Artificial Intelligence 348 (2025) 104390

Contents lists available at ScienceDirect

Artificial Intelligence

journal homepage: www.elsevier.com/locate/artint

On the design of truthful mechanisms for the capacitated facility location problem with two and more facilities ☆

Gennaro Auricchio [a,*], Zihe Wang [b], Jie Zhang [c]

[a] Department of Mathematics, University of Padua, Via Trieste 63, Padua, 35100, Italy
[b] Renmin University of China, 59 Zhongguancun St, Haidian District, Beijing, 100872, China
[c] Department of Computer Science, University of Bath, Claverton Down, Bath, BA2 7AY, England, United Kingdom

ARTICLE INFO

Keywords:
Mechanism design
Facility location problem
Worst-case analysis

ABSTRACT

In this paper, we explore the Mechanism Design aspects of the m-Capacitated Facility Location Problem (m-CFLP) on a line, focusing on two frameworks. In the first framework, the number of facilities is arbitrary, all facilities share the same capacity, and the number of agents matches the total capacity of the facilities. In the second framework, we need to locate two facilities, each with a capacity equal to at least half the number of agents. For both frameworks, we propose truthful mechanisms with bounded approximation ratios in terms of Social Cost (SC) and Maximum Cost (MC). When $m > 2$, our results stand in contrast to the impossibility results known for the classical m-Facility Location Problem, where capacity constraints are absent. Moreover, all the proposed mechanisms are optimal with respect to MC and either optimal or near-optimal with respect to the SC among anonymous mechanisms. We then establish lower bounds on the approximation ratios that any truthful and deterministic mechanism achieves with respect to SC and MC for both frameworks. Lastly, we run several numerical experiments to empirically evaluate the performances of our mechanisms with respect to the SC or the MC. Our empirical analysis shows that our proposed mechanisms outperform all previously proposed mechanisms applicable in this setting.

1. Introduction

Mechanism design is a subfield of artificial intelligence that provides the tools to develop systems or algorithms capable of achieving desired outcomes in the presence of self-interested agents. This framework is essential in decision-making processes, where preventing strategic, self-serving behavior that could distort the final outcome is critical. Indeed, aggregating the reports of multiple agents is a key problem in various applied frameworks, ranging from facility location problems [2–4] and voting [5,6] to auction [7,8] and machine learning ensemble models [9–11], where the ultimate goal is to optimize a suitable objective function, be it the revenue of the decision maker or the social welfare attained by the community. However, selecting an outcome or making a decision based on only the agents' reported information often leads to undesirable manipulation, as the agents might act strategically and misrepresent their information to deviate from the final decision toward an outcome that they prefer. This issue is especially

☆ A previous version of this paper appeared in the proceedings of the 33rd International Joint Conference on Artificial Intelligence [1]. In this enhanced version, we conducted a comprehensive comparison between our mechanisms and other mechanisms known in the literature.

* Corresponding author.
E-mail address: gennaro.auricchio@unipd.it (G. Auricchio).

https://doi.org/10.1016/j.artint.2025.104390
Received 21 October 2024; Received in revised form 21 June 2025; Accepted 23 June 2025

Available online 7 July 2025

important in social choice problems, such as electing a winner [5,6] or eliciting the position of public facilities [2,12]. For this reason, one of the most important properties that an aggregation algorithm should possess is *truthfulness*, which ensures that no agent can benefit from misreporting their private information. The definition and study of efficient truthful algorithms, or *mechanisms*, able to aggregate the private information of a group of strategic agents is the defining issue of Mechanism Design [13]. Unfortunately, the truthfulness requirement often conflicts with the goal of optimizing the social objective, so that to prevent any agent manipulation the mechanism is forced to select a suboptimal choice. To measure this efficiency loss, Nisan and Ronen introduced the concept of *approximation ratio*, which quantifies the worst-case ratio between the social objective achieved by a truthful mechanism and the optimal social objective across all possible agents' reports [13].

A classic problem in this context is the m-Facility Location Problem (m-FLP). In its simplest form, the m-FLP involves placing m facilities for n self-interested agents, each aiming to have a facility positioned as close as possible to their location. Moreover, since each facility can serve an unlimited number of agents, the mechanism only needs to determine the locations of the facilities, leaving the agents free to choose their preferred facilities without concerns of capacity constraints. In this paper, we study the m-Capacitated Facility Location Problem (m-CFLP) on the line [14]. The m-CFLP is a natural extension of the m-Facility Location Problem in which every facility has a capacity limit, i.e. each facility has a maximum number of agents that can serve. Considering facilities with capacity constraints is a natural approach for modeling scenarios where facilities offer limited resources, as it happens in distribution planning [15] and telecommunication network design [16,17]. For example, facilities can represent servers, with agents as tasks awaiting execution. Alternatively, facilities could be grocery stores, while agents represent customers needing service.

Our main result shows that, from a mechanism design perspective, the study of the m-FLP and m-CFLP differs significantly: enforcing a capacity limit to each facility allows us to elude the impossibility results known for the m-FLP [12,18]. In particular, this is the case when we have m facilities with equal capacity k and the number of agents is $n = km$. For this class of problems, we characterize both the upper and lower bounds of the approximation ratio of anonymous, truthful, and deterministic mechanisms for $m \geq 2$. These two bounds have finite values and do coincide, making the bounds tight. It is also noteworthy that the study of the m-CFLP is contingent upon the specifics of the problem. In fact, even in case $m = 2$, the properties of any mechanism depend on factors such as whether different facilities have different capacities [3], whether the total capacity is greater than the number of agents [19], or whether the capacity of each facility is lower than a critical threshold [3]. For this reason, the few results that provide tight lower bounds on the approximation ratio are limited to very specific settings [3]. The second main contribution of this paper presents a framework for addressing the 2-CFLP, unifying the majority of known results for this problem. Additionally, it establishes both an upper and lower bound on the achievable approximation ratio, which coincide in most cases.

Our contribution and paper overview. In this paper, we study two relevant frameworks for the m-CFLP from a Mechanism Design perspective. First, we study the m-CFLP with equi-capacitated facilities and no spare capacity, i.e. the m-CFLP in which all the facilities have the same capacity, namely k, and the total capacity of the facilities equals the number of agents. We present two truthful and anonymous mechanisms, the Propagating Median Mechanism (PMM) and the Propagating InnerPoint Mechanism (PIPM). We show that both the PMM and the PIPM have a bounded approximation ratio with respect to the Social Cost (SC) and the Maximum Cost (MC), regardless of the value of m. This result stands in contrast with the classic results for the m-FLP, according to which no mechanism can be deterministic, anonymous, truthful, and achieve a finite approximation ratio when $m > 2$, even on the line [12,18]. We then present three lower bounds for the approximation ratio for the m-CFLP with equi-capacitated facilities and no spare capacities. In particular: (i) no truthful and deterministic mechanism achieves an approximation ratio lower than 2 with respect to MC. Thus, the PMM and the PIPM are optimal with respect to this metric. (ii) When $k > 3$, no truthful and deterministic mechanism can achieve an approximation ratio lower than 3 with respect to the SC. (iii) No truthful, deterministic, and anonymous mechanism achieves an approximation ratio with respect to SC lower than $\left(\frac{k(m-1)}{2} + 1\right)$, if m is odd, or lower than $\left(\frac{km}{2} - 1\right)$ if m is even. In particular, the PMM and the PIPM are the best possible truthful, deterministic, and anonymous mechanisms for odd and even m, respectively.

We then study the 2-CFLP with abundant capacities, in which we have two facilities capable of accommodating at least half of the agents, that is $c_1, c_2 \geq \left\lfloor \frac{n}{2} \right\rfloor$. This framework has been studied under further assumptions: (i) in [3], the authors studied the case in which n is even, $c_1 = c_2 = \frac{n}{2}$, and proposed the InnerPoint (IM) Mechanism, (ii) in [19], the author studied the case in which n is odd and $c_1 = \left\lceil \frac{n}{2} \right\rceil$, $c_2 = \left\lfloor \frac{n}{2} \right\rfloor$, and proposed the InnerChoice (IC) Mechanisms, and (iii) in [19], the author also studied the case in which n is arbitrary but $c_1 = c_2$, and proposed the InnerGap (IG) Mechanism. However, this is the first time that a study of a framework encompassing all these different cases has been conducted. We propose the Extended InnerGap (EIG) Mechanism, which generalizes and includes IM, IC, and IG. The EIG is strong Group Strategyproof, thus truthful, and attains a bounded approximation ratio with respect to the SC and the MC. We then provide a lower bound on the approximation ratio of any truthful mechanism with respect to the SC and the MC and show that the EIG is optimal with respect to the MC. Moreover, the EIG is optimal with respect to the SC whenever $n \geq \bar{c} + \sqrt{\bar{c}}$, where $\bar{c} = \max\{c_1, c_2\}$. In Table 1, we summarize our findings in terms of lower and upper bounds for the cases we study.

A previous version of this paper was published in the proceedings of the 33rd International Joint Conference on Artificial Intelligence [1]. In this enhanced version, we have improved the presentation and simplified most of the proofs. Additionally, we conducted a comprehensive comparison between our mechanisms and other established methods in the literature, assuming agents' positions are generated by a probability distribution. Our results show that both the PMM and the PIMP outperform all other known mechanisms for locating more than one capacitated facility. Furthermore, these two mechanisms are the only ones whose performance improves as the number of facilities increases, whereas the performance of other mechanisms declines if the number of facilities increases.

Table 1
Each row contains the Lower and Upper Bounds (LB and UB, respectively) with respect to the Social and Maximum Cost for a different class of problems for the m-CFLP. The value $\bar{c}$ is the maximum capacity of the facilities. The LB column contains the lower bounds for the class of truthful and deterministic mechanisms. The LB^* column contains the lower bounds for the class of mechanisms that are truthful, deterministic, and anonymous.

	Social Cost			Maximum Cost	
	LB	LB*	UB	LB	UB
$c_j = k$ $n = km$	3	$\frac{k(m-1)}{2} + 1$ (m odd) $\frac{km}{2} - 1$ (m even)	$\frac{k(m-1)}{2} + 1$ (m odd) $\frac{km}{2} - 1$ (m even)	2	2
$c_1, c_2 \geq \left\lfloor \frac{n}{2} \right\rfloor$ $c_1 + c_2 \geq n$	3	$n - \bar{c} - 1$	$\max\{n - \bar{c}, \frac{\bar{c}}{n-\bar{c}}\} - 1$	2	2

Related works. The m-Facility Location Problem (m-FLP) and its variants are relevant problems in several applied fields such as disaster relief [20], supply chain management [21], healthcare [22], clustering [23], and public facilities accessibility [24]. The Mechanism Design study of the m-FLP was first explored by Procaccia and Tennenholtz, who laid the foundation of this field in their pioneering work [2]. While this first work approached the problem under the assumption that agents live on a line, several following works extended the study to encompass a wider variety of geometrical domains, such as trees [25,26], circles [27,28], general graphs [29,30], and generic metric spaces [31,32]. Likewise, other studies extended the framework proposed by Procaccia and Tennenholtz by considering agents' preferences that are not single peaked. For example, in [33] the authors considered the case in which agents' preferences are symmetric and double-peaked, whereas in [34] it has been considered the case in which facilities are heterogeneous, thus agents have different levels of appreciation for different facilities. Other works considered the obnoxious facility location problem, where agents dislike the facility, hence they would like to have it located as far as possible from their real location [35,36].

However, all the positive results on the Facility Location Problem are limited to cases where the mechanism designer needs to place 2 facilities or the mechanism is not deterministic. Indeed, it has been shown that no mechanism capable of placing more than two uncapacitated facilities on a line can achieve a bounded approximation ratio while being deterministic, anonymous, and truthful [12]. This result was later also extended to higher dimensional spaces endowed with the Manhattan metric [18].

The m-Capacitated Facility Location Problem (m-CFLP) is a natural extension of the m-FLP, in which each facility has a maximum number of agents that it can serve. While the algorithmic aspects of this problem have been widely considered in the literature [37,14,38], the Mechanism Design aspects of the m-CFLP have received relatively little attention until recently. There are two game theoretical frameworks to describe the m-CFLP:

1. In the first game theoretical framework, the mechanism decides both the positions of the facilities and the set of agents that are accommodated by the facility. In this way, the mechanism designer is able to ensure that no facility is overloaded with agents. This framework has been introduced in [3], where the authors studied various truthful mechanisms (such as the InnerPoint Mechanism and the Extended Endpoint Mechanism) and studied their approximation ratios. Notably, only mechanisms capable of locating two facilities achieve a bounded approximation ratio. A more theoretical analysis of the problem has been then presented in [19], where the author demonstrated that no mechanism can locate more than two capacitated facilities while being truthful, anonymous, and Pareto optimal.
2. In the second game theoretical framework, the mechanism only decides where the facilities are located, then how agents get accommodated by the facilities, and thus their utility, is decided via a First-Come-First-Served game. This framework has been introduced in [39], where the authors studied the problem of locating a single facility unable to serve all the agents taking part in the eliciting routine. This initial work was later complemented by [40] where the authors extended the study to the case in which there are two or more facilities to locate and by [41] where agents are supported over higher dimensional spaces.

In this paper, we focus on the first framework, as it is the most extensively studied, while the second framework is designed for cases where the total capacity of the facilities is lower than the total number of agents. Lastly, it is worth mentioning that the m-FLP and the m-CFLP have been studied from a Bayesian Mechanism Design viewpoint [42,43]. In particular, in [44] the authors show that it is always possible to retrieve a percentile mechanism [4] whose expected cost converges to the optimal expected cost when the number of agents increases. In [45] the authors study how to select the best extended ranking mechanism [3] to solve m-CFLP when agents are distributed according to a known probability distribution. We refer the reader to [46] for a comprehensive survey of the Mechanism Design aspects of Facility Location Problems.

2. Preliminaries

In this section, we introduce the two m-CFLP frameworks studied in this paper and establish the main notation. Throughout the paper, we assume that the agents lay on a line and denote with $\vec{x} := (x_1, \dots, x_n) \in \mathbb{R}^n$ the vector containing their positions. Moreover, we assume $m > 1$, since the 1-CFLP is equivalent to the classic 1-FLP.

The m-CFLP with equi-capacitated facilities and no spare capacity. In the first framework, we have m facilities whose capacity is the same, *i.e.* $c_j = k > 1$ for every $j \in [m]$, and the total capacity of the facilities equals the total number of agents, hence $n = mk$. We call this framework the m-CFLP with *equi-capacitated facilities* and *no spare capacity*. Since in this setting all the facilities have the same capacity, a facility location is defined by two objects: (i) a m-dimensional vector $\vec{y} = (y_1, \dots, y_m)$ whose entries are the positions of the facilities on the line, and (ii) an agents-to-facility matching $\mu \subset [n] \times [m]$ that determines how the agents are assigned to facilities, i.e. $(i, j) \in \mu$ if and only if the agent at x_i is assigned to y_j. Due to the capacity constraints, the degree of every vertex $j \in [m]$ according to μ is at most k. Moreover, since every agent is assigned to only one facility, the degree of $i \in [n]$ according to μ is 1.

The 2-CFLP with abundant capacities. In the second framework we consider, we have two facilities whose capacities, namely c_1 and c_2 are such that $\left\lfloor \frac{n}{2} \right\rfloor \le c_2, c_1 \le n - 1$. We call this framework 2-CFLP with *abundant capacities*. Since the facilities may have different capacities, eliciting two positions, namely y_1 and y_2, and an agents-to-facility matching μ is not sufficient, as we need to also specify the capacity of each facility. Thus, in this framework, a facility location is defined by three objects: (i) a vector $\vec{y} = (y_1, y_2) \in \mathbb{R}^2$ whose entries are the positions of the facilities, (ii) a permutation $\pi : [2] \to [2]$ that specifies the capacity of each facility, so that if $\pi(1) = j$ the facility at y_1 has capacity c_j and the facility built at y_2 has capacity c_i with $i \neq j$ $i, j \in [2]$, and (iii) a matching $\mu \subset [n] \times [2]$ that determines how the agents are assigned to facilities. Notice that, owing to the nature of the problem, the degree of every vertex $j \in [2]$ according to μ must be at most $c_{\pi(j)}$, while the degree of every $i \in [n]$ according to μ is 1.

Mechanism design framework for the m-CFLP. In both frameworks, given the positions of the facilities $\vec{y}$ and a matching μ, we define the cost of an agent positioned in x_i as $c_{i,\mu}(x_i, \vec{y}) = |x_i - y_{j_i}|$, where (i, j_i) is the unique edge in μ adjacent to i. Finally, a cost function is a map $C_\mu : \mathbb{R}^n \times \mathbb{R}^m \to [0, +\infty)$ that associates to $(\vec{x}, \vec{y})$ the overall cost of placing the facilities at $\vec{y}$ and assigning the agents positioned at $\vec{x}$ according to μ. For both frameworks, given a vector $\vec{x} \in \mathbb{R}^n$ containing the agents' positions, the *m-Capacitated Facility Location Problem* with respect to the cost C, consists in finding the locations for m facilities and a matching μ that minimize the value of the function $\vec{y} \to C_\mu(\vec{x}, \vec{y})$. Throughout the paper, we consider

- the *Social Cost* (SC), defined as the sum of all the agents' costs, i.e.,

$$SC_\mu(\vec{x}, \vec{y}) = \sum_{i \in [n]} c_{i,\mu}(x_i, \vec{y}) = \sum_{i \in [n]} |x_i - y_{j_i}|,$$

  where (i, j_i) is the only edge in μ that is adjacent to i or, equivalently, j_i is the unique index in $[m]$ for which $(i, j_i) \in \mu$ and
- the *Maximum Cost* (MC), defined as the maximum cost among all agents' costs, i.e.

$$MC_\mu(\vec{x}, \vec{y}) := \max_{i \in [n]} c_{i,\mu}(x_i, \vec{y}) = \max_{i \in [n]} |x_i - y_{j_i}|,$$

  where (i, j_i) is the only edge in μ that is adjacent to i or, equivalently, j_i is the unique index in $[m]$ for which $(i, j_i) \in \mu$.

In what follows, we omit μ from the indexes of c and C if it is clear from the context which matching we are considering.

A mechanism for the m-CFLP is a function f that takes the private information of n self-interested agents as input and returns a facility location. Thus, for the m-CFLP with equi-capacitated facilities and no spare capacity, the mechanism returns a set of locations $\vec{y}$ and a matching μ between the agents and the facilities. A mechanism f is said to be *truthful* (or *strategy-proof*) if, for every agent, its cost is minimized when it reports its true position, i.e., for any $x_i' \in \mathbb{R}$, we have

$$c_i(x_i, f(\vec{x})) \le c_i(x_i, f(\vec{x}_{-i}, x_i'))$$

where x_i is the agent's real position and $\vec{x}_{-i}$ is the vector $\vec{x}$ without its i-th component. A mechanism f is strong Group Strategyproof (GSP) if no group of agents can misreport their positions in such a way that (i) the cost of every agent in the group after manipulating is less than or equal to the cost they would get by reporting truthfully, (ii) at least one of the agents in the group incurs a strictly lower cost after the group manipulation.

Albeit a truthful mechanism prevents agents from misreporting their position, their output is usually suboptimal. To evaluate this efficiency loss, we consider the approximation ratio of the mechanism introduced in [13]. Given a truthful mechanism f, its approximation ratio with respect to the SC is defined as

$$ar_{SC}(f) := \sup_{\vec{x} \in \mathbb{R}^n} \frac{SC_f(\vec{x})}{SC_{opt}(\vec{x})},$$

where $SC_f(\vec{x})$ is the SC of the solution returned by f and $SC_{opt}(\vec{x})$ is the optimal SC achievable on instance $\vec{x}$. Likewise, the approximation ratio of f with respect to the MC is the highest ratio between the MC achieved by f and the optimal MC, that is

$$ar_{MC}(f) := \sup_{\vec{x} \in \mathbb{R}^n} \frac{MC_f(\vec{x})}{MC_{opt}(\vec{x})}.$$

3. The m-CFLP with equi-capacitated facilities and no spare capacity

In this section, we focus on the Mechanism Design aspects of the m-CFLP with equi-capacitated facilities and no spare capacity, i.e. given the number of facilities m and their capacity k, it holds $n = mk$. Given $j \in [m]$ and a vector $\vec{x}$ containing the agents' positions ordered from left to right, i.e. $x_i \le x_{i+1}$ for every $i \in [n-1]$, we define I_j as follows

$$I_j = \{x_{(j-1)k+i} \text{ where } i \in [k]\}. \tag{1}$$

Notice that, since every facility has the same capacity, that is $c_1 = c_2 = \cdots = c_m$, the optimal solution to the m-CFLP with respect to the SC places the facilities at the positions $\vec{y} = (y_1, \dots, y_m)$, where each y_j is a median of set I_j, and then it assigns every agent whose position is in I_j to y_j.

Likewise, the optimal solution with respect to the MC places the facilities at $y_j = \frac{x_{(j-1)k+1} + x_{jk}}{2}$ and then it assigns the agents in I_j to y_j.

3.1. The propagating median mechanism

We now introduce and study our first truthful mechanism for the m-CFLP with equi-capacitated facilities and no spare capacity, the Propagating Median Mechanism (PMM).

Mechanism 1 *(Propagating Median Mechanism (PMM)).* Let n be the total number of agents, m the number of facilities to place, and $k = \frac{n}{m} \in \mathbb{N}$ the capacity of each facility. Let us set $r = \left\lfloor \frac{m+1}{2} \right\rfloor$. The routine of the PMM is as follows:

(i) First, we locate the facility y_r at one of the medians of I_r, i.e.

$$y_r = x_{k(r-1)+\left\lfloor \frac{k+1}{2} \right\rfloor}.$$

(ii) Second, we determine the positions of the other facilities via the following iterative routine. For any $l \ge r$, let us be given the position of the l-th facility, namely y_l. Then the position of the $(l+1)$-th facility is

$$y_{l+1} = \max\{x_{kl+1}, x_{kl} + d_l\},$$

where d_l is the distance between y_l and $x_{kl} := \max_{x \in I_l} x$. Similarly, given $l \le r$ and the position of the l-th facility, namely y_l, the $(l-1)$-th facility is placed at $y_{l-1} = \min\{x_{k(l-1)}, x_{k(l-1)+1} - d_l\}$, where d_l is the distance between y_l and $x_{k(l-1)+1} := \min_{x \in I_l} x$.

(iii) Finally, all the agents in I_j are assigned to y_j.

Let $\vec{y} = (y_1, \dots, y_m)$ be the positions of the facilities returned by the PMM on a given instance $\vec{x}$. It is easy to see that the entries of $\vec{y}$ are non-decreasing, i.e. $y_j \le y_{j+1}$ for every $j \in [m-1]$. Moreover, the PMM assigns every agent to its closest facility, so that

$$\min_{\ell \in [m]} |x_i - y_\ell| = |x_i - y_j|,$$

where $j \in [m]$ is the only index for which it holds $x_i \in I_j$.

Theorem 1. *The PMM is truthful.*

Proof. Toward a contradiction, let x_i be the real position of an agent able to manipulate by reporting x_i' instead of its real position x_i. We denote with $\vec{y}$ and I_j the positions of the facilities returned by the PMM and the sets in (1) on the truthful input, respectively. Similarly, we denote with $\vec{y}'$ and I_j' the positions of the facilities returned by the PMM and the sets in (1) when the agent at x_i reports x_i', respectively. Since the other case is symmetric, we assume that $y_r \le x_i$, where $r = \left\lfloor \frac{m+1}{2} \right\rfloor$ and $y_r = x_{k(r-1)+\left\lfloor \frac{k+1}{2} \right\rfloor}$, is a median of I_r.

We first show that no agent in I_r can manipulate. Toward a contradiction, let us assume that $x_i \in I_r$ is the real position of an agent able to manipulate the mechanism. If $x_i = y_r$, the cost of the agent is null, thus it cannot benefit by misreporting. Without loss of generality, let us assume that $y_r < x_i$, as the other case follows by a symmetric argument. Let us denote with x_i' a report that increases the utility of the agent. If $x_i' < y_r$, we have that $x_i' \in I_\ell'$ with $\ell \le r$. In this case, we have that $y_r' \le y_r$, thus $y_l' \le \cdots \le y_r' \le y_r < x_i$, which means that the manipulating agent is assigned to a facility that is not closer than y_r, hence its cost does not decrease after the manipulation. Finally, let us consider the case $y_r < x_i'$. In this case, we have that $y_r' = y_r$, since the median of I_r is the same regardless of whether the manipulating agent reports truthfully or not. Thus, if $x_i' \in I_r'$, it will still be assigned to $y_r' = y_r$, which brings no benefit to the manipulative agent. So it must be that $x_i' \notin I_r'$, then $x_i' \in I_\ell'$, where $\ell > r$, thus the manipulating agent is assigned to $y_\ell' \ge y_{r+1}'$, since $\ell > r$. Let us denote with x_{rk}' the position of the (rk)-th agent from the left in the manipulated instance (x_i', x_{-i}). Since $x_i' > x_i$, we have $x_{rk}' \ge x_{rk} \ge x_i$, hence

$$y_{r+1}' = \max\{x_{rk+1}', x_{rk}' + |y_r' - x_{rk}'|\} \ge x_{rk}' + |y_r' - x_{rk}'|$$

$$\geq x_i + |y_r - x'_{rk}| \geq x_i + |y_r - x_i|,$$

so the cost of being assigned to y'_ℓ is no less than the cost of being assigned to y_r.

To conclude the proof, we need to consider the case in which $x_i \in I_j$, with $j > r$. We have two cases to analyze, depending on whether $y_j = x_{k(j-1)+1}$ or $y_j = x_{k\ell} + |y_\ell - x_{k\ell}|$, for an index $r < \ell < j$.

If $y_j = x_{k(j-1)+1}$, we have that the facility is placed on the leftmost agent of I_j, thus this case is analogous to the case in which $x_i \in I_r$.

Let us then consider the case in which $y_j = x_{k\ell} + |y_\ell - x_{k\ell}|$ for $\ell < j$. We break this case into two subcases: (i) the manipulating agent is assigned to the j-th facility after manipulating and (ii) the manipulating agent is assigned to another facility after it manipulates.

Let us consider the first subcase. By definition of PMM, there is only one case in which a single agent in I_j can alter the position of y_j while still being assigned to the j-th facility: the position of the manipulative agent is the leftmost position in I_j and all the other positions in I_j are on the right of y_j. In this case, however, the manipulative agent can only move the j-th facility further to the right, which increases its cost.

Let us then consider the second subcase: the agent at x_i manipulates in such a way that it is assigned to the ℓ'-th facility, i.e. $y'_{\ell'}$. If $\ell' > j$, we have that $x'_i > x_i$, $x'_t = x_t$ for every $t < i$, and

$$y_j = \max\{x_{k\ell} + |y_\ell - x_{k\ell}|, x_{k(j-1)+1}\} \leq \max\{x_{k\ell} + |y_\ell - x_{k\ell}|, x'_{k(j-1)+1}\} = y'_j \leq y'_{\ell'},$$

hence the agent cannot benefit from the misreport in this case since $x_i \leq y_j \leq y'_{\ell'}$.

To conclude, we consider the case $\ell' < j$. If $\ell' \leq r$, we have that $y'_{\ell'} \leq y'_r \leq y_r \leq y_\ell$, thus the cost of the agent does not decrease since we have

$$|x_i - y_{\ell'}| \geq |x_i - y_r| \geq |x_i - y_j|.$$

If $r < \ell' \leq \ell$, we have $x'_t = x_t$ for every $t \leq k(\ell' - 1)$. Since

$$y'_{\ell'} = \max\{x'_{k(\ell'-1)+1}, x'_{k(\ell'-1)} + d'_{(\ell'-1)}\} = \max\{x'_{k(\ell'-1)+1}, x_{k(\ell'-1)} + d_{(\ell'-1)}\}$$

and $x'_{k(\ell-1)+1} \leq x_{k(\ell-1)+1}$, we have $y'_{\ell'} \leq y_{\ell'} \leq y_\ell$, thus

$$|x_i - y'_{\ell'}| \geq |x_i - y_{\ell'}| \geq |x_i - y_j|.$$

Finally, if $\ell < \ell' < j$, we have $x'_t = x_t$ for every $t \leq k(\ell' - 1)$, thus $y_{\ell'} = y_j = x_{k\ell} + |y_\ell - x_{k\ell}| = x'_{k\ell} + |y'_\ell - x'_{k\ell}| = y'_{\ell'}$, hence the cost of the manipulative agent is unchanged. $\square$

Although the PMM is truthful, it is not strong GSP, as the following example shows.

Example 1. Let us fix $n = 9$, $m = 3$, and $k = 3$. Let us consider the following instance: $x_1 = x_2 = x_3 = 0$, $x_4 = x_5 = 1$, $x_6 = 2$, $x_7 = 2.5$, and $x_8 = x_9 = 4$. The PMM places the facilities at $y_1 = 0$, $y_2 = 1$, and $y_3 = 3$. In this instance, the agents at x_6 and x_7 can collude: indeed, if x_6 reports $x'_6 = 1$, the PMM places the facilities at $y_1 = 0$, $y_2 = 1$, and $y_3 = 2.5$, thus the cost of x_7 decreases while the cost incurred by agent x_6 is unchanged.

To conclude, we provide an analysis of the approximation ratio of the PMM with respect to the SC and MC. In particular, we prove that ar_{SC}(PMM) and ar_{MC}(PMM) are finite.

Theorem 2. *It holds* $ar_{SC}(PMM) = k\left\lfloor \frac{m}{2} \right\rfloor + 1$.

Proof. Let $r = \left\lfloor \frac{m+1}{2} \right\rfloor$ and $\vec{x}$ be a vector containing all the agents' reports ordered from left to right. Let $SC_{opt}(\vec{x})$ be the optimal SC for $\vec{x}$, then it holds that

$$SC_{opt}(\vec{x}) = \sum_{j \in [m]} SC_{opt}(I_j),$$

where $SC_{opt}(I_j)$ is the SC of the agents whose reports are in I_j according to the optimal solution. Similarly, let $SC_{\text{PMM}}(\vec{x})$ be the SC of instance $\vec{x}$ according to the output of PMM, then it holds that

$$SC_{\text{PMM}}(\vec{x}) = \sum_{j \in [m]} SC_{\text{PMM}}(I_j),$$

where $SC_{\text{PMM}}(I_j)$ is the SC of the agents whose reports are in I_j according to the output of PMM.

We denote with $\vec{y}$ the vector containing the locations of the facilities returned by the PMM and define $J \subset [m]$ as the set of indexes $j \in [m]$ such that $y_j \in [x_{(j-1)k+1}, x_{jk}]$. We notice that J is non-empty since $r \in J$ by definition of the PMM. We notice that $SC_{opt}(I_r) = SC_{\text{PMM}}(I_r)$. Let us now consider $j \in J$ such that $j \neq r$. Since $y_j \in [x_{(j-1)k+1}, x_{jk}]$, we have that

$$SC_{\text{PMM}}(I_j) \le (k-1)|x_{(j-1)k+1} - x_{jk}| \le (k-1)SC_{opt}(I_j).$$

Let us now consider $j \notin J$ and, without loss of generality, let us assume that $r < j$ since the other case is symmetric. Since $j \notin J$, there exists an index $\ell \in J$ such that $y_j = x_{k\ell} + |y_\ell - x_{k\ell}|$. If $\ell \ne r$, we have that

$$SC_{\text{PMM}}(I_j) \le k|y_\ell - x_{k\ell}| \le kSC_{opt}(I_\ell)$$

since, for every $x_t \in I_j$, we have $x_{k\ell} \le x_t \le y_j$. Similarly, if $\ell = r$, we have that $SC_{\text{PMM}}(I_j) \le k|y_r - x_{kr}|$. Therefore, it holds

$$\begin{aligned} SC_{\text{PMM}}(\vec{x}) \le & \sum_{J \ni j \ne r} (k\lambda_j + k - 1)SC_{opt}(I_j) + SC_{opt}(I_r) \\ & + k\gamma_r|y_r - x_{kr}| + k\gamma_l|y_r - x_{k(r-1)+1}| \\ \le & \sum_{j \in J, j \ne r} (k\lambda_j + k - 1)SC_{opt}(I_j) + (k\Gamma + 1)SC_{opt}(I_r) \end{aligned}$$

where

1. λ_j for $j \in J$ and $j \ne r$, is the number of $i \in [m]$ such that $i \ne j$ and $y_i = x_{jk} + |y_j - x_{jk}|$ if $j > r$ and the number of $i \in [m]$ such that $i \ne j$ and $y_i = x_{jk} - |y_j - x_{jk}|$ if $j < r$,
2. γ_l is the number of $i \in [m]$ such that $i < r$ and $y_i = x_{k(r-1)+1} - |y_r - x_{k(r-1)+1}|$,
3. γ_r is the number of $i \in [m]$ such that $i > r$ and $y_i = x_{kr} + |y_r - x_{kr}|$, and
4. $\Gamma = \max\{\gamma_r, \gamma_l\}$.

If we set $K_r = (k\Gamma + 1)$ and $K_j = (k\lambda_j + k - 1)$ for every $j \in J$ such that $j \ne r$, we have

$$ar_{SC}(\text{PMM}) \le \frac{\sum_{j \in J} K_j SC_{opt}(I_j)}{\sum_{j \in J} SC_{opt}(I_j)},$$

thus $ar_{SC}(\text{PMM}) \le \max_{j \in J}\{K_j\}$. Since, $\lambda_j \le \left\lfloor \frac{m}{2} \right\rfloor - 1$ and $\Gamma \le \left\lfloor \frac{m}{2} \right\rfloor$, we have $K_j \le k\left\lfloor \frac{m}{2} \right\rfloor + 1$ for every $j \in J$, therefore $ar_{SC}(\text{PMM}) \le k\left\lfloor \frac{m}{2} \right\rfloor + 1$.

Finally, to prove that $ar_{SC}(\text{PMM}) = \left(k\left\lfloor \frac{m}{2} \right\rfloor + 1\right)$, consider the following instance: $x_1 = \cdots = x_{kr-1} = 0$ and $x_{kr} = \cdots = x_n = 1$. The optimal cost of this instance is 1. The PMM places the facilities as it follows $y_1 = \cdots = y_r = 0$ and $y_{r+1} = \cdots = y_m = 2$, thus the Social Cost of the mechanism is $n - (k\left\lfloor \frac{m+1}{2} \right\rfloor - 1) = k\left\lfloor \frac{m}{2} \right\rfloor + 1$. □

Through a similar argument, we retrieve the approximation ratio of PMM with respect to the MC.

Theorem 3. *It holds $ar_{MC}(PMM) = 2$.*

Proof. First, we rewrite the optimal and the mechanism MC as the maximum of m different costs. Indeed, we have $MC_{opt}(\vec{x}) = \max_{j \in [m]}\left\{MC_{opt}(I_j)\right\}$ and $MC_{\text{PMM}}(\vec{x}) = \max_{j \in [m]}\left\{MC_{\text{PMM}}(I_j)\right\}$, respectively, where $MC_{opt}(I_j)$ is the MC of the agents whose reports are in I_j according to the optimal solution and $MC_{\text{PMM}}(I_j)$ is the MC of the agents whose reports are in I_j according to the output of the PMM. Denoted with $\vec{y}$ the output of the PMM, we define J as the set of indexes such that $y_j \in [x_{(j-1)k+1}, x_{jk}]$. By definition of PMM, $r \in J$, thus J is non-empty. If $j \in J$, using the same argument used to prove Theorem 2, we retrieve $MC_{\text{PMM}}(I_j) \le 2MC_{opt}(I_j)$. If $j \notin J$ and $j > r$, there exists an index $\ell \in J$ such that $r \le \ell < j$ and $y_j = x_{k\ell} + |y_\ell - x_{k\ell}|$. For every $x_t \in I_j$, we have $x_{k\ell} \le x_t$ and $|x_t - y_j| \le |x_{k\ell} - y_j| = |y_\ell - x_{k\ell}| \le MC_{\text{PMM}}(I_r)$. Similarly, if $j \notin J$ and $j < r$, we conclude that there exists $j < \ell \le r$ such that $|x_t - y_j| \le MC_{\text{PMM}}(I_\ell)$, for every $x_t \in I_j$. Then, we have $\max_{j \in [m]} MC_{\text{PMM}}(I_j) \le 2\max_{j \in J} MC_{opt}(I_j)$ and $\max_{j \in J} MC_{opt}(I_j) \le \max_{j \in [m]} MC_{opt}(I_j)$, thus $ar_{MC}(PMM) \le 2$. Hence the approximation ratio with respect to the MC is less than 2. To prove that this bound is tight, consider the instance used in the proof of Theorem 2. □

3.2. The propagating InnerPoint mechanism

We now present our second truthful mechanism for the m-CFLP with equi-capacitated facilities and no spare capacity, the Propagating InnerPoint Mechanism (PIPM). The routine of the PIPM is similar to the routine of the PMM, the main difference lies in how the algorithm determines the positions of the first facilities. Indeed, the PMM places a facility at the median of $I_{\left\lfloor \frac{m+1}{2} \right\rfloor}$, while the PIPM places two facilities: one at the maximum value of $I_{\left\lfloor \frac{m}{2} \right\rfloor}$ and one at the minimum value of $I_{\left\lfloor \frac{m}{2} \right\rfloor + 1}$.

Mechanism 2 *(Propagating InnerPoint Mechanism)*. Let n be the total number of agents, m be the number of facilities to place and $k = \frac{n}{m} \in \mathbb{N}$ be the capacity of each facility. Let us set $r = \left\lfloor \frac{m}{2} \right\rfloor$. The mechanism runs as follows:

(i) First, we locate the facilities y_r and y_{r+1} at the positions x_{rk} and x_{rk+1}, respectively.
(ii) To place the other facilities, we run an iterative routine. For $l \geq r+1$, given the position of the l-th facility, namely y_l, we place the $(l+1)$-th facility at the position $y_{l+1} = \max\{x_{kl+1}, x_{kl} + d_l\}$, where d_l is the distance between y_l and x_{kl}. For $l \leq r$, we run a similar iterative routine. Given the position of the l-th facility y_l, the $(l-1)$-th facility is placed at $\min\{x_{k(l-1)}, x_{k(l-1)+1} - d_l\}$, where d_l is the distance between y_l and $x_{k(l-1)+1}$.
(iii) Finally, all the agents in I_j are assigned to y_j.

Due to the similarities between the definitions of the PMM and the PIPM, it is possible to adapt the arguments used in the proof of Theorem 1, 2, and 3 to this mechanism. In particular, the PIPM is truthful and achieves a bounded approximation ratio with respect to both the SC and MC.

Theorem 4. *The PIPM is truthful. Moreover, we have that* $ar_{SC}(PIPM) = k\left\lceil \frac{m}{2} \right\rceil - 1$ *and* $ar_{MC}(PIPM) = 2$.

Proof. We divide the proof into three pieces: in the first one we show that PIPM is truthful, in the second one we compute the approximation ratio of PIPM with respect to the Social Cost, and in the third one we compute the approximation ratio of PIPM with respect to the Maximum Cost.

PIPM is truthful. Toward a contradiction, let x_i be the real position of an agent able to manipulate. We denote with x_i' the position that the agent uses to manipulate the mechanism. We denote with $\{y_j\}_{j=1,\dots,m}$ the positions of the facilities returned by the PIPM on the truthful input and with y_j' the positions of the facilities returned by the PIPM when x_i reports x_i'. Notice that the output of the PIPM is ordered in non-decreasing order, that is $y_1 \leq y_2 \leq \cdots \leq y_m$. Finally, we recall that

$$I_j = \{x_{k(j-1)+1}, x_{k(j-1)+2}, \dots, x_{k(j-1)+k}\}$$

and that, according to the PIPM, every agent in I_j is assigned to the j-th facility. By a symmetry argument, without loss of generality, we assume that $y_{r+1} \leq x_i$.

First, we show that if the agent at x_i is able to manipulate, then $x_i \notin I_{r+1}$. Toward a contradiction, let us assume that $x_i \in I_{r+1}$. Then, if $x_i = y_{r+1}$, the cost of the agent is null, thus it cannot benefit by misreporting. Thus, it must be that $x_i > y_{r+1}$. If $y_{r+1} < x_i' < x_i$, the output of the mechanism does not change. If $x_i' < y_{r+1}$, we have that $y_r' \leq y_{r+1}' \leq y_{r+1}$ and, since $x_i' \in I_\ell$ with $\ell \leq r+1$, we have that the facility to which the manipulating agent is assigned is further to the left than y_{r+1} so its cost is increased after the manipulation. Finally, let us consider the case $x_i' > x_i$. In this case, we have that $y_{r+1}' = y_{r+1}$, since the $(rk+1)$-th report from the left is the same regardless of whether the manipulating agent reports truthfully or not. Thus, if $x_i' \in I_{r+1}'$, it will still be assigned to $y_{r+1}' = y_{r+1}$, which brings no benefit to the manipulative agent. If $x_i' \notin I_{r+1}'$, then $x_i' \in I_\ell'$, where $\ell > r+1$, thus the manipulating agent is assigned to $y_\ell' \geq y_{r+2}'$, since $\ell > r+1$. Let us denote with $x_{(r+1)k}'$ the position of the $((r+1)k)$-th agent from the left in the manipulated instance (x_i', x_{-i}). Notice that, since $x_i' > x_i$, we have $x_{(r+1)k}' \geq x_{(r+1)k} \geq x_i$. Moreover, owing to the fact that $x_{(r+1)k}' \geq x_{r(k+1)} \geq y_{r+1}$, we infer that

$$\begin{aligned}
y_{r+2}' &= \max\{x_{(r+1)k+1}', x_{(r+1)k}' + |y_{r+1}' - x_{(r+1)k}'|\} \\
&\geq x_{(r+1)k}' + |y_{r+1}' - x_{(r+1)k}'| \\
&> x_i + |y_{r+1} - x_{(r+1)k}'| \\
&\geq x_i + |y_{r+1} - x_{(r+1)k}| \\
&\geq x_i + |y_{r+1} - x_i|,
\end{aligned}$$

hence the cost of being assigned to y_ℓ' is always greater or equal to the cost of being assigned to y_{r+1}. We then conclude that $x_i \notin I_{r+1}$.

Lastly, let us consider the case in which $x_i \in I_j$, with $j > r+1$. We have two cases to analyze, depending on whether $y_j = x_{k(j-1)+1}$ or $y_j = x_{kr'} + |y_{r'} - x_{kr'}|$, with $r' < j$. If $y_j = x_{k(j-1)+1}$, we have that the facility is placed at the position of the leftmost agent in I_j, thus this case is analogous to the case in which $x_i \in I_{r+1}$.

Finally, let us consider the case in which $y_j = x_{kr'} + |y_{r'} - x_{kr'}|$. We break this case into two subcases: (i) the manipulative agent is still assigned to the j-th facility after manipulating, and (ii) the manipulative agent is assigned to another facility after it manipulates. Let us consider the first scenario. By definition of PIPM, there is only one case in which a single agent in I_j can manipulate the position of y_j while still being assigned to the j-th facility: the position of the manipulative agent is the leftmost position in I_j and all the other agents in I_j are on the right of y_j. In this case, however, the manipulative agent can only move the j-th facility further to the right, which increases its cost.

Let us then consider the second possible case: the manipulative agent reports in such a way that it is assigned to the ℓ-th facility, i.e. y_ℓ'. Without loss of generality, let us assume that $\ell < j$. Since the routine that determines the positions of the facilities of the PIPM is the same as the one used by the PMM, we can adapt the argument used to prove Theorem 2 to infer that the agent is unable to lower its cost also in this case.

The approximation ratio of the PIPM with respect to the Social Cost. For the sake of simplicity, let us consider the case in which m is even so that $r = \frac{m}{2}$ is an integer. The case in which m is odd is similar. Notice that, since $n = km$, n is also even. Given instance $\vec{x}$, we recall that

$$SC_{opt}(\vec{x}) = \sum_{j=1}^{m} SC_{opt}(I_j)$$

and

$$SC_{\text{PIPM}}(\vec{x}) = \sum_{j=1}^{m} SC_{\text{PIPM}}(I_j)$$

hold, where $SC_{opt}(I_j)$ is the Social Cost of the agents whose position is in I_j according to the optimal solution, while $SC_{\text{PIPM}}(I_j)$ is the Social Cost of the agents whose position is in I_j according to the output of the PIPM. Notice that, by definition of PIPM, agents whose position is in I_j are assigned to the same facility. Let $\vec{y} = (y_1, \dots, y_m)$ be the positions at which the PIPM places the facilities when the input is $\vec{x}$. For every $j \in [m]$ we have two cases: either $y_j \in [x_{(j-1)k+1}, x_{jk}]$ or $y_j \notin [x_{(j-1)k+1}, x_{jk}]$. We denote with $J \subset [m]$ the set of indexes j for which it holds $y_j \in [x_{(j-1)k+1}, x_{jk}]$. Notice that, by definition of PIPM, we have $r, r+1 \in J$, thus J is always not empty.

Let us assume that $j \in J$, so that $y_j \in [x_{(j-1)k+1}, x_{jk}]$. In this case, as previously shown in the proof of Theorem 2, we have that

$$SC_{\text{PIPM}}(I_j) \le (k-1)|x_{k(j-1)+1} - x_{kj}| \le (k-1) SC_{opt}(I_j).$$

Let us now consider the case in which $y_j \notin [x_{(j-1)k+1}, x_{jk}]$. By definition of y_j, we must have that $x_{jk} < y_j = x_{r'k} + |x_{r'k} - y_{r'}|$, where $r' < j$. Since $x_{r'k} \le x_{(j-1)k+1} \le \dots \le x_{jk} < y_j = x_{r'k} + |x_{r'k} - y_{r'}|$, we have that

$$|y_j - x_{(j-1)k+l}| < |x_{r'k} - y_{r'}| \le SC_{opt}(I_{r'})$$

for every $l = 1, 2, \dots, k$. In particular, we have $SC_{\text{PIPM}}(I_j) \le k SC_{opt}(I_{r'})$. Therefore, we have that

$$\begin{aligned} \frac{SC_{\text{PIPM}}(\vec{x})}{SC_{opt}(\vec{x})} = \frac{\sum_{j=1}^{m} SC_{\text{PIPM}}(I_j)}{\sum_{j=1}^{m} SC_{opt}(I_j)} &\le \frac{\sum_{j\in J}(t_j k + k - 1) SC_{opt}(I_j)}{\sum_{j=1}^{m} SC_{opt}(I_j)} \\ &\le \frac{\sum_{j\in J}(t_j k + k - 1) SC_{opt}(I_j)}{\sum_{j\in J} SC_{opt}(I_j)} \end{aligned} \tag{2}$$

where $J \subset [m]$ is the non-empty set of indexes j for which it holds $y_j \in [x_{(j-1)k+1}, x_{jk}]$, while t_j is the number of sets I_ℓ such that $y_\ell = y_j$. Let us now rewrite (2) as

$$\frac{SC_{\text{PIPM}}(\vec{x})}{SC_{opt}(\vec{x})} \le \sum_{j\in J} \beta_j \alpha_j,$$

where $\beta_j = (t_j k + k - 1)$ and $\alpha_j = \frac{SC_{opt}(I_j)}{\sum_{j\in J} SC_{opt}(I_j)}$. Since $\sum_{j\in J} \alpha_j = 1$, we have that

$$\frac{SC_{\text{PIPM}}(\vec{x})}{SC_{opt}(\vec{x})} \le \max_{j\in J} \beta_j.$$

Finally, since $r, r+1 \in J$, we have that the maximum possible value of t_j is $r-1$, hence $\max \beta_j \le (r-1)k + k - 1 = \frac{m}{2}k - 1 = \frac{n}{2} - 1$. We then infer that the approximation ratio of PIPM is less than or equal to $\frac{n}{2} - 1$.

To prove that the bound is tight, consider the following instance: $x_1 = x_2 = \dots = x_{\frac{n}{2}-1} = 1$, $x_{\frac{n}{2}} = 2$, $x_{\frac{n}{2}+1} = 3$, and $x_{\frac{n}{2}+2} = x_{\frac{n}{2}+3} = \dots = x_n = 4$. It is easy to see that the optimal cost of this instance is 2. On this instance, the PIPM places y_r at 2, y_{r+1} at 3, all the facilities y_ℓ with $\ell < r$ at 0, and all the other ones at 5. The cost of the mechanism is then $2(\frac{n}{2} - 1)$, thus the approximation ratio of the mechanism is greater or equal to $\frac{n}{2} - 1$, which concludes the proof.

The approximation ratio of the PIPM with respect to the Maximum Cost. Let us now consider the Maximum Cost. Again, we have that

$$MC_{opt}(\vec{x}) = \max_{j\in[m]} \left\{ MC_{opt}(I_j) \right\}$$

and

$$MC_{\text{PIPM}}(\vec{x}) = \max_{j\in[m]} \left\{ MC_{\text{PIPM}}(I_j) \right\},$$

where $MC_{opt}(I_j)$ is the Maximum Cost of the agents whose position is in I_j according to the optimal solution and $MC_{\text{PIPM}}(I_j)$ is the Maximum Cost of the agents whose position is in I_j according to the output of the mechanism. If $j \in J$, using the same argument used for the Social Cost, we retrieve $MC_{\text{PIPM}}(I_j) \le 2 MC_{opt}(I_j)$. If $j \notin J$, we can use the same argument used to prove Theorem 3 to show that there exists an index $\ell \in J$ such that $\ell < j$ and $|x_\ell - y_j| \le MC_{\text{PIPM}}(I_\ell)$ for every $x_\ell \in I_j$. Then, we have

$$\frac{\max_{j\in[m]} \left\{ MC_{\text{PIPM}}(I_j) \right\}}{\max_{j\in[m]} \left\{ MC_{opt}(I_j) \right\}} \le \frac{\max_{j\in[m]} \left\{ MC_{\text{PIPM}}(I_j) \right\}}{\max_{j\in J} \left\{ MC_{opt}(I_j) \right\}} \le \frac{\max_{j\in J} \left\{ MC_{\text{PIPM}}(I_j) \right\}}{\max_{j\in J} \left\{ MC_{opt}(I_j) \right\}}$$

$$\leq \frac{\max_{j\in J}\left\{2MC_{opt}(I_j)\right\}}{\max_{j\in J}\left\{MC_{opt}(I_j)\right\}} \leq 2.$$

Hence the approximation ratio with respect to the Maximum Cost is no more than 2. To prove that this bound is tight, it suffices to consider the same instance used for the Social Cost case. □

Albeit $ar_{MC}(\text{PMM}) = ar_{MC}(\text{PIPM})$, the approximation ratios of the two mechanisms with respect to the SC are different. Indeed, we have that

$$ar_{SC}(\text{PMM}) < ar_{SC}(\text{PIPM}) \qquad \text{if } m \text{ is odd}$$
$$ar_{SC}(\text{PIPM}) < ar_{SC}(\text{PMM}) \qquad \text{if } m \text{ is even.}$$

Finally, it is easy to adapt Example 1 to show that PIPM is not strong GSP.

3.3. Lower bounds for the approximation ratio

To conclude the section, we study the lower bounds for the approximation ratio of truthful mechanisms for the m-CFLP with equi-capacitated facilities and no spare capacity. First, we show that 2 is the best approximation ratio for any truthful and deterministic mechanism with respect to the MC.

Theorem 5. *No deterministic truthful mechanism for the m-CFLP with equi-capacitated facilities and no spare capacity can achieve an approximation ratio with respect to the Maximum Cost that is lower than* 2.

Proof. Let k be the capacity of the m facilities, hence $n = mk$ is the total number of agents. Toward a contradiction, let M be a truthful and deterministic mechanism such that $ar_{MC}(M) = 2 - \delta$ where $\delta > 0$. Let us consider the following instance: $x_1 = 0$ and $x_2 = \cdots = x_n = 2$. It is easy to see that the optimal MC is 1. Since there is no spare capacity, we have that there exists a facility, namely y, that serves the agent in 0 and $k-1$ agents located at 1. Since $ar_{MC}(M) = 2 - \delta$, we must have that the position of the facility y lies in the interval $(\delta, 2-\delta)$.

Let us now consider the instance $x_1 = y$ and $x_2 = \cdots = x_n = 2$. By the same argument used before, we have that there exists a facility, namely y', that serves both the agent located at y and $k-1$ agents located at 2. Owing to the fact that $ar_{MC}(M) = 2 - \delta$, we infer that $y < y' < 2$, but this is impossible, as it would mean that an agent located at y would benefit by reporting 0 instead of its true position. We therefore conclude that $ar_{MC}(M) \geq 2$. □

We now move to the lower bound for the Social Cost.

Theorem 6. *No deterministic truthful mechanism for the m-CFLP with equi-capacitated facilities and no spare capacity can achieve an approximation ratio with respect to the Social Cost that is lower than* 3 *whenever* $k > 3$.

Proof. Given m facilities with capacity k, let $n = mk$ be the number of agents and let M be a truthful mechanism. Without loss of generality, we consider the case in which the output of the mechanism locates the facilities at the positions $\vec{y} = (y_1, \dots, y_m)$ such that

$$\min_{i=1,\dots,n}\{x_i\} \leq y_1 \leq y_2 \leq \cdots \leq y_m \leq \max_{i=1,\dots,n}\{x_i\}$$

for every input $\vec{x} = (x_1, x_2, \dots, x_n)$. Indeed, if M does not satisfy this property, we have that, up to a reordering, changing the positions of the facilities from y_i to $\bar{y}_i$ defined as

$$\bar{y}_i = \max\left\{\min\left\{y, \max_{i=1,\dots,n}\{x_i\}\right\}, \min_{i=1,\dots,n}\{x_i\}\right\}$$

and leaving the agents-to-facility matching unchanged decreases the cost of every agent and thus the overall Social Cost of the mechanism.

Let us consider the following instance: $x_1 = \cdots = x_{k+1} = 0$ and $x_{k+2} = \cdots = x_n = 2$. Let $\vec{y} = (y_1, y_2, \dots, y_m)$ be the facilities positions returned by the mechanism M on such instance and let γ be the agents-to-facility matching produced by M. It is easy to see that the Social Cost induced by the output of M is always greater than the following quantity

$$SC_{\vec{y}}(\vec{x}) := \min_{\pi} \sum_{i=1}^{n}\sum_{j=1}^{m} |x_i - y_j|\pi_{i,j}, \tag{3}$$

where π is a feasible agents-to-facility matching, that is

- $\pi_{i,j} \in \{0,1\}$,
- $\sum_{i=1}^{n} \pi_{i,j} = k$ for every $j = 1, \dots, m$, and
- $\sum_{j=1}^{m} \pi_{i,j} = 1$ for every $i = 1, \dots, n$.

Owing to a classic result on the 1-dimensional matching problem, we have that the solution to problem (3) is monotone, hence every agent in I_j is assigned to y_j. In particular, we infer the following lower bound

$$SC_M(\vec{x}) \geq SC_{\vec{y}}(\vec{x}) = \sum_{j=1}^{m} \sum_{x_i \in I_j} |x_i - y_j| \geq \sum_{x_i \in I_j} |x_i - y_j| \tag{4}$$

for every input $\vec{x} = (x_1, \dots, x_n)$ and every $j = 1, \dots, m$. Therefore, from now on, we assume that the agents-to-facility matching returned by the mechanism assigns every agent in I_j in y_j, as otherwise the SC attained by the mechanism would be larger.

We now split the proof of the theorem into two steps. First, we show that $y_2 = y_3 = \cdots = y_m = 2$ and then we prove our lower bound by constructing a couple of *ad hoc* instances.

Step 1: Owing to the bound in (4), we infer that if $y_2 = 0$, we have that $ar_{SC}(M) \geq k-1$, as $SC_M(\vec{x}) \geq 2(k-1)$ and the optimal Social Cost is equal to 2. Let us now consider the case $y_2 \in (0,2)$. Without loss of generality, let us assume that one of the agents located at 0 is assigned to the facility located at y_2. Indeed, since there are $k+1$ agents located at 0, not all of them can be assigned to y_1. Moreover, we can assume that at least one of the agents located at 0 is assigned to y_2 because if it was assigned to y_j with $j > 2$, the total Social Cost would be increased since $y_2 \leq y_j$. We therefore assume that one of the agents located at 0, namely x_{k+1}, gets assigned to y_2.

Consider now the instance $x_1 = \cdots = x_k = 0$, $x'_{k+1} = y$, and $x_{k+2} = \cdots = x_n = 2$ and let us denote with y'_j the positions of the facilities according to M. Owing to the truthfulness of the mechanism, we must have that $y'_2 = y_2 = y$ and that agent the agent located at y is assigned to y'_2, as otherwise the agent located at y would be able to manipulate the mechanism by reporting 0 over its real position y. Owing to the bound in (4), we then infer that $ar_{SC}(M) \geq k-1$.

Step 2: We thus conclude that if $ar_{SC}(M) < k-1$, it must be the case that $y_2 = 2$, hence $y_2 = y_3 = \cdots = y_m = 2$ and $y_1 < 2$. If $y_1 > 0$, we can leverage truthfulness to move every agent located at 0 to the position of y_1 without changing the output of the mechanism. Thus, up to a translation and a scalar factor, we can consider the case in which $y_1 = 0$, $y_2 = \cdots = y_m = 2$, $x_1 = \cdots = x_{k+1} = 0$ and $x_{k+2} = \cdots = x_n = 2$.

To complete the proof, consider an agent located at 0 that is not assigned to a facility located at 2. For the sake of simplicity, we denote that agent by x_1. Given $\epsilon > 0$, let us now consider the instance $x'_1 = 1-\epsilon$, $x'_2 = \cdots = x'_{k+1} = 0$, and $x'_{k+2} = \cdots = x'_n = 2$. Let y' be the position of the facility at which x_1 is assigned. If $y' = 2$, the agent x_1 can manipulate by reporting 0 over its true position x'_1 as it would result in it being assigned to a facility located at 0. If $y' = 0$, then the approximation ratio of the mechanism is at least equal to

$$\frac{2+1-\epsilon}{1+\epsilon}$$

which converges to 3 when $\epsilon \to 0$. Indeed, even if we locate the remaining m facilities in the best possible way and assign the other agents in the most efficient way, an agent located at 0 necessarily shares a facility with an agent located at 2, thus the conclusion. Finally, if $y' \in (0,2)$, by the same argument used in Step 1, we can consider the instance $x''_1 = y'$, $x''_2 = \cdots = x''_{k+1} = 0$ and $x''_{k+2} = \cdots = x''_n = 2$ and infer that the approximation ratio is at least equal to $k-1$. We therefore conclude that $ar_{SC}(M) \geq 3$. □

Finally, we present a lower bound for the approximation ratio with respect to the SC of deterministic, anonymous, and truthful mechanisms. We recall that a mechanism M is anonymous if every agent's outcome depends only on its reports, i.e. two agents swapping two different reports cause the mechanism to swap their outcomes as well.

Theorem 7. *No deterministic, anonymous, and truthful mechanism for the m-CFLP with equi-capacitated facilities and no spare capacity can achieve an approximation ratio with respect to the Social Cost that is lower than* $\left(\frac{k(m-1)}{2}+1\right)$ *if m is odd or lower than* $\left(\frac{km}{2}-1\right)$ *if m is even.*

Proof. Let M be a deterministic, anonymous, and truthful mechanism. Let us consider the following instance $x_1 = \cdots = x_{kr+1} = 0$ and $x_{kr+2} = \cdots = x_n = 1$, where $r = \left\lfloor \frac{m}{2} \right\rfloor$. The optimal cost of this instance is 1. First, we show that, according to M, the locations of the facilities serving the agents at 0 are all placed at the same distance from 0. Toward a contradiction, let us assume that M places two facilities at two positions, namely y and y', such that $|y-0| \neq |y'-0|$ and that both facilities serve at least an agent that reported 0. Without loss of generality, let us assume that $|y-0| = |y| < |y'| = |y'-0|$. Let us denote by x_i one of the agents who reported 0 that is assigned to y. Let us now consider the instance (y, x_{-i}). Since M is truthful, we must have that the mechanism places a facility at y and that the agent at y is assigned to it, as otherwise it could misreport by reporting 0. Let us now denote with x_j one of the agents in 0 that is assigned to y'. Since M is anonymous, if x_j reports y, it is assigned to y, which is closer to 0 than y', which contradicts the truthfulness of M. In particular, we infer that all agents placed at 0 incur the same cost. Similarly, all agents placed at 1 incur the same cost. Since there is no spare capacity, there exists at least one facility that serves an agent placed at 0 and an agent placed at 1, let us denote with $\lambda \in \mathbb{R}$ its position on the line. Then, the total cost of the mechanism is

$$C = |\lambda|\left(k\left\lfloor\frac{m}{2}\right\rfloor + 1\right) + |1-\lambda|\left(n - k\left\lfloor\frac{m}{2}\right\rfloor - 1\right).$$

Finally, we notice that $C \geq \left(\frac{k(m-1)}{2}+1\right)$ if m is odd and $C \geq \left(\frac{km}{2}-1\right)$ if m is even, which concludes the proof. □

Table 2
Frameworks under which the mechanisms operate when $c_1, c_2 \geq \left\lfloor \frac{n}{2} \right\rfloor$. From right to left, the column tells us whether the mechanism is capable of working (1) for every number of agents n, (2) when the total capacity is larger than the number of agents, and (3) when the two facilities have different capacities.

	$\forall n \in \mathbb{N}$	$n < c_1 + c_2$	$c_1 \neq c_2$
EIG	Yes	Yes	Yes
IG	Yes	Yes	No
IC	No	No	Yes
IM	No	No	No

Since the PMM and the PIPM are anonymous, the lower bound in Theorem 7 is tight. In fact, PMM achieves the lower bound for odd values of m, while PIPM achieves the lower bound for even values of m. Therefore, for the m-CFLP with equi-capacitated and no spare capacity, PMM and PIPM are the best anonymous, deterministic, and truthful mechanisms for odd and even m, respectively.

4. The 2-CFLP with abundant capacities

We now consider the case in which we have to place two facilities capable of accommodating at least half of the agents. In particular, we present the Extended InnerGap (EIG) mechanism, a truthful mechanism that generalizes and includes mechanisms that operate under further assumptions: the InnerPoint (IM) Mechanism, which works under the assumption that n is even and $c_1 = c_2 = \frac{n}{2}$ [3], the InnerGap (IG) Mechanism, which works under the assumption that $c_1 = c_2$ [19], and the InnerChoice (IC) Mechanism, which works under the assumption that n is odd and $c_1 = \left\lceil \frac{n}{2} \right\rceil$ and $c_2 = \left\lfloor \frac{n}{2} \right\rfloor$ [19]. We summarize in Table 2 how the EIG extends the other three mechanisms.

We show that EIG achieves a finite approximation ratio with respect to the SC and the MC and corroborate these results by providing lower bounds on the approximation ratio achievable by truthful and deterministic mechanisms with respect to SC and MC. As a consequence, we infer the approximation ratio of the IC and IG mechanisms, which to the best of our knowledge were previously unknown in the literature.

Mechanism 3 *(Extended InnerGap Mechanism).* Let $\bar{c} := \max\{c_1, c_2\}$ and let $\vec{x} \in \mathbb{R}^n$ be the vector containing the agents' report ordered from left to right. Let us fix $y_1 = x_{n-\bar{c}}$, $y_2 = x_{\bar{c}+1}$, and $z = \frac{y_1+y_2}{2}$, let n_1 be the number of agents in $[y_1, z] \cap \{x_i\}_{i\in[n]}$ and n_2 be the number of agents in $(z, y_2] \cap \{x_i\}_{i\in[n]}$. Then the output of the EIG over $\vec{x}$ is (i) to place the facility with the largest capacity at y_1 and the other at y_2 if $n_1 \geq n_2$; or (ii) to place the facility with the lowest capacity at y_1 and the other at y_2 if $n_2 > n_1$. In both cases, every agent is assigned to the facility closer to its reported position.

Theorem 8. *The EIG is strong GSP, hence truthful.*

Proof. Let $\vec{x}$ be the vector containing the agents' true positions. We denote with $y_1 \leq y_2$ the positions of the facilities according to the EIG on the truthful input. Let $I := \{x_{i_1}, \dots, x_{i_s}\}$ be the real positions of the agents that form a coalition able to manipulate the output of the EIG. Without loss of generality, we assume that I is minimal, that is no subset of the agents in I can collude. We recall that the EIG places the two facilities: one at the $(n-\bar{c})$-th agent's report from the left, namely y_1, and one at the $(\bar{c}+1)$-th agent's report from the left, namely y_2. Since I is minimal, none of the agents whose true position coincides with y_1 or y_2 takes part in the group manipulation. Let us now consider a coalition of agents I that is able to lower the cost of an agent, whose real location is x_{i_1}, without increasing the cost of the other agents in I. Let us assume first that $x_{i_1} < y_1$, hence $y_1' < y_1$. If $y_1' < y_1$, it must be the case that at least one agent whose real position, namely x_t, was on the right of y_1 reports a position on the left of y_1, i.e. $x_t' \in I_1'$, where x_t' is the misreport of an agent in I whose real position was x_t. If that agent was assigned to y_2 in the truthful input, it must be that $|x_t - y_2| \leq |x_t - y_1| < |x_t - y_1'|$, since $y_1' < y_1 \leq x_t$. Thus, the agent at x_t is increasing its cost, which contradicts $x_t \in I$. Similarly, if x_t was assigned to y_1 according to the truthful input, its cost still increases after the manipulation since $y_1' < y_1$, which concludes the proof.

By a symmetry argument, we get to the same conclusion when $x_{i_1} > y_2$.

Finally, let us assume that $y_1 < x_{i_1} < y_2$. Again, since the cost of agent x_{i_1} decreases, we must have that at least one of the two facilities moves closer to x_{i_1}. Without loss of generality, let us assume that after the manipulation the position of the leftmost facility, namely y_1' is such that $y_1 < y_1'$. By the same argument used before, we have that at least one agent whose real position is on the left of y_1 reports a position that is on the right of y_1, which increases its cost regardless of whether it is assigned to y_1' or y_2'. We therefore conclude that the EIG is strong GSP. □

The EIG mechanism determines the facility position using the same routine used by a percentile mechanism [4]. However, the percentile mechanisms are not strong GSP in general, while the EIG mechanism is, as we show in the next example. This difference is due to the fact that the EIG forces the agents to use a specific facility, while the percentile mechanism does not.

Example 2. Let us consider the following FLP problem with 5 agents and 2 facilities. We have that $x_1 = 0$, $x_2 = 1$, $x_3 = x_4 = 2$, and $x_5 = 4$. Let us consider the percentile mechanism induced by the percentile vector $\vec{p} = (0.25, 0.75)$. By definition of the percentile mechanism, we have that the facilities will be placed at the position of the $(\lfloor (5-1)0.25 \rfloor + 1)$-th and $(\lfloor (5-1)0.75 \rfloor + 1)$-th agents from the left, i.e. x_2 and x_4 in the truthful input. Notice that the $(0.25, 0.75)$-percentile mechanism places the facilities at the same position as the EIG mechanism if both the facilities have capacity equal to 3. However, if we use the percentile mechanism, the agent x_4 and x_1 can collude: indeed, if x_4 reports 0 instead of 2, the new input is $(0, 0, 1, 2, 4)$, thus the facilities are placed at 0 and 2 which reduces the cost of the agent at 0 and leaves the cost of the agent at 2 unchanged.

Notice that, if we used the EIG mechanism to locate the facilities, the agent x_4 would be forced to be assigned to the facility at 0, which prevents the group manipulation.

4.1. The approximation ratio of the EIG

In this section, we study the approximation ratio of the EIG mechanism and prove that the EIG attains the best possible approximation ratio when $n \geq \bar{c} + \sqrt{\bar{c}}$.

Theorem 9. *It holds that* $ar_{SC}(EIG) = \max\{(n - \bar{c} - 1), (\frac{\bar{c}}{n-\bar{c}} - 1)\}$. *Moreover, it holds that* $ar_{MC}(EIG) = 2$.

Proof. We divide the proof into two parts, one dedicated to the Maximum Cost and the other dedicated to the Social Cost.

Approximation ratio for the MC. Let us denote with I_i the set of agents that are assigned to the facility with capacity c_i according to the optimal solution and, without loss of generality, we assume that all the agents in I_1 are placed to the left of the agents in I_2. The optimal MC is then

$$\frac{1}{2} \max\{|\min\{I_1\} - \max\{I_1\}|, |\min\{I_2\} - \max\{I_2\}|\}.$$

Let $y_1 \leq y_2$ be the position at which the mechanism places the two facilities. Then the MC of the EIG is lower or equal to the MC attained by assigning all the agents in I_i to the facility at y_i. Finally, since $x_{n-\bar{c}} \in I_1$ and $x_{\bar{c}+1} \in I_2$, we infer that

$$\begin{aligned} MC_{EIG}(\vec{x}) &\leq \max\{\max_{x \in I_1} |x - x_{n-\bar{c}}|, \max_{x \in I_2} |x - x_{\bar{c}+1}|\} \\ &\leq \max\{|x_1 - \max_{x \in I_1}\{x\}|, |\min_{x \in I_2}\{x\} - x_n|\} \\ &\leq 2MC_{opt}(\vec{x}), \end{aligned}$$

thus $ar_{MC}(EIG) \leq 2$. Lastly, let us define $\vec{x}$ as $x_1 = \cdots = x_{\bar{c}+1} = 0$, and $x_{\bar{c}+2} = \cdots = x_n = 1$. The optimal MC is 0.5, while the MC attained by the EIG mechanism is 1, hence $ar_{MC}(EIG) = 2$.

Approximation ratio for the SC. Let $\vec{x}$ be the vector containing the agents' reports sorted in non-decreasing order, that is $x_i \leq x_{i'}$ for every $i \leq i'$. Since we are placing two facilities, the optimal solution splits the agents into two continuous sets, one served by the facility with capacity c_1 and the other served by the facility with capacity c_2. We denote with I_i the set of agents assigned to the facility with capacity c_i according to the optimal solution. Without loss of generality, let us assume that $a \leq b$ for every $a \in I_1$ and $b \in I_2$, thus if we denote with y_i the position of the facility serving the agents in I_i, it holds $y_1 \leq y_2$. Let us now denote with $y_1^{(M)} \leq y_2^{(M)}$ the positions at which the EIG mechanism places the two facilities and let us denote with $I_i^{(M)}$ the set of agents assigned to the facility located at $y_i^{(M)}$ according to the output of the EIG. By definition, all the agents in $I_1^{(M)}$ are on the left of the agents in $I_2^{(M)}$. Moreover, it holds that

$$SC_{EIG}(\vec{x}) = SC_{EIG}(I_1^{(M)}) + SC_{EIG}(I_2^{(M)}),$$

where $SC_{EIG}(I_i^{(M)})$ is the SC of the agents in $I_i^{(M)}$ according to the output of EIG. Since EIG assigns the agents to their closest facility, we have that $SC_{EIG}(\vec{x}) \leq SC_{EIG}(I_1) + SC_{EIG}(I_2)$, where $SC_{EIG}(I_i)$ is the Social Cost of the agents in I_i if they are assigned to a facility placed at $y_i^{(M)}$ by the mechanism.

Let n_1 be the number of agents in I_1. Then, by definition of EIG, the value $SC_{EIG}(I_1)$ is the cost of a mechanism for the 1-FLP that, given in input the reports of n_1 agents, locates the facility at the position of the $(n - \bar{c})$-th agent to the left. This mechanism has an approximation ratio equal to

$$AR_1 \leq \max\left\{\frac{n - \bar{c} - 1}{n_1 - (n - \bar{c} - 1)}, \frac{n_1 - (n - \bar{c})}{n - \bar{c}}\right\}.$$

Indeed, given m the position of the median agent in $\{x_1, \dots, x_{n_1}\}$, it is easy to see that the worst ratio between the SC of EIG and the optimal SC is the one in which all the agents are positioned at m or at $x_{n-\bar{c}}$. If $m = x_{n-\bar{c}}$, there is nothing to prove. If $m \neq x_{n-\bar{c}}$, we have

$$SC_{opt}(x_1, \dots, x_{n_1}) \geq \min\left\{n - \bar{c}, n_1 - (n - \bar{c} - 1)\right\}|x_{n-\bar{c}} - m|.$$

Likewise, we have

$$SC_{EIG}(I_1) \leq \max\left\{n-\bar{c}-1, n_1-(n-\bar{c})\right\}|x_{n-\bar{c}}-m|,$$

hence

$$AR_1 \leq \frac{\max\{n-\bar{c}-1, n_1-(n-\bar{c})\}}{\min\{n-\bar{c}, n_1-(n-\bar{c}-1)\}} = \max\left\{\frac{n-\bar{c}-1}{n_1-(n-\bar{c}-1)}, \frac{n_1-(n-\bar{c})}{n-\bar{c}}\right\}.$$

Since $n_1 \in \{n-\bar{c}, \bar{c}\}$, it holds $\frac{n-\bar{c}-1}{n_1-(n-\bar{c}-1)} \leq n-\bar{c}-1$ and $\frac{n_1-(n-\bar{c})}{n-\bar{c}} \leq \frac{2\bar{c}-n}{n-\bar{c}}$, thus

$$SC_{EIG}(I_1) \leq \max\left\{\frac{2\bar{c}-n}{n-\bar{c}}, n-\bar{c}-1\right\}SC_{opt}(I_1).$$

Through a similar argument, we infer that

$$SC_{EIG}(I_2) \leq AR_2 \cdot SC_{opt}(I_2),$$

where

$$AR_2 = \max\left\{\frac{\bar{c}-n_1}{(n-n_1)-(\bar{c}-n_1)}, \frac{(n-n_1)-(\bar{c}-n_1+1)}{\bar{c}-n_1+1}\right\}.$$

Again, it is easy to see that

$$\frac{\bar{c}-n_1}{(n-n_1)-(\bar{c}-n_1)} = \frac{\bar{c}-n_1}{n-\bar{c}} \leq \frac{2\bar{c}-n}{n-\bar{c}}$$

and

$$\frac{(n-n_1)-(\bar{c}-n_1+1)}{\bar{c}-n_1+1} = \frac{n-\bar{c}-1}{\bar{c}-n_1+1} \leq n-\bar{c}-1,$$

thus

$$SC_{EIG}(I_2) \leq \max\left\{\frac{2\bar{c}-n}{n-\bar{c}}, n-\bar{c}-1\right\}SC_{opt}(I_2).$$

Finally, since it holds $AR_1, AR_2 \leq \max\{n-\bar{c}-1, \frac{\bar{c}}{n-\bar{c}}-1\}$, we conclude that

$$ar_{SC}(EIG) \leq \max\{n-\bar{c}-1, \frac{\bar{c}}{n-\bar{c}}-1\}.$$

Lastly, consider the following two instances. In the first one we have $x_1 = \cdots = x_{n-\bar{c}-1} = 0$, $x_{n-\bar{c}} = 1$, and $x_i = 5$ for every other $i \in [n]$. The optimal solution has a cost equal to 1, while EIG has a cost equal to $n-\bar{c}-1$, hence $ar_{SC}(EIG) \geq n-\bar{c}-1$. In the second instance we have $x_1 = \cdots = x_{n-\bar{c}} = 0$, $x_{n-\bar{c}+1} = \cdots = x_{\bar{c}} = 1$, and $x_i = 2$ for all the other $i \in [n]$. The optimal cost is $\min\{n-\bar{c}, \bar{c}-(n-\bar{c})\} = \min\{n-\bar{c}, 2\bar{c}-n\}\}$, while the mechanism cost is $(2\bar{c}-n)$, thus we have that $ar_{SC}(EIG) \geq \max\{1, \frac{2\bar{c}-n}{n-\bar{c}}\} = \max\{1, \frac{\bar{c}}{n-\bar{c}}-1\}$, which concludes the proof. □

We now provide a lower bound on the approximation ratio with respect to both the MC and SC of any truthful and deterministic mechanism for this framework. Our results show that the EIG is optimal or almost optimal for both costs.

Theorem 10. *Let M be a truthful and deterministic mechanism that places two facilities with capacity $c_1, c_2 \geq \left\lfloor \frac{n}{2} \right\rfloor$. Then, we have that $ar_{MC}(M) \geq 2$ and $ar_{SC}(M) \geq 3$. If M is also anonymous, and $n > 5$, then we have that $ar_{SC}(M) \geq (n-\bar{c}-1)$.*

Proof. Lower bound with respect to the Maximum Cost. Toward a contradiction, let us assume that M is a truthful mechanism such that $ar_{MC}(M) < 2$. Let us consider the following instance $x_1 = -2$ and $x_2 = \cdots = x_{n-1} = 0$ and $x_n = 2$. It is easy to see that the optimal MC is equal to 1.

Since we have two facilities to locate, it must be the case that at least two agents located at different positions share a common facility. Without loss of generality, we can assume that the agent x_1 located at -2 and one agent located at 0 share the same facility. Indeed, the case in which the agent x_n located at 2 and one agent located at 0 share the same facility is symmetric, while the case in which the agents x_1 and x_n, located at 2 and -2 respectively, shares the same facility entails a larger MC.

Let us then assume that M assigns the agent located at -2 and an agent located at 0 to the same facility y. Since $ar_{MC}(M) < 2$, we must have that the position of the facility y lies in the interval $(-2, 0)$, in particular $-2 < y < 0$.

Let us now consider the instance $x_1 = y$, $x_2 = \cdots = x_{n-1} = 0$ and $x_n = 2$. By the same argument used in the proof of Theorem 5, we can assume that the agent located at y and at least one agent located at 0 are both served by the same facility, namely y' that must be located between y and 0, i.e. $y < y' < 0$. However, owing to the truthfulness of M, we get a contradiction, hence $ar_{MC}(M) \geq 2$.

Lower bound with respect to the Social Cost. Let us consider the following instance: $x_1 = \cdots = x_{\bar{c}+1} = 0$ and $x_{\bar{c}+2} = \cdots = x_n = 2$. By the same argument used in the proof of Theorem 6, we have that the mechanism does not place both facilities at either 0 or 2 and no facility can be located in $(0, 2)$, as otherwise the approximation ratio would be larger than 3. We then have that one facility is located at 0 and the other one at 2. We assume that the facility with capacity $\bar{c}$ is located at 0 and the other one at 2, as otherwise the approximation ratio would be larger. Finally, let us assume that x is one of the agents located at 0 that is assigned to the facility located at 0. Owing to the truthfulness of the mechanism we have that if this agent was located at $1-\epsilon$, it would be still assigned to the facility at 0. Therefore the cost of the mechanism would be $3-\epsilon$, while the optimal cost would be $1+\epsilon$, which concludes the proof.

Lower bound for anonymous mechanisms with respect to the Social Cost. Let us consider the following instance: $x_1=\cdots=x_{\bar{c}+1}=0$ and $x_{\bar{c}+2}=\cdots=x_n=1$. The optimal Social Cost of this instance is 1. By the same argument used in Theorem 7, any truthful and anonymous mechanism places the two facilities at the same distance from 0. Denoted with λ the distance of the facilities from 0, we have that the Social Cost of M is equal to

$$|\lambda|(\bar{c}+1)+|1-\lambda|(n-\bar{c}-1).$$

Since we have that $\bar{c}\geq\left\lfloor\frac{n}{2}\right\rfloor$, we have that $\bar{c}+1\geq n-\bar{c}-1$ we have that $SC_M(\vec{x})\geq n-\bar{c}-1$, thus the approximation ratio of any truthful, anonymous, and deterministic mechanism is larger than $(n-\bar{c}-1)$. □

In particular, the EIG is the best truthful, anonymous, and deterministic mechanism whenever $n\geq\bar{c}+\sqrt{\bar{c}}$.

4.2. The EIG and previous mechanisms

To conclude, we show that the EIG mechanism extends and includes three already-known mechanisms. In particular, (i) when n is an even number and $c_1=c_2=\frac{n}{2}$, the EIG mechanism coincides with the InnerPoint Mechanism, presented in [3]. (ii) When $n=2k+1$ is odd, $c_1=k+1$, and $c_2=k$, the EIG mechanism coincides with the InnerChoice Mechanism, presented in [19]. (iii) When $c_1=c_2$, the EIG mechanism coincides with the InnerGap Mechanism, presented in [19].

Owing to the results presented in the previous section, we are able to characterize the approximation ratios of the IC and EG mechanisms, which, to the best of our knowledge, were previously unknown. For the sake of clarity, we divide the section into three subsections, one dedicated to each mechanism. Notice that the lower bounds proposed in Theorem 10 apply to these mechanisms as well.

Lastly, we notice that the only other mechanism known that is not extended by the EIG is the Extended Endpoint Mechanism (EEM). However, the approximation ratio of EEM is equal to $\frac{3n}{2}$, which is larger than the one attained by the EIG, making it suboptimal for this framework [3].

4.2.1. The innerpoint mechanism

Given an even number $n=2k$ and two facilities whose capacities are $c_1=c_2=k$, the routine of the Innerpoint Mechanism (IM) is as follows: (i) Given $\vec{x}=(x_1,\dots,x_n)$ the vector containing the agents' reports ordered from left to right. (ii) We set $y_1=x_k$ and $y_2=x_{k+1}$. Since $c_1=c_2$, we do not need to specify the capacity of the facility placed at y_1 and y_2. (iii) Lastly, every agent is assigned to its closest facility.

It is easy to see that the IM is the IG when $c_1=c_2=\frac{n}{2}$, thus all the results presented for the EIG and IG do apply to this case. In particular, we have the following result.

Theorem 11. *The Innerpoint Mechanism is strong GSP. Moreover, we have that $ar_{SC}(IM)=\frac{n}{2}-1$ and $ar_{MC}(IM)=2$.*

Proof. The strong GSP follows from Theorem 8. The results on the approximation ratios follow from Theorem 9. Indeed, since $\bar{c}=\frac{n}{2}$, we have

$$\max\left\{n-\frac{n}{2}-1,\frac{\frac{n}{2}}{n-\frac{n}{2}-1}\right\}=\frac{n}{2}-1,$$

and therefore $ar_{SC}(IM)=\frac{n}{2}-1$. □

4.2.2. The InnerChoice mechanism

Given an odd number $n=2k+1$ and two facilities whose capacities are $c_1=k+1$ and $c_2=k$, the routine of the IC mechanism is as follows: (i) Given $\vec{x}=(x_1,\dots,x_n)$ the vector containing the agents' reports ordered from left to right, i.e. $x_i\leq x_{i+1}$, we define $\delta_1=|x_{k+1}-x_k|$ and $\delta_2=|x_{k+2}-x_{k+1}|$. (ii) If $\delta_1\leq\delta_2$, we locate the facility with capacity c_1 at x_k and the other one at x_{k+2}. Otherwise, we locate the facility with capacity c_1 at x_{k+2} and the other one at x_k. (iii) Lastly, every agent is assigned to its closest facility.

Since $\bar{c}=k+1$, we have that $x_{n-\bar{c}}=x_k$ and $x_{\bar{c}+1}=x_{k+2}$, hence, for every $\vec{x}\in\mathbb{R}^n$, the output of EIG and IC are the same, thus the two mechanisms do coincide. It was shown in [19] that the IC is truthful, however, owing to Theorem 8, we infer that IC is also strong GSP.

Theorem 12. *The IC is strong Group Strategyproof.*

Proof. Since the EIG mechanism and the IC mechanism do coincide when $n=2k+1$, $c_1=k+1$, and $c_2=k$, it follows directly from Theorem 8. □

Similarly, we extend the results on the approximation ratio of the IC mechanism with respect to the SC and MC.

Theorem 13. *Let n be an odd number, then $ar_{MC}(IC) = 2$. Moreover, if $n > 5$, it holds $ar_{SC}(IC) = k - 1 = \frac{n-3}{2}$, otherwise $ar_{SC}(IC) = 1$.*

Proof. The approximation ratio of the IC with respect to the Maximum Cost follows trivially from Theorem 9.

Let us consider the approximation ratio with respect to the Social Cost. First, let us assume that $n > 5$. In this case, we have that $\bar{c} = k + 1$. From Theorem 9, we have that

$$ar_{SC}(IC) = \max\left\{n - k - 2, \frac{k+1}{n-k-2}\right\} = \max\left\{k - 1, \frac{k+1}{k-1}\right\} = k - 1,$$

since $n = 2k + 1$, which concludes the proof.

Finally, we consider the case $n = 5$, the case $n = 3$ is similar. Since $c_1 = 3$ and $c_2 = 2$, the optimal solution splits the agents' positions as $I_1 = \{x_1, x_2, x_3\}$ and $I_2 = \{x_4, x_5\}$ or as $I_1 = \{x_1, x_2\}$ and $I_2 = \{x_3, x_4, x_5\}$. Notice that, in both cases, placing the facilities at x_2 and x_4 would result in an optimal solution as long as we can correctly select where to place c_1 and c_2. Finally, we notice that if we place c_1 at x_2, the Social Cost of the solution is $|x_1 - x_2| + |x_2 - x_3| + |x_4 - x_5|$, while if we place c_1 at x_4, the cost is $|x_1 - x_2| + |x_3 - x_4| + |x_4 - x_5|$. It is then easy to see that the optimal solution places c_1 at x_2 if and only if $|x_2 - x_3| \le |x_3 - x_4|$. Since this is the routine that defines the IC mechanism, we conclude that the IC is optimal when $n = 5$, $c_1 = 3$, and $c_2 = 2$. □

4.2.3. The InnerGap mechanism

Given a number of agents n and two facilities whose capacities are $c_1 = c_2 = k \ge \frac{n}{2}$, the routine of the IG mechanism is as follows: (i) Let $\vec{x} = (x_1, \dots, x_n)$ be the vector containing the agents' reports ordered from left to right, i.e. $x_i \le x_{i+1}$. (ii) We set $y_1 = x_{n-k}$ and $y_2 = x_{k+1}$. Since $c_1 = c_2$, we do not need to specify the capacity of the facility placed at y_1 and y_2. (iii) Lastly, every agent is assigned to its closest facility.

Since $\bar{c} = c_1 = c_2 = k$, we have that $x_{n-\bar{c}} = x_{n-k}$ and $x_{\bar{c}+1} = x_{k+1}$, hence, for every $\vec{x} \in \mathbb{R}^n$, the output of EIG and IG are the same, thus the two mechanisms do coincide. It was shown in [19] that the IG is truthful, however, owing to Theorem 8, we have that IG is strong GSP.

Theorem 14. *The IG is strong Group Strategyproof.*

Proof. Since the routine of the IG is the same as the one of the EIG, the results follow from Theorem 8. □

Similarly, we extend the results on the approximation ratio of the IG mechanism with respect to the SC and MC.

Theorem 15. *Let n be an odd number, then $ar_{MC}(IG) = 2$. Moreover, it holds $ar_{SC}(IG) = \max\{(n - k - 1), (\frac{k}{n-k} - 1)\}$.*

Proof. Since the routine of the IG is the same as the one of the EIG, the result follows from Theorem 9. □

5. Numerical experiments

In this section, we complement our theoretical study of the m-CFLP by conducting several numerical experiments. We consider a Bayesian framework, in which we assume a prior distribution λ over agent preference profiles, allowing different agents to have different distributions describing their positions [42]. Noticeably, these assumptions are also common in automated mechanism design [47,48]. Indeed, in many practical contexts, such as facility location or product design, a prior model λ (e.g., derived from observations) is typically available. For these reasons, in our experiments, we sample preference profiles from different distributions and consider the average cost that different mechanisms induce.

We focus on the m-CFLP with equi-capacitated facilities and no spare capacity for three reasons: (i) first, the case in which $m > 2$ is the most interesting case since very few mechanisms are known for this framework due to the impossibility results about the classic Facility Location Problem [12]; (ii) second, the routine of the EIG is similar to the routine of the Extended Ranking Mechanisms presented in [45], hence the numerical results would be in line with what has been already studied in [45] and thus less interesting; and (iii) in this framework we have the largest amount of mechanisms to compare with, while in the 2-CFLP with abundant capacities our mechanism generalizes the vast majority of known mechanisms.

Within this framework, the aim of our experiments is twofold

- First, we want to assess how reliable the worst-case bound given in Theorem 2 and 4 are and whether, in practice, the costs induced by the PIPM and PMM are closer to the optimal cost.
- Second, we want to compare the costs attained by the PIPM and PMM with the performances of other known truthful mechanisms.

5.1. Experiment details

In this section, we report the details of our experiments.

Evaluation metrics. To compare the mechanisms and assess their efficiency when the agents' preferences are sampled from a distribution, we consider two different metrics.

The first metric is the *Bayesian approximation ratio*, which quantifies the closeness between the expected Social (Maximum) Cost induced by the mechanism and the expected optimal SC (MC) when agents' positions are sampled from a probability distribution λ [42]. Formally, the Bayesian approximation of a mechanism M with respect to the Social Cost is given by

$$B_{ar,SC}(M) := \frac{\mathbb{E}_{\vec{X}\sim\lambda}[SC_M(\vec{X})]}{\mathbb{E}_{\vec{X}\sim\lambda}[SC_{opt}(\vec{X})]}, \tag{5}$$

while the Bayesian approximation ratio with respect to the Maximum Cost is defined as

$$B_{ar,MC}(M) := \frac{\mathbb{E}_{\vec{X}\sim\lambda}[MC_M(\vec{X})]}{\mathbb{E}_{\vec{X}\sim\lambda}[MC_{opt}(\vec{X})]}, \tag{6}$$

where $\vec{X}$ is an n-dimensional random vector distributed according to λ.

The second metric is the *Average-Case approximation ratio*, which measures the average ratio of the Social Cost induced by the mechanism and the optimal Social Cost when agents' positions are sampled from a probability distribution λ [49]. Formally, given a mechanism M, the Average-Case approximation ratio with respect to the Social Cost is defined as:

$$AVG_{ar,SC}(M) = \mathbb{E}_{\vec{X}\sim\lambda}\left[\frac{SC_M(\vec{X})}{SC_{opt}(\vec{X})}\right]. \tag{7}$$

Similarly, we denote with $AVG_{ar,MC}$ the Average-Case approximation ratio with respect to the Maximum cost.

It is important to note that the Bayesian approximation ratio and the Average-Case approximation ratio are distinct measures: $B_{ar,SC}$ evaluates the expected percentage loss in Social Cost, while $AVG_{ar,SC}$ measures the average percentage loss in Social Cost.

Experiment setup. Throughout our experiments, we consider both the Social and Maximum Cost. We consider two types of agents and thus divide our experiments into two different batches.

In the first batch of experiments, we assume that agents are independent and identically distributed. In this case, we sample the agents' preferences from three probability distributions: the uniform distribution $\mathcal{U}[0,1]$, the standard normal distribution $\mathcal{N}(0,1)$, and the exponential distribution Exp(1). Owing to the results in [45], the Bayesian approximation ratio is scale-invariant, therefore testing the mechanisms over the standard Gaussian distribution is the same as testing over a generic Gaussian distribution $\mathcal{N}(\mathfrak{m},\sigma^2)$. The same holds for the Average Case approximation ratio, making our experiments comprehensive.

In the second batch of experiments, we assume that agents are independent, but not identically distributed. In this case, we assume that different agents are distributed according to different uniform distributions. To describe the agents' different distributions, we thus introduce the set $\Theta = [0,1]$ and assume that an agent whose type is θ is distributed uniformly on $[0,\theta]$. To sample the agents type θ, we consider a probability measure η whose density is $f_\eta(\theta) = 3\theta^2$. From our assumptions, the joint distribution describing the agents' position and type has density

$$f_\lambda(x,\theta) = \mathbb{I}_{[0,\theta]}(x)3\theta^2, \tag{8}$$

where $\mathbb{I}_{[0,\theta]}$ is the indicator function of the set $[0,\theta]$, which is defined as

$$\mathbb{I}_{[0,\theta]} = \begin{cases} 1 & \text{if } x \in [0,\theta], \\ 0 & \text{otherwise.} \end{cases}$$

Lastly, the number of facilities m ranges in the set $\{5,10,15\}$, while the capacity of each facility k ranges in $\{3,4,5,6,7\}$. Every empirical expected value is computed by taking the average on 100000 instances. For the sake of exposition, we omit the confidence interval of each computed value, since its spawn is always less than 0.005.

Benchmark mechanisms. In our experiments, we compare the PIPM and the PMM with other truthful mechanisms known in the literature. Since we are considering the case m-CFLP with no spare capacity and we are interested in studying cases in which $m > 2$, the only other mechanisms capable of handling such instances are the Ranking Mechanisms (RM) [3] and the Extended Ranking Mechanisms (ERM) [45]. These mechanisms have a similar routine: both the RM and the ERM place the facilities between one or two positions, namely y_1 and y_2. The points y_1 and y_2 are selected as two subsequent agents positions, so that $y_1 = x_i$ and $y_2 = x_{i+1}$ for a suitable index i.

Following the asymptotic study proposed in [45] and [44], we are able to select the optimal value of i depending on the number of agents n, the capacity of the facilities k, the number of facilities m, and the probability distribution λ. In particular, owing to the fact that the total capacity of the facilities matches the number of agents, the routines of the mechanisms we consider are as follows

- the best ERM is the median mechanism, namely Med, defined as

$$Med(\vec{x}) := (y_1, y_2, \dots, y_m) = (s, s, \dots, s),$$

Table 3
Table containing all the Bayesian approximation ratios computed for each test instance considered when the agents are distributed according to a normal distribution. For both the Social Cost and the Maximum Cost, we report in bold text the minimal value for each experiment. We denote with Med the Median mechanism (which is the best ERM) and denote with $RankMech$ the best Ranking Mechanism, defined as in (9).

		Social Cost				Maximum Cost			
		PMM	$PIPM$	Med	$RankMech$	PMM	$PIPM$	Med	$RankMech$
$k=5$	$m=3$	**1.27**	1.49	2.51	2.46	**1.77**	1.82	2.69	2.77
	$m=4$	**1.31**	1.44	3.20	2.94	1.75	**1.63**	3.09	2.99
	$m=5$	**1.34**	1.47	3.85	3.73	**1.72**	1.73	3.39	3.40
	$m=6$	**1.37**	1.48	4.49	4.25	**1.74**	1.81	3.64	3.57
	$m=7$	**1.39**	1.49	5.12	4.97	**1.77**	1.78	3.85	3.85
$k=10$	$m=3$	**1.16**	1.48	2.38	2.46	**1.71**	1.79	2.53	2.67
	$m=4$	**1.21**	1.37	3.01	2.89	1.74	**1.47**	2.85	2.81
	$m=5$	**1.21**	1.42	3.63	3.63	**1.69**	1.70	3.09	3.15
	$m=6$	**1.23**	1.45	4.23	4.12	**1.69**	1.84	3.29	3.26
	$m=7$	**1.26**	1.42	4.82	4.79	**1.74**	1.76	3.46	3.48
$k=15$	$m=3$	**1.12**	1.50	2.37	2.50	**1.68**	1.79	2.47	2.62
	$m=4$	**1.15**	1.37	3.00	2.92	1.70	**1.43**	2.76	2.73
	$m=5$	**1.16**	1.43	3.61	3.65	**1.69**	1.71	2.98	3.05
	$m=6$	**1.18**	1.46	4.21	4.14	**1.71**	1.86	3.16	3.14
	$m=7$	**1.20**	1.41	4.80	4.80	**1.74**	1.76	3.31	3.35

Table 4
Table containing all the Average Case approximation ratios computed for each test instance considered when the agents are distributed according to a normal distribution. For both the Social Cost and the Maximum Cost, we report in bold text the minimal value for each experiment. We denote with Med the Median mechanism (which is the best ERM) and denote with $RankMech$ the best Ranking Mechanism, defined as in (9).

		Social Cost				Maximum Cost			
		PMM	$PIPM$	Med	$RankMech$	PMM	$PIPM$	Med	$RankMech$
$k=5$	$m=3$	**1.28**	1.51	2.59	2.52	**1.76**	1.83	2.77	2.89
	$m=4$	**1.33**	1.46	3.29	3.01	1.74	**1.63**	3.23	3.11
	$m=5$	**1.35**	1.49	3.94	3.82	**1.71**	1.72	3.57	3.60
	$m=6$	**1.38**	1.50	4.60	4.35	**1.74**	1.80	3.86	3.78
	$m=7$	**1.40**	1.50	5.23	5.08	**1.76**	1.77	4.11	4.11
$k=10$	$m=3$	**1.17**	1.50	2.41	2.50	**1.69**	1.79	2.57	2.74
	$m=4$	**1.22**	1.39	3.05	2.93	1.72	**1.46**	2.92	2.87
	$m=5$	**1.22**	1.44	3.67	3.67	**1.67**	1.69	3.18	3.25
	$m=6$	**1.24**	1.46	4.28	4.16	**1.68**	1.83	3.40	3.37
	$m=7$	**1.27**	1.43	4.87	4.83	**1.73**	1.74	3.58	3.61
$k=15$	$m=3$	**1.13**	1.52	2.39	2.52	**1.67**	1.80	2.49	2.68
	$m=4$	**1.16**	1.38	3.03	2.95	1.69	**1.41**	2.80	2.78
	$m=5$	**1.16**	1.44	3.64	3.68	**1.67**	1.70	3.04	3.12
	$m=6$	**1.19**	1.47	4.24	4.17	**1.70**	1.85	3.23	3.21
	$m=7$	**1.21**	1.42	4.83	4.84	**1.72**	1.75	3.39	3.44

where s is the median of the input vector $\vec{x}$. The agents-to-facility matching assigns every agent in I_j to y_j.

- Owing to the specific structure of the probability measures, we have that the best RM, namely RM_{best}, locates $\lfloor \frac{m}{2} \rfloor$ facilities at $x_{k(\lfloor \frac{m}{2} \rfloor)}$ and the remaining facilities at $x_{k(\lfloor \frac{m}{2} \rfloor)+1}$, that is

$$RM_{best}(\vec{x}) := (\underbrace{x_{k(\lfloor \frac{m}{2} \rfloor)}, \dots, x_{k(\lfloor \frac{m}{2} \rfloor)}}_{\lfloor \frac{m}{2} \rfloor}, \underbrace{x_{k(\lfloor \frac{m}{2} \rfloor)+1}, \dots, x_{k(\lfloor \frac{m}{2} \rfloor)+1}}_{\lceil \frac{m}{2} \rceil}). \quad (9)$$

  As per the median mechanism, the agents–to-facility assigns every agent in I_j to y_j. In particular, every agent on the left of $x_{k(\lfloor \frac{m}{2} \rfloor)}$ is assigned to a facility located at $x_{k(\lfloor \frac{m}{2} \rfloor)}$, while the others are assigned to a facility located at $x_{k(\lfloor \frac{m}{2} \rfloor)+1}$.

Lastly, the optimal solution to every instance of m-CFLP with no spare capacity is computed as described at the beginning of Section 3.

Table 5
Uniform Distribution - Bayesian approximation ratio.

Bayesian approximation ratio									
		Social Cost				Maximum Cost			
		PMM	$PIPM$	Med	$RankMech$	PMM	$PIPM$	Med	$RankMech$
$k = 5$	$m = 3$	**1.38**	1.79	3.11	3.06	**1.63**	1.87	2.96	3.18
	$m = 4$	**1.49**	1.84	4.17	3.75	**1.65**	1.80	3.65	3.45
	$m = 5$	**1.56**	1.82	5.20	5.00	**1.66**	1.79	4.27	4.40
	$m = 6$	**1.59**	1.80	6.25	5.83	**1.67**	1.79	4.90	4.73
	$m = 7$	**1.62**	1.80	7.29	7.02	**1.68**	1.78	5.49	5.57
$k = 10$	$m = 3$	**1.24**	1.84	3.00	3.13	**1.47**	1.86	2.87	3.29
	$m = 4$	**1.32**	1.87	4.00	3.80	**1.51**	1.78	3.60	3.51
	$m = 5$	**1.38**	1.80	5.00	5.00	**1.53**	1.78	4.30	4.60
	$m = 6$	**1.41**	1.75	6.00	5.80	**1,55**	1.78	4.98	4.89
	$m = 7$	**1.43**	1.73	7.00	6.94	**1.58**	1.76	5.64	5.88
$k = 15$	$m = 3$	**1.17**	1.90	3.01	3.21	**1.37**	1.87	2.86	3.37
	$m = 4$	**1.24**	1.91	4.02	3.88	**1.42**	1.80	3.62	3.56
	$m = 5$	**1.28**	1.83	5.02	5.09	**1.45**	1.80	4.35	4.75
	$m = 6$	**1.32**	1.77	6.03	5.89	**1.49**	1.80	5.07	5.01
	$m = 7$	**1.35**	1.75	7.03	7.04	**1.52**	1.77	5.77	6.11
Average Case approximation ratio									
		Social Cost				Maximum Cost			
		PMM	$PIPM$	Med	$RankMech$	PMM	$PIPM$	Med	$RankMech$
$k = 5$	$m = 3$	**1.39**	1.80	3.19	3.11	**1.64**	1.87	3.05	3.29
	$m = 4$	**1.50**	1.85	4.25	3.80	**1.65**	1.81	3.77	3.55
	$m = 5$	**1.56**	1.82	5.28	5.06	**1.66**	1.80	4.42	4.55
	$m = 6$	**1.60**	1.80	6.33	5.90	**1.67**	1.80	5.07	4.88
	$m = 7$	**1.62**	1.80	7.37	7.09	**1.69**	1.79	5.68	5.76
$k = 10$	$m = 3$	**1.25**	1.86	3.03	3.16	**1.47**	1.87	2.91	3.34
	$m = 4$	**1.33**	1.88	4.03	3.83	**1.51**	1.79	3.65	3.55
	$m = 5$	**1.38**	1.80	5.03	5.03	**1.52**	1.79	4.37	4.55
	$m = 6$	**1.42**	1.75	6.04	5.83	**1.56**	1.79	5.05	4.97
	$m = 7$	**1.44**	1.74	7.03	6.97	**1.58**	1.77	5.73	7.06
$k = 15$	$m = 3$	**1.18**	1.91	3.03	3.23	**1.37**	1.88	2.88	3.41
	$m = 4$	**1.25**	1.92	4.04	3.90	**1.42**	1.81	3.65	3.59
	$m = 5$	**1.29**	1.83	5.04	5.11	**1.45**	1.80	4.40	4.81
	$m = 6$	**1.32**	1.77	6.05	5.91	**1.49**	1.81	5.12	5.07
	$m = 7$	**1.35**	1.75	7.05	7.06	**1.52**	1.78	5.84	6.18

5.2. Experimental results

From our results, we observe that the PMM and the PIPM consistently outperform the other known mechanisms by a significant margin (≥ 0.5), regardless of whether we consider the Social or the Maximum Cost and regardless of whether we consider the Bayesian approximation ratio or the Average Case approximation ratio. The same is not true for both the Median Mechanism and the best Ranking Mechanism.

In Table 3 and 4, we report our results for the Normal Distribution with respect to the Bayesian approximation ratio and Average Case approximation ratio, respectively. If we consider the Social Cost, we observe that the PMM attains the lowest Bayesian approximation ratio, ranging from a maximum of 1.39 to a minimum of 1.12, and the lowest Average Case approximation ratio, ranging from a maximum of 1.40 to a minimum of 1.13, in most of the settings considered. Notably, both the Bayesian and Average Case approximation ratios of the PMM and PIMP hint that the cost entailed by the PMM and PIPM are much lower than what the worst-case approximation ratio suggests. We believe that the reason why the PMM outperforms the PIPM lies in the fact that the PIPM, unlike the PMM, starts its routine by placing two facilities side by side, which is rarely optimal [44].

If we consider the Maximum Cost, we still have that the PMM outperforms the other mechanisms for most of the considered settings. However, in some instances, the PIPM attains a lower value with respect to both metrics.

Our experiments also suggest that the Bayesian and Average Case approximation ratios of both the PMM and PIPM remain stable as the number of facilities increases. In particular, both the Bayesian and Average Case approximation ratios decrease as the number of facilities m increases, while the same is not true for the Median and the best Ranking Mechanism. To give a graphical representation of this, we plot the Bayesian and Average Case approximation ratio for the case $k = 10$ in Fig. 1. We observe that both the PMM and PIPM attain a low Bayesian and Average Case approximation ratio that stays constant as the number of facilities increases. On the contrary, the Bayesian and Average Case approximation ratio attained by the Median mechanism and the best Ranking Mechanism appears to grow linearly as the number of facilities increases.

We have repeated the same experiment with the Exponential (see Table 5) and Uniform distribution (see Table 6), however, we observe no substantial changes in the behavior of the mechanisms. In particular, the PMM and the PIPM outperform the other baseline

Table 6
Results for the Exponential Distribution.

Bayesian approximation ratio

		Social Cost				Maximum Cost			
		PMM	$PIPM$	Med	$RankMech$	PMM	$PIPM$	Med	$RankMech$
$k=5$	$m=3$	**1.28**	1.70	2.35	2.34	**1.85**	2.00	2.48	2.70
	$m=4$	**1.36**	1.47	2.93	2.72	1.85	**1.72**	2.70	2.70
	$m=5$	**1.34**	1.51	3.49	3.41	**1.34**	1.72	2.96	3.10
	$m=6$	**1.37**	1.47	4.02	3.83	**1.76**	1.82	3.10	3.10
	$m=7$	**1.38**	1.49	4.55	4.44	**1.80**	1.82	3.28	3.87
$k=10$	$m=3$	**1.21**	1.78	2.22	2.28	**1.84**	2.00	2.30	2.51
	$m=4$	**1.26**	1.42	2.76	2.66	1.84	**1.65**	2.51	2.51
	$m=5$	**1.22**	1.44	3.27	3.28	1.73	**1.65**	2.67	2.80
	$m=6$	**1.24**	1.41	3.77	3.68	**1.73**	1.83	2.80	2.80
	$m=7$	**1.26**	1.43	4.26	4.24	**1.78**	1.83	2.91	3.01
$k=15$	$m=3$	**1.16**	1.83	2.20	2.30	**1.81**	2.00	2.27	2.44
	$m=4$	**1.20**	1.43	2.74	2.67	1.81	**1.65**	2.44	2.44
	$m=5$	**1.17**	1.44	3.25	3.29	1.75	**1.65**	2.58	2.69
	$m=6$	**1.19**	1.41	3.75	3.69	**1.75**	1.85	2.69	2.69
	$m=7$	**1.20**	1.42	4.23	4.24	**1.78**	1.85	2.79	2.87

Average Case approximation ratio

		Social Cost				Maximum Cost			
		PMM	$PIPM$	Med	$RankMech$	PMM	$PIPM$	Med	$RankMech$
$k=5$	$m=3$	**1.31**	1.75	2.49	2.48	**1.81**	2.00	2.63	2.92
	$m=4$	**1.39**	1.51	3.12	2.87	1.81	**1.68**	2.92	2.92
	$m=5$	**1.37**	1.55	3.71	3.62	1.73	**1.68**	3.27	3.46
	$m=6$	**1.40**	1.50	4.27	4.05	**1.73**	1.78	3.46	3.46
	$m=7$	**1.40**	1.52	4.82	4.71	**1.76**	1.78	3.70	3.84
$k=10$	$m=3$	**1.23**	1.81	2.28	2.35	**1.81**	2.00	2.36	2.61
	$m=4$	**1.27**	1.44	3.12	2.73	1.81	**1.59**	2.61	2.61
	$m=5$	**1.23**	1.46	3.71	3.38	1.68	**1.59**	2.80	2.96
	$m=6$	**1.25**	1.43	4.27	3.79	**1.68**	1.80	2.96	2.96
	$m=7$	**1.27**	1.44	4.82	4.37	**1.75**	1.80	3.09	3.21
$k=15$	$m=3$	**1.16**	1.86	2.24	2.34	**1.78**	2.00	2.31	2.50
	$m=4$	**1.20**	1.44	2.80	2.73	1.78	**1.60**	2.50	2.50
	$m=5$	**1.18**	1.45	3.32	3.35	1.72	**1.60**	2.67	2.80
	$m=6$	**1.20**	1.42	3.83	3.76	**1.72**	1.82	2.79	2.79
	$m=7$	**1.21**	1.43	4.32	4.32	**1.75**	1.82	2.91	3.00

mechanisms according to both metrics, the PMM performs the best when we consider the Social Cost, and the same remains true for the Maximum Cost for most of the considered instances.

In Table 7, we report our results for the case in which agents are not identically distributed. Again, we observe no major nor significant differences between this case and the case in which agents are identically distributed, which confirms the validity of both the PMM and PIPM.

6. Conclusion and future works

In this paper, we investigated two frameworks for the m-CFLP from a Mechanism Design perspective. First, we considered the m-CFLP with equi-capacitated facilities and no spare capacity. We proposed two truthful mechanisms: the Propagating Median Mechanism (PMM) and the Propagating InnerPoint Mechanism (PIPM). Both the mechanisms have bounded approximation ratios with respect to the Social and Maximum Costs. We then established lower bounds on the approximation ratio of any truthful and deterministic mechanism for the m-CFLP with equi-capacitated facilities and no spare capacity. Notably, both PMM and PIPM achieved optimal approximation ratios for the Maximum Cost. Additionally, we demonstrated that PMM and PIPM achieve the minimum possible approximation ratio for the Social Cost among truthful, deterministic, and anonymous mechanisms. In the second framework, we considered the case in which we have two facilities to place and both facilities can accommodate at least half of the agents. We proposed the Extended InnerGap mechanism, which is strong Group Strategyproof, achieves a finite approximation ratio, is optimal with respect to the MC and almost optimal with respect to the SC.

In future research avenues, we aim to improve the lower bounds for non-anonymous mechanisms concerning the Social Cost, to explore higher-dimensional scenarios for agent placements, and to adapt existing randomized mechanisms to enhance approximation ratios results for this problem class [2].

Table 7
Results for Non-Identically distributed agents.

Bayesian approximation ratio									
		Social Cost				Maximum Cost			
		PMM	$PIPM$	Med	$RankMech$	PMM	$PIPM$	Med	$RankMech$
$k=5$	$m=3$	**1.27**	1.49	2.51	2.46	**1.77**	1.82	2.69	2.78
	$m=4$	**1.31**	1.44	3.20	2.94	1.76	**1.63**	3.09	2.99
	$m=5$	**1.33**	1.47	3.85	3.73	1.73	**1.72**	3.38	3.40
	$m=6$	**1.37**	1.49	4.50	4.26	**1.75**	1.81	3.64	3.57
	$m=7$	**1.39**	1.49	5.12	4.97	**1.77**	1.78	3.85	3.85
$k=10$	$m=3$	**1.17**	1.49	2.38	2.46	**1.71**	1.79	2.53	2.67
	$m=4$	**1.21**	1.38	3.01	2.89	1.74	**1.47**	2.85	2.81
	$m=5$	**1.21**	1.43	3.63	3.62	**1.69**	**1.69**	3.09	3.15
	$m=6$	**1.24**	1.46	4.22	4.11	**1.70**	1.83	3.29	3.26
	$m=7$	**1.26**	1.42	4.82	4.79	**1.74**	1.76	3.46	3.49
$k=15$	$m=3$	**1.12**	1.50	2.37	2.50	**1.68**	1.79	2.47	2.62
	$m=4$	**1.15**	1.37	3.00	2.92	1.70	**1.43**	2.76	2.73
	$m=5$	**1.16**	1.43	3.61	3.65	**1.69**	1.71	2.98	3.04
	$m=6$	**1.19**	1.46	4.21	4.14	**1.71**	1.86	3.16	3.15
	$m=7$	**1.21**	1.42	4.80	4.80	**1.74**	1.75	3.31	3.35
Average Case approximation ratio									
		Social Cost				Maximum Cost			
		PMM	$PIPM$	Med	$RankMech$	PMM	$PIPM$	Med	$RankMech$
$k=5$	$m=3$	**1.28**	1.52	2.59	2.53	**1.76**	1.83	2.78	2.90
	$m=4$	**1.33**	1.46	3.29	3.01	1.75	**1.63**	3.23	3.11
	$m=5$	**1.35**	1.49	3.95	3.82	**1.71**	1.72	3.57	3.60
	$m=6$	**1.38**	1.50	4.60	4.35	**1.74**	1.80	3.86	3.79
	$m=7$	**1.40**	1.50	5.24	5.08	**1.76**	1.77	4.11	4.11
$k=10$	$m=3$	**1.17**	1.50	2.41	2.50	**1.69**	1.79	2.57	2.74
	$m=4$	**1.22**	1.39	3.05	2.98	1.73	**1.46**	2.92	2.87
	$m=5$	**1.22**	1.44	3.67	3.67	**1.67**	1.68	3.18	3.25
	$m=6$	**1.24**	1.46	4.27	4.16	**1.68**	1.83	3.40	3.37
	$m=7$	**1.27**	1.43	4.87	4.87	**1.73**	1.74	3.58	3.62
$k=15$	$m=3$	**1.13**	1.52	2.39	2.52	**1.67**	1.80	2.49	2.68
	$m=4$	**1.16**	1.38	3.03	2.95	1.69	**1.41**	2.80	2.78
	$m=5$	**1.17**	1.44	3.64	3.68	**1.68**	1.70	3.04	3.12
	$m=6$	**1.19**	1.47	4.25	4.17	**1.70**	1.85	3.24	3.22
	$m=7$	**1.21**	1.42	4.83	4.84	**1.72**	1.75	3.40	3.44

CRediT authorship contribution statement

Gennaro Auricchio: Writing – review & editing, Writing – original draft, Methodology, Conceptualization. **Zihe Wang:** Writing – review & editing, Writing – original draft, Methodology, Conceptualization. **Jie Zhang:** Writing – review & editing, Writing – original draft, Methodology, Conceptualization.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Zihe Wang reports financial support was provided by National Natural Science Foundation of China (Grant No. 62172422). Jie Zhang reports financial support was provided by Leverhulme Trust. Jie Zhang reports financial support was provided by Engineering and Physical Sciences Research Council (EP/W014912/1). If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

Zihe Wang was partially supported by the National Natural Science Foundation of China (Grant No. 62172422). Jie Zhang was partially supported by a Leverhulme Trust Research Project Grant (2021 – 2024) and the EPSRC grant (EP/W014912/1).

Data availability

No data was used for the research described in the article.

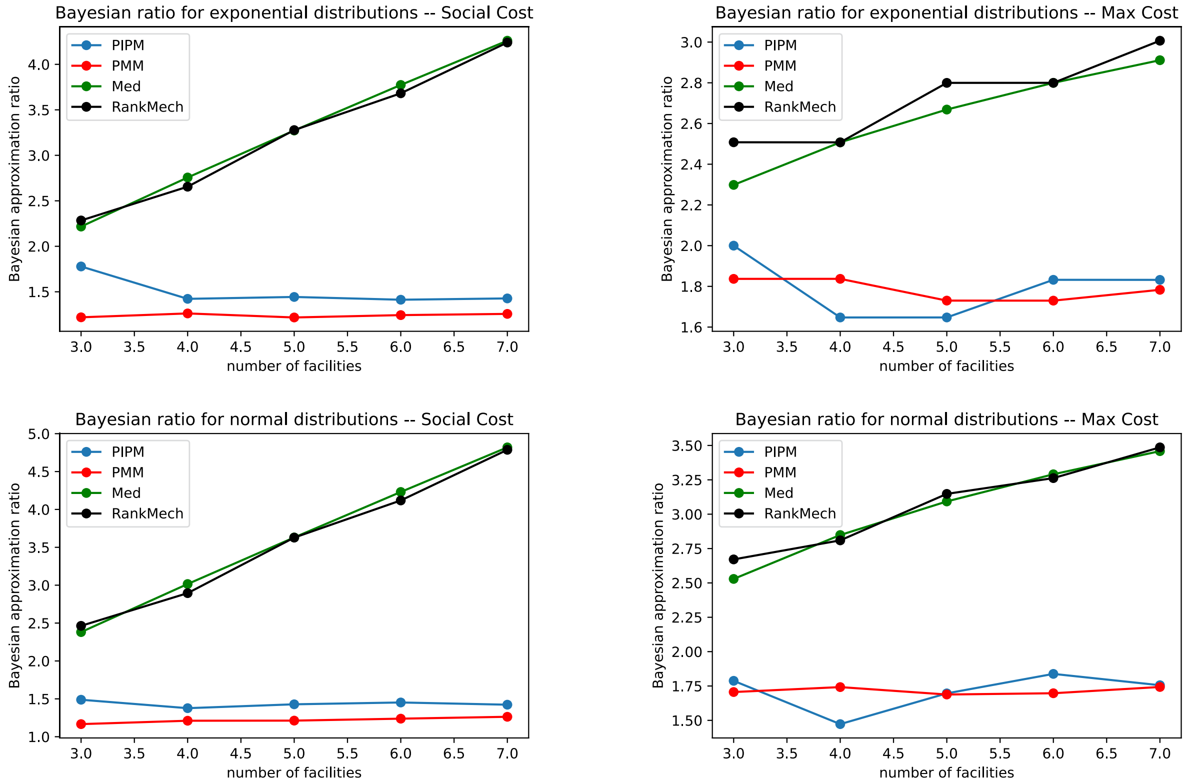

Fig. 1. Bayesian approximation of the Propagating Median Mechanism (PMM), Propagating InnerPoint Mechanism (PIPM), Median mechanism (Med), and the best Ranking Mechanism (RankMech) when agents are distributed according to a different distribution. In the first column, we report the results for the Social Cost, while in the second column, we report the results for the Maximum Cost. In the first row, we report the results for the exponential distribution, while in the second row, we report the results for the normal distribution.

References

[1] G. Auricchio, Z. Wang, J. Zhang, Facility location problems with capacity constraints: two facilities and beyond, in: K. Larson (Ed.), Proceedings of the Thirty-Third International Joint Conference on Artificial Intelligence, IJCAI-24, International Joint Conferences on Artificial Intelligence Organization, 2024, pp. 2651–2659, Main Track.
[2] A.D. Procaccia, M. Tennenholtz, Approximate mechanism design without money, ACM Trans. Econ. Comput. 1 (2013) 1–26.
[3] H. Aziz, H. Chan, B. Lee, B. Li, T. Walsh, Facility location problem with capacity constraints: algorithmic and mechanism design perspectives, in: AAAI, vol. 34, 2020, pp. 1806–1813.
[4] X. Sui, C. Boutilier, T. Sandholm, Analysis and optimization of multi-dimensional percentile mechanisms, in: Proceedings of the Twenty-Third International Joint Conference on Artificial Intelligence, 2013, pp. 367–374.
[5] R. Meir, Voting and mechanism design, in: Strategic Voting, Springer, 2018, pp. 47–66.
[6] T. Börgers, D. Smith, Robust mechanism design and dominant strategy voting rules, Theor. Econ. 9 (2014) 339–360.
[7] R.B. Myerson, Optimal auction design, Math. Oper. Res. 6 (1981) 58–73.
[8] D.P. Bertsekas, Auction algorithms, Encycl. Optim. 1 (2009) 73–77.
[9] Z.-H. Zhou, Ensemble Methods: Foundations and Algorithms, CRC Press, 2012.
[10] Z. Guo, T.J. Norman, E.H. Gerding, Maddm: multi-advisor decision making based on maximizing the dynamic utility, in: Proceedings of the 22nd International Conference on Autonomous Agents and Multiagent Systems, 2023, pp. 2013–2021.
[11] L.I. Kuncheva, Combining Pattern Classifiers: Methods and Algorithms, John Wiley & Sons, 2014.
[12] D. Fotakis, C. Tzamos, On the power of deterministic mechanisms for facility location games, ACM Trans. Econ. Comput. 2 (2014) 1–37.
[13] N. Nisan, A. Ronen, Algorithmic mechanism design, in: STOC, 1999, pp. 129–140.
[14] M. Pal, T. Tardos, T. Wexler, Facility location with nonuniform hard capacities, in: FOCS, IEEE, 2001, pp. 329–338.
[15] Y. Pochet, L.A. Wolsey, Lot-size models with backlogging: strong reformulations and cutting planes, Math. Program. 40 (1988) 317–335.
[16] T. Boffey, Location problems arising in computer networks, J. Oper. Res. Soc. 40 (1989) 347–354.
[17] P. Chardaire, Hierarchical two level location problems, in: Telecommunications Network Planning, 1999, pp. 33–54.
[18] T. Walsh, Strategy proof mechanisms for facility location in Euclidean and Manhattan space, arXiv preprint, arXiv:2009.07983, 2020.
[19] T. Walsh, Strategy proof mechanisms for facility location with capacity limits, in: L.D. Raedt (Ed.), IJCAI-22, 2022, pp. 527–533.
[20] B. Balcik, B.M. Beamon, Facility location in humanitarian relief, Int. J. Logist. 11 (2008) 101–121.
[21] M. Melo, S. Nickel, F.S. da Gama, Facility location and supply chain management – a review, Eur. J. Oper. Res. 196 (2009) 401–412.
[22] A. Ahmadi-Javid, P. Seyedi, S.S. Syam, A survey of healthcare facility location, Comput. Oper. Res. 79 (2017) 223–263.
[23] T. Hastie, R. Tibshirani, J.H. Friedman, J.H. Friedman, The Elements of Statistical Learning: Data Mining, Inference, and Prediction, vol. 2, Springer, 2009.
[24] O.H. Barda, J. Dupuis, P. Lencioni, Multicriteria location of thermal power plants, Eur. J. Oper. Res. 45 (1990) 332–346.
[25] M. Feldman, Y. Wilf, Strategyproof facility location and the least squares objective, in: EC, ACM, 2013, pp. 873–890.
[26] A. Filimonov, R. Meir, Strategyproof facility location mechanisms on discrete trees, in: AAMAS, ACM, 2021, pp. 510–518.

[27] P. Lu, X. Sun, Y. Wang, Z.A. Zhu, Asymptotically optimal strategy-proof mechanisms for two-facility games, in: EC, ACM, 2010, pp. 315–324.
[28] P. Lu, Y. Wang, Y. Zhou, Tighter Bounds for Facility Games, WINE, vol. 5929, Springer, 2009, pp. 137–148.
[29] N. Alon, M. Feldman, A.D. Procaccia, M. Tennenholtz, Strategyproof approximation of the minimax on networks, Math. Oper. Res. 35 (2010) 513–526.
[30] E. Dokow, M. Feldman, R. Meir, I. Nehama, Mechanism design on discrete lines and cycles, in: EC, ACM, 2012, pp. 423–440.
[31] R. Meir, Strategyproof facility location for three agents on a circle, in: SAGT, in: Lecture Notes in Computer Science, vol. 11801, Springer, 2019, pp. 18–33.
[32] P. Tang, D. Yu, S. Zhao, Characterization of group-strategyproof mechanisms for facility location in strictly convex space, in: EC, ACM, 2020, pp. 133–157.
[33] A. Filos-Ratsikas, M. Li, J. Zhang, Q. Zhang, Facility location with double-peaked preferences, Auton. Agents Multi-Agent Syst. 31 (2017) 1209–1235.
[34] P. Serafino, C. Ventre, Heterogeneous facility location without money, Theor. Comput. Sci. 636 (2016) 27–46.
[35] D. Ye, L. Mei, Y. Zhang, Strategy-proof mechanism for obnoxious facility location on a line, in: International Computing and Combinatorics Conference, Springer, 2015, pp. 45–56.
[36] L. Mei, D. Ye, G. Zhang, Mechanism design for one-facility location game with obnoxious effects, in: International Workshop on Frontiers in Algorithmics, Springer, 2016, pp. 172–182.
[37] J. Brimberg, E. Korach, M. Eben-Chaim, A. Mehrez, The capacitated p-facility location problem on the real line, Int. Trans. Oper. Res. 8 (2001) 727–738.
[38] K. Aardal, P.L. van den Berg, D. Gijswijt, S. Li, Approximation algorithms for hard capacitated k-facility location problems, Eur. J. Oper. Res. 242 (2015) 358–368.
[39] H. Aziz, H. Chan, B.E. Lee, D.C. Parkes, The capacity constrained facility location problem, Games Econ. Behav. 124 (2020) 478–490.
[40] G. Auricchio, H.J. Clough, J. Zhang, On the capacitated facility location problem with scarce resources, in: The 40th Conference on Uncertainty in Artificial Intelligence, Barcellona, 2024.
[41] G. Auricchio, H.J. Clough, J. Zhang, Mechanism design for locating facilities with capacities with insufficient resources, arXiv preprint, arXiv:2407.18547, 2024.
[42] J.D. Hartline, T. Roughgarden, Simple versus optimal mechanisms, in: Proceedings of the 10th ACM Conference on Electronic Commerce, 2009, pp. 225–234.
[43] J.D. Hartline, B. Lucier, Bayesian algorithmic mechanism design, in: Proceedings of the Forty-Second ACM Symposium on Theory of Computing, 2010, pp. 301–310.
[44] G. Auricchio, J. Zhang, The k-facility location problem via optimal transport: a Bayesian study of the percentile mechanisms, in: International Symposium on Algorithmic Game Theory, Springer, 2024, pp. 147–164.
[45] G. Auricchio, J. Zhang, M. Zhang, Extended ranking mechanisms for the m-capacitated facility location problem in Bayesian mechanism design, in: Proceedings of the 23rd International Conference on Autonomous Agents and Multiagent Systems, 2024, pp. 87–95.
[46] H. Chan, A. Filos-Ratsikas, B. Li, M. Li, C. Wang, Mechanism design for facility location problems: a survey, in: 30th International Joint Conference on Artificial Intelligence, International Joint Conferences on Artificial Intelligence Organization, 2021, pp. 4356–4365.
[47] W. Cook, A.M. Gerards, A. Schrijver, É. Tardos, Sensitivity theorems in integer linear programming, Math. Program. 34 (1986) 251–264.
[48] A.D. Procaccia, J.S. Rosenschein, Average-case tractability of manipulation in voting via the fraction of manipulators, in: Proceedings of the 6th International Joint Conference on Autonomous Agents and Multiagent Systems, 2007.
[49] X. Deng, Y. Gao, J. Zhang, Beyond the worst-case analysis of random priority: smoothed and average-case approximation ratios in mechanism design, Inf. Comput. 285 (2022) 104920, https://doi.org/10.1016/J.IC.2022.104920.