Dimensionality and Coordination in Voting: The Distortion of STV

Ioannis Anagnostides,[1] Dimitris Fotakis,[2] Panagiotis Patsilinakos[2]

[1] Carnegie Mellon University
[2] National Technical University of Athens
ianagnos@cs.cmu.edu, fotakis@cs.ntua.gr, patsilinak@corelab.ntua.gr

Abstract

We study the performance of voting mechanisms from a utilitarian standpoint, under the recently introduced framework of metric-distortion, offering new insights along two main lines. First, if d represents the doubling dimension of the metric space, we show that the distortion of STV is $O(d \log \log m)$, where m represents the number of candidates. For doubling metrics this implies an exponential improvement over the lower bound for general metrics, and as a special case it effectively answers a question left open by Skowron and Elkind (AAAI '17) regarding the distortion of STV under low-dimensional Euclidean spaces. More broadly, this constitutes the first nexus between the performance of any voting rule and the "intrinsic dimensionality" of the underlying metric space. We also establish a nearly-matching lower bound, refining the construction of Skowron and Elkind. Moreover, motivated by the efficiency of STV, we investigate whether natural learning rules can lead to low-distortion outcomes. Specifically, we introduce simple, deterministic and decentralized exploration/exploitation dynamics, and we show that they converge to a candidate with $O(1)$ distortion.

1 Introduction

Aggregating the preferences of individual entities into a collective decision lies at the foundations of voting theory, and has recently found a myriad of applications in areas such as information retrieval, recommender systems, and machine learning (Lu and Boutilier 2014). A common hypothesis in the literature of social choice asserts that agents only provide an *order* of preferences over a (finite) set of alternatives, without indicating a precise *measure* of each preference. However, this assertion might seem misaligned with many classical models in economic theory (von Neumann and Morgenstern 1944) which espouse a *utilitarian* framework to represent agents' preferences. This raises the following concern: What is the loss in utilitarian efficiency of a mechanism eliciting only ordinal information?

This question was raised by Procaccia and Rosenschein (2006), introducing the concept of *distortion*, and has since led to a substantial body of work. In this paper we focus on the refined notion of *metric distortion* (Anshelevich, Bhardwaj, and Postl 2015), wherein agents and candidates are associated with points in some *metric space*, and preferences are being determined based on the proximity in the underlying metric (see Section 2 for a formal definition). Importantly, this framework offers a quantitative "benchmark" for comparing different voting rules commonly employed in practice. Indeed, one of the primary considerations of our work lies in characterizing the performance of the *single transferable vote*[1] mechanism (henceforth STV).

STV is a widely-popular iterative voting system employed in the national elections of several countries, including Australia, Ireland, and India, as well as in many other preference aggregation tasks; e.g., in the Academy Awards. To be more precise, STV proceeds in an iterative fashion: In each round, agents vote for their most preferred candidate—among the *active* ones, while the candidate who enjoyed the least amount of support in the current round gets eliminated. This process is repeated for $m - 1$ rounds, where m represents the number of (initial) alternatives, and the last surviving candidate is declared the winner of STV. As an aside, notice that this process is generally non-deterministic due to the need for a tie-breaking mechanism; as in (Skowron and Elkind 2017), we will work with the *parallel universe model* of Conitzer, Rognlie, and Xia (2009), wherein a candidate is said to be an STV winner if it survives under *some* sequence of eliminations.

In this context, Skowron and Elkind (2017) were the first to analyze the distortion of STV under metric preferences. Specifically, they showed that the distortion of STV in general metric spaces is always $O(\log m)$, while they also gave a nearly-matching lower bound in the form of $\Omega(\sqrt{\log m})$. Interestingly, a careful examination of their lower bound reveals the existence of a high-dimensional submetric, as depicted in Figure 1, and it is a well-known fact in the theory of metric embeddings that such objects cannot be *isometrically* embedded into low-dimensional[2] Euclidean spaces (Matoušek 2002). As a result, Skowron and Elkind (2017) left open the following intriguing question:

Question 1. *What is the distortion of* STV *under low-dimensional Euclidean spaces?*

[1]For consistency with prior work STV will represent throughout this paper the *single-winner* variant of the system, which is sometimes referred to as *instant-runoff voting* (IRV).

[2]We say that a Euclidean space is low-dimensional if its dimension d is bounded by a "small" universal constant, i.e. $d = O(1)$.

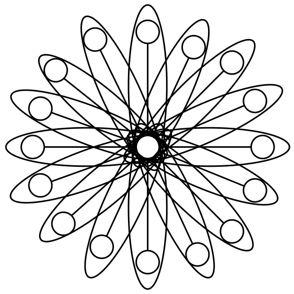


Figure 1: A high-dimensional metric in the form of a "star".

Mechanism	Lower Bound	Dimension
Plurality	$2m-1$	1
Borda	$2m-1$	1
Copeland	5	2
Veto	$2n-1$	1
Approval	$2n-1$	1

Table 1: The Euclidean dimension required to construct a (tight) lower bound for several common voting rules; these results appear in (Anshelevich, Bhardwaj, and Postl 2015). We should note that for Copeland the metric constructed in (Anshelevich, Bhardwaj, and Postl 2015) is not Euclidean, but can be easily modified to be one.

Needless to say that the performance of voting rules in low-dimensional spaces has been a subject of intense scrutiny in spatial voting theory, under the premise that voters and candidates are typically embedded in subspaces with small dimension (Arrow 1990; Enelow and Hinich 1984). For example, recent experimental work by Elkind et al. (2017) evaluates several voting rules in a 2-dimensional Euclidean space, motivated by the fact that preferences are typically crystallized on the basis of a few crucial dimensions; e.g., economic policy and healthcare. Indeed, in the so-called Nolan Chart—a celebrated political spectrum diagram—political views are charted along two axes, expanding upon the traditional one-dimensional representation; to quote from the work of Elkind et al. (2017):

> *"...the popularity of the Nolan Chart [...] indicates that two dimensions are often sufficient to provide a good approximation of voters' preferences."*

Thus, it is natural to ask whether we can refine the analysis of STV under low-dimensional spaces. In fact, as part of a broader agenda analogous questions can be raised for other mechanisms as well. However, it is interesting to point out that for many voting rules analyzed within the framework of distortion there exist low-dimensional lower bounds; some notable examples are given in Table 1. In contrast, our work will separate STV from the mechanisms in Table 1, effectively addressing Question 1. Importantly, we shall provide a characterization well-beyond Euclidean spaces, to metrics with "intrinsically" low dimension.

The next consideration of our work is directly motivated by the efficiency of STV compared to the plurality rule, and in particular the strategic implications of this discrepancy. A good starting point for this discussion stems from the fact that in many fundamental preference aggregation settings alternatives are chosen by inefficient mechanisms, and in many cases any reform faces insurmountable impediments. For example, in political elections the voting mechanism is typically dictated by electoral laws, or even the constitution (Lijphart 1992). As a result, understanding the behavior of strategic agents when faced with inefficient mechanisms is of paramount importance (Brill and Conitzer 2015; Zuckerman et al. 2011). A rather orthogonal way of viewing this is whether autonomous agents can converge to admissible social choices through natural learning rules; this begs the question:

Question 2. *To what extent can strategic behavior improve efficiency in voting?*

We stress that although in the absence of any information it might be unclear how agents can engage in strategic behavior, in most applications of interest agents have plenty of prior information before they cast their votes, e.g. through polls, surveys, forecasts, prior elections, or even early voting. Indeed, there is a prolific line of work which studies population dynamics for agents that cast their votes in response to the information they possess (Restrepo, Rael, and Hyman 2009), as well as the role of information in shaping public policy (Larcinese 2003).

To address such considerations we propose a natural model wherein agents act iteratively based on some partial feedback on the other voters' preferences. We explain how STV can be very naturally cast in this framework, while we establish the existence of simple and decentralized *coordination* dynamics converging to a near-optimal alternative.

Overview of Results

Our first contribution is to relate the distortion of STV to the dimensionality of the underlying metric space. Specifically, our first insight is to employ the following fundamental concept from metric geometry:

Definition 1.1 (Doubling Dimension)**.** The *doubling constant* of a metric space $(\mathcal{M}, \mathrm{dist})$ is the least integer $\lambda \geq 1$ such that for all $x \in \mathcal{M}$ and for all $r > 0$, every ball $\mathcal{B}(x, 2r)$ can be covered by the union of at most λ balls of the form $\mathcal{B}(s, r)$, where $s \in \mathcal{M}$; that is, there exists a subset $\mathcal{S} \subseteq \mathcal{M}$ with $|\mathcal{S}| \leq \lambda$ such that

$$\mathcal{B}(x, 2r) \subseteq \bigcup_{s \in \mathcal{S}} \mathcal{B}(s, r). \tag{1}$$

The *doubling dimension* is defined as $\dim(\mathcal{M}) := \log_2 \lambda$.

This concept generalizes the standard notion of dimension since $\dim(\mathbb{R}^d) = \Theta(d)$ when $\mathbb{R}^d$ is endowed with the ℓ_p norm. Moreover, it is clear that for a finite metric space $(\mathcal{M}, \mathrm{dist}), \dim(\mathcal{M}) \leq \log_2 |\mathcal{M}|$; for example, this is essentially tight for the high-dimensional metric of Figure 1.

The concept of doubling dimension was introduced by Larman (1967) and Assouad (1983), and was first used in algorithm design by Clarkson (1997) in the context of the nearest neighbors problem. Nevertheless, we are not aware of any prior characterization that leverages the doubling dimension in the realm of voting theory. In the sequel, it will be assumed that $(\mathcal{M}, \mathrm{dist}(\cdot,\cdot))$ stands for the metric space induced by the set of candidates and voters. In this context, our first main contribution is the following theorem:

Theorem 1.2. *If d is the doubling dimension of $\mathcal{M}$, then the distortion of* STV *is $O(d \log\log m)$.*

For *doubling metrics*[3] this theorem already implies an exponential improvement in the distortion over the $\Omega(\sqrt{\log m})$ lower bound for general metrics. Moreover, it addresses as a special case Question 1:

Corollary 1.3. *The distortion of* STV *under low-dimensional Euclidean spaces is $O(\log\log m)$.*

To the best of our knowledge, this is the first result that relates the performance of any voting rule to the "intrinsic dimensionality" of the underlying metric space. It also corroborates the experimental findings of Elkind et al. (2017) regarding the superiority of STV on the 2-dimensional Euclidean plane. More broadly, we suspect that our characterization applies for a wide range of iterative voting rules, to which STV serves as a canonical example. We should note that the $O(\log\log m)$ factor appears to be an artifact of our analysis. Indeed, we put forward the following conjecture:

Conjecture 1.4. *If d is the doubling dimension of $\mathcal{M}$, then the distortion of* STV *is $O(d)$.*

Verifying this conjecture in light of our result might be of small practical importance, but nonetheless we believe that it can be established by extending our techniques. In fact, for one-dimensional spaces we actually confirm this conjecture, proving that the distortion of STV on the line is $O(1)$ in Theorem 3.1. It should be noted, however, that the underlying phenomenon is inherently different once we turn our attention to higher-dimensional spaces. In addition, to complement our positive results we refine the lower bound of Skowron and Elkind (2017), showing an $\Omega(\sqrt{d})$ lower bound, where d represents the doubling dimension of the submetric induced by the set of candidates $\mathcal{M}_C$. Thus, it should be noted that there are still small gaps left to be bridged in future research.

Other Notions of Dimension. An important advantage of the doubling dimension is that it subsumes other commonly-used notions of dimension. Most notably, Karger and Ruhl (2002) have introduced a concept of dimension based on the *growth rate* of a (finite) metric space, and it is known (Gupta, Krauthgamer, and Lee 2003, Proposition 1.2) that the doubling dimension can only be a factor of 4 larger than the growth rate of Karger and Ruhl. Moreover, a similar statement applies for the *local density* of an unweighted graph, another natural notion of volume that has been employed in the analysis of a graph's bandwidth (Feige 2000).

[3] A doubling metric refers to a metric space with doubling dimension upper-bounded by some universal constant.

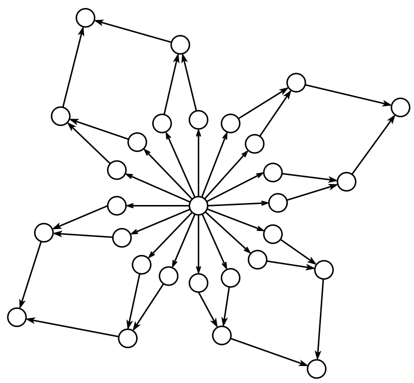
High-Level Intuition. In this paragraph we briefly attempt to explain why the distortion of STV depends on the "covering dimension" of the underlying metric space. First, we have to describe the technique developed by Skowron and Elkind (2017). Specifically, their method for deriving an upper bound for the distortion of an iterative voting rule consists of letting a substantial fraction of agents reside within close proximity to the optimal candidate, and then analyze how the support of these agents propagates throughout the evolution of the iterative process. More precisely, the overall distance covered immediately implies an upper bound on the distortion (see Lemma 2.2). The important observation is that the underlying dimension drastically affects this phenomenon. In particular, when a large fraction of agents lies in a low-dimensional ball supporting many different candidates, we can infer that their (currently) second most-preferred alternatives ought to be "close"—for most of the agents—by a covering argument (and the triangle inequality). This directly circumscribes the propagation of the support, as hinted in Figure 2b, juxtaposed to the phenomenon in high dimensions in Figure 2a. We stress that we shall make use of this basic skeleton developed by Skowron and Elkind (2017). We also remark that we recover their $O(\log m)$ bound under general metrics through a simpler analysis, which incidentally reveals a very clean recursive structure; this argument will be directly invoked for the proof of our main theorem.

The next theme of our work is motivated by the performance of STV, and in particular offers a preliminary answer to Question 2. Specifically, to formally address such questions we first propose a natural iterative model: In each day every agent has to select a *single* candidate, and at the end of the round agents are informed about the (plurality) scores of all the candidates (cf., see (Borodin et al. 2019)). This process is repeated for sufficiently many days, and it is assumed that the candidate who enjoyed the largest amount of support in the ultimate day will eventually prevail. Observe that in this scenario truthful engagement appears to be very unrealistic since agents would endeavor to adapt their support based on the popularity of each candidate; for example, it would make little sense to squander one's vote (at least towards the last stages) to an unpopular candidate. More broadly, there is an interesting nexus between distortion and stability, as we elaborate in Section 4, emphasizing on a connection with the notion of *core* in cooperative game theory (Proposition 4.2).

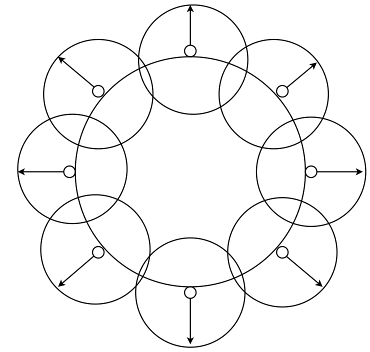
In this context, STV already suggests a particularly natural strategic engagement, improving exponentially over the outcome of the truthful dynamics. Yet, it yields superconstant distortion due to the greedy aspect of the induced dynamics. We address this issue by designing a simple and decentralized exploration/exploitation scheme:

Theorem 1.5. *There exist simple, deterministic and distributed dynamics that converge to a candidate with $O(1)$ distortion.*

We elaborate on the proposed dynamics, as well as on all the aforementioned issues in Section 4.



(a) The propagation of the support in high dimensions.



(b) The propagation of the support in low dimensions.

Figure 2: The impact of the underlying dimension on STV.

Related Work

The framework of distortion under *metric preferences* was introduced by Anshelevich, Bhardwaj, and Postl (2015) (see also (Anshelevich et al. 2018)). Specifically, they observed a lower bound of 3 for any deterministic mechanism, while they also showed—among others—that Copeland's method, a very popular voting system, always incurs distortion at most 5, with the bound being tight for certain instances. This threshold was subsequently improved by Munagala and Wang (2019), introducing a novel (deterministic) mechanism with distortion $2 + \sqrt{5}$, while the same bound was independently obtained by Kempe (2020) through an approach based on LP duality. The lower bound of 3 was only recently matched by PLURALITYMATCHING, a mechanism introduced by Gkatzelis, Halpern, and Shah (2020). All of the aforementioned results apply under arbitrary metric spaces. Several special cases have also attracted attention in the literature. For one-dimensional spaces, Feldman, Fiat, and Golomb (2016) establish several improved bounds, while a comprehensive characterization in a *distributed* setting was recently given by Filos-Ratsikas and Voudouris (2021). The interested reader is referred to the concise survey of Anshelevich et al. (2021) for detailed accounts on the rapidly growing literature on the subject. Moreover, for related research beyond the framework of distortion we refer to (Gershkov, Moldovanu, and Shi 2019), and references therein.

Our considerations in Section 4 are related to the seminal work of Brânzei et al. (2013) (see also the extensive follow-up work, such as (Obraztsova et al. 2015)), viewing voting from the standpoint of *price of anarchy* (PoA). In particular, the authors study the discrepancy between the plurality scores under truthfulness, and under worst-case limit points of *best-response* dynamics. Instead, we argue that the utilitarian performance of a voting rule—in terms of distortion—offers a very compelling alternative to study this discrepancy, similarly to the original formulation of PoA in the context of routing games (Koutsoupias and Papadimitriou 1999), while going beyond best-response dynamics is very much in line with the modern approach in the context of learning in games (Cesa-Bianchi and Lugosi 2006). Finally, we stress that Question 2 has already received extensive attention in the literature (cf., see (Brill and Conitzer 2015; Zuckerman et al. 2011) and references therein), but it was not addressed within the framework of (metric) distortion.

2 Preliminaries

A *metric space* is a pair $(\mathcal{M}, \mathrm{dist}(\cdot, \cdot))$, where $\mathrm{dist} : \mathcal{M} \times \mathcal{M} \mapsto \mathbb{R}$ is a *metric* on $\mathcal{M}$, i.e., (i) $\forall x, y \in \mathcal{M}, \mathrm{dist}(x, y) = 0 \iff x = y$ (identity of indiscernibles), (ii) $\forall x, y \in \mathcal{M}, \mathrm{dist}(x, y) = \mathrm{dist}(y, x)$ (symmetry), and (iii) $\forall x, y, z \in \mathcal{M}, \mathrm{dist}(x, y) \leq \mathrm{dist}(x, z) + \mathrm{dist}(z, y)$ (triangle inequality). Now consider a set of n voters $V = \{1, 2, \dots, n\}$, and a set of m candidates C; we will reference candidates with lowercase letters such as a, b, w, x. Voters and candidates are associated with points in a finite metric space $(\mathcal{M}, \mathrm{dist})$, while it is assumed that $\mathcal{M}$ is the (finite) set induced by the set of voters and candidates. The goal is to select a candidate x who minimizes the *social cost*: $\mathrm{SC}(x) = \sum_{i=1}^{n} \mathrm{dist}(i, x)$. This task would be trivial if we had access to the agents' distances from all the candidates. However, in the *metric distortion* framework every agent i provides only a *ranking* (a total order) σ_i over the points in C according to the *order* of i's distances from the candidates, with ties broken arbitrarily. We also define $\sigma := (\sigma_1, \dots, \sigma_n)$, while we will sometimes use $\mathsf{top}(i)$ to represent i's most preferred candidate.

A *deterministic social choice rule* is a function that maps an *election* in the form of a 3-tuple $\mathcal{E} = (V, C, \sigma)$ to a single candidate $a \in C$. We will measure the performance of f for a given input of preferences σ in terms of its *distortion*; namely, the worst-case approximation ratio it provides with respect to the social cost:

$$\mathsf{distortion}(f; \sigma) = \sup \frac{\mathrm{SC}(f(\sigma))}{\min_{a \in C} \mathrm{SC}(a)}, \tag{2}$$

where the supremum is taken over all metrics consistent with the voting profile. The distortion of a social choice rule f is the maximum of $\mathsf{distortion}(f; \sigma)$ over all possible input preferences σ.

We define the *open ball* on the metric space $(\mathcal{M}, \mathrm{dist})$ with center $x \in \mathcal{M}$ and radius $r > 0$ as $\mathcal{B}(x, r) := \{z \in \mathcal{M} : \mathrm{dist}(z, x) < r\}$. The following covering lemma will be useful for the analysis of STV in doubling metrics.

Lemma 2.1. *Consider a metric space $(\mathcal{M}, \mathrm{dist})$ with doubling constant $\lambda \geq 1$. Then, for any $x \in \mathcal{M}$ and $r > 0$, the ball $\mathcal{B}(x, r)$ can be covered by at most $\lambda^{\lceil \log(r/\epsilon) \rceil}$ balls of radius at most ϵ.*

When unspecified, the $\log(\cdot)$ will always be implied to the base 2. We conclude this section with a useful lemma observed by Skowron and Elkind (2017):

Lemma 2.2 ((Skowron and Elkind 2017))**.** *Consider two distinct candidates $a, b \in C$. If $r := \mathrm{dist}(a, b)/h$ for some parameter $h > 0$, and at most γn agents reside in $\mathcal{B}(a, r)$ for some $\gamma \in [0, 1)$, then*

$$\frac{\mathrm{SC}(b)}{\mathrm{SC}(a)} \leq 1 + \frac{h}{1 - \gamma}. \tag{3}$$

3 STV in Doubling Metrics

In this section we refine the analysis of STV based on the intrinsic dimensionality of the underlying metric space; our main result is Theorem 3.2.

STV on the Line

As a warm-up, we analyze the performance of STV on the line. In particular, we establish the following result:

Theorem 3.1. *The distortion of* STV *on the line is at most* 15.

The proof is deferred to the full version of our paper. Before we proceed to our main result, a few remarks are in order. First of all, we did not pursue optimizing the constant in the theorem, although this might be an interesting avenue for future research. It should also be noted that Theorem 3.1 already implies a stark separation between STV and PLURALITY, as the latter is known to admit a one-dimensional $\Omega(m)$ lower bound (recall Table 1).

Main Result

Moving on to the main result of this section, we will prove the following theorem:

Theorem 3.2. *If d is the doubling dimension of $\mathcal{M}$, then the distortion of* STV *is $O(d \log \log m)$.*

The first ingredient of the proof is a recursive pattern regarding the propagation of the support during STV, incidentally leading to a much simpler analysis in general metrics. In this context, the main technical challenge under doubling metrics lies in maintaining the appropriate invariance. We address this by essentially identifying a subset of the domain with a sufficient degree of regularity. We should also note that the second part of the proof makes use of the technique devised by Skowron and Elkind (2017).

Proof of Theorem 3.2. Let $w \in C$ be the winner of STV under some sequence of eliminations, and $x \in C$ be the candidate who minimizes the social cost. Moreover, let $r := \mathrm{dist}(x, w)/(4h + 7)$, where h is defined as $h := 1 + \lceil \log_2(6\lambda^{\log \mathcal{H}_m + 1}) \rceil = \Theta(d \log \log m)$. If γ represents the fraction of the voters in $\mathcal{B}(x, r)$, we will establish that $\gamma \leq 2/3$.

For the sake of contradiction, let us assume that $\gamma > 2/3$. Our argument will characterize the propagation of the support of the voters in $\mathcal{B}(x, r)$. In particular, we proceed in the following two phases:

Phase I. Our high-level strategy is to essentially employ our argument for general metrics, but not for the entire set of voters in $\mathcal{B}(x, r)$. Instead, we will establish the existence of a set with a helpful invariance, which still contains most of the voters. More precisely, we first consider a covering $\{\mathcal{B}(z_j, r_j)\}_{j=1}^{\mu}$ of the ball $\mathcal{B}(x, r)$, where the radius of every ball is at most $\epsilon \times r$ for some parameter $\epsilon \in (0, 1)$. The balls that do not contain any voter may be discarded for the following argument. We let $\mathcal{S}^{(0)}$ be the union of these balls. We know from Lemma 2.1 that $\mu = \mu(\epsilon; \lambda) \leq \lambda^{\log(1/\epsilon)+1}$. For Phase I we assume that more than M candidates remain active in STV, where $M := 6\mu$, while $\epsilon := 1/\mathcal{H}_m$ ($\mathcal{H}_m$ denotes the m-th harmonic number).

Let us consider a round $t = 1, \ldots, m-M$ of STV. In particular, let $a \in C$ be the candidate who is eliminated at round t. Observe that if a is not supported by any voter residing in $\mathcal{B}(x, r)$, the support of these agents remains invariant under round t. Thus, let us focus on the contrary case. Specifically, if there exists a ball in the covering which contains exclusively supporters of candidate a, we shall remove every such ball from the current covering, updating analogously the set $\mathcal{S}^{(t)}$. Given that we are at round t, we can infer that the number of such supporters is at most $n/(m - t + 1) < n/M$. Thus, since we can only remove μ balls from the initial covering, it follows that the set $\mathcal{S} := \mathcal{S}^{(t)}$ with $t = m - M$ contains strictly more than $2n/3 - n\mu/M = n/2$ voters.

Next, we will argue about the propagation of the support for the voters in $\mathcal{S}$ during the first $m-M$ rounds of STV. By construction of the set $\mathcal{S}$, we have guaranteed the following invariance: Whenever a candidate a supported by voters in $\mathcal{S}$ gets eliminated, every supporter of a from $\mathcal{S}$ lies within a ball of radius at most ϵ with agents championing a different candidate. Now, let us define $\overline{D}^{(t)}$ as follows:

$$\overline{D}^{(t)} := \frac{1}{\gamma' n} \sum_{i \in \mathcal{S}} \mathrm{dist}(i, \mathsf{top}(i; t)), \tag{4}$$

where γ' represents the fraction of the voters in $\mathcal{S}$ and $\mathsf{top}(i; t)$ is i's top active candidate at round t. Consider two voters i, j supporting two candidates a, b respectively. We will show that $\mathrm{dist}(i, b) \leq \mathrm{dist}(i, a) + 2\,\mathrm{dist}(i, j)$, and similarly, $\mathrm{dist}(j, a) \leq \mathrm{dist}(j, b) + 2\,\mathrm{dist}(i, j)$. Indeed, successive applications of the triangle inequality imply that

$$\begin{aligned}\mathrm{dist}(i, b) &\leq \mathrm{dist}(i, j) + \mathrm{dist}(j, b)\\ &\leq \mathrm{dist}(i, j) + \mathrm{dist}(j, a)\\ &\leq \mathrm{dist}(i, j) + \mathrm{dist}(j, i) + \mathrm{dist}(i, a)\\ &= \mathrm{dist}(i, a) + 2\,\mathrm{dist}(i, j).\end{aligned}$$

Thus, if the voters i and j happen to reside within a ball of radius at most ϵ, we can infer that $\mathrm{dist}(i,b) \leq \mathrm{dist}(i,a)+4\epsilon$. As a result, we can inductively conclude that

$$\overline{D}^{(t)} \leq \overline{D}^{(t-1)} + \frac{1}{\gamma'}\frac{4\epsilon r}{m-t+1} \leq \overline{D}^{(t-1)} + \frac{8\epsilon r}{m-t+1},$$

in turn implying that

$$\overline{D}^{(m-M)} \leq 8(\epsilon r)\mathcal{H}_m. \tag{5}$$

In particular, for $\epsilon = 1/\mathcal{H}_m$ this implies that during the first phase the agents in $\mathcal{S}$ support candidates within $O(1) \times r$ distance from x.

Phase II. At the beginning of the second phase there are M remaining candidates. Let us denote with $\mathcal{B}_j := \mathcal{B}(x, (2j-1) \times r)$. In this phase we will argue about the entire set of voters in $\mathcal{B}(x,r)$. Let $m_1 \leq M$ be the number of candidates supported by voters in $\mathcal{B}(x,r)$ at the start of the second phase. Our previous argument for Phase I implies that every such candidate will reside in $\mathcal{B}_7$; this follows directly by applying the triangle inequality. Let us denote with m_j the number of candidates residing outside $\mathcal{B}_{4+2j}$ for $j \geq 2$ at the round the last candidate from $\mathcal{B}_{3+2j}$ gets eliminated.

By the pigeonhole principle, we can infer that there exists a candidate a in $\mathcal{B}_7$ who enjoys the support of at least $\gamma n/m_1$ voters. Moreover, observe that the triangle inequality implies that no voter will support a candidate outside $\mathcal{B}_8$ as long as candidate a remains active. Thus, at the round a gets eliminated we can deduce that $(1-\gamma)n/m_2 \geq \gamma n/m_1 \iff m_2 \leq m_1 \times (1-\gamma)/\gamma$, where we used that the number of candidates in every subset can only decrease during STV. Inductively, we can infer that

$$m_h \leq \left(\frac{1-\gamma}{\gamma}\right)^{h-1} m_1 < \left(\frac{1}{2}\right)^{h-1} M \leq 1, \tag{6}$$

for $h = \lceil \log_2 M \rceil + 1$, where we used that $\gamma > 2/3$. This implies that the winner of STV should lie within B_{4+2h}, i.e. $\mathrm{dist}(x,w)/r < 4h+7$, which is a contradiction since $\mathrm{dist}(x,w) = (4h+7) \times r$. Thus, the theorem follows directly from Lemma 2.2. □

The Lower Bound

We also refine the $\Omega(\sqrt{\log m})$ lower bound (Skowron and Elkind 2017, Theorem 4) based on the doubling dimension of the submetric induced by the set of candidates $\mathcal{M}_C$. In particular, we establish the following theorem:

Theorem 3.3 (Lower Bound for STV). *For any $\lambda \geq 2$ there exists a metric space induced by the set of candidates $(\mathcal{M}_C, \mathrm{dist})$, with $d = \Theta(\log \lambda)$ being the doubling dimension of $\mathcal{M}_C$, and a voting profile such that the distortion of* STV *is $\Omega(\sqrt{d})$.*

The construction follows similarly to that of Skowron and Elkind (2017), and it is included in the full version.

4 Coordination Dynamics

In this section we explore whether natural and distributed learning dynamics can converge to social choices with near-optimal distortion. We should point out that there is a concrete connection between such considerations and the results of the previous section, as it will be revealed in detail shortly. First, let us commence with the following observation:

Observation 4.1. *Consider a voting instance under a metric space so that some candidate $a \in C$ has distortion at least $\mathfrak{D}$.*[4] *Then, there exists a candidate $x \neq a$ and subset $W \subseteq V$ such that*

1. *Every agent in W strictly prefers x to a;*
2. *$|W|/n \geq 1 - 2/(\mathfrak{D}+1)$.*

This statement essentially tells us that candidates with large distortion are inherently unstable, in the sense that there will exist a large "coalition" of voters that strictly prefer a different outcome. Interestingly, this observation implies a connection between (metric) distortion and the notion of *core* in cooperative game theory. To be more precise, we will say that a set of coalitions $\mathcal{W}$ is *α-large*, with $\alpha \in [0,1]$, if it contains every coalition $W \subseteq V$ such that $|W|/n \geq \alpha$; a candidate a is said to be in the core if there does *not* exist a coalition $W \in \mathcal{W}$ such that every agent in W (strictly) prefers a different alternative.[5] In this context, the following proposition follows directly from Observation 4.1:

Proposition 4.2. *Consider a voting instance under a metric space so that some candidate $a \in C$ has distortion at least $\mathfrak{D}$. Then, candidate a cannot be in the core with respect to an α-large set of coalitions, as long as $\alpha \leq 1 - 2/(\mathfrak{D}+1)$.*

As a result, it is interesting to study the strategic behavior and the potential coordination dynamics that may arise in the face of an inefficient voting system.

The Model

We consider the following abstract model: For some given voting system, agents are called upon to cast their votes for a series of T days or rounds, where T is sufficiently large. After the end of each day, voters are informed about the results of the round, and the winner is determined based on the results of the *ultimate* day. This is essentially an iterative implementation of a given voting rule, in place of the *one-shot* execution typically considered, and it is introduced to take into account external information typically accumulated before the actual voting (e.g. through polls). For concreteness, we will assume that the voting rule employed in each day is simply the PLURALITY mechanism, not least due to its popularity both in theory and in practice.

Before we describe and analyze natural dynamics in this model, let us first note that if all the voters engage truthfully throughout this game, the victor will coincide with the plurality winner, and as we know there are instances for which this candidate may have $\Omega(m)$ distortion. As a result, Observation 4.1 implies that there will be a large coalition with

[4] That is, $\mathrm{SC}(a)/\min_{x \in C} \mathrm{SC}(x) \geq \mathfrak{D}$.

[5] Considering only "large" coalitions is standard in the literature; cf., (Brandt et al. 2016).

a $1-\Theta(1/m)$ fraction of the voters that strictly prefer a different outcome. Indeed, the lower bound of PLURALITY is built upon $m-1$ clusters of voters formed arbitrarily close, while a different extreme party with roughly the same plurality score could eventually prevail. However, the access to additional information renders this scenario rather unrealistic given that we expect some type of adaptation or coordination mechanism from the agents.

A Greedy Approach

Let us denote with $n_a^{(t)}$ the plurality score of candidate a at round $t \in [T]$. A particularly natural approach for an agent to engage in this scenario consists of maintaining a time-varying parameter $\theta^{(t)}$, which will essentially serve as the "temperature". Then, at some round $t > 1$ agent i will support the candidate b for which $b \succeq_i a$ for all $a, b \in C^{(t)}$, where $C^{(t)} := \{a \in C : n_a^{(t-1)} \geq \theta^{(t)}\}$.[6] That is, agents only consider candidates who exceeded some level of support during the previous day. Then, the temperature parameter is updated accordingly, for example with some small constant increment $\theta^{(t+1)} := \theta^{(t)} + \epsilon$, for some $\epsilon > 0$. In this context, observe that for a sufficiently small ϵ these dynamics will converge to an STV winner (based on the parallel universe model). This implies that the greedy tactics already offer an exponential improvement—in terms of the utilitarian efficiency—compared to the truthful dynamics. Nevertheless, the lower bound for STV (Theorem 3.3) suggests that we have to design a more careful adaptive rule in order to attain $O(1)$ distortion.

Exploration/Exploitation

The inefficiency of the previous approach—and subsequently of STV—stems from the greedy nature of the iterative process: Agents may choose to dismiss candidates prematurely. For example, this becomes apparent by inspecting the elimination pattern in the lower bound of Theorem 3.3. In light of this, the remedy we propose—and what arguably occurs in many practical scenarios—is an exploration phase. In particular, voters initially do not possess any information about the preferences of the rest of the population. Thus, they may attempt to explore several alternatives in order to evaluate the viability of each candidate; while doing so, agents will endeavor to somehow indicate or favor their own preferences. After the exploration phase, agents will leverage the information they have learnt to adapt their support. More concretely, we consider the following dynamics:

1. Exploration phase: In each round $t \in [m]$ every agent i maintains a list $\mathcal{L}_i^{(t)}$, initialized as $\mathcal{L}_i^{(1)} := \emptyset$. If $C_i^{(t)} := C \setminus \mathcal{L}_i^{(t)}$, then at round t agent i shall vote for the candidate $a \in C_i^{(t)}$ such that $a \succeq_i b$ for all $b \in C_i^{(t)}$. Next, agent i updates her list accordingly: $\mathcal{L}_i^{(t+1)} := \mathcal{L}_i^{(t)} \cup \{a\}$;
2. Exploitation phase: Every agent supports the first candidate[7] within her list that managed to accumulate—over all prior rounds—at least $n/2$ votes.

So, the winner is determined in the first round t for which there is some candidate lying in the top t positions of at least half of the voters' rankings. We shall refer to this iterative process as COORDINATION dynamics.

Theorem 4.3. COORDINATION *dynamics lead to a candidate with distortion at most* 11.

Proof. Let w be the winner under COORDINATION, and x be the candidate who minimizes the social cost. For $r := \mathrm{dist}(x, w)/5$, we consider the sequence of balls $\{\mathcal{B}_i\}_{i=1}^3$ such that $\mathcal{B}_i := \mathcal{B}(x, (2i-1)r)$ for $i = 1, 2, 3$. If γ is the fraction of the voters in $\mathcal{B}_1$, we will argue that $\gamma \leq 1/2$.

For the sake of contradiction, let us assume that $\gamma > 1/2$. Let t be the first round for which a voter in $\mathcal{B}_1$ supports a candidate outside $\mathcal{B}_3$. Then, it follows by the triangle inequality that the list of this voter just after round $t-1$ included all the candidates in $\mathcal{B}_2$. This in turn implies that by round $t-1$ every agent in $\mathcal{B}_1$ had already voted for all candidates in $\mathcal{B}_1$. Given that $\gamma > 1/2$, we can conclude that no agent from $\mathcal{B}_1$ voted for w during the exploitation phase.

Now let us consider the first round for which some candidate $a \in C$ accumulated at least $n/2$ votes, which clearly happens during the exploration phase. Then, at the exact same round at least $n/2$ agents have a in their list; this follows since agents vote for different candidates during the exploration phase, and a candidate is always included in the list once voted for. As a result, our tie-breaking assumption implies that there will be a candidate with the support of at least $n/2$ agents during the exploitation phase. But our previous argument shows that this candidate cannot be w, which is an obvious contradiction. As a result, we have shown that $\gamma \leq 1/2$, and the theorem follows by Lemma 2.2. $\square$

Future Directions. In conclusion, let us briefly mention some intriguing open problems related to the results of this section. We have attempted to argue that candidates with small distortion may arise through natural learning rules. This was motivated in part by Observation 4.1, which implies the instability of outcomes with large distortion. However, the converse of this statement is not quite true: Although there is always a candidate with distortion at most 3 (Gkatzelis, Halpern, and Shah 2020), there might be a subset with at least half of the voters that strictly prefer a different outcome (a.k.a. Condorcet's paradox). Still, there might be an appropriate notion of stability which ensures that near-optimal candidates are in some sense stable. In spirit, this is very much pertinent to the main result of Gkatzelis, Halpern, and Shah (2020) concerning the existence of an *undominated* candidate, leading to the following question:

Question 3. *Are there deterministic and distributed learning rules which converge to a candidate with distortion* 3*?*

[6]The definition of the set $C^{(t)}$ for $t > 1$ is subject to $|C^{(t)}| \geq |C^{(t-1)}| - 1$, i.e. agents never disregard more than 1 candidate in the course of a single round; otherwise, the guarantee we state for the dynamics does not hold due to some pathological instances.

[7]For simplicity, it is assumed that in case multiple such agents exist we posit some arbitrary but *common* among all agents tie-breaking mechanism.

References

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

Anshelevich, E.; Bhardwaj, O.; and Postl, J. 2015. Approximating Optimal Social Choice under Metric Preferences. In Bonet, B.; and Koenig, S., eds., *Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence, January 25-30, 2015*, 777–783. AAAI Press.

Anshelevich, E.; Filos-Ratsikas, A.; Shah, N.; and Voudouris, A. A. 2021. Distortion in Social Choice Problems: The First 15 Years and Beyond. In *Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence, IJCAI 2021*, 4294–4301.

Arrow, K. 1990. *Advances in the Spatial Theory of Voting*. Cambridge University Press.

Assouad, P. 1983. Plongements lipschitziens dans $\mathbb{R}^n$. *Bulletin de la Société Mathématique de France*, 111: 429–448.

Borodin, A.; Lev, O.; Shah, N.; and Strangway, T. 2019. Primarily about Primaries. In *The Thirty-Third AAAI Conference on Artificial Intelligence, AAAI 2019*, 1804–1811. AAAI Press.

Brandt, F.; Conitzer, V.; Endriss, U.; Lang, J.; and Procaccia, A. D. 2016. *Handbook of Computational Social Choice*. USA: Cambridge University Press, 1st edition.

Brânzei, S.; Caragiannis, I.; Morgenstern, J.; and Procaccia, A. D. 2013. How Bad Is Selfish Voting? In desJardins, M.; and Littman, M. L., eds., *Proceedings of the Twenty-Seventh AAAI Conference on Artificial Intelligence*. AAAI Press.

Brill, M.; and Conitzer, V. 2015. Strategic Voting and Strategic Candidacy. In Bonet, B.; and Koenig, S., eds., *Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence*, 819–826. AAAI Press.

Cesa-Bianchi, N.; and Lugosi, G. 2006. *Prediction, learning, and games*. Cambridge University Press.

Clarkson, K. L. 1997. Nearest Neighbor Queries in Metric Spaces. In Leighton, F. T.; and Shor, P. W., eds., *Proceedings of the Twenty-Ninth Annual ACM Symposium on the Theory of Computing, 1997*, 609–617. ACM.

Conitzer, V.; Rognlie, M.; and Xia, L. 2009. Preference Functions That Score Rankings and Maximum Likelihood Estimation. In *Proceedings of the 21st International Jont Conference on Artifical Intelligence*, IJCAI'09, 109–115. Morgan Kaufmann Publishers Inc.

Elkind, E.; Faliszewski, P.; Laslier, J.; Skowron, P.; Slinko, A.; and Talmon, N. 2017. What Do Multiwinner Voting Rules Do? An Experiment Over the Two-Dimensional Euclidean Domain. In Singh, S. P.; and Markovitch, S., eds., *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence, February 4-9, 2017*, 494–501. AAAI Press.

Enelow, J. M.; and Hinich, M. J. 1984. *The Spatial Theory of Voting*. Cambridge Books. Cambridge University Press.

Feige, U. 2000. Approximating the Bandwidth via Volume Respecting Embeddings. *Journal of Computer and System Sciences*, 60(3): 510–539.

Feldman, M.; Fiat, A.; and Golomb, I. 2016. On Voting and Facility Location. In Conitzer, V.; Bergemann, D.; and Chen, Y., eds., *Proceedings of the 2016 ACM Conference on Economics and Computation, EC '16*, 269–286. ACM.

Filos-Ratsikas, A.; and Voudouris, A. A. 2021. Approximate Mechanism Design for Distributed Facility Location. In *Algorithmic Game Theory - 14th International Symposium, SAGT 2021*, volume 12885 of *Lecture Notes in Computer Science*, 49–63. Springer.

Gershkov, A.; Moldovanu, B.; and Shi, X. 2019. Voting on multiple issues: What to put on the ballot? *Theoretical Economics*, 14(2): 555–596.

Gkatzelis, V.; Halpern, D.; and Shah, N. 2020. Resolving the Optimal Metric Distortion Conjecture. In *61st IEEE Annual Symposium on Foundations of Computer Science, FOCS 2020*, 1427–1438. IEEE.

Gupta, A.; Krauthgamer, R.; and Lee, J. R. 2003. Bounded Geometries, Fractals, and Low-Distortion Embeddings. In *44th Symposium on Foundations of Computer Science (FOCS 2003), 11-14 October 2003*, 534–543. IEEE Computer Society.

Karger, D. R.; and Ruhl, M. 2002. Finding nearest neighbors in growth-restricted metrics. In Reif, J. H., ed., *Proceedings on 34th Annual ACM Symposium on Theory of Computing, May 19-21, 2002*, 741–750. ACM.

Kempe, D. 2020. An Analysis Framework for Metric Voting based on LP Duality. In *The Thirty-Fourth AAAI Conference on Artificial Intelligence, AAAI 2020*, 2079–2086. AAAI Press.

Koutsoupias, E.; and Papadimitriou, C. H. 1999. Worst-case Equilibria. In Meinel, C.; and Tison, S., eds., *STACS 99, 16th Annual Symposium on Theoretical Aspects of Computer Science*, volume 1563 of *Lecture Notes in Computer Science*, 404–413. Springer.

Larcinese, V. 2003. The Instrumental Voter Goes to the News-Agent: Demand for Information, Election Closeness, and the Media. UFAE and IAE Working Papers 579.03, Unitat de Fonaments de l'Anàlisi Econòmica (UAB) and Institut d'Anàlisi Econòmica (CSIC).

Larman, D. G. 1967. A New Theory of Dimension. *Proceedings of the London Mathematical Society*, s3-17(1): 178–192.

Lijphart, A. 1992. Democratization and Constitutional Choices in Czecho-Slovakia, Hungary and Poland 1989-91. *Journal of Theoretical Politics*, 4(2): 207–223.

Lu, T.; and Boutilier, C. 2014. Effective sampling and learning for mallows models with pairwise-preference data. *J. Mach. Learn. Res.*, 15(1): 3783–3829.

Matoušek, J. 2002. *Embedding Finite Metric Spaces into Normed Spaces*, 355–400. New York, NY: Springer New York.

Munagala, K.; and Wang, K. 2019. Improved Metric Distortion for Deterministic Social Choice Rules. In Karlin, A.; Immorlica, N.; and Johari, R., eds., *Proceedings of the 2019 ACM Conference on Economics and Computation, EC 2019*, 245–262. ACM.

Obraztsova, S.; Markakis, E.; Polukarov, M.; Rabinovich, Z.; and Jennings, N. R. 2015. On the Convergence of Iterative Voting: How Restrictive Should Restricted Dynamics Be? In Bonet, B.; and Koenig, S., eds., *Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence*, 993–999. AAAI Press.

Procaccia, A. D.; and Rosenschein, J. S. 2006. The Distortion of Cardinal Preferences in Voting. In Klusch, M.; Rovatsos, M.; and Payne, T. R., eds., *Cooperative Information Agents X, 10th International Workshop, CIA 2006*, volume 4149 of *Lecture Notes in Computer Science*, 317–331. Springer.

Restrepo, J. M.; Rael, R. C.; and Hyman, J. M. 2009. Modeling the Influence of Polls on Elections: A Population Dynamics Approach. *Public Choice*, 140(3/4): 395–420.

Skowron, P. K.; and Elkind, E. 2017. Social Choice Under Metric Preferences: Scoring Rules and STV. In Singh, S. P.; and Markovitch, S., eds., *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence, February 4-9, 2017*, 706–712. AAAI Press.

von Neumann, J.; and Morgenstern, O. 1944. *Theory of Games and Economic Behavior*. Princeton University Press.

Zuckerman, M.; Faliszewski, P.; Conitzer, V.; and Rosenschein, J. S. 2011. An NTU Cooperative Game Theoretic View of Manipulating Elections. In Chen, N.; Elkind, E.; and Koutsoupias, E., eds., *Internet and Network Economics - 7th International Workshop, WINE 2011*, volume 7090 of *Lecture Notes in Computer Science*, 363–374. Springer.