Mechanism Design for Constrained Heterogeneous Facility Location

Maria Kyropoulou, Carmine Ventre, and Xiaomeng Zhang

University of Essex, UK
{maria.kyropoulou,c.ventre,xzhangao}@essex.ac.uk

Abstract. The facility location problem has emerged as the benchmark problem in the study of the trade-off between incentive compatibility without transfers and approximation guarantee, a research area also known as approximate mechanism design without money. One limitation of the vast literature on the subject is the assumption that agents and facilities have to be located on the same physical space. We here initiate the study of constrained heterogeneous facility location problems, wherein selfish agents can either like or dislike the facility and facilities can be located on a given feasible region of the Euclidean plane. In our study, agents are assumed to be located on a real segment, and their location together with their preferences towards the facilities can be part of their private type. Our main result is a characterization of the feasible regions for which the optimum is incentive-compatible in the settings wherein agents can only lie about their preferences or about their locations. The stark contrast between the two findings is that in the former case any feasible region can be coupled with incentive compatibility, whilst in the second, this is only possible for feasible regions where the optimum is constant.

Keywords: Mechanism design without money · facility location · incentive compatibility.

1 Introduction

Deciding where to locate a public facility, like a school, in order to serve a group of strategic agents, is a fundamental problem that has received a great deal of attention. Under such a setting, the city council, or some other public authority, needs to elicit private information from the concerned (local) people, or agents, without using money, and choose the location of the school based on that information. The authority defines the rules of choosing the location with the objective to maximize the *social welfare*, i.e., the total satisfaction of the agents. However, agents might misreport their private information in an attempt to maximize their own individual utility, which is usually captured using some distance measure between their ideal location for the facility, commonly considered to be (part of) their private information (a.k.a., *type*), and the location of the facility itself. The absence of money makes it very challenging to align the incentives of the authority with those of the individual agents.

The field of mechanism design [11] focuses on the implementation of desired outcomes in strategic settings. A primary designer goal that has been extensively studied is that of truthfulness, which informally states that an agent should be able to optimize her own individual utility by reporting truthfully her private information. However, achieving this is not always compatible with maintaining a high social welfare [7, 15]. Monetary compensations have been commonly used as a means towards aligning the incentives of the individuals with those of society, however, the use of payments is not always allowed due to ethical [11], legal (e.g., organ donations), or even just practical reasons. With this motivation in mind, researchers have started turning their attention to possible ways of achieving truthfulness without the use of payments, i.e., designing truthful (or *strategy-proof*, SP for short) mechanisms that do not use monetary transfers.

Mechanism design without money has been examined from the point of view of exact and approximate solutions. Exact mechanism design without money has a rich history in social choice literature (cf., e.g., [10]), while Procaccia and Tennenholtz [13] were the first to consider achieving truthfulness (or strategy-proofness) without using payments, by sacrificing the optimality of the solution and settling for just an approximation; their work has given rise to what is now known as approximate mechanism design without money. In a nutshell, the objective is that of finding the best approximation guarantee which guarantees strategy-proofness for a given optimization problem.

However, in many settings (such as the school location discussed above) the mechanism designer has some control on the set of feasible solutions (e.g., the area in the city where a school can be built) and would arguably be more interested in leveraging this power to marry strategy-proofness and optimality. In this paper, we initiate the investigation of this research direction and ask whether we can achieve strategy-proofness without using payments by restricting the feasibility of the solution space. As a case study, we consider a facility location problem similarly to Procaccia and Tennenholtz [13]. In our model, the agents are located on a single-dimensional space, just like in [13], while the facility can be located in a feasible region in $\mathbb{R}^2$. Contrary to their approach, we keep the requirement for an optimal solution and study how the shape of the feasible region for the facility location can impact the incentives. We are interested in the following general question:

What is the biggest feasible region for the facility that would allow for the optimal solution to be implemented in a truthful way?

We consider this question in the setting of *heterogeneous preferences* [18, 1, 19], where the facility is not commonly believed to be desirable by the agents; some agents might find it attractive and wish to have it located as close to them as possible, but others might have different views and desire to be far away from them. This preference might also be part of the private information of each agent (in addition/place of their location). When the facility is a school, for example, it is reasonable to expect that families with small children will want to reside close to a school, yet others might prefer to live as far away as possible from it in order to avoid possible noise and traffic.

We define the utility of the agents to be quadratic in the distance between the agent and the facility. On the one hand, the literature on facility location in higher dimensional spaces has twists in the definition of distance in order to make this study feasible. On the other, in many problems, one dimension is not rich enough to fully describe preferences. For example, Barberá et al. [2] mention the city block metric, i.e., the shortest path between two points on a multidimensional grid, as a possible appropriate metric. Our model captures real-life scenarios wherein the agents are environmentally conscious or have resources that are depleted quadratically in the distance (as, e.g., power consumption in wireless communication [12]). In our city planner motivating example, quadratic costs align with environmentally conscious agents, i.e., agents who suffer quadratically in the distance to the facility due to the pollution caused by the travel to cover that distance.

1.1 Our contribution

We examine whether restricting the solution space for the facility location problem can be used as a means of achieving the optimal social welfare in a strategy-proof way. Our findings show a dichotomy result in the sense that all or nothing can be done in the setting of heterogeneous facility location for exact optimal solutions.[1]

Specifically, we consider two different settings, where either the preferences of the agents are private information but their locations are publicly known (unknown preferences case), or the opposite, i.e., the locations of the agents are private information but their preferences are publicly known (unknown locations case).

In Section 3 we treat the case of unknown preferences and we show that the optimal mechanism (the one that maximizes the social welfare) is *group strategy-proof (GSP)* no matter the feasible region. GSP is a stronger requirement than strategy-proofness as it does not even allow for profitable deviations of coalitions of agents. Technically, this is proved by reducing the optimization problem of maximizing the social welfare to the geometric problem of selecting a point in the feasible region which is at maximum/minimum (depending on the shape of the instance) distance from a carefully defined point on the line where the agents reside. This point, which we call β, is a snapshot of the instance and is what the coalitions of agents can manipulate (together with the rule to choose the point in the feasible region). The proof identifies key properties that must be satisfied by a successful manipulation and then observes how those are incompatible with optimality.

We then handle the case of unknown locations in Section 4, where we need to distinguish between different cases depending on the majority of the preferences

[1] Note that depending on the feasible region, the optimal solution might not be well defined, e.g. if the feasible region is $\mathbb{R}^2$ and all agents dislike the facility. Our results implicitly assume that the optimization problem is well-defined and focus on coupling it with incentive considerations.

(which is, in this setting, public knowledge). In each of these cases, we show that in order to be able to implement the optimal solution in a strategy-proof way it would have to hold that the optimal social welfare is constant, i.e., there is a unique point in the feasible region that maximizes the sum of utilities. From the conceptual point of view, this result shows that in the unknown location setting the power coming from the restriction of the feasible region is null as to obtain a strategy-proof optimum, the incentives have to disappear altogether. This is a quite strong negative characterizing result, which paves the way for future research where the interplay between approximation guarantee and the feasible region is considered (see conclusions). From the technical point of view, the proof of this result adopts an iterative approach which identifies several instances showing that the optimum must be the same for both the minimum and the maximum possible value of β. However, while for the case wherein the preferences are homogenous, the argument uses SP and optimality constraints to establish the shape of the feasible region in the limit, the proof for heterogeneous instances requires a more careful step-by-step argument to prove that the optimum is constant.

1.2 Related work

The facility location problem has been studied by many diverse research communities previously. We here discuss some of the most fundamental research directions that have been explored in the context of facility location.

Relevant research from a Social Choice perspective has mostly focused on the problem of locating a single facility on the line. In his seminal paper [10], Moulin characterizes the class of generalized median voter schemes as the only deterministic *SP* mechanisms for *single-peaked* agents on the line. Schummer and Vohra [17] extend the result of Moulin to trees and *continuous graphs*. Dokow et al. [3] prove that for small discrete graphs there are anonymous *SP* mechanisms, contrarily to the case of continuous cycles studied in [17]. They prove that *SP* mechanisms on *discrete large cycles* are nearly-dictatorial in that all agents can affect the outcome to a certain extent.

Facility location has also been one of the fundamental problems in the field of Mechanism Design without money. The work of Procaccia and Tennenholtz on facility location in [14] initiates the study of approximate mechanism design without money, where they suggest the idea of sacrificing a factor of the approximation guarantee as a means to obtain strategy-proofness. For the 2-facility location problem, they propose the Two-Extremes algorithm, that places the two facilities in the leftmost and rightmost location of the instance, and prove that it is group strategy-proof and $(n-2)$-approximate, where n is the number of agents. Furthermore, they provide a lower bound of $3/2$ on the approximation ratio of any *SP* algorithm for the facility location problem on the line and conjecture a lower bound of $\Omega(n)$. The latter conjecture has been recently proven by Fotakis et al. [6]. Their main result is the characterization of deterministic *SP* mechanisms with *bounded approximation ratio* for the 2-facility location problem on the line. They show that there exist only two such algorithms: (i) a mechanism that admits a unique dictator or (*ii*) the Two-Extremes mechanism proposed in [14]. Lu et al. [9], improve several bounds studied in [14]. Particularly, as regards deterministic algorithms they prove a better (w.r.t. [14]) lower bound of $2 - \mathcal{O}(\frac{1}{n})$. Furthermore, they prove a 1.045 lower bound for randomized mechanisms for the 2-facility location problem on the line and present a randomized $n/2$-approximate mechanism.

Our work falls under the category of exact (as opposed to approximate) mechanism design without money. We consider the restriction of the feasibility space so that optimality and strategy-proofness are not mutually exclusive. To the best of our knowledge, this work is the first in the facility location literature to distinguish the region of the agents' locations and the feasible region for the facility. However, similar restrictions have been studied in the judgement aggregation literature, see [4]. We study the case of heterogeneous preferences [18, 1, 19], and distinguish between cases where agents can only misreport their locations or their preferences (but not both). A similar distinction has been considered in [5].

In [8], Lu et al. consider general metric spaces for the 2-facility game. They give an $\Omega(n)$ lower bound for the approximation of deterministic strategy-proof mechanisms and prove that a constant approximation ratio can be achieved by a natural randomized mechanism, the so-called *Proportional* Mechanism.

2 Model and Preliminaries

We assume to have k agents, located on the segment $[0, \ell]$; we say that agent i is located at x_i in that segment. We let $\mathbf{x} = (x_1, \ldots, x_k)$. We need to locate a facility on a given *feasible region* $\Gamma \subseteq \mathbb{R}^2$. (We assume for simplicity that, in the larger space, the segment has the second coordinate equal to 0; our results hold no matter this choice, given that this coordinate is known in either the setting considered.) Each agent might like the facility or dislike it; we let $p_i \in \{-1, 1\}$ denote the preference of agent i with the meaning that if agent i likes (dislikes, respectively) the facility then $p_i = 1$ ($p_i = -1$, respectively). In this sense, our model is heterogeneous, in that not all the agents will have the same opinion of the usefulness of the facility. We let $\mathbf{p} = (p_1, \ldots, p_k)$, let m denote the number of agents whose preference is 1, and let n the number of agents whose preference is -1, so that $k = m + n$. Agent i has a *utility* which depends on her location x_i and preference p_i, and the location $f = (x_f, y_f)$ of the facility in Γ, that is,

$$u_i((x_i, p_i), f) = \begin{cases} (x_i - x_f)^2 + y_f^2 & \text{if } p_i = -1 \\ \lambda - (x_i - x_f)^2 - y_f^2 & \text{if } p_i = 1 \end{cases},$$

where λ is a constant which guarantees that the utilities are not negative. Intuitively, an agent who likes the facility wants to be close to f, while an agent who dislikes it wants to be far from it. Our definition of utility captures that and is similar in spirit to the one in [1]; in our definition, however, the utility is quadratic in the distance between x_i and f (see Section 1 for a relevant discussion).

We study this problem from a mechanism design perspective. That is, we assume that the agents have a private *type* t_i, and we consider the two extreme cases of type being either the preference or the location of each agent, i.e. $t_i \in \{x_i, p_i\}$. A mechanism $\mathcal{M}$ collects reports from the agents, which are potentially different *bids* b_i, and on this input returns a location for the facility in $\mathit{\Gamma}$. With a slight abuse of notation, we assume that the bid of agent i to the mechanism is completed with the public part of $\{x_i, p_i\}$. Our objective is to design a *truthful mechanism* (a.k.a., strategy-proof, SP for short) $\mathcal{M}$, i.e., a mechanism such that for any t_i, b_i and $\mathbf{b}_{-i} = (b_j)_{j \neq i}$,

$$u_i(t_i, \mathcal{M}(t_i, \mathbf{b}_{-i})) \geq u_i(t_i, \mathcal{M}(b_i, \mathbf{b}_{-i})).$$

A stronger requirement is for the mechanism to be *group strategy-proof* (GSP, for short). A mechanism $\mathcal{M}$ is GSP if for any profile $\mathbf{b}$ and any coalition $C \subseteq [k]$, there is no joint deviation $\mathbf{b}'_C = (b'_i)_{i \in C}$ of the agents in C such that no agent in C loses and at least one gains, that is, for all $\mathbf{b}$, for all $C \subseteq [k]$ and for all $\mathbf{b}'_C$ there exists $i \in C$ such that

$$u_i(b_i, \mathcal{M}(\mathbf{b})) > u_i(b_i, \mathcal{M}(\mathbf{b}'_C, \mathbf{b}_{-C}))$$

or for all $i \in C$,

$$u_i(b_i, \mathcal{M}(\mathbf{b})) = u_i(b_i, \mathcal{M}(\mathbf{b}'_C, \mathbf{b}_{-C})),$$

where $\mathbf{b}_{-C} = (b_i)_{i \notin C}$. We restrict the focus in this work on optimal mechanisms for the *social welfare*, that is, we want $\mathcal{M}$ to find, on input an instance $\mathbf{b} = (b_1, \ldots, b_k)$, the point $f^\star = \arg\max_{f \in \mathit{\Gamma}} SW(\mathbf{b}, f)$, where $SW(\mathbf{b}, f) = \sum_{i=1}^{k} u_i(b_i, f)$. Clearly, optimality depends on the choice of the feasible region $\mathit{\Gamma}$, but we omit this dependence when referring to optimal mechanisms for clarity of exposition.

3 Unknown Preferences

We begin by introducing some notations, that allow a more useful formulation of the social welfare and ultimately a geometric characterization of the optimum.

Fix a profile $\mathbf{b}$. Recall that m denote the number of agents whose preference is 1 and n the number of agents whose preference is -1. Note that in the case of unknown preferences, m and n depend on the agents' strategies. We let $\gamma = m - n$, $s_p = \sum_{i:p_i=p} x_i$, for $p \in \{-1, 1\}$ and $\delta = s_1 - s_{-1}$. Tedious calculations can verify that we can rewrite $SW(\mathbf{b}, (x_f, y_f))$ as follows:

$$\begin{cases} -\gamma\left(x_f - \dfrac{\delta}{\gamma}\right)^2 - \gamma y_f^2 + \dfrac{\delta^2}{\gamma} + m\lambda - \delta & \text{if } \gamma \neq 0; \\ 2\delta x_f + m\lambda - \delta & \text{if } \gamma = 0. \end{cases}$$

Therefore, for a given instance $\mathbf{b}$, optimizing the social welfare is equivalent to finding a point (i.e., the x_f and y_f) in $\mathit{\Gamma}$ that maximizes the equations above, in the respective cases. This amounts to choosing a point in Γ with maximum/minimum x-coordinate (depending on the sign of δ) when the number of players with preference 1 is equal to the number of players whose preference is -1. In the case in which $m \neq n$ ($\gamma \neq 0$), however, we need to maximize the quadratic equation. This is equivalent to finding the point in Γ that either maximizes (when $\gamma < 0$) or minimizes (when $\gamma > 0$) the (square of the) distance from the point $\beta = \left(\frac{\delta}{\gamma}, 0\right)$.

Below, we will let $d(\cdot, \cdot)$ denote the distance between two points. Moreover, we will let $m, n, s_1, s_{-1}, p_i, \delta$ denote the parameters for the instance in which agent i is truthful and $f = (x_f, y_f)$ be the output of the mechanism. $x_\beta = \delta/\gamma$ naturally corresponds to the x-coordinate of point β defined above. We add a prime symbol to denote aspects of the instance where agents misreport their type.

We are now ready to prove the first part of our dichotomy.

Theorem 1. *For all $\Gamma \subseteq \mathbb{R}^2$, the optimum mechanism is GSP.*

Proof. Assume that there exists $\Gamma \subseteq \mathbb{R}^2$ such that OPT mechanism is not GSP. This means that there is a coalition C that by joint deviation manages to change the outcome from f to f' while no agent in C loses and at least one gains. We extend the notations above with a c and $-c$ symbol to restrict the respective quantities to the agents inside and outside C, respectively; so, for example, γ_c (γ_{-c}, resp.) denotes the difference between the number of agents inside (outside, resp.) C with preference 1 and those with preference -1.

We begin by showing that if there is a profitable deviation for the coalition, then $x_f \neq x_{f'}$. First, consider the case in which the agents in C have heterogeneous preferences, that is, there are agents in C with either preference. Since the mechanism is not GSP then there exist agents i and j in C such that $p_i = 1$ and $p_j = -1$ and it holds that:

$$d^2(x_i, f) > d^2(x_i, f')$$
$$d^2(x_j, f) \leq d^2(x_j, f')$$

where we assumed w.l.o.g. that i is the agent in C for whom the inequality is strict (at least one such agent must be in C). By simple algebraic manipulations, we conclude that

$$(x_j - x_i)(x_f - x_{f'}) > 0.$$

This implies that $x_f \neq x_{f'}$.

Consider now the case in which all the agents in C have the same preference, so that $\gamma_c \neq 0$. When $x_f = x_{f'}$, we shall prove that:

$$\gamma_c(\gamma_c + \gamma_{-c}) \leq 0, \tag{1}$$
$$\gamma_c(-\gamma_c + \gamma_{-c}) \geq 0. \tag{2}$$

Consider the case in which $\gamma_c > 0$ (the case $\gamma_c < 0$ is symmetric). Since the coalition finds it profitable to change the output from f to f' then, since $x_f = x_{f'}$, it must be the case that $|y_{f'}| < |y_f|$. But then since the optimum for the original instance chooses f but not f' it cannot be that $\gamma > 0$. We then have that $\gamma_c \gamma \leq 0$, thus proving (1). Similarly, given that the optimum for the modified instance (in which the agents in C lie) returns f' and not f it cannot be that $\gamma' < 0$ and then $\gamma_c \gamma' \geq 0$; observe that $\gamma' = -\gamma_c + \gamma_{-c}$. Summing up (1) and (2), we have $\gamma_c^2 \leq 0 \Rightarrow \gamma_c = 0$ — a contradiction.

It now remains to argue only about the case $x_f \neq x_{f'}$. Contradicting this case as well will prove that there is no profitable deviation for a coalition, as desired. Let ϵ be

$$\epsilon = \frac{x_f + x_{f'}}{2} + \frac{x_f + x_{f'}}{2} \frac{y_f - y_{f'}}{x_f - x_{f'}}. \tag{3}$$

Intuitively, the point $(\epsilon, 0)$ is the intersection of the x-axis with the perpendicular crossing the middle of the line segment connecting f and f'. As $x_f \neq x_{f'}$, this intersection must exist. Note that ϵ is an important parameter to determine where β and β' are in the cases in which $\gamma \neq 0$ and $\gamma' \neq 0$. In fact, $(\epsilon, 0)$ partitions the points on the x-axis according to the facility they are closer to. So, for example, by definition of the optimum, β and f must be on the same side of $(\epsilon, 0)$ for $\gamma > 0$.

When $\gamma_c = 0$, we have

$$\delta_c (x_f - x_{f'}) < 0. \tag{4}$$

Indeed, since $\gamma_c = 0$ then the coalition is heterogeneous and $m_c = n_c$. Assume that $x_{f'} < x_f$ (the opposite case being symmetric). Since the coalition prefers f' over f, then it must be that $x_j \leq \epsilon \leq x_l$, for every $j, l \in C$ such that $p_j = 1$ and $p_l = -1$, with at least one j, l for which the inequality is strict (for otherwise, the coalition would not deviate to change the outcome from f to f'). But then this means that $\delta_c < 0$.

When $\gamma_c \neq 0$, we can prove:

$$\gamma_c (x_{\beta_c} - \epsilon)(x_f - x_{f'}) < 0. \tag{5}$$

Indeed, consider the case $\gamma_c > 0$, i.e., $m_c > n_c > 0$. Assume that $x_{f'} < x_f$; as argued above, all the agents in the coalition with preference 1 (-1, resp.) must be to the left (right, resp.) of ϵ and at least one must have a location different from ϵ. Choose now a subset S of n_c agents in the coalition with preference 1, including one with location not ϵ (if any). We can then conclude that $\sum_{j \in S} x_j - \sum_{l \in C, p_l = -1} x_l < 0$. But then, observing that $\sum_{j \in C \setminus S, p_j = 1} x_j \leq (m_c - n_c)\epsilon$, we can conclude that $x_{\beta_c} < \epsilon$ and prove (5). (The remaining cases can be proved with the same argument mutatis mutandis.)

In order to conclude the proof of the theorem, we will now show a contradiction with (either) (4) and (5); we will consider three different cases depending on the values of γ and γ'. Note that $\delta' = -\delta_c + \delta_{-c}$ and so:

$$x_\beta = \frac{\delta_c + \delta_{-c}}{\gamma_c + \gamma_{-c}} \text{ and } x'_\beta = \frac{-\delta_c + \delta_{-c}}{-\gamma_c + \gamma_{-c}}.$$

Case $\gamma \neq 0$ and $\gamma' \neq 0$. Let us only discuss here $\gamma > 0, \gamma' > 0$; the other cases can be proved with the same argument. By the definition of optimum for positive values of γ and γ', we have:

$$(x_\beta - \epsilon)(x_f - x_{f'}) \geq 0 \text{ and } (x_{\beta'} - \epsilon)(x_f - x_{f'}) \leq 0.$$

Therefore, for $x_f > x_{f'}$ we have

$$\frac{\delta_c + \delta_{-c}}{\gamma_c + \gamma_{-c}} \geq \epsilon \text{ and } \frac{-\delta_c + \delta_{-c}}{-\gamma_c + \gamma_{-c}} \leq \epsilon.$$

By simple algebraic manipulations, we can conclude that $\delta_c - \epsilon\gamma_c \geq 0$. Similarly, we can show that $x_f < x_{f'}$ yields $\delta_c - \epsilon\gamma_c \leq 0$. But this contradicts (5), when $\gamma_c \neq 0$ and (4) in the case $\gamma_c = 0$.

Case $\gamma = 0$ and $\gamma' = 0$. We begin by observing that, by summing up the conditions on $\gamma = 0, \gamma' = 0$, we get that $\gamma_c = 0$ (and then (5) does not hold in this case). By the definition of optimum, we have :

$$\begin{aligned}(\delta_c + \delta_{-c})(x_f - x_{f'}) &\geq 0 \\ (-\delta_c + \delta_{-c})(x_f - x_{f'}) &\leq 0.\end{aligned}$$

From the two inequalities, we get $\delta_c(x_f - x_{f'}) \geq 0$, which contradicts (4).

Case $\gamma = 0$ (exclusive) or $\gamma' = 0$. We here discuss only $\gamma > 0, \gamma' = 0$; the remaining cases can be proved in the same manner. By the definition of optimum, we have:

$$\begin{aligned}(x_\beta - \epsilon)(x_f - x_{f'}) &\geq 0 \\ (-\delta_c + \delta_{-c})(x_f - x_{f'}) &\leq 0.\end{aligned}$$

As $\gamma' = -\gamma_c + \gamma_{-c} = 0$, we have $\gamma_c = \gamma_{-c}$ and then since $\gamma = -\gamma_c + \gamma_{-c} > 0$ we can conclude that $\gamma_c > 0$ (and so (4) does not hold here). Thus, from the first inequality, we have:

$$\begin{aligned}&\frac{1}{2\gamma_c}(\delta_c + \delta_{-c} - 2\epsilon\gamma_c)(x_f - x_{f'}) \geq 0 \\ &\Rightarrow (\delta_c + \delta_{-c} - 2\epsilon\gamma_c)(x_f - x_{f'}) \geq 0.\end{aligned}$$

Combined with the second inequality above, we have

$$(\delta_c - \epsilon\gamma_c)(x_f - x_{f'}) \geq 0,$$

which contradicts (5).

4 Unknown locations

We now prove the second part of our dichotomy result. Given that the value of $\gamma = m - n$ changes combinatorially the optimum and that, in the case of unknown locations, γ is known to the designer, our analysis needs to differentiate all the three cases about the relative order between m and n.

4.1 Case $m > n$

We begin with all the instances where the number of 1's is bigger than the number of -1's. For two points α and ζ on the plane, we let $C(\alpha, \zeta)$ denote the points in the interior of the circle centred in α of radius $d^2(\alpha, \zeta)$; formally,

$$C(\alpha,\zeta) = \left\{(x,y)\middle|(x-x_\alpha)^2 + (y-y_\alpha)^2 < d^2(\alpha,\zeta)\right\}.$$

We denote by $\bar{C}(\alpha, \zeta)$ the points of the circle including those on the circumference.

Theorem 2. *When $m > n = 0$, if the optimum mechanism is strategy-proof, then Γ is such that the optimum is constant.*

Proof. We are going to show, through a sequence of instances, that for the optimum to be strategy-proof on those instances, Γ must have a certain shape. We can then observe that given such a shape, the optimum is constant no matter the instance. Specifically, we will prove that there is $f \in \Gamma$ such that

$$\Big[C(\beta_{\min}, f) \cup C(\beta_{\max}, f)\Big] \cap \Gamma = \emptyset, \tag{6}$$

where $\beta_{\max} = (l, 0)$ and $\beta_{\min} = (0, 0)$.

We let f be the optimum of the instance in which $m-2$ agents are on $l/2$, one agent i is on $x_L = l/2 - \frac{1}{m+1}$ and the last one, named j, is on $x_R = l/2 + \frac{1}{m+1}$, where $\delta = \frac{1}{m+1}$. Observe that $x_\beta = l/2$, $x_L = l/2\xi$ and $x_R = l/2\zeta$, where $\xi = \frac{m}{m+1} < 1$ and $\zeta = \frac{m+2}{m+1} > 1$.

We now will argue that there is a sequence of instances that prove in the limit that $C(\beta_{\min}, f) \cap \Gamma = \emptyset$. We set $\mathbf{x}'$ to be as $\mathbf{x}$ except that x'_L to be $\xi^2(l/2)$. Note that in $\mathbf{x}'$, the value used to calculate the optimum, denoted β', is exactly x'_L. Consider the case in which the true type of i is x_L and she misreports to x'_L. By strategy-proofness, we must then have that $f' \notin C((x_L, 0), f) = C(\beta', f)$. We can now iterate the reasoning above and define instance $\mathbf{x}^r$, as $\mathbf{x}$ expect that $x^r_L = (l/2)\xi^r$, observe that the point β^r used to determine the optimum is x^r_l and conclude that $C(\beta^r, f) \cap \Gamma = \emptyset$. Given that $\xi < 1$, in the limit we have that $C(\beta_{\min}, f) \cap \Gamma = \emptyset$.

We can use the same argument now on the right of β to conclude that $C(\beta_{\max}, f) \cap \Gamma = \emptyset$.

We adopt a different argument to account for the case in which there are agents who dislike the facility.

Theorem 3. *When $m > n > 0$, if the optimum mechanism is strategy-proof, then Γ is such that the optimum is constant.*

Proof. The proof uses the same structure of the proof above for $n = 0$; in particular, we will prove (6), where $\beta_{\max} = \left(\frac{ml}{m-n}, 0\right)$ and $\beta_{\min} = \left(\frac{-nl}{m-n}, 0\right)$. From this we can conclude that the optimum is unique, no matter where β is.

Let S^+ and S^- denote the set of agents with preference 1 and -1, respectively. We are going to use the following instances to establish this structural property of Γ. We let $\mathbf{x}^{+(j)}$, $0 \leq j \leq m$, be an instance where j (of the m) agents in S^+ are on l and the n agents in S^- plus the $m-j$ remaining agents in S^+ are on 0. We also denote with $\mathbf{x}^{-(j)}$, $0 < j \leq n$, an instance where j (of the n) agents in S^- are on l and the m agents in S^+ plus the $n-j$ remaining agents in S^- are on 0. The proof will use three steps to establish that the optimum on $\mathbf{x}^{+(m)}$ is the same as the optimum on $\mathbf{x}^{-(n)}$; this yields (6) given that these are the instances defining $\beta_{\max}$ and $\beta_{\min}$, respectively.

Step 1: From $\mathbf{x}^{+(m)}$ to $\mathbf{x}^{+(n+1)}$. (Note that this step is not needed if $m = n+1$.) The idea is to use optimality and strategy-proofness constraint of agents, with preference $+1$, true type l and misreport 0. Let f be the optimum of $\mathbf{x}^{+(m)}$; by definition (given that $\gamma > 0$) $C(\beta_{\max}, f) \cap \Gamma = \emptyset$. Let $\beta' = \left(\frac{(m-1)l}{m-n}, 0\right)$ be the point used to calculate the optimum f' on $\mathbf{x}^{+(m-1)}$. We are going to prove that $f' = f$. By definition of optimum,

$$f' \in \bar{C}(\beta', f) \cap \Gamma = \bar{C}(\beta', f) \setminus C(\beta_{max}, f)$$

and then $x_{f'} \leq x_f$. Let j be the agent changing her true location l in $\mathbf{x}^{+(m)}$ to a misreport 0 in order to obtain $\mathbf{x}^{+(m-1)}$; recall that $p_j = 1$. Then by the strategy-proofness constraint of agent i, $f' \notin C((l,0), f)$. Since $l < x_{\beta'}$ this means that the only feasible point for f' is f, i.e., $f = f'$. We can now reiterate the argument above, and conclude that f is the optimum also for $\mathbf{x}^{+(1)}$.

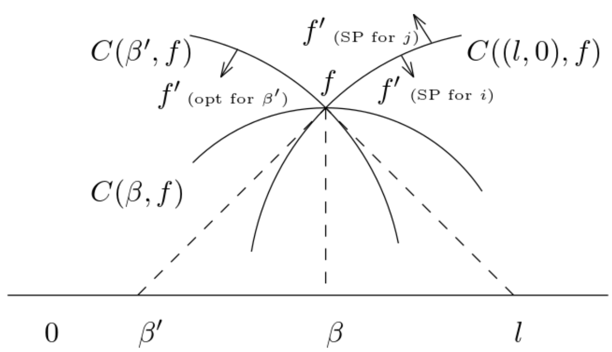

Fig. 1. Proving that $f = f'$ in Step 2 of the proof; the distances plotted are squared Euclidean distances.

Step 2: From $\mathbf{x}^{+(n+1)}$ to $\mathbf{x}^{+(0)}$. The proof for this step uses similar ideas to the one above, but uses the strategy-proofness constraint twice, once for a bidder of either preference, with true type l and misreport 0. Let us denote with $\beta = \left(\frac{(m-n+1)l}{m-n}, 0\right)$ the point used to calculate f, the optimum on $\mathbf{x}^{+(n+1)}$; note that $\beta' = (l, 0)$ is instead used to calculate f', the optimum on $\mathbf{x}^{+(n)}$. We prove that

$f = f'$. (See Fig. 1 for a graphical representation of the argument.) The key observation is that β' is also used to calculate the optimum for the instance $\mathbf{x}'$ defined as $\mathbf{x}^{+(n+1)}$ except that an agent i in S^- moves from 0 to l; recall that $p_i = -1$. If we consider the case in which the true type of i is l and 0 is a misreport then we have the following the strategy-proofness constraint for i: the facility f' must not be further from i's true location (i.e., $(l, 0)$) than the facility she will get by lying (f calculated with β). As before, by denoting with j the agent of preference 1 changing her location from l (true type) to 0 (lie) we conclude that f' must not be closer to j's true location ($(l, 0)$) than f'. By combining the two conditions we get that

$$f' \in \bar{C}((l, 0), f) \setminus C((l, 0), f).$$

The proof concludes by using the same argument used in Step 1 for optimality: indeed, we can observe that the only intersection between $\bar{C}((l, 0), f) \setminus C((l, 0), f)$ and $\bar{C}(\beta', f)$ is f'. We can now reiterate the argument above, and conclude that f is the optimum also for $\mathbf{x}^{+(0)}$.

Step 3: From $\mathbf{x}^{+(0)}$ to $\mathbf{x}^{-(n)}$. The proof of the last part uses a very similar argument to the one used in Step 2, with the difference that this time the two strategy-proofness constraints use the case that the true location of agents of either preference is 0 and they misreport to l. We omit the details.

4.2 Case $m < n$

The arguments used for the case $m > n$ can be used in a very similar manner to prove the following claim. The only change is in the definition of optimum (maximum distance from Γ as opposed to minimum distance as for $m > n$) and, consequently, the constraints on the shape of the feasible region Γ. We omit the details.

Theorem 4. *When $m < n$, if the optimum mechanism is strategy-proof, then Γ is such that the optimum is constant.*

4.3 Case $m = n$

We now complete our proof for the case in which $m = n > 1$; we leave the arguably less interesting case of two agents open for future research.

Theorem 5. *When $m = n > 1$, if the optimum mechanism is strategy-proof, then Γ is such that the optimum is constant.*

Proof. Recall that when $m = n$, the optimum requires to choose the point in Γ with maximum/minimum x-coordinate (depending on the sign of δ). We will then prove that for all $f, f' \in \Gamma$, it holds $x_{f'} = x_f$.

Consider the instance $\mathbf{x}$ comprised of four agents, where

$$\begin{gathered} 0 < x_1 < x_2 < x_3 < x_4 < l \\ p_1 = p_3 = 1, p_2 = p_4 = -1, \\ x_4 - x_3 = x_3 - x_2 = x_2 - x_1 = \xi \\ x_3 < l - 2\xi \\ x_2 > 2\xi, \end{gathered}$$

for some suitable $\xi > 0$. Let f be the optimum for instance $\mathbf{x}$. Observe that δ has negative sign and therefore f has minimum x-coordinate in Γ (recall that $\gamma = 0$ here). Assume by contradiction that the claim is not true and let f' be a point in Γ such that $x_{f'} > x_f$. Let ϵ be the quantity defined in (3) for f and f'.

If $\epsilon < x_3$, then the third agent has an incentive to declare $x_3' = l$ so that $\delta' > 0$ (this is guaranteed by the definition of the instance) and the optimum becomes f', which is closer to x_3 than f. Similarly, when $\epsilon > x_2$, the second agent has an incentive to declare $x_2' = 0$ to change the sign of δ. Observe that one of these two conditions on ϵ must be true, since $x_3 \neq x_2$. Thus, wherever ϵ is, there at least exists one agent who has incentives to lie – a contradiction.

5 Conclusions

We have introduced a new perspective in the research on mechanism design without money. Whereas the quality of the solutions, in terms of their approximation guarantee, has been used as a way to obtain truthfulness (or strategy-proofness), we propose here to use the feasibility of the solution space as a way to get incentive-compatibility. The former is usually detrimental to the designer, who instead might well be in charge of defining feasibility. Just as one aims at the best possible approximation, here we would aim at having the largest possible set of feasible solutions.

In addition to this conceptual contribution, our work has given a set of involved technical contributions showing a dichotomy in the case study of heterogeneous facility location problem. Whilst any feasible region can be used to design optimal GSP mechanisms when agents can lie about their preferences, very little can be done for SP mechanisms facing agents who can misreport their location.

Our work leaves a number of compelling open questions. Even only for the variant of facility location considered, one might wonder to what extent the two sides of our dichotomy generalize. For the positive side of the coin of unknown preferences, we wonder whether a similar theorem holds in the case in which the agents are located on a bidimensional subset of $\mathbb{R}^2$ rather than a segment; possible interesting case studies include agents located on the boundary of a circle, a region expressed by a quadratic function, or even an arbitrary shape. The negative part of our dichotomy for unknown locations is reminiscent of the known characterization for collusion-resistant mechanisms with money [16]. Differently from that, our constant-outcome characterization is qualified by optimality. The natural next question is then to relax the requirement of optimality to (constant) approximations. One question of interest could be: What approximation guarantee allows us to achieve truthfulness for every possible combination of positive and negative preferences? More generally, our research agenda can be applied to other mechanism design optimization problems studied in the literature; how many feasible solutions can we allow to get truthfulness?

References

1. Eleftherios Anastasiadis and Argyrios Deligkas. Heterogeneous facility location games. In *AAMAS*, pages 623–631, 2018.
2. Salvador Barberà, Faruk Gul, and Ennio Stacchetti. Generalized median voter schemes and committees. *Journal of Economic Theory*, 61(2):262 – 289, 1993.
3. Elad Dokow, Michal Feldman, Reshef Meir, and Ilan Nehama. Mechanism design on discrete lines and cycles. In *ACM EC*, pages 423–440, 2012.
4. Ulle Endriss. Judgment aggregation. In F. Brandt, V. Conitzer, U. Endriss, J. Lang, and A. D. Procaccia, editors, *Handbook of Computational Social Choice*, chapter 17. Cambridge University Press, 2016.
5. Chi Kit Ken Fong, Minming Li, Pinyan Lu, Taiki Todo, and Makoto Yokoo. Facility location games with fractional preferences, 2018.
6. D. Fotakis and C. Tzamos. On the power of deterministic mechanisms for facility location games. In *ICALP*, pages 449–460, 2013.
7. Allan Gibbard. Manipulation of voting schemes: A general result. *Econometrica*, 41(4):587–601, 1973.
8. Pinyan Lu, Xiaorui Sun, Yajun Wang, and Zeyuan Allen Zhu. Asymptotically optimal strategy-proof mechanisms for two-facility games. In *ACM EC*, pages 315–324, 2010.
9. Pinyan Lu, Yajun Wang, and Yuan Zhou. Tighter bounds for facility games. In *WINE*, pages 137–148, 2009.
10. H. Moulin. On strategy-proofness and single-peakedness. *Public Choice*, 35:437–455, 1980.
11. Noam Nisan, Tim Roughgarden, Eva Tardos, and Vijay V. Vazirani. *Algorithmic Game Theory*. Cambridge University Press, 2007.
12. Paolo Penna and Carmine Ventre. Sharing the cost of multicast transmissions in wireless networks. In *SIROCCO*, pages 255–266, 2004.
13. Ariel D. Procaccia and Moshe Tennenholtz. Approximate mechanism design without money. *ACM Trans. Economics and Comput.*, 1(4):18, 2013.
14. Ariel D. Procaccia and Moshe Tennenholtz. Approximate mechanism design without money. *ACM Trans. Economics and Comput.*, 1(4):18, 2013.
15. Mark Allen Satterthwaite. Strategy-proofness and Arrow's conditions: Existence and correspondence theorems for voting procedures and social welfare functions. *Journal of Economic Theory*, 10(2):187 – 217, 1975.
16. James Schummer. Manipulation through bribes. *J. Economic Theory*, 91(2):180–198, 2000.
17. James Schummer and Rakesh V. Vohra. Strategy-proof location on a network. *Journal of Economic Theory*, 104:405–428, 2002.
18. Paolo Serafino and Carmine Ventre. Heterogeneous facility location without money on the line. In *ECAI - Including PAIS*, pages 807–812, 2014.
19. Shaokun Zou and Minming Li. Facility location games with dual preference. In *AAMAS*, pages 615–623, 2015.