Optimal Group Manipulation in Facility Location Problems

Xin Sui(✉) and Craig Boutilier

Department of Computer Science, University of Toronto,
6 King's College Road, Toronto, ON, Canada
{xsui,cebly}@cs.toronto.edu

Abstract. We address optimal group manipulation in multi-dimensional, multi-facility location problems. We focus on two families of mechanisms, *generalized median* and *quantile* mechanisms, evaluating the difficulty of group manipulation of these mechanisms. We show 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 show that optimal manipulation is NP-hard, but can be formulated as a mixed integer linear or second-order cone program, under the L_1- and L_2-norms, respectively. Despite this hardness result, empirical evaluation shows that multi-facility manipulation can be computed in reasonable time with our formulations.

1 Introduction

Mechanism design deals with the design of protocols to elicit the preferences of self-interested agents to achieve some social objective [22]. An important property in mechanism design is *strategy-proofness*, which requires that there is no incentive for an individual agent to misreport their preferences. While much work in mechanism design deals with settings where monetary transfers can be used to facilitate strategy-proofness [6,18,31], many problems do not admit payments for a variety of reasons [28].

The *Gibbard-Satterthwaite theorem* [16,27] shows that under fairly broad conditions, one cannot construct mechanisms that achieve strategy-proofness in general. However, one can impose restrictions on the preference domain to escape this impossibility result. A widely used restriction is *single-peakedness* [4]. In single-peaked domains, each agent has a single, most-preferred *ideal* point in the outcome space, and (loosely) her preference for outcomes decreases with as the distance of that outcome from the ideal increases. In such settings, strategy-proofness is guaranteed by the classic *median mechanism* and its generalizations for single outcomes [2,24], or *quantile mechanisms* [29] for multiple outcomes. Applications of such models include facility location, voting, product design, customer segmentation, and many others.

While these mechanisms are *individual strategy-proof*, they are not *group strategy-proof*—a group of agents may jointly misreport their preferences to induce a more preferred outcome that makes some group members better off

T. Walsh (Ed.): ADT 2015, LNAI 9346, pp. 505–520, 2015.
DOI: 10.1007/978-3-319-23114-3_30

without harming others. In this paper, we consider the *group manipulation problem* for *facility location problems* (*FLPs*) with multiple facilities in multi-dimensional spaces, with a focus on quantile mechanisms (QMs) (and to some extent *generalized median mechanisms* (*GMMs*)). Since these mechanisms are both transparent and (individual) strategy-proof for general multi-dimensional, multi-FLPs, we seek to understand the difficulty of their group manipulation problems.

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 computational 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 [5]). By contrast, we show that multi-FLPs are NP-hard by reduction from the geometric p-median problem [23] under both norms. Despite this, we provide formulations of 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 group manipulations (or prove that none exists) for multi-FLPs of reasonable size rather effectively, despite the theoretical NP-hardness of the problem.

2 Background and Notation

We begin by defining FLPs, the quantile mechanism, the group manipulation problem we consider, and provide a brief discussion of related work.

2.1 Facility Location and Group Manipulation

A d-dimensional, m-facility *facility location problem (FLP)* involves selecting m facilities in some d-dimensional subspace $\mathbb{S} \subseteq \mathbb{R}^d$ (we omit mention of $\mathbb{S}$ subsequently, assuming all locations fall in $\mathbb{S}$). We assume a set of *agents* $N = \{1, \ldots, n\}$, each with an *ideal location* or *type* $t_i \in \mathbb{R}^d$, which determines her *cost* $s_i(x, t_i)$ for using a facility located at x (we sometimes refer to this as facility x). Given a *location vector* $\mathbf{x} = (x_1, \ldots, x_m)$, $x_j \in \mathbb{R}^d$, of m facilities, we assume each agent uses her most preferred facility, defining $s_i(\mathbf{x}, t_i) = \min_{j \leq m} s_i(x_j, t_i)$. Given the ideal points of all agents, our goal is to select an outcome that implements some social choice function (e.g., minimize social cost, ensure Pareto efficiency, etc.). Below we equate cost with L_1 or L_2 distance. A *mechanism* for an FLP is a function f that accepts as input the *reported* ideal points of the n agents and returns a location vector $\mathbf{x}$.

FLPs can be interpreted literally, naturally modeling the placement of homogeneous facilities (e.g., warehouses, public projects) in a geographic space, where agents use the least cost or closest facility. Voting is often modeled this way, where candidates are ordered along each of several dimensions (e.g., stance on environment, fiscal policy, etc.), voters have ideal points in this space, and one elects one or more candidates to a legislative body. Product design, customer segmentation, and other problems can be modeled as FLPs.

Even without explicit distance functions, it is often natural to assume agent (ordinal) preferences are *single-peaked*: an agent's preferences are constrained so that outcomes become less preferred as they are "moved away" from her ideal point (or *peak*). When preferences are single-peaked, the classic *median mechanism* and its generalizations [2,24] guarantee strategy-proofness. Sui *et al.* [29] develop *quantile mechanisms (QMs)* which extend these mechanisms to the multi-facility, multi-dimensional case. We focus here on QMs.

Definition 1. *[29] Let* $\mathbf{q} = \langle \mathbf{q}_1; \ldots; \mathbf{q}_m \rangle$ *be a* $m \times d$ *matrix, where each* $\mathbf{q}_j = \{q_j^1, \ldots, q_j^d\}$ *is a d-vector in the unit cube. A* $\mathbf{q}$*-quantile mechanism* $f_{\mathbf{q}}$ *asks each agent* i *to report her ideal location (or peak)* t_i*. The mechanism locates each facility* j *at the* q_j^k*th quantile among the* n *reported peaks in each dimension* k *independently.*

Example 1. Consider the two-dimensional quantile mechanism in which $\mathbf{q} = \langle 0.25, 0.75; 1.0, 0.5 \rangle$. Given a peak profile of 5 agents $\mathbf{t} = ((1,4),(2,7),(4,2),(7,9),(8,3))$, the $\mathbf{q}$-quantile mechanism will locate the first facility at the intersection of quantile 0.25 in the first dimension and quantile 0.75 in the second dimension, i.e., $(2,7)$ in this example; and the second facility at $(8,4)$.

We note that quantile mechanisms are special case of *generalized median mechanisms (GMMs)* [2,24] when applied to single-FLPs, and can be interpreted as applying a specific form of GMM to the selection of each of the m facilities. As such, QMs are (individual) strategy-proof [29]. However, the characterization of Barberà et al. [2] shows that no (anonymous) mechanism can offer *group* strategy-proofness for multi-dimensional, multi-FLPs in general.[1] The main reason that group manipulation is possible is that a group of manipulators can submit a joint misreport of their ideal locations in which each of them increases her cost in some dimensions but decreases it in others, thereby achieving a lower total cost.

In this paper, we investigate the computational problem of finding just such a group manipulation. Specifically, we consider: (a) the formulation of the *optimal group manipulation problem* as mathematical programs of various types; (b) the computational complexity of this problem; and (c) how much manipulators might gain given optimal manipulations, under different cost functions, when GMMs/QMs are used.[2]

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

[2] Barberà et al.'s [2] characterizations do not preclude the existence of group strategy-proof mechanisms when specific cost functions are used. However, it is still meaningful to study the group manipulation of GMMs and QMs due to their simplicity and intuitive nature, their (individual) strategy proofness, and their flexibility. Indeed, these are the only "natural" such mechanisms for multi-dimensional, multi-FLPs of which we are aware.

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 and at least one is strictly better-off. Our objective of minimizing the sum of costs is the natural one, which represents the social welfare of the manipulators. While one general problem is whether there exists a joint misreport such that no one is worse-off, our optimization version subsumes the former problem.[3] We formalize this as follows:

Definition 2. *Let $N = S \cup M$, where S is a set of sincere agents and M is a set of manipulators with type vectors t_S and t_M. Let $f_{\mathbf{q}}$ be a QM with quantile matrix $\mathbf{q}$. Let $\mathbf{x_q} = f_{\mathbf{q}}(t_M, t_S)$ be the location vector chosen by $f_{\mathbf{q}}$ if all agents report their peaks truthfully, and $\mathbf{x}'_{\mathbf{q}} = f_{\mathbf{q}}(t'_M, t_S)$ be the location vector chosen given some misreport t'_M by the manipulators M. The* optimal group manipulation problem *is to find a joint misreport t'_M for the agents in M satisfying:*

$$t'_M = \arg\min \sum\nolimits_{i \in M} s_i\left(\boldsymbol{x}'_q, t_i\right) \tag{1}$$

$$\begin{aligned} \text{s.t.}\quad & s_i\left(\boldsymbol{x}'_q, t_i\right) \le s_i\left(\boldsymbol{x}_q, t_i\right), \quad \forall i \in M \\ & s_i\left(\boldsymbol{x}'_q, t_i\right) < s_i\left(\boldsymbol{x}_q, t_i\right), \quad \textit{for some } i \in M \end{aligned}$$

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

2.2 Related Work

There has been extensive study of the manipulation problem in other social choice, especially in the contect of voting. While the Gibbard-Satterthwaite theorem shows that strategy-proof mechanisms do not exist in general, Bartholdi et al. [3] demonstrated that manipulation of certain voting rules can be computationally difficult. This spawned an important line of research into the complexity of manipulation for many voting rules—collectively this can be viewed as proposing the use of computational complexity as a barrier to practical manipulation; see, for example, [8,13] for an excellent survey. Recent work has shown that when preferences 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—becomes polynomial time solvable for many voting rules [12]. Our work is similar in its objective to this approach, with a key difference being that in voting outcomes are discrete and atomic, whereas we deal with a continuous, multi-dimensional space.

[3] NP-hardness refers to the corresponding decision problem (as is colloquially understood for optimization problems): is there a misreport that gives the manipulators total cost less than epsilon (for any fixed epsilon). This implies NP-hardness of existence (set cost to truthful cost).

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. Recent work has 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 [7,14,19,26,32,33]. The complete information assumption has also been challenged, and manipulation given probabilistic knowledge of other agent's preferences has been studied in equilibrium [1,21] and from an optimization perspective [20].

3 Group Manipulation for Single-FLPs

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, and finally describing a second-order cone programming formulation under the L_2-norm.

3.1 Group Manipulation Specification

Recall from Definition 2 that a group manipulation is a set of misreports by the manipulating coalition M such that no manipulators is worse off and at least one is better off. The optimization formulation of this problem in Eq. (1) 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 large search space. Fortunately, we can decrease the search space dramatically by only considering *viable* locations for manipulator misreports. We first define *viability*:

Definition 3. *Let $f_{\mathbf{q}}$ be a QM with quantile matrix $\mathbf{q}$. A location $x \in \mathbb{R}^d$ is* viable *for a manipulating coalition M if there exists a joint misreport t'_M s.t. $x = f_{\boldsymbol{q}}(t'_M, t_S)$, where t_S is the report from the sincere agents $S = N \backslash M$. We say t'_M* implements *x in this case.*

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

Proposition 1. *For single-FLPs, let t'_M be a group manipulation and $x'_{\mathbf{q}}$ be a viable location implemented by t'_M under mechanism $f_{\mathbf{q}}$. Then $x'_{\mathbf{q}}$ is also implemented by the group manipulation $t^*_M = \{x'_{\mathbf{q}}, \ldots, x'_{\mathbf{q}}\}$.*

Proof (Sketch). We provide a sketch of proof for $d = 2$, but the analysis can be easily generalized. Consider an arbitrary group manipulation t'_M, which implements location $x'_{\mathbf{q}} = f_{\mathbf{q}}(t'_M, t_S) \in \mathbb{R}^2$ (as shown in Fig. 1). Let us denote the

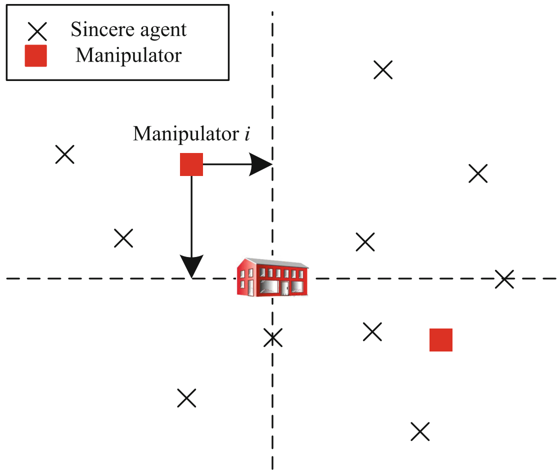

Fig. 1. Each manipulator can move her misreport to $x'_{\mathbf{q}}$ without changing the outcome.

misreport of each manipulator by $t'_i = (t_i^{'1}, t_i^{'2}), \forall i \in M$ and the location by $x'_{\mathbf{q}} = (x_{\mathbf{q}}^{'1}, x_{\mathbf{q}}^{'2})$.

Pick an arbitrary manipulator $i \in M$, and assume w.l.o.g. that $t_i^{'1} \leq x_q^{'1}$ and $t_i^{'2} \geq x_q^{'2}$. We construct another group manipulation $t_M^{''}$ by changing the misreport of manipulator i to $x'_{\mathbf{q}}$. Recall that the mechanism $f_{\mathbf{q}}$ locates the facility at a specified quantile, so we have:

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

Repeating this procedure over all manipulators completes our proof.

Proposition 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. Therefore, we can reformulate the optimal group manipulation problem (Definition 2) as follows:

Definition 4. *Let $f_{\mathbf{q}}$ be a QM with quantile matrix $\boldsymbol{q}$. Let $x_{\mathbf{q}} = f_{\mathbf{q}}(t_M, t_S)$ and $x'_{\mathbf{q}} = f_{\mathbf{q}}(t'_M, t_S)$ be the location chosen by $f_{\mathbf{q}}$ under truthful reports and misreport t'_M, resp. Optimal group manipulation can be reformulated as:*

$$\min_{x \in \mathbb{R}^d} \sum\nolimits_{i \in M} s_i\left(x, t_i\right) \tag{2}$$

$$\text{s.t.} \quad s_i\left(x, t_i\right) \leq s_i\left(x_{\mathbf{q}}, t_i\right), \quad \forall i \in M \tag{3}$$

$$s_i\left(x, t_i\right) < s_i\left(x_{\mathbf{q}}, t_i\right), \quad \textit{for some } i \in M \tag{4}$$

$$x \textit{ is a viable location under } f_{\mathbf{q}} \tag{5}$$

In the sequel, our specific formulations of the problem will rely on Definition 4. We can also safely omit the constraints in Eq. 4, as they can easily be checked after the fact given the optimized location vector—if no manipulator is strictly better off, then a group manipulation obviously cannot exist.

3.2 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., $s_i(x, t_i) = \sum_{k \leq d} |x^k - t_i^k|$ for any location $x \in \mathbb{R}^d$. Let $x = (x^1, \ldots, x^d)$ 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. (2), and the constraints Eq. (3), as follows:

$$\min_{x} \sum\nolimits_{i \in M} c_i \tag{6}$$

$$s.t. \quad c_i = \sum\nolimits_{k \leq d} |x^k - t_i^k|, \quad \forall i \in M \tag{7}$$

$$0 \leq c_i \leq u_i, \qquad \forall i \in M \tag{8}$$

where $u_i = s_i(x_{\mathbf{q}}, t_i)$ is the cost of manipulator i under a truthful report t_M by the manipulators.

This formulation contains absolute values in the nonlinear constraints (7). 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 (7) as follows:

$$-D_i^k \leq t_i^k - x^k \leq D_i^k, \quad \forall i \in M, \forall k \leq d \tag{9}$$

$$D_i^k \geq 0, \qquad \forall i \in M, \forall k \leq d \tag{10}$$

$$c_i = \sum\nolimits_{k \leq d} D_i^k, \qquad \forall i \in M \tag{11}$$

Finally, we need constraints that guarantee the new location x is viable. Recall that a QM locates the facility at a specified quantile of reported peaks in each dimension independently, and by Proposition 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 $\mathbf{q} = (q^1, \ldots, q^d)$ be the quantile vector (for single-FLPs, we have a single vector rather than a full matrix), and let

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

If we let $\bar{x}_S^k = \{\bar{x}_1^k, \ldots, \bar{x}_{|S|}^k\}$ denote the ordered coordinates of the reports of agents S in the kth dimension, we have:

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

This lemma ensures that we can use the following boundary constraints as to enforce viability (see Eq. (5)):

$$\bar{x}^k_{\perp^k} \leq x^k \leq \bar{x}^k_{\top^k}, \qquad \forall k \leq d \tag{12}$$

To summarize, we can formulate the optimal group manipulation under the L_1-norm as an LP. The objective function (6) minimizes the sum of costs over all manipulators. Constraints (8)–(11) guarantee that no manipulators is worse-off, and constraints (12) ensure that the optimized location induced by the misreport is viable. The LP has $O(d|M|)$ variables. We state this result formally in the following theorem:

Theorem 1. *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 (6) and constraints (8)–(12).*

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

3.3 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 (11) to incorporate Euclidean distances as follows:

$$(c_i)^2 \geq \sum\nolimits_{k \leq d} \left(D^k_i\right)^2, \qquad \forall i \in M \tag{13}$$

Constraint (13), combined with the objective function (6) and constraints (8)–(10) and (12), constitutes a second-order cone program (SOCP) under the L_2-norm:

Theorem 2. *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 (6) and constraints (8)–(10) and (12)–(13).*

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

Remark 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.

4 Group Manipulation for Multi-FLPs

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.

4.1 The Complexity of Group Manipulation

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

Theorem 3. *Optimal group manipulation for multi-facility location under either the L_1- or L_2-norms is NP-hard.*

This hardness result is proved using a reduction from the geometric p-median problem, which is known to be NP-hard under both L_1- and L_2-distance [23]. Given a set of points in the d-dimensional space ($d \geq 2$), the *geometric p-median* is a set P of p points that minimizes the sum of distances between each given point and its closest point in P. A complete proof is provided in a longer version of this paper. The rough intuition is as follows. Considering the optimal group manipulation problem for $(p+1)$ facilities, our proof assumes no sincere agents, and constructs a manipulator location profile and QM $f_{\mathbf{q}}$ such that all $(p+1)$ facilities are located at a single extreme position by $f_{\mathbf{q}}$ given truthful reports. However, the optimal group manipulation induces the mechanism to "spread out" p of the facilities to the benefit of a subset of the manipulators, without harming those who would use the original position. This constitutes an optimal solution to the p-median problem for the p non-extreme locations. As such, an algorithm for optimal group manipulation can be used to solve the p-median problem.

While this implies that worst-case instances may be difficult to solve, it does not mean that instances arising in practice can't be solved efficiently. We now describe formulations of optimal group manipulation for multi-FLPs as integer programs that may support practical solution. Our formulations are quite compact, and combined with the empirical evaluation in Sect. 5, suggest that optimal group manipulations can be found reasonably quickly.

4.2 MILP Formulation Under the L_1-norm

We first describe our mixed integer linear programming (MILP) formulation of optimal group manipulation under the L_1-norm. Due to space limitations, we defer certain technical details and proofs to a longer version of this paper. The following result is analogous to Proposition 1 for single-FLPs.

Proposition 2. *Let t'_M be a group manipulation and $\mathbf{x} = \{(x_1^1, \ldots, x_1^d), \ldots, (x_m^1, \ldots, x_m^d)\}$ be a viable location vector implemented by t'_M. Let $\mathbf{X}^k = \{x_1^k, \ldots, x_m^k\}$ denote the set of coordinates of these facilities in the kth dimension. Then there exists a group manipulation t^*_M that implements $\mathbf{x}$, where $t^*_i \in \prod_{k \leq d} \mathbf{X}^k, \forall i \in M$.*

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. The precise misreports at these intersection positions must be coordinated to guarantee a viable location vector (see below).

Let $\mathbf{x} = \{(x_1^1, \ldots, x_1^d), \ldots, (x_m^1, \ldots, x_m^d)\}$ represent the location vector to be optimized. Let c_i be the cost of manipulator i given outcome $\mathbf{x}$, c_{ij} be the cost of manipulator i w.r.t. facility j, and I_{ij} be an indicator variable whose value is 1 iff the closest facility for manipulator i is j. We can formulate the objective Eq. (2), and the constraints Eq. (3), as follows:

$$\min_{\mathbf{x}\in(\mathbb{R}^{\mathbf{d}})^{\mathbf{m}}} \sum\nolimits_{i\in M} c_i \tag{14}$$

$$s.t. \qquad c_i = \sum\nolimits_{j\le m} I_{ij}\cdot c_{ij}, \qquad \forall i\in M \tag{15}$$

$$\sum\nolimits_{j\le m} I_{ij} = 1, \qquad \forall i\in M \tag{16}$$

$$I_{ij}\in\{0,1\}, \qquad \forall i\in M, \forall j\le m \tag{17}$$

$$0\le c_i\le u_i, \qquad \forall i\in M, \forall j\le m \tag{18}$$

$$c_{ij}\ge 0, \qquad \forall i\in M, \forall j\le m \tag{19}$$

where $u_i = s_i(\mathbf{x_q}, t_i)$ is the cost of manipulator i under a truthful report t_M by the manipulators.

Since both I_{ij} and c_{ij} are variables in constraint (15), we must linearize these quadratic terms by introducing additional variables. Let O_{ij} be some upper bound on the product of I_{ij} and c_{ij}. We can then replace the constraint (15) by

$$c_i = \sum\nolimits_{j\le m} O_{ij}, \qquad \forall i\in M \tag{20}$$

$$O_{ij}\ge c_{ij} + (I_{ij}-1)U, \quad \forall i\in M, \forall j\le m \tag{21}$$

$$O_{ij}\ge 0, \qquad \forall i\in M, \forall j\le m \tag{22}$$

where U is any upper bound on manipulator cost.

Let D^k_{ij} be an upper bound on the distance between manipulator i and facility j in the kth dimension. We have:

$$-D^k_{ij}\le t^k_i - x^k_j \le D^k_{ij}, \quad \forall i\in M, \forall j\le m, \forall k\le d \tag{23}$$

$$D^k_{ij}\ge 0, \qquad \forall i\in M, \forall j\le m, \forall k\le d \tag{24}$$

$$c_{ij} = \sum\nolimits_{k\le d} D^k_{ij}, \qquad \forall i\in M, \forall j\le m \tag{25}$$

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

$$\perp^k_j = \min\{z\in\mathbb{Z}^+ : z + |M| \ge q^k_j\cdot n\} \text{ and}$$
$$\top^k_j = \max\{z\in\mathbb{Z}^+ : |S| + |M| - z \ge (1-q^k_j)\cdot n\}$$

and $\bar{x}^k_S = \{\bar{x}_1,\ldots,\bar{x}_{|S|}\}$ be the ordered coordinates of the reports of sincere agents in S in the kth dimension. We break $[\bar{x}^k_{\perp^k_j}, \bar{x}^k_{\top^k_j}]$ into several (ordered) close and open intervals: $[\bar{x}^k_{\perp^k_j}, \bar{x}^k_{\perp^k_j}], (\bar{x}^k_{\perp^k_j}, \bar{x}^k_{\perp^k_j+1}), \ldots, (\bar{x}^k_{\top^k_j-1}, \bar{x}^k_{\top^k_j}), [\bar{x}^k_{\top^k_j}, \bar{x}^k_{\top^k_j}]$ (see Fig. 2 for an illustration). Let Δ^k_j index these intervals ($0\le\Delta^k_j < 2|M|+1$), and let $I_{\Delta^k_j}$ be an indicator variable whose value is 1 iff the coordinate of facility j is contained in the Δ^k_jth interval in the kth dimension. We then have:

$$\sum\nolimits_{\Delta^k_j} I_{\Delta^k_j} = 1, \qquad \forall j\le m, \forall k\le d \tag{26}$$

$$\sum\nolimits_{\Delta^k_j} I_{\Delta^k_j}\bar{x}^k_l \le x^k_j \le \sum\nolimits_{\Delta^k_j} I_{\Delta^k_j}\bar{x}^k_r, \qquad \forall j\le m, \forall k\le d \tag{27}$$

$$I_{\Delta^k_j}\in\{0,1\}, \qquad \forall j\le m, \forall k\le d \tag{28}$$

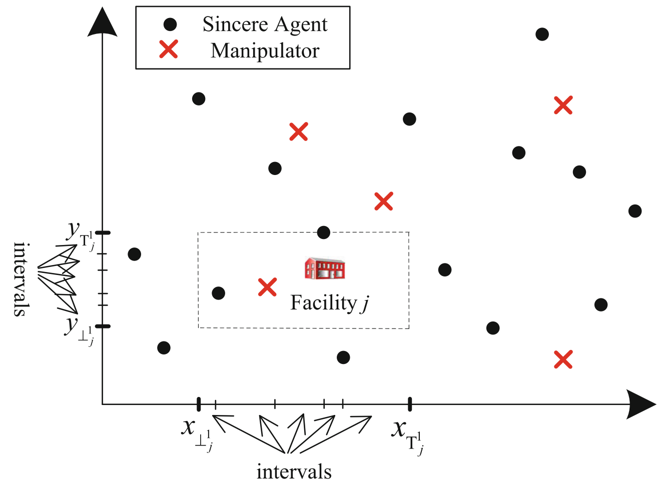

Fig. 2. For each facility in each dimension, the boundaries are split into small intervals, each bounded by one/two sincere agents.

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

For each interval, we pre-compute the number of sincere agents that lie to the left of and right of it (including equality) in each dimension k, which we denote by $L_{\Delta_j^k}$ and $R_{\Delta_j^k}$, respectively. 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 (29) and (30) are integral). Given a quantile matrix $\mathbf{q}$, the location vector $\mathbf{x}$ to be optimized is viable if the following constraints are satisfied:

$$\sum\nolimits_{\Delta_j^k} I_{\Delta_j^k} L_{\Delta_j^k} + \sum\nolimits_{j' \leq_{\mathbf{q}} j} \sum\nolimits_{i \in M} T_{ij'}^k \geq n q_j^k, \quad \forall j, \forall k \tag{29}$$

$$\sum\nolimits_{\Delta_j^k} I_{\Delta_j^k} R_{\Delta_j^k} + \sum\nolimits_{j' \geq_{\mathbf{q}} j} \sum\nolimits_{i} T_{ij'}^k \geq n(1 - q_j^k), \forall j, \forall k \tag{30}$$

$$\sum\nolimits_{j \leq m} T_{ij}^k = 1, \qquad \forall i \in M, \forall k \leq d \tag{31}$$

$$T_{ij}^k \in [0, 1], \qquad \forall i \in M, \forall j \leq m, \forall k \leq d \tag{32}$$

The LHS of constraint (29) 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 $j' \leq_{\mathbf{q}} j$ denotes the facility j' to the left of j in the kth dimension, (i.e., $q_{j'}^k \leq q_j^k$). According to $f_{\mathbf{q}}$, this number should be greater than or equal to nq_j^k. Constraints (30) are similar, but used to count from the right. Constraints (31) and (32) ensure that each manipulator reports the location of one facility on each dimension.

To summarize, we can formulate optimal group manipulation for multi-FLPs under the L_1-norm as a MILP with $O(dm|M|)$ binary and continuous variables:

Theorem 4. *The optimal group manipulation problem for multi-facility location under the L_1-norm can be formulated as a mixed integer linear program with objective function (14) and constraints (16)–(32).*

The final step is to construct a misreport profile t'_M that implements the location vector optimized above. By Proposition 2, we can arbitrarily choose a set of manipulators of size exactly $\sum_i T^k_{ij}$ for each target facility j in each dimension k.

4.3 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 (25) as follows:

$$(c_{ij})^2 \geq \sum\nolimits_{k \leq d} \left(D^k_{ij}\right)^2, \qquad \forall i \in M, \forall j \leq m \tag{33}$$

Using this we obtain the following result:

Theorem 5. *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 (14), and constraints (16)–(24) and (26)–(33).*

5 Empirical Evaluation

In this section, we evaluate the efficiency of the formulations outlined above. We provide empirical results only for multi-facility problems here (since the optimal manipulation problem for single-FLPs is poly-time solvable), testing the efficiency of the MILP/ MISOCP described in Sect. 4.

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 $\mathbf{q} = \{0.3, 0.4; 0.8, 0.7\}$. The second is a four-dimensional, three-facility location problem under the L_1-norm, where the quantile matrix used is $\mathbf{q} = \{0.1, 0.6, 0.4, 0.9; 0.4, 0.2, 0.8, 0.6; 0.7, 0.8, 0.3, 0.4\}$. 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 manipulators).

For the two-dimensional problem, we use preference data from the Dublin west constituency in the 2002 Irish General Election. Since the data includes only voter *rankings* over the set of candidates, the ideal location of each voter is

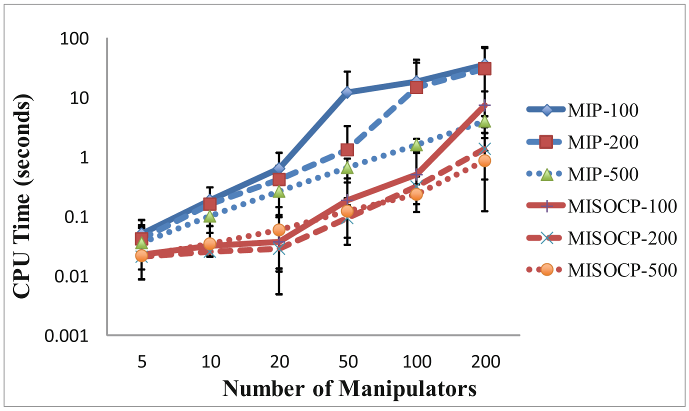

Fig. 3. Time to compute an optimal manipulation (y-axis is log-scale, x-axis is approx. log-scale). Error bars show sample st. dev.

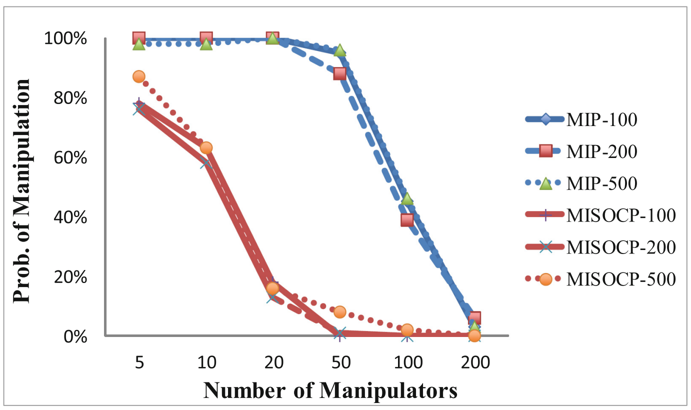

Fig. 4. Probability of a manipulation existing (y-axis is log-scale, x-axis is approx. log-scale).

unknown. Fortunately, recent analysis suggests that this data is approximately single-peaked in two-dimensions [30], and a spatial model using L_2 distance can be used to explain voter preferences [17]. We fit this data to a two-dimensional spatial model, and estimate the voter peaks and candidate positions in the underlying latent space so constructed (Details are provided in a longer version of this paper.) For the four-dimensional problem, we use a synthetic data set in which the peaks of both sincere agents and manipulators are randomly generated from a uniform distribution on the unit cube.

For each instance, the MILP/MISOCP is solved using CPLEX 12.51, on a 2.9GHz, quad-core machine with 8GB memory. Figure 3 shows the average computation time required to find the optimal group manipulation (or show that no group manipulation exists) for both models. We see that our formulations admit very effective solution—for small problems, the optimal group manipulation is 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 s (on average). The performance of our formulations is also very stable (see error bars in the figure).

We illustrate the probability of manipulation for both problems in Fig. 4. For 2D problems, the probability of manipulation decreases from around 80 % to 0 quickly, 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 manipulator. This is not surprising since, as the number of manipulator 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.

6 Conclusion

In this paper, we addressed the optimal group manipulation problem in multi-dimensional, multi-facility location problems. Specifically, we analyzed the computational problems of manipulating 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 shows that our MILPs/MISOCPs formulation for multi-FLPs scales well, despite the NP-hardness result.

Our work 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 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 [17,25,30] has shown that agent preferences are often not exactly single-peaked, but may be approximately so under some forms of approximation [9–11,15]. The theoretical and empirical evaluation of group manipulation in such settings would be extremely valuable.

References

1. Ballester, M.Á., Rey-Biel, P.: Does uncertainty lead to sincerity? simple and complex voting mechanisms. Soc. Choice Welf. **33**(3), 477–494 (2009)
2. Barberà, S., Gul, F., Stacchetti, E.: Generalized median voter schemes and committees. J. Econ. Theory **61**(2), 262–289 (1993)

3. Bartholdi III, J.J., Tovey, C.A., Trick, M.A.: The computational difficulty of manipulating an election. Soc. Choice Welf. **6**(3), 227–241 (1989)
4. Black, D.: On the rationale of group decision-making. J. Polit. Econ. **56**(1), 23–34 (1948)
5. Boyd, S., Vandenberghe, L.: Convex Optimization. Cambridge University Press, Cambridge (2004)
6. Clarke, E.H.: Multipart pricing of public goods. Public Choice **11**(1), 17–33 (1971)
7. Conitzer, V., Sandholm, V.: Nonexistence of voting rules that are usually hard to manipulate. In: Proceedings of the Twenty-first National Conference on Artificial Intelligence (AAAI 2006), pp. 627–634, Boston (2006)
8. Conitzer, V., Sandholm, T., Lang, J.: When are elections with few candidates hard to manipulate? J. ACM **54**(3), 1–33 (2007)
9. Erdélyi, G., Lackner, M., Pfandler, A.: The complexity of nearly single-peaked consistency. In: Proceedings of the Fourth International Workshop on Computational Social Choice (COMSOC 2012), Kraków, Poland (2012)
10. Escoffier, B., Lang, J., Öztürk, M.: Single-peaked consistency and its complexity. In: Proceedings of the Eighteenth European Conference on Artificial Intelligence (ECAI 2008), Patras, Greece, pp. 366–370 (2008)
11. Faliszewski, P., Hemaspaandra, E., Hemaspaandra, L.A.: The complexity of manipulative attacks in nearly single-peaked electorates. In: Proceedings of the Thirteenth Conference on Theoretical Aspects of Rationality and Knowledge (TARK 2011), Groningen, The Netherlands, pp. 228–237 (2011)
12. Faliszewski, P., Hemaspaandra, E., Hemaspaandra, L.A., Rothe, J.: 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, pp. 118–127. ACM (2009)
13. Faliszewski, P., Procaccia, A.D.: AI's war on manipulation: are we winning? AI Mag. **31**(4), 53–64 (2010)
14. Friedgut, E., Kalai, G., Nisan, N.: Elections can be manipulated often. In: Proceedings of the 49th Annual IEEE Symposium on the Foundations of Computer Science (FOCS 2008), Philadelphia, pp. 243–249 (2008)
15. Galand, L., Cornaz, D., Spanjaard, O.: Bounded single-peaked width and proportional representation. In: Proceedings of the Fourth International Workshop on Computational Social Choice (COMSOC 2012), Kraków, Poland (2012)
16. Gibbard, A.: Manipulation of voting schemes: a general result. Econometrica **41**(4), 587–601 (1973)
17. Gormley, I.C., Murphy, T.B.: A latent space model for rank data. In: Airoldi, E.M., Blei, D.M., Fienberg, S.E., Goldenberg, A., Xing, E.P., Zheng, A.X. (eds.) ICML 2006. LNCS, vol. 4503, pp. 90–102. Springer, Heidelberg (2007)
18. Groves, T.: Incentives in teams. Econometrica **41**, 617–631 (1973)
19. Isaksson, M., Kindler, G., Mossel, E.: The geometry of manipulationa quantitative proof of the gibbard-satterthwaite theorem. Combinatorica **32**(2), 221–250 (2012)
20. Lu, T., Tang, P., Procaccia, A.D., Boutilier, C.: Bayesian vote manipulation: optimal strategies and impact on welfare. In: Proceedings of the Twenty-eighth Conference on Uncertainty in Artificial Intelligence (UAI 2012), Catalina, CA, pp. 543–553 (2012)
21. Majumdar, D., Sen, A.: Ordinally Bayesian incentive compatible voting rules. Econometrica **72**(2), 523–540 (2004)
22. Mas-Colell, A., Whinston, M.D., Green, J.R.: Microeconomic Theory. Oxford University Press, New York (1995)

23. Megiddo, N., Supowit, K.J.: On the complexity of some common geometric location problems. SIAM J. Comput. **13**(1), 182–196 (1984)
24. Moulin, H.: On strategy-proofness and single peakedness. Public Choice **35**(4), 437–455 (1980)
25. Poole, K.T., Rosenthal, H.: A spatial model for legislative roll call analysis. Am. J. Polit. Sci. **29**(2), 357–384 (1985)
26. Procaccia, A.D., Rosenschein, J.S.: Junta distributions and the average-case complexity of manipulating elections. J. Artif. Intell. Res. **28**, 157–181 (2007)
27. Satterthwaite, M.A.: 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)
28. Schummer, J., Vohra, R.V.: Mechanism design without money. In: Nisan, N., Roughgarden, T., Tardos, E., Vazirani, V.V. (eds.) Algorithmic Game Theory, pp. 243–265. Cambridge University Press, Cambridge (2007)
29. Sui, X., Boutilier, C., Sandholm, T.: Analysis and optimization of multidimensional percentile mechanisms. In: Proceedings of the Twenty-third International Joint Conference on Artificial Intelligence (IJCAI 2013), Beijing (2013)
30. Sui, X., Francois-Nienaber, A., Boutilier, C.: Multi-dimensional single-peaked consistency and its approximations. In: Proceedings of the Twenty-third International Joint Conference on Artificial Intelligence (IJCAI 2013), Beijing (2013)
31. Vickrey, W.: Counterspeculation, auctions, and competitive sealed tenders. J. Finance **16**(1), 8–37 (1961)
32. Walsh, T.: 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 2009), Pasadena, CA, pp. 324–329 (2009)
33. Xia, L., Conitzer, V.: A sufficient condition for voting rules to be frequently manipulable. In: Proceedings of the Ninth ACM Conference on Electronic Commerce (EC 2008), Chicago, pp. 99–108 (2008)