Metric Distortion in Peer Selection

Javier Cembrano[1] , **Golnoosh Shahkarami**[2]
[1]Max Planck Institut für Informatik
[2]Max Planck Institut für Informatik, Universität des Saarlandes
{jcembran, gshahkar}@mpi-inf.mpg.de

Abstract

In the *metric distortion* problem, a set of voters and candidates lie in a common metric space, and a committee of k candidates is to be elected. The goal is to select a committee with a small social cost, defined as an increasing function of the distances between voters and selected candidates, but a voting rule only has access to voters' ordinal preferences. The *distortion* of a rule is then defined as the worst-case ratio between the social cost of the selected set and the optimal set, over all possible preferences and consistent distances.

We initiate the study of metric distortion when voters and candidates coincide, which arises naturally in peer selection, and provide tight results for various social cost functions on the line metric. We consider both *utilitarian* and *egalitarian* social cost, given by the sum and maximum of the individual social costs, respectively. For utilitarian social cost, we show that the simple voting rule that selects the k middle agents achieves a distortion that varies between 1 and 2 as k varies between 1 and n when the cost of an individual is the sum of their distances to all selected candidates (additive aggregation). When the cost of an individual is their distance to their qth closest candidate (q-cost), we provide positive results for $q = k = 2$ but mostly show that negative results for general elections carry over to our restricted setting: No constant distortion is possible when $q \leq k/2$ and no distortion better than $3/2$ is possible for $q \geq k/2 + 1$. For egalitarian social cost, a rule that selects extreme agents achieves the best-possible distortion of 2 for additive cost and q-cost with $q > k/3$, whereas no constant distortion is possible for $q \leq k/3$. Our results suggest that having a common set of voters and candidates allows for better constants compared to the general setting, but cases in which no constant is possible in general remain hard under this restriction.

1 Introduction

A fundamental problem in social choice is the aggregation of individual preferences, expressed as rankings over a set of candidates, into a social preference consisting of a subset of elected candidates. For centuries, social choice theorists have proposed several axioms to capture desirable properties that these aggregation or *voting* rules should guarantee, usually leading to strong impossibility results [Arrow, 1963; de Condorcet, 1785; Gibbard, 1973; Satterthwaite, 1975].

As an alternative approach, attempting to quantify the extent to which a certain voting rule is able to faithfully translate the voter preferences into the selected committee, Procaccia and Rosenschein [2006] introduced the notion of *distortion* of a rule. The underlying assumption is that a voter's (dis)affinity with a candidate can be represented by a certain cost, and voters' rankings are the expression of these cardinal preferences. The cost of a committee for a voter is then defined by aggregating the costs of the committee members, and the overall *social cost* of the committee by aggregating the costs for all voters. The distortion then corresponds to the worst-case ratio between the social cost of the selected committee and that of the optimal committee, over all possible preferences and consistent costs.

The study of the distortion of voting rules has usually focused on two ways of modeling the social cost: utilitarian and egalitarian [Caragiannis and Procaccia, 2011; Goel *et al.*, 2018; Caragiannis *et al.*, 2017]. In the utilitarian case, the social cost is defined as the sum of the individual costs of the voters, ensuring that all voters' costs contribute equally to the objective. In contrast, the egalitarian social cost considers the maximum individual cost among all voters, aiming to capture a notion of fairness where no voter is excessively disadvantaged.

In voting theory, it is common to assume that voters' preferences are not fully arbitrary but enjoy some structural properties. A relevant line of work has indeed sought structural restrictions that are natural and have powerful implications, such as single-peaked [Black, 1948] or single-crossing [Mirrlees, 1971]; see Elkind *et al.* [2022] for a survey. A rather general framework among these is that of *spatial* or *metric voting*, where voters and candidates are assumed to lie in a common low-dimensional metric space and voters' costs correspond to their distance to each candidate [Aziz, 2020; Jessee, 2012; Enelow and Hinich, 1984; Merrill and

Grofman, 1999]. For instance, a line metric is commonly employed to capture political affinity on the left-right spectrum, whereas geographical distances are represented in a two-dimensional space.

This structural assumption naturally fits in the metric distortion framework: The distances to candidates fully define the social cost of a committee, but the voting rules only receive their expression as preference rankings. Since preferences are restricted in this model, improved bounds on the distortion of voting rules have been established. Notably, a tight distortion bound of 3 has been established for any single-winner deterministic voting rule [Anshelevich *et al.*, 2018; Kizilkaya and Kempe, 2022; Gkatzelis *et al.*, 2020]. Extending distortion to multi-winner elections requires defining how a voter's cost is aggregated over the selected committee. Two ways have been considered in the literature: the *additive cost*, where a voter's cost is the sum of their distances to all members of the committee [Babashah *et al.*, 2024], and the *q-cost*, where the cost is determined by their distance to their q-th closest committee member [Caragiannis *et al.*, 2022b; Chen *et al.*, 2020].

Work on metric distortion has so far focused on the case where voters and candidates constitute disjoint sets, which constitutes a natural model for large-scale elections. However, in many decision-making scenarios, a group of agents aims to elect a subset of their own members. One can think, for example, of a political organization selecting a committee. Each member ranks others according to their political affinity and the organization aims to select a committee that represents the variety of preferences of its members. Since the voting rule only receives ordinal preferences, a small distortion constitutes a suitable objective to ensure a close-to-optimal outcome under this limited information. In general, this situation arises in the context of *peer selection*, where individuals evaluate each other to choose a group for governance, leadership, or resource allocation. Further examples include academic hiring and promotions, student representative elections, self-organized committees in cooperatives, and local governance selection.

While peer selection rules have been extensively studied in other contexts, particularly in terms of the effect of strategic behavior [e.g. Holzman and Moulin, 2013; Alon *et al.*, 2011; Caragiannis *et al.*, 2022a], little is known regarding their ability to accurately reflect agents' ordinal preferences. On the other hand, previous work on metric distortion for single-selection has often parameterized an election via its *decisiveness*, corresponding to the maximum ratio between a voter's distance to their top choice and to any other candidate [Anshelevich and Postl, 2017; Gkatzelis *et al.*, 2020]. These works have motivated this parameter with the fact that it becomes zero in the peer selection setting as each agent becomes their own top choice.

However, directly considering a common set of voters and candidates constitutes a structural modification to the problem that has not been considered so far.

1.1 Our Contributions and Techniques

We initiate the study of metric distortion when the set of voters and candidates coincide and bound the distortion achievable by voting rules selecting k out of n agents on the line metric for several social costs; see Table 1 for a summary of our results.

We start by observing a simple yet strong property of metric voting on the line with a single set of voters and candidates that follows from previous work [Elkind and Faliszewski, 2014; Babashah *et al.*, 2024]: We can fully compute the order of the agents from their rankings. This constitutes a powerful tool for the design of our mechanisms, as in the following we can always take this order as given.

Utilitarian Additive Cost. We first consider the utilitarian social cost, in which the social cost of a committee is defined as the sum of all individual costs. Intuitively, selecting k consecutive agents results in lower utilitarian social cost. In Section 3.1, we focus on the case of additive aggregation: The cost of a committee for a voter is given by the sum of all distances from the candidates to this voter. As a natural extension of the optimal rules for one or two agents, which select the median and closest-to-median agents, we consider a rule called MEDIAN ALTERNATION that selects k middle agents. We show that MEDIAN ALTERNATION provides a distortion of at most $\frac{2}{k}\left(n - \sqrt{2n\left\lfloor\frac{n-k}{2}\right\rfloor}\right)$, which is close to 1 when k is small compared to n and approaches 2 as k goes to n. Despite its simplicity, the analysis of this rule holds significant challenge. In short, we reduce any metric to another with only two locations and show the distortion for this class of metrics. We show that this reduction is possible without improving the distortion by establishing the existence of a non-improving direction of movement for each agent.

Utilitarian q-Cost. In Section 3.2, we consider utilitarian q-cost, where the cost of a committee for an agent is given by the agent's distance to their qth closest candidate in the committee. In Theorem B.4, we show that no voting rule can provide a constant distortion when $q \leq \frac{k}{2}$, implying that this known impossibility from the setting with disjoint voters and candidates and a general metric space [Caragiannis *et al.*, 2022b] remains in place in our restricted setting. To prove this bound, we partition all but q agents into $\left\lfloor\frac{k}{q}\right\rfloor \geq 2$ sets and consider two metrics that differ in the position of the remaining q agents: relatively close to the other agents in one metric; very far in the other. Intuitively, selecting these q agents leads to an unbounded distortion in the former case but is necessary for a bounded distortion in the latter. For $q > \frac{k}{2}$, the existence of rules with distortion 3 follows from a general result by Caragiannis *et al.* [2022b]. We provide a lower bound, that varies between $\frac{3}{2}$ and 2 as q varies between $\frac{k}{2}+1$ and 2, by considering three different metrics consistent with the same rankings and showing that, in one of them, there are q agents in one extreme that cannot be consistently selected. We finally take a closer look at the case with $k = q = 2$, where a best-possible distortion of 2 can be achieved by selecting the median agents when k is even. For odd k, we show that a rule selecting a *couple* of agents—a pair of agents who prefer each other over all other agents—among the five middle agents achieves an improved distortion of $\frac{4}{3}$, which is again best-possible. The FAVORITE COUPLE rule leverages two key principles: (1) selecting agents close to the median

	additive	q-cost: $q \le \frac{k}{3}$	$\frac{k}{3} < q \le \frac{k}{2}$	$\frac{k}{2} < q \le k$	$q = k = 2$
utilitarian	$1+\sqrt{1+\frac{2}{k}}$; $\frac{7}{3}+\frac{4}{k}\left(\sqrt{2}-\frac{4}{3}\right)$ [BKSS]	—— ∞ [CSV] ——		3; 3 [CSV]	
	1; $\frac{2}{k}\left(n-\sqrt{2n\left\lfloor\frac{n-k}{2}\right\rfloor}\right)$ [T. 3.1]	—— ∞ [T. B.4] ——		$2-\frac{k-q}{4q-k-3}^{(*)}$; 3 [T. B.5, CSV]	2; 2 (n even) [P. B.7] $\frac{4}{3}$; $\frac{4}{3}$ (n odd) [T. 3.2]
egalitarian		∞ [CSV]	—— 3; 3 [CSV] ——		
	$\frac{3}{2}-\frac{1}{k}$; $\frac{3}{2}-\frac{1}{2(k-1)}$ (k even) [T. 4.2] $\frac{3}{2}-\frac{1}{k}$; $\frac{3}{2}-\frac{1}{k(k-1)}$ (k odd) [T. 4.2]	∞ [T. C.2]	—— 2; 2 [T. C.3] ——		

Table 1: Our and previous bounds on the distortion that voting rules can achieve in different settings. Values before and after the semicolon represent lower and upper bounds for the corresponding setting, respectively. Lower bounds take worst-case number of agents n. When the lower bound can be made arbitrarily large, we just write ∞ for simplicity. The number in square brackets refers to the theorem (T.) or proposition (P.) where this bound is shown; the letters in square brackets refer to the paper where a bound is taken from: BKSS is Babashah *et al.* [2024] and CSV is Caragiannis *et al.* [2022b]. In particular, bounds in gray correspond to the previously studied setting with disjoint voters and candidates for comparison, either under a general metric [CSV] or under the line metric [BKSS]. The upper bound for utilitarian q-cost marked with $(*)$ is only valid when $q \ge \frac{k}{2}+1$, which is slightly stronger than the general condition $q > \frac{k}{2}$ on that column.

so as to balance overall distances from agents on each side of the median, and (2) selecting consecutive agents with a small distance between them since this distance is part of the cost of all agents. This intuition of selecting consecutive agents that are as close to each other while also being close to the median in principle holds for larger k, but determining how tightly a group of k agents is clustered based solely on ordinal rankings remains a challenge.

Egalitarian Additive Cost. In Section 4, we turn our attention to the egalitarian social cost, where we focus on the maximum cost of a committee for a voter. We consider the simple k-EXTREMES rule, which selects half of the committee from each extreme. On an intuitive level, this constitutes a natural rule in this setting as it avoids that extreme voters are excessively disadvantaged. For the additive setting, we show in Section 4.1 that k-EXTREMES achieves an optimal distortion up to $O\left(\frac{1}{k}\right)$ terms. In particular, the optimal distortion of 1 is attained for $k = 2$, and distortions of $\frac{3}{2} - \frac{1}{2(k-1)}$ and $\frac{3}{2} - \frac{1}{k(k-1)}$ are achieved for even and odd $k \ge 3$, respectively, almost matching a lower bound of $\frac{3}{2} - \frac{1}{k}$. The worst-case instances involve $k+1$ agents in one extreme, a single agent in the other extreme, and k agents in the middle, which are selected in the optimal committee but cannot be detected by any rule when considering two symmetric distance metrics.

Egalitarian q-Cost. In Appendix C.4, we show that k-EXTREMES attains a distortion of 2 for q-cost as long as $q > \frac{k}{3}$. To do so, we prove that the social cost of the set selected by this rule is at most the distance from the agent closest to the center to their nearest extreme, and bound the social cost of the optimal set from below by half of this distance. We provide a matching lower bound by revisiting the instance used for the additive case. Finally, we show that no constant distortion is possible when $q \le \frac{k}{3}$, again implying that the general impossibility result of Caragiannis *et al.* [2022b] still holds in our setting. In the worst-case instances, we partition the agents into $\left\lfloor \frac{k}{q} \right\rfloor$ sets and consider two symmetric distance metrics where all but one set are placed at a unit distance from one another and two sets in one extreme are at the same location. We show that no rule can pick q agents from each location.

1.2 Further Related Work

Distortion of voting rules was first introduced by Procaccia and Rosenschein [2006]. Since then, extensive research has been conducted to establish lower and upper bounds on the distortion of different rules under various scenarios, both within the metric and non-metric frameworks. For a comprehensive survey, we refer to Anshelevich *et al.* [2021].

Single-Winner Voting. In the non-metric framework, Caragiannis and Procaccia [2011] showed that the distortion of any voting rule is at least $\Omega(m^2)$ and that simple rules such as Plurality achieve a distortion of at most $O(m^2)$, where m is the number of candidates.

In the metric framework, Anshelevich *et al.* [2018] established a general lower bound of 3 on the distortion of any deterministic voting rule. They also analyzed the distortion of common voting rules such as Majority, Borda, and Copeland, showing that the latter achieves the lowest distortion of 5 among them. Goel *et al.* [2017] disproved a conjecture by Anshelevich *et al.* regarding a better-than-5 distortion of the Ranked Pairs rule and introduced the notion of *fairness ratio* of a rule, which captures the egalitarian social cost as a special case. These results were later improved by Munagala and Wang [2019], who extended the analysis to uncovered set rules and reduced the upper bound to 4.236. Gkatzelis *et al.* [2020] closed the gap by improving this bound to 3, and showed the validity of this bound in terms of fairness ratio and thus egalitarian social cost.

Randomized voting rules have also been extensively explored in the metric framework [Pulyassary and Swamy, 2021; Fain *et al.*, 2019]. The best-known upper bound for a randomized voting rule was recently obtained by Charikar *et al.* [2024], who showed that a carefully designed randomization over existing and novel voting rules achieves a distortion of at most 2.753. As of lower bounds, Charikar and Ramakrishnan [2022] disproved a conjecture by Goel *et al.* [2017] regarding the existence of a randomized voting rule with distortion 2, by constructing instances whose distortion approaches 2.113 as the number of candidates grows.

Multi-Winner Voting. In the study of metric distortion for multi-winner voting, various objective functions have been proposed to capture the cost incurred by each voter for the elected committee [Elkind *et al.*, 2017; Faliszewski *et al.*, 2017]. A foundational result by Goel *et al.* [2018] showed that, for the additive cost function, iterating a single-winner voting rule with distortion δ for k rounds produces a k-winner committee with the same distortion. Chen *et al.* [2020] studied the 1-cost objective in the metric framework when each voter casts a vote for a single candidate. They proposed a deterministic rule with a tight distortion of 3 and a randomized rule with a distortion of $3 - \frac{2}{m}$. More generally, Caragiannis *et al.* [2022b] introduced the q-cost objective, where a voter's cost for a committee is determined by the distance to their q-th closest member. They showed that the distortion is unbounded for $q \leq \frac{k}{3}$ and linear in n for $\frac{k}{3} < q \leq \frac{k}{2}$. For $q > \frac{k}{2}$, they presented a non-polynomial voting rule that achieves a distortion of 3 and a polynomial rule with a distortion of 9. They discussed how these upper bounds for $q > \frac{k}{2}$ and the unbounded distortion for $q \leq \frac{k}{3}$ carry over to egalitarian social cost, but interestingly showed that a constant distortion is possible for this objective when $\frac{k}{3} < q \leq \frac{k}{2}$. Kizilkaya and Kempe [2022] later proposed a polynomial-time rule with a distortion of 3. Recently, Babashah *et al.* [2024] studied the distortion of multi-winner elections with additive cost on the line, devising a rule with a distortion of roughly $\frac{7}{3}$. Caragiannis *et al.* [2017] studied distortion in multi-winner voting for the non-metric framework, defining a voter's utility for a committee as the highest utility derived from any of its members. They proposed a rule achieving a distortion of $1 + \frac{m(m-k)}{k}$ for deterministic committee selection when selecting k out of m candidates.

Restricted Voting Settings. A specialized setting in metric voting considers single-peaked and 1-Euclidean preferences, where both voters and alternatives are embedded on the real line [Black, 1948; Moulin, 1980; Miyagawa, 2001; Fotakis *et al.*, 2016; Fotakis and Gourvès, 2022; Voudouris, 2023; Ghodsi *et al.*, 2019]. In particular, the work of Fotakis *et al.* [2024] investigated the distortion of deterministic algorithms for k-committee selection on the line under the 1-cost objective, leveraging additional distance queries.

2 Preliminaries

We let $\mathbb{N}$ denote the strictly positive integers and, for $n \in \mathbb{N}$, we write $[n] = \{1, \ldots, n\}$ for the first n. A *linear order* $\succ$ on a set S is a complete, transitive, and antisymmetric binary relation on S; we denote the set of all linear orders on $[n]$ by $\mathcal{L}(n)$.

An instance of a committee election, or simply an *election* is described by the triple $\mathcal{E} = (A, k, \succ)$, where:

- $A = [n]$ is the set of agents,
- $k \in \mathbb{N}$ is the number of agents to be selected for the committee, and
- $\succ = (\succ_1, \succ_2, \ldots, \succ_n) \in \mathcal{L}^n(n)$ comprises the agents' preference profiles, where $\succ_a \in \mathcal{L}(n)$ is a linear order on $[n]$ for every $a \in [n]$.

We let $\binom{A}{k} = \{S \subseteq A \mid |S| = k\}$ denote the feasible committees for a given election; i.e., the set of all subsets of A of size k.

Line metric. A *distance metric* on A is a function $d \colon A \times A \to \mathbb{R}_+$ satisfying (i) $d(a, b) = 0$ if and only if $a = b$, (ii) $d(a, b) = d(b, a)$ for every $a, b \in A$, and (iii) $d(a, c) \leq d(a, b) + d(b, c)$ for every $a, b, c \in A$. In this paper, we focus on the line metric: We associate each agent $a \in A$ with a position $x_a \in (-\infty, \infty)$, and the metric d is defined by $d(a, b) = |x_a - x_b|$ for every $a, b \in A$. A metric d is said to be *consistent* with a ranking profile $\succ \in \mathcal{L}^n(n)$, denoted as $d \rhd \succ$, if for every triple of agents $a, b, c \in A$, the condition $d(a, b) < d(a, c)$ holds whenever $b \succ_a c$.[1] Since d is fully defined by the position vector $x \in (-\infty, \infty)^A$, we often refer directly to this vector being consistent with a ranking profile $\succ \in \mathcal{L}^n(n)$ and denote it by $x \rhd \succ$. Likewise, we often exchange d by x in the definitions that follow. Finally, for a fixed election $\mathcal{E} = (A, k, \succ)$, consistent vector of locations $x \in (-\infty, \infty)^n$, and interval $I = (y, z)$ with $y < z$, we let $A(I) = \{a \in A \mid x_a \in I\}$ denote the agents with locations in I. When I is a single point $\bar{x}$, we write $A(\bar{x})$ for the agents located at this point.

Social cost. For a certain set of agents A, a committee size $k \in \mathbb{N}$, and a *candidate-aggregation function* $h \colon \mathbb{R}^k_+ \to \mathbb{R}_+$, the *cost* of $S \in \binom{A}{k}$ for agent $a \in A$ is simply $\mathrm{SC}(S, a; d) = h((d(a, b))_{b \in S})$. For a set of agents A, a committee size $k \in \mathbb{N}$, and a *voter-aggregation function* $g \colon \mathbb{R}^n_+ \to \mathbb{R}$, the *social cost* of $S \in \binom{A}{k}$ is $\mathrm{SC}(S, A; d) = g((\mathrm{SC}(S, a; d))_{a \in A})$. In this paper, we study a handful of candidate- and voter-aggregation functions. In terms of the voter-aggregation function $g \colon \mathbb{R}^n \to \mathbb{R}_+$, we focus on the *utilitarian social cost*, given by $g(y) = \sum_{i \in [n]} y_i$, and the *egalitarian social cost*, given by $g(y) = \max\{y_i \mid i \in [n]\}$. In terms of the candidate-aggregation function $h \colon \mathbb{R}^k_+ \to \mathbb{R}_+$, we focus on the *additive social cost*, given by $h(y) = \sum_{i \in [k]} y_i$, and the

[1] Note that this definition allows for agent-dependent tie-breaking; i.e., when $d(a, b) = d(a, c)$ agent a can rank either $b \succ_a c$ or $c \succ_a b$, independently of other agents. This assumption makes the problem in principle harder, so that our upper bounds on the distortion remain valid if a common tie-breking rule is employed, and it allows us to construct simpler examples for lower bounds. It is not hard to see that the same lower bounds can be obtained without the assumption: Whenever a metric has ties, distances can be perturbed by a small enough constant ε so that there are no longer ties and the distortion does not improve.

q-cost, given by $h(y) = \tilde{y}_q$, where $\tilde{y}$ is the vector with the entries of y sorted in increasing order. Thus, for example, the 1-*cost* is given by $h(y) = \min\{y_i \mid i \in [k]\}$; and the k-*cost* is given by $h(y) = \max\{y_i \mid i \in [k]\}$.

Voting rules and distortion. For $n, k \in \mathbb{N}$ with $n \geq k$, an (n,k)-*voting rule* is a function f that takes a preference profile $\succ \in \mathcal{L}^n(n)$ and returns a subset $S \in \binom{[n]}{k}$, to which we often refer as a *committee*. For an election $\mathcal{E} = ([n], k, \succ)$ and a metric d, the *distortion* $\text{dist}(S, \mathcal{E}; d)$ *of* $S \subseteq A$ *under* d is the ratio between the social cost of the committee and the minimum social cost of any committee; i.e.,

$$\text{dist}(S, \mathcal{E}; d) = \frac{\text{SC}(S, A; d)}{\min_{S' \in \binom{A}{k}} \text{SC}(S', A; d)}.$$

For an election $\mathcal{E} = (A, k, \succ)$, the *distortion* $\text{dist}(S, \mathcal{E})$ *of a committee* $S \subseteq A$ is then defined as the worst-case distortion over all metrics consistent with the ranking profile $\succ$; i.e.,

$$\text{dist}(S, \mathcal{E}) = \sup_{d \rhd \succ} \text{dist}(S, \mathcal{E}; d).$$

Finally, for an (n,k)-voting rule f, the distortion of f is defined as the worst-case distortion of its output across all possible elections; i.e.,

$$\text{dist}(f) = \sup_{\succ \in \mathcal{L}^n(n)} \text{dist}(f(\succ), ([n], k, \succ)).$$

Throughout the paper, we study the distortion that voting rules can achieve under different social costs.

2.1 Computing the Order From an Election

An essential property in line metric settings is the ability to determine the order of agents based on their preferences. This result has been established in prior work. Specifically, Elkind and Faliszewski [2014] and Babashah *et al.* [2024] demonstrate that if the preference lists of voters are pairwise distinct, it is possible to uniquely determine their ordering on the line, along with the ordering of non-Pareto-dominated alternatives. While their setting differentiates between voters and alternatives, this result naturally extends to our context, where agents serve as both voters and candidates.

For simplicity, whenever we fix an election throughout the paper we will assume w.l.o.g. that the agents are already ordered, i.e., that the permutation π stated in the lemma is the identity. Hence, we denote the ordered agents by $1, \ldots, n$ and informally refer to this order as *from left to right*.

3 Utilitarian Social Cost

Using Lemma A.1, we know that the order of agents can be fully determined from the preference profile $\succ$. This allows us to compute the *median agent*, which is optimal when selecting one agent ($k = 1$) under the utilitarian objective.

For larger committee sizes ($k > 1$), it becomes necessary to define a way to aggregate voters' distances to the selected agents. In this section, we study two aggregation rules: one that considers the sum of all distances to selected agents in Section 3.1, and one that considers the distance to the qth closest agent in Section 3.2.

3.1 Utilitarian Additive Cost

In this section, we focus on the *utilitarian additive* objective for committee selection. This objective aims to minimize the *utilitarian additive social cost*, which is defined as the total distance from all agents to the selected committee. Formally, the *utilitarian additive social cost* of a committee $S' \in \binom{A}{k}$ is given by

$$\text{SC}(S', A; d) = \sum_{a \in A} \sum_{b \in S'} d(a, b).$$

The cost of each agent $a \in A$ is the sum of their distances to all members of the selected committee S', and the overall social cost is the sum of these individual costs across all agents in A.

It is not hard to see that the optimal committee can be directly computed from the preferences for committee sizes $k = 1$ and $k = 2$. This was already discussed for $k = 1$, while for $k = 2$ the optimal committee depends on the parity of n. If n is even, it consists of the two median agents. If n is odd, it consists of the median agent and the agent closest to them. In any case, these agents can be identified directly from the input preference profile $\succ$, without knowledge of the underlying metric. This results in a voting rule with a distortion of 1.

For selecting a committee of size $k \geq 2$, we consider the following voting rule.

Voting Rule 1 (Median Alternation)**.** *Compute the order of the agents* $1, \ldots, n$ *and return* $S = \left\{ \left\lfloor \frac{n-k}{2} \right\rfloor + 1, \ldots, \left\lfloor \frac{n}{2} \right\rfloor + 1, \ldots, \left\lfloor \frac{n+k}{2} \right\rfloor \right\}$.

Not that the rule selects k agents, leaving $\left\lfloor \frac{n-k}{2} \right\rfloor$ unselected agents on the left extreme and $n - \left\lfloor \frac{n+k}{2} \right\rfloor$ unselected agents on the right extremes. These values are equal if $n - k$ is even; the latter is one unit larger if $n - k$ is odd. On an intuitive level, the rule can be understood as constructed by going through the rank list of the median(s) agent(s), selecting agents in the order reported by them but alternating between those to their left and to their right. This ensures a balanced representation of agents on both sides.

Regarding the selection of larger committees, an important ingredient for our results is that an optimal committee selecting consecutive agents always exists. We state this in Appendix B.1. We now present our main result in terms of utilitarian additive social cost, regarding the distortion guaranteed by MEDIAN ALTERNATION. The complete proof is deffered to Appendix Appendix B.2.

Theorem 3.1. *The distortion of* MEDIAN ALTERNATION *is at most* $\frac{2}{k}\left(n - \sqrt{2\left\lfloor \frac{n-k}{2} \right\rfloor n}\right)$ *for utilitarian additive social cost.*

The distortion stated in the theorem ranges between 1 and 2, except for the case where $k = n - 1$ and k is odd, in which it is equal to $\frac{2n}{n-1}$, making it marginally greater than 2. The bound is equal to $\frac{2}{k}\left(n - \sqrt{(n-k)n}\right)$ if $n - k$ is even and to $\frac{2}{k}\left(n - \sqrt{(n-k-1)n}\right)$ if $n - k$ is odd, so that it is better for even values than for neighboring odd values, with more prominent differences for small k. Besides these parity differences, the bound takes values closer to 1 when k is small and closer to 2 as k approaches n. Figure 3 illustrates the bound for $n = 100$ and k between 2 and $n - 1$.

In order to prove Theorem 3.1, we will show that we can reduce any metric to another one where all agents are in one out of two locations. As a first step, we prove that an agent (or set of agents at the same location) can always be moved in one direction such that the distortion does not improve, as long as they do not pass through other agents' locations. The proof of this lemma relies on the linearity of the objective function: If moving an agent or set of agents to the right has a certain effect on the social cost, moving them to the left has the opposite effect. Then, the ratio between the social cost of any two fixed committees must not improve in one of these directions. Since the committee selected by MEDIAN ALTERNATION remains fixed as long as the order of agents does not change, and changing the optimal set can only lead to a worse distortion, the result follows.

3.2 Utilitarian q-Cost

In this section, we study the distortion of voting rules in the context of utilitarian q-cost, in which the cost of a committee S' for an agent is given by its distance to the qth closest agent in S', and the social cost of a committee is the sum of its cost for all agents. Formally, for a set of agents A, a committee size k, a committee $S' \in \binom{A}{k}$, and a distance metric d, the social cost of the committee is given by

$$\text{SC}(S', A; d) = \sum_{a \in A} \tilde{d}(a)_q,$$

where $\tilde{d}(a) \in \mathbb{R}_+^S$ contains the values $\{d(a, s) \mid s \in S'\}$ in increasing order.

Similarly to the classic setting with disjoint voters and candidates, the distortion of voting rules heavily depends on the value of q. Indeed, a result by Caragiannis *et al.* [2022b] directly implies the existence of (n, k)-voting rules with distortion 3 for q-cost whenever $q > \frac{k}{2}$, since their result holds in a more general setting with disjoint voters and candidates and general distance metrics. We complement this result in Appendix B.2 by providing a lower bound that ranges from $\frac{3}{2}$ and 2 as q varies between $\lceil \frac{k}{2} \rceil + 1$ and k. For $q \leq \frac{k}{2}$, Caragiannis *et al.* show that no rule provides a bounded distortion. We show that this impossibility still holds in our setting. We study the case where $q = k = 2$ in further detail and achieve the best-possible distortions of $\frac{4}{3}$ and 2 for odd and even n, respectively, through natural voting rules that are able to leverage the different objectives involved in the problem.

A Voting Rule for $k = 2$

In this section, we focus on the special case of utilitarian q-cost when $q = k = 2$. In this setting, the social cost of a committee S for an agent a is determined by the distance to the *farthest* agent in the committee S'.

On an intuitive level, the goal is to select agents that are both close to each other and close to the median agent(s). In particular, it is not hard to see that the optimal committee always consists of two consecutive agents: For any committee of non-consecutive agents, replacing the most extreme agent among the selected one with another closer to the median cannot decrease the social cost.

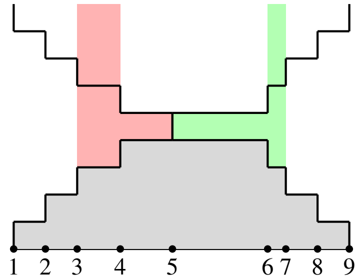


Figure 1: Stair diagram for $n = 9$. The red area corresponds to the committee $\{3, 4\}$; the green area to $\{6, 7\}$.

A visual aid for computing the social cost of a committee is what we call *stair diagrams*, illustrated in Figure 1. The area below both *staircases* is a cost that every committee must incur. A specific committee $\{s_1, s_2\}$ must incur, in addition, a cost equal to the area of the rectangle whose basis is the line segment between both selected candidates and whose height is n (and potentially an additional area to reach this point from the median). Lemma Appendix B.3 bounds the social cost of any committee from below and provides intuition about this objective.

Odd number of agents We first focus on odd values of n. For $n = 3$, it is easy to see that the optimal set corresponds to the median agent and the agent that the median prefers among the others, which yields a simple rule with distortion 1. For $n \geq 5$ we introduce a voting rule called FAVORITE COUPLE. For an election $\mathcal{E} = (A, k, \succ)$, we say that agents $a, b \in A$ are a *couple* if they rank each other above all other agents; i.e., if $b \succ_a c$ and $a \succ_b c$ for every $c \in A \setminus \{a, b\}$. Note that each agent can take part in at most one couple. FAVORITE COUPLE selects the closest couple to the median when restricting to the five middle agents.

Voting Rule 2 (FAVORITE COUPLE)**.** *For a preference profile $\succ$, compute the order from left to right $1, \ldots, n$ and let $m = \frac{n+1}{2}$ be the median agent. If there is a couple among the sets $\{m-1, m\}$ and $\{m, m+1\}$, return it. Else, return $\{m+1, m+2\}$ if $m+2 \succ_m m-2$ and return $\{m-2, m-1\}$ otherwise.*

On an intuitive level, this voting rule selects two consecutive agents who are both close to each other and to the median agent. The restriction to middle agents is necessary; simply choosing an arbitrary couple can lead to a distortion of up to 2. For example, this is the case if there are n agents with distances $d(a, a+1) = 1 + a\varepsilon$ for all $a \in [n]$ and a small $\varepsilon > 0$, as the only couple is $\{1, 2\}$ with a social cost of approximately $\frac{n^2}{2}$, while the committee consisting of the median agent and any neighbor is close to $\frac{n^2}{4}$. We now show that this rule provides the best-possible distortion of $\frac{4}{3}$ for an odd number of agents.

Theorem 3.2. *For every odd $n \geq 5$, FAVORITE COUPLE achieves a distortion of $\frac{4}{3}$ for utilitarian 2-cost. Moreover, there exists $n \in \mathbb{N}$ such that, for every $(n, 2)$-voting rule f,* *we have* $\text{dist}(f) \geq \frac{4}{3}$ *for utilitarian* 2*-cost.*

Even number of agents When n is even, we show that the voting rule that selects the two median agents attains the best-possible distortion of 2. Detailed discussion can be found in Appendix B.5.

4 Egalitarian Social Cost

In this section, we study the worst-case distortion achievable by voting rules in the context of peer selection with egalitarian social cost. Recall that, in this case, given a set of agents A, a committee size k, and a distance metric d, the social cost of a committee $S' \in \binom{A}{k}$ corresponds to the maximum cost of this committee for some agent $a \in A$:

$$\text{SC}(S', A; d) = \max\{\text{SC}(S', a; d) \mid a \in A\}.$$

We will start the section with the simple case $k = 1$, where $S' = \{s\}$ for some $s \in A$ and thus $\text{SC}(S', a; d)$ is simply $d(a, s)$ for every $a \in A$. In Section 4.1 and appendix C.4 we explore the case of general committee size under additive and q-cost candidate-aggregation functions, respectively.

Proposition 4.1. *For every* $n \in \mathbb{N}$*, any* $(n, 1)$*-voting rule has distortion* 2 *for egalitarian social cost. There exists* $n \in \mathbb{N}$ *such that, for every* $(n, 1)$*-voting rule* f*,* $\text{dist}(f) \geq 2$ *for egalitarian social cost.*

4.1 Egalitarian Additive Social Cost

In this section, we study voting rules in the context of egalitarian additive social cost, defined as the maximum over agents of the sum of the distances from the agent to all selected candidates. That is, for a set of agents A, a committee size k, and a distance metric d, the social cost of a committee $S' \in \binom{A}{k}$ is

$$\text{SC}(S', A; d) = \max\left\{\sum_{s \in S'} d(a, s) \;\middle|\; a \in A\right\}.$$

We begin with a simple observation: When $k = 2$ candidates are to be selected, a simple rule selecting both extreme candidates achieves the best-possible distortion of 1. Intuitively, this voting rule makes sense because, for any selected committee, (1) the cost of the committee is maximized for one of the extreme agents, and (2) the sum of the costs of the committee for both extreme agents is fixed (and equal to two times the distance between them). Thus, selecting both extreme agents ensures they incur the same cost and minimizes the maximum cost between them. This rule and its distortion will be covered as a special case of the rule and result we introduce in what follows. For larger k, the above intuition about the cost of any committee being maximized for the extreme agents remains true. We state this property, which will be exploited in the development and analysis of a voting rule guaranteeing a constant distortion, in Appendix C.2.

The rule, which we denote k-EXTREMES, simply returns the $\lfloor \frac{k}{2} \rfloor$ agents closest to one extreme and the $\lceil \frac{k}{2} \rceil$ agents closest to the other extreme.

Voting Rule 3 (k-EXTREMES)**.** *For a preference profile* $\succ$*, compute the order of agents from left to right* $1, \dots, n$ *and return* $S = \{1, \dots, \lfloor \frac{k}{2} \rfloor\} \cup \{n - \lceil \frac{k}{2} \rceil + 1, \dots, n\}$.

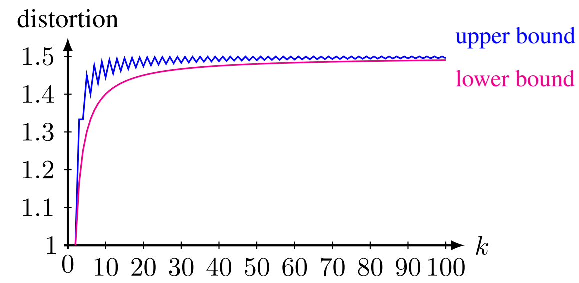


Figure 2: Distortion of k-EXTREMES and lower bound stated in Theorem 4.2 for $k \in \{2, \dots, 99\}$.

The following theorem states the distortion of this voting rule. It captures the previously claimed distortion of 1 for $k = 2$, and it approaches $\frac{3}{2}$ as k grows. This is best possible up to $O\left(\frac{1}{k}\right)$ terms, which vanish as k grows. The upper and lower bounds stated in this theorem are depicted in Figure 2.

Theorem 4.2. *For every* $n, k \in \mathbb{N}$ *with* $n \geq k \geq 2$*,* k*-*EXTREMES *has a distortion for egalitarian additive social cost of at most* $\frac{3}{2} - \frac{1}{2(k-1)}$ *if* k *is even and at most* $\frac{3}{2} - \frac{1}{k(k-1)}$ *if* k *is odd. Conversely, for every* $k \in \mathbb{N}$ *with* $k \geq 3$ *there exists* $n \in \mathbb{N}$ *with* $n \geq k$ *such that, for every* (n, k)*-voting rule* f*,* $\text{dist}(f) \geq \frac{3}{2} - \frac{1}{k}$ *for egalitarian additive social cost.*

5 Discussion

In this work, we have introduced the study of metric distortion in committee elections where voters and candidates coincide and provided a first step towards an understanding of this setting by focusing on the line metric. Our results span a variety of social costs and include both analyses of voting rules and constructions of negative instances to provide impossibility results. Although most of our results are tight, an intriguing gap remains for utilitarian q-cost when q is greater than $\frac{k}{2}$. We believe that rules with a distortion better than the current upper bound of 3 exist and their design may benefit from the insights provided by our rule for $q = k = 2$.

The study of the distortion of voting rules in more general metric spaces constitutes another interesting direction for future work. As the lower bounds presented in this work remain valid and constant upper bounds for q-cost would still be attainable due to the general result by Caragiannis *et al.* [2022b], the design of voting rules providing a small distortion beyond the line in the case of additive cost is the main open question in this regard.

Another challenge in the design of elections is preventing strategic behavior. A mild assumption in the context of peer selection, adopted by the growing literature on impartial selection, is that agents' primary concern is whether they are selected themselves, and a voting rule is deemed impartial if an agent cannot affect this fact by changing their reported preferences. On the other hand, a rule is called strategyproof in the voting literature if no agent can misreport their preferences and lead to a better outcome with respect to their actual preferences. Designing impartial and strategyproof voting rules with bounded distortion for peer selection constitutes an interesting challenge for future work in the area.

References

Noga Alon, Felix Fischer, Ariel Procaccia, and Moshe Tennenholtz. Sum of us: Strategyproof selection from the selectors. In *Proceedings of the 13th Conference on Theoretical Aspects of Rationality and Knowledge*, pages 101–110, 2011.

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

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

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*, pages 4294–4301, 2021.

Kenneth J. Arrow. *Social Choice and Individual Values*. Cowles Foundation, 2nd edition, 1963.

Haris Aziz. Justifications of welfare guarantees under normalized utilities. *ACM SIGecom Exchanges*, 17(2):71–75, 2020.

Negar Babashah, Hasti Karimi, Masoud Seddighin, and Golnoosh Shahkarami. Distortion of multi-winner elections on the line metric: The polar comparison rule. *arXiv preprint arXiv:2411.13720*, 2024.

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

Ioannis Caragiannis and Ariel D Procaccia. Voting almost maximizes social welfare despite limited communication. *Artificial Intelligence*, 175:1655–1671, 2011.

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.

Ioannis Caragiannis, George Christodoulou, and Nicos Protopapas. Impartial selection with additive approximation guarantees. *Theory of Computing Systems*, 66:721–742, 2022.

Ioannis Caragiannis, Nisarg Shah, and Alexandros A Voudouris. The metric distortion of multiwinner voting. *Artificial Intelligence*, 313:103802, 2022.

Moses Charikar and Prasanna Ramakrishnan. Metric distortion bounds for randomized social choice. In *Proceedings of the 2022 Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 2986–3004, 2022.

Moses Charikar, Prasanna Ramakrishnan, Kangning Wang, and Hongxun Wu. Breaking the metric voting distortion barrier. In *Proceedings of the 2024 Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 1621–1640, 2024.

Xujin Chen, Minming Li, and Chenhao Wang. Favorite-candidate voting for eliminating the least popular candidate in a metric space. In *Proceedings of the 34th AAAI Conference on Artificial Intelligence*, pages 1894–1901, 2020.

Nicolas de Condorcet. *Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix*. Cambridge University Press, 1785. Reprint 2016.

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

E Elkind, P Faliszewski, P Skowron, and A Slinko. Properties of multiwinner voting rules. *Social Choice and Welfare*, 48:599–632, 2017.

Edith Elkind, Martin Lackner, and Dominik Peters. Preference restrictions in computational social choice: A survey. *arXiv preprint arXiv:2205.09092*, 2022.

James M Enelow and Melvin J Hinich. *The spatial theory of voting: An introduction*. CUP Archive, 1984.

Brandon Fain, Ashish Goel, Kamesh Munagala, and Nina Prabhu. Random dictators with a random referee: Constant sample complexity mechanisms for social choice. In *Proceedings of the 33rd AAAI Conference on Artificial Intelligence*, pages 1893–1900, 2019.

Piotr Faliszewski, Piotr Skowron, Arkadii Slinko, and Nimrod Talmon. Multiwinner voting: A new challenge for social choice theory. *Trends in computational social choice*, 74(2017):27–47, 2017.

Dimitris Fotakis and Laurent Gourvès. On the distortion of single winner elections with aligned candidates. *Autonomous Agents and Multi-Agent Systems*, 36(2), 2022.

Dimitris Fotakis, Laurent Gourvès, and Jérôme Monnot. Conference program design with single-peaked and single-crossing preferences. In *Proceedings of the 12th International Conference on Web and Internet Economics*, pages 221–235, 2016.

Dimitris Fotakis, Laurent Gourvès, and Panagiotis Patsilinakos. On the distortion of committee election with 1-euclidean preferences and few distance queries. *arXiv preprint arXiv:2408.11755*, 2024.

Mohammad Ghodsi, Mohamad Latifian, and Masoud Seddighin. On the distortion value of the elections with abstention. In *Proceedings of the 33rd AAAI Conference on Artificial Intelligence*, pages 1981–1988, 2019.

Allan Gibbard. Manipulation of voting schemes: A general result. *Econometrica*, 41(4):587–601, 1973.

Vasilis Gkatzelis, Daniel Halpern, and Nisarg Shah. Resolving the optimal metric distortion conjecture. In *Proceedings of the 2020 IEEE 61st Annual Symposium on Foundations of Computer Science*, pages 1427–1438, 2020.

Ashish Goel, Anilesh K Krishnaswamy, and Kamesh Munagala. Metric distortion of social choice rules: Lower bounds and fairness properties. In *Proceedings of the 2017 ACM Conference on Economics and Computation*, pages 287–304, 2017.

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*, 2018.

Ron Holzman and Hervé Moulin. Impartial nominations for a prize. *Econometrica*, 81(1):173–196, 2013.

Stephen A Jessee. *Ideology and spatial voting in American elections*. Cambridge University Press, 2012.

Fatih Erdem Kizilkaya and David Kempe. Plurality veto: A simple voting rule achieving optimal metric distortion. In *Proceedings of the 31st International Joint Conference on Artificial Intelligence*, pages 349–355, 2022.

Samuel Merrill and Bernard Grofman. *A unified theory of voting: Directional and proximity spatial models*. Cambridge University Press, 1999.

James A Mirrlees. An exploration in the theory of optimum income taxation. *The review of economic studies*, 38(2):175–208, 1971.

Eiichi Miyagawa. Locating libraries on a street. *Social Choice and Welfare*, 18(3):527–541, 2001.

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

Kamesh Munagala and Kangning Wang. Improved metric distortion for deterministic social choice rules. In *Proceedings of the 20th ACM Conference on Economics and Computation*, pages 245–262, 2019.

Ariel D Procaccia and Jeffrey S Rosenschein. The distortion of cardinal preferences in voting. In *International Workshop on Cooperative Information Agents, CIA*, pages 317–331, 2006.

Haripriya Pulyassary and Chaitanya Swamy. On the randomized metric distortion conjecture. *arXiv preprint arXiv:2111.08698*, 2021.

Mark A. Satterthwaite. Strategy-proofness and Arrow's conditions: Existence and correspondence theorems for voting procedures and social welfare functions. *J. Econ. Theory*, 10:187–217, 1975.

Alexandros A. Voudouris. Tight distortion bounds for distributed metric voting on a line. *Operations Research Letters*, 51(3):266–269, 2023.

A Lemma from Section 2

Lemma A.1 (Elkind and Faliszewski [2014], Babashah *et al.* [2024])**.** *For every election $\mathcal{E} = ([n], k, \succ)$, we can compute a permutation $\pi\colon [n] \to [n]$ of the agents such that, for any consistent position vector $x \in (-\infty, \infty)^n$ with $x \rhd \succ$, we have either $x_{\pi(1)} \leq x_{\pi(2)} \leq \ldots x_{\pi(n-1)} \leq x_{\pi(n)}$ or $x_{\pi(n)} \leq x_{\pi(n-1)} \leq \cdots \leq x_{\pi(2)} \leq x_{\pi(1)}$.*

B Proofs Deferred from Section 3

B.1 Lemma B.1

Lemma B.1. *For any election $\mathcal{E} = (A, k, \succ)$ and consistent metric $d \rhd \succ$, there exists $i \in [n-k+1]$ such that, defining $S^* = \{i, i+1, \ldots, i+k-1\}$, we have $SC(S^*, A; d) = \min\left\{SC(S', A; d) \mid S' \in \binom{A}{k}\right\}$.*

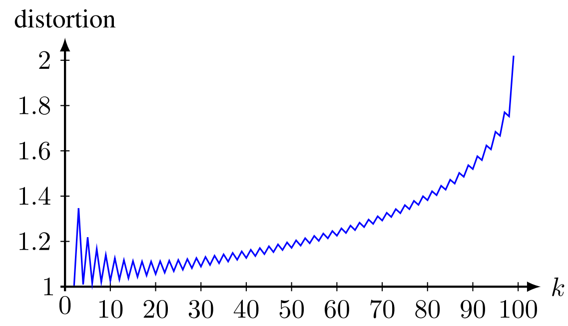


Figure 3: Distortion of MEDIAN ALTERNATION stated in Theorem 3.1 for $n = 100$ and $k \in \{2, \ldots, 99\}$.

Proof. Let $\mathcal{E} = (A, k, \succ)$ with $A = [n]$ and d be as in the statement, and let also $x \rhd \succ$ be a consistent position vector defining d. The result is trivial if $k = 1$, so we assume that $k \geq 2$ in what follows. We first observe that, by Lemma 2 in Babashah *et al.* [2024], $\text{SC}(a, A; d) \leq \text{SC}(b, A; d)$ holds for $a, b \in A$ are such that either (1) $a, b \geq \frac{n+1}{2}$ and $a - \frac{n+1}{2} \leq b - \frac{n+1}{2}$, or (2) $a, b \leq \frac{n+1}{2}$ and $\frac{n+1}{2} - a \leq \frac{n+1}{2} - b$. In simple words, if two agents lie on the same side of the median agent(s), the agent closer to them has a lower cost. Thus, there exist $S^* \in \binom{A}{k}$ that minimizes the social cost such that $\{m_1, m_2\} \subseteq S^*$, where $m_1 = \left\lfloor \frac{n+1}{2} \right\rfloor$ and $m_2 = \left\lceil \frac{n+1}{2} \right\rceil$ denote the median agent(s) (note that $m_1 = m_2$ if n is odd). Now, suppose that S^* is not consecutive. Since $m_1, m_2 \in S^*$, there exists an agent $a \notin S^*$ and $b \in S^*$ such that either (1) $a, b \geq \frac{n+1}{2}$ and $a - \frac{n+1}{2} \leq b - \frac{n+1}{2}$, or (2) $a, b \leq \frac{n+1}{2}$ and $\frac{n+1}{2} - a \leq \frac{n+1}{2} - b$. But then, using the result by Babashah *et al.* again, we obtain that $\text{SC}((S^* \setminus \{b\}) \cup \{a\}, A; d) \leq \text{SC}(S^*, A; d)$; i.e., we can exchange b by a and the social cost of the committee does not increase. By repeating this procedure, we reach a committee with consecutive agents and minimum social cost, as claimed in the statement. ☐

B.2 Proof of Theorem 3.1

Theorem 3.1. *The distortion of* MEDIAN ALTERNATION *is at most $\frac{2}{k}\left(n - \sqrt{2\left\lfloor \frac{n-k}{2} \right\rfloor n}\right)$ for utilitarian additive social cost.*

In order to prove Theorem 3.1, we will show that we can reduce any metric to another one where all agents are in one out of two locations. As a first step, we prove that an agent (or set of agents at the same location) can always be moved in one direction such that the distortion does not improve, as long as they do not pass through other agents' locations. To this end, for a position vector $x \in (-\infty, \infty)^n$, a position $\bar{x} \in (-\infty, \infty)$ such that $A(\bar{x}) \neq \emptyset$, and $\delta > 0$, we define the *shifted position vectors* $x^-(\bar{x}, \delta), x^+(\bar{x}, \delta) \in (-\infty, \infty)^n$ as follows:

$$
\begin{aligned}
x_a^-(\bar{x},\delta) &= x_a - \delta \text{ for every } a \in A(\bar{x}),\\
x_a^-(\bar{x},\delta) &= x_a \text{ for every } a \in A \setminus A(\bar{x}),\\
x_a^+(\bar{x},\delta) &= x_a + \delta \text{ for every } a \in A(\bar{x}),\\
x_a^+(\bar{x},\delta) &= x_a \text{ for every } a \in A \setminus A(\bar{x}).
\end{aligned}
$$

Lemma B.2. *Let $\mathcal{E} = (A, k, \succ)$ be an election with $A = [n]$, let $S \in \binom{A}{k}$ be the committee selected by* MEDIAN ALTERNATION *on this election, and let $x \in (-\infty, \infty)^n$ with $x \rhd \succ$ be a consistent position vector. Let $\bar{x} \in (-\infty, \infty)$ be such that $A(\bar{x}) \neq \emptyset$, let $\delta > 0$ be such that $A((\bar{x}-\delta, \bar{x}+\delta)) = A(\bar{x})$ and let $x^- = x^-(\bar{x},\delta)$ and $x^+ = x^+(\bar{x},\delta)$. Then, for all preference profiles $\succ^-, \succ^+$ such that $x^- \rhd \succ^-$ and $x^+ \rhd \succ^+$, at least one of the following inequalities holds:*

$$
\begin{aligned}
\operatorname{dist}(S,(A,k,\succ^-);x^-) &\geq \operatorname{dist}(S,\mathcal{E};x), \quad \textit{or}\\
\operatorname{dist}(S,(A,k,\succ^+);x^+) &\geq \operatorname{dist}(S,\mathcal{E};x).
\end{aligned}
$$

The proof of this lemma relies on the linearity of the objective function: If moving an agent or set of agents to the right has a certain effect on the social cost, moving them to the left has the opposite effect. Then, the ratio between the social cost of any two fixed committees must not improve in one of these directions. Since the committee selected by MEDIAN ALTERNATION remains fixed as long as the order of agents does not change, and changing the optimal set can only lead to a worse distortion, the result follows. We now proceed with the formal proof.

Proof of Lemma B.2. Let $\mathcal{E} = (A, k, \succ)$, S, x, $\bar{x}$, δ, x^-, x^+, $\succ^-$, and $\succ^+$ be as in the statement. We denote by d, d^-, and d^+ the distance metrics associated to x, x^-, and x^+, respectively.

We first consider an arbitrary committee $S' \in \binom{A}{k}$ and compute the difference between the social cost of this committee under metric d and under both of the other metrics. From the definition of the additive social cost, for any $a \in A$ such that $x_a < \bar{x}$ we have that

$$
\begin{aligned}
\mathrm{SC}(S',a;x^-) &= \sum_{b \in S' \cap A(\bar{x})} d^-(a,b) + \sum_{b \in S' \setminus A(\bar{x})} d^-(a,b)\\
&= \sum_{b \in S' \cap A(\bar{x})} (d(a,b) - \delta) + \sum_{b \in S' \setminus A(\bar{x})} d(a,b)\\
&= \mathrm{SC}(S',a;x) - \delta\,|S' \cap A(\bar{x})|. \qquad (1)
\end{aligned}
$$

Similarly, for any $a \in A$ such that $x_a > \bar{x}$ we have that

$$
\begin{aligned}
\mathrm{SC}(S',a;x^-) &= \sum_{b \in S' \cap A(\bar{x})} d^-(a,b) + \sum_{b \in S' \setminus A(\bar{x})} d^-(a,b)\\
&= \sum_{b \in S' \cap A(\bar{x})} (d(a,b) + \delta) + \sum_{b \in S' \setminus A(\bar{x})} d(a,b)\\
&= \mathrm{SC}(S',a;x) + \delta\,|S' \cap A(\bar{x})|. \qquad (2)
\end{aligned}
$$

Finally, for every a with $x_a = \bar{x}$, i.e., $a \in A(\bar{x})$, we have that

$$
\begin{aligned}
\mathrm{SC}(S',a;x^-) &= \sum_{b \in S' \cap A((-\infty,\bar{x}))} d^-(a,b) \qquad (3)\\
&\quad + \sum_{b \in S' \cap A((\bar{x},+\infty))} d^-(a,b)\\
&= \sum_{b \in S' \cap A((-\infty,\bar{x}))} (d(a,b) - \delta) \qquad (4)\\
&\quad + \sum_{b \in S' \cap A((\bar{x},+\infty))} (d^-(a,b) + \delta)\\
&= \mathrm{SC}(S',a;x) + \qquad (5)\\
&\quad \delta\left(|S' \cap A((\bar{x},+\infty))| - |S' \cap A((-\infty,\bar{x}))|\right). \qquad (6)
\end{aligned}
$$

Combining eqs. (1), (2) and (6), we obtain from the definition of utilitarian social cost that

$$
\begin{aligned}
\mathrm{SC}(S',A;x^-) &= \sum_{a \in A} \mathrm{SC}(S',a;d^-)\\
&= \mathrm{SC}(S',A;x) -\\
&\quad \delta\,|S' \cap A(\bar{x})|\,\big(|A(-\infty,\bar{x})| - |A(\bar{x},+\infty)|\big) -\\
&\quad \delta\,|A(\bar{x})|\,\big(|S' \cap A((-\infty,\bar{x}))| - |S' \cap A((\bar{x},+\infty))|\big).
\end{aligned}
$$

One can proceed analogously for d^+ to obtain

$$
\begin{aligned}
\mathrm{SC}(S',A;x^+) &= \mathrm{SC}(S',A;x) +\\
&\quad \delta\,|S' \cap A(\bar{x})|\,\big(|A(-\infty,\bar{x})| - |A(\bar{x},+\infty)|\big) +\\
&\quad \delta\,|A(\bar{x})|\,\big(|S' \cap A((-\infty,\bar{x}))| - |S' \cap A((\bar{x},+\infty))|\big).
\end{aligned}
$$

Hence, there exists a value $\Delta(S')$, that only depends on the committee δ, such that

$$
\begin{aligned}
\mathrm{SC}(S',A;x^-) &= \mathrm{SC}(S',A;x) - \Delta(S'),\\
\mathrm{SC}(S',A;x^+) &= \mathrm{SC}(S',A;x) + \Delta(S'). \qquad (7)
\end{aligned}
$$

We let S^* denote an optimal committee for the metric d in what follows, i.e., a committee such that $\mathrm{SC}(S^*,A;x) = \min\left\{\mathrm{SC}(S',A;x) \mid S' \in \binom{A}{k}\right\}$. We observe that

$$
\begin{aligned}
\operatorname{dist}(S,(A,k,\succ^-);x^-) &= \frac{\mathrm{SC}(S,A;x^-)}{\min_{S' \in \binom{A}{k}} \mathrm{SC}(S',A;x^-)}\\
&\geq \frac{\mathrm{SC}(S,A;x^-)}{\mathrm{SC}(S^*,A;x^-)} = \frac{\mathrm{SC}(S,A;x) - \Delta(S)}{\mathrm{SC}(S^*,A;x) - \Delta(S^*)}, \qquad (8)
\end{aligned}
$$

and

$$
\begin{aligned}
\operatorname{dist}(S,(A,k,\succ^-);x^+) &= \frac{\mathrm{SC}(S,A;x^+)}{\min_{S' \in \binom{A}{k}} \mathrm{SC}(S',A;x^+)}\\
&\geq \frac{\mathrm{SC}(S,A;x^+)}{\mathrm{SC}(S^*,A;x^+)} = \frac{\mathrm{SC}(S,A;x) + \Delta(S)}{\mathrm{SC}(S^*,A;x) + \Delta(S^*)}. \qquad (9)
\end{aligned}
$$

If either $\mathrm{SC}(S^*,A;x) = \Delta(S^*)$ or $\mathrm{SC}(S^*,A;x) = -\Delta(S^*)$ holds, the distortion becomes unbounded in one of the new instances and the result follows directly. Otherwise, it follows from the simple property stated in the following claim.

Claim B.1. *For any values $y, z \in \mathbb{R}_+$ and $w \in (-z, z)$, we have either $\frac{y+w}{z+w} \geq \frac{y}{z}$ or $\frac{y-w}{z-w} \geq \frac{y}{z}$.*

Proof. Suppose towards a contradiction that both $\frac{y+w}{z+w} < \frac{y}{z}$ and $\frac{y-w}{z-w} < \frac{y}{z}$ hold. Since $w < z$, the first inequality is equivalent to

$$z(y+w) < y(z+w) \iff zw < yw.$$

Since $w > -z$, the second inequality is equivalent to

$$z(y-w) < y(z-w) \iff yw < zw.$$

As the inequalities contradict each other, we conclude. □

Applying these properties to inequalities (8) and (9), we obtain that either

$$\begin{aligned}\text{dist}(S, (A, k, \succ^-); x^+) &\geq \frac{\text{SC}(S, A; x) + \Delta(S)}{\text{SC}(S^*, A; x) + \Delta(S^*)} \\ &\geq \frac{\text{SC}(S, A; x)}{\text{SC}(S^*, A; x)} = \text{dist}(S, \mathcal{E}; x)\end{aligned}$$

or

$$\begin{aligned}\text{dist}(S, (A, k, \succ^-); x^-) &\geq \frac{\text{SC}(S, A; x) - \Delta(S)}{\text{SC}(S^*, A; x) - \Delta(S^*)} \\ &\geq \frac{\text{SC}(S, A; x)}{\text{SC}(S^*, A; x)} = \text{dist}(S, \mathcal{E}; x)\end{aligned}$$

holds, concluding the proof. □

We can use the previous lemma to conclude that, for every election and consistent metric, MEDIAN ALTERNATION selects a committee such that, under another metric with only two locations, the distortion does not improve. Indeed, we can iterating the argument in Lemma B.2 to move (sets of) agents in non-extreme positions in their non-improving direction. This procedure terminates with all agents in one of the original extreme positions x_1 or x_n and that the distortion has not improved. The following lemma formally states this fact.

Lemma B.3. *Let $\mathcal{E} = (A, k, \succ)$ be an election with $A = [n]$, let $S \in \binom{A}{k}$ be the committee selected by* MEDIAN ALTERNATION *on this election, and let $x \in (-\infty, \infty)^n$ with $x \rhd \succ$ be a consistent position vector. Then, there exists a position vector $x' \in (-\infty, \infty)^n$ such that $x'_a \in \{x_1, x_n\}$ for every $a \in A$ and $\text{dist}(S', (A, k, \succ'); x') \geq \text{dist}(S, \mathcal{E}, x)$, where $\succ'$ is any preference profile such that $x' \rhd \succ'$ and $S' \in \binom{A}{k}$ is the committee selected by* MEDIAN ALTERNATION *on the election $(A, k, \succ')$.*

Proof. Let $\mathcal{E} = (A, k, \succ)$ and x be as in the statement, where, as usual, x_1 and x_n represent the positions of the two extreme agents. To construct x' as claimed in the statement, we iteratively move agents toward the positions of the extreme agents using Lemma B.2. Specifically, we initialize $x' = x$ and, as long as $x'_a \in (x_1, x_n)$ for some $a \in A$, we fix $\bar{x} = x_a$, we define

$$\delta^* = \max\{\delta > 0 \mid A((\bar{x} - \delta, \bar{x} + \delta)) = A(\bar{x})\},$$

and we update $x'_b \leftarrow x'_b \pm \delta^*$ for every $b \in A(\bar{x})$ and the sign that ensures not increasing the distortion $\text{dist}(S, A; x')$ of S. Note that the definition of δ^* ensures both the existence of this sign, due to Lemma B.2, and the fact that the number of different positions $|\{y \in (-\infty, \infty) \mid \exists a \in [n] : x'_a = y\}|$ is reduced in each step. Thus, the procedure terminates with a vector $x' \in (-\infty, \infty)$ such that (1) $x'_a \in \{x_1, x_n\}$ for every $a \in A$, and (2) the distortion of S under the resulting metric has not decreased. Note that, since the order of the agents has not been changed besides ties, we have either $S' = S$ if the committee selected by MEDIAN ALTERNATION has not changed or $S' \neq S$ but $\text{SC}(S', A, x') = \text{SC}(S, A; x')$ if the committee has changed due to a different tie-breaking. □

We now proceed with the proof of Theorem 3.1.

Proof of Theorem 3.1. Let $\mathcal{E} = (A, k, \succ)$ be an arbitrary election, where $A = [n]$ is the set of agents. Let $d \rhd \succ$ be any consistent distance metric induced by positions $x \in (-\infty, \infty)^n$, and let S denote the committee selected by MEDIAN ALTERNATION on this election. From Lemma B.3, we know that there exists a new position vector $x' \in (-\infty, \infty)^n$ and associated election $\mathcal{E}' = (A, k, \succ')$, with $x' \rhd \succ'$, such that where all agents are positioned at the two extreme positions of the original instance and the distortion in $\mathcal{E}'$ is at least as bad as the distortion in $\mathcal{E}$; i.e., $x'_a \in \{x_1, x_n\}$ for every $a \in A$ and $\text{dist}(S', (A, k, \succ'); x') \geq \text{dist}(S, \mathcal{E}, x)$, where S' denotes the committee selected by MEDIAN ALTERNATION on $\mathcal{E}'$. Thus, it suffices to compute the distortion for this election $\mathcal{E}'$ to bound the distortion of the voting rule. As usual, we denote by d' the metric induced by the position vector x'.

We partition the set of agents into two groups, $A = A_1 \dot{\cup} A_n$, where

$$A_1 = \{a \in A \mid x'_a = x_1\} \text{ and } A_n = \{a \in A \mid x'_a = x_n\}$$

denote the sets of agents located at positions x_1 and x_n under the position vector x', respectively. We let $S_1 = S' \cap A_1$ and $S_2 = S' \cap A_2$ denote the agents selected by MEDIAN ALTERNATION on $\mathcal{E}'$ from agents in A_1 and A_2, respectively. Then, the social cost of S' is given by

$$\begin{aligned}\text{SC}(S', A; d') &= \sum_{a \in A_1} \sum_{b \in S'} d'(x_1, x_b) + \sum_{a \in A_n} \sum_{b \in S'} d'(x_n, x_b) \\ &= |A_1| \cdot |S_n| \cdot d'(x_1, x_n) + |A_n| \cdot |S_1| \cdot d'(x_1, x_n).\end{aligned}$$

On the other hand, the optimal committee S^* clearly minimizes the total social cost by selecting as many agents as possible from the larger group between A_1 and A_n, as this cost is only incurred by agents in the smaller set. We suppose that $|A_n| \geq |A_1|$ w.l.o.g. We have two cases: either $|A_n| \geq k$ or $|A_n| < k$. In the former case,

$$\text{SC}(S^*, A; d) = |A_1| \cdot k \cdot d'(x_1, x_n),$$

while in the latter case,

$$\text{SC}(S^*, A; d) = |A_1| \cdot |A_n| \cdot d'(x_1, x_n) + |A_n| \cdot (k - |A_n|) \cdot d'(x_1, x_n).$$

Since $|A_n| \geq |A_1|$ implies

$$|A_1| \cdot |A_n| \cdot d'(x_1, x_n) + |A_n| \cdot (k - |A_n|) \cdot d'(x_1, x_n) \geq |A_1| \cdot k \cdot d'(x_1, x_n),$$

the social cost induced by S^* is smaller when $|A_n| \geq k$ and it suffices to bound the distortion in this case. Therefore,

$$\mathrm{dist}(f) \leq \frac{\mathrm{SC}(S, A; d')}{\mathrm{SC}(S^*, A; d')} \quad (10)$$

$$= \frac{|A_1| \cdot |S_n| \cdot d'(x_1, x_n) + |A_n| \cdot |S_1| \cdot d'(x_1, x_n)}{|A_1| \cdot k \cdot d'(x_1, x_n)}$$

$$= \frac{|A_1| \cdot |S_n| + |A_n| \cdot |S_1|}{|A_1| \cdot k}. \quad (11)$$

If $|S_n| = k$, we obtain $\mathrm{dist}(f) = 1$. In what follows, we thus assume $S_1 \neq \emptyset$. From the definition of the MEDIAN ALTERNATION voting rule, we know that $|A_n| - |S_n| = |A_1| - |S_1|$ if $n-k$ is even, and either $|A_n| - |S_n| = |A_1| - |S_1| + 1$ or $|A_n| - |S_n| = |A_1| - |S_1| - 1$ if $n-k$ is odd. Since the distortion increases in $|S_1|$ for fixed n and k due to the assumption that $|A_n| \geq |A_1|$, the worst case is $|A_n| - |S_n| = |A_1| - |S_1| + 1$ when $n-k$ is odd, so we restrict to it in what follows. For ease of notation, we define a value $\chi \in \{0, 1\}$, such that $\chi = 0$ if $n-k$ is even and $\chi = 1$ if $n-k$ is odd, so that we can express the previous equations simply as

$$|A_n| - |S_n| = |A_1| - |S_1| + \chi.$$

From this equality, alongside $|A_1| + |A_n| = n$ and $|S_1| + |S_n| = k$, we can express all $|A_1|$, $|S_1|$, and $|S_n|$ in terms of $|A_n|$ as follows:

$$|A_1| = n - |A_n|, \qquad |S_1| = \frac{n+k+\chi}{2} - |A_n|,$$

$$|S_n| = |A_n| - \frac{n-k-\chi}{2}.$$

Replacing in inequality (11), we obtain

$$\mathrm{dist}(f) \leq \frac{(n - |A_n|)\big(|A_n| - \frac{n-k-\chi}{2}\big) + |A_n|\big(\frac{n+k+\chi}{2} - |A_n|\big)}{(n - |A_n|)k}$$

$$= \frac{1}{k}\left(2|A_n| - \frac{(n-k-\chi)n}{2(n - |A_n|)}\right)$$

$$= h(|A_n|), \quad (12)$$

where we have defined a function $h\colon \left\{\left\lceil \frac{n}{2} \right\rceil, \dots, n-1\right\} \to \mathbb{R}$, which evaluated at $|A_n|$ gives the last expression. Its first and second derivatives are given by

$$h'(y) = \frac{1}{k}\left(2 - \frac{(n-k-\chi)n}{2(n-y)^2}\right),$$

$$h''(y) = -\frac{(n-k-\chi)n}{k(n-y)^3}.$$

Since $h''(y) \leq 0$ for every y in the domain of h, an upper bound for the value of h is given by its value at y^*, where y^* is such that

$$h'(y^*) = 0 \iff y^* = n - \frac{1}{2}\sqrt{(n-k-\chi)n}.$$

Combining this fact with inequality (12), we conclude that

$$\mathrm{dist}(f) \leq h(y^*) =$$

$$\frac{1}{k}\left(2\left(n - \frac{1}{2}\sqrt{(n-k-\chi)n}\right) - \frac{(n-k-\chi)n}{2 \cdot \frac{1}{2}\sqrt{(n-k-\chi)n}}\right)$$

$$= \frac{2}{k}\left(n - \sqrt{(n-k-\chi)n}\right),$$

which is the same as the expression in the statement. □

Impossibility Results

In this section, we provide two strong impossibilities regarding distortion bounds for q-cost, analyzing the cases with $q \leq \frac{k}{2}$ and with $q \geq \left\lceil \frac{k}{2} \right\rceil + 1$ separately.

We begin with a strong impossibility for the case where we focus, for each agent, on their qth closest selected agent with $q \leq \frac{k}{2}$. We show that no constant distortion is possible in this setting, regardless of the number of agents to select.

Theorem B.4. *For every $k \in \mathbb{N}$ with $k \geq 2$ and $q \in \mathbb{N}$ with $q \leq \frac{k}{2}$, there exists $n \in \mathbb{N}$ with $n \geq k$ such that, for every (n,k)-voting rule f, $\mathrm{dist}(f)$ is unbounded for utilitarian q-cost.*

Proof. We let k and q be as in the statement, fix $n \in \mathbb{N}$ to a large value, in particular with $n \geq 2k + q$ (we will ultimately take the limit $n \to \infty$), and consider an arbitrary (n,k)-voting rule f. We denote $p = \left\lfloor \frac{k}{q} \right\rfloor \geq 2$ and partition the agents into $p+1$ sets $A = \dot{\bigcup}_{i=1}^{p} A_i \cup B$, such that $|A_i| \in \left\{\left\lfloor \frac{n-q}{p} \right\rfloor, \left\lceil \frac{n-q}{p} \right\rceil\right\}$ for every $i \in [p]$ and $|B| = q$. Note that this is possible since

$$p\left\lfloor \frac{n-q}{p} \right\rfloor + q \leq n \leq p\left\lceil \frac{n-q}{p} \right\rceil + q.$$

We consider the profile $\succ \in \mathcal{L}^n(n)$, where

(i) $b \succ_a c$ whenever $a \in A_i, b \in A_j, c \in A_\ell$ for some $i, j, \ell \in [p]$ with $|i-j| < |i-\ell|$;

(ii) $b \succ_a c$ whenever $a \in A_i, b \in A_j, c \in B$ for some $i, j \in [p]$;

(iii) $b \succ_a c$ whenever $a, b \in B, c \in A_i$ for some $i \in [p]$;

(iv) $b \succ_a c$ whenever $a \in B, b \in A_i, c \in A_j$ for some $i, j \in [p]$ with $i > j$;

and the remaining pairwise comparisons are arbitrary. We consider the election $\mathcal{E} = (A, k, \succ)$ with $A = [n]$.

In what follows, we distinguish whether f selects all q agents in B or not and construct appropriate distance metrics to show that, in either case, the distortion can be arbitrarily large. Intuitively, if f selects B we will consider this set to be relatively close to A_p, so that picking q agents from each set $A_1, \dots, A_p$ would give a much lower social cost. On the contrary, if f does not select B, we will place this set extremely far from all others, so that the social cost of the selected set is huge compared to the social cost of a committee containing B.

Formally, we first consider the case with $B \subseteq S$ and define the distance metric d_1 on A given by the following positions $x \in (-\infty, \infty)^n$: $x_a = i - 1$ for every $a \in A_i$ and every

$i \in [p]$, and $x_a = 2(p-1)$ for every $a \in B$. It is not hard to see that $d_1 \rhd \succ$; see Figure 4 for an illustration. Since $B \subseteq S$, we have that $\left|S \cap \bigcup_{i \in [p]} A_i\right| \le k - q$. Hence, from an averaging argument, there exists $j \in [p]$ with

$$|S \cap A_j| \le \frac{k-q}{p} = \frac{q}{k}(k-q) < q.$$

From the definition of q-cost, we thus have

$$\text{SC}(S, a; d_1) \ge \min\{d_1(a,b) \mid b \in A \backslash A_j\} \ge 1 \quad \text{for every } a \in A_j. \tag{13}$$

On the other hand, consider the set $S = \cup_{i \in [p]} S_i$, where $S_i \subseteq A_i$ and $|S_i| \ge q$ for every $i \in [p]$. Note that this set exists because $pq = k$ and

$$|A_i| \ge \left\lfloor \frac{n-q}{p} \right\rfloor \ge \left\lfloor \frac{2k}{k} q \right\rfloor \ge q,$$

where we used our assumption $n \ge 2k + q$. From the definition of q-cost, we have that $\text{SC}(S, a; d_1) = 0$ for every $a \in A_i$ and every $i \in [p]$. For each $a \in B$, we have $\text{SC}(S, a; d_1) = p - 1$. Combining these facts with inequality (13), we obtain

$$\text{dist}(f(\succ), \mathcal{E}) \ge \frac{\text{SC}(S, A; d_1)}{\text{SC}(S, A; d_1)} \ge \frac{|A_j|}{(p-1)|B|} \ge \left\lfloor \frac{n-q}{p} \right\rfloor \frac{1}{(p-1)q} = \left\lfloor \frac{(n-q)q}{k} \right\rfloor \cdot \frac{1}{k-q}.$$

We now consider the case with $B \not\subseteq S$ and define the distance metric d_2 on A given by the following positions $x \in (-\infty, \infty)^n$: $x_a = i - 1$ for every $a \in A_i$ and every $i \in [p]$, and $x_a = p - 1 + Mn$ for every $a \in B$. It is not hard to see that $d_2 \rhd \succ$; see Figure 4 for an illustration. Since $B \not\subseteq S$, we have that $|S \cap B| < q$ and thus, by the definition of q-cost, we have

$$\text{SC}(S, a; d_2) \ge \min\{d_2(a,b) \mid b \in A \backslash B\} \ge Mn \quad \text{for every } a \in B. \tag{14}$$

On the other hand, consider the set $T = B \cup \cup_{i \in [p-1]} T_i$, where $T_i \subseteq A_i$ and $|T_i| \ge q$ for every $i \in [p-1]$. Note that this set exists because $(p-1)q = k - q$ and

$$|A_i| \ge \left\lfloor \frac{n-q}{p} \right\rfloor \ge \left\lfloor \frac{2k}{k} q \right\rfloor \ge q,$$

where we used our assumption $n \ge 2k + q$. From the definition of q-cost, we have that $\text{SC}(T, a; d_2) = 0$ for every $a \in A_i$ and every $i \in [p-1]$ and $\text{SC}(T, a; d_2) = 0$ for every $a \in B$. For each $a \in A_p$, we have $\text{SC}(T, a; d_2) = 1$. Combining these facts with inequality (14), we obtain

$$\text{dist}(f(\succ), \mathcal{E}) \ge \frac{\text{SC}(S, A; d_2)}{\text{SC}(T, A; d_2)} \ge \frac{Mn|B|}{|A_p|} \ge \frac{1}{\lceil \frac{n-q}{p} \rceil} Mnq = \frac{1}{\lceil \frac{(n-q)q}{k} \rceil} Mnq.$$

Since $\text{dist}(f(\succ), \mathcal{E}) \ge \left\lfloor \frac{(n-q)q}{k} \right\rfloor \cdot \frac{1}{k-q}$ if $B \subseteq S$ and $\text{dist}(f(\succ), \mathcal{E}) \ge \frac{1}{\lceil \frac{(n-q)q}{k} \rceil} Mnq$ otherwise, we conclude that

$$\text{dist}(f) \ge \min\left\{ \left\lfloor \frac{(n-q)q}{k} \right\rfloor \cdot \frac{1}{k-q}, \frac{1}{\lceil \frac{(n-q)q}{k} \rceil} Mnq \right\},$$

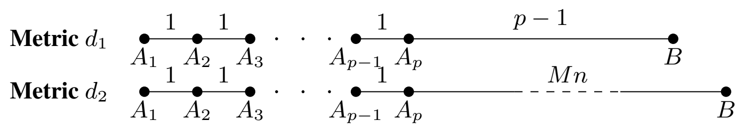


Figure 4: Metrics considered in the proof of Theorem B.4. In this and all similar figures throughout the paper, the (sets of) agents are represented by circles, with the identity of the agents or sets below them, and the distances between them are written on top of the corresponding line segments. All figures consider indistinguishable metrics for a certain preference profile of the agents and thus any voting rule must select the same subsets for any of these metrics.

which can be unbounded by taking n and M arbitrarily large. □

Next, we prove a lower bound of $2 - \frac{k-q}{4q-k-3}$ for the distortion of any voting rule for utilitarian q-cost when $\lceil \frac{k}{2} \rceil < q \le k$ and $k \ge 3$.

Theorem B.5. *For every $k \in \mathbb{N}$ with $k \ge 3$ and $q \in \mathbb{N}$ with $\frac{k}{2} + 1 \le q \le k$, there exists $n \in \mathbb{N}$ with $n \ge k$ such that, for every (n,k)-voting rule f, $\text{dist}(f)$ is at least $2 - \frac{k-q}{4q-k-3}$ for utilitarian q-cost.*

Proof. We let k and q be as in the statement and fix $n = 2(3q - k - 2)$, and consider an arbitrary (n,k)-voting rule f. We partition the agents into four sets $A = \dot{\bigcup}_{i=1}^{4} A_i$ such that $|A_1| = |A_4| = q - 1$ and $|A_2| = |A_3| = 2q - k - 1$. Note that all these values lie between 1 and $q - 1$. Indeed, this is trivial for $|A_1|$ and $|A_4|$, whereas for $|A_2|$ and $|A_3|$ we have $2q - k - 1 \ge 2\left(\frac{k}{2} + 1\right) - k - 1 = 1$ and $2q - k - 1 \le 2q - q - 1 = q - 1$, where we have used that q lies between $\frac{k}{2} + 1$ and k.

We consider the profile $\succ \in \mathcal{L}^n(n)$, where

(i) $b \succ_a c$ whenever $a \in A_i, b \in A_j, c \in A_\ell$ for some $i, j, \ell \in [4]$ with $|i - j| < |i - \ell|$;

(ii) $b \succ_a c$ whenever $a \in A_2, b \in A_1, c \in A_3$;

(iii) $b \succ_a c$ whenever $a \in A_3, b \in A_4, c \in A_2$;

and the remaining pairwise comparisons are arbitrary. We consider the election $\mathcal{E} = (A, k, \succ)$ with $A = [n]$.

In what follows, we distinguish whether f selects q or more agents from $A_1 \cup A_2$, from $A_3 \cup A_4$, or from none of them, and construct appropriate distance metrics to show that, in either case, the distortion is at least the one claimed in the statement. Intuitively, if f selects less than q agents from both $A_1 \cup A_2$ and from $A_3 \cup A_4$, we will consider $A_1 \cup A_2$ on one extreme and $A_3 \cup A_4$ on the other, so that picking q agents from any of these sets would lead to a lower social cost. If f selects q or more agents from $A_1 \cup A_2$ we will consider a metric where A_1 lies in one extreme, A_2 in the middle, and both A_3 and A_4 in the other extreme, so that picking all agents from A_4 would lead to a lower social cost. If f selects q or more agents from $A_3 \cup A_4$, we will construct a symmetric instance.

Formally, we first consider the case with $|S \cap (A_1 \cup A_2)| < q$ and $|S \cap (A_3 \cup A_4)| < q$ and define the distance metric d_1 on A by the following positions $x \in (-\infty, \infty)^n$: $x_a = 0$

for every $a \in A_1 \cup A_2$ and $x_a = 2$ for every $a \in A_3 \cup A_4$. It is not hard to check that $d_1 \rhd \succ$; see Figure 5.(b) for an illustration. It is clear that $\text{SC}(S, a; d_1) = 2$ for every $a \in A$. If we consider the alternative committee $S' = A_1 \cup A_2 \in \binom{A}{k}$, we have $\text{SC}(S', a; d_1) = 0$ for every $a \in A_1 \cup A_2$ and $\text{SC}(S', a; d_1) = 2$ for every $a \in A_3 \cup A_4$. We obtain

$$\text{dist}(f(\succ), \mathcal{E}) \geq \frac{\text{SC}(S, A; d_1)}{\text{SC}(S', A; d_1)} = \frac{2 \cdot n}{2\frac{n}{2}} = 2.$$

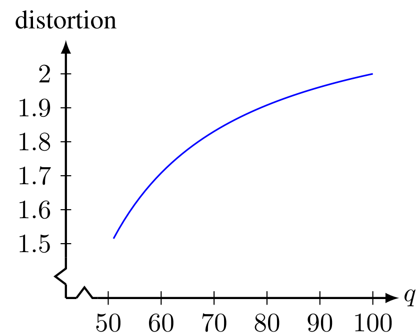
If $|S \cap (A_3 \cup A_4)| \geq q$, we define the distance metric d_2 on A by the following positions $x \in (-\infty, \infty)^n$: $x_a = 0$ for every $a \in A_1 \cup A_2$, $x_a = 1$ for every $a \in A_3$, and $x_a = 2$ for every $a \in A_4$. It is not hard to check that $d_2 \rhd \succ$; see Figure 5.(b) for an illustration. Since $|S \cap (A_1 \cup A_2 \cup A_3)| \leq (k-q) + |A_3| = q - 1 < q$, we have that $\text{SC}(S, a; d_2) = 2$ for every $a \in A_1 \cup A_2$. Furthermore, since both $|A_3| < q$ and $|A_4| < q$, we have that $\text{SC}(S, a; d_2) = 1$ for every $a \in A_3 \cup A_4$. If we consider an alternative committee $S' \subseteq A_1 \cup A_2 \in \binom{A}{k}$, which exists due to $|A_1 \cup A_2| = 3q - k - 2 \geq q$, we have $\text{SC}(S', a; d_2) = 0$ for every $a \in A_1 \cup A_2$, $\text{SC}(S', a; d_2) = 1$ for every $a \in A_3$, and $\text{SC}(S', a; d_2) = 2$ for every $a \in A_4$. Thus, we obtain

$$\begin{aligned}
\text{dist}(f(\succ), \mathcal{E}) &\geq \frac{\text{SC}(S, A; d_2)}{\text{SC}(S', A; d_2)} \\
&= \frac{2|A_1 \cup A_2| + |A_3 \cup A_4|}{|A_3| + 2|A_4|} \\
&= \frac{3(3q - k - 2)}{(2q - k - 1) + 2(q - 1)} \\
&= 2 - \frac{k - q}{4q - k - 3}.
\end{aligned}$$

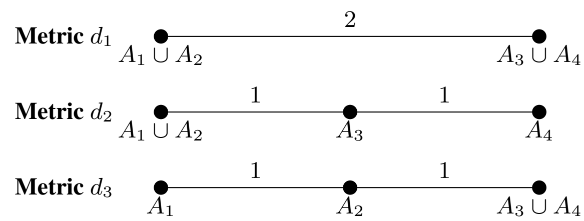
Analogously, if $|S \cap (A_1 \cup A_2)| \geq q$, we define the distance metric d_3 on A by the following positions $x \in (-\infty, \infty)^n$: $x_a = 0$ for every $a \in A_1$, $x_a = 1$ for every $a \in A_2$, and $x_a = 2$ for every $a \in A_3 \cup A_4$. It is not hard to check that $d_3 \rhd \succ$; see Figure 5.(b) for an illustration. Since $|S \cap (A_2 \cup A_3 \cup A_4)| \leq (k - q) + |A_2| = q - 1 < q$, we have that $\text{SC}(S, a; d_3) = 2$ for every $a \in A_3 \cup A_4$. Furthermore, since both $|A_1| < q$ and $|A_2| < q$, we have that $\text{SC}(S, a; d_3) = 1$ for every $a \in A_1 \cup A_2$. If we consider an alternative committee $S' \subseteq A_3 \cup A_4 \in \binom{A}{k}$, which exists due to $|A_3 \cup A_4| = 3q - k - 2 \geq q$, we have $\text{SC}(S', a; d_3) = 0$ for every $a \in A_3 \cup A_4$, $\text{SC}(S', a; d_3) = 1$ for every $a \in A_2$, and $\text{SC}(S', a; d_3) = 2$ for every $a \in A_1$. Thus, we obtain

$$\begin{aligned}
\text{dist}(f(\succ), \mathcal{E}) &\geq \frac{\text{SC}(S, A; d_3)}{\text{SC}(S', A; d_3)} \\
&= \frac{2|A_3 \cup A_4| + |A_1 \cup A_2|}{|A_2| + 2|A_1|} \\
&= \frac{3(3q - k - 2)}{(2q - k - 1) + 2(q - 1)} \\
&= 2 - \frac{k - q}{4q - k - 3}.
\end{aligned}$$

Since $\text{dist}(f(\succ), \mathcal{E}) \geq 2 - \frac{k-q}{4q-k-3}$ regardless of $f(\succ)$, we conclude that $\text{dist}(f) \geq 2 - \frac{k-q}{4q-k-3}$. □



(a) Lower bounds for $k = 100$, $q \in \{51, \dots, 100\}$.



(b) Metrics considered in the proof.

Figure 5: Lower bound on the distortion of any rule for utilitarian q-cost stated in Theorem B.5, and metrics used to prove it.

The lower bound provided in this theorem increases in q and varies between $\frac{3}{2} + \frac{3}{2(k+1)}$ for $q = \frac{k}{2} + 1$ and 2 for $q = k$; Figure 5.(a) illustrates it for $k = 100$ and q between 51 and 100.

B.3 Proof of Lemma B.6

Lemma B.6. *Let $\mathcal{E} = (A, k, \succ)$ be an election and $d \rhd \succ$ a consistent metric. Then, for every committee $S' = \{s_1, s_2\} \in \binom{A}{2}$,*

$$SC(S', A; d) \geq \begin{cases} \sum_{i=1}^{\frac{n-1}{2}} d(i, n-i+1) + \frac{n-1}{2} \cdot d(s_1, s_2) + SC\big(S', \{\frac{n+1}{2}\}; d\big) & \textit{if } n \textit{ is odd,} \\ \sum_{i=1}^{\frac{n}{2}} d(i, n-i+1) + \frac{n}{2} \cdot d(s_1, s_2) & \textit{if } n \textit{ is even.} \end{cases}$$

Proof. Let $\mathcal{E} = (A, k, \succ)$ with $A = [n]$ and d be as in the statement and $S' = \{s_1, s_2\} \in \binom{A}{k}$ an arbitrary committee. We assume that $s_1 < s_2$ w.l.o.g.. Let $i \in \{1, \dots, \lfloor \frac{n}{2} \rfloor\}$ be a fixed agent. If $i \leq s_1 < s_2 \leq n - i + 1$, we have that the cost of the committee for agents i and $n - i + 1$ is at least

$$\begin{aligned}\text{SC}(S', i; d) + \text{SC}(S', n-i+1; d) &= d(i, s_2) + d(s_1, n-i+1) \\ &\geq d(i, n-i+1) + d(s_1, s_2).\end{aligned}$$

Similarly, if $s_2 < i$, we have

$$\begin{aligned}\text{SC}(S', i; d) + \text{SC}(S', n-i+1; d) &= d(s_1, i) + d(s_1, n-i+1) \\ &\geq d(i, n-i+1) + d(s_1, s_2),\end{aligned}$$

and if $s_1 > n - i + 1$,

$$\text{SC}(S', i; d) + \text{SC}(S', n-i+1; d) = d(i, s_2) + d(n-i+1, s_2)$$

$$\geq d(i, n-i+1) + d(s_1, s_2).$$

Summing up over all agents, we obtain

$$\text{SC}(S', A; d) = \sum_{i=1}^{\frac{n}{2}} (\text{SC}(S', i; d) + \text{SC}(S', n-i+1; d)) \geq \sum_{i=1}^{\frac{n}{2}} d(i, n-i+1) + \frac{n}{2} d(s_1, s_2)$$

if n is even, and

$$\text{SC}(S', A; d) = \sum_{i=1}^{\frac{n-1}{2}} (\text{SC}(S', i; d) + \text{SC}(S', n-i+1; d)) + \text{SC}\left(S', \frac{n+1}{2}; d\right) \geq \sum_{i=1}^{\frac{n-1}{2}} d(i, n-i+1) + \frac{n-1}{2} d(s_1, s_2) + \text{SC}\left(S', \frac{n+1}{2}; d\right)$$

if n is odd. □

B.4 Proof of Theorem 3.2

Proof. We consider an arbitrary election $\mathcal{E} = (A, k, \succ)$ with $n \geq 5$ and $A = [n]$, and a consistent metric $d \rhd \succ$. We denote the five middle agents by $a_1, \ldots, a_5$ from left to right, with a_3 being the median agent. We let S denote the committee selected by FAVORITE COUPLE and S^* denote the optimal committee for the metric d. We analyze two main cases, depending on whether the rule selects the median agent or not.

Case 1: $a_3 \in S$ w.l.o.g., we assume that $a_2 \succ_{a_3} a_4$, which implies that the selected committee is $S = \{a_2, a_3\}$. This implies that agents a_2 and a_3 form a couple, and both $d(a_2, a_3) \leq d(a_1, a_2)$ and $d(a_2, a_3) \leq d(a_3, a_4)$ hold. Therefore,

$$d(a_1, a_5) \geq 3 \cdot d(a_2, a_3), \quad d(a_2, a_4) \geq 2 \cdot d(a_2, a_3). \quad (15)$$

For each $i \leq \frac{n-1}{2}$, the joint cost of S for agents i and $n-i+1$ is given by

$$\text{SC}(S, i; d) + \text{SC}(S, n-i+1; d) = d(i, a_3) + d(a_2, n-1+1) = d(i, n-i+1) + d(a_2, a_3).$$

Since the median agent incurs a cost of $\text{SC}(A, a_3; d) = d(a_2, a_3)$, we obtain:

$$\text{SC}(S, A; d) = \sum_{i=1}^{\frac{n-3}{2}} d(i, n-i+1) + d(a_2, a_4) + \left(\frac{n-1}{2}\right) d(a_2, a_3) + d(a_2, a_3) = \sum_{i=1}^{\frac{n-3}{2}} d(i, n-i+1) + \left(\frac{n+1}{2}\right) d(a_2, a_3) + d(a_2, a_4).$$

On the other hand, by Lemma B.6, we have:

$$\text{SC}(S^*, A; d) \geq \sum_{i=1}^{\frac{n-1}{2}} d(i, n-i+1) + \text{SC}(\{a_3\}, A; d) \geq \sum_{i=1}^{\frac{n-3}{2}} d(i, n-i+1) + d(a_2, a_4) + d(a_2, a_3),$$

where we used, for the second inequality, that the cost of the median agent is at least $d(a_2, a_3)$ due to the assumption that $a_2 \succ_{a_3} a_4$. Thus, the distortion is:

$$\begin{aligned} \text{dist}(f) &= \frac{\text{SC}(S, A; d)}{\text{SC}(S^*, A; d)} \\ &\leq \frac{\sum_{i=1}^{\frac{n-3}{2}} d(i, n-i+1) + \left(\frac{n+1}{2}\right) d(a_2, a_3) + d(a_2, a_4)}{\sum_{i=1}^{\frac{n-3}{2}} d(i, n-i+1) + d(a_2, a_3) + d(a_2, a_4)} \\ &\leq \frac{\left(\frac{n-3}{2}\right) \cdot 3 \cdot d(a_2, a_3) + \left(\frac{n+1}{2}\right) \cdot d(a_2, a_3) + 2 \cdot d(a_2, a_3)}{\left(\frac{n-3}{2}\right) \cdot 3 \cdot d(a_2, a_3) + d(a_2, a_3) + 2 \cdot d(a_2, a_3)} \\ &= \frac{\frac{4n-8}{2} + 2}{\frac{3n-9}{2} + 3} = \frac{4n-4}{3n-3} = \frac{4}{3}, \end{aligned}$$

where the second inequality follows from inequalities (15) and the fact that $d(i, n-i+1) \geq d(1, 5)$ for every $i \leq \frac{n-3}{2}$. This concludes the proof for this case.

Case 2: $a_3 \notin S$ In this case, we either have $S = \{a_1, a_2\}$ or $S = \{a_4, a_5\}$; we assume the former w.l.o.g.. From the definition of FAVORITE COUPLE, this implies that $\{a_2, a_3\}$ and $\{a_3, a_4\}$ are not couples, so we must have $a_1 \succ_{a_2} a_3$ and $a_5 \succ_{a_4} a_3$. It also implies that $a_1 \succ_{a_3} a_5$, since $\{a_4, a_5\}$ would be selected otherwise. In terms of distances:

$$d(a_2, a_3) \geq d(a_1, a_2), \quad d(a_3, a_4) \geq d(a_4, a_5), \quad d(a_3, a_5) \geq d(a_1, a_3). \quad (16)$$

Similarly as before, the social cost of the selected committee is

$$\begin{aligned} \text{SC}(S, A; d) &= \sum_{i=1}^{\frac{n-3}{2}} d(i, n-i+1) + \left(\frac{n-3}{2}\right) d(a_1, a_2) + d(a_1, a_2) + d(a_1, a_3) + d(a_1, a_4) \\ &= \sum_{i=1}^{\frac{n-3}{2}} d(i, n-i+1) + \left(\frac{n+3}{2}\right) d(a_1, a_2) + d(a_2, a_3) + d(a_2, a_4). \end{aligned}$$

We now consider two cases depending on whether a_3 is in the optimal committee.

Case 2.1: $a_3 \in S^*$. If the median agent is selected in the optimal committee, we have from Lemma B.6 that

$$\text{SC}(S^*, A; d) \geq \sum_{i=1}^{\frac{n-3}{2}} d(i, n-i+1) + d(a_2, a_4) + \left(\frac{n-1}{2} + 1\right) \min\{d(a_2, a_3), d(a_3, a_4)\}. \quad (17)$$

We now claim that $\left(\frac{n-1}{2}+1\right)\min\{d(a_2,a_3),d(a_3,a_4)\} \geq \frac{3}{2}d(a_1,a_3)$. Indeed, if we have $\min\{d(a_2,a_3),d(a_3,a_4)\} = d(a_2,a_3)$, this holds because $\frac{n-1}{2}+1 \geq 3$ and, due to inequalities (B.4), $3d(a_2,a_3) \geq \frac{3}{2}d(a_1,a_3)$. If $\min\{d(a_2,a_3),d(a_3,a_4)\} = d(a_3,a_4)$, this holds because $\frac{n-1}{2}+1 \geq 3$ and, due to inequalities (B.4), $3d(a_3,a_4) \geq \frac{3}{2}d(a_3,a_5) \geq \frac{3}{2}d(a_1,a_3)$.

Replacing in inequality (17), we obtain

$$\mathrm{SC}(S^*,A;d) \geq \sum_{i=1}^{\frac{n-3}{2}} d(i,n-i+1) + d(a_2,a_4) + \frac{3}{2}\cdot d(a_1,a_2) + \frac{3}{2}\cdot d(a_2,a_3).$$

Thus, the distortion is

$$\begin{aligned}\mathrm{dist}(f) &= \frac{\mathrm{SC}(S,A;d)}{\mathrm{SC}(S^*,A;d)}\\ &\leq \frac{\sum_{i=1}^{\frac{n-3}{2}} d(i,n-i+1) + \left(\frac{n+3}{2}\right)d(a_1,a_2) + d(a_2,a_3) + d(a_2,a_4)}{\sum_{i=1}^{\frac{n-3}{2}} d(i,n-i+1) + d(a_2,a_4) + \frac{3}{2}\cdot d(a_1,a_2) + \frac{3}{2}\cdot d(a_2,a_3)}\\ &\leq \frac{\left(\frac{n-3}{2}\right)\cdot 4\cdot d(a_1,a_2) + \left(\frac{n+3}{2}\right)d(a_1,a_2) + d(a_1,a_2) + 2\cdot d(a_1,a_2)}{\left(\frac{n-3}{2}\right)\cdot 4\cdot d(a_1,a_2) + 2\cdot d(a_1,a_2) + \frac{3}{2}\cdot d(a_1,a_2) + \frac{3}{2}\cdot d(a_1,a_2)}\\ &\leq \frac{\left(\frac{n-3}{2}\right)\cdot 4 + \left(\frac{n+3}{2}\right) + 1 + 2}{\left(\frac{n-3}{2}\right)\cdot 4 + 2 + \frac{3}{2} + \frac{3}{2}}\\ &= \frac{4n-12+n+3+2+4}{4n-12+4+3+3} = \frac{5n-3}{4n-2} \leq \frac{5}{4} \leq \frac{4}{3},\end{aligned}$$

where the second inequality follows by applying inequalities (B.4) and the fact that $d(i,n-i+1) \geq d(1,5)$ for every $i \leq \frac{n-3}{2}$. We conclude the distortion bound of $\frac{4}{3}$ for this case.

Case 2.2: $a_3 \notin S^*$**.** We begin by rewriting the social cost of S more conveniently as

$$\begin{aligned}&\mathrm{SC}(S,A;d)\\ &= \sum_{i=1}^{\frac{n-3}{2}} d(i,n-i+1) + \left(\frac{n-3}{2}\right)d(a_1,a_2) + d(a_1,a_2) + d(a_1,a_3) + d(a_1,a_4)\\ &= \sum_{i=1}^{\frac{n-3}{2}} d(i,n-i+1) + \left(\frac{n+1}{2}\right)d(a_1,a_2) + d(a_1,a_3) + d(a_2,a_3) + d(a_3,a_4)\\ &\leq \sum_{i=1}^{\frac{n-3}{2}} d(i,n-i+1) + \left(\frac{n+1}{2}\right)d(a_1,a_2) + d(a_3,a_5) + d(a_2,a_3) + d(a_3,a_5),\end{aligned}$$

where the last inequality follows from inequalities (B.4). We distinguish two further cases to bound the social cost of the optimal committee from below, depending on whether the optimal committee selects agents from the left or from the right side of the median.

Case 2.2.1: $S^* \subseteq \{a_4,a_5,\dots,n\}$ If the optimal committee selects an agent on the right side of the median agent, its social cost satisfies

$$\begin{aligned}\mathrm{SC}(S^*,A;d) &\geq \sum_{i=1}^{\frac{n-3}{2}} d(i,n-i+1) + d(a_2,a_5) + d(a_3,a_5)\\ &= \sum_{i=1}^{\frac{n-3}{2}} d(i,n-i+1) + d(a_2,a_3) + 2d(a_3,a_5).\end{aligned}$$

Thus, the distortion is

$$\begin{aligned}\mathrm{dist}(f) &= \frac{\mathrm{SC}(S,A;d)}{\mathrm{SC}(S^*,A;d)}\\ &\leq \frac{\sum_{i=1}^{\frac{n-3}{2}} d(i,n-i+1) + \left(\frac{n+1}{2}\right)d(a_1,a_2) + d(a_2,a_3) + 2d(a_3,a_5)}{\sum_{i=1}^{\frac{n-3}{2}} d(i,n-i+1) + d(a_2,a_3) + 2d(a_3,a_5)}\\ &\leq \frac{\left(\frac{n-3}{2}\right)\cdot 4\cdot d(a_1,a_2) + \left(\frac{n+1}{2}\right)d(a_1,a_2) + d(a_1,a_2) + 2\cdot 2\cdot d(a_1,a_2)}{\left(\frac{n-3}{2}\right)\cdot 4\cdot d(a_1,a_2) + d(a_1,a_2) + 2\cdot 2\cdot d(a_1,a_2)}\\ &\leq \frac{\left(\frac{n-3}{2}\right)\cdot 4 + \left(\frac{n+1}{2}\right) + 1 + 4}{\left(\frac{n-3}{2}\right)\cdot 4 + 1 + 4}\\ &= \frac{4n-12+n+1+2+8}{4n-12+2+8} = \frac{5n-1}{4n-2} \leq \frac{4}{3},\end{aligned}$$

where we used inequalities (B.4) for the second inequality.

Case 2.2.2: $S^* \subseteq \{1,\dots,a_1,a_2\}$**.** If $S^* = S$, the distortion is trivially 1 and we conclude. Otherwise, the social cost of S^* satisfies

$$\mathrm{SC}(S^*,A;d) \geq \sum_{i=1}^{\frac{n-3}{2}} d(i,n-i+1) + d(a_1,a_2) + d(a_1,a_3) + d(a_1,a_4).$$

Thus, the distortion is

$$\begin{aligned}\mathrm{dist}(f) &= \frac{\mathrm{SC}(S,A;d)}{\mathrm{SC}(S^*,A;d)}\\ &\leq \frac{\sum_{i=1}^{\frac{n-3}{2}} d(i,n-i+1) + \left(\frac{n-3}{2}\right)d(a_1,a_2) + d(a_1,a_2) + d(a_1,}{\sum_{i=1}^{\frac{n-3}{2}} d(i,n-i+1) + d(a_1,a_2) + d(a_1,a_3) + d(a_1}\\ &\leq \frac{\left(\frac{n-3}{2}\right)\cdot 4d(a_1,a_2) + \left(\frac{n-3}{2}\right)d(a_1,a_2) + d(a_1,a_2) + 2d(a_1,a_2}{\left(\frac{n-3}{2}\right)\cdot 4d(a_1,a_2) + d(a_1,a_2) + 2d(a_1,a_2) + 3d(a_1}\\ &\leq \frac{4n-12+n-3+2+4+6}{4n-12+2+4+6} = \frac{5n-3}{4n} < \frac{4}{3},\end{aligned}$$

where we used inequalities (B.4) for the second inequality. This concludes the proof of the distortion of FAVORITE COUPLE.

For the lower bound, we fix $n=5$ and an arbitrary $(n,2)$-voting rule f, consider the profile $\succ \in \mathcal{L}^5(5)$ defined as

$$\begin{aligned}&1 \succ_1 2 \succ_1 3 \succ_1 4 \succ_1 5,\\ &2 \succ_2 1 \succ_2 3 \succ_2 4 \succ_2 5,\\ &3 \succ_3 2 \succ_3 1 \succ_3 4 \succ_3 5,\\ &4 \succ_4 5 \succ_4 3 \succ_4 2 \succ_4 1,\\ &5 \succ_5 4 \succ_5 3 \succ_5 2 \succ_5 1,\end{aligned}$$

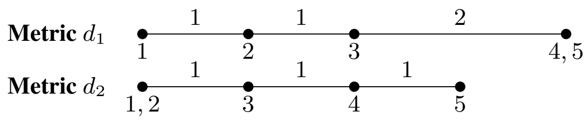


Figure 6: Metrics considered in the proof of Theorem 3.2.

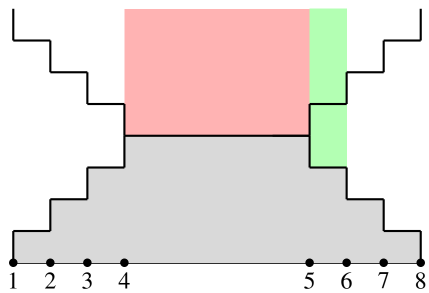


Figure 7: Stair diagram for $n = 8$. The red area corresponds to the committee $\{4, 5\}$; the green area to $\{5, 6\}$.

and consider the election $\mathcal{E} = (A, 2, \succ)$ with $A = [5]$. We distinguish two cases depending on the set of agents $S = f(\succ)$ selected by the rule.

Suppose first that $S = \{1, 2\}$. We take the distance metric d_1 on A given by positions $x_1 = 0$, $x_2 = 1$, $x_3 = 2$, and $x_4 = x_5 = 4$. It is not hard to check that $d_1 \rhd \succ$; see Figure 6 for an illustration. Since $\mathrm{SC}(\{1, 2\}, A; d_1) = 12$, and $\mathrm{SC}(\{4, 5\}, A; d_1) = 9$, we obtain

$$\mathrm{dist}(f(\succ), \mathcal{E}) \geq \frac{\mathrm{SC}(S, A; d_1)}{\min_{S' \in \binom{A}{2}} \mathrm{SC}(S', A; d_1)} \geq \frac{12}{9} = \frac{4}{3}.$$

If $S \in \{\{2, 3\}, \{3, 4\}, \{4, 5\}\}$, we consider the distance metric d_2 on A given by positions $x_1 = x_2 = 0$, $x_3 = 1$, $x_4 = 2$, and $x_5 = 3$. It is not hard to check that $d_2 \rhd \succ$; see Figure 6 for an illustration. Since $\mathrm{SC}(\{2, 3\}, A; d_2) = \mathrm{SC}(\{3, 4\}, A; d_2) = 8$ and $\mathrm{SC}(\{4, 5\}, A; d_2) = 10$, whereas $\mathrm{SC}(\{1, 2\}, A; d_2) = 6$, we obtain

$$\mathrm{dist}(f(\succ), \mathcal{E}) \geq \frac{\mathrm{SC}(S, A; d_2)}{\min_{S' \in \binom{A}{2}} \mathrm{SC}(S', A; d_2)} \geq \frac{8}{6} = \frac{4}{3}.$$

Since $\mathrm{dist}(f(\succ), \mathcal{E}) \geq \frac{4}{3}$ in all these cases and sets of non-consecutive agents can only induce a larger social cost, we conclude that $\mathrm{dist}(f) \geq \frac{4}{3}$. □

B.5 Even number of agents

Even number of agents When n is even, we show that the voting rule that selects the two median agents attains the best-possible distortion of 2.

Proposition B.7. *For an even number of agents n, the voting rule that selects the two median agents achieves a distortion of 2 for utilitarian 2-cost. Moreover, there exists $n \in \mathbb{N}$ such that, for every $(n, 2)$-voting rule f, we have $\mathrm{dist}(f) \geq 2$ for utilitarian 2-cost.*

Proof. We consider an arbitrary election $\mathcal{E} = (A, k, \succ)$ with even $n \geq 4$ and $A = [n]$, and a consistent metric $d \rhd \succ$. Note that the assumption $n \geq 4$ is w.l.o.g. since, for $n = 2$, a distortion of 1 is trivially achieved. We let $m_1 = \frac{n}{2}$ and $m_2 = \frac{n}{2} + 1$ denote the left and right median, respectively, $S = \{m_1, m_2\}$ denote the committee selected by the rule, and S^* denote the optimal committee for the metric d. The social cost of S is

$$\mathrm{SC}(S, A; d) = \sum_{i=1}^{\frac{n}{2}} d(i, n-i+1) + \frac{n}{2} d(m_1, m_2),$$

whereas Lemma B.6 implies a lower bound on the social cost of the optimal committee of

$$\mathrm{SC}(S^*, A; d) \geq \sum_{i=1}^{\frac{n}{2}} d(i, n-i+1).$$

Thus, the distortion of the voting rule is

$$\mathrm{dist}(f) = \frac{\mathrm{SC}(S, A; d)}{\mathrm{SC}(S^*, A; d)} \leq \frac{\sum_{i=1}^{\frac{n}{2}} d(i, n-i+1) + \frac{n}{2} d(m_1, m_2)}{\sum_{i=1}^{\frac{n}{2}} d(i, n-i+1)}$$

$$\leq \frac{2 \sum_{i=1}^{\frac{n}{2}} d(i, n-i+1)}{\sum_{i=1}^{\frac{n}{2}} d(i, n-i+1)} = 2,$$

where the second inequality follows from the fact that $d(m_1, m_2) \leq d(i, n-i+1)$ for any $i \leq \frac{n}{2}$. Thus, the voting rule achieves a distortion of at most 2.

For the lower bound, we fix $n = 4$ and an arbitrary $(n, 2)$-voting rule f, consider the profile $\succ \in \mathcal{L}^4(4)$ defined as

$$\begin{aligned} &1 \succ_1 2 \succ_1 3 \succ_1 4, \\ &2 \succ_2 1 \succ_2 3 \succ_2 4, \\ &3 \succ_3 4 \succ_3 2 \succ_3 1, \\ &4 \succ_4 3 \succ_4 2 \succ_4 1, \end{aligned}$$

and consider the election $\mathcal{E} = (A, 2, \succ)$ with $A = [4]$. We distinguish three cases depending on the set of agents $S = f(\succ)$ selected by the rule.

Suppose first that $S = \{3, 4\}$. We take the distance metric d_1 on A given by positions $x_1 = x_2 = 0$, $x_3 = 1$, and $x_4 = 2$. It is not hard to check that $d_1 \rhd \succ$; see Figure 8 for an illustration. Since $\mathrm{SC}(\{3, 4\}, A; d_1) = 6$, and $\mathrm{SC}(\{1, 2\}, A; d_1) = 3$, we obtain

$$\mathrm{dist}(f(\succ), \mathcal{E}) \geq \frac{\mathrm{SC}(S, A; d_1)}{\min_{S' \in \binom{A}{2}} \mathrm{SC}(S', A; d_1)} \geq \frac{6}{3} = 2.$$

If $S = \{1, 2\}$, we consider the distance metric d_2 on A given by positions $x_1 = 0$, $x_2 = 1$, and $x_3 = x_4 = 2$. It is not hard to check that $d_2 \rhd \succ$; see Figure 8 for an illustration. Since $\mathrm{SC}(\{1, 2\}, A; d_2) = 6$ and $\mathrm{SC}(\{3, 4\}, A; d_2) = 3$, we obtain

$$\mathrm{dist}(f(\succ), \mathcal{E}) \geq \frac{\mathrm{SC}(S, A; d_2)}{\min_{S' \in \binom{A}{2}} \mathrm{SC}(S', A; d_2)} \geq \frac{6}{3} = 2.$$

Finally, if $S = \{2, 3\}$, we consider the distance metric d_3 on A given by positions $x_1 = x_2 = 0$ and $x_3 = x_4 = 2$. It

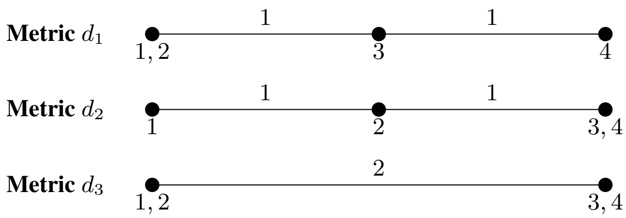


Figure 8: Metrics considered in the proof of Proposition B.7 and Proposition 4.1.

is not hard to check that $d_3 \rhd \succ$; see Figure 8 for an illustration. Since $\text{SC}(\{2,3\}, A; d_3) = 8$ and $\text{SC}(\{1,2\}, A; d_3) = \text{SC}(\{3,4\}, A; d_3) = 4$, we obtain

$$\text{dist}(f(\succ), \mathcal{E}) \geq \frac{\text{SC}(S, A; d_3)}{\min_{S' \in \binom{A}{2}} \text{SC}(S', A; d_3)} \geq \frac{8}{4} = 2.$$

Since $\text{dist}(f(\succ), \mathcal{E}) \geq 2$ in all these cases and sets of non-consecutive agents can only induce a larger social cost, we conclude that $\text{dist}(f) \geq 2$. ☐

C Proofs Deferred from Section 4

C.1 Proof of Proposition 4.1

Proposition 4.1. *For every $n \in \mathbb{N}$, any $(n,1)$-voting rule has distortion 2 for egalitarian social cost. There exists $n \in \mathbb{N}$ such that, for every $(n,1)$-voting rule f, $\text{dist}(f) \geq 2$ for egalitarian social cost.*

Proof. Fix $n \in \mathbb{N}$ and an $(n,1)$-voting rule f arbitrarily. Let $\succ \in \mathcal{L}^n(n)$ be any preference profile on $A = [n]$ and let s be the agent that f outputs for this profile, i.e., $S = f(\succ)$ and $S = \{s\}$. We denote the agents by $\{1, \ldots, n\}$ from left to right, and we let $d \rhd \succ$ be any consistent distance metric. It is clear that, on the one hand, we have

$$\begin{aligned}\text{SC}(\{s\}, A; d) &= \max\{d(a,s) \mid a \in A\} \\ &\leq \max\{d(a,b) \mid a, b \in A\} = d(1,n). \qquad (18)\end{aligned}$$

On the other hand, for every agent $b \in A$ we have that $d(1,b) + d(b,n) = d(1,n)$ and, therefore, $\max\{d(1,b), d(b,n)\} \geq \frac{d(1,n)}{2}$. This implies

$$\begin{aligned}\min_{S' \in \binom{A}{1}} \text{SC}(S', A; d) &= \min_{b \in A} \max\{d(a,b) \mid a \in A\} = \\ &\min_{b \in A} \max\{d(1,b), d(b,n)\} \geq \frac{d(1,n)}{2}. \qquad (19)\end{aligned}$$

Combining inequalities (18) and (19), we directly obtain that $\text{dist}(f) \leq 2$.

For the second claim, we denote $S = f(\succ)$, and we fix $n = 4$ and an arbitrary $(n,1)$-voting rule f, consider the profile $\succ \in \mathcal{L}^4(4)$ defined as

$$\begin{aligned}&1 \succ_1 2 \succ_1 3 \succ_1 4, \\ &2 \succ_2 1 \succ_2 3 \succ_2 4, \\ &3 \succ_3 4 \succ_3 2 \succ_3 1, \\ &4 \succ_4 3 \succ_4 2 \succ_4 1,\end{aligned}$$

and consider the election $\mathcal{E} = (A, 1, \succ)$ with $A = [4]$. We distinguish two cases depending on the agent selected by f.

Suppose first that $S \in \{1,2\}$. We take the distance metric d_1 on A given by positions $x_1 = x_2 = 0$, $x_3 = 1$, and $x_4 = 2$. It is not hard to check that $d_1 \rhd \succ$; see Figure 8 for an illustration. Since $\text{SC}(\{1\}, A; d_1) = 2$, $\text{SC}(\{2\}, A; d_1) = 2$, and $\text{SC}(\{3\}, A; d_1) = 1$, we obtain

$$\begin{aligned}\text{dist}(f(\succ), \mathcal{E}) &\geq \frac{\text{SC}(S, A; d_1)}{\min_{a \in A} \text{SC}(\{a\}, A; d_1)} \\ &\geq \frac{\text{SC}(\{2\}, A; d_1)}{\text{SC}(\{3\}, A; d_1)} = 2.\end{aligned}$$

Similarly, if $S \in \{3,4\}$, we consider the distance metric d_2 on A given by positions $x_1 = 0$, $x_2 = 1$, $x_3 = x_4 = 2$. It is not hard to check that $d_2 \rhd \succ$; see Figure 8 for an illustration. Since $\text{SC}(\{3\}, A; d_2) = 2$, $\text{SC}(\{4\}, A; d_2) = 2$, and $\text{SC}(\{2\}, A; d_2) = 1$, we obtain

$$\text{dist}(f(\succ), \mathcal{E}) \geq \frac{\text{SC}(S, A; d_2)}{\min_{a \in A} \text{SC}(\{a\}, A; d_2)} \geq \frac{\text{SC}(\{3\}, A; d_2)}{\text{SC}(\{2\}, A; d_2)} = 2.$$

Since $\text{dist}(f(\succ), \mathcal{E}) \geq 2$ both when $S \in \{1,2\}$ and when $S \in \{3,4\}$, we conclude that $\text{dist}(f) \geq 2$. ☐

C.2 Lemma C.1

Lemma C.1. *For every set of agents $A = [n]$, committee size k, committee $S' \in \binom{A}{k}$, and distance metric d, it holds that*

$$SC(S', A; d) = \max\{SC(S', 1; d), SC(S', n; d)\}.$$

Proof. Let $A = [n]$, k, S', and d be as in the statement, and recall that we refer to the agents sorted from left to right by $\{1, \ldots, n\}$. We suppose towards a contradiction that there exists $a \in A$ such that $\text{SC}(S', a; d) > \max\{\text{SC}(S', 1; d), \text{SC}(S', n; d)\}$; i.e.,

$$\sum_{s \in S'} d(a,s) > \max\left\{\sum_{s \in S'} d(1,s), \sum_{s \in S'} d(s,n)\right\}. \qquad (20)$$

We now distinguish two cases. If a has at least as many agents in S' weakly to its left as strictly to its right; i.e., $|\{s \in S' \mid s \leq a\}| \geq |\{s \in S' \mid s > a\}|$, then

$$\begin{aligned}\sum_{s \in S'} d(s,n) = &\sum_{s \in S' : s \leq a} (d(a,s) + d(a,n)) \\ &+ \sum_{s \in S' : s > a} (d(a,s) - (d(a,s) - d(s,n))) \\ &\geq \sum_{s \in S' : s \leq a} (d(a,s) + d(a,n)) \\ &+ \sum_{s \in S' : s > a} (d(a,s) - d(a,n)) \\ &= \sum_{s \in S'} d(a,s) \\ &+ (|\{s \in S' : s \leq a\}| - |\{s \in S' : s > a\}|) d(a,n) \\ &\geq \sum_{s \in S'} d(a,s),\end{aligned}$$

a contradiction to inequality (20). Analogously, if $|\{s \in S' \mid s \leq a\}| < |\{s \in S' \mid s > a\}|$, then

$$\begin{aligned}
&\sum_{s\in S'} d(1,s) \\
&= \sum_{s\in S': s>a} (d(1,a) + d(a,s)) \\
&+ \sum_{s\in S': s\leq a} (d(a,s) - (d(a,s) - d(1,s))) \\
&\geq \sum_{s\in S': s>a} (d(1,a) + d(a,s)) \\
&+ \sum_{s\in S': s\leq a} (d(a,s) - d(1,a)) \\
&= \sum_{s\in S'} d(a,s) \\
&+ (|\{s \in S' : s > a\}| - |\{s \in S' : s \leq a\}|) d(1,a) \\
&\geq \sum_{s\in S'} d(a,s),
\end{aligned}$$

a contradiction to inequality (20). □

C.3 Proof of Theorem 4.2

Theorem 4.2. *For every $n, k \in \mathbb{N}$ with $n \geq k \geq 2$, k-*EXTREMES *has a distortion for egalitarian additive social cost of at most $\frac{3}{2} - \frac{1}{2(k-1)}$ if k is even and at most $\frac{3}{2} - \frac{1}{k(k-1)}$ if k is odd. Conversely, for every $k \in \mathbb{N}$ with $k \geq 3$ there exists $n \in \mathbb{N}$ with $n \geq k$ such that, for every (n,k)-voting rule f, $\mathrm{dist}(f) \geq \frac{3}{2} - \frac{1}{k}$ for egalitarian additive social cost.*

Proof. We first show the bound on the distortion of k-EXTREMES. We fix $n, k \in \mathbb{N}$ with $n \geq k \geq 2$, a linear order $\succ$ on $A = [n]$, and a consistent distance metric $d \rhd \succ$. We write $\mathcal{E} = (A, k, \succ)$ for the corresponding election and denote k-EXTREMES by f and the outcome by S in this part of the proof for compactness.

We claim that, if d is such that $\mathrm{SC}(S,1;d) < \mathrm{SC}(S,n;d)$, there exists an alternative distance metric d' with $\mathrm{SC}(S,1;d') \geq \mathrm{SC}(S,n;d')$ and $\mathrm{dist}(f(\succ),\mathcal{E};d') \geq \mathrm{dist}(f(\succ),\mathcal{E};d)$. Indeed, consider such d defined by positions $x \in (-\infty,\infty)^n$, and let d' be defined by positions $x' \in (-\infty,\infty)^n$, where $x'_a = x_{n+1-a}$ for every $a \in [n]$. Since f selects $\lfloor \frac{k}{2} \rfloor$ agents closest to the left-most agent and the $\lceil \frac{k}{2} \rceil$ agents closest to the right-most agent, we have

$$\mathrm{SC}(S,1;d') \geq \mathrm{SC}(S,n;d) > \mathrm{SC}(S,1;d) \geq \mathrm{SC}(S,n;d').$$

Furthermore, this chain of inequalities combined with Lemma C.1 imply that $\mathrm{SC}(S,A;d') \geq \mathrm{SC}(S,A;d)$. Since $\min\{\mathrm{SC}(S',A;d') \mid S' \in \binom{A}{k}\} = \min\{\mathrm{SC}(S',A;d) \mid S' \in \binom{A}{k}\}$, this yields $\mathrm{dist}(f(\succ),\mathcal{E};d') \geq \mathrm{dist}(f(\succ),\mathcal{E};d)$, so the claim follows. Thanks to this claim, we can assume in what follows that $\mathrm{SC}(S,1;d) \geq \mathrm{SC}(S,n;d)$ and thus, by Lemma C.1, $\mathrm{SC}(S,A;d) = \mathrm{SC}(S,1;d)$.

We distinguish three cases depending on the distances from agent 1 to other agents and show the claimed distortion for each of them. We first suppose that $d\big(1, \lfloor \frac{k}{2} \rfloor\big) \leq \frac{d(1,n)}{2}$. In this case,

$$\begin{aligned}
\mathrm{SC}(S,1;d) &= \sum_{s=1}^{\lfloor k/2 \rfloor} d(1,s) + \sum_{s=n-\lceil k/2 \rceil+1}^{n} d(1,s) \\
&\leq \left(\left\lfloor \frac{k}{2} \right\rfloor - 1\right) \frac{d(1,n)}{2} + \left\lceil \frac{k}{2} \right\rceil d(1,n) \\
&= \left(k + \left\lceil \frac{k}{2} \right\rceil - 1\right) \frac{d(1,n)}{2},
\end{aligned}$$

where we used the assumption $d\big(1, \lfloor \frac{k}{2} \rfloor\big) \leq d/2$ and the fact that $d(1,1) = 0$ for the inequality. From Lemma C.1 and **??** we know that $\mathrm{SC}(S',A;d) \geq \frac{kd(1,n)}{2}$ for any $S' \in \binom{A}{k}$, so we obtain

$$\begin{aligned}
\mathrm{dist}(f(\succ),\mathcal{E}) &= \frac{\mathrm{SC}(S,1;d)}{\min_{S' \in \binom{A}{k}} \mathrm{SC}(S',A;d)} \\
&\leq \frac{\big(k + \lceil \frac{k}{2} \rceil - 1\big) \frac{d(1,n)}{2}}{\frac{kd(1,n)}{2}} = \frac{3}{2} - \frac{2 - k \bmod 2}{2k},
\end{aligned}$$

which is smaller than $\frac{3}{2} - \frac{1}{2(k-1)}$ for even $k \geq 2$ and smaller than $\frac{3}{2} - \frac{1}{k(k-1)}$ for odd $k \geq 3$. Thus, we conclude the result in this case.

We next suppose that $d\big(1, \lfloor \frac{k}{2} \rfloor\big) > \frac{d(1,n)}{2}$ and $\sum_{s=2}^{\lfloor k/2 \rfloor} d(1,s) \leq \frac{k-2-k \bmod 2}{k-1} \cdot \frac{kd(1,n)}{4}$. In a similar way as before, we now have

$$\begin{aligned}
\mathrm{SC}(S,1;d) &= \sum_{s=1}^{\lfloor k/2 \rfloor} d(1,s) + \sum_{s=n-\lceil k/2 \rceil+1}^{n} d(1,s) \\
&\leq \frac{k-2-k \bmod 2}{k-1} \cdot \frac{kd(1,n)}{4} + \left\lceil \frac{k}{2} \right\rceil d(1,n) \\
&= \left(3k - \frac{k-(k-2)k \bmod 2}{k-1}\right) \frac{d(1,n)}{4},
\end{aligned}$$

where the inequality follows from the assumption $\sum_{s=2}^{\lfloor k/2 \rfloor} d(1,s) \leq \frac{k-2-k \bmod 2}{k-1} \cdot \frac{kd(1,n)}{4}$ and the fact that $d(1,1) = 0$. From Lemma C.1 and **??** we know that $\mathrm{SC}(S',A;d) \geq \frac{kd(1,n)}{2}$ for any $S' \in \binom{A}{k}$, so we obtain

$$\begin{aligned}
\mathrm{dist}(f(\succ),\mathcal{E}) &= \frac{\mathrm{SC}(S,1;d)}{\min_{S' \in \binom{A}{k}} \mathrm{SC}(S',A;d)} \\
&\leq \frac{\big(3k - \frac{k-(k-2)k \bmod 2}{k-1}\big) \frac{d(1,n)}{4}}{\frac{kd(1,n)}{2}} \\
&= \frac{3}{2} - \frac{k-(k-2)k \bmod 2}{2k(k-1)},
\end{aligned}$$

which corresponds to the expression in the statement.

We finally consider the case with $d\big(1, \lfloor \frac{k}{2} \rfloor\big) > \frac{d(1,n)}{2}$ and $\sum_{s=2}^{\lfloor k/2 \rfloor} d(1,s) > \frac{k-2-k \bmod 2}{k-1} \cdot \frac{kd(1,n)}{4}$. Since the distance between 1 and the right-most point among $\{2, \dots, \lfloor \frac{k}{2} \rfloor\}$, namely $d(1, \lfloor \frac{k}{2} \rfloor)$, is at least its average distance to points within this set, we know that

$$d\left(1, \left\lfloor \frac{k}{2} \right\rfloor\right) \geq \frac{1}{\lfloor \frac{k}{2} \rfloor - 1} \sum_{s=2}^{\lfloor k/2 \rfloor} d(1,s)$$

$$\geq \frac{1}{\lfloor \frac{k}{2} \rfloor - 1} \cdot \frac{k - 2 - k \bmod 2}{k-1} \cdot \frac{kd(1,n)}{4} = \frac{kd(1,n)}{2(k-1)}. \quad (21)$$

Let now $S' \in \binom{A \setminus \{1\}}{k-1}$ be any set of $k-1$ agents without 1. Since $\{2, \ldots, \lfloor \frac{k}{2} \rfloor\}$ are the closest agents to 1, we know that $\frac{1}{k-1} \sum_{s \in S'} d(1,s) \geq \frac{1}{\lfloor k/2 \rfloor - 1} \sum_{s=2}^{\lfloor k/2 \rfloor} d(1,s)$. Rearranging this expression and using our assumption once again, we obtain

$$\sum_{s \in S'} d(1,s) \geq \frac{k-1}{\lfloor \frac{k}{2} \rfloor - 1} \sum_{s=2}^{\lfloor k/2 \rfloor} d(1,s) \geq \frac{kd(1,n)}{2},$$

where we used inequality (21) for the last inequality. For any committee $S' \in \binom{A}{k}$, this implies that $\mathrm{SC}(S', 1; d) \geq \frac{kd(1,n)}{2}$, **??** implies that $\mathrm{SC}(S', 1; d) \geq \mathrm{SC}(S', n; d)$, and Lemma C.1 implies that $\mathrm{SC}(S', A; d) = \mathrm{SC}(S', 1; d)$. Therefore,

$$\min_{S' \in \binom{A}{k}} \mathrm{SC}(S', A; d) = \min_{S' \in \binom{A}{k}} \mathrm{SC}(S', 1; d) = \sum_{s=2}^{k} d(1,s); \quad (22)$$

i.e., the optimal set in this case corresponds to $\{1, \ldots, k\}$. Combining the previous expressions, we obtain the following chain of inequalities:

$$\begin{aligned} &\mathrm{dist}(f(\succ), \mathcal{E}) \\ &= \frac{\mathrm{SC}(S, 1; d)}{\min_{S' \in \binom{A}{k}} \mathrm{SC}(S', A; d)} \\ &= 1 + \frac{\mathrm{SC}(S, 1; d) - \min_{S' \in \binom{A}{k}} \mathrm{SC}(S', A; d)}{\min_{S' \in \binom{A}{k}} \mathrm{SC}(S', A; d)} \\ &\leq 1 + \frac{2}{kd(1,n)} \left(\sum_{s \in S} d(1,s) - \sum_{s=2}^{k} d(1,s) \right) \\ &= 1 + \frac{2}{kd(1,n)} \left(\sum_{s=n-\lceil k/2 \rceil + 1}^{n} d(1,s) - \sum_{s=\lfloor k/2 \rfloor + 1}^{k} d(1,s) \right) \\ &\leq 1 + \frac{2}{kd(1,n)} \cdot \left\lceil \frac{k}{2} \right\rceil \left(d(1,n) - d\left(1, \left\lfloor \frac{k}{2} \right\rfloor \right) \right) \\ &\leq 1 + \frac{2}{kd(1,n)} \cdot \left\lceil \frac{k}{2} \right\rceil \left(d(1,n) - \frac{kd(1,n)}{2(k-1)} \right) \\ &= \frac{3}{2} - \frac{k - (k-2)k \bmod 2}{2k(k-1)}. \end{aligned}$$

Indeed, the first inequality follows from equality (22) and the fact that $\mathrm{SC}(S', A; d) \geq \frac{kd(1,n)}{2}$ for every $S' \in \binom{A}{k}$ due to **??**, the third equality from the definition of f, the second inequality from simple bounds on $d(1,s)$ for different values of s, and the last inequality from inequality (21). The other equalities come from simple calculations. Since the last expression again corresponds to the expression in the statement, we conclude.

For the lower bound, we consider any $k \in \mathbb{N}$ with $k \geq 3$, we fix $n = 2(k+1)$, and consider an arbitrary (n,k)-voting rule f. We partition the agents into four sets $A = \dot{\bigcup}_{i=1}^{4} A_i$ such that $A_1 = \{1\}$, $A_4 = \{n\}$ and $|A_2| = |A_3| = k$. We consider the profile $\succ \in \mathcal{L}^n(n)$, where $S = f(\succ)$, and

(i) $b \succ_a c$ whenever $a \in A_i, b \in A_j, c \in A_\ell$ for some $i, j, \ell \in [4]$ with $|i-j| < |i-\ell|$;

(ii) $1 \succ_a b$ whenever $a \in A_2, b \in A_3 \cup A_4$;

(iii) $n \succ_a b$ whenever $a \in A_3, b \in A_1 \cup A_2$;

and the remaining pairwise comparisons are arbitrary. We consider the election $\mathcal{E} = (A, k, \succ)$ with $A = [n]$.

In what follows, we distinguish whether f selects more agents from $A_1 \cup A_2$ or from $A_3 \cup A_4$ and construct appropriate distance metrics to show that, in either case, the distortion is at least the one claimed in the statement. Intuitively, if f selects more agents from $A_1 \cup A_2$ we will consider a metric where these sets lie on one extreme, $A_4 = n$ on the other extreme, and all agents A_3 in the middle, so that picking all agents from A_3 would lead to a much lower social cost. In the opposite case, we will construct a symmetric instance.

Formally, we first consider the case with $|S \cap (A_1 \cup A_2)| \geq \frac{k}{2}$ and define the distance metric d_1 on A by the following positions $x \in (-\infty, \infty)^n$: $x_a = 0$ for every $a \in A_1 \cup A_2$, $x_a = 1$ for every $a \in A_3$, and $x_n = 2$. It is not hard to check that $d_1 \rhd \succ$; see Figure 9 for an illustration. Since $|S \cap (A_1 \cup A_2)| \geq \frac{k}{2}$, we obtain

$$\begin{aligned} \mathrm{dist}(f(\succ), \mathcal{E}) &\geq \frac{\mathrm{SC}(S, A; d_1)}{\mathrm{SC}(A_3, A; d_1)} \geq \frac{\mathrm{SC}(S, n; d_1)}{\mathrm{SC}(A_3, n; d_1)} \\ &\geq \frac{(k-1) + |S \cap (A_1 \cup A_2)|}{k} \geq \frac{3}{2} - \frac{1}{k}. \end{aligned}$$

Conversely, if $|S \cap (A_3 \cup A_4)| \geq \frac{k}{2}$, we define the distance metric d_2 on A by the following positions $x \in (-\infty, \infty)^n$: $x_1 = 0$, $x_a = 1$ for every $a \in A_2$, and $x_a = 2$ for every $a \in A_3 \cup A_4$. It is not hard to check that $d_2 \rhd \succ$; see Figure 9 for an illustration. Since $|S \cap (A_3 \cup A_4)| \geq \frac{k}{2}$, we obtain

$$\begin{aligned} \mathrm{dist}(f(\succ), \mathcal{E}) &\geq \frac{\mathrm{SC}(S, A; d_2)}{\mathrm{SC}(A_2, A; d_2)} \geq \frac{\mathrm{SC}(S, 1; d_2)}{\mathrm{SC}(A_2, 1; d_2)} \\ &\geq \frac{(k-1) + |S \cap (A_3 \cup A_4)|}{k} \geq \frac{3}{2} - \frac{1}{k}. \end{aligned}$$

Since $\mathrm{dist}(f(\succ), \mathcal{E}) \geq \frac{3}{2} - \frac{1}{k}$ regardless of $f(\succ)$, we conclude that $\mathrm{dist}(f) \geq \frac{3}{2} - \frac{1}{k}$. □

C.4 Egalitarian q-Cost

We now turn our attention to the q-cost aggregation function of candidates, so that the social cost is the maximum over agents of the distance from each agent to its qth closest candidate; i.e.,

$$\mathrm{SC}(S', A; d) = \max \left\{ \tilde{d}(a)_q \mid a \in A \right\}$$

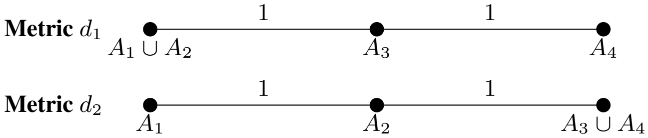


Figure 9: Metrics considered in the proof of Theorem 4.2.

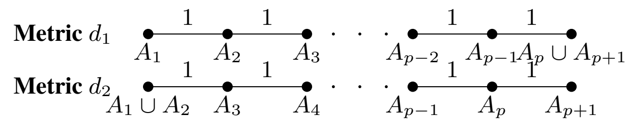


Figure 10: Metrics considered in the proof of Theorem C.2.

for a set of agents A, a committee size k, a committee $S' \in \binom{A}{k}$, and a distance metric d, where $\tilde{d}(a) \in \mathbb{R}_+^{S'}$ contains the values $\{d(a,s) \mid s \in S'\}$ in increasing order.

We begin by showing that no voting rule can guarantee a constant distortion for q-cost when $q \leq \frac{k}{3}$. This implies that the unbounded distortion for this objective, previously established in the setting of disjoint voters and candidates [Caragiannis *et al.*, 2022b], also holds in our setting.

Theorem C.2. *For every $k, q \in \mathbb{N}$ with $\frac{k}{3} \geq q$, there exists $n \in \mathbb{N}$ with $n \geq k$ such that, for every (n,k)-voting rule f, $\mathrm{dist}(f)$ is unbounded for egalitarian q-cost.*

Proof. We let $k, q \in \mathbb{N}$ with $\frac{k}{3} \geq q$ be arbitrary, define $p = \lfloor \frac{k}{q} \rfloor$, and take $n = (p+1)q$. We partition the agents into $p+1 \geq 4$ sets $A = \dot{\bigcup}_{i\in[p+1]} A_i$ such that $|A_i| \geq q$ for every $i \in [p+1]$; note that this is possible since $(p+1)q \leq (\frac{k}{q}+1)q = k+q = n$. We consider any fixed (n,k)-voting rule f and the profile $\succ \in \mathcal{L}^n(n)$, where $S = f(\succ)$, and

(i) $b \succ_a c$ whenever $a \in A_i, b \in A_j, c \in A_\ell$ with $|i-j| < |i-\ell|$ for some $i,j,\ell \in [p+1]$;

(ii) $b \succ_a c$ whenever $a \in A_i, b \in A_1, c \in A_j$ with $|i-1| = |i-j|$ for some $i, j \in [p]$;

(iii) $b \succ_a c$ whenever $a \in A_i, b \in A_{p+1}, c \in A_j$ with $|i-(p+1)| = |i-j|$ for some $i,j \in \{2,\dots,p+1\}$;

and the remaining pairwise comparisons are arbitrary. We consider the election $\mathcal{E} = (A, k, \succ)$ with $A = [n]$. Since $(p+1)q > \frac{k}{q}q = k$, we know that there exists $j \in [p+1]$ such that $|S \cap A_j| < q$. We distinguish two cases depending on the identity of j.

If $j \notin \{p, p+1\}$, we consider the distance metric d_1 on A given by the following positions $x \in (-\infty,\infty)^n$: $x_a = i-1$ for every $a \in A_i$ and $i \in [p]$, and $x_a = p-1$ for every $a \in A_{p+1}$. It is not hard to see that $d_1 \rhd \succ$; see Figure 10 for an illustration. Since $|S \cap A_j| < q$ for some $j \notin \{p,p+1\}$, we have that $\mathrm{SC}(S, A_j; d_1) = 1$. On the other hand, we can define an alternative committee $S' = \bigcup_{i\in[p]} S'_i$ such that $|S'_i \cap A_i| \geq q$ for every $i \in [p]$, which is possible because $|A_i| \geq q$ for every $i \in [p]$ and $pq \leq \frac{k}{q}q = k$. Since $\mathrm{SC}(S', A; d_1) = 0$ and $\mathrm{dist}(f(\succ), \mathcal{E}) \geq \frac{\mathrm{SC}(S,A;d_1)}{\mathrm{SC}(S',A;d_1)}$, we conclude that $\mathrm{dist}(f(\succ), \mathcal{E})$ is unbounded.

If $j \in \{p, p+1\}$, we consider the distance metric d_2 on A given by the following positions $x \in (-\infty,\infty)^n$: $x_a = 0$ for every $a \in A_q$, and $x_a = i-2$ for every $a \in A_i$ and $i \in \{2,\dots,p+1\}$. It is not hard to see that $d_2 \rhd \succ$; see Figure 10 for an illustration. Since $|S \cap A_j| < q$ for some $j \in \{p, p+1\}$, we have that $\mathrm{SC}(S, A_j; d_2) = 1$. On the other hand, we can define an alternative committee $S' = \bigcup_{i\in\{2,\dots,p+1\}} S'_i$ such that $|S'_i \cap A_i| \geq q$ for every $i \in \{2,\dots,p+1\}$, which is possible because $|A_i| \geq q$ for every $i \in \{2,\dots,p+1\}$ and $pq \leq \frac{k}{q}q = k$. Since $\mathrm{SC}(S', A; d_2) = 0$ and $\mathrm{dist}(f(\succ), \mathcal{E}) \geq \frac{\mathrm{SC}(S,A;d_2)}{\mathrm{SC}(S',A;d_2)}$, we conclude that $\mathrm{dist}(f(\succ), \mathcal{E})$ is unbounded.

Since $\mathrm{dist}(f(\succ), \mathcal{E})$ is unbounded regardless of $f(\succ)$, we conclude that $\mathrm{dist}(f)$ is unbounded. □

In the context of egalitarian q-cost for $q > \frac{k}{3}$, much better results are possible. The case with $q > \frac{k}{2}$ behaves similarly to the setting where a single candidate is to be selected: Any voting rule achieves a distortion of 2 and this is best possible. When $\frac{k}{3} < q \leq \frac{k}{2}$, the best-possible distortion a voting rule can achieve is again 2, but not any rule does so. We show that k-EXTREMES attains it.

Theorem C.3. *Let $n, k, q \in \mathbb{N}$ be such that $n \geq k \geq 2$ and $q > \frac{k}{3}$. If $q > \frac{k}{2}$, any (n,k)-voting rule has distortion 2 for egalitarian q-cost. If $q > \frac{k}{3}$, k-EXTREMES has distortion 2 for egalitarian q-cost. For every $k, q \in \mathbb{N}$ with $q > \frac{k}{3} \geq 1$, there exists $n \in \mathbb{N}$ with $n \geq k$ such that, for every (n,k)-voting rule f, $\mathrm{dist}(f) \geq 2$.*

Proof. Let $n, k \in \mathbb{N}$ be such that $n \geq k \geq 2$. Let first $q \in \mathbb{N}$ be such that $q > \frac{k}{2}$. Let f be any (n,k)-voting rule and let $\succ \in \mathcal{L}^n(n)$ be an arbitrary preference profile on $A = [n]$. We denote, as usual, agents by $\{1,\dots,n\}$ from left to right, $S = f(\succ)$, and we let $d \rhd \succ$ be any consistent distance metric. For a committee $S' \in \binom{A}{k}$, we let $\tilde{d}(S', a) \in \mathbb{R}_+^k$ denote the vector with the values $\{d(a,s) \mid s \in S'\}$ in increasing order. It is clear that

$$\begin{aligned}\mathrm{SC}(S, A; d) &= \max\{\tilde{d}(S,a)_q \mid a \in A\}\\ &\leq \max\{d(a,b) \mid a, b \in A\} = d(1,n). \qquad (23)\end{aligned}$$

On the other hand, for every committee $S' \in \binom{A}{k}$, if we denote the agents in S' in increasing order by $s_1, \dots, s_k$ we have that $s_q > s_{k-q}$ because $q > \frac{k}{2}$. This implies that, for every committee $S' \in \binom{A}{k}$, we have

$$\tilde{d}(S',1)_q + \tilde{d}(S',n)_q = s(1, s_q) + d(s_{k-q}, n) > d(1,n),$$

and thus $\max\{\tilde{d}(S',1)_q, \tilde{d}(S',n)_q\} \geq \frac{d(1,n)}{2}$. Therefore,

$$\begin{aligned}\min_{S'\in\binom{A}{k}} \mathrm{SC}(S', A; d) &= \min_{S'\in\binom{A}{k}} \max\{\tilde{d}(S',a)_q \mid a \in A\}\\ &\geq \min_{S'\in\binom{A}{k}} \max\{\tilde{d}(S',1)_q, \tilde{d}(S',n)_q\}\\ &\geq \frac{d(1,n)}{2}. \qquad (24)\end{aligned}$$

Combining inequalities (23) and (24), we directly obtain that $\mathrm{dist}(f) \leq 2$.

Let now $q \in \mathbb{N}$ be such that $\frac{k}{3} < q \leq \frac{k}{2}$, $\succ \in \mathcal{L}^n(n)$ be an arbitrary preference profile on $A = [n]$, and $d \rhd \succ$ be a consistent distance metric; we consider the election $\mathcal{E} = (A, k, \succ)$. We denote the outcome of k-EXTREMES for this profile by S for compactness. We denote agents by $\{1, \dots, n\}$ from left to right and, for $S' \in \binom{A}{k}$, we let $\tilde{d}(S', a) \in \mathbb{R}^k_+$ denote the vector with the values $\{d(a, s) \mid s \in S'\}$ in increasing order. We finally let $a^* \in \arg\max\{\min\{d(1, a), d(a, n)\} \mid a \in A\}$ denote the agent with maximum distance from both extreme agents, assume w.l.o.g. that this is its distance to 1, i.e., $d(1, a^*) \leq d(a^*, n)$, and write $d^* = d(1, a^*)$ for this distance. Observe that

$$\min\{d(a^*, n), d(1, a^* + 1)\} \geq \frac{d(1, n)}{2}. \tag{25}$$

Indeed, $d(a^*, n) \geq \frac{d(1,n)}{2}$ follows directly from the inequality $d(1, a^*) \leq d(a^*, n)$ and the equality $d(1, a^*) + d(a^*, n) = d(1, n)$. Having $d(1, a^* + 1) < \frac{d(1,n)}{2}$ would imply $\min\{d(1, a^* + 1), d(a^* + 1, n)\} > d^*$, a contradiction to the definition of a^*.

We first tackle two simple cases. If $a^* < q$, i.e., there are less than q agents between 1 and a^*, then for any committee $S' \in \binom{A}{k}$ we have $\mathrm{SC}(S', A; d) \geq \mathrm{SC}(S', 1; d) \geq d(1, a^* + 1) \geq \frac{d(1,n)}{2}$, where the second inequality follows from inequality (25). Since $\mathrm{SC}(S', A; d) \leq d(1, n)$ holds for any committee $S' \in \binom{A}{k}$, we know that in particular $\mathrm{SC}(S, A; d) \leq d(1, n)$ and thus $\mathrm{dist}(f) \leq 2$. Similarly, if $n - a^* < q$, i.e., there are less than q agents between $a^* + 1$ and n, then for any committee $S' \in \binom{A}{k}$ we have $\mathrm{SC}(S', A; d) \geq \mathrm{SC}(S', 1; d) \geq d(a^*, n) \geq \frac{d(1,n)}{2}$, where the second inequality follows from inequality (25). As before, $\mathrm{dist}(f) \leq 2$ thus follows directly.

If none of the previous cases hold, we have both $a^* \geq q$ and $n - a^* \geq 2$, so that from the definition of k-EXTREMES we have $|S \cup \{1, \dots, a^*\}| = \lfloor \frac{k}{2} \rfloor \geq q$ and $|S \cup \{a^* + 1, \dots, n\}| = \lceil \frac{k}{2} \rceil \geq q$. This implies that

$$\mathrm{SC}(S, A; d) \leq \max\{d(1, a^*), d(a^* + 1, n)\} \leq d^*. \tag{26}$$

We claim that, for every $S' \in \binom{A}{k}$, we have $\mathrm{SC}(S', A; d) \geq \frac{d^*}{2}$. Together with inequality (26), this would immediately imply $\mathrm{dist}(f) \leq 2$ and conclude the proof. To prove this fact, suppose for the sake of contradiction that $\mathrm{SC}(S', A; d) < \frac{d^*}{2}$ for some $S' \in \binom{A}{k}$. This is equivalent to the fact that

$$\mathrm{SC}(S', a; d) < \frac{d^*}{2} \iff \left| S' \cup \left\{ b \in A : d(a, b) < \frac{d^*}{2} \right\} \right| \geq q$$

for every $a \in A$. Since the sets $\left\{ b \in A \mid d(a, b) < \frac{d^*}{2} \right\}$ for $a \in \{1, a^*, n\}$ are disjoint, we conclude that $|S'| \geq 3q > k$, a contradiction.

For the lower bound, we consider the same instances as in the proof of Theorem 4.2; we repeat the construction for completeness. Naturally, the proof of the lower bound in the end differs from the additive case. We consider any $k \in \mathbb{N}$ with $k \geq 2$, we fix $n = 2(k + 1)$, and consider an arbitrary (n, k)-voting rule f. We partition the agents into four sets $A = \dot{\bigcup}_{i=1}^{4} A_i$ such that $A_1 = \{1\}$, $A_4 = \{n\}$ and $|A_2| = |A_3| = k$. We consider the profile $\succ \in \mathcal{L}^n(n)$, where $S = f(\succ)$, and

(i) $b \succ_a c$ whenever $a \in A_i, b \in A_j, c \in A_\ell$ for some $i, j, \ell \in [4]$ with $|i - j| < |i - \ell|$;

(ii) $1 \succ_a b$ whenever $a \in A_2, b \in A_3 \cup A_4$;

(iii) $n \succ_a b$ whenever $a \in A_3, b \in A_1 \cup A_2$;

and the remaining pairwise comparisons are arbitrary. We consider the election $\mathcal{E} = (A, k, \succ)$ with $A = [n]$.

In what follows, we distinguish whether f selects more agents from $A_1 \cup A_2$ or from $A_3 \cup A_4$ and construct appropriate distance metrics to show that, in either case, the distortion is at least the one claimed in the statement. Intuitively, if f selects more agents from $A_1 \cup A_2$ we will consider a metric where these sets lie on one extreme, $A_4 = n$ on the other extreme, and all agents A_3 in the middle. This way, the selected committee gives twice the social cost as picking all agents from A_3 In the opposite case, we will construct a symmetric instance.

Formally, we first consider the case with $|S \cap (A_1 \cup A_2)| \geq \frac{k}{2}$ and define the distance metric d_1 on A by the following positions $x \in (-\infty, \infty)^n$: $x_a = 0$ for every $a \in A_1 \cup A_2$, $x_a = 1$ for every $a \in A_3$, and $x_n = 2$. It is not hard to check that $d_1 \rhd \succ$; see Figure 9 for an illustration. Since $|S \cap (A_1 \cup A_2)| \geq \frac{k}{2}$, we obtain $\mathrm{SC}(S, n; d_1) = 2$ and thus

$$\mathrm{dist}(f(\succ), \mathcal{E}) \geq \frac{\mathrm{SC}(S, A; d_1)}{\mathrm{SC}(A_3, A; d_1)} \geq \frac{\mathrm{SC}(S, n; d_1)}{\mathrm{SC}(A_3, n; d_1)} \geq 2.$$

Conversely, if $|S \cap (A_3 \cup A_4)| \geq \frac{k}{2}$, we define the distance metric d_2 on A by the following positions $x \in (-\infty, \infty)^n$: $x_1 = 0$, $x_a = 1$ for every $a \in A_2$, and $x_a = 2$ for every $a \in A_3 \cup A_4$. It is not hard to check that $d_2 \rhd \succ$; see Figure 9 for an illustration. Since $|S \cap (A_3 \cup A_4)| \geq \frac{k}{2}$, we obtain $\mathrm{SC}(S, 1; d_2) = 2$ and thus

$$\mathrm{dist}(f(\succ), \mathcal{E}) \geq \frac{\mathrm{SC}(S, A; d_2)}{\mathrm{SC}(A_2, A; d_2)} \geq \frac{\mathrm{SC}(S, 1; d_2)}{\mathrm{SC}(A_2, 1; d_2)} \geq 2.$$

Since $\mathrm{dist}(f(\succ), \mathcal{E}) \geq 2$ regardless of $f(\succ)$, we conclude that $\mathrm{dist}(f) \geq 2$. $\square$