry R. Green. *Microeconomic Theory*. Oxford University Press, New York, 1995.

Jordi Massó and Inés Moreno de Barreda. On strategy-proofness and symmetric single-peakedness. *Games and Economic Behavior*, 72(2):467–484, 2011.

Frank McSherry and Kunal Talwar. Mechanism design via differential privacy. In *Proceedings of the 48th Annual IEEE Symposium on Foundations of Computer Science (FOCS-07)*, pages 94–103, Providence, RI, 2007.

Nimrod Megiddo and Kenneth J. Supowit. On the complexity of some common geometric location problems. *SIAM Journal on Computing*, 13(1):182–196, 1984.

Hervé Moulin. On strategy-proofness and single peakedness. *Public Choice*, 35(4):437–455, 1980.

Eitan Muller and Mark A. Satterthwaite. The equivalence of strong positive association and strategy-proofness. *Journal of Economic Theory*, 14:412–418, 1977.

Roger B Myerson. Incentive compatibility and the bargaining problem. *Econometrica: journal of the Econometric Society*, pages 61–73, 1979.

Roger B. Myerson. Optimal auction design. *Mathematics of Operations Research*, 6:58–73, 1981.

Noam Nisan and Amir Ronen. Algorithmic mechanism design. In *Proceedings of the Thirty-first Annual ACM Symposium on Theory of Computing (STOC-99)*, pages 129–140. ACM, 1999.

Noam Nisan and Amir Ronen. Computationally feasible VCG mechanisms. In *Proceedings of the Second ACM Conference on Electronic Commerce (EC'00)*, pages 242–252, Minneapolis, MI, 2000.

Noam Nisan and Ilya Segal. The communication requirements of efficient allocations and supporting prices. *Journal of Economic Theory*, 129(1):192–224, 2006.

Noam Nisan, Tim Roughgarden, Eva Tardos, and Vijay V. Vazirani, editors. *Algorithmic Game Theory*. Cambridge University Press, Cambridge, 2007.

Martin J. Osborne and Ariel Rubinstein. *A Course in Game Theory*. MIT Press, Cambridge, 1994.

David C. Parkes. ibundle: An efficient ascending price bundle auction. In *Proceedings of the First ACM Conference on Electronic Commerce (EC'99)*, pages 148–157, Denver, 1999.

David C. Parkes. Computational mechanism design. In *Lecture notes of Tutorials at 10th Conference on Theoretical Aspects of Rationality and Knowledge (TARK-05)*. Institute of Mathematical Sciences, University of Singapore, 2008.

David C. Parkes and Lyle H. Ungar. Iterative combinatorial auctions: Theory and practice. In *Proceedings of the Seventeenth National Conference on Artificial Intelligence (AAAI-00)*, pages 74–81, Austin, TX, 2000.

R. L. Plackett. The analysis of permutations. *Applied Statistics*, 24:193–202, 1975.

Keith T. Poole and Howard Rosenthal. A spatial model for legislative roll call analysis. *American Journal of Political Science*, 29(2):357–384, 1985.

Areil D. Procaccia and Jeffrey S. Rosenschein. Junta distributions and the average-case complexity of manipulating elections. *Journal of Artificial Intelligence Research*, 28:157–181, 2007.

Ariel D. Procaccia and Moshe Tennenholtz. Approximate mechanism design without money. In *Proceedings of the Tenth ACM Conference on Electronic Commerce (EC'09)*, pages 177–186, Stanford, CA, 2009.

Michael H. Rothkopf, Aleksander Pekeč, and Ronald M. Harstad. Computationally manageable combinatorial auctions. *Management Science*, 44(8):1131–1147, 1998.

Michael Saks and Lan Yu. Weak monotonicity suffices for truthfulness on convex domains. In *Proceedings of the Sixth ACM Conference on Electronic Commerce (EC'05)*, pages 286–293, Vancouver, 2005.

Tuomas Sandholm. Automated mechanism design: A new application area for search algorithms. In *Proceedings of the International Conference on Principles and Practice of Constraint Programming (CP-03)*, Kinsale, Ireland, 2003.

Tuomas Sandholm and Felix Brandt. On the existence of unconditionally privacy-preserving auction protocols. *ACM Transactions on Information and System Security*, 11(2), 2008. Article 6.

Mark A. Satterthwaite. Strategy-proofness and Arrow's conditions: Existence and correspondence theorems for voting procedures and social welfare functions. *Journal of Economic Theory*, 10:187–217, 1975.

James Schummer and Rakesh V. Vohra. Mechanism design without money. In Noam Nisan,

Tim Roughgarden, Eva Tardos, and Vijay V. Vazirani, editors, *Algorithmic Game Theory*, pages 243–265. Cambridge University Press, 2007.

Amartya Sen. Social choice. *The New Palgrave Dictionary of Economics*, 1987.

David F. Shanno. Conditioning of quasi-newton methods for function minimization. *Mathematics of Computation*, 24(111):647–656, 1970.

David F. Shanno and Paul C. Kettler. Optimal conditioning of quasi-newton methods. *Mathematics of Computation*, 24(111):657–664, 1970.

Roger N. Shepard. Stimulus and response generalization: A stochastic model relating generalization to distance in psychological space. *Psychometrika*, 22(4):325–345, 1959.

Lawrence V. Snyder and Mark S. Daskin. Reliability models for facility location: the expected failure cost case. *Transportation Science*, 39(3):400–416, 2005.

Michael A Trick. Recognizing single-peaked preferences on a tree. *Mathematical Social Sciences*, 17(3):329–334, 1989.

Yehuda Vardi and Cun-Hui Zhang. The multivariate l1-median and associated data depth. *Proceedings of the National Academy of Sciences*, 97(4):1423–1426, 2000.

William Vickrey. Counterspeculation, auctions, and competitive sealed tenders. *Journal of Finance*, 16(1):8–37, 1961.

Toby Walsh. Where are the really hard manipulation problems? the phase transition in manipulating the veto rule. In *Proceedings of the Twenty-first International Joint Conference on Artificial Intelligence (IJCAI-09)*, pages 324–329, Pasadena, CA, 2009.

Lirong Xia and Vincent Conitzer. A sufficient condition for voting rules to be frequently manipulable. In *Proceedings of the Ninth ACM Conference on Electronic Commerce (EC'08)*, pages 99–108, Chicago, 2008. doi: http://doi.acm.org/10.1145/1386790.1386810.

Martin Zinkevich, Avrim Blum, and Tuomas Sandholm. On polynomial-time preference elicitation with value queries. In *ACM Conference on Electronic Commerce*, pages 176–185, San Diego, 2003.