Mechanism Design for Connecting Regions Under Disruptions

Hau Chan[1], Jianan Lin[2], Zining Qin[3], Chenhao Wang[4,3]*

[1]University of Nebraska-Lincoln
[2]Rensselaer Polytechnic Institute
[3]Guangdong Provincial Key Laboratory IRADS, BNU-HKBU United International College
[4]Beijing Normal University-Zhuhai

Abstract

Man-made and natural disruptions such as planned constructions on roads, suspensions of bridges, and blocked roads by trees/mudslides/floods can often create obstacles that separate two connected regions. As a result, the traveling and reachability of agents from their respective regions to other regions can be affected. To minimize the impact of the obstacles and maintain agent accessibility, we initiate the problem of constructing a new pathway (e.g., a detour or new bridge) connecting the regions disconnected by obstacles from the mechanism design perspective. In the problem, each agent in their region has a private location and is required to access the other region. The cost of an agent is the distance from their location to the other region via the pathway. Our goal is to design strategyproof mechanisms that elicit truthful locations from the agents and approximately optimize the social or maximum cost of agents by determining locations in the regions for building a pathway. We provide a characterization of all strategyproof and anonymous mechanisms. For the social and maximum costs, we provide upper and lower bounds on the approximation ratios of strategyproof mechanisms.

1 Introduction

In modern societies, various types of infrastructures are constructed to connect regions to facilitate the traveling or reachability of agents from their corresponding regions to other regions (Amekudzi, Thomas-Mobley, and Ross 2007; Forkenbrock and Foster 1990; Narayanaswami 2017). These types of infrastructures include highways, streets, roads, bridges, and transportation systems. For instance, using the road infrastructure, an agent from a region can drive to reach another region effectively.

Unfortunately, these infrastructures can sometimes be interrupted either temporarily or permanently due to man-made or natural disruptions (Boakye et al. 2022; Faturechi and Miller-Hooks 2015; Gu et al. 2020; Serdar, Koç, and Al-Ghamdi 2022). For example, man-made disruptions can refer to the planned large construction project of a road, the construction of a transportation hub (e.g., a subway station), the suspension of bridges (e.g., due to accidents), or the interruption of an area due to public activities (e.g., parades, temporary markets, or sporting events). In addition, natural disruptions can be in the form of the aftermath of disasters (e.g., earthquakes and storms), where roads and bridges are damaged or blocked by large trees, mudslides, or floods.

These disruptions can often result in obstacles that disconnect any two regions and affect the traveling and reachability of agents. Therefore, our goal is to determine the best way to construct new routes/pathways connecting the disconnected regions in order to minimize the impact of the obstacles and maintain the accessibility of the agents. In temporary disruptions with obstacles (e.g., road constructions, public events, or large trees on roads), the new pathways can be viewed as detours connecting the regions so that agents can continue to access other regions before the removal of the obstacles. In permanent disruptions with obstacles (e.g., the suspensions of bridges or roads), the new pathways can be regarded as part of newly added roads or bridges connecting the regions. With the new pathways, the agents from their corresponding regions can still travel and reach the other regions, overcoming the obstacles due to disruptions.

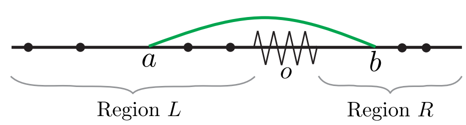

Figure 1: An obstacle o disconnects the agents (denoted as solid points) in two regions, and a new pathway (a, b) connects them.

In Figure 1, we provide an example of the suspension of a structurally deficient bridge (modeled simply as a line with an obstacle on it). The bridge originally connected Region L and Region R. With the bridge suspension, the bridge now becomes an obstacle (denoted by o on the line), and the two regions are now disconnected. Each agent located in their respective region needs to access (e.g., for work, school, or other daily routine) the other region divided by the obstacle. As agents cannot cross the obstacle directly, our goal is to maintain the accessibility of the agents to the regions (e.g., from Region L to Region R) by building a pathway or new bridge (denoted as a green line with endpoints a and b in Figure 1) connecting Region L and Region R.

Existing optimization literature has considered building

*Corresponding author. chenhwang@bnu.edu.cn

optimal pathways between two disconnected regions, aiming to minimize the maximum distance between any two points from the two regions (see, e.g., (Cai, Xu, and Zhu 1999; Kim, Shin, and Chwa 1998; Kim and Shin 2001; Tan 2000, 2002)). While these studies designed polynomial algorithms for building optimal bridges between different types of convex polygons (more details in related work), there are two main assumptions that make the current optimization literature not ideal for capturing real-world situations under disruptions involving agents. First, existing literature assumes that the agents are located in all of the points in the regions. However, in many real-world situations, agents' locations consist only of a subset of discrete points in the regions. Second, existing literature assumes that each agent's location is public information. However, agent locations might not be known in advance and require elicitation (Nisan et al. 2007; Procaccia and Tennenholtz 2013). Therefore, our goal is to build optimal pathways to account for agents' locations to connect them to the respective regions.

Our Contribution

We initiate the mechanism design study of building (approximately) optimal pathways between two regions disconnected by obstacles under disruptions to connect agents from their respective regions to other regions. We focus on a basic setting where a line segment (denoted by an interval $[0, 1]$) connecting two regions is separated by an obstacle o (see Figure 1).[1] Agents in the regions are denoted by sets N_1 and N_2, depending on whether their locations are points on the left-hand side or the right-hand side of the obstacle (i.e., $x_i \in [0, o)$ or $x_i \in (o, 1]$ for any agent i in N_1 or N_2).

We aim to design mechanisms to elicit agent locations and build a pathway/edge (a, b) that connects the two disconnected regions, where the left endpoint a is in $[0, o)$ and the right endpoint b is in $(o, 1]$. Given an edge (a, b), the cost of an agent at $x_i \in [0, o)$ is $|x_i - a| + k(b - a) + 1 - b$ with k being a non-negative multiplication factor, that is, the distance from their location to the (farthest) endpoint on the other region, passing through edge (a, b). The cost of an agent at $x_i \in (o, 1]$ is defined similarly as $|x_i - b| + k(b - a) + a$. We consider adding edges that minimize two different objectives: the social cost (i.e., the total cost of all agents) and the maximum cost (i.e., the maximum cost among all agent costs). When $k \geq 1$, a mechanism that returns an optimal solution $(o - \epsilon, o + \epsilon)$ is group strategyproof for $\epsilon \to 0$. Therefore, we only need to focus on the situation when $k \in [0, 1)$.[2]

We first provide a characterization of all strategyproof and anonymous mechanisms as two-dimensional generalized median mechanisms by showing that the agent preferences over the locations on where to build the pathway are two-dimensional single-peaked. The single-peakedness means that agents have preferences over a set of options (i.e., pathway locations in our setting) that can be ordered, so that each agent has the most preferred option (called the peak) and their preference for other options decreases as they move away from this peak. See more details in Section 3.

For the social cost, we derive an optimal solution on where to build a pathway and show that the mechanism that returns the optimal solution is group strategyproof. For the maximum cost, we show that there is a unique optimal solution on where to build a pathway and the optimal solution is not strategyproof. We show that a deterministic group strategyproof mechanism, TWOEXTREME, that simply connects two agent locations nearest to the obstacle has an approximation ratio of $\frac{2}{1+k}$. We provide an improved mechanism, TWOEXTREMERESTRICT, by not allowing the endpoints of the edge to be too close to the obstacle (see Theorem 5). On the other hand, we show that no deterministic strategyproof mechanism has an approximation ratio less than $\frac{2}{1+\sqrt{k}}$. Moreover, we design a randomized group strategyproof mechanism RANDMAXCOST that has an approximation ratio of $\max\left(\frac{4-2k}{3-k}, \frac{1+k}{1+k^2}\right)$. We also provide a lower bound $\frac{6+6k}{5+7k}$ for any randomized strategyproof mechanisms. See Figure 2 for an illustration of the above bounds.

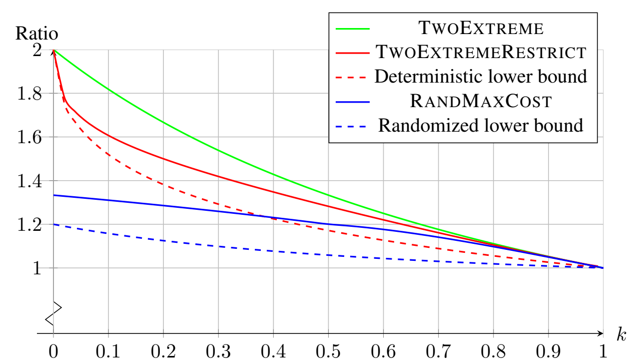

Figure 2: An illustration of upper and lower bounds for the maximum cost when $k \in [0, 1)$. The upper bounds of our mechanisms are depicted in solid lines, and the lower bounds are in dashed lines.

All of our results apply to the setting where the obstacle is a closed subinterval of $[0, 1]$ because such a subinterval can shrink to a point as in our setting.

Organization. We present the model in Section 2. We provide a characterization of all strategyproof mechanisms in Section 3. We study the social cost in Section 4 and the maximum cost in Section 5. All omitted proofs are in Appendix.

Related Works

While no existing mechanism design literature considers our setting, we discuss the most related optimization studies on building optimal bridges connecting two regions and adding edges to discrete networks to improve network parameters. We also discuss the related works in the approximate mechanism design without money paradigm.

[1]The line space has been extensively studied in mechanism design of facility location problems for modeling geographic regions and other real-world non-geographic situations (Chan et al. 2021; Procaccia and Tennenholtz 2013).

[2]In various situations, the social planner can determine the value of k appropriately. For instance, $k >> 1$ can be set for constructing a temporary detour. When creating a new road or bridge (to replace an older one), the social planner can set $k < 1$ by making it wider or having higher speed limits.

Bridge-building. Existing optimization literature has considered the problem of building an optimal bridge to connect two disconnected regions. Cai et al. (1999) introduced the problem of adding a line segment to connect two disjoint convex polygonal regions in a plane, such that the length of the longest path from a point in one polygon, passing through the bridge, to a point in another region is minimized. They proposed an $O(n^2 \log n)$-time algorithm, where n is the maximum number of extreme points of the polygons. Later, (Bhattacharya and Benkoczi 2001) proposed a linear-time algorithm that improves the $O(n^2 \log n)$-time algorithm in (Cai, Xu, and Zhu 1999). Tan (2000) independently presented an alternative linear-time algorithm for the above setting and further generalized it to an $O(n^2)$-time algorithm for bridging two convex polyhedra in space. (Kim and Shin 2001) provided algorithms to find an optimal bridge between two convex polygons, two simple non-convex polygons, and one convex and one simple non-convex polygons in $O(n)$, $O(n^2)$, and $O(n \log n)$, respectively. Later, Tan (2002) provided an $O(n \log^3 n)$-time algorithm for the settings of two simple non-convex polygons. Kim et al. (1998) proposed a linear-time algorithm to compute an optimal bridge between two parallel lines separated by an obstacle to minimize the length of the longest path connecting two points on the lines. However, all of the above-mentioned works focus on all points in the regions. Our work focuses on a finite subset of points, which are the agents' locations, and the mechanism perspective in which agents' locations are private.

Edge addition on networks. Existing optimization studies have examined adding edges to discrete networks (with nodes and edges) to minimize the diameter or average shortest distances between pairs of nodes of a network (see, e.g., (Demaine and Zadimoghaddam 2010; Meyerson and Tagiku 2009; Papagelis, Bonchi, and Gionis 2011; Perumal, Basu, and Guan 2013). However, all these optimization studies on discrete networks do not consider disconnected regions that are continuous and assume agents occupy all nodes/vertices of the network. Moreover, they do not consider the mechanism design perspective.

Mechanism design. Our considered mechanism design setting is within the paradigm of approximate mechanism design without money, initialized by Procaccia and Tennenholtz (Procaccia and Tennenholtz 2013) who used facility location problems (FLPs) as case studies. This paradigm investigates the design of approximately optimal strategyproof mechanisms through the lens of the approximation ratio. In a typical setting of FLPs, the agents report their private locations on the real line to a mechanism. The mechanism determines the locations for building facilities to minimize some objectives involving the costs of agents, where the cost of each agent is their distance to the facilities. Following their work, variations of FLPs have been introduced and studied (see, e.g., (Dokow et al. 2012; Feldman and Wilf 2013; Filos-Ratsikas and Voudouris 2021; Lin 2020; Mei et al. 2019; Meir 2019)). We note that the case $k = 0$ of our setting is equivalent to a 2-FLP problem where each agent i has two locations, x_i and the endpoint of the other region (0 or 1), whose cost is the total distance from their two locations to the two facilities (which are now represented as a pathway with $k = 0$). We refer readers to a survey on mechanism design for FLPs (Chan et al. 2021). The most relevant mechanism design work to ours is the work of (Chan and Wang 2023) in which they considered modifying the structure of regions by adding a shuttle or road to improve the distances of the agents to a prelocated facility. In contrast, they do not consider two regions separated by an obstacle.

2 Model

Let $N = \{1, \ldots, n\}$ be the set of agents located in an interval $[0, 1]$. The location profile of agents is denoted as $\mathbf{x} = (x_1, \ldots, x_n)$. There is an obstacle located at point $o \in (0, 1)$. Provided that no agent is at o, this obstacle partitions the agents into $N = (N_1, N_2)$ according to their regions, where $N_1 = \{i \in N \mid x_i < o\}$ is the set of agents on the left region, and $N_2 = \{i \in N \mid x_i > o\}$ is the set of agents on the right region. The agents on one region are required to access the other region. Due to the obstacle, the agents cannot pass through it and reach the other region directly. Hence, we want to build a new edge (a, b) that connects the two regions with $a \in [0, o)$ and $b \in (o, 1]$. The length of the edge is $k(b-a)$, where k is a positive constant.

A deterministic mechanism $f : \mathbb{R}^n \to \mathbb{R}^2$ is a function that takes the agent location profile $\mathbf{x}$ as input and returns an edge $f(\mathbf{x})$. Given an edge $f(\mathbf{x}) = (a, b)$, the *cost* of each agent $i \in N_1$ on the left region is the distance to the right endpoint 1 through the edge,

$$cost(a, b, x_i) = |x_i - a| + k(b - a) + (1 - b).$$

Similarly, the *cost* of each agent $i \in N_2$ on the right region is the distance to the left endpoint 0 through the edge,

$$cost(a, b, x_i) = |x_i - b| + k(b - a) + a.$$

A randomized mechanism is a function f from $\mathbb{R}^n$ to probability distributions over $\mathbb{R}^2$. If $f(\mathbf{x}) = P$ is a probability distribution, the cost of agent $i \in N$ is defined as the expected cost $cost(P, x_i) = \mathbb{E}_{(a,b)\sim P}[cost(a, b, x_i)]$.

A mechanism f is *strategyproof* if no agent can decrease their cost by misreporting the location within their region. Formally, f is strategyproof if for any $i \in N, \mathbf{x}$ and x'_i with x_i, x'_i on the same region, $cost(f(x_i, \mathbf{x}_{-i}), x_i) \leq cost(f(x'_i, \mathbf{x}_{-i}), x_i)$, where $\mathbf{x}_{-i}$ is the location profile of the agents in $N \setminus \{i\}$. Further, f is called *group strategyproof* if no group of agents can misreport simultaneously so that all agents in the group are better off. That is, for any $S \subseteq N, \mathbf{x}, \mathbf{x}'_S$, there exists an agent $i \in S$ such that $cost(f(\mathbf{x}), x_i) \leq cost(f(\mathbf{x}'_S, \mathbf{x}_{-S}), x_i)$. A mechanism f is *anonymous* if the outcomes are invariant under permutation of agents, i.e., $f(x_1, \ldots, x_n) = f(x_{\pi(1)}, \ldots, x_{\pi(n)})$ for every profile $\mathbf{x}$ and every permutation of agents $\pi : N \to N$. Since a non-anonymous mechanism is based on the identity of the agents and is much less interesting, we focus only on anonymous mechanisms.

Our goal is to design (group) strategyproof mechanisms with good performance guarantees under two objectives: minimizing the social cost and minimizing the maximum cost. The social cost with respect to an edge (a, b) and location profile $\mathbf{x}$ is $SC(a, b, \mathbf{x}) = \sum_{i \in N} cost(a, b, x_i)$. The maximum cost with respect to an edge (a, b) and location profile $\mathbf{x}$ is $MC(a, b, \mathbf{x}) = \max_{i \in N} cost(a, b, x_i)$. A mechanism f is α-approximation ($\alpha \geq 1$) for the objective function $\Delta \in \{SC, MC\}$ if $\Delta(f(\mathbf{x}), \mathbf{x}) \leq \alpha \cdot \min_{(a,b) \in \mathbb{R}^2} \Delta(a, b, \mathbf{x})$ for all location profiles $\mathbf{x} \in \mathbb{R}^n$.

We remark that when the constant coefficient is $k \geq 1$, there is a trivial solution $(o - \epsilon, o + \epsilon)$ for some fixed value $\epsilon > 0$. As $k \geq 1$, every agent wants the edge to be as short as possible, that is, ϵ approaches 0. Then, a mechanism that returns the fixed solution $(o - \epsilon, o + \epsilon)$ is clearly group strategyproof and (almost) optimal for both objectives when $\epsilon \to 0$. Therefore, in the remainder of this paper, we assume that $k \in [0, 1)$. Because it is natural for each agent to only misreport locations within their own region, our model assumes that agents cannot misreport their locations in other regions. However, it is worth noting that all of our mechanisms (Mechanism 1-4) are still strategyproof and retain the same approximation even without this assumption.

3 Characterizing Strategyproof Mechanisms

In this section, we show that the preference profile of agents is multi-dimensional single-peaked, and the generalized median mechanisms compose the class of all anonymous strategyproof mechanisms.

We start with some necessary definitions. Let D be a set of possible outcomes. A one-dimensional axis A on D is any strict ordering $<_A$ of the outcomes in D. A multi-dimensional axis $A^m = \langle A_1, \ldots, A_m \rangle$ on D is a collection of m distinct axes, each being a one-dimensional axis on D.

Definition 1 ((Barberà, Gul, and Stacchetti 1993)). Let A^m be an m-dimensional axis on the set D of possible outcomes. An agent i's preference $\succeq_i$ is *m-dimensional single-peaked with respect to A^m* if: (1) there is a single most-preferred outcome (peak) $p_i \in D$, and (2) for any two outcomes $\alpha, \beta \in D$, $\alpha \succeq_i \beta$ whenever $\beta <_{A_t} \alpha <_{A_t} p_i$ or $p_i <_{A_t} \alpha <_{A_t} \beta$ for all axes $A_t, t = 1, \ldots, m$.

Then, a preference profile is called *m-dimensional single-peaked* if there exists an m-dimensional axis A^m such that every agent preference is m-dimensional single-peaked with respect to A^m. While the preference profile in our problem is not one-dimensional single-peaked, it is indeed two-dimensional single-peaked.

Theorem 1. *For any instance of our problem, the preference profile of agents is 2-dimensional single-peaked.*

Proof. In our problem, an outcome is a shortcut edge (a, b) with $a \in [0, o)$ and $b \in (o, 1]$, and it uniquely corresponds to a point (a, b) in the 2-dimensional xy-coordinate system. Thus, the set of all possible outcomes can be represented by a set $D = \{(x, y) \in \mathbb{R}^2 \mid 0 \leq x < o < y \leq 1\}$ in a plane, as shown in Figure 3.

Now we show that every agent is 2-dimensional single-peaked with respect to the collection of x-axis and y-axis. For any agent $i \in N_1$, the single most-preferred outcome (peak) is $p_i = (x_i, 1) \in D$. For any two outcomes (points)

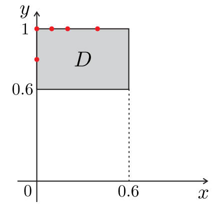

Figure 3: An illustration of the set D when the obstacle is at $o = 0.6$. For location profile $\mathbf{x} = (0.1, 0.2, 0.4, 0.8, 1)$, the agent peaks are $(0.1, 1), (0.2, 1), (0.4, 1), (0, 0.8), (0, 1)$, denoted as red points.

$(x, y), (x', y') \in D$, if they satisfy (1) $x' < x < x_i$ or $x_i < x < x'$, and (2) $y' < y < 1$, then we have

$$c(x, y, x_i) = |x_i - x| + k(y - x) + 1 - y$$
$$< |x_i - x'| + k(y' - x') + 1 - y' = c(x', y', x_i),$$

implying that agent i has a smaller cost under (x, y) than that under (x', y') and the agent prefers (x, y). On the other hand, for any agent $j \in N_2$, the single most-preferred outcome is $p_j = (0, x_j) \in D$. For any two outcomes $(x, y), (x', y') \in D$, if they satisfy (1) $y' < y < x_j$ or $x_j < y < y'$, and (2) $0 < x < x'$, similarly, it is easy to see that agent j has a smaller cost under (x, y) than that under (x', y'), and thus the agent prefers (x, y). Hence, both conditions in Definition 1 are satisfied for every agent, and the preference profile is 2-dimensional single-peaked. □

For (one-dimensional) single-peaked preferences in a one-dimensional space, a mechanism is a *generalized median mechanism* if there exists $n + 1$ constants $b_1, \ldots, b_{n+1} \in \mathbb{R} \cup \{-\infty, +\infty\}$ such that the outcome is $\text{med}(p_1, \ldots, p_n, b_1, \ldots, b_{n+1})$, where med is the median function, and $p_1, \ldots, p_n$ is the most-preferred outcome of the n agents. For m-dimensional single-peaked preferences in an m-dimensional space, an *m-dimensional generalized median mechanism* can be decomposed into m independent one-dimensional generalized median mechanisms, with the t-th mechanism determining the coordinate of the outcome on the t-th dimension, for all $t = 1, \ldots, m$ (see, e.g., (Sui 2015)). Barberà *et al.* (1993) provide a characterization result: a mechanism for multi-dimensional single-peaked preferences in a multi-dimensional space is strategyproof and anonymous if and only if it is a multi-dimensional generalized median mechanism. This characterization applies to our problem by Theorem 1.

Corollary 1. *A mechanism for our problem is strategyproof and anonymous if and only if it is a 2-dimensional generalized median mechanism.*

4 Social Cost

In this section, we consider the social cost. We start with some necessary notations. For each point $x \in [0, o)$, define $N_{1l}(x) = \{i \in N_1 | x_i \leq x\}$ to be the set of agents in N_1 on the left of x, and $N_{1r}(x) = \{i \in N_1 | x_i > x\}$ to be the set of agents in N_1 on the right of x. For each point $x \in (o, 1]$, define $N_{2l}(x) = \{i \in N_2 | x_i < x\}$ to be the set of agents in

N_2 on the left of x, and $N_{2r}(x) = \{i \in N_2 | x_i \geq x\}$ to be the set of agents in N_2 on the right of x.

We consider a mechanism that determines the location of the two endpoints separately. First, the mechanism fixes b and moves a from 0 towards the obstacle as long as the social cost is decreasing. Second, it moves b from 1 towards the obstacle as long as the social cost is decreasing.

Mechanism 1 (OPTSOCCOST)**.** Given location profile $\mathbf{x}$, let X_L be the set of all points $x \in [0, o)$ that satisfy

$$|N_{1l}(x) \cup N_2| \cdot (1-k) < |N_{1r}(x)| \cdot (1+k),$$

and let $a^* = \sup X_L$ if X_L is non-empty and $a^* = 0$ otherwise. Let X_R be the set of all points $x \in (0, 1]$ that satisfy

$$|N_{2r}(x) \cup N_1| \cdot (1-k) < |N_{2l}(x)| \cdot (1+k),$$

and let $b^* = \inf X_R$ if X_R is non-empty and $b^* = 0$ otherwise. Return (a^*, b^*).

Note that for every $x \in [0, o)$ and fixed b, if we move the left endpoint a from x to $x + \epsilon$ for some sufficiently small value $\epsilon > 0$, the cost of the agents in $N_{1l}(x) \cup N_2$ increases by $(1-k)\epsilon$, and the cost of the agents in $N_{1r}(x)$ decreases by $(1+k)\epsilon$. Thus, if we move a from 0 towards the obstacle as long as the social cost is decreasing, it approaches the supremum of X_L. Symmetrically, we move b towards the obstacle and it approaches the infimum of X_R. Since the supremum of X_L and the infimum of X_R can only be the agent locations, the mechanism is in polynomial time.

When $k = 0$, OPTSOCCOST is exactly the median mechanism that selects the x-coordinate (resp. y-coordinate) of the outcome to be the median of the x-coordinates (resp. y-coordinates) of n agent peaks.

Theorem 2. *Mechanism 1 is group strategyproof and optimal for the social cost.*

Proof. For the group strategyproofness, we consider a group of agents $S \subseteq N_1 \cup N_2$. Let $f(\mathbf{x}) = (a, b)$ be the outcome when all agents report true locations, and $f(\mathbf{x}'_S, \mathbf{x}_{-S}) = (a', b')$ be the outcome when the agents in S misreport $\mathbf{x}'_S$. Assume w.l.o.g. that $|a - a'| \geq |b - b'|$. We show that at least one agent in the group cannot gain by misreporting.

If $a' < a$, the agents in $N_{1r}(a)$ cannot gain and they are not in the group. Since the agents in N_2 cannot change the location of a, by the definition of the mechanism, an agent located at a must be in the group and misreport a location to the left of a. The cost of this agent decreases by at most $(1-k)|b' - b| - (1+k)(a - a') \leq 0$, indicating that this agent can never gain.

If $a' > a$, by the mechanism, there exists at least one agent in $N_{1l}(a)$ and in group S who misreport the location to the right of a. However, the cost of such an agent in $N_{1l}(a) \cap S$ will decrease by at most $(1-k)|b' - b| - (1-k)(a' - a) \leq 0$, and the agent cannot gain by misreporting.

For the optimality, let (a^*, b^*) be the solution returned by the mechanism, and let (a, b) be an optimal solution. Assume w.l.o.g. that $|a - a^*| \geq |b - b^*|$. If $a < a^*$, we consider a new solution $(a + \epsilon, b)$, where $\epsilon > 0$ is a sufficiently small number so that there is no agent in the interval $(a, a + \epsilon)$. Compared with solution (a, b), each agent in $N_{1l}(a) \cup N_2$ increases the cost by $(1-k)\epsilon$, and each agent in $N_{1r}(a)$ decreases the cost by $(1+k)\epsilon$. By the definition of a^* and the fact that $a < a^*$, we have $|N_{1l}(a) \cup N_2| \cdot (1-k) < |N_{1r}(a)| \cdot (1+k)$, indicating that the social cost of (a, b) is larger than that of $(a+\epsilon, b)$, which contradicts the optimality.

If $a > a^*$, we change the solution (a, b) to (a^*, b). Each agent in $N_{1l}(a^*) \cup N_2$ decreases the cost by $(1-k)(a-a^*)$, and the increase of cost for every agent in $N_{1r}(a^*)$ is at most $(1+k)(a - a^*)$. By the definition of $a^* = \sup X_L$ and the fact that X_L is an open set, we have $|N_{1l}(a^*) \cup N_2| \cdot (1-k) \geq |N_{1r}(a^*)| \cdot (1+k)$. It indicates that the social cost of (a^*, b) is no more than that of (a, b), and (a^*, b) is also optimal.

Then we change the solution (a^*, b) to (a^*, b^*). By a symmetric analysis, we know that $b > b^*$ is impossible, and if $b \leq b^*$, the social cost of (a^*, b^*) is no more than that of (a^*, b). Hence, (a^*, b^*) is optimal. □

5 Maximum Cost

In this section, we consider the maximum cost. We first characterize the unique optimal solution and show that the mechanism that returns the optimal solution is not strategyproof. We then study deterministic and randomized mechanisms.

The Optimal Maximum Cost

Given location profile $\mathbf{x}$, let $x_l = \min\{x_i | i \in N_1\}$ and $x_r = \max\{x_i | i \in N_1\}$ be the two extreme agent locations to the left, and let $y_l = \min\{x_j | j \in N_2\}$ and $y_r = \max\{x_j | j \in N_2\}$ be the two extreme agent locations to the right.

Lemma 1. *In any optimal solution (a, b) for maximum cost, we have $a \leq \frac{x_l + x_r}{2}$ and $b \geq \frac{y_l + y_r}{2}$.*

Proof. Suppose that $a > \frac{x_l + x_r}{2}$, which indicates that $cost(a, b, x_r) < cost(a, b, x_l)$. Clearly, the maximum cost is attained by x_l or y_l or y_r. We consider a solution $(a - \epsilon, b)$ that moves the left endpoint to a sufficiently small positive value $\epsilon < a - \frac{x_l + x_r}{2}$. Then the cost of the agents at x_l, y_l, y_r decreases, and thus the maximum cost decreases, which contradicts the optimality. Therefore, it must be $a \leq \frac{x_l + x_r}{2}$. By a symmetric analysis, we can prove $b \geq \frac{y_l + y_r}{2}$. □

Lemma 2. *In any optimal solution (a, b) for maximum cost, either $a = \frac{x_l + x_r}{2}$, or $b = \frac{y_l + y_r}{2}$.*

Proof. Suppose that $a \neq \frac{x_l + x_r}{2}$ and $b \neq \frac{y_l + y_r}{2}$. By Lemma 1, we have $a < \frac{x_l + x_r}{2}$ and $b > \frac{y_l + y_r}{2}$. It is easy to see that the maximum cost is attained by either x_r or y_l. We consider a solution $(a+\epsilon, b-\epsilon)$ that moves the two endpoints towards the obstacle by a sufficiently small value $\epsilon > 0$. The cost of both agents at x_r and y_l decreases by $(1+k)\epsilon - (1-k)\epsilon = 2k\epsilon$, and thus the maximum cost decreases, which contradicts the optimality. Therefore, it must be $a = \frac{x_l + x_r}{2}$, or $b = \frac{y_l + y_r}{2}$, or both hold. □

Now we describe the algorithm OPTMAXCOST to derive the optimal maximum cost. Given location profile $\mathbf{x}$, if $1 - y_r \geq x_l$, define $a^* = \frac{x_l + x_r}{2}$ and $b^* = \frac{y_l - x_l}{2} + \frac{1}{2}$. If $1 - y_r < x_l$, define $a^* = \frac{x_r - y_r}{2} + \frac{1}{2}$ and $b^* = \frac{y_l + y_r}{2}$. Return (a^*, b^*).

Notice that when $1 - y_r = x_l$, the mechanism simply returns the two midpoints $(\frac{x_l+x_r}{2}, \frac{y_l+y_r}{2})$.

Theorem 3. OPTMAXCOST *returns the unique optimal solution for maximum cost. In addition, the optimal maximum cost is attained by both the agents located at x_r and y_l.*

Proof. We focus only on the case when $1 - y_r \geq x_l$, as the other case is symmetric. The mechanism returns $(a^*, b^*) = (\frac{x_l+x_r}{2}, \frac{y_l-x_l}{2} + \frac{1}{2})$. The maximum cost under (a^*, b^*) is attained by the agents at x_l, x_r and y_l, that is,

$$\begin{aligned} cost(a^*, b^*, x_l) &= \frac{x_r - x_l}{2} + k(b^* - a^*) + \frac{1}{2} - \frac{y_l - x_l}{2} \\ &= \frac{1}{2} - \frac{y_l + x_l}{2} + k(b^* - a^*) + \frac{x_l + x_r}{2} \\ &= cost(a^*, b^*, y_l), \end{aligned}$$

which is no less than $cost(a^*, b^*, y_r)$ since $b^* \geq \frac{y_l+y_r}{2}$.

Suppose that (a, b) is an optimal solution. By Lemma 2, we have either $a = \frac{x_l+x_r}{2}$ or $b = \frac{y_l+y_r}{2}$. When $a = \frac{x_l+x_r}{2}$, if $b < b^*$, then the agents at x_l and x_r have a larger cost in (a, b) than that in (a^*, b^*), implying that (a, b) is not optimal for maximum cost. If $b > b^*$, then the agent at y_l has a larger cost in (a, b) than that in (a^*, b^*), and thus (a, b) is not optimal. Therefore, (a^*, b^*) is the unique optimal solution.

When $b = \frac{y_l+y_r}{2}$, we have $a \leq \frac{x_l+x_r}{2}$ by Lemma 1, and the maximum cost induced by (a, b) is at least

$$\begin{aligned} cost(a, b, x_r) &= x_r - a + k(\frac{y_l + y_r}{2} - a) + 1 - \frac{y_l + y_r}{2} \\ &\geq x_r - (1+k)\frac{x_l + x_r}{2} + 1 - (1-k)\frac{y_l + y_r}{2} \\ &\geq cost(a^*, b^*, y_l), \end{aligned}$$

where the equation $cost(a, b, x_r) = cost(a^*, b^*, y_l)$ holds only if $1 - y_r = x_l$ and $a = \frac{x_l+x_r}{2}$, that is, $(a, b) = (a^*, b^*)$. Therefore, (a^*, b^*) is the unique optimal solution. □

However, OPTMAXCOST is not strategyproof. Consider a location profile $\mathbf{x} = (0, 0.2, 0.8, 1)$, and the obstacle is at 0.5. The mechanism returns the two midpoints $(\frac{x_l+x_r}{2}, \frac{y_l+y_r}{2}) = (0.1, 0.9)$, and the cost of the agent at 0.2 is $0.2 + 0.8k$. If this agent misreports the location as 0.4, then the outcome of the mechanism becomes $(0.2, 0.9)$, and the cost of this agent with true location 0.2 decreases to $0.1 + 0.7k < 0.2 + 0.8k$.

Deterministic Mechanisms

We consider designing deterministic strategyproof mechanisms with good performance guarantees. We first present a simple mechanism that connects the two extreme agent locations x_r and y_l (following the idea that the optimal maximum cost is attained by both x_r and y_l). We then improve the mechanism by restricting the endpoints of the edge from being too close to the obstacle.

Mechanism 2 (TWOEXTREME)**.** Given location profile $\mathbf{x}$, return (x_r, y_l).

This mechanism falls in the class of generalized median mechanisms. Because x_r is the rightmost x-coordinate of the n peaks and y_l is the leftmost y-coordinate of the n peaks, setting $b_1 = -\infty$ and $b_2 = \cdots = b_{n+1} = +\infty$ gives $x_r = \text{med}(0, \ldots, 0, (x_i)_{i \in N_1}, b_1, \ldots, b_{n+1})$, and setting $b_1 = \cdots = b_n = -\infty$ and $b_{n+1} = +\infty$ gives $y_l = \text{med}(1, \ldots, 1, (x_j)_{j \in N_2}, b_1, \ldots, b_{n+1})$. Thus it is strategyproof by the characterization in Corollary 1.

Theorem 4. *Mechanism 2 is group strategyproof and $\frac{2}{1+k}$-approximation for maximum cost.*

Proof. For the group strategyproofness, we consider a group of agents $S \subseteq N_1 \cup N_2$. Let $f(\mathbf{x}) = (x_r, y_l)$ be the outcome when all agents report true locations, and $f(\mathbf{x}'_S, \mathbf{x}_{-S}) = (x'_r, y'_l)$ be the outcome when the agents in S misreport $\mathbf{x}'_S$. Assume w.l.o.g. that $|x_r - x'_r| \geq |y_l - y'_l|$. If $x'_r < x_r$, an agent located at x_r must be in the group and misreport a location to the left of x_r. The cost of this agent decreases by at most $(1-k)|y'_l - y_l| - (1+k)(x_r - x'_r) \leq 0$, indicating that this agent can never gain. If $x'_r > x_r$, the cost of any agent in N_1 decreases by at most $(1-k)|y'_l - y_l| - (1-k)(x'_r - x_r) \leq 0$, indicating that the agent in N_1 can never gain and they are not in the group S. However, by the definition of the mechanism, other agents in N_2 are not able to induce an outcome (x'_r, y'_l) with $x'_r \neq x_r$, giving a contradiction.

For the approximation, we consider an arbitrary instance with location profile $\mathbf{x}$. Assume w.l.o.g. that $1 - y_r \geq x_l$. By Theorem 3, the optimal solution is $(a^*, b^*) = (\frac{x_l+x_r}{2}, \frac{y_l-x_l}{2} + \frac{1}{2})$, and the optimal maximum cost is

$$\begin{aligned} &cost(a^*, b^*, x_r) = cost(a^*, b^*, y_l) = cost(a^*, b^*, x_l) \\ &= \frac{x_l + x_r}{2} - x_l + k(b^* - \frac{x_l + x_r}{2}) + 1 - b^* \\ &= \frac{1}{2} - \frac{y_l - x_r}{2} + k(\frac{1}{2} + \frac{y_l - 2x_l - x_r}{2}). \end{aligned}$$

Now we consider the solution (x_r, y_l) returned by the mechanism. Clearly, the maximum cost is attained by x_l or y_r. Since $1 - y_r \geq x_l$, the cost of the agent at x_l is

$$\begin{aligned} &cost(x_r, y_l, x_l) = x_r - x_l + k(y_l - x_r) + 1 - y_l \\ &\geq x_r + k(y_l - x_r) + y_r - y_l = cost(x_r, y_l, y_r), \end{aligned}$$

implying that the maximum cost is attained by x_l. Further,

$$\begin{aligned} &\frac{cost(x_r, y_l, x_l)}{cost(a^*, b^*, x_r)} = 2 \cdot \frac{x_r - x_l + k(y_l - x_r) + 1 - y_l}{1 - y_l + x_r + k(1 + y_l - 2x_l - x_r)} \\ &= 2 \cdot \frac{1 + (1-k)x_r - x_l - (1-k)y_l}{1 + (1-k)x_r + k(1 - 2x_l) - (1-k)y_l}. \end{aligned}$$

Note that the assumption $1 - y_r \geq x_l$ implies $x_l \leq \frac{1}{2}$, and thus $\frac{1+(1-k)x_r - x_l}{1+(1-k)x_r + k(1-2x_l)} \leq 1$. Hence, $\frac{cost(x_r, y_l, x_l)}{cost(a^*, b^*, x_r)}$ is decreasing with y_l, and by the fact $y_l \geq x_r$, it gives

$$\begin{aligned} \frac{cost(x_r, y_l, x_l)}{cost(a^*, b^*, x_r)} &\leq 2 \cdot \frac{1 + (1-k)x_r - x_l - (1-k)x_r}{1 + (1-k)x_r + k(1 - 2x_l) - (1-k)x_r} \\ &= 2 \cdot \frac{1 - x_l}{1 + k - 2kx_l} \leq \frac{2}{1+k}, \end{aligned}$$

where the last inequality follows from that $\frac{1-x_l}{1+k-2kx_l}$ is decreasing with x_l. Therefore, Mechanism 2 is $\frac{2}{1+k}$-approximation for maximum cost. □

Remark 1. We remark that it may be of interest to consider other two-extreme mechanisms that always return (x_l, y_r), (x_l, y_l) or (x_r, y_r). Although these mechanisms are also group strategyproof, their approximation ratio is 2 and it cannot be improved for any $k \in [0, 1)$.

The analysis for Mechanism 2 is tight for any $k \in [0, 1]$. Consider an instance with $x_l = 0$, $1 - y_r \geq x_l$, and $y_l = x_r + \epsilon$ for some sufficiently small positive number ϵ. The optimal solution is $(\frac{x_l+x_r}{2}, \frac{y_l-x_l+1}{2}) = (\frac{x_r}{2}, \frac{y_l+1}{2})$ by Theorem 3, and the optimal maximum cost is $\frac{x_r}{2} + k \cdot \frac{1+y_l-x_r}{2} + 1 - \frac{1+y_l}{2} = \frac{1+k}{2} - \frac{1-k}{2}\epsilon$. Mechanism 2 returns (x_r, y_l), and the induced maximum cost is $x_r - x_l + k(y_l - x_r) + 1 - y_l = 1 - (1-k)\epsilon$. The ratio is $\frac{1-(1-k)\epsilon}{(1+k)/2-(1-k)\epsilon/2} \to \frac{2}{1+k}$, when ϵ approaches 0. Hence, the worst case happens when x_r and y_l are very close to the obstacle o. To solve this case, we need the endpoints of the edge to keep some distance from the obstacle, which inspires the following mechanism.

Mechanism 3 (TWOEXTREMERESTRICT). Given location profile $\mathbf{x}$, return (a, b) with $a = \min(x_r, o - oc)$ and $b = \max(y_l, o + c - oc)$, where o is the obstacle location and c is a value in $[0, 1]$ set to be $\frac{1+k^2-\sqrt{k^4-k^3+3k^2+k}}{1-k^2}$.

Note that TWOEXTREME is a special case of Mechanism 3 when $c = 0$. Since the value of c above is optimally set, Mechanism 3 has an improved approximation ratio than the $\frac{2}{1+k}$-approximation of TWOEXTREME. Mechanism 3 is also in the class of generalized median mechanisms. Indeed, we can set $|N_1|$ phantoms at point $(o - oc, 0)$, $|N_2|$ phantoms at $(1, o+c-co)$, and one phantom at $(0, 1)$ in Euclidean plane, and then a (resp. b) is the x-coordinate (resp. y-coordinate) median of the $2n+1$ points consisting of the $n+1$ phantoms and the n peaks of agents.

Theorem 5. *Mechanism 3 is group strategyproof and the approximation ratio for maximum cost is at most two times the maximum of the four numbers*

$$\frac{1-(1-k)c}{1+k-(1-k)c}, \frac{k(2c-c^2)+1-c^2}{2-2c+2ck}, \frac{1+2ck}{2-(1-k)c} \text{ and } c.$$

Next, we complete the results by a lower bound.

Theorem 6. *No deterministic strategyproof mechanism has an approximation ratio less than* $\frac{2}{1+\sqrt{k}}$ *for the maximum cost, for any* $k \in [0, 1)$.

Proof. Consider the instance with location profile $\mathbf{x} = (0, 1-\epsilon, 1)$, where the obstacle is located between $1-\epsilon$ and 1, and $\epsilon > 0$ is a sufficiently small positive number. For convenience, we will ignore the terms with respect to ϵ in the following calculations. Let f be a strategyproof mechanism, and it returns $f(\mathbf{x}) = (a, 1)$ for some $a \in [0, 1)$.

If $a \geq \frac{\sqrt{k}}{1+\sqrt{k}}$, consider another instance with location profile $\mathbf{x}' = (0, a, 1)$. The optimal solution is $(\frac{a}{2}, 1)$ by Theorem 3, and the optimal maximum cost is $\frac{a}{2} + k(1 - \frac{a}{2})$. By the strategyproofness, f must return $f(\mathbf{x}') = (a, 1)$, as otherwise the agent located at a can decrease the cost by misreporting a location $1-\epsilon$ (resulting in a location profile $\mathbf{x}$ and an outcome $(a, 1)$). Then the maximum cost induced by the mechanism f is $cost(a, 1, 0) = a + k(1-a)$. Thus, the approximation ratio of f is at least

$$\frac{a + k(1-a)}{\frac{a}{2} + k(1-\frac{a}{2})} = 2 - \frac{k}{k + \frac{(1-k)a}{2}} \geq \frac{2}{1+\sqrt{k}}.$$

If $a < \frac{\sqrt{k}}{1+\sqrt{k}}$, consider another instance with location profile $\mathbf{x}'' = (a, 1-\epsilon, 1)$. The optimal solution is $(\frac{1}{2}, 1)$ by Theorem 3, and the optimal maximum cost is $\frac{1+k}{2}$. Again, by the strategyproofness, f must return $f(\mathbf{x}'') = (a, 1)$, as otherwise the agent located at a can decrease the cost by misreporting a location 0 (resulting in a location profile $\mathbf{x}$ and an outcome $(a, 1)$). Then the maximum cost induced by f is $1 - a + k(1-a)$. Thus, the ratio of f is at least

$$\frac{(k+1)(1-a)}{(k+1)/2} = 2(1-a) \geq 2(1 - \frac{\sqrt{k}}{1+\sqrt{k}}) = \frac{2}{1+\sqrt{k}}.$$

□

Although generally there is a gap between our upper bound and the lower bound $\frac{2}{1+\sqrt{k}}$, they are matching when $k = 0$ or $k \to 1$. The largest gap is about 0.127, which happens when $k \approx 0.311$.

Randomized Mechanisms

Next, we consider randomized mechanisms. Inspired by the worst-case instance of TWOEXTREME that returns (x_r, y_l), we have the following randomized mechanism.

Mechanism 4 (RANDMAXCOST). Given location profile $\mathbf{x}$, return (a, b) as (x_r, y_l) with probability p and $(\frac{x_r}{2}, \frac{y_l+1}{2})$ with probability $1-p$, where $p = \max\left(\frac{1+k}{3-k}, \frac{k+k^2}{1+k^2}\right)$.

Theorem 7. *Mechanism 4 is a randomized group strategyproof mechanism, and the approximation ratio for maximum cost is* $\max\left(\frac{4-2k}{3-k}, \frac{1+k}{1+k^2}\right)$.

Theorem 8. *No randomized strategyproof mechanism has an approximation ratio less than* $\frac{6+6k}{5+7k}$ *for the maximum cost, for any* $k \in [0, 1)$.

The largest gap between the upper and lower bounds is about 0.162, which happens when $k \approx 0.249$.

6 Conclusion

We studied a novel mechanism design setting for connecting two regions disconnected by obstacles under disruptions by adding a pathway to minimize the social cost and the maximum cost of the agents. We first characterize all of the strategyproof and anonymous mechanisms as 2-dimensional generalized median mechanisms. For the social cost and maximum cost, we derived optimal solutions on where to add the pathway and designed strategyproof mechanisms.

For the open directions, an immediate direction is to examine whether one can improve the gaps between the upper and lower bounds. Moreover, it would be interesting to consider more general settings, including other types of regions (e.g., convex regions), more than two regions, more than one obstacle, or more than one pathway.

Acknowledgments

Hau Chan is supported by the National Institute of General Medical Sciences of the National Institutes of Health [P20GM130461], the Rural Drug Addiction Research Center at the University of Nebraska-Lincoln, and the National Science Foundation under grants IIS:RI #2302999 and IIS:RI #2414554. The work is supported in part by the Guangdong Provincial Key Laboratory IRADS (2022B1212010006, R0400001-22). Chenhao Wang is supported by NSFC under grant 12201049, and by UIC (R0200008-23, R0700036-22). The content is solely the responsibility of the authors and does not necessarily represent the official views of the funding agencies.

References

Amekudzi, A. A.; Thomas-Mobley, L.; and Ross, C. 2007. Transportation planning and infrastructure delivery in major cities and megacities. *Transportation Research Record*, 1997(1): 17–23.

Barberà, S.; Gul, F.; and Stacchetti, E. 1993. Generalized median voter schemes and committees. *Journal of Economic Theory*, 61(2): 262–289.

Bhattacharya, B.; and Benkoczi, R. 2001. On computing the optimal bridge between two convex polygons. *Information processing letters*, 79(5): 215–221.

Boakye, J.; Guidotti, R.; Gardoni, P.; and Murphy, C. 2022. The role of transportation infrastructure on the impact of natural hazards on communities. *Reliability Engineering & System Safety*, 219: 108184.

Cai, L.; Xu, Y.; and Zhu, B. 1999. Computing the optimal bridge between two convex polygons. *Information processing letters*, 69(3): 127–130.

Chan, H.; Filos-Ratsikas, A.; Li, B.; Li, M.; and Wang, C. 2021. Mechanism Design for Facility Location Problems: A Survey. In *Proceedings of the 30th International Joint Conference on Artificial Intelligence (IJCAI)*, 4356–4365.

Chan, H.; and Wang, C. 2023. Mechanism Design for Improving Accessibility to Public Facilities. In *Proceedings of the 22nd International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, 2116–2124.

Demaine, E. D.; and Zadimoghaddam, M. 2010. Minimizing the diameter of a network using shortcut edges. In *Proceedings of the 12th International Scandinavian Workshop on Algorithm Theory (SWAT)*, 420–431. Springer.

Dokow, E.; Feldman, M.; Meir, R.; and Nehama, I. 2012. Mechanism design on discrete lines and cycles. In *Proceedings of the 13th ACM Conference on Electronic Commerce (EC)*, 423–440.

Faturechi, R.; and Miller-Hooks, E. 2015. Measuring the performance of transportation infrastructure systems in disasters: A comprehensive review. *Journal of infrastructure systems*, 21(1): 04014025.

Feldman, M.; and Wilf, Y. 2013. Strategyproof facility location and the least squares objective. In *Proceedings of the 14th ACM Conference on Electronic Commerce (EC)*, 873–890.

Filos-Ratsikas, A.; and Voudouris, A. A. 2021. Approximate mechanism design for distributed facility location. In *Proceedings of the 14th International Symposium on Algorithmic Game Theory (SAGT)*, 49–63. Springer.

Forkenbrock, D. J.; and Foster, N. S. 1990. Economic benefits of a corridor highway investment. *Transportation Research Part A: General*, 24(4): 303–312.

Gu, Y.; Fu, X.; Liu, Z.; Xu, X.; and Chen, A. 2020. Performance of transportation network under perturbations: Reliability, vulnerability, and resilience. *Transportation Research Part E: Logistics and Transportation Review*, 133: 101809.

Kim, B.-J.; Shin, C.-S.; and Chwa, K.-Y. 1998. Linear algorithms for computing a variant segment center. In *Proceedings of Korean Information Science Society Conference (KISS)*, volume A, 708–710. (in Korean).

Kim, S. K.; and Shin, C.-S. 2001. Computing the optimal bridge between two polygons. *Theory of Computing Systems*, 34(4): 337–352.

Lin, J. 2020. Nearly Complete Characterization of 2-Agent Deterministic Strategyproof Mechanisms for Single Facility Location in L p Space. In *International Conference on Combinatorial Optimization and Applications*, 411–425.

Mei, L.; Li, M.; Ye, D.; and Zhang, G. 2019. Facility location games with distinct desires. *Discrete Applied Mathematics*, 264: 148–160.

Meir, R. 2019. Strategyproof facility location for three agents on a circle. In *Proceedings of 12th International Symposium on Algorithmic Game Theory (SAGT)*, 18–33. Springer.

Meyerson, A.; and Tagiku, B. 2009. Minimizing Average Shortest Path Distances via Shortcut Edge Addition. In *Proceedings of the 12th International Workshop and 13th International Workshop on Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques (APPROX / RANDOM)*, volume 5687, 272–285.

Narayanaswami, S. 2017. Urban transportation: innovations in infrastructure planning and development. *The International Journal of Logistics Management*, 28(1): 150–171.

Nisan, N.; Roughgarden, T.; Tardos, É.; and Vazirani, V. V., eds. 2007. *Algorithmic Game Theory*. Cambridge University Press. ISBN 9780511800481.

Papagelis, M.; Bonchi, F.; and Gionis, A. 2011. Suggesting ghost edges for a smaller world. In *Proceedings of the 20th ACM international conference on Information and knowledge management (CIKM)*, 2305–2308.

Perumal, S.; Basu, P.; and Guan, Z. 2013. Minimizing eccentricity in composite networks via constrained edge additions. In *MILCOM 2013-2013 IEEE Military Communications Conference (MILCOM)*, 1894–1899. IEEE.

Procaccia, A. D.; and Tennenholtz, M. 2013. Approximate mechanism design without money. *ACM Transactions on Economics and Computation (TEAC)*, 1(4): 1–26.

Serdar, M. Z.; Koç, M.; and Al-Ghamdi, S. G. 2022. Urban transportation networks resilience: indicators, disturbances, and assessment methods. *Sustainable Cities and Society*, 76: 103452.

Sui, X. 2015. *Mechanism Design for Multi-dimensional Facility Location Problems: A Computational and Informational Perspective*. University of Toronto (Canada).

Tan, X. 2000. On optimal bridges between two convex regions. *Information processing letters*, 76(4-6): 163–168.

Tan, X. 2002. Finding an optimal bridge between two polygons. *International Journal of Computational Geometry & Applications*, 12(03): 249–261.

A Remarks on single-peaked preferences

We remark that the preference profile of agents is not one-dimensional single-peaked, though we have shown it is two-dimensional single-peaked.

Proposition 1. *For any instance of our problem, the preference profile of agents is not one-dimensional single-peaked.*

Proof. Let the obstacle be located at $o = 0.5$. Suppose for contradiction that there is a one-dimensional axis A on D so that every agent is one-dimensional single-peaked with respect to A. Consider an agent $i \in N_1$ located at $x_i = 0$. This agent has a peak at $(0, 1)$ and a preference

$$(0,1) \succeq_i (0,0.9) \succeq_i (0.2,1) \succeq_i (0.2,0.9) \tag{1}$$

over several possible outcomes.

If both outcomes $(0.2, 1), (0.2, 0.9)$ are on the same region of $(0, 1)$ with respect to axis A, by the single-peakedness and (1), the ordering can only be $(0,1) <_A (0,0.9) <_A (0.2,1) <_A (0.2,0.9)$, or $(0.2,0.9) <_A (0.2,1) <_A (0,0.9) <_A (0,1)$, or $(0,0.9) <_A (0,1) <_A (0.2,1) <_A (0.2,0.9)$, or $(0.2,0.9) <_A (0.2,1) <_A (0,1) <_A (0,0.9)$. Then we consider an agent $i' \in N_2$ located at $x_{i'} = 0.9$ whose peak is $(0, 0.9)$. In all of the four cases, by the strict ordering $<_A$ and the single-peakedness of i', agent i' should prefer $(0.2, 1)$ to $(0.2, 0.9)$. However, this is not true because

$$cost(0.2, 1, x_{i'}) = 0.3 + 0.8k > 0.2 + 0.7k = cost(0.2, 0.9, x_{i'}),$$

giving a contradiction.

If the outcomes $(0.2, 1), (0.2, 0.9)$ are on different regions of $(0, 1)$ with respect to axis A, by the single-peakedness of agent i and (1), the ordering can only be $(0.2,1) <_A (0,1) <_A (0,0.9) <_A (0.2,0.9)$, or $(0.2,1) <_A (0,0.9) <_A (0,1) <_A (0.2,0.9)$, or $(0.2,0.9) <_A (0,1) <_A (0,0.9) <_A (0.2,1)$, or $(0.2,0.9) <_A (0,0.9) <_A (0,1) <_A (0.2,1)$. Then we consider an agent $i'' \in N_1$ located at $x_{i''} = 0.2$ whose peak is $(0.2, 1)$. In all of the four cases, by the strict ordering $<_A$ and the single-peakedness of i'', agent i'' should prefer $(0, 0.9)$ to $(0.2, 0.9)$. However, this is not true because

$$cost(0, 0.9, x_{i''}) = 0.3 + 0.9k > 0.1 + 0.7k = cost(0.2, 0.9, x_{i''}),$$

giving a contradiction.

Therefore, it is impossible for every agent preference to be one-dimensional single-peaked with respect to axis A. □

As an example of 2-dimensional generalized median mechanisms, the MEDIAN mechanism selects the x-coordinate (resp. y-coordinate) of the outcome to be the median of the x-coordinates (resp. y-coordinates) of n agent peaks. Note that the peak of an agent $i \in N_1$ is $(x_i, 1)$, and the peak of an agent $j \in N_2$ is $(0, x_j)$. That is, the MEDIAN mechanism returns

$$\begin{aligned}&\big(\text{med}(0, \ldots, 0, (x_i)_{i \in N_1}, b_1, \ldots, b_{n+1}),\\ &\text{med}((x_j)_{j \in N_2}, 1, \ldots, 1, b_1, \ldots, b_{n+1})\big),\end{aligned}$$

where function med has $2n + 1$ entries, $b_1 = \cdots = b_{\lceil \frac{n}{2} \rceil} = -\infty$, and $b_{\lceil \frac{n}{2} \rceil + 1} = \cdots = b_{n+1} = +\infty$.

While this mechanism is of interests in facility location problems (Procaccia and Tennenholtz 2013; Sui 2015), it does not perform as well as our mechanisms proposed for both social cost and maximum cost. We give some examples to illustrate it. Consider the instance with location profile $\mathbf{x} = (0, \epsilon, 1 - \epsilon, 1)$, where the obstacle o satisfies $1 - \epsilon < o < 1$ and $\epsilon > 0$ is a sufficiently small positive number. For the maximum cost, the solution returned by the MEDIAN mechanism is $(\epsilon, 1)$, and the optimal solution is $(\frac{1-\epsilon}{2}, 1)$. The approximation ration is at least

$$\frac{MC(\epsilon, 1, \mathbf{x})}{MC(\frac{1-\epsilon}{2}, 1, \mathbf{x})} = \frac{k(1-\epsilon) + 1 - 2\epsilon}{k(1 - \frac{1-\epsilon}{2}) + \frac{1-\epsilon}{2}} \to 2,$$

which is larger than the approximation ratio $\frac{2}{1+k}$ of our TWOINNEREXTREME mechanism for any $k \in (0, 1)$.

For the social cost, we note that when $k = 0$ the MEDIAN mechanism is exactly our OPTSOCCOST mechanism, and thus is optimal. When $k > \frac{1}{2}$, consider profile $\mathbf{x}$. The solution returned by the MEDIAN mechanism is again $(\epsilon, 1)$, and the optimal solution is $(a \to 1 - \epsilon, 1)$. The approximation ratio is at least

$$\frac{SC(\epsilon, 1, \mathbf{x})}{SC(1-\epsilon, 1, \mathbf{x})} = \frac{4k(1-\epsilon) + \epsilon + \epsilon + 1 - 2\epsilon}{4k\epsilon + 2(1-\epsilon) + (1 - 2\epsilon)} \to \frac{4k+1}{3},$$

which is worse than our mechanism OPTSOCCOST. When $0 < k \le \frac{1}{2}$, consider another instance with one agent located at 0, t agents at ϵ, and t agents at 1 ($t = 1, 2, 3, \ldots$), where the total number of agents is $n = 2t + 1 > \frac{1}{k}$ when t is sufficiently large, and $\epsilon > 0$ is a sufficiently small positive number. The obstacle o satisfies $0 < o < \epsilon$. Thus, the location profile is $\mathbf{x}' = (0, \overbrace{\epsilon, \ldots, \epsilon}^{t}, \overbrace{1, \ldots, 1}^{t})$. The MEDIAN mechanism outputs the solution $(0, 1)$, while the optimal solution is $(0, \epsilon)$. The approximation ratio is at least

$$\frac{SC(0, 1, \mathbf{x}')}{SC(0, \epsilon, \mathbf{x}')} = \frac{k(2t+1) + t(1-\epsilon)}{k\epsilon(2t+1) + (t+1)(1-\epsilon)} \to \frac{t + k(2t+1)}{t+1} = 1 + \frac{k(2t+1) - 1}{t+1} > 1,$$

which is worse than our mechanism OPTSOCCOST.

B Other Two-Extreme Mechanisms

The TWOINNEREXTREME mechanism that returns (x_r, y_l) is proven to be group strategyproof and $\frac{2}{1+k}$-approximation for the maximum cost in Theorem 4. We remark that other two-extreme mechanisms that return (x_l, y_r), (x_l, y_l) or (x_r, y_r) are also group strategyproof, but the approximation ratio is 2.

The group strategyproofness follows from a similar analysis as in the proof of Theorem 4. For the approximation ratio, we focus only on the instances with $1 - y_r \geq x_l$, as other instances are symmetric. The optimal solution is $(a^*, b^*) = (\frac{x_l+x_r}{2}, \frac{y_l-x_l}{2} + \frac{1}{2})$, and the optimal maximum cost is attained by x_l, x_r and y_l simultaneously. For the mechanism that returns (x_l, y_r), clearly the maximum cost is attained by x_r or y_l. The cost of the agent at x_r is

$$\begin{aligned} cost(x_l, y_r, x_r) &= 2(a^* - x_l) + k(y_r - x_l) + 1 - y_r \\ &\leq 2(a^* - x_l) + 2k(b^* - a^*) + 2(1 - b^*) \\ &= 2 \cdot cost(a^*, b^*, x_r), \end{aligned}$$

where the last inequality comes from the fact that $b^* = \frac{1}{2} + \frac{y_l - x_l}{2} \leq \frac{1+y_r}{2}$. The cost of the agent at y_l is

$$\begin{aligned} cost(x_l, y_r, y_l) &= y_r - y_l + k(y_r - x_l) + x_l \\ &\leq 2(b^* - y_l) + 2k(b^* - a^*) + x_l \\ &\leq 2(b^* - y_l) + 2k(b^* - a^*) + 2a^* \\ &= 2 \cdot cost(a^*, b^*, y_l). \end{aligned}$$

For the mechanism that returns (x_l, y_l), the maximum cost is attained by x_r, that is,

$$\begin{aligned} cost(x_l, y_l, x_r) &= 2(a^* - x_l) + k(y_l - x_l) + 1 - y_l \\ &\leq 2(a^* - x_l) + 2k(b^* - a^*) + 2(1 - b^*) \\ &= 2 \cdot cost(a^*, b^*, x_r), \end{aligned}$$

For the mechanism that returns (x_r, y_r), we have $cost(x_r, y_r, x_l) \leq cost(x_r, y_l, x_l) \leq \frac{2}{1+k} cost(a^*, b^*, x_r)$ by Theorem 4, and

$$\begin{aligned} cost(x_r, y_r, y_l) &= y_r - y_l + k(y_r - x_r) + x_r \\ &\leq 2(b^* - y_l) + 2k(b^* - a^*) + 2a^* \\ &= 2 \cdot cost(a^*, b^*, y_l). \end{aligned}$$

Therefore, the approximation ratio of all above mechanisms is 2.

The 2-approximation for the two-extreme mechanisms that return (x_l, y_r), (x_l, y_l) or (x_r, y_r) cannot be improved, for any $k \in [0, 1]$. Consider any instance with $x_l = 0$, $x_r = 1 - \epsilon$ for some sufficiently small positive number ϵ, and $y_l = y_r = 1$, and the obstacle is between $1 - \epsilon$ and 1. The optimal solution is $(\frac{x_l+x_r}{2}, 1)$ by Theorem 3, and the optimal maximum cost is $\frac{1-\epsilon}{2} + k(1 - \frac{1-\epsilon}{2}) = \frac{1+k}{2} - \frac{1-k}{2}\epsilon$. However, the maximum cost induced by $(x_l, y_r) = (x_l, y_l) = (0, 1)$ is $1 - \epsilon + k$. We have $\frac{1-\epsilon+k}{(1+k)/2-(1-k)\epsilon/2} \to 2$, when ϵ approaches 0. A symmetric instance shows that the 2-approximation analysis for mechanism (x_r, y_r) is also tight.

C Proof of Theorem 8

Proof. Suppose that f is a randomized strategyproof mechanism with approximation ratio $r < \frac{6+6k}{5+7k}$. Consider the instance with location profile $\mathbf{x} = (\frac{1}{3}, \frac{2}{3}, 1)$, where the obstacle is located at $1 - \epsilon$ and $\epsilon > 0$ is a sufficiently small positive number. For convenience, we will ignore the terms with respect to ϵ in the following calculations. The optimal solution is $(\frac{1}{3}, 1)$, and the optimal maximum cost is $\frac{1}{3} + \frac{2k}{3}$. Let P be the distribution of the left endpoint returned by mechanism f. For any realization $s \sim P$, the maximum cost is $\frac{1}{3} + |s - \frac{1}{3}| + k(1 - s)$ (attained by either the agent at $\frac{1}{3}$ or the agent at $\frac{2}{3}$), and the expected maximum cost is

$$MC = \frac{1}{3} + \mathbb{E}[|s - \frac{1}{3}|] + k - k\mathbb{E}[s].$$

By the approximation ratio of r, we have

$$\frac{\frac{1}{3} + \mathbb{E}[|s - \frac{1}{3}|] + k - k\mathbb{E}[s]}{\frac{1}{3} + \frac{2k}{3}} \leq r, \tag{2}$$

$$\Rightarrow \mathbb{E}[|s - \frac{1}{3}|] \leq r(\frac{1}{3} + \frac{2k}{3}) + k(\mathbb{E}[|s - \frac{1}{3}|] + \frac{1}{3}) - \frac{1}{3} - k, \tag{3}$$

$$\Rightarrow \mathbb{E}[|s - \frac{2}{3}|] \geq \frac{1}{3} - \mathbb{E}[|s - \frac{1}{3}|] \geq \frac{1}{3} - \frac{r(\frac{1}{3} + \frac{2k}{3}) + \frac{k}{3} - \frac{1}{3} - k}{1 - k}. \tag{4}$$

Next, we consider the instance with location profile $\mathbf{x} = (\frac{1}{3}, 1-2\epsilon, 1)$, where the obstacle is located at $1-\epsilon$. Again we ignore the terms with respect to ϵ for simplicity. The optimal solution is $(\frac{1}{2}, 1)$, and the optimal maximum cost is $\frac{1}{2}+\frac{k}{2}$. Let P' be the distribution of the left endpoint returned by mechanism f, and $s' \sim P'$. By the strategyproofness, we have

$$\mathbb{E}_{s'\sim P'}[|s'-\frac{2}{3}|] + k(1-\mathbb{E}[s']) \geq \mathbb{E}[|s-\frac{2}{3}|] + k(1-\mathbb{E}[s]), \tag{5}$$

as otherwise the agent located at $\frac{2}{3}$ in the first instance would like to misreport the location as $1-2\epsilon$ and decrease the cost.

If $\mathbb{E}[s'] \leq \mathbb{E}[s]$, then by (2) we have

$$\frac{(1-k)\mathbb{E}[s']+k}{\frac{1}{3}+\frac{2k}{3}} \leq \frac{\mathbb{E}[s]+k-k\mathbb{E}[s]}{\frac{1}{3}+\frac{2k}{3}} \leq r$$
$$\Rightarrow\ \mathbb{E}[s'] \leq \frac{(\frac{1}{3}+\frac{2k}{3})r-k}{1-k}.$$

The maximum cost induced by the mechanism is at least

$$\begin{aligned}\frac{1}{2}+\mathbb{E}[|s'-\frac{1}{2}|]+k(1-\mathbb{E}[s']) &\geq 1+k-(1+k)\mathbb{E}[s'] \\ &\geq 1+k-(1+k)\frac{(\frac{1}{3}+\frac{2k}{3})r-k}{1-k}\end{aligned}$$

Recall that the optimal maximum cost is $\frac{1+k}{2}$. Hence, the approximation ratio is at least

$$2-\frac{2(\frac{1}{3}+\frac{2k}{3})r-2k}{1-k} = \frac{2-\frac{2+4k}{3}r}{1-k},$$

which can be easily verified to be larger than $\frac{6+6k}{5+7k}$ for any $k \in [0,1)$, given that $r < \frac{6+6k}{5+7k}$. Therefore, it contradicts the approximation ratio.

If $\mathbb{E}[s'] > \mathbb{E}[s]$, then by (5) we have $\mathbb{E}[|s'-\frac{2}{3}|] \geq \mathbb{E}[|s-\frac{2}{3}|]$. The maximum cost induced by the mechanism is at least

$$\begin{aligned}\frac{1}{2}+\mathbb{E}[|s'-\frac{1}{2}|]+k(1-\mathbb{E}[s']) &\geq \frac{1}{2}+\mathbb{E}[|s'-\frac{1}{2}|]+k(\frac{1}{2}-\mathbb{E}[|s'-\frac{1}{2}|]) \\ &= \frac{1+k}{2}+(1-k)\mathbb{E}[|s'-\frac{1}{2}|] \\ &\geq \frac{1+k}{2}+(1-k)(\mathbb{E}[|s'-\frac{2}{3}|]-\frac{1}{6}) \\ &\geq \frac{1+k}{2}+(1-k)(\mathbb{E}[|s-\frac{2}{3}|]-\frac{1}{6}) \\ &\geq \frac{1+k}{2}+(1-k)(\frac{1}{3}-\frac{r(\frac{1}{3}+\frac{2k}{3})+\frac{k}{3}-\frac{1}{3}-k}{1-k})-\frac{1-k}{6} \\ &= \frac{2+k}{3}-r(\frac{1}{3}+\frac{2k}{3})-\frac{k}{3}+\frac{1}{3}+k \\ &= 1+k-r(\frac{1}{3}+\frac{2k}{3}),\end{aligned}$$

where the last inequality comes from (4). Then the approximation ratio is at least

$$\frac{1+k-r(\frac{1}{3}+\frac{2k}{3})}{\frac{1}{2}+\frac{k}{2}},$$

which is strictly larger than r when $r < \frac{6+6k}{5+7k}$. This gives a contradiction. □

D Proof of Theorem 7

Lemma 3. *Mechanism 4 is group strategyproof.*

Proof. We show that the mechanism is group strategyproof whenever $p \geq \frac{1+k}{3-k}$. Consider a group of agents $S \subseteq N_1 \cup N_2$. Let $f(\mathbf{x}) = (a,b)$ be the outcome when all agents report true locations, and $f(\mathbf{x}'_S, \mathbf{x}_{-S}) = (a',b')$ be the outcome when the agents in S misreport $\mathbf{x}'_S$, where a, b, a', b' are random variables that follow the distributions given in the mechanism. Assume w.l.o.g.

that $|\mathbb{E}[a]-\mathbb{E}[a']| \geq |\mathbb{E}[b]-\mathbb{E}[b']|$, which will cause $|\mathbb{E}[a]-\mathbb{E}[a']| \neq 0$. We show that at least one agent in the group cannot gain by misreporting.

Case 1. When $\mathbb{E}[a'] < \mathbb{E}[a]$, then it must be $x'_r < x_r$, and the agent located at x_r is in the group. Under the solution (a, b), the cost of the agent at x_r is

$$cost(a, b, x_r) = x_r - \mathbb{E}[a] + k(\mathbb{E}[b] - \mathbb{E}[a]) + (1 - \mathbb{E}[b]).$$

Under the solution (a', b'), the cost of the agent at x_r is

$$cost(a', b', x_r) = x_r - \mathbb{E}[a'] + k(\mathbb{E}[b'] - \mathbb{E}[a']) + (1 - \mathbb{E}[b']).$$

Since $|\mathbb{E}[a]-\mathbb{E}[a']| \geq |\mathbb{E}[b]-\mathbb{E}[b']|$, it follows that

$$cost(a', b', x_r) - cost(a, b, x_r) = (1 + k)(\mathbb{E}[a] - \mathbb{E}[a']) - (1 - k)(\mathbb{E}[b'] - \mathbb{E}[b]) \geq 0,$$

indicating that this agent cannot gain.

Case 2. When $\mathbb{E}[a'] > \mathbb{E}[a], \mathbb{E}[b'] \leq \mathbb{E}[b]$, there exists at least one agent $i \in S \cap N_1$. It is clear that any agent located at $[0, \frac{x_r}{2}]$ cannot gain because the change of the endpoints in both sides do not benefit this agent. For an agent $i \in N_1$ located at $(\frac{x_r}{2}, x_r]$, under the solution (a, b), the cost of the agent at x_r is

$$\begin{aligned} cost(a, b, x_i) &= p \cdot (x_r - x_i + k(\mathbb{E}[b] - x_r)) + (1 - p) \cdot (x_i - \frac{x_r}{2} + k(\mathbb{E}[b] - \frac{x_r}{2})) + 1 - \mathbb{E}[b] \\ &= p \cdot (x_r - x_i - kx_r) + (1 - p) \cdot (x_i - \frac{x_r}{2} - \frac{kx_r}{2}) + 1 - (1 - k)\mathbb{E}[b] \end{aligned}$$

Under the solution (a', b'), the cost of this agent is

$$cost(a', b', x_i) = p \cdot (x'_r - x_i - kx'_r) + (1 - p) \cdot (|x_i - \frac{x'_r}{2}| - \frac{kx'_r}{2}) + 1 - (1 - k)\mathbb{E}[b'].$$

Since $\mathbb{E}[b'] \leq \mathbb{E}[b]$ and $|x_i - \frac{x'_r}{2}| \geq x_i - \frac{x'_r}{2}$, it follows that

$$\begin{aligned} cost(a', b', x_i) - cost(a, b, x_i) &\geq p \cdot (1 - k)(x'_r - x_r) - (1 - p) \cdot \frac{1 + k}{2}(x'_r - x_r) \\ &\geq \frac{1 + k}{3 - k}(1 - k)(x'_r - x_r) - \frac{2 - 2k}{3 - k} \cdot \frac{1 + k}{2}(x'_r - x_r) = 0, \end{aligned}$$

indicating that this agent cannot gain.

Case 3. When $\mathbb{E}[a'] > \mathbb{E}[a], \mathbb{E}[b'] > \mathbb{E}[b]$, the agent located at y_l must be in the group and misreport a location on the right of y_l. Still we calculate the cost of agent at y_l. Under solution (a, b), it is

$$cost(a, b, y_l) = \mathbb{E}[b] - y_l + k(\mathbb{E}[b] - \mathbb{E}[a]) + \mathbb{E}[a]$$

And under solution (a', b'), it is

$$cost(a', b', y_l) = \mathbb{E}[b'] - y_l + k(\mathbb{E}[b'] - \mathbb{E}[a']) + \mathbb{E}[a']$$

So we have

$$cost(a', b', y_l) - cost(a, b, y_l) = (1 + k)(\mathbb{E}[b'] - \mathbb{E}[b]) + (1 - k)(\mathbb{E}[a'] - \mathbb{E}[a]) > 0,$$

and thus this agent cannot decrease the cost. □

Then we prove the approximation ratio.

Proof. For the approximation, given any instance with location profile $\mathbf{x}$, we assume w.l.o.g. that $1 - y_r \geq x_l$. By Theorem 3, the optimal solution is $(a, b) = (\frac{x_l + x_r}{2}, \frac{y_l - x_l}{2} + \frac{1}{2})$, and the optimal maximum cost is

$$cost(a, b, x_l) = a - x_l + k(b - a) + 1 - b = \frac{1}{2}[1 + (1 - k)x_r + k(1 - 2x_l) - (1 - k)y_l].$$

Now we consider the solution returned by Mechanism 4. We discuss the 2 realizations of the probability distribution.

- (x_r, y_l) with probability p. By the analysis in the proof of Theorem 4 and the assumption $1 - y_r \geq x_l$, the maximum cost is attained by x_l, that is,

$$MC(x_r, y_l, \mathbf{x}) = cost(x_r, y_l, x_l) = x_r - x_l + k(y_l - x_r) + 1 - y_l.$$

- $(\frac{x_r}{2}, \frac{1+y_l}{2})$ with probability $1 - p$. The maximum cost is attained by x_r or y_l, where both costs are equal to

$$MC(\frac{x_r}{2}, \frac{1 + y_l}{2}, \mathbf{x}) = \frac{x_r}{2} + k(\frac{1 + y_l}{2} - \frac{x_r}{2}) + \frac{1 - y_l}{2}.$$

Then the expected maximum cost is

$$
\begin{aligned}
& p \cdot (x_r - x_l + k(y_l - x_r) + 1 - y_l) + (1-p) \cdot \left(\frac{x_r}{2} + k(\frac{1+y_l}{2} - \frac{x_r}{2}) + \frac{1-y_l}{2} \right) \\
= & \frac{(1+p)(1-k)(x_r - y_l) + 1 + p + (1-p)k - 2px_l}{2}.
\end{aligned}
$$

Therefore, the ratio between the expected maximum cost and the optimal maximum cost is

$$
\begin{aligned}
& \frac{(1+p)(1-k)(x_r - y_l) + 1 + p + (1-p)k - 2px_l}{1 + k + (1-k)(x_r - y_l) - 2kx_l} \\
\leq & \frac{1 + p + (1-p)k - 2px_l}{1 + k - 2kx_l} && (6) \\
\leq & \max \left(\frac{1 + p + (1-p)k}{1+k}, \frac{1 + p + (1-p)k - p}{1 + k - k} \right) && (7) \\
= & 1 + \max \left(\frac{p(1-k)}{1+k}, (1-p)k \right), && (8)
\end{aligned}
$$

where (6) is because $\frac{1+p+(1-p)k-2px_l}{1+k-2kx_l}$ is no more than $1+p$, and (7) comes from the facts that $1 - y_r \geq x_l$ and $x_l \leq 0.5$.

Though setting $p = \frac{k^2+k}{1+k^2}$ minimizes the bound in (8), recall that the mechanism is group strategyproof only when $p \geq \frac{1+k}{3-k}$. Hence,we set $p = \max \left(\frac{k^2+k}{1+k^2}, \frac{1+k}{3-k} \right)$. In this way, when $k < 0.5$, the ratio is at most $\frac{4-2k}{3-k}$, and when $k > 0.5$, the ratio is at most $\frac{1+k}{1+k^2}$. □

E Proof of Theorem 5

In this proof, we first treat c as a parameter in interval $[0, 1]$ and do not specify its value. Then we select the best value of c to minimize the approximation ratio.

Proof of the group strategyproofness. Let (a, b) be the output. Note that a and b are independent random variables, and $(1-c)o$ and $o + c(1-o)$ are two constants only related to k. Consider a group of agents $S \subseteq N_1 \cup N_2$, and let (a', b') be the output when the agents in S misreport. Assume w.l.o.g. that $|a' - a| \geq |b' - b|$ and $|a' - a| > 0$.

When $a < x_r$, since $|a' - a| > 0$, an agent located at x_r must be in the group and misreport a location to the left of x_r, implying that $a' < a < x_r$. The cost of this agent decreases by at most $(1-k)|b' - b| - (1+k)(a - a') \leq 0$, indicating that this agent can never gain.

When $a = x_r$, it is either the case when an agent at x_r misreports to its left so that $a' < a$, or the case when some agent in N_1 misreports to the right of x_r so that $a' > a$. In both cases, the cost of this agent decreases by at most $(1-k)|b' - b| - (1-k)|a' - a| \leq 0$, indicating that this agent can never gain.

Proof of the approximation ratio. Given any instance with location profile $\mathbf{x}$, we assume w.l.o.g. that $1 - y_r \geq x_l$. By Theorem 3, the optimal solution is $(a^*, b^*) = (\frac{x_l + x_r}{2}, \frac{y_l - x_l}{2} + \frac{1}{2})$, and the optimal maximum cost is

$$
OPT = cost(a^*, b^*, x_l) = a^* - x_l + k(b^* - a^*) + 1 - b^* = \frac{1}{2}[1 + (1-k)x_r + k(1 - 2x_l) - (1-k)y_l].
$$

We discuss four cases with respect to the output of the mechanism.

Case 1. $x_r \leq o(1-c), y_l \geq o + c(1-o)$. The output is $a = x_r, b = y_l$, and the maximum cost must be achieved by x_l, because when $1-y_r \ge x_l$ the cost at y_r is no more than the cost at x_l. We have

$$\begin{aligned}
\frac{cost(a,b,x_l)}{2\cdot OPT} &= \frac{x_r - x_l + k(y_l - x_r) + 1 - y_l}{1+(1-k)x_r + k(1-2x_l) - (1-k)y_l} \\
&= \frac{1+(1-k)x_r - x_l - (1-k)y_l}{1+(1-k)x_r + k(1-2x_l) - (1-k)y_l} \qquad (9)\\
&\le \frac{1+(1-k)(1-c)o - x_l - (1-k)y_l}{1+(1-k)(1-c)o + k(1-2x_l) - (1-k)y_l} \\
&\le \frac{1+(1-k)(1-c)o - x_l - (1-k)(o+c(1-o))}{1+(1-k)(1-c)o + k(1-2x_l) - (1-k)(o+c(1-o))} \\
&= \frac{1-(1-k)c - x_l}{1+k-(1-k)c-2kx_l} \qquad (10)\\
&\le \frac{1-(1-k)c}{1+k-(1-k)c}.
\end{aligned}$$

The first inequality is because (9) is no more than 1, and the last inequality is because (10) is no more than $\frac{1}{2k}$.

Case 2. $x_r \ge o(1-c), y_l \le o + c(1-o)$. The output is $a = o(1-c), b = o + c(1-o)$. Note that the cost at y_r is either at most the cost at y_l or at most that at x_l (since $1 - y_r \ge x_l$). Thus the maximum cost is achieved by at least one of the agents at x_l, x_r, y_l. First, we consider the cost at x_l, and we can assume $x_l \le o(1-c)$; otherwise we have $cost(a,b,x_l) \le cost(a,b,x_r)$ and it reduces to consider the cost at x_r. We have

$$\begin{aligned}
\frac{cost(a,b,x_l)}{2\cdot OPT} &= \frac{o(1-c) - x_l + k(o+c(1-o) - o(1-c)) + 1 - o - c(1-o)}{1+(1-k)x_r + k(1-2x_l) - (1-k)y_l} \\
&= \frac{1-(1-k)c - x_l}{1+(1-k)x_r + k(1-2x_l) - (1-k)y_l} \\
&\le \frac{1-(1-k)c - x_l}{1+(1-k)(1-c)o + k(1-2x_l) - (1-k)(o+c(1-o))} \\
&= \frac{1-(1-k)c - x_l}{1+k(1-2x_l) - (1-k)c} \\
&\le \frac{1-(1-k)c}{1+k-(1-k)c}.
\end{aligned}$$

Second, for the cost at x_r, we have

$$\begin{aligned}
\frac{cost(a,b,x_r)}{2\cdot OPT} &= \frac{x_r - o(1-c) + kc + 1 - (o+c(1-o))}{1+(1-k)x_r + k(1-2x_l) - (1-k)y_l} \\
&= \frac{x_r - 2o(1-c) + 1 - (1-k)c}{1+(1-k)x_r + k(1-2x_l) - (1-k)y_l} \\
&\le \frac{x_r - 2o(1-c) + 1 - (1-k)c}{1+(1-k)x_r + k(1-2x_l) - (1-k)(o+c(1-o))} \\
&\le \frac{o - 2o(1-c) + 1 - (1-k)c}{1+(1-k)o + k(1-2x_l) - (1-k)(o+c(1-o))} \\
&\le \frac{o - 2o(1-c) + 1 - (1-k)c}{1+(1-k)o + k(1-2\min(o,1-o)) - (1-k)(o+c(1-o))} \\
&\le \max\left(\frac{1-(1-k)c}{1+k-(1-k)c}, \frac{1+2ck}{2-(1-k)c}, c\right).
\end{aligned}$$

The second last inequality is because $x_l \le o$ and $x_l \le 1 - y_r \le 1 - o$. For the last inequality, we regard o as a variable, and it is easy to find that when $0 \le o \le 0.5$, we have

$$\frac{cost(a,b,x_r)}{2\cdot OPT} \le \frac{o - 2o(1-c) + 1 - (1-k)c}{1+(1-k)o + k(1-2o) - (1-k)(o+c(1-o))},$$

and when $0.5 < o \le 1$, we have

$$\frac{cost(a,b,x_r)}{2\cdot OPT} \le \frac{o - 2o(1-c) + 1 - (1-k)c}{1+(1-k)o + k(1-2(1-o)) - (1-k)(o+c(1-o))}.$$

Since both expressions on the right hand side are monotone with respect to o (possibly increasing or decreasing), the upper bound must be attained by the maximum of the three cases when $o = 0, 0.5, 1$, establishing the inequality.

Last, for the cost at y_l we have

$$\begin{aligned}
\frac{cost(a,b,y_l)}{2\cdot OPT} &= \frac{o(1-c)+kc+o+c(1-o)-y_l}{1+(1-k)x_r+k(1-2x_l)-(1-k)y_l} \\
&\le \frac{o(1-c)+kc+o+c(1-o)-y_l}{1+(1-k)o(1-c)+k(1-2x_l)-(1-k)y_l} \\
&\le \frac{o(1-c)+kc+o+c(1-o)-o}{1+(1-k)o(1-c)+k(1-2x_l)-(1-k)o} \\
&\le \frac{o(1-c)+kc+o+c(1-o)-o}{1+(1-k)o(1-c)+k(1-2\min(o,1-o))-(1-k)o} \\
&\le \max\left(c, \frac{1+2ck}{2-(1-k)c}, \frac{1-(1-k)c}{1+k-(1-k)c}\right).
\end{aligned}$$

For the last inequality, we regard o as a variable, and it is easy to find that when $0 \le o \le 0.5$, we have

$$\frac{cost(a,b,y_l)}{2\cdot OPT} \le \frac{o(1-c)+kc+o+c(1-o)-o}{1+(1-k)o(1-c)+k(1-2o)-(1-k)o},$$

and when $0.5 < o \le 1$, we have

$$\frac{cost(a,b,y_l)}{2\cdot OPT} \le \frac{o(1-c)+kc+o+c(1-o)-o}{1+(1-k)o(1-c)+k(1-2(1-o))-(1-k)o}.$$

Since both expressions on the right hand side are monotone with respect to o, the upper bound must be attained by the maximum of the three cases when $o = 0, 0.5, 1$, establishing the inequality.

Case 3. $x_r \le o(1-c), y_l \le o + c(1-o)$. The output is $a = x_r, b = o + c(1-o)$, and the maximum cost is achieved by x_l or y_l. First, we consider the cost at x_l, and we have

$$\begin{aligned}
\frac{cost(a,b,x_l)}{2\cdot OPT} &= \frac{x_r-x_l+k(o+c(1-o)-x_r)+1-o-c(1-o)}{1+(1-k)x_r+k(1-2x_l)-(1-k)y_l} \\
&= \frac{1+(1-k)x_r-x_l-(1-k)(o+c(1-o))}{1+(1-k)x_r+k(1-2x_l)-(1-k)y_l} \\
&\le \frac{1+(1-k)(1-c)o-x_l-(1-k)(o+c(1-o))}{1+(1-k)(1-c)o+k(1-2x_l)-(1-k)y_l} \\
&\le \frac{1+(1-k)(1-c)o-x_l-(1-k)(o+c(1-o))}{1+(1-k)(1-c)o+k(1-2x_l)-(1-k)(o+c(1-o))} \\
&= \frac{1-(1-k)c-x_l}{1-(1-k)c+k(1-2x_l)} \\
&\le \frac{1-(1-k)c-0}{1-(1-k)c+k(1-0)} \\
&= \frac{1-(1-k)c}{1+k-(1-k)c}.
\end{aligned}$$

Second, for the cost at y_l, we have

$$\begin{aligned}\frac{cost(a,b,y_l)}{2\cdot OPT} &= \frac{x_r + k(o+c(1-o)-x_r)+o+c(1-o)-y_l}{1+(1-k)x_r+k(1-2x_l)-(1-k)y_l}\\ &= \frac{(1-k)x_r - y_l + (1+k)(o+c(1-o))}{1+(1-k)x_r+k(1-2x_l)-(1-k)y_l}\\ &\le \frac{(1-k)o(1-c) - y_l + (1+k)(o+c(1-o))}{1+(1-k)o(1-c)+k(1-2x_l)-(1-k)y_l}\\ &\le \frac{(1-k)o(1-c) - o + (1+k)(o+c(1-o))}{1+(1-k)o(1-c)+k(1-2x_l)-(1-k)o}\\ &\le \frac{(1-k)o(1-c) - o + (1+k)(o+c(1-o))}{1-(1-k)oc+k(1-2\min(1-o,o(1-c)))}\\ &\le \max\left(c, \frac{k(2c-c^2)+1-c^2}{2-2c+2ck}, \frac{1-(1-k)c}{1+k-(1-k)c}\right).\end{aligned}$$

The second last inequality is because $x_l \le x_r \le o(1-c)$ and $x_l \le 1-y_r \le 1-o$. For the last inequality, we regard o as a variable, and it is easy to find that when $0 \le o \le \frac{1}{2-c}$, we have

$$\frac{cost(a,b,y_l)}{2\cdot OPT} \le \frac{o-2oc+c+kc}{1-(1-k)oc+k(1-2o(1-c))},$$

and when $\frac{1}{2-c} < o \le 1$, we have

$$\frac{cost(a,b,y_l)}{2\cdot OPT} \le \frac{o-2oc+c+kc}{1-(1-k)oc+k(1-2(1-o))}.$$

Since both expressions on the right hand side are monotone with respect to o, the upper bound must be attained by the maximum of the three cases when $o = 0, \frac{1}{2-c}, 1$, establishing the inequality.

Case 4. $x_r \ge o(1-c), y_l \ge o+c(1-o)$. The output is $a = o(1-c), b = y_l$, and the maximum cost is achieved by x_l or x_r. First, we consider $cost(a,b,x_l)$, and we can assume $x_l \le o(1-c)$, as otherwise $cost(x_l) \le cost(x_r)$. We have

$$\begin{aligned}\frac{cost(a,b,x_l)}{2\cdot OPT} &= \frac{o(1-c)-x_l+k(y_l-o(1-c))+1-y_l}{1+(1-k)x_r+k(1-2x_l)-(1-k)y_l}\\ &= \frac{1+(1-k)(1-c)o-x_l-(1-k)y_l}{1+(1-k)x_r+k(1-2x_l)-(1-k)y_l}\\ &\le \frac{1+(1-k)(1-c)o-x_l-(1-k)y_l}{1+(1-k)o(1-c)+k(1-2x_l)-(1-k)y_l}\\ &\le \frac{1+(1-k)(1-c)o-x_l-(1-k)(o+c(1-o))}{1+(1-k)(1-c)o+k(1-2x_l)-(1-k)(o+c(1-o))}\\ &= \frac{1-x_l-(1-k)c}{1+k(1-2x_l)-(1-k)c}\\ &\le \frac{1-(1-k)c}{1+k-(1-k)c}.\end{aligned}$$

Second, for the cost at x_r, we have

$$\begin{aligned}
\frac{cost(a,b,x_r)}{2\cdot OPT} &= \frac{x_r - o(1-c) + k(y_l - o(1-c)) + 1 - y_l}{1 + (1-k)x_r + k(1-2x_l) - (1-k)y_l} \\
&= \frac{1 + x_r - (1+k)(1-c)o - (1-k)y_l}{1 + (1-k)x_r + k(1-2x_l) - (1-k)y_l} \\
&\le \frac{1 + o - (1+k)(1-c)o - (1-k)y_l}{1 + (1-k)o + k(1-2x_l) - (1-k)y_l} \\
&\le \frac{1 + o - (1+k)(1-c)o - (1-k)(o + c(1-o))}{1 + (1-k)o + k(1-2x_l) - (1-k)(o + c(1-o))} \\
&\le \frac{1 - o + 2co - (1-k)c}{1 + k(1 - 2\min(o, (1-c)(1-o))) - (1-k)c(1-o)} \\
&\le \max\left(\frac{1-(1-k)c}{1+k-(1-k)c}, \frac{k(2c-c^2)+1-c^2}{2-2c+2ck}, c\right).
\end{aligned}$$

The second last inequality is because $x_l \le x_r \le o$ and $x_l \le 1 - y_r \le 1 - y_l \le 1 - o - c(1-o)$. For the last inequality, we regard o as a variable, and it is easy to find that when $0 \le o \le \frac{1-c}{2-c}$, we have

$$\frac{cost(a,b,x_r)}{2\cdot OPT} \le \frac{1 - o + 2co - (1-k)c}{1 + k(1-2o) - (1-k)c(1-o)},$$

and when $\frac{1-c}{2-c} < o \le 1$, we have

$$\frac{cost(a,b,x_r)}{2\cdot OPT} \le \frac{1 - o + 2co - (1-k)c}{1 + k(1 - 2(1-c)(1-o)) - (1-k)c(1-o)}.$$

Since both expressions on the right hand side are monotone with respect to o, the upper bound must be attained by the maximum of the three cases when $o = 0, \frac{1-c}{2-c}, 1$, establishing the inequality.

According to the four cases above, the ratio is

$$\frac{MC(a,b,\mathbf{x})}{OPT} \le 2\cdot\max\left(\frac{1-(1-k)c}{1+k-(1-k)c}, \frac{k(2c-c^2)+1-c^2}{2-2c+2ck}, \frac{1+2ck}{2-(1-k)c}, c\right). \tag{11}$$

We need to select a proper value of c so that the right hand side of (11) is minimized. Fixing k, note that $\frac{1-(1-k)c}{1+k-(1-k)c}$ is decreasing with c, and $\frac{1+2ck}{2-(1-k)c}$ is increasing with c. Consider the equation

$$\frac{1-(1-k)c}{1+k-(1-k)c} = \frac{1+2ck}{2-(1-k)c}.$$

The only solution is

$$c = \frac{1 + k^2 - \sqrt{k^4 - k^3 + 3k^2 + k}}{1 - k^2}.$$

Furthermore, when $c = \frac{1+k^2-\sqrt{k^4-k^3+3k^2+k}}{1-k^2}$, we have $\frac{1-(1-k)c}{1+k-(1-k)c} = \frac{k(2c-c^2)+1-c^2}{2-2c+2ck}$ and $\frac{1-(1-k)c}{1+k-(1-k)c} \ge c$. Hence, it minimizes the right hand side of (11).

F Another randomized mechanism for maximum cost

We present another randomized mechanism that is at most 1.441-approximation for maximum cost.

Mechanism 5. Given location profile $\mathbf{x}$, let a be x_r and $\frac{x_r}{2}$ with probabilities $\frac{1+k}{3-k}$ and $\frac{2(1-k)}{3-k}$, respectively. Let b be y_l and $\frac{1+y_l}{2}$ with probabilities $\frac{1+k}{3-k}$ and $\frac{2(1-k)}{3-k}$, respectively. Return (a, b).

Lemma 4. *Mechanism 5 is group strategyproof.*

Proof. We consider a group of agents $S \subseteq N_1 \cup N_2$. Let $f(\mathbf{x}) = (a, b)$ be the outcome when all agents report true locations, and $f(\mathbf{x}'_S, \mathbf{x}_{-S}) = (a', b')$ be the outcome when the agents in S misreport $\mathbf{x}'_S$, where a, b, a', b' are random variables that follow the distributions given in the mechanism. Assume w.l.o.g. that $|\mathbb{E}[a] - \mathbb{E}[a']| \ge |\mathbb{E}[b] - \mathbb{E}[b']|$. We show that at least one agent in the group cannot gain by misreporting.

When $\mathbb{E}[a'] < \mathbb{E}[a]$, then it must be $x_r' < x_r$, and the agent located at x_r is in the group. Under the solution (a, b), the cost of the agent at x_r is

$$cost(a, b, x_r) = x_r - \mathbb{E}[a] + k(\mathbb{E}[b] - \mathbb{E}[a]) + (1 - \mathbb{E}[b]).$$

Under the solution (a', b'), the cost of the agent at x_r is

$$cost(a', b', x_r) = x_r - \mathbb{E}[a'] + k(\mathbb{E}[b'] - \mathbb{E}[a']) + (1 - \mathbb{E}[b']).$$

Since $|\mathbb{E}[a] - \mathbb{E}[a']| \geq |\mathbb{E}[b] - \mathbb{E}[b']|$, it follows that

$$cost(a', b', x_r) - cost(a, b, x_r) = (1 + k)(\mathbb{E}[a] - \mathbb{E}[a']) + (1 - k)(\mathbb{E}[b] - \mathbb{E}[b']) \geq 0,$$

indicating that this agent cannot gain.

When $\mathbb{E}[a'] > \mathbb{E}[a]$, there exists at least one agent $i \in S \cap N_1$. If $\mathbb{E}[b'] \leq \mathbb{E}[b]$, it is clear that any agent located at $[0, \frac{x_r}{2}]$ cannot gain because the change of the endpoints in both regions do not benefit this agent. For an agent $i \in N_1$ located at $(\frac{x_r}{2}, x_r]$, under the solution (a, b), the cost of the agent at x_r is

$$\begin{aligned} cost(a, b, x_i) &= \frac{1+k}{3-k}(x_r - x_i + k(\mathbb{E}[b] - x_r)) + \frac{2-2k}{3-k}(x_i - \frac{x_r}{2} + k(\mathbb{E}[b] - \frac{x_r}{2})) + 1 - \mathbb{E}[b] \\ &= \frac{1+k}{3-k}(x_r - x_i - kx_r) + \frac{2-2k}{3-k}(x_i - \frac{x_r}{2} - \frac{kx_r}{2}) + 1 - (1-k)\mathbb{E}[b] \end{aligned}$$

Under the solution (a', b'), the cost of this agent is

$$cost(a', b', x_i) = \frac{1+k}{3-k}(x_r' - x_i - kx_r') + \frac{2-2k}{3-k}(|x_i - \frac{x_r'}{2}| - \frac{kx_r'}{2}) + 1 - (1-k)\mathbb{E}[b'].$$

Since $\mathbb{E}[b'] \leq \mathbb{E}[b]$, it follows that

$$cost(a', b', x_i) - cost(a, b, x_i) \geq \frac{1+k}{3-k}(1-k)(x_r' - x_r) - \frac{2-2k}{3-k} \cdot \frac{1+k}{2}(x_r' - x_r) = 0,$$

indicating that this agent cannot gain. If $\mathbb{E}[b'] > \mathbb{E}[b]$, then the agent located at y_l must be in the group and misreport a location on the right of y_l. It is easy to see that

$$cost(a', b', y_l) - cost(a, b, y_l) = (1 + k)(\mathbb{E}[b'] - \mathbb{E}[b]) + (1 - k)(\mathbb{E}[a'] - \mathbb{E}[a]) \geq 0,$$

and thus this agent cannot decrease the cost. □

Now we prove the approximation ratio.

Theorem 9. *Mechanism 5 is a randomized group strategyproof mechanism. The approximation ratio for maximum cost is* $\frac{4-2k}{3-k}$ *when* $k \in [0, \kappa]$*, and is* $\frac{11+2k^3-9k^2}{9+k^2-6k}$ *when* $k \in [\kappa, 1)$*, where* $\kappa = \frac{9-\sqrt{73}}{4} \approx 0.114$.

Proof. Given any instance with location profile $\mathbf{x}$, we assume w.l.o.g. that $1 - y_r \geq x_l$. By Theorem 3, the optimal solution is $(a, b) = (\frac{x_l+x_r}{2}, \frac{y_l-x_l}{2} + \frac{1}{2})$, and the optimal maximum cost is

$$cost(a, b, x_l) = a - x_l + k(b - a) + 1 - b = \frac{1}{2}[1 + (1-k)x_r + k(1 - 2x_l) - (1-k)y_l].$$

Now we consider the solution returned by Mechanism 5. We discuss the 4 realizations of the probability distribution.

- (x_r, y_l) with probability $\frac{(1+k)^2}{(3-k)^2}$. By the analysis in the proof of Theorem 4 and the assumption $1 - y_r \geq x_l$, the maximum cost is attained by x_l, that is,

$$MC(x_r, y_l, \mathbf{x}) = cost(x_r, y_l, x_l) = x_r - x_l + k(y_l - x_r) + 1 - y_l.$$

- $(x_r, \frac{y_l+1}{2})$ with probability $\frac{2(1-k)(1+k)}{(3-k)^2}$. The maximum cost is attained by x_l or y_l. The cost of x_l is $x_r - x_l + k(\frac{y_l+1}{2} - x_r) + 1 - \frac{y_l+1}{2}$, and the cost of y_l is $\frac{1-y_l}{2} + k(\frac{y_l+1}{2} - x_r) + x_r$. It is easy to see that the cost of y_l is no less than the cost of x_l. Hence, the maximum cost is attained by y_l, that is,

$$MC(x_r, \frac{y_l+1}{2}, \mathbf{x}) = cost(x_r, \frac{y_l+1}{2}, y_l) = \frac{1-y_l}{2} + k(\frac{y_l+1}{2} - x_r) + x_r.$$

- $(\frac{x_r}{2}, y_l)$ with probability $\frac{2(1-k)(1+k)}{(3-k)^2}$. The maximum cost is attained by x_r, that is,

$$MC(\frac{x_r}{2}, y_l, \mathbf{x}) = cost(\frac{x_r}{2}, y_l, x_r) = \frac{x_r}{2} + k(y_l - \frac{x_r}{2}) + 1 - y_l.$$

- $(\frac{x_r}{2}, \frac{1+y_l}{2})$ with probability $\frac{4(1-k)^2}{(3-k)^2}$. The maximum cost is attained by x_r or y_l, where both costs are equal to

$$MC(\frac{x_r}{2}, \frac{1+y_l}{2}, \mathbf{x}) = \frac{x_r}{2} + k(\frac{1+y_l}{2} - \frac{x_r}{2}) + \frac{1-y_l}{2}.$$

Therefore, the expected maximum cost of the solution returned by the mechanism is

$$\begin{aligned}
&\frac{(1+k)^2}{(3-k)^2} \cdot (x_r - x_l + k(y_l - x_r) + 1 - y_l) + \frac{2(1-k)(1+k)}{(3-k)^2} \cdot \left(\frac{1-y_l}{2} + k(\frac{y_l+1}{2} - x_r) + x_r\right) \\
&+ \frac{2(1-k)(1+k)}{(3-k)^2} \cdot \left(\frac{x_r}{2} + k(y_l - \frac{x_r}{2}) + 1 - y_l\right) + \frac{4(1-k)^2}{(3-k)^2} \cdot \left(\frac{x_r}{2} + k(\frac{1+y_l}{2} - \frac{x_r}{2}) + \frac{1-y_l}{2}\right) \\
=& \frac{(1+k)^2}{(3-k)^2} \cdot (x_r - x_l + k(y_l - x_r) + 1 - y_l) + \frac{2(1-k)(1+k)}{(3-k)^2} \cdot \left(\frac{3-3y_l}{2} + k(\frac{3y_l+1}{2} - \frac{3x_r}{2}) + \frac{3x_r}{2}\right) \\
&+ \frac{4(1-k)^2}{(3-k)^2} \cdot \left(\frac{x_r}{2} + k(\frac{1+y_l}{2} - \frac{x_r}{2}) + \frac{1-y_l}{2}\right) \\
=& \frac{(1+k)^2}{(3-k)^2} \cdot (x_r - x_l + k(y_l - x_r) + 1 - y_l) + \frac{2(1-k)(1+k)}{(3-k)^2} \cdot \left(\frac{3(k-1)(y_l - x_r)}{2} + \frac{3+k}{2}\right) \\
&+ \frac{4(1-k)^2}{(3-k)^2} \cdot \left(\frac{x_r}{2} + k(\frac{1+y_l}{2} - \frac{x_r}{2}) + \frac{1-y_l}{2}\right) \\
=& \frac{(1+k)(2-k)}{3-k} + \frac{2(1-k)(3-k)x_r - 2(1-k)(3-k)y_l - (1+k)^2 x_l}{(3-k)^2} \\
=& \frac{(1+k)(2-k) + 2(1-k)(x_r - y_l)}{3-k} - \frac{(1+k)^2 x_l}{(3-k)^2}.
\end{aligned}$$

Then, the ratio between the expected maximum cost and the optimal maximum cost is

$$2 \cdot \frac{\frac{(1+k)(2-k)+2(1-k)(x_r-y_l)}{3-k} - \frac{(1+k)^2 x_l}{(3-k)^2}}{1 + (1-k)x_r + k(1-2x_l) - (1-k)y_l}. \tag{12}$$

Considering $x_r - y_l$ as a variable of the function in (12), the derivative with respect to this variable is always non-negative for any $k \in [0, 1)$, which implies that the maximum possible value is achieved when $x_r = y_l$. Then the ratio becomes

$$2 \cdot \frac{\frac{(1+k)(2-k)}{3-k} - \frac{(1+k)^2 x_l}{(3-k)^2}}{1 + k - 2kx_l} = \frac{2}{(3-k)^2} \cdot \frac{(1+k)(2-k)(3-k) - (1+k)^2 x_l}{1 + k - 2kx_l}. \tag{13}$$

Let $\kappa = \frac{9-\sqrt{73}}{4} \approx 0.114$ be the root of the equation $\frac{(1+k)(2-k)(3-k)}{1+k} = \frac{(1+k)^2}{2k}$.

- When $k \in [0, \kappa]$, we have $\frac{(1+k)(2-k)(3-k)}{1+k} \leq \frac{(1+k)^2}{2k}$, the maximum value of the ratio in (13) is achieved when $x_l = 0$, that is
$$\frac{2}{(3-k)^2} \cdot \frac{(1+k)(2-k)(3-k)}{1+k} = \frac{4-2k}{3-k}. \tag{14}$$
- When $k \in [\kappa, 1]$, we have $\frac{(1+k)(2-k)(3-k)}{1+k} \geq \frac{(1+k)^2}{2k}$, and the ratio in (13) is increasing with x_l. Since $1 - y_r \geq x_l$, x_l is upper bounded by $\frac{1}{2}$. Letting $x_l = \frac{1}{2}$, the maximum value of the ratio in (13) is
$$\frac{2(1+k)(2-k)(3-k) - (1+k)^2}{(3-k)^2} = \frac{11 + 2k^3 - 9k^2}{9 + k^2 - 6k}. \tag{15}$$

□

The maximum possible value of the approximation ratio over all $k \in [0, 1)$ is $9 - 6\sqrt[3]{2} \approx 1.441$, which is attained by $k = 3 - 2\sqrt[3]{2}$. Hence, generally we can say that Mechanism 5 is 1.441-approximation for any $k \in [0, 1)$, and in particular, it is $\frac{4}{3}$-approximation when $k = 0$, and nearly optimal when k approaches 1. Compared with the approximation ratio $\frac{2}{1+k}$ of the deterministic TwoInnerExtreme, this randomized one improves when $k \leq 0.396$, but is worse for any larger k.