On the Distortion of Committee Election with 1-Euclidean Preferences and Few Distance Queries

Dimitris Fotakis[1,2], **Laurent Gourvès**[3], **Panagiotis Patsilinakos**[3,4]
[1]National Technical University of Athens, Greece
[2]Archimedes/Athena RC, Greece
[3]Université Paris-Dauphine, Université PSL, CNRS, LAMSADE, 75016, Paris, France
[4]Athens University of Economics and Business, Greece
fotakis@cs.ntua.gr, laurent.gourves@dauphine.fr, patsilinak@mail.ntua.gr

Abstract

We consider committee election of $k \geq 3$ (out of $m \geq k+1$) candidates, where the voters and the candidates are associated with locations on the real line. Each voter's cardinal preferences over candidates correspond to her distance to the candidate locations, and each voter's cardinal preferences over committees is defined as her distance to the nearest candidate elected in the committee. We consider a setting where the true distances and the locations are unknown. We can nevertheless have access to degraded information which consists of an order of candidates for each voter. We investigate the best possible distortion (a worst-case performance criterion) w.r.t. the social cost achieved by deterministic committee election rules based on ordinal preferences submitted by n voters and few additional distance queries. We show that for any $k \geq 3$, the best possible distortion of any deterministic algorithm that uses at most $k-3$ distance queries cannot be bounded by any function of n, m and k. We present deterministic algorithms for k-committee election with distortion of $O(n)$ with $O(k)$ distance queries and $O(1)$ with $O(k \log n)$ distance queries.

1 Introduction

Electing a set of representatives based on the preferences submitted by voters is a central problem in social choice. In typical applications, the voters express *ordinal* preferences over the set of candidates, which are consistent with their *cardinal* preferences, but do not include any quantitative information about the *strength* of each preference. An important reason for resorting to ordinal information has to do with the cognitive difficulty of quantifying preferences. However, crucial information may be lost when voters summarize cardinal to ordinal preferences.

Procaccia and Rosenschein (2006) introduced the framework of *utilitarian distortion* as a means to quantify the efficiency loss, due to the fact that elections are based on ordinal information only, and to investigate the sensitivity of different voting rules to the absence of cardinal information. The distortion of a voting rule is the worst-case approximation ratio of its social welfare to the optimal social welfare achievable when cardinal information is available. Previous work has quantified the best possible distortion of single and multiwinner voting rules (often assuming normalized cardinal utilities, see e.g., (Boutilier et al. 2015; Caragiannis and Procaccia 2011; Caragiannis et al. 2017)).

Motivated by the frequent use of spatial preferences in social choice (e.g., (Enelow and Hinich 1984)), Anshelevich et al. (2018) introduced the framework of *metric distortion*, where the voters and the candidates are associated with locations in a metric space. The voters' cardinal preferences over candidates correspond to their distance to the candidate locations. The voters rank the candidates in increasing order of distance and submit this information to the voting rule. Without knowledge of the voter and candidate locations and distances, the voting rule aims to minimize the sum of distances (a.k.a. the *social cost*) of the voters to the candidate elected. Distortion is now defined w.r.t. the social cost. There has been significant interest in analyzing the metric distortion of prominent voting rules (see e.g., (Anshelevich et al. 2018; Goel, Krishnaswamy, and Munagala 2017; Kempe 2020; Anagnostides, Fotakis, and Patsilinakos 2022)) and in designing voting rules with optimal metric distortion for single-winner elections (Gkatzelis, Halpern, and Shah 2020; Kizilkaya and Kempe 2023).

There has not been much previous work on the metric distortion of multiwinner voting, where we elect a committee of $k \geq 2$ (out of $m \geq k+1$) candidates based on ordinal preferences submitted by n voters. As before, the voters' cardinal preferences correspond to their distance to the candidate locations. However, there are different ways to define the voter cardinal preferences over k-committees, resulting in different types and desirable properties of multiwinner elections (see e.g., (Elkind et al. 2017; Faliszewski et al. 2017)).

Goel, Hulett, and Krishnaswamy (2018) and Chen, Li, and Wang (2020) were the first to consider the metric distortion of committee elections with the cost of each voter for a committee defined as the sum of her distances to all committee members. Goel, Hulett, and Krishnaswamy (2018) proved that the best possible distortion in this setting is equal to the best possible distortion of single-winner voting. Chen, Li, and Wang (2020) proved that single-vote rules achieve a best possible distortion for the case where $k = m - 1$.

Caragiannis, Shah, and Voudouris (2022) considered the metric distortion of k-committee election with the cost of each voter for a committee defined as her distance to the q-th nearest member. They proved an interesting trichotomy: the distortion is unbounded if $q \leq k/3$, $\Theta(n)$ if $q \in (k/3, k/2]$, and equal to the best possible metric distortion of single-winner election if $q > k/2$. For the most interesting case where $q = 1$ and each voter's cost is her distance to the nearest committee member, their results imply that the distortion is $\Theta(n)$ if $k = 2$, and unbounded for all $k \geq 3$, with their lower bounds proven for the real line.

Subsequently, Burkhardt et al. (2024) and Pulyassary (2022) considered the metric distortion of classical clustering problems, such as k-median (which corresponds to k-committee election with $q = 1$) and k-center, for $k \geq 2$, in a setting where the clustering algorithm receives only ordinal information about demand points' locations and may query few distances. They focused on the case where the voter and the candidate locations coincide (a.k.a. *peer selection*), and asked about the minimum number of distance queries required for constant distortion. For k-median, Pulyassary (2022) proved that $O(1)$ distortion can be achieved deterministically with $O(n\text{poly}(\log n))$ distance queries and by a randomized algorithm with $O(nk)$ queries. Burkhardt et al. (2024) gave a randomized $O(1)$-distortion algorithm with $O(k^4 \log^5 n)$ queries. As for k-center, Burkhardt et al. (2024) showed how to implement the classical 2-approximate greedy algorithm with $k(k-1)/2$ queries and presented a deterministic 4-distortion algorithm with only $2k$ distance queries. Burkhardt et al. (2024) also proved lower bounds showing that in general metric spaces, their query bounds are not far from best possible.

Motivation and Objective. In this work, we study the metric distortion of k-committee elections where the cost of each voter for a committee is defined as her distance to the nearest member (i.e., we have $q = 1$). Our setting is conceptually close to (and strongly motivated by) the prominent rules of Chamberlin and Courant (1983) and Monroe (1995), which aim to elect a diverse committee that best reflects the preferences of the entire population of voters (Elkind et al. 2017; Faliszewski et al. 2017).

Our approach is rather orthogonal to (Burkhardt et al. 2024; Pulyassary 2022). We consider the more general (and more demanding) setting where the sets (and the locations) of voters and candidates may be different, and focus on deterministic rules and on the simplest (but nevertheless interesting and challenging enough) case of 1-*Euclidean preferences*, where the voters and the candidates are embedded in the real line. As in (Burkhardt et al. 2024; Pulyassary 2022) (and also motivated by the success of Amanatidis et al. (2021, 2022a,b) in improving the utilitarian distortion for single-winner elections and one-sided matchings with cardinal queries), we aim to shed light on the following:

Question 1. *How many distance queries are required for a bounded (or even constant) distortion in k-committee election with 1-Euclidean preferences, for $k \geq 3$?*

The case of 1-Euclidean preferences in particularly interesting because it allows for a maximum possible exploitation of ordinal preferences towards achieving low distortion with a small number of distance queries. Moreover, the lower bounds of Burkhardt et al. (2024) are based on tree metrics and do not apply to 1-Euclidean preferences. As for the upper bounds of Burkhardt et al. (2024) and Pulyassary (2022), though very strong and informative about the power of distance queries in k-median and k-center, there are two key difficulties towards applying them to our setting where the voter and candidate locations do not coincide: (i) in general metric spaces, we do not know how to extract (even approximate) information about candidate-to-candidate or voter-to-voter distances from "regular" voter-to-candidate distance queries (which request information present in voter cardinal preferences); and (ii) to the best of our understanding, the algorithms of Burkhardt et al. (2024); Pulyassary (2022) require ordinal information about how candidate and/or voter locations are ranked in increasing order of distance to certain candidate locations; we do not know how such ordinal information can be extracted from voter ordinal preferences, if the voter and the candidate locations are different.

Contribution and Techniques. We consider the general metric distortion setting, where the voter and the candidate locations may be different, focus on the simplest case of the line metric (and deal with difficulties (i) and (ii) above), and provide almost best possible answers to Question 1.

In Section 3, we review three different query types (voter-to-candidate, candidate-to-candidate and voter-to-voter). We show that the answer to queries of the second and the third types can be obtained from a small constant number of voter-to-candidate distance queries. Thus, we can rely on the more convenient candidate-to-candidate distance queries.

In Section 4, we lower bound the number of distance queries required for bounded distortion. We show that for any $k \geq 3$, the distortion of any deterministic rule that uses at most $k-3$ distance queries and selects k out of $m \geq 2(k-1)$ candidates on the real line is not bounded by any function of n, m and k (Theorem 1). Our construction shows that a bounded distortion is not possible if we restrict distance queries to few top candidates of each voter.

In Section 5, we asymptotically match our lower bound with a greedy voting rule, which uses at most $6(k-3)+3$ queries and achieves a distortion of at most $5n$ (Theorem 2). It is based on the classical greedy algorithm for k-center (Williamson and Shmoys 2010, Section 2.2) (as it happens with Polar-Opposites in (Caragiannis, Shah, and Voudouris 2022) and the k-center algorithm in (Burkhardt et al. 2024, Section 3.1)). For a query-efficient implementation of greedy, we consider the intervals defined by pairs of elected candidates that are consecutive on the real line, and compute the most distant candidate in each such interval using the voters' ordinal preferences and at most 3 distance queries. Interestingly, greedy achieves a distortion of at most 5 for the egalitarian cost, where (as in the k-center objective) we aim to minimize the maximum voter cost.

In Section 6, we show how to achieve low distortion with a small number of distance queries by selecting a small representative set of candidates and focusing on the restricted instance induced by them (Theorem 3). To demonstrate the usefulness of this reduction, in Section 7, we exploit a generalization of the greedy rule. Our construction for selecting a small representative set of candidates is inspired by the notion of *coresets*, extensively used for k-median in computational geometry (see e.g., (Frahling and Sohler 2005)). Our construction uses $O(k \log n)$ distance queries and computes a set of $O(k \log n)$ representative candidates that allow for a distortion of 5 (Theorem 4). The idea is to maintain a hierarchical partitioning of the candidate axis into intervals, so that the contribution of the voters associated with each interval to the social cost is bounded.

The proofs and the technical details omitted from this extended abstract due to space constraints can be found at (Fotakis, Gourvès, and Patsilinakos 2024).

Related Work. Metric distortion was introduced in (Anshelevich et al. 2018), where the distortion of popular voting rules for single-winner elections was studied. Subsequent work analyzed the metric distortion of popular voting rules, such as STV (Anagnostides, Fotakis, and Patsilinakos 2022). Munagala and Wang (2019) and Kempe (2020) presented deterministic rules with distortion $2 + \sqrt{5}$, breaking the barrier of 5 achieved by Copeland. Gkatzelis, Halpern, and Shah (2020) introduced Plurality Matching and proved that it achieves an optimal distortion of 3 in general metric spaces (see also (Kizilkaya and Kempe 2023)). Anshelevich and Zhu (2021) studied the distortion of single and multiwinner elections with known candidate locations. The survey of Anshelevich et al. (2021) provides a detailed overview.

Boutilier et al. (2015) and Caragiannis et al. (2017) studied the best possible utilitarian distortion of single and multiwinner elections, respectively. Amanatidis et al. (2021) significantly improved on the best possible utilitarian distortion for single-winner elections using cardinal information. They introduced single-winner voting rules with distortion $O(m^{1/(\ell+1)})$ using $O(n\ell \log m)$ value queries. Subsequently, Amanatidis et al. (2022a,b) significantly improved on the best possible utilitarian distortion for one-sided matchings using algorithms that resort to a small number of value queries per voter. Interestingly, our query bounds are linear in the size k of the committee and only logarithmic in the number n of voters (instead of linear in n).

The main result of Fotakis, Gourvès, and Monnot (2016) implies that the metric distortion of a maximization version of k-committee election with 1-Euclidean preferences is 3, for $k \in \{1, 2\}$, and at most $\frac{2k-1}{2k-3}$, for any $k \geq 3$. I.e., the distortion for the maximization version of k-committee election on the line tends to 1 as the committee size k increases.

Multiwinner voting is a significant research direction in social choice and has been studied from many different viewpoints, e.g., proportional representation (Aziz et al. 2017; Peters and Skowron 2020), axiomatic justification (Elkind et al. 2017), core-stability in restricted domains (Pierczyński and Skowron 2022). Selection of a single candidate or a committee of candidates based on 1-Euclidean preferences submitted by voters (or agents) is a typical setting in social choice and mechanism design and has been the topic of previous work (see e.g., (Miyagawa 2001; Procaccia and Tennenholtz 2013; Fotakis and Tzamos 2014; Feldman, Fiat, and Golomb 2016) for mechanism design, and (Fotakis and Gourvès 2022) and a few references in (Anshelevich et al. 2021) for social choice and distortion).

2 Model and Notation

We consider a set $\mathcal{C} = \{c_1, \dots, c_m\}$ of m candidates and a set $\mathcal{V} = \{v_1, \dots, v_n\}$ of n voters. We assume that they are located on the real line $\mathbb{R}$, i.e., each candidate c_i (resp. voter v_j) is associated with a location $x(c_i) \in \mathbb{R}$ (resp. $x(v_j) \in \mathbb{R}$). For brevity, we usually let c_i (resp. v_j) denote both the candidate (resp. the voter) and their location $x(c_i)$ (resp. $x(v_j)$). We always index candidates in increasing order of their coordinates, i.e., $c_1 < c_2 < \cdots < c_m$, that is the order in which they appear on the candidate axis from left to right. We let $\mathcal{C}[c, c'] = \mathcal{C} \cap [c, c']$ be the set (or interval) of candidates in $\mathcal{C}$ between c and c' on the candidate axis.

For each voter v, we let her L_1 distance to the candidate locations quantify her cardinal preferences over $\mathcal{C}$. I.e., v's cost for being represented by a candidate c is

$$\mathrm{cost}_v(c) = d(v, c) = |x(v) - x(c)| = |v - c| \,.$$

For a voter v and a set $S \subseteq \mathcal{C}$ of candidates, we let $d(v, S) = \min_{c \in S}\{d(v, c)\} = \min_{c \in S}\{|v - c|\}$. Motivated by the Chamberlin and Courant (1983) rule for k-committee election, we assume that each voter v is represented by (or is assigned to) her nearest candidate in any given set S of elected candidates. Formally, for any $S \subseteq \mathcal{C}$, we let $\mathrm{cost}_v(S) = d(v, S) = \min_{c \in S}\{d(v, c)\}$ be the cost experienced by v from the set S of elected candidates.

Problem Definition. The problem of **k-Committee Election** is to select a candidate set (a.k.a. committee) $S \subseteq \mathcal{C}$, with $|S| = k \leq m - 1$, that minimizes the (utilitarian) *social cost* $\mathrm{SC}(S) = \sum_{v \in \mathcal{V}} \mathrm{cost}_v(S)$ of the voters. We also consider the *egalitarian cost* $\mathrm{EC}(S) = \max_{v \in \mathcal{V}} \mathrm{cost}_v(S)$ of the voters for a k-committee S of elected candidates. We often refer to $(\mathcal{C}, \mathcal{V})$, where $\mathcal{C}$ is the set of candidates and $\mathcal{V}$ is the set of voters, along with their locations on the real line (which are assumed fixed, but unknown to the voting rule), as an *instance* of k-Committee Election.

Committee Election with 1-Euclidean Preferences and Distance Queries. k-Committee Election can be solved in $O(nk \log n)$ time, by dynamic programming (Hassin and Tamir 1991), if we have access to the voter and the candidate locations on the real line (or to all voter-candidate distances). However, in our setting, every voter v provides only a ranking $\succ_v$ over the set $\mathcal{C}$ of candidates that is consistent with the function $\mathrm{cost}_v : \mathcal{C} \to \mathbb{R}_{\geq 0}$. Namely, for every two candidates c and c', $c \succ_v c'$ (i.e., v prefers c to c') if and only if $d(v, c) < d(v, c')$. As usual in relevant literature (see e.g., (Anshelevich et al. 2021, Section 2)), we assume that for every voter v, $\succ_v$ is a strict total order, i.e., that for every pair of candidates c and c', $d(v, c) \neq d(v, c')$.

Our committee election rules receive a ranking profile $\vec{\succ} = (\succ_1, \dots, \succ_n)$ consisting of a strict total order $\succ_j$ over $\mathcal{C}$ for each voter $v_j \in \mathcal{V}$. We only consider 1-*Euclidean* ranking profiles $\vec{\succ}$, in the sense that all $\succ_v$ in $\vec{\succ}$ are consistent with a cost function cost_v computed w.r.t. some fixed (and common) collection of voter and candidate locations on the real line. Under the assumption that total orders $\succ_v$ are strict, 1-Euclidean ranking profiles are *single-peaked* (Black 1948) and *single-crossing* (Karlin 1968; Mirrlees 1971), properties that have received significant attention in computational social choice (see e.g., (Escoffier, Lang, and Öztürk 2008; Elkind and Faliszewski 2014; Elkind, Lackner, and Peters 2022) and the references therein). Doignon and Falmagne (1994) and Elkind and Faliszewski (2014) show that given a ranking profile $\vec{\succ}$, we can verify if $\vec{\succ}$ is 1-Euclidean and compute in polynomial time a strict ordering of the candidates on the real line, from left to right, that is consistent with $\vec{\succ}$. We refer to such an ordering (which is unique up to symmetry under a mild assumption) as the *candidate axis*.

A *deterministic rule* R for k-committee election receives a 1-Euclidean ranking profile $\vec{\succ} = (\succ_1, \dots, \succ_n)$ over a set $\mathcal{C}$ of m candidates, the desired committee size k and a non-negative integer q. Then, using $\vec{\succ}$ and information about the distance of at most q candidate pairs on the real line, R computes a committee $R(\vec{\succ}, k, q) = S \subseteq \mathcal{C}$ with k candidates. Our committee election rules assume availability of the candidate axis corresponding to $\vec{\succ}$ and may ask distance queries *adaptively*. We assume that the responses to all distance queries are consistent with a fixed collection of voter and candidate locations on $\mathbb{R}$ that result in $\vec{\succ}$.

Distortion. We evaluate the performance of a committee election rule R (often called *rule* or *algorithm*, for brevity) for given ranking profile $\vec{\succ}$, committee size k and query number q in terms of its *distortion* (Boutilier et al. 2015; Procaccia and Rosenschein 2006), i.e, the worst-case approximation ratio that R achieves w.r.t. the social cost:

$$\mathrm{dist}(R, \vec{\succ}, k, q) = \sup \frac{\mathrm{SC}(R(\vec{\succ}, k, q))}{\min_{S:|S|=k} \mathrm{SC}(S)}\,, \qquad (1)$$

where the supremum is taken over all collections of voter and candidate locations on the real line that are consistent with $\vec{\succ}$ and with the responses to the q candidate distance queries asked by R. The distortion of a deterministic k-committee rule R is the maximum of $\mathrm{dist}(R, \vec{\succ}, k, q)$ over all linear ranking profiles $\vec{\succ}$ with n voters and m candidates. We sometimes also consider the distortion w.r.t. the egalitarian cost $\mathrm{EC}(S)$ by explicitly referring to it.

Notation. We let $\mathrm{top}(v)$ be the top candidate of voter v in $\succ_v$. A candidate's c cluster $\mathrm{Cluster}(c)$ consists of all voters in $\mathcal{V}$ with c as their top candidate. We say that a candidate c is *active* if $\mathrm{Cluster}(c) \neq \emptyset$, i.e., there is some voter v with c as her top choice. We assume *non-degenerate* ranking profiles $\vec{\succ}$, where $n \geq k + 1$ and all candidates are located between the leftmost and the rightmost active candidate. We note that the candidate axis, determined from $\vec{\succ}$ by the algorithm of Elkind and Faliszewski (2014), is guaranteed to be unique (up to symmetry) for non-degenerate ranking profiles $\vec{\succ}$.

We always assume that the candidate set $\mathcal{C}$ given as input to our algorithms consists of *active candidates only*. An instance is *candidate-restricted*, if all candidates are active and all voters are moved to the location of their top candidate. Assuming that all voters are colocated with their top candidates (and then removing inactive candidates) increases the distortion by a factor of 3 (see also Theorem 3).

The analysis of our distortion bounds sometimes refer to candidate-restricted instances. However, our algorithms work without assuming anything about voter and candidate locations and our distortion bounds hold against an optimal solution for the original instance, where candidates may be inactive and candidate and voter locations may be different.

There is a delicate issue that restricts the use of candidate-restricted instances in our algorithms (and with which our algorithms carefully deal): When we use a ranking $\succ_v$ in an algorithm, we have to take care of the fact that $\succ_v$ may be different from the ranking $\succ_{\mathrm{top}(v)}$, where the candidates are ranked in increasing order of distance to $\mathrm{top}(v)$ (because the locations of v and $\mathrm{top}(v)$ may be different). The difficulty of deducing useful information about the rankings $\succ_c$ at candidate locations c from a voter ranking profile $\vec{\succ}$ imposes a significant difference between our setting and the clustering setting in (Burkhardt et al. 2024; Pulyassary 2022).

3 Distance Query Types

Before analyzing the distortion of committee election rules with distance queries, we discuss different types of them:

Regular queries. Given a voter $v \in \mathcal{V}$ and a candidate $c \in \mathcal{C}$, we ask for the distance $d(v, c) = |v - c|$.

Candidate queries. Given two candidates $c, c' \in \mathcal{C}$, we ask for the distance $d(c, c') = |c - c'|$.

Voter queries. Given two voters $v, v' \in V$, we ask for the distance $d(v, v') = |v - v'|$.

Regular queries ask for information available in the voter cardinal preferences $\mathrm{cost}_v : \mathcal{C} \to \mathbb{R}_{\geq 0}$. We can show how to simulate candidate queries and voter queries with at most six and two regular queries, respectively. Therefore, as long as we care about the asymptotics of the number of queries, we may use these types of queries interchangeably. Hence, we state and analyze our committee election rules assuming access to candidate queries.

4 Lower Bound on the Number of Queries Required for Bounded Distortion

For any $k \geq 2$ and $m \geq k + 1$, the candidate axis can be reconstructed from $m - 1$ distance queries. Then, we can find the optimal k-committee for the corresponding candidate-restricted instance using dynamic programming. This implies a distortion of 3 (for both the social and the egalitarian cost) using $m - 1$ distance queries. We next show that:

Theorem 1. *For any $k \geq 3$, the distortion of any deterministic k-committee election rule that uses at most $k - 3$ distance queries and selects k out of at least $2(k - 1)$ candidates on the real line cannot be bounded by any function of n, m and k (for both the social and the egalitarian cost).*

Proof Sketch. For every $k \geq 3$, we construct a family of $2(k - 1)$ instances on the real line with $k - 1$ candidate pairs each that cannot be distinguished with fewer than $k - 2$ distance queries. The distances are chosen so that in any committee with a bounded distortion, both candidates of a particular pair must be chosen, along with one candidate from each of the remaining pairs.

For the construction, we consider $m = 2(k - 1)$ candidates, $c_1 < c_2 < \cdots < c_{2k-3} < c_{2k-2}$. We let D sufficiently large, so that $D^2 \gg \max\{2D + 1, k\}$, and an $\epsilon \in (0, 1/k)$ sufficiently small used for tie breaking. In the

$$c_1 \overset{1}{\text{—}} c_2 \overset{D^2}{\text{————}} c_3 \overset{1}{\text{—}} c_4 \overset{D^2+\epsilon}{\text{————}} c_5 \overset{1}{\text{—}} c_6 \overset{D^2+2\epsilon}{\text{————}} c_7 \overset{1}{\text{—}} c_8 \overset{D^2+3\epsilon}{\text{————}} c_9 \overset{1}{\text{—}} c_{10}$$

Figure 1: The basic instance used in the lower bound of Theorem 1 for $k = 6$.

basic instance, we let $d(c_{2i-1}, c_{2i}) = 1$, for all $i \in [k-1]$, and let $d(c_{2i}, c_{2i+1}) = D^2 + (i-1)\epsilon$, for all $i \in [k-2]$ (see Fig. 1). There are $n = m$ voters, each with a different top candidate. The voters are colocated with their top candidate in the basic instance and its variants.

We construct a family of $2(k-1)$ different variants of the basic instance, by moving candidate c_j, $j = 1, \ldots, 2(k-1)$, by D, while keeping every other candidate at her original location. Specifically, in the j-th variant, if j is odd, we increase the distance $d(c_j, c_{j+1})$ from 1, in the basic instance, to $D+1$, by moving candidate c_j by D to the left. In the j-th variant, if j is even, we increase the distance $d(c_{j-1}, c_j)$ from 1, in the basic instance, to $D+1$, by moving candidate c_j by D to the right. All other candidates maintain the locations that they have in the basic instance.

Intuitively, the basic instance (and each variant) consists of $k-1$ essentially isolated candidate pairs. In each variant, the candidates of exactly one pair are far away from each other (so for a bounded distortion, we need to identify this pair and elect both candidates), while the candidates of the remaining pairs are quite close to each other (so we may elect any of them). Since the $2(k-1)$ variants are symmetric otherwise, any distance query that discovers that the distance of a candidate pair is as in the basic instance can exclude at most two variants (from the list of all possible instances used in this proof). Therefore, any deterministic rule requires at least $k-2$ distance queries in the worst case, before it is able to identify the candidate pair at distance D to each other. $\square$

An interesting question is if bounded distortion is possible with distance queries restricted to the few top candidates of each voter. In the proof of Theorem 1, we can embed a large group of candidates, located extremely close to each other, in each location c_i in the basic instance and its variants (and also have a voter colocated with each of them). The (many) candidates of each group are ranked first by each voter in the same group. Hence, unless we are allowed to query distances to candidates further down the voters' rankings, we cannot get any useful information about these instances.

5 Bounded Distortion with $\Theta(k)$ Queries

We present a simple greedy rule for k-committee election with bounded distortion and $\Theta(k)$ distance queries, thus asymptotically matching the lower bound of Theorem 1.

We show that the classical 2-approximate greedy algorithm for k-center can be implemented with few distance queries. The greedy algorithm (Williamson and Shmoys 2010, Section 2.2) iteratively maintains a set S of candidates, starting with any candidate, and adding the candidate c with maximum distance $d(c, S)$ to the current set S in each iteration. Algorithm 1 shows how greedy is applied to 1-Euclidean instances.

Algorithm 1: Greedy for k-committee election

Input: Candidates $\mathcal{C} = \{c_1, \ldots, c_m\}$, $k \in \{2, \ldots, m-1\}$, distance function $d : \mathcal{C} \times \mathcal{C} \to \mathbb{R}_{\geq 0}$.
Output: Set $S \subseteq \mathcal{C}$ of k candidates.

$S \leftarrow \{c_1, c_m\}$ {pick leftmost and rightmost candidates}
while $|S| < k$ **do**
  $\hat{c} \leftarrow \arg\max_{c \in \mathcal{C}}\{d(c, S)\}$
  $S \leftarrow S \cup \{\hat{c}\}$
end while
return S

To implement Algorithm 1 with distance queries (see Algorithm 2), we need to compute the most distant candidate in $\mathcal{C}$ to current candidate set $S = \{c_1, \ldots, c_\ell\}$, while $\ell < k$. For convenience, we let the candidates in S be indexed as they appear on the candidate axis, from left to right, and c_1 (resp. c_ℓ) is the leftmost (resp. rightmost) candidate in S.

Algorithm 2 maintains a set $\hat{C}$ with $\ell - 1$ candidate-distance pairs $(\hat{c}_i, \delta_i)$, where for each $i \in [\ell-1]$, $\hat{c}_i$ is the most distant candidate in the interval $\mathcal{C}[c_i, c_{i+1}]$ to the interval's endpoints $c_i, c_{i+1} \in S$ and $\delta_i = d(\hat{c}_i, \{c_i, c_{i+1}\})$.

This information is provided by the Distant-Candidate algorithm. Every time a new candidate c, lying between $c_i, c_{i+1} \in S$ on the axis, is added to S, the most distant candidates $\hat{c}_i$ and his distance δ_i to $\{c_i, c\}$ and $\hat{c}_{i+1}$ and his distance δ_{i+1} to $\{c, c_{i+1}\}$ are computed by two calls to Distant-Candidate and are added to $\hat{C}$ (Algorithm 2, step 10).

In the next iteration, the most distant candidate in $\mathcal{C}$ to current S is computed (step 4) and added to S (step 5). For step 4, we observe that the most distant candidate to S corresponds to the pair $(c, \delta) \in \hat{C}$, where c has maximum distance δ to his neighbors in S among all other $(c', \delta') \in \hat{C}$.

Distant-Candidate. The Distant-Candidate algorithm gets two candidates c, c', makes at most 3 distance queries, and returns a pair $(\hat{c}, \delta)$, where $\hat{c} \in \mathcal{C}[c, c']$ is the most distant candidate to $\mathcal{C}[c, c']$'s endpoints and $\delta = d(\hat{c}_i, \{c, c'\})$.

For the intuition, we initially assume access to the rankings $\succ_c$ and $\succ_{c'}$, where all candidates in $\mathcal{C}$ are listed in increasing order of distance to c and c', respectively.

Due to the linear structure of $\mathcal{C}$, the most distant candidate $\hat{c} = \arg\max_{c'' \in \mathcal{C}[c,c']}\{d(c'', \{c, c'\})\}$ can be computed as follows: Starting with c and moving from left to right on the candidate interval $\mathcal{C}[c, c']$, we find the rightmost candidate $c_l \in \mathcal{C}[c, c']$ that prefers c to c' and the leftmost candidate $c_r \in \mathcal{C}[c, c']$ that prefers c' to c. We note that c_l and c_r can be found using only ordinal information, they are next to each other on the candidate axis, and $\hat{c}$ must be either c_l or c_r. Then, $d(c_l, \{c, c'\}) = d(c_l, c)$ and $d(c_r, \{c, c'\}) = d(c_r, c')$. Hence, $\hat{c}$ is c_l, if $d(c_l, c) > d(c_r, c')$, and c_r otherwise, which can be determined by 2 queries $d(c_l, c)$ and $d(c_r, c')$.

Algorithm 2: Query-efficient implementation of Greedy

Input: Candidates $\mathcal{C} = \{c_1, \ldots, c_m\}$, $k \in \{2, \ldots, m-1\}$, voter ranking profile $\vec{\succ} = (\succ_1, \ldots, \succ_n)$
Output: Set $S \subseteq \mathcal{C}$ of k candidates

1: $S \leftarrow \{c_1, c_m\}$ {pick leftmost and rightmost candidates}
2: $\hat{C} \leftarrow \{\text{Distant-Candidate}(\mathcal{C}[c_1, c_m])\}$
3: **while** $|S| < k$ **do**
4: Let c be s.t. $(c, \delta) \in \hat{C}$ and $\delta \geq \delta'$ for all $(c', \delta') \in \hat{C}$
5: $S \leftarrow S \cup \{c\}$
6: $\hat{C} \leftarrow \hat{C} \setminus \{(c, \delta)\}$
7: **if** $|S| < k$ **then**
8: Let c_i be the rightmost candidate in S on c's left
9: Let c_{i+1} be the leftmost candidate in S on c's right
10: $\hat{C} \leftarrow \hat{C} \cup \{\text{Distant-Candidate}(\mathcal{C}[c_i, c])\} \cup \{\text{Distant-Candidate}(\mathcal{C}[c, c_{i+1}])\}$
11: **end if**
12: **end while**
13: **return** S

With some more care, we can implement Distant-Candidate so that it computes $(\hat{c}, d(\hat{c}, \{c, c'\}))$ based on the ordinal information submitted by the voters (i.e., it uses a ranking $\succ_v$, submitted by voter $v \in \text{Cluster}(c)$, instead of the ranking $\succ_c$ above). Since $\succ_v$ and $\succ_c$ may differ (because the locations of v and c may be different), we use the information provided by a 3rd distance query.

Theorem 2. *For any $k \geq 3$, Algorithm 2 achieves a distortion of at most $5n$ for the social cost (and at most 5 for the egalitarian cost) for k-Committee Election on the real line using at most $6k - 15$ candidate distance queries.*

Proof. Algorithm 2 adapts Algorithm 1 to the case where candidate locations are not known. It calls Distant-Candidate, which computes the most distant candidate $\hat{c}_i$ in each interval $\mathcal{C}[c_i, c_{i+1}]$ defined by the candidates already elected in S. The Distant-Candidate algorithm is called once in step 2 and $2(k-3)$ times in step 10 (twice in each while-loop iteration, for $|S| = 3, \ldots, k-1$). So, the total number of distance queries is at most $6(k-3)+3$.

The correctness of Algorithm 2 (i.e., the fact that in each iteration, the candidate c with maximum $d(c, S)$ is added to S) follows from the discussion above. The distortion bound for the egalitarian cost uses that Algorithm 1 is 2-approximate for the egalitarian cost in candidate-restricted instances (Williamson and Shmoys 2010, Theorem 2.3). Based on this, we can show an upper bound of 5 on the distortion for the egalitarian cost in general instances. The bound of $5n$ on the distortion for the utilitarian social cost holds because for any $S \subseteq \mathcal{C}$, $\text{EC}(S) \leq \text{SC}(S) \leq n\,\text{EC}(S)$. We can also show that the clusters created by Algorithm 2 give $\Omega(n)$ distortion for the social cost. $\square$

6 Low Distortion via Good Subsets

We next analyze the distortion achieved by the optimal k-committee of the candidate-restricted instance induced by a small representative set of candidates.

We say that a candidate subset $\mathcal{C}' \subseteq \mathcal{C}$ is (ℓ, β)**-good**, for some $\ell \geq k$ and some $\beta \geq 1$, if $|\mathcal{C}'| = \ell$ and $\text{SC}(\mathcal{C}') \leq \beta\,\text{SC}(S^*)$, where S^* is an optimal k-committee for the original instance. Namely, $\mathcal{C}'$ is *ℓ-sparse*, in the sense that $\mathcal{C}'$ includes $\ell \leq m$ candidates (ideally $\ell \ll m$), and is *β-good*, in the sense that representing each voter by her top candidate in $\mathcal{C}'$ imposes a social cost at most β times the optimal social cost. The original set $\mathcal{C}$ of candidates is $(m, 1)$-good, while any k-committee with distortion β is (k, β)-good.

Given an (ℓ, β)-good set of candidates $\mathcal{C}'$, we let $\mathcal{C}'_{\text{cr}} = \{(c_1, n_1), \ldots, (c_\ell, n_\ell)\}$ denote the candidate-restricted instance induced by $\mathcal{C}'$. Then, $c_1 < \cdots < c_\ell$ denote the locations of candidates in $\mathcal{C}'$ on the line, and $n_i = |\text{Cluster}(c_i)|$ is the number of voters with c_i as their top candidate in $\mathcal{C}'$. We maintain that $n_1 + \cdots + n_\ell = n$ and that each $n_i > 0$ (the latter by removing inactive candidates from $\mathcal{C}'$).

The following shows that an optimal k-committee for the candidate-restricted instance $\mathcal{C}'_{\text{cr}}$ induced by an (ℓ, β)-good set $\mathcal{C}'$ achieves a distortion of $1+2\beta$ for the original instance.

Theorem 3. *Let $(\mathcal{C}, \mathcal{V})$ be an instance of the k-Committee Election, let $\mathcal{C}' \subseteq \mathcal{C}$ be an (ℓ, β)-good set, let $\mathcal{C}'_{cr}$ be the candidate-restricted instance induced by $\mathcal{C}'$ and let S (resp. S^*) be an optimal k-committee for $\mathcal{C}'_{cr}$ (resp. for $(\mathcal{C}, \mathcal{V})$). Then,* $\text{SC}(S) \leq (1+2\beta)\text{SC}(S^*)$.

Proof. For each voter v (with her location v as in the original instance), we let $\text{top}'(v) \in \mathcal{C}'$ be v's top candidate in $\mathcal{C}'$. Then, by the triangle inequality, $d(v, S) \leq d(v, \text{top}'(v)) + d(\text{top}'(v), S)$. Summing up over all voters $v \in \mathcal{V}$, we obtain:

$$\text{SC}(S) \leq \text{SC}(\mathcal{C}') + \text{SC}(\mathcal{C}'_{\text{cr}}, S)\,, \tag{2}$$

where $\text{SC}(\mathcal{C}') = \sum_{v \in \mathcal{V}} d(v, \text{top}'(v)) = \sum_{v \in \mathcal{V}} d(v, \mathcal{C}')$ and $\text{SC}(\mathcal{C}'_{\text{cr}}, S) = \sum_{v \in V} d(\text{top}'(v), S) = \sum_{i=1}^{\ell} n_i\, d(c_i, S)$ is the social cost of S for the candidate-restricted instance $\mathcal{C}'_{\text{cr}}$.

We observe that $\text{SC}(\mathcal{C}'_{\text{cr}}, S) \leq \text{SC}(\mathcal{C}'_{\text{cr}}, S^*)$, because in the candidate-restricted instance $\mathcal{C}'_{\text{cr}}$, we can replace candidates in $S^* \setminus \mathcal{C}'$ with candidates in S without increasing the social cost. Moreover, since $d(\text{top}'(v), S^*) \leq d(\text{top}'(v), v) + d(v, S^*)$, we obtain that $\text{SC}(\mathcal{C}'_{\text{cr}}, S^*) \leq \text{SC}(\mathcal{C}') + \text{SC}(S^*)$.

Combined with the observations above, (2) implies that:

$$\text{SC}(S) \leq 2\text{SC}(\mathcal{C}') + \text{SC}(S^*) \leq (1+2\beta)\text{SC}(S^*)\,,$$

where the second inequality follows from the hypothesis that $\mathcal{C}'$ is a (ℓ, β)-good set of candidates. $\square$

As soon as we have the distances between all active candidates in an (ℓ, β)-good set $\mathcal{C}'$, which requires $\ell - 1$ distance queries, an optimal k-committee for the candidate-restricted instance $\mathcal{C}'_{\text{cr}}$ induced by $\mathcal{C}'$ can be computed in polynomial time by dynamic programming.

7 Hierarchical Partitioning for Good Subsets

Next, we use hierarchical partitioning of the candidate axis (see Algorithm 3) and compute a $(O(k \log n), O(1))$-good set of candidates with $O(k \log n)$ distance queries.

In Algorithm 3, each triple (or *interval*) in $\mathcal{I}$ consists of a candidate interval $\mathcal{C}[c_a, c_b]$, the number n_{ab} of voters v with $\text{top}(v) \in \mathcal{C}[c_a, c_b]$, and the interval length $d(c_a, c_b)$.

Algorithm 3: Hierarchical partitioning of $\mathcal{C}[c_1, c_m]$

Input: Candidates $\mathcal{C} = \{c_1, \dots, c_m\}$, $k \in \{2, \dots, m-1\}$, voter ranking profile $\vec{\succ} = (\succ_1, \dots, \succ_n)$
Output: Partitioning $\mathcal{I}$ of $\mathcal{C}$ into $O(k \log n)$ intervals.

1: Let $S = \{c^1, \dots, c^k\}$ be the result of Algorithm 2
2: $\mathcal{I} \leftarrow \{\, (\mathcal{C}[c_a^1, c_b^1], n_1, d(c_a^1, c_b^1)), \dots,$
$\qquad (\mathcal{C}[c_a^k, c_b^k], n_k, d(c_a^k, c_b^k)) \,\}$
$\quad$ {Start with the partitioning of $(\mathcal{C}, \mathcal{V})$ induced by S}
3: $\delta^* \leftarrow \max_{i \in [k]} \{d(c_a^i, c_b^i)\}$
4: **while** $|\mathcal{I}| \le 7k(\log_2(5nk) + 2)$ **do**
5: $\quad$ Let $(\mathcal{C}[c_a, c_b], n_{ab}, d(c_a, c_b)) \in \mathcal{I}$ with $|\mathcal{C}[c_a, c_b]| \ge 4$ and maximum weight $\mathrm{wt}(c_a, c_b) = n_{ab} d(c_a, c_b)$
6: $\quad$ **if** $\mathrm{wt}(c_a, c_b) \le \delta^*/(5k)$ **then break**
7: $\quad$ $\mathcal{I} \leftarrow \big(\mathcal{I} \setminus \{(\mathcal{C}[c_a, c_b], n_{ab}, d(c_a, c_b))\}\big)$
$\qquad \cup \text{Partitioning}(\mathcal{C}[c_a, c_b])$
8: **end while**
9: **return** $\mathcal{I}$

We refer to $\mathrm{wt}(c_a, c_b) = n_{ab} d(c_a, c_b)$ as the *weight* of interval $\mathcal{C}[c_a, c_b]$. Algorithm 3 computes a partitioning $\mathcal{I}$ of the candidate axis $\mathcal{C}[c_1, c_m]$ into at most $7k(\log_2(5nk) + 2)$ intervals by repeatedly splitting the interval in $\mathcal{I}$ with largest weight (and at least 4 candidates) into two intervals.

Algorithm 3 starts with the partitioning induced by the k-committee $S = \{c^1, \dots, c^k\}$ computed by Algorithm 2 (we let $c^1 < c^2 < \cdots < c^k$). In step 2, for each $i \in [k]$, $\mathcal{C}[c_a^i, c_b^i]$ is the interval that includes all candidates in $\mathcal{C}$ closer to c^i than to any other candidate in S (c_a^i, resp. c_b^i, is the leftmost, resp. the rightmost, such candidate), n_i is the number of voters v with $\mathrm{top}(v) \in \mathcal{C}[c_a^i, c_b^i]$, and $d(c_a^i, c_b^i)$ is the length of $\mathcal{C}[c_a^i, c_b^i]$. We let $\delta^* = \max_{i \in [k]} \{d(c_a^i, c_b^i)\}$. Then, $\mathrm{SC}(S^*) \ge \delta^*/5$, where S^* is an optimal k-committee for the original instance, because S has distortion 5 for the egalitarian cost and we assume that all candidates in $\mathcal{C}$ are active.

For the interval split, in step 4, we use the Partitioning algorithm, which is a modification of Distant-Candidate used in Section 5. The Partitioning algorithm uses at most 4 distance queries and splits each interval $\mathcal{C}[c_a, c_b]$ into two intervals $(\mathcal{C}[c_a, c_l], n_{al}, d(c_a, c_l))$ and $(\mathcal{C}[c_r, c_b], n_{rb}, d(c_r, c_b))$, where c_l (resp. c_r) is the rightmost (resp. leftmost) candidate in $\mathcal{C}[c_a, c_b]$ on the left (resp. right) of the interval's midpoint $(c_a + c_b)/2$. The correctness essentially follows from the discussion about Distant-Candidate in Section 5.

The collection of intervals $\mathcal{I}$ maintained by Algorithm 3 is always a partitioning of $\mathcal{C}$. To obtain an $(O(k \log n), O(1))$-good set $\mathcal{C}'(\mathcal{I}) \subseteq \mathcal{C}$ from $\mathcal{I}$, we include in $\mathcal{C}'(\mathcal{I})$ the endpoints c_a and c_b of (resp. all candidates in) each interval $\mathcal{C}[c_a, c_b]$ in $\mathcal{I}$ with more than (resp. at most) 3 candidates. Since $|\mathcal{I}| = O(k \log n)$, $\mathcal{C}'(\mathcal{I})$ consists of $O(k \log n)$ candidates. The following shows that $\mathrm{SC}(\mathcal{C}'(\mathcal{I})) \le 2\,\mathrm{SC}(S^*)$.

Theorem 4. *Let $\mathcal{I}$ be the partitioning of $\mathcal{C}$ computed by Algorithm 3. Then, the resulting set $\mathcal{C}'(\mathcal{I}) \subseteq \mathcal{C}$ is a $(O(k \log n), 2)$-good set of candidates.*

Proof. We first upper bound $\mathrm{SC}(\mathcal{C}'(\mathcal{I}))$. We call an interval $\mathcal{C}[c_a, c_b] \in \mathcal{I}$ *expensive*, if $\mathcal{C}[c_a, c_b] \cap S^* \neq \emptyset$ (i.e., $\mathcal{C}[c_a, c_b]$ includes an optimal candidate), and *cheap* otherwise. For each voter v with $\mathrm{top}(v)$ in a cheap interval $\mathcal{C}[c_a, c_b]$, $\mathrm{cost}_v(\mathcal{C}') \le \mathrm{cost}_v(S^*)$, because the interval's endpoints $c_a, c_b \in \mathcal{C}'$, while v's nearest candidate in S^* is outside $\mathcal{C}[c_a, c_b]$. Therefore, the total contribution to $\mathrm{SC}(\mathcal{C}'(\mathcal{I}))$ of all voters associated with cheap intervals in $\mathcal{I}$ is at most their contribution to $\mathrm{SC}(S^*)$. The contribution to $\mathrm{SC}(\mathcal{C}'(\mathcal{I}))$ of the voters v with $\mathrm{top}(v)$ in an expensive interval $\mathcal{C}[c_a, c_b]$ is at most their contribution to $\mathrm{SC}(S^*)$ plus $\mathrm{wt}(c_a, c_b)$. Since each expensive interval includes a candidate of S^*, there are at most k expensive intervals in $\mathcal{I}$. Adapting an argument of (Frahling and Sohler 2005), we next show that as soon as $|\mathcal{I}| > 7k(\log_2(5nk) + 2)$, each interval in $\mathcal{I}$ has weight at most $\mathrm{SC}(S^*)/k$. Thus, at this point, the additional cost due to the total weight of expensive intervals is at most $\mathrm{SC}(S^*)$.

We refer to an interval $\mathcal{C}[c_a, c_b]$ as *light*, if $\mathrm{wt}(c_a, c_b) \le \mathrm{SC}(S^*)/k$, and as *heavy*, otherwise. Splitting an interval $\mathcal{C}[c_a, c_b]$ results in two intervals $\mathcal{C}[c_a, c_l]$ and $\mathcal{C}[c_r, c_b]$ each of length (resp. weight) at most half the length (resp. weight) of $\mathcal{C}[c_a, c_b]$. Hence, splitting a light interval replaces it with two light intervals in $\mathcal{I}$. Moreover, since Algorithm 3 always splits the heaviest interval in $\mathcal{I}$, from the first iteration that it splits a light interval and on, all intervals in $\mathcal{I}$ are light.

A heavy interval $\mathcal{C}[c_a, c_b]$ is *close*, if $d(\{c_a, c_b\}, S^*) < d(c_a, c_b)$, and *far* otherwise. An interval $\mathcal{C}[c_a, c_b]$ is *encountered* by Algorithm 3, if there is a point during its execution where $\mathcal{C}[c_a, c_b] \in \mathcal{I}$. We can show that the total number of far (resp. close) heavy intervals encountered by Algorithm 3 is at most $2k(\log_2(5nk) + 1)$ (resp. $5k(\log_2(5nk) + 1)$).

Since the total number of heavy intervals encountered by Algorithm 3 is at most $7k(\log_2(5nk) + 1)$, the first split of a light interval happens no later than iteration $7k(\log_2(5nk) + 1) + 1$. At that point all intervals in $\mathcal{I}$ are light and $|\mathcal{I}| \le 7k(\log_2(5nk) + 2) = O(k \log n)$ (we recall that $n \ge k$). $\square$

Since the Partitioning algorithm uses at most 4 distance queries per split, the total number of queries is $O(k \log n)$. Combining Theorem 3 with Theorem 4, we obtain that:

Theorem 5. *There is a polynomial-time deterministic rule for k-Committee Election that uses $O(k \log n)$ distance queries and achieves a distortion of at most* 5.

8 Directions for Further Research

Our work leaves several interesting open questions. First, in the proof of Theorem 4, dependence on $\log n$ seems necessary in order to obtain enough information about the locations of near optimal candidates. Hence, we conjecture a lower bound of $\Omega(k \log n)$ on the number of distance queries required for constant distortion for 1-Euclidean preferences.

Our results crucially exploit the 1-Euclidean structure. It would be interesting if bounded distortion can be achieved with few distance queries for the case where the voters and the candidates are embedded in $\mathbb{R}^d$ and the voters provide a ranking of the candidates in each dimension.

For general metric spaces, it would be interesting if constant distortion can be achieved with $O(k \log n)$ distance queries for *perturbation-stable* instances (e.g., (Makarychev and Makarychev 2021)), where the different clusters of voters are somewhat easier to identify.

Acknowledgements

This work has been partially supported by project MIS 5154714 of the National Recovery and Resilience Plan Greece 2.0 funded by the European Union under the NextGenerationEU Program, and by the Hellenic Foundation for Research and Innovation (H.F.R.I.) under the "First Call for H.F.R.I. Research Projects to support Faculty members and Researchers and the procurement of high-cost research equipment grant", project BALSAM, HFRI-FM17-1424.

Laurent Gourvès is supported by Agence Nationale de la Recherche (ANR), project THEMIS ANR-20-CE23-0018.

Panagiotis Patsilinakos has been partially supported by the framework of H.F.R.I call "Basic research Financing (Horizontal support of all Sciences)" under the National Recovery and Resilience Plan Greece 2.0 funded by the European Union under the NextGenerationEU (H.F.R.I. Project Number: 15877). Most of this work was performed while Panagiotis Patsilinakos was with the National Technical University of Athens.

References

Amanatidis, G.; Birmpas, G.; Filos-Ratsikas, A.; and Voudouris, A. A. 2021. Peeking behind the ordinal curtain: Improving distortion via cardinal queries. *Artificial Intelligence*, 296: 103488.

Amanatidis, G.; Birmpas, G.; Filos-Ratsikas, A.; and Voudouris, A. A. 2022a. Don't Roll the Dice, Ask Twice: The Two-Query Distortion of Matching Problems and Beyond. In *Proc. of the 35th Conference on Advances in Neural Information Processing Systems (NeurIPS 2022)*.

Amanatidis, G.; Birmpas, G.; Filos-Ratsikas, A.; and Voudouris, A. A. 2022b. A Few Queries Go a Long Way: Information-Distortion Tradeoffs in Matching. *Journal of Artificial Intelligence Research*, 74.

Anagnostides, I.; Fotakis, D.; and Patsilinakos, P. 2022. Dimensionality and Coordination in Voting: The Distortion of STV. In *Proc. of the 36th AAAI Conference on Artificial Intelligence (AAAI 2022)*, 4776–4784. AAAI Press.

Anshelevich, E.; Bhardwaj, O.; Elkind, E.; Postl, J.; and Skowron, P. 2018. Approximating optimal social choice under metric preferences. *Artificial Intelligence*, 264: 27–51.

Anshelevich, E.; Filos-Ratsikas, A.; Shah, N.; and Voudouris, A. A. 2021. Distortion in Social Choice Problems: The First 15 Years and Beyond. In *Proc. of the 30th International Joint Conference on Artificial Intelligence (IJ-CAI 2021)*, 4294–4301.

Anshelevich, E.; and Zhu, W. 2021. Ordinal Approximation for Social Choice, Matching, and Facility Location Problems Given Candidate Positions. *ACM Transactions in Economics and Computation*, 9(2): 9:1–9:24.

Aziz, H.; Brill, M.; Conitzer, V.; Elkind, E.; Freeman, R.; and Walsh, T. 2017. Justified representation in approval-based committee voting. *Social Choice and Welfare*, 48(2): 461–485.

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

Boutilier, C.; Caragiannis, I.; Haber, S.; Lu, T.; Procaccia, A. D.; and Sheffet, O. 2015. Optimal social choice functions: A utilitarian view. *Artificial Intelligence*, 227: 190–213.

Burkhardt, J.; Caragiannis, I.; Fehrs, K.; Russo, M.; Schwiegelshohn, C.; and Shyam, S. 2024. Low-Distortion Clustering with Ordinal and Limited Cardinal Information. In *Proc. of the 38th AAAI Conference on Artificial Intelligence (AAAI 2024)*, 9555–9563. AAAI Press.

Caragiannis, I.; Nath, S.; Procaccia, A. D.; and Shah, N. 2017. Subset selection via implicit utilitarian voting. *Journal of Artificial Intelligence Research*, 58: 123–152.

Caragiannis, I.; and Procaccia, A. D. 2011. Voting almost maximizes social welfare despite limited communication. *Artificial Intelligence*, 175(9-10): 1655–1671.

Caragiannis, I.; Shah, N.; and Voudouris, A. A. 2022. The metric distortion of multiwinner voting. *Artificial Intelligence*, 313: 103802.

Chamberlin, J. R.; and Courant, P. N. 1983. Representative Deliberations and Representative Decisions: Proportional Representation and the Borda Rule. *American Political Science Review*, 77(3): 718–733.

Chen, X.; Li, M.; and Wang, C. 2020. Favorite-candidate voting for eliminating the least popular candidate in a metric space. In *Proc. of the 34th AAAI Conference on Artificial Intelligence (AAAI 2020)*, 1894–1901. AAAI Press.

Doignon, J.-P.; and Falmagne, J.-C. 1994. A polynomial time algorithm for unidimensional unfolding representations. *Journal of Algorithms*, 16(2): 218–233.

Elkind, E.; and Faliszewski, P. 2014. Recognizing 1-Euclidean Preferences: An Alternative Approach. In *Proc. of the 7th Symposium on Algorithmic Game Theory (SAGT 2014)*, 146–157.

Elkind, E.; Faliszewski, P.; Skowron, P.; and Slinko, A. 2017. Properties of multiwinner voting rules. *Social Choice and Welfare*, 48(3): 599–632.

Elkind, E.; Lackner, M.; and Peters, D. 2022. Preference Restrictions in Computational Social Choice: A Survey. arXiv:2205.09092.

Enelow, J. M.; and Hinich, M. J. 1984. *The spatial theory of voting: An introduction*. Cambridge University Press.

Escoffier, B.; Lang, J.; and Öztürk, M. 2008. Single-peaked consistency and its complexity. In *Proc. of the 18th European Conference on Artificial Intelligence (ECAI 2008)*, 366–370.

Faliszewski, P.; Skowron, P.; Slinko, A.; and Talmon, N. 2017. *Multiwinner voting: A new challenge for social choice theory*, volume 74 of *Trends in Computational Social Choice*, 27–47. Lulu Publisher.

Feldman, M.; Fiat, A.; and Golomb, I. 2016. On Voting and Facility Location. In *Proc. of the 17th ACM Conference on Economics and Computation (EC 2016)*, 269–286. ACM.

Fotakis, D.; and Gourvès, L. 2022. On the distortion of single winner elections with aligned candidates. *Autonomous Agents and Multi-Agent Systems*, 36(2): 37.

Fotakis, D.; Gourvès, L.; and Monnot, J. 2016. Conference Program Design with Single-Peaked and Single-Crossing Preferences. In *Proc. of the 12th Conference on Web and Internet Economics (WINE 2016)*, volume 10123 of *Lecture Notes in Computer Science*, 221–235. Springer.

Fotakis, D.; Gourvès, L.; and Patsilinakos, P. 2024. On the Distortion of Committee Election with 1-Euclidean Preferences and Few Distance Queries. arXiv:2408.11755.

Fotakis, D.; and Tzamos, C. 2014. On the Power of Deterministic Mechanisms for Facility Location Games. *ACM Transactions on Economics and Computation*, 2(4): 15:1–15:37.

Frahling, G.; and Sohler, C. 2005. Coresets in dynamic geometric data streams. In *Proc. of the 37th ACM Symposium on Theory of Computing (STOC 2005)*, 209–217. ACM.

Gkatzelis, V.; Halpern, D.; and Shah, N. 2020. Resolving the optimal metric distortion conjecture. In *Proc. of the 61st IEEE Symposium on Foundations of Computer Science (FOCS 2020)*, 1427–1438. IEEE.

Goel, A.; Hulett, R.; and Krishnaswamy, A. K. 2018. Relating Metric Distortion and Fairness of Social Choice Rules. In *Proc. of the 13th Workshop on Economics of Networks, Systems and Computation (NetEcon@SIGMETRICS 2018)*, 4:1. ACM.

Goel, A.; Krishnaswamy, A. K.; and Munagala, K. 2017. Metric distortion of social choice rules: Lower bounds and fairness properties. In *Proc. of the 18th ACM Conference on Economics and Computation (EC 2017)*, 287–304.

Hassin, R.; and Tamir, A. 1991. Improved complexity bounds for location problems on the real line. *Operations Research Letters*, 10(7): 395–402.

Karlin, S. 1968. *Total Positivity*. Stanford University Press.

Kempe, D. 2020. An Analysis Framework for Metric Voting based on LP Duality. In *Proc. of the 34th AAAI Conference on Artificial Intelligence (AAAI 2020)*, 2079–2086. AAAI Press.

Kizilkaya, F. E.; and Kempe, D. 2023. Generalized Veto Core and a Practical Voting Rule with Optimal Metric Distortion. In *Proc. of the 24th ACM Conference on Economics and Computation (EC 2023)*, 913–936. ACM.

Makarychev, K.; and Makarychev, Y. 2021. *Perturbation Resilience*, chapter 5, 95–119. Beyond the Worst-Case Analysis of Algorithms. Cambridge University Press.

Mirrlees, J. A. 1971. An exploration in the theory of optimal income taxation. *Review of Economic Studies*, 38: 175––208.

Miyagawa, E. 2001. Locating Libraries on a Street. *Social Choice and Welfare*, 18: 527–541.

Monroe, B. L. 1995. Fully proportional representation. *Americal Political Science Review*, 89(4): 925–940.

Munagala, K.; and Wang, K. 2019. Improved Metric Distortion for Deterministic Social Choice Rules. In *Proc. of 20th ACM Conference on Economics and Computation (EC 2019)*, 245–262. ACM.

Peters, D.; and Skowron, P. 2020. Proportionality and the Limits of Welfarism. In *Proc. of the 21st ACM Conference on Economics and Computation (EC 2020)*, 793–794. ACM.

Pierczyński, G.; and Skowron, P. 2022. Core-Stable Committees Under Restricted Domains. In *Proc. of the 18th Conference on Web and Internet Economics (WINE 2022)*, volume 13778 of *Lecture Notes in Computer Science*, 311–329. Springer.

Procaccia, A. D.; and Rosenschein, J. S. 2006. The Distortion of Cardinal Preferences in Voting. In *Proc. of the 10th Workshop on Cooperative Information Agents (CIA 2006)*, 317–331.

Procaccia, A. D.; and Tennenholtz, M. 2013. Approximate Mechanism Design without Money. *ACM Transactions on Economics and Computation*, 1(4): 18:1–18:26.

Pulyassary, H. 2022. *Algorithm Design for Ordinal Settings*. Master's thesis, University of Waterloo, Waterloo, Ontario, Canada. Available at http://hdl.handle.net/10012/18668.

Williamson, D.; and Shmoys, D. 2010. *The Design of Approximation Algorithms*. Cambridge University Press.