Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence (AAAI-17)

Social Choice Under Metric Preferences: Scoring Rules and STV

Piotr Skowron
University of Oxford
United Kingdom
p.k.skowron@gmail.com

Edith Elkind
University of Oxford
United Kingdom
elkind@cs.ox.ac.uk

Abstract

We consider voting under metric preferences: both voters and candidates are associated with points in a metric space, and each voter prefers candidates that are closer to her to ones that are further away. In this setting, it is often desirable to select a candidate that minimizes the sum of distances to the voters. However, common voting rules operate on voters' preference rankings and therefore may be unable to identify the best candidate. A relevant measure of the quality of a voting rule is then its *distortion*, defined as the worst-case ratio between the performance of a candidate selected by the rule and that of an optimal candidate. Anshelevich, Bhardwaj and Postl (2015) show that some popular rules such as Borda and Plurality do badly in this regard: their distortion scales linearly with the number of candidates. On the positive side, Anshelevich et al. identify a few voting rules whose distortion is bounded by a constant; however, these rules are rarely used in practice. In this paper, we analyze the distortion of two widely used (classes of) voting rules, namely, scoring rules and Single Transferable Vote (STV). We show that all scoring rules have super-constant distortion, answering a question that was left open by Anshelevich et al.; however, we identify a scoring rule whose distortion is asymptotically better than that of Plurality and Borda. For STV, we obtain an upper bound of $O(\ln m)$, where m is the number of candidates, as well as a super-constant lower bound; thus, STV is a reasonable, though not a perfect rule from this perspective.

1 Introduction

Voting rules aggregate preferences of multiple agents over a set of available candidates, enabling the agents to choose an option that reflects their collective opinion. Often, voters' preferences are determined by the candidate's positions on several issues, such as the levels of taxation or military spending. In this case, each voter and each candidate can be identified with a point in the *issue space*, and voters tend to prefer candidates who are close to them to the ones that are further away. This setting can be formally modeled by embedding the input election into a *metric space*, i.e., a set of points S endowed with a distance measure d: voters' preferences are consistent with this embedding if voter v prefers candidate a to candidate b whenever $d(v,a) < d(v,b)$. The spatial model of preferences has received a considerable amount of attention in the social choice literature due to its intuitive appeal (Davis and Hinich 1966; Plott 1967; Enelow and Hinich 1984; 1990; McKelvey, Ordeshook, and others 1990; Merrill and Grofman 1999; Schofield 2007). Recently Elkind et al. (2017) used the spacial model to illustrate principles governing different multi-winner election rules.

Now, when preferences are driven by distances, it is natural to measure the *quality* of a candidate c by computing the sum of distances or the maximum distance from c to the voters; these two approaches correspond to, respectively, the utilitarian and the egalitarian social welfare associated with c. We may then want to select a candidate that optimizes the relevant notion of welfare. Of course, this task is not difficult if we are given access to voters' and candidates' locations. However, typically voters are unable to precisely pinpoint their position with respect to each issue, and even the issue space itself may not be known to the designer of the aggregation function. Thus, it is more realistic to expect the voters to simply provide their rankings of candidates, which are determined by the underlying metric space: each voter ranks the candidates by the distance from her. We can then apply one of the many commonly used voting rules (Zwicker 2015) to select an election winner.

Of course, we cannot expect a voting rule that operates on ranked ballots to always identify an alternative that maximizes the social welfare. Consider for instance, the following example, which can be found in the work of Anshelevich, Bhardwaj, and Postl (2015): voters and candidates are positioned on the real line, with candidate a at -1, candidate b at 1, $n+1$ voters at $-.0001$, and n voters at 1. Then the utilitarian social cost associated with a and b is approximately $3n$ and n, respectively. However, any rule that operates on ranked ballots would see that the majority of voters prefer a to b, so any sensible deterministic rule should favor a. Thus, some amount of *distortion* is unavoidable. Nevertheless, we may want to identify a voting rule with the best possible distortion, measured as the ratio of the social cost provided by the optimal candidate and the social cost provided by a candidate selected by the rule, or to bound the distortion of a given voting rule.

Anshelevich, Bhardwaj, and Postl (2015) investigate the latter question for a number of well-known voting rules. Their first result is disappointing: Plurality and the Borda rule, which are arguably the most popular voting rules, fare very poorly in this regard. Specifically, the distortion of these rules scales linearly with the number of candidates m. For other popular rules, such as k-Approval and Veto, distortion cannot even be bounded as a function of m and scales linearly with the number of voters. Anshelevich et al. then consider a number of other voting rules, and identify some rules whose distortion can be bounded by a small constant; notably, this list includes the Copeland rule and any rule that selects from the uncovered set. However, while these voting rules are familiar to (computational) social choice theorists, they are rarely used in practice. Of course, the results of Anshelevich et al. may lead to these rules becoming more popular. However, meanwhile, it remains an important task to assess the distortion provided by voting rules that are currently used by decision-making bodies.

In this paper, we consider two (families of) rules that arguably satisfy this criterion, namely, scoring rules and Single Transferable Vote (STV). Scoring rules can be seen as a generalization of both Plurality and the Borda rule: each position in a voter's preference list is associated with a numerical score, and the rule selects a candidate with the maximum sum of scores. Scoring rules are used in a variety of applications because of their simplicity and intuitive appeal: for instance, Eurovision winners are chosen by a scoring rule. STV is an iterative elimination rule that is used to elect members of governing bodies (at local or national level) in several countries, including Australia, New Zealand, United Kingdom and United States.

Our Contribution We prove bounds on distortion of scoring rules and STV in the utilitarian setting (i.e., where the goal is to select a candidate that minimizes the sum of distances to the voters). Anshelevich, Bhardwaj, and Postl (2015) leave open the question of whether there exists a scoring rule whose distortion is bounded by a constant. We answer this question in the negative, by showing that the distortion of every m-candidate scoring rule is at least $1 + 2\sqrt{\ln m - 1}$. The technique used in our proof can be applied to get stronger bounds for specific families of rules: for instance, we recover the results of Anshelevich et al. for Plurality, Veto, k-Approval and the Borda rule, and obtain a bound of $\Omega\left(\frac{m}{\ln m}\right)$ for the harmonic rule (see Section 2 for definitions). One may wonder if our result can be strengthened to obtain a linear lower bound for all scoring rules (recall that the known bounds for Plurality and the Borda rule are linear). We demonstrate that this is not the case, by showing that the distortion of the harmonic rule is $O\left(\frac{m}{\sqrt{\ln m}}\right)$. Thus, not all scoring rules are equally bad from the perspective of distortion. For STV our upper and lower bounds are very close to each other: the distortion of STV is upper-bounded by $O(\ln m)$ and lower-bounded by $\Omega(\sqrt{\ln m})$. Thus, while STV does not perform quite as well as the Copeland rule, its distortion is much better than that of Plurality or Borda.

Related Work The notion of distortion was proposed by Procaccia and Rosenschein (2006) for a more general model, where voters have utilities for all candidates and the goal is to select a candidate that maximizes the total utility. This model was further explored by Caragiannis and Procaccia (2011) and Boutilier et al. (2015); in particular, Boutilier et al. introduce the harmonic rule, which plays an important role in our analysis. Anshelevich, Bhardwaj, and Postl (2015) initiated the study of distortion in the metric model and proved upper and lower bounds on distortion of a variety of voting rules, both in the utilitarian and in the egalitarian model. Subsequently, Anshelevich and Postl (2016) extended this analysis to randomized voting rules. Feldman, Fiat, and Golomb (2016) analyze distortion under the additional constraint of strategyproofness; in particular, they propose a universally truthful randomized mechanism whose distortion does not exceed 2.

2 Preliminaries

Given a positive integer k, let $[k] = \{1, \dots, k\}$.

An *election* is a triple $E = (C, V, \mathcal{P})$, where $C = \{c_1, \dots, c_m\}$ is a set of m *candidates*, $V = \{v_1, \dots, v_n\}$ is a set of n *voters*, and $\mathcal{P} = (\succ_1, \dots, \succ_n)$ is a list of voters' *preference rankings*; throughout the paper we assume that each voter's preference ranking is a total order over C. We write $a \succ_i b$ to indicate that voter v_i strictly prefers candidate a to candidate b. We write $\mathrm{pos}_v(c)$ to denote the position of candidate c in the preference ranking of voter v: if $\mathrm{pos}_v(c) = 1$ then c is v's most preferred candidate.

Voting Rules A *voting rule* is a mapping $\mathcal{R}$ that, given an election $E = (C, V, \mathcal{P})$, outputs a non-empty subset W of C. The candidates in W are called the *election winners under* $\mathcal{R}$. In this paper we consider a class of voting rules known as *scoring rules* as well as a rule called *Single Transferable Vote*.

A vector $\mathbf{w} = (\omega_1, \dots, \omega_m) \in \mathbb{Q}^m$ is a *weight vector* if $\omega_1 \geq \dots \geq \omega_m \geq 0$. Given an election $(C, V, \mathcal{P})$ with $|C| = m$ and a weight vector $\mathbf{w}$, the $\mathbf{w}$-*score* of a candidate $c \in C$ is computed as $\sum_{v \in V} \omega_{\mathrm{pos}_v(c)}$. A collection $\mathcal{W} = \{\mathbf{w}^m\}_{m=1,\dots}$, where $\mathbf{w}^m$ is a weight vector in $\mathbb{Q}^m$, defines a *scoring rule* $\mathcal{R}_\mathcal{W}$: given an election $(C, V, \mathcal{P})$ with $|C| = m$, it computes the $\mathbf{w}^m$-scores of all candidates in C and outputs the candidates with the maximum $\mathbf{w}^m$-score.

Some popular scoring rules are

- the Borda rule $\mathcal{R}^B$, which is defined by $\omega_j^m = m - j$;
- the Plurality rule $\mathcal{R}^P$, which is defined by $\omega_j^m = 1$ if $j = 1$ and $\omega_j^m = 0$ otherwise;
- the k-Approval rule $\mathcal{R}^{A(k)}$, which is defined by $\omega_j^m = 1$ if $j \leq k$ and $\omega_j^m = 0$ otherwise;
- the Veto rule $\mathcal{R}^V$, which is defined by $\omega_j^m = 1$ if $j < m$ and $\omega_j^m = 0$ if $j = m$;
- the harmonic rule $\mathcal{R}^H$, which is defined by $\omega_j^m = \frac{1}{j}$.

The Single Transferable Vote (STV) is an iterative rule that works as follows. In each round one of the candidates with the lowest Plurality score is removed from the set of candidates and from the rankings of the voters; the Plurality scores are then recalculated. After $m - 1$ rounds, only one candidate survives; this candidate is declared to be a winner. Note that this procedure is non-deterministic: in any given round, there may be multiple candidates with the lowest Plurality score. There are several ways to make it deterministic; in this work, we will consider the so-called *parallel-universe model* (Conitzer, Rognlie, and Xia 2009): in this model, a candidate is said to be an STV winner if it survives after $m-1$ rounds for *some* sequence of choices at each elimination step.

This procedure is illustrated by the following example.

Example 1. Consider the following preference profile:

$$
\begin{array}{ll}
1\colon c_1 \succ_1 c_2 \succ_1 c_3 \succ_1 c_4, & 2\colon c_1 \succ_2 c_4 \succ_2 c_3 \succ_2 c_2,\\
3\colon c_2 \succ_3 c_4 \succ_3 c_1 \succ_3 c_3, & 4\colon c_2 \succ_4 c_4 \succ_4 c_3 \succ_4 c_1,\\
5\colon c_3 \succ_5 c_2 \succ_5 c_4 \succ_5 c_1, & 6\colon c_3 \succ_6 c_2 \succ_6 c_1 \succ_6 c_4,\\
7\colon c_1 \succ_7 c_2 \succ_7 c_4 \succ_7 c_3, & 8\colon c_4 \succ_8 c_1 \succ_8 c_2 \succ_8 c_3.
\end{array}
$$

In the first round c_4 is eliminated. In the second round, after c_4 is removed, the Plurality scores of the candidates c_1, c_2, and c_3, are equal to 4, 2, and 2, respectively. If we choose to eliminate c_3, then in the next round the scores of c_1 and c_2 are equal to 4, so either of them can be eliminated. Alternatively, we can eliminate c_2 in the second round; in the next round, the score of c_1 is 5, while the score of c_3 is 3, so c_3 is eliminated. Thus, the set of STV winners is $\{c_1, c_2\}$.

Distortion A *metric space* is a pair $\mathcal{M} = (S, d)$, where S is a set of points and $d : S \times S \to \mathbb{R}$ is a distance function. A common example of a metric space is the space $\mathbb{R}^k$ for some $k \in \mathbb{N}$ together with the Euclidean distance function $d((x_1, \dots, x_k), (y_1, \dots, y_k)) = (\sum_{i\in[k]} |x_i - y_i|^k)^{1/k}$.

Consider an election $E = (C, V, \mathcal{P})$ and a metric space $\mathcal{M} = (S, d)$. We say that E is $\mathcal{M}$*-consistent* if voters and candidates can be associated with points in S and voters' preferences over the candidates are governed by d, i.e., $V \subseteq S$, $C \subseteq S$ and for each voter v_i and every pair of candidates $a, b \in C$ we have $a \succ_i b$ whenever $d(v_i, a) < d(v_i, b)$. Given an $\mathcal{M}$-consistent election $E = (C, V, \mathcal{P})$, for each candidate $c \in C$ we can compute the sum of distances from c to all voters, i.e., its *utilitarian cost*:

$$q_{\mathcal{M}}(c, E) = \sum_{v\in V} d(v, c).$$

Elements of the set $\operatorname{argmin}_{c\in C} q_{\mathcal{M}}(c, E)$ can be seen as optimal candidates. The *distortion* of a voting rule $\mathcal{R}$ relates the utilitarian cost of the output of this rule to that of an optimal candidate.

Definition 1. *Given a metric space $\mathcal{M} = (S, d)$ and an $\mathcal{M}$-consistent election $E = (C, V, \mathcal{P})$, the* distortion of $\mathcal{R}$ on E *is the ratio*

$$\operatorname{dist}^{\mathcal{M}}(\mathcal{R}, E) = \frac{\max_{c\in\mathcal{R}(E)} q_{\mathcal{M}}(c, E)}{\min_{c\in C} q_{\mathcal{M}}(c, E)}.$$

Let $\mathcal{E}^{\mathcal{M}}_m$ be the set of all $\mathcal{M}$-consistent elections with m candidates. We set

$$\operatorname{Dist}^{\mathcal{M}}_m(\mathcal{R}) = \max_{E\in\mathcal{E}^{\mathcal{M}}_m} \operatorname{dist}^{\mathcal{M}}(\mathcal{R}, E).$$

We note that we measure distortion as a function of the number of candidates and let the number of voters vary. We believe that this is a reasonable approach, as the number of candidates is usually much smaller than the number of voters. In all our lower bound proofs, the number of voters is polynomially related to the number of candidates

When $\mathcal{M} = \mathbb{R}$ with the usual Euclidean distance, we omit it from the notation, i.e., we write $\operatorname{Dist}_m(\mathcal{R})$ in place of $\operatorname{Dist}^{\mathbb{R}}_m(\mathcal{R})$.

3 Distortion of Scoring Rules

Anshelevich, Bhardwaj, and Postl (2015) show that for Plurality and Borda it holds that $\operatorname{Dist}_m \geq 2m - 1$, and for k-approval with $k > 1$ and the Veto rule the distortion cannot be bounded as a function of m. We will now prove that no scoring rule has bounded distortion with respect to $\mathcal{M} = \mathbb{R}$.

Theorem 1. *For every scoring rule $\mathcal{R}_{\mathcal{W}}$ we have $\operatorname{Dist}_m(\mathcal{R}_{\mathcal{W}}) \geq 1 + 2\sqrt{\ln m - 1}$.*

Proof. Suppose that $\mathcal{W} = (\mathbf{w}^m)_{m=1,\dots}$. Fix the number of candidates m and let $\omega_i = \omega^m_i$ for $i \in [m]$. We can assume without loss of generality that $\omega_1 = 1$, $\omega_m = 0$. For every $z = 1, \dots, m$, let

$$F_z = \omega_1 + \dots + \omega_z, \qquad L_z = \omega_{m-z+1} + \dots + \omega_m,$$

and set

$$f_z = \frac{F_z}{z}, \qquad \ell_z = \frac{L_z}{z}.$$

We will construct an m-candidate instance where all candidates and all voters are located in $\mathbb{R}$; the description of the instance depends on the parameters z, n_1, and n_2 whose values will be chosen later.

There are z candidates, $0 < z < m$, at $x = 1$; we denote these candidates by $a_1, \dots, a_z$ and write $A = \{a_1, \dots, a_z\}$. The remaining $m - z$ candidates are at $x = -1$; we denote them by $b_1, \dots, b_{m-z}$, and write $B = \{b_1, \dots, b_{m-z}\}$. There are $z \cdot n_1$ voters at $x = 0$ and $z \cdot n_2$ voters at $x = 1$.

We assume that all voters at $x = 0$ prefer candidates in B to candidates in A, all voters rank the candidates in B in the same way, and in aggregate the voters are indifferent among candidates in A. Specifically, for each $i = 1, \dots, z$ there are n_1 voters at $x = 0$ who rank the candidates as

$$b_1 \succ \dots \succ b_{m-z} \succ a_i \succ \dots \succ a_z \succ a_1 \succ \dots \succ a_{i-1}.$$

and n_2 voters at $x = 1$ who rank the candidates as

$$a_i \succ \dots \succ a_z \succ a_1 \succ \dots \succ a_{i-1} \succ b_1 \succ \dots \succ b_{m-z}.$$

For this instance, the sum of distances from voters to any candidate in A is zn_1, while the sum of distances from voters to any candidate in B is $zn_1 + 2zn_2$, so as long as z, n_1 and n_2 are all positive, it is optimal to select an arbitrary candidate in A. However, the score of b_1 is

$$zn_1 + zn_2\omega_{z+1},$$

whereas the score of any candidate in A is

$$\begin{aligned}&n_1(\omega_{m-z+1} + \dots + \omega_m) + n_2(\omega_1 + \dots + \omega_z)\\ &= n_1 L_z + n_2 F_z.\end{aligned}$$

Thus, if

$$zn_1 + zn_2\omega_{z+1} > n_1 L_z + n_2 F_z, \quad (1)$$

our rule outputs b_1 and its distortion is at least

$$\frac{zn_1 + 2zn_2}{zn_1} = 1 + 2\frac{n_2}{n_1}.$$

We will now show that we can choose $z \in \{1, \dots, m-1\}$ and positive integers n_2, n_1 so that $\frac{n_2}{n_1} \geq \sqrt{\ln m - 1}$ and condition (1) is satisfied.

It is immediate that if $\omega_2 = 1$, we can choose $z = 1$, $n_1 = 1$, and $n_2 = n - 1$: equation (1) then becomes $n > n - 1$, and the distortion of our rule is at least $1 + 2(n-1)$, so it is not bounded as a function of m; in particular, we can choose n so that $n - 1 > \sqrt{\ln m - 1}$ (this generalizes the argument for k-Approval and Veto from the work of Anshelevich, Bhardwaj, and Postl (2015)). Thus, from now on we assume $\omega_2 < 1$.

In this case, it will be convenient to rewrite condition (1) as follows:

$$z + z \cdot \omega_{z+1} \cdot \frac{n_2}{n_1} > F_z \cdot \frac{n_2}{n_1} + L_z$$

or, equivalently,

$$\frac{n_2}{n_1} < \frac{z - L_z}{F_z - z \cdot \omega_{z+1}};$$

note that the denominator of the fraction in the right-hand side is not zero since we assume $\omega_1 = 1$, $\omega_2 < 1$. Let

$$R(z) = \frac{z - L_z}{F_z - z \cdot \omega_{z+1}}.$$

To complete the proof, it remains to argue that we can choose $z \in \{1, \dots, m-1\}$ so that $R(z) \geq \sqrt{\ln m - 1}$. Indeed, for this value of z the quantity $R(z)$ is a rational fraction $\frac{p}{q}$, and we can set $n_2 = p$, $n_1 = q$ to obtain a profile on which the distortion of our rule is at least $1+2\sqrt{\ln m - 1}$. For readability, we prove this fact in a separate lemma.

Lemma 1. *There exists a $z \in \{1, \dots, m-1\}$ such that $R(z) \geq \sqrt{\ln m - 1}$.*

Proof. Let $\lambda = 1/\sqrt{\ln m - 1}$. We will consider two cases.

Case 1: $\omega_2 > 1 - \lambda$.

In this case we can set $z = 1$. Indeed, for $z = 1$ we have $z - L_z = 1$, $F_z - z \cdot \omega_{z+1} = \omega_1 - \omega_2 < \lambda$, so $R(1) \geq 1/\lambda$, which is what we need to prove.

Case 2: $\omega_2 \leq 1 - \lambda$.

Note that in this case we have $z - L_z \geq z\lambda$ for $z = 1, \dots, m-1$.

We will argue that there exists a $z \in \{1, \dots, m-1\}$ such that $f_z - \omega_{z+1} \leq \lambda^2$; we then obtain

$$R(z) \geq \frac{z\lambda}{z\lambda^2} = \frac{1}{\lambda}.$$

Indeed, suppose that for each $z \in \{1, \dots, m-1\}$ we have $f_z - \omega_{z+1} > \lambda^2$. We have

$$\begin{aligned}(z+1) \cdot f_{z+1} &= F_{z+1} = F_z + \omega_{z+1} = z \cdot f_z + \omega_{z+1} \\ &< (z+1) f_z - \lambda^2;\end{aligned}$$

dividing both sides by $z + 1$ gives $f_{z+1} < f_z - \frac{\lambda^2}{z+1}$. Inductively, this implies

$$f_m < 1 - \lambda^2 \left(\frac{1}{2} + \dots + \frac{1}{m}\right).$$

But $\frac{1}{2} + \dots + \frac{1}{m} > \ln m - 1$, and hence

$$\lambda^2 \left(\frac{1}{2} + \dots + \frac{1}{m}\right) > 1,$$

whereas $f_m = \frac{1}{m}(\omega_1 + \dots + \omega_m)$ is necessarily positive, a contradiction. Thus, $f_z - \omega_{z+1} \leq \lambda^2$ for some $z \in \{1, \dots, m-1\}$, and the proof is complete. □

As argued above, Lemma 1 implies that we can pick z, n_1, and n_2 so that the distortion of our rule on the resulting instance is at least $1 + 2\sqrt{\ln m - 1}$. □

We remark that for many scoring rules we can use the construction in the proof of Theorem 1 to obtain lower bounds that are much stronger than $1 + 2\sqrt{\ln m - 1}$. For instance, for the Borda rule we can take $z = 1$ and obtain

$$\mathrm{Dist}_m(\mathcal{R}^B) \geq 1 + 2(m-1)$$

and for Plurality we can take $z = m - 1$ and obtain

$$\mathrm{Dist}_m(\mathcal{R}^P) \geq 1 + 2(m-1),$$

thereby recovering the results of Anshelevich, Bhardwaj, and Postl (2015). Also, for the harmonic rule we can take $z = m - 1$ and obtain

$$\mathrm{Dist}_m(\mathcal{R}^H) = \Omega\left(\frac{m}{\ln m}\right).$$

In particular, our lower bound is linear for Plurality and Borda and sublinear for the harmonic rule. We will now show that, indeed, the harmonic rule always provides sublinear distortion; our result holds for arbitrary metric spaces.

Theorem 2. *For every metric space $\mathcal{M}$ we have*

$$\mathrm{Dist}_m^{\mathcal{M}}(\mathcal{R}^H) = O\left(m(\ln m)^{-1/2}\right).$$

Proof. For the proof, it will be convenient to extend the definition of the weight vector to non-integer "indices" and set $\omega_t = 1/t$ for all $t > 0$. As in the proof of Theorem 1, given a positive integer z, we write $F_z = 1 + \frac{1}{2} + \ldots + \frac{1}{z}$.

Consider a metric space $\mathcal{M} = (S, d)$ and an $\mathcal{M}$-consistent election $E = (C, V, \mathcal{P})$. Let w denote a candidate selected by the harmonic rule, and let o be an optimal candidate, i.e., $o \in \operatorname{argmin} q_{\mathcal{M}}(c, E)$. Set $d = d(o, w)$. Let

$$U = \{v \in V : d(v, o) < d/6\}, \quad u = |U|.$$

We have

$$\begin{aligned}\frac{\sum_{v \in V} d(v, w)}{\sum_{v \in V} d(v, o)} &\leq \frac{\sum_{v \in V} (d(v, o) + d(o, w))}{\sum_{v \in V} d(v, o)} \\ &= 1 + \frac{nd}{\sum_{v \in V} d(v, o)} \leq 1 + \frac{nd}{\sum_{v \in V \setminus U} d(v, o)} \\ &\leq 1 + \frac{nd}{(n-u)d/6} = 1 + \frac{6n}{n-u}.\end{aligned}$$

It remains to argue that $n - u = n \cdot \Omega\left(\frac{\sqrt{\ln m}}{m}\right)$. Let

$$G = \{c \in C : d(c, o) < d/3\}, \quad z = |G|.$$

For each $v \in U$, let $s(v)$ be the score that o receives from v; we have

$$s(v) = \omega_{\mathrm{pos}_v(o)} = \frac{1}{\mathrm{pos}_v(o)}.$$

Also, let

$$\sigma(v) = F_{\mathrm{pos}_v(o)} = \omega_1 + \cdots + \omega_{\mathrm{pos}_v(o)}.$$

Fix a voter $v \in U$. For each $c \in G$ we have

$$d(v, c) \le d(v, o) + d(o, c) < d/6 + d/3 = d/2,$$
$$d(v, w) \ge d(o, w) - d(v, o) > d - d/6 = 5d/6,$$

so v prefers each candidate from G over w. On the other hand, for each $a \in C \setminus G$ we have

$$d(v, a) \ge d(a, o) - d(v, o) > d/3 - d/6 = d/6 > d(v, o),$$

so v prefers o to each candidate from $C \setminus G$. Consequently, v's preference order is of the following form:

$$v\colon \underbrace{c_{i_1} \succ \ldots \succ c_{i_\ell}}_{\text{subset of } G \setminus \{o\}} \succ o \succ \ldots \succ w \succ \ldots$$

Thus, the total score of w is at most

$$u\omega_{z+1} + (n - u),$$

whereas the total score of o is at least

$$\sum_{v \in U} s(v).$$

Moreover, the total score that the candidates in G get from a voter $v \in U$ is at least $\sigma(v)$, so by the pigeonhole principle the total score of some candidate in G is at least

$$\frac{1}{z} \sum_{v \in U} \sigma(v).$$

Since w is the winner under the harmonic rule, we have

$$u \cdot \omega_{z+1} + (n - u) \ge \sum_{v \in U} s(v), \text{ and} \tag{2}$$

$$u \cdot \omega_{z+1} + (n - u) \ge \frac{1}{z} \sum_{v \in U} \sigma(v). \tag{3}$$

Let $p = \frac{1}{u} \sum_{v \in U} \mathrm{pos}_v(o)$. By the inequality between the harmonic mean and the arithmetic mean, we have

$$\frac{u}{\sum_{v \in U} s(v)} = \frac{u}{\sum_{v \in U} \frac{1}{\mathrm{pos}_v(o)}} \le \frac{\sum_{v \in U} \mathrm{pos}_v(o)}{u} = p,$$

so $\sum_{v \in U} s(v) \ge u/p = u \cdot \omega_p$, and inequality (2) implies

$$u \cdot \omega_{z+1} + (n - u) \ge u \cdot \omega_p. \tag{4}$$

As we have $\mathrm{pos}_v(o) \le z$ for each $v \in U$ and $\frac{F_x}{x}$ is a decreasing function of x, we obtain $F_{\mathrm{pos}_v(o)} \ge \mathrm{pos}_v(o) \frac{F_z}{z}$ for each $v \in U$ and hence

$$\frac{\sum_{v \in U} \sigma(v)}{u \cdot p} = \frac{\sum_{v \in U} \sigma(v)}{\sum_{v \in U} \mathrm{pos}_v(o)} = \frac{\sum_{v \in U} F_{\mathrm{pos}_v(o)}}{\sum_{v \in U} \mathrm{pos}_v(o)} \ge \frac{F_z}{z}.$$

Therefore, inequality (3) implies

$$u \cdot \omega_{z+1} + (n - u) \ge \frac{up \cdot F_z}{z^2}. \tag{5}$$

Now, if $p \le z(\ln z)^{-1/2}$ it holds that $\omega_p \ge (\ln z)^{1/2}/z$, whereas if $p > z(\ln z)^{-1/2}$, we have $F_z > \ln z$, and hence $p \cdot F_z/z^2 > (\ln z)^{1/2}/z$. Together with inequalities (4) and (5), this implies

$$\frac{u}{z+1} + (n - u) \ge u \frac{\sqrt{\ln z}}{z},$$

or, equivalently,

$$n \ge u \left(1 + \frac{\sqrt{\ln z}}{z} - \frac{1}{z+1}\right).$$

As $z \le m$, it follows that

$$u \le \frac{n}{1 + \frac{\sqrt{\ln m}}{m} - \frac{1}{m+1}},$$

and, consequently,

$$n - u \ge n \cdot \frac{\frac{\sqrt{\ln m}}{m} - \frac{1}{m+1}}{1 + \frac{\sqrt{\ln m}}{m} - \frac{1}{m+1}} = n \cdot \Omega\left(\frac{\sqrt{\ln m}}{m}\right),$$

which is what we had to prove. □

4 Distortion of Single Transferable Vote

In this section, we focus on STV and show that it has fairly low distortion. Specifically, we demonstrate that its distortion grows at most logarithmically with m; note that this is a much better upper bound than the bound for the harmonic rule provided by Theorem 2. However, we also show that STV is not as good as the Copeland rule or Uncovered Set, by proving a non-constant lower bound.

Theorem 3. *For every metric space $\mathcal{M}$ we have $\mathrm{Dist}_m^{\mathcal{M}}(\mathrm{STV}) = O(\ln m)$.*

Proof. Let $\mathcal{M} = (S, d)$, and consider an $\mathcal{M}$-consistent election $E = (C, V, \mathcal{P})$. Pick $o \in \operatorname{argmin} q_{\mathcal{M}}(c, E)$, $w \in \mathrm{STV}(E)$, and set $d = d(o, w)$. Fix $\gamma > \frac{2}{3}$. Let

$$x = 2\left\lceil \log_{\frac{\gamma}{1-\gamma}} \left(m \cdot \frac{2\gamma - 1}{3\gamma - 2}\right)\right\rceil + 1;$$

note that x is an odd integer and $x = O(\ln m)$. Set $r = \frac{d}{2x}$.

For $i = 1, \ldots, x + 1$, let $\mathcal{B}_i$ be a ball with center o and radius $(2i - 1)r$ (see Figure 1). Note that $w \in \mathcal{B}_{x+1} \setminus \mathcal{B}_x$. We will now argue that $\mathcal{B}_1$ contains at most γn voters.

For the sake of contradiction, assume that this is not the case, i.e., that $|\mathcal{B}_1 \cap V| > \gamma n$. Fix an elimination sequence that results in w being the last surviving candidate. We will say that a candidate c is *supported by a voter* v at some stage of the STV elimination procedure if c is the closest not-yet-removed candidate to v. For each $i \in [x]$, let ℓ_{i-1} be the last candidate from $\mathcal{B}_{i-1}$ to be removed by STV, and let y_i denote the number of candidates in $\mathcal{B}_i \setminus \mathcal{B}_{i-1}$ just before ℓ_{i-1} is removed.

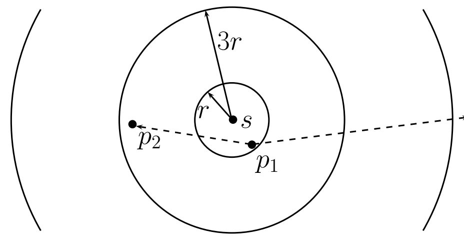

Figure 1: The sequence of balls from the the proof of Theorem 3.

Consider $i \leq x-2$. For every voter $v \in \mathcal{B}_1$ and every candidate $c \notin \mathcal{B}_{i+1}$ it holds that

$$\begin{aligned} d(v,\ell_i) &\leq d(v,o)+d(o,\ell_i) \leq r+(2i-1)r \\ &= (2i+1)r - r < d(c,o)-d(v,o) \leq d(v,c). \end{aligned}$$

Hence, just before ℓ_i is removed, each voter in $\mathcal{B}_1$ supports a candidate in $\mathcal{B}_{i+1}$. Thus, from the pigeonhole principle we infer that at this moment there exists a candidate $c \in \mathcal{B}_{i+1}$ that is supported by more than $\frac{\gamma n}{y_{i+1}+1}$ voters from $\mathcal{B}_1$. Consequently, when STV decides to remove c, all surviving candidates in $\mathcal{B}_{i+3} \setminus \mathcal{B}_{i+2}$ are supported by more than $\frac{\gamma n}{y_{i+1}+1}$ voters. None of these voters is in $\mathcal{B}_1$, as all voters in $\mathcal{B}_1$ prefer c to every $c' \in \mathcal{B}_{i+3} \setminus \mathcal{B}_{i+2}$: indeed, for $v \in \mathcal{B}_1$ we have

$$\begin{aligned} d(v,c) &\leq d(v,o)+d(o,c) \leq 2(i+1)r \\ &< d(c',o)-d(v,o) \leq d(v,c'). \end{aligned}$$

Thus, we get

$$y_{i+3}\frac{\gamma n}{y_{i+1}+1} < n(1-\gamma),$$

and hence $y_{i+1} > \frac{\gamma}{1-\gamma}y_{i+3} - 1$. Set $\xi = \frac{\gamma}{1-\gamma}$. Recall that $w \in \mathcal{B}_{x+1} \setminus \mathcal{B}_x$, so $y_{x+1} \geq 1$. We have

$$\begin{aligned} y_1 &> \xi y_3 - 1 \geq \xi^2 y_5 - \xi - 1 \geq \dots \\ &\geq \xi^{\frac{x-1}{2}} - \xi^{\frac{x-1}{2}-1} - \dots - 1 = \xi^{\frac{x-1}{2}} - \frac{1-\xi^{\frac{x-1}{2}}}{1-\xi} \\ &= \xi^{\frac{x-1}{2}}\left(1+\frac{1}{1-\xi}\right) - \frac{1}{1-\xi} \geq \xi^{\frac{x-1}{2}} \cdot \frac{2-\xi}{1-\xi} \\ &= \left(\frac{\gamma}{1-\gamma}\right)^{\frac{x-1}{2}} \cdot \frac{3\gamma-2}{2\gamma-1} \geq m\frac{2\gamma-1}{3\gamma-2} \cdot \frac{3\gamma-2}{2\gamma-1} = m. \end{aligned}$$

Thus, we obtain $y_1 > m$, a contradiction. We conclude that $\mathcal{B}_1$ contains at most γn voters. Let us now assess the distortion:

$$\begin{aligned} \frac{\sum_{v \in V} d(v,w)}{\sum_{v \in V} d(v,o)} &\leq \frac{\sum_{v \in V}(d(v,o)+d(o,w))}{\sum_{v \in V} d(v,o)} \\ &= 1 + \frac{nd}{\sum_{v \in V} d(v,o)} \leq 1 + \frac{nd}{\sum_{v \in V \setminus \mathcal{B}_1} d(v,o)} \\ &\leq 1 + \frac{nd}{n(1-\gamma)r} = 1 + \frac{2x}{1-\gamma}. \end{aligned}$$

Since γ is a constant and $x = O(\ln m)$, the distortion is upper-bounded by $O(\ln m)$. $\square$

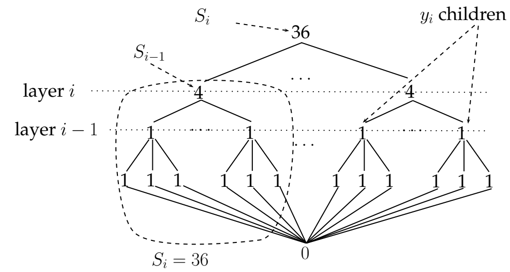

Figure 2: The metric space in the proof of Theorem 4 for $h = 4$. The number in a node denotes the number of voters in that node.

Our next theorem establishes a lower bound of the distortion of STV.

Theorem 4. *There exists a metric space $\mathcal{M}$ such that* $\mathrm{Dist}_m^{\mathcal{M}}(\mathrm{STV}) = \Omega(\sqrt{\ln m})$.

Proof. Given a positive integer h, we construct a perfectly balanced tree of height h and then connect all leaves to one additional node (see Figure 2). We say that all leaves belong to the first layer; for $i > 1$ layer i consists of parents of the nodes at level $i-1$. For $i = 2,\dots,h$, each node at level i has y_i children, where $y_i = 2^i + 2^{i-2} - 2$. We denote by z_i the number of nodes at level i.

We define the length of each edge of our graph to be one, and define the distance between a pair of nodes to be the length of a shortest path between these nodes.

We place one candidate in each node, including the node that is connected to all leaves. As we have $y_i = 2^i + 2^{i-2} - 2$ for $2 \leq i \leq h$, and hence $y_i \leq 2^{i+1}$ for $i \in [h]$, the total number of leaves can be upper-bounded as

$$z_1 \leq 2^{h+1} \cdot 2^h \cdot \dots \cdot 2 = 2^{\frac{(h+1)(h+2)}{2}} \leq 2^{(h+1)^2};$$

as the degree of each internal node is at least 2, this implies that the number of candidates m is at most $2 \cdot 2^{(h+1)^2}$. From this, we conclude that $h \geq \sqrt{\log_2 m} - 2$.

The voters' positions are defined as follows. We place one voter in each leaf node. Let S_i denote the total number of voters in a subtree rooted in a level-i node; we have $S_1 = 1$. Now, for each $i = 2,\dots,h$, we compute S_{i-1} and place S_{i-1} voters in each node of layer i. Thus, we have $S_i = S_{i-1}(y_i+1)$. Note that there are exactly z_1 voters in the bottom layer.

There are $z_i S_{i-1}$ voters at level i, and we have $z_i = z_{i+1}y_{i+1}$. Thus,

$$\frac{z_{i+1}S_i}{z_i S_{i-1}} = \frac{S_i}{y_{i+1}S_{i-1}} = \frac{y_i+1}{y_{i+1}} = \frac{2^i+2^{i-2}-1}{2^{i+1}+2^{i-1}-2} = \frac{1}{2},$$

i.e., layer i contains twice as many voters as layer $i+1$. As layer 1 has z_1 voters, the number of voters in layer i is equal to $z_1 2^{-(i-1)}$.

Let c_0 be the candidate located in the node connected to all the leaves. STV would first remove c_0, as no voter ranks it first, and every other candidate is ranked first by at least one voter. STV can then remove all candidates located in the leaves, one by one: initially, each such candidate is ranked first by exactly one voter, and no leaf candidate gains additional votes as other leaf candidates are removed. Inductively, suppose that STV has removed all candidates in layers $1, \ldots, i-1$, and all other candidates are still present. Then a candidate in layer i is ranked first among the remaining candidates by the S_i voters in the respective subtree, and each candidate in layer j, $j > i$, is ranked first by the $S_{j-1} \geq S_i$ voters who are located in the same node as that candidate. Thus, STV can remove the candidates in layer i one by one. We conclude that the root of the tree can be selected as the winner. As there is a voter in each leaf, the total distance of the voters to the root, which we will denote by d_{stv}, is at least $z_1 h$.

In contrast, the total distance of the voters to the candidate in the node connecting all the leaves can be upper-bounded as

$$\begin{aligned} d_{\mathrm{bot}} &= z_1 + \frac{z_1}{2}2 + \ldots + \frac{z_1}{2^{h-1}}h = z_1 \sum_{i=1}^{h} i 2^{-(i-1)} \\ &= 4z_1(1-(h+1)2^{-h} + h2^{-h-1}) \leq 4z_1. \end{aligned}$$

Thus, we can lower-bound the distortion in our example as

$$\frac{d_{\mathrm{stv}}}{d_{\mathrm{bot}}} \geq \frac{h}{4} \geq \frac{\sqrt{\log_2 m} - 2}{4}.$$

This completes the proof. □

5 Conclusions

We have obtained upper and lower bounds on the distortion of scoring rules and STV. For STV, our bounds provide us with a fairly clear understanding of its distortion, telling us that this is an acceptable rule for a moderate number of candidates. We note, however, that our lower bound in Theorem 4 does not apply to $\mathbb{R}$, and thus STV may perform better when voters and candidates belong to a low-dimensional Euclidean space; understanding whether this is a case is a topic for future research. For scoring rules, our results are much less conclusive. While we have shown that no scoring rule provides constant distortion, our lower bound does not rule out the possibility that some scoring rule performs as least as well as STV. In particular, the lower bound of Theorem 1 is tight for the scoring rule associated with the family of weight vectors given by $\omega_j^m = 1 - \sqrt{\frac{j-1}{m-1}}$; it would be interesting to obtain an upper bound on the distortion of this rule.

Acknowledgments

The authors were supported by ERC-StG 639945.

References

Anshelevich, E., and Postl, J. 2016. Randomized social choice functions under metric preferences. In *Proceedings of the 25th International Joint Conference on Artificial Intelligence (IJCAI-2016)*, 46–59.

Anshelevich, E.; Bhardwaj, O.; and Postl, J. 2015. Approximating optimal social choice under metric preferences. In *Proceedings of the 29th Conference on Artificial Intelligence (AAAI-2015)*, 777–783.

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(9–10):1655–1671.

Conitzer, V.; Rognlie, M.; and Xia, L. 2009. Preference functions that score rankings and maximum likelihood estimation. In *Proceedings of the 21st International Joint Conference on Artificial Intelligence (IJCAI-2009)*, 109–115.

Davis, O. A., and Hinich, M. J. 1966. A mathematical model of preference formation in a democratic society. In Bernd, J., ed., *Mathematical Applications in Political Science II*. Southern Methodist University Press. 175–208.

Elkind, E.; Faliszewski, P.; Laslier, J. F.; Skowron, P.; Slinko, A.; and Talmon, N. 2017. What do multiwinner voting rules do? An experiment over the two-dimensional euclidean domain. In *Proceedings of the 31st Conference on Artificial Intelligence (AAAI-2017)*. To appear.

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

Enelow, J. M., and Hinich, M. J. 1990. *Advances in the spatial theory of voting*. Cambridge University Press.

Feldman, M.; Fiat, A.; and Golomb, I. 2016. On voting and facility location. In *Proceedings of the 17th ACM Conference on Economics and Computation (ACM EC-2016)*, 269–286.

McKelvey, R. D.; Ordeshook, P. C.; et al. 1990. A decade of experimental research on spatial models of elections and committees. *Advances in the spatial theory of voting* 99–144.

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

Plott, C. R. 1967. A notion of equilibrium and its possibility under majority rule. *The American Economic Review* 57(4):787–806.

Procaccia, A. D., and Rosenschein, J. S. 2006. The distortion of cardinal preferences in voting. In *Proceedings of the 10th International Workshop on Cooperative Information Agents (CIA-2006)*, 317–331.

Schofield, N. 2007. *The spatial model of politics*. Routledge.

Zwicker, W. S. 2015. Introduction to the theory of voting. In Brandt, F.; Conitzer, V.; Endriss, U.; Lang, J.; and Procaccia, A. D., eds., *Handbook of Computational Social Choice*. Cambridge University Press. chapter 2.