Communication, Distortion, and Randomness in Metric Voting

David Kempe
University of Southern California

Abstract

In distortion-based analysis of social choice rules over metric spaces, voters and candidates are jointly embedded in a metric space. Voters rank candidates by non-decreasing distance. The mechanism, receiving only this ordinal (comparison) information, must select a candidate approximately minimizing the sum of distances from all voters to the chosen candidate. It is known that while the Copeland rule and related rules guarantee distortion at most 5, the distortion of many other standard voting rules, such as Plurality, Veto, or k-approval, grows unboundedly in the number n of candidates.

An advantage of Plurality, Veto, or k-approval with small k is that they require less communication from the voters; all deterministic social choice rules known to achieve constant distortion require voters to transmit their complete rankings of all candidates. This motivates our study of the tradeoff between the distortion and the amount of communication in deterministic social choice rules.

We show that any one-round deterministic voting mechanism in which each voter communicates only the candidates she ranks in a given set of k positions must have distortion at least $\frac{2n-k}{k}$; we give a mechanism achieving an upper bound of $O(n/k)$, which matches the lower bound up to a constant. For more general communication-bounded voting mechanisms, in which each voter communicates b bits of information about her ranking, we show a slightly weaker lower bound of $\Omega(n/b)$ on the distortion.

For randomized mechanisms, Random Dictatorship achieves expected distortion strictly smaller than 3, almost matching a lower bound of $3 - \frac{2}{n}$ for any randomized mechanism that only receives each voter's top choice. We close this gap, by giving a simple randomized social choice rule which only uses each voter's first choice, and achieves expected distortion $3 - \frac{2}{n}$.

1   Introduction

In *voting* or *social choice*, there is a set of n alternatives (such as political candidates or courses of action) from which a group (such as a country or an organization) wants to select a winner. Each voter submits a *ranking* (or *preference order*) of the candidates, and the *mechanism* (or *social choice rule*) chooses a winner based on these submitted rankings.

Contrasting and complementing the traditional "axiomatic" approach of comparing social choice rules, an approach that has recently gained in popularity is to consider social choice as an *optimization problem* with the goal of selecting the "best" candidate for the population (Boutilier et al. 2015; Caragiannis and Procaccia 2011; Procaccia 2010; Procaccia and Rosenschein 2006). Each voter has a utility (or cost) for each candidate, and the mechanism's goal is to optimize the aggregate utility or cost of all voters. However, as articulated in (Boutilier and Rosenschein 2016; Anshelevich, Bhardwaj, and Postl 2015), the social choice rule has to optimize with crucial information missing: a voter can only communicate her[1] *ranking* according to the utility/cost. In other words, the mechanism receives only *ordinal* information — which candidate is preferred over which other candidate — even though it needs to optimize a *cardinal* objective function. From an optimization perspective, this means that the mechanism should simultaneously optimize over all utility/cost functions that are *consistent* with the reported rankings, in that they would give rise to the observed rankings. The worst-case ratio (over all cost/utility functions) between the mechanism's cost/utility and that of the optimum candidate for the specific function is called the mechanism's *distortion*. (Formal definitions of all concepts and terms are given in Section 2.)

In applying this general framework, an important question is what class of cost/utility functions to consider. A natural approach was suggested in (Anshelevich, Bhardwaj, and Postl 2015) (see also the expanded/improved journal version (Anshelevich et al. 2018) and general overview (Anshelevich 2016)): all candidates and voters are jointly embedded in a metric space, and the cost of voter v for candidate x is their metric distance $d(v,x)$. The assumption that voters rank candidates by non-decreasing distance in a latent space[2] dates back to earlier work on so-called *single-peaked*

---

[1] For consistency and clarity, we will always refer to voters using female and candidates using male pronouns.

[2] It is an interesting "philosophical" question whether these distances are "real," e.g., in the sense that voters could articulate or quantify them. Whether this is the case will clearly depend on

preferences (see, e.g., (Black 1948; Moulin 1980)), though much of the earlier work focuses on the special case when the metric is the line. Using the framework of distortion and metric costs, (Anshelevich, Bhardwaj, and Postl 2015; Anshelevich et al. 2018) showed a remarkable separation. While many commonly used voting rules (such as Plurality, Veto, k-approval, Borda count) have either unbounded distortion or distortion linear in the number n of candidates, and indeed all score-based rules have distortion $\omega(1)$ (in terms of the number of candidates), *uncovered-set rules* have distortion at most 5. To describe uncovered-set rules, consider a tournament graph G on the n candidates which contains the directed edge (x, y) iff at least as many voters prefer x to y as vice versa. The uncovered set of G is the set of all candidates with paths of length at most 2 to all other candidates (Moulin 1986); an example of such a candidate is the candidate x with maximum outdegree, which is selected by the Copeland rule. (Anshelevich et al. 2018) show that any candidate in the uncovered set of G has distortion at most 5, and also show a lower bound of 3 on the distortion of *every* deterministic voting mechanism.

One advantage of some of the mechanisms with large distortion — such as Plurality, Veto, or k-approval with small k — is that they require little communication from the voters. A voter under Plurality only needs to share her first choice; similarly, a voter under Veto only needs to share her last choice. This observation raises the question of whether high distortion is *inherently* a consequence of limited communication between voters and the mechanism.

The answer to the preceding question is clearly "No:" there are simple *randomized* mechanisms achieving constant distortion. Perhaps the simplest is Random Dictatorship: "Return the first choice of a uniformly random voter." This mechanism is known to have distortion strictly smaller than 3 (Anshelevich and Postl 2016), a smaller distortion than any deterministic mechanism can achieve. However, despite the frequent mathematical appeal and elegance of randomized algorithms and mechanisms, most organizations are leery of using randomization for making important decisions;[3] hence, we consider determinism a very desirable property in the design of voting mechanisms. Considering the following three properties: (1) low distortion, (2) low communication, (3) determinism, it is known that any two can be achieved simultaneously:

- Random Dictatorship satisfies (1), (2).

- Uncovered-set mechanisms satisfy (1), (3).
- Plurality and many other mechanisms satisfy (2), (3).

The big-picture question we investigate in this article is the tradeoff between all *three* of these desirable properties.

1.1 Our Models and Results

We only consider the goal of minimizing the *average* (or total) metric distance of all voters from the winning candidate.[4] Our main result, proved in Section 4, is essentially a negative answer to the question of whether any voting mechanisms can simultaneously have all three desirable properties. We consider a model in which each voter communicates b bits of information about her ranking to the mechanism, in a single round.[5] Associated with each b-bit string μ is a subset Π_μ of rankings. The Π_μ form a disjoint cover of all possible rankings. If they did not form a cover, some voters might not have any message to send, making the mechanism ill-defined. And if the Π_μ were not disjoint, then it is not clear how a voter with multiple possible messages μ would make the (non-deterministic) choice which one to send; in particular, this choice could depend on the actual metric distances, and it might require much more subtle definitions to place meaningful restrictions on a mechanism to not exploit such information. Each voter communicates the (unique) μ such that her permutation is in Π_μ. Under this model, in Section 4, we prove the following lower bound:

Theorem 1.1 *Every one-round deterministic voting mechanism in which each voter sends only a b-bit string to the mechanism has distortion at least $\frac{2n-4}{b} - 1$.*

Most mechanisms with limited communication are of a fairly specific form: voters can communicate only their choices in a (small) set K of k positions of their ranking, typically at the top or bottom of their ballots. (Either giving the candidate for each such position, or specifying them as a set, as in k-approval.) For such restricted mechanisms, a simpler proof (in Section 3) gives a lower bound that is stronger by a factor $\Theta(\log n)$:

Theorem 1.2 *Any deterministic one-round social choice rule which receives, from each voter, no information about candidates outside positions K in her ranking, has distortion at least $\frac{2n-|K|}{|K|}$.*

The proof of Theorem 1.2 is significantly easier and cleaner than the proof of Theorem 1.1, while still containing some of the key ideas. Therefore, we present the proof of Theorem 1.2 *before* that of Theorem 1.1.

Theorem 1.2 provides a slight generalization of Theorem 1 of the recent work (Fain et al. 2019), who prove linear distortion for the special case when K consists of the top k positions. In fact, (Fain et al. 2019) show linear lower

---

the application, and will impact which models and types of mechanisms are the "right" ones. A more in-depth discussion can be found in the full version of this paper.

   [3]A reader taking issue with this statement may want to think about his/her own academic or industry research department. Even though these are likely among the most savvy organizations in terms of understanding randomization, decision making procedures practically never involve randomization, except the occasional coin flip to break a tie.

The reasons for such a preference in most organization likely include an aversion to variance or to low-probability undesirable events; naturally, one can envision guarantees between deterministic and expectation bounds, such as the bounds on the squared distortion in (Fain et al. 2019).

---

   [4](Anshelevich, Bhardwaj, and Postl 2015) and several follow-up articles studied both the average and median distance.

   [5]Analyzing the distortion of multi-round deterministic mechanisms with limited communication is a very interesting direction for future work.

bounds on the *expected squared distortion* of randomized mechanisms; this directly implies the same bounds for deterministic mechanisms.

The fact that the lower bound of Theorem 1.2 is stronger than that of Theorem 1.1 by a factor of $\Theta(\log n)$ is discussed in more detail in Section 4. Closing this $\Theta(\log n)$ gap is an interesting direction for future work, discussed in Section 7.

The reason we consider Theorem 1.1 our main contribution is that it helps us pinpoint the source of high distortion. Several recent works have shown lower bounds on the distortion of different specific classes of social choice rules, such as score-based rules (Anshelevich et al. 2018) or the above-mentioned top-k ballots (Fain et al. 2019). Our result implies that regardless of the intricacy of the mechanism, low communication (within the context studied here) and determinism are enough to force high distortion. Communication as a measure of complexity is fairly natural, as evidenced by the mechanisms typically used in practice for large numbers of alternatives. Communication can also be regarded as a proxy for cognitive effort imposed on the voters, although admittedly, the computation of a message μ in a general b-bit bounded mechanism may still require the voter to first determine her full ranking of all candidates.

The results of Theorems 1.1 and 1.2 are lower bounds, raising the question of how small one can make a mechanism's distortion when communication is limited. In Section 5, we address this question, proving the following theorem.

Theorem 1.3 *There is a one-round deterministic social choice rule which, given only each voter's top k candidates (in order), selects a candidate with distortion at most $\frac{79n}{k}$.*

The deterministic social choice rule of Theorem 1.3 is a generalization of the Copeland rule to top-k ballots. Up to constant factors,[6] the bounds of Theorems 1.2 and 1.3 match. Closing the gap between the upper and lower bound is likely difficult, as even for $k = n$, the best known lower bound of 3 does not match the best current upper bound of $2 + \sqrt{5} \approx 4.23$ due to (Munagala and Wang 2019); whether there is a deterministic mechanism with metric distortion 3 is a well-known open question. Notice also that Theorem 1.3 implies that knowing each voter's ranking for a constant fraction of candidates is sufficient to achieve constant distortion, a fact that may not be a priori obvious.

The main focus in this article is on *deterministic* mechanisms: as discussed earlier, the *Random Dictatorship* mechanism has distortion (slightly) smaller than 3, achieving small distortion and low communication simultaneously. (Gross, Anshelevich, and Xia 2017) prove a nearly matching lower bound: they show that every randomized social choice rule in which each voter only communicates her top $k < n/2$ candidates must have distortion at least $3 - \frac{2}{\lfloor n/k \rfloor}$. However, even for $k = 1$, this leaves a gap between the upper bound of essentially 3 for Random Dictatorship and the

---

[6]An application of Corollary 5.3 of (Kempe 2020) gives an upper bound of $\frac{12n}{k}$, which, however, is still far from matching the lower bound.

lower bound of $3 - \frac{2}{n}$. Recently, (Fain et al. 2019) shrunk this gap: they proved that the Random Oligarchy mechanism — which samples three voters and outputs a majority of first-place votes if it exists, and otherwise the choice of a random voter among the three — achieves expected distortion close to $3 - \frac{2}{n}$, though there still remains a small gap between the upper and lower bounds. As an additional result, in Section 6, we close this remaining gap:

Theorem 1.4 *There is a simple randomized social choice rule in which each voter only communicates her first-choice candidate, and which achieves distortion at most $3 - \frac{2}{n}$.*

Due to space constraints, the proofs of several results, as well as a more in-depth discussion of related work, are omitted from this version. They can be found in the full version on the arXiv (Kempe 2019).

1.2 Related Work

The impact of communication on social choice rules has been an area of active research; see, for instance, (Boutilier and Rosenschein 2016) for an overview, and the classic paper (Conitzer and Sandholm 2002) on *vote elicitation*. Most of the focus in this line of past work has been on the number of bits that need to be communicated in order to compute the outcome of a *particular* social choice rule, rather than on proving lower bounds arising due to limited communication when the social choice rule is not pre-specified.

Several very recent papers have explicitly considered the tradeoff between communication and distortion in social choice, both in deterministic and randomized settings.

Most immediately related is the recent paper (Fain et al. 2019). Their focus is on mechanisms with extremely low communication which achieve low expected squared distortion, a measure somewhere between expected distortion and deterministic distortion. They prove that the Random Referee mechanism, which asks two randomly chosen voters for their top choices, and asks a third voter to choose between these two choices, achieves constant expected squared distortion. Notice that this mechanism elicits different information from different voters. Theorem 1 of (Fain et al. 2019) shows that this is unavoidable, in that any mechanism that only obtains top-k lists (for constant k), even from all voters, must have linear expected squared distortion, implying the same result for the distortion of deterministic mechanisms.

Another very related recent paper is (Mandal et al. 2019), studying the communication-distortion tradeoff in a setting where the voters have utilities (instead of costs) for the candidates, and these utilities are only assumed to be non-negative and normalized, but do not need to satisfy any other properties (such as being derived from a metric). The other major modeling difference between our work and (Mandal et al. 2019) is that they assume that agents compute their message μ to the mechanism directly from their utility vector, rather than the ranking. In particular, the mechanism can be designed to allow voters to express the strength of their preferences, albeit in possibly coarse form. As a result, their results are not directly comparable to ours, although some of the techniques and preliminary results are similar.

Another related recent piece of work is on approval-based voting, due to (Pierczyński and Skowron 2019), who — among other results — analyze the distortion of approval-based voting. They consider mechanisms in which voters approve all candidates within a given *distance* of themselves, not bounded by a number of candidates. They show that under certain technical assumptions (which are fairly restrictive), this mechanism gives constant distortion, as low as 3 depending on exact parameter values.

Low communication complexity of voter preferences is also the focus of the recent paper (Bentert and Skowron 2019). They study the more "traditional" goal of implementing given voting rules with low communication (Boutilier and Rosenschein 2016), but are interested in *approximate* implementation of these rules. To make approximation meaningful, they focus on score-based rules, which naturally assign each candidate a score (such as Borda Count, Plurality, or MiniMax). Then, the quality of approximation is the ratio between the score of the winner under full information vs. the score of the winner under limited communication. They focus on mechanisms in which each voter is asked to rank a small subset of candidates; this subset is either the voter's top k candidates (a deterministic mechanism) or a random subset of k candidates (a randomized mechanism). Given that the goal in (Bentert and Skowron 2019) is the approximate implementation of specific scoring-based voting rules rather than achieving low distortion, the results are not directly comparable.

2 Preliminaries

2.1 Voters, Candidates, and Social Choice Rules

There are n candidates, which we always denote by lower-case letters at the end of the alphabet. Sets of candidates are denoted by uppercase letters, and X is the set of all candidates. The preference order (or ranking) of voter v over the candidates is a bijection $\pi_v : \{1, \ldots, n\} \to X$, mapping positions i to the candidate $x = \pi_v(i)$ which voter v ranks in position i. We say that v (strictly) *prefers* x to y iff $\pi_v^{-1}(x) < \pi_v^{-1}(y)$. The set of all voters is denoted by V. We write $\mathcal{S}_n$ for the set of all possible rankings $\pi : \{1, \ldots, n\} \to X$, and $\mathcal{P} = (\pi_v(i))_{v \in V, i \in \{1, \ldots, n\}}$ for the rankings of all voters, which we call the *vote profile*.

In the traditional full-information view, a *social choice rule* (we use the terms *mechanism* or *voting mechanism* interchangeably) $f : \mathcal{S}_n^V \to X$ is given the rankings of all voters, i.e., $\mathcal{P}$, and produces as output one *winning* candidate $w = f(\mathcal{P})$. For most of this article, we are interested only in *deterministic* social choice rules f.

2.2 Communication-bounded mechanisms

Our main contribution is to consider communication-bounded social choice rules. As in the standard model described above, we still only consider deterministic *single-round* mechanisms, i.e., each voter can only send a single message to the mechanism. However, this message is now also restricted to be at most b bits long.

This induces $M = 2^b$ sets $\Pi_1, \Pi_2, \ldots, \Pi_M$ of rankings; when the mechanism receives a message μ from voter v,

all it learns is that $\pi_v \in \Pi_\mu$. As discussed in the introduction, we assume that the Π_μ form a disjoint partition of $\mathcal{S}_n$, i.e., they are pairwise disjoint and cover all rankings: $\bigcup_{\mu=1}^{M} \Pi_\mu = \mathcal{S}_n$. The fact that M is a power of 2 is not relevant anywhere in our proofs, so we also consider mechanisms with arbitrary numbers M of sets.

Definition 2.1 *An* M*-communication bounded social choice rule consists of pairwise disjoint sets* $\Pi_1, \Pi_2, \ldots, \Pi_M \subseteq \mathcal{S}_n$ *with* $\bigcup_{\mu=1}^{M} \Pi_\mu = \mathcal{S}_n$, *and a deterministic mapping* $f : \{1, \ldots, M\}^V \to X$.

Communication-bounded social choice rules that are used in practice, such as Plurality, Veto, k-approval, and combinations thereof, are of a specific form: there is a set K of k positions, and voters can communicate the set of candidates they have in positions in K, possibly with an ordering, but cannot communicate any additional information about their ranking of candidates in positions outside K. For such mechanisms, we will be able to prove stronger lower bounds on the distortion, and with a significantly simpler proof. We define them formally as follows:

Definition 2.2 *A* k*-entry social choice rule is an* M*-communication bounded social choice rule with the following additional restriction on the sets* $\Pi_1, \Pi_2, \ldots, \Pi_M$*: there exists a set* $K \subseteq \{1, \ldots, n\}$ *of at most* k *positions such that if* π, π' *agree for all positions in* K, *i.e.,* $\pi(i) = \pi'(i)$ *for all* $i \in K$, *then* $\pi \in \Pi_\mu$ *if and only if* $\pi' \in \Pi_\mu$.

2.3 Metric Space and Distortion

The key modeling contribution of the metric-based distortion (Anshelevich, Bhardwaj, and Postl 2015) objective is to assume that all voters and candidates are embedded in a pseudo-metric space d. $d(v, x)$ denotes the distance between voter v and candidate x. Being a pseudo-metric, it satisfies non-negativity and the triangle inequality $d(v, x) \leq d(v, y) + d(v', y) + d(v', x)$ for all voters v, v' and candidates x, y. For our upper bounds, we explicitly allow the distance between candidates and voters (and thus also between pairs of candidates or pairs of voters) to be 0; however, for improved flow, we still refer to d as a *metric*. In our lower-bound constructions, all distances will be strictly positive; that is, we do not exploit the increased generality for negative results.

We say that a vote profile $\mathcal{P}$ is *consistent* with the metric d, and write $d \sim \mathcal{P}$, if $\pi_v(x) < \pi_v(y)$ whenever $d(v, x) < d(v, y)$. That is, $\mathcal{P}$ is consistent with d iff all voters rank candidates by non-decreasing distance from themselves. In case of ties among distances, i.e., $d(v, x) = d(v, y)$, several vote profiles are consistent with d. None of our results depend on any tie breaking assumptions.

The *cost* of candidate x is $C(x) = \sum_v d(v, x)$, i.e., the sum of distances of x to all voters.

The social choice rule is handicapped by not knowing the metric d, instead only observing the consistent vote profile $\mathcal{P}$ (or some limited information about it, when communication is restricted). Due to this handicap, and possibly other

suboptimal choices, it will typically not choose the candidate of smallest cost. The *distortion* of f is the worst-case ratio between the cost of the candidate chosen by f, and the minimum cost of any candidate (determined with knowledge of the actual distances d). Formally,

$$\rho(f) \;=\; \max_{\mathcal{P}} \; \sup_{d:d\sim\mathcal{P}} \frac{C(f(\mathcal{P}))}{\min_x C(x)}.$$

We can think of the distortion in terms of a game between the social choice rule and an adversary. First, the adversary chooses the vote profile $\mathcal{P}$. Then, the social choice rule, knowing only $\mathcal{P}$ (or part of that information, in case of communication restrictions), chooses a winning candidate $w = f(\mathcal{P})$. Then, the adversary chooses a metric d consistent with $\mathcal{P}$ that maximizes the ratio between the cost of the candidate chosen by f and the optimum candidate for d.

3 The Restricted Lower Bound

In this section, we establish the lower bound of Theorem 1.2, restated here formally.

Theorem 3.1 *Every one-round deterministic k-entry social choice rule has distortion at least $\frac{2n-k}{k}$.*

Proof. Let $K = \{\kappa_1 < \kappa_2 < \cdots < \kappa_k\}$. Because *every* deterministic social choice rule has distortion at least 3 (Anshelevich, Bhardwaj, and Postl 2015), we only need to consider the case where $2n - k > 3k$, i.e., $k < n/2$. We will prove the theorem by induction on n, with the base case $n = 2$ holding because the only such case with $k < n/2$ is $k = 0$, where the mechanism receives no information about any voter's preferences, and hence has unbounded distortion.

First, we consider the case $n \in K$. We designate one candidate $\hat{x}$ who is "infinitely" far from all voters, and thus ranked last by all of them. The mechanism clearly cannot choose $\hat{x}$ as a winner. This reduces the problem to one of $n - 1$ candidates, and a set $K' = K \setminus \{n\}$ of $k - 1$ positions at which voters specify their ranking. By induction hypothesis, applied to this instance, the distortion is lower-bounded by $\frac{2(n-1)-(k-1)}{k-1} = \frac{2n-k-1}{k-1} > \frac{2n-k}{k}$; the inequality holds because $k < n$.

For the remainder of the proof, we can assume that $n \notin K$, i.e., voters do not specify their least favorite candidate. In this case, we will not need to use the induction hypothesis for $n - 1$. For each subset $S \subseteq X, |S| = k$ of k candidates, and each ordering $\sigma : \{1, \ldots, k\} \to S$, we say that a voter v has *type* (S, σ) if she puts the candidates from S in the positions K, in the order given by σ. That is, v has type (S, σ) iff $\pi_v(\kappa_i) = \sigma(i)$ for $i = 1, \ldots, k$. There are $t = \binom{n}{k} \cdot k!$ types of voters. We define a vote profile which has exactly a $1/t$ fraction of voters of type (S, σ), for each type. Throughout, we will talk about fractions, rather than numbers, of voters, so that the total adds up to 1.

Let w be the candidate chosen by the social choice rule for this input. w is well-defined as a function of all voters' types, because (1) for each voter v, the message sent by v

is uniquely determined by her ranking of candidates in positions in K, and (2) the mechanism's output is a deterministic function of only the messages sent by the voters.

We now define a metric space. Let ϵ be a very small constant (we will let $\epsilon \to 0$), and $0 < \epsilon_1 < \epsilon_2 < \cdots < \epsilon_n < \epsilon$. Consider a voter v of type (S, σ). We distinguish two cases:

1. In the first case, $w \notin S$. Let π_v be any ordering that puts the candidates in S in positions K in the order σ, and which additionally has $\pi_v(n) = w$, i.e., candidate w is in the last position in v's ranking. Apart from this, π_v is arbitrary. By construction, a voter v with ranking π_v has type (S, σ). We now set the distance between v and the candidate w to 1, and the distance from v to every candidate $\pi_v(i)$ (for $i < n$) to $\epsilon + \epsilon_i$. These distances are consistent with the ranking π_v.

2. In the second case, $w \in S$. Again, let π_v be any permutation that puts the candidates in S in positions K in the order σ (ensuring that π_v is consistent with v having type (S, σ)). This time, the position of w in π is prescribed by S, σ, and we let the remaining positions of π_v be arbitrary. Voter v has distance exactly $\frac{1}{2} + \epsilon + \epsilon_i$ from each candidate $\pi_v(i)$, including the case when $\pi_v(i) = w$. Again, v ranks the candidates in the order given by π_v.

We now verify that these distances satisfy the triangle inequality. Consider voters v, v' and candidates x, y. We will show that $d(v, y) \le d(v, x) + d(v', x) + d(v', y)$, by distinguishing two cases for y:

1. In the first case, $y = w$. Then, $\frac{1}{2} + \epsilon \le d(v, y) \le 1$. Either the distance $d(v', y) = 1$, in which case the triangle inequality holds obviously, or $d(v', y) \ge \frac{1}{2} + \epsilon$, in which case our definition ensures that $d(v', x) \ge \frac{1}{2} + \epsilon$ as well. In either case, the triangle inequality holds.

2. In the second case, $y \neq w$, so either $\epsilon < d(v, y) < 2\epsilon$ or $\frac{1}{2} + \epsilon < d(v, y) < \frac{1}{2} + 2\epsilon$, depending on the case of the definition. Because all distances are lower-bounded by ϵ, the triangle inequality clearly holds if $d(v, y) < 2\epsilon$. In the other case $\frac{1}{2} + \epsilon < d(v, y)$, we have that $\frac{1}{2} + \epsilon < d(v, x)$, which together with $\epsilon < d(v', x)$ again ensures that the triangle inequality holds.

Recall that w is selected by the social choice rule under the given rankings. Each voter of type (S, σ) with $w \notin S$ has cost 1 for w, and cost at most 2ϵ for any candidate $x \neq w$. Each voter of type (S, σ) with $w \in S$ has cost at least $\frac{1}{2}$ for w, and cost at most $\frac{1}{2} + 2\epsilon$ for each candidate $x \neq w$.

Of the t types (S, σ), exactly $\binom{n-1}{k-1} \cdot k!$ have $w \in S$. Thus, the cost of candidate w is at least $\frac{1}{t} \cdot (\frac{1}{2} \cdot \binom{n-1}{k-1} \cdot k! + 1 \cdot (t - \binom{n-1}{k-1} \cdot k!))$, while the cost of any other candidate is at most $\frac{1}{t} \cdot (2\epsilon + \frac{1}{2} \cdot \binom{n-1}{k-1} \cdot k!)$. Letting $\epsilon \to 0$, the distortion approaches $1 + \frac{2\left(\frac{n!}{(n-k)!} - \frac{k \cdot (n-1)!}{(n-k)!}\right)}{\frac{k \cdot (n-1)!}{(n-k)!}} = \frac{2n-k}{k}$. $\blacksquare$

4 The General Lower Bound

In this section, we prove the more general lower bound of Theorem 1.1. The bound applies to all M-communication bounded social choice rules, but is slightly weaker than

that of Theorem 3.1. To gain some insight into general communication-bounded social choice rules, we begin with an easy proposition, independently obtained as Lemma 4.1 in (Mandal et al. 2019).

Proposition 4.1 *Assume that there exists a set Π_μ containing two rankings π, π' with $\pi(1) \neq \pi'(1)$, i.e., there is a μ which does not uniquely specify the voter's top-ranked candidate. Then, the corresponding social choice rule has unbounded distortion.*

Theorem 4.2 *Let f be any one-round M-communication bounded social choice rule on n candidates. Then, f must have distortion at least $\frac{2n-4}{\ln M} - 1$.*

Proof. The high-level idea of the proof is to use induction on the number of candidates, to show that when communication is "sufficiently bounded," any social choice rule must have high distortion. After completing the proof by induction, we would like to apply the result to n candidates, and "sufficiently bounded" must then include M-communication bounded. Therefore, the relationship between the number of candidates in the induction proof and the bound on communication depends on n, M, and to avoid notational ambiguity, we will use different variable names for the induction. Specifically, we use ν for the number of candidates within the induction proof, and M_ν for the upper bound on communication.

Let $\gamma = 1 - M^{-1/(n-2)}$. We will prove by induction on ν that every M_ν-communication bounded social choice rule on ν candidates with $M_\nu \leq \frac{1}{(1-\gamma)^{\nu-2}}$ has distortion at least $\frac{2}{\gamma} - 1$.

The base case $\nu = 2$ is easy: the communication bound is $M_2 \leq \frac{1}{(1-\gamma)^{2-2}} = 1$, so the voters cannot communicate any preference. By Proposition 4.1, the social choice rule has unbounded distortion. For the induction step, we distinguish two cases:

1. In the first case, we assume that for each candidate x, at least a $1 - \gamma$ fraction of all sets Π_μ contain a ranking $\pi_\mu \in \Pi_\mu$ that ranks x last, i.e., $\pi_\mu(\nu) = x$. Then, we consider a vote profile with M_ν voters in which for each $\mu = 1, \ldots, M_\nu$, exactly one voter submits μ.

   Let w be the candidate chosen by f. Consider the following metric space: For every voter v who submitted μ such that there is a ranking $\pi_\mu \in \Pi_\mu$ ranking w last, we define the distance between v and w to be 1, and the distance from all other candidates[7] to be 0. For all other voters, the distance to all candidates is $\frac{1}{2}$. Said differently, all candidates $x \neq w$ are at distance 0 from each other, and at distance 1 from w. All voters who could possibly rank w last are in the same location as the candidates different from w, while all other voters are halfway between w and the other candidates.

   Then, the cost of w is at least $\gamma \cdot \frac{1}{2} + (1 - \gamma) \cdot 1 = 1 - \frac{\gamma}{2}$, while the cost of each other candidate is at most $\gamma \cdot \frac{1}{2} +$

   ---
   [7]Ties could be broken by using small $\epsilon_i \to 0$ without affecting the final result.

$(1 - \gamma) \cdot 0 = \frac{\gamma}{2}$. Thus, the distortion of the mechanism is at least $\frac{2}{\gamma} - 1$, completing the proof directly.

2. Otherwise, let x be a candidate such that at most a $1 - \gamma$ fraction of all sets Π_μ contain a ranking $\pi_\mu \in \Pi_\mu$ that ranks x last. Define $M_{\nu-1}$ to be the number of such sets, and assume w.l.o.g. (by renumbering) that $\Pi_1, \Pi_2, \ldots, \Pi_{M_{\nu-1}}$ are all the sets which contain at least one ranking with x in the last position. By the assumption in this part of the proof, we have that $M_{\nu-1} \leq \lfloor (1 - \gamma) \cdot M_\nu \rfloor$. We will only construct instances in which all voters rank x last; thus, no voter communicates any message $\mu > M_{\nu-1}$.

   No mechanism with finite distortion can select x as a winner, by the same argument as in the preceding case. (That is, the metric puts x at distance 1 from all voters, and all other candidates at distance 0 from all voters.) As a result, we obtain an instance with $\nu - 1$ candidates, only $(\nu - 1)!$ remaining possible rankings, and — crucially — only $M_{\nu-1} \leq (1 - \gamma) \cdot M_\nu$ remaining sets of rankings. We can therefore apply the induction hypothesis for $\nu - 1$, and conclude that the mechanism's distortion is at least $\frac{2}{\gamma} - 1$.

To show that we can apply the inductive claim with $\nu = n$ in the end, observe that $M_n = M = M^{(n-2)/(n-2)} = \frac{1}{(1-\gamma)^{n-2}}$.

It remains to show that $\frac{2}{\gamma} - 1 \geq \frac{2n-4}{\ln M} - 1$. To do so, we rewrite γ by using the Taylor expansion of $t^{1/(n-2)}$ around $t = 1$, then apply straightforward bounds:

$$\gamma = 1 - M^{-1/(n-2)}$$

$$= \frac{1}{n-2} \sum_{k=1}^{\infty} \frac{1}{k} \cdot (1 - 1/M)^k \cdot \prod_{j=1}^{k-1} \left(1 - \frac{1}{j \cdot (n-2)} \right)$$

$$\leq \frac{1}{n-2} \sum_{k=1}^{\infty} \frac{1}{k} \cdot (1 - 1/M)^k = \frac{1}{n-2} \cdot \ln M.$$

Substituting this bound for γ into the distortion completes the proof. ∎

To compare the bound of Theorem 4.2 with that of Theorem 3.1, observe that when voters get to specify the candidates in each of k (given) positions in a ranking, this generates a partition of $\mathcal{S}_n$ into $M = \binom{n}{k} \cdot k! = \frac{n!}{(n-k)!}$ sets: one for each subset and order within that subset. These sets of rankings do in fact form a disjoint cover. For the "interesting" range $k \leq n/2$, we can simply bound $(n/2)^k \leq M \leq n^k$, so we get that $\ln M \approx k \ln n$. This shows that the lower bound of Theorem 4.2 is weaker than that of Theorem 3.1 by a factor of $\Theta(\log n)$.

5 A Near-Matching Upper Bound

In this section, we give a k-entry social choice rule whose distortion — up to constants — matches the lower bound of Theorem 3.1. This shows that the lower bound of Theorem 3.1 is essentially tight. Not surprisingly, the mechanism is a variation on uncovered set mechanisms, which are the

only type of mechanism known to achieve constant distortion even with access to the full vote profile.

In our mechanism, each voter communicates her top k choices. We say that voter v *prefers* x over y if either: (1) both x and y are among her top k, and she ranks x higher than y, or (2) x is among her top k, and y is not.

As in uncovered set mechanisms like Copeland, we construct a *comparison graph* G among the n candidates. Define $\alpha = \frac{k}{3n}$. For each ordered pair x, y, the graph G contains a directed edge (x, y) if and only if at least an α fraction of all voters prefer x over y. Notice that because $\alpha \leq \frac{1}{2}$, it is possible that G contains both (x, y) and (y, x). Similarly, it is possible that for a pair $\{x, y\}$, G contains neither (x, y) nor (y, x); for instance, this will happen if no voter ranks either x or y among her top k candidates.

Let S_2 be the set of candidates x such that at least a 2α fraction of voters rank x among their top k candidates. (We will show in the proof of Lemma 5.2 that S_2 is not empty.) The winner w returned by $\mathcal{M}$ is a candidate in the induced graph $G[S_2]$ with largest outdegree; notice that edges leaving S_2 are *not* counted.

Theorem 5.1 $\mathcal{M}$ *has distortion at most* $\frac{79n}{k}$.

We begin with a lemma showing the key structural property of the winning candidate w.

Lemma 5.2 *In G, for every candidate x, there is a directed path of length at most 3 from w to x.*

Proof. Similar to the definition of S_2, let S_3 be the set of candidates x such that at least a 3α fraction of the voters ranks x somewhere among their top k candidates. By the Pigeon Hole Principle, because each voter ranks a $\frac{k}{n}$ fraction of candidates in her top k, and $\alpha = \frac{k}{3n}$, at least one candidate occurs in a 3α fraction of top-k lists. In particular, S_3 (and thus S_2) is non-empty.

Each candidate $x \in S_3$ has a directed edge to each candidate $y \notin S_2$. This is because x appears in at least a 3α fraction of top-k lists, while y appears in at most a 2α fraction. In particular, at least an α fraction of voters rank x, but not y, in their top-k lists, and thus prefer x to y.

The induced graph $G[S_2]$ contains, for each pair $x, y \in S_2$, at least one of the edges (x, y) or (y, x). This is because of the (at least) 2α fraction of voters with x in their lists, at least an α fraction rank y higher in their lists, or at least an α fraction rank y lower (or not in their lists). Hence, $G[S_2]$ is a supergraph of a tournament graph.

Because w has maximum degree in $G[S_2]$, it also has maximum degree in at least one tournament subgraph of $G[S_2]$. It is well known (see, e.g., (Moulin 1986; Anshelevich, Bhardwaj, and Postl 2015)) that the maximum-degree node in a tournament graph is in the uncovered set, i.e., it has a path of length at most 2 to every other node. Those paths still exist in the supergraphs $G[S_2]$ and G. Thus, w has a directed path of length 2 in G to every candidate $x \in S_2$.

Let $y \in S_3$ be arbitrary. By the preceding two paragraphs, y has a directed edge to each $x \notin S_2$, and w has a directed

path of length at most 2 to y. In summary, w has a directed path of length at most 3 to each candidate x. ∎

Lemma 5.3, upper-bounds the cost ratio of two candidates x, y when x has a directed path of length at most ℓ to y.

Lemma 5.3 *Let w, z be two candidates such that there is a directed path of length at most ℓ edges from w to z in G. Then, $C(w) \leq (1 + \frac{3^\ell - 1}{\alpha}) \cdot C(z)$.*

Remark 5.4 By using Lemma 6 of (Anshelevich, Bhardwaj, and Postl 2015) repeatedly, one can immediately obtain an upper bound of $(1 + O(1/\alpha))^\ell$. However, since we are typically interested in the case $\alpha = o(1)$, the resulting bound is asymptotically weaker than that of Lemma 5.3.

The exponential dependence on ℓ is an artifact of our relatively simple proof. Applying Corollary 5.3 from (Kempe 2020) instead would yield an improved bound of $\frac{\ell}{\alpha} + 1$ or $\frac{\ell + 1}{\alpha} - 1$, depending on whether ℓ is even or odd.

Proof of Theorem 5.1. By Lemma 5.2, w has a path of length at most 3 in G to every candidate x; in particular, to the optimum candidate $x = x^*$. Thus, by Lemma 5.3 with $\ell = 3$, $C(w) \leq (1 + \frac{26}{\alpha}) \cdot C(x^*)$. Substituting $\alpha = \frac{k}{3n}$ and bounding $1 \leq \frac{n}{k}$ now completes the proof. ∎

6 An Optimal Randomized Algorithm

In this section, we give a simple randomized mechanism which achieves an expected distortion of $3 - \frac{2}{n}$. This closes the remaining gap between the distortion bound for the Random Oligarchy mechanism (Fain et al. 2019) and the lower bound of $3 - \frac{2}{n}$ on the distortion of *every* randomized mechanism which only knows each voter's first choice, proved by (Gross, Anshelevich, and Xia 2017). The mechanism $\mathcal{M}$ is as follows:

- With probability $\frac{1}{n-1}$, select a candidate using the *Proportional to Squares* mechanism. That is, for each candidate x, let ν_x be the fraction of voters who rank x first. Select candidate x with probability $\frac{\nu_x^2}{\sum_y \nu_y^2}$.

- With the remaining probability $\frac{n-2}{n-1}$, select a candidate using the *Random Dictatorship* mechanism. That is, choose a voter uniformly at random, and return her first choice. Notice that this mechanism selects candidate x with probability exactly ν_x.

Theorem 6.1 *The expected distortion of $\mathcal{M}$ is at most $3 - \frac{2}{n}$.*

7 Conclusions

As we already discussed in the introduction and Section 4, there is a gap of $\Theta(\log n)$ in the lower bound on distortion we achieve for k-entry social choice rules and more general M-communication bounded social choice rules. If the stronger lower bound holds more generally, a proof will likely require a deeper understanding of the combinatorial structure of partitions of $\mathcal{S}_n$. An intriguing alternative is that there may be a mechanism in which voters communicate only $\Theta(1)$ bits of information per candidate, but which nonetheless achieves constant distortion.

Throughout, we assumed that all voters use the same "encoding" in communicating with the mechanism. For both k-entry social choice rules and M-communication bounded rules, one could consider relaxing this uniformity — although voting mechanisms which treat voters differently *a priori* are typically not widely accepted. For k-entry social choice rules, our lower-bound proof can be directly adapted to give the same lower bound so long as only few voters get to specify which candidate they rank last. The proof does not carry over directly when many, but not all, voters can specify their bottom-ranked candidate. For M-communication bounded rules, it is much less clear how to deal with arbitrarily differing encodings.

A further generalization would be to let voters *choose* which encoding to use, or which subset of positions to fill in. Mechanisms allowing such a choice by the voters would be "non-deterministic," because there is not a unique message any more for each ranking. This raises the issue of how a voter would determine which of many possible messages to send. The specific choice of message may encode additional (e.g., cardinal) information about the voter's ranking. It would require some subtlety to define a model to rule out the revelation of a lot of cardinal information, while still allowing voters non-trivial choices.

Here, we only considered single-round mechanisms. In many settings, including the implementation of social choice rules (Boutilier and Rosenschein 2016; Segal 2010), multiple rounds of communication can lead to significantly (including exponentially) lower overall communication. (Gross, Anshelevich, and Xia 2017; Fain et al. 2019) studied *randomized* multi-round voting mechanisms with the explicit goal of reducing the required communication, while achieving low distortion. In the case of randomized mechanisms, receiving $\log_2 n$ bits of information from each voter is enough to achieve distortion $3 - \frac{2}{n}$, so the room for improvement with multiple rounds is limited. However, for *deterministic* mechanisms, there is potential for significant improvement: a natural question is whether one might even achieve constant distortion with only $O(\log n)$ (or $O(\mathrm{polylog}(n))$) communication from each voter.

References

Anshelevich, E., and Postl, J. 2016. Randomized social choice functions under metric preferences. In *Proc. 25th Intl. Joint Conf. on Artificial Intelligence*, 46–59.

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.; Bhardwaj, O.; and Postl, J. 2015. Approximating optimal social choice under metric preferences. In *Proc. 29th AAAI Conf. on Artificial Intelligence*, 777–783.

Anshelevich, E. 2016. Ordinal approximation in matching and social choice. *ACM SIGecom Exchanges* 15(1):60–64.

Bentert, M., and Skowron, P. 2019. Comparing election methods where each voter ranks only few candidates. Preprint available on arXiv 1901.10848.

Black, D. 1948. On the rationale of group decision making. *J. Political Economy* 56:23–34.

Boutilier, C., and Rosenschein, J. S. 2016. Incomplete information and communication in voting. In Brandt, F.; Conitzer, V.; Endriss, U.; Lang, J.; and Procaccia, A. D., eds., *Handbook of Computational Social Choice*. Cambridge University Press. chapter 10, 223–257.

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.

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

Conitzer, V., and Sandholm, T. 2002. Vote elicitation: Complexity and strategy-proofness. In *Proc. 17th AAAI Conf. on Artificial Intelligence*, 392–397.

Fain, B.; Goel, A.; Munagala, K.; and Prabhu, N. 2019. Random dictators with a random referee: Constant sample complexity mechanisms for social choice. In *Proc. 33rd AAAI Conf. on Artificial Intelligence*, 1893–1900.

Gross, S.; Anshelevich, E.; and Xia, L. 2017. Vote until two of you agree: Mechanisms with small distortion and sample complexity. In *Proc. 31st AAAI Conf. on Artificial Intelligence*, 544–550.

Kempe, D. 2019. Communication, distortion, and randomness in metric voting. arXiv: 1911.08129.

Kempe, D. 2020. An analysis framework for metric voting based on LP duality. In *Proc. 34th AAAI Conf. on Artificial Intelligence*.

Mandal, D.; Procaccia, A. D.; Shah, N.; and Woodruff, D. P. 2019. Efficient and thrifty voting by any means necessary. In *Proc. 33rd Advances in Neural Information Processing Systems*.

Moulin, H. 1980. On strategy-proofness and single peakedness. *Public Choice* 35:437–455.

Moulin, H. 1986. Choosing from a tournament. *Social Choice and Welfare* 3(4):271–291.

Munagala, K., and Wang, K. 2019. Improved metric distortion for deterministic social choice rules. In *Proc. 20th ACM Conf. on Economics and Computation*, 245–262.

Pierczyński, G., and Skowron, P. 2019. Approval-based elections and distortion of voting rules. Preprint available on arXiv 1901.06709.

Procaccia, A. D., and Rosenschein, J. S. 2006. The distortion of cardinal preferences in voting. In *Proc. 10th Intl. Workshop on Cooperative Inform. Agents X*, 317–331.

Procaccia, A. D. 2010. Can approximation circumvent Gibbard-Satterthwaite? In *Proc. 24th AAAI Conf. on Artificial Intelligence*, 836–841.

Segal, I. R. 2010. Nash implementation with little communication. *Theoretical Economics* 5:51–71.