
Tight Bounds On The Distortion of Randomized and Deterministic Distributed Voting

MohammadAli Abam*
abam@sharif.edu

Davoud Kareshki*
davood.kareshki12@sharif.edu

Marzieh Nilipour*
marzie.nilipour@sharif.edu

MohammadHossein Paydar*
mohammad.pay@sharif.edu

Masoud Seddighin†
m.seddighin@teias.institute

Abstract

We study metric distortion in distributed voting, where n voters are partitioned into k groups, each selecting a local representative, and a final winner is chosen from these representatives (or from the entire set of candidates). This setting models systems like U.S. presidential elections, where state-level decisions determine the national outcome. We focus on four cost objectives from Anshelevich *et al.* [1]: avg-avg, avg-max, max-avg, and max-max. We present improved distortion bounds for both deterministic and randomized mechanisms, offering a near-complete characterization of distortion in this model.

For deterministic mechanisms, we reduce the upper bound for avg-max from 11 to 7, establish a tight lower bound of 5 for max-avg (improving on $2 + \sqrt{5}$), and tighten the upper bound for max-max from 5 to 3. For randomized mechanisms, we consider two settings: (i) only the second stage is randomized, and (ii) both stages may be randomized. In case (i), we prove tight bounds: $5 - 2/k$ for avg-avg, 3 for avg-max and max-max, and 5 for max-avg. In case (ii), we show tight bounds of 3 for max-avg and max-max, and nearly tight bounds for avg-avg and avg-max within $[3 - 2/n, 3 - 2/(kn^*)]$ and $[3 - 2/n, 3]$, respectively, where n^* denotes the largest group size.

1 Introduction

In social choice theory, a voting rule is a function that takes agents' preferences over alternatives and selects one as the final outcome. Preferences are usually represented as ranked lists, and the goal is to design a voting rule that best reflects these preferences.

How can we evaluate whether a voting rule is appropriate? There are both *axiomatic* and *quantitative* benchmarks for assessing outcomes [2, 3, 4]. In this paper, we focus on one of the most prominent quantitative measures: *distortion*. The idea is simple: each agent has hidden numerical values—either *costs* or *utilities*—for the alternatives, and their ordinal rankings reflect these values. Suppose our ultimate goal is to optimize an objective function ϕ , such as social cost, maximum cost, or total utility, based on these hidden values. However, since the voting rule only has access to the agents' ordinal preferences—not the numerical values—it may select a suboptimal outcome with respect to ϕ . The

*Computer Engineering Department, Sharif University of Technology, Tehran, Iran

†Computer Science Department, Tehran Institute for Advanced Studies (TeIAS), Tehran, Iran

Cost objective	General metric	Line metric
avg-avg	$[7, 11]$ [1]	7 [1]
avg-max	$[2 + \sqrt{5}, 11]$ [1]	$2 + \sqrt{5}$ [25]
max-avg	$[2 + \sqrt{5}, 5]$ [1]	$2 + \sqrt{5}$ [25]
max-max	$[3, 5]$ [1]	3 [1]

Table 1: Known results for deterministic distributed mechanisms under various cost objectives.

distortion of a voting rule captures how far its chosen outcome can be from the optimal one in the worst case. It is defined as the ratio between the value of ϕ for the selected outcome and the value of ϕ for the optimal alternative in the worst case.

Since its introduction by Procaccia and Rosenschein [4], distortion has been a consistent focus of investigation—not only in voting, but also in related social choice problems such as facility location [5, 6, 7, 8] and matching [7, 9, 10, 11]. Still, the core of the literature lies in voting, with particularly rich results when costs of agents form a metric space [10, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21]. For a comprehensive overview of distortion in voting, we refer to the survey by Anshelevich *et al.* [22].

In this paper, we study the distortion in the metric setting when the voting process is distributed. Unlike centralized voting, in many large-scale scenarios, outcomes emerge via a two-stage manner: decisions are made locally within separate groups of agents, the local outcomes are then aggregated into a final outcome. A notable example is the U.S. presidential election, where each state selects a winner, and the national outcome is determined by a weighted aggregation of the state-level results. More formally, a distributed voting mechanism is a pair (f_{in}, f_{ov}) , where

1. f_{in} is an *in-group* voting rule that selects a local winner for each group based solely on the preferences of agents within that group.
2. Assuming R is the set of local winners, f_{ov} is an *over-group* voting rule that selects the final winner based on the preferences of R over all alternatives [1] or local winners [23].

The study of distortion in distributed voting was pioneered by Filos-Ratsikas *et al.* [24], who extended the notion of distortion to the utility-based distributed scenario. Later, Anshelevich *et al.* [1] investigated distributed voting in the metric cost setting. In the context of distributed voting, since decisions occur in two stages, it is natural to define separate cost objectives for each level. Building on this, Anshelevich *et al.* [1] introduced four standard objectives combining *average* and *maximum* costs within and across the groups: avg-avg, max-avg, avg-max, and max-max. In the deterministic setting, they proved constant upper and lower bounds for all objectives in both general and line metric spaces, summarized in Table 1. Later, Voudouris [25] focused on the line metric and proposed two simple mechanisms for the avg-max and max-avg objectives. These mechanisms achieve an upper bound of $2 + \sqrt{5}$, closing the corresponding gap derived by [1].

As shown in Table 1, distributed voting on the line metric is well-understood, with tight distortion bounds already achieved. We therefore turn to general metric spaces and explore whether randomization can also improve distortion in the distributed setting. This work presents the first formal investigation into randomized distributed mechanisms within the metric setting.

In this paper, we make significant progress on the distortion of distributed voting mechanisms in two main directions. First, we improve the existing distortion bounds of *deterministic* mechanisms with respect to the avg-max, max-avg, and max-max objectives. Second, we explore rules that incorporate *randomized* mechanisms—either in the second stage only, or in both stages—referred to as *rand-det* and *rand-rand*, respectively. The output of a randomized mechanism is a probability distribution over the alternatives, rather than a single winner. For both the *rand-det* and *rand-rand* mechanisms, we prove tight bounds for almost all of the objectives.

1.1 Further Related Work

The most relevant studies to our work [1, 25] are discussed in Section 1. Here, we briefly review other related studies. Since Procaccia and Rosenschein’s seminal work [4], research on distortion in

social choice problems has expanded, covering utilitarian settings [26, 27, 28, 29, 30], metric settings [12, 14, 31, 32, 16, 17, 33, 20], and combined approaches [34].

In deterministic case, Anshelevich *et al.* [12] pioneered the study of distortion for the metric framework. Using a simple example, they show that the distortion of any deterministic voting rule is at least 3. Gkatzelis *et al.* [14] proposed an elegant and intricate voting rule, *Plurality Matching*, which achieves a tight distortion of 3. Next, Kizilkaya and Kempe [32] attained the same upper bound with a simpler voting rule, *Plurality Veto*. Filos-Ratsikas *et al.* [24] pioneered distortion analysis in distributed voting under the utilitarian framework. Their work extended to other social choice problems, including facility location [35], aiming to select a single location from a set of alternatives. More recently, Voudouris [36] investigated the distributed distortion in obnoxious voting, where alternatives are undesirable.

Unlike deterministic voting rules, randomized rules can achieve distortion below 3. Anshelevich and Postl [15] proved that the metric distortion of *Random Dictatorship* is at most $3 - 2/n$, with n agents, and establish a lower bound of 2 for any randomized voting rule. Kempe [37] improved the upper bound for *Random Dictatorship* to $3 - 2/m$, where m is the number of candidates. Charikar and Ramakrishnan [16] further raised the lower bound for any randomized rule to 2.112. Recently, Charikar *et al.* [33] reduced the upper bound to 2.753. In the context of distributed voting, Filos-Ratsikas and Voudouris [23] investigated randomized mechanisms under the utilitarian framework, establishing distortion bounds in various cases.

1.2 Our Contributions

Our results provide improved upper and lower bounds on the distortion of distributed mechanisms across various combinations of deterministic and randomized voting rules and cost objectives. As summarized in Table 2, most of our bounds are tight—despite the fact that our proposed mechanisms are simple. In addition to general metric spaces, we also analyze the Euclidean setting and derive corresponding bounds under this restriction.

Randomized Mechanisms. Previous work on metric distortion in the distributed setting has focused exclusively on deterministic voting rules [1, 25, 38, 36]. In this paper, we take a significant step toward understanding randomized mechanisms in distributed voting. We study two natural classes of randomized mechanisms—*rand-det* and *rand-rand*—within general metric spaces, and analyze their performance with respect to the four objectives. See Table 2 for an overview of our results.

rand-det mechanisms, defined as pairs (f_{in}, f_{ov}) , where f_{in} is a deterministic voting rule and f_{ov} is a randomized one. We derive tight distortion bounds with respect to the all objectives in Section 3.

- **max-max, avg-max:** For both objectives, we derive a tight distortion bound of 3. The lower bound is established through a basic example within a single group on a line metric, simplifying the max-max and avg-max objectives to max. The upper bound is proven by a distributed mechanism that first selects a representative for each group with the *Plurality Matching* rule and then chooses the final winner uniformly at random.
- **max-avg:** We establish a tight distortion bound of 5. The lower bound is proven using a line metric and a novel tool we introduce, called the *Bias Tournament*, which may be of independent interest. For the upper bound, we show that applying a deterministic in-group rule with a distortion at most $\alpha \geq 3$, followed by the *Random Dictatorship* rule, achieves an overall distortion of at most $\alpha + 2$. Since the best achievable value of α is 3 (via the *Plurality Matching* rule), this yields a matching upper bound of 5.
- **avg-avg:** We prove a tight distortion bound of $5 - 2/k$. Obtaining this bound for the avg-avg objective is the most challenging aspect of the *rand-det* mechanisms. The lower bound construction, though similar to that of the max-avg objective, requires a more delicate analysis to extract the $2/k$ improvement. Once again, we employ the *Bias Tournament* and model the metric space via shortest-path distances in a graph.

In Section 4, we analyze *rand-rand* mechanisms, defined as pairs (f_{in}, f_{ov}) comprising of two randomized voting rules, and derive tight or near-tight distortion bounds. All the upper bounds are obtained via a distributed mechanism that initially applies the *Random Dictatorship* rule within

	Objective	Distortion	
		lower bound	upper bound
det-det	avg-avg	7 [1]	11 [1]
	avg-max	$2 + \sqrt{5}$ [1]	7 (Corollary 5.3)
	max-avg	5 (Theorem 5.4)	5 [1]
	max-max	3 [1]	3 (Theorem 5.1)
rand-det	avg-avg	$5 - \frac{2}{k}$ (Theorem 3.8)	$5 - \frac{2}{k}$ (Corollary 3.4)
	avg-max	3 (Theorem 3.6)	3 (Theorem 3.5)
	max-avg	5 (Theorem 3.7)	5 (Corollary 3.2)
	max-max	3 (Theorem 3.6)	3 (Theorem 3.5)
rand-rand	avg-avg	$3 - \frac{2}{n}$ (Theorem 4.9)	$3 - \frac{2}{kn^*}$ (Theorem 4.4)
	avg-max	$3 - \frac{2}{n}$ (Theorem 4.7)	3 (Theorem 4.2)
	max-avg	3 (Theorem 4.6)	3 (Theorem 4.3)
	max-max	3 (Theorem 4.6)	3 (Theorem 4.1)

Table 2: An overview of our results for various cost objectives in general metric spaces, with gray-text results indicating those derived from prior work. n^* denotes the size of the largest group. Thus, the bound of $3 - 2/n$ for the avg-avg objective in rand-rand is tight when all group sizes are equal.

each group and then randomly selects the final winner from the chosen representatives with uniform probability.

- **max-max, max-avg:** For both objectives, we establish a tight distortion bound of 3. We construct a shared example consisting of k single-voter groups to establish the lower bound, even when the metric space is a line. In this scenario, the max-max and max-avg objectives both simplify to the max objective.
- **avg-max:** We establish a lower bound of $3 - \frac{2}{n}$, which nearly matches our upper bound of 3. The lower bound is proven with an instance where the number of candidates and voters are equal ($n = m$) and there is only a single group ($k = 1$). Additionally, we conclude a lower bound for the max objective in the centralized setting: we show that any randomized rule must have a distortion of at least $3 - \varepsilon$ for any constant $\varepsilon > 0$. This is particularly interesting since even deterministic rules are known to have an upper bound of 3 for max [14].
- **avg-avg:** We establish a nearly tight distortion bound slightly below 3. For an instance with k single-voter groups on a tree graph, we prove lower bound of $3 - \frac{2}{n}$. We further derive an upper bound of $3 - \frac{2}{kn^*}$, where n^* denotes the largest group size. When all groups are of equal size, it yields matching upper and lower bounds. Notably, deriving these bounds is the most challenging aspect of analyzing rand-rand, mechanisms.

Deterministic Mechanisms. We consider det-det mechanisms, defined as pairs (f_{in}, f_{ov}) comprising of two independently deterministic voting rules, in Section 5. We resolve the previously known gaps for the max-avg and max-max objectives and provide an enhanced upper bound for the avg-max objective. In this section, we adopt a setting akin to [1], where f_{ov} selects a winner from the set of *all candidates*, not solely those chosen in the first stage.

- **avg-max.** We improve the upper bound from **11** to **7**. Anshelevich *et al.* [1] show that combining the in-group and over-group voting rules with distortions α and β , respectively, yields an overall distortion of $\alpha + \beta + \alpha\beta$. With their best known values ($\alpha = 3, \beta = 2$), this gives 11. We prove that if the in-group rule, f_{in} , merely satisfies the property of *Pareto efficiency*, then the overall distortion is at most $2\beta + 3$, which is *independent* of α . This results in a tighter upper bound of 7, and shows the dominant role of the over-group rule.
- **max-avg.** We improve the lower bound from $2 + \sqrt{5}$ to **5**. Interestingly, our lower-bound instance is based on a metric constructed via shortest-path in a graph, rather than a line or Euclidean. This confirms that the upper bound from [1] is indeed tight.
- **max-max.** We improve the upper bound on distortion from **5** to **3** for general metric spaces. Although a bound of 3 was previously known for the *line metric*, the general case remained open. We show that a distributed mechanism same as the *Arbitrary Dictator*, proposed by [1], actually achieves distortion 3 for any metric space.

Bias Tournament. The Bias Tournament is a directed graph with one node per candidate. For any pair of candidates c_1 and c_2 , we add a directed edge from c_1 to c_2 if, in a group containing only two voters with preferences (c_1, c_2, \dots) and (c_2, c_1, \dots) —where the remaining candidates are ordered identically according to a fixed permutation σ over all candidates—the in-group rule f_{in} deterministically selects c_1 as the winner. This construction captures the bias in f_{in} ’s tie-breaking behavior across candidate pairs. To analyze the implications of these biases, we use a well-known property of tournaments: any tournament on m nodes contains at least one node with in-degree at least $\lceil (m-1)/2 \rceil$. This fact allows us to identify a *losing* candidate—one frequently defeated in pairwise comparisons—and use it as the basis for constructing an instance with high distortion, which helps us to establish the lower bounds for the rand-det and det-det mechanisms.

Centralized Setting. Interestingly, one of our lower bound constructions, originally developed for distributed mechanisms also applies to randomized centralized setting. In particular, we analyze the max cost objective in this setting. Previously, Gkatzelis *et al.* [14] established an upper bound of 3 for max using *Plurality Matching*, as defined in Section 2. In our work, we revisit the avg-max objective for the rand-rand mechanisms and prove a distortion lower bound. We then reuse the same construction to show that the distortion of max under centralized voting is at least $3 - \varepsilon$ for any constant $\varepsilon > 0$, thus matching the known upper bound up to an arbitrarily small gap.

Euclidean Space. Our lower-bound constructions for the avg-avg objective rely on general metric spaces that do not reduce to simpler structures such as the line or Euclidean space. This raises the natural question of whether similarly strong bounds can be achieved in more structured settings. In Section 6, we address this by constructing novel instances within a Euclidean hyper-simplex, obtaining lower bounds of $\sqrt{5} - \varepsilon$ for the rand-rand mechanisms and $2 + \sqrt{5} - \varepsilon$ for the rand-det ones, where ε is arbitrarily small. Specifically, for rand-rand, we fix an integer l and construct an instance with $l + 2$ candidates and $k = l + 1$ single-voter groups embedded in \mathbb{R}^{l+1} . Similarly, for rand-det, we construct an instance with $2m$ candidates and m groups of two voters each, embedded in \mathbb{R}^{m+1} (where m is fixed).

2 Basic Notations

An instance of the *distributed voting problem* is a tuple $(\mathcal{V}, \mathcal{C}, \mathcal{G}, \pi, \mathbf{d})$. Here, \mathcal{V} is a set of n voters, and \mathcal{C} is a set of m candidates. We use v_i to denote the i -th voter and symbols a, b, c, \mathbf{o} , and \mathbf{w} to refer to candidates. In particular, \mathbf{w} typically denotes the winner and \mathbf{o} refers to the optimal candidate. \mathcal{G} is a partition of voters into k groups, such that each voter belongs to exactly one group. For each group $g \in \mathcal{G}$, n_g is the number of voters in g . We denote the optimal candidate for group g by \mathbf{o}_g . We say an instance is *symmetric*, if all groups have equal sizes. Each voter v_i has a strict ranking π_i over the set of candidates, representing their ordinal preferences. A preference profile is the collection of preferences from all voters, denoted by $\pi = (\pi_1, \dots, \pi_n)$. These preferences arise from a cost function based on the underlying metric space \mathbf{d} . For each voter v and candidate c , $\mathbf{d}(v, c)$ denotes the cost of candidate c for voter v . The voters rank all candidates in increasing order of cost, which means they prefer those who are closer. The distance function \mathbf{d} satisfies the standard metric properties of *non-negativity*, *symmetry*, and the *triangle inequality*.

We denote the top-ranked candidate of each voter v as $\text{top}(v)$. For a candidate c and a group $g \in \mathcal{G}$, $v^*(c, g) = \arg \max_{v \in g} \mathbf{d}(v, c)$ denotes the farthest voter from c within group g . Similarly, $v^{**}(c) = \arg \max_{v \in \mathcal{V}} \mathbf{d}(v, c)$ represents the farthest voter from c across all voters. We also use $\text{cost}_g(c)$ to denote the cost of candidate c restricted to group g , and define $g^*(c) = \arg \max_{g \in \mathcal{G}} (1/n_g) \sum_{v \in g} \mathbf{d}(v, c)$ that is, the group in which candidate c incurs the highest average cost.

Given an instance \mathcal{I} , various cost objectives can be considered to evaluate the final winner. In the distributed voting, we can be even more flexible by applying different objectives at each stage. Following [1], we consider four cost objectives:

- **Average of averages (avg-avg):** Average the costs within each group, then average across all groups; $\text{avg-avg}(c \mid \mathcal{I}) = \frac{1}{k} \sum_{g \in \mathcal{G}} \frac{1}{n_g} \sum_{v \in g} \mathbf{d}(v, c) = \frac{1}{k} \sum_{g \in \mathcal{G}} \text{cost}_g(c)$.
- **Average of maxima (avg-max):** Find the most dissatisfied voter in each group, then average their costs; $\text{avg-max}(c \mid \mathcal{I}) = \frac{1}{k} \sum_{g \in \mathcal{G}} \max_{v \in g} \mathbf{d}(v, c) = \frac{1}{k} \sum_{g \in \mathcal{G}} \mathbf{d}(v^*(c, g), c)$.

- **Maximum of averages (max-avg):** Compute the average cost in each group; return the worst among them; $\text{max-avg}(c \mid \mathcal{I}) = \max_{g \in \mathcal{G}} (\frac{1}{n_g} \sum_{v \in g} d(v, c)) = \text{cost}_{g^*(c)}(c)$.
- **Maximum of maxima (max-max):** Return the cost of the most dissatisfied voter overall; $\text{max-max}(c \mid \mathcal{I}) = \max_{g \in \mathcal{G}} \max_{v \in g} d(v, c) = d(v^{**}(c), c)$.

For simplicity, when the objective is clear from context, we simply write $\text{cost}(c \mid \mathcal{I})$. When the instance is also clear, we omit it entirely.

A voting rule f maps a preference profile π to a winning candidate \mathbf{w} . This rule may be *deterministic* or involve *randomization*. Next, we define a distributed voting mechanism $\Psi = (f_{in}, f_{ov})$, which consists of two stages:

- **Stage 1:** Each group g independently selects a *representative* candidate \mathbf{w}_g by applying the in-group rule f_{in} to the preferences of its members. Let $R = \{\mathbf{w}_g \mid g \in \mathcal{G}\}$.
- **Stage 2:** The final outcome is chosen by applying the over-group rule f_{ov} to the (centralized) instance $(R, R, \pi^R, \mathbf{d})$, where R acts as *both the set of candidates and the set of voters*, and π^R denotes the preferences of representatives over one another.

For each group g , the rule f_{in} has access only to local information: the group size (n_g) and the preference profile of its voters (π^g) over all candidates. In contrast, f_{ov} receives the preferences of the selected representatives together with the sizes of all groups.

Distortion. Given a distributed mechanism $\Psi = (f_{in}, f_{ov})$, which takes an instance $\mathcal{I} = (\mathcal{V}, \mathcal{C}, \mathcal{G}, \pi, \mathbf{d})$ as input and outputs a winner $\mathbf{w} = \Psi(\mathcal{I})$, the expected cost of \mathbf{w} is defined as $E[\text{cost}(\mathbf{w} \mid \mathcal{I})] = \sum_{g \in \mathcal{G}} \Pr(\mathbf{w} = \mathbf{w}_g) \cdot E[\text{cost}(\mathbf{w}_g \mid \mathcal{I})]$, and the expected cost of \mathbf{w}_g is defined as $E[\text{cost}(\mathbf{w}_g \mid \mathcal{I})] = \sum_{c \in \mathcal{C}} \Pr(\mathbf{w}_g = c) \cdot \text{cost}(c \mid \mathcal{I})$, where $\text{cost}(\cdot)$ denotes one of the four objectives defined earlier. Now, we define the distortion of Ψ as

$$D(\Psi) = \sup_{\mathcal{I}} \frac{E[\text{cost}(\Psi(\mathcal{I}) \mid \mathcal{I})]}{\min_{c \in \mathcal{C}} \text{cost}(c \mid \mathcal{I})}.$$

Definition 2.1 (Pareto Efficiency). *A voting rule is Pareto efficient if, for any pair of candidates x and y , if all voters prefer x to y , then the rule does not select y as the winner.*

We now present several voting rules (for centralized settings) and mechanisms (for distributed settings) frequently used in this work:

- **Plurality Matching rule (f_{pm}):** Introduced by [14], Plurality Matching is a deterministic voting rule that guarantees a distortion of 3, the best possible among all deterministic rules in the metric setting. Given an instance \mathcal{I} with voters \mathcal{V} , candidates \mathcal{C} , and preference profile σ , f_{pm} considers each candidate c and constructs a bipartite domination graph $G(c)$ with voters on both sides and an edge (v_i, v_j) if and only if voter v_i prefers c to top-ranked candidate of voter v_j (i.e., $c \succeq_{v_i} \text{top}(v_j)$). The rule then selects any candidate c for which its corresponding graph $G(c)$ admits a perfect matching (which is guaranteed to exist).
- **Pareto-efficient Plurality Matching rule (f_{pm-par}):** We now introduce a variant of f_{pm} that ensures Pareto efficiency while preserving the optimal distortion of 3. f_{pm-par} operates as follows: it starts with a candidate c_1 chosen by f_{pm} (one whose corresponding domination graph $G(c_1)$ admits a perfect matching M) and iteratively refines this choice;
 - If c_1 is Pareto efficient—no other candidate c_2 is preferred by all voters over c_1 —then c_1 is returned as the final winner.
 - Otherwise, there must exist a candidate c_2 that Pareto dominates c_1 (i.e., every voter prefers c_2 to c_1). A key insight is that if c_2 Pareto dominates c_1 , then $G(c_2)$ is also admits the same perfect matching M as $G(c_1)$.
 - The voting rule replaces c_1 with c_2 and iterates until a candidate satisfying Pareto efficiency is found.
- **Random Dictatorship rule (f_{rd}):** A randomized voting rule in which a voter is selected uniformly at random, and the outcome is that voter's top-ranked candidate [15, 37]. In the first stage of the **rand-rand** mechanism used to establish upper bounds, we apply this rule.

- **Uniform selection rule** (f_{ur}): A randomized voting rule that selects the winner uniformly at random from the set of candidates. This rule is used in the second stage of both the **rand-det** and **rand-rand** mechanisms that establish upper bounds.
- **Arbitrary Dictator mechanism** (m_{ad}): Introduced by [1], this mechanism operates in two stages. First, each group selects a representative by arbitrarily choosing a voter and taking her top-ranked candidate. Second, the final winner is determined by arbitrarily selecting one of the representatives. This mechanism is employed to analyze the upper bound of the max-max objective in **det-det**.
- **α -in- β -over mechanism** ($m_{\alpha\beta}$): Proposed by [1], the α -in- β -over mechanism operates in two deterministic stages, first applying an in-group voting rule with distortion at most α , followed by selecting a final winner using an over-group rule with distortion at most β .

Definition 2.2 (Promotion). *Given an order σ over a set of candidates and a candidate $c \in \sigma$, the operation $\sigma \uparrow c$ returns a new preference σ' in which c is moved to the top, and the relative order of all other candidates remains unchanged. When multiple \uparrow operations are applied in sequence, they are evaluated from left to right. $\sigma \uparrow b \uparrow a$ first moves b to the top of σ , then moves a to the top of $\sigma \uparrow b$.*

Using the promote operation, we define the Bias Tournament—a special complete directed graph (tournament) over the candidates—which is crucial for establishing lower bounds of the **rand-det** and **det-det** mechanisms.

Definition 2.3 (Bias Tournament). *Let f be a deterministic voting rule, C a set of candidates, and σ an ordering of C . The Bias Tournament $\mathcal{T}(f, C, \sigma)$ is a complete directed graph where each vertex corresponds to a candidate in C . For every pair of distinct candidates u and w , there is a directed edge from u to w if and only if f selects u as the winner in a two-voter election with preferences $\pi_1 = \sigma \uparrow w \uparrow u$ and $\pi_2 = \sigma \uparrow u \uparrow w$.*

Example 1. *Let $C = \{c_1, c_2, c_3\}$ and a deterministic rule f that selects the candidate with the smallest index among those ranked first by at least one voter. Suppose: (i) between c_1 and c_2 , the winner is c_1 ; (ii) between c_1 and c_3 , the winner is c_1 ; (iii) between c_2 and c_3 , the winner is c_2 . Then $\mathcal{T}(f, C, \sigma)$ contains edges $c_1 \rightarrow c_2$, $c_1 \rightarrow c_3$, and $c_2 \rightarrow c_3$.*

Basic Observations. We present several preliminary observations that lay the groundwork for proving our main theorems. For clarity and consistency, we fix an instance $\mathcal{I} = (\mathcal{V}, \mathcal{C}, \mathcal{G}, \pi, \mathbf{d})$ throughout this section. The proof of Observation 2.6 is deferred to the appendix.

Observation 2.1. *In distributed voting with single-voter groups, let $\Psi = (f_{in}, f_{ov})$ be a distributed mechanism with finite distortion. Within each group, f_{in} must select the top-ranked candidate of each voter as the group representative.*

Observation 2.2. *For any deterministic voting rule f and an ordering σ over C , there exists a candidate with in-degree at least $\lceil \frac{m-1}{2} \rceil$ in $\mathcal{T}(f, C, \sigma)$.*

Observation 2.3. *Since \mathbf{o}_g is the optimal candidate in group g , we have $\text{cost}_g(\mathbf{o}_g) \leq \text{cost}_g(c)$ for any candidate c , including \mathbf{o} . This holds for all objectives considered.*

Observation 2.4. *Since $\text{cost}(\cdot)$ is defined as the maximum over $\text{cost}_g(\cdot)$ under the max-avg and max-max objectives, it follows that $\text{cost}_g(\mathbf{o}) \leq \text{cost}(\mathbf{o})$, for each group g .*

Observation 2.5. *For **rand-det** mechanism $\Psi = (f_{in}, f_{ur})$ with output \mathbf{w} , the expected cost of the mechanism is given by $\mathbb{E}[\text{cost}(\mathbf{w})] = \frac{1}{k} \sum_{g \in \mathcal{G}} \text{cost}(\mathbf{w}_g)$.*

Observation 2.6. *For **rand-rand** mechanism $\Psi = (f_{rd}, f_{ur})$ with output \mathbf{w} , the expected cost of the mechanism is given by $\mathbb{E}[\text{cost}(\mathbf{w})] = \frac{1}{k} \sum_{g \in \mathcal{G}} \frac{1}{n_g} \sum_{v \in g} \text{cost}(\text{top}(v))$.*

Observation 2.7. *For the max-avg objective and any group g , we have $\text{cost}_g(\mathbf{o}) \leq \text{cost}(\mathbf{o})$, as implied directly by the definition of max-avg.*

Observation 2.8. *Since $\text{top}(v)$ denotes the candidate closest to voter v , it follows that $d(v, \text{top}(v)) \leq d(v, c)$ for any candidate c .*

Observation 2.9. *For every voter v and every candidate c , we have $d(v, c) \leq d(v^{**}(c), c)$.*

Observation 2.10. *For every group g , every voter $v \in g$, and every candidate c , we have $d(v, c) \leq d(v^*(c, g), c)$.*

Observation 2.11. *Consider a distributed mechanism $\Psi = (f_{in}, f_{ov})$, where f_{in} is a deterministic rule with distortion at most α . By the definition of centralized distortion, for any group $g \in \mathcal{G}$, we obtain $\text{cost}_g(\mathbf{w}_g) \leq \alpha \cdot \text{cost}_g(\mathbf{o}_g)$.*

Observation 2.12. Consider a *det-det* mechanism $\Psi = (f_{in}, f_{ov})$, where f_{ov} has distortion at most β with respect to the *avg* objective. By the definition of centralized distortion, for any group $g \in \mathcal{G}$, we have $\frac{1}{k} \sum_{g \in \mathcal{G}} d(w, w_g) \leq \beta \cdot \frac{1}{k} \sum_{g \in \mathcal{G}} d(o, w_g)$.

3 Distortion Bounds for rand-det

This section examines *rand-det* mechanisms, defined as pairs (f_{in}, f_{ov}) , where f_{in} is a deterministic voting rule and f_{ov} is a randomized one. We establish distortion bounds under all considered objectives. For brevity, all proofs are deferred to the appendix.

3.1 Upper Bounds

Let f_α be a deterministic voting rule with distortion at most α , and let f_{par} be any deterministic voting rule that satisfies Pareto efficiency. We analyze the mechanisms (f_α, f_{ur}) and (f_{par}, f_{ur}) .

For the *max-avg* and *avg-avg* objectives, we show that the mechanism (f_α, f_{ur}) achieves distortion at most $\alpha + 2$ and $\alpha + 2 - 2/k$, respectively. Since f_{pm-par} achieves the best-known distortion of 3 and also satisfies Pareto efficiency, we instantiate our theorems with the mechanism (f_{pm-par}, f_{ur}) to obtain the tightest bounds.

Finally, for the *avg-max* and *max-max* objectives, we prove that the mechanism (f_{par}, f_{ur}) achieves distortion at most 3, a consequence of Pareto efficiency.

Theorem 3.1. For the *max-avg* objective in general metric spaces, we have $D((f_\alpha, f_{ur})) \leq \alpha + 2$.

By using f_{pm} instead of f_α as the in-group voting rule (with $\alpha = 3$) and applying Theorem 3.1, we conclude that $\Psi = (f_{pm}, f_{ur})$ is a *rand-det* mechanism that satisfies the bound stated in Corollary 3.2.

Corollary 3.2 (of theorem 3.1). For the *max-avg* objective in general metric spaces, there exists a *rand-det* mechanism with distortion at most 5.

Theorem 3.3. For the *avg-avg* objective in general metric spaces, we have $D((f_\alpha, f_{ur})) \leq \alpha + 2 - \frac{2}{k}$.

Once again, by using f_{pm} instead of f_α as the in-group voting rule (with $\alpha = 3$) and applying Theorem 3.3, we conclude that $\Psi = (f_{pm}, f_{ur})$ is a *rand-det* mechanism that satisfies the bound stated in Corollary 3.4.

Corollary 3.4 (of theorem 3.3). For the *avg-avg* objective in general metric spaces, there exists a *rand-det* mechanism with distortion at most $5 - 2/k$.

For the *avg-max* and *max-max* objectives, we derive an upper bound of 5 following an argument analogous to the proof of Theorem 3.1. Now, we improve the upper bound of 5 to 3 by applying the property of Pareto efficiency.

Theorem 3.5. For the *avg-max* and *max-max* objectives in general metric spaces, we have $D((f_{par}, f_{ur})) \leq 3$.

3.2 Lower Bounds

Now, we establish lower bounds on the distortion of *rand-det* mechanisms. Specifically, Theorem 3.6 provides lower bounds for the *max-max* and *avg-max* objectives, Theorem 3.7 covers the *max-avg* objective, and Theorem 3.8 addresses the *avg-avg* objective. It is worth noting that all the lower bounds in this section are derived from symmetric instances and apply in that setting as well.

Theorem 3.6. For the *avg-max* and *max-max* objectives, the distortion of any *rand-det* mechanism is at least 3, even when the metric space is a line.

For the *max-avg*, and *avg-avg*, objectives, we use the Bias Tournament to establish the lower bounds stated in Theorems 3.7 and 3.8.

Theorem 3.7. For the *max-avg* objective, the distortion of any *rand-det* mechanism is at least 5, even when the metric space is a line.

We now establish the lower bound for the *avg-avg* objective in general metric spaces.

Theorem 3.8. For general metric spaces and the *avg-avg* objective, the distortion of any *rand-det* mechanism is at least $5 - \frac{2}{k}$.

4 Distortion Bounds for rand-rand

This section examines the rand-rand mechanisms, which are pairs (f_{in}, f_{ov}) composed of two randomized voting rules f_{in} and f_{ov} . We establish distortion bounds under all considered objectives.

4.1 Upper Bounds

Throughout this section, we analyze the mechanism (f_{rd}, f_{ur}) and show that despite its simplicity, it achieves tight (or nearly tight) distortion bounds for various cost objectives. In particular, for the max-max objective (indeed the max objective), we establish that choosing the top candidate of any voter yields a distortion of at most 3. We begin with the simplest case, max-max, and move towards the most intricate avg-avg.

Theorem 4.1. *For the max-max objective in general metric spaces, we have $D((f_{rd}, f_{ur})) \leq 3$.*

Theorem 4.2. *For the avg-max objective in general metric spaces, we have $D((f_{rd}, f_{ur})) \leq 3$.*

For the max-avg objective, the key insight is to show that for any voter v , $\text{cost}_{g^*(\text{top}(v))}(\text{top}(v)) \leq 2d(v, o) + \text{cost}(o)$. This crucial inequality is the foundation for proving the desired upper bound.

Theorem 4.3. *For the max-avg objective in general metric spaces, we have $D((f_{rd}, f_{ur})) \leq 3$.*

Theorem 4.4. *For the avg-avg objective in general metric spaces, we have $D((f_{rd}, f_{ur})) \leq 3 - 2/kn^*$ where n^* represents the maximum value of n_g over all groups.*

As a corollary of Theorem 4.4, we conclude that $\Psi = (f_{rd}, f_{ur})$ is a rand-rand mechanism that satisfies the bound stated in Corollary 4.5, particularly in the symmetric case ($kn^* = n$).

Corollary 4.5 (of theorem 4.4). *For the avg-avg objective in general metric spaces, there exists a rand-rand mechanism with distortion at most $3 - \frac{2}{n}$, provided that the groups are symmetric.*

4.2 Lower Bounds

In this section, all results are derived from the symmetric instances and apply to that setting as well.

Theorem 4.6. *For the max-avg and max-max objectives, the distortion of any rand-rand mechanism is at least 3, even when the metric space is a line.*

Theorem 4.7. *For the avg-max objective, the distortion of any rand-rand mechanism is at least $3 - \frac{2}{n}$ (equivalently, $3 - \frac{2}{m}$), even when the metric space is a line.*

In the centralized setting, where all voters belong to a single group ($k = 1$), the avg-max objective simplifies to the max objective. Thus, the instance and analysis from the proof of Theorem 4.7, which already considers the single-group case, apply directly. As the number of voters n increases, the distortion approaches 3. Therefore, for any constant $\varepsilon > 0$, an instance can be constructed with a sufficiently large number of voters ($n > \frac{2}{\varepsilon}$) such that the distortion of any randomized voting rule exceeds $3 - \varepsilon$. This yields the following corollary.

Corollary 4.8 (of Theorem 4.7). *For the max objective in the centralized setting, the distortion of any randomized voting rule is at least $3 - \varepsilon$, for any constant $\varepsilon > 0$, even when the metric is a line.*

Theorem 4.9. *For general metric spaces and the avg-avg objective, the distortion of any rand-rand mechanism is at least $3 - \frac{2}{n}$ (equivalently $3 - \frac{2}{k}$).*

5 Resolving the Distortion Bounds for det-det

In this section, we study the distortion of deterministic distributed mechanisms, providing both lower and upper bounds in general metric spaces. Anshelevich *et al.* [1] proposed the α -in- β -over mechanism $(m_{\alpha\beta})$, which allows any candidate—not just representatives—to be selected as the final winner. We adopt their approach here.

5.1 Upper Bounds

We establish improved upper bounds for deterministic mechanisms with respect to the avg-max and max-max objectives. Let f_β be a deterministic voting rule with distortion at most β , and let f_{par}

be a deterministic voting rule that satisfies the property of Pareto efficiency. We improve the best known upper bound for the avg-max objective from 11, as proved by [1], to 7 by applying mechanism (f_{par}, f_{β}) . For the max-max objective, we improve the previous upper bound of 5 to 3 using m_{ad} . We begin with the simplest case, max-max, and move towards the slightly more intricate avg-max.

Theorem 5.1. *For the max-max objective in general metric spaces, we have $D(m_{ad}) \leq 3$.*

Theorem 5.2. *For the avg-max objective in general metric spaces, we have $D((f_{par}, f_{\beta})) \leq 2\beta + 3$.*

We can apply f_{pm-par} as both the in-group and over-group voting rules. Assuming each voter is at a distance of 0 from her top choice, this yields a distortion of $\beta = 2$, as shown in [14]. Combined with Theorem 5.2, we conclude that $\Psi = (f_{pm-par}, f_{pm-par})$ is a **det-det** mechanism that satisfies the bound in Corollary 5.3.

Corollary 5.3 (of Theorem 5.2). *For the avg-max objective in general metric spaces, there exists a det-det mechanism with distortion at most 7.*

5.2 Lower Bounds

In this section, we present a lower bound of 5 on the distortion of any **det-det** mechanisms under the max-avg objective, improving upon the previous bound of $2 + \sqrt{5}$, which achieved by [1]. Following the framework in that paper, our analysis applies the over-group voting rule across the set of candidates (C), rather than the group representatives (R).

Theorem 5.4. *For general metric spaces and the max-avg objective, the distortion of any det-det mechanism is at least 5.*

6 An Extension of Lower Bounds for rand-rand and rand-det

In this section, we focus on the Euclidean metric and establish lower bounds on the avg-avg distortion for both the rand-rand and rand-det mechanisms, as presented in Theorems 6.1 and 6.2, respectively.

Theorem 6.1. *For the avg-avg objective in Euclidean space, the distortion of any rand-rand mechanism is at least $\sqrt{5} - \varepsilon$, for every constant $\varepsilon > 0$.*

Theorem 6.2. *For the avg-avg objective in Euclidean space, the distortion of any rand-det mechanism is at least $2 + \sqrt{5} - \varepsilon$, for every constant $\varepsilon > 0$.*

7 Discussion and Open Problems

In this paper, we have initiated the study of metric distortion in single-winner distributed voting under randomized mechanisms (rand-rand and rand-det) for many different objectives. We also have improved upon previous results for deterministic mechanisms (det-det). Although our work presents an almost complete picture in the distortion of distributed voting problem, it reveals several promising directions for future research. A significant challenge about our work leaves open lies in analyzing the det-rand mechanisms, where random decisions in the first stage are followed by deterministic ones in the second. Our primary tool, the Bias Tournament technique, is incompatible with the randomized first stage of det-rand. Currently, our understanding is confined to basic results inherited from rand-rand (for the lower bounds) and det-det (for the upper bounds). To obtain tighter results, we need more refined arguments, a promising direction for future work.

Within the scope of our work, another possible direction could be to close the remaining narrow gaps between the lower and upper bounds presented in Table 2, particularly for the avg-avg and avg-max objectives in det-det, as well as the avg-avg and avg-max objective in rand-rand. Another direction is to examine the distortion of the avg-avg objective for all proposed mechanisms, in more structured spaces—Euclidean and line metrics—where results may differ from those in general metric spaces. Another natural extension is to investigate distributed mechanisms in a cardinal setting, where agents have access to exact distances, instead of solely the ordinal rankings induced by those distances.

Going beyond the single-winner voting, one could study the distortion of distributed mechanisms that select committees comprising a specified number of alternatives. Another intriguing direction for future research is to investigate how agents' strategic behavior impacts distributed distortion. The goal could be to understand whether it is possible to design distributed mechanisms that are both strategyproof and capable of achieving low distortion.

References

- [1] Elliot Anshelevich, Aris Filos-Ratsikas, and Alexandros A Voudouris. The distortion of distributed metric social choice. *Artificial Intelligence*, 308:103713, 2022.
- [2] Kenneth J Arrow, Amartya Sen, and Kotaro Suzumura. *Handbook of social choice and welfare*. 2010.
- [3] Felix Brandt, Vincent Conitzer, Ulle Endriss, Jérôme Lang, and Ariel D Procaccia. *Handbook of computational social choice*. 2016.
- [4] Ariel D Procaccia and Jeffrey S Rosenschein. The distortion of cardinal preferences in voting. In *Proceedings of International Workshop on Cooperative Information Agents (CIA)*, pages 317–331, 2006.
- [5] Michal Feldman, Amos Fiat, and Iddan Golomb. On voting and facility location. In *Proceedings of ACM Conference on Economics and Computation (EC)*, pages 269–286, 2016.
- [6] Hau Chan, Aris Filos-Ratsikas, Bo Li, Minming Li, and Chenhao Wang. Mechanism design for facility location problems: A survey. In *Proceedings of International Joint Conference on Artificial Intelligence (IJCAI)*, pages 4356–4365, 2021.
- [7] Elliot Anshelevich and Wennan Zhu. Ordinal approximation for social choice, matching, and facility location problems given candidate positions. *ACM Transactions on Economics and Computation*, 9(2):1–24, 2021.
- [8] Panagiotis Kanellopoulos, Alexandros A Voudouris, and Rongsen Zhang. Truthful two-facility location with candidate locations. In *Proceedings of International Symposium on Algorithmic Game Theory (SAGT)*, pages 365–382, 2023.
- [9] Georgios Amanatidis, Georgios Birmas, Aris Filos-Ratsikas, and Alexandros A Voudouris. A few queries go a long way: Information-distortion tradeoffs in matching. *Journal of Artificial Intelligence Research*, 74:227–261, 2022.
- [10] Nima Anari, Moses Charikar, and Prasanna Ramakrishnan. Distortion in metric matching with ordinal preferences. In *Proceedings of ACM Conference on Economics and Computation (EC)*, pages 90–110, 2023.
- [11] Mohamad Latifian and Alexandros A Voudouris. The distortion of threshold approval matching. In *Proceedings of International Joint Conference on Artificial Intelligence (IJCAI)*, pages 2851–2859, 2024.
- [12] Elliot Anshelevich, Onkar Bhardwaj, Edith Elkind, John Postl, and Piotr Skowron. Approximating optimal social choice under metric preferences. *Artificial Intelligence*, 264:27–51, 2018.
- [13] Kamesh Munagala and Kangning Wang. Improved metric distortion for deterministic social choice rules. In *Proceedings of the 2019 ACM Conference on Economics and Computation (EC)*, pages 245–262, 2019.
- [14] Vasilis Gkatzelis, Daniel Halpern, and Nisarg Shah. Resolving the optimal metric distortion conjecture. In *Proceedings of Symposium on Foundations of Computer Science (FOCS)*, pages 1427–1438, 2020.
- [15] Elliot Anshelevich and John Postl. Randomized social choice functions under metric preferences. *Journal of Artificial Intelligence Research*, 58:797–827, 2017.
- [16] Moses Charikar and Prasanna Ramakrishnan. Metric distortion bounds for randomized social choice. In *Proceedings of Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 2986–3004, 2022.
- [17] Ioannis Caragiannis, Nisarg Shah, and Alexandros A Voudouris. The metric distortion of multiwinner voting. *Artificial Intelligence*, 313:103802, 2022.
- [18] Ioannis Anagnostides, Dimitris Fotakis, and Panagiotis Patsilinakos. Metric-distortion bounds under limited information. *Journal of Artificial Intelligence Research*, 74:1449–1483, 2022.
- [19] Fatih Erdem Kizilkaya and David Kempe. Generalized veto core and a practical voting rule with optimal metric distortion. In *Proceedings of ACM Conference on Economics and Computation (EC)*, pages 913–936, 2023.
- [20] Amirreza Bagheridelouee, Marzie Nilipour, Masoud Seddighin, and Maziar Shamsipour. Metric distortion under public-spirited voting. In *Proceedings of International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, pages 2144–2146, 2024.

- [21] Mohammad Ghodsi, Mohamad Latifian, and Masoud Seddighin. On the distortion value of the elections with abstention. In *Proceedings of the AAAI Conference on Artificial Intelligence (AAAI)*, volume 33, pages 1981–1988, 2019.
- [22] Elliot Anshelevich, Aris Filos-Ratsikas, Nisarg Shah, and Alexandros A Voudouris. Distortion in social choice problems: The first 15 years and beyond. In *Proceedings of International Joint Conference on Artificial Intelligence (IJCAI)*, pages 4294–4301, 2021.
- [23] Aris Filos-Ratsikas and Alexandros A Voudouris. Revisiting the distortion of distributed voting. *Theory of Computing Systems*, 68(5):1138–1159, 2024.
- [24] Aris Filos-Ratsikas, Evi Micha, and Alexandros A Voudouris. The distortion of distributed voting. *Artificial Intelligence*, 286:103343, 2020.
- [25] Alexandros A Voudouris. Tight distortion bounds for distributed metric voting on a line. *Operations Research Letters*, 51(3):266–269, 2023.
- [26] Craig Boutilier, Ioannis Caragiannis, Simi Haber, Tyler Lu, Ariel D Procaccia, and Or Sheffet. Optimal social choice functions: A utilitarian view. *Artificial Intelligence*, 227:190–213, 2015.
- [27] Ioannis Caragiannis and Ariel D Procaccia. Voting almost maximizes social welfare despite limited communication. *Artificial Intelligence*, 175(9-10):1655–1671, 2011.
- [28] Soroush Ebadian, Aris Filos-Ratsikas, Mohamad Latifian, and Nisarg Shah. Explainable and efficient randomized voting rules. In *Proceedings of Conference on Neural Information Processing Systems (NeurIPS)*, volume 36, pages 23034–23046, 2023.
- [29] Soroush Ebadian, Anson Kahng, Dominik Peters, and Nisarg Shah. Optimized distortion and proportional fairness in voting. *ACM Transactions on Economics and Computation*, 12(1):1–39, 2024.
- [30] Mark Bedaywi, Bailey Flanigan, Mohamad Latifian, and Nisarg Shah. The distortion of public-spirited participatory budgeting. In *Proceedings of the AAAI Conference on Artificial Intelligence (AAAI)*, volume 39, pages 13605–13613, 2025.
- [31] Michal Jaworski and Piotr Skowron. Evaluating committees for representative democracies: the distortion and beyond. In *Proceedings of International Joint Conference on Artificial Intelligence (IJCAI)*, pages 196–202, 2020.
- [32] Fatih Erdem Kizilkaya and David Kempe. Plurality veto: A simple voting rule achieving optimal metric distortion. In *Proceedings of International Joint Conference on Artificial Intelligence (IJCAI)*, pages 349–355, 2022.
- [33] Moses Charikar, Prasanna Ramakrishnan, Kangning Wang, and Hongxun Wu. Breaking the metric voting distortion barrier. *Journal of the ACM*, 71(6):1–33, 2024.
- [34] Vasilis Gkatzelis, Mohamad Latifian, and Nisarg Shah. Best of both distortion worlds. In *Proceedings of ACM Conference on Economics and Computation (EC)*, pages 738–758, 2023.
- [35] Aris Filos-Ratsikas, Panagiotis Kanellopoulos, Alexandros A Voudouris, and Rongsen Zhang. The distortion of distributed facility location. *Artificial Intelligence*, 328:104066, 2024.
- [36] Alexandros A Voudouris. Metric distortion of obnoxious distributed voting. *Information Processing Letters*, 189:106559, 2025.
- [37] David Kempe. Communication, distortion, and randomness in metric voting. In *Proceedings of the AAAI Conference on Artificial Intelligence (AAAI)*, volume 34, pages 2087–2094, 2020.
- [38] Georgios Amanatidis, Elliot Anshelevich, Christopher Jerrett, and Alexandros A Voudouris. Metric distortion under group-fair objectives. In *Proceedings of International Symposium on Algorithmic Game Theory (SAGT)*, pages 465–482, 2025.

NeurIPS Paper Checklist

1. Claims

Question: Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope?

Answer: [\[Yes\]](#)

Justification: We provide a proof sketch for all our claims in the main paper, with the full proofs detailed in the Appendix.

Guidelines:

- The answer NA means that the abstract and introduction do not include the claims made in the paper.
- The abstract and/or introduction should clearly state the claims made, including the contributions made in the paper and important assumptions and limitations. A No or NA answer to this question will not be perceived well by the reviewers.
- The claims made should match theoretical and experimental results, and reflect how much the results can be expected to generalize to other settings.
- It is fine to include aspirational goals as motivation as long as it is clear that these goals are not attained by the paper.

2. Limitations

Question: Does the paper discuss the limitations of the work performed by the authors?

Answer: [\[NA\]](#)

Justification: The gap between our upper and lower bounds for various settings is detailed in Table 2.

Guidelines:

- The answer NA means that the paper has no limitation while the answer No means that the paper has limitations, but those are not discussed in the paper.
- The authors are encouraged to create a separate "Limitations" section in their paper.
- The paper should point out any strong assumptions and how robust the results are to violations of these assumptions (e.g., independence assumptions, noiseless settings, model well-specification, asymptotic approximations only holding locally). The authors should reflect on how these assumptions might be violated in practice and what the implications would be.
- The authors should reflect on the scope of the claims made, e.g., if the approach was only tested on a few datasets or with a few runs. In general, empirical results often depend on implicit assumptions, which should be articulated.
- The authors should reflect on the factors that influence the performance of the approach. For example, a facial recognition algorithm may perform poorly when image resolution is low or images are taken in low lighting. Or a speech-to-text system might not be used reliably to provide closed captions for online lectures because it fails to handle technical jargon.
- The authors should discuss the computational efficiency of the proposed algorithms and how they scale with dataset size.
- If applicable, the authors should discuss possible limitations of their approach to address problems of privacy and fairness.
- While the authors might fear that complete honesty about limitations might be used by reviewers as grounds for rejection, a worse outcome might be that reviewers discover limitations that aren't acknowledged in the paper. The authors should use their best judgment and recognize that individual actions in favor of transparency play an important role in developing norms that preserve the integrity of the community. Reviewers will be specifically instructed to not penalize honesty concerning limitations.

3. Theory assumptions and proofs

Question: For each theoretical result, does the paper provide the full set of assumptions and a complete (and correct) proof?

Answer: [Yes]

Justification: [TODO]

Guidelines:

- The answer NA means that the paper does not include theoretical results.
- All the theorems, formulas, and proofs in the paper should be numbered and cross-referenced.
- All assumptions should be clearly stated or referenced in the statement of any theorems.
- The proofs can either appear in the main paper or the supplemental material, but if they appear in the supplemental material, the authors are encouraged to provide a short proof sketch to provide intuition.
- Inversely, any informal proof provided in the core of the paper should be complemented by formal proofs provided in appendix or supplemental material.
- Theorems and Lemmas that the proof relies upon should be properly referenced.

4. Experimental result reproducibility

Question: Does the paper fully disclose all the information needed to reproduce the main experimental results of the paper to the extent that it affects the main claims and/or conclusions of the paper (regardless of whether the code and data are provided or not)?

Answer: [NA]

Justification: [TODO]

Guidelines:

- The answer NA means that the paper does not include experiments.
- If the paper includes experiments, a No answer to this question will not be perceived well by the reviewers: Making the paper reproducible is important, regardless of whether the code and data are provided or not.
- If the contribution is a dataset and/or model, the authors should describe the steps taken to make their results reproducible or verifiable.
- Depending on the contribution, reproducibility can be accomplished in various ways. For example, if the contribution is a novel architecture, describing the architecture fully might suffice, or if the contribution is a specific model and empirical evaluation, it may be necessary to either make it possible for others to replicate the model with the same dataset, or provide access to the model. In general, releasing code and data is often one good way to accomplish this, but reproducibility can also be provided via detailed instructions for how to replicate the results, access to a hosted model (e.g., in the case of a large language model), releasing of a model checkpoint, or other means that are appropriate to the research performed.
- While NeurIPS does not require releasing code, the conference does require all submissions to provide some reasonable avenue for reproducibility, which may depend on the nature of the contribution. For example
 - (a) If the contribution is primarily a new algorithm, the paper should make it clear how to reproduce that algorithm.
 - (b) If the contribution is primarily a new model architecture, the paper should describe the architecture clearly and fully.
 - (c) If the contribution is a new model (e.g., a large language model), then there should either be a way to access this model for reproducing the results or a way to reproduce the model (e.g., with an open-source dataset or instructions for how to construct the dataset).
 - (d) We recognize that reproducibility may be tricky in some cases, in which case authors are welcome to describe the particular way they provide for reproducibility. In the case of closed-source models, it may be that access to the model is limited in some way (e.g., to registered users), but it should be possible for other researchers to have some path to reproducing or verifying the results.

5. Open access to data and code

Question: Does the paper provide open access to the data and code, with sufficient instructions to faithfully reproduce the main experimental results, as described in supplemental material?

Answer: [NA]

Justification: [TODO]

Guidelines:

- The answer NA means that paper does not include experiments requiring code.
- Please see the NeurIPS code and data submission guidelines (<https://nips.cc/public/guides/CodeSubmissionPolicy>) for more details.
- While we encourage the release of code and data, we understand that this might not be possible, so “No” is an acceptable answer. Papers cannot be rejected simply for not including code, unless this is central to the contribution (e.g., for a new open-source benchmark).
- The instructions should contain the exact command and environment needed to run to reproduce the results. See the NeurIPS code and data submission guidelines (<https://nips.cc/public/guides/CodeSubmissionPolicy>) for more details.
- The authors should provide instructions on data access and preparation, including how to access the raw data, preprocessed data, intermediate data, and generated data, etc.
- The authors should provide scripts to reproduce all experimental results for the new proposed method and baselines. If only a subset of experiments are reproducible, they should state which ones are omitted from the script and why.
- At submission time, to preserve anonymity, the authors should release anonymized versions (if applicable).
- Providing as much information as possible in supplemental material (appended to the paper) is recommended, but including URLs to data and code is permitted.

6. Experimental setting/details

Question: Does the paper specify all the training and test details (e.g., data splits, hyper-parameters, how they were chosen, type of optimizer, etc.) necessary to understand the results?

Answer: [NA]

Justification: [TODO]

Guidelines:

- The answer NA means that the paper does not include experiments.
- The experimental setting should be presented in the core of the paper to a level of detail that is necessary to appreciate the results and make sense of them.
- The full details can be provided either with the code, in appendix, or as supplemental material.

7. Experiment statistical significance

Question: Does the paper report error bars suitably and correctly defined or other appropriate information about the statistical significance of the experiments?

Answer: [NA]

Justification: [TODO]

Guidelines:

- The answer NA means that the paper does not include experiments.
- The authors should answer "Yes" if the results are accompanied by error bars, confidence intervals, or statistical significance tests, at least for the experiments that support the main claims of the paper.
- The factors of variability that the error bars are capturing should be clearly stated (for example, train/test split, initialization, random drawing of some parameter, or overall run with given experimental conditions).
- The method for calculating the error bars should be explained (closed form formula, call to a library function, bootstrap, etc.)

- The assumptions made should be given (e.g., Normally distributed errors).
- It should be clear whether the error bar is the standard deviation or the standard error of the mean.
- It is OK to report 1-sigma error bars, but one should state it. The authors should preferably report a 2-sigma error bar than state that they have a 96% CI, if the hypothesis of Normality of errors is not verified.
- For asymmetric distributions, the authors should be careful not to show in tables or figures symmetric error bars that would yield results that are out of range (e.g. negative error rates).
- If error bars are reported in tables or plots, The authors should explain in the text how they were calculated and reference the corresponding figures or tables in the text.

8. Experiments compute resources

Question: For each experiment, does the paper provide sufficient information on the computer resources (type of compute workers, memory, time of execution) needed to reproduce the experiments?

Answer: [NA]

Justification: [TODO]

Guidelines:

- The answer NA means that the paper does not include experiments.
- The paper should indicate the type of compute workers CPU or GPU, internal cluster, or cloud provider, including relevant memory and storage.
- The paper should provide the amount of compute required for each of the individual experimental runs as well as estimate the total compute.
- The paper should disclose whether the full research project required more compute than the experiments reported in the paper (e.g., preliminary or failed experiments that didn't make it into the paper).

9. Code of ethics

Question: Does the research conducted in the paper conform, in every respect, with the NeurIPS Code of Ethics <https://neurips.cc/public/EthicsGuidelines>?

Answer: [Yes]

Justification: [TODO]

Guidelines:

- The answer NA means that the authors have not reviewed the NeurIPS Code of Ethics.
- If the authors answer No, they should explain the special circumstances that require a deviation from the Code of Ethics.
- The authors should make sure to preserve anonymity (e.g., if there is a special consideration due to laws or regulations in their jurisdiction).

10. Broader impacts

Question: Does the paper discuss both potential positive societal impacts and negative societal impacts of the work performed?

Answer: [NA]

Justification: [TODO]

Guidelines:

- The answer NA means that there is no societal impact of the work performed.
- If the authors answer NA or No, they should explain why their work has no societal impact or why the paper does not address societal impact.
- Examples of negative societal impacts include potential malicious or unintended uses (e.g., disinformation, generating fake profiles, surveillance), fairness considerations (e.g., deployment of technologies that could make decisions that unfairly impact specific groups), privacy considerations, and security considerations.

- The conference expects that many papers will be foundational research and not tied to particular applications, let alone deployments. However, if there is a direct path to any negative applications, the authors should point it out. For example, it is legitimate to point out that an improvement in the quality of generative models could be used to generate deepfakes for disinformation. On the other hand, it is not needed to point out that a generic algorithm for optimizing neural networks could enable people to train models that generate Deepfakes faster.
- The authors should consider possible harms that could arise when the technology is being used as intended and functioning correctly, harms that could arise when the technology is being used as intended but gives incorrect results, and harms following from (intentional or unintentional) misuse of the technology.
- If there are negative societal impacts, the authors could also discuss possible mitigation strategies (e.g., gated release of models, providing defenses in addition to attacks, mechanisms for monitoring misuse, mechanisms to monitor how a system learns from feedback over time, improving the efficiency and accessibility of ML).

11. Safeguards

Question: Does the paper describe safeguards that have been put in place for responsible release of data or models that have a high risk for misuse (e.g., pretrained language models, image generators, or scraped datasets)?

Answer: [NA]

Justification: **[TODO]**

Guidelines:

- The answer NA means that the paper poses no such risks.
- Released models that have a high risk for misuse or dual-use should be released with necessary safeguards to allow for controlled use of the model, for example by requiring that users adhere to usage guidelines or restrictions to access the model or implementing safety filters.
- Datasets that have been scraped from the Internet could pose safety risks. The authors should describe how they avoided releasing unsafe images.
- We recognize that providing effective safeguards is challenging, and many papers do not require this, but we encourage authors to take this into account and make a best faith effort.

12. Licenses for existing assets

Question: Are the creators or original owners of assets (e.g., code, data, models), used in the paper, properly credited and are the license and terms of use explicitly mentioned and properly respected?

Answer: [NA]

Justification: **[TODO]**

Guidelines:

- The answer NA means that the paper does not use existing assets.
- The authors should cite the original paper that produced the code package or dataset.
- The authors should state which version of the asset is used and, if possible, include a URL.
- The name of the license (e.g., CC-BY 4.0) should be included for each asset.
- For scraped data from a particular source (e.g., website), the copyright and terms of service of that source should be provided.
- If assets are released, the license, copyright information, and terms of use in the package should be provided. For popular datasets, paperswithcode.com/datasets has curated licenses for some datasets. Their licensing guide can help determine the license of a dataset.
- For existing datasets that are re-packaged, both the original license and the license of the derived asset (if it has changed) should be provided.

- If this information is not available online, the authors are encouraged to reach out to the asset’s creators.

13. **New assets**

Question: Are new assets introduced in the paper well documented and is the documentation provided alongside the assets?

Answer: [NA]

Justification: **[TODO]**

Guidelines:

- The answer NA means that the paper does not release new assets.
- Researchers should communicate the details of the dataset/code/model as part of their submissions via structured templates. This includes details about training, license, limitations, etc.
- The paper should discuss whether and how consent was obtained from people whose asset is used.
- At submission time, remember to anonymize your assets (if applicable). You can either create an anonymized URL or include an anonymized zip file.

14. **Crowdsourcing and research with human subjects**

Question: For crowdsourcing experiments and research with human subjects, does the paper include the full text of instructions given to participants and screenshots, if applicable, as well as details about compensation (if any)?

Answer: [NA]

Justification: **[TODO]**

Guidelines:

- The answer NA means that the paper does not involve crowdsourcing nor research with human subjects.
- Including this information in the supplemental material is fine, but if the main contribution of the paper involves human subjects, then as much detail as possible should be included in the main paper.
- According to the NeurIPS Code of Ethics, workers involved in data collection, curation, or other labor should be paid at least the minimum wage in the country of the data collector.

15. **Institutional review board (IRB) approvals or equivalent for research with human subjects**

Question: Does the paper describe potential risks incurred by study participants, whether such risks were disclosed to the subjects, and whether Institutional Review Board (IRB) approvals (or an equivalent approval/review based on the requirements of your country or institution) were obtained?

Answer: [NA]

Justification: **[TODO]**

Guidelines:

- The answer NA means that the paper does not involve crowdsourcing nor research with human subjects.
- Depending on the country in which research is conducted, IRB approval (or equivalent) may be required for any human subjects research. If you obtained IRB approval, you should clearly state this in the paper.
- We recognize that the procedures for this may vary significantly between institutions and locations, and we expect authors to adhere to the NeurIPS Code of Ethics and the guidelines for their institution.
- For initial submissions, do not include any information that would break anonymity (if applicable), such as the institution conducting the review.

16. **Declaration of LLM usage**

Question: Does the paper describe the usage of LLMs if it is an important, original, or non-standard component of the core methods in this research? Note that if the LLM is used only for writing, editing, or formatting purposes and does not impact the core methodology, scientific rigorousness, or originality of the research, declaration is not required.

Answer: [NA]

Justification: **[TODO]**

Guidelines:

- The answer NA means that the core method development in this research does not involve LLMs as any important, original, or non-standard components.
- Please refer to our LLM policy (<https://neurips.cc/Conferences/2025/LLM>) for what should or should not be described.

A Proofs for Section 2 (Basic Notations)

Observation 2.6. For *rand-rand* mechanism $\Psi = (f_{rd}, f_{ur})$ with output \mathbf{w} , the expected cost of the mechanism is given by $E[\text{cost}(\mathbf{w})] = \frac{1}{k} \sum_{g \in \mathcal{G}} \frac{1}{n_g} \sum_{v \in g} \text{cost}(\text{top}(v))$.

Proof. By the definitions of the *Random Dictatorship* rule (f_{rd}) and the uniform selection rule (f_{ur}), we have

$$\begin{aligned} E[\text{cost}(\mathbf{w})] &= \sum_{g \in \mathcal{G}} \Pr(\mathbf{w} = \mathbf{w}_g) \cdot E[\text{cost}(\mathbf{w}_g)] \\ &= \frac{1}{k} \sum_{g \in \mathcal{G}} E[\text{cost}(\mathbf{w}_g)] \\ &= \frac{1}{k} \sum_{g \in \mathcal{G}} \sum_{v \in g} \Pr(\mathbf{w}_g = \text{top}(v)) \cdot \text{cost}(\text{top}(v)) \\ &= \frac{1}{k} \sum_{g \in \mathcal{G}} \frac{1}{n_g} \sum_{v \in g} \text{cost}(\text{top}(v)). \end{aligned}$$

□

B Proofs for Section 3 (Distortion Bounds for rand-det)

Theorem 3.1. For the *max-avg* objective in general metric spaces, we have $D((f_\alpha, f_{ur})) \leq \alpha + 2$.

Proof. Consider an instance $\mathcal{I} = (\mathcal{V}, \mathcal{C}, \mathcal{G}, \pi, d)$ and *rand-det* mechanism $\Psi = (f_\alpha, f_{ur})$. We have

$$\begin{aligned} E[\text{cost}(\mathbf{w})] &= \sum_{g \in \mathcal{G}} \frac{1}{k} \cdot \text{cost}(\mathbf{w}_g) && \text{(Observation 2.5)} \\ &= \sum_{g \in \mathcal{G}} \frac{1}{k} \cdot \frac{1}{n_{g^*}(\mathbf{w}_g)} \sum_{v \in g^*(\mathbf{w}_g)} d(v, \mathbf{w}_g) && \text{(Definition of max-avg)} \\ &\leq \sum_{g \in \mathcal{G}} \frac{1}{k} \cdot \frac{1}{n_{g^*}(\mathbf{w}_g)} \sum_{v \in g^*(\mathbf{w}_g)} \left(d(v, \mathbf{o}) + d(\mathbf{o}, \mathbf{w}_g) \right) && \text{(Triangle Inequality)} \\ &= \sum_{g \in \mathcal{G}} \frac{1}{k} \cdot \frac{1}{n_{g^*}(\mathbf{w}_g)} \sum_{v \in g^*(\mathbf{w}_g)} d(v, \mathbf{o}) \\ &\quad + \sum_{g \in \mathcal{G}} \frac{1}{k} \cdot \frac{1}{n_{g^*}(\mathbf{w}_g)} \sum_{v \in g^*(\mathbf{w}_g)} d(\mathbf{o}, \mathbf{w}_g) \\ &\leq \sum_{g \in \mathcal{G}} \frac{1}{k} \cdot \text{cost}(\mathbf{o}) + \sum_{g \in \mathcal{G}} \frac{1}{k} \cdot \frac{1}{n_{g^*}(\mathbf{w}_g)} \sum_{v \in g^*(\mathbf{w}_g)} d(\mathbf{o}, \mathbf{w}_g) && \text{(Observation 2.4)} \\ &= \sum_{g \in \mathcal{G}} \frac{1}{k} \cdot \text{cost}(\mathbf{o}) + \sum_{g \in \mathcal{G}} \frac{1}{k} \cdot d(\mathbf{o}, \mathbf{w}_g) \\ &\leq \text{cost}(\mathbf{o}) + \sum_{g \in \mathcal{G}} \frac{1}{k} \cdot \frac{1}{n_g} \sum_{v \in g} \left(d(v, \mathbf{o}) + d(v, \mathbf{w}_g) \right) && \text{(Triangle Inequality)} \\ &= \text{cost}(\mathbf{o}) + \sum_{g \in \mathcal{G}} \frac{1}{k} \cdot \frac{1}{n_g} \sum_{v \in g} d(v, \mathbf{o}) \\ &\quad + \sum_{g \in \mathcal{G}} \frac{1}{k} \cdot \frac{1}{n_g} \sum_{v \in g} d(v, \mathbf{w}_g) \\ &\leq \text{cost}(\mathbf{o}) + \sum_{g \in \mathcal{G}} \frac{1}{k} \cdot \text{cost}(\mathbf{o}) + \sum_{g \in \mathcal{G}} \frac{1}{k} \cdot \text{cost}_g(\mathbf{w}_g) \end{aligned}$$

$$\begin{aligned}
&= 2\text{cost}(\mathbf{o}) + \sum_{g \in \mathcal{G}} \frac{1}{k} \cdot \text{cost}_g(\mathbf{w}_g) \\
&\leq 2\text{cost}(\mathbf{o}) + \sum_{g \in \mathcal{G}} \frac{1}{k} \cdot \alpha \cdot \text{cost}_g(\mathbf{o}_g) && (\text{Observation 2.11}) \\
&\leq 2\text{cost}(\mathbf{o}) + \alpha \sum_{g \in \mathcal{G}} \frac{1}{k} \cdot \text{cost}_g(\mathbf{o}) && (\text{Observation 2.3}) \\
&\leq 2\text{cost}(\mathbf{o}) + \alpha \sum_{g \in \mathcal{G}} \frac{1}{k} \cdot \text{cost}(\mathbf{o}) && (\text{Observation 2.4}) \\
&= (\alpha + 2)\text{cost}(\mathbf{o}).
\end{aligned}$$

□

Theorem 3.3. *For the avg-avg objective in general metric spaces, we have $D((f_\alpha, f_{ur})) \leq \alpha + 2 - \frac{2}{k}$.*

Proof. Consider an instance $\mathcal{I} = (\mathcal{V}, \mathcal{C}, \mathcal{G}, \pi, \mathbf{d})$ and rand-det mechanism $\Psi = (f_\alpha, f_{ur})$. We have

$$\begin{aligned}
\mathbb{E}[\text{cost}(\mathbf{w})] &= \sum_{g \in \mathcal{G}} \frac{1}{k} \cdot \text{cost}(\mathbf{w}_g) && (\text{Observation 2.5}) \\
&= \frac{1}{k} \sum_{g \in \mathcal{G}} \frac{1}{k} \sum_{g' \in \mathcal{G}} \text{cost}_{g'}(\mathbf{w}_g) \\
&= \frac{1}{k} \sum_{g \in \mathcal{G}} \frac{1}{k} \sum_{g' \in \mathcal{G}} \frac{1}{n_{g'}} \sum_{v \in g'} \mathbf{d}(v, \mathbf{w}_g) && (\text{Definition of avg-avg}) \\
&\leq \frac{1}{k} \sum_{g \in \mathcal{G}} \frac{1}{k} \sum_{g' \in \mathcal{G}, g \neq g'} \frac{1}{n_{g'}} \sum_{v \in g'} \left(\mathbf{d}(v, \mathbf{o}) + \mathbf{d}(\mathbf{o}, \mathbf{w}_g) \right) \\
&\quad + \frac{1}{k} \sum_{g \in \mathcal{G}} \frac{1}{k} \sum_{g' \in \mathcal{G}, g = g'} \frac{1}{n_{g'}} \sum_{v \in g'} \mathbf{d}(v, \mathbf{w}_g) && (\text{Triangle Inequality}) \\
&= \frac{1}{k} \sum_{g \in \mathcal{G}} \frac{1}{k} \sum_{g' \in \mathcal{G}, g \neq g'} \frac{1}{n_{g'}} \sum_{v \in g'} \left(\mathbf{d}(v, \mathbf{o}) + \mathbf{d}(\mathbf{o}, \mathbf{w}_g) \right) \\
&\quad + \frac{1}{k} \sum_{g \in \mathcal{G}} \frac{1}{k} \cdot \text{cost}_g(\mathbf{w}_g) \\
&\leq \frac{1}{k} \sum_{g \in \mathcal{G}} \frac{1}{k} \sum_{g' \in \mathcal{G}, g \neq g'} \frac{1}{n_{g'}} \sum_{v \in g'} \left(\mathbf{d}(v, \mathbf{o}) + \mathbf{d}(\mathbf{o}, \mathbf{w}_g) \right) \\
&\quad + \frac{1}{k} \sum_{g \in \mathcal{G}} \frac{1}{k} \cdot \alpha \cdot \text{cost}_g(\mathbf{o}_g) && (\text{Observation 2.11}) \\
&\leq \frac{1}{k} \sum_{g \in \mathcal{G}} \frac{1}{k} \sum_{g' \in \mathcal{G}, g \neq g'} \frac{1}{n_{g'}} \sum_{v \in g'} \left(\mathbf{d}(v, \mathbf{o}) + \mathbf{d}(\mathbf{o}, \mathbf{w}_g) \right) \\
&\quad + \frac{1}{k} \sum_{g \in \mathcal{G}} \frac{1}{k} \cdot \alpha \cdot \text{cost}_g(\mathbf{o}) && (\text{Observation 2.3}) \\
&= \frac{1}{k} \sum_{g \in \mathcal{G}} \frac{1}{k} \sum_{g' \in \mathcal{G}, g \neq g'} \frac{1}{n_{g'}} \sum_{v \in g'} \left(\mathbf{d}(v, \mathbf{o}) + \mathbf{d}(\mathbf{o}, \mathbf{w}_g) \right) \\
&\quad + \frac{\alpha}{k} \cdot \text{cost}(\mathbf{o}) && (\text{cost}(\mathbf{o}) = \sum_{g \in \mathcal{G}} \frac{1}{k} \text{cost}_g(\mathbf{o})) \\
&= \frac{1}{k} \sum_{g \in \mathcal{G}} \frac{1}{k} \sum_{g' \in \mathcal{G}, g \neq g'} \frac{1}{n_{g'}} \sum_{v \in g'} \mathbf{d}(v, \mathbf{o})
\end{aligned}$$

$$\begin{aligned}
& + \frac{1}{k} \sum_{g \in \mathcal{G}} \frac{1}{k} \sum_{g' \in \mathcal{G}, g \neq g'} \frac{1}{n_{g'}} \sum_{v \in g'} d(o, w_g) + \frac{\alpha}{k} \cdot \text{cost}(o) \\
& = \frac{1}{k} \sum_{g \in \mathcal{G}} \frac{1}{k} \sum_{g' \in \mathcal{G}, g \neq g'} \frac{1}{n_{g'}} \sum_{v \in g'} d(v, o) \\
& + \frac{1}{k} \sum_{g \in \mathcal{G}} \frac{k-1}{k} \cdot d(o, w_g) + \frac{\alpha}{k} \cdot \text{cost}(o) \\
& \leq \frac{1}{k} \sum_{g \in \mathcal{G}} \frac{1}{k} \sum_{g' \in \mathcal{G}, g \neq g'} \frac{1}{n_{g'}} \sum_{v \in g'} d(v, o) \\
& + \frac{1}{k} \sum_{g \in \mathcal{G}} \frac{k-1}{k} \cdot \frac{1}{n_g} \sum_{v \in g} (d(v, o) + d(v, w_g)) + \frac{\alpha}{k} \cdot \text{cost}(o) \quad (\text{Triangle Inequality}) \\
& = \frac{1}{k} \sum_{g \in \mathcal{G}} \frac{1}{k} \sum_{g' \in \mathcal{G}, g \neq g'} \frac{1}{n_{g'}} \sum_{v \in g'} d(v, o) \\
& + \frac{1}{k} \sum_{g \in \mathcal{G}} \frac{k-1}{k} \cdot (\text{cost}_g(o) + \text{cost}_g(w_g)) + \frac{\alpha}{k} \cdot \text{cost}(o) \quad (\text{Definition of } \text{cost}_g(\cdot)) \\
& \leq \frac{1}{k} \sum_{g \in \mathcal{G}} \frac{1}{k} \sum_{g' \in \mathcal{G}, g \neq g'} \frac{1}{n_{g'}} \sum_{v \in g'} d(v, o) \\
& + \frac{1}{k} \sum_{g \in \mathcal{G}} \frac{k-1}{k} \cdot (\text{cost}_g(o) + \alpha \cdot \text{cost}_g(o_g)) + \frac{\alpha}{k} \cdot \text{cost}(o) \quad (\text{Observation 2.11}) \\
& \leq \frac{1}{k} \sum_{g \in \mathcal{G}} \frac{1}{k} \sum_{g' \in \mathcal{G}, g \neq g'} \frac{1}{n_{g'}} \sum_{v \in g'} d(v, o) \\
& + \frac{1}{k} \sum_{g \in \mathcal{G}} \frac{k-1}{k} \cdot (\text{cost}_g(o) + \alpha \cdot \text{cost}_g(o)) + \frac{\alpha}{k} \cdot \text{cost}(o) \quad (\text{Observation 2.3}) \\
& = \frac{1}{k} \sum_{g \in \mathcal{G}} \frac{1}{k} \sum_{g' \in \mathcal{G}, g \neq g'} \frac{1}{n_{g'}} \sum_{v \in g'} d(v, o) \\
& + \frac{(k-1)(\alpha+1)}{k} \cdot \text{cost}(o) + \frac{\alpha}{k} \cdot \text{cost}(o) \quad (\text{Definition of avg-avg}) \\
& = \frac{1}{k} \sum_{g \in \mathcal{G}} \frac{1}{k} \sum_{g' \in \mathcal{G}, g \neq g'} \text{cost}_{g'}(o) + \frac{\alpha k + k - 1}{k} \cdot \text{cost}(o) \\
& = \frac{1}{k} \sum_{g \in \mathcal{G}} \frac{1}{k} \sum_{g' \in \mathcal{G}} \text{cost}_{g'}(o) - \frac{1}{k} \sum_{g \in \mathcal{G}} \frac{1}{k} \sum_{g' \in \mathcal{G}, g = g'} \text{cost}_{g'}(o) \\
& + \frac{\alpha k + k - 1}{k} \cdot \text{cost}(o) \\
& = \text{cost}(o) - \frac{1}{k} \sum_{g \in \mathcal{G}} \frac{1}{k} \cdot \text{cost}_g(o) + \frac{\alpha k + k - 1}{k} \cdot \text{cost}(o) \\
& = \text{cost}(o) - \frac{1}{k} \cdot \text{cost}(o) + \frac{\alpha k + k - 1}{k} \cdot \text{cost}(o) \\
& = (\alpha + 2 - \frac{2}{k}) \text{cost}(o).
\end{aligned}$$

□

Theorem 3.5. *For the avg-max and max-max objectives in general metric spaces, we have $D((f_{par}, f_{ur})) \leq 3$.*

Proof. We present a proof for the avg-max objective. A similar argument can be used to prove the result for the max-max objective as well.

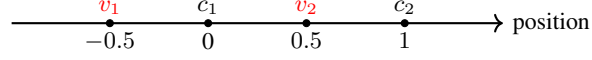


Figure 1: An example used in the proof of Theorem 3.6.

Consider an instance $\mathcal{I} = (\mathcal{V}, \mathcal{C}, \mathcal{G}, \pi, \mathbf{d})$ and **rand-det** mechanism $\Psi = (f_{par}, f_{ur})$. By the property of Pareto efficiency, for each group g , there exists a voter $v_g \in g$ who prefers \mathbf{w}_g to \mathbf{o} . Therefore, we have

$$\begin{aligned}
\mathbb{E}[\text{cost}(\mathbf{w})] &= \sum_{g \in \mathcal{G}} \frac{1}{k} \cdot \text{cost}(\mathbf{w}_g) && \text{(Observation 2.5)} \\
&= \sum_{g \in \mathcal{G}} \frac{1}{k} \cdot \frac{1}{k} \sum_{g' \in \mathcal{G}} d(v^*(\mathbf{w}_g, g'), \mathbf{w}_g) && \text{(Definition of avg-max)} \\
&\leq \sum_{g \in \mathcal{G}} \frac{1}{k} \cdot \frac{1}{k} \sum_{g' \in \mathcal{G}} \left(d(v^*(\mathbf{w}_g, g'), \mathbf{o}) + d(\mathbf{o}, \mathbf{w}_g) \right) && \text{(Triangle Inequality)} \\
&\leq \sum_{g \in \mathcal{G}} \frac{1}{k} \cdot \frac{1}{k} \sum_{g' \in \mathcal{G}} \left(d(v^*(\mathbf{o}, g'), \mathbf{o}) + d(\mathbf{o}, \mathbf{w}_g) \right) && \text{(Observation 2.10)} \\
&= \text{cost}(\mathbf{o}) + \sum_{g \in \mathcal{G}} \frac{1}{k} \cdot d(\mathbf{o}, \mathbf{w}_g) \\
&\leq \text{cost}(\mathbf{o}) + \frac{1}{k} \sum_{g \in \mathcal{G}} \left(d(\mathbf{o}, v_g) + d(v_g, \mathbf{w}_g) \right) && \text{(Triangle Inequality)} \\
&\leq \text{cost}(\mathbf{o}) + \frac{2}{k} \sum_{g \in \mathcal{G}} d(v_g, \mathbf{o}) && (d(v_g, \mathbf{w}_g) \leq d(v_g, \mathbf{o})) \\
&\leq \text{cost}(\mathbf{o}) + \frac{2}{k} \sum_{g \in \mathcal{G}} d(v^*(\mathbf{o}, g), \mathbf{o}) && \text{(Observation 2.10)} \\
&= 3\text{cost}(\mathbf{o}).
\end{aligned}$$

□

Theorem 3.6. *For the avg-max and max-max objectives, the distortion of any **rand-det** mechanism is at least 3, even when the metric space is a line.*

Proof. Consider a **rand-det** mechanism $\Psi = (f_{in}, f_{ov})$. We construct an instance with candidates $\mathcal{C} = \{c_1, c_2\}$ and voters $\mathcal{V} = \{v_1, v_2\}$ in a single group. c_1 and c_2 are located at positions 0 and 1, respectively. v_1 and v_2 with preference profiles $\pi_1 = (c_1, c_2)$ and $\pi_2 = (c_2, c_1)$, are also positioned at -0.5 and 0.5 , respectively. Refer to Figure 1 for a visual illustration. Without loss of generality, assume that Ψ selects c_2 as the representative of the group, and thus the final winner is c_2 . Since there is only one group, the avg-max and max-max objectives both simplify to max. Thus, we have $\text{cost}(c_1) = \frac{1}{2}$ and $\text{cost}(c_2) = \frac{3}{2}$. Clearly, c_1 is the optimal candidate. The distortion of Ψ is

$$D(\Psi) \geq \frac{\text{cost}(c_2)}{\text{cost}(c_1)} = 3.$$

□

Theorem 3.7. *For the max-avg objective, the distortion of any **rand-det** mechanism is at least 5, even when the metric space is a line.*

Proof. Consider a **rand-det** mechanism $\Psi = (f_{in}, f_{ov})$. We construct an instance with candidates $\mathcal{C} = \{c_1, c_2, c_3, c_4\}$, and voters $\mathcal{V} = \{v_1, v_2, v_3, v_4\}$, all located along a line metric. The voters are partitioned into two groups, $g_1 = \{v_1, v_2\}$ and $g_2 = \{v_3, v_4\}$. Let σ be an arbitrary ordering of the candidates. Without loss of generality, assume that c_1 is a candidate with in-degree at least

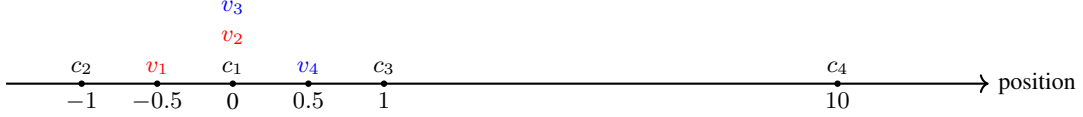


Figure 2: An example used in case 1 of Theorem 3.7. Different voter groups are distinguished by distinct colors.

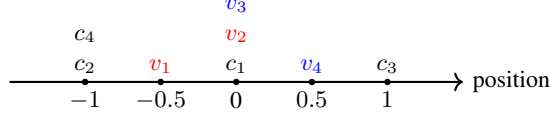


Figure 3: An example used in case 2 of Theorem 3.7. Different voter groups are distinguished by distinct colors.

$\lceil \frac{m-1}{2} \rceil = 2$ in tournament $\mathcal{T}(f_{in}, \mathcal{C}, \sigma)$, such a candidate is guaranteed to exist by Observation 2.2. Suppose c_2 and c_3 are two candidates with directed edges toward c_1 , meaning that c_1 is the "losing" candidate while both c_2 and c_3 "defeat" it in the tournament. We may further assume that $c_2 \succ_{\sigma} c_3$. Now, consider the following construction on the line metric:

- Voters v_2 and v_3 are located at positions 0, while voters v_1 and v_4 are located at -0.5 and 0.5 , respectively.
- Candidates c_2 , c_1 , and c_3 are located at positions -1 , 0 , and 1 , respectively. The position of candidate c_4 depends on the ordering σ , ensuring the input to $\mathcal{T}(f_{in}, \mathcal{C}, \sigma)$ remains valid. We analyze three cases based on the relative ordering of c_2, c_3, c_4 :
 - **Case 1:** If $c_2 \succ_{\sigma} c_3 \succ_{\sigma} c_4$, candidate c_4 is located at position 10. Refer to Figure 2 for an illustration.
 - **Case 2:** If $c_2 \succ_{\sigma} c_4 \succ_{\sigma} c_3$, candidate c_4 is located at position -1 . This case is illustrated in Figure 3
 - **Case 3:** If $c_4 \succ_{\sigma} c_2 \succ_{\sigma} c_3$, candidate c_4 is located at 1. A visual representation of this case can be found in Figure 4

Note that when a voter is equidistant from two candidates, multiple preference profiles may be consistent with the underlying metric space. According to $\mathcal{T}(f_{in}, \mathcal{C}, \sigma)$, we can determine the group representatives:

- A directed edge from c_2 to c_1 , implies that c_2 is selected as the representative for group g_1 .
- Similarly, a directed edge from c_3 to c_1 , means c_3 is the representative for group g_2 .

By the definition of the max-avg objective, we have $\text{cost}(c_2) = \text{cost}(c_3) = \frac{5}{4}$ and $\text{cost}(c_1) = \frac{1}{4}$. Thus, c_1 is the optimal candidate in all cases. The mechanism must select the final winner from the group representatives, c_2 or c_3 . Finally, we derive the distortion of mechanism Ψ as:

$$\begin{aligned} D(\Psi) &\geq \min \left(\frac{\text{cost}(c_2)}{\text{cost}(o)}, \frac{\text{cost}(c_3)}{\text{cost}(o)} \right) \\ &= \frac{\frac{5}{4}}{\text{cost}(c_1)} \\ &= 5. \end{aligned}$$

□

Theorem 3.8. *For general metric spaces and the avg-avg objective, the distortion of any rand-det mechanism is at least $5 - \frac{2}{k}$.*

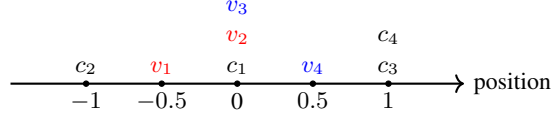


Figure 4: An example used in case 3 of Theorem 3.7. Different voter groups are distinguished by distinct colors.

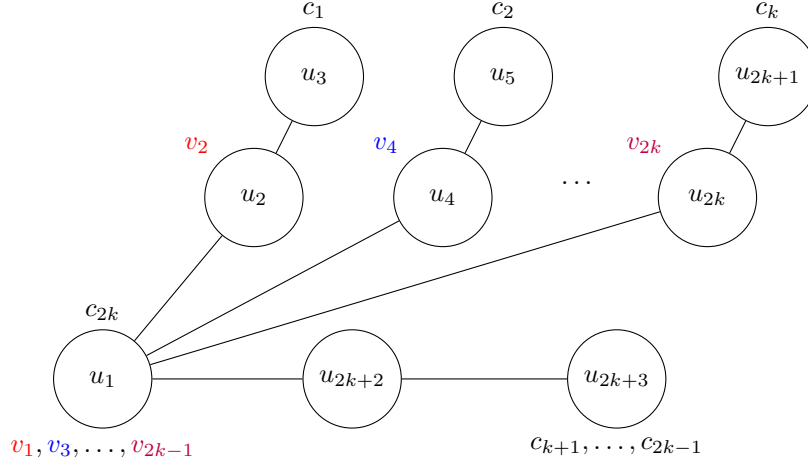


Figure 5: Tree graph used in the proof of Theorem 3.8. Different voter groups are distinguished by distinct colors.

Proof. Consider a **rand-det** mechanism $\Psi = (f_{in}, f_{ov})$. We construct an instance with a set of $m = 2k$ candidates, $\mathcal{C} = \{c_1, c_2, \dots, c_{m=2k}\}$, a set of $n = 2k$ voters, $\mathcal{V} = \{v_1, v_2, \dots, v_{n=2k}\}$, and k groups $g_i = \{v_{2i-1}, v_{2i}\}$ for $1 \leq i \leq k$. Let σ be an arbitrary ordering of the candidates. Without loss of generality, assume that c_{2k} is a candidate with in-degree at least $\lceil \frac{m-1}{2} \rceil = k$ in the tournament $\mathcal{T}(f_{in}, \mathcal{C}, \sigma)$, such a candidate is guaranteed to exist by Observation 2.2. Further, suppose c_1, c_2, \dots, c_k are k candidates that have directed edges toward c_{2k} in the tournament, meaning that c_{2k} is the "losing" candidate.

We construct a connected graph G with $2k + 3$ vertices, denoted $u_1, u_2, \dots, u_{2k+3}$, where the shortest-path distances in G define the underlying metric space \mathbf{d} . Each voter and candidate is placed on one of the vertices (a single vertex may host multiple entities). The graph G is constructed as follows (see Figure 5 for an illustration):

- Place candidate c_{2k} at vertex u_1 .
- For each $1 \leq i \leq k + 1$, add an edge between u_1 and u_{2i} , and another edge between u_{2i} and u_{2i+1} . This forms $k + 1$ branches extending from the central vertex u_1 .
- For each $1 \leq i \leq k$, place voter v_{2i-1} at vertex u_1 , voter v_{2i} at vertex u_{2i} , and candidate c_i at vertex u_{2i+1} .
- For each $k + 1 \leq i \leq 2k - 1$, place candidate c_i at vertex u_{2k+3} .

Pairwise distances between the candidates and voters are presented in Tables 3 and 4. Moreover, the preference profiles in Table 5 induced by the shortest-path distances in graph G , are consistent with the metric space \mathbf{d} . Note that multiple preference profiles may be consistent with \mathbf{d} .

According to $\mathcal{T}(f_{in}, \mathcal{C}, \sigma)$, the representative of group g_i is candidate c_i for any $1 \leq i \leq k$. Thus, the mechanism must select one of these representatives as the final winner. By the definition of the avg-avg objective, we have $\text{cost}(c_{2k}) = \frac{1}{2}$, and $\text{cost}(c_i) = \frac{5-\frac{2}{k}}{2}$, for all $1 \leq i \leq k$. Thus, c_{2k} is the optimal candidate. It follows that the distortion of mechanism Ψ :

$d(\cdot, \cdot)$	c_i
v_{2i-1}	2
v_{2i}	1
v_{2j-1}	2
v_{2j}	3

Table 3: For any $1 \leq i, j \leq k$ with $i \neq j$, the shortest-path distances in graph G between candidates c_1, c_2, \dots, c_k and the voters used in the proof of Theorem 3.8.

$d(\cdot, \cdot)$	c_i	c_{2k}
v_{2j-1}	2	0
v_{2j}	3	1

Table 4: For any $k+1 \leq i \leq 2k-1$ and $1 \leq j \leq k$, the shortest-path distances in graph G between candidates $c_{k+1}, c_{k+2}, \dots, c_{2k-1}, c_{2k}$ and the voters used in the proof of theorem 3.8.

$$\begin{aligned}
D(\Psi) &\geq \frac{\min_{1 \leq i \leq k} (\text{cost}(c_i))}{\text{cost}(o)} \\
&= \frac{\frac{5 - \frac{2}{k}}{2}}{\text{cost}(c_{2k})} \\
&= 5 - \frac{2}{k}.
\end{aligned}$$

□

C Proofs for Section 4 (Distortion Bounds for rand-rand)

Theorem 4.1. *For the max-max objective in general metric spaces, we have $D((f_{rd}, f_{ur})) \leq 3$.*

Proof. Consider an instance $\mathcal{I} = (\mathcal{V}, \mathcal{C}, \mathcal{G}, \pi, d)$ and rand-rand mechanism $\Psi = (f_{rd}, f_{ur})$. For any voter $v \in \mathcal{V}$, we have

$$\begin{aligned}
\text{cost}(\text{top}(v)) &= d(v^{**}(\text{top}(v)), \text{top}(v)) && \text{(Definition of max-max)} \\
&\leq d(v^{**}(\text{top}(v)), o) + d(o, \text{top}(v)) && \text{(Triangle Inequality)} \\
&\leq d(v^{**}(\text{top}(v)), o) + d(v, o) + d(v, \text{top}(v)) && \text{(Triangle Inequality)} \\
&\leq d(v^{**}(o), o) + d(v, o) + d(v, \text{top}(v)) && \text{(Observation 2.9)} \\
&\leq d(v^{**}(o), o) + d(v, o) + d(v, o) && \text{(Observation 2.8)} \\
&= d(v^{**}(o), o) + 2d(v, o) \\
&\leq 3d(v^{**}(o), o) && \text{(Observation 2.9)} \\
&= 3\text{cost}(o) && (\text{cost}(o) = d(v^{**}(o), o)).
\end{aligned}$$

Combining this with Observation 2.6, we have

$$\begin{aligned}
\mathbb{E}[\text{cost}(w)] &= \frac{1}{k} \sum_{g \in \mathcal{G}} \frac{1}{n_g} \sum_{v \in g} \text{cost}(\text{top}(v)) \\
&\leq 3\text{cost}(o).
\end{aligned}$$

□

Theorem 4.2. *For the avg-max objective in general metric spaces, we have $D((f_{rd}, f_{ur})) \leq 3$.*

Voter	Preference Profile
v_{2i-1}	$\sigma \uparrow c_i \uparrow c_{2k}$
v_{2i}	$\sigma \uparrow c_{2k} \uparrow c_i$

Table 5: The preference profiles of the voters within each group g_i ($1 \leq i \leq k$), used in the proof of theorem 3.8.

Proof. Consider an instance $\mathcal{I} = (\mathcal{V}, \mathcal{C}, \mathcal{G}, \pi, \mathbf{d})$ and rand-rand mechanism $\Psi = (f_{rd}, f_{ur})$. By definition of the avg-max objective for any voter $v \in \mathcal{V}$, we have

$$\text{cost}(\text{top}(v)) = \frac{1}{k} \sum_{g \in \mathcal{G}} \text{cost}_g(\text{top}(v)).$$

Now, for any groups $g, g' \in \mathcal{G}$ and any voter $v \in g'$, we have

$$\begin{aligned}
\text{cost}_g(\text{top}(v)) &= d(v^*(\text{top}(v), g), \text{top}(v)) && \text{(Definition of } \text{cost}_g(.)) \\
&\leq d(v^*(\text{top}(v), g), v) + d(v, \text{top}(v)) && \text{(Triangle Inequality)} \\
&\leq d(v^*(\text{top}(v), g), v) + d(v, o) && \text{(Observation 2.8)} \\
&\leq d(v^*(\text{top}(v), g), o) + d(v, o) + d(v, o) && \text{(Triangle Inequality)} \\
&\leq d(v^*(o, g), o) + 2d(v, o) && \text{(Observation 2.10)} \\
&= \text{cost}_g(o) + 2d(v, o) && \text{(Definition of } \text{cost}_g(.)) \\
&\leq \text{cost}_g(o) + 2d(v^*(o, g'), o) && \text{(Observation 2.10)} \\
&= \text{cost}_g(o) + 2\text{cost}_{g'}(o) && \text{(Definition of } \text{cost}_g(.)).
\end{aligned}$$

Combining this with Observation 2.6, we obtain

$$\begin{aligned}
\mathbb{E}[\text{cost}(\mathbf{w})] &= \frac{1}{k} \sum_{g \in \mathcal{G}} \frac{1}{n_g} \sum_{v \in g} \frac{1}{k} \sum_{g' \in \mathcal{G}} \text{cost}_{g'}(\text{top}(v)) \\
&\leq \frac{1}{k} \sum_{g \in \mathcal{G}} \frac{1}{n_g} \sum_{v \in g} \frac{1}{k} \sum_{g' \in \mathcal{G}} (2\text{cost}_g(o) + \text{cost}_{g'}(o)) \\
&= \frac{1}{k} \sum_{g \in \mathcal{G}} \left(2\text{cost}_g(o) + \frac{1}{k} \sum_{g' \in \mathcal{G}} \text{cost}_{g'}(o) \right) \\
&= \frac{1}{k} \sum_{g \in \mathcal{G}} 2\text{cost}_g(o) + \text{cost}(o) && \text{(Definition of } \text{cost}(.)) \\
&= 3\text{cost}(o) && \text{(Definition of } \text{cost}(.)).
\end{aligned}$$

□

Theorem 4.3. For the max-avg objective in general metric spaces, we have $D((f_{rd}, f_{ur})) \leq 3$.

Proof. Consider an instance $\mathcal{I} = (\mathcal{V}, \mathcal{C}, \mathcal{G}, \pi, \mathbf{d})$ and rand-rand mechanism $\Psi = (f_{rd}, f_{ur})$. By the definition of the max-avg objective for any voter $v \in \mathcal{V}$, we have

$$\begin{aligned}
\text{cost}(\text{top}(v)) &= \max_{g \in \mathcal{G}} \text{cost}_g(\text{top}(v)) \\
&= \text{cost}_{g^*(\text{top}(v))}(\text{top}(v)) \\
&= \frac{1}{n_{g^*(\text{top}(v))}} \sum_{v' \in g^*(\text{top}(v))} d(v', \text{top}(v)) && \text{(Definition of } \text{cost}_g(.)).
\end{aligned}$$

Thus, we have

$$\begin{aligned}
\text{cost}_{g^*(\text{top}(v))}(\text{top}(v)) &= \frac{1}{n_{g^*(\text{top}(v))}} \sum_{v' \in g^*(\text{top}(v))} d(v', \text{top}(v)) \\
&\leq \frac{1}{n_{g^*(\text{top}(v))}} \sum_{v' \in g^*(\text{top}(v))} \left(d(v, \text{top}(v)) + d(v, v') \right) && \text{(Triangle Inequality)}
\end{aligned}$$

$$\begin{aligned}
&\leq \frac{1}{n_{g^*(\text{top}(v))}} \sum_{v' \in g^*(\text{top}(v))} \left(d(v, o) + d(v, v') \right) && \text{(Observation 2.8)} \\
&\leq \frac{1}{n_{g^*(\text{top}(v))}} \sum_{v' \in g^*(\text{top}(v))} \left(d(v, o) + d(v, o) + d(o, v') \right) && \text{(Triangle Inequality)} \\
&= 2d(v, o) + \frac{1}{n_{g^*(\text{top}(v))}} \sum_{v' \in g^*(\text{top}(v))} d(o, v') \\
&= 2d(v, o) + \text{cost}_{g^*(\text{top}(v))}(o) \\
&\leq 2d(v, o) + \text{cost}(o) && \text{(Observation 2.7).}
\end{aligned}$$

Combining this with Observation 2.6, we obtain

$$\begin{aligned}
\mathbb{E}[\text{cost}(\mathbf{w})] &= \frac{1}{k} \sum_{g \in \mathcal{G}} \frac{1}{n_g} \sum_{v \in g} \text{cost}(\text{top}(v)) \\
&= \frac{1}{k} \sum_{g \in \mathcal{G}} \frac{1}{n_g} \sum_{v \in g} \text{cost}_{g^*(\text{top}(v))}(\text{top}(v)) && \text{(Definition of cost(.))} \\
&\leq \frac{1}{k} \sum_{g \in \mathcal{G}} \frac{1}{n_g} \sum_{v \in g} (2d(v, o) + \text{cost}(o)) \\
&= \text{cost}(o) + \frac{1}{k} \sum_{g \in \mathcal{G}} \frac{1}{n_g} \sum_{v \in g} 2d(v, o) \\
&= \text{cost}(o) + \frac{2}{k} \sum_{g \in \mathcal{G}} \text{cost}_g(o) && \text{(Definition of cost}_g\text{(.))} \\
&\leq \text{cost}(o) + \frac{2}{k} \sum_{g \in \mathcal{G}} \text{cost}(o) && \text{(Observation 2.7)} \\
&= 3\text{cost}(o).
\end{aligned}$$

□

Theorem 4.4. *For the avg-avg objective in general metric spaces, we have $D((f_{rd}, f_{ur})) \leq 3 - 2/kn^*$ where n^* represents the maximum value of n_g over all groups.*

Proof. Consider an instance $\mathcal{I} = (\mathcal{V}, \mathcal{C}, \mathcal{G}, \pi, d)$ and rand-rand mechanism $\Psi = (f_{rd}, f_{ur})$. By the definition of the avg-avg objective for any voter $v \in \mathcal{V}$, we have

$$\text{cost}(\text{top}(v)) = \frac{1}{k} \sum_{g' \in \mathcal{G}} \text{cost}_{g'}(\text{top}(v)).$$

For any voter $v \in \mathcal{V}$ and group g' , we have

$$\begin{aligned}
\text{cost}_{g'}(\text{top}(v)) &= \frac{1}{n_{g'}} \sum_{v' \in g'} d(v', \text{top}(v)) && \text{(Definition of cost}_g\text{(.))} \\
&\leq \frac{1}{n_{g'}} \sum_{v' \in g'} \left(d(v, \text{top}(v)) + d(v', v) \right) && \text{(Triangle Inequality)} \\
&\leq \frac{1}{n_{g'}} \sum_{v' \in g'} \left(d(v, o) + d(v', v) \right) && \text{(Observation 2.8)} \\
&= d(v, o) + \frac{1}{n_{g'}} \sum_{v' \in g'} d(v', v).
\end{aligned}$$

Combining this with Observation 2.6, we obtain

$$\mathbb{E}[\text{cost}(\mathbf{w})] = \frac{1}{k} \sum_{g \in \mathcal{G}} \frac{1}{n_g} \sum_{v \in g} \frac{1}{k} \sum_{g' \in \mathcal{G}} \text{cost}_{g'}(\text{top}(v)) \quad \text{(Definition of cost(.))}$$

$$\begin{aligned}
&\leq \frac{1}{k} \sum_{g \in \mathcal{G}} \frac{1}{n_g} \sum_{v \in g} \frac{1}{k} \sum_{g' \in \mathcal{G}} \left(d(v, o) + \frac{1}{n_{g'}} \sum_{v' \in g'} d(v', v) \right) \\
&= \frac{1}{k} \sum_{g \in \mathcal{G}} \frac{1}{n_g} \sum_{v \in g} d(v, o) + \frac{1}{k} \sum_{g \in \mathcal{G}} \frac{1}{n_g} \sum_{v \in g} \frac{1}{k} \sum_{g' \in \mathcal{G}} \frac{1}{n_{g'}} \sum_{v' \in g'} d(v', v) \\
&= \text{cost}(o) + \frac{1}{k} \sum_{g \in \mathcal{G}} \frac{1}{n_g} \sum_{v \in g} \frac{1}{k} \sum_{g' \in \mathcal{G}} \frac{1}{n_{g'}} \sum_{v' \in g'} d(v', v) \quad (\text{Definition of avg-avg}) \\
&\leq \text{cost}(o) + \frac{1}{k} \sum_{g \in \mathcal{G}} \frac{1}{n_g} \sum_{v \in g} \frac{1}{k} \sum_{g' \in \mathcal{G}} \frac{1}{n_{g'}} \sum_{v' \in g', v' \neq v} \left(d(v, o) + d(o, v') \right) \quad (\text{Triangle Inequality}) \\
&= \text{cost}(o) + \frac{1}{k} \sum_{g \in \mathcal{G}} \frac{1}{n_g} \sum_{v \in g} \frac{1}{k} \sum_{g' \in \mathcal{G}} \frac{1}{n_{g'}} \sum_{v' \in g'} \left(d(v, o) + d(o, v') \right) \\
&\quad - \frac{1}{k} \sum_{g \in \mathcal{G}} \frac{1}{n_g} \sum_{v \in g} \frac{1}{k} \sum_{g' \in \mathcal{G}, g'=g} \frac{1}{n_{g'}} \sum_{v' \in g', v'=v} \left(d(v, o) + d(o, v') \right) \\
&= \text{cost}(o) + \frac{1}{k} \sum_{g \in \mathcal{G}} \frac{1}{n_g} \sum_{v \in g} \frac{1}{k} \sum_{g' \in \mathcal{G}} \frac{1}{n_{g'}} \sum_{v' \in g'} \left(d(v, o) + d(o, v') \right) \\
&\quad - \frac{1}{k} \sum_{g \in \mathcal{G}} \frac{1}{n_g} \sum_{v \in g} \frac{1}{k} \cdot \frac{1}{n_g} \cdot 2d(v, o) \\
&= \text{cost}(o) + 2\text{cost}(o) - \frac{1}{k} \sum_{g \in \mathcal{G}} \frac{1}{n_g} \sum_{v \in g} \frac{1}{k} \cdot \frac{1}{n_g} \cdot 2d(v, o) \quad (\text{Definition of avg-avg}) \\
&= 3\text{cost}(o) - \frac{1}{k} \sum_{g \in \mathcal{G}} \frac{1}{n_g} \sum_{v \in g} \frac{1}{k} \cdot \frac{1}{n_g} \cdot 2d(v, o) \\
&= 3\text{cost}(o) - \frac{2}{k^2} \sum_{g \in \mathcal{G}} \frac{1}{n_g} \cdot \text{cost}_g(o) \quad (\text{Definition of cost}_g(.)) \\
&\leq 3\text{cost}(o) - \frac{2}{k^2} \sum_{g \in \mathcal{G}} \frac{1}{n^*} \cdot \text{cost}_g(o) \\
&= \left(3 - \frac{2}{kn^*}\right) \text{cost}(o) \quad (\text{Definition of avg-avg}).
\end{aligned}$$

□

Theorem 4.6. *For the max-avg and max-max objectives, the distortion of any rand-rand mechanism is at least 3, even when the metric space is a line.*

Proof. Consider any rand-rand mechanism $\Psi = (f_{in}, f_{ov})$. We construct an instance with candidates $\mathcal{C} = \{c_1, c_2, c_3\}$ and voters $\mathcal{V} = \{v_1, v_2\}$, where each voter belongs to a distinct group: $v_1 \in g_1$ and $v_2 \in g_2$. The instance is constructed on the line metric as follows (refer to Figure 6):

- Candidates c_1 , c_2 , and c_3 are located at positions -1 , 0 , and 1 , respectively.
- Voters v_1 and v_2 are located at positions -0.5 and 0.5 , respectively.
- The preference profile of each voter is $\pi_1 = (c_1, c_2, c_3)$ for v_1 and $\pi_2 = (c_3, c_2, c_1)$ for v_2 .

Trivially, the preference profiles are consistent with the distances in Figure 6. According to Observation 2.1, candidates c_1 and c_3 are chosen as the representatives of groups g_1 and g_2 , respectively, and then mechanism Ψ must select one of them as the final winner. We have $\text{cost}(c_1) = \text{cost}(c_3) = \frac{3}{2}$ and $\text{cost}(c_2) = \frac{1}{2}$. Clearly, c_2 is the optimal candidate. The distortion of Ψ is:

$$D(\Psi) \geq \frac{\min(\text{cost}(c_1), \text{cost}(c_3))}{\text{cost}(o)}$$

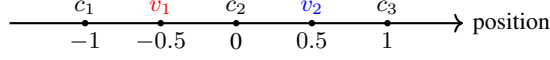


Figure 6: An example used in the proof of Theorem 4.6.

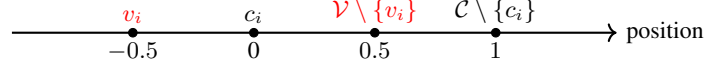


Figure 7: Configuration of the candidates and voters in instance \mathcal{I}_i used in the proof of Theorem 4.7.

$$\begin{aligned}
 &= \frac{\frac{3}{2}}{\text{cost}(c_2)} \\
 &= 3.
 \end{aligned}$$

□

Theorem 4.7. *For the avg-max objective, the distortion of any rand-rand mechanism is at least $3 - \frac{2}{n}$ (equivalently, $3 - \frac{2}{m}$), even when the metric space is a line.*

Proof. We construct an instance with a set of m candidates, $\mathcal{C} = \{c_1, c_2, \dots, c_m\}$, and a set of $n = m$ voters, $\mathcal{V} = \{v_1, v_2, \dots, v_n\}$, all belonging to a single group. Each voter v_i has a preference profile π_i that ranks candidates cyclically, starting with c_i as their top choice:

$$\pi_i = (c_i, c_{i+1}, \dots, c_m, c_1, c_2, \dots, c_{i-1}).$$

Now, we construct m instances, denoted $\mathcal{I}_1, \mathcal{I}_2, \dots, \mathcal{I}_m$, on the line metric. Across all instances, the voter set \mathcal{V} , candidate set \mathcal{C} , and the preference profile π are identical; they differ only in the arrangement of voters and candidates within the underlying metric space.

For any instance \mathcal{I}_i , where $1 \leq i \leq m$, configuration of the line metric is as follows (see Figure 7):

- Candidate c_i is located at position 0, while all other candidates are located at 1.
- Voter v_i is located at position -0.5 , and all other voters are located at 0.5 .

It is straightforward to verify that each constructed instance is consistent with the specified preference profile. Since there is a single group, the avg-max objective simplifies to max. For each instance \mathcal{I}_i , we have $\text{cost}(c_i) = \frac{1}{2}$, and $\text{cost}(c_j) = \frac{3}{2}$, for $1 \leq j \leq m$ (where $i \neq j$). Clearly, c_i is the optimal candidate in \mathcal{I}_i .

Now, consider any rand-rand mechanism Ψ . Let p_i denote the probability that Ψ selects c_i as the winner, where $\sum_{i=1}^m p_i = 1$. For instance \mathcal{I}_i , the mechanism's expected cost is

$$\mathbb{E}[\text{cost}(\Psi(\mathcal{I}_i))] = \sum_{j=1}^m p_j \cdot \text{cost}(c_j) = \sum_{j \neq i} p_j \cdot \frac{3}{2} + p_i \cdot \frac{1}{2} = \frac{3}{2}(1 - p_i) + \frac{p_i}{2} = \frac{3}{2} - p_i.$$

It follows that

$$\frac{\mathbb{E}[\text{cost}(\Psi(\mathcal{I}_i))]}{\text{cost}(c_i)} = \frac{\frac{3}{2} - p_i}{\frac{1}{2}} = 3 - 2p_i.$$

Since the total probability must sum up to 1, there exists some index i such that $p_i \leq \frac{1}{m} = \frac{1}{n}$. Finally, we obtain the following lower bound on the distortion of Ψ

$$D(\Psi) \geq 3 - \frac{2}{m} = 3 - \frac{2}{n}.$$

□

Theorem 4.9. *For general metric spaces and the avg-avg objective, the distortion of any rand-rand mechanism is at least $3 - \frac{2}{n}$ (equivalently $3 - \frac{2}{k}$).*

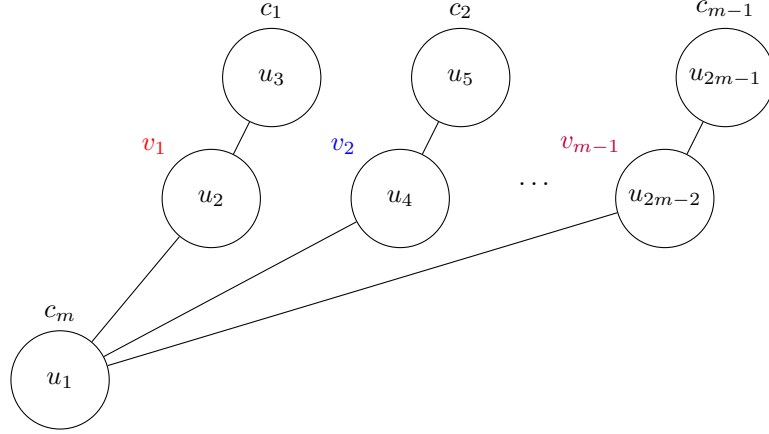


Figure 8: Tree graph used in the proof of Theorem 4.9. Different voter groups are distinguished by distinct colors.

$d(\cdot, \cdot)$	c_i	c_j	c_m
v_i	1	3	1

Table 6: For any $1 \leq i, j \leq k$ with $i \neq j$, the shortest-path distances between candidates and voters, derived from tree graph G in the proof of Theorem 4.9.

Proof. Consider a **rand-rand** mechanism Ψ . We construct an instance with a set of $m = k + 1$ candidates, $\mathcal{C} = \{c_1, c_2, \dots, c_{m=k+1}\}$, a set of $n = k$ voters $\mathcal{V} = \{v_1, v_2, \dots, v_{n=k}\}$, and k single-voter groups, $g_i = \{v_i\}$ for $1 \leq i \leq k$. Each voter v_i has a preference profile π_i , with c_i as the top choice, immediately followed by c_m , denoted as $\pi_i = \sigma \uparrow c_m \uparrow c_i$, where σ is an ordering of the candidates.

We now construct a connected graph G with $n + m$ vertices, denoted u_1, u_2, \dots, u_{n+m} , where the shortest-path distances in G define the underlying metric space \mathbf{d} . Each voter and candidate is placed on one of the vertices. The construction of G is as follows (see Figure 8 for an illustration):

- Place candidate c_m at vertex u_1 .
- For each $1 \leq i \leq k$, place voter v_i at vertex u_{2i} , and candidate c_i at vertex u_{2i+1} .
- For each $1 \leq i \leq k$, add an edge between u_1 and u_{2i} , and another edge between u_{2i} and u_{2i+1} . This forms k branches extending from the central vertex u_1 .

See Table 6 for the corresponding distances. For all $1 \leq i \leq k$, we have $\text{top}(v_i) = c_i$. Therefore, the representative of group g_i is candidate c_i (by Observation 2.1). Consequently c_m is not the representative of any group, and thus cannot be the winner of mechanism Ψ . According to the definition of the **avg-avg** objective, we have $\text{cost}(c_i) = \frac{3(n-1)+1}{n} = 3 - \frac{2}{n}$ for $1 \leq i < m$, and $\text{cost}(c_m) = 1$. Clearly, c_m is the optimal candidate. We obtain the lower bound on the distortion of Ψ as follows:

$$\begin{aligned}
 D(\Psi) &\geq \min_{1 \leq i < m} \left(\frac{\text{cost}(c_i)}{\text{cost}(c_m)} \right) \\
 &= \frac{3 - \frac{2}{n}}{1} \\
 &= 3 - \frac{2}{n}.
 \end{aligned}$$

□

D Proofs for Section 5 (Resolving the Distortion Bounds for **det-det**)

Theorem 5.1. *For the max-max objective in general metric spaces, we have $D(m_{ad}) \leq 3$.*

Proof. Consider an instance $\mathcal{I} = (\mathcal{V}, \mathcal{C}, \mathcal{G}, \pi, d)$ and the *Arbitrary Dictator* mechanism ($\Psi = m_{ad}$), which selects the top-ranked candidate of an arbitrary voter v as the final winner; that is, $w = \text{top}(v)$. We follow a strategy roughly analogous to that in Theorem 4.1.

$$\begin{aligned}
\text{cost}(\text{top}(v)) &= d(v^{**}(\text{top}(v)), \text{top}(v)) && \text{(Definition of max-max)} \\
&\leq d(v^{**}(\text{top}(v)), v) + d(v, \text{top}(v)) && \text{(Triangle Inequality)} \\
&\leq d(v^{**}(\text{top}(v)), v) + d(v, o) && \text{(Observation 2.8)} \\
&\leq d(v^{**}(\text{top}(v)), o) + d(o, v) + d(v, o) && \text{(Triangle Inequality)} \\
&= d(v^{**}(\text{top}(v)), o) + 2d(v, o) \\
&\leq d(v^{**}(\text{top}(v)), o) + 2d(v^{**}(o), o) && \text{(Observation 2.9)} \\
&\leq 3d(v^{**}(o), o) && \text{(Observation 2.9)} \\
&= 3\text{cost}(o) && \text{(Definition of max-max).}
\end{aligned}$$

□

Theorem 5.2. *For the avg-max objective in general metric spaces, we have $D((f_{par}, f_\beta)) \leq 2\beta + 3$.*

Proof. Consider an instance $\mathcal{I} = (\mathcal{V}, \mathcal{C}, \mathcal{G}, \pi, d)$ and a **det-det** mechanism $\Psi = (f_{par}, f_\beta)$. By the property of Pareto efficiency, for each group g , there exists a voter $v_g \in g$ who prefers the representative w_g over the optimal candidate o . We have

$$\begin{aligned}
\text{cost}(w) &= \frac{1}{k} \sum_{g \in \mathcal{G}} d(v^*(w, g), w) && \text{(Definition of avg-max)} \\
&\leq \frac{1}{k} \sum_{g \in \mathcal{G}} d(v^*(w, g), o) + d(o, w) && \text{(Triangle Inequality)} \\
&\leq \frac{1}{k} \sum_{g \in \mathcal{G}} d(v^*(o, g), o) + d(o, w) && \text{(Observation 2.10)} \\
&= \text{cost}(o) + \frac{1}{k} \sum_{g \in \mathcal{G}} d(o, w) && \text{(Definition of avg-max)} \\
&\leq \text{cost}(o) + \frac{1}{k} \sum_{g \in \mathcal{G}} d(o, w_g) + d(w, w_g) && \text{(Triangle Inequality)} \\
&\leq \text{cost}(o) + (\beta + 1) \cdot \frac{1}{k} \sum_{g \in \mathcal{G}} d(o, w_g) && \text{(Observation 2.12)} \\
&\leq \text{cost}(o) + (\beta + 1) \cdot \frac{1}{k} \sum_{g \in \mathcal{G}} d(o, v_g) + d(v_g, w_g) && \text{(Triangle Inequality)} \\
&\leq \text{cost}(o) + 2(\beta + 1) \cdot \frac{1}{k} \sum_{g \in \mathcal{G}} d(o, v_g) && (d(v_g, w_g) \leq d(v_g, o)) \\
&\leq \text{cost}(o) + 2(\beta + 1) \cdot \frac{1}{k} \sum_{g \in \mathcal{G}} d(o, v^*(o, g)) && \text{(Observation 2.10)} \\
&= (2\beta + 3)\text{cost}(o) && \text{(Definition of avg-max).}
\end{aligned}$$

□

Theorem 5.4. *For general metric spaces and the max-avg objective, the distortion of any **det-det** mechanism is at least 5.*

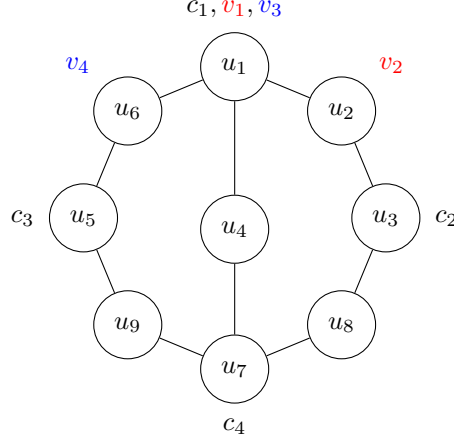


Figure 9: An illustration of the graph G used in the proof of Theorem 5.4. Different voter groups are distinguished by distinct colors.

$d(\cdot, \cdot)$	c_1	c_2	c_3	c_4
v_1	0	2	2	2
v_2	1	1	3	3
v_3	0	2	2	2
v_4	1	3	1	3

Table 7: The shortest-path distances between candidates and voters, as derived from the graph G in the proof of theorem 5.4.

Proof. Consider a **det-det** mechanism $\Psi = (f_{in}, f_{ov})$. We construct an instance with a set of candidates $\mathcal{C} = \{c_1, c_2, c_3, c_4\}$, a set of voters $\mathcal{V} = \{v_1, v_2, v_3, v_4\}$, and 2 groups $g_1 = \{v_1, v_2\}$ and $g_2 = \{v_3, v_4\}$. Let σ be an arbitrary ordering of the candidates. Without loss of generality, assume that c_1 has an in-degree of at least $\lceil \frac{m-1}{2} \rceil = 2$ in the tournament $\mathcal{T}(f_{in}, \mathcal{C}, \sigma)$, such a candidate is guaranteed to exist by Observation 2.2. Further, let c_2 and c_3 be the two candidates with directed edges toward c_1 in the tournament.

We now construct a connected graph G with 9 vertices, denoted u_1, u_2, \dots, u_9 , where the shortest-path distances in G define the underlying metric space d . Each voter and candidate is placed at one of the vertices (a single vertex may host multiple entities). The graph configuration is as shown in Figure 9.

Pairwise distances between the candidates and voters are presented in Table 7. Now, we define the preference profiles in Table 8, which are consistent with both the shortest-path distances in graph G and the input of tournament $\mathcal{T}(f_{in}, \mathcal{C}, \sigma)$. Note that multiple profiles may be consistent.

Since candidates c_2 and c_3 defeat c_1 in $\mathcal{T}(f_{in}, \mathcal{C}, \sigma)$, they must serve as the representative of groups g_1 and g_2 , respectively. Thus, the set of representatives is $R = \{c_2, c_3\}$.

At the second stage of the distributed voting process, we consider two instances:

- \mathcal{I}_1 : The preference profiles of c_2 and c_3 are (c_2, c_1, c_4, c_3) and (c_3, c_1, c_4, c_2) , respectively.
- \mathcal{I}_2 : The preference profiles of c_2 and c_3 are (c_2, c_4, c_1, c_3) and (c_3, c_4, c_1, c_2) , respectively.

It is straightforward to verify that both \mathcal{I}_1 and \mathcal{I}_2 are consistent with the metric space defined earlier. In both cases, the preference profiles of c_2 and c_3 follow the pattern

$$(c_2, c_a, c_b, c_3) \quad \text{and} \quad (c_3, c_a, c_b, c_2),$$

where (c_a, c_b) is some ordering of (c_1, c_4) .

If the rule selects the first-ranked or fourth-ranked candidate, then c_1 is not chosen. If the rule selects the second-ranked candidate (c_a), then in \mathcal{I}_2 we again exclude c_1 . If the rule selects the third-ranked

Voter	Preference Profile
v_1	$\sigma \uparrow c_2 \uparrow c_1$
v_2	$\sigma \uparrow c_1 \uparrow c_2$
v_3	$\sigma \uparrow c_3 \uparrow c_1$
v_4	$\sigma \uparrow c_1 \uparrow c_3$

Table 8: The voter preference profiles used in the proof of theorem 5.4.

candidate (c_b), then in \mathcal{I}_1 we exclude c_1 . Therefore, in every case, there exists an instance in which the mechanism selects the winner from the set $\{c_2, c_3, c_4\}$.

By the definition of the max-avg objective, we have $\text{cost}(c_1) = \frac{1}{2}$ and $\text{cost}(c_2) = \text{cost}(c_3) = \text{cost}(c_4) = \frac{5}{2}$. Thus, c_1 is the optimal candidate. The distortion of mechanism Ψ is obtained as follows:

$$\begin{aligned}
D(\Psi) &\geq \min\left(\frac{\text{cost}(c_2), \text{cost}(c_3), \text{cost}(c_4)}{\text{cost}(c_1)}\right) \\
&\geq \frac{\frac{5}{2}}{\frac{1}{2}} \\
&= 5.
\end{aligned}$$

□

E Proofs for Section 6 (An Extension of Lower Bounds for rand-rand and rand-det)

Theorem 6.1. *For the avg-avg objective in Euclidean space, the distortion of any rand-rand mechanism is at least $\sqrt{5} - \varepsilon$, for every constant $\varepsilon > 0$.*

Proof. Consider any rand-rand mechanism Ψ . Let l be a positive integer. Consider an instance in $(l+1)$ -dimensional Euclidean space, \mathbb{R}^{l+1} , with $l+2$ candidates, denoted c_1, c_2, \dots, c_{l+2} , and $k = l+1$ groups, each with a single voter v_i for $1 \leq i \leq l+1$. We construct the instance as follows:

- Let q_i be the point in \mathbb{R}^{l+1} whose i -th coordinate is 1 and all other coordinates are 0, for $1 \leq i \leq l+1$.
- Place candidate c_i at point q_i for each $1 \leq i \leq l+1$.
- The final candidate, c_{l+2} , is placed at the centroid of the other candidates; $\left(\frac{1}{l+1}, \frac{1}{l+1}, \dots, \frac{1}{l+1}\right)$.
- In the i -th group, the single voter v_i is positioned at the midpoint between their corresponding candidate, c_i and the centroid candidate, c_{l+2} . Indeed, each voter v_i is located at a point where the i -th coordinate is $\frac{l+2}{2(l+1)}$ and all other coordinates are $\frac{1}{2(l+1)}$. Note that each voter's preference profile is structured so that the top-ranked candidate of voter v_i is c_i , consistent with the underlying Euclidean space.

In particular, in 3D space ($l = 2$), the instance lies within an equilateral triangle with vertices $(1, 0, 0)$, $(0, 1, 0)$, and $(0, 0, 1)$, as illustrated in Figure 10.

For all $1 \leq i \leq l+1$, we have

$$\begin{aligned}
d(c_i, v_i) &= d(c_{l+2}, v_i) \\
&= \sqrt{\left(\frac{1}{2(l+1)}\right)^2 l + \left(\frac{l}{2(l+1)}\right)^2} \\
&= \sqrt{\frac{l}{4(l+1)}}.
\end{aligned}$$

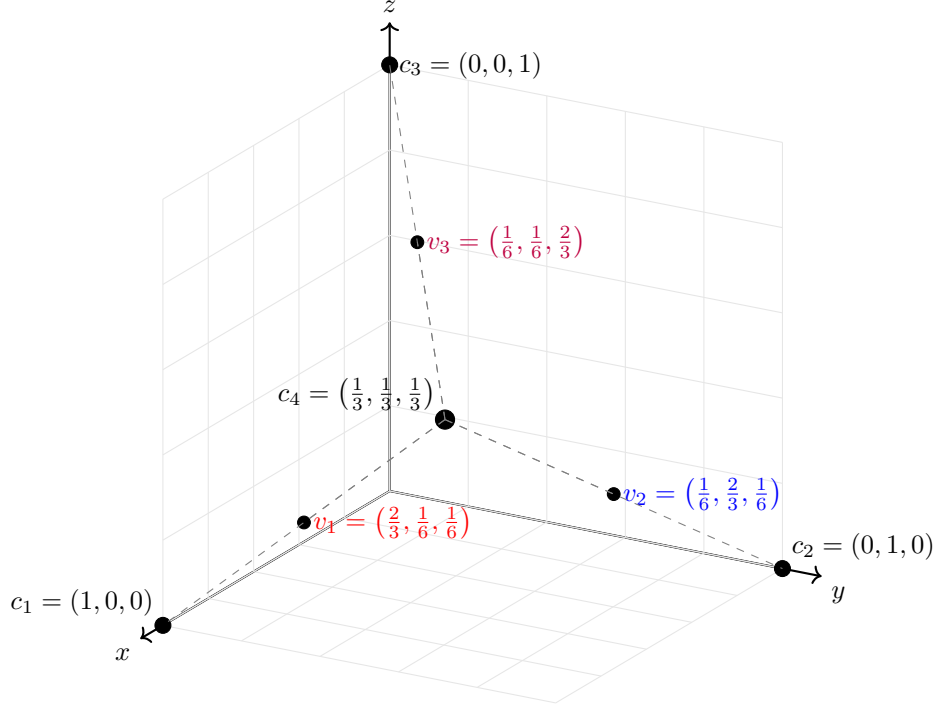


Figure 10: A 3 – dimensional Euclidean model ($l = 2$) illustrating the geometric lower bound construction for the rand-rand mechanisms under the avg-avg objective. Candidates c_1 , c_2 , and c_3 are located at the unit basis vectors, with $c_4 = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ at the centroid of the triangle they form. Voters v_1 , v_2 , and v_3 are positioned at the midpoints between the centroid c_4 and their top-ranked candidate c_i . Different group voters are indicated via distinct colors.

Moreover, for all $1 \leq i, j \leq l + 1$ (where $i \neq j$), we have

$$\begin{aligned} d(c_i, v_j) &= \sqrt{\left(\frac{1}{2(l+1)}\right)^2 (l-1) + \left(\frac{l+2}{2(l+1)}\right)^2 + \left(\frac{2l+1}{2l+2}\right)^2} \\ &= \sqrt{\frac{5l+4}{4l+4}}. \end{aligned}$$

By the definition of the avg-avg objective, the cost of each candidate is the average distance to all $l + 1$ voters. Thus, we conclude that

$$\begin{aligned} \text{cost}(c_{l+2}) &= \frac{1}{2} \sqrt{\frac{l}{l+1}}, \\ \text{cost}(c_i) &= \frac{l \sqrt{\frac{5l+4}{4l+4}} + \frac{1}{2} \sqrt{\frac{l}{l+1}}}{l+1} \quad (1 \leq i \leq l+1). \end{aligned}$$

Clearly, the optimal candidate is c_{l+2} . According to Observation 2.1, the representative of the i -th group is candidate c_i . Finally, the mechanism selects the winner from among the candidates c_1, c_2, \dots, c_{l+1} . A lower bound on the distortion of the mechanism Ψ is obtained as follows ($1 \leq i \leq l+1$):

$$\begin{aligned} D(\Psi) &\geq \frac{\text{cost}(c_i)}{\text{cost}(c_{l+2})} \\ &= \frac{\frac{l}{l+1} \sqrt{\frac{5l+4}{4l+4}} + \frac{1}{2(l+1)} \sqrt{\frac{l}{l+1}}}{\frac{1}{2} \sqrt{\frac{l}{l+1}}}. \end{aligned}$$

As $l \rightarrow \infty$, the ratio approaches $\sqrt{5} \approx 2.236$. Therefore, for any $\varepsilon > 0$, we can construct an instance with distortion greater than $\sqrt{5} - \varepsilon$. \square

Theorem 6.2. *For the avg-avg objective in Euclidean space, the distortion of any rand-det mechanism is at least $2 + \sqrt{5} - \varepsilon$, for every constant $\varepsilon > 0$.*

Proof. Consider a rand-det mechanism $\Psi = (f_{in}, f_{ov})$, a set of $2m$ candidates $\mathcal{C} = \{c_1, c_2, \dots, c_{2m}\}$, and an arbitrary ordering σ over them. By Observation 2.2, the tournament $\mathcal{T}(f_{in}, \mathcal{C}, \sigma)$, must have a candidate with in-degree at least $\lceil \frac{2m-1}{2} \rceil = m$. Without loss of generality, let c_{m+1} be such a candidate and let c_1, c_2, \dots, c_m be m candidates that have directed edges toward c_{m+1} in the tournament. We now construct the following instance with $k = m$ groups in $(m+1)$ -dimensional Euclidean space:

- Let q_i be the point in \mathbb{R}^{m+1} whose i -th coordinate is 1 and all other coordinates are 0, for $1 \leq i \leq m+1$.
- Place candidate c_i at point q_i for each $1 \leq i \leq m$ and candidate c_{m+1} at the centroid $\left(\frac{1}{m+1}, \frac{1}{m+1}, \dots, \frac{1}{m+1}\right)$.
- Place candidate c_i at point q_{m+1} for each $m+2 \leq i \leq 2m$.
- In the i -th group ($1 \leq i \leq m$), there are two voters:
 1. Voter v_{2i-1} is located at the centroid, which is the same position as candidate c_{m+1} .
 2. Voter v_{2i} is located exactly at the midpoint between candidates c_{m+1} and c_i , with coordinates equal to $\frac{m+2}{2(m+1)}$ in the i -th dimension and $\frac{1}{2(m+1)}$ in all other dimensions.
- The ordinal preferences of the v_{2i-1} and v_{2i} are defined as $\pi_{2i-1} = \sigma \uparrow c_i \uparrow c_{m+1}$ and $\pi_{2i} = \sigma \uparrow c_{m+1} \uparrow c_i$. These preferences are consistent with the underlying Euclidean space:
 1. The distance from v_{2i-1} to all candidates except c_{m+2} is equal.
 2. The distance from v_{2i-1} to c_{m+2} is zero.
 3. Voter v_{2i} is closer to candidates c_{m+1} and c_i than to any other candidates, and is equidistant from all remaining ones.
 4. The distance from v_{2i} to candidates c_{m+1} and c_i is equal.

When $m = 2$ the instance lies within an equilateral triangle with vertices at $(1, 0, 0)$, $(0, 1, 0)$, and $(0, 0, 1)$, as shown in Figure 11. For all $1 \leq i, j \leq m$, we have

$$\begin{aligned} d(c_i, c_{m+1}) &= d(c_i, v_{2j-1}) \\ &= \sqrt{\left(\frac{1}{m+1}\right)^2 m + \left(\frac{m}{m+1}\right)^2} \\ &= \sqrt{\frac{m}{m+1}}, \end{aligned}$$

for all $1 \leq i, j \leq m$ (where $i \neq j$), we have

$$\begin{aligned} d(c_i, v_{2j}) &= \sqrt{\left(\frac{1}{2(m+1)}\right)^2 (m-1) + \left(\frac{m+2}{2(m+1)}\right)^2 + \left(\frac{2m+1}{2m+2}\right)^2} \\ &= \sqrt{\frac{5m+4}{4m+4}}, \end{aligned}$$

and for all $1 \leq i \leq m$, we have

$$\begin{aligned} d(c_i, v_{2i}) &= \sqrt{\left(\frac{1}{2(m+1)}\right)^2 m + \left(\frac{m}{2(m+1)}\right)^2} \\ &= \sqrt{\frac{m}{4(m+1)}}. \end{aligned}$$

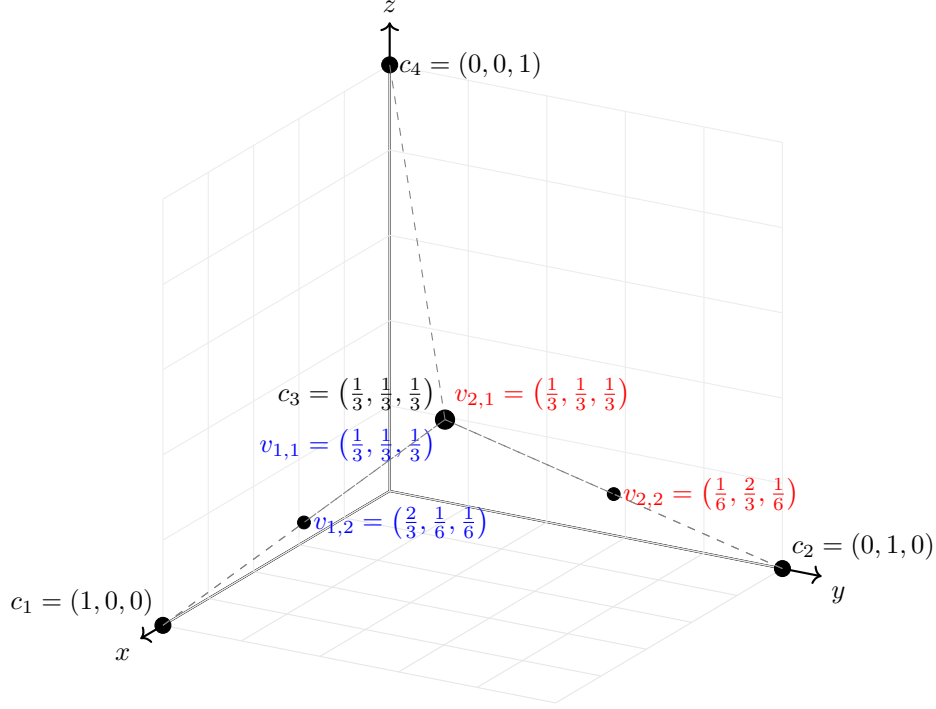


Figure 11: An illustration of the constructed instance when $m = 2$. The candidates are positioned at the corners and centroid of the 3D simplex (i.e., the equilateral triangle embedded in \mathbb{R}^3). Candidate c_3 is placed at the centroid, representing the candidate with high in-degree in $\mathcal{T}(f_{in}, \mathcal{C}, \sigma)$. Each group contains two voters: v_{2i-1} is located at the centroid, while v_{2i} is placed at the midpoint between c_3 and c_i for $i = 1, 2$. Different group voters are indicated via distinct colors.

By the definition of the avg-avg objective, we conclude that

$$\begin{aligned} \text{cost}(c_{m+1}) &= \frac{m \left(\frac{0 + \sqrt{\frac{m}{4(m+1)}}}{2} \right)}{m} = \frac{1}{4} \sqrt{\frac{m}{m+1}}, \\ \text{cost}(c_i) &= \frac{\frac{\sqrt{\frac{m}{m+1}} + \sqrt{\frac{5m+4}{4m+4}}}{2} (m-1) + \frac{\sqrt{\frac{m}{m+1}} + \sqrt{\frac{m}{4(m+1)}}}{2}}{m} \quad (1 \leq i \leq m). \end{aligned}$$

Clearly, the optimal candidate is c_{m+1} . By the definition of $\mathcal{T}(f_{in}, \mathcal{C}, \sigma)$, the representative of group i ($1 \leq i \leq m$) is c_i . Therefore, the mechanism selects the final winner from among the candidates c_1, c_2, \dots, c_m . A lower bound on the distortion of the mechanism Ψ is obtained as follows ($1 \leq i \leq m$):

$$\begin{aligned} D(\Psi) &\geq \frac{\text{cost}(c_i)}{\text{cost}(c_{m+1})} \\ &= 2 + \frac{m-1}{m} \sqrt{\frac{5m+4}{m}} + \frac{1}{m} \quad (1 \leq i \leq m), \end{aligned}$$

As $m \rightarrow \infty$, the ratio approaches $2 + \sqrt{5} \approx 4.236$. Therefore, for any $\varepsilon > 0$, we can construct an instance with distortion greater than $2 + \sqrt{5} - \varepsilon$. \square