Mechanism Design without Money for Heterogeneous and Distributed Facility Location Problems

Rongsen Zhang

A thesis submitted for the degree of Doctor of Philosophy

School of Computer Science and Electronic Engineering

University of Essex

March, 2024

Abstract

Algorithmic Mechanism Design (AMD) is an interdisciplinary field that bridges economics and computer science and aims to design mechanisms for systems with self-interested agents. This study focuses on motivating truthful information disclosure while optimizing social goals. Traditional payment-based mechanisms cannot be effectively applied to some of these problems, but we can utilize approximate mechanisms to obtain truthfulness without recourse to payments.

One of the well-established problems in approximate mechanism design without money is the facility location problem. In this context, agents possess private positions along a line, and the goal is to determine the location of a public facility while motivating truthful disclosures and achieving optimal social outcomes.

This thesis presents contributions across three key dimensions of facility location problems: (1) Heterogeneous Two-Facility Location Problem: Addressing a discrete setting where agents occupy nodes on a line graph and possess private preferences for two facilities, the research introduces deterministic strategy-proof mechanisms with improved approximation ratios, surpassing existing approaches. (2) Two-Facility Location with Candidate Positions: Investigating another variant, where agents have private positions and known preferences for two facilities, the study identifies deterministic strategy-proof mechanisms that achieve the best possible approximation ratios for social and maximum costs.

(3) Distributed Facility Location Problem: A set of agents with positions on the line of real numbers are partitioned into disjoint districts, we designed deterministic distributed mechanisms that satisfy various criteria of interest and achieve the best possible distortion bounds. The research analyzes two mechanism classes: Unrestricted, where agents directly provide truthful positions, and strategy-proof, designed to incentivize honesty. The study establishes tight bounds for various social objectives, including average social cost, max cost, and other fairness-inspired criteria.

In summary, approximate mechanism design without money addresses complex challenges in multi-agent systems, creating mechanisms that promote truthfulness and optimize societal objectives. This thesis introduces innovative mechanisms and provides comprehensive insights into facility location problems from three distinct perspectives.

Acknowledgements

As I conclude this doctoral journey, profound gratitude is extended to my esteemed supervisors, Dr Panagiotis Kanellopoulos and Dr Alexandros A. Voudouris, whose guidance and unwavering support have been instrumental in completing this thesis. Their expertise, insightful feedback, and commitment to academic excellence have shaped the intellectual rigour of my research, leaving an indelible mark on my academic journey. I am privileged to have had them as mentors, and their influence will resonate in my pursuits beyond academia.

Special thanks go to the members of the termly panel meetings, Dr Haider Raza and Dr Ali Zulfiqar at the University of Essex, as well as my thesis examiners, Dr Maria Kyropoulou at the University of Essex and Dr Eduard Eiben at the Royal Holloway University of London. Their valuable insights and comments on my research, presentation, and thesis structure are appreciated.

To my parents and entire family, your enduring support has strengthened me throughout this challenging endeavour. This thesis stands as much as a testament to your contributions as it does to my academic pursuits. Your sacrifices and unwavering encouragement have been instrumental in reaching this milestone.

A special person deserves a unique acknowledgement during academic challenges and research endeavours. My dearest girlfriend, Xunyi Chen. Your patience during late nights and encouragement during self-doubt have been the pillars of this thesis. Your belief in me has been a constant source of inspiration, and your comforting presence has turned the toughest days into manageable ones. Your love has been my sanctuary, and I am profoundly grateful for your unwavering support.

In this academic odyssey, I have had the privilege of forging friendships with Jiayi Bai, Henlin Chen, Eva Christodoulaki, Jiacheng Dong, Daniel Gettings, Ouwen Hong, Zhirui Huang, Shize Liu, Xinyu Li, Zhuxin Ran, Shuang Sun, Zicheng Wu, Yi Wang, Lyu Yang, Yukun Zhang, Ruikai Zhong, and Hao Zhou. Meeting you during this pursuit has been fortuitous and enriching.

Lastly, special appreciation is extended to everyone at the Centre for Computational Finance and Economic Agents (CCFEA) for fostering such a pleasant environment. Your collective contribution has added to the richness of this academic journey.

Contents

Abstract **ii**

Acknowledgements **v**

1 Introduction **1**

1.1 Algorithmic Mechanism Design 1

1.2 Facility Location Problems and Games 4

1.2.1 Two-facility Location Games 5

1.2.2 Multiple facilities 6

1.2.3 Obnoxious Facility Location 7

1.2.4 Heterogeneous Facility Location 8

1.2.5 Facility Location with Candidate Locations 10

1.2.6 Distributed Facility Location 12

1.3 Structure of the Thesis 13

1.4 List of Publications 16

2 Preliminaries **18**

3 On Discrete Truthful Heterogeneous Two-Facility Location **21**

3.1 Definitions and notation 21

3.2 Overview of Contribution 22

3.3 Social cost: A general constant upper bound 24

3.4 Social cost: A tight bound for instances with three agents 34

3.5 Maximum cost 38

3.5.1 Improving the upper bound 38

3.5.2 A tight lower bound 49

4 Heterogeneous Two-Facility Location with Candidate Locations **52**

4.1 Definitions and notation 52

4.2 Overview of contribution 54

4.3 Social cost . . . 57
4.3.1 Doubleton instances . . . 57
4.3.2 Singleton instances . . . 63
4.3.3 General instances . . . 69
4.4 Max cost . . . 75
4.4.1 Doubleton instances . . . 75
4.4.2 Singleton instances . . . 77
4.4.3 General instances . . . 84
4.5 Allowing same facility locations . . . 85
4.5.1 Social cost . . . 87
4.5.2 Max cost . . . 92

5 Settling the Approximation Ratio of Distributed Facility Location 95

5.1 Definitions and notation . . . 95
5.1.1 Social objectives and strategyproofness . . . 96
5.1.2 Useful observations . . . 97
5.2 Overview of Contribution . . . 100
5.3 Average social cost . . . 101
5.3.1 Unrestricted mechanisms . . . 101
5.4 Max cost . . . 109
5.5 Average-of-Max . . . 111
5.5.1 Unrestricted mechanisms . . . 111
5.5.2 Strategyproof mechanisms . . . 112
5.6 Max-of-Average . . . 118
5.6.1 Unrestricted mechanisms . . . 118
5.6.2 Strategyproof mechanisms . . . 120

6 Conclusion and Directions for Future Work 125

References 128

Chapter 1

Introduction

Algorithmic game theory (AGT) [Roughgarden, 2010] is a multidisciplinary area in the intersection of computer science and economics. Its central goal is to analyze and design algorithms when input is given from selfish, strategic agents. In problems considered in this area, the agents aim to maximize their own profit, possibly by misreporting their input, which may lead to system inefficiency. Typical objectives include the study of the existence of stable outcomes (known as *equilibria*), the analysis of system efficiency in these outcomes, and the design of mechanisms that lead to efficient equilibria that may also enjoy other properties. The focus of this thesis is on a particular class of interesting problems in AGT (known as facility location games) and the design of efficient mechanisms for them that have the property to incentivize the strategic agents report their input truthfully.

1.1 Algorithmic Mechanism Design

A subarea of AGT is that of Algorithmic Mechanism Design (AMD), where the main objective is to design mechanisms (define the rules of the game) for multiple self-interested participants, ensuring that their actions lead to equilibra with good mechanism performance and no participant has incentive to deviate to affect the mechanism performance. Key goals include revenue maximization and social welfare maximization. In AMD we want strategyproofness in combination with approximately optimal system efficiency, rather than strategyproofness combined with other axioms as in classic Mechanism Design that has been studied in Economics. Moreover, it prioritizes computational efficiency, dismissing mechanisms that can't be efficiently implemented in polynomial time as unsuitable solutions.

As a result, we tend to exclude classical economic mechanisms, such as the Vickrey-Clark-Groves(VCG) Auction [Vickrey, 1961]. In this auction, the highest bidder receives the item but pays an amount equal to the second highest bidder, thus incentivizing honest bidding. It extends this principle to the case where there are multiple items seeking to optimize the total value of the players while requiring a fee that reflects the value of the forgoing allocation. However, despite the theoretical beauty of the VCG mechanism (it is an optimal mechanism which means the worst-case results from this mechanism are comparable to the best results for all mechanisms and can make sure each participant doesn't have any incentive to deviate), its general applicability is limited by practical factors. One of the most significant limitations is computational issues. For example, determining the optimal allocation in combinatorial auctions is known to be NP-hard, indicating immense computational complexity [Roughgarden, 2010]. While methods such as approximation algorithms can provide solutions to optimization problems, using such approximations introduces a potential drawback. It may compromise the strategyproofness of the mechanism, which means that participants may not find it in their best interest to be completely honest [Nisan and Ronen, 1999]. Generally, AMD can be divided into two main parts, whether involving money [Nisan and Ronen, 1999] or not [Procaccia and Tennenholtz, 2013]. The field of mechanism design with money has permeated various domains, including *combinatorial auctions* [Sandholm, 2002, Dobzinski et al., 2006, Mu'Alem and Nisan, 2008, Assadi and Singla, 2020, Assadi et al., 2021], *keyword search auctions* [Aggarwal et al., 2006, Zhou et al., 2008], *sponsored search auctions* [Qin et al., 2015, Roberts et al., 2016] and *social welfare maximization*[Briest et al., 2005, Filos-Ratsikas et al., 2014, Filos-Ratsikas and Miltersen, 2014, Huang et al., 2019, Rahmattalabi et al., 2021, Banerjee et al., 2022].

Payments are easily challenged when it comes to social choice settings. For example, even if it is the same thing, different people will have different inner prices for it, and we need to have something else to quantify them. Or in elections, the introduction of money will cause corruption. In addition, Schummer and Vohra [2007] states that 'money cannot be used as a medium of compensation in many important contexts' because of ethical considerations (such as in political decision-making) or legal considerations (such as in organ donation). So, it is natural to ask whether it is possible to design some mechanisms that do not require payment; these mechanisms are called *strategyproof* in the social choice literature. Subsequently, approximate mechanism design without money has been applied in many different settings like in *fair division* [Cole et al., 2013, Amanatidis et al., 2017, 2023, Bei et al., 2020], *voting* [Aziz and Lam, 2021, Feldman et al., 2016, Filos-Ratsikas and Voudouris, 2021, Anshelevich et al., 2022], and *applications in kidney exchange* [Caragiannis et al., 2011, Ashlagi et al., 2015, McElfresh et al., 2019, Freedman et al., 2020].

1.2 Facility Location Problems and Games

The facility location problem (FLP) has been studied extensively in various different disciplines, including operations research, computational geometry, and approximation algorithms [Hochbaum, 1982, Williamson and Shmoys, 2011]. In its fundamental formulation, the facility location problem comprises a set of potential facility points and a set of demand points. The task is to choose a subset of facilities to open, aiming to minimize the total or max distance from each demand point to its nearest facility.

The facility location problem has also been studied from the perspective of game theory. When the objective is to decide where to place a single facility given *single-peaked* preferences of a set of agents (such as when the agents are positioned on a line and report their private positions), Moulin [1980] showed that the class of strategyproof mechanisms (i.e., those that incentivize the agents to report their positions truthfully) consists of generalized median rules, which choose to place the facility at the median of a set of points that includes the positions of the agents and a set of (possibly) empty positions.

More recently, Procaccia and Tennenholtz [2013] considered the single-facility location game on the line from the perspective of approximate mechanism design without money, aiming to identify strategyproof mechanisms that also achieve a good approximation of the social objective functions, such as the social cost (total distance of agents from the chosen facility location) or the maximum cost.

Procaccia and Tennenholtz showed that the optimal social cost can always be achieved by the Median mechanism, which is strategyproof and operates as follows: if the total number of all agents is odd, it selects the median location of the agents' positions; if the total number is even, it selects the left median of all agents' positions. For the maximum cost, Procaccia and Tennenholtz showed that the deterministic strategyproof mechanism Two-Extreme is 2-approximate. This mechanism operates as follows: it selects the leftmost (or rightmost) agent's location to be the facility location. Procaccia and Tennenholtz also showed that the approximation ratio of any strategyproof deterministic mechanism for the maximum cost is at least 2.

Procaccia and Tennenholtz also considered a randomized strategyproof mechanisms to get a better bound for the maximum cost. They showed that the mechanism which selects the leftmost agent's location with probability $1/4$, the rightmost agent's location with probability $1/4$, and their midpoint with probability $1/2$ achieves an upper bound of $3/2$, which is the best possible among all randomized mechanisms.

Alon et al. [2010] designed strategyproof mechanism for more general metrics (circles and general graphs) to extend this problem. Tang et al. [2020a] characterized group-strategyproof mechanisms for the single facility location game in strictly convex space and proved (almost) tight bound for the social cost of $n/2$ and maximum cost of 2.

Beyond the single-facility location game, many variants that consider multiple facilities, constrained spaces, or heterogeneous preferences have been studied. We discuss these below.

1.2.1 Two-facility Location Games

In the truthful *two-facility* location problem, the objective is to locate two facilities. The individual cost of an agent is equal to the distance from the closest facility. For the social cost, Procaccia and Tennenholtz [2013] showed a constant lower bound of $3/2$ on the approximation ratio of all deterministic mechanisms and a linear upper bound of $n-2$ by using the Two-Extremes mechanism. Later, Lu et al. [2009] improved the deterministic lower bound from $3/2$ to 2, proposed a randomized strategyproof mechanism with approximation ratio of $n/2$ and proved the social cost approximation ratio of any strategyproof randomized mechanism is at least 1.045. Furthermore, Lu et al. [2010] showed an asymptotically linear lower bound of $\Omega(n)$ for deterministic mechanisms, before Fotakis and Tzamos [2014] finally showed that the exact bound is $n-2$. Lu et al. [2010] also proposed the randomized Proportional mechanism which is 4-approximate.

For the maximum cost objective, Procaccia and Tennenholtz [2013] showed that the Two-Extremes mechanism is 2-approximate and that the approximation of any deterministic strategyproof mechanism is at least 2. Further, by using randomization, Procaccia and Tennenholtz proved that the approximation ratio of any randomized strategyproof mechanism is at least $3/2$ and designed a randomized strategyproof mechanism with approximation ratio of $5/3$.

1.2.2 Multiple facilities

Escoffier et al. [2011] focused on locating $n-1$ facilities in general metric spaces and trees by using deterministic and randomized strategyproof mechanisms with respect to the utilitarian and egalitarian objectives. As a result, they showed linear upper bounds and constant lower bounds.

Fotakis and Tzamos [2014] considered *k-facility* location games, where n strategic agents report their locations in a metric space and a mechanism maps them to k facilities. Their main result is an elegant characterization of deterministic strategyproof mechanisms with a bounded approximation ratio for 2-Facility

Location on the line. Then, they showed that for every $k \geq 3$, there do not exist any deterministic anonymous strategyproof mechanisms with a bounded approximation ratio for k-Facility location on the line, even for simple instances with $k+1$ agents. Moreover, building on the characterization for the line, they show that there do not exist any deterministic strategyproof mechanisms with a bounded approximation ratio for 2-Facility location and instances with $n \geq 3$ agents located in a star.

For any given concave cost functions, Fotakis and Tzamos [2016] showed that the randomized group strategyproof mechanism Equal-Cost achieves a bounded approximation ratio for all k and n. In particular, its approximation ratio is at most 2 for the max cost and at most n for the social cost. Their result implies an interesting separation between deterministic mechanisms, whose approximation ratio for max cost jumps from 2 to unbounded when k increases from 2 to 3, and randomized mechanisms, whose approximation ratio remains at most 2 for all k.

1.2.3 Obnoxious Facility Location

The *obnoxious* facility location problem is a classic research direction in the literature on algorithms and optimization which is first studied by Church and Garfinkel [1978]. They dealt with the problem of locating a point on a network so as to maximize the sum of its weighted distances to the nodes.Cheng et al. [2013b] studied this problem on a network (path, circle or tree). Ye et al. [2015] focus on designing strategyproof mechanisms on the real line with two objective functions (maximizing the sum of squares of distances and maximizing the sum of distances). Ibara and Nagamochi [2012] completely characterized the class of (group) strategyproof mechanisms with exactly two candidates in the general metric and showed that there exists a 4-approximation group strategyproof mechanism in any metric space.

1.2.4 Heterogeneous Facility Location

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, whereas a single person might want the opposite. Many settings have been proposed to model the different preferences the agents may have about the facilities.

Dual preferences. 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, 2013a], 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.

Approval preferences. Serafino and Ventre [2016] studied a discrete version, where the line is a discrete graph, the agents occupy nodes of the line (which is common knowledge) and have *approval* preferences over two facilities (that is, an agent either likes a facility or is indifferent about it), which can only be placed at *different* nodes of the line. Given facility locations, the cost of an agent is defined as the total distance from the facilities she approves. Serafino and Ventre presented bounds on the approximation ratio of deterministic and randomized mechanisms in terms of both social objectives of interest. 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 $3/2$. 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.

Chen et al. [2020] studied a continuous setting with agents that have approval preferences over the facilities. The authors consider 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 improve some results of Chen et al.. Deligkas et al. [2023] considered a similar preference model, but with the difference that the goal is to locate just one of the facilities (and, more generally,

k out of m), and not all of them.

Fractional preferences. Fong et al. [2018] proposed a fractional preference model for the facility location game with two facilities that serve the same purpose, where each agent, besides his location information, has also a fractional preference to indicate how well they prefer the facilities. The preference for each facility is in the range $[0, 1]$ such that the sum of the preferences for all facilities is 1. The utility of an agent is a function of the distance of the agent from the facilities multiplied by the preferences. Duan et al. [2021] later generalized the minimum distance setting by allowing for private fractional preferences over the two facilities.

1.2.5 Facility Location with Candidate Locations

Sui and Boutilier [2015] were among the first to consider truthful facility location problems with candidate locations (referred to as constrained facility location), with a focus on achieving approximate strategyproofness by bounding the incentives of the agents to manipulate; for multiple facilities, they considered only the homogeneous case where each agent's individual cost is the distance to the closest facility. Feldman et al. [2016] considered a candidate selection problem with a fixed set of candidates, a model which translates into a single-facility location problem where the facility can only be placed at a location from a given set of discrete candidate locations. They focused on the social cost objective and, among other results, proved that the Median mechanism that places the facility at the location closest to the position reported by the median agent, achieves an upper bound of 3; they also showed that this is the best possible bound among deterministic mechanisms.

Tang et al. [2020b] considered a setting with two facilities, which can be placed at locations chosen from a set of candidate ones, allowing the facilities to be placed even at the same location. The positions of the agents are assumed to be private information and an agent's individual cost is defined as her distance from the closest facility. They proved an upper bound of $2n - 3$ for the social cost objective and a tight bound of 3 for the maximum cost. They also considered the case of a single facility and the max cost objective, for which they showed a bound of 3 (extending the work of Feldman et al. [2016] who only focused on the social cost). Walsh [2021] considered a similar setting, where one or more facilities can only be placed at different subintervals of the line, and showed bounds on the approximation ratio of strategyproof mechanisms for many social objective functions, beyond the classic ones. Zhao et al. [2023] studied a slightly different setting, in which the agents have known approval preferences over the two different facilities and their individual costs are defined as their distance from the farthest facility among the ones they approve. For homogeneous agent preferences, they showed a tight bound of 3 for both the social cost and the max cost objectives, while for general, heterogeneous preferences, they showed an upper bound of $2n + 1$ for the social cost, and an upper bound of 9 for the maximum cost. Their results for general preferences were recently improved by Lotfi and Voudouris [2023] to 11 and 5 for the social cost and the max cost, respectively.

Xu et al. [2021] considered a setting where two facilities must be located so that there is a minimum distance between them (not at specific given candidate locations). They showed results for two types of individual costs. The first one is the total distance (assuming that the facilities play a different role, and thus the agents are interested in both of them) and showed that, for any minimum distance requirement, the optimal solution for the social cost or the maximum cost can be attained by a strategyproof mechanism. The second one is the minimum distance (assuming that the facilities are of the same type, and thus the agents are interested only in their closest one), and showed that the approximation ratio of strategyproof mechanisms is unbounded. They also considered the case where the facility is obnoxious and showed a bound that depends on the minimum distance parameter.

Gai et al. [2022] considered the setting of obnoxious facility location games with candidate locations. For obnoxious single facility location games under social utility objective, they presented a tight bound of 3. For obnoxious heterogeneous two-facility location games, they proved an upper bound of 2 and lower bound of $\frac{3}{2}$ for the social cost objective.

1.2.6 Distributed Facility Location

Filos-Ratsikas and Voudouris [2021] considered a distributed facility location setting in which the agents are partitioned into districts and a mechanism works in two steps: It first chooses a representative location for each district of agents, and then chooses one of the representatives as the final facility location. Filos-Ratsikas and Voudouris focused on the social cost objective and showed a tight bound of 3 among all possible deterministic mechanisms. Similar distributed settings were previously considered in the more general context of utilitarian and metric social choice [Filos-Ratsikas et al., 2020, Anshelevich et al., 2022], aiming to model settings where decisions are first made at a local level, among disjoint sets of agents, and then these decisions are aggregated into a collective outcome, such as in elections.

1.3 Structure of the Thesis

In chapter 3, we revisit the discrete heterogeneous two-facility location problem that was first considered by Serafino and Ventre [2016], 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.

Contribution. 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. In particular, for the social cost, we show an upper bound of $17/4$ and a lower bound of $4/3$ (the lower bound is tight for instances with three agents). For the max cost, we show a tight bound of 2. Our results have been published as [Kanellopoulos et al., 2023b].

In chapter 4, we study a truthful two-facility location problem in which a set of agents have private positions on the line of real numbers and known approval preferences over two facilities. Given the locations of the two facilities, the cost of an agent is the total distance from the facilities she approves. The goal is to decide where to place the facilities from a given finite set of candidate locations so as to (a) approximately optimize desired social objectives, and (b) incentivize the agents to truthfully report their private positions.

Contribution. We focus on the class of deterministic strategyproof mechanisms and pinpoint the ones with approximation ratio in terms of the social cost (i.e., the total cost of the agents) and the max cost. In particular, for the social cost of doubleton instances (in which all agents approve both facilities), we show that the best possible approximation ratio of strategyproof mechanisms is between $1+\sqrt{2}$ and 3. For singleton instances (in which each agent approves one facility), we show a lower bound of 3. Then, for general instances, we show an upper bound of 7. For the max cost of doubleton instances, we show that the best possible approximation ratio is between 2 and 3. We show a tight bound of 3 for singleton and general instances, respectively. Our results have been published as [Kanellopoulos et al., 2023a].

In chapter 5, we study the distributed facility location problem that was first considered by Filos-Ratsikas and Voudouris [2021], where a set of agents with positions on the line of real numbers are partitioned into disjoint districts, and the goal is to choose a point to satisfy certain criteria, such as optimize an objective function or avoid strategic behavior. A mechanism in our distributed setting works in two steps: For each district it chooses a point that is representative of the positions reported by the agents in the district, and then decides one of these representative points as the final output. We consider two classes of mechanisms: Unrestricted mechanisms which assume that the agents directly provide their true positions as input, and strategyproof mechanisms which deal with strategic agents and aim to incentivize them to truthfully report their positions.

Contribution. For both classes, we show tight bounds on the best possible approximation in terms of several minimization social objectives, including the average social cost and the max cost, as well as other fairness-inspired objectives that are tailor-made for the distributed setting, in particular, the max-of-average and the average-of-max. Our results have been published as [Filos-Ratsikas et al., 2024].

1.4 List of Publications

My primary research interest lies in the intersection of Algorithmic Mechanism Design , specifically focusing on mechanisms for distributed and heterogeneous facility location problems without money. This research addresses the design and evaluation of mechanisms that encourage truthful reporting of information while optimizing social welfare objectives. Notable contributions in this field are detailed in a series of papers, where I explore strategic behavior in systems without payments, propose new approximate mechanisms, and provide thorough analyses on these mechanisms.

My contributions on these topics can be best summarized by the following papers[1]:

1. **'On Discrete Truthful Heterogeneous Two-Facility Location'**
   Panagiotis Kanellopoulos, Alexandros A. Voudouris, **Rongsen Zhang**. *In Proceedings of the Thirty-First International Joint Conference on Artificial Intelligence(IJCAI 2022)*

2. **'Settling the Distortion of Distributed Facility Location'**
   Aris Filos-Ratsikas, Panagiotis Kanellopoulos, Alexandros A. Voudouris, **Rongsen Zhang**. *In proceedings of the 2023 International Conference on Autonomous Agents and Multiagent Systems(AAMAS 2023)*

[1] The order of authors is alphabetical.

3. **’Mechanism Design for Heterogeneous and Distributed Facility Location Problems’(Extend Abstract)**

   Rongsen Zhang. *In proceedings of the 2023 International Conference on Autonomous Agents and Multiagent Systems(AAMAS 2023)*

4. **’Truthful Two-Facility Location with Candidate Locations’**

   Panagiotis Kanellopoulos, Alexandros A. Voudouris, **Rongsen Zhang**. *In proceedings of the 16th International Symposium on Algorithmic Game Theory (SAGT 2023)*

5. **’On discrete truthful heterogeneous two-facility location’**

   Panagiotis Kanellopoulos, Alexandros A Voudouris, **Rongsen Zhang**. *Artificial Intelligence 328 (2024) p. 104066. Elsevier, 2024*

6. **’The distortion of distributed facility location’**

   Aris Filos-Ratsikas, Panagiotis Kanellopoulos, Alexandros A Voudouris, **Rongsen Zhang**. *SIAM Journal on Discrete Mathematics 37.2 (2023) pp. 779–799. SIAM, 2023*

Chapter 2

Preliminaries

We consider *facility location problems* in this thesis. An instance I of a facility location problem consists of a set N of $n \geq 2$ *agents* and a set F consisting of at most two *facilities*, f_1 and potentially f_2 as well. Each agent is associated with a position on a line (which can either be a discrete graph or continuous). When $|F| = 2$, the agents may also have *approval preferences* $\mathbf{p}_i \in \{0, 1\}^{|F|}$ over the two facilities: If $p_{ij} = 1$, agent $i \in N$ approves facility $j \in \{1, 2\}$. When $|F| = 1$, we assume that p_{ij}s' are trivially set to 1. Then, we denote by N_j the set of agents that approve facility f_j if $|F| \neq 1$, i.e., $i \in N_j$ if $p_{ij} = 1$. Clearly, the two sets N_1 and N_2 need not be disjoint if there are agents that approve both facilities.

A *feasible solution* is $\mathbf{w} = (w_j)_{j \in |F|}$ determines the location w_j of each facility $f_j \in F$. The locations can be chosen from a set of *candidate locations*; this set can either be finite (i.e., it consists of a given set of points or nodes on the line) or infinite (i.e., it consists of any point on the line of real numbers). When there are two facilities to be located, we also require that they are located at different points, that is, $w_1 \neq w_2$. Given a feasible solution $\mathbf{w}$ for instance I, the cost of any agent i in instance I is the (*total*) distance from her position to the location of the facility or facilities she approves, i.e.,

$$\text{cost}_i(\mathbf{w}|I) = \sum_{j \in F} p_{ij} \cdot d(i, j),$$

where $d(x, y) = |x - y|$ denotes the distance between any two points x and y on the line, $|F|$ denotes the number of facilities in the setting, $|F|$ is at most 2. Since the line is a special metric space, the distances satisfy the *triangle inequality*, which states that $d(x, y) \leq d(x, z) + d(z, y)$ for any three points x, y and z on the line, with the equality being true when $z \in [x, y]$.

A *mechanism* M takes as input an instance I(not only the reported private information, but also the public information), and outputs a feasible solution; we will denote by $M(I)$ the output of M when given information about I as input. To determine the quality of the solutions computed by mechanisms, we consider two natural objective social functions that have been considered extensively within the truthful facility location literature. Given an instance I, the *social cost* of a feasible solution $\mathbf{w}$ is the total cost of all the agents, i.e.,

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

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

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

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

For any social objective $f \in \{\mathrm{SC}, \mathrm{MC}\}$, the f*-approximation ratio* 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.,

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

A mechanism M is said to be *strategyproof* if the solution $M(I)$ it returns when given as input any instance I is such that there is no agent i with incentive to misreport their private information (it will be the preference $\mathbf{p}_i'$ or location $\mathbf{x}_i'$) to decrease her individual cost, that is,

$$\mathrm{cost}_i(M(I)|I) \leq \mathrm{cost}_i(M(\mathbf{x}_i'/\mathbf{p}_i', I_{-i})|I),$$

where I_{-i} is any instance in which all agents besides i report the same private information as in I.

In this thesis we will consider different types of facility location problems with different assumptions about the private information of the agents and how the mechanisms can operate. We will refine the basic definitions presented in the following chapters. In all cases, our goals are to (a) design mechanisms with an as low f-approximation ratio as possible (close to 1), and (b) incentivise the agents truthfully report their private information.

Chapter 3

On Discrete Truthful Heterogeneous Two-Facility Location

3.1 Definitions and notation

We consider the discrete two-facility location problem in this chapter.

Besides the basic condition in chapter 2, the instance I in this chapter consists of a line graph with $m \geq n$ nodes. Hence, each agent occupies a node x_i of the line, such that different agents occupy different nodes. We define y_j is the position of the leftmost median agent in N_j. The position profile $\mathbf{x}$ consists not only of the position of agents but also of the possible empty nodes. The position profile is assumed to be common knowledge and each agent i also has private approval preference. As $\mathbf{x}$ and $\mathbf{p}$ include all the information related to an instance, we denote $I = (\mathbf{x}, \mathbf{p})$.

Given a feasible solution $\mathbf{w} = (w_1, w_2)$, the cost of any agent i in instance I is her total distance from the facilities she approves, i.e.,

$$cost_i(\mathbf{w}|I) = \sum_{j\in\{1,2\}} p_{ij} \cdot d(i,j),$$

A mechanism M is said to be strategyproof in this chapter if the solution $M(I)$ it computes when given as input the instance $I = (\mathbf{x}, \mathbf{p})$ is such that no agent i has incentive to report a false preference $\mathbf{p}_i' \neq \mathbf{p}_i$

$$cost_i(M(I)|I) \leq cost_i(M(\mathbf{x}, (\mathbf{p}_i', \mathbf{p}_{-i}))|I),$$

where $(\mathbf{p}_i', \mathbf{p}_{-i})$ is the preference profile according to which agent i's preference is $\mathbf{p}_i^{'}$, while the preference of any other agent is the same as in $\mathbf{p}$.

3.2 Overview of 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 only 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 shown by Serafino and Ventre for the social and the maximum cost both follow by the same deterministic mechanism, named TwoExtremes, which works along the lines of the mechanism considered by Procaccia and Tennenholtz [2013] for homogeneous 2-facility location. TwoExtremes 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. There are two main reasons for the deficiency of TwoExtremes: (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 take care of these two reasons: We place the facilities closer to median agents (without breaking strategyproofness), and exploit the information about the agent 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 mechanisms, which follows by two instances with only three agents and no empty nodes. Motivated by this lower bound construction, we then focus on instances with three agents, and design the 3-agent Priority-Dictatorship mechanism that achieves the best-possible bound of $4/3$.

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

Table 3.1: 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 shown by Serafino and Ventre [2016]. The lower bound of $4/3$ marked with a $\star$ is tight for instances with three agents.

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], is given in Table 3.1.

3.3 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 $n - 1$ of Serafino and Ventre [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

Mechanism 1: FIXED-OR-MEDIAN-NEAREST-EMPTY (FMNE)

Input: Instance I with n agents ;
Output: Feasible solution $\mathbf{w} = (w_1, w_2)$;
if there are no empty nodes **then** // FIXED part
  $w_1 \leftarrow \lfloor n/2 \rfloor$;
  $w_2 \leftarrow \lfloor n/2 \rfloor + 1$;
else // MEDIAN-NEAREST-EMPTY part
  for $j \in \{1, 2\}$ **do**
    $y_j \leftarrow$ position of the leftmost median agent in N_j;
  $w_1 \leftarrow y_1$;
  $w_2 \leftarrow$ nearest empty node to y_2, breaking ties in favor of the rightmost such empty node;

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 Mechanism 1.

Theorem 3.3.1. *FMNE is strategyproof.*

Proof. Consider an arbitrary instance I. 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 $y'_1 = y_1$.

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. □

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 3.3.2 ([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*

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

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 facility 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 $\frac{n_j^2}{4}$. □

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 3.3.3. *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 $|n/2+1-i|$. 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 $\mathbf{w}$ computed by the mechanism is bounded by

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

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

By Lemma 3.3.2, $\text{SC}^*(I) \geq \frac{1}{4}\left(n_1^2 + n_2^2 - 2\right)$, and thus the approximation ratio is bounded by

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

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 $\mathbf{w}$ computed by the mechanism is bounded by

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

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

By Lemma 3.3.2, $\mathrm{SC}^*(I) \geq \frac{1}{4}\left(n_1^2 + n_2^2 - \mathbf{1}\left\{n_1 \text{ odd}\right\} - \mathbf{1}\left\{n_2 \text{ odd}\right\}\right)$, and thus the approximation ratio is bounded by

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

To prove the claim it suffices to show that $(n_1 - n_2)^2 + n_1^2 + n_2^2 + b^2 + 4b \geq 2n_1 + 2n_1 + 1 + 3(\mathbf{1}\left\{n_1 \text{ odd}\right\} + \mathbf{1}\left\{n_2 \text{ odd}\right\})$. 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 3.3.4. *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 $\max\{f(3,3), f(3,4)\} = \max\{13/4, 73/23\} \leq 13/4$. □

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

Theorem 3.3.5. *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

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

Now, let us focus on the social cost of the solution $\mathbf{w}$ 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 $\mathbf{w} = (y_1, x_e)$, and

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

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). Observe that

- For every agent $i \in M$, there exists a unique agent in $j \in L$ such that

$$\begin{aligned} d(x_j, x_e) &= d(x_j, x_i) + d(x_i, x_e) \\ &= d(x_j, y_2) + d(x_i, y_2) + d(x_i, x_e). \end{aligned}$$

- For every agent $i \in R$, there exists a unique agent $j \in L$ such that

$$\begin{aligned} d(x_j, x_e) + d(x_i, x_e) &= d(x_j, y_2) + d(y_2, x_e) + d(x_i, x_e) \\ &= d(x_j, y_2) + d(x_i, y_2). \end{aligned}$$

Hence, we have that

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

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:

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

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

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

By putting everything together, we have

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

By Lemma 3.3.2, we have $\text{SC}^*(I) \geq \frac{1}{4}\left(n_1^2 + n_2^2 - 2\right)$, and thus the approximation ratio is bounded by

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

The bound of $17/4$ follows by applying Lemma 3.3.4 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 3.3.6. *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 $17/4$. □

3.4 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 Ventre [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.

Theorem 3.4.1. *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 $w_1 = (1, 3)$ and $w_2 = (2, 3)$ have social cost 4 in I_2, whereas an optimal solution (such as $w^* = (1, 2)$) has social cost 3; the theorem follows. □

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 $x_\ell < x_c < x_r$. Without loss of generality, we assume that c is closer to r than to ℓ, that is, $x_r - x_c \leq x_c - x_\ell$. 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 Mechanism 2 for a formal description. We first show that the mechanism is strategyproof.

Theorem 3.4.2. 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

Mechanism 2: PRIORITY-DICTATORSHIP

```
Input: Instance I with three agents ℓ, c, and r such that x_ℓ < x_c < x_r ;
Output: Feasible solution w ;
if c ∈ N_1 \ N_2 then
    if r ∈ N_2 then
        w ← (x_c, x_r);
    else
        w ← (x_c, x_ℓ);
else if c ∈ N_2 \ N_1 then
    if r ∈ N_1 then
        w ← (x_r, x_c);
    else
        w ← (x_ℓ, x_c);
else
    if r ∈ N_2 then
        w ← (x_c, x_c + 1);
    else
        w ← (x_c + 1, x_c);
```

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. □

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

Theorem 3.4.3. *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 $\alpha = \beta = 1$.

The proof is now complete. □

3.5 Maximum cost

We now turn our attention to the maximum cost. For this objective, Serafino and Ventre [2016] showed an upper bound of 3 on the approximation ratio of the TwoExtemes 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.

3.5.1 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, 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

Mechanism 3: α-LEFT-RIGHT

Input: Instance I with n agents;
Output: Feasible solution $\mathbf{w} = (w_1, w_2)$;
$L \leftarrow$ 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**
  $w_1 \leftarrow$ median node of line defined by $N(X)$ (ties in favor of nodes farther from α);
  $w_2 \leftarrow$ median node 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**
    $X \leftarrow L$;
  $w_\ell \leftarrow$ median node of line defined by $N(X)$ (ties in favor of nodes farther from α);
  $w_{3-\ell} \leftarrow \beta \in \{\alpha, \alpha + 1\} \setminus X$;
// (case 3): Both parts include agents from N_1 and N_2
else
  $w_1 \leftarrow$ rightmost node of L;
  $w_2 \leftarrow$ leftmost node of R;

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 Mechanism 3 for a formal description. We first show that every α-LEFT-RIGHT is strategyproof.

Theorem 3.5.1. *For any $\alpha \in \{1, \ldots, m-1\}$, 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 $N(Y)$. 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 cases 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), and 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. □

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 3.5.2. *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

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

(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

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

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 $a+b-2x$, while the cost of the agent at node y is $2y-a-b$. 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 $a+b-2x$, the cost of the agent at y is $y-a$, while the cost of the agent at z is $z-b$. As the sum of costs equals $y+z-2x$, 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 3.5.3. *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 $\mathbf{w}$ 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, $\text{MC}(\mathbf{w}|I) \leq \frac{m}{2}$. 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 3.5.2(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 $\text{MC}(\mathbf{w}|I) = |x - y_L| + \frac{m}{2} + 1 - x$.

  If $x > y_L$, we have $\text{MC}(\mathbf{w}|I) = \frac{m}{2} + 1 - y_L \leq \frac{m}{2}$. 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 $\text{MC}(\mathbf{w}|I) = \frac{m}{2} + 1 + y_L - 2x \leq \frac{3(m+2)}{4} - 2x$. Since agent i and the agent at node m both approve facility 2, by Lemma 3.5.2(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, $\text{MC}(\mathbf{w}|I) \leq \frac{m}{2}$ (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 3.5.2(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, $\text{MC}(\mathbf{w}|I) = \frac{m}{2} - x + \frac{m}{2} + 1 - x = m + 1 - 2x$. As the agent at node m approves some facility that is also approved by i, by Lemma 3.5.2(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. □

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 3.5.4. *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 $\mathbf{w}$ 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 3.5.2(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 $\text{MC}(\mathbf{w}|I) \leq \max\left\{\frac{m+1}{4}, \frac{m-3}{2}\right\}$. Since

$\mathrm{MC}^*(I) \geq \frac{m-1}{4}$, 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 a most $\frac{m+3}{2} - 1 = \frac{m+1}{2}$, and thus $\mathrm{MC}(\mathbf{w}|I) \leq \max\left\{\frac{m+1}{4}, \frac{m+1}{2}\right\} = \frac{m+1}{2}$. Since $\mathrm{MC}^*(I) \geq \frac{m-x}{2} \geq \frac{m+1}{4}$, 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 $\mathrm{MC}(\mathbf{w}|I) = |x - y_L| + \frac{m+3}{2} - x$.

  If $x \leq y_L$, then since $y_L \leq \frac{m+3}{4}$, we have that $\mathrm{MC}(\mathbf{w}|I) = \frac{m+3}{2} + y_L - 2x \leq \frac{3(m+3)}{4} - 2x$. As node m is occupied by an agent that approves facility 2, by Lemma 3.5.2(a), we have that $\mathrm{MC}^*(I) \geq \frac{m-x}{2}$, 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 $\mathbf{w}$ is 2, while $\mathrm{MC}^*(I) \geq 1$, 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 $\mathrm{MC}^*(I) \geq 1$, and so the approximation ratio is at most 2. Hence, assume that $y_L \geq 2$. Then, we have $\mathrm{MC}(\mathbf{w}|I) = \frac{m+3}{2} - y_L \leq \frac{m-1}{2}$. Since $x \leq \frac{m+1}{2}$ and the agent at node m approves facility 2, by Lemma 3.5.2(a), we have that $\mathrm{MC}^*(I) \geq \frac{m-1}{4}$, 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, $MC(\mathbf{w}|I) \leq \frac{m+3}{2} - 1 = \frac{m+1}{2}$. 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 3.5.2(a) gives $\text{MC}^*(I) \geq \frac{y-1}{2} \geq \frac{m+1}{4}$, 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, $\text{MC}(\mathbf{w}|I) = \frac{m+1}{2} - x + \frac{m+3}{2} - x = m - 2x + 2$. 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 3.5.2(a), $\text{MC}^*(I) \geq y - x \geq \frac{m+3}{2} - x$, 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 $y \in \left[\frac{m+3}{2}, m\right)$ that approves facility 1. If $x = 1$, then $\text{MC}(\mathbf{w}|I) \leq m$ and, by Lemma 3.5.2(b), $\text{MC}^*(I) \geq \left\lceil\frac{y+m-2}{3}\right\rceil \geq \left\lceil\frac{m}{2} - \frac{1}{6}\right\rceil = \frac{m+1}{2}$; hence, the approximation ratio is at most 2. If $x \geq 2$, it suffices to use the bound $\text{MC}^*(I) \geq \frac{m-x}{2}$ implied by Lemma 3.5.2(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. □

3.5.2 A tight lower bound

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

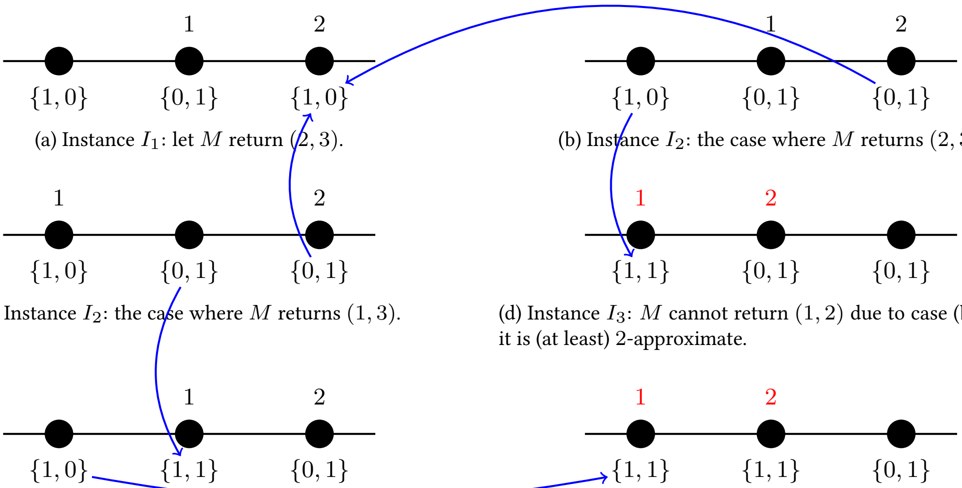
Proof. Suppose that there exists a strategyproof mechanism M with approximation ratio strictly smaller than 2. We will reach a contradiction by examining a series of instances, all of which involve three agents and no empty nodes; see also Figure 3.1.

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 $\text{MC}((2, 3)|I_1) = \text{MC}((2, 1)|I_1) = 1$; 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. Without loss of generality, let us assume that M returns the solution $(2, 3)$.

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 $\mathrm{MC}((1,2)|I_3) = 1$, 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.

Otherwise, when 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 $\mathrm{MC}((2,3)|I_4) = \mathrm{MC}((1,2)|I_4) = 1$; 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 $\mathrm{MC}((1,2)|I_5) = 1$; 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. $\square$



(a) Instance I_1: let M return $(2, 3)$.

(b) Instance I_2: the case where M returns $(2, 3)$.

(c) Instance I_2: the case where M returns $(1, 3)$.

(d) Instance I_3: M cannot return $(1, 2)$ due to case (b); hence, it is (at least) 2-approximate.

(e) Instance I_4: M must return either $(1, 2)$ or $(2, 3)$ since it is better than 2-approximate. It cannot, however, return $(1, 2)$ due to case (c).

(f) Instance I_5: M cannot return $(1, 2)$ due to case (e); hence, it is (at least) 2-approximate.

Figure 3.1: The instances used in the proof of Theorem 3.5.5. Each instance has 3 agents and no empty nodes. The agent preferences appear below each node, while the facility assignment appears above the nodes. A black font denotes the mechanism's assignment, while a red font denotes an optimal but excluded assignment. Blue arrows denote how instances are related when a single agent's preferences change.

Chapter 4

Heterogeneous Two-Facility Location with Candidate Locations

4.1 Definitions and notation

We consider the two-facility location problem with candidate locations in this chapter.

Besides the private position profile $\mathbf{x}$ and known approval preference profile $\mathbf{p}$,there is a set of $m \geq 2$ candidate locations C where the facilities can be located in this chapter. To be concise, we denote an instance using the tuple $I = (\mathbf{x}, \mathbf{p}, C)$.

A *feasible solution* in this chapter is a pair $\mathbf{w} = (w_1, w_2) \in C^2$ of candidate locations with $w_1 \neq w_2$, where the two facilities can be placed; that is, for each $j \in [2]$, F_j is placed at w_j. Given a feasible solution $\mathbf{w}$, the cost of any agent i in instance I is her total distance from the facilities she approves, i.e.,

$$cost_i(\mathbf{w}|I) = \sum_{j\in\{1,2\}} p_{ij} \cdot d(x_i, w_j),$$

A mechanism is said to be *strategyproof* if the solution $M(I)$ it returns when given as input any instance $I = (\mathbf{x}, \mathbf{p}, C)$ is such that there is no agent i with incentive to misreport a position $x_i' \neq x_i$ to decrease her individual cost, that is,

$$cost_i(M(I)|I) \leq cost_i(M((x_i', \mathbf{x}), \mathbf{p}, C)|I),$$

where $(x_i', \mathbf{x}_{-i})$ is the position profile obtained by $\mathbf{x}$ when only agent i reports a different position x_i'.

Finally, let us introduce some further notation and terminology that will be useful. Any agent that approves both facilities belongs to the intersection $N_1 \cap N_2$ and has a *doubleton preference*. Any agent that approves one facility belongs to either $N_1 \setminus N_2$ or $N_2 \setminus N_1$ and has a *singleton preference*. Besides *general* instances (with agents that have any type of approval preferences), we will also pay particular attention to the following two classes of instances:

- *Doubleton:* All agents have a doubleton preference, that is, $N_1 \cap N_2 = N$;
- *Singleton:* All agents have a singleton preference, that is, $N_1 \cap N_2 = \varnothing$.

We will also denote by m_j, ℓ_j, and r_j the median[1], leftmost, and rightmost, respectively, agent in N_j. In addition, for any agent i we denote by $t(i)$ and $s(i)$ the closest and the second closest, respectively, candidate location to i.

[1]Without loss of generality, we break potential ties in favor of the leftmost median agent.

	Social cost	Max cost
Doubleton	$[1+\sqrt{2}, 3]$	$[2, 3]$
Singleton	3	3
General	$[3, 7]$	3

Table 4.1: An overview of the bounds that we show in this paper on the approximation ratio of deterministic strategyproof mechanisms for the different combinations of social objectives functions (social cost and max cost) and agent preferences (doubleton, singleton, or general).

4.2 Overview of contribution

Our goal is to design mechanisms that take as input the positions reported by the agents, and, using also the available information about the preferences of the agents, decide where to place the two facilities, so that (a) a social objective function is (approximately) optimized, and (b) the agents are incentivized to truthfully report their positions. As in previous work, we consider the well-known *social cost* (the total individual cost of the agents) and the *max cost* (the maximum individual cost over all agents) as our social objective functions. We treat separately the class of instances in which all agents approve both facilities (to which we refer as *doubleton*), the class of instances in which all agents approve one facility (to which we refer as *singleton*), and the general class of all possible instances. For all possible combinations of objectives and types of preferences, we design *deterministic* strategyproof mechanisms with small, constant approximation ratios. An overview of our results is given in Table 4.1.

In Section 4.3 we consider the social cost and show the following results:

- For doubleton instances (in which all agents approve both facilities), we show that the best possible approximation ratio of strategyproof mechanisms is between $1+\sqrt{2}$ and 3. Our upper bound follows by a mechanism, which places the facilities at the two candidate locations closest to the median agent; this is the natural extension of the MEDIAN mechanism which achieves the best possible approximation ratio of 3 for the single-facility location problem [Feldman et al., 2016]. These results can be found in Section 4.3.1.

- For singleton instances (in which each agent approves one facility), we first observe that no strategyproof mechanism can achieve an approximation ratio better than 3; this follows from the fact that the problem is now a generalization of the single-facility location problem. The main technical difficulty, which does not allow us to simply treat a singleton instance as two separate single-facility location problems (one for each facility), is that the facilities cannot be placed at the same location. We circumvent this difficulty and show a tight upper bound of 3 by considering a mechanism that places each facility at the available candidate location closest to the median agent among those that approve it. To decide the order in which the facilities are placed, we first perform a voting step that allows the agents that approve each facility to decide if they prefer the closest or second-closest candidate location to the respective median agent; this is necessary since just blindly choosing the order of placing the facilities leads to a mechanism with a rather large approximation ratio. These results are presented in Section 4.3.2.

- For general instances, we show an upper bound of 7 by considering a mechanism which switches between two cases depending on the cardinalities of the sets of agents with different preferences. In particular, when there is a large number of agents that approve both facilities, we run the simple median mechanism we used for doubleton instances by ignoring the other agents. Otherwise, we run a mechanism that places the facility that is approved by most agents at the location closest to the median of the agents that approve *only* it, while the other facility is placed at the available location that is closest to the median of the agents that approve it. These results are presented in Section 4.3.3. Our bound of 7 for general instances significantly improves upon the bound of 22 that Lotfi and Voudouris [2023] showed via a reduction between the model in which the individual cost of an agent is the distance to the farthest facility among the ones she approves and our model (in which the individual cost of an agent is the distance to all the facilities she approves).

In Section 4.4, we turn our attention to the max cost objective and show the following results:

- For doubleton instances, we show that the best possible approximation ratio is between 2 and 3. Our upper bound follows by a simple mechanism that places the facilities at the available candidate locations closest to the leftmost agent; see Section 4.4.1.
- For singleton instances, we show a tight bound of 3 by considering a mechanism that places the two facilities at the candidate locations closest to some agents among those that approve. The main difficulty here is to decide which agents to pick. In particular, after placing the first facility at the candidate location closest to one of the agents that approve it (such as the leftmost), we then need to dynamically decide whether the second facility can be placed closer to the leftmost or rightmost among the agents that approve it, or neither of them. This again is done by a voting-like procedure that is used to decide the order of the agents that approve the second facility relative to the two candidate locations that are closest to where the first facility has been placed.

- For general instances, we show a tight bound of 3 by splitting the class of all instances into those that consist of at least one agent that approves both facilities (in which case we employ the mechanism for doubleton instances) and the remaining instances which are singleton (and we employ the corresponding mechanism).

Finally, in Section 4.5, we consider a slightly simpler model in which the two facilities are allowed to be placed at the same candidate location. For this model, we manage to show improved, tight bounds on the approximation ratio of deterministic mechanisms for doubleton and general instances for both the social and the max cost (the problem is not interesting for singleton instances). This is possible because we can now avoid possible misreports by agents with doubleton preferences, which in turn allows us to consider a class of mechanisms that is not strategyproof when the facilities are constrained to be placed at different locations.

4.3 Social cost

In this section we will focus on the social cost. We will show that the best possible approximation ratio of strategyproof mechanisms is between $1+\sqrt{2}$ and 3 for doubleton instances, exactly 3 for singleton instances, and between 3 and 7 for general instances.

4.3.1 Doubleton instances

We start with the case of doubleton instances in which all agents approve both facilities. Recall that for the single-facility location problem, Feldman et al. [2016] showed that the best possible approximation ratio of 3 is achieved by the Me-

DIAN mechanism, which places the facility at the candidate location closest to the position reported by the median agent m. We can generalize this mechanism by placing the two facilities at the two candidate locations that are closest to the position reported by m; that is, F_1 is placed at $w_1 = t(m)$ and F_2 is placed at $w_2 = s(m)$; see Mechanism 4. It is not hard to show that this is a strategyproof mechanism; the median agent minimizes her cost and any other agent would have to become the median agent to manipulate the outcome which could only lead to placing the facilities farther away. We next show that the mechanism achieves an approximation ratio of at most 3, but cannot do better; we remark that this result has also been independently shown by Gai et al. [2024] when all agents are of type-II in their model.

Mechanism 4: MEDIAN

Input: Reported positions of agents with doubleton preferences;
Output: Facility locations $\mathbf{w} = (w_1, w_2)$;
$m \leftarrow$ median agent in $N_1 \cap N_2$;
$w_1 \leftarrow t(m)$;
$w_2 \leftarrow s(m)$;

Theorem 4.3.1. *For doubleton instances, the approximation ratio of the* MEDIAN *mechanism is at most* 3*, and this is tight.*

Proof. Let $\mathbf{o} = (o_1, o_2)$ be an optimal solution. Since the position of the median agent minimizes the total distance of all agents, we have that

$$\sum_{i \in N} d(i, m) \leq \sum_{i \in N} d(i, x)$$

for any point x of the line (including o_1 and o_2), and thus

$$2 \sum_{i \in N} d(i, m) \leq \sum_{i \in N} d(i, o_1) + \sum_{i \in N} d(i, o_2) = \text{SC}(\mathbf{o}).$$

Also, since $t(m)$ and $s(m)$ are the two closest candidate locations to m, we have that $d(m, t(m)) \leq d(m, x)$ for any candidate location x, and there exists $o \in \{o_1, o_2\}$ such that $d(m, s(m)) \leq d(m, o)$; let $\tilde{o} \in \{o_1, o_2\} \setminus \{o\}$. Therefore, using these facts and the triangle inequality, we obtain

$$\begin{aligned}
\text{SC}(\mathbf{w}) &= \sum_{i \in N} \Big(d(i, t(m)) + d(i, s(m)) \Big) \\
&\leq 2 \sum_{i \in N} d(i, m) + \sum_{i \in N} d(m, t(m)) + \sum_{i \in N} d(m, s(m)) \\
&\leq \text{SC}(\mathbf{o}) + \sum_{i \in N} d(m, \tilde{o}) + \sum_{i \in N} d(m, o) \\
&\leq \text{SC}(\mathbf{o}) + 2 \sum_{i \in N} d(i, m) + \sum_{i \in N} d(i, \tilde{o}) + \sum_{i \in N} d(i, o) \\
&\leq 3 \cdot \text{SC}(\mathbf{o}).
\end{aligned}$$

The analysis of the mechanism is tight due to the following instance: There are four candidate locations at 0, ε, $1 - \varepsilon$, and 1, for some infinitesimal $\varepsilon > 0$. There are also two agents positioned at $1/2 - \varepsilon$ and 1, respectively. Let the first agent be the median one (in case the second agent is the median, there is a symmetric instance). Then, the two facilities are placed at 0 and ε for a social cost of approximately 3, whereas the optimal solution is to place the facilities at $1 - \varepsilon$ and 1 for a social cost of approximately 1, leading to a lower bound of nearly 3. $\square$

We next show a lower bound of $1 + \sqrt{2}$ on the approximation ratio of any strategyproof mechanism.

Theorem 4.3.2. *For doubleton instances, the approximation ratio of any strategyproof mechanism is at least* $1 + \sqrt{2} - \delta$, *for any* $\delta > 0$.

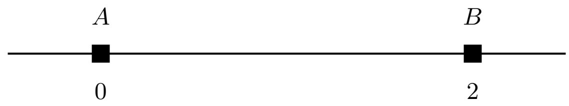
Proof. Let $\varepsilon > 0$ be an infinitesimal. We will consider instances with four candidate locations, two in the ε-neighborhood of 0 (for example, $-\varepsilon$ and ε) and two in the ε-neighborhood of 2 (for example, $2-\varepsilon$ and $2+\varepsilon$). To simplify the calculations in the remainder of the proof, we will assume that there can be candidate locations at the same point of the line, so that we have two candidate locations at 0 and two at 2.

First, consider the following generic instance I with the aforementioned candidate locations: There is at least one agent at 0, at least one agent at 2, while each remaining agent is arbitrarily located at a location from $\{0, 1-\varepsilon, 1+\varepsilon, 2\}$. We make the following observation: Any solution returned by a strategyproof mechanism when given as input I must also be returned when given as input any of the following two instances:

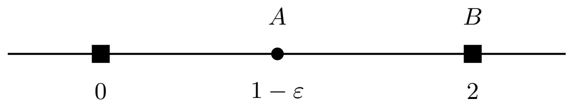
- J_1: Same as I with the difference that an agent j_1 has been moved from 0 to $1-\varepsilon$.
- J_2: Same as I with the difference that an agent j_2 has been moved from 2 to $1+\varepsilon$.

Suppose towards a contradiction that this is not true for J_1; similar arguments can be used for J_2. We consider the following cases:

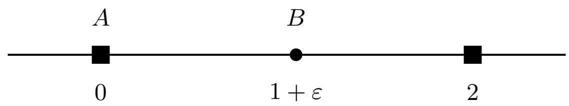
- Both facilities are placed at 2 in I. If this is not done in J_1, then j_1 can misreport her position as $1-\varepsilon$ in I so that the instance becomes J_1 and at least one facility moves to her true position 0.
- Both facilities are placed at 0 in I. If this is not done in J_1, then j_1 can misreport her position as 0 in J_1 so that the instance becomes I and both facilities move to 0 which is closer to her true position $1-\varepsilon$, a contradiction.



(a) The initial instance I_1.



(b) The instance obtained after moving the agents in A from 0 to $1 - \varepsilon$ in Case 1 and Case 3.



(c) The instance obtained after moving the agents in B from 2 to $1 + \varepsilon$ in Case 2.

Figure 4.1: The instances used in the proof of the lower bound of $1 + \sqrt{2}$ in terms of the social cost for doubleton instances (Theorem 4.3.2). Set A consists of αn agents and set B consists of $(1 - \alpha)n$ agents; all of them approve both facilities. Rectangles represent candidate locations; recall that we assume that there are two candidate locations arbitrarily close to 0 and two candidate locations arbitrarily close to 2.

- One facility is placed at 0 and the other is placed at 2 in I. Observe that it cannot be the case that both facilities are placed at 0 in J_1 since that would mean that j_1 can misreport her position in I as $1 - \varepsilon$ so that the instance becomes J_1 and both facilities move to her true position 0. So, the only possibility of having a different solution in I and J_1 is that both facilities are placed at 2 in J_1. But then, j_1 can misreport her true position as 0 in J_1 so that the instance becomes I and one of the facilities moves to 0 which is closer to her true position.

Hence, the same solution must be computed by the mechanism when given I or J_1 as input.

Now, consider an arbitrary strategyproof mechanism and let $\alpha = \sqrt{2} - 1$; note that α is such that $\frac{1+\alpha}{1-\alpha} = \frac{1}{\alpha} = 1 + \sqrt{2}$. Let I_1 be the following instance with the aforementioned candidate locations: αn agents are at 0 and $(1 - \alpha)n$ agents are at 2. See Figure 4.1a. We consider the following cases depending on the solution returned by the mechanism when given I_1 as input:

Case 1: The mechanism places both facilities at 0. We consider the sequence of instances obtained by moving one by one the αn agents that are positioned at 0 in I_1 to $1-\varepsilon$; see Figure 4.1b. By the observation above, the mechanism must return the same solution for any two consecutive instances of this sequence (essentially, the first one is of type I and the second one is of type J_1), which means that the mechanism must eventually return the same solution for all of them. Therefore, the mechanism must place both facilities at 0 in the last instance of this sequence, where αn agents are at $1-\varepsilon$ and the remaining $(1-\alpha)n$ agents are at 2. This solution has social cost $2\alpha n + 4(1-\alpha)n = 2(2-\alpha)n$. However, the solution that places both facilities at 2 has social cost $2\alpha n$, leading to an approximation ratio of $\frac{2}{\alpha} - 1 > 1 + \sqrt{2}$.

Case 2: The mechanism places both facilities at 2. Similarly to Case 1 above, we now consider the sequence of instances obtained by moving one by one the $(1-\alpha)n$ agents that are positioned at 2 in I_1 to $1+\varepsilon$; see Figure 4.1c. Again, by the observation above, the mechanism must return the same solution for any two consecutive instances of this sequence (the first one is of type I and the second one is of type J_2), which means that the mechanism must eventually return the same solution for all of them. Therefore, the mechanism must place both facilities at 2 in the last instance of this sequence, where αn agents are at 0 and the remaining $(1-\alpha)n$ agents are at $1+\varepsilon$. This solution has social cost $4\alpha n + 2(1-\alpha)n = 2(1+\alpha)n$. However, the solution that places both facilities at 0 has social cost $2(1-\alpha)n$, leading to an approximation ratio of $\frac{1+\alpha}{1-\alpha} = 1 + \sqrt{2}$.

Case 3: The mechanism places one facility at 0 and the other at 2. We consider the same sequence of instances as in Case 1. This results in that the mechanism must place one facility at 0 and the other at 2 when given as input the instance where αn agents are at $1-\varepsilon$ while the remaining $(1-\alpha)n$ agents are at 2. This solution has social cost $2\alpha n + 2(1-\alpha)n = 2n$. However, the solution that places both facilities at 2 has social cost $2\alpha n$, leading to an approximation ratio of $\frac{1}{\alpha} = 1+\sqrt{2}$. □

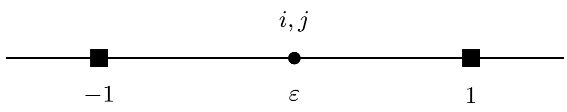
4.3.2 Singleton instances

It is not hard to observe that our two-facility problem with singleton instances is more general than the single-facility location problem studied by Feldman et al. [2016]; indeed, there are singleton instances in which all agents approve the same facility, and thus the location of the other facility does not affect the social cost nor the approximation ratio. Consequently, we cannot hope to achieve an approximation ratio better than 3. For completeness, we include here a slightly different proof of the lower bound of 3 for all strategyproof mechanisms with instances that involve agents that approve different facilities. Recall that, for singleton instances, $N_1 \cap N_2 = \varnothing$.

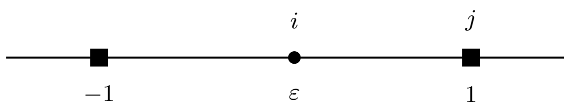
Theorem 4.3.3. *For singleton instances, the approximation ratio of any strategyproof mechanism is at least $3-\delta$, for any $\delta > 0$.*

Proof. Let $\varepsilon > 0$ be an infinitesimal and consider an instance I_1 with two candidate locations at -1 and 1, and two agents positioned at $\varepsilon > 0$ such that one of them approves F_1 while the other approves F_2; see Figure 4.2a. There are two possible solutions, $(-1, 1)$ or $(1, -1)$. Without loss of generality, suppose that $(1, -1)$ is the solution chosen by an arbitrary strategyproof mechanism.

Next, consider instance I_2, which is the same as I_1, with the only difference that the agent that approves F_2 is moved from ε to 1; see Figure 4.2b. To maintain strategyproofness, the solution $(1, -1)$ must be returned in I_2 as well; otherwise,



(a) Instance I_1.



(b) Instance I_2.

Figure 4.2: The two instances used in the proof of the lower bound of 3 in terms of the social cost for the general case (Theorem 4.3.3). Agent i approves F_1 and agent j approves F_2. Rectangles represent candidate locations.

the moving agent would have decreased her cost in I_1 from $1 + \varepsilon$ to $1 - \varepsilon$. This solution has social cost $1 - \varepsilon + 2 = 3 - \varepsilon$, whereas the other solution $(-1, 1)$ has social cost just $1 + \varepsilon$, leading to a lower bound of $3 - \delta$, for any $\delta > 0$. □

Since there is an adaptation of the MEDIAN mechanism that achieves an approximation ratio of 3 for doubleton instances (see Theorem 4.3.1), one might wonder if there is a variant that can do so for singleton instances as well. In particular, the natural extension of MEDIAN is to place F_1 at the candidate location closest to the (leftmost) median agent m_1 of N_1, and F_2 at the available candidate location closest to the (leftmost) median agent m_2 of N_2. While this seems like a good idea at first glance, the following example shows that it fails to achieve the desired approximation ratio bound.

Example 4.3.4. Consider an instance with two candidate locations at 0 and 2. For some $x \geq 1$, there are $2x + 1$ agents that approve only F_1 such that $x + 1$ of them are located at $1 - \varepsilon$ and the other x are located at 2. There are also $2x + 1$ agents that approve only F_2 and are all located at 0. According to the definition of the mechanism, F_1 is placed at 0 (which is the candidate location closest to the median agent in N_1), and then F_2 is placed at 2 as 0 is now occupied and 2 is available. This solution has social cost approximately $(x + 2x) + 4x = 7x$, whereas the solution that places F_1 at 2 and F_2 at 0 has social cost approximately

x, leading to an approximation ratio of nearly 7.

The issue with the aforementioned variant of the Median mechanism is the order in which it decides to place the facilities. If it were to place F_2 first and F_1 second then it would have made the optimal choice in the example. However, there is a symmetric example that would again lead to a lower bound of approximately 7. So, the mechanism needs to be able to dynamically determine the order in which it places F_1 and F_2. This brings us to the following idea: We will again place each facility one after the other at the closest candidate location to the median among the agents that approve it. However, the facility that is placed first (and thus has priority in case the median agents of N_1 and N_2 are closer to the same candidate location) is the one with stronger majority in terms of the number of agents that approve it who are closer to the top choice of the median agent rather than her second choice; ties are broken in favor of the facility that is approved by most agents, which is assumed to be F_1 without loss of generality. We refer to this mechanism as Proportional-Majority-Median; see Mechanism 5 for a more formal description.

Mechanism 5: Proportional-Majority-Median

Input: Reported positions of agents with singleton preferences;
Output: Facility locations $\mathbf{w} = (w_1, w_2)$;
for $j \in [2]$ **do**
  $m_j \leftarrow$ median agent in N_j;
  $S_j \leftarrow$ set of agents in N_j (weakly) closer to $t(m_j)$ than to $s(m_j)$;
if $2|S_1| - |N_1| \geq 2|S_2| - |N_2|$ **then**
  $j \leftarrow 1$;
else
  $j \leftarrow 2$;
$w_j \leftarrow t(m_j)$;
if $t(m_{3-j})$ is available **then**
  $w_{3-j} \leftarrow t(m_{3-j})$;
else
  $w_{3-j} \leftarrow s(m_{3-j})$;

We first show that this mechanism is strategyproof; it is not hard to observe that this must be true as the mechanism is a composition of variants of two simple strategyproof mechanisms (median plus majority voting).

Theorem 4.3.5. Proportional-Majority-Median *is strategyproof.*

Proof. Clearly, if $t(m_1) \neq t(m_2)$ then no agent has incentive to deviate as then the facilities are placed at $t(m_1)$ and $t(m_2)$ independently of whether $2|S_j| - |N_j| \geq 2|S_{3-j}| - |N_{3-j}|$ or not for $j \in [2]$. So, it suffices to consider the case where $t(m_1) = t(m_2)$ and $2|S_j| - |N_j| \geq 2|S_{3-j}| - |N_{3-j}|$ for some $j \in [2]$, leading to $w_j = t(m_j)$ and $w_{3-j} = s(m_{3-j})$, solving ties in favor of F_1.

- m_j and any agent $i \in N_j$ that is closer to $t(m_j)$ than to $s(m_j)$ have no incentive to deviate as $t(m_j)$ is the best choice for them.
- Any agent $i \in N_j$ that is closer to $s(m_j)$ than to $t(m_j)$ has no incentive to deviate, as going closer to $t(m_j)$ can only increase the quantity $2|S_j| - |N_j|$ and cannot change the outcome.
- m_{3-j} and any agent $i \in N_{3-j}$ that is closer to $t(m_{3-j})$ than to $s(m_{3-j})$ have no incentive to deviate as moving closer to $s(m_{3-j})$ would decrease the quantity $2|S_{3-j}| - |N_{3-j}|$ and would not change the outcome.
- Any agent $i \in N_{3-j}$ that is closer to $s(m_{3-j})$ than to $t(m_{3-j})$ has no incentive to deviate as $s(m_{3-j})$ is the best choice for her.

So, the mechanism is strategyproof. □

Next, we show the upper bound of 3 on the approximation ratio.

Theorem 4.3.6. *The approximation ratio of* Proportional-Majority-Median *is at most* 3.

Proof. Let $\mathbf{o} = (o_1, o_2)$ be an optimal solution; without loss of generality, we can assume that $w_1 < w_2$ and $o_1 < o_2$. We consider the following two cases:

Case 1: $t(m_1) \neq t(m_2)$. Then, we have that $w_1 = t(m_1)$ and $w_2 = t(m_2)$. By the properties of the median, for any $j \in [2]$, we have that

$$\sum_{i \in N_j} d(i, m_j) \leq \sum_{i \in N_j} d(i, x)$$

for any point x of the line, including o_j. Also, by the definition of $t(m_j)$, we have that $d(m_j, t(m_j)) \leq d(m_j, x)$ for any candidate location x, again including o_j. Therefore, using these facts and the triangle inequality, we obtain

$$\begin{aligned}
\text{SC}(\mathbf{w}) &= \sum_{j \in [2]} \sum_{i \in N_j} d(i, t(m_j)) \\
&\leq \sum_{j \in [2]} \sum_{i \in N_j} d(i, m_j) + \sum_{j \in [2]} \sum_{i \in N_j} d(m_j, t(m_j)) \\
&\leq \sum_{j \in [2]} \sum_{i \in N_j} d(i, m_j) + \sum_{j \in [2]} \sum_{i \in N_j} d(m_j, o_j) \\
&\leq 2 \cdot \sum_{j \in [2]} \sum_{i \in N_j} d(i, m_j) + \sum_{j \in [2]} \sum_{i \in N_j} d(i, o_j) \\
&\leq 3 \cdot \sum_{j \in [2]} \sum_{i \in N_j} d(i, o_j) \\
&= 3 \cdot \text{SC}(\mathbf{o}).
\end{aligned}$$

Case 2: $t(m_1) = t(m_2)$. We can without loss of generality focus on the case where $2|S_1| - |N_1| \geq 2|S_2| - |N_2|$; the case where the inequality is the other way around can be handled using similar arguments. So, $w_1 = t(m_1)$ and $w_2 = s(m_2)$. Note that $|S_1| \geq |N_1|/2$ and $|S_2| \geq |N_2|/2$. If m_2 is closer to $s(m_2)$ than to o_2, then we can repeat the arguments of Case 1 to obtain an upper bound of 3. So, it suffices to focus on the case where m_2 is closer to o_2 than to $s(m_2)$, which means that $o_2 = t(m_2)$, and thus $o_1 < w_1 = o_2 < w_2$. Now, observe the following:

- Since m_1 is closer to w_1 than to o_1, we can move the agents in S_1 at $\frac{o_1+w_1}{2}$ and the remaining $|N_1| - |S_1|$ agents at o_1. Doing this, the approximation ratio cannot decrease (as we move towards o_1 and either towards w_1 at the same rate or away from w_1), m_1 remains the median agent of N_1 since $|S_1| \geq |N_1|/2$, and it is still true that $t(m_1) = w_1$.

- We have that m_2 is closer to o_2 than to w_2. If $m_2 \leq o_2$, then we can move the agents of N_2 as follows: each agent in S_2 is moved at o_2 and the remaining $|N_2| - |S_2|$ agents of N_2 (who are closer to w_2 than to o_2) at $\frac{o_2+w_2}{2}$. Doing this, the approximation ratio cannot decrease, m_2 remains the median agent of N_2 as $|S_2| \geq |N_2|/2$, and clearly, it is still true that $s(m_2) = w_2$.

  If $m_2 > o_2$, then we can move the agents of N_2 as follows: $|N_2|/2$ agents at o_2, $|S_2| - |N_2|/2$ agents at m_2, and the remaining $|N_2| - |S_2|$ agents (who are closer to w_2 than to o_2) at $\frac{o_2+w_2}{2}$. Doing this, the approximation ratio cannot decrease, m_2 remains the median agent of N_2 as $|S_2| \geq |N_2|/2$, and clearly, it is still true that $s(m_2) = w_2$.

  It is not hard to observe that the first case ($m_2 \leq o_2$) is worse in terms of approximation ratio than the second case ($m_2 > o_2$) as more agents are exactly at their optimal location. So, it suffices to consider this one.

Based on the above, in the worst case, we have

$$\begin{aligned}\text{SC}(\mathbf{w}) &= (|N_1| - |S_1|)(w_1 - o_1) + |S_1|\frac{w_1 - o_1}{2} \\ &\quad + |S_2|(w_2 - o_2) + (|N_2| - |S_2|)\frac{w_2 - o_2}{2}.\end{aligned}$$

and

$$\text{SC}(\mathbf{o}) = |S_1|\frac{w_1 - o_1}{2} + (|N_2| - |S_2|)\frac{w_2 - o_2}{2}.$$

Hence,

$$\begin{aligned}\frac{\text{SC}(\mathbf{w})}{\text{SC}(\mathbf{o})} &= 1 + 2 \cdot \frac{(|N_1| - |S_1|)(w_1 - o_1) + |S_2|(w_2 - o_2)}{|S_1|(w_1 - o_1) + (|N_2| - |S_2|)(w_2 - o_2)} \\ &\leq 1 + 2 \cdot \frac{|N_1| - |S_1| + |S_2|}{|S_1| + |N_2| - |S_2|} \\ &\leq 3,\end{aligned}$$

where the last inequality holds as $2|S_1| - |N_1| \geq 2|S_2| - |N_2|$ and the first inequality follows since $|S_1| \geq |N_1|/2$, $|S_2| \geq |N_2|/2$, and $w_2 - o_2 \leq o_2 - o_1 = w_1 - o_1$; the last is true as $s(m_2) = w_2$ and thus m_2, who is located at $o_2 = w_1$ in this worst-case instance, is closer to w_2 than to o_1. □

4.3.3 General instances

To tackle the general case, we consider the following mechanism. Let $j^* = \arg\max_{j \in [2]} |N_j \setminus N_{3-j}|$.

- If $|N_1 \cap N_2| \geq |N_{j^*} \setminus N_{3-j^*}|$, then run the Median mechanism with input the agents of $N_1 \cap N_2$ (ignoring all other agents).
- Otherwise, choose w_{j^*} to be the candidate location closest to the median m_{j^*}

of $N_{j^*} \setminus N_{3-j^*}$ (we slightly abuse notation here as m_{j^*} would normally be the median of N_{j^*}), and w_{3-j^*} to be the available candidate location closest to the median m_{3-j^*} of N_{3-j^*}; we refer to this mechanism as Alternate-Median.

Note that Median was shown to be strategyproof in Section 4.3.1. As for Alternate-Median, it is strategyproof since agents in $N_{j^*} \setminus N_{3-j^*}$ have no incentive to misreport and affect the choice of m_{j^*} and w_{j^*}, while agents in N_{3-j^*} cannot affect the choice of w_{j^*} and have no incentive to misreport and affect the choice of m_{3-j^*} and w_{3-j^*}; any misreport can only push the median, and the corresponding nearest location, farther away. Since the two cases are independent (the cardinalities of the sets of agents with different approval preferences are known), the mechanism combining Median and Alternate-Median is strategyproof.

We will bound the approximation ratio of the mechanism with the following two theorems which bound the approximation ratio of the mechanism in the two cases. Without loss of generality, to simplify our notation, let $j^* = 1$.

Theorem 4.3.7. *For general instances with $|N_1 \cap N_2| \geq |N_1 \setminus N_2|$, the approximation ratio of* Median *(mechanism 4) is at most 7.*

Proof. By Theorem 4.3.1, we have that

$$\sum_{i \in N_1 \cap N_2} \sum_{j \in [2]} d(i, w_j) \leq 3 \cdot \sum_{i \in N_1 \cap N_2} \sum_{j \in [2]} d(i, o_j).$$

For the agents in $N_1 \setminus N_2$, by the triangle inequality and since $|N_1 \setminus N_2| \leq$

$|N_1 \cap N_2|$, we have

$$\begin{aligned}\sum_{i \in N_1 \setminus N_2} d(i, w_1) &\leq \sum_{i \in N_1 \setminus N_2} d(i, o_1) + |N_1 \setminus N_2| \cdot d(w_1, o_1) \\ &\leq \sum_{i \in N_1 \setminus N_2} d(i, o_1) + \sum_{i \in N_1 \cap N_2} d(w_1, o_1) \\ &\leq \sum_{i \in N_1 \setminus N_2} d(i, o_1) + \sum_{i \in N_1 \cap N_2} \Big(d(i, w_1) + d(i, o_1) \Big).\end{aligned}$$

Similarly, for the agents in $N_2 \setminus N_1$, since $|N_2 \setminus N_1| \leq |N_1 \setminus N_2| \leq |N_1 \cap N_2|$, we have

$$\sum_{i \in N_2 \setminus N_1} d(i, w_2) \leq \sum_{i \in N_2 \setminus N_1} d(i, o_2) + \sum_{i \in N_1 \cap N_2} \Big(d(i, w_2) + d(i, o_2) \Big).$$

By combining these, we have

$$\begin{aligned}\mathrm{SC}(\mathbf{w}) &\leq 3 \cdot \mathrm{SC}(\mathbf{o}) + \sum_{i \in N_1 \cap N_2} \sum_{j \in [2]} \Big(d(i, w_j) + d(i, o_j) \Big) \\ &\leq 3 \cdot \mathrm{SC}(\mathbf{o}) + 4 \cdot \sum_{i \in N_1 \cap N_2} \sum_{j \in [2]} d(i, o_j) \\ &\leq 7 \cdot \mathrm{SC}(\mathbf{o}).\end{aligned}$$

Therefore, the approximation ratio is at most 7. □

Theorem 4.3.8. *For general instances with $|N_1 \cap N_2| \leq |N_1 \setminus N_2|$, the approximation ratio of* Alternate-Median *is at most* 7.

Proof. We consider the following cases:

Case 1: $t(m_1) \neq t(m_2)$**.** Then, we have that $w_1 = t(m_1)$ and $w_2 = t(m_2)$. By the properties of the median, we have that

$$\sum_{i \in N_1 \setminus N_2} d(i, m_1) \leq \sum_{i \in N_1 \setminus N_2} d(i, x)$$

and

$$\sum_{i \in N_2} d(i, m_2) \leq \sum_{i \in N_2} d(i, x)$$

for any point x of the line, including o_1 and o_2. Also, by the definition of $t(m_j)$ for $j \in [2]$, we have that $d(m_j, w_j) \leq d(m_j, x)$ for any candidate location x, including o_j. Therefore, using these facts and the triangle inequality, we bound the contribution of the different types of agents to the social cost of $\mathbf{w}$. In particular, for the agents of $N_1 \setminus N_2$, we have

$$\begin{aligned}
\sum_{i \in N_1 \setminus N_2} d(i, w_1) &\leq \sum_{i \in N_1 \setminus N_2} \Big(d(i, m_1) + d(m_1, w_1) \Big) \\
&\leq \sum_{i \in N_1 \setminus N_2} \Big(d(i, m_1) + d(m_1, o_1) \Big) \\
&\leq \sum_{i \in N_1 \setminus N_2} \Big(2 \cdot d(i, m_1) + d(i, o_1) \Big) \\
&\leq 3 \cdot \sum_{i \in N_1 \setminus N_2} d(i, o_1).
\end{aligned}$$

Similarly, for the agents of N_2, we have

$$\sum_{i \in N_2} d(i, w_2) \leq 3 \cdot \sum_{i \in N_2} d(i, o_2).$$

For the agents of $N_1 \cap N_2$ in terms of w_1, using the triangle inequality, we obtain

$$\sum_{i \in N_1 \cap N_2} d(i, w_1) \leq \sum_{i \in N_1 \cap N_2} d(i, o_1) + |N_1 \cap N_2| \cdot d(w_1, o_1)$$

$$
\begin{aligned}
&\leq \sum_{i \in N_1 \cap N_2} d(i, o_1) + \sum_{i \in N_1 \setminus N_2} d(w_1, o_1) \\
&= \sum_{i \in N_1 \cap N_2} d(i, o_1) + \sum_{i \in N_1 \setminus N_2} \Big(d(i, w_1) + d(i, o_1) \Big) \\
&\leq \sum_{i \in N_1 \cap N_2} d(i, o_1) + 4 \cdot \sum_{i \in N_1 \setminus N_2} d(i, o_1).
\end{aligned}
$$

By putting everything together, we obtain an upper bound of 7.

Case 2: $t(m_1) = t(m_2)$**.** In this case, we have that $w_1 = t(m_1) = t(m_2)$ and $w_2 = s(m_2)$. Clearly, if $d(m_2, w_2) \leq d(m_2, o_2)$, we get an upper bound of 7, similarly to Case 1. So, we can assume that $d(m_2, w_2) > d(m_2, o_2)$, which combined with the fact that $w_2 = s(m_2)$, implies that $o_2 = t(m_2) = w_1$. For the agents in $N_1 \setminus N_2$, since $w_1 = t(m_1)$, we have a 3-approximation guarantee (using the same arguments as above):

$$
\sum_{i \in N_1 \setminus N_2} d(i, w_1) \leq 3 \cdot \sum_{i \in N_1 \setminus N_2} d(i, o_1).
$$

For the agents in $N_1 \cap N_2$ in terms of w_1, similarly to Case 1, we have

$$
\begin{aligned}
\sum_{i \in N_1 \cap N_2} d(i, w_1) &\leq \sum_{i \in N_1 \cap N_2} \Big(d(i, o_1) + d(o_1, w_1) \Big) \\
&= \sum_{i \in N_1 \cap N_2} d(i, o_1) + |N_1 \cap N_2| \cdot d(o_1, w_1).
\end{aligned}
$$

For the agents in $N_2 = (N_2 \setminus N_1) \cup (N_1 \cap N_2)$ in terms of w_2, since $d(m_2, w_2) \leq d(m_2, o_1)$, $w_1 = o_2$, and m_2 minimizes the total distance of the agents in N_2 from any other point of the line, by the triangle inequality, we have

$$
\sum_{i \in N_2} d(i, w_2) \leq \sum_{i \in N_2} \Big(d(i, m_2) + d(m_2, w_2) \Big)
$$

$$
\begin{aligned}
&\le \sum_{i \in N_2} \Big(d(i, m_2) + d(m_2, o_1) \Big) \\
&\le \sum_{i \in N_2} \Big(d(i, m_2) + d(m_2, o_2) + d(o_1, o_2) \Big) \\
&\le \sum_{i \in N_2} \Big(2d(i, m_2) + d(i, o_2) + d(o_1, w_1) \Big) \\
&\le 3 \cdot \sum_{i \in N_2} d(i, o_2) + |N_2| \cdot d(o_1, w_1).
\end{aligned}
$$

So, by putting everything together and using the fact that $|N_2| = |N_2 \setminus N_1| + |N_2 \cap N_1|$, we have

$$
\begin{aligned}
\text{SC}(\mathbf{w}) &\le 3 \cdot \sum_{i \in N_1 \setminus N_2} d(i, o_1) + 3 \cdot \sum_{i \in N_2} d(i, o_2) + \sum_{i \in N_1 \cap N_2} d(i, o_1) \\
&\quad + |N_1 \cap N_2| \cdot d(o_1, w_1) + |N_2| \cdot d(o_1, w_1) \\
&\le 3 \cdot \text{SC}(\mathbf{o}) + \Big(|N_2 \setminus N_1| + 2 \cdot |N_1 \cap N_2| \Big) \cdot d(o_1, w_1).
\end{aligned}
$$

Since $w_1 = t(m_1)$, half of the agents in $N_1 \setminus N_2$ suffer a cost of at least $d(o_1, w_1)/2$ in the optimal solution. Also, all the agents of $N_1 \cap N_2$ suffer a cost of at least $d(o_1, o_2)/2 = d(o_1, w_1)/2$, and thus

$$
\text{SC}(\mathbf{o}) \ge \left(\frac{|N_1 \setminus N_2|}{4} + \frac{|N_1 \cap N_2|}{2} \right) d(o_1, w_1).
$$

Hence, since $|N_2 \setminus N_1| \le |N_1 \setminus N_2|$, the approximation ratio is at most

$$
3 + 4 \cdot \frac{|N_2 \setminus N_1| + 2|N_1 \cap N_2|}{|N_1 \setminus N_2| + 2|N_1 \cap N_2|} \le 7.
$$

Consequently, the approximation ratio is overall at most 7. □

Using Theorem 4.3.7 and Theorem 4.3.8, we obtain the following result.

Corollary 4.3.9. *For general instances, there is a strategyproof mechanism with approximation ratio at most* 7.

4.4 Max cost

In this section, we turn our attention to the max cost objective for which we show that the best possible approximation ratio of strategyproof mechanisms is between 2 and 3 for doubleton instances, and exactly 3 for singleton and general preferences.

4.4.1 Doubleton instances

For the upper bound, we consider a simple mechanism that places both facilities at the candidate locations that are closest to the leftmost agent ℓ. We refer to this mechanism as Leftmost; see Mechanism 6. It is not hard to show that this mechanism is strategyproof and that it achieves an approximation ratio of 3.

Mechanism 6: Leftmost

Input: Reported positions of agents;
Output: Facility locations $\mathbf{w} = (w_1, w_2)$;
$\ell \leftarrow$ leftmost agent in $N_1 \cap N_2$;
$w_1 \leftarrow t(\ell)$;
$w_2 \leftarrow s(\ell)$;

Theorem 4.4.1. *For doubleton instances,* Leftmost *is strategyproof and achieves an approximation ratio of at most* 3.

Proof. For the strategyproofness of the mechanism, consider any agent i; recall that i approves both facilities. To affect the outcome, agent i would have report a position that lies at the left of ℓ. However, changing the leftmost agent position can only lead to placing the facilities at locations farther away from i, and hence i has no incentive to misreport.

For the approximation ratio, let $\mathbf{o} = (o_1, o_2)$ be an optimal solution. Clearly, there exist $x \in \{o_1, o_2\}$ and $y \in \{o_1, o_2\} \setminus \{x\}$ such that $d(\ell, w_1) \leq d(\ell, x)$ and $d(\ell, w_2) \leq d(\ell, y)$. Let i be the (rightmost) agent who determines the max cost of the mechanism. Using the triangle inequality, we have

$$\begin{aligned}
\text{MC}(\mathbf{w}) &= d(i, w_1) + d(i, w_2) \\
&\leq \Big(d(i, x) + d(\ell, x) + d(\ell, w_1)\Big) + \Big(d(i, y) + d(\ell, y) + d(\ell, w_2)\Big) \\
&\leq 3 \max_{j \in N} \Big(d(j, x) + d(j, y)\Big) \\
&= 3 \cdot \text{MC}(\mathbf{o}).
\end{aligned}$$

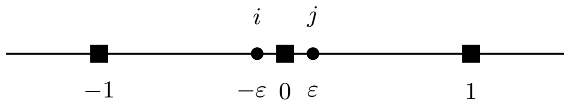
Therefore, the approximation ratio is at most 3. □

We next show a slightly weaker lower bound of 2 on the approximation ratio of any strategyproof mechanism.

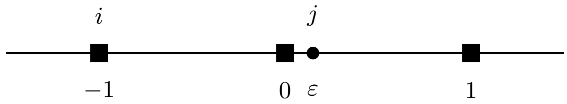
Theorem 4.4.2. *For doubleton instances, the approximation ratio of any deterministic strategyproof mechanism is at least $2 - \delta$, for any $\delta > 0$.*

Proof. Consider the following instance I_1: There are three candidate locations at -1, 0, and 1 and two agents (that approve both facilities) positioned at $-\varepsilon$ and ε, respectively, for some infinitesimal $\varepsilon > 0$. Since there are two facilities to be located, at least one of them must be placed at -1 or 1; see Figure 4.3a. Without loss of generality, let us assume that a facility is placed at 1.

Now, consider the instance I_2, which is the same as I_1 with the only difference that the agent at $-\varepsilon$ has been moved to -1; see Figure 4.3b. To maintain strate-



(a) Instance I_1.



(b) Instance I_2.

Figure 4.3: The two instances used in the proof of the lower bound of 2 in terms of the max cost for doubleton instances (Theorem 4.4.2). Both agents i and j approve both facilities. Rectangles represent candidate locations.

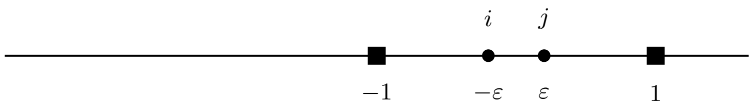
gyproofness, a facility must be placed at 1 in I_2 as well; otherwise, the agent at $-\varepsilon$ in I_1 would misreport her location as -1 to affect the outcome and decrease her cost. So, in I_2, any strategyproof mechanism either places one facility at -1 and one facility at 1, for a max cost of 2, or one facility at 0 and one facility at 1, for a max cost of 3. However, placing one facility at -1 and one facility at 0 leads to max cost $1+\varepsilon$, and thus an approximation ratio of at least $2-\delta$, for any $\delta > 0$. □

4.4.2 Singleton instances

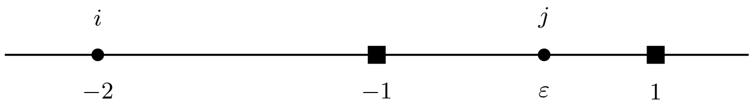
As argued at the beginning of Section 4.3.2, instances in which all agents approve one of the facilities are equivalent to having just this one facility to place. Consequently, by the work of Tang et al. [2020b], we cannot hope to achieve an approximation ratio better than 3 for singleton instances. For completeness, we include a simple proof of this lower bound here.

Theorem 4.4.3. *For singleton instances, the approximation ratio of any strategyproof mechanism is at least $3-\delta$, for any $\delta > 0$.*

Proof. Consider the following instance I_1: There are two candidate locations at -1 and 1 and two agents approving only F_1 positioned at $-\varepsilon$ and ε, respectively, for some infinitesimal $\varepsilon > 0$; see Figure 4.4a. Without loss of generality, we can



(a) Instance I_1.



(b) Instance I_2.

Figure 4.4: The two instances used in the proof of the lower bound of 3 in terms of the max cost for singleton instances (Theorem 4.4.3). Both agents i and j approve facility F_1. Rectangles represent candidate locations.

assume that F_1 is placed at 1 and F_2 at -1.

Now, consider the instance I_2, which is the same as I_1 with the only difference that the agent at $-\varepsilon$ has been moved to -2; see Figure 4.4b. To maintain strategyproofness, F_1 must be placed at 1 in I_2 as well; otherwise, the agent at $-\varepsilon$ in I_1 would misreport her position as -2 to decrease her cost from $1+\varepsilon$ to $1-\varepsilon$. This leads to a max cost of 3, when a max cost of $1+\varepsilon$ is possible by placing F_1 at -1. Therefore, the approximation ratio is at least $3-\delta$, for any $\delta > 0$. $\square$

A first idea towards an upper bound could be to place F_1 at the closest candidate location to the leftmost agent ℓ_1 of N_1, and F_2 at the closest available candidate location to an agent of N_2, such as the leftmost agent ℓ_2 or the rightmost agent r_2. While these mechanisms are clearly strategyproof, it is not hard to observe that they cannot achieve a good enough approximation ratio.

Example 4.4.4. If we place F_1 at $t(\ell_1)$ and F_2 at $t(r_2)$ or $s(r_2)$ depending on availability, then consider the following instance: There are three candidate locations at 0, 2, and 6. There is an agent ℓ_1 that approves F_1 at $1+\varepsilon$, an agent ℓ_2 that approves F_2 at 1, and another agent r_2 that approves F_2 at $3+\varepsilon$, for some infinitesimal $\varepsilon > 0$. So, we place F_1 at 2 and F_2 at 6 for a max cost of 5 (determined by r_2). On the other hand, we could place F_1 at 0 and F_2 at 2 for a max cost of approximately 1, leading to an approximation ratio of 5. Clearly, if we chose ℓ_2 instead of r_2 to determine the location of F_2, there is a symmetric instance leading again to the same lower bound.

The above example illustrates that it is not always a good idea to choose a priori ℓ_2 or r_2 to determine where to place F_2, especially when the closest candidate location to them might not be available after placing F_1. Instead, we need to carefully decide whether ℓ_2 or r_2 or neither of them is the best one to choose where to place F_2. We make this decision as follows: We "ask" ℓ_2 and r_2 to "vote" over two candidate locations; the candidate location L that is the closest at the left of $t(\ell_1)$ (where F_1 is placed) and the candidate location R that is the closest at the right of $t(\ell_1)$. If ℓ_2 and r_2 "agree", then they are both on the same side of the midpoint of the interval defined by L and R, and thus depending on whether they are on the left side (agree on L) or the right side (agree on R), we allow r_2 or ℓ_2, respectively, to make the choice of where to place F_2. If they "disagree", they are on different sides of the interval's midpoint, so neither ℓ_2 nor r_2 should make a choice of where to place F_2; in this case, the closest of L and R to $t(\ell_1)$ is a good candidate location to place F_2. This idea is formalized in Mechanism 7, which we call Vote-for-Priority.

We first show that Vote-for-Priority is strategyproof.

Theorem 4.4.5. *For singleton instances,* Vote-for-Priority *is strategyproof.*

Proof. Observe that no agent in N_1 has incentive to misreport as facility F_1 is located at the closest candidate location to ℓ_1; indeed, ℓ_1 is content while no agent would like to misreport to become the leftmost agent of N_1 as then F_1 will either remain at the same location or could be moved farther away. For the agents of N_2, we consider each case separately depending on which is the true profile. Denote by w_2 the location of F_2 when the agents report their positions

Mechanism 7: VOTE-FOR-PRIORITY

```
Input: Reported positions of agents with singleton preferences;
Output: Facility locations w = (w_1, w_2);
ℓ_1 ← leftmost agent in N_1;
ℓ_2 ← leftmost agent in N_2;
r_2 ← rightmost agent in N_2;
w_1 ← t(ℓ_1);
L ← closest candidate location at the left of w_1;
R ← closest candidate location at the right of w_1;
// (case 1) ℓ_2 and r_2 agree that L is closer, so r_2 gets to choose
if ℓ_2 and r_2 are both closer to L than to R then
    if t(r_2) is available then
        w_2 ← t(r_2);
    else
        w_2 ← s(r_2);
// (case 2) ℓ_2 and r_2 agree that R is closer, so ℓ_2 gets to choose
else if ℓ_2 and r_2 are both closer to R than to L then
    if t(ℓ_2) is available then
        w_2 ← t(ℓ_2);
    else
        w_2 ← s(ℓ_2);
// (case 3) ℓ_2 and r_2 disagree, so choose the closest of L and R to w_1
else
    w_2 ← arg min_{x∈{L,R}} {|w_1 − x|};
```

truthfully.

(Case 1) Clearly, r_2 has no incentive to deviate. Consider an agent $i \in N_2$, other than r_2, that deviates and misreports a position g.

- If $g \leq r_2$, then the location F_2 is still w_2.
- If $g \in \left(r_2, \frac{L+R}{2}\right]$, agent i becomes the rightmost agent but we are still in Case 1. So, the location of F_2 becomes the closest available location to g, which is either w_2 or some candidate location at the right of w_2. This means that the cost of i either remains the same or increases, and thus i has no incentive to misreport such a position.
- If $g > \frac{L+R}{2}$, agent i becomes the rightmost agent and the location of F_2 is

determined by Case 3, i.e., becomes $y = \arg\min_{x\in\{L,R\}}\{|w_1 - x|\}$. Since the true rightmost agent r_2 is closer to L than to R, it holds that $w_2 \leq L \leq y$. This again means that the cost of i either remains the same or increases, and thus i has no incentive to misreport such a position.

(Case 2) This is symmetric to Case 1.

(Case 3) Observe that any deviation that still leads to Case 3 does not affect the outcome of the mechanism as $w_2 = \arg\min_{x\in\{L,R\}}\{|w_1 - x|\}$. Hence, no agent $i \in N_2 \setminus \{\ell_2, r_2\}$ can affect the outcome as any possible misreported position can either be at the left of ℓ_2 or the right of r_2, which means that we are still in Case 3. Now, let us assume that r_2 misreports so that the location of F_2 is determined by Case 1. Since, in that case, all agents are closer to L than to R, and there are no other available candidate locations in the interval $[L, R]$ (since w_1 is occupied by F_1), F_2 can only be placed at some location $y \leq L$, which is clearly not better for r_2. A symmetric argument for ℓ_2 shows that again no agent can misreport. $\square$

Next, we show that Vote-for-Priority achieves an approximation ratio of at most 3.

Theorem 4.4.6. *For singleton instances, the approximation ratio of* Vote-for-Priority *is at most* 3.

Proof. If the max cost of the mechanism is due to an agent $i \in N_1$, the choice $w_1 = t(\ell_1)$ implies that $d(\ell_1, w_1) \leq d(\ell_1, o_1)$, and thus, by the triangle inequality, we have that

$$\begin{aligned}\text{MC}(\mathbf{w}) &= d(i, w_1)\\ &\leq d(i, o_1) + d(\ell_1, o_1) + d(\ell_1, w_1)\end{aligned}$$

$$\leq d(i, o_1) + 2 \cdot d(\ell_1, o_1)$$
$$\leq 3 \cdot \text{MC}(\mathbf{o}).$$

So, we now focus on the case where the max cost of the mechanism is determined by an agent in N_2 and we may assume that $w_2 \neq o_2$ as otherwise the claim holds trivially. Due to the symmetry of Case 1 and Case 2, it suffices to bound the approximation ratio in Case 1 and in Case 3. In any of these cases, if there is an agent of N_2 that is closer to w_2 than to o_2, then, similarly to above, by applying the triangle inequality, we can again show that the approximation ratio is at most 3. Thus, we will assume that all agents of N_2 are closer to o_2 than to w_2, which means that $w_1 = o_2$. To see that, note that, in Case 1, o_2 has to be unavailable as it must hold $t(r_2) = o_2$, while in Case 3, w_2 is the closest candidate location among L and R to w_1 (and thus ℓ_2 cannot be at the left of L if L is chosen and r_2 cannot be at the right of R if R is chosen). Due to this, o_1 cannot be w_1 and we have the following two possibilities:

- If $o_1 \leq L$, then $d(\ell_1, o_1) \geq d(\ell_1, L)$.
- If $o_1 \geq R$, then $d(\ell_1, o_1) \geq d(\ell_1, R)$.

(Case 1) Since $t(r_2) = w_1 = o_2$ and w_2 is the closest available candidate location to r_2, it has to be the case that $w_2 = L$. Let $i \in \{\ell_2, r_2\}$ be the agent of N_2 that gives the max cost.

- If $o_1 \leq L$, then due to the triangle inequality, and the facts that $o_2 = w_1$ and $d(\ell_1, o_1) \geq d(\ell_1, L)$, we have

$$\text{MC}(\mathbf{w}) = d(i, L) \leq d(i, o_2) + d(\ell_1, o_2) + d(\ell_1, L)$$

$$
\begin{aligned}
&= d(i, o_2) + d(\ell_1, w_1) + d(\ell_1, L) \\
&\leq d(i, o_2) + d(\ell_1, o_1) + d(\ell_1, o_1) \\
&\leq 3 \cdot \mathrm{MC}(\mathbf{o}).
\end{aligned}
$$

- If $o_1 \geq R$, then due to the triangle inequality, and the facts that $o_2 = w_1$, $d(i, L) \leq d(i, R)$ and $d(\ell_1, o_1) \geq d(\ell_1, R)$, we have

$$
\begin{aligned}
\mathrm{MC}(\mathbf{w}) = d(i, L) \leq d(i, R) &\leq d(i, o_2) + d(\ell_1, o_2) + d(\ell_1, R) \\
&= d(i, o_2) + d(\ell_1, w_1) + d(\ell_1, R) \\
&\leq d(i, o_2) + d(\ell_1, o_1) + d(\ell_1, o_1) \\
&\leq 3 \cdot \mathrm{MC}(\mathbf{o}).
\end{aligned}
$$

(Case 3) Without loss of generality, let us assume that $w_2 = R$; the case where $w_2 = L$ is symmetric. So, $d(R, w_1) = d(R, o_2) \leq d(L, o_2) = d(L, w_1)$. Since $r_2 \leq R$, the max cost of the mechanism is determined by agent ℓ_2.

- If $o_1 \geq R$, then due to the triangle inequality, and the facts that $w_1 = o_2$ and $d(\ell_1, o_1) \geq d(\ell_1, R)$, we have

$$
\begin{aligned}
\mathrm{MC}(\mathbf{w}) = d(\ell_2, R) &\leq d(\ell_2, o_2) + d(\ell_1, o_2) + d(\ell_1, R) \\
&= d(\ell_2, o_2) + d(\ell_1, w_1) + d(\ell_1, R) \\
&\leq d(\ell_2, o_2) + d(\ell_1, o_1) + d(\ell_1, o_1) \\
&\leq 3 \cdot \mathrm{MC}(\mathbf{o}).
\end{aligned}
$$

- If $o_1 \leq L$, then due to the triangle inequality, and the facts that $d(R, o_2) \leq$

$d(L, o_2)$, $w_1 = o_2$ and $d(\ell_1, o_1) \geq d(\ell_1, L)$, we have

$$\begin{aligned}
\text{MC}(\mathbf{w}) = d(\ell_2, R) &\leq d(\ell_2, o_2) + d(R, o_2) \\
&\leq d(\ell_2, o_2) + d(L, o_2) \\
&\leq d(\ell_2, o_2) + d(\ell_1, o_2) + d(\ell_1, L) \\
&= d(\ell_2, o_2) + d(\ell_1, w_1) + d(\ell_1, L) \\
&\leq d(\ell_2, o_2) + d(\ell_1, o_1) + d(\ell_1, o_1) \\
&\leq 3 \cdot \text{MC}(\mathbf{o}).
\end{aligned}$$

This completes the proof. □

4.4.3 General instances

To tackle the general case, we consider a mechanism that runs Leftmost in case the instance consists of at least one agent with doubleton preference, and Vote-for-Priority in case the instance is singleton. It is not hard to observe that Leftmost is strategyproof even when there are agents with singleton preference; its decision is fully determined by the leftmost agent with doubleton preference and the input of any other agent is ignored. Hence, the mechanism is overall strategyproof. We will now show that Leftmost still achieves an approximation ratio of at most 3 when it is applied, which will allow us to show an overall bound of 3.

Theorem 4.4.7. *For instances with at least one agent with doubleton preference, the approximation ratio of* Leftmost *is at most* 3.

Proof. We consider cases depending on the preference of the agent i that determines the max cost of the mechanism. Let ℓ be the leftmost agent in $N_1 \cap N_2$, and recall that $w_1 = t(\ell)$ and $w_2 = s(\ell)$.

(Case 1) The max cost is determined by an agent $i \in N_1 \setminus N_2$. Then, by the triangle inequality and since $d(\ell, w_1) \leq d(\ell, o_2)$, we have

$$\begin{aligned} \text{MC}(\mathbf{w}) = d(i, w_1) &\leq d(i, o_1) + d(\ell, o_1) + d(\ell, w_1) \\ &\leq d(i, o_1) + d(\ell, o_1) + d(\ell, o_2) \\ &\leq 2 \cdot \text{MC}(\mathbf{o}). \end{aligned}$$

(Case 2) The max cost is determined by an agent $i \in N_2 \setminus N_1$. Since $w_2 = s(\ell)$, there exists $x \in \{o_1, o_2\}$ such that $d(\ell, w_2) \leq d(\ell, x) \leq \text{MC}(\mathbf{o})$. Hence, by the triangle inequality, we have

$$\text{MC}(\mathbf{w}) = d(i, w_2) \leq d(i, o_2) + d(\ell, o_2) + d(\ell, w_2) \leq 3 \cdot \text{MC}(\mathbf{o}).$$

(Case 3) The max cost is determined by an agent $i \in N_1 \cap N_2$. Then, following the proof of Theorem 4.4.1 for doubleton instances, we can show an upper bound of 3. $\square$

By combining Theorem 4.4.7 and Theorem 4.4.6, we obtain the following result.

Corollary 4.4.8. *For general instances, there is a strategyproof mechanism with approximation ratio at most* 3*.*

4.5 Allowing same facility locations

In this last section we explore the simpler model in which *the two facilities can be placed at the same candidate location.* We show tight bounds on the approximation ratio of deterministic mechanisms for doubleton and general instances (we

	Social cost	Max cost
Doubleton	$1+\sqrt{2}$	2
General	3	3

Table 4.2: Overview of the tight bounds for the model where the two facilities are allowed to be placed at the same candidate location.

will not consider singleton instances separately as the approximation ratio turns out to be exactly the same as for general instances). Our results for this model are summarized in Table 4.2.

All the mechanisms we will consider in this section place the facilities at the closest locations to some fixed agents that approve them. In particular, given the positions reported by the agents, for some $q_j \in [n]$, we place each facility F_j at $t(i_j)$, where i_j is the q_j-th ordered agent in N_j. It is not hard to verify that all such mechanisms are strategyproof. Indeed, to change the outcome of the mechanism, an agent in N_j would have to report a position that changes the q_j-th ordered agent in N_j, but this would mean that the facilities that this agent approves might move farther away from the true position of the agent.

Before we continue we remark that the fact that facilities can be placed at the same location is crucial for our mechanisms to be strategyproof since this eliminates possible misreports by the q_j-th ordered agents who determine where the facilities are placed. To be more specific, suppose that we try to adapt this mechanism for the main model that we considered in the previous sections in which the two facilities can only be placed at different locations. Then, in case $t(i_1) = t(i_2)$ we would have to resolve this collision somehow, for example by giving priority to one of these agents, say i_1, and placing F_1 at $w_1 = t(i_1)$ and then F_2 at some other location w_2 that is a function of i_2 such as $s(i_2)$. However, if i_1 approves both facilities, it might be the case that w_2 is not close to her position, and thus she prefers to misreport that she is closer to $s(i_1)$ rather than $t(i_1)$, leading to F_1 being placed at $s(i_1)$ and then F_2 at $t(i_2) = t(i_1)$. Such misreports cannot happen when facilities are allowed to be placed at the same location.

4.5.1 Social cost

We start with the case of doubleton instances for which we show a tight bound of $1+\sqrt{2}$. The lower bound follows by observing that the proof of Theorem 4.3.2 holds even when facilities are allowed to be placed at the same location; in particular, in the proof of that theorem we made the simplification that there are two candidate location at 2 and -2, thus having capacity for both facilities. For the upper bound, first observe that the MEDIAN mechanism from Section 4.3.1 can also be adapted to the current model (by placing both facilities to the location closest to the median agent), but it is not hard to show that it still cannot achieve an approximation ratio better than 3. To improve upon the bound of 3, we consider a family of mechanisms, which, for a parameter $\alpha \in (0, 1/2)$, place one facility at the candidate location closest to the position reported by the αn-leftmost agent, and the other facility at the candidate location closest to the position reported by the $(1-\alpha)n$-leftmost agent[2]. We refer to such mechanisms as α-STATISTIC; see Mechanism 8 for a description. It is not hard to observe that, for any $\alpha \in (0, 1/2)$, the mechanism is strategyproof since it falls within the class of mechanisms we described earlier with $q_1 = \alpha n$ and $q_2 = (1-\alpha)n$. We now focus on bounding the approximation ratio.

Theorem 4.5.1. *For doubleton instances, the approximation ratio of*

[2]Formally, it would be the $\lceil \alpha n \rceil$-leftmost and the $\lceil (1-\alpha)n \rceil$-leftmost agent, respectively, and we require that $\lceil \alpha n \rceil < \lceil (1-\alpha)n \rceil$. This can be guaranteed by creating an identical number of copies for each agent and running the mechanism on the modified instance; the approximation ratio for the modified instance is exactly the same as for the original instance. We omit the ceilings to make the exposition clearer.

Mechanism 8: α-STATISTIC

Input: Reported positions of agents with doubleton preferences;
Output: Facility locations $\mathbf{w} = (w_1, w_2)$;
$i \leftarrow \alpha n$-leftmost agent;
$j \leftarrow (1-\alpha)n$-leftmost agent;
$w_1 \leftarrow t(i)$;
$w_2 \leftarrow t(j)$;

$\left(\sqrt{2}-1\right)$-STATISTIC *is at most* $1+\sqrt{2}$.

Proof. We have $\alpha = \sqrt{2}-1$ and note that $\frac{1+\alpha}{1-\alpha} = \frac{1}{\alpha} = 1+\sqrt{2}$. If o is a location that minimizes the total distance from the agent positions, then for any x and y such that $o \leq x \leq y$ or $y \leq x \leq o$, it holds that $\sum_{i\in N} d(i,o) \leq \sum_{i\in N} d(i,x) \leq \sum_{i\in N} d(i,y)$. Hence, since the individual cost of each agent is the sum of distances from both facilities, there exists an optimal solution $\mathbf{o} = (o_1, o_2)$ such that $o_1 = o_2 = o$. Without loss of generality, we assume that $w_1 \leq w_2$, and it must be the case that $w_1 \neq o$ or $w_2 \neq o$ since otherwise the approximation ratio would be 1. We consider the following cases:

Case 1: $w_1 < o = w_2$ **(the case** $w_1 = o < w_2$ **is symmetric).**

By the definition of the mechanism, there is a set S of αn agents that are closer to w_1 than to o. Hence, we have

$$
\begin{aligned}
\text{SC}(\mathbf{w}) &= \sum_{i\in S} d(i,w_1) + \sum_{i\notin S} d(i,w_1) + \sum_{i\in N} d(i,o) \\
&\leq \sum_{i\in S} d(i,o) + \sum_{i\notin S} \Big(d(i,o) + d(w_1,o)\Big) + \sum_{i\in N} d(i,o) \\
&= \text{SC}(\mathbf{o}) + (1-\alpha)n \cdot d(w_1,o)
\end{aligned}
$$

and

$$\text{SC}(\mathbf{o}) \geq 2 \cdot \alpha n \cdot \frac{d(w_1, o)}{2} = \alpha n \cdot d(w_1, o).$$

Therefore, the approximation ratio is at most $1 + \frac{1-\alpha}{\alpha} = \frac{1}{\alpha} = 1 + \sqrt{2}$.

Case 2: $w_1 < o < w_2$**.**

By the definition of the mechanism, there is a set S_1 of αn agent that are closer to w_1 than to o, i.e., $d(i, w_1) \leq d(i, o)$ for every $i \in S_1$, and thus $d(i, o) \geq d(o, w_1)/2$. Similarly, there is another set S_2 of αn agents that are closer to w_2 than to o, i.e., $d(i, w_2) \leq d(i, o)$ for every $i \in S_2$, and thus $d(i, o) \geq d(o, w_2)/2$. By combining these facts with the triangle inequality, we have

$$\begin{aligned}
\text{SC}(\mathbf{w}) &= \sum_{i \in S_1} \Big(d(i, w_1) + d(i, w_2) \Big) + \sum_{i \in S_2} \Big(d(i, w_1) + d(i, w_2) \Big) \\
&\quad + \sum_{i \notin S_1 \cup S_2} \Big(d(i, w_1) + d(i, w_2) \Big) \\
&\leq \sum_{i \in S_1} \Big(2 \cdot d(i, o) + d(o, w_2) \Big) + \sum_{i \in S_2} \Big(2 \cdot d(i, o) + d(o, w_1) \Big) \\
&\quad + \sum_{i \notin S_1 \cup S_2} \Big(2 \cdot d(i, o) + d(o, w_1) + d(o, w_2) \Big) \\
&= \text{SC}(\mathbf{o}) + (1 - \alpha) n \Big(d(o, w_1) + d(o, w_2) \Big).
\end{aligned}$$

We can also bound the optimal social cost as follows:

$$\text{SC}(\mathbf{o}) \geq 2 \cdot \alpha n \frac{d(o, w_1)}{2} + 2 \cdot \alpha n \frac{d(o, w_2)}{2} = \alpha n \Big(d(o, w_1) + d(o, w_2) \Big)$$

Consequently, the approximation ratio is at most $1 + \frac{1-\alpha}{\alpha} = 1/\alpha = 1 + \sqrt{2}$.

Case 3: $o < w = w_1 = w_2$ **(the case** $w_1 = w_2 = w < o$ **is symmetric).**

By the definition of the mechanism, there is a set S of $(1-\alpha)n$ agents that are closer to w than to o. Hence, by the triangle inequality, we have

$$\begin{aligned}
\text{SC}(\mathbf{w}) &= 2 \cdot \sum_{i \in S} d(i,w) + 2 \cdot \sum_{i \notin S} d(i,w) \\
&\leq 2 \cdot \sum_{i \in S} d(i,o) + 2 \cdot \sum_{i \notin S} \Big(d(i,o) + d(o,w) \Big) \\
&= \text{SC}(\mathbf{o}) + 2\alpha n \cdot d(o,w).
\end{aligned}$$

and

$$\text{SC}(\mathbf{o}) \geq 2 \cdot (1-\alpha)n \cdot \frac{d(o,w)}{2} = (1-\alpha)n \cdot d(o,w)$$

Therefore, the approximation ratio is at most $1 + \frac{2\alpha}{1-\alpha} = \frac{1+\alpha}{1-\alpha} = 1 + \sqrt{2}$.

Case 4: $o < w_1 < w_2$ (the case $w_1 < w_2 < o$ is symmetric).

Clearly, since $o < w_1 < w_2$, $d(o,w_2) = d(o,w_1) + d(w_1,w_2)$. By the definition of the mechanism, there is a set S of $(1-\alpha)n$ agents who are closer to w_1 than to o, i.e., $d(i,w_1) \leq d(i,o)$ for every $i \in S$, and thus $d(i,o) \geq d(o,w_1)/2$. Also, there is a set $T \subset S$ of αn agents who are closer to w_2 than to w_1, i.e., $d(i,w_2) \leq d(i,w_1) \leq d(i,o)$ for every $i \in T$, and thus $d(i,o) \geq d(o,w_1) + d(w_1,w_2)/2$. By combining these two facts with the triangle inequality, we have

$$\begin{aligned}
\text{SC}(\mathbf{w}) = &\sum_{i \notin S} \Big(d(i,w_1) + d(i,w_2) \Big) + \sum_{i \in S \setminus T} \Big(d(i,w_1) + d(i,w_2) \Big) \\
&+ \sum_{i \in T} \Big(d(i,w_1) + d(i,w_2) \Big) \\
&\leq \sum_{i \notin S} \Big(2 \cdot d(i,o) + d(o,w_1) + d(o,w_2) \Big) + \sum_{i \in S \setminus T} \Big(2 \cdot d(i,o) + d(o,w_2) \Big)
\end{aligned}$$

$$
\begin{aligned}
&+ 2 \cdot \sum_{i \in T} d(i,o) \\
&= \text{SC}(\mathbf{o}) + \alpha\Big(d(o,w_1) + d(o,w_2)\Big) + (1-2\alpha)d(o,w_2) \\
&= \text{SC}(\mathbf{o}) + \alpha d(o,w_1) + (1-\alpha)d(o,w_2) \\
&= \text{SC}(\mathbf{o}) + d(o,w_1) + (1-\alpha)d(w_1,w_2).
\end{aligned}
$$

For the optimal social cost, we have

$$
\begin{aligned}
\text{SC}(\mathbf{o}) &\geq 2|S \setminus T|\frac{d(o,w_1)}{2} + 2|T|\left(d(o,w_1) + \frac{d(w_1,w_2)}{2}\right) \\
&= d(o,w_1) + \alpha d(w_1,w_2).
\end{aligned}
$$

Hence, the approximation ratio is at most $1 + \frac{1-\alpha}{\alpha} = 1/\alpha = 1 + \sqrt{2}$. □

For general instances we show a tight bound of 3. The lower bound follows by the fact that when all agents have singleton preferences, then the problem reduces to two independent single-facility location problems, and the best possible approximation ratio for each of them is 3 [Feldman et al., 2016]; alternatively, one can verify that the proof of Theorem 4.3.3 holds even when the facilities can be placed at the same location. For the upper bound, we consider the Two-Medians mechanism, which independently places each facility F_j at the location closest to the median agent $m_j \in N_j$.

Theorem 4.5.2. *For general instances, the approximation ratio of* Two-Medians *is at most* 3.

Proof. Using the fact that the median agent m_j minimizes the total distance of all the agents in N_j, the fact that $w_j = t(m_j)$, and the triangle inequality, we have

$$\begin{aligned}
\text{SC}(\mathbf{w}) &= \sum_{j\in[2]}\sum_{i\in N_j} d(i,w_j) \\
&\leq \sum_{j\in[2]}\sum_{i\in N_j} \Big(d(i,m_j) + d(m_j,w_j)\Big) \\
&\leq \sum_{j\in[2]}\sum_{i\in N_j} \Big(d(i,m_j) + d(m_j,o_j)\Big) \\
&\leq \sum_{j\in[2]}\sum_{i\in N_j} \Big(2\cdot d(i,m_j) + d(i,o_j)\Big) \\
&\leq 3\cdot \sum_{j\in[2]}\sum_{i\in N_j} d(i,o_j),
\end{aligned}$$

and thus the approximation ratio is at most 3. □

4.5.2 Max cost

We now consider the max cost and start by showing a tight bound of 2 for doubleton instances. The lower bound follows by a sequence of instances similar to those in the proof of Theorem 4.4.2 but just with two candidate locations.

Theorem 4.5.3. *For doubleton instances, the approximation ratio of any deterministic strategyproof mechanism is at least $2-\delta$, for any $\delta > 0$.*

Proof. Consider an arbitrary deterministic mechanisms and the following instance I_1: There are two candidate locations at -1 and 1 and two agents (that approve both facilities) positioned at $-\varepsilon$ and ε, respectively, for some infinitesimal $\varepsilon > 0$.

First, suppose that the mechanism places both facilities at one of the two locations, say -1. Then, consider the instance I_2 in which the agent at ε in I_1 moves to 1 in I_2, while the other agent remains at $-\varepsilon$. The mechanism must still place both facilities at -1 in I_2 since otherwise the agent that moved would decrease her cost. However, $\mathrm{MC}(-1,-1) \approx 6$ and $\mathrm{MC}(1,1) \approx 2$, leading to an approximation ratio of at least 3.

Second, support that the mechanism places one facility at -1 and the other at 1. Then, consider the instance I_3 in which the agent at ε in I_1 moves to 2 in I_3, while the other agent remains at $-\varepsilon$. The mechanism must either still output the solution $(-1,1)$ or the solution $(-1,-1)$, but it cannot output $(1,1)$ as then the agent that moved would decrease her cost. However, $\mathrm{MC}(-1,1) \approx 4$, $\mathrm{MC}(-1,-1) = 6$, and $\mathrm{MC}(1,1) = 2$, leading to an approximation ratio of at least 2. □

For the upper bound, we consider the mechanism that places F_1 at the candidate location closest to the leftmost agent ℓ and F_2 at the candidate location closest to the rightmost agent r. We refer to this mechanism as Leftmost-Rightmost; see Mechanism 9.

Mechanism 9: Leftmost-Rightmost

Input: Reported positions of agents with doubleton preferences;
Output: Facility locations $\mathbf{w} = (w_1, w_2)$;
$\ell \leftarrow$ leftmost agent in $N_1 \cap N_2$;
$r \leftarrow$ rightmost agent in $N_1 \cap N_2$;
$w_1 \leftarrow t(\ell)$;
$w_2 \leftarrow t(r)$;

Theorem 4.5.4. *For doubleton instances, the approximation ratio of* Leftmost-Rightmost *is at most* 2.

Proof. Let $i \in \{\ell, r\}$ be the agent that determines the max cost of the mechanism, and $j \in \{\ell, r\} \setminus \{i\}$. Let $\mathbf{o} = (o_1, o_2)$ be an optimal solution. Since $w_1 = t(\ell)$

and $w_2 = t(r)$, by the triangle inequality and the definition of $t(\cdot)$, we have

$$\begin{aligned}
\text{MC}(\mathbf{w}) &= d(i,t(i)) + d(i,t(j)) \\
&\leq d(i,t(i)) + d(i,o_2) + d(j,o_2) + d(j,t(j)) \\
&\leq d(i,o_1) + d(i,o_2) + d(j,o_2) + d(j,o_1) \leq 2 \cdot \text{MC}(\mathbf{o}).
\end{aligned}$$

Therefore, the approximation ratio is at most 2 in any case. □

For general instances, it is not hard to obtain a tight upper bound of 3. The lower bound follows again by the fact that with singleton preferences the problem is equivalent to two independent single-facility location problems, while the upper bounds follows by the variant of the Leftmost mechanism that places F_j at the leftmost agent $\ell_j \in N_j$.

Theorem 4.5.5. *For general instances, the approximation ratio of* Leftmost *is at most* 3.

Proof. Let i be the agent that determines the max cost of the mechanism. By the triangle inequality and the definition of $t(\cdot)$, we have

$$\begin{aligned}
\text{MC}(\mathbf{w}) = \sum_{j\in[2]:i\in N_j} d(i,w_j) &\leq \sum_{j\in[2]:i\in N_j} \Big(d(i,o_j) + d(\ell_j,o_j) + d(\ell_j,w_j) \Big) \\
&\leq \sum_{j\in[2]:i\in N_j} \Big(d(i,o_j) + 2 \cdot d(\ell_j,o_j) \Big) \\
&\leq 3 \cdot \text{MC}(\mathbf{o}).
\end{aligned}$$

Hence the approximation ratio is at most 3. □

Chapter 5

Settling the Approximation Ratio of Distributed Facility Location

5.1 Definitions and notation

We consider *one-facility* location problems in this chapter.

With the basic definition in chapter 2, the instance I in this chapter is a tuple $I = (N, \mathbf{x}, D)$, where

- N is a set of n *agents*.
- $\mathbf{x} = (x_i)_{i \in N}$ is a vector containing the *position* $x_i \in \mathbb{R}$ of agent i on the line of real numbers.
- $D = \{d_1, ..., d_k\}$ is a set of $k \geq 1$ *districts*. Each district $d \in D$ contains a set $N_d \subseteq N$ of agents such that $N_d \cap N_{d'} = \varnothing$ and $\bigcup_{d \in D} N_d = N$. By $n_d = |N_d|$ we denote the number of agents in d; when $n_d := \lambda := n/k$ for every $d \in D$, we say that the districts are *symmetric*.

A *distributed mechanism* M is used to decide the location of a facility based on the positions reported by the agents and the composition of the districts. In particular, given an instance I, a distributed mechanism works by implementing the following two steps:

- Step 1: For each district $d \in D$, using only the positions of the agents in N_d, the mechanism chooses a *representative location* $y_d \in \mathbb{R}$ for the district.
- Step 2: Given the size and the representative locations of the districts, the mechanism outputs a single location $M(I) \in \{y_d\}_{d \in D}$ as the *winner*.

If a location z is chosen, then the distance $\delta(x_i, z) = |x_i - z|$ between the position x_i of agent i and z is the *individual cost* of agent i for z.

5.1.1 Social objectives and strategyproofness

We design deterministic distributed mechanisms that satisfy various criteria of interest and achieve the best possible approximation ratio bounds. First, we aim to design distributed mechanisms to approximately optimize social objectives that are functions of the distances between the chosen locations and the positions of the agents. Following the work of Anshelevich et al. [2022], we focus on the following objectives:

- The *Average* cost (or *average social cost*) of location z is the average total individual cost of all agents for z:

$$\frac{1}{n}\sum_{i \in N} \delta(x_i, z) = \frac{1}{n}\sum_{d \in D}\sum_{i \in N_d} \delta(x_i, z).$$

- The *Max* cost of location z is the maximum individual cost over all agents for z:

$$\max_{i \in N} \delta(x_i, z) = \max_{d \in D} \max_{i \in N_d} \delta(x_i, z).$$

- The *Average-of-Max* cost of location z is the average sum over each district of the maximum individual cost therein:

$$\frac{1}{k}\sum_{d\in D}\left\{\max_{i\in N_d}\delta(x_i,z)\right\}.$$

- The *Max-of-Average* cost of location z is the maximum over each district of the average total individual cost therein:

$$\max_{d\in D}\left\{\frac{1}{n_d}\sum_{i\in N_d}\delta(x_i,z)\right\}.$$

To simplify our notation, whenever the social objective is clear from context, we will use cost$(w|I)$ to denote the cost of $w\in\mathbb{R}$ according to the objective function at hand in instance I. Whenever I is clear from context, we will drop it from notation and simply write cost(w); this will mostly be done in the proofs of our upper bounds.

Another goal is to design mechanisms that are resilient to strategic manipulation, that is, they do not allow the agents to unilaterally affect the outcome in their favor (i.e., lead to a location with smaller individual cost) by reporting false positions. Formally, a mechanism is *strategyproof* if for any pair of instances $I=(N,(\mathbf{x}_{-i},x_i),D)$ and $J=(N,(\mathbf{x}_{-i},x_i'),D)$ that differ in the position of a single agent i, it holds that $\delta(x_i,M(I))\leq\delta(x_i,M(J))$.

5.1.2 Useful observations

Before we proceed with the presentation of our main technical results in the upcoming sections, we first state some useful properties. The bounds on the approximation ratio of some of our mechanisms will follow by characterizing worst-case instances, and for that we will need the inequality

$$\frac{\alpha+\gamma}{\beta+\gamma} < \frac{\alpha}{\beta}, \tag{5.1}$$

which holds for any $\alpha > \beta \geq 0$ and $\gamma > 0$.

Another useful observation is that any distributed mechanism with finite approximation ratio with respect to any of the social objectives that we consider must be cardinally unanimous. Formally, a mechanism is *cardinally-unanimous* if it chooses the representative location of a district to be z whenever all agents in the district are positioned at z.

Lemma 5.1.1. *Any distributed mechanism that achieves finite approximation ratio with respect to any social objective $F \in$ {Average, Max, Average-of-Max, Max-of-Average} must be cardinally-unanimous.*

Proof. Let M be a distributed mechanism that is *not* cardinally-unanimous. Consequently, there must exist a location z such that when all the agents of a district are positioned at z, the mechanism decides the representative location of the district to be some $y \neq z$. Now, consider an instance in which all agents (no matter which district they belong to) are positioned at z. Given the behavior of the mechanism, y is the representative location of all districts, and thus it must be the winner. However, cost$(w) = 0$ and cost$(y) > 0$ for any social objective F, and thus the approximation ratio is unbounded. So, to achieve finite approximation ratio, any mechanism must be cardinally-unanimous. □

We next show that a class of p-Statistic-of-q_d-Statistic distributed mechanisms is strategyproof. Let $p \in [k]$ and $q_d \in [n_d]$ for any district d. The p-Statistic-of-q_d-Statistic mechanism first chooses the representative location of each district

d to be the position of the q_d-th ordered agent therein, and then outputs the p-th ordered representative location as the winner. For example, if $p = \lfloor (k+1)/2 \rfloor$ and $q_d = \lfloor (n_d+1)/2 \rfloor$, the mechanism selects the position of the (leftmost) median agent in each district to be its representative location and then selects the (leftmost) median representative location as the winner. All strategyproof mechanisms that achieve the best possible approximation ratio for the various social objectives we consider are members of this class. The next lemma shows that any such mechanism is strategyproof, and will allow us to only focus on bounding the approximation ratio in the next sections.

Lemma 5.1.2. *Let $p \in [k]$ and $q_d \in [n_d]$ for $d \in D$. Then, the p-Statistic-of-q_d-Statistic mechanism is strategyproof.*

Proof. Let M be the p-Statistic-of-q_d-Statistic mechanism. Consider any instance $I = (N, \mathbf{x}, D)$ and let $w = M(I)$ be the location chosen by M. Let i be any agent that belongs to some district $d \in D$ that is represented by y. If the position of i is the final winner, then i clearly has no incentive to deviate. So, without loss of generality, assume that the winner is some location $w > x_i$. Observe that to affect the outcome of the mechanism, agent i must first be able to affect the representative of d. We distinguish between the following cases.

- If $y < x_i$, then agent i would have to report a position $x_i' < y$ to change the representative of d, but such a position cannot affect the final winner as the order of representatives remains the same (w would still be at the right of the representative for district d).

- If $y > w$, then agent i would have to report a position $x_i' > y$ to change the representative of d to x_i'. However, this again cannot affect the final winner as the order of representatives remains the same (w would still be at the left of the representative for district d).

- If $y \in [x_i, w]$, then agent i could potentially affect the outcome by reporting a position $x_i' > w$ to change the order of representatives, but this would lead to a higher individual cost as the new winner x_i' would be farther away.

Hence, agent i has no incentive to deviate, thus proving that the mechanism is strategyproof. □

5.2 Overview of Contribution

Our first contribution is the design of a novel mechanism that achieves an approximation ratio of 2 for the average cost; as mentioned above, this settles a question left open in the work of Filos-Ratsikas and Voudouris [2021] in the affirmative, matching their lower bound of 2. For the remaining objectives we provide mechanisms as well as lower bounds establishing that these mechanisms achieve the best possible approximation ratio. The precise bounds are shown in the first column of Table 5.1. Quite interestingly, and perhaps unexpectedly, our mechanism for the Average-of-Max objective is *optimal*, that is, it achieves an approximation ratio of 1. This demonstrates that for this particular objective, the distributed nature of the decision making does not influence the quality of the decision at all, and stands in contrast to the results of Anshelevich et al. [2022] for the same objective in the discrete setting.

Next, we consider strategyproof mechanisms, i.e., mechanisms that do not incentivize the agents to misreport their locations. This type of mechanisms were considered by Filos-Ratsikas and Voudouris [2021] who settled their approximation ratio for the social cost. For the remaining three objectives, strategyproof

	Unrestricted	Strategyproof
Average	2 (Section 5.3)	$3^{\star}$
Max	2 (Section 5.4)	2 (Section 5.4)
Average-of-Max	1 (Section 5.5.1)	$1+\sqrt{2}$ (Section 5.5.2)
Max-of-Average	2 (Section 5.6.1)	$1+\sqrt{2}$ (Section 5.6.2)

Table 5.1: Overview of our tight approximation ratio bounds for deterministic distributed mechanisms. The bound of 3 for the social cost and the class of strategyproof mechanisms marked with a $\star$ is due to Filos-Ratsikas and Voudouris [2021].

mechanisms have not been previously studied, not even in the discrete setting of Anshelevich et al. [2022]. We show tight bounds by carefully composing centralized statistics mechanisms for choosing the district representatives and the final location; in particular, depending on the objective at hand, we appropriately choose the values of two parameters p and q to define mechanisms that work by choosing the position of the q-th agent in a district as its representative, and then select the p-th representative as the output location. Our results for strategyproof mechanisms are shown in the second column of Table 5.1.

5.3 Average social cost

5.3.1 Unrestricted mechanisms

We begin with the social cost (Sum) objective. In previous work, Filos-Ratsikas and Voudouris [2021] showed that the Median-of-Medians mechanism (that is, the $\lfloor \lambda/2 \rfloor$-Statistic-of-$\lfloor k/2 \rfloor$-Statistic mechanism) has approximation ratio at most 3, and this is best possible strategyproof mechanism. For the class of unrestricted mechanisms, they showed a lower bound of 2, thus leaving a gap between 2 and 3. Here, we complete the picture by showing a tight bound of 2 for unrestricted mechanisms. We do this by considering the Weighted-Median-of-TruncatedAvg mechanism which works as follows: For each district, the mechanism considers a set of $n_d/2$ agents ranging from the $(n_d/4+1)$-th leftmost to the $(3n_d/4)$-th leftmost[1], and chooses their average as the representative location of the district. Then, it chooses the median representative location as the final location. See Mechanism 10 for a detailed description.

Mechanism 10: WEIGHTED-MEDIAN-OF-TRUNCATEDAVG

for each district $d \in D$ **do**

  $S_d := \{i \in N_d : i$ is at least the $(n_d/4+1)$-th and at most the $(3 \cdot n_d/4)$-th leftmost agent$\}$;

  $y_d := \frac{\sum_{i \in S_d} x_i}{|S_d|}$;

return $w :=$ Median$_{d \in D}\{y_d^{n_d}\}$;

To bound the approximation ratio of WEIGHTED-MEDIAN-OF-TRUNCATEDAVG, we characterize the structure of worst-case instances, where the approximation ratio of the mechanism is maximized and is strictly larger than 1. For such an instance I, let w be the location chosen by the mechanism when given as input a worst-case instance, and denote by o the optimal location; since the objective is the average social cost, o is the position of the median agent (or any point between the positions of the median agents in case of an even total number of agents). Without loss of generality, we assume that $w < o$; the case $w > o$ is symmetric.

We first show that there are cases where, starting from an instance with approximation ratio strictly larger than 1, moving particular agents to appropriate intervals, leads to new instances that have strictly worse approximation ratio. This transformation will be useful when characterizing the worst-case instances for the mechanism.

[1] For simplicity, we present the mechanism assuming that the number of agents in each district is a multiple of 4; extending the description of the mechanism and the proof is straightforward.

Lemma 5.3.1. *Let I and J be two instances that differ on the position of a single agent i, such that w is the location chosen by the mechanism when given any of the two instances as input, and o is the optimal location for I. The approximation ratio of the mechanism when given J as input is strictly larger than its approximation ratio when given I as input in the following cases:*

(a) i is positioned at $x_i < o$ in I, and at $x_i' \in (x_i, o]$ in J;

(b) i is positioned at $x_i > o$ in I, and at $x_i' \in [o, x_i)$ in J.

Proof. Since the optimal location o' for J satisfies the inequality $\text{cost}(o'|J) \leq \text{cost}(o|J)$, it suffices to show that

$$\frac{\text{cost}(w|I)}{\text{cost}(o|I)} < \frac{\text{cost}(w|J)}{\text{cost}(o|J)},$$

which would then imply that

$$\frac{\text{cost}(w|I)}{\text{cost}(o|I)} < \frac{\text{cost}(w|J)}{\text{cost}(o'|J)}.$$

For (a), we have that $\delta(x_i, w) \leq \delta(x_i, x_i') + \delta(x_i', w)$ by the triangle inequality, and also $\delta(x_i, o) = \delta(x_i, x_i') + \delta(x_i', o)$; recall our assumption that $w < o$. So,

$$\frac{\text{cost}(w|I)}{\text{cost}(o|I)} = \frac{\sum_{j \neq i} \delta(x_j, w) + \delta(x_i, w)}{\sum_{j \neq i} \delta(x_j, o) + \delta(x_i, o)} \leq \frac{\sum_{j \neq i} \delta(x_j, w) + \delta(x_i, x_i') + \delta(x_i', w)}{\sum_{j \neq i} \delta(x_j, o) + \delta(x_i, x_i') + \delta(x_i', o)}.$$

Since the approximation ratio of the mechanism when given I as input is strictly larger than 1 and the distances are non-negative, we can apply Inequality (5.1) with $\alpha = \sum_{j \neq i} \delta(x_j, w) + \delta(x_i', w)$, $\beta = \sum_{j \neq i} \delta(x_j, o) + \delta(x_i', o)$ and $\gamma =$

$\delta(x_i, x'_i)$, to obtain

$$\frac{\text{cost}(w|I)}{\text{cost}(o|I)} < \frac{\sum_{j\neq i}\delta(x_j, w) + \delta(x'_i, w)}{\sum_{j\neq i}\delta(x_j, o) + \delta(x'_i, o)} = \frac{\text{cost}(w|J)}{\text{cost}(o|J)}.$$

For (b), observe that $\delta(x_i, w) = \delta(x_i, x'_i) + \delta(x'_i, w)$ and $\delta(x_i, o) = \delta(x_i, x'_i) + \delta(x'_i, o)$. Therefore, the desired inequality again follows by appropriately applying Inequality (5.1). □

We are now ready to show that the worst-case instance I has the following properties:

- At least $k/2$ districts have representative w (Lemma 5.3.2);
- o can be the only other district representative and all agents in such districts are positioned at o (Lemma 5.3.3).

Lemma 5.3.2. *In the worst-case instance I, there are no district representatives to the left of w.*

Proof. Suppose towards a contradiction that the worst-case instance I is such that there is a district d with representative $y < w$. Since y is an average of some agent positions in d, there is a set of agents $S \subseteq S_d$ with $x_i \leq w$ for every $i \in S$. We move each agent $i \in S$ to a new position x'_i such that $x_i < x'_i \leq w$ and the truncated average of the agents in d becomes w. Clearly, the outcome of the mechanism, as well as the optimal location, remain the same in the new instance; w is still the median representative, and the position of the overall median agent did not change. By Lemma 5.3.1(a) and since $w < o$, moving any agent $i \in S$ to $x'_i \leq w$ leads to a new instance with strictly larger approximation ratio, which contradicts the fact that we start from a worst-case instance. □

Lemma 5.3.3. *In the worst-case instance I, besides w, the only other district representative can be o, and all agents in such districts are positioned on o.*

Proof. Suppose towards a contradiction that the worst-case instance I is such that there exists a district d with representative $y \notin \{w, o\}$. We move every agent $i \in N_d$ from x_i to $x'_i = o$. Hence, the truncated average of the agents in d changes from y to o. By Lemma 5.3.2 and since w is the (weighted) median representative, we have that at least half of the multiset defined by district representatives coincide with w. Consequently, the outcome of the mechanism is not affected when we move the agents of d. The optimal location also remains the same as the median agent location does not change. By Lemma 5.3.1, the approximation ratio of the new instance we obtain after moving each agent i (irrespective of whether $x_i < o$ or $x_i > o$) is strictly larger than the approximation ratio of instance I, contradicting the fact that it is a worst-case instance. □

We also argue that it suffices to focus on the case where the worst-case instance I consists of just two districts that are in fact symmetric; this will simplify the last part of our proof.

Lemma 5.3.4. *There exists a worst-case instance with two symmetric districts, one with representative w and one with representative o.*

Proof. Consider any worst-case instance, and let D_w and D_o denote the sets of districts that have representative w and o, respectively. We first argue that $\sum_{d \in D_w} n_d = \sum_{d \in D_o} n_d$. Note that since w is a median among all copies of representatives, we have $\sum_{d \in D_w} n_d \geq \sum_{d \in D_o} n_d$. Let us assume that $\sum_{d \in D_w} n_d > \sum_{d \in D_o} n_d$; we will reach a contradiction by creating a new instance, with strictly larger approximation ratio, that has one additional district with $\sum_{d \in D_w} n_d - \sum_{d \in D_o} n_d$ agents positioned at o. Clearly, in this new instance the mechanism again outputs w, while the optimal location remains o. Since the agents in the newly added district contribute 0 to the optimal cost and strictly greater than 0 to the social cost of w, the approximation ratio is strictly larger.

Now, since $\sum_{d \in D_w} n_d = \sum_{d \in D_o} n_d$ and all agents in districts with representative o are positioned at o (by Lemma 5.3.3), we can redistribute the agents in districts with representative o in a different set of districts, so that for any district $d \in D_w$ there is a dedicated district $d' \in D_o$ with $n_d = n_{d'}$. Note that w and o remain the same in this instance and so does the approximation ratio. We can then, without loss of generality, limit our focus on worst-case instances with just two symmetric districts, one with representative w and one with representative o. □

Having shown that it suffices to consider a worst-case instance with two symmetric districts, where district d_w has representative w while district d_o has all agents positioned at o, we now argue about the agent positions in d_w; recall that each district has size $\lambda = n/2$ in this case. Let ℓ and r be the locations of the $(\lambda/4+1)$- and $3\lambda/4$-leftmost agent, respectively, in d_w (i.e., the leftmost and rightmost location among agents in S_{d_w}). Clearly, it holds that $\ell \leq w \leq r$. We argue that $r \leq o$, and that all agents not in S_{d_w} are either at ℓ or at o.

Lemma 5.3.5. *In district d_w, $r \leq o$.*

Proof. Suppose towards a contradiction that the worst-case instance I is such that $r > o$ in d_w, and thus $\ell < w$. Let L be the set of agents in S_{d_w} that are positioned to the left of or at w, and R the set of agents in S_{d_w} that are positioned to the right of o. By the definition of w, for any set $Q \subseteq L$, we have

$$\begin{aligned} w &= \frac{2}{\lambda}\left(\sum_{i\in L} x_i + \sum_{i\in R} x_i + \sum_{i\in S_{d_w}\setminus(L\cup R)} x_i\right) \\ &= \frac{2}{\lambda}\left(\sum_{i\in L} x_i + \sum_{i\in R}(x_i - o) + \sum_{i\in R} o + \sum_{i\in S_{d_w}\setminus(L\cup R)} x_i\right) \\ &= \frac{2}{\lambda}\left(\sum_{i\in L\setminus Q} x_i + \sum_{i\in Q}\left(x_i + \frac{1}{|Q|}\sum_{j\in R}(x_j - o)\right) + \sum_{i\in R} o + \sum_{i\in S_{d_w}\setminus(L\cup R)} x_i\right). \end{aligned}$$

Consequently, there must exist a set $L_< \subseteq L$ such that $x_i + \frac{1}{|L_<|}\sum_{j\in R}(x_j - o) \leq w < o$ for every $i \in L_<$; if no such set exists, then the last expression above would be strictly larger than w. We obtain a new instance J by moving all agents in R from x_i to $x'_i = o$ and all agents in $L_<$ from x_i to $x'_i = x_i + \frac{1}{|L_<|}\sum_{j\in R}(x_j - o)$. Clearly, w is still the representative of d_w and o the optimal location. By Lemma 5.3.1, since all agents that moved are closer to o in J that in I, J must have approximation ratio strictly larger than I, a contradiction. □

Lemma 5.3.6. *In district d_w, the $\lambda/4$ leftmost agents are positioned at ℓ and the $\lambda/4$ rightmost agents are positioned at o.*

Proof. Assume otherwise and note that all these agents are not in S_{d_w} and, hence, do not affect w. By repeatedly applying Lemma 5.3.1 and moving each agent i with $x_i < \ell$ to ℓ and each agent i with $x_i > r$ at o, we reach an instance with strictly larger approximation ratio; a contradiction. □

We are finally ready to prove the main result of this section.

Theorem 5.3.7. *For Average, the approximation ratio of* Weighted-Median-of-TruncatedAvg *is at most* 2.

Proof. By Lemmas 5.3.3, 5.3.4 and 5.3.6, we have that the 2λ agents in the worst-case instance I are distributed on the line as follows: $\lambda/4$ agents are positioned at ℓ, $5\lambda/4$ agents are positioned at o (λ agents from d_o and $\lambda/4$ agents from d_w), and $\lambda/2$ agents are positioned in $[\ell, r]$. We partition the $\lambda/2$ agents in S_{d_w} into two sets: $L = \{i \in S_{d_w} : x_i \le w\}$ and $R = \{i \in S_{d_w} : x_i > w\}$. Since $r \le o$ (due to Lemma 5.3.5) and $w = \sum_{i \in L \cup R} x_i$ (by definition), the optimal cost is

$$\begin{aligned}
\text{cost}(o|I) &= \frac{1}{2\lambda}\left(\frac{\lambda}{4}(o-\ell) + \sum_{i\in L}(o-x_i) + \sum_{i\in R}(o-x_i)\right) \\
&= \frac{1}{2\lambda}\left(\frac{\lambda}{4}(o-\ell) + \frac{\lambda}{2}(o-w)\right) \\
&= \frac{1}{2\lambda}\left(\frac{\lambda}{4}(w-\ell) + \frac{3\lambda}{4}(o-w)\right). \qquad (5.2)
\end{aligned}$$

Similarly, the cost of the mechanism is

$$\text{cost}(w|I) = \frac{1}{2\lambda}\left(\frac{\lambda}{4}(w-\ell) + \sum_{i\in L}(w-x_i) + \sum_{i\in R}(x_i-w) + \frac{5\lambda}{4}(o-w)\right). \qquad (5.3)$$

By the definition of w, $\sum_{i\in L}(w-x_i) = \sum_{i\in R}(x_i-w)$. Also, again by definition, $|L| \ge 1$. If $R = \varnothing$, it must be the case that $\ell = w = r$, and the approximation ratio is at most $5/3$ as Equations (5.2) and (5.3) are simplified to $\text{cost}(o|I) = 3(o-w)/8$ and $\text{cost}(w|I) = 5(o-w)/8$, respectively. Hence, in the rest of the proof we will assume that $|R| \ge 1$.

Since $x_i \le o$ for each agent $i \in R$ and $|L| + |R| = \lambda/2$, we have

$$\sum_{i\in L}(w-x_i) = \sum_{i\in R}(x_i-w) \le |R|(o-w) \Leftrightarrow o-w \ge \frac{\sum_{i\in L}(w-x_i)}{\lambda/2-|L|}.$$

Similarly, as $x_i \geq \ell$ for each agent $i \in L$, we obtain

$$\sum_{i \in L} (w - x_i) \leq |L|(w - \ell) \Leftrightarrow w - \ell \geq \frac{\sum_{i \in L} (w - x_i)}{|L|}.$$

Let $o - w = \frac{\sum_{i \in L}(w - x_i)}{\frac{\lambda}{2} - |L|} + \xi_1$ and $w - \ell = \frac{\sum_{i \in L}(w - x_i)}{|L|} + \xi_2$, where $\xi_1, \xi_2 \geq 0$. Therefore, Equations (5.2) and (5.3) can be rewritten as

$$\text{cost}(o|I) = \frac{1}{2\lambda}\left(\frac{\lambda}{4}\left(\frac{\sum_{i \in L}(w - x_i)}{|L|} + \xi_2\right) + \frac{3\lambda}{4}\left(\frac{\sum_{i \in L}(w - x_i)}{\frac{\lambda}{2} - |L|} + \xi_1\right)\right)$$

$$\text{cost}(w|I) = \frac{1}{2\lambda}\left(\frac{\lambda}{4}\left(\frac{\sum_{i \in L}(w - x_i)}{|L|} + \xi_2\right) + \frac{5\lambda}{4}\left(\frac{\sum_{i \in L}(w - x_i)}{\frac{\lambda}{2} - |L|} + \xi_1\right)\right) + \frac{1}{2\lambda}\left(2\sum_{i \in L}(w - x_i)\right).$$

It is not hard to see that, unless the approximation ratio is at most $5/3$ and the claim holds trivially, the ratio is maximized when $\xi_1 = \xi_2 = 0$. We can then obtain the following upper bound on the approximation ratio.

$$\frac{\text{cost}(w|I)}{\text{cost}(o|I)} \leq \frac{\frac{\lambda}{4|L|} + \frac{5\lambda}{2\lambda - 4|L|} + 2}{\frac{\lambda}{4|L|} + \frac{3\lambda}{2\lambda - 4|L|}} \leq 2,$$

where the last inequality follows since $\frac{\lambda}{4|L|} + \frac{5\lambda}{2\lambda - 4|L|} + 2 \leq 2\left(\frac{\lambda}{4|L|} + \frac{3\lambda}{2\lambda - 4|L|}\right) \Leftrightarrow (\lambda - 4|L|)^2 \geq 0$. This concludes the proof. □

5.4 Max cost

We now consider the Max cost objective, for which we show a tight bound of 2 for both unrestricted and strategyproof mechanisms. For the upper bound, we consider the Arbitrary mechanism, which chooses the representative of each district to be the position of any agent therein, and then chooses any representative as the final winner. See Mechanism 11 for a specific implementation of this mechanism using the position of the leftmost agent from each district as the district representative, and then the leftmost representative as the final winner. Clearly, ARBITRARY is equivalent to some p-Statistic-of-q-Statistic mechanism depending on the choices within and over districts; for example, the particular implementation of ARBITRARY as Mechanism 11 is equivalent to 1-Statistic-of-1-Statistic.

Mechanism 11: ARBITRARY (LEFTMOST-OF-LEFTMOST)

---

for each district d **do**
  $y_d := \min_{i \in N_d}\{x_i\}$;
return $w := \min_d\{y_d\}$;

Theorem 5.4.1. *For Max, the approximation ratio of* ARBITRARY *is at most* 2.

Proof. Given any instance, let ℓ and r denote the positions of the leftmost and the rightmost agent, respectively. Clearly, the optimal location is $o = \frac{r-\ell}{2}$, and thus $\text{cost}(o) = \frac{r-\ell}{2}$. On the other hand, the ARBITRARY mechanism will necessarily return the location of some agent as the winner w, and hence $\text{cost}(w) \leq r - \ell$; the claim follows. □

We also show a matching lower bound for all mechanisms, thus completing the picture.

Theorem 5.4.2. *For Max, the approximation ratio of any mechanism (unrestricted or strategyproof) is at least* 2.

Proof. Consider any mechanism and the following instance I with two districts. The agents in the first district are all positioned at -1, while the agents in the second district are all positioned at 1. Due to unanimity (Lemma 5.1.1), the representatives of the two districts must be -1 and 1, respectively. Hence, the winner is either -1 or 1. However, $\text{cost}(-1|I) = \text{cost}(1|I) = 2$, whereas $\text{cost}(0|I) = 1$, leading to a approximation ratio of 2. $\square$

5.5 Average-of-Max

Here, we focus on the Average-of-Max objective; recall that this objective is the average sum over each district of the maximum agent cost therein. For unrestricted mechanisms, we show that it is possible to compute the optimal location (and thus achieve a approximation ratio of 1), whereas, for strategyproof mechanisms, we show a tight approximation ratio bound of $1 + \sqrt{2}$.

5.5.1 Unrestricted mechanisms

We will show that the MEDIAN-OF-MIDPOINTS mechanism optimizes the Average-of-Max objective. This mechanism chooses the representative of each district to be the midpoint of the interval defined by the positions of the agents therein, and then chooses the median representative (breaking ties in favor of the leftmost median in case there are two) as the final winner. See Mechanism 12 for a detailed description.

Mechanism 12: MEDIAN-OF-MIDPOINTS

```
for each district d do
    y_d := 1/2 · ( max_{i∈N_d} x_i + min_{i∈N_d} x_i ) ;
return w := Median_{d∈D}{y_d} ;
```

Theorem 5.5.1. *For Average-of-Max, the approximation ratio of* MEDIAN-OF-MIDPOINTS *is* 1.

Proof. For any district d, let ℓ_d and r_d be the (positions of the) leftmost and rightmost agents therein, respectively. The Average-of-Max cost of an arbitrary location z is

$$\begin{aligned}\frac{1}{k}\sum_{d\in D}\max_{i\in N_d}\delta(x_i,z) &= \frac{1}{k}\sum_{d\in D}\max\{\delta(\ell_d,z),\delta(r_d,z)\}\\ &= \frac{1}{k}\sum_{d\in D}\delta\left(\frac{\ell_d+r_d}{2},z\right)+\frac{1}{k}\sum_{d\in D}\frac{r_d-\ell_d}{2}.\end{aligned}$$

Since the second term is a constant in terms of z, the above expression is minimized when the first term is minimized, which is done when z is chosen to minimize the average distance from the midpoints of the intervals defined by the agents in each district. Consequently, it suffices to choose the median district midpoint as the winner. This is precisely the definition of MEDIAN-OF-MIDPOINTS. □

5.5.2 Strategyproof mechanisms

For strategyproof mechanisms, we will show a tight bound of $1+\sqrt{2}$. For the upper bound, we consider the $\left(1-1/\sqrt{2}\right)k$-LEFTMOST-OF-RIGHTMOST mechanism, which chooses the representative of each district to be the position of the rightmost agent therein, and then chooses the $\lceil\left(1-1/\sqrt{2}\right)k\rceil$-th leftmost representative as the final winner. See Mechanism 13 for a detailed description. Clearly, the mechanism is an implementation of p-Statistic-of-q-Statistic with $p=\lceil\left(1-1/\sqrt{2}\right)k\rceil$ and $q_d=n_d$, and is thus strategyproof. So, it suffices to show that it achieves a approximation ratio of at most $1+\sqrt{2}$.

Theorem 5.5.2. *For the Average-of-Max cost, the approximation ratio of* $\left(1-1/\sqrt{2}\right)k$-LEFTMOST-OF-RIGHTMOST *is at most* $1+\sqrt{2}$*.*

Proof. Let w be the location chosen by the mechanism when given some in-

Mechanism 13: $\left(1-1/\sqrt{2}\right)$ k-LEFTMOST-OF-RIGHTMOST

for each district $d \in D$ **do**
  $y_d :=$ rightmost agent;
return $w := \lceil \left(1-1/\sqrt{2}\right) k \rceil$-th leftmost representative;

stance as input, and o an optimal location. For each district d, let i_d be the most distant agent from w, and i_d^* the most distant agent from o. So, $\text{cost}(w|I) = \frac{1}{k}\sum_{d \in D} \delta(i_d, w)$, and $\text{cost}(o|I) = \frac{1}{k}\sum_{d \in D} \delta(i_d^*, o) \geq \frac{1}{k}\sum_{d \in D} \delta(j, o)$ for any agent $j \in N_d$. We consider the following two cases depending on the relative positions of w and o.

Case 1: $o < w$**.**

Let $S = \{d \in D : y_d \geq w\}$ be the set of district representatives to the right of w. By the definition of w, we have that $|S| = k + 1 - \lceil \left(1-1/\sqrt{2}\right) k \rceil = 1 + \lfloor k/\sqrt{2} \rfloor \geq \frac{k}{\sqrt{2}}$. Since $o < w \leq y_d$ for every $d \in S$ and $y_d \in N_d$, we have that

$$\begin{aligned}
\text{cost}(o) &\geq \frac{1}{k}\sum_{d \in S} \delta(y_d, o) \\
&\geq \frac{1}{k} \cdot |S| \cdot \delta(w, o) \\
&\geq \frac{1}{\sqrt{2}} \cdot \delta(w, o) \Leftrightarrow \delta(w, o) \leq \sqrt{2} \cdot \text{cost}(o).
\end{aligned}$$

By the triangle inequality and since $i_d \in N_d$, we have

$$\begin{aligned}
\text{cost}(w) &= \frac{1}{k}\sum_{d \in D} \delta(i_d, w) \\
&\leq \frac{1}{k}\sum_{d \in D} \delta(i_d, o) + \frac{1}{k}\sum_{d \in D} \delta(w, o)
\end{aligned}$$

$$\leq \text{cost}(o) + \delta(w, o)$$
$$\leq (1 + \sqrt{2}) \cdot \text{cost}(o).$$

Case 2: $w < o$**.**

We partition the districts into a set L that includes $\lceil (1 - \frac{1}{\sqrt{2}})k \rceil$ districts from the one with the leftmost representative until the one with the $\lceil (1 - \frac{1}{\sqrt{2}})k \rceil$-th leftmost representative (that is, w), and a set R that includes the remaining districts. By definition, we have that $|R|/|L| = (k - \lceil (1 - \frac{1}{\sqrt{2}})k \rceil)/(\lceil (1 - \frac{1}{\sqrt{2}})k \rceil) \leq 1 + \sqrt{2}$. For every district d, let ℓ_d and r_d be the leftmost and rightmost agents in d, respectively. We make the following observations:

- For every $d \in L$, since y_d is the rightmost agent of d and $y_d \leq w < o$, it must be the case that $i_d = i_d^* = \ell_d$. Due to the positions of ℓ_d, w and o, we have that $\delta(\ell_d, o) = \delta(\ell_d, w) + \delta(w, o)$.
- For every $d \in R$, by the triangle inequality, we have that $\delta(i_d, w) \leq \delta(i_d, o) + \delta(w, o)$. Since $\delta(i_d, o) \leq \delta(i_d^*, o)$ by the definition of i_d^*, we further have that $\delta(i_d, w) \leq \delta(i_d^*, o) + \delta(w, o)$.

Hence,

$$\begin{aligned}
\text{cost}(w) &= \frac{1}{k} \sum_{d \in D} \delta(i_d, w) = \frac{1}{k} \sum_{d \in L} \delta(\ell_d, w) + \frac{1}{k} \sum_{d \in R} \delta(i_d, w) \\
&\leq \frac{1}{k} \sum_{d \in L} \Big(\delta(\ell_d, w) + \delta(w, o) \Big) - \frac{|L|}{k} \cdot \delta(w, o) + \frac{1}{k} \sum_{d \in R} \Big(\delta(i_d^*, o) + \delta(w, o) \Big) \\
&= \text{cost}(o) + \frac{|R| - |L|}{k} \cdot \delta(w, o).
\end{aligned}$$

Since $y_d \leq w < o$ for every $d \in L$ and $y_d \in N_d$, we have that

$$\text{cost}(o) \geq \frac{1}{k}\sum_{d\in L}\delta(y_d, o) \geq \frac{|L|}{k}\cdot\delta(w,o) \Leftrightarrow \delta(w,o) \leq \frac{k}{|L|}\cdot\text{cost}(o).$$

Therefore, we obtain

$$\text{cost}(w) \leq \text{cost}(o) + \frac{|R|-|L|}{|L|}\cdot\text{cost}(o) = \frac{|R|}{|L|}\cdot\text{cost}(o) \leq (1+\sqrt{2})\cdot\text{cost}(o),$$

as desired. □

We also show a matching lower bound on the approximation ratio of all strategyproof mechanisms.

Theorem 5.5.3. *For Average-of-Max, the approximation ratio of any strategyproof mechanism is at least $1+\sqrt{2}-\varepsilon$, for any $\varepsilon > 0$.*

Proof. Assume towards a contradiction that there exists some $\varepsilon > 0$ and a strategyproof mechanism with approximation ratio strictly smaller than $1+\sqrt{2}-\varepsilon$. Without loss of generality, we assume that when there are two symmetric districts with different representatives, we choose the leftmost as the final winner. We will prove the statement by showing some properties about the behavior of strategyproof mechanisms in particular instances.

Property (P1): We claim that there is a district with two agents such that the mechanism chooses some agent position as the district representative. Consider a district d with one agent positioned at x and one agent positioned at $y > x$. If the mechanism chooses the representative to be x or y, then we are done. Otherwise, suppose that the representative is chosen to be some $z \notin \{x, y\}$. Due to strategyproofness, z must also be the representative of the district d' where any of the two agents has been moved to z; otherwise, in the single-district instance consisting of d', the agent that is moved would have incentive to report that she is positioned as in d to change the outcome to z.

Property (P2): By Property (P1) there exists a district with two agents such that the mechanism chooses the district representative to be the position of one of the agents; without loss of generality we assume that the agents are positioned at 0 and 1. We claim that the representative of this district must be 1 as otherwise the approximation ratio would be at least 3. Indeed, suppose otherwise that the representative is 0, and consider the following instance I_1 with two districts:

- In the first district, there is an agent at 0 and an agent at 1. By the above discussion, the representative is 0.
- In the second district, there are two agents at $1/2$. Due to unanimity, the representative is $1/2$ (otherwise the approximation ratio would be infinite due to Lemma 5.1.1).

Since there are only two districts and two different representatives, the overall winner is 0. But,

$$\text{cost}(0|I_1) = \frac{1}{2}\Big((1-0) + (1/2-0)\Big) = 3/4$$

and

$$\text{cost}(1/2|I_1) = \frac{1}{2}\Big((1-1/2) + (1/2-1/2)\Big) = 1/4,$$

leading to a approximation ratio of 3.

Property (P3): Let $\alpha < \beta$ be two (large) integers such that $\beta/\alpha = 1+\sqrt{2}-\delta$, for some arbitrarily small $\delta > 0$. We claim that in instances with $\alpha+\beta$ districts such that $1/2$ is the representative of α districts and 1 is the representative of β districts, the overall winner must be 1 as otherwise the approximation ratio would be $\beta/\alpha = 1+\sqrt{2}-\delta$. Indeed, suppose that the winner is $1/2$ in such a case, and consider the following instance I_2 with $\alpha+\beta$ districts:

- In α districts, there are two agents at $1/2$.
- In β districts, there are two agents at 1.

Due to unanimity (Lemma 5.1.1), the representatives are $1/2$ and 1, respectively, and the overall winner is $1/2$ by assumption. Then, $\text{cost}(1/2|I_2) = \frac{1}{2}\cdot\beta/2$ and $\text{cost}(1|I_2) = \frac{1}{2}\cdot\alpha/2$. So, the approximation ratio is at least $\beta/\alpha = 1+\sqrt{2}-\delta$.

Reaching a contradiction: Now, we consider the following instance I_3 with $\alpha+\beta$ districts:

- In α districts, there are two agents at $1/2$. Due to unanimity the representative of all these districts is $1/2$.
- In β districts, there is one agent at 0 and one agent at 1. By property (P2), the representative of all these districts is 1.

Since $1/2$ is the representative of α districts and 1 is the representative of β districts, by property (P3), the overall winner is 1. We have that

$$\text{cost}(1|I_3) = \frac{1}{2}\left(\frac{\alpha}{2}+\beta\right)$$

and

$$\text{cost}(1/2|I_3) = \frac{1}{2}\cdot\frac{\beta}{2}.$$

That is, the approximation ratio is at least $2+\frac{\alpha}{\beta} > 2+\frac{1}{1+\sqrt{2}} = 1+\sqrt{2}$; a contradiction. $\square$

5.6 Max-of-Average

We now turn our attention to the last objective, Max-of-Average, which is the maximum over each district of the average total individual cost therein. We show a tight bound of 2 for unrestricted mechanisms and a tight bound of $1+\sqrt{2}$ for strategyproof mechanisms.

5.6.1 Unrestricted mechanisms

Since the lower bound of 2 for the Max cost objective holds even when there is a single agent in each district, it extends to the case of Max-of-Average as well. For the upper bound, we consider the ARBITRARY-OF-AVG mechanism, which chooses the representative of each district to be the average of the positions of the agents in the district, and then chooses an arbitrary representative (e.g., the leftmost) as the final winner. See Mechanism 14 for a detailed description.

Mechanism 14: ARBITRARY-OF-AVG

for each district $d \in D$ **do**
  $y_d := \frac{\sum_{i \in N_d} x_i}{n_d}$;
return $w := \min_{d \in D} y_d$;

Theorem 5.6.1. *For Max-of-Average, the approximation ratio of* ARBITRARY-OF-AVG *is at most* 2.

Proof. Let w be the location chosen by the mechanism when given some instance as input, and o an optimal location; without loss of generality, we assume that $w < o$. Denote by d^* a district that defines the cost of w, that is, $d^* \in \arg\max_{d \in D} \frac{1}{n_d} \sum_{i \in N_d} \delta(x_i, w)$. Also, denote by d_w a district that has representative w, that is,

$$w = \frac{1}{n_{d_w}} \sum_{i \in N_{d_w}} x_i \Leftrightarrow \frac{1}{n_{d_w}} \sum_{i \in N_{d_w}} (w - x_i) = 0.$$

By the triangle inequality, we have that

$$\begin{aligned} \text{cost}(w) &= \frac{1}{n_{d^*}} \sum_{i \in N_{d^*}} \delta(x_i, w) \\ &\leq \frac{1}{n_{d^*}} \sum_{i \in N_{d^*}} \delta(x_i, o) + \frac{1}{n_{d^*}} \sum_{i \in N_{d^*}} \delta(w, o) \\ &\leq \text{cost}(o) + \delta(w, o). \end{aligned}$$

By the definition of d_w, we have that

$$\begin{aligned} \delta(w, o) &= o - w \\ &= o - w + \frac{1}{n_{d_w}} \sum_{i \in N_{d_w}} (w - x_i) \\ &= \frac{1}{n_{d_w}} \sum_{i \in N_{d_w}} (o - x_i) \\ &\leq \frac{1}{n_{d_w}} \sum_{i \in N_{d_w}} \delta(x_i, o) \\ &\leq \text{cost}(o), \end{aligned}$$

where the inequality follows since $\delta(x_i, o) = o - x_i$ when $x_i \leq o$ and $\delta(x_i, o) = x_i - o \geq o - x_i$ when $x_i \geq o$. Therefore, we obtain that $\text{cost}(w) \leq 2 \cdot \text{cost}(o)$, as desired. □

5.6.2 Strategyproof mechanisms

We now turn out attention to strategyproof mechanisms and show a tight bound of $1+\sqrt{2}$. For the upper bound, we consider the RIGHTMOST-OF-$\left(1 - 1/\sqrt{2}\right) n_d$-LEFTMOST mechanism, which chooses the representative of each district d to be the position of the $\lceil \left(1 - 1/\sqrt{2}\right) n_d \rceil$-th leftmost agent therein, and then chooses the rightmost representative as the final winner. See Mechanism 15 for a detailed description. This mechanism is an implementation of p-Statistic-of-q_d-Statistic with $p = k$ and $q_d = \lceil \left(1 - 1/\sqrt{2}\right) n_d \rceil$, and is thus strategyproof. So, it suffices to show that it achieves a approximation ratio of at most $1 + \sqrt{2}$.

Mechanism 15: RIGHTMOST-OF-$\left(1 - 1/\sqrt{2}\right) n_d$-LEFTMOST

---

for each district $d \in D$ **do**
  $y_d := \lceil \left(1 - 1/\sqrt{2}\right) n_d \rceil$-th leftmost agent;
return $w :=$ rightmost representative;

Theorem 5.6.2. *For the Max-of-Average cost, the approximation ratio of* RIGHTMOST-OF-$\left(1 - 1/\sqrt{2}\right) n_d$-LEFTMOST *is at most* $1 + \sqrt{2}$.

Proof. Let w be the location chosen be the mechanism when given some instance as input, and o an optimal location. Denote by d^* a district that gives the max average sum for w, and by d_w a district with representative w. Also, for any district d, we denote by $\text{cost}_d(x) = \frac{1}{n_d} \sum_{i \in N_d} \delta(i, x)$ the average total distance of the agents in d from location x, and let o_d be the location that minimizes this distance (that is, o_d is the median agent of d). Clearly, by definition, we have that $\text{cost}(w) = \text{cost}_{d^*}(w)$, and $\text{cost}_d(o) \leq \text{cost}(o)$ for every district d. We consider the following two cases:

Case 1: $o < w$.

By the definition of d^* and the triangle inequality, we have

$$\begin{aligned}\text{cost}(w) &= \frac{1}{n_{d^*}} \sum_{i \in N_{d^*}} \delta(i,w) \\ &\leq \frac{1}{n_{d^*}} \sum_{i \in N_{d^*}} \delta(i,o) + \frac{1}{n_{d^*}} \sum_{i \in N_{d^*}} \delta(o,w) \\ &\leq \text{cost}(o) + \delta(o,w).\end{aligned}$$

Let $S = \{i \in N_{d_w} : x_i \geq w\}$ be the set of agents that are positioned at the right of (or exactly at) w in d_w. Since $o < w$, $\delta(i,o) \geq \delta(w,o)$ for every $i \in S$. Also, by the definition of w, $|S| = n_{d_w} + 1 - \lceil \left(1 - 1/\sqrt{2}\right) n_{d_w} \rceil = 1 + \lfloor \frac{1}{\sqrt{2}} \cdot n_{d_w} \rfloor \geq \frac{1}{\sqrt{2}} \cdot n_{d_w}$. Hence,

$$\begin{aligned}&\text{cost}_{d_w}(o) = \frac{1}{n_{d_w}} \sum_{i \in N_{d_w}} \delta(i,o) \geq \frac{1}{n_{d_w}} \cdot |S| \cdot \delta(w,o) \geq \frac{1}{\sqrt{2}} \cdot \delta(w,o) \\ &\Leftrightarrow \delta(w,o) \leq \sqrt{2} \cdot \text{cost}_{d_w}(o) \leq \sqrt{2} \cdot \text{cost}(o).\end{aligned}$$

By combining everything together, we obtain a bound of $1 + \sqrt{2}$.

Case 2: $w < o$.

We consider the following two subcases:

- $o_{d^*} \leq w < o$. By the monotonicity of the (average) social cost[2] for the agents in district d^*, we have that $\text{cost}_{d^*}(o_{d^*}) \leq \text{cost}_{d^*}(w) \leq \text{cost}_{d^*}(o)$, and thus $\text{cost}(w) \leq \text{cost}(o)$.

- $w < o_{d^*}$. Since w is the rightmost representative, it must be the case that $y_{d^*} \leq w < o_{d^*}$. So, again by the monotonicity of the (average) social cost

[2]It is a well-known fact that the social cost objective is monotone in the locations. In particular, for any set of agents S, if $y_1 \in \arg\min_x \sum_{i \in S} \delta(i,x)$, then $\sum_{i \in S} \delta(i,y_1) \leq \sum_{i \in S} \delta(i,y_2) \leq \sum_{i \in S} \delta(i,y_3)$ for any $y_1 \leq y_2 \leq y_3$ or $y_3 \leq y_2 \leq y_1$.

within the district d^*, we have that $\text{cost}_{d^*}(o_{d^*}) \leq \text{cost}_{d^*}(w) \leq \text{cost}_{d^*}(y_{d^*})$. We will argue that $\text{cost}_{d^*}(y_{d^*}) \leq (1+\sqrt{2}) \cdot \text{cost}_{d^*}(o_{d^*})$.

Let L be the set that includes $\lceil (1-\frac{1}{\sqrt{2}})n_{d^*} \rceil$ agents of d^* from the leftmost to the $\lceil (1-\frac{1}{\sqrt{2}})n_{d^*} \rceil$-th leftmost agent (that is, y_{d^*}), and the set R that includes the remaining agents. By definition, we have that $|R|/|L| = (n_{d^*} - \lceil (1-\frac{1}{\sqrt{2}})n_{d^*} \rceil)/\lceil (1-\frac{1}{\sqrt{2}})n_{d^*} \rceil \leq 1+\sqrt{2}$. Now, observe that

- For every agent $i \in L, i \leq y_{d^*}$, and thus $\delta(i, o_{d^*}) = \delta(i, y_{d^*}) + \delta(y_{d^*}, o_{d^*})$.
- For every agent $i \in R$, $i \geq y_{d^*}$, and thus $\delta(i, y_{d^*}) \leq \delta(i, o_{d^*}) + \delta(y_{d^*}, o_{d^*})$.

Hence,

$$\begin{aligned}
\text{cost}_{d^*}(y_{d^*}) &= \frac{1}{n_{d^*}} \sum_{i \in N_{d^*}} \delta(i, y_{d^*}) \\
&= \frac{1}{n_{d^*}} \sum_{i \in L} \delta(i, y_{d^*}) + \frac{1}{n_{d^*}} \sum_{i \in R} \delta(i, y_{d^*}) \\
&\leq \frac{1}{n_{d^*}} \sum_{i \in L} \delta(i, y_{d^*}) + \frac{1}{n_{d^*}} \sum_{i \in R} \Big(\delta(i, o_{d^*}) + \delta(y_{d^*}, o_{d^*}) \Big) \\
&= \frac{1}{n_{d^*}} \sum_{i \in L} \Big(\delta(i, y_{d^*}) + \delta(y_{d^*}, o_{d^*}) \Big) + \frac{1}{n_{d^*}} \sum_{i \in R} \delta(i, o_{d^*}) \\
&+ \frac{|R| - |L|}{n_{d^*}} \cdot \delta(y_{d^*}, o_{d^*}) \\
&= \text{cost}_{d^*}(o_{d^*}) + \frac{|R| - |L|}{n_{d^*}} \cdot \delta(y_{d^*}, o_{d^*}).
\end{aligned}$$

Since $y_{d^*} < o_{d^*}$, we also have that $\text{cost}_{d^*}(o) \geq \frac{1}{n_{d^*}} \cdot |L| \cdot \delta(y_{d^*}, o_{d^*})$, and thus

$$\begin{aligned}
\text{cost}_{d^*}(y_{d^*}) &\leq \text{cost}_{d^*}(o_{d^*}) + \frac{|R| - |L|}{|L|} \cdot \text{cost}_{d^*}(o_{d^*}) \\
&= \frac{|R|}{|L|} \cdot \text{cost}_{d^*}(o_{d^*})
\end{aligned}$$

$$\leq (1+\sqrt{2}) \cdot \text{cost}_{d^*}(o_{d^*}).$$

From this, we finally get that

$$\text{cost}_{d^*}(w) \leq (1+\sqrt{2}) \cdot \text{cost}_{d^*}(o_{d^*}) \leq (1+\sqrt{2}) \cdot \text{cost}_{d^*}(o),$$

and thus $\text{cost}(w) \leq (1+\sqrt{2})\text{cost}(o)$. □

Finally, we show a matching lower bound for any strategyproof mechanism.

Theorem 5.6.3. *For Max-of-Average, the approximation ratio of any strategyproof mechanism is at least $1+\sqrt{2}-\varepsilon$, for any $\varepsilon > 0$.*

Proof. Suppose towards a contradiction that there is a strategyproof mechanism with approximation ratio strictly smaller than $1+\sqrt{2}-\varepsilon$, for any $\varepsilon > 0$. We will reach a contradiction by showing several properties that any strategyproof mechanism must satisfy when given particular instances with symmetric districts consisting of $\lambda = (2+\sqrt{2})x$ agents, where x is an arbitrarily large integer.

Property (P1): Consider a district with $(1+\sqrt{2})x$ agents at 0 and x agents at 1.[3] We claim that the mechanism must choose 0 as the representative of this district as otherwise the approximation ratio would be at least $1+\sqrt{2}$. Indeed, suppose that the representative is some $y \neq 0$. By moving one of the agents at 1 to y, we obtain a new district whose representative must still be y; otherwise, in the instance that consists only of this new district, the agent at y would have incentive to misreport her position as 1, thus leading to the representative (and the final winner) to change to y. By induction, we obtain that y must be the representative of the district with $(1+\sqrt{2})x$ agents at 0 and x agents at y. In the

[3] To be precise, since the number of agents must be an integer, we would need to have $\lceil (1+\sqrt{2})x \rceil$ agents at 0. We simplify our notation by dropping the ceilings, but it should be clear that this does not affect our arguments.

instance I that consists of only the latter district, the winner is y with $\text{cost}(y|I) = \frac{1}{\lambda} \cdot (1+\sqrt{2})x \cdot |y|$, whereas $\text{cost}(0|I) = \frac{1}{\lambda} \cdot x \cdot |y|$, leading to a approximation ratio of at least $1+\sqrt{2}$.

Property (P2): Consider a district with x agents at 1 and $(1+\sqrt{2})x$ agents at 2. We claim that the mechanism must choose 2 as the representative of this district as otherwise the approximation ratio would be at least $1+\sqrt{2}$. This follows by arguments similar to those for property (P1).

Reaching a contradiction: Consider the following instance J with two districts:

- In the first district, there are $(1+\sqrt{2})x$ agents at 0 and x agents at 1.
- In the second district, there are x agents at 1 and $(1+\sqrt{2})x$ agents at 2.

By properties (P1) and (P2), the representatives of the two districts must be 0 and 2, respectively, and thus one of these two locations is chosen as the final winner. However, $\text{cost}(0|J) = \text{cost}(2|J) = \frac{1}{\lambda} \cdot \Big(2(1+\sqrt{2})x + x\Big)$, while $\text{cost}(1|J) = \frac{1}{\lambda} \cdot (1+\sqrt{2})x$, leading to a approximation ratio of $2 + \frac{1}{1+\sqrt{2}} = 1+\sqrt{2}$. $\square$

Chapter 6

Conclusion and Directions for Future Work

This thesis examined the intersection of artificial intelligence and economics through the lens of mechanism design in facility location problems. By exploring sophisticated theoretical models across three distinctive chapters, this work not only contributes to advancing the current understanding of approximate mechanism design without monetary incentives but also sets the stage for resolving complex, real-world issues in strategic multi-agent environments.

In Chapter 3, we achieve a small constant bound for the social cost, significantly improving upon the previously established linear bound. It is important to note, however, that a considerable gap remains between the established lower bound of $\frac{4}{3}$ and the upper bound of $\frac{17}{4}$. This discrepancy calls for further exploration, particularly in terms of randomized mechanisms and the maximum cost objective, given the existing disparity between the lower bound of $\frac{4}{3}$ and the upper bound of $\frac{3}{2}$, as identified by Serafino and Ventre [2016].

In Chapter 4, we studied a truthful two-facility location problem with candidate locations and showed tight bounds on the best possible approximation ratio of deterministic strategyproof mechanisms in terms of the social cost and the max cost.

In Chapter 5, we considered two classes of mechanisms: *Unrestricted* mechanisms which assume that the agents directly provide their true positions as input, and *strategyproof* mechanisms which deal with strategic agents and aim to incentivize them to truthfully report their positions. For both classes, we show tight bounds on the best possible approximation in terms of several minimization social objectives, including the well-known average social cost (average total distance of agents from the chosen point) and max cost (maximum distance among all agents from the chosen point), as well as other fairness-inspired objectives that are tailor-made for the distributed setting, in particular, the max-of-average and the average-of-max.

While our research has yielded valuable insights, it is essential to acknowledge the numerous open questions and future research directions that remain to be explored. To guide forthcoming inquiries, we pose the following pivotal questions:

Randomized Mechanisms: Investigating the role of randomized mechanisms in achieving improved approximation guarantees presents a promising research direction. Randomized strategies may offer solutions where deterministic approaches are limited, especially in scenarios involving higher complexity and greater agent diversity. Exploring these possibilities could lead to more robust models capable of handling unpredictability and variability in agent behaviors and preferences.

New Information Assumptions: Altering the dynamics of public and private information within the models studied, particularly those in Chapter 3 and 4, could have significant impacts. This exploration might facilitate the development of more versatile and adaptable strategies in mechanism design, catering to diverse and evolving environmental constraints. It would be worthwhile to examine how different information assumptions affect the strategic interactions and outcomes within these models.

New Social Objectives: Defining and implementing additional meaningful social objectives within the frameworks discussed in Chapter 5 could provide deeper insights into fairness and efficiency, especially in distributed and decentralized systems. This effort could help in designing mechanisms that not only optimize cost but also enhance societal welfare in a more balanced manner.

Complex Topological Extensions: Extending these models to more complex topological structures such as trees or regular graphs poses theoretical and practical challenges. Addressing these challenges could significantly impact network design and infrastructure development, providing insights into more efficient and equitable facility placement across diverse geographical and social landscapes.

Real-World Implementations and Applications: Bridging the gap between theoretical models and practical implementations, especially in fields such as urban planning and public goods allocation, remains crucial. Focused studies on transforming theoretical models into actionable strategies could directly benefit society by improving the allocation of public resources and enhancing the quality of urban and rural infrastructures.

Generally, the future research directions proposed here are designed to leverage the solid theoretical foundation established by this work, aiming to inspire innovative solutions that can be adapted to various real-world contexts. Continued exploration of these questions is expected to yield substantial contributions to the fields of economics, computer science, and beyond, potentially leading to significant societal impacts.

Bibliography

Gagan Aggarwal, Ashish Goel, and Rajeev Motwani. Truthful auctions for pricing search keywords. In *Proceedings of the 7th ACM Conference on Electronic Commerce*, pages 1–7, 2006.

Noga Alon, Michal Feldman, Ariel D Procaccia, and Moshe Tennenholtz. Strategyproof approximation of the minimax on networks. *Math. Oper. Res.*, 35(3): 513–526, 2010.

Georgios Amanatidis, Georgios Birmpas, George Christodoulou, and Evangelos Markakis. Truthful allocation mechanisms without payments: Characterization and implications on fairness. In *Proceedings of the 2017 ACM Conference on Economics and Computation*, pages 545–562, 2017.

Georgios Amanatidis, Georgios Birmpas, Federico Fusco, Philip Lazos, Stefano Leonardi, and Rebecca Reiffenhäuser. Allocating indivisible goods to strategic agents: Pure nash equilibria and fairness. *Mathematics of Operations Research*, 2023.

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, and Alexandros A. Voudouris. The distortion of distributed metric social choice. *Artificial Intelligence*, 308:103713, 2022.

Itai Ashlagi, Felix Fischer, Ian A Kash, and Ariel D Procaccia. Mix and match: A strategyproof mechanism for multi-hospital kidney exchange. *Games and Economic Behavior*, 91:284–296, 2015.

Sepehr Assadi and Sahil Singla. Improved truthful mechanisms for combinatorial auctions with submodular bidders. *ACM SIGecom Exchanges*, 18(1):19–27, 2020.

Sepehr Assadi, Thomas Kesselheim, and Sahil Singla. Improved truthful mechanisms for subadditive combinatorial auctions: Breaking the logarithmic barrier. In *Proceedings of the 2021 ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 653–661. SIAM, 2021.

Haris Aziz and Alexander Lam. Obvious manipulability of voting rules. In *Algorithmic Decision Theory: 7th International Conference, ADT 2021, Toulouse, France, November 3–5, 2021, Proceedings 7*, pages 179–193. Springer, 2021.

Siddhartha Banerjee, Vasilis Gkatzelis, Artur Gorokh, and Billy Jin. Online nash social welfare maximization with predictions. In *Proceedings of the 2022 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1–19. SIAM, 2022.

Xiaohui Bei, Guangda Huzhang, and Warut Suksompong. Truthful fair division without free disposal. *Social Choice and Welfare*, 55:523–545, 2020.

Patrick Briest, Piotr Krysta, and Berthold Vöcking. Approximation techniques for utilitarian mechanism design. In *Proceedings of the thirty-seventh annual ACM symposium on Theory of computing*, pages 39–48, 2005.

Ioannis Caragiannis, Aris Filos-Ratsikas, and Ariel D Procaccia. An improved 2-agent kidney exchange mechanism. In *Internet and Network Economics: 7th*

International Workshop, WINE 2011, Singapore, December 11-14, 2011. Proceedings 7, pages 37–48. Springer, 2011.

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 (COCOA)*, 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, 2013a.

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

Richard L Church and Robert S Garfinkel. Locating an obnoxious facility on a network. *Transportation science*, 12(2):107–118, 1978.

Richard Cole, Vasilis Gkatzelis, and Gagan Goel. Mechanism design for fair division: allocating divisible items without payments. In *Proceedings of the fourteenth ACM conference on Electronic commerce*, pages 251–268, 2013.

Argyrios Deligkas, Aris Filos-Ratsikas, and Alexandros A. Voudouris. Heterogeneous facility location with limited resources. *Games and Economic Behavior*, 139:200–215, 2023.

Shahar Dobzinski, Noam Nisan, and Michael Schapira. Truthful randomized mechanisms for combinatorial auctions. In *Proceedings of the thirty-eighth annual ACM symposium on Theory of computing*, pages 644–652, 2006.

Longteng Duan, Zifan Gong, Minming Li, Chenhao Wang, and Xiaoying Wu. Facility location with fractional preferences and minimum distance. In *Proceedings of the 27th International Conference on Computing and Combinatorics (COCOON)*, pages 499–511, 2021.

Bruno Escoffier, Laurent Gourvès, Nguyen Kim Thang, Fanny Pascual, and Olivier Spanjaard. Strategy-proof mechanisms for facility location games with many facilities. In *ADT*, pages 67–81, 2011.

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.

Aris Filos-Ratsikas and Peter Bro Miltersen. Truthful approximations to range voting. In *Proceedings of the 10th International Conference on Web and Internet Economics (WINE)*, pages 175–188, 2014.

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.

Aris Filos-Ratsikas, Søren Kristoffer Stiil Frederiksen, and Jie Zhang. Social welfare in one-sided matchings: Random priority and beyond. In *International Symposium on Algorithmic Game Theory*, pages 1–12. Springer, 2014.

Aris Filos-Ratsikas, Evi Micha, and Alexandros A. Voudouris. The distortion of distributed voting. *Artificial Intelligence*, 286:103343, 2020.

Aris Filos-Ratsikas, Panagiotis Kanellopoulos, Alexandros A. Voudouris, and Rongsen Zhang. The distortion of distributed facility location. *Artif. Intell.*, 328:104066, 2024.

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.

Dimitris Fotakis and Christos Tzamos. On the power of deterministic mechanisms for facility location games. *ACM Trans. Econ. Comput.*, 2(4):15:1–15:37, 2014.

Dimitris Fotakis and Christos Tzamos. Strategyproof facility location for concave cost functions. *Algorithmica*, 76(1):143–167, 2016.

Rachel Freedman, Jana Schaich Borg, Walter Sinnott-Armstrong, John P Dickerson, and Vincent Conitzer. Adapting a kidney exchange algorithm to align with human values. *Artificial Intelligence*, 283:103261, 2020.

Ling Gai, Mengpei Liand, and Chenhao Wang. Obnoxious facility location games with candidate locations. In *Proceedings of the 16th International Conference on Algorithmic Aspects in Information and Management (AAIM)*, pages 96–105, 2022.

Ling Gai, Mengpei Liang, and Chenhao Wang. Two-facility-location games with mixed types of agents. *Applied Mathematics and Computation*, 466:128479, 2024.

Dorit S Hochbaum. Heuristics for the fixed cost median problem. *Mathematical programming*, 22:148–162, 1982.

Xumin Huang, Peichun Li, and Rong Yu. Social welfare maximization in container-based task scheduling for parked vehicle edge computing. *IEEE Communications Letters*, 23(8):1347–1351, 2019.

Ken Ibara and Hiroshi Nagamochi. Characterizing mechanisms in obnoxious facility game. In *COCOA*, pages 301–311, 2012.

Panagiotis Kanellopoulos, Alexandros A. Voudouris, and Rongsen Zhang. Truthful two-facility location with candidate locations. In *Algorithmic Game Theory - 16th International Symposium, SAGT 2023, Egham, UK, September 4-7, 2023, Proceedings*, pages 365–382, 2023a.

Panagiotis Kanellopoulos, Alexandros A Voudouris, and Rongsen Zhang. On discrete truthful heterogeneous two-facility location. *SIAM Journal on Discrete Mathematics*, 37(2):779–799, 2023b.

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.

Mohammad Lotfi and Alexandros A. Voudouris. On truthful constrained heterogeneous facility location with max-variant cost. *CoRR*, abs/2309.05379, 2023.

Pinyan Lu, Yajun Wang, and Yuan Zhou. Tighter bounds for facility games. In *Proceedings of the 5th International Workshop Internet and Network Economics (WINE)*, pages 137–148, 2009.

Pinyan Lu, Xiaorui Sun, Yajun Wang, and Zeyuan Allen Zhu. Asymptotically optimal strategy-proof mechanisms for two-facility games. In *Proceedings of the 11th ACM Conference on Electronic Commerce (EC)*, pages 315–324, 2010.

Duncan C McElfresh, Hoda Bidkhori, and John P Dickerson. Scalable robust kidney exchange. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 33, pages 1077–1084, 2019.

Hervé Moulin. On strategy-proofness and single peakedness. *Public Choice*, 35 (4):437–455, 1980.

Ahuva Mu'Alem and Noam Nisan. Truthful approximation mechanisms for restricted combinatorial auctions. *Games and Economic Behavior*, 64(2):612–631, 2008.

Noam Nisan and Amir Ronen. Algorithmic mechanism design. In *Proceedings of the thirty-first annual ACM symposium on Theory of computing*, pages 129–140, 1999.

Ariel D. Procaccia and Moshe Tennenholtz. Approximate mechanism design without money. *ACM Transactions on Economics and Computation*, 1(4):18:1–18:26, 2013.

Tao Qin, Wei Chen, and Tie-Yan Liu. Sponsored search auctions: Recent advances and future directions. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 5(4):1–34, 2015.

Aida Rahmattalabi, Shahin Jabbari, Himabindu Lakkaraju, Phebe Vayanos, Max Izenberg, Ryan Brown, Eric Rice, and Milind Tambe. Fair influence maximization: A welfare optimization approach. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 35, pages 11630–11638, 2021.

Ben Roberts, Dinan Gunawardena, Ian A Kash, and Peter Key. Ranking and tradeoffs in sponsored search auctions. *ACM Transactions on Economics and Computation (TEAC)*, 4(3):1–21, 2016.

Tim Roughgarden. Algorithmic game theory. *Communications of the ACM*, 53 (7):78–86, 2010.

Tuomas Sandholm. Algorithm for optimal winner determination in combinatorial auctions. *Artificial intelligence*, 135(1-2):1–54, 2002.

James Schummer and Rakesh V Vohra. Mechanism design without money. *Algorithmic game theory*, 10:243–299, 2007.

Paolo Serafino and Carmine Ventre. Heterogeneous facility location without money. *Theoretical Computer Science*, 636:27–46, 2016.

Xin Sui and Craig Boutilier. Approximately strategy-proof mechanisms for (constrained) facility location. In *Proceedings of the 2015 International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, pages 605–613, 2015.

Pingzhong Tang, Dingli Yu, and Shengyu Zhao. Characterization of group-strategyproof mechanisms for facility location in strictly convex space. In *EC*, pages 133–157, 2020a.

Zhongzheng Tang, Chenhao Wang, Mengqi Zhang, and Yingchao Zhao. Mechanism design for facility location games with candidate locations. In *Proceedings of the 14th International Conference on Combinatorial Optimization and Applications (COCOA)*, pages 440–452, 2020b.

William Vickrey. On the prevention of gerrymandering. *Political Science Quarterly*, 76(1):105–110, 1961.

Toby Walsh. Strategy proof mechanisms for facility location at limited locations. In *Proceedings of the 8th Pacific Rim International Conference on Artificial Intelligence (PRICAI)*, pages 113–124, 2021.

David P Williamson and David B Shmoys. *The design of approximation algorithms*. Cambridge university press, 2011.

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.

Deshi Ye, Lili Mei, and Yong Zhang. Strategy-proof mechanism for obnoxious facility location on a line. In *COCOON*, pages 45–56, 2015.

Qi Zhao, Wenjing Liu, Qingqin Nong, and Qizhi Fang. Constrained heterogeneous facility location games with max-variant cost. *Journal on Combinatorial Optimization*, 45(3):90, 2023.

Yunhong Zhou, Deeparnab Chakrabarty, and Rajan Lukose. Budget constrained bidding in keyword auctions and online knapsack problems. In *Proceedings of the 17th international conference on world wide web*, pages 1243–1244, 2008.

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.