

ON DISCRETE TRUTHFUL HETEROGENEOUS TWO-FACILITY LOCATION*

PANAGIOTIS KANELLOPOULOS[†], ALEXANDROS A. VOUDOURIS[†], AND RONGSEN ZHANG[†]

Abstract. We revisit the discrete heterogeneous two-facility location problem, in which there is a set of agents that occupy nodes of a line graph, and have private approval preferences over two facilities. When the facilities are located at some nodes of the line, each agent suffers a cost that is equal to her total distance from the facilities she approves. The goal is to decide where to locate the two facilities, so as to (a) incentivize the agents to truthfully report their preferences, and (b) achieve a good approximation of the minimum total (social) cost or the maximum cost among all agents. For both objectives, we design deterministic strategyproof mechanisms with approximation ratios that significantly outperform the state-of-the-art, and complement these results with (almost) tight lower bounds.

Key words. Facility location, Mechanism design, Approximation

MSC codes. 68Q25, 68W25, 91A10

1. Introduction. In the classic truthful single-facility location problem, there is a set of agents with private positions on the line of real numbers, and a public facility (such as a library or a park) whose location we need to decide. This decision must be made so as to (a) incentivize the agents to truthfully reveal their positions (an agent would be willing to lie if this leads to the facility being located closer to her true position), and (b) optimize a social objective (such as the total distance of the agents from the facility location, or the maximum distance). Since the celebrated paper of Procaccia and Tennenholtz [2013], this problem and its many variants have been studied extensively in the literature on *approximate mechanism design without money*. For a comprehensive introduction to the various different facility location models that have been considered over the years, we refer the interested reader to the recent survey of Chan et al. [2021].

A recent stream of papers have focused on *heterogeneous facility location* problems, with multiple facilities (typically, two) that are different in nature (e.g., a school and a bar). As such, the agents care both for the location and the types of the facilities, aiming for the facilities they like the most to be as close to their position as possible. To give an example, a family would like to be closer to a school than to a bar, a single person might want to be closer to the bar, and a young couple might want to be close to both facilities. Many settings have been proposed to model the different preferences the agents may have about the facilities (see the discussion in the related work). With few exceptions, all of these models assume that both facilities can be placed at any point of the real line, even at the same one. Serafino and Ventre [2016] deviated from these assumptions and studied a discrete version of the problem. In this model, the line is a *discrete* graph, and each node of this graph is either occupied by a single agent, or is empty. This information is assumed to be common knowledge. There are also two facilities (such as the school and the bar in our example above) which can only be placed at *different* nodes of the line. Each agent has a private

*A preliminary version of this paper appears in *Proceedings of the 31st International Joint Conference on Artificial Intelligence (IJCAI 2022)*.

[†]School of Computer Science and Electronic Engineering, University of Essex, United Kingdom (panagiotis.kanellopoulos@essex.ac.uk, alexandros.voudouris@essex.ac.uk, rz19109@essex.ac.uk).

approval preference over the facilities, that is, she either wants to be close to a facility or is indifferent about its location. Given the locations of the facilities, the cost of an agent is then defined as the total distance from the facilities she approves.

Serafino and Ventre presented bounds on the approximation ratio of deterministic and randomized strategyproof mechanisms in terms of the *social cost* (i.e., the total cost of agents) and the *maximum cost*. In particular, for the social cost, they showed that the best possible approximation ratio of deterministic mechanisms is between $9/8$ and $n - 1$, where n is the number of agents. In contrast, they designed a randomized mechanism that always outputs a solution with minimum expected social cost. For the maximum cost objective, they showed that the best possible approximation ratio of deterministic mechanisms is between $3/2$ and 3 , and that of randomized mechanisms is between $4/3$ and $3/2$. In this paper, we focus exclusively on deterministic strategyproof mechanisms when agents are interested in the outcome (that is, approve at least one facility)¹, and improve upon the bounds of Serafino and Ventre for both objectives in this case.

1.1. Our contribution. The main technical difficulty in designing deterministic strategyproof mechanisms with low approximation ratio in terms of the social cost is the constraint of locating the two facilities at different nodes. If each agent approves just a single facility, then locating each facility to the median agent among those that approve it would be a strategyproof mechanism with minimum social cost. However, in general, there might exist agents that approve both facilities, in which case the medians for the two facilities could coincide, and any choice of how to break this tie could lead to some agent having incentive to misreport.

The upper bounds of Serafino and Ventre (for their more general model) follow by the TWOEXTREMES mechanism, which locates one of the facilities at the node occupied by the leftmost agent among those that approve it, and the other facility at the node occupied by the rightmost agent among those that approve it; in case of a collision, one of the facilities is moved a node to the left or the right. While this mechanism successfully places the two facilities without providing incentives to the agents, it can be quite inefficient. There are two main reasons for this deficiency: (i) the boundary agents (leftmost and rightmost) among those that approve a facility may be rather far away from the median such agent, whose node would be the ideal location for the facility, and (ii), it does not exploit the available information about the position of the agents in any way. Our improved mechanisms, for the setting we consider, take care of these two limitations: We place the facilities closer to median agents (without breaking strategyproofness), and exploit the information about the positions of the agents.

For the social cost, we design the FIXED-OR-MEDIAN-NEAREST-EMPTY (FMNE) mechanism with an approximation ratio of at most $17/4 = 4.25$. The mechanism switches between two cases based on the structure of the line: If there are no empty nodes, it fixes the locations of the facilities to be the two central nodes of the line; otherwise, if there are empty nodes, it locates one of the facilities at the position of the median agent among those that approve it, and the other facility at one of the nearest empty nodes to the median agent among those that approve it. We complement this result with an improved lower bound of $4/3$ on the approximation ratio of all strategyproof mechanisms, which follows by two instances with only three agents and no empty nodes. Motivated by this lower bound construction, we then

¹Serafino and Ventre [2016] consider a more general model where agents may not approve any facility.

Table 1: An overview of our bounds on the approximation ratio of deterministic strategyproof mechanisms for the social cost and the maximum cost. The bounds in parentheses are the previously best known ones implied by Serafino and Ventre [2016]. The lower bound of $4/3$ marked with a \star is tight for instances with three agents.

	Lower bound	Upper bound
Social cost	$4/3^\star$ (9/8)	$17/4$ ($n - 1$)
Maximum cost	2 (3/2)	2 (3)

focus on instances with three agents, and design the 3-agent PRIORITY-DICTATORSHIP mechanism that achieves the best-possible bound of $4/3$.

For the maximum cost, we design a parameterized class of mechanisms α -LEFT-RIGHT (α -LR), each member of which partitions the line into two parts, from the first node to node α , and from node $\alpha + 1$ to the last node. Then, the decision about the locations of the facilities is based on the preferences of the agents included in the two parts. We show that all mechanisms of the class are strategyproof, and there are members with approximation ratio at most 2. In particular, when the size m of the line is an even number, the bound is achieved by $m/2$ -LR, and when m is odd, it is achieved by $(m + 1)/2$ -LR. Finally, we show a tight lower bound of 2 on the approximation ratio of all strategyproof mechanisms, using a construction involving a sequence of five instances with three agents and no empty nodes.

An overview of our bounds, and how they compare to the previously best ones shown by Serafino and Ventre [2016], as applied to our model, is given in Table 1.

1.2. Related work. As already mentioned above, the survey of Chan et al. [2021] nicely discusses the many different facility models that have been considered over the years in the literature on approximate mechanism design without money. Here, we will mainly discuss papers on heterogeneous facility location models that are closely related to ours. The vast majority of the related literature has mainly focused on continuous settings, where the facilities can be located at any point of the line, even the same one. Some notable exceptions include the paper of Serafino and Ventre [2016], a few papers on characterizations of onto strategyproof mechanisms for the single-facility location problem in discrete lines, cycles [Dokow et al., 2012] and trees [Filimonov and Meir, 2021], and some papers related to truthful single-winner voting on a line metric (e.g., [Feldman et al., 2016, Filos-Ratsikas and Voudouris, 2021]). Discrete facility location problems have also been studied within the distortion literature on social choice problems, where the main source of inefficiency is due to having limited information; see the recent survey of Anshelevich et al. [2021] for an overview.

The first heterogeneous facility location model, combining elements from the classic single-facility location problem and the obnoxious single-facility location problem [Cheng et al., 2011, 2013], was independently proposed and studied by Feigenbaum and Sethuraman [2015] and Zou and Li [2015]. In this setting, there are two facilities to be located on the real line, and the agents have *dual* preferences over the facilities; that is, an agent likes or dislikes a facility. The authors showed bounds on the approximation ratio of deterministic and randomized strategyproof mechanisms for different cases depending on whether the positions or the preferences of the agents are their private information (and can thus lie about them). Kyropoulou et al. [2019] considered an extension of this model, where the location space of the two facilities is

a constrained region of the Euclidean space.

Chen et al. [2020] studied a setting with agents that have *optional* (or, *approval*) preferences over the facilities; that is, an agent either likes a facility or is indifferent about it. The authors considered two different cost functions of the agents, one that is equal to the distance from the closest facility that the agent approves, and one that is equal to the distance from the farthest such facility. Li et al. [2020] studied an extension of this setting in more general metrics (beyond the line), and designed mechanisms that significantly improved some the linear (in the number of agents) upper bound of Chen et al. to a constant. Deligkas et al. [2022] considered a similar approval preference model, but with the difference that the goal is to locate just one of the two facilities (and, more generally, k out of m). Elkind et al. [2022] also considered the case of locating only a subset of the available facilities, but based on approval preferences that are induced by the possible locations of the facilities, as opposed to being exogenous.

Anastasiadis and Deligkas [2018] considered a model that combines dual and optional preferences, in the sense that the agents can like, dislike or be indifferent about a facility. Fong et al. [2018] studied a setting with *fractional* preferences, where each agent has a weight in $[0, 1]$ for each facility indicating how much she likes it. Finally, Xu et al. [2021] focused on the problem where the locations of the locations of the two facilities must satisfy a minimum distance requirement.

2. Preliminaries. We consider the discrete two-facility location problem. An instance I of this problem consists of a set N of $n \geq 2$ *agents*, two *facilities*, and a line graph with $m \geq n$ nodes. Each agent occupies a node x_i of the line, such that different agents occupy different nodes. By \mathbf{x} we denote the *position profile* consisting of the positions of all agents (i.e., the nodes they occupy) as well as the positions of possible empty nodes; the position profile is assumed to be *common knowledge*; in many cases, we enumerate the nodes from 1 to m , so that we can compare the positions of the agents. Every agent i also has a private *approval preference* $\mathbf{t}_i \in \{0, 1\}^2$ over the two facilities such that $t_{i1} + t_{i2} \geq 1$: If $t_{ij} = 1$, agent $i \in N$ approves facility $j \in \{1, 2\}$; otherwise, she is indifferent to it. By $\mathbf{t} = (\mathbf{t}_i)_{i \in N}$ we denote the *preference profile* consisting of the preferences of all agents. Given \mathbf{t} , it will be useful to denote by N_j the set of agents that approve facility $j \in \{1, 2\}$. Clearly, the two sets N_1 and N_2 need not be disjoint if there are agents that approve both facilities. As \mathbf{x} and \mathbf{t} implicitly include all the information related to an instance, we denote $I = (\mathbf{x}, \mathbf{t})$.

A *feasible solution* $\mathbf{z} = (z_1, z_2)$ determines the node z_j where each facility $j \in \{1, 2\}$ is located, so that $z_1 \neq z_2$. Given a feasible solution \mathbf{z} , the cost of any agent i in instance I is her *total* distance from the facilities she approves, i.e.,

$$\text{cost}_i(\mathbf{z}|I) = \sum_{j \in \{1, 2\}} t_{ij} \cdot d(i, j),$$

where $d(i, j) = |x_i - z_j|$ is the distance between agent i and facility j .

A *mechanism* takes as input an instance and outputs a feasible solution. A mechanism M is said to be *strategyproof* if the solution $M(I)$ it computes when given as input the instance $I = (\mathbf{x}, \mathbf{t})$ is such that no agent i has incentive to report a false preference $\mathbf{t}'_i \neq \mathbf{t}_i$ to decrease her cost, i.e.,

$$\text{cost}_i(M(I)|I) \leq \text{cost}_i(M(\mathbf{x}, (\mathbf{t}'_i, \mathbf{t}_{-i}))|I),$$

where $(\mathbf{t}'_i, \mathbf{t}_{-i})$ is the preference profile according to which agent i 's preference is \mathbf{t}'_i , while the preference of any other agent is the same as in \mathbf{t} .

We consider two well-known social objectives, which are functions of feasible solutions. Given an instance I , the *social cost* of a feasible solution \mathbf{z} is the total cost of all the agents, i.e.,

$$\text{SC}(\mathbf{z}|I) = \sum_{i \in N} \text{cost}_i(\mathbf{z}|I).$$

The *max cost* of \mathbf{z} is the maximum cost among all agents, i.e.,

$$\text{MC}(\mathbf{z}|I) = \max_{i \in N} \text{cost}_i(\mathbf{z}|I).$$

Let $\text{SC}^*(I) = \min_{\mathbf{z}} \text{SC}(\mathbf{z}|I)$ be the minimum possible social cost for instance I , achieved by any feasible solution. Similarly, let $\text{MC}^*(I) = \min_{\mathbf{z}} \text{MC}(\mathbf{z}|I)$ be the minimum possible maximum cost for I .

For any social objective $f \in \{\text{SC}, \text{MC}\}$, the *f-approximation ratio* $\rho_f(M)$ of a mechanism M is the worst-case ratio (over all possible instances) between the objective value of the solution computed by M over the minimum possible objective value among all feasible solutions, i.e.,

$$\rho_f(M) = \sup_I \frac{f(M(I)|I)}{f^*(I)}.$$

Our goal is to design strategyproof mechanisms with an as low f -approximation ratio as possible (close to 1).

3. Social cost: A general constant upper bound. We start with the social cost objective. For general instances with n agents, we design the strategyproof mechanism FIXED-OR-MEDIAN-NEAREST-EMPTY (FMNE) with approximation ratio $17/4$, thus greatly improving upon the previous bound of $n - 1$ that follows from Serafino and Ventrè [2016]. Our mechanism exploits the known information about the position profile, and distinguishes between two cases depending on whether the given instance contains empty nodes or not. If there are no empty nodes, FMNE locates the facilities next to each other at central nodes of the line (in particular, nodes $\lfloor n/2 \rfloor$ and $\lfloor n/2 \rfloor + 1$); this is the FIXED part of the mechanism. If there are empty nodes, FMNE locates facility 1 at the node occupied by the median agent among those that approve facility 1, and facility 2 at the empty node that is nearest to the node occupied by the median agent among those that approve facility 2; this is the MEDIAN-NEAREST-EMPTY part of the mechanism. See Algorithm 3.1 for the formal definition.

THEOREM 3.1. *FMNE is strategyproof.*

Proof. Consider an arbitrary instance I . Recall that the positions of the agents are known. The mechanism is clearly strategyproof if there are no empty nodes in I as the locations of the facilities are fixed and independent of the preferences of the agents. So, it remains to consider the case where I contains empty nodes. Let i be an arbitrary agent. We switch between the following three cases:

Agent i approves only facility 1 ($i \in N_1 \setminus N_2$). Suppose without loss of generality that $x_i \leq y_1$. Any misreport of agent i can only lead to a median y'_1 among the agents that approve facility 1 which is farther away from x_i . In particular, if i misreports that she approves only facility 2, then $y'_1 \geq y_1$, whereas if i misreports that she approves both facilities, then $y'_1 = y_1$.

Algorithm 3.1 FIXED-OR-MEDIAN-NEAREST-EMPTY (FMNE)**Input:** Instance I with n agents**Output:** Feasible solution $\mathbf{z} = (z_1, z_2)$ ▷ **FIXED part****if** there are no empty nodes **then** $z_1 \leftarrow \lfloor n/2 \rfloor$ $z_2 \leftarrow \lfloor n/2 \rfloor + 1$ ▷ **MEDIAN-NEAREST-EMPTY part****else****for** $j \in \{1, 2\}$ **do** $y_j \leftarrow$ position of the leftmost median agent in N_j **end for** $z_1 \leftarrow y_1$ $z_2 \leftarrow$ nearest empty node to y_2 , breaking ties in favor of rightmost such node**end if**

Agent i approves only facility 2 ($i \in N_2 \setminus N_1$). Suppose without loss of generality that facility 2 is positioned at some empty node e with $x_e > y_2$. Denote by y'_2 the median node occupied by the agents that approve facility 2 when i misreports.

- If agent i misreports that she approves both facilities, then $y'_2 = y_2$, and hence the position of facility 2 as well as the cost of agent i remain the same.
- If $x_i \leq y_2$ and agent i misreports that she only approves facility 1, then $y'_2 \geq y_2$. As a result, either e continues to be the nearest empty node to y'_2 and the cost of i remains exactly the same, or another empty node e' with $x_{e'} > x_e$ becomes the nearest empty node to y'_2 , and the cost of i strictly increases.
- If $x_i > y_2$ and agent i misreports that she only approves facility 1, then $y'_2 \leq y_2$. As a result, either e continues to be the nearest empty node to y'_2 , or another empty node e' with $x_{e'} < y_2 < x_e$ becomes the nearest empty node to y'_2 . In any case, the cost of i does not decrease.

Agent i approves both facilities ($i \in N_1 \cap N_2$). Since the cost of agent i is the sum of costs she derives from the two facilities and we decide where to locate each facility independently from the other facility, the same arguments for the previous two cases show that no possible misreport can lead to a strictly lower cost. \square

To argue about the approximation ratio of FMNE, we will distinguish between instances with and without empty nodes. In our proofs, we exploit the following lower bounds on the optimal social cost; we include the proof for completeness. Here, $\mathbf{1}\{X\}$ is equal to 1 if the event X is true, and 0 otherwise.

LEMMA 3.2 ([Serafino and Ventre, 2016]). *For any instance I in which there are n_j agents that approve facility $j \in \{1, 2\}$, it holds that*

$$\begin{aligned} \text{SC}^*(I) &\geq \frac{1}{4} \left(n_1^2 + n_2^2 - \mathbf{1}\{n_1 \text{ odd}\} - \mathbf{1}\{n_2 \text{ odd}\} \right) \\ &\geq \frac{1}{4} \left(n_1^2 + n_2^2 - 2 \right). \end{aligned}$$

Proof. We argue about each facility $j \in \{1, 2\}$ independently. When n_j is odd, the optimal allocation can, at best, have facility j at one of these n_j nodes and have

two agents at distance i , for $i \in \{1, \dots, \frac{n_j-1}{2}\}$ from j ; the total cost due to facility j is then $\frac{n_j^2-1}{4}$. When n_j is even, the optimal allocation can, at best, place facility j at one of the n_j nodes and have two agents at distance i , for $i \in \{1, \dots, \frac{n_j}{2}-1\}$, and an agent at distance $\frac{n_j}{2}$; the total cost due to facility j in this case is then $\frac{n_j^2}{4}$. \square

For instances without empty nodes and $n \geq 5$, we will show that the approximation ratio of FMNE (in particular, its FIXED part) is at most 3; note that for $n \leq 4$, the TWOEXTREMES mechanism of Serafino and Ventre [2016] is 3-approximate.

THEOREM 3.3. *For any instance with $n \geq 5$ agents and no empty nodes, the SC-approximation ratio of FMNE is at most 3.*

Proof. Consider an instance I and recall that N_j denotes the set of agents that approve facility j . Let $n_1 = |N_1|$, $n_2 = |N_2|$, and $b = |N_1 \cap N_2|$; clearly, it holds that $n = n_1 + n_2 - b$.

We first consider the case where n is even, i.e., $n \geq 6$. For any agent i , with $i \in \{1, \dots, n\}$, the maximum distance of i to a facility is $\lfloor n/2 + 1 - i \rfloor$. Furthermore, each of the b agents that approve both facilities faces an added cost of at most $n/2$ due to the distance to the agent's nearest facility. Therefore, the total cost of the solution \mathbf{z} computed by the mechanism is bounded by

$$\begin{aligned} \text{SC}(\mathbf{z}|I) &\leq 2 \sum_{i=1}^{n/2} i + b \cdot \frac{n}{2} \\ &= \frac{n}{2} \cdot \left(\frac{n}{2} + 1 \right) + b \cdot \frac{n}{2} \\ &= \frac{n_1^2 + n_2^2 + b^2 + 2n_1n_2 - 2bn_1 - 2bn_2 + 2n_1 + 2n_2 - 2b}{4} + \frac{bn_1 + bn_2 - b^2}{2} \\ &\leq \frac{n_1^2 + n_2^2 + 2n_1n_2 + 2n_1 + 2n_2}{4}, \end{aligned}$$

where the second equality holds since $n = n_1 + n_2 - b$, while the last inequality holds since $b \geq 0$.

By Lemma 3.2, $\text{SC}^*(I) \geq \frac{1}{4}(n_1^2 + n_2^2 - 2)$, and thus the approximation ratio is bounded by

$$\frac{\text{SC}(\mathbf{z}|I)}{\text{SC}^*(I)} \leq \frac{n_1^2 + n_2^2 + 2n_1n_2 + 2n_1 + 2n_2}{n_1^2 + n_2^2 - 2}.$$

To prove the claim, it suffices to show that, when $n_1 + n_2 \geq 6$, it holds that $n_1^2 + n_2^2 + 2n_1n_2 + 2n_1 + 2n_2 \leq 3n_1^2 + 3n_2^2 - 6$, i.e., $(n_1 - n_2)^2 + n_1^2 + n_2^2 \geq 2n_1 + 2n_2 + 6$. Observe that, when $n_1 + n_2 \geq 6$, it holds that $n_1^2 + n_2^2 \geq 3(n_1 + n_2) \geq 2n_1 + 2n_2 + 6$; the claim follows.

We now consider the case where $n \geq 5$ is odd; the analysis is slightly more involved, but follows along similar lines. Observe that the maximum distance of any agent i positioned at some of the first $(n-1)/2$ nodes from a facility (in particular, facility 2) is $(n+1)/2 - i$, while the maximum distance of any agent i positioned at some of the last $(n+1)/2$ nodes from a facility (in this case, facility 1) is $i - (n-1)/2$. Furthermore, each of the b agents that approve both facilities faces an added cost of at most $(n-1)/2$. So, the total cost of the solution \mathbf{z} computed by the mechanism is

bounded by

$$\begin{aligned}
\text{SC}(\mathbf{z}|I) &\leq \sum_{i=1}^{(n-1)/2} i + \sum_{i=1}^{(n+1)/2} i + b \cdot \frac{n-1}{2} \\
&= \frac{(n+1)^2}{4} + b \cdot \frac{n-1}{2} \\
&= \frac{n_1^2 + n_2^2 + b^2 + 2n_1n_2 - 2bn_1 - 2bn_2 + 2n_1 + 2n_2 - 2b + 1}{4} \\
&\quad + \frac{bn_1 + bn_2 - b^2 - b}{2} \\
&= \frac{n_1^2 + n_2^2 + 2n_1n_2 + 2n_1 + 2n_2 + 1 - b^2 - 4b}{4},
\end{aligned}$$

where, again, the second equality holds since $n = n_1 + n_2 - b$.

By Lemma 3.2, $\text{SC}^*(I) \geq \frac{1}{4}(n_1^2 + n_2^2 - \mathbf{1}\{n_1 \text{ odd}\} - \mathbf{1}\{n_2 \text{ odd}\})$, and thus the approximation ratio is bounded by

$$\frac{\text{SC}(\mathbf{z}|I)}{\text{SC}^*(I)} \leq \frac{n_1^2 + n_2^2 + 2n_1n_2 + 2n_1 + 2n_2 + 1 - b^2 - 4b}{n_1^2 + n_2^2 - \mathbf{1}\{n_1 \text{ odd}\} - \mathbf{1}\{n_2 \text{ odd}\}}.$$

To prove the claim it suffices to show that $(n_1 - n_2)^2 + n_1^2 + n_2^2 + b^2 + 4b \geq 2n_1 + 2n_2 + 1 + 3(\mathbf{1}\{n_1 \text{ odd}\} + \mathbf{1}\{n_2 \text{ odd}\})$. If $b \geq 1$, then $n_1 + n_2 \geq 6$, and the claim follows since $(n_1 - n_2)^2 + n_1^2 + n_2^2 \geq 2(n_1 + n_2 + 1)$ holds in this case. Otherwise, when $b = 0$, then exactly one of n_1, n_2 is odd and it suffices to show that $(n_1 - n_2)^2 + n_1^2 + n_2^2 \geq 2n_1 + 2n_2 + 4$. Since $(n_1 - n_2)^2 \geq 1$, this always holds if $n_1 + n_2 \geq 5$. \square

For instances with at least one empty node, we will show that the approximation ratio of FMNE (in particular, its MEDIAN-NEAREST-EMPTY part) is $17/4$ for any $n \geq 6$; observe that the TWOEXTREMES mechanism of Serafino and Ventre [2016] achieves an approximation ratio of at most 4 when $n \leq 5$. Our proof for the approximation ratio of FMNE in this case relies on the following technical lemma.

LEMMA 3.4. *Let $f(x, y) = \frac{y^2 + 4xy + 2y + 1}{x^2 + y^2 - 2}$. For non-negative integers x, y such that $x + y \geq 6$, it holds $f(x, y) \leq 13/4$.*

Proof. First, observe that $f(x, y)$ can be written as $f(x, y) = 1 + \frac{-x^2 + 4xy + 2y + 3}{x^2 + y^2 - 2}$. It suffices to limit our attention to the values of x, y for which $f(x, y) > 3$, i.e., to these x, y such that $\frac{-x^2 + 4xy + 2y + 3}{x^2 + y^2 - 2} > 2$. By rearranging, we obtain $-x^2 + 4xy + 2y + 3 > 2x^2 + 2y^2 - 4$, and, therefore, $-x^2 + 2y + 7 > 2(x - y)^2$.

Let $y = x + k$ for some integer k and rewrite the last inequality as $-x^2 + 2x + 7 > 2(k^2 - k)$. Clearly, for $x \geq 4$ the inequality never holds as the left-hand-side term is negative and the right-hand-side term is always non-negative. Hence, we obtain that $f(x \geq 4, y) < 13/4$. For the remaining values of x , i.e., when $x \in \{0, 1, 2, 3\}$, recall that $x + y \geq 6$, i.e., $2x + k \geq 6$. When $x = 0$, it must be $k \geq 6$ and the inequality does not hold, as $7 < 60$. When $x = 1$, we have $k \geq 4$ and, again, the inequality does not hold, as $8 < 24$. For $x = 2$, we obtain $k \geq 2$ and the inequality becomes $7 > 2(k^2 - k)$, which holds only when $k = 2$; in this case, $f(2, 4) = 19/6 < 13/4$. Finally, for $x = 3$, we have that $k \geq 0$ and the inequality becomes $4 > 2(k^2 - k)$ which holds for $k \in \{0, 1\}$. The proof follows by observing that $\max\{f(3, 3), f(3, 4)\} = \max\{13/4, 73/23\} \leq 13/4$. \square

We are now ready to prove the bound for instances with empty nodes.

THEOREM 3.5. *For any instance with $n \geq 6$ and at least one empty node, the SC-approximation ratio of FMNE is at most $17/4$.*

Proof. Consider any instance I . We first argue a bit about the optimal social cost of I . A solution that minimizes the social cost locates each facility $j \in \{1, 2\}$ to the node y_j occupied by a median agent in N_j . However, this solution might not be feasible if $y_1 = y_2$, and so the optimal social cost can only be larger. We have that

$$(3.1) \quad \text{SC}^*(I) \geq \sum_{i \in N_1} d(x_i, y_1) + \sum_{i \in N_2} d(x_i, y_2).$$

Now, let us focus on the social cost of the solution \mathbf{z} computed by the mechanism. Let e be the empty node where facility 2 is located; without loss of generality, we can assume that $x_e > y_2$. Combined with the fact that facility 1 is located at y_1 , we have that $\mathbf{z} = (y_1, x_e)$, and

$$\text{SC}(\mathbf{z}|I) = \sum_{i \in N_1} d(x_i, y_1) + \sum_{i \in N_2} d(x_i, x_e).$$

The first term appears in the lower bound of the optimal social cost given by Inequality (3.1), so all we need to do is bound the second term of the above expression.

We partition the set N_2 into three sets L , M and R depending on the positions of the agents in N_2 compared to y_2 and x_e , as follows:

- $L = \{i \in N_2 : x_i \leq y_2\}$;
- $M = \{i \in N_2 : x_i \in (y_2, x_e)\}$;
- $R = \{i \in N_2 : x_i > x_e\}$.

By the definition of median, we have that $|L| \geq |M| + |R|$; in particular, this is an equality if $n_2 = |N_2|$ is even, and a strict inequality if n_2 is odd (as L also includes the median agent in this case). Due to this, we can create disjoint pairs of agents as follows:

- We match every agent $i \in M$ to a unique agent $j \in L$ such that

$$d(x_j, x_e) = d(x_j, x_i) + d(x_i, x_e) = d(x_j, y_2) + d(x_i, y_2) + d(x_i, x_e).$$

- We match every agent $i \in R$ to a unique agent $j \in L$ such that

$$d(x_j, x_e) + d(x_i, x_e) = d(x_j, y_2) + d(y_2, x_e) + d(x_i, x_e) = d(x_j, y_2) + d(x_i, y_2).$$

Hence, we have that

$$\begin{aligned} \sum_{i \in N_2} d(x_i, x_e) &= \sum_{i \in L} d(x_i, x_e) + \sum_{i \in M} d(x_i, x_e) + \sum_{i \in R} d(x_i, x_e) \\ &\leq \sum_{i \in N_2} d(x_i, y_2) + d(y_2, x_e) \mathbf{1}\{n_2 \text{ odd}\} + 2 \cdot \sum_{i \in M} d(x_i, x_e). \end{aligned}$$

Next, we will bound the second and third terms of the above expression. Since each agent occupies a different node, we can upper-bound the total distance of the agents in M as follows:

$$\begin{aligned} &d(y_2, x_e) \mathbf{1}\{n_2 \text{ odd}\} + 2 \cdot \sum_{i \in M} d(x_i, x_e) \\ &\leq d(y_2, x_e) \mathbf{1}\{n_2 \text{ odd}\} + 2 \cdot \left(d(y_2, x_e) - 1 + d(y_2, x_e) - 2 + \dots + d(y_2, x_e) - |M| \right) \\ &= -|M|^2 + (2d(y_2, x_e) - 1)|M| + d(y_2, x_e) \mathbf{1}\{n_2 \text{ odd}\}. \end{aligned}$$

Now observe that $d(y_2, x_e) > |M|$ (since all agents in M are between y_2 and e); thus, the last expression in the above derivation is an increasing function in terms of $|M|$. It is clearly also an increasing function in terms of $d(y_2, x_e)$. Since $|M| \leq \frac{1}{2}(n_2 - \mathbf{1}\{n_2 \text{ odd}\})$ and $d(y_2, x_e) \leq n_1 + 1 + |M| \leq n_1 + 1 + \frac{1}{2}(n_2 - \mathbf{1}\{n_2 \text{ odd}\})$, by doing calculations and also using the fact that $\mathbf{1}\{n_2 \text{ odd}\} \leq 1$, we obtain

$$d(y_2, x_e)\mathbf{1}\{n_2 \text{ odd}\} + 2 \cdot \sum_{i \in M} d(x_i, x_e) \leq \frac{1}{4} \left(n_2^2 + 4n_1n_2 + 2n_2 + 1 \right).$$

By putting everything together, we have

$$\begin{aligned} \text{SC}(\mathbf{z}|I) &\leq \sum_{i \in N_1} d(x_i, y_1) + \sum_{i \in N_2} d(x_i, y_2) + \frac{1}{4} \left(n_2^2 + 4n_1n_2 + 2n_2 + 1 \right) \\ &\leq \text{SC}^*(I) + \frac{1}{4} \left(n_2^2 + 4n_1n_2 + 2n_2 + 1 \right). \end{aligned}$$

By Lemma 3.2, we have $\text{SC}^*(I) \geq \frac{1}{4} (n_1^2 + n_2^2 - 2)$, and thus the approximation ratio is bounded by

$$\frac{\text{SC}(\mathbf{z}|I)}{\text{SC}^*(I)} \leq 1 + \frac{n_2^2 + 4n_1n_2 + 2n_2 + 1}{n_1^2 + n_2^2 - 2}.$$

The bound of $17/4$ follows by applying Lemma 3.4 with $x = n_1$ and $y = n_2$. \square

We conclude this section by showing that our analysis of the approximation ratio of FMNE is tight.

LEMMA 3.6. *There exists an instance with $n \geq 5$ and no empty nodes such that the SC-approximation ratio of FMNE is at least 3, and an instance with $n \geq 6$ and at least one empty node such that the SC-approximation ratio of FMNE is at least $17/4$.*

Proof. For the FIXED part of FMNE consider the following instance I_1 with 5 agents and no empty nodes. The first two agents approve only facility 2, and the last three agents approve only facility 1. The mechanism outputs the solution $(2, 3)$, that is, it locates facility 1 at the second node and facility 2 at the third node. The social cost of this solution is $\text{SC}((2, 3)|I_1) = 9$. However, an optimal solution is $(4, 2)$ with social cost $\text{SC}^*(I_1) = 3$, leading to an approximation ratio of 3.

For the MEDIAN-NEAREST-EMPTY part of FMNE consider the following instance I_2 with 6 agents and one empty node. The first three nodes are occupied by agents that approve only facility 2, the next three nodes are occupied by agents that approve only facility 1, and the last node is empty. The mechanism outputs the solution $(5, 7)$, that is, it locates facility 1 at node 5 and facility 2 at the empty node. This solution has social cost $\text{SC}((5, 7)|I_2) = 17$. In contrast, an optimal solution is $(5, 2)$ with $\text{SC}^*(I_2) = 4$, leading to an approximation ratio of $17/4$. \square

4. Social cost: A tight bound for instances with three agents. In this section, we restrict to instances with three agents (and possibly many empty nodes). We show a tight bound of $4/3$ on the approximation ratio of strategyproof mechanisms. In particular, we present a rather simple instance without empty nodes showing that the approximation ratio of any strategyproof mechanism is at least $4/3$; this improves upon the previous lower bound of $9/8$ shown by Serafino and Ventrè [2016]. We complement this result by designing a mechanism that achieves the bound of $4/3$ when given as input any instance with three agents.

Algorithm 4.1 PRIORITY-DICTATORSHIP

Input: Instance I with three agents ℓ , c , and r such that $x_\ell < x_c < x_r$
Output: Feasible solution \mathbf{z}

```

if  $c \in N_1 \setminus N_2$  then
  if  $r \in N_2$  then
     $\mathbf{z} \leftarrow (x_c, x_r)$ 
  else
     $\mathbf{z} \leftarrow (x_c, x_\ell)$ 
  end if
else if  $c \in N_2 \setminus N_1$  then
  if  $r \in N_1$  then
     $\mathbf{z} \leftarrow (x_r, x_c)$ 
  else
     $\mathbf{z} \leftarrow (x_\ell, x_c)$ 
  end if
else
  if  $r \in N_2$  then
     $\mathbf{z} \leftarrow (x_c, x_c + 1)$ 
  else
     $\mathbf{z} \leftarrow (x_c + 1, x_c)$ 
  end if
end if

```

THEOREM 4.1. *The SC-approximation ratio of any strategyproof mechanism is at least $4/3$.*

Proof. We consider two instances with three agents and no empty nodes. In the first instance I_1 , all agents approve both facilities. Clearly, any mechanism must locate a facility to the first or the third node (or, perhaps, both). Without loss of generality, suppose the mechanism locates facility 2 at the third node.

In the second instance I_2 , the first two agents approve both facilities, while the third agent approves only facility 2 (that is, the only difference between I_1 and I_2 is the preference of the third agent). Since facility 2 is located at the third node in I_1 , the same must happen in I_2 ; otherwise, agent 3 would have cost at least 1 in I_2 and incentive to misreport that she approves both facilities, thus changing I_2 to I_1 , and decreasing her cost to 0. However, both possible feasible solutions $\mathbf{z}_1 = (1, 3)$ and $\mathbf{z}_2 = (2, 3)$ have social cost 4 in I_2 , whereas an optimal solution (such as $\mathbf{z}^* = (1, 2)$) has social cost 3; the theorem follows. \square

Next, we design the 3-agent mechanism PRIORITY-DICTATORSHIP, which is strategyproof and has an approximation ratio of at most $4/3$. Consider any instance with three agents; for convenience, we call the agents ℓ , c , and r and let $x_\ell < x_c < x_r$. To be concise, we describe the mechanism (and prove its properties) for the case where c is at least as close to r as to ℓ , that is, $x_r - x_c \leq x_c - x_\ell$; in the other case, it would suffice to swap ℓ and r in the discussion that follows. Our mechanism gives priority to the central agent over the right agent, and does not take into account the preference of the left agent at all. In particular, the mechanism locates at x_c one of the facilities that agent c approves, and decides the location of the other facility based on the preference of agent r . See Algorithm 4.1 for a formal description. We first show that the mechanism is strategyproof.

THEOREM 4.2. PRIORITY-DICTATORSHIP *is strategyproof*.

Proof. Consider any instance with three agents ℓ , c and r such that $x_\ell < x_c < x_r$. Since the preference of agent ℓ is not taken into account, ℓ cannot affect the outcome and thus has no incentive to misreport. In addition, the mechanism always locates at x_c one of the facilities that c approves, and if c approves both facilities, the other facility is located at $x_c + 1$; hence, the cost of c is always minimized. Finally, to see why agent r also has no incentive to misreport, it suffices to observe that in all cases (where the preference of agent c is fixed) the location of the facility that r approves is either independent of her preference, or is closer to her position than if she misreports. As an example, if $c \in N_1 \setminus N_2$, facility 1 is located at x_c independently from the report of r , and facility 2 is located at x_r if r approves it. Hence, agent r minimizes her cost by being truthful. The same holds for the remaining two cases. \square

Next, we show the upper bound of $4/3$ on the approximation ratio of the mechanism.

THEOREM 4.3. *For instances with three agents, the SC-approximation ratio of PRIORITY-DICTATORSHIP is at most $4/3$.*

Proof. Consider any instance I with three agents ℓ , c and r . We distinguish between the three cases considered by the mechanism.

- If $c \in N_1 \setminus N_2$, c is a median agent for facility 1.
 - If $r \in N_2$, the outcome is (x_c, x_r) , and the approximation ratio is 1 as r is a median agent for facility 2.
 - If $r \in N_1 \setminus N_2$, the outcome is (x_c, x_ℓ) , and the approximation ratio is again 1 as either ℓ is a median agent for facility 2, or all agents approve only facility 1.
- If $c \in N_2 \setminus N_1$, due to symmetry to the above case, the approximation ratio is again 1.
- If $c \in N_1 \cap N_2$, c is a median agent for both facilities. The approximation ratio is 1 in the following cases: (a) c is the *unique* median for both facilities, which happens when ℓ and r approve the same set of facilities; (b) r is a median for the facility located at $x_c + 1$ (so that this facility is located in-between median agents), which happens when ℓ and r approve a single (different) facility, or when $\ell \in N_1 \setminus N_2$ and $r \in N_1 \cap N_2$. So, we can consider the remaining three cases. Let $\alpha = x_c - x_\ell \geq 1$ and $\beta = x_r - x_c \geq 1$.
 - $\ell \in N_1 \cap N_2$, $c \in N_1 \cap N_2$, $r \in N_2 \setminus N_1$. One possible optimal solution is (x_ℓ, x_c) with social cost $2\alpha + \beta$. The solution $(x_c, x_c + 1)$ computed by the mechanism has social cost $2\alpha + \beta + 1$. Hence, the approximation ratio is $\frac{2\alpha + \beta + 1}{2\alpha + \beta} = 1 + \frac{1}{2\alpha + \beta}$. As this is a non-increasing function in terms of α and β , it attains its maximum value of $4/3$ for $\alpha = \beta = 1$.
 - $\ell \in N_2 \setminus N_1$, $c \in N_1 \cap N_2$, $r \in N_1 \cap N_2$. One possible optimal solution is (x_r, x_c) with social cost $\alpha + 2\beta$. The solution $(x_c, x_c + 1)$ computed by the mechanism has social cost $\alpha + 2\beta + 1$. Hence, the approximation ratio is $\frac{\alpha + 2\beta + 1}{\alpha + 2\beta} = 1 + \frac{1}{\alpha + 2\beta}$. This is again a non-increasing function in terms of α and β , and thus attains its maximum value of $4/3$ for $\alpha = \beta = 1$.
 - $\ell \in N_1 \cap N_2$, $c \in N_1 \cap N_2$, $r \in N_1 \setminus N_2$. One possible optimal solution is (x_c, x_ℓ) with social cost $2\alpha + \beta$. The solution $(x_c + 1, x_c)$ computed by the mechanism has social cost $2\alpha + \beta + 1$. Hence, the approximation ratio is $\frac{2\alpha + \beta + 1}{2\alpha + \beta} = 1 + \frac{1}{2\alpha + \beta}$, which is maximized once again to $4/3$ for

Algorithm 5.1 α -LEFT-RIGHT (α -LR)

Input: Instance I with n agents
Output: Feasible solution $\mathbf{z} = (z_1, z_2)$
 $L \leftarrow$ left part of line from node 1 to node α
 $N(L) \leftarrow$ agents that occupy nodes in L
 $R \leftarrow$ right part of line from node $\alpha + 1$ to node m
 $N(R) \leftarrow$ agents that occupy nodes in R
 ▷ (Case 1): Each part includes agents that approve only one, different facility
if $\exists X, Y \in \{L, R\}$: $N_1 = N(X)$ and $N_2 = N(Y)$ **then**
 $z_1 \leftarrow$ median of line defined by $N(X)$ (ties in favor of nodes farther from α)
 $z_2 \leftarrow$ median of line defined by $N(Y)$ (ties in favor of nodes farther from α)
 ▷ (Case 2): One part includes agents that approve only one facility
else if $\exists \ell \in \{1, 2\}$, $X \in \{L, R\}$: $N_\ell \subseteq N(X)$ **then**
 if N_ℓ is empty **then**
 $X \leftarrow L$
 end if
 $z_\ell \leftarrow$ median of line defined by $N(X)$ (ties in favor of nodes farther from α)
 $z_{3-\ell} \leftarrow \beta \in \{\alpha, \alpha + 1\} \setminus X$
 ▷ (Case 3): Both parts include agents from N_1 and N_2
else
 $z_1 \leftarrow \alpha$ (rightmost node of L)
 $z_2 \leftarrow \alpha + 1$ (leftmost node of R)
end if

$$\alpha = \beta = 1.$$

The proof is now complete. \square

5. Maximum cost. We now turn our attention to the maximum cost. For this objective, Serafino and Ventrè [2016] showed an upper bound of 3 on the approximation ratio of the TWOEXTREMES mechanism, and a lower bound of $3/2$ on the approximation ratio of any strategyproof mechanism. We improve both bounds by showing a tight bound of 2 when the agents approve at least one facility.

5.1. Improving the upper bound. To achieve the improved upper bound of 2, we consider a class of mechanisms that use only the part of the line that is occupied, from the first to the last occupied node, with possible empty nodes in-between; with some abuse of notation, we denote by m the size of exactly this part of the line. These mechanisms, termed α -LEFT-RIGHT (α -LR, for short), are parameterized by an integer $\alpha \in \{1, \dots, m - 1\}$, and their general idea is as follows: They partition the line into two parts depending on the value of α , and then decide where to locate the facilities based on the preferences of the agents occupying nodes in these two parts. See Algorithm 5.1 for a formal description. We first show that every α -LR mechanism is strategyproof.

THEOREM 5.1. *For any $\alpha \in \{1, \dots, m - 1\}$, mechanism α -LR is strategyproof.*

Proof. Consider any instance. We distinguish between the three cases considered by the mechanism.

True preferences are as in Case 1. The mechanism locates facility 1 at the median node of the line defined by $N(X)$, and facility 2 at the median node of the

line defined by $N(X)$. It suffices to show that any agent $i \in N_1$ has no incentive to deviate; the case $i \in N_2$ is symmetric. If i is the unique agent in $N(X)$, then she occupies the median node of the line defined by $N(X)$, where facility 1 is located, and thus has zero cost. So, we can assume that there is some agent in $N(X) \setminus \{i\}$. If agent i misreports by approving either just facility 2 or both facilities, then we transition to Case 2 with $N_1 \subseteq N(X)$, meaning that the location of facility 1 remains the median of the line defined by $N(X)$. So, agent i cannot decrease her cost, and has no incentive to deviate.

True preferences are as in Case 2. Suppose that $N_1 \subseteq N(L)$, while N_2 has agents in both L and R ; all other cases that fall under Case 2 are symmetric. So, the mechanism locates facility 1 at the median node of the line defined by $N(L)$, and facility 2 at the leftmost node of R . Consider any agent i , and switch between all possible preferences of i :

- $i \in N_1$. Since all agents in $N(R)$ approve only facility 2, we can never transition to Case 3 when i misreports. Also, agent i is indifferent between Case 1 and Case 2 if she only approves facility 1, and prefers Case 2 to Case 1 if she approves both facilities (since her position is closer to the leftmost node of R than to the median node of $N(R)$). So, agent i has no incentive to misreport.
- $i \in (N_2 \setminus N_1) \cap N(L)$. Similarly to the above case, we can never transition to Case 3 when i misreports. Now, i strictly prefers Case 2 to Case 1, as she wants facility 2 to be located at the leftmost node of R , so her cost is minimized, and has no incentive to misreport.
- $i \in (N_2 \setminus N_1) \cap N(R)$. Note that i approves only facility 2. If she misreports that she approves both facilities, then we transition to Case 3, where the location of facility 2 remains the same. If she misreports that she approves only facility 1, then either the outcome remains the same if there is another agent in $N(R)$, or we transition to a symmetric case of Case 2, where $N_2 \subseteq N(L)$ (while N_1 has agents in both L and R), thus changing the location of facility 2 from the leftmost node of R to the median node of the line defined by $N(L)$. As this would increase the cost of agent i , she has no incentive to misreport.

True preferences are as in Case 3. Since each of $N(L)$ and $N(R)$ contains agents from both N_1 and N_2 , the mechanism locates facility 1 at the rightmost node of L , and facility 2 at the leftmost node of R . Consider any agent $i \in N_\ell \cap N(X)$, where $\ell \in \{1, 2\}$ and $X \in \{L, R\}$. Observe that if for each facility $j \in \{1, 2\}$ there exists some agent in $N(X) \setminus \{i\}$ that approves j , then agent i cannot affect the outcome; no matter what i reports, we are still in Case 3. So, we can assume that for some $j \in \{1, 2\}$, all agents in $N(X) \setminus \{i\}$ approve only facility j . Since we are in Case 3, we can also assume that $j \neq \ell$ (of course, agent i might also approve j). To change the case considered by the mechanism, i must completely agree with the other agents in $N(X)$ and report that she approves only facility j . This leads to a symmetric case of Case 2, where $N_\ell \subseteq N(\{L, R\} \setminus X)$ (and N_j contains agents in both L and R). Hence, facility ℓ is located at the median node of the line defined by $N(\{L, R\} \setminus X)$ and facility j is still located at either α or $\alpha + 1$. Clearly, the cost of agent i can only increase as facility ℓ has moved farther away. \square

Next, we focus on the approximation ratio of α -LR mechanisms for the max cost. We distinguish between cases where the size m of the line is an even or odd number, and show that there are values of α such that α -LR achieves an approximation ratio

of at most 2. Before we do this, we prove a lemma providing lower bounds on the optimal max cost of a given instance, which we will use extensively.

LEMMA 5.2. *Let I be an instance. The following are true:*

(a) *If there are two agents positioned at x and $y > x$, and $q \in \{0, 1, 2\}$ is the number of facilities they both approve, then*

$$\text{MC}^*(I) \geq q \cdot \frac{y - x}{2}.$$

(b) *If there is an agent positioned at x that approves both facilities, an agent positioned at $y > x$ that approves facility 1, and an agent positioned at $z > y$ that approves facility 2, then*

$$\text{MC}^*(I) \geq \left\lceil \frac{y + z - 2x}{3} \right\rceil.$$

Proof. To show the two properties, we use the fact that the optimal cost for an instance is at least the optimal cost when we restrict to any subset of agents and any subset of facilities, in which case we aim to balance the cost of all the agents involved. We have:

(a) We begin with the case $q = 1$ as the claim holds trivially when $q = 0$. Clearly, if we place the facility before x or after y , the claim holds. So, let us assume that we place the facility at node a such that $x \leq a \leq y$. The cost of the agent at node x is then (at least) $a - x$, while the cost of the agent at node y is (at least) $y - a$. The claim follows since it cannot be that both $a - x$ and $y - a$ are strictly less than $\frac{y-x}{2}$.

When $q = 2$, the claim follows if at least one facility is placed before x or after y . Let us assume that we place the facilities at nodes a and b such that $x \leq \min\{a, b\} < \max\{a, b\} \leq y$. The cost of the agent at node x is then (at least) $a + b - 2x$, while the cost of the agent at node y is (at least) $2y - a - b$. The claim follows since it cannot be that both $a + b - 2x$ and $2y - a - b$ are strictly less than $y - x$.

(b) In this case, we want to locate facility 1 at some node $a \in [x, y]$ and facility 2 at some node $b \in [x, z]$, such that the maximum cost among the three agents is minimized. The cost of the agent at node x is then (at least) $a + b - 2x$, the cost of the agent at y is (at least) $y - a$, while the cost of the agent at z is (at least) $z - b$. As the sum of costs is (at least) $y + z - 2x$, it cannot be the case that all three costs are strictly less than $\frac{y+z-2x}{3}$. The claim follows since any cost must be an integer. \square

We are now ready to bound the approximation ratio of particular α -LR mechanisms. We start with instances where m is an even number, for which we use $\alpha = m/2$; that is, we partition the line into two parts of equal size.

THEOREM 5.3. *When m is even, the MC-approximation ratio of $m/2$ -LR is at most 2.*

Proof. Consider any instance I . We distinguish between the three cases considered by the mechanism.

Case 1. Since the agents in $N(X)$ approve only facility 1 and the agents in $N(Y)$ approve only facility 2, locating facility 1 at the median node of the line defined by $N(X)$, and facility 2 at the median node of the line defined by $N(Y)$ is the optimal solution.

Case 2. Suppose that $N_1 \subseteq N(L)$ and that N_2 contains agents in both $N(L)$ and $N(R)$; this is one of the symmetric instances captured by Case 2. The mechanism

locates facility 1 at the median node y_L (with $1 \leq y_L \leq \lfloor \frac{m+2}{4} \rfloor$) of the line defined by $N(L)$, and facility 2 at node $\frac{m}{2} + 1$ (the leftmost node of R). We distinguish between the following cases depending on the preferences of the agents with the maximum cost for the solution \mathbf{z} computed by the mechanism.

- **The cost of the mechanism is equal to the cost of an agent that approves a single facility.** As all agents that approve facility 1 are in $N(L)$, and facility 1 is located at the median of the line defined by $N(L)$, the cost of any agent that approves only facility 1 can be at most $\max\{\lfloor \frac{m+2}{4} \rfloor - 1, \frac{m}{2} - \lfloor \frac{m+2}{4} \rfloor\} \leq \frac{m}{4}$. Since facility 2 is located at node $\frac{m}{2} + 1$, the cost of any agent that approves only facility 2 can be at most $\frac{m}{2} + 1 - 1 = \frac{m}{2}$. Hence, $\text{MC}(\mathbf{z}|I) \leq \frac{m}{2}$. As N_2 contains at least one agent in $N(L)$, there exists at least one agent at a node $x \leq \frac{m}{2}$ that approves facility 2. By applying Lemma 5.2(a) with x and $y = m$, we have that $\text{MC}^*(I) \geq \frac{m-x}{2} \geq \frac{m}{4}$, yielding that the approximation ratio is at most 2.
- **The cost of the mechanism is equal to the cost of an agent that approves both facilities.** Since we are in Case 2 with $N_1 \subseteq N_L$, let $x \leq m/2$ be the position of the agent i that approves both facilities and has the maximum cost among all such agents. The cost of agent i , and thus of the mechanism, is $\text{MC}(\mathbf{z}|I) = |x - y_L| + \frac{m}{2} + 1 - x$. If $x > y_L$, we have $\text{MC}(\mathbf{z}|I) = \frac{m}{2} + 1 - y_L \leq \frac{m}{2}$. As in the case where the cost is due to an agent that approves a single facility, we have $\text{MC}^*(I) \geq \frac{m}{4}$, and thus the approximation ratio is at most 2. Otherwise, if $x \leq y_L$, we have $\text{MC}(\mathbf{z}|I) = \frac{m}{2} + 1 + y_L - 2x \leq \frac{3(m+2)}{4} - 2x$. Since agent i and the agent at node m both approve facility 2, by Lemma 5.2(a), we have that $\text{MC}^*(I) \geq \frac{m-x}{2}$. Hence, the approximation ratio is at most $\frac{3m+6-8x}{2m-2x}$. As this is a non-increasing function in terms of x , it attains its maximum value of $\frac{3m-2}{2m-2}$ for $x = 1$. For every $m \geq 2$, it holds that $\frac{3m-2}{2m-2} \leq 2$.

Case 3. Recall that the mechanism locates facility 1 at $\frac{m}{2}$ (rightmost node of L), and facility 2 at $\frac{m}{2} + 1$ (leftmost node of R). Without loss of generality, we can assume that the agent at node 1 approves facility 1 (and possibly also facility 2). We switch between the following two subcases:

- **The cost of the mechanism is equal to the cost of an agent that approves a single facility.** Then, $\text{MC}(\mathbf{z}|I) \leq \frac{m}{2}$ (the distance between node 1 and node $\frac{m}{2} + 1$). As we are in case 3, there exists an agent at some node $y \geq \frac{m}{2} + 1$ that approves facility 1, and by our assumption that the agent at node 1 approves facility 1, Lemma 5.2(a) gives $\text{MC}^*(I) \geq \frac{y-1}{2} \geq \frac{m}{4}$. So, the approximation ratio is at most 2.
- **The cost of the mechanism is equal to the cost of an agent that approves both facilities.** Without loss of generality, let $x \leq m/2$ be the position of the agent i that has the maximum cost among all agents that approve both facilities. Then, $\text{MC}(\mathbf{z}|I) = \frac{m}{2} - x + \frac{m}{2} + 1 - x = m + 1 - 2x$. As the agent at node m approves some facility that is also approved by i , by Lemma 5.2(a), we get $\text{MC}^*(I) \geq \frac{m-x}{2}$. The approximation ratio is $2 \cdot \frac{m+1-2x}{m-x}$, which is a non-increasing function in terms of x , and attains its maximum value of 2 for $x = 1$.

In any case, the approximation ratio of the mechanism is 2, and the theorem follows. \square

For instances with odd m , we use $\alpha = (m + 1)/2$. The proof of the following

theorem is similar in structure with the previous theorem for even m , but is slightly more complicated.

THEOREM 5.4. *When m is odd, the MC-approximation ratio of $(m+1)/2$ -LR is at most 2.*

Proof. Consider any instance I . We distinguish between the three cases considered by the mechanism.

Case 1. Since the agents in $N(X)$ approve only facility 1, and the agents in $N(Y)$ approve only facility 2, locating facility 1 at the median node of the line defined by $N(X)$, and facility 2 at the median node of the line defined by $N(Y)$ is the optimal outcome.

Case 2. Suppose that $N_1 \subseteq N(L)$ and that N_2 contains agents in both $N(L)$ and $N(R)$; this is one of the symmetric instances captured by Case 2. The mechanism locates facility 1 at the median node y_L (with $1 \leq y_L \leq \lfloor \frac{m+3}{4} \rfloor$) of the line defined by $N(L)$, and facility 2 at node $\frac{m+1}{2} + 1 = \frac{m+3}{2}$ (the leftmost node of R). We distinguish between the following cases depending on the preferences of the agents with the maximum cost for the solution \mathbf{z} computed by the mechanism.

- **The cost of the mechanism is equal to the cost of an agent that approves a single facility.** As all agents that approve facility 1 are in $N(L)$, and facility 1 is located at the median of the line defined by $N(L)$, the cost of any agent that approves only facility 1 can be at most $\max\{\lfloor \frac{m+3}{4} \rfloor - 1, \frac{m+1}{2} - \lfloor \frac{m+3}{4} \rfloor\} \leq \frac{m+1}{4}$. Let $x \leq \frac{m+1}{2}$ be the position of the leftmost agent that approves facility 2. Since the agent at node m also approves facility 2, by Lemma 5.2(a), we have that $\text{MC}^*(I) \geq \frac{m-x}{2}$. We now distinguish between two subcases, based on the value of x .

If $x = \frac{m+1}{2}$, the maximum cost among agents that approve facility 2 is at most $m - \frac{m+3}{2} = \frac{m-3}{2}$, and thus $\text{MC}(\mathbf{z}|I) \leq \max\{\frac{m+1}{4}, \frac{m-3}{2}\}$. Since $\text{MC}^*(I) \geq \frac{m-1}{4}$, the approximation ratio is at most 2.

Otherwise, if $x \leq \frac{m-1}{2}$, the maximum cost among agents that approve facility 2 is at most $\frac{m+3}{2} - 1 = \frac{m+1}{2}$, and thus $\text{MC}(\mathbf{z}|I) \leq \max\{\frac{m+1}{4}, \frac{m+1}{2}\} = \frac{m+1}{2}$. Since $\text{MC}^*(I) \geq \frac{m-x}{2} \geq \frac{m+1}{4}$, the approximation ratio is again at most 2.

- **The cost of the mechanism is equal to the cost of an agent that approves both facilities.** Since we are in Case 2 with $N_1 \subseteq N_L$, let $x \leq \frac{m+1}{2}$ be the position of the agent i that approves both facilities and has the maximum cost among all such agents. The cost of agent i , and thus of the mechanism, is $\text{MC}(\mathbf{z}|I) = |x - y_L| + \frac{m+3}{2} - x$.

If $x \leq y_L$, then since $y_L \leq \frac{m+3}{4}$, we have that $\text{MC}(\mathbf{z}|I) = \frac{m+3}{2} + y_L - 2x \leq \frac{3(m+3)}{4} - 2x$. As node m is occupied by an agent that approves facility 2, by Lemma 5.2(a), we have that $\text{MC}^*(I) \geq \frac{m-x}{2}$, and thus the approximation ratio is $\frac{3m+9-8x}{2m-2x}$. This is a non-increasing function of $x \geq 1$, and attains its maximum value of $\frac{3m+1}{2m-2}$ for $x = 1$. For every $m \geq 5$, it holds that $\frac{3m+1}{2m-2} \leq 2$. When $m = 3$, for $x \leq y_L$ to be true, it has to be the case that $x = y_L = 1$; so, the cost of agent i for \mathbf{z} is 2, while $\text{MC}^*(I) \geq 1$, leading to an approximation ratio of at most 2.

Otherwise, if $x > y_L$, for $y_L = 1$ to be possible, it would have to be the case that $m = 3$ and $x = 2$; then, the cost of agent i is 2, while $\text{MC}^*(I) \geq 1$, and so the approximation ratio is at most 2. Hence, assume that $y_L \geq 2$. Then, we have $\text{MC}(\mathbf{z}|I) = \frac{m+3}{2} - y_L \leq \frac{m-1}{2}$. Since $x \leq \frac{m+1}{2}$ and the agent at node m approves facility 2, by Lemma 5.2(a), we have that $\text{MC}^*(I) \geq \frac{m-1}{4}$, and

the approximation ratio is at most 2.

Case 3. Recall that in this case the mechanism locates facility 1 at $\frac{m+1}{2}$, and facility 2 at $\frac{m+3}{2}$. Without loss of generality, we assume that the agent at node 1 approves facility 1 (and possibly also facility 2). We switch between two subcases:

- **The cost of the mechanism is equal to the cost of an agent that approves a single facility.** Then, $MC(\mathbf{z}|I) \leq \frac{m+3}{2} - 1 = \frac{m+1}{2}$. As we are in case 3, there exists an agent at some node $y \geq \frac{m+3}{2}$ that approves facility 1, and by our assumption that the agent at node 1 approves facility 1, Lemma 5.2(a) gives $MC^*(I) \geq \frac{y-1}{2} \geq \frac{m+1}{4}$, and the approximation ratio is at most 2.
- **The cost of the mechanism is equal to the cost of an agent that approves both facilities.** Without loss of generality, let $x \leq \frac{m+1}{2}$ be the position of the agent i that has the maximum cost among agents that approve both facilities. Then, $MC(\mathbf{z}|I) = \frac{m+1}{2} - x + \frac{m+3}{2} - x = m - 2x + 2$. Since we are in case 3, in $N(R)$, there exists an agent that approves facility 1 and an agent that approves facility 2. Consider the following two subcases:
If there is an agent j at some node $y \geq \frac{m+3}{2}$ that approves both facilities, then, by Lemma 5.2(a), $MC^*(I) \geq y - x \geq \frac{m+3}{2} - x$, and the approximation ratio is at most $2 \cdot \frac{m-2x+2}{m-2x+3} \leq 2$.

Otherwise, if there is no agent in $N(R)$ that approves both facilities, suppose that the agent at node m approves facility 2, and there exists an agent at some node $y \in [\frac{m+3}{2}, m)$ that approves facility 1. If $x = 1$, then $MC(\mathbf{z}|I) \leq m$ and, by Lemma 5.2(b), $MC^*(I) \geq \lceil \frac{y+m-2}{3} \rceil \geq \lceil \frac{m}{2} - \frac{1}{6} \rceil = \frac{m+1}{2}$; hence, the approximation ratio is at most 2. If $x \geq 2$, it suffices to use the bound $MC^*(I) \geq \frac{m-x}{2}$ implied by Lemma 5.2(a), to get an upper bound of $2 \cdot \frac{m-2x+2}{m-x}$ on the approximation ratio. This is a non-increasing function of x , and thus attains its maximum value of 2 for $x = 2$.

In any case, the approximation ratio is at most 2, and the proof is complete. \square

5.2. A tight lower bound for deterministic mechanisms. We conclude the presentation of our technical results with a tight lower bound of 2 on the approximation ratio of any strategyproof mechanism with respect to the maximum cost objective.

THEOREM 5.5. *The MC-approximation ratio of any strategyproof mechanism is at least 2.*

Proof. Suppose there exists a strategyproof mechanism M with approximation ratio strictly smaller than 2. We will reach a contradiction by examining a series of instances, all of which involve three agents and no empty nodes; see also Figure 1.

We begin with instance I_1 , in which the first and third agents approve only facility 1, while the second agent approves only facility 2. Clearly, M must return either $(2, 3)$ or $(2, 1)$ as $MC((2, 3)|I_1) = MC((2, 1)|I_1) = 1$; any solution where facility 1 is not placed at the second node has maximum cost 2, and returning such a solution would contradict the assumption that the approximation ratio of M is strictly smaller than 2. Let us assume that M returns the solution $(2, 3)$; the other case is similar.

Next, consider instance I_2 , in which the first agent approves only facility 1, while the remaining agents approve only facility 2. M must output either $(2, 3)$ or $(1, 3)$ due to strategyproofness. Indeed, any solution where facility 2 is not placed at the third node leads to a cost of at least 1 for the third agent. But then, that agent would misreport that she only approves facility 1, thus leading to instance I_1 , and obtain a

cost of 0 for the resulting solution $(2, 3)$.

If M returns $(2, 3)$ for instance I_2 , consider instance I_3 , in which the first agent approves both facilities, while the other two agents approve facility 2. M must return the optimal solution $(1, 2)$ with $\text{MC}((1, 2)|I_3) = 1$, since any other solution leads to a maximum cost of at least 2. In this case, however, the first agent in I_2 would misreport that she approves both facilities to reduce her cost from 1 to 0; this contradicts the assumption that M is strategyproof.

If M returns $(1, 3)$ for I_2 , consider instance I_4 , in which the first agent approves only facility 1, the second agent approves both facilities, and the last agent approves only facility 2. There are two optimal solutions in I_4 , $(2, 3)$ and $(1, 2)$, with $\text{MC}((2, 3)|I_4) = \text{MC}((1, 2)|I_4) = 1$; any other solution has maximum cost 2. Out of these solutions, $(1, 2)$ would give the second agent in I_2 incentive to misreport that she approves both facilities to reduce her cost from 1 to 0. Hence, M must return $(2, 3)$ when given as input I_4 . To conclude the proof, consider instance I_5 , in which the first two agents approve both facilities, while the third agent approves only facility 2. The optimal solution is $(1, 2)$ with $\text{MC}((1, 2)|I_5) = 1$; any other solution has maximum cost of at least 2. But then, the first agent in I_4 has incentive to misreport that she approves both facilities to reduce her cost from 1 to 0; this again contradicts the fact that M is strategyproof. \square

6. Conclusion and open problems. In this paper, we revisited the discrete truthful heterogeneous two-facility location problem, and showed bounds on the approximation ratio of deterministic strategyproof mechanisms when agents approve at least one facility. Our results greatly improve upon the previous best-known ones for this case, both with respect to the social cost as well as the maximum cost. There are still many open questions and directions for future research.

While we were able to show a bound for the social cost that is a small constant, thus improving upon the linear bound that was previously known for a more general setting, we were unable to close the gap between the lower bound of $4/3$ and the upper bound of $17/4$. Besides deterministic mechanisms, it would also be interesting to focus on randomized mechanisms and the maximum cost objective, for which there is a gap between the lower bound of $4/3$ and the upper bound of $3/2$ shown by Serafino and Ventrè [2016]. Let us also remark that improving upon the linear bound for the social cost in the particular setting studied by Serafino and Ventrè poses an interesting challenge; our lower bounds clearly extend to that setting, while we believe that our mechanism for the maximum cost achieves the same approximation guarantee requiring only slight modifications.

Going beyond the particular model studied here, there are many extensions to be considered. One could study settings with more than just two facilities, settings where the positions of the agents are their private information and can report empty nodes as their positions, settings with different heterogeneous preferences such as fractional or obnoxious ones, and also settings with more general location graphs such as trees or regular graphs.

References.

- Eleftherios Anastasiadis and Argyrios Deligkas. Heterogeneous facility location games. In *Proceedings of the 17th International Conference on Autonomous Agents and MultiAgent Systems (AAMAS)*, pages 623–631, 2018.
- Elliot Anshelevich, Aris Filos-Ratsikas, Nisarg Shah, and Alexandros A. Voudouris. Distortion in social choice problems: The first 15 years and beyond. In *Proceedings*

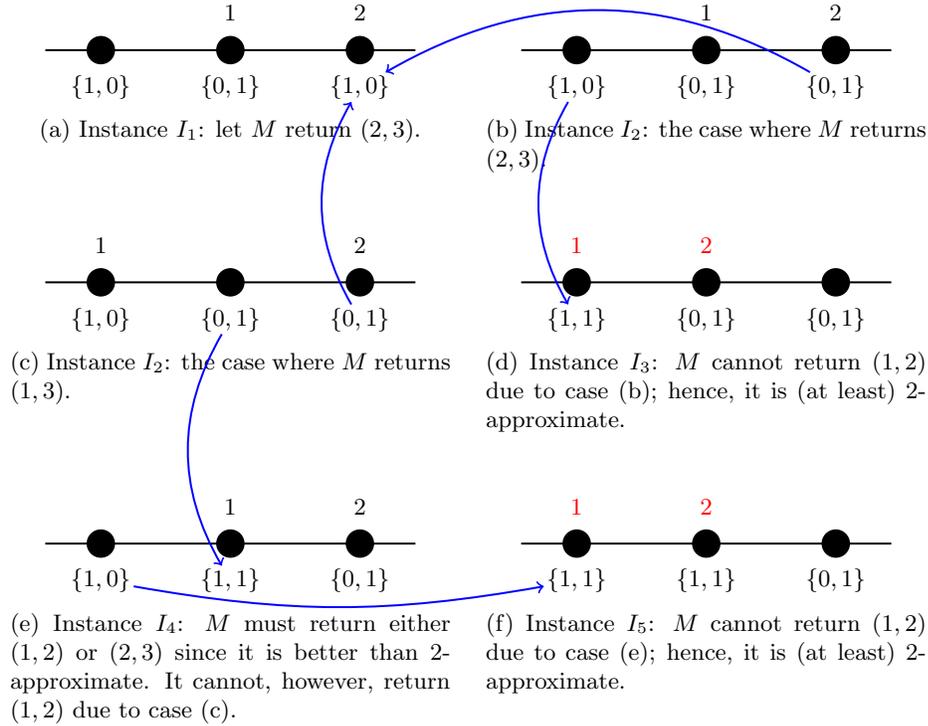


Fig. 1: The instances used in the proof of Theorem 5.5. Each instance has 3 agents and no empty nodes. The agent preferences appear below each node, while the facility assignment appears above the nodes. In particular, the (black) numbers above the nodes in (a), (b), (c) and (e) denote the mechanism's output in the corresponding instances, while the (red) numbers above the nodes in (d) and (f) denote an optimal but excluded assignment. The (blue) arrows denote how instances are related when a single agent's preferences change.

of the 30th International Joint Conference on Artificial Intelligence (IJCAI), pages 4294–4301, 2021.

Hau Chan, Aris Filos-Ratsikas, Bo Li, Minming Li, and Chenhao Wang. Mechanism design for facility location problems: A survey. In *Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence (IJCAI)*, pages 4356–4365, 2021.

Zhihuai Chen, Ken C. K. Fong, Minming Li, Kai Wang, Hongning Yuan, and Yong Zhang. Facility location games with optional preference. *Theoretical Computer Science*, 847:185–197, 2020.

Yukun Cheng, Wei Yu, and Guochuan Zhang. Mechanisms for obnoxious facility game on a path. In *Proceedings of the Combinatorial 5th International Conference Optimization and Applications (COCO)*, pages 262–271, 2011.

Yukun Cheng, Wei Yu, and Guochuan Zhang. Strategy-proof approximation mechanisms for an obnoxious facility game on networks. *Theoretical Computer Science*, 497:154–163, 2013.

- Argyrios Deligkas, Aris Filos-Ratsikas, and Alexandros A. Voudouris. Heterogeneous facility location with limited resources. In *Proceedings of the 36th AAAI Conference on Artificial Intelligence (AAAI)*, pages 4966–4974, 2022.
- Elad Dokow, Michal Feldman, Reshef Meir, and Ilan Nehama. Mechanism design on discrete lines and cycles. In *Proceedings of the 13th ACM Conference on Electronic Commerce (EC)*, pages 423–440, 2012.
- Edith Elkind, Minming Li, and Houyu Zhou. Facility location with approval preferences: Strategyproofness and fairness. In *Proceedings of the 2022 International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS)*, pages 391–399, 2022.
- Itai Feigenbaum and Jay Sethuraman. Strategyproof mechanisms for one-dimensional hybrid and obnoxious facility location models. In *AAAI Workshop on Incentive and Trust in E-Communities*, volume WS-15-08, 2015.
- Michal Feldman, Amos Fiat, and Iddan Golomb. On voting and facility location. In *Proceedings of the 2016 ACM Conference on Economics and Computation (EC)*, pages 269–286, 2016.
- Alina Filimonov and Reshef Meir. Strategyproof facility location mechanisms on discrete trees. In *Proceedings of the 20th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, pages 510–518, 2021.
- Aris Filos-Ratsikas and Alexandros A. Voudouris. Approximate mechanism design for distributed facility location. In *Proceedings of the 14th International Symposium on Algorithmic Game Theory (SAGT)*, pages 49–63, 2021.
- Chi Kit Ken Fong, Minming Li, Pinyan Lu, Taiki Todo, and Makoto Yokoo. Facility location games with fractional preferences. In *Proceedings of the 32nd AAAI Conference on Artificial Intelligence (AAAI)*, pages 1039–1046, 2018.
- Maria Kyropoulou, Carmine Ventre, and Xiaomeng Zhang. Mechanism design for constrained heterogeneous facility location. In *Proceedings of the 12th International Symposium on Algorithmic Game Theory (SAGT)*, pages 63–76, 2019.
- Minming Li, Pinyan Lu, Yuhao Yao, and Jialin Zhang. Strategyproof mechanism for two heterogeneous facilities with constant approximation ratio. In *Proceedings of the 29th International Joint Conference on Artificial Intelligence (IJCAI)*, pages 238–245, 2020.
- Ariel D. Procaccia and Moshe Tennenholtz. Approximate mechanism design without money. *ACM Transactions on Economics and Computation*, 1(4):18:1–18:26, 2013.
- Paolo Serafino and Carmine Ventre. Heterogeneous facility location without money. *Theoretical Computer Science*, 636:27–46, 2016.
- Xinping Xu, Bo Li, Minming Li, and Lingjie Duan. Two-facility location games with minimum distance requirement. *Journal of Artificial Intelligence Research*, 70:719–756, 2021.
- Shaokun Zou and Minming Li. Facility location games with dual preference. In *Proceedings of the 2015 International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, pages 615–623, 2015.