

A Study of Strategic Facility Location : Off-line Mechanisms for Collinear Instances in $\mathbb{R}^2$

Yuki Crivelli[†]

Bachelor Research Project, EPFL
Supervisor: Ola Nils Anders Svensson

January 5, 2026

Abstract

In this work, we present the current bounds and best-known mechanisms (as of end of 2025) for the *Strategic Facility Location Problem*. This problem involves a set of agents who each report their location, while a mechanism chooses a facility position in a metric space to minimize a social cost function depending on distances between agents and the facility. The strategic dimension arises because agents may misreport their locations in order to bring the facility closer to themselves. Hence, the goal is to design **strategyproof** mechanisms that remove any incentive to lie. We review existing results for various objectives and settings, and we focus on the special case of collinear agents in the two-dimensional Euclidean plane under the maximum cost objective, where we present a new truthful mechanism that achieves a $\sqrt{2}$ -approximation — improving upon all previously known mechanisms for this setting. Then, we try an extension of this mechanism to bigger dimensions and more importantly, we design UPA, a mechanism for general $\mathbb{R}^2$ instances, that performs nicely by outputting out of the line for collinear profiles. We also discuss randomization and prediction techniques to nuance the results and look for new directions to explore.

[†]yuki.crivelli@epfl.ch

Contents

1	Introduction & Preliminaries	3
1.1	Motivation and Problem Setting	3
1.2	Description, Preliminaries and Notation	4
2	Background and Related Work	4
2.1	Deterministic Mechanisms	4
2.1.1	1D instances ($\mathbb{R}$)	4
2.1.2	2D instances ($\mathbb{R}^2$)	6
2.2	Randomized Mechanisms	7
2.2.1	1D instances ($\mathbb{R}$)	7
2.2.2	2D instances ($\mathbb{R}^2$)	9
2.3	Summary of Known Information	11
3	New Setting : Collinear Instances in the Plane	12
3.1	Apex Mechanism	12
3.2	Strategyproofness	13
3.3	Approximation Ratio	15
3.4	Optimality	16
3.5	Randomization cannot do better !	18
3.6	Discussion about Scale-invariance and Randomization	19
4	Generalization to d Dimensions	19
4.1	Negative Result of Intuitive Generalization	19
4.2	Trying to go beyond the naive generalization	20
4.2.1	Rethinking mechanisms in $\mathbb{R}^2 \rightarrow \mathbb{R}^3$	21
4.2.2	Using higher-dimensional outcome spaces	21
5	Use of Apex for General 2D Mechanisms	21
5.1	Impossibility Result : Deterministic	22
5.2	Randomized	23
5.2.1	The Two-agent Case	23
5.2.2	Three Agents or more	24
5.3	Uniform Pairwise Apex (UPA)	25
5.3.1	Bounds	25
5.3.2	A Surprising Behaviour	29
5.4	Discussion about the conjecture in Balkanski et al. (2024)	29
5.5	A Word about Predictions	30
6	Discussion	30
6.1	Recap of our Contribution	30
6.2	Conceptual Takeaways	31
6.3	Limitations and Critical Assessment	31
7	Future Work	31
8	Conclusion	32
A	Deterministic Lower Bound Without Scale-Invariance	33

1 Introduction & Preliminaries

1.1 Motivation and Problem Setting

Motivation for Strategic Facility Location. The Facility Location problem is an old and well-studied model for decision-making. It asks a simple question: given a set of positions for agents in a metric space, where should we place a facility (a hospital, server, or public service center) so that it serves everyone as well as possible? In its algorithmic form, the problem asks to minimize a *social cost function* — for example, the maximum distance one has to travel in order to access the facility.

However, in real-world scenarios, agents are often **strategic**, i.e., they can misreport their position to bring the facility closer to them. We can think of a governmental service that decides to place a hospital to serve a set of cities and asks each mayor to report the location of their town hall. Each mayor could exaggerate its position to theoretically attract the hospital toward their city.

We therefore want to find **strategyproof (SP)** mechanisms (also called **truthful** or **incentive compatible**) that never reward lying. This means that every agent’s best choice (the one that minimizes their distance to the facility) is to tell the truth. Studying Strategic Facility Location sits at the intersection of several domains such as optimization and economics, and captures how collective decisions can be made both *fairly* and *efficiently*.

Motivation for Learning-Augmented Mechanisms. Historically, mechanism design aimed to achieve the best possible **approximation guarantee**, i.e., the worst-case ratio between a mechanism’s social cost and the optimal result under complete information. With the rise of machine learning, and in order to bring theoretical results closer to real-world applications, we now study **learning-augmented mechanisms** that have access to a prediction of some information. Our goal is to find a trade-off between following the prediction when it is accurate (**consistency**) and maintaining strong approximation guarantees when the prediction is poor (**robustness**).

For example, in a smart-city setting, the government might already have predictions about future population concentrations. A learning-augmented mechanism would place the hospital according to this prediction while keeping a reasonable guarantee even if the prediction is inaccurate. This model bridges the gap between *theoretical guarantees* and *practical, data-informed decision-making*, and has become the focus of much recent work in Strategic Facility Location.

Motivation for Off-Line Mechanisms. Most of the literature on SP facility location focuses on one-dimensional or two-dimensional instances — that is, all agents and the facility lie on a line or in a plane. We now want to study different metric spaces for the agents and the facility. For example, our main result uses a two-dimensional setting (the facility can be placed anywhere in $\mathbb{R}^2$) but the agents lie on a line (W.L.O.G., $y_i = 0 \forall i \in \{1, \dots, n\}$, i.e., all agents are on the x-axis).

The question is whether mechanisms are always better off returning on-line locations, and this is an open question that we address. Intuitively, one could think that any mechanism’s projection onto the line would perform better than the mechanism itself, but this projection may lose its strategyproofness. More generally, we are interested in studying strategic facility location instances where the metric space of the agents is not the same as the metric space of the facility. As an illustrative example, consider a line of soldiers positioned along a front who occasionally need to restock ammunition. Each soldier radios their position, and a truck carrying ammunition must be sent to a location where all can resupply. For safety, the truck is likely parked *off the shooting line*, yet it should be placed so as to minimize the maximum walking distance any soldier faces. This scenario intuitively shows that an off-line placement can be preferable, both practically (safer) and geometrically (more balanced).

1.2 Description, Preliminaries and Notation

Formally, the *Strategic Facility Location Problem* can be stated as follows. We work with a set of n agents in some metric space (often $\mathbb{R}$ or $\mathbb{R}^2$) that are strategic, i.e., they can misreport their location to bring the facility closer to them. We define by $p_i \in \mathbb{R}^d$ the reported location of agent i , and with $P \in \mathbb{R}^{dn}$ the vector of all positions. We will sometimes denote the truthful position of an agent with p_i and her reported position with $\tilde{p}_i$ (upon specification). When we have a prediction of the optimal location for the facility, it will be denoted by $\hat{o} \in \mathbb{R}^d$. Note that this prediction depends on P and that if it is completely accurate, then $\hat{o} = o$, where o is the real optimal location for the facility.

We want to design a mechanism that outputs $f(P, \hat{o})$ the location for the facility ($f(P)$ if no prediction, and sometimes written simply f). If not mentioned otherwise, every mechanism that we discuss is **strategyproof**, **anonymous** and **unanimous**. Anonymity is the condition that the output of the mechanism does not depend on the agents identity, i.e., swapping two agents will not change the location of the facility. Unanimity means that if all agents are located at the same point ($p_i = p_j \forall i, j \in \{1, \dots, n\}$), then the location is also placed at this same point.

The objective is to minimize some social cost function $\mathbf{C}$, that depends on the location of the agents and the facility : $C(f(P, \hat{o}), P)$, or simply $C(f, P)$. In strategic facility location, the two common studied social cost functions are the Egalitarian function, i.e., the maximum distance any agent has to walk : $C_{Egalitarian}(f, P) = \max_{p_i \in P} d(p_i, f)$ and the Utilitarian function, i.e., the average distance walked by the agents : $C_{Utilitarian}(f, P) = \frac{1}{n} \sum_{i=1}^n d(p_i, f)$. The distance metric is usually the Euclidean distance, and this is the one we will use. $d(x, y) = \sqrt{\sum_{i=1}^d (x_i - y_i)^2}$, where d is the dimension and x_i and y_i the respective value of the agent in the i 's dimension.

2 Background and Related Work

In this first section, we will discuss two of the most important papers concerning Strategic Facility Location, namely **Learning-Augmented Mechanism Design: Leveraging Predictions for Facility Location** by Agrawal et al. [2022] and **Randomized Strategic Facility Location with Predictions** by Balkanski et al. [2024].

Note that from now on, if not said otherwise, we will use the **Egalitarian** social cost function as an objective to minimize, i.e., we want to minimize the largest distance an agent has to walk to reach the facility. This is a choice and the study of the Utilitarian social cost function is also very interesting, however this allows us to focus more on off-line mechanisms and randomization, reducing the number of possible settings.

2.1 Deterministic Mechanisms

We will start by discussing the deterministic mechanisms, mostly presented in Agrawal et al. [2022] and previous work. We also present the known bounds for both the one-dimensional and two-dimensional instances, without and with predictions.

2.1.1 1D instances ($\mathbb{R}$)

We consider the one-dimensional setting where each agent reports a location $\vec{x} = (x_1, \dots, x_n) \in \mathbb{R}^n$. Following the convention of Procaccia and Tennenholtz [2013], we let

$$x_1 = \min_i x_i, \quad x_n = \max_i x_i,$$

i.e., x_1 is the leftmost agent and x_n the rightmost one. This simplification is without loss of generality, since any deterministic mechanism we study is anonymous, and therefore invariant

under permutations of the agents.

Mechanism 2.1 (Median Mechanism). Given reported locations $x_1, \dots, x_n \in \mathbb{R}$, the mechanism outputs

$$f(x) = \text{median}(x_1, \dots, x_n).$$

The median mechanism is well known to be strategyproof. Moreover, it achieves a 2-approximation for the maximum-cost objective. Indeed, the output always lies inside the segment $[x_1, x_n]$, so for any agent i ,

$$|x_i - f(x)| \leq x_n - x_1.$$

Let $D = x_n - x_1$ denote the diameter of the instance. The optimal facility location is $o(x) = \frac{x_1 + x_n}{2}$, yielding an optimal maximum distance of $D/2$. Thus the median achieves an approximation ratio

$$\frac{D}{D/2} = 2.$$

Theorem 2.2 (Procaccia and Tennenholtz [2013]). *For the egalitarian objective on $\mathbb{R}$, no deterministic strategyproof mechanism can achieve an approximation ratio strictly better than 2.*

Hence, the median mechanism is *optimal* among deterministic strategyproof mechanisms in one dimension.

Learning-augmented setting. We now assume the availability of a prediction $\hat{o}$ of the optimal facility location. The goal is to follow the prediction when it is accurate while maintaining good performance when the prediction is adversarial. Agrawal et al. [2022] introduce a deterministic mechanism achieving the optimal consistency–robustness tradeoff.

Mechanism 2.3 (MinMaxP [Agrawal et al., 2022]). Given reports $x_1, \dots, x_n$ and a prediction $\hat{o} \in \mathbb{R}$, the mechanism outputs

$$f(x, \hat{o}) = \begin{cases} \hat{o}, & \text{if } \hat{o} \in [x_1, x_n], \\ x_1, & \text{if } \hat{o} < x_1, \\ x_n, & \text{if } \hat{o} > x_n. \end{cases}$$

In words, MinMaxP outputs the prediction if it lies between the leftmost and rightmost agents, and otherwise outputs the closer endpoint.

Theorem 2.4 (Agrawal et al. [2022]). *For the one-dimensional maximum-cost objective, MinMaxP is anonymous, deterministic, strategyproof, 1-consistent, and 2-robust.*

When the prediction is accurate, we have $\hat{o} \in [x_1, x_n]$, and $f(x, \hat{o}) = \hat{o}$, giving consistency 1. When the prediction is arbitrarily wrong, the facility always lies in $[x_1, x_n]$, so the worst-case maximum distance is at most $x_n - x_1 = 2 \cdot \text{OPT}$, giving robustness 2, as can be observed in Figure 1.

Proposition 2.5 (Chan et al. [2025]). *Among all deterministic strategyproof mechanisms for the one-dimensional maximum-cost objective, MinMaxP achieves the optimal consistency–robustness tradeoff: no such mechanism can be both $(1 - \varepsilon)$ -consistent and $(2 - \varepsilon)$ -robust for any $\varepsilon > 0$.*

The proof in Chan et al. [2025] relies on viewing approximation as a function of the *prediction error* $\eta = |\hat{o} - o(x)|$. They show that MINMAXP achieves

$$\alpha(\eta) = 1 + \min(1, \eta),$$

and that this curve is unimprovable by any deterministic SP mechanism. Taking $\eta = 0$ yields consistency $C = \alpha(0) = 1$, while letting $\eta \rightarrow \infty$ yields robustness $R = \lim_{\eta \rightarrow \infty} \alpha(\eta) = 2$.

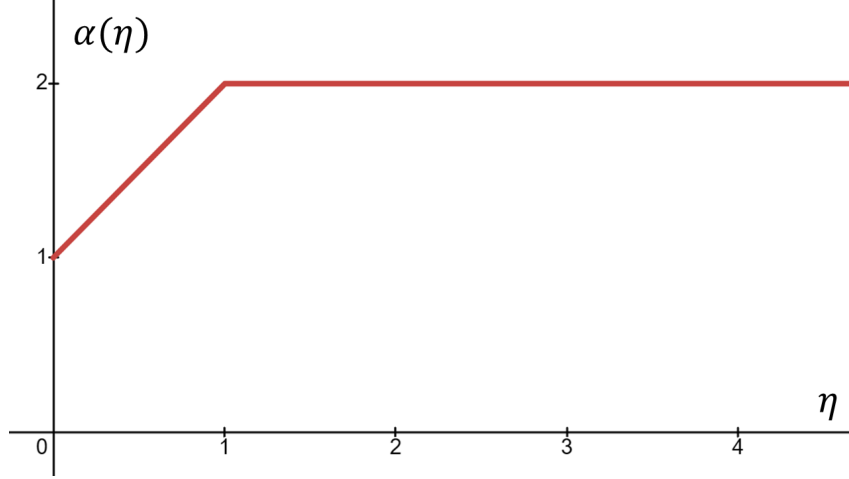


Figure 1: Approximation ratio of MINMAXP as a function of the error factor

2.1.2 2D instances ($\mathbb{R}^2$)

The two-dimensional setting is considerably more constrained than the one-dimensional case. A classical result of Peters et al. [1993] shows that any continuous, anonymous, strategyproof mechanism over $(\mathbb{R}^2, \|\cdot\|_2)$ must act *coordinatewise*: such a mechanism is exactly a generalized coordinate-wise median with respect to some orthogonal basis. This parallels Moulin’s characterization on the line, but the two-dimensional setting admits no further freedom.

Mechanism 2.6 (Coordinate-wise Median (CM)). Given reports $p_i = (x_i, y_i) \in \mathbb{R}^2$, define

$$f(P) = (\text{median}(x_1, \dots, x_n), \text{median}(y_1, \dots, y_n)).$$

CM is anonymous, unanimous, and strategyproof [Peters et al., 1993]. For the egalitarian objective $C(f, P) = \max_i \|p_i - f\|_2$, the output of CM always lies inside the minimal bounding rectangle $[x_1, x_n] \times [y_1, y_n]$, where $x_1 \leq \dots \leq x_n$ and $y_1 \leq \dots \leq y_n$ are the sorted coordinates. Since the optimal facility $o(P)$ lies in the *Euclidean* diameter- D ball centered at the midpoint of the diameter of the instance, one obtains the well-known bound

$$C(f(P), P) \leq 2 \cdot C(o(P), P).$$

A matching lower bound of 2 for all deterministic SP mechanisms is proved in Procaccia and Tennenholtz [2013]. Hence CM is *optimal* among deterministic SP mechanisms in $\mathbb{R}^2$ under the maximum-cost objective.

With predictions. We now consider the learning-augmented setting of Chan et al. [2025], who study deterministic, anonymous, SP mechanisms on $(\mathbb{R}^2, \|\cdot\|_p)$ for $p \geq 2$, equipped with a prediction $\hat{o}$ of the optimal facility.

The natural deterministic extension of MinMaxP to $\mathbb{R}^2$ is the *Minimum Bounding Box* mechanism, which applies MinMaxP coordinate-by-coordinate.

Mechanism 2.7 (Minimum Bounding Box (MBB) [Chan et al., 2025]). Let $x_1 \leq \dots \leq x_n$ and $y_1 \leq \dots \leq y_n$ be the sorted coordinates. Given prediction $\hat{o} = (\hat{o}_x, \hat{o}_y) \in \mathbb{R}^2$, output

$$f(P, \hat{o}) = (\max\{x_1, \min\{x_n, \hat{o}_x\}\}, \max\{y_1, \min\{y_n, \hat{o}_y\}\}).$$

MBB is deterministic, anonymous, unanimous, and strategyproof. Its performance guarantees in ℓ_p^2 are:

$$\text{consistency} = 1, \quad \text{robustness} = 1 + 2^{1/p}.$$

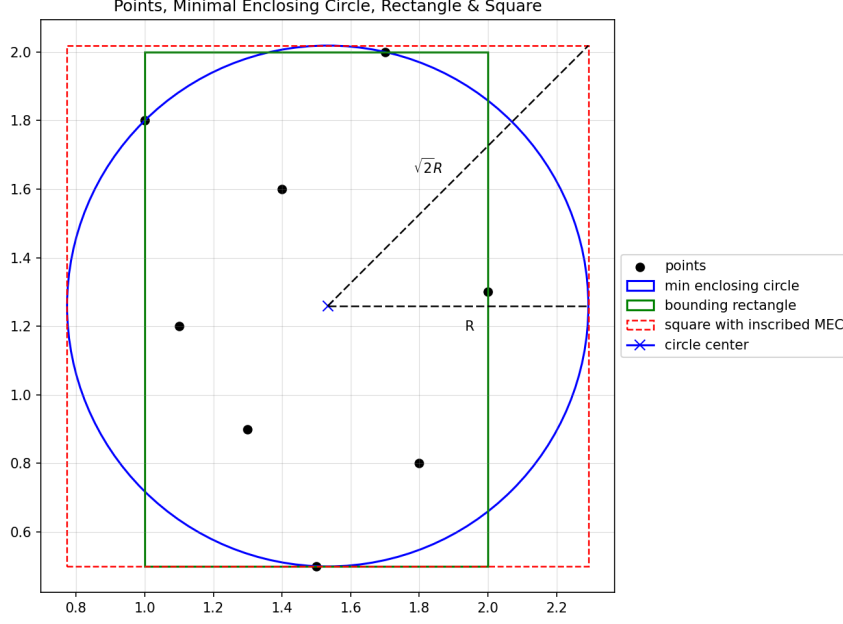


Figure 2: Visual proof for $1 + \sqrt{2}$ -robustness

In particular, for the Euclidean norm ($p = 2$), robustness is

$$R = 1 + \sqrt{2},$$

as can be seen in Figure 2.

Remarkably, Chan et al. [2025] prove a matching lower bound showing that MBB is *optimal* for the consistency-robustness tradeoff.

Theorem 2.8 (Lower bound on deterministic SP mechanisms in $\mathbb{R}^2$ [Chan et al., 2025]). *For the egalitarian objective in $(\mathbb{R}^2, \|\cdot\|_p)$ with $p \geq 2$, any deterministic strategyproof mechanism with consistency $C < 2$ must satisfy*

$$R \geq 1 + 2^{1/p}.$$

In particular, in the Euclidean plane ($p = 2$),

$$C < 2 \implies R \geq 1 + \sqrt{2}.$$

This theorem immediately implies that MBB achieves the *best possible* robustness $1 + \sqrt{2}$ among all deterministic strategyproof mechanisms with consistency 1. Thus MBB is optimal in the learning-augmented deterministic setting in $\mathbb{R}^2$.

Summary. For deterministic mechanisms in $\mathbb{R}^2$:

- Without predictions, CM achieves a 2-approximation and this is optimal.
- With predictions, MBB is 1-consistent and $(1 + \sqrt{2})$ -robust, and Theorem 2.8 shows this robustness is optimal for any mechanism with consistency < 2 .

2.2 Randomized Mechanisms

2.2.1 1D instances ($\mathbb{R}$)

We now turn to the randomized setting for one-dimensional instances, i.e., $p_1, \dots, p_n \in \mathbb{R}$ and distances $d(p, f) = |p - f|$. As before, our goal is to minimize the *egalitarian* objective

$$C(f, P) = \max_i |p_i - f|.$$

Randomization dramatically improves the achievable worst-case guarantee: while any deterministic strategyproof (SP) mechanism must suffer a factor-2 approximation, a randomized SP mechanism can guarantee a $3/2$ -approximation, and this is known to be optimal.

Without Predictions : The fundamental randomized mechanism for the maximum-cost objective is the **Left–Right–Middle (LRM)** mechanism, introduced by Procaccia and Tennenholtz [2013]. Let $p_1 = \min_i p_i$ and $p_n = \max_i p_i$ denote the leftmost and rightmost agent.

Mechanism 2.9 (LRM [Procaccia and Tennenholtz, 2013]). Given reports $p_1, \dots, p_n$, define

$$f(P) = \begin{cases} p_1 & \text{with probability } \frac{1}{4}, \\ p_n & \text{with probability } \frac{1}{4}, \\ \frac{p_1 + p_n}{2} & \text{with probability } \frac{1}{2}. \end{cases}$$

In words, LRM chooses an endpoint with probability $1/4$ each to maintain strategyproofness, and outputs the midpoint with probability $1/2$ to reduce the expected maximum distance.

Theorem 2.10 (Procaccia and Tennenholtz [2013]; see also Balkanski et al. [2024]). *LRM is anonymous and strategyproof and satisfies*

$$\mathbb{E}[C(f(P), P)] \leq \frac{3}{2} C(o(P), P).$$

Moreover, no randomized strategyproof mechanism can achieve an approximation ratio strictly smaller than $3/2$.

Thus LRM is *optimal* among all randomized SP mechanisms for 1D instances:

$$\boxed{\alpha_{\text{rand}}^*(1\text{D}) = \frac{3}{2}.$$

With Predictions : We now consider the learning-augmented setting, where a prediction $\hat{o} \in \mathbb{R}$ of the optimal facility location is available. Our objective is to design a mechanism that is simultaneously *consistent* (i.e., performs optimally when $\hat{o} = o(P)$) and *robust* (i.e., still performs well when $\hat{o}$ is adversarial).

In the deterministic setting, Agrawal et al. [2022] show that the **MinMaxP** mechanism is 1-consistent and 2-robust, and that this tradeoff is optimal among deterministic SP mechanisms (see also Chan et al. [2025] for a tight characterization). Randomization allows improved robustness ($3/2$ instead of 2), at the price of reduced consistency.

A natural learning-augmented randomized mechanism is obtained by mixing LRM with MinMaxP according to a *trust parameter* $\delta \in [0, 1]$.

Mechanism 2.11 (δ -Mixture of MinMaxP and LRM). Given reports P and a prediction $\hat{o}$, output

$$f_\delta(P, \hat{o}) = \begin{cases} \text{MinMaxP}(P, \hat{o}) & \text{with probability } \delta, \\ \text{LRM}(P) & \text{with probability } 1 - \delta. \end{cases}$$

Since both MinMaxP and LRM are SP, the mixture f_δ is SP for all $\delta \in [0, 1]$.

Consistency. When $\hat{o} = o(P)$, MinMaxP is perfectly accurate (1-consistent), whereas LRM has approximation ratio $3/2$. Thus the consistency of f_δ equals

$$C(\delta) = \delta \cdot 1 + (1 - \delta) \cdot \frac{3}{2} = \frac{3}{2} - \frac{\delta}{2}.$$

Robustness. When the prediction is adversarial, MinMaxP is 2-robust and LRM is $3/2$ -robust, so the robustness of f_δ is

$$R(\delta) = \delta \cdot 2 + (1 - \delta) \cdot \frac{3}{2} = \frac{3}{2} + \frac{\delta}{2}.$$

Thus f_δ provides a continuous tradeoff between consistency and robustness:

$$(C(\delta), R(\delta)) = \left(\frac{3}{2} - \frac{\delta}{2}, \frac{3}{2} + \frac{\delta}{2} \right), \quad \delta \in [0, 1].$$

At the extremes:

$$f_0 = \text{LRM} : (C, R) = \left(\frac{3}{2}, \frac{3}{2} \right), \quad f_1 = \text{MinMaxP} : (C, R) = (1, 2).$$

Proposition 2.12. *The mixture mechanism f_δ is anonymous, strategyproof, and achieves consistency $C(\delta) = \frac{3}{2} - \frac{\delta}{2}$ and robustness $R(\delta) = \frac{3}{2} + \frac{\delta}{2}$.*

Optimality Discussion. The robustness lower bound of $3/2$ for randomized SP mechanisms [Balkanski et al., 2024] implies that no learning-augmented randomized mechanism can achieve $R < 3/2$, even with perfect confidence in the prediction. Similarly, no mechanism can be more than 1-consistent. The mixture mechanism interpolates between these two tight endpoints:

$$(C, R) \in [1, \frac{3}{2}] \times [\frac{3}{2}, 2],$$

and every known randomized SP mechanism lies on or above this curve. This suggests—though a complete impossibility theorem is not yet known—that f_δ achieves the optimal consistency–robustness tradeoff in the randomized 1D setting.

2.2.2 2D instances ($\mathbb{R}^2$)

Randomization in two-dimensional Euclidean space is far less understood than in one dimension. For deterministic mechanisms, the best-possible approximation is 2 (achieved by the coordinate-wise median), whereas for randomized mechanisms, the landscape is significantly more complex. Prior to recent work, essentially no truthful randomized mechanism was known to beat the deterministic factor 2.

Recently, Tang et al. [2020] introduced the *Centroid mechanism*, and more recently, Balkanski et al. [2024] developed new randomized learning-augmented mechanisms that leverage predictions about the *extreme agents*, i.e., those agents that determine the minimum enclosing circle. The present subsection summarizes the state of the art.

Classical Prediction-Free Setting. The best-known truthful randomized mechanism without predictions is due to Tang et al. [2020], and was re-analyzed in Balkanski et al. [2024] (see [Tang et al., 2020, Balkanski et al., 2024]).

Mechanism 2.13 (Centroid Mechanism [Tang et al., 2020]). Given reported locations $p_1, \dots, p_n \in \mathbb{R}^2$:

- Let $G = \frac{1}{n} \sum_{i=1}^n p_i$ be the centroid of all agents.
- With probability $\frac{1}{2}$, output G .
- With probability $\frac{1}{2n}$, output each agent location p_i .

This mechanism is truthful in expectation and unanimous (it returns the common point when all agents coincide). Since the centroid always lies inside the minimum enclosing circle, its performance is governed by how far the centroid may be displaced from the optimal center.

Theorem 2.14 (Tang et al. 2020). *The Centroid mechanism is strategyproof in expectation and achieves a*

$$(2 - 1/n)\text{-approximation}.$$

This remains the best-known prediction-free randomized guarantee in $\mathbb{R}^2$, and no matching lower bound is known. In fact, Balkanski et al. [2024] highlight a substantial gap:

$$1.118 \leq \alpha_{\text{rand}}^*(\mathbb{R}^2) \leq 2 - \frac{1}{n}$$

where 1.118 is the first nontrivial lower bound for truthful randomized mechanisms in Euclidean 2D without predictions. This lower bound is proved in [Balkanski et al., 2024].

Theorem 2.15 (Lower bound for randomized truthful mechanisms [Balkanski et al., 2024]). *No truthful randomized mechanism in $\mathbb{R}^2$ can achieve an approximation ratio strictly better than*

$$1.118.$$

Thus the best-known upper and lower bounds leave a wide gap.

Learning-Augmented Setting: Predictions of Extreme Agents. Balkanski et al. [2024] consider a prediction about the *identities* of the extreme agents, i.e., those agents that lie on the circumference of the minimum enclosing circle:

$$\hat{e} = (e_1, \dots, e_k), \quad \text{where } \{e_1, \dots, e_k\} = \arg \max_{i \in [n]} d(p_i, o(P)).$$

Since $o(P)$ is the center of the minimum enclosing circle, a correct prediction of $\hat{e}$ identifies the agents whose locations determine the optimal facility.

Let k denote the number of extreme agents (always 2 or 3 except for degenerate cases). Balkanski et al. [2024] define the following mechanism.

Mechanism 2.16 (Centroid-on-Extreme-Agents [Balkanski et al., 2024]). Given a prediction $\hat{e} = (e_1, \dots, e_k)$:

- Compute the centroid

$$G = \frac{1}{k} \sum_{j=1}^k p_{e_j}.$$

- With probability $\frac{1}{2}$, output G .
- With probability $\frac{1}{2k}$, output each p_{e_j} .

Two main performance results follow, depending on the number of extreme agents.

Theorem 2.17 (Two extreme agents [Balkanski et al., 2024], Theorem 6). *If $k = 2$ and the prediction of extreme agents is correct, then Mechanism 2.16 is truthful in expectation, achieves 1.5 consistency, and 2 robustness.*

Theorem 2.18 (Three extreme agents [Balkanski et al., 2024], Theorem 7). *If $k = 3$ and the prediction of extreme agents is correct, then Mechanism 2.16 is truthful in expectation, achieves*

$$1.67\text{-consistency}$$

and 2 robustness. The 1.67 bound arises from geometric properties of the Euler line of a triangle.

Since $k = 3$ is the generic case (the probability of $k \geq 4$ is 0 under perturbations [de Berg et al., 2008]), this yields the following overall theorem.

Theorem 2.19 (Learning-augmented randomized mechanism in 2D [Balkanski et al., 2024]). *With predictions of the identities of extreme agents, there exists a truthful randomized mechanism with*

$$1.67 \text{ consistency}, \quad 2 \text{ robustness}.$$

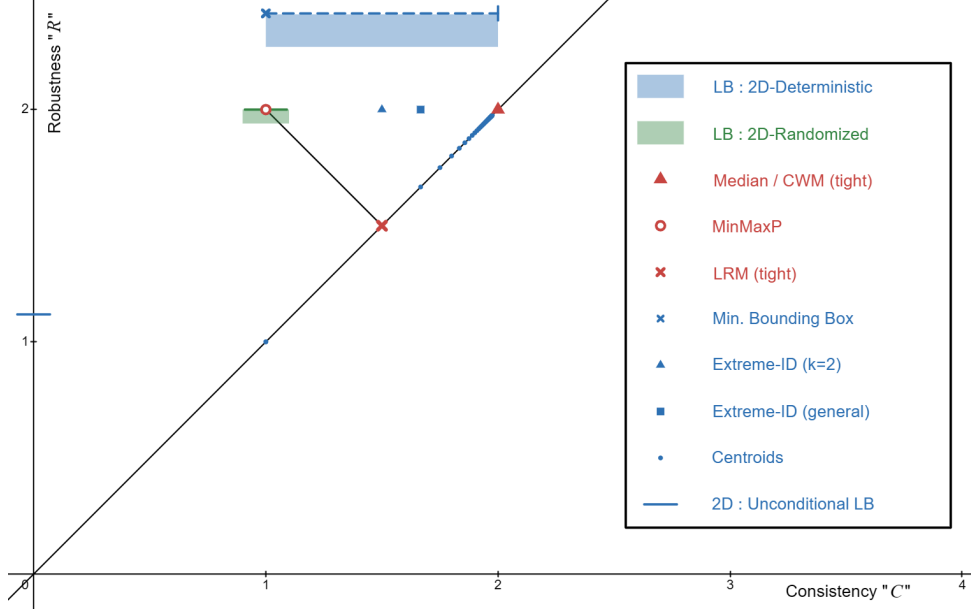


Figure 3: Current Knowledge lower bounds and mechanisms

Lower Bounds in the Learning-Augmented Setting. Balkanski et al. [2024] also show that the lower bound of Theorem 2.15 remains valid with predictions, unless the prediction is extremely strong. Moreover, they prove that no deterministic mechanism can simultaneously guarantee both

$$\text{consistency} < 2 \quad \text{and} \quad \text{robustness} < 1 + \sqrt{2},$$

even with the strongest prediction (the exact optimal facility location). Likewise, no truthful randomized mechanism can be 1-consistent and strictly better than 2-robust.

Discussion. The current best-known picture for randomized facility location in $\mathbb{R}^2$ can be summarized as follows:

$$1.118 \leq \alpha_{\text{rand}}^*(\mathbb{R}^2) \leq 2 - \frac{1}{n},$$

with the lower bound due to Balkanski et al. [2024], and the upper bound achieved by the Centroid mechanism of Tang et al. [2020]. With predictions of extreme agents, one can improve consistency to 1.67, while maintaining 2 robustness.

Finding the exact optimal randomized SP approximation ratio in $\mathbb{R}^2$ remains a major open problem.

2.3 Summary of Known Information

Figure 3 summarizes what is currently known in the consistency–robustness landscape for strategyproof facility location, in both 1D and 2D, with and without predictions.

In one dimension, the picture is essentially closed. Deterministically, the median is tight at $(C, R) = (2, 2)$. Randomization improves robustness to $3/2$, and LRM exactly hits this point; no mechanism can go below it. With predictions, MinMaxP reaches $(C, R) = (1, 2)$, guaranteeing the robustness of the median but improving the result when an accurate prediction is given. All randomized learning-augmented mechanisms fall on or above the straight line connecting MinMaxP and LRM, and no point below this line is achievable.

In two dimensions, things are far less resolved. Deterministically, the coordinate-wise median sits at $(C, R) = (2, 2)$ and is tight. With predictions, the Minimum Bounding Box mechanism

achieves $(1, 1 + \sqrt{2})$, and the blue shaded region in the plot shows that no deterministic SP mechanism can do better: consistency < 2 forces robustness at least $1 + \sqrt{2}$.

Randomization in 2D is still wide open. The best unconditional lower bound is 1.118 (horizontal blue line), and the best known mechanisms—centroid-based ones—sit well above it, near robustness 2. Learning-augmented EXTREME-ID mechanisms improve consistency (around 1.5–1.67) but still hit robustness 2. The green bar shows the only general randomized lower bound we have: nothing can drop below 1.118 unless the prediction is extremely strong.

Overall, Figure 3 illustrates the contrast between 1D and 2D. In 1D, the frontier is fully understood: all mechanisms lie exactly on the tight Pareto curve. In 2D, we merely know a few isolated points (CM, MBB, centroid, Extreme-ID), and large regions of the plane remain unexplored.

3 New Setting : Collinear Instances in the Plane

Our contribution starts by considering so called **collinear instances** in $\mathbb{R}^2$. This setting assumes a 2D space in which the facility can be placed, but we know that the agents are collinear. Without loss of generality and as a simplification, we will only consider instances in which all the agents y position is zero, i.e., $y_i = 0 \forall i$. However, since we are working in the two-dimensional space, the facility can be located out of the line. Our goal is to understand how the right to add a dimension to the solution can modify the known mechanisms for the line and change the distances.

The motivation behind this specific case is that it appears in some optimization problems including a waiting line and a facility that can be placed anywhere in the plane. We could think of a line of workers (for example railworkers, collinear) and we want to place a food truck somewhere in the plane (not necessarily on the railway) in order to minimize the maximum distance a worker has to walk in order to get his meal.

We will later see and discuss how we can use our work on this specific instance to tackle the general strategic facility location problem.

3.1 Apex Mechanism

The mechanism we present is called the APEX MECHANISM, because of its shape. The logic behind it is to "trust" the agents as to tell the truth, but punish them if they lie. For the following proofs, we will denote by x_i the x-axis position of an agent and by $\tilde{x}_i$ the **reported** location of an agent. We start by collecting all reported locations $\langle \tilde{x}_1, \tilde{x}_2, \dots, \tilde{x}_n \rangle$ and we define $x_1 = \min_i \{\tilde{x}_i\}$ as well as $x_n = \max_i \{\tilde{x}_i\}$, i.e., the leftmost and rightmost agents' x position. Once computed, we place our facility at the location :

$$f = \left(\frac{x_1 + x_n}{2}, \frac{x_n - x_1}{2} \right),$$

as can be seen in Figure 4.

This mechanism does, as wanted, satisfy the properties :

- **Anonymous** : The output does not depend on the agents identity but only on their position. Permuting the order in which the positions are given will not change the output. It is important to make the difference between anonymity and equal-influence (our mechanism does not give the same weight to all positions, but it is anonymous).
- **Unanimous** : If all agents are on the same point $p = (x, 0)$, then the output will be on $f = (\frac{x+x}{2}, \frac{x-x}{2}) = (x, 0)$, i.e., at the agents location.
- **Strategyproof** : See next section...

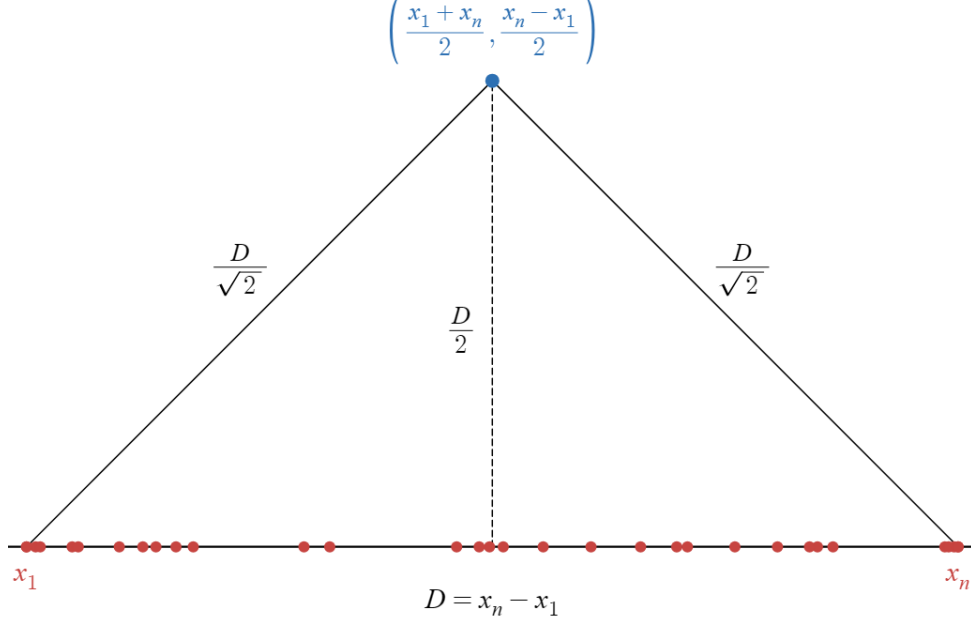


Figure 4: Visualization of APEX MECHANISM with collinear instances in $\mathbb{R}^2$

3.2 Strategyproofness

Lemma 3.1. *The Apex Mechanism is strategyproof.*

In this section, we will use the notation x'_1 and x'_n for the new leftmost resp. rightmost points after the lie. We call "Extreme agents" one of the two agents whose x location was rightmost or leftmost and we call mid agent any other agent, i.e., agent i such that $\exists j \exists k : i \neq j$ and $i \neq k, x_j \leq x_i \leq x_k$.

We want to prove the strategyproofness of our mechanism by cases, showing that no agent can approach the facility to them by lying about their position.

Note that in order to compare distances, we might compare the square of the distances instead, as the result is the same.

Extreme agents We will only consider the rightmost agent and see what happens if they misreport their position. This simplification is without loss of generality because of the symmetry of the problem. Indeed, if the leftmost agent reports $\tilde{x}_1 = x_1 \pm \varepsilon$, it symmetrically has the same effect on the facility location than the rightmost agent reporting $\tilde{x}_n = x_n \mp \varepsilon$.

We can already compute the square of the truthful distance of the rightmost agent n to the facility :

$$d^2(p_n, f) = \left(x_n - \frac{x_n + x_1}{2}\right)^2 + \left(\frac{x_n - x_1}{2}\right)^2 = \frac{x_n^2 - 2x_n x_1 + x_1^2}{2}$$

Stays rightmost First, we consider the case in which the rightmost agent misreports, but they remain the rightmost reported location.

$$\begin{cases} x'_n = x_n + \varepsilon, & \varepsilon \in \mathbb{R} \\ x_i \leq x'_n, & \forall i \neq n \end{cases}$$

The new location will be $f' = (\frac{x'_n + x_1}{2}, \frac{x'_n - x_1}{2}) = (\frac{x_n + x_1 + \varepsilon}{2}, \frac{x_n - x_1 + \varepsilon}{2})$. We can therefore compute the new squared distance and compare it to the truthful one :

$$d^2(p_n, f') = \left(\frac{x_n - x_1 - \varepsilon}{2}\right)^2 + \left(\frac{x_n - x_1 + \varepsilon}{2}\right)^2$$

$$\begin{aligned}
&= \frac{x_n^2 + x_1^2 + \varepsilon^2 - 2x_1x_n - 2x_n\varepsilon + 2x_1\varepsilon + x_n^2 + x_1^2 + \varepsilon^2 - 2x_1x_n + 2x_n\varepsilon - 2x_1\varepsilon}{4} \\
&= \frac{x_n^2 + x_1^2 + \varepsilon^2 - 2x_1x_n}{2}
\end{aligned}$$

We can now compare the squared distances between the truthful and the misreported scenarios:

$$d^2(p_n, f') - d^2(p_n, f) = \frac{\varepsilon^2}{2} > 0$$

Since the new distance is always bigger than the truthful one, we see that an extreme agent does not gain by lying and staying extreme on her side.

Becomes some mid agent $x'_n = x_p$ where $p \neq n$.

We can think of this as if the rightmost agent stayed rightmost (case 1.a), but it had reported $\tilde{x}_n = x_p$, i.e., exactly at the same point as the second rightmost agent, which is equivalent to staying rightmost in the limit. We therefore already saw that this cannot improve her distance to the facility location.

Becomes leftmost agent $x'_n = x_p$ and $x'_1 = x_1 - \varepsilon$, $\varepsilon > 0 \in \mathbb{R}$, where $x_p = x_n - \delta$ is the location of the second rightmost agent. ($\delta > 0$)

As in point 1.a, we can compute the position f' of the new facility location and compare the distance $d^2(p_n, f')$ to the truthful distance $d^2(p_n, f)$.

$$\begin{aligned}
f' &= \left(\frac{x_p + x_1 - \varepsilon}{2}, \frac{x_p - x_1 + \varepsilon}{2} \right) \\
d^2(p_n, f') &= \left(x_n - \frac{x_p + x_1 - \varepsilon}{2} \right)^2 + \left(\frac{x_p - x_1 + \varepsilon}{2} \right)^2 \\
&= \left(\frac{x_n + x_1 - \varepsilon + \delta}{2} \right)^2 + \left(\frac{x_n - x_1 + \varepsilon - \delta}{2} \right)^2
\end{aligned}$$

Let $\xi = \varepsilon - \delta$:

$$d^2(p_n, f') = \left(\frac{x_n + x_1 - \xi}{2} \right)^2 + \left(\frac{x_n - x_1 + \xi}{2} \right)^2$$

By the same type of reasoning than in 1.a, we can easily see that $d^2(p_n, f') - d^2(p_n, f) = \frac{\xi^2}{2} > 0$, making the truthful distance smaller than the modified (lied) one.

Therefore, we have proved that an extreme agent does not have an interest in lying to approach the facility to them.

Mid agent We now consider all the agents that are not extreme and we prove, by cases, that they do not have an interest in lying.

Stays a mid agent If a mid agent lies but stays a mid agent, then it will have no effect on the location of the facility, as f only depends on the position of the extremes.

Becomes an extreme agent (W.L.O.G. it becomes rightmost)

$$x'_n = x_n + \varepsilon, \varepsilon > 0 \in \mathbb{R}$$

The new facility location can be written as

$$f' = \left(\frac{x_n + \varepsilon + x_1}{2}, \frac{x_n + \varepsilon - x_1}{2} \right)$$

and we can, with the same type of reasoning as in parts 1.a and 1.c, compute the difference between the new squared distance and the truthful squared distance :

$$\begin{aligned} d^2(p_i, f) &= \left(\frac{2x_i - x_n - x_1}{2} \right)^2 + \left(\frac{x_n - x_1}{2} \right)^2 \\ &= \frac{4x_i^2 + x_n^2 + x_1^2 - 4x_i x_n - 4x_i x_1 + 2x_n x_1 + x_n^2 - 2x_n x_1 + x_1^2}{4} \\ &= \frac{2x_i^2 + x_n^2 + x_1^2 - 2x_i x_n - 2x_i x_1}{2} \\ d^2(p_i, f') &= \left(\frac{2x_i - x_n - x_1 - \varepsilon}{2} \right)^2 + \left(\frac{x_n - x_1 + \varepsilon}{2} \right)^2 \\ &= x_i^2 - x_i x_1 - x_i x_n - x_i \varepsilon + \frac{1}{2} x_1^2 + \frac{1}{2} x_n^2 + x_n \varepsilon + \frac{1}{2} \varepsilon^2 \end{aligned}$$

We can now compute the difference between the distances :

$$d^2(p_i, f') - d^2(p_i, f) = \underbrace{\frac{1}{2} \varepsilon^2}_{>0} + \underbrace{\varepsilon}_{>0} \underbrace{(x_n - x_i)}_{\geq 0} > 0$$

Therefore, we can conclude that no agent never has an interest in lying about her true position, making our mechanism **strategyproof**. $\square$

3.3 Approximation Ratio

Lemma 3.2. *The Apex Mechanism achieves $\sqrt{2}$ -approximation on collinear instances in $\mathbb{R}^2$.*

Now that we can assume all the reported locations to be truthful, we want to compute the approximation ratio that our mechanism achieves. For this section, we will be using the following notations : $a = x_1$, $b = x_n$, $D = b - a$ and $m = \frac{a+b}{2}$. Note that the agents are located at $(x_i, 0)$, with $x_i \in [a, b]$.

We first remark that any single disk that covers both extremes a and b (and therefore all other agents) must have a diameter of at least D , i.e., a radius of $D/2$. Placing the facility at $(m, 0)$ achieves that every distance to agents is $\leq D/2$ (with equality $D/2$ for the extremes). Hence, $\text{OPT} = D/2$.

For any agent $x \in [a, b]$,

$$d^2((x, 0), f) = (x - m)^2 + (D/2)^2.$$

Since $|x - m| \leq D/2$, we get that

$$d^2((x, 0), f) \leq (D/2)^2 + (D/2)^2 = D^2/2,$$

with equality at $x = a$ and $x = b$. The worst case distance under our rule is therefore :

$$\max_i \{d((x_i, 0), f)\} = \sqrt{D^2/2} = D/\sqrt{2}.$$

The approximation ratio that our mechanism achieves can therefore be computed as :

$$\frac{\text{MaxCost}}{\text{OPT}} = \frac{D/\sqrt{2}}{D/2} = \sqrt{2}$$

This concludes the discussion on the approximation ratio of the APEX MECHANISM, that is $\sqrt{2}$. $\square$

3.4 Optimality

In this subsection we present an alternative (and much shorter) proof of the $\sqrt{2}$ lower bound for deterministic strategyproof mechanisms on collinear instances, under an additional regularity assumption: *scale-invariance*. Although we do not ultimately rely on scale-invariance in our main results, the argument below provides a clean geometric explanation of why the APEX Mechanism is optimal. The general proof, which removes this assumption, is given in Appendix A.

Definition 3.3 (Scale-invariance). A deterministic mechanism f is scale-invariant if for every $\lambda > 0$ and every instance $S = \{(0, 0), (D, 0)\}$,

$$f(\{(0, 0), (\lambda D, 0)\}) = \lambda f(\{(0, 0), (D, 0)\}).$$

Intuitively, scale-invariance means that dilating the instance simply dilates the output. This is satisfied by essentially all “natural” geometric mechanisms, including the APEX MECHANISM, coordinate-wise medians, and projection-based rules.

Lower Bound Under Scale-Invariance Consider the two-agent instance $S_D = \{(0, 0), (D, 0)\}$ and let

$$f(S_D) = (x, y).$$

Scale-invariance implies that for every D the output lies on a fixed ray from the origin. Thus there exist constants $(\alpha, \beta) \neq (0, 0)$ such that

$$f(S_D) = (D\alpha, D\beta) \quad \forall D > 0.$$

Now consider the right agent located at $(D, 0)$. If she misreports her position as $(D', 0)$, the mechanism outputs

$$f(S_{D'}) = (D'\alpha, D'\beta).$$

Strategyproofness requires that the deviation cannot reduce her distance to the facility:

$$\|(D'\alpha, D'\beta) - (D, 0)\| \geq \|(D\alpha, D\beta) - (D, 0)\| \quad \forall D'.$$

The left-hand side is minimized exactly when $(D'\alpha, D'\beta)$ is the orthogonal projection of $(D, 0)$ onto the ray $\text{span}\{(\alpha, \beta)\}$. Hence strategyproofness forces

$$(D\alpha, D\beta) = \text{proj}_{\text{span}(\alpha, \beta)}(D, 0).$$

A simple computation shows that this projection condition is equivalent to

$$(D\alpha)^2 + (D\beta)^2 = D(D\alpha),$$

or, dividing by D^2 ,

$$\alpha^2 + \beta^2 = \alpha.$$

Multiplying again by D^2 and writing $(x, y) = (D\alpha, D\beta)$, we obtain the circle identity

$$x^2 + y^2 = Dx.$$

Thus every scale-invariant SP mechanism must output a point (x, y) lying on the circle of radius $D/2$ centered at $(D/2, 0)$ (see Fig. 5):

$$(x - D/2)^2 + y^2 = (D/2)^2.$$

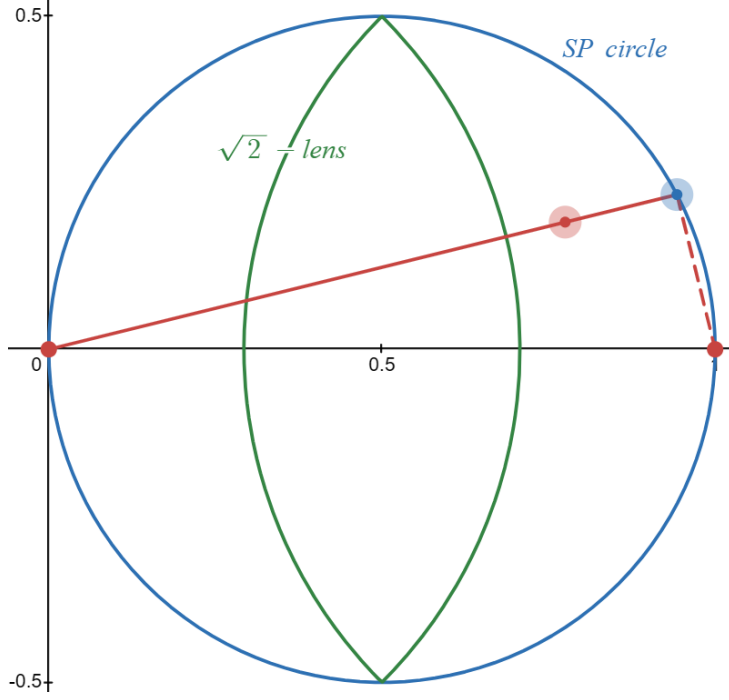


Figure 5: Visual proof of the $\sqrt{2}$ -lens being included in the SP circle

Approximating the optimum. The optimal facility location for S_D is $(D/2, 0)$ with maximum distance $D/2$. Any mechanism constrained to lie on the above circle has maximum distance

$$r = \sqrt{x^2 + y^2} = \sqrt{(x - D)^2 + y^2}.$$

This distance is minimized exactly when the ray (α, β) makes a 45° angle with the horizontal axis, yielding

$$r = \frac{D}{\sqrt{2}}.$$

Hence the approximation ratio is at least

$$\frac{D/\sqrt{2}}{D/2} = \sqrt{2},$$

with equality achieved at

$$f(S_D) = \left(\frac{D}{2}, \frac{D}{2}\right),$$

which is precisely the output of the APEX MECHANISM. Therefore:

Theorem 3.4. *Under scale-invariance, no deterministic strategyproof mechanism can achieve an approximation ratio strictly smaller than $\sqrt{2}$ on collinear instances in $\mathbb{R}^2$. The Apex Mechanism attains this bound and is thus optimal.*

Why Scale-Invariance Is Not Needed The argument above relies crucially on scale-invariance to deduce that the set of possible outputs lies on a fixed ray. In Appendix A we show that this assumption is in fact unnecessary: for deterministic mechanisms in Euclidean spaces, *strategyproofness together with convexity of the range implies continuity* [Peters et al., 1993]. Continuity alone is sufficient to carry out the geometric “projection argument” without any scale assumptions.

Lemma 3.5 (Convex range suffices). *Let f be a deterministic, anonymous, unanimous strategyproof mechanism with convex range. Then f is continuous, and the projection argument used in Appendix A applies without assuming scale-invariance. In particular, the same $\sqrt{2}$ lower bound holds for all such mechanisms.*

Thus scale-invariance provides a clean geometric shortcut, but is not required for the correctness of the lower bound. The full derivation without any scale assumptions appears in Appendix A.

3.5 Randomization cannot do better !

In this subsection, we are going to argue that no randomized mechanism can achieve a strictly better than $\sqrt{2}$ guarantee in our setting, i.e., **collinear instances in the 2D-plane**. Specifically, we will consider the two-agent instance $\{(0,0), (1,0)\}$ and show the following theorem.

Theorem 3.6. *No randomized mechanism defined in $\mathbb{R}^2$ with an assumption of collinear agents can guarantee a better than $\sqrt{2}$ -approximation.*

Our proof uses an assumption of scale-invariance of the mechanisms, which is usual for natural mechanisms like APEX or any other mechanism discussed in this paper. On our two-agent instance, any randomized mechanism outputs a mixture of points $(a_1, b_1), \dots, (a_n, b_n)$, each assigned to a probability p_i , where $0 < p_i \leq 1$ and $\sum_i p_i = 1$.

We consider the possible misreports of the right agents located at $(1,0)$ as $(x,0)$. Because of scale-invariance, for any misreport $(x,0)$ our mechanism has to output in (xa_i, xb_i) with a probability p_i for each of the points. We can therefore compute the function

$$f(x) = \sum_{i=1}^n p_i \sqrt{(1 - xa_i)^2 + (xb_i)^2}$$

that represents the expected distance of the output to the true location of right agent, i.e., $(1,0)$ as a function of the misreport x . Strategyproofness implies that the right agent's best choice is to report truthfully, i.e., $f(x)$ should be minimized at $x = 1$.

Under this condition, we want to minimize the function

$$\sum_{i=1}^n 2p_i \sqrt{\max\{a_i, 1 - a_i\}^2 + b_i^2}$$

that computes the average approximation ratio of our mechanism. Therefore, our problem (once the points (a_i, b_i) have been fixed) comes down to an optimization problem over the weights p_i :

$$\min_p S(p) = \min_p \sum_i p_i s_i$$

subject to

$$C(p) = \sum_i p_i c_i = 0, \quad \sum_i p_i = 1, \quad p_i \geq 0,$$

where

$$s_i = 2\sqrt{\max\{a_i, 1 - a_i\}^2 + b_i^2}, \quad c_i = \frac{a_i^2 + b_i^2 - a_i}{\sqrt{(1 - a_i)^2 + b_i^2}}.$$

Using the extreme point principle, i.e., *A linear function on a compact convex polytope attains its minimum at an extreme point*, we can only analyze the case $p_1 = 1$.

As this comes back to the deterministic case, we have already proved that APEX is optimal and that no mechanism can guarantee a better than $\sqrt{2}$ -approximation on this instance, generalizing this lower bound to any randomized mechanism in our setting. $\square$

3.6 Discussion about Scale-invariance and Randomization

We can note that in several proofs along this paper, we use translation-invariance ($T(f(x)) = f(T(x))$) and rotation-invariance ($R(f(x)) = f(R(x))$). One could argue that APEX MECHANISM breaks either rotation invariance or translation-invariance. Indeed, if we consider the input $S = \{(0, 0), (1, 0)\}$, APEX outputs $(1/2, 1/2)$. Let T be the translation by 1 to the left and R be the rotation of π rad around the origin. It is easy to see that $R(S) = T(S) = \{(0, 0), (-1, 0)\}$ and therefore the output on this new instance should be $R(1/2, 1/2) = (-1/2, -1/2)$ and $T(1/2, 1/2) = (-1/2, 1/2)$ at the same time, which cannot happen deterministically.

We have two arguments to discuss this question. Firstly, rotation-invariance is not necessarily required in our specific setting, since we are working on a restricted domain of collinear instances, rather than on arbitrary point sets in $\mathbb{R}^2$. Translation-invariance and reflection-invariance can be argued to be the isometries that we want to preserve in our setting.

If we still want to keep rotation-invariance, this is the time when **randomization** can become helpful. Instead of outputting

$$\left(\frac{x_1 + x_n}{2}, \frac{x_n - x_1}{2} \right),$$

one can decide to output on the mixture

$$f(x) = \begin{cases} \left(\frac{x_1 + x_n}{2}, \frac{x_n - x_1}{2} \right), & \text{w.p. } 1/2 \\ \left(\frac{x_1 + x_n}{2}, -\frac{x_n - x_1}{2} \right), & \text{w.p. } 1/2 \end{cases}.$$

This small change allows us to keep the definitions with rotation-invariance and translation-invariance **in distribution**. We will later see that randomization turns out to be helpful when trying to import the results in this setting to general $\mathbb{R}^2$ mechanisms.

4 Generalization to d Dimensions

Now that we have a satisfying understanding of the gain that we get by adding one dimension to linear profiles, we want to study this same effect on higher dimensions. We first focus on the attempt of using an additional dimension to achieve "lifted" strategyproof mechanisms for planar profiles. Specifically, we look at the intuitive generalization of APEX to the $\mathbb{R}^3$ output but planar (in $\mathbb{R}^2$) instances.

4.1 Negative Result of Intuitive Generalization

Lemma 4.1 (Failure of lifted 1-center mechanisms in $\mathbb{R}^2 \rightarrow \mathbb{R}^3$). *Consider the following deterministic mechanism for the max-cost objective. Given a profile of agents $x_1, \dots, x_n \in \mathbb{R}^2$, let*

$$c \in \mathbb{R}^2 \quad \text{and} \quad r \geq 0$$

be respectively the center and radius of the smallest enclosing circle of the reported points. The mechanism outputs the facility

$$f(x_1, \dots, x_n) = (c, r) \in \mathbb{R}^3.$$

Then this mechanism is not strategyproof.

Proof. Consider the following instance :

$$\{(0, 1, 0), (0, -1, 0), (1, 0, 0)\},$$

and consider the misreports of the third agent (the one in $(1,0,0)$) as $\tilde{x}_3 = (x,0,0)$ for some $x \geq 1$ (she misreports to the right, along the x-axis). Then, we can easily compute the center of the smallest enclosing circle for this instance as

$$c = \left(\frac{x^2 - 1}{2x}, 0, 0 \right).$$

Let's call $c(x) := c_x = \frac{x^2 - 1}{2x}$. We can also compute the radius of the circle,

$$r(x) = x - \frac{x^2 - 1}{2x} = \frac{x^2 + 1}{2x}.$$

Therefore, the distance from the agent's true location in $(1,0,0) \in \mathbb{R}^3$ and the output in $(c(x), 0, r(x)) \in \mathbb{R}^3$ is

$$d(x) := \sqrt{(1 - c(x))^2 + r(x)^2} = \sqrt{\left(\frac{x^2 - 2x - 1}{2x} \right)^2 + \left(\frac{x^2 + 1}{2x} \right)^2} = \sqrt{\frac{x^4 - 2x^3 + 2x^2 + 2x + 1}{2x^2}}.$$

Since the square root function is increasing, minimizing $d(x)$ is equivalent to minimizing the expression inside of the square root. We get

$$d(x)^2 \propto \frac{x^4 - 2x^3 + 2x^2 + 2x + 1}{x^2} = x^2 - 2x + 2 + \frac{2}{x} + \frac{1}{x^2} =: f(x).$$

To minimize $f(x)$, we get it's derivative and find the real solutions :

$$\begin{aligned} \frac{df}{dx}(x) &= 2x - 2 - \frac{2}{x^2} - \frac{2}{x^3} \\ &\rightarrow 2x^4 - 2x^3 - 2x - 2 = 0 \\ &\rightarrow x^4 - x^3 - x - 1 = (x^2 + 1)(x^2 - x - 1) = 0, \end{aligned}$$

whose only real solution that is ≥ 1 is $x = \frac{1+\sqrt{5}}{2}$. □

Hence the intuitive generalization of APEX to the dimension above does not keep its strategyproofness. Note that the same type of result appears if we decide to output in $(c, ar) \in \mathbb{R}^3$, with $a > 0$ some constant of "height". As a grows, the minimizer of the distance function $d(x)$ tends to $x = 1$, but reaches it only at the limit, when the height of the output is infinitely big, yielding a degenerate mechanism.

4.2 Trying to go beyond the naive generalization

Lemma 4.1 shows that the most natural way of extending the Apex idea to planar instances does not work. Lifting the center of the smallest enclosing circle by a multiple of its radius already fails to be strategyproof for three agents. This highlights an important difference between the one-dimensional and two-dimensional settings. In $\mathbb{R}^2$, agents can influence the geometry of the instance in several directions, which makes manipulation much easier.

This negative result does not rule out all possible uses of higher-dimensional outputs. Instead, it suggests that we need to rethink how extra dimensions are used. In this section, we briefly discuss alternative directions that could still be explored.

4.2.1 Rethinking mechanisms in $\mathbb{R}^2 \rightarrow \mathbb{R}^3$

The failure of the lifted 1-center construction comes from the fact that the facility location depends smoothly on the center and radius of the smallest enclosing circle. As shown in the previous section, this smooth dependence allows agents to slightly change the geometry of the instance and strictly decrease their cost.

One possible idea is to move away from global geometric objects such as the smallest enclosing circle, and instead rely on simpler, more rigid structures. For example, one may consider mechanisms that:

- select a specific pair of agents according to a deterministic rule,
- place the facility at the midpoint of this pair in the plane,
- and lift the facility in the third dimension by a height that depends on the distance between the two selected agents.

Mechanisms of this form resemble Apex in spirit, since they reduce the geometry to a single segment, even though the agents live in $\mathbb{R}^2$. Such constructions were tested and were found not to be strategyproof in general. However, they illustrate a potentially useful idea: reducing a two-dimensional instance to a one-dimensional structure before using the extra dimension.

Overall, if one hopes to beat the classical 2-approximation bound using outputs in $\mathbb{R}^3$, it is unlikely that this can be achieved by a simple, smooth lifting of standard geometric centers.

4.2.2 Using higher-dimensional outcome spaces

A more speculative direction is to allow the facility to be placed in a higher-dimensional space $\mathbb{R}^{f(2)}$ (where $f(2) \geq 3$), while agents remain in $\mathbb{R}^2$. At first glance, additional dimensions may seem useless, since any component orthogonal to the plane of the agents only increases distances. However, the APEX mechanism shows that extra dimensions can still be useful as a way to control incentives, rather than to improve the geometric optimum itself.

One idea is to use several independent “height” coordinates, each linked to a different feature of the reported instance. For example, these coordinates could depend on distances in different directions, or on distances to different extreme agents. The goal would be to make the projection of the facility onto the plane stable, while deviations are penalized through increases in one or more **orthogonal** coordinates.

Another way to view higher-dimensional outputs is as a deterministic analogue of smoothing. By spreading incentive penalties across multiple dimensions, it may become harder for a single agent to find a deviation that reduces her overall distance.

That said, increasing the dimension of the outcome space also introduces new difficulties. Any such mechanism must still guarantee a bounded approximation ratio with respect to the planar optimum, and avoid creating new forms of manipulability. At present, it is unclear whether allowing more than one additional dimension can overcome the incentive issues identified in Lemma 4.1, or whether these issues persist in all higher dimensions.

5 Use of Apex for General 2D Mechanisms

APEX MECHANISM is an interesting result, but it relies on the collinearity assumption of the agents. What we want is a mechanism that works for any $\mathbb{R}^2$ instance, that acts “Apex-like” for collinear profiles, but keeps strategyproofness.

5.1 Impossibility Result : Deterministic

As discussed in section 3.6, translation-invariance and rotation-invariance can help us to understand how Apex cannot be generalized to any instance in $\mathbb{R}^2$. What we mean by *any instance in $\mathbb{R}^2$* is that our agents can be located anywhere in the plane, i.e., they're not assumed to be collinear. In this setting, no known mechanism can achieve better than a 2-approximation without predictions. We want to study an import of APEX to this setting, so as to construct a mechanism that uses APEX if the agents happen to be collinear (without assumption of such condition) and something else otherwise.

In this first subsection, we will show that this is impossible deterministically. In this objective, we will consider the two agent instance

$$S = \{(0, 0), (1, 0)\}.$$

Since our mechanism has to output APEX if the agents are collinear and any two agents are collinear, it must output either $(1/2, 1/2)$ or $(1/2, -1/2)$, deterministically. We will consider the following lemma :

Lemma 5.1. *For any two-agent instance and a deterministic Apex output, there exists a small enough radius $r > 0$ such that any misreport of any of the two agents in a ball of radius r around her original position does not switch the output from one side of the line connecting the two agents to the other.*

This is a very standard assumption / consequence of 1-Lipschitzness. The intuition is that for two very similar two-agent instances, the output cannot go from the *upper Apex* to the *lower Apex* and reversely. Using this lemma, it is very easy to construct a counterexample that shows how APEX cannot be used deterministically for any collinear $\mathbb{R}^2$ input.

Once again, consider the instance

$$S = \{(0, 0), (1, 0)\}$$

and let the output of APEX be located in

$$\begin{aligned} \text{case 1: } & (1/2, 1/2) \\ \text{case 2: } & (1/2, -1/2). \end{aligned}$$

In case 1, the left-agent can decide to misreport by dragging her position from $(0, 0)$ to $(0, -1)$. By our continuity assumption, this action drags the output from $(1/2, 1/2)$ to $(0, 0)$, hence the output is **on** the left-agent's true position, giving her an incentive to lie. In the second case, by a similar reasoning, the left-agent has an incentive to lie and misreport at $(0, 1)$, which once again brings the output **on** her true location.

Therefore, there cannot be a strategyproof mechanism defined over all $\mathbb{R}^2$ instances that outputs APEX deterministically when the agents are collinear.

A more generalized version of this counter example is to consider any two-agent instance

$$S = \{(x_1, y_1), (x_2, y_2)\}$$

and their midpoint in

$$(m_x, m_y).$$

We can now consider rotative movements of the agents around the midpoint. Note that we are not studying strategyproofness here since it is a concept that applies (in this paper) only on one agent's displacement, and not to *group movements*. However, we can still study what happens when agents move one-by-one. For example, consider the following set of rotations, performed one after the other :

$$\{R_{\pi/2}^{(1)}, R_{\pi/2}^{(2)}, R_{\pi/2}^{(1)}, R_{\pi/2}^{(2)}\},$$

where $R_\theta^{(i)}$ represents the rotation of agent i by an angle of θ around the midpoint (m_x, m_y) . We can easily see, using our continuity lemma, that the output of a deterministic APEX simply switches sides after these four rotations. However, since we simply performed two π rotations around the midpoint of (x_1, y_1) and (x_2, y_2) , we have the exact same instance as the original one, but the agents have swapped their positions. By anonymity and since the instance is the same up to swaps of the agents, we should have the exact same output as the original one, and therefore we find a contradiction. $\square$

5.2 Randomized

We first note that the impossibility proofs that we have used previously to show that any deterministic SP mechanism **must** output on the line whenever the input is collinear do not work for the randomized scenario. We can for example consider the following mechanism :

Mechanism 5.2 (Balanced Apex). On a collinear instance in $\mathbb{R}^2$, perform a change of basis as to get the instance $\{p_1, \dots, p_n\}$, where $p_i = (x_i, y_i) = (x_i, 0)$. Output the mixture

$$\begin{cases} (D/2, D/2), & \text{w.p. } 1/2 \\ (D/2, -D/2), & \text{w.p. } 1/2 \end{cases}$$

If we try to construct the same rotation-argument that we had used for the deterministic case, we find that the rotation simply modifies the output from

$$\begin{cases} (D/2, D/2), & \text{w.p. } 1/2 \\ (D/2, -D/2), & \text{w.p. } 1/2 \end{cases}$$

to

$$\begin{cases} (D/2, -D/2), & \text{w.p. } 1/2 \\ (D/2, D/2), & \text{w.p. } 1/2, \end{cases}$$

which is exactly the same output up to the order. This is the main reason that tells us that randomization might find its usefulness in this setting.

5.2.1 The Two-agent Case

We start with this warm-up case, where we have two agents $\{p_1, p_2\} = \{(x_1, y_1), (x_2, y_2)\}$, where $x_1, x_2, y_1, y_2 \in \mathbb{R}$. From the axioms of Euclidean geometry, we know that two points are **always** collinear, so we can simply try to use BALANCED APEX MECHANISM.

Up to a change of basis and a scaling, any two-agent instance can be brought to the $\{(0, 0), (1, 0)\}$ instance. We will study the incentive of left-agent about lying to show that this mechanism on a two-agent instance is strategyproof. Consider the report of left agent to be in (x, y) . By easy computations, one can see that BALANCED APEX will output at the points

$$\begin{cases} (x_1, y_1) = \left(\frac{x+y+1}{2}, \frac{-x+y+1}{2} \right), & \text{w.p. } 1/2 \\ (x_2, y_2) = \left(\frac{x-y+1}{2}, \frac{-x+y-1}{2} \right), & \text{w.p. } 1/2. \end{cases}$$

The distance from left agent's true location and the facility is therefore

$$f(x, y) = \frac{1}{2} \|(0, 0), (x_1, y_1)\| + \frac{1}{2} \|(0, 0), (x_2, y_2)\| = \frac{1}{2} \sqrt{x_1^2 + y_1^2} + \frac{1}{2} \sqrt{x_2^2 + y_2^2}.$$

By developing and rearranging the terms, we obtain

$$f(x, y) = \frac{1}{2} \sqrt{\frac{x^2 + (y+1)^2}{2}} + \frac{1}{2} \sqrt{\frac{x^2 + (y-1)^2}{2}},$$

i.e., half the sum of the distances from (x, y) to $(0, 1)$ and $(0, -1)$. By the triangle inequality, we know that this function is minimized exactly on the segment $[(0, 1), (0, -1)]$. Specifically, the point $(0, 0)$ is on this segment and is therefore a minimizer. Hence, the left agent does not strictly gain by reporting another point than her true location.

5.2.2 Three Agents or more

What we want for three agents and more is a mechanism that uses BALANCED APEX when agents are collinear but something else otherwise. Intuitively, it is good to think that an *almost collinear* instance should yield an *almost Apex* behaviour to the mechanism. We don't want the "Apex case" output to be far from the "non-Apex" one, we would like a continuous mechanism. Several ideas can be studied, and it is still an open question to know whether such mechanisms can exist while keeping strategyproofness.

A first idea is to identify a "best fitting" line for the reported agent locations, project all agents onto this line, and apply the APEX MECHANISM to the resulting one-dimensional instance. The motivation is that if agents are nearly collinear, the chosen line should be close to the true underlying structure, yielding a continuous extension of Apex beyond perfectly collinear profiles.

This approach can be instantiated in several ways, for example by choosing:

- the line minimizing squared distances (principal component analysis),
- the line minimizing maximum distance (1-center style),
- or a line inferred via a learned predictor trained on typical instances.

However, this approach is fundamentally incompatible with strategyproofness. The choice of the line depends discontinuously on the reported locations, and a single agent can beneficially manipulate the selected line by deviating slightly, thereby influencing the projection of *all* agents. As a result, the induced mechanism fails to be strategyproof even for small numbers of agents.

Multi-line Balanced Apex. To avoid privileging a single line, a more symmetric construction was considered. For every unordered pair of agents, the line determined by that pair is considered. All agents are projected onto this line, and the two Apex locations induced by the projected instance are computed. The mechanism then outputs the uniform mixture over all such Apex points.

This construction enjoys several desirable properties:

- it is anonymous and symmetric,
- it agrees exactly with the Apex mechanism on collinear instances,
- and it varies continuously as instances approach collinearity.

Despite these appealing features, the mechanism is not strategyproof. Although each individual "project-and-Apex" component is strategyproof in one dimension, strategyproofness does not compose under averaging when the components themselves depend on the reported profile. A single agent can influence many lines simultaneously, creating correlated shifts in the support of the output distribution that result in a profitable deviation.

5.3 Uniform Pairwise Apex (UPA)

The most interesting result that we achieve with this Apex idea is the following mechanism.

Mechanism 5.3 (Uniform Pairwise Apex (UPA)). Consider all unordered pairs of agents $\{i, j\}$. For each pair, consider the two Apex points induced by (x_i, x_j) . The mechanism outputs a uniform distribution over all such Apex points.

Lemma 5.4. *Uniform Pairwise Apex is (individually) strategyproof.*

Proof. Since we want to study individual strategyproofness, we can fix agent i and split the outcome distribution into pairs that do and pairs that do not involve agent i . For the outcomes that are computed with a pair that does not involve i , they are completely independent on the displacement / reported location of this agent, therefore they do not have any influence on the strategyproofness of the mechanism relatively to this agent.

For the outcomes computed with $(i, \star)$ pairs, for any $\star$, the computation of the output falls down to a BALANCED APEX, which is strategyproof (Mechanism 5.2). Since our mechanism is a simple mixture over constant or strategyproof mechanisms with respect to a fixed agent, it is strategyproof with respect to this same agent. The operation can be done for any agent, hence our mechanism is strategyproof. $\square$

We have therefore constructed a strategyproof mechanism that does not necessarily output on the line containing the agents if they are collinear, but we now want to study its approximation guarantees.

5.3.1 Bounds

Theorem 5.5. *Uniform Pairwise Apex has a (maybe not tight) $\sqrt{5}$ -ratio upper bound for its approximation guarantee.*

Proof. Refer to Figure 6 for a visual example. Consider any set of agents $\{x_1, \dots, x_n\}$, $x_i \in \mathbb{R}^2$ and their smallest enclosing circle of diameter D . We want to study the approximation ratio of an arbitrary agent i . For any output point, consider the two agents j and k that were used to compute this output point. Note that i can be one of these two agents. Since both j and k are in the circle, we have that

$$\|x_j - x_k\| \leq D.$$

Consider the midpoint between x_j and x_k and call it m . We define by $a_{1,2}$ the Apex point (notation when an equation holds for both Apex points). We have that

$$\|a_{1,2} - m\| = \frac{\|x_j - x_k\|}{2} \leq \frac{D}{2},$$

i.e., the distances between the midpoint of the two agents and the Apex points are smaller than the radius of the circle. As said before, both agents are in the circle. By basic laws of Euclidean geometry, this means that the midpoint m is also in this circle. Since x_i , the agent that we want to study, is also inside the circle, we have that

$$\|x_i - m\| \leq D.$$

We could use *triangle inequality* to obtain a first upper bound but this yields a 3-approximation bound, which is loose. We will rather use the orthogonality of the system to get a tighter bound. We want to study the distance $\|x_i - a_{1,2}\|$, so we first see that the segment from m to $a_{1,2}$ is orthogonal to the line connecting the two agents j and k . We can decompose the vector $x_i - m$ into a component parallel to this line, that we call v and a component perpendicular, that we

will consider that the mechanism outputs only one Apex per pair, i.e., it is not balanced, as this simplifies notations. However, this does not change the result as all cost functions are the same with the symmetrical point.

$(n-1)$ of the Apex points — the ones whose pair contains the extreme agent — will be placed in $(1, 1)$ and the $\frac{(n-1)(n-2)}{2}$ remaining ones — the ones whose pair consists of two in-cluster agents — will be placed in $(0, 0)$. The $(1, 1)$ points are at a distance $\sqrt{1+1} = \sqrt{2}$ of the extreme agent and the points in $(0, 0)$ will be at a distance 2 of the extreme agent. Therefore, the expected cost from the extreme agent to the facility location is

$$\begin{aligned} & \left(\frac{(n-1)(n-2)}{2} 2 + (n-1)\sqrt{2} \right) \cdot \frac{2}{n(n-1)} \\ &= \frac{2(n-2+\sqrt{2})}{n}. \end{aligned}$$

Since the optimal facility location (in $(1, 0)$) achieves a 1-approximation, our mechanism achieves a $\frac{2(n-2+\sqrt{2})}{n}$ approximation. $\square$

We will now show that this lower bound is tight for collinear instances, i.e., UPA can always guarantee a $\frac{2(n-2+\sqrt{2})}{n}$ -approximation ratio if the agents are on a same line.

Theorem 5.7. *On collinear profiles, UPA never performs worse than $\frac{2(n-2+\sqrt{2})}{n}$.*

Proof. We assume that all agents lie on the x -axis and we let

$$0 = x_1 \leq x_2 \leq \dots \leq x_{n-1} \leq x_n = D,$$

which is obtained after translation, rotation or/and scaling. Here the optimal approximation ratio is $D/2$, the midpoint. For any pair (i, j) with $i < j$, we define :

$$\begin{aligned} \text{midpoint: } m_{ij} &= \frac{x_i + x_j}{2} \\ \text{pair distance: } d_{ij} &= x_j - x_i. \end{aligned}$$

UPA samples the two apexes for each pair, but since they have the same cost vector with any agent on a collinear profile, we will ignore the $\pm$ and the just analyze the pair.

For a fixed pair (i, j) , the Apex locations are $(m_{ij}, \pm d_{ij}/2)$. For any agent at position x , her distance to that Apex is

$$\sqrt{(x - m_{ij})^2 + (d_{ij}/2)^2}.$$

Note that for fixed m_{ij} and d_{ij} (i.e., fixed pair (i, j)), the function

$$x \mapsto \sqrt{(x - m_{ij})^2 + (d_{ij}/2)^2}$$

is convex in x and symmetric around m_{ij} , hence its maximum over the interval $[0, D]$ is achieved at one of the endpoints $x \in \{0, D\}$. Therefore the maximum cost for pair (i, j) is

$$C_{ij} = \max \left\{ \sqrt{m_{ij}^2 + (d_{ij}/2)^2}, \sqrt{(D - m_{ij})^2 + (d_{ij}/2)^2} \right\}.$$

Equivalently, if we let $r(m) := \max \{m, D - m\}$, we can write

$$C_{ij} = \sqrt{r(m_{ij})^2 + (d_{ij}/2)^2}.$$

Thus the expected max-cost of UPA is

$$\mathbb{E} \left[\max_k d(x_k, \text{UPA}(x)) \right] = \frac{1}{\binom{n}{2}} \sum_{i < j} C_{ij}.$$

If we fix all agents except some agent t , whose position is $x \in [0, D]$ and we consider the contribution to the total sum over pairs that involve t , we have :

$$\Phi_t(x) := \sum_{j \neq t} C(x, x_j),$$

where $C(x, x_j)$ denotes the max-cost value C_{ij} computed from the pair (x, x_j) . We claim that $\Phi_t(x)$ is convex on $[0, D]$. Indeed, for each fixed x_j , we define

$$m(x) := \frac{x + x_j}{2}, \quad d(x) := |x - x_j|.$$

Then,

$$C(x, x_j) = \sqrt{r(m(x))^2 + (d(x)/2)^2}.$$

Now we check convexity ingredient by ingredient :

- $m(x)$ is affine in x .
- $r(m) = \max\{m, D - m\}$ is the max of two affine functions, hence **convex** in m , hence **convex** in x by composition with an affine map.
- $d(x) = |x - x_j|$ is convex in x , hence $d(x)/2$ also **convex** in x .
- The function $f(u, v) = \sqrt{u^2 + v^2}$ is **convex** on $\mathbb{R}^2$ (since it's a norm) and **coordinatewise nondecreasing** for $u, v \geq 0$.
- Therefore, by the standard composition rule for convex functions, we have that if f is convex and nondecreasing in each argument on the relevant domain and $u(x), v(x)$ are convex and nonnegative, then $x \mapsto f(u(x), v(x))$ is **convex**.

We apply this characterization with $u(x) = r(m(x)) \geq 0$, $v(x) = d(x)/2 \geq 0$. This shows that each term $C(x, x_j)$ is convex in x , making their sum $\Phi_t(x)$ also **convex** in x . Hence the maximum of $\Phi_t(x)$ is reached at one of the endpoints, i.e.,

$$\Phi_t(x) \geq \max \{ \Phi_t(0), \Phi_t(D) \}.$$

So, one of the endpoints satisfies : moving the agent from her interior location x to this endpoint does not decrease the expectation.

By repeatedly applying the "push to an endpoint" step, we can transform any collinear instance into a profile where all agents lie in $\{0, D\}$ without decreasing the expected max-cost under UPA. Thus the worst-case collinear instance for UPA is the profile where we have k agents at $x = 0$ and $n - k$ agents at $x = D$. For such a profile, if a sampled pair is $(0, 0)$ or (D, D) , then $d = 0$, Apex is exactly at the endpoint, and the max-cost is the opposite point, at a distance D , hence ratio = $\frac{D}{D/2} = 2$. If a sample pair is $(0, D)$, then $m = D/2$, $d = D$, hence the Apex is at midpoint raised by half the distance, which achieves a $\sqrt{2}$ ratio. Hence the expected ratio for a $k/(n - k)$ split is

$$\text{AR}(k) = 2 \cdot \frac{\binom{k}{2} + \binom{n-k}{2}}{\binom{n}{2}} + \sqrt{2} \cdot \frac{k(n-k)}{\binom{n}{2}}.$$

This is maximized when the split is the most imbalanced, i.e., $k = 1$ or $k = n - 1$, yielding exactly

$$\text{AR}_{\text{collinear}}^{\max} = \frac{2(n - 2 + \sqrt{2})}{n},$$

which matches our "cluster + 1 outlier" example, making our lower bound also an upper bound, for collinear profiles. $\square$

Beyond collinear instances, we conjecture that the $\sqrt{5}$ approximation ratio does not characterize the worst-case performance of UPA on any arbitrary planar input in $\mathbb{R}^2$. This belief is supported by empirical and structural evidence. Since UPA decomposes into an average of pairwise Apex outcomes and that the maximum of this average seems to be reached in degenerate cases like the "cluster + 1 outlier" example, any dispersion of the cluster or introduction of additional extreme agents seems to increase the fraction of "good pairs", decreasing the average. Experiments for small n corroborate this idea.

Some examples however show that a $\frac{2(n-2+\sqrt{2})}{n}$ upper bound would be broken in some cases. We can however believe that UPA might achieve a worst-case approximation ratio of $2 - \Theta(1/n)$ on all instances in $\mathbb{R}^2$, strictly improving over the classical 2 bound and having a similar result to the currently best known bound of $2 - 1/n$ achieved using centroids. Importantly, this improvement would be achieved without using predictions, positioning UPA as a simple and robust mechanism with the strongest known worst-case guarantee in this setting. What we believe is that the worst-case of UPA is a slowly growing function ρ_n of n , and that it converges to some $\rho_\infty < \sqrt{2}$. We expect the following behaviour :

$$\frac{2(n - 2 + \sqrt{2})}{n} \leq \rho_n \leq 2, \quad \forall n \in \mathbb{N}.$$

5.3.2 A Surprising Behaviour

Even though the construction of UPA comes from a reasonably intuitive generalization of APEX that follows the money burning principle, the behaviour of this mechanism is surprising. If we consider n agents on the segment from $(0, 0)$ to $(1, 0)$ that placed evenly, i.e., the pairwise distance between two neighbouring agents is always the same, then our UPA outputs a mixture of uniformly distributed points in the diamond shaped area, as can be seen in Figure 7 :

$$(0, 0) \rightarrow (1/2, 1/2) \rightarrow (1, 0) \rightarrow (1/2, -1/2).$$

The expected max cost on this instance can be computed as

$$\mathbb{E}[\text{max_cost}] = 8 \int_0^{1/2} \int_0^x \frac{\sqrt{(1-x)^2 + y^2}}{1/2} dy dx \approx 1.399,$$

which beats the 1.5 achieved by centroids on this instance. This means that our mechanism does not perform poorly even with large values of n , which is surprising knowing its unnatural behaviour. Deeper understanding of this mechanism is left for future work.

5.4 Discussion about the Conjecture in Balkanski et al. [2024]

We want to study the conjecture cited in Balkanski et al. [2024], namely :

Conjecture 5.8. For a two-agent instance in $\mathbb{R}^2$, strategyproof mechanisms are always better off returning a facility on the line containing the agents' locations.

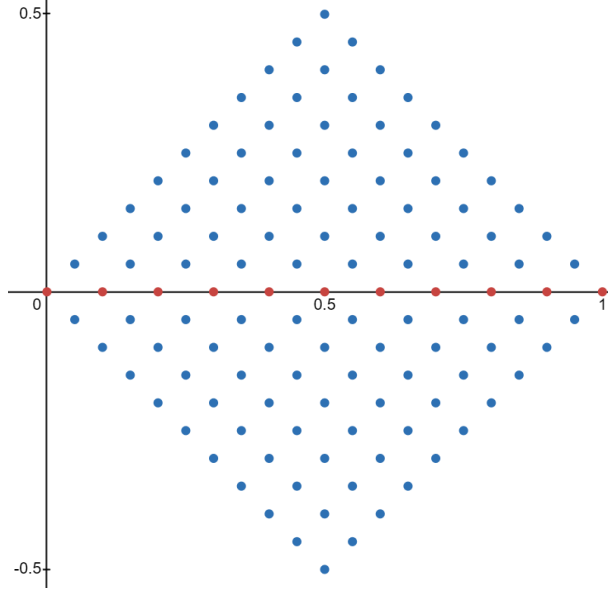


Figure 7: An example use of UPA on an evenly distributed collinear profile

What's known currently is that the best known mechanism for this profile is called centroids and it achieves $2 - 1/n = 2 - 1/2 = 3/2$ as an approximation. Note that for a two-agent instance, centroids is exactly equivalent to running LRM, that we discussed previously. This returns an on-line mixture of outputs. For $n = 2$, UPA achieves

$$\frac{2(n - 2 + \sqrt{2})}{n} = \frac{2 \cdot \sqrt{2}}{2} = \sqrt{2} \approx 1.414,$$

which beats the 1.5 ratio achieved by staying on the line. Moreover, we can note that for two agents, any truthful-in-expectation mechanism whose outcome is constrained to lie on the line inherits the one-dimensional lower bound of 1.5 for the max-cost objective. Since $\text{UPA} \equiv \text{APEX}$ achieves $\sqrt{2} < 3/2$ by placing the facility off this line, off-line outcomes can strictly improve truthful approximation, contradicting the strong form of the conjecture. $\square$

5.5 A Word about Predictions

Throughout this paper, we study facility location mechanisms in the plane under the assumption that agents may lie on a line, but the mechanism must operate on all instances in $\mathbb{R}^2$. This can be interpreted as a prediction problem : we may expect (or predict) that instances are close to collinear, but must remain robust to arbitrary deviations.

The APEX mechanism exploits this structure optimally when the prediction is correct, achieving a $\sqrt{2}$ approximation ratio. The UNIFORM PAIRWISE APEX (UPA) extends this idea to general 2D instances, achieving a $\sqrt{5}$ approximation when the prediction fails. However, for "almost collinear" instances, UPA is "almost" upper bounded by $\frac{2(n-2+\sqrt{2})}{n}$, that tends to 2 as n grows. Thus, although UPA does not explicitly take a prediction as an input, its analysis reveals a consistency-robustness structure, performances improve significantly when the instance is collinear, but remains controlled on non-collinear profiles.

6 Discussion

6.1 Recap of our Contribution

We have defined the **Collinear in $\mathbb{R}^2$** model : agents lie on a plane but we know/assume that they are collinear, without loss of generality we say that $y_i = 0$ for every agent i . We have

studied the APEX MECHANISM, its strategyproofness, its approximation ratio guarantee of $\sqrt{2}$ and that it's optimal amongst deterministic SP under this setting.

We have showed that under scale-invariance, no mechanism can beat APEX in this collinear setting, even using randomization.

We have discussed the generalization of APEX to higher dimensions, which appears to be non-trivial already for coplanar instances in $\mathbb{R}^{f(2)}$, $f(2) \geq 3$.

Lastly, we have explored the use of APEX on general instances, and after several tries, no mechanism that strictly collapses to APEX when the agents are collinear was found, but UNIFORM PAIRWISE APEX (UPA) provides a good insight of how off-line outputs can help to achieve better approximation on collinear profiles, and conjecture 5.8 was disproved using this mechanism.

6.2 Conceptual Takeaways

The main conceptual takeaway is that *Adding a dimension to the outcome can strictly improve truthful approximation*. This can be seen via our APEX model that strictly beats the best known 1.5-approximation and achieves $\sqrt{2}$.

The unsuccessful trials of generalizing APEX to higher dimensions enlighten how they differ from $\mathbb{R}$ and allow much more complex cost vectors, as the literature had discussed previously.

6.3 Limitations and Critical Assessment

This work made several structural and modeling choices that could have been done in different manners. Here are a couple of points that we may reconsider with more time and other objectives.

- A limitation of the present work is that several impossibility and lower-bound arguments rely on the assumption of scale-invariance. This choice was made since this assumption is standard in facility-location mechanisms, but it may exclude some peculiar models whose behaviour depends on absolute-distances rather than ratios.
- The strategyproofness definition and all the work focuses only on individual strategyproofness. For example, APEX gives incentives to lie for a group of agents. Extending Apex-like constructions to group-strategyproofness appears to be challenging and remains open.
- Throughout this work, we have derived all the results under the maximum (egalitarian) cost objective. Other objectives, such as utilitarian cost, can be considered for mere studies.

7 Future Work

Generalize Apex to all of $\mathbb{R}^2$

We defined UPA that behaves almost like APEX for collinear profiles, but it still outputs a mixture of more than just the two Apex points found with the extreme agents. We would like to find a mechanism that truly collapses to APEX when the instance happens to be collinear while keeping its strategyproofness, or to describe a general impossibility result that holds under randomization.

Understanding UPA

As mentioned previously, there are several points on which we could work to deepen our understanding of UPA. For example, there is a need to close the gap between the general $\sqrt{5}$ upper bound and the stronger-looking empirical bounds. We expect this mechanism to perform

no worse than 2 on any instance but this remains open.

One could also find worst-case constructions for UPA for non-collinear profiles, or prove a tighter general upper bound.

Learning-Augmented Mechanisms Tailored to "Almost Collinear"

Since predictions were not the main focus of this paper's contribution, we could try to improve this "almost collinear" idea using learning-augmented mechanisms. A formalization of this notion (distance to best-fitting line, thickness, PCA residual, etc.) and a design or impossibility result for the question : *Can we construct a learning-augmented SP mechanism that is $\sqrt{2}$ -consistent on perfectly collinear and smoothly degrades with "non-collinearity" ?*

Higher-Dimensional Outcome Spaces

Is there a construction that uses $\mathbb{R}^d \rightarrow \mathbb{R}^{d+1}$ while keeping its strategyproofness ? We could also allow larger dimensions : $\mathbb{R}^d \rightarrow \mathbb{R}^{f(d)}$ with some function f that satisfies $f(d) \geq d + 1$. To understand more broadly the behaviour of mechanisms in $\mathbb{R}^2$ that is known to have a crucially different geometry than $\mathbb{R}$, one could explore what geometric objects replace APEX (sphere/circle center + height, smallest enclosing ball, etc.) and where the manipulation arises.

8 Conclusion

This thesis studied strategyproof facility location under the maximum cost objective in a geometric setting where agents are collinear in $\mathbb{R}^2$ but the facility may be placed off the line. In this model, we introduced the APEX mechanism and showed that it is strategyproof and achieves a $\sqrt{2}$ approximation ratio, strictly improving over the best known guarantees when the facility is restricted to lie on the agents' line. We further established that this guarantee is optimal among deterministic strategyproof mechanisms, and that randomization does not yield an improvement under standard regularity assumptions.

Beyond the collinear setting, we explored the possibility of extending Apex-like ideas to more general instances. We argued that mechanisms that deterministically collapse to APEX on collinear profiles face fundamental obstacles, and we proposed the UNIFORM PAIRWISE APEX mechanism as a global alternative. UPA remains strategyproof, achieves a bounded approximation on all instances, and preserves the benefits of off-line outputs on collinear profiles. As a consequence, this work disproves the conjecture that strategyproof mechanisms are always better off returning a facility on the agents' line when such a line exists.

Overall, our results demonstrate that enlarging the outcome space can strictly improve truthful approximation guarantees, while also revealing inherent geometric and incentive-theoretic limits. Understanding how these ideas extend to randomized mechanisms, learning-augmented settings, and higher dimensions remains an interesting direction for future work.

Appendix A Deterministic Lower Bound Without Scale-Invariance

Theorem A.1. *No deterministic strategyproof mechanism can guarantee a strictly smaller than $\sqrt{2}$ approximation for collinear instances in $\mathbb{R}^2$.*

We focus on the two-point instance $S_D = \{(0, 0), (D, 0)\}$; a lower bound on this instance implies the same bound for all collinear instances.

A regularity assumption: convex range and continuity. Our proof of Lemma A.3 invokes continuity of the mechanism when taking limits of the form $D' \rightarrow D^+$. Continuity does not follow from strategyproofness alone in general domains; however, for Euclidean (single-peaked, distance-based) preferences, a classical theorem of Peters et al. [1993] shows that *strategyproofness together with convexity of the range of the mechanism is equivalent to continuity*. Formally, in the setting of voting schemes on $\mathbb{R}^m$ with Euclidean preferences, they prove:

Lemma A.2 ([Peters et al., 1993]). *A deterministic social choice function $f : (\mathbb{R}^m)^n \rightarrow \mathbb{R}^m$ is strategyproof and continuous if and only if it is strategyproof and its range $\text{Range}(f)$ is convex.*

Convexity of the range is a mild and natural regularity assumption in facility location: many canonical strategyproof mechanisms (dictatorships, projection mechanisms, coordinatewise medians, and the APEX MECHANISM) all have convex ranges. For example, the APEX MECHANISM has range $\{(x, y) : y \geq 0\}$, a half-plane, which is convex. Hence the continuity assumption we rely on below is satisfied for all mechanisms we study, and is standard in the Euclidean facility-location literature.

We conjecture that Lemma A.3 holds for *all* strategyproof mechanisms even without assuming convex range, but establishing this in full generality remains open. In this work we adopt the standard convex-range assumption, which by Lemma A.2 ensures continuity of f at all profiles.

Step 1. A geometric constraint implied by SP.

Lemma A.3. *Let $f(S_D) = (x, y)$ be the output of a deterministic, anonymous and unanimous SP mechanism on $S_D = \{(0, 0), (D, 0)\}$. Then*

$$x^2 + y^2 \geq Dx \quad \text{and} \quad (x - D)^2 + y^2 \geq D(D - x).$$

Proof. We prove the first inequality; the second is symmetric. Let agent 2 be at $(D, 0)$ and suppose agent 2 misreports $(D', 0)$, with the mechanism output $f(S_{D'}) = (x', y')$. Strategyproofness gives $\|(x', y') - (D, 0)\| \geq \|(x, y) - (D, 0)\|$ for all D' . Taking $D' \rightarrow D^+$ and assuming continuity, a rightward deviation cannot strictly decrease the distance. Geometrically, at (x, y) the vector $(x - D, y)$ cannot form an acute angle with the direction of deviation; equivalently, (x, y) must be the orthogonal projection of $(D, 0)$ onto the line through the origin and (x, y) . Writing this projection condition yields

$$x = \frac{Dx^2}{x^2 + y^2}, \quad y = \frac{Dxy}{x^2 + y^2},$$

which (in the nondegenerate case) is $x^2 + y^2 = Dx$. If $x^2 + y^2 < Dx$, agent 2 could improve by a slight rightward deviation, contradicting SP; hence $x^2 + y^2 \geq Dx$. $\square$

Step 2. From the SP circles directly to the $\sqrt{2}$ lens. Let $r_1 = \sqrt{x^2 + y^2}$ and $r_2 = \sqrt{(x - D)^2 + y^2}$ be the distances to the two agents. The optimal max-cost on S_D is $D/2$, attained at the midpoint $(D/2, 0)$. Define the $\sqrt{2}$ -lens

$$L = \left\{ (x, y) : r_1 \leq D/\sqrt{2} \text{ and } r_2 \leq D/\sqrt{2} \right\} = \left\{ \begin{array}{l} x^2 + y^2 \leq D^2/2, \\ (x - D)^2 + y^2 \leq D^2/2 \end{array} \right\}.$$

Lemma A.4. *For any SP outcome (x, y) on S_D , either $\max\{r_1, r_2\} \geq D/\sqrt{2}$, or else $(x, y) = (D/2, \pm D/2)$.*

Proof. Suppose $\max\{r_1, r_2\} < D/\sqrt{2}$. Then $x^2 + y^2 < D^2/2$ and $(x - D)^2 + y^2 < D^2/2$. Adding gives

$$2x^2 - 2Dx + 2y^2 + D^2 < D^2 \implies x^2 + y^2 < Dx,$$

which contradicts Lemma A.3. Hence $\max\{r_1, r_2\} \geq D/\sqrt{2}$ unless both lens inequalities hold at equality and Lemma A.3 is tight, i.e.,

$$x^2 + y^2 = D^2/2 = Dx \quad \text{and} \quad (x - D)^2 + y^2 = D^2/2 = D(D - x).$$

Solving gives $(x, y) = (D/2, \pm D/2)$. □

Step 3. Conclude the lower bound. By Lemma A.4, every SP outcome satisfies $\max\{r_1, r_2\} \geq D/\sqrt{2}$, with equality only at $(D/2, \pm D/2)$. Since the optimal max-cost is $D/2$, any deterministic SP mechanism has approximation ratio

$$\frac{\max\{r_1, r_2\}}{D/2} \geq \frac{D/\sqrt{2}}{D/2} = \sqrt{2}.$$

The APEX MECHANISM outputs $(D/2, \pm D/2)$ on S_D and thus attains ratio $\sqrt{2}$, completing the proof of Theorem A.1. □

References

- Priyank Agrawal, Eric Balkanski, Vasilis Gkatzelis, Tingting Ou, and Xizhi Tan. Learning-augmented mechanism design: Leveraging predictions for facility location. In *arXiv preprint arXiv:2204.01120*, 2022.
- Eric Balkanski, Vasilis Gkatzelis, and Golnoosh Shahkarami. Randomized strategic facility location with predictions. In *arXiv preprint arXiv:2409.07142*, 2024.
- Hau Chan, Jianan Lin, and Chenhao Wang. Strategyproof mechanisms for facility location with prediction under the maximum cost objective. In *arXiv preprint arXiv:2509.00439*, 2025.
- Xu Chenyang and Lu Pinyan. Mechanism design with predictions. In *Proceedings of the Thirty-First International Joint Conference on Artificial Intelligence (IJCAI-22)*, 2022.
- Mark de Berg, Otfried Cheong, Marc van Kreveld, and Mark Overmars. Computational geometry: Algorithms and applications. In *Springer*, 2008. ISBN 978-3-540-77974-2.
- Sumit Goel and Wade Hann-Caruthers. Coordinate-wise median: Not bad, not bad, pretty good. In *arXiv preprint arXiv:2509.00439*, 2021.
- Thodoris Lykouris and Sergei Vassilvitskii. Competitive caching with machine learned advice. In *Thodoris Lykouris and Sergei Vassilvitskii. 2021. Competitive Caching with Machine Learned Advice. J. ACM 68, 4, Article 24 (July 2021), 25 pages*, 2021.
- Nimrod Megiddo. The weighted euclidean 1-center problem. In *MATHEMATICS OF OPERATIONS RESEARCH Vol. 8, No. 4, November 1983*, 1983.
- Hans Peters, Hans Van der Stel, and Ton Storcken. Range convexity, continuity, and strategy-proofness of voting schemes. In *ZOR - Methods and Models of Operations Research (1993) 38:213-229*, 1993.
- A. D. Procaccia and M. Tennenholtz. Approximate mechanism design without money. In *ACM Trans. Econ. Comp. 1, 4, Article 18 (December 2013), 26 pages*, 2013.
- Pingzhong Tang, Dingli Yu, and Shengyu Zhao. Characterization of group-strategyproof mechanisms for facility location in strictly convex space. In *Proceedings of the 21st ACM conference on Economics and Computation (EC '20)*, 2020.