# On the Distortion of Single Winner Elections with Aligned Candidates

Dimitris Fotakis, Laurent Gourvès

# On the Distortion of Single Winner Elections with Aligned Candidates

Dimitris Fotakis[1] Laurent Gourvès[2*]

1. National Technical University of Athens, Greece
fotakis@cs.ntua.gr
2. Université Paris Dauphine-PSL, CNRS, LAMSADE, 75016, Paris, France
laurent.gourves@dauphine.fr, *corresponding author

**Abstract**

We study the problem of selecting a single element from a set of candidates on which a group of agents has some spatial preferences. The exact distances between agent and candidate locations are unknown but we know how agents rank the candidates from the closest to the farthest. Whether it is desirable or undesirable, the winning candidate should either minimize or maximize its aggregate distance to the agents. The goal is to understand the optimal distortion, which evaluates how good an algorithm that determines the winner based only on the agent rankings performs against the optimal solution. We give a characterization of the distortion in the case of latent Euclidean distances such that the candidates are aligned, but the agent locations are not constrained. This setting generalizes the well-studied setting where both agents and candidates are located on the real line. Our bounds on the distortion are expressed with a parameter which relates, for every agent, the distance to her best candidate to the distance to any other alternative.

**Keywords:** Distortion, Single Winner Election, Obnoxious Facility

## 1 Introduction

The problem of electing a set of representatives is central in social choice theory. Some voters (a.k.a. agents) express their preferences over a set of candidates and one has to aggregate the voters' preferences to identify the winners (see e.g., [41]). In typical voting scenarios, the voters can only express *ordinal* preferences over the candidates, which are consistent and summarize their *cardinal* preferences. The reason for having ordinal data instead of cardinal data is that determining the numerical values is often a cognitively difficult task. For example, the voters and the candidates may occupy points in an unknown metric space, and every agent's true cost for a candidate is the distance between them. Though it is hard to obtain the exact distances to the candidates, it is undoubtedly easier for an agent to rank them from the closest to the farthest.

In a recent stream of articles (see, for example, [4] for a recent survey), researchers study problems where some agents have latent distances over a set of candidates but these distances are unknown. Nevertheless, each agent has reported a ranking of the candidates, from the closest to the farthest. Though these rankings are consistent with the latent distance function, we are not guaranteed to find the candidates whose aggregate distance to the agents is minimum, even if we aim to choose a single candidate [5].

Similar to the approximation ratio [39], the *distortion* measures the worst-case performance of an algorithm due to lack of cardinal information [37, 9]. The intriguing question of determining the best distortion for selecting a single winner (called the *metric distortion problem*) has attracted a lot of attention. For this problem, the Copeland voting rule has distortion 5 [2]. This result has been improved to $2 + \sqrt{5} \approx 4.236$ in [36]. Subsequently, Gkatzelis et al. [24] proposed a deterministic algorithm with distortion 3, which is optimal because no deterministic algorithm has distortion less than 3 [3, 2].

When randomization is possible, Anshelevich and Postl gave a lower bound of 2 and an algorithm whose expected distortion is $3 - 2/n$ for any number $n$ of voters [5]. Another randomized algorithm with distortion $3 - 2/m$ has been proposed by Kempe, where $m$ is the number of candidates [29, 30]. Therefore, the case $m = 2$ is resolved. Determining the best possible distortion for randomized algorithms is considered as a major open problem [4]. Recently, Chakirar and Ramakrishnan [13] gave improved lower bounds on the distortion for randomized algorithms (namely, $2.02613$ for $m = 3$, $2.04957$ for $m = 4$, and up to $2.11264$ when $m \to \infty$) using a family of metrics called $(0, 1, 2, 3)$-*metrics*. In their article, Chakirar and Ramakrishnan also resolved the case $m = 3$ (any election with 3 candidates has a randomized algorithm that guarantees distortion at most $2.02613$) and proposed nearly matching upper bounds for $(0, 1, 2, 3)$-metrics.

More insight into the problem can be gained when more information on the instance is available. In this respect, $\alpha$-*decisiveness*, where $\alpha$ is a real in $[0, 1]$, plays a key role [5]. This parameter captures how much more the agents prefer their best candidate to any other alternative. In an $\alpha$-decisive instance, every agent's distance to her closest candidate is at most the distance to her second closest candidate multiplied by $\alpha$. Then, every agent is co-located with her top choice when $\alpha = 0$. For the other extreme ($\alpha = 1$), $\alpha$-decisiveness does not constrain the agents' locations at all.

The algorithm of Gkatzelis et al. has distortion $2 + \alpha$ for $\alpha$-decisive instances with at least 3 candidates [24]. The deterministic lower bound of 3, which relies on a two-candidate instance, can be extended to show that when the number of candidates $m$ is at least 2, no deterministic algorithm has $\alpha$-distortion less than $1+2\alpha$. The upper and lower bounds do not match anymore under the $\alpha$-decisiveness framework, but Gkatzelis et al. proposed a lower bound which approaches $2 + \alpha$ when the number of candidates $m$ tends to infinity [24]. When $m = 2$, the deterministic algorithm which outputs the top choice of a majority of agents has distortion $1 + 2\alpha$ [5, 24]. Regarding randomized algorithms parameterized by $\alpha$, the best lower and upper bounds, for any number of candidates $m$, are $2 + \alpha - 2(1 - \alpha)/m$ and $2 + \alpha - 2/m$, respectively [24].

Besides $\alpha$-decisiveness, the metric distortion problem has been studied in the well-known case where agents and candidates are located on a real number line. The locations are unknown but the agents rank the candidates from the closest to the farthest. The preferences induced by this setting (a.k.a. 1-Euclidean because the distances are Euclidean and there is only one dimension) possess nice properties (namely, *single-peakedness* [8] and *single-crossingness* [28, 35]) which can be favorably exploited by an algorithm.[1] Anshelevich and Postl proposed a randomized algorithm with an optimal distortion of $1 + \alpha$ for $\alpha$-decisive instances on a line [5]. They exploit the possibility to efficiently identify a set of (at most) two candidates which are consecutive on the line and to which the optimum must belong. Regarding deterministic algorithms, the aforementioned lower bound of $1 + 2\alpha$ deriving from the lower bound of 3, applies to the case where agents and candidates are on a line. On the contrary, the candidates are not aligned in the lower bound approaching $2 + \alpha$ presented in [24].

Elections share similarities with $k$-median and facility location problems [11, 40, 23]. The goal is to choose a subset of candidate locations where desirable facilities (e.g., schools) can be built. The total distance to some given agent set has to be minimized, assuming that each agent is connected to the nearest facility. Sometimes, the candidate to be selected is undesirable (e.g., a garbage depot or a candidate to leave a group of people). In this case, one wants to select a candidate of *maximum* total distance to the agents (see [17] for a recent survey on *obnoxious* facility location). Obnoxious facility location problems have previously received attention from several viewpoints. In a "pure" optimization framework one wants to choose the location of the facilities and the true distances are accessible (see e.g., [38, 27] and the references therein). In the field of algorithmic mechanism design, the agents may misreport their preferences over the set of candidates so as maximize their individual distance to the winner(s). The authors of [16, 33, 34] pursue the goal of designing (group) strategyproof mechanisms[2] with the best possible approximation ratio. Recently, Chen et al. [15] studied the distortion of algorithms in a setting where the location of the candidates is known but the location of every agent is private.[3] They resolved the deterministic case for which the best distortion is 3. For randomized algorithms, a general lower bound of 1.5 is given, together with upper bounds for well studied special cases. In particular, they proposed two randomized mechanisms for building a single facility on the real line. The first mechanism is strategyproof and its distortion is 2. The second mechanism is not strategyproof but its distortion is

[1]We shall see that these properties hold under the mild assumption that no agent is equidistant from two distinct candidates.
[2]There is no incentive for a single agent or a group of agents to misreport their true rankings.
[3]In the present work, the location of the agents and the candidates are private.

| | $\bar{\alpha} < \frac{1}{3}$ | $\frac{1}{3} \leq \bar{\alpha} \leq \sqrt{2} - 1$ | $\sqrt{2} - 1 < \bar{\alpha}$ |
|---|---|---|---|
| $m > 2$ | 1 | $\frac{3\bar{\alpha} - \bar{\alpha}^2}{2 - 3\bar{\alpha} - \bar{\alpha}^2}$ | $1 + 2\bar{\alpha}$ |
| $m = 2$ | $1 + 2\bar{\alpha}$ | | |

Table 1: Distortion of $\bar{\alpha}$-decisive instances for selecting an undesirable candidate ($\sqrt{2} - 1 \approx 0.41$).

lower: $13/7$.

## Contribution and Organization

We consider the metric distortion problem in $\alpha$-decisive instances (defined in Section 2). The distances between agent and candidate locations are unknown but every agent has reported a strict preference over the candidate set. The influence of $\alpha$-decisiveness on the agents' locations is clear when $\alpha = 0$ or $\alpha = 1$, but no previous work precisely explains (to our best knowledge) how $\alpha$-decisiveness rules the agents' locations when $\alpha \in (0, 1)$. Our first contribution is to fill this gap by showing that agents lie inside some spheres under Euclidean distances (Lemma 1). This characterization is interesting on its own and we exploit it in the rest of the article.

Our second contribution is the definition of a domain which generalizes the well-studied case where both agents and candidates are located on the real line (1-Euclidean). In this generalization called AC for "Aligned Candidates" and defined in Section 3, the candidates are aligned but the agent locations are not constrained. As for the 1-Euclidean case, the distances in the AC setting are Euclidean. As an application of AC, one can think of a straight road that crosses a region. The agents can be located anywhere in the region but the candidates must be along the road. One can also interpret the AC domain from an electoral perspective: every candidate lies on a left right political axis while the voters' ideological positions are more complex and require more dimensions.

We demonstrate that, as for the 1-Euclidean domain and under the same mild assumption, the agent preferences remain single-peaked and single-crossing under the AC domain (Corollary 1). Hence, when one wants to select a desirable candidate from which the agents want to be as close as possible, the set of potential optima can be reduced to two contiguous candidates, as for the 1-Euclidean case [19, 5]. Since the metric distortion problem is resolved when $m = 2$ by selecting the candidate supported by a majority of agents [5, 24], we get a deterministic algorithm with distortion at most $1 + 2\alpha$ for any number of aligned candidates (Corollary 2). This is the best possible ratio because the aforementioned lower bound of $1 + 2\alpha$ applies to the setting of aligned candidates.

Afterwards, we investigate the distortion of choosing a single *undesirable* candidate (Section 4). The aim is to determine the candidate that *maximizes* the total distance to the agents. As opposed to the setting studied in [15], we do not consider that the location of the candidates are public. We generalize the notion of $\alpha$-decisiveness to the case of selecting an undesirable candidate (Definition 1). Namely, an instance is $\bar{\alpha}$-decisive, for some $\bar{\alpha} \in [0, 1]$, if every agent prefers her best candidate (now, this is the farthest candidate) at least $1/\bar{\alpha}$ times more than her second best (the second farthest). Though this definition reads similar to that of decisiveness for a desirable facility, $\bar{\alpha}$-decisiveness constrains the instances in a very different way (see the discussion of Section 4.1). We obtain tight bounds on the distortion of undesirable single winner election by deterministic algorithms, as a function of $\bar{\alpha}$, in two cases. When there are only two candidates and the latent distance function $d$ is a metric ($d$ is not necessarily Euclidean), we show in Section 4.2 that the simple algorithm which outputs the candidate ranked last by a majority of agents has distortion $1 + 2\bar{\alpha}$ (Theorem 1) and this is the best possible ratio (Proposition 1).

Section 4.3 deals with the AC domain with any number of candidates. As for the case of selecting a desirable candidate, the set of possible optima of the undesirable case with aligned candidates reduces to (at most) two elements which can be efficiently identified from the preference profile. However, these possible optima are totally different (for instance non-consecutive), except when $m = 2$. All our bounds for selecting an undesirable alternative are tight and summarized in Table 1. Regarding these bounds, note that $\frac{3\bar{\alpha} - \bar{\alpha}^2}{2 - 3\bar{\alpha} - \bar{\alpha}^2} = 1$ when $\bar{\alpha} = 1/3$, $\frac{3\bar{\alpha} - \bar{\alpha}^2}{2 - 3\bar{\alpha} - \bar{\alpha}^2} = 1 + 2\bar{\alpha}$ when $\bar{\alpha} = \sqrt{2} - 1$, and $\frac{3\bar{\alpha} - \bar{\alpha}^2}{2 - 3\bar{\alpha} - \bar{\alpha}^2} < 1 + 2\bar{\alpha}$ for all $\bar{\alpha} \in [1/3, \sqrt{2} - 1)$. Since $\bar{\alpha} \in [0, 1]$, the distortion is always below 3, which is consistent with the results of [15].

In fact, all our bounds on the distortion, for both selecting a desirable or undesirable candidate, are best possible and derive from the same simple algorithm: identify a set of two candidates containing the optimum and return the one that is supported by a majority of agents. We conclude in Section 5 with directions for future work.

## 2 The Model

There are some agents (a.k.a. voters) $\mathcal{N} = \{1, \ldots, n\}$, some candidates $\mathcal{C} = \{f_1, \ldots, f_m\}$, and a distance function $d : (\mathcal{N} \cup \mathcal{C})^2 \to [0, \infty)$. The function $d$ is a metric. For all $x, y, z \in \mathcal{N} \cup \mathcal{C}$, $d$ satisfies the following axioms.

- *identity:* $d(x, y) = 0 \Leftrightarrow x = y$,
- *symmetry:* $d(x, y) = d(y, x)$,
- *triangle inequality:* $d(x, z) \leq d(x, y) + d(y, z)$).

We will also consider the special case of Euclidean distances. In a $\delta$-dimensional Euclidean space, where $\delta$ is a positive integer, every point $p$ has $\delta$ coordinates $(p_1, \ldots, p_\delta)$ and the distance between two points $p$ and $q$ is defined as $\sqrt{\sum_{i=1}^{\delta}(p_i - q_i)^2}$.

Neither the location of any element of $\mathcal{N} \cup \mathcal{C}$ nor the distance between any two elements of $\mathcal{N} \cup \mathcal{C}$ is known.[4] Instead, every agent $i \in \mathcal{N}$ expresses a strict preference $\succ_i$ over $\mathcal{C}$ (also called ranking): $f \succ_i f'$ means that $f$ is closer to agent $i$ than $f'$. The preferences are consistent with $d$, namely

$$\forall i \in \mathcal{N},\ d(i, f) < d(i, f') \Longrightarrow f \succ_i f'. \tag{1}$$

For the moment we deliberately conceal how agents rank equidistant candidates, if such a case occurs.[5] However, we will see later in Section 3.1 that this aspect can have an impact on the validity of some properties of the preference profile , and whether such properties can be exploited by an algorithm.

As a notation, let $\succ$ be shorthand for the preference profile $(\succ_i)_{i \in \mathcal{N}}$. The input is an election $\langle \mathcal{N}, \mathcal{C}, \succ \rangle$ and we want to select a single candidate out of $\mathcal{C}$, called the *winner*. A standard interpretation is that every member of $\mathcal{C}$ is a desirable electoral candidate on which the agents have *spatial preferences* [21], and agents want to be as close as possible to the winner. The objective function, to be minimized, is the total sum of agent distances to the output $f$, i.e., $\min_{f \in \mathcal{C}} \sum_{i \in \mathcal{N}} d(i, f)$. Another interpretation is that one wants to build a desirable facility (e.g., a school) and every element of $\mathcal{C}$ is a candidate place.

**Example 1.** *Consider an instance with 3 candidates $\{f_1, f_2, f_3\}$, and 4 agents $\{a, b, c, d\}$ having the following preferences:*

$$\begin{array}{ccccc} f_1 & \succ_a & f_3 & \succ_a & f_2 \\ f_2 & \succ_b & f_3 & \succ_b & f_1 \\ f_3 & \succ_c & f_1 & \succ_c & f_2 \\ f_2 & \succ_d & f_3 & \succ_d & f_1 \end{array}$$

*The latent distance are unknown but they can, for example, derive from the locations depicted on Figure 1.*

Later in this article (Section 4), we consider the reversed problem of selecting an *undesirable* candidate from which the agents want to be as far as possible. Elements of this related problem are postponed and, for the moment, we concentrate on the standard minimization problem.

A *social choice function* $\mathcal{A}$, or simply *algorithm*, is a function of $\succ$. If $\mathcal{A}$ is deterministic, then it outputs a member of $\mathcal{C}$. If $\mathcal{A}$ is randomized, then it outputs a probability distribution over $\mathcal{C}$. As the distances are unknown, we cannot

[4]Note that the location of the candidates are public in [6, 15].

[5]Some previous works do not explicitly specify how to deal with ties probably because the input must contain strict preferences and ties are thus implicitly excluded. However, in [19, 2], the authors clearly state that no agent is equidistant from two candidates. In [24], the authors mention that candidates that are equidistant to an agent can be ranked arbitrarily by the agent.

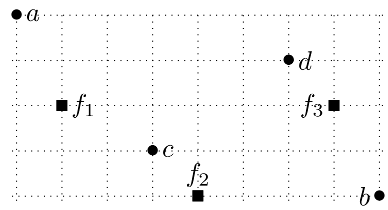


Figure 1: Illustration of Example 1 where the latent distances are Euclidean.

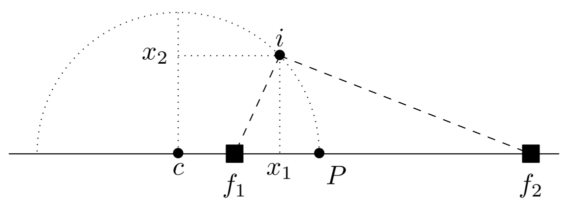


Figure 2: We have $\alpha \in (0,1)$, $d(c, f_1) = \frac{\alpha^2}{1-\alpha^2} d(f_1, f_2)$, and $d(f_1, P) = \frac{\alpha}{1+\alpha} d(f_1, f_2)$. Points $i$ such that $d(i, f_1) = \alpha d(i, f_2)$ describe a $(\delta - 1)$-dimensional sphere of center $c$ and radius $d(c, P) = \frac{\alpha}{1-\alpha^2} d(f_1, f_2)$.

expect $\mathcal{A}$ to output the optimal candidate. Nevertheless one can try to minimize its *distortion*. The distortion of a deterministic algorithm $\mathcal{A}$ is the maximum value taken by the ratio[6]

$$\frac{\sum_{i \in \mathcal{N}} d(i, \mathcal{A}(\succ))}{\sum_{i \in \mathcal{N}} d(i, \mathsf{opt}(\succ))} \tag{2}$$

over all possible elections $\langle \mathcal{N}, \mathcal{C}, \succ \rangle$, where $d$ is consistent with $\succ$, $\mathcal{A}(\succ)$ denotes the output of $\mathcal{A}$, and $\mathsf{opt}(\succ)$ denotes the optimum [37, 9].

Bounds on the distortion can also be expressed with a parameter called $\alpha$-decisiveness. Quoting Anshelevich and Postl [5], it is *a measure of how strongly an agent feels about her top preference relative to her second choice*. Formally, an instance is $\alpha$-decisive with $\alpha \in [0, 1]$ if $d(i, top(i)) \leq \alpha d(i, sec(i))$ holds for all $i \in \mathcal{N}$, where $top(i)$ and $sec(i)$ are the first and second elements of $\succ_i$, respectively. It follows from the definition that $d(i, top(i)) \leq \alpha d(i, f)$ holds for all $i \in \mathcal{N}$ and $f \in \mathcal{C} \setminus \{top(i)\}$. If $\alpha = 1$, then the instance is not constrained by the $\alpha$-decisiveness. The other extreme value ($\alpha = 0$) puts a hard constraint on the instance because it forces every agent to be located on her top choice. When $\alpha \in (0, 1)$, we note that points $i$ satisfying $d(i, top(i)) = \alpha d(i, sec(i))$ belong to a $(\delta - 1)$-dimensional sphere whose center is *not* $top(i)$ and the sphere's radius depends on $\alpha$.

**Lemma 1.** *Given $\alpha \in (0, 1)$ and two candidates $f_1$ and $f_2$, points $i$ such that $d(i, f_1) = \alpha d(i, f_2)$ describe a $(\delta - 1)$-dimensional sphere of radius $\frac{\alpha}{1-\alpha^2} d(f_1, f_2)$ and center $c$ such that $f_1$ belongs to the line segment $\overline{cf_2}$ and $d(c, f_1) = \frac{\alpha^2}{1-\alpha^2} d(f_1, f_2)$.*

*Proof.* We will suppose w.l.o.g. that $d(f_1, f_2) = 1$ since one can rescale the instance if it is not the case.

Consider a point $i$ located on the $(\delta - 1)$-dimensional sphere of radius $\frac{\alpha}{1-\alpha^2}$ and center $c$. There is also a point $P$ on the line segment $\overline{f_1 f_2}$ such that $d(f_1, P) = \frac{\alpha}{1+\alpha}$ and $d(P, f_2) = 1 - \frac{\alpha}{1+\alpha} = \frac{1}{1+\alpha}$. See Figure 2 for an illustration ($\delta = 2$).

Without loss of generality, the coordinates of $c$, $i$, $f_1$, $P$ and $f_2$ are $(0, \ldots, 0)$, $(x_1, \ldots, x_\delta)$, $(\frac{\alpha^2}{1-\alpha^2}, 0, \ldots, 0)$, $(\frac{\alpha}{1-\alpha^2}, 0, \ldots, 0)$, and $(1 + \frac{\alpha^2}{1-\alpha^2}, 0, \ldots, 0)$, respectively. Moreover, one can reason in the 2-dimensional subspace

[6]Technically, we consider that the distortion is 1 when both its numerator and denominator are 0. The distortion is infinite when its numerator is positive and its denominator is 0.

where $i$, $c$, $P$, $f_1$, and $f_2$ lie, thus assuming $x_j = 0$ for all $j \in \{3, \ldots, \delta\}$. Since $i$ is on the sphere, it holds that

$$\sum_{j=1}^{\delta} x_j^2 = x_1^2 + x_2^2 = \left(\frac{\alpha}{1-\alpha^2}\right)^2. \tag{3}$$

We have $d(i, f_1)^2 = \left(x_1 - \frac{\alpha^2}{1-\alpha^2}\right)^2 + x_2^2 = x_1^2 - 2x_1\frac{\alpha^2}{1-\alpha^2} + \left(\frac{\alpha^2}{1-\alpha^2}\right)^2 + x_2^2$. Use (3) to get that

$$\begin{aligned} d(i, f_1)^2 &= \left(\frac{\alpha}{1-\alpha^2}\right)^2 - \frac{2x_1\alpha^2}{1-\alpha^2} + \left(\frac{\alpha^2}{1-\alpha^2}\right)^2 \\ &= \frac{\alpha^2+\alpha^4}{(1-\alpha^2)^2} - \frac{2x_1\alpha^2}{1-\alpha^2}. \end{aligned} \tag{4}$$

We have $d(i, f_2)^2 = \left(1 + \frac{\alpha^2}{1-\alpha^2} - x_1\right)^2 + x_2^2 = \left(\frac{1}{1-\alpha^2}\right)^2 - \frac{2x_1}{1-\alpha^2} + x_1^2 + x_2^2$. Use (3) to get that $d(i, f_2)^2 = \left(\frac{1}{1-\alpha^2}\right)^2 - \frac{2x_1}{1-\alpha^2} + \left(\frac{\alpha}{1-\alpha^2}\right)^2 = \frac{1+\alpha^2}{(1-\alpha^2)^2} - \frac{2x_1}{1-\alpha^2}$. Thus,

$$\alpha^2 d(i, f_2)^2 = \frac{\alpha^2+\alpha^4}{(1-\alpha^2)^2} - \frac{2x_1\alpha^2}{1-\alpha^2}. \tag{5}$$

We deduce from (4) and (5) that $d(i, f_1) = \alpha d(i, f_2)$.

Now let us show that $d(i, f_1) = \alpha d(i, f_2)$ implies (3). We consider two cases: $x_1 \geq 0$ and $x_1 < 0$. When $x_1 \geq 0$, $d(i, f_1)^2 = \alpha^2 d(i, f_2)^2$ can be rewritten as follows.

$$\begin{aligned} \left(x_1 - \frac{\alpha^2}{1-\alpha^2}\right)^2 + x_2^2 &= \alpha^2\left(\left(1 + \frac{\alpha^2}{1-\alpha^2} - x_1\right)^2 + x_2^2\right) \\ x_1^2 - 2\frac{x_1\alpha^2}{1-\alpha^2} + \left(\frac{\alpha^2}{1-\alpha^2}\right)^2 + x_2^2 &= \alpha^2\left(\left(\frac{1}{1-\alpha^2}\right)^2 - 2\frac{x_1}{1-\alpha^2} + x_1^2 + x_2^2\right) \\ x_1^2 + \left(\frac{\alpha^2}{1-\alpha^2}\right)^2 + x_2^2 &= \frac{\alpha^2}{(1-\alpha^2)^2} + \alpha^2\left(x_1^2 + x_2^2\right) \\ (1-\alpha^2)\left(x_1^2 + x_1^2\right) &= (1-\alpha^2)\left(\frac{\alpha}{1-\alpha^2}\right)^2 \\ x_1^2 + x_1^2 &= \left(\frac{\alpha}{1-\alpha^2}\right)^2 \Leftrightarrow (3) \end{aligned}$$

In the above equations, we use the facts that $1 + \frac{\alpha^2}{1-\alpha^2} = \frac{1}{1-\alpha^2}$ and $1 - \alpha^2 > 0$ because $\alpha \in (0, 1)$. Using the same arguments, $d(i, f_1)^2 = \alpha^2 d(i, f_2)^2$ can be rewritten as follows when $x_1 < 0$.

$$\begin{aligned} \left(x_1 + \frac{\alpha^2}{1-\alpha^2}\right)^2 + x_2^2 &= \alpha^2\left(\left(\frac{1}{1-\alpha^2} + x_1\right)^2 + x_2^2\right) \\ x_1^2 + 2\frac{x_1\alpha^2}{1-\alpha^2} + \left(\frac{\alpha^2}{1-\alpha^2}\right)^2 + x_2^2 &= \alpha^2\left(\left(\frac{1}{1-\alpha^2}\right)^2 + 2\frac{x_1}{1-\alpha^2} + x_1^2 + x_2^2\right) \\ x_1^2 + \left(\frac{\alpha^2}{1-\alpha^2}\right)^2 + x_2^2 &= \frac{\alpha^2}{(1-\alpha^2)^2} + \alpha^2\left(x_1^2 + x_2^2\right) \\ (1-\alpha^2)\left(x_1^2 + x_1^2\right) &= (1-\alpha^2)\left(\frac{\alpha}{1-\alpha^2}\right)^2 \Leftrightarrow (3) \end{aligned}$$

In both cases, assuming $d(i, f_1) = \alpha d(i, f_2)$ implies that $i$ is on the sphere. $\square$

Lemma 1 implies that in an $\alpha$-decisive instance, the agents are located in a $(\delta - 1)$-dimensional ball. This fact will be exploited later in the article.

In the following, $\mathcal{I}_\alpha$ denotes the set of all $\alpha$-decisive instances. Thus, $\mathcal{I}_\alpha \subseteq \mathcal{I}_{\alpha'}$ holds for $0 \le \alpha \le \alpha' \le 1$. The $\alpha$-*distortion* of an algorithm $\mathcal{A}$ is defined as the largest value reached by ratio (2) over all instances of $\mathcal{I}_\alpha$.

# 3 Aligned Candidates

This section introduces a domain called "Aligned Candidates" but we need a detour before giving its formal definition.

## 3.1 Single-Peakedness and Single-Crossingness

In the $\delta$-Euclidean domain, there is a mapping $\mathtt{x} : \mathcal{N} \cup \mathcal{C} \rightarrow \mathbb{R}^\delta$, a distance $d(a,b) = \sqrt{\sum_{k=1}^{\delta}(\mathtt{x}_k(a) - \mathtt{x}_k(b))^2}$ where $\delta$ is a positive integer, and the preference of every agent follows from her distance to the candidates (closer is better). The 1-Euclidean domain is a well-studied special case where agents and candidates are located on the real line [5, 14, 19]. As an application, one can think of a street along which the candidates and the agents are located.

The 1-Euclidean domain is often said to be *single-crossing* and *single-peaked* (see for example [19, 14, 2] and references therein). These properties guarantee the existence of a Condorcet winner and many problems that are hard in a general election are tractable when single-crossingness or single-peakedness is satisfied.

Let $[k]$ denote $\{1, 2, \dots, k\}$ for every positive integer $k$. Being single-crossing [28, 35] means that there exists an ordering of the agents, say $1, 2, \dots, n$, having the following property: for every pair $f, f' \in \mathcal{C}$ satisfying $f \succ_1 f'$, there exists an index $\ell \in [n]$ such that $\{i \in \mathcal{N} : f \succ_i f'\} = [\ell]$.

A preference order is single-peaked [8] if it satisfies the following property for some linear order $\rhd$ over $\mathcal{C}$: for each three items $f_a, f_b, f_c \in \mathcal{C}$ such that $f_a \rhd f_b \rhd f_c$ or $f_c \rhd f_b \rhd f_a$, $f_a \succ_i f_b$ implies $f_b \succ_i f_c$. A preference profile is single-peaked if all its preferences are single-peaked w.r.t. the same linear order (also called axis).

Quoting [14], the argument supporting that the 1-Euclidean domain is single-crossing and single-peaked is that *the left-to-right ordering of the candidates along the Euclidean representation is single-peaked, and the left-to-right ordering of the agents along the Euclidean representation is single-crossing.* However we shall see two examples which demonstrate that the argument goes with a mild assumption (see [19, 2] for similar discussions).

Take a real line and suppose four candidates $f_1$, $f_2$, $f_3$ and $f_4$ have coordinates 0, 1, 3 and 4, respectively. There are also four co-located agents $A$, $B$, $C$, and $D$ whose common coordinate is 2. If, as done in [24], the equidistant candidates can be ranked arbitrarily by the agents, then the corresponding preference profile can be $\succ_A$: $f_2\, f_3\, f_1\, f_4$, $\succ_B$: $f_2\, f_3\, f_4\, f_1$, $\succ_C$: $f_3\, f_2\, f_1\, f_4$, $\succ_D$: $f_3\, f_2\, f_4\, f_1$, but it is *not* single-crossing.

In another example on a real line, three candidates $f_1$, $f_2$, and $f_3$ are co-located and four (not necessarily co-located) agents have preferences $\succ_A$: $f_1\, f_2\, f_3$, $\succ_B$: $f_1\, f_3\, f_2$, $\succ_C$: $f_2\, f_1\, f_3$, and $\succ_D$: $f_2\, f_3\, f_1$, which are *not* single-peaked.

In both examples, the difficulty originates from the presence of equidistant candidates. However, the 1-Euclidean domain is single-crossing and single-peaked under the following additional assumption [19, 2].

**Assumption 1.** *No agent is equidistant from two distinct candidates.*

Under this assumption, no tie-breaking rule is necessary and (1) suffices for fully deriving every agent's preference over $\mathcal{C}$ from her distance to the candidates. An immediate consequence of Assumption 1 is that candidates must occupy distinct locations. However, several agents can be co-located, and agents can be co-located with a candidate.

## 3.2 Definition and Properties of the AC Domain

The "Aligned Candidates" domain (AC domain in short) is at the same time a special case of the $\delta$-Euclidean domain and a generalization of the 1-Euclidean domain. Under the AC domain, all the candidates are on a line (called the candidate line thereafter). However, the agents are not constrained to be on the candidate line.

Concretely, there is a mapping $\mathtt{x} : \mathcal{N} \cup \mathcal{C} \rightarrow \mathbb{R}^\delta$ where $\delta \ge 1$, the distance $d$ is Euclidean, and we impose w.l.o.g. that $\mathtt{x}_k(f) = 0$ holds for all $f \in \mathcal{C}$ and $k \in \{2, \dots, \delta\}$.

For every agent $i \in \mathcal{N}$, let $\pi(i)$ denote a point such that $\mathtt{x}(\pi(i))$ is the orthogonal projection of $\mathtt{x}(i)$ onto the candidate line.

**Lemma 2.** *Under the AC domain,* $d(i,f) < d(i,f') \iff d(\pi(i),f) < d(\pi(i),f')$ *and* $d(i,f) = d(i,f') \iff d(\pi(i),f) = d(\pi(i),f')$ *hold for every* $i, f, f' \in \mathcal{N} \times \mathcal{C} \times \mathcal{C}$.

*Proof.* Consider the two dimensional space where $i$ and $\mathcal{C}$ lie. The Pythagorean theorem gives $d(i,f)^2 = d(i,\pi(i))^2 + d(\pi(i),f)^2$ and $d(i,f')^2 = d(i,\pi(i))^2 + d(\pi(i),f')^2$. We deduce that $d(i,f) < d(i,f') \iff d(i,f)^2 < d(i,f')^2 \iff d(\pi(i),f)^2 < d(\pi(i),f')^2 \iff d(\pi(i),f) < d(\pi(i),f')$. Moreover, $d(i,f) = d(i,f') \iff d(i,f)^2 = d(i,f')^2 \iff d(i,f)^2 - d(i,\pi(i))^2 = d(i,f')^2 - d(i,\pi(i))^2 \iff d(\pi(i),f)^2 = d(\pi(i),f')^2 \iff d(\pi(i),f) = d(\pi(i),f')$. ☐

It follows from Lemma 2 that locations $\mathtt{x}(i)$ and $\mathtt{x}(\pi(i))$ induce the same preference over $\mathcal{C}$. Therefore the AC domain has the same properties as the 1-Euclidean domain, but it also requires the same precaution regarding single-peakedness and single-crossingness.

**Corollary 1.** *The AC domain is single-crossing and single-peaked under Assumption 1.*

## 3.3 Selecting a Desirable Candidate Under the AC Domain

This section is devoted to the following result which is stated as corollary because its proof (given below) follows from previous results and Corollary 1.

**Corollary 2.** *Under the AC domain and Assumption 1, there exists a polynomial time deterministic algorithm with* $\alpha$*-distortion at most* $1 + 2\alpha$ *for any number of candidates.*

One can determine in polynomial time whether $\succ$ is single-peaked [7, 22], single-crossing [20, 10], or 1-Euclidean [18, 31, 19]. The proofs are constructive, relying on algorithms polynomial in $|\mathcal{N}|$ and $|\mathcal{C}|$. If $\succ$ is single-crossing, then the property holds for a *unique* ordering $\sqsubset$ of the agents (unique up to the reversal of the ordering, or the rearrangement of the agents having identical preferences) [19]. However, for single-peakedness, a consistent axis is not necessarily unique. Nevertheless, Elkind and Faliszewski observed that a part of $\rhd$ can be guessed [19, Proposition 2]. Namely, one can deduce from a single-peaked profile $\succ$ the ordering between the top candidates of the first and the last agents in $\sqsubset$. Anshelevich and Postl exploited this property to demonstrate that in the 1-Euclidean domain, one can reduce the set of possible optimal candidates to three consecutive alternatives [5, Lemma 7]. This result consists of considering the median agent in $\sqsubset$, her top candidate $f_X$, and the candidates $f_Y$ and $f_Z$ which are directly to the left and right of $f_X$. Afterwards, the set of possible optimal candidates is reduced to two consecutive alternatives (either $f_Y$ or $f_Z$ is removed) by comparing the number of agents who prefer $f_Y$ to $f_X$ with the number of agents who prefer $f_Z$ to $f_X$ [5, Lemma 8].

Since the AC domain has the same properties as the 1-Euclidean domain (Lemma 2), and both single-peakedness and single-crossingness are satisfied under Assumption 1 (Corollary 1), one can identify in polynomial time a set of two consecutive candidates $\{f_\ell, f_r\}$ which must include the optimum. Now we can exploit the fact that there exists a deterministic algorithm with distortion at most $1 + 2\alpha$ when there are only two candidates [5, 24]. Indeed, the algorithm which outputs the candidate of $\{f_\ell, f_r\}$ supported by a majority of agents has $\alpha$-distortion at most $1 + 2\alpha$.

This ratio of $1 + 2\alpha$ is best possible since the lower bound provided in [3, 2] for deterministic algorithms, which relies on a two-candidate instance (hence with aligned candidates), can be extended to a lower bound of $1 + 2\alpha$ for $\alpha$-decisive instances. Therefore, $1 + 2\alpha$ is the best possible distortion for a deterministic algorithm under the AC domain. The same goes for the 1-Euclidean domain because of the lower bound. The fact that $1 + 2\alpha$ is the best possible $\alpha$-distortion in both cases (for deterministic algorithms) relies on the possibility to reduce the set of possible optima to two candidates when only $\succ$ is known.

Because Lemma 2 uses orthogonal projections, one may believe that the result for the AC domain immediately reduces to the 1-Euclidean case. The reason would be that replacing an agent by her orthogonal projection onto the candidate line only increases the distortion, without changing the preferences or breaking the $\alpha$-decisiveness. However, the following example shows that this intuition is not correct.

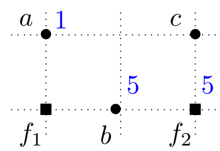


Figure 3: Illustration of Example 2.

**Example 2.** *(See Figure 2 for an illustration.) There are two candidates $f_1$ and $f_2$ located at $(0,0)$ and $(2,0)$, respectively. Points $a$ and $b$ are located at $(0,1)$ and $(1-\epsilon,0)$, respectively, with $1 \gg \epsilon > 0$. Suppose there is 1 agent on $a$, $5$ agents on $b$, and $5$ agents on $f_2$.*

A majority of agents prefers $f_1$ to $f_2$. However, $f_1$ is suboptimal and has distortion $\frac{16-5\epsilon}{5+5\epsilon+\sqrt{5}} \approx 2.21$. If the agent on $a$ is projected on her top choice $f_1$, then $f_1$ remains suboptimal and the distortion drops to $\frac{15-5\epsilon}{7+5\epsilon} \approx 2.14$.

On the contrary, there exist examples where projecting an agent on her top choice increases the distortion. Take the same instance and suppose there are 6 agents on $b$, 4 agents on $f_2$ and 1 agent on point $c$ which is located at $(2,1)$. The distortion increases if the agent on $c$ is projected on her top choice $f_2$.

In the next section, we explore the same problem but the goal is to select an *undesirable* candidate.

# 4 Selecting an Undesirable Candidate

This section departs from the previous one because we consider the problem of selecting an undesirable candidate (e.g., build a garbage depot or choose a candidate to leave a group of people) [38, 27, 16, 33, 34, 15]. In this case, one wants to select a candidate of *maximum* total distance to the agents. We still suppose that the agents have declared their preferences from the closest candidate to the farthest. However the distortion of a deterministic algorithm $\mathcal{A}$ for a preference profile $\succ$ is now defined as the maximum value taken by the ratio

$$\frac{\sum_{i\in\mathcal{N}} d(i,\mathsf{opt}(\succ))}{\sum_{i\in\mathcal{N}} d(i,\mathcal{A}(\succ))} \tag{6}$$

over all possible elections $\langle \mathcal{N},\mathcal{C},\succ\rangle$, where $d$ is consistent with $\succ$, $\mathcal{A}(\succ)$ denotes the output of $\mathcal{A}$, and $\mathsf{opt}(\succ)$ denotes the optimum. Compared to (2), the ratio is just reversed in order to keep its value above one.[7]

Our results rely on a parameter which is similar to $\alpha$. The purpose of $\alpha$-decisiveness is to quantify how good the best candidate is compared to the second best candidate (namely, the closest and second closest candidates, respectively). As we are now interested in selecting a candidate that should be as far as possible from the agents, there is a need to adapt the notion of $\alpha$-decisiveness. Let $last(i)$ and $stlast(i)$ be the last and the second to last elements of $\succ_i$. Thus, $d(i,last(i))$ is equal to $\max_{f\in\mathcal{C}} d(i,f)$.

**Definition 1.** *An instance is $\bar{\alpha}$-decisive if $\bar{\alpha}d(i,last(i)) \geq d(i,stlast(i))$ holds for all $i \in \mathcal{N}$, where $\bar{\alpha} \in [0,1]$.*

This means that $\bar{\alpha}d(i,last(i)) \geq d(i,f)$ holds for all $f \in \mathcal{C} \setminus \{last(i)\}$. Lemma 1 and Definition 1 imply that every agent is in a $(\delta-1)$-dimensional ball.

In the sequel, $\mathcal{I}_{\bar{\alpha}}$ denotes the set of all $\bar{\alpha}$-decisive instances. We say that the $\bar{\alpha}$-distortion of an algorithm $\mathcal{A}$ is the largest value reached by (6) over all instances of $\mathcal{I}_{\bar{\alpha}}$.

## 4.1 Discussion on Decisiveness

Though they look similar, the $\alpha$-decisiveness and the $\bar{\alpha}$-decisiveness do not constrain the instances in the same way.

[7] Again, we consider that the distortion is 1 when both its numerator and denominator are 0; it is infinite if the denominator is 0 but the numerator is positive.

Let us illustrate this fact with a simple example consisting of three aligned candidates $\{f_1, f_2, f_3\}$ such that $f_2$ is between $f_1$ and $f_3$, and one agent $i$ whose preference is $f_1 \succ_i f_2 \succ_i f_3$. Thus, $f_1$ and $f_3$ are the closest and farthest candidates, respectively. The $\alpha$-decisiveness gives:

$$\begin{aligned} d(i, f_1) &\leq \alpha d(i, f_2) && (7) \\ d(i, f_1) &\leq \alpha d(i, f_3) && (8) \end{aligned}$$

According to Lemma 1, Inequality (7) forces $i$ to belong to a $(\delta - 1)$-dimensional ball of radius $\frac{\alpha}{1-\alpha^2} d(f_1, f_2)$ and center $c$ such that $c$, $f_1$ and $f_2$ are aligned, and $d(c, f_1) = \frac{\alpha^2}{1-\alpha^2} d(f_1, f_2)$. Inequality (8) forces $i$ to belong to a $(\delta - 1)$-dimensional ball of radius $\frac{\alpha}{1-\alpha^2} d(f_1, f_3)$ and center $c'$ such that $c'$, $f_1$ and $f_3$ are aligned, and $d(c', f_1) = \frac{\alpha^2}{1-\alpha^2} d(f_1, f_3)$. Since $\alpha \in [0, 1]$, $f_1$ belongs to both balls. The first ball is included into the second one because $d(f_1, f_2) < d(f_1, f_3)$. In other words, in comparison to Inequality (7), Inequality (8) does not put any additional constraint on the location of $i$.

Now the $\bar{\alpha}$-decisiveness for the same example gives:

$$\begin{aligned} d(i, f_1) &\leq \bar{\alpha} d(i, f_3) && (9) \\ d(i, f_2) &\leq \bar{\alpha} d(i, f_3) && (10) \end{aligned}$$

Inequality (9) means that $i$ belongs to a $(\delta - 1)$-dimensional ball of radius $\frac{\bar{\alpha} d(f_1, f_3)}{1-\bar{\alpha}^2}$ and center $k$ such that $k$, $f_1$ and $f_3$ are aligned, and $d(k, f_1) = \frac{\bar{\alpha}^2 d(f_1, f_3)}{1-\bar{\alpha}^2}$. Inequality (10) forces $i$ to belong to a $(\delta - 1)$-dimensional ball of radius $\frac{\bar{\alpha}}{1-\bar{\alpha}^2} d(f_2, f_3)$ and center $k'$ such that $k'$, $f_2$ and $f_3$ are aligned, and $d(k, f_2) = \frac{\bar{\alpha}^2}{1-\bar{\alpha}^2} d(f_2, f_3)$. The ball associated with Inequality (9) is *not* necessarily included into the one associated with Inequality (10), so $i$ is located in the intersection of the two balls. As opposed to $\alpha$-decisiveness, one cannot guarantee that one of the three candidates belongs to both balls.

The given differences between $\alpha$-decisiveness and $\bar{\alpha}$-decisiveness have the following consequence. Suppose the locations of the candidates are given, together with a consistent preference profile $\succ$. For every $\alpha \in [0, 1]$, it is always possible to locate the agents such that the instance is $\alpha$-decisive, and the agents' preferences for the candidates are $\succ$: place every agent on her top choice. On the contrary, not every value of $\bar{\alpha}$ is possible if one imposes a preference profile consistent with a given location of the candidates. For example, one can observe that $\mathcal{I}_{\bar{\alpha}}$ is empty when $\bar{\alpha} = 0$ and the instance contains at least 3 candidates having distinct locations. To see this, consider an agent $i$, her farthest candidate $f_1$, and two other candidates $f_2$ and $f_3$. By the $\bar{\alpha}$-decisiveness, we have $\bar{\alpha} \geq \frac{\max(d(i,f_2), d(i,f_3))}{d(i,f_1)}$. Since $f_2$ and $f_3$ occupy distinct locations, $\max(d(i, f_2), d(i, f_3))$ is strictly positive; so is $\bar{\alpha}$. Nevertheless, $\mathcal{I}_{\bar{\alpha}}$ is non-empty for all $\bar{\alpha} \in [0, 1]$ when $m = 2$.

## 4.2 Two Candidate Instances

This section is devoted to the $\bar{\alpha}$-distortion in the case of only two candidates ($m = 2$) and the function $d$ is a metric. Thus, $d$ satisfies identity, symmetry, and the triangle inequality, but $d$ is not necessarily Euclidean. Preferences satisfy (1), but Assumption 1 is not made.[8] Ties between equidistant candidates can be broken arbitrarily.

We provide matching upper and lower bounds on the $\bar{\alpha}$-distortion of deterministic algorithms.

**Theorem 1.** *When $m = 2$, the deterministic algorithm which outputs the candidate appearing in the last position of a majority of agents (break ties arbitrarily) has $\bar{\alpha}$-distortion at most $1 + 2\bar{\alpha}$.*

*Proof.* Let $\mathcal{C} = \{f_1, f_2\}$. Let $\mathcal{N}_i$ be the set of agents who are closer to $f_i$ than $f_{3-i}$, with $i \in \{1, 2\}$. Suppose w.l.o.g. that $|\mathcal{N}_2| \leq |\mathcal{N}_1|$. Thus, the algorithm outputs $f_2$. Let us suppose $f_1$ is the optimal choice. One can upper bound the distortion $\frac{\sum_{i \in \mathcal{N}} d(i,f_1)}{\sum_{i \in \mathcal{N}} d(i,f_2)}$ as follows.

The instance being $\bar{\alpha}$-decisive, it holds that

$$\bar{\alpha} d(i, f_2) \geq d(i, f_1), \ \forall i \in \mathcal{N}_1. \tag{11}$$

[8] The two candidates can even be co-located

Since $(\mathcal{N}_1, \mathcal{N}_2)$ is a partition of $\mathcal{N}$, we have that

$$\sum_{i\in\mathcal{N}} d(i,f_1) = \sum_{i\in\mathcal{N}_1} d(i,f_1) + \sum_{i\in\mathcal{N}_2} d(i,f_1).$$

Use the triangle inequality to obtain

$$\sum_{i\in\mathcal{N}} d(i,f_1) \leq \sum_{i\in\mathcal{N}_1} d(i,f_1) + \sum_{i\in\mathcal{N}_2} \left(d(i,f_2) + d(f_1,f_2)\right).$$

Use (11) to get that

$$\sum_{i\in\mathcal{N}} d(i,f_1) \leq \bar{\alpha} \sum_{i\in\mathcal{N}_1} d(i,f_2) + \sum_{i\in\mathcal{N}_2} d(i,f_2) + |\mathcal{N}_2| d(f_1,f_2). \tag{12}$$

The fact that $|\mathcal{N}_2| \leq |\mathcal{N}_1|$ implies

$$|\mathcal{N}_2| d(f_1,f_2) \leq |\mathcal{N}_1| d(f_1,f_2). \tag{13}$$

Use the triangle inequality for every agent $i \in \mathcal{N}_1$ to get that

$$|\mathcal{N}_1| d(f_1,f_2) = \sum_{i\in\mathcal{N}_1} d(f_1,f_2) \leq \sum_{i\in\mathcal{N}_1} \left(d(i,f_1) + d(i,f_2)\right).$$

Inequality (13) becomes

$$|\mathcal{N}_2| d(f_1,f_2) \leq \sum_{i\in\mathcal{N}_1} \left(d(i,f_1) + d(i,f_2)\right).$$

Use (11) to obtain

$$|\mathcal{N}_2| d(f_1,f_2) \leq \bar{\alpha} \sum_{i\in\mathcal{N}_1} d(i,f_2) + \sum_{i\in\mathcal{N}_1} d(i,f_2). \tag{14}$$

Plug (14) into (12) to get that

$$\begin{aligned} \sum_{i\in\mathcal{N}} d(i,f_1) &\leq \sum_{i\in\mathcal{N}_2} d(i,f_2) + 2\bar{\alpha} \sum_{i\in\mathcal{N}_1} d(i,f_2) + \sum_{i\in\mathcal{N}_1} d(i,f_2) \\ &= \sum_{i\in\mathcal{N}} d(i,f_2) + 2\bar{\alpha} \sum_{i\in\mathcal{N}_1} d(i,f_2). \end{aligned}$$

Since $\sum_{i\in\mathcal{N}_1} d(i,f_2) \leq \sum_{i\in\mathcal{N}} d(i,f_2)$, it follows that $\sum_{i\in\mathcal{N}} d(i,f_1) \leq (1+2\bar{\alpha}) \sum_{i\in\mathcal{N}} d(i,f_2)$. Thus, the distortion is upper bounded by $1+2\bar{\alpha}$. □

**Proposition 1.** *When $m = 2$, any deterministic algorithm has $\bar{\alpha}$-distortion at least $1 + 2\bar{\alpha}$.*

*Proof.* Suppose there are two candidates $f_1$ and $f_2$, and two agents. Agent 1 has preference order $f_1 \succ_1 f_2$ and agent 2's preference order is $f_2 \succ_2 f_1$. Suppose $f_2$ is output (the case $f_1$ is symmetric). There is a consistent 1-Euclidean instance (both candidates and agents are on a line). The location of $f_1$, $f_2$, agent 1 and agent 2, are 0, $\bar{\alpha}+1$, $\bar{\alpha}$, and $\bar{\alpha}+1$, respectively.

The instance is $\bar{\alpha}$-decisive, and the distortion is $1+2\bar{\alpha}$. □

## 4.3 Aligned Candidates

We consider in this section the $\bar{\alpha}$-distortion under the AC domain. We have already seen that the AC domain is single-peaked under Assumption 1 (Corollary 1). One property of the single-peaked domain is that at most two candidates appear in the last positions of the agents' rankings, corresponding to the candidates on the extremities of the axis [7, 22]. The previous property is actually satisfied by the AC domain under the following weaker version of Assumption 1.[9]

[9] Under Assumption 2, there is a unique leftmost candidate and a unique rightmost candidate on the candidate line. The least preferred candidate of every agent (i.e., the farthest) must be one of them.

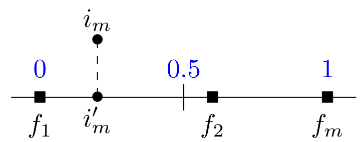


Figure 4: Illustration of the proof of Theorem 2.

**Assumption 2.** *Candidates occupy distinct locations.*

In this section, we consider the AC domain under Assumption 2, and equidistant candidates can be ranked arbitrarily by the agents.

If one candidate appears in the last position of all agents, then it must be optimal. Otherwise, there are two candidates in the last positions of the preferences. These candidates –let us denote them by $f_1$ and $f_m$– occupy the leftmost and rightmost positions of the candidate line. Thus, it is easy to identify $f_1$ and $f_m$ from the preference profile. Let us first observe that the optimal solution for maximizing the sum of the agents' distance must be either $f_1$ or $f_m$ under the AC domain.[10]

**Lemma 3.** *There exists $f^* \in \{f_1, f_m\}$ such that $\sum_{i \in \mathcal{N}} d(i, f^*) \geq \sum_{i \in \mathcal{N}} d(i, f)$ holds for all $f \in \mathcal{C}$.*

*Proof.* Take an agent $i \in \mathcal{N}$ and let $\pi(i)$ be its orthogonal projection onto the candidate line. The distance between $i$ and any candidate $f \in \mathcal{C}$ is $\sqrt{d(i, \pi(i))^2 + d(\pi(i), f)^2}$, where $d(i, \pi(i))^2$ is constant while $d(\pi(i), f)^2$ varies with the position of $f$ on the real line. The function $x \mapsto \sqrt{\kappa + x^2}$ being convex ($\kappa$ is a non-negative constant), the distance between $i$ and $f$ is a convex function of $f$'s location on the line. Since the sum of convex functions is also convex, we deduce that $\sum_{i \in \mathcal{N}} d(i, f)$ is a convex function of $f$'s location on the line. Therefore, $\sum_{i \in \mathcal{N}} d(i, f)$ finds its maximum on one of its extremities, namely $f_1$ or $f_m$. □

As for the case of selecting a desirable candidate, it follows that the number of optimal candidates to the problem of selecting an undesirable candidate under the AC domain can be reduced to 2, but these candidates are not the same.

We are going to characterize the best $\bar{\alpha}$-distortion for deterministic algorithms when selecting an undesirable candidate under the AC domain. When $m = 2$, the results of Section 4.2 imply that the best $\bar{\alpha}$-distortion for the AC domain is $1 + 2\bar{\alpha}$. From now on we consider the case of $m > 2$ candidates.

First we note that a distortion of 1 is possible when $\bar{\alpha}$ is small enough. Indeed, Theorem 2 states that all the agents agree on which candidate is the farthest when $\bar{\alpha} < 1/3$. Afterwards, Theorem 3 provides lower bounds when $1/3 \leq \bar{\alpha}$. We conclude with matching upper bounds (Theorem 4) achieved by the algorithm which outputs the farthest candidate for a majority of agents.

**Theorem 2.** *For $m > 2$ aligned candidates and $\bar{\alpha} < 1/3$, all agents agree on which candidate is the farthest.*

*Proof.* The extreme candidates are $f_1$ and $f_m$. By contradiction, suppose there is an agent $i_1$ for which $f_1 = last(i_1)$ and another agent $i_m$ for which $f_m = last(i_m)$. Since $m > 2$, there must be a third candidate $f_2$ on the line segment $\overline{f_1 f_m}$. Let us suppose w.l.o.g. that $d(f_1, f_2) \geq d(f_2, f_m)$.

For the sake of simplicity, we assume that the coordinates of $f_1$ and $f_m$ on the candidate line are 0 and 1, respectively. The coordinate of $f_2$ is in $[0.5, 1)$. Let $i'_m$ be the orthogonal projection of $i_m$ onto the line where the candidates lie. See Figure 4 for an illustration.

Suppose $d(f_1, i'_m) > 1/4$. Thus, $d(i'_m, f_m) = 1 - d(f_1, i'_m) < 3/4$. Since $\bar{\alpha} < 1/3$, the $\bar{\alpha}$-decisiveness gives $d(i_m, f_m) > 3d(i_m, f_1)$. In other words, $\sqrt{d(i_m, i'_m)^2 + d(i'_m, f_m)^2} > \sqrt{9d(i_m, i'_m)^2 + 9d(f_1, i'_m)^2}$. It follows that $d(i'_m, f_m)^2 > 8d(i_m, i'_m)^2 + 9d(f_1, i'_m)^2$. Moreover, $d(i'_m, f_m)^2 > 9d(f_1, i'_m)^2$ because $d(i_m, i'_m) \geq 0$. This is in contradiction with the hypotheses $3/4 > d(i'_m, f_m)$ and $d(i'_m, f_1) > 1/4$.

[10] See, for example, [16] and references therein for a similar result on a real line or a path.

Now suppose $d(f_1, i'_m) \le 1/4$ holds. The $\bar{\alpha}$-decisiveness, together with $\bar{\alpha} < 1/3$ give $d(i_m, f_m) > 3d(i_m, f_2)$. Combined with $d(i_m, f_2)+d(f_2, f_m) \ge d(i_m, f_m)$ which follows from the triangle inequality, we get that $d(f_2, f_m) > 2d(i_m, f_2)$. Since $i'_m$ is the orthogonal projection of $i_m$, we know that $d(i_m, f_2) \ge d(i'_m, f_2)$, leading to

$$d(f_2, f_m) > 2d(i'_m, f_2). \tag{15}$$

Since $f_2$ is in $[0.5, 1)$, we have

$$0.5 \ge d(f_2, f_m) \tag{16}$$

and $d(f_1, f_2) \ge 0.5$. This last inequality with $d(f_1, i'_m) \le 0.25$ give

$$d(i'_m, f_2) = d(f_1, f_2) - d(f_1, i'_m) \ge 0.25. \tag{17}$$

Inequalities (15), (16), and (17) lead to a contradiction.

To conclude, if $f_2 \in (0, 0.5]$, then switch the role of $i_m$ and $f_m$ with $i_1$ and $f_1$. □

Note that Theorem 2 cannot be extended to the case $\bar{\alpha} = 1/3$ because of the following 1-Euclidean instance. Consider a real line with three candidates at coordinates 0, 0.5 and 1, respectively, and two agents at coordinates 0.25 and 0.75, respectively.

**Theorem 3.** *In the presence of $m > 2$ aligned candidates, any deterministic algorithm has $\bar{\alpha}$-distortion at least $\frac{3\bar{\alpha}-\bar{\alpha}^2}{2-3\bar{\alpha}-\bar{\alpha}^2}$ when $1/3 < \bar{\alpha} \le \sqrt{2}-1$, and at least $1+2\bar{\alpha}$ when $\sqrt{2}-1 < \bar{\alpha} \le 1$.*

*Proof.* Suppose there are $m$ candidates $f_1, \dots, f_m$, and two agents. The candidates are on a line, placed by ascending index from left to right. Thus, the extremities are $f_1$ and $f_m$. Agent 1 has preference order $f_1 \succ_1 f_2 \succ_1 \cdots \succ_1 f_{m-1} \succ_1 f_m$ and agent 2's preference order is $f_m \succ_2 f_{m-1} \succ_2 \cdots \succ_2 f_2 \succ_2 f_1$. Suppose $f' \in \mathcal{C} \setminus \{f_1\}$ is output (the case $f' \in \mathcal{C} \setminus \{f_m\}$ is symmetric).

Let us describe a consistent 1-Euclidean instance: each element $e$ (i.e., candidate, agent or point) has a coordinate $\mathtt{x}(e) \in \mathbb{R}$. See Figure 5 for an illustration.

Suppose $\mathtt{x}(f_1) = 2\delta$ for some $\delta > 0$ and $\mathtt{x}(f_m) = 1+\bar{\alpha}$. Candidates $f_2, \dots, f_{m-1}$ are on the line segment $\overline{Pf_m}$ and their coordinates are in the interval $[2\bar{\alpha}-\delta, 2\bar{\alpha})$. There is a point $P$ such that $\mathtt{x}(P) = \bar{\alpha}$ where agent 1 is located. The instance is $\bar{\alpha}$-decisive for agent 1 because her farthest candidate is $f_m$ and $d(P, f) \le \bar{\alpha}d(P, f_m) = \bar{\alpha}$ holds for all $f \in \mathcal{C} \setminus \{f_m\}$. The position of agent 2 depends on whether $1/3 < \bar{\alpha} \le \sqrt{2}-1$ or $\sqrt{2}-1 < \bar{\alpha} \le 1$.

- Case $1/3 < \bar{\alpha} \le \sqrt{2}-1$. Agent 2 is on a point $Q$ of coordinate $\frac{2\bar{\alpha}-2\bar{\alpha}\delta}{1-\bar{\alpha}}$, between coordinates $2\bar{\alpha}$ and $1+\bar{\alpha}$ such that $f_1$ and $f_2$ are agent 2's farthest and second farthest candidates, respectively. Since $\bar{\alpha} \le \sqrt{2}-1$, we do have $\mathtt{x}(Q) \le \mathtt{x}(f_m)$. The fact that $f_2$ is the second farthest candidate imposes $d(Q, f_m) \le d(Q, f_2) \Leftrightarrow 1+\bar{\alpha}-\mathtt{x}(Q) \le \mathtt{x}(Q)-2\bar{\alpha}+\delta \Leftrightarrow (1+3\bar{\alpha}-\delta)/2 \le \mathtt{x}(Q)$. Since $\bar{\alpha} > 1/3$, there exists $\delta > 0$ such that $\mathtt{x}(Q) = \frac{2\bar{\alpha}-2\bar{\alpha}\delta}{1-\bar{\alpha}} \ge (1+3\bar{\alpha}-\delta)/2$. The instance is $\bar{\alpha}$-decisive for agent 2 because $\bar{\alpha}(\mathtt{x}(Q)-2\delta) = \mathtt{x}(Q)-(2\bar{\alpha}-\delta)$.

The output $f'$ of the algorithm is in $\{f_2, \dots, f_m\}$. The total distance to the agents is maximized (this is the worst case for deriving a lower bound) when $f_m$ is output because $f_m$ is within $\{f_2, \dots, f_m\}$, the only candidate outside $[P, Q]$. We have $d(P, f_m)+d(Q, f_m) = 1+1+\bar{\alpha}-\frac{2\bar{\alpha}-2\bar{\alpha}\delta-\delta}{1-\bar{\alpha}}$. The optimal choice is $f_1$ and $d(P, f_1)+d(Q, f_1) = \bar{\alpha}-2\delta+\frac{2\bar{\alpha}-2\bar{\alpha}\delta-\delta}{1-\bar{\alpha}}-2\delta$. Therefore, the $\bar{\alpha}$-distortion tends to $\frac{3\bar{\alpha}-\bar{\alpha}^2}{2-3\bar{\alpha}-\bar{\alpha}^2}$ when $\delta$ goes to 0.

- Case $\sqrt{2}-1 < \bar{\alpha} \le 1$. Agent 2 is co-located with $f_m$. The instance is $\bar{\alpha}$-decisive for agent 2 because her farthest candidate ($f_1$) is at distance $1+\bar{\alpha}-2\delta$, and her second farthest candidate ($f_2$) is at distance at most $1+\bar{\alpha}-(2\bar{\alpha}-\delta) = 1-\bar{\alpha}+\delta$. Since $\bar{\alpha} > \sqrt{2}-1$, there exists $\delta > 0$ such that $\bar{\alpha}(1+\bar{\alpha}-2\delta) \ge 1-\bar{\alpha}+\delta$.

By construction, the output $f'$ of the algorithm is on the line segment $\overline{Pf_m}$ whose length is 1. Thus, $\sum_{i\in\mathcal{N}} d(i, f') = d(P, f_m) = 1$ for all $f' \in \mathcal{C} \setminus \{f_1\}$. The optimal choice is $f_1$, and $\sum_{i\in\mathcal{N}} d(i, f_1) = \bar{\alpha}-2\delta+1+\bar{\alpha}-2\delta$. Therefore, the $\bar{\alpha}$-distortion tends to $1+2\bar{\alpha}$ when $\delta$ goes to 0. □

Now we turn our attention to matching upper bounds. Regarding the following theorem, note that $\frac{3\bar{\alpha}-\bar{\alpha}^2}{2-3\bar{\alpha}-\bar{\alpha}^2} = 1$ when $\bar{\alpha} = 1/3$, $\frac{3\bar{\alpha}-\bar{\alpha}^2}{2-3\bar{\alpha}-\bar{\alpha}^2} = 1+2\bar{\alpha}$ when $\bar{\alpha} = \sqrt{2}-1$, and $\frac{3\bar{\alpha}-\bar{\alpha}^2}{2-3\bar{\alpha}-\bar{\alpha}^2} < 1+2\bar{\alpha}$ for all $\bar{\alpha} \in [1/3, \sqrt{2}-1)$.

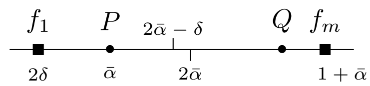


Figure 5: Illustration of the proof of Theorem 3.

**Theorem 4.** *When there are $m > 2$ aligned candidates, the $\bar{\alpha}$-distortion of the algorithm which outputs the candidate that a majority of agents places in last position (break ties arbitrarily) is 1 if $\bar{\alpha} < \frac{1}{3}$, $\frac{3\bar{\alpha}-\bar{\alpha}^2}{2-3\bar{\alpha}-\bar{\alpha}^2}$ if $\frac{1}{3} \leq \bar{\alpha} < \sqrt{2}-1$, and $1+2\bar{\alpha}$ if $\sqrt{2}-1 \leq \bar{\alpha}$.*

*Proof.* There are $m$ candidates $f_1, \dots, f_m$ with $m > 2$. Let us first observe that the $\bar{\alpha}$-distortion is at most $1+2\bar{\alpha}$ for all possible value of $\bar{\alpha}$. This is due to the proof of Theorem 1 which can be reproduced: identify $f_1$ and $f_2$ in the proof of Theorem 1 with the extreme candidates $f_1$ and $f_m$. This gives us the desired upper bound for the case $\sqrt{2}-1 \leq \bar{\alpha}$.

When $\bar{\alpha} < 1/3$, it is immediate from Theorem 2 that returning the last candidate of all preference orders has $\bar{\alpha}$-distortion 1.

From now on, we suppose that $1/3 \leq \bar{\alpha} < \sqrt{2}-1$, and the agents do not agree on which candidate is the farthest. Since the preferences are single-peaked, and the element to be output is undesirable, every agent considers either $f_1$ or $f_m$ (the leftmost and rightmost candidates) as her farthest candidate.

Every candidate $f' \in \mathcal{C}\setminus\{f_1, f_m\}$ (there is at least one such candidate) is on the line segment $\overline{f_1 f_m}$. As explained in the discussion (Section 4.1), the presence of $f' \in \mathcal{C}\setminus\{f_1, f_m\}$ combined with the $\bar{\alpha}$-decisiveness add some constraints on where the agents can be located. Therefore, the more candidates there are in $\mathcal{C}\setminus\{f_1, f_m\}$, the smaller is the set of possible instances. Since the $\bar{\alpha}$-distortion derives from a worst-case analysis over all possible instances, any upper bound for the case $m = 3$ applies to the case $m > 2$.

Therefore, we can restrict ourselves to the case of 3 aligned candidates $f_\ell$, $f_b$, and $f_r$, where $\ell$, $b$ and $r$ stand for *left*, *between*, and *right*, respectively.

Following Lemma 3, the optimum is either $f_l$ or $f_r$. Suppose w.l.o.g. that $f_r$ is returned by the algorithm whereas $f_\ell$ is the optimum. Concretely, $n_1$ agents declare that $f_r$ is the farthest, $n_2$ agents declare that $f_\ell$ is the farthest, $n = n_1 + n_2$, and $n_1 \geq n_2$.

The algorithm only uses the candidates appearing in last position of the agents' preferences. Therefore, we can suppose w.l.o.g. that in an instance with largest possible distortion, the $n_1$ agents for which $f_r$ is the farthest candidate are co-located. Indeed, if two agents $i$ and $i'$ have distinct locations but they agree on their farthest candidate, then moving $i$ to the location of $i'$ cannot increase the distortion (because the largest possible distortion is already reached) or decrease it (otherwise moving $i'$ to the location of $i$ would increase the distortion). Using similar arguments, we can also suppose w.l.o.g. that the $n_2$ agents for which $f_\ell$ is the farthest candidate are co-located.

In all, the worst case distortion appears in a 3 candidate instance where $n_1$ agents, all located on a point that we denote by $P_1$, declare that $f_r$ is their farthest candidate, and $n_2 = n - n_1$ agents, all located on a point that we denote by $P_2$, declare that $f_\ell$ is their farthest candidate. Since $f_\ell$ and $f_r$ occupy distinct locations, $d(f_\ell, f_r)$ is positive and we suppose w.l.o.g. that $d(f_\ell, f_r) = 1$ (rescale the instance if it is not the case). Thus, we can suppose that $\mathtt{x}_1(f_\ell) = 0$, $\mathtt{x}_1(f_r) = 1$, and $0 < \mathtt{x}_1(f_b) < 1$.

The $\bar{\alpha}$-decisiveness gives the following constraints:

$$\bar{\alpha} d(P_1, f_r) \geq d(P_1, f_\ell) \quad (18)$$
$$\bar{\alpha} d(P_1, f_r) \geq d(P_1, f_b) \quad (19)$$
$$\bar{\alpha} d(P_2, f_\ell) \geq d(P_2, f_r) \quad (20)$$
$$\bar{\alpha} d(P_2, f_\ell) \geq d(P_2, f_b) \quad (21)$$

Following Lemma 1, Inequalities (18) and (19) impose that $P_1$ lies in the intersection of two $(\delta-1)$-dimensional balls. Similarly, Inequalities (20) and (21) force $P_2$ to be in the intersection of two $(\delta-1)$-dimensional balls. See Figure 6 for an illustration.

In order to upper bound the distortion, we shall analyze the situation where $P_1$ and $P_2$ are as close as possible to $f_r$ (the output of the algorithm) and, at the same time, as far as possible from $f_\ell$ (the optimum).

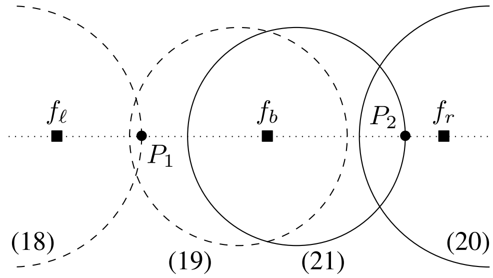


Figure 6: Illustration of Theorem 4. The dashed spheres correspond to Inequalities (18) and (19), respectively. The solid spheres correspond to Inequalities (20) and (21), respectively.

Concerning $P_1$, it consists of sliding $f_b$ towards $f_r$ on the line segment $\overline{f_\ell f_r}$ as much as possible, i.e., until Inequalities (18) and (19) leave a single feasible point, and by the fact that (18) and (19) are two balls, this point must be on the line segment $\overline{f_\ell f_b}$.

Using $d(f_\ell, P_1) = \mathtt{x}_1(P_1)$, $d(P_1, f_r) = d(f_\ell, f_r) - d(f_\ell, P_1) = 1 - \mathtt{x}_1(P_1)$, and (18), we get that $\mathtt{x}_1(P_1) \leq \frac{\bar{\alpha}}{1+\bar{\alpha}}$, so we can fix

$$\mathtt{x}_1(P_1) = \frac{\bar{\alpha}}{1+\bar{\alpha}},$$

and $\mathtt{x}_j(P_1) = 0$ for all $j \neq 1$. It follows that $d(P_1, f_r) = 1 - \frac{\bar{\alpha}}{1+\bar{\alpha}}$. Using (19), we get that $\bar{\alpha}(1 - \frac{\bar{\alpha}}{1+\bar{\alpha}}) \geq d(P_1, f_b) = \mathtt{x}_1(f_b) - \mathtt{x}_1(P_1) = \mathtt{x}_1(f_b) - \frac{\bar{\alpha}}{1+\bar{\alpha}}$. This gives us $\frac{2\bar{\alpha}}{1+\bar{\alpha}} \geq \mathtt{x}_1(f_b)$ so we can fix

$$\mathtt{x}_1(f_b) = \frac{2\bar{\alpha}}{1+\bar{\alpha}}.$$

Since $f_b$ is on the candidate line, $\mathtt{x}_j(f_b) = 0$ for all $j \neq 1$.

Now we can observe that $P_2$ cannot be co-located with $f_r$. Indeed, by contradiction, (21) gives $\bar{\alpha} d(f_r, f_\ell) \geq d(f_r, f_b)$, i.e., $\bar{\alpha} \geq 1 - \mathtt{x}_(f_b) = 1 - \frac{2\bar{\alpha}}{1+\bar{\alpha}}$ which is equivalent to $\bar{\alpha}^2 + 2\bar{\alpha} - 1 \geq 0$. However, $\bar{\alpha}^2 + 2\bar{\alpha} - 1 \geq 0$ is not valid because $\bar{\alpha} < \sqrt{2} - 1$.

Therefore, $f_r$ is outside the ball defined by Inequality (21). Since we analyze the situation where $P_2$ is as close as possible to $f_r$, as far as possible from $f_\ell$, and within the ball defined by Inequality (21), we deduce that $P_2$ must be on the line segment $\overline{f_b f_r}$ and at the boundary of the ball defined by Inequality (21). In other words,

$$\begin{aligned}
\bar{\alpha} d(P_2, f_\ell) &= d(P_2, f_b) \\
\bar{\alpha} \mathtt{x}_1(P_2) &= \mathtt{x}_1(P_2) - \mathtt{x}_1(f_b) \\
\bar{\alpha} \mathtt{x}_1(P_2) &= \mathtt{x}_1(P_2) - \frac{2\bar{\alpha}}{1+\bar{\alpha}} \\
\mathtt{x}_1(P_2) &= \frac{2\bar{\alpha}}{1-\bar{\alpha}^2}
\end{aligned}$$

and $\mathtt{x}_j(P_2) = 0$ for all $j \neq 1$. One can verify that $P_2$ is in the ball defined by Inequality (20).

$$\begin{aligned}
\bar{\alpha} d(P_2, f_\ell) &\geq d(P_2, f_r) \\
\bar{\alpha} \frac{2\bar{\alpha}}{1-\bar{\alpha}^2} &\geq 1 - \frac{2\bar{\alpha}}{1-\bar{\alpha}^2} \\
3\bar{\alpha}^2 + 2\bar{\alpha} - 1 &\geq 0
\end{aligned}$$

This last inequality holds because $\bar{\alpha} \geq 1/3$.

By considering $n_1$ agents on $P_1$ of coordinates $(\frac{\bar{\alpha}}{1+\bar{\alpha}}, 0, \ldots, 0)$, and $n_2$ agents on $P_2$ of coordinates $(\frac{2\bar{\alpha}}{1-\bar{\alpha}^2}, 0, \ldots, 0)$, we get an upper bound on the $\bar{\alpha}$-distortion of

$$\frac{n_1 \frac{\bar{\alpha}}{1+\bar{\alpha}} + n_2 \frac{2\bar{\alpha}}{1-\bar{\alpha}^2}}{n_1(1 - \frac{\bar{\alpha}}{1+\bar{\alpha}}) + n_2(1 - \frac{2\bar{\alpha}}{1-\bar{\alpha}^2})}.$$

It finds its maximum when $n_2$ is maximum, namely $n_1 = n_2 = n/2$. We get a ratio of $\frac{3\bar{\alpha}-\bar{\alpha}^2}{2-3\bar{\alpha}-\bar{\alpha}^2}$. □

# 5 Conclusion and Future Work

We considered the problem of selecting a single candidate on the basis of the agents' rankings when the latent distances derive from a configuration where the candidates are aligned. In both cases (desirable or undesirable), the set of optima can be reduced to two elements, and choosing the one supported by a majority of agents leads to the best possible deterministic distortion (as a function of $\alpha$ or $\bar{\alpha}$).

A natural next step is to consider the distortion of randomized algorithms. For a desirable facility on a line, the best distortion of randomized algorithms is $1 + \alpha$ [5], but the algorithm assumes knowledge of $\alpha$. In Proposition 2 of the appendix, we show a lower bound of $\frac{1+2\bar{\alpha}}{1+\bar{\alpha}}$ on the distortion of randomized algorithms for an undesirable facility. An interesting question is whether the lower bounds above can be matched by randomized algorithms that do not assume any knowledge of $\alpha$ (resp., $\bar{\alpha}$) whatsoever.

Other interesting research directions regarding obnoxious facility location in general metrics include either to extend the main result of [24] to the selection of an undesirable candidate (i.e., without assuming that the candidates are aligned) or, as in [6, 15], to consider the case where the location of the candidates is known, but the agent locations are unknown.

The amount of information available influences distortion [26, 30]. Improved upper bounds on the distortion may be obtained for aligned candidates, if we could exploit either the full cardinal preferences profile of a selected agent or the cardinal values of the agents' top few candidates. An interesting direction would be to query the agent cardinal preferences, so as to gain insight into the instance (e.g., [1, 32] follow this direction) or to learn the $\alpha$ (resp., $\bar{\alpha}$) parameter. Can these queries help us to drop Assumption 1 or 2?

We focused on a social cost function defined as the sum of the agents' distances to the winner. One can think of other objective functions such as the maximum or median distance over the agents [5]. Finally, a possible future work concerns the problem of selecting multiple winners [12, 25].

**Acknowledgements:** We thank anonymous reviewers for their valuable comments on the preliminary version of this work. Laurent Gourvès is supported by Agence Nationale de la Recherche (ANR), project THEMIS ANR-20-CE23-0018.

# References

[1] Georgios Amanatidis, Georgios Birmpas, Aris Filos-Ratsikas, and Alexandros A. Voudouris. Peeking behind the ordinal curtain: Improving distortion via cardinal queries. In *Proceedings of the 34th AAAI Conference on Artificial Intelligence*, pages 1782–1789, 2020.

[2] Elliot Anshelevich, Onkar Bhardwaj, Edith Elkind, John Postl, and Piotr Skowron. Approximating optimal social choice under metric preferences. *Artificial Intelligence*, 264:27–51, 2018.

[3] Elliot Anshelevich, Onkar Bhardwaj, and John Postl. Approximating optimal social choice under metric preferences. In *Proceedings of the 29th AAAI Conference on Artificial Intelligence*, pages 777–783, 2015.

[4] 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 the 30th International Joint Conference on Artificial Intelligence, IJCAI 2021, Virtual Event / Montreal, Canada, 19-27 August 2021*, pages 4294–4301, 2021.

[5] Elliot Anshelevich and John Postl. Randomized social choice functions under metric preferences. *Journal of Artificial Intelligence Research*, 58:797–827, 2017.

[6] Elliot Anshelevich and Wennan Zhu. Ordinal approximation for social choice, matching, and facility location problems given candidate positions. In *Web and Internet Economics - 14th International Conference, WINE 2018, Oxford, UK, December 15-17, 2018, Proceedings*, pages 3–20, 2018.

[7] John Bartholdi and Michael A. Trick. Stable matching with preferences derived from a psychological model. *Operations Research Letters*, 5(4):165–169, 1986.

[8] Duncan Black. On the rationale of group decision-making. *Journal of Political Economy*, 56(1):23–34, 1948.

[9] 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.

[10] Robert Bredereck, Jiehua Chen, and Gerhard J. Woeginger. A characterization of the single-crossing domain. *Social Choice and Welfare*, 41(4):989–998, 2013.

[11] Jaroslaw Byrka, Thomas W. Pensyl, Bartosz Rybicki, Aravind Srinivasan, and Khoa Trinh. An improved approximation for $k$-median and positive correlation in budgeted optimization. *ACM Transactions on Algorithms*, 13(2):23:1–23:31, 2017.

[12] Ioannis Caragiannis, Swaprava Nath, Ariel D. Procaccia, and Nisarg Shah. Subset selection via implicit utilitarian voting. *Journal of Artificial Intelligence Research*, 58:123–152, 2017.

[13] Moses Charikar and Prasanna Ramakrishnan. Metric distortion bounds for randomized social choice. In *Proceedings of the 2022 ACM-SIAM Symposium on Discrete Algorithms, SODA 2022, Virtual Conference / Alexandria, VA, USA, January 9 - 12, 2022*, pages 2986–3004, 2022.

[14] Jiehua Chen, Kirk Pruhs, and Gerhard J. Woeginger. The one-dimensional euclidean domain: finitely many obstructions are not enough. *Social Choice and Welfare*, 48(2):409–432, 2017.

[15] Xujin Chen, Minming Li, and Chenhao Wang. Favorite-candidate voting for eliminating the least popular candidate in a metric space. In *The Thirty-Fourth AAAI Conference on Artificial Intelligence, AAAI 2020, The Thirty-Second Innovative Applications of Artificial Intelligence Conference, IAAI 2020, The Tenth AAAI Symposium on Educational Advances in Artificial Intelligence, EAAI 2020, New York, NY, USA, February 7-12, 2020*, pages 1894–1901, 2020.

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

[17] Richard L. Church and Zvi Drezner. Review of obnoxious facilities location problems. *Computers & Operations Research*, 138:105468, 2022.

[18] Jean-Paul Doignon and Jean-Claude Falmagne. A polynomial time algorithm for unidimensional unfolding representations. *Journal of Algorithms*, 16(2):218–233, 1994.

[19] Edith Elkind and Piotr Faliszewski. Recognizing 1-Euclidean preferences: An alternative approach. In *Proceedings of the 7th International Symposium on Algorithmic Game Theory (SAGT 2014)*, pages 146–157, 2014.

[20] Edith Elkind, Piotr Faliszewski, and Arkadii M. Slinko. Clone structures in voters' preferences. In *Proceedings of the 13th ACM Conference on Electronic Commerce, EC 2012, Valencia, Spain, June 4-8, 2012*, pages 496–513, 2012.

[21] James M. Enelow and Melvin J. Hinich. *The spatial theory of voting: An introduction*. Cambridge University Press, 1984.

[22] Bruno Escoffier, Jérôme Lang, and Meltem Öztürk. Single-peaked consistency and its complexity. In *ECAI 2008 - 18th European Conference on Artificial Intelligence, Patras, Greece, July 21-25, 2008, Proceedings*, pages 366–370, 2008.

[23] Michal Feldman, Amos Fiat, and Iddan Golomb. On voting and facility location. In *Proceedings of the 2016 ACM Conference on Economics and Computation, EC '16, Maastricht, The Netherlands, July 24-28, 2016*, pages 269–286, 2016.

[24] Vasilis Gkatzelis, Daniel Halpern, and Nisarg Shah. Resolving the optimal metric distortion conjecture. In *61st IEEE Annual Symposium on Foundations of Computer Science, FOCS 2020, Durham, NC, USA, November 16-19, 2020*, pages 1427–1438, 2020.

[25] Ashish Goel, Reyna Hulett, and Anilesh K. Krishnaswamy. Relating metric distortion and fairness of social choice rules. In *Proceedings of the 13th Workshop on Economics of Networks, Systems and Computation, NetEcon@SIGMETRICS 2018, Irvine, CA, USA, June 18, 2018*, page 4:1, 2018.

[26] Stephen Gross, Elliot Anshelevich, and Lirong Xia. Vote until two of you agree: Mechanisms with small distortion and sample complexity. In *Proceedings of the 31st AAAI Conference on Artificial Intelligence*, pages 544–550. AAAI Press, 2017.

[27] Sara Hosseini and Ameneh Moharerhaye Esfahani. *Obnoxious Facility Location*, pages 315–345. Physica-Verlag HD, Heidelberg, 2009.

[28] Samuel Karlin. *Total Positivity*. Stanford University Press, 1968.

[29] David Kempe. An analysis framework for metric voting based on LP duality. In *Proceedings of the 34th AAAI Conference on Artificial Intelligence (AAAI 2020)*, pages 2079–2086. AAAI Press, 2020.

[30] David Kempe. Communication, distortion, and randomness in metric voting. In *Proceedings of the 34th AAAI Conference on Artificial Intelligence (AAAI 2020)*, pages 2087–2094. AAAI Press, 2020.

[31] Vicki Knoblauch. Recognizing one-dimensional euclidean preference profiles. *Journal of Mathematical Economics*, 46(1):1–5, 2010.

[32] Debmalya Mandal, Nisarg Shah, and David P. Woodruff. Optimal communication-distortion tradeoff in voting. In *EC '20: The 21st ACM Conference on Economics and Computation, Virtual Event, Hungary, July 13-17, 2020*, pages 795–813, 2020.

[33] Lili Mei, Deshi Ye, and Guochuan Zhang. Mechanism design for one-facility location game with obnoxious effects on a line. *Theoretical Computer Science*, 734:46–57, 2018.

[34] Lili Mei, Deshi Ye, and Yong Zhang. Approximation strategy-proof mechanisms for obnoxious facility location on a line. *Journal of Combinatorial Optimization*, 36(2):549–571, 2018.

[35] James A. Mirrlees. An exploration in the theory of optimal income taxation. *Review of Economic Studies*, 38:175—208, 1971.

[36] 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 2019*, pages 245–262, 2019.

[37] Ariel D. Procaccia and Jeffrey S. Rosenschein. The distortion of cardinal preferences in voting. In *Cooperative Information Agents X, 10th International Workshop, CIA 2006, Edinburgh, UK, September 11-13, 2006, Proceedings*, pages 317–331, 2006.

[38] Arie Tamir. Obnoxious facility location on graphs. *SIAM Journal on Discrete Mathematics*, 4(4):550–567, 1991.

[39] Vijay V. Vazirani. *Approximation algorithms*. Springer, 2001.

[40] Zvi Drezner and Horst W. Hamacher. *Facility Location: Applications and Theory*. Springer, 2004.

[41] William S. Zwicker. Introduction to the theory of voting. In *Handbook of Computational Social Choice*, pages 23–56. Cambridge University Press, 2016.

# 6 Appendix

## 6.1 About Randomization

In the following result, the $\bar{\alpha}$-distortion of a *randomized* algorithm $\mathcal{A}$ under preference profile $\succ$ is the worst case value that $\frac{\sum_{i\in\mathcal{N}} d(i,\mathsf{opt}(\succ))}{\mathbb{E}_{f\sim\mathcal{A}(\succ)}[\sum_{i\in\mathcal{N}} d(i,f)]}$ takes.

**Proposition 2.** *When $m = 2$, any randomized algorithm has $\bar{\alpha}$-distortion at least $\frac{1+2\bar{\alpha}}{1+\bar{\alpha}}$.*

*Proof.* Suppose there are two candidates $f_1$ and $f_2$, and two agents. Agent 1 has preference order $f_1 \succ_1 f_2$ and agent 2's preference order is $f_2 \succ_2 f_1$. Suppose $f_1$ and $f_2$ are output with probability $p$ and $1-p$, respectively, with $p \le 1-p$ (the case $p > 1-p$ is symmetric).

Consider an instance where the candidates and the agents are on a line. The location of $f_1$, $f_2$, agent 1 and agent 2, are 0, $\bar{\alpha}+1$, $\bar{\alpha}$, and $\bar{\alpha}+1$, respectively.

The instance is consistent with the preference profile and $\bar{\alpha}$-decisive. The distortion is $\frac{1+2\bar{\alpha}}{p(1+2\bar{\alpha})+1-p} = \frac{1+2\bar{\alpha}}{1+2\bar{\alpha}p}$. The largest value that $p$ can take is $1/2$, giving a lower bound of $\frac{1+2\bar{\alpha}}{1+\bar{\alpha}}$. □