

Comparing Election Methods Where Each Voter Ranks Only Few Candidates

Matthias Bentert,¹ Piotr Skowron²

¹Algorithmics and Computational Complexity, Faculty IV, TU Berlin, Berlin, Germany
matthias.bentert@tu-berlin.de

²Faculty of Mathematics, Informatics, and Mechanics, University of Warsaw, Warsaw, Poland
p.skowron@mimuw.edu.pl

Abstract

Election rules are formal processes that aggregate voters' preferences, typically to select a single winning candidate. Most of the election rules studied in the literature require the voters to rank the candidates from the most to the least preferred one. This method of eliciting preferences is impractical when the number of candidates to be ranked is large. We ask how well certain election rules (focusing on positional scoring rules and the Minimax rule) can be approximated from partial preferences collected through one of the following procedures: (i) randomized—we ask each voter to rank a random subset of ℓ candidates, and (ii) deterministic—we ask each voter to provide a ranking of her ℓ most preferred candidates (the ℓ -truncated ballot). We establish theoretical bounds on the approximation ratios and complement our theoretical analysis with computer simulations. We find that it is usually better to use the randomized approach.

1 Introduction

An election rule is a function that takes as input a collection of voters' preferences over a given set of m candidates and returns a single candidate, called the winner. There is a large variety of election rules known in the literature (we refer the reader to the survey by Zwicker (2015) for an overview); most of them require the voters to provide strict linear orders over the candidates. Yet, it is often hard, or even infeasible for a voter to provide such a preference ranking, especially when the set of candidates is large. Indeed, it is often believed that a voter can rank at most five to nine candidates (Miller 1956).

In this paper, we ask how the quality of decisions made through voting depends on the amount of information available. Specifically, our goal is to assess the quality of outcomes of elections when each voter can be asked to rank at most $\ell < m$ candidates. We compare two ways of eliciting preferences. In the first approach—which we call *randomized*—we ask each voter to rank a random subset of ℓ candidates. In the second approach—which we call *deterministic*—we ask each voter to provide the ranking of her top ℓ most preferred candidates (the so-called ℓ -truncated

ballot). We investigate how well positional scoring rules and the Minimax rule (a.k.a., Simpson rule) can be approximated by algorithms that use one of the two elicitation methods.

Our Contribution

1. In Section 3.1 we identify a class Sep_ℓ of positional scoring rules that, for a given ℓ , can be well approximated using the randomized approach. Sep_2 consists of a single rule, namely the Borda count; the number of rules in Sep_ℓ grows exponentially with ℓ . We theoretically prove approximation guarantees for the rules from Sep_ℓ —these guarantees are more likely to be accurate when the number of voters is large. In Section 3.2 we provide an analogous analysis for the Minimax rule.
2. In Section 4 we prove upper-bounds on the approximation ratios of algorithms that use ℓ -truncated ballots. We show that the algorithm that minimizes the maximal regret of Lu and Boutilier (2011) (we recall this algorithm in Section 4.1) matches our upper-bounds (for Minimax our analysis is tight up to a small constant factor).
3. We ran computer simulations to verify how the approximation ratio depends on the particular distribution of voters preferences (Section 5). Our experiments confirm that in most cases (with the exception of very unstructured preferences) the randomized approach is superior. We also show that usually only a couple of hundreds of voters are required to achieve a reasonably good approximation.

Related Work

Our work contributes to the broad literature on handling incomplete information in voting—for a survey on this topic, we refer the reader to the book chapter by Boutilier and Rosenschein (2015). Specifically, our research is closely related to the idea of minimizing the maximal regret (Lu and Boutilier 2011). Therein, for a partial preference profile P , the goal is to select a candidate c such that the score of c in the worst possible completion of P is maximized. Our paper complements this literature by (1) providing an accurate analysis of the approximation ratios for various methods (which allows to better judge suitability of different methods for handling incomplete information), and (2) by analyzing two natural methods of preference elicitation.

Algorithms for minimizing the maximal regret interpret the missing information in the most pessimistic way. Other approaches include assuming the missing pairwise preferences to be distributed uniformly (e.g. Xia and Conitzer (2011)) or similarly to the already observed pairwise preferences (Doucette (2014; 2015)).

Our work is also closely related to the literature on distortion (Procaccia and Rosenschein 2006; Caragiannis and Procaccia 2011; Boutilier et al. 2015). There, an underlying utility model is assumed and the goal is to estimate how well various voting rules that only have access to ordinal preferences approximate optimal winners, i.e., candidates that maximize the total utility of the voters. The definition of distortion has also been adapted to social welfare functions (where the goal is to output a ranking of candidates rather than a single winner) (Benadé, Procaccia, and Qiao 2019) and to participatory budgeting (Benade et al. 2017). Some works also study distortion assuming a certain structure of the underlying utility model (for example, that it can be represented as a metric space) (Anshelevich et al. 2018; Anshelevich and Postl 2017; Feldman, Fiat, and Golomb 2016; Goel, Krishnaswamy, and Munagala 2017; Gross, Anshelevich, and Xia 2017; Pierczynski and Skowron 2019).

Another related line of research concerns ordinal peer grading (Raman and Joachims 2014; Caragiannis, Krimpas, and Voudouris 2015; 2016; Caragiannis et al. 2019). The main difference between these works and ours is that there, the authors assume that a ground truth exists. Here, we do not: The voters’ preferences can vary not only because the voters have imperfect information of what is the “correct” outcome, but also since they put focus on various aspects—different voters can have different opinions of what is “correct”; the goal is to aggregate (subjective) preferences rather than to discover the ground truth (which might not even exist). Our results both apply when the existence of a ground truth is assumed and when it is not assumed and hence is more widely applicable.

Our randomized algorithms are similar to the one proposed by Hansen (2016). Yet, the work of Hansen differs from ours in several important aspects. First, in Hansen’s algorithm (called Random Pairs) each voter is asked to compare a predetermined number of random pairs of candidates (the pairs are sampled uniformly at random); in our case the voter does not compare pairs but rather ranks predetermined subsets of candidates. Second and more importantly, Hansen does not ask how well certain rules can be approximated by schemes that require less information, but rather analyzes how often his method is able to find a Condorcet winner. Our approach is different—we investigate a number of rules, and for each of these rules we ask how well this rule can be approximated having only partial knowledge of the voters’ preferences. Third, Hansen’s work is based only on simulations and does not contain theoretical results.

A similar approach to ours has been recently applied to STV (Ayadi et al. 2019) and to other rules whose definitions are based on the notion of score (Ayadi, Ben Amor, and Lang 2019). In particular, the work of Ayadi, Ben Amor, and Lang (2019) complements our experimental analysis for the Borda and Minimax rules for truncated ballots.

2 Preliminaries

An election is a pair $E = (V, C)$, where $V = \{v_1, v_2, \dots, v_n\}$ and $C = \{c_1, c_2, \dots, c_m\}$ denote the sets of n voters and m candidates, respectively. Each voter v_i is endowed with a *preference ranking* over the candidates, which is a total ordering of the candidates and which we denote by \succ_i . For each candidate $c \in C$ by $\text{pos}_i(c)$ we denote the position of c in v_i ’s preference ranking.

A *voting rule* is a function that, for a given election E , returns a subset of candidates, which we call tied *winning candidates*. Below we describe several (classes of) voting rules that we will focus on in this paper.

For an integer t we use $[t]$ to denote the set $\{1, 2, \dots, t\}$. A *positional scoring function* is a mapping $\lambda: [m] \rightarrow \mathbb{R}$ that assigns to each position a real value: intuitively, $\lambda(p)$ is a score that a voter assigns to a candidate that she ranks as her p -th most preferred one. For each positional scoring function λ we define the λ -score of a candidate c as $\text{sc}_\lambda(c) = \sum_{v_i \in V} \lambda(\text{pos}_i(c))$, and the corresponding election rule selects the candidate(s) with the highest λ -score. Examples of common positional scoring rules include:

Borda rule: Based on a linear decreasing positional scoring function, the Borda rule is defined as $\beta(p) = m - p$.

Plurality rule: Being equivalent to the 1-approval rule, the positional scoring function for the Plurality rule assigns a score of one to the first position and zero to all others.

Another important class of voting rules origins from the Condorcet criterion. It says that if there exists a candidate c that is preferred to any other candidate by a majority of voters, then the voting rule should select c . We focus on one particular rule satisfying the Condorcet criterion (we chose a rule picking the candidates that maximize a certain score so that the standard definition of approximation applies):

Minimax rule (a.k.a., the Simpson rule). For an election $E = (V, C)$ and two candidates $c, c' \in C$, we define $\text{sc}_{\text{MM}}(c, c') = |\{v_i \in V \mid c \succ_i c'\}|$ as the number of voters who prefer c to c' and

$$\text{sc}_{\text{MM}}(c) = \min_{c' \neq c} \{\text{sc}_{\text{MM}}(c, c')\}.$$

The rule then selects the candidates with the highest sc_{MM} score.

We use a natural definition of approximation.

Definition 1. We say that \mathcal{A} is an α -approximation algorithm for a rule \mathcal{R} if for each election E it holds that:

$$\frac{\text{score}_{\mathcal{R}}(\mathcal{A}(E))}{\max_{w \in \mathcal{R}(E)} \text{score}_{\mathcal{R}}(w)} \geq \alpha,$$

where $\text{score}_{\mathcal{R}}$ is a function representing the score \mathcal{R} awards each candidate, $\mathcal{R}(E)$ is the set of winners returned by \mathcal{R} , and $\mathcal{A}(E)$ is the candidate returned by \mathcal{A} .

Later on, we will consider algorithms that have access only to certain parts of the input instances. In such cases the above definition still applies. For example, let $\text{trunc}(E, \ell)$ denote the ℓ truncated instance obtained from E , that is, a partial election which for each voter only contains her preferences ranking from E , truncated to the top ℓ positions.

Algorithm 1: Algorithm α -PSF-ALG for positional scoring functions.

```

1 foreach candidate  $c$  do
2    $\text{total\_score}(c) \leftarrow 0$ 
3    $\text{ranked}(c) \leftarrow 0$ 
4 foreach voter  $v$  do
5    $S_v \leftarrow$  random set of  $\ell$  candidates
6   ask  $v$  to rank  $S_v$ 
7   foreach  $c \in S_v$  do
8     if  $c$  is ranked  $i$ -th among  $S_v$  then
9        $\text{total\_score}(c) \leftarrow \text{total\_score}(c) + \alpha_i$ 
10       $\text{ranked}(c) \leftarrow \text{ranked}(c) + 1$ 
11 foreach candidate  $c$  do
12    $\text{score}(c) \leftarrow \text{total\_score}(c) \cdot \frac{n}{\text{ranked}(c)}$ 
13 return candidate  $c$  with maximal  $\text{score}(c)$ 

```

Then, we say that \mathcal{A} is an α -approximation algorithm for \mathcal{R} for ℓ -truncated instances, when for each election instance E it holds that:

$$\frac{\text{score}_{\mathcal{R}}(\mathcal{A}(\text{trunc}(E, \ell)))}{\max_{w \in \mathcal{R}(E)} \text{score}_{\mathcal{R}}(w)} \geq \alpha.$$

3 Randomized Approach

In this section we explore a randomized approach, where each voter is asked to rank a random subset of candidates.

3.1 Scoring Rules

We start our analysis by looking at the class of positional scoring rules. For the sake of simplicity, we will assume throughout this section that n is divisible by m . We first present an algorithm that estimates the score of each candidate and picks the candidate with the highest score. The algorithm is parameterized with a natural number $\ell \leq m$ and a vector of ℓ reals $\alpha = (\alpha_1, \dots, \alpha_\ell)$ —for a fixed vector α we will call the algorithm α -PSF-ALG. This algorithm asks each voter to rank a random set of ℓ candidates. We say that a candidate c is ranked by a voter v if c belongs to the set of ℓ candidates that v was asked to rank. If c is the i -th most preferred among the candidates ranked by a voter, then c receives the score of α_i from the voter. Such scores are summed up for each candidate, normalized by the number of voters who ranked the respective candidate, and the candidate with the highest total score is declared the winner. Pseudocode of this algorithm is given in Algorithm 1.

We will show that for some positional scoring rules, by choosing the vector α carefully, we can find good approximations of winning candidates with high probability. First, through Theorem 1 we establish a relation between positional scoring functions λ and vectors α that should be used to assess λ ; the formula is not intuitive, and we will discuss it later on. In particular, we will explain which positional scoring functions can be well approximated using this approach, that is, we will discuss the structure of the class of scoring functions which are covered by the following theorem.

Theorem 1 (\star^1). For a non-increasing vector of ℓ reals $\alpha = (\alpha_1, \dots, \alpha_\ell)$ consider the positional scoring function

$$\lambda_\alpha(p) = \frac{1}{\binom{m-1}{\ell-1}} \cdot \sum_{i=1}^{\ell} \alpha_i \binom{p-1}{i-1} \cdot \binom{m-p}{\ell-i}.$$

For a candidate $c \in C$ that is ranked by at least one voter, we denote by X_c the random variable describing the total normalized score that c was assigned by α -PSF-ALG. Then, the expected value $\mathbb{E}(X_c)$ is equal to the λ_α -score of c , and the probability that the score computed by α -PSF-ALG for c differs from its expected value by a multiplicative factor of $1 \pm \epsilon$ is upper-bounded by $2 \exp\left(-\frac{\epsilon^2 \mathbb{E}(X_c)}{3}\right)$, that is,

$$\mathbb{P}\left(|X_c - \mathbb{E}(X_c)| \geq \epsilon \mathbb{E}(X_c)\right) \leq 2 \exp\left(-\frac{\epsilon^2 \ell \text{sc}_{\lambda_\alpha}(c)}{6m\alpha_1}\right).$$

Now, let us discuss the form of positional scoring functions $\lambda_\alpha(p)$ used in Theorem 1. First, observe that for $\ell = 2$, if we set $\alpha_1 = 1$ and $\alpha_2 = 0$, then we have that $\lambda_\alpha(p) = \binom{p-1}{0} \cdot \binom{m-p}{2-1} = m-p = \beta(p)$. This means that by asking each voter to rank only two candidates, we can correctly (in expectation) assess the Borda scores of the candidates.

Corollary 2. For a candidate c the expected value of the score computed by Algorithm (1, 0)-SEP-ALG for c is the Borda score of c .

Unfortunately, not every positional scoring function can be efficiently assessed while asking each voter to rank only few candidates and we will now describe the class of all positional scoring functions which can be computed correctly in expectation by our algorithm for any fixed ℓ . Since each positional scoring function is based on some m -dimensional vector $\beta = (\beta_1, \beta_2, \dots, \beta_m)$ which can be expressed as $\sum_{i=1}^m \eta_i \cdot \beta_i$, where $\eta_1 = (1, 0, \dots)$, $\eta_2 = (0, 1, 0, \dots)$ and so on, these η -vectors form a basis of the linear space of positional scoring functions.

Let $\text{Sep}_\ell = \{\lambda_\alpha : \alpha \in \mathbb{R}^\ell\}$ be the set of all positional scoring functions that can be computed (correctly in expectation) by our algorithm for a fixed ℓ . Since it holds for each two ℓ -element vectors $\alpha, \alpha' \in \mathbb{R}^\ell$ that $\lambda_{\alpha+\alpha'}(p) =$

$$\sum_{i=1}^{\ell} (\alpha_i + \alpha'_i) \binom{p-1}{i-1} \binom{m-p}{\ell-i} = \sum_{i=1}^{\ell} \alpha_i \binom{p-1}{i-1} \binom{m-p}{\ell-i} + \sum_{i=1}^{\ell} \alpha'_i \binom{p-1}{i-1} \binom{m-p}{\ell-i} =$$

$\lambda_\alpha(p) + \lambda_{\alpha'}(p)$, we have that Sep_ℓ is a linear space too.

Thus, Sep_ℓ is an ℓ -dimensional linear subspace of the m -dimensional space of all positional scoring functions, and so we can compactly describe it by providing ℓ scoring functions forming a basis of Sep_ℓ . Figure 1 visually illustrates the scoring functions forming a basis for $\ell \in \{2, 4\}$. In other words, for a given value of ℓ , we can use Theorem 1 to correctly compute (in expectation) all scoring functions which can be obtained as linear combinations of the scoring functions depicted in Figure 1.

¹Due to space constraints, we defer all proof details to the full

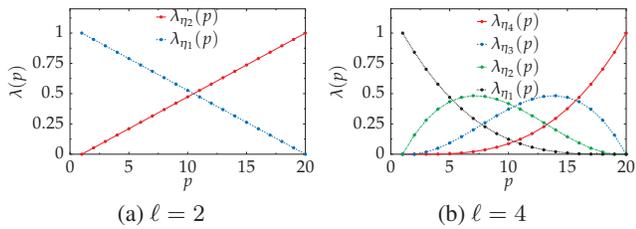


Figure 1: A basis of the space $Sep_\ell = \{\lambda_\alpha : \alpha \in \mathbb{R}^\ell\}$ for two different values of ℓ . Each plot illustrates ℓ positional scoring functions that span the space of all positional scoring functions which can be correctly assessed by algorithm α -PSF-ALG while asking each voter to rank only ℓ candidates.

Theorem 1 can be also applied to positional scoring rules that do not belong to the aforementioned linear space. Let us explain that using the example of the harmonic scoring rule, defined through the positional scoring function $\lambda_{\text{har}}(p) = 1/p$ (our choice of the harmonic rule is due to its particularly appealing theoretical properties (Boutlier et al. 2012; Anshelevich et al. 2018)). By solving a simple linear program we can find coefficients for the functions from the basis of Sep_ℓ that induce a function from Sep_ℓ that is closest to λ_{har} , according to the following distance:

$$d(\lambda_1, \lambda_2) = \max_{p \in [m]} |\lambda_1(p) - \lambda_2(p)|.$$

For example, for $m = 50$ candidates, using $\ell = 5$ comparisons we can arbitrarily well approximate a scoring rule based on the vector of scores which is of distance 0.15 from the harmonic vector. For $\ell = 10$, we can arbitrarily well approximate a scoring rule based on a vector within a distance of 0.03 from the harmonic vector. For $\ell = 20$ the distance drops to 0.0003. This shows that our algorithm can be efficiently used for approximating e.g., the harmonic rule.

Finally, let us give some intuition regarding the probabilities assessed in Theorem 1. For 21 candidates and n voters the Borda score of a winning candidate is at least $10n$. When assessing the score of a winning candidate with $\ell = 2$, to have an approximation within a factor of 1.01 with probability at least 0.999, we need about 72 thousands voters. For one million voters, this probability drops below $4/10^{42}$.

3.2 Minimax Rule

We will now investigate whether the Minimax rule can be well approximated when each voter is only asked to rank a few candidates. We will use an algorithm similar to Algorithm 1: Each voter v ranks a subset of candidates S_v and whenever two candidates $c, c' \in S_v$ are ranked by a voter v , we use her preference list to estimate $\text{sc}_{\text{MM}}(c, c')$. This algorithm is formalized in Algorithm 2.

Theorem 3 (\star). *For each candidate $c \in C$ the probability that the total normalized score computed by Algorithm 2*

version of the paper (Bentert and Skowron 2019) (affected results are marked with a \star).

Algorithm 2: Randomized Algorithm for Minimax.

```

1 foreach candidates  $c, c' \in C$  do
2    $S[c, c'] \leftarrow 0$ 
3 foreach voter  $v$  do
4    $S_v \leftarrow$  random set of  $\ell$  candidates
5   ask  $v$  to rank  $S_v$ 
6   foreach  $c \in S_v$  do
7     foreach  $c' \in S_v \setminus \{c\}$  do
8       if  $c \succ_i c'$  then
9          $S[c, c'] \leftarrow S[c, c'] + 1$ 
10 foreach candidate  $c$  do
11    $S[c] \leftarrow \min \left\{ \frac{n \cdot S[c, c']}{S[c, c'] + S[c', c]} : S[c, c'] + S[c', c] > 0 \right\}$ 
12 return candidate  $c$  with maximal  $S[c]$ 

```

for c differs from the true Minimax score of c by a multiplicative factor of at least $1 \pm \epsilon$ is upper-bounded by

$$m \exp\left(-\frac{\epsilon^2 \ell^2 \text{sc}_{\text{MM}}(c, c_{\min})}{6m^2}\right)$$

4 Deterministic Approach

When not asking each voter about each candidate, one always has to decide whether each voter is asked about random candidates or about specific ones. On the one hand, asking about specific positions in preference rankings allows one to focus on the top ones that seem to contain more relevant information; especially when the goal is to select the winner, who—intuitively—is more likely to appear in top positions. On the other hand, asking voters about random candidates might be more advantageous as the input may contain dependencies between candidates that are not known a priori.

In this section we investigate the case when each voter is asked about her ℓ most preferred candidates. We will describe an algorithm that approximates the true score at least as well as any other algorithm and analyze its performance for Borda and for Minimax. We will then show general lower bounds on the approximation ratio, and prove that the approximation ratio for the Borda rule matches this lower-bound, and the ratio for Minimax almost does.

4.1 The Best Approximation Algorithm for ℓ -Truncated Elections

Let us start by describing the algorithm that for each ℓ -truncated instance gives the best possible approximation guarantee, that is, the best approximation of the true winner in the worst-case full preference profile that induces the given ℓ -truncated instance. We mention that the idea of this algorithm is very similar to the one behind the algorithms for minimizing the maximal regret (Lu and Boutlier 2011), yet the analysis of the approximation ratio of the algorithm is new in this paper.

Consider an election E and let E_ℓ be the ℓ -truncated instance obtained from E . Observe that when given E_ℓ and

choosing a winner, the worst case occurs if the picked winner is ranked at the very last position by all voters that did not rank this candidate in E among the first ℓ positions, and the true winner (that our algorithm did not pick) is ranked at position $\ell + 1$ by each voter who did not rank this candidate.

For each candidate c we compute two scores: The worst possible score that c is guaranteed to get (denoted by $\text{worst}(c)$) and the best possible score that c can get (denoted by $\text{best}(c)$). Let w, b_1 , and b_2 be the candidates with the highest worst, the highest best, and the second highest best score, respectively. If any candidate $c \neq b_1$ is declared winner, then we can guarantee an approximation ratio of $\text{worst}(c)/\text{best}(b_1)$, which is clearly maximized by $c = w$.

If candidate b_1 is declared winner, then we can guarantee an approximation ratio of $\text{worst}(b_1)/\text{best}(b_2)$. The optimal approximation ratio is therefore achieved by an algorithm that computes all the possible scores and then declares w or b_1 winning depending on which one guarantees a better result.

Interestingly, while our algorithm provides the best possible approximation, it can select a candidate that is not a possible winner, i.e., that is not a winner in any profile consistent with the truncated ballot at hand.

Example 1. Consider the following instance with 5 candidates A, B, C, D, E , four voters v_1, v_2, \dots, v_4 , and $\lambda = (3, 1, 1, 1, 0)$.

$$\begin{aligned} v_1 &: B \succ A \\ v_2 &: C \succ A \\ v_3 &: D \succ A \\ v_4 &: E \succ A \end{aligned}$$

It holds that $\text{worst}(A) = 4$, $\text{worst}(B) = \text{worst}(C) = \text{worst}(D) = \text{worst}(E) = 3$, $\text{best}(A) = 4$ and $\text{best}(B) = \text{best}(C) = \text{best}(D) = \text{best}(E) = 6$ and therefore declaring a winning achieves the best approximation ratio. However, in any election that is consistent with the given truncated election, at least two candidates in $\{B, C, D, E\}$ get at least 5 points while A always gets 4 points. Thus, A is not a possible winner.

4.2 Positional Scoring Rules: Approximation Guarantees for ℓ -Truncated Elections

In this section, we continue our analysis of the algorithm from Section 4.1. We will now prove guarantees that each of the two rules (i.e., choosing the candidate with highest worst, respectively best, score) provide for positional scoring rules when only having access to ℓ -truncated elections.

Theorem 4 (\star). Let \mathcal{R} be a positional scoring rule defined by the scoring function $\lambda(i) = \alpha_i$. The algorithm from Section 4.1 for ℓ -truncated elections gives an approximation guarantee of

$$\frac{\sum_{i=1}^{\ell} \alpha_i}{m\alpha_{\ell+1} + \frac{\alpha_1 - \alpha_{\ell+1}}{\alpha_1} \sum_{i=1}^{\ell} \alpha_i}.$$

Theorem 4 gives a very general result that applies to any positional scoring rule. For instance, for k -approval we get the approximation ratio of ℓ/m .

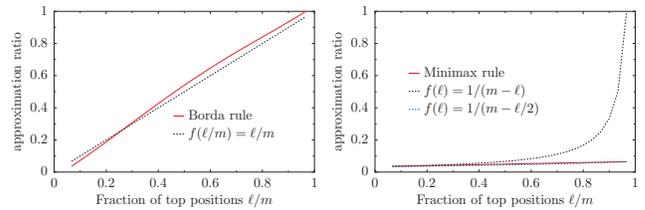


Figure 2: The approximation ratios for the deterministic algorithms from Section 4.1 for the Borda rule (left) and for the Minimax rule (right). The plot was generated for $m = 30$. Note that the plotted line for the Minimax rule is almost indistinguishable from the plotted line for $1/(m - \ell/2)$

Corollary 5 (\star). The algorithm from Section 4.1 for k -approval with ℓ -truncated elections, $k > \ell$, gives the approximation guarantee of ℓ/m .

For the Borda rule, we get the approximation of $\frac{\ell}{m + \frac{\ell}{m-1} \cdot \ell}$, which looks similarly to $\frac{\ell}{m}$ (see the left-hand side plot in Figure 2).

Corollary 6 (\star). The algorithm from Section 4.1 gives an approximation guarantee of $\frac{\ell}{m + \frac{\ell}{m-1} \cdot \ell}$ for Borda with ℓ -truncated elections.

We conclude by providing intuitive explanation of instances that match the bound from Theorem 4. In these instances all candidates get roughly the same worst score and there are two candidates a, b that also get average worst score but only appear as few times as possible in the first ℓ positions. If candidate a is declared the winner by any rule then she gets 0 points from all voters, that did not rank her in the first ℓ positions and b gets $m - \ell - 1$ points from these voters. Otherwise the winning candidate gets 0 points from all voters that did not rank her in the first ℓ positions and a gets $m - \ell - 1$ points whenever she is not ranked first or second. For such constructed instances the approximation ratio of our (optimal) algorithm is exactly as given in Theorem 4.

4.3 Minimax Rule

Let us now move to the analysis for the Minimax rule. We start by showing that no deterministic algorithm for Minimax can guarantee a better approximation ratio than $\frac{1}{m-\ell}$.

Theorem 7 (\star). There exists no (deterministic) algorithm for ℓ -truncated elections \mathcal{F} that is a $\frac{1}{m-\ell-\epsilon}$ -approximation algorithm for the Minimax rule for any $\epsilon > 0$.

Theorem 7 already shows that with ℓ -truncated ballots Minimax cannot be well approximated. In particular, the bound for the Minimax rule is much worse than for scoring-based rules. We do not know whether the bound from Theorem 7 is tight. Yet, we can show that a simplified variant of the algorithm from Section 4.1 that computes the maximum worst score of each candidate and declares the one with the highest score winner, achieves an approximation ratio which is lower-bounded by $\frac{1}{m-\ell/2}$, which means that for reasonably small ℓ it (almost) matches the upper bound

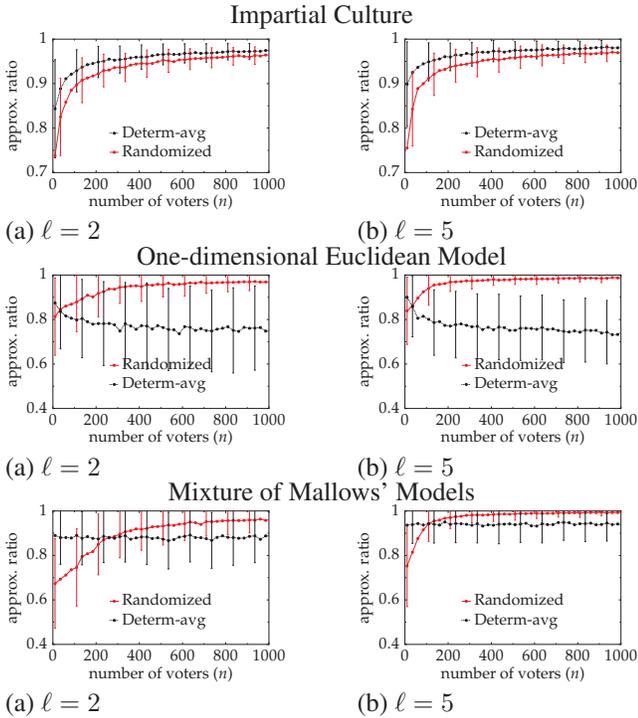


Figure 3: Approximation ratio for the two algorithms for the Borda rule assessed through computer simulations.

from Theorem 7 (see the right-hand side plot in Figure 2 for the comparison of these two bounds).

Theorem 8 (\star). *The algorithm from Section 4.1 approximates Minimax in ℓ -truncated elections within a factor of*

$$\frac{1}{(m - \ell) \cdot \left(1 + \frac{\ell^2}{m^2 - \ell^2 - m + \ell}\right)} \geq \frac{1}{m - \ell/2}.$$

5 Experimental Evaluation

In Sections 3 and 4 we have assessed the worst-case guarantees of our approximation algorithms. In this section we investigate how these guarantees depend on particular distributions of the voters' preferences. We tested the following distributions over preference rankings:

Impartial Culture (IC). In the Impartial Culture model each ranking over the candidates has the same probability.

One-dimensional Euclidean Model (1D). First, we associate each voter and each candidate with a point from the interval $[0, 1]$ —these points are sampled independently and uniformly at random. Then, each voter ranks the candidates according to her distance, preferring the ones which are closer to those which are farther.

Mixture of Mallows' Models (MMM). In the Mallows' model (Mallows 1957) we are given a reference ranking π and a real value $\phi \in [0, 1]$. The probability of sampling a ranking τ is proportional to $\phi^{d_K(\pi, \tau)}$, where $d_K(\pi, \tau)$ is the number of swaps of adjacent candidates that are required to turn ϕ into τ . We used a uniform mixture of

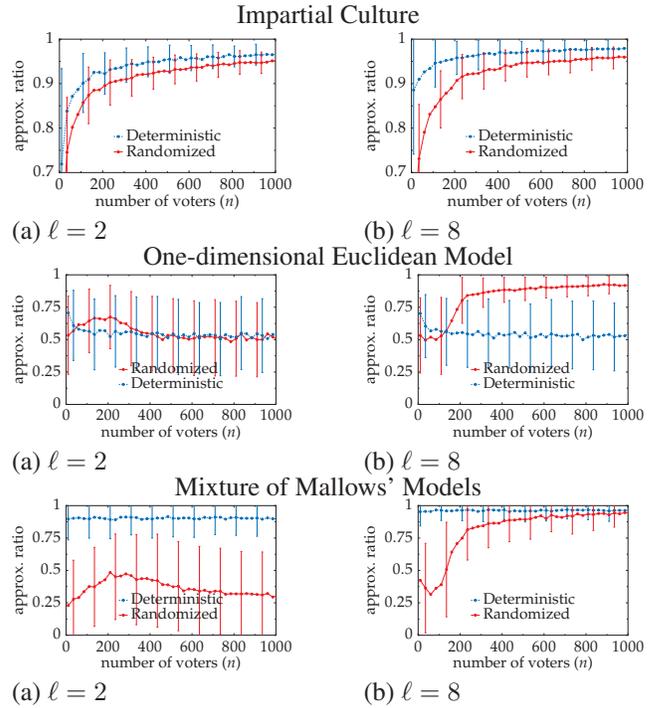


Figure 4: Approximation ratio for the two algorithms for the Minimax rule assessed through computer simulations.

three Mallows' models, where for each of the three models we drew the reference ranking π and the real value ϕ uniformly at random.

For each distributions \mathcal{D} over preferences and for each approximation algorithm \mathcal{A} we ran computer simulations as follows: We set the number m of candidates to 50 and tested for $\ell \in \{2, 5, 8\}$ and n ranging from 10 to 1000 in steps of 25. For each combination of values of (ℓ, n) we ran 500 independent experiments, each time computing the ratio $r(\mathcal{A}, \mathcal{D})$ between the score of the candidate returned by algorithm \mathcal{A} to the score of the optimal candidate. The averages of these ratios and the corresponding standard deviations for the Borda and the Minimax rules are depicted in Figure 3 and Figure 4.

5.1 Approximation Algorithms for Borda

We empirically tested how well the two algorithms that we analyzed theoretically in the previous sections approximate the Borda rule. Specifically, we implemented Algorithm 1—which we will refer to as RANDOMIZED, and the algorithm described in Section 4.1. We also checked two other deterministic heuristics that appear simple and intuitive:

1. The variant of the deterministic algorithm that always picks the candidate with the highest worst score.
2. An algorithm we call DETER-AVG that, for each voter v_i and candidate c_j assigns to c_j the score
 - (a) $\beta(\text{pos}_i(c_j))$ if $\text{pos}_i(c_j) \leq \ell$,
 - (b) the average score of the unranked positions $\sum_{p=\ell+1}^m \beta(p)/(m - \ell)$, otherwise.

The three deterministic algorithms were almost indistinguishable in our simulations—DETER-AVG was slightly better than the other two. Thus, we only present the results for DETER-AVG and RANDOMIZED. We found the following:

1. For preferences with no or with little structure (IC), the deterministic algorithm gives better results. For preferences with more structure (those obtained from ID and MMM models), the randomized algorithm significantly outperforms the deterministic ones.
2. For each preference distribution that we tested the randomized algorithm gives high quality approximations unless the number of voters is small. Our results suggest to ask each voter to rank a random subsets of alternatives when the goal is to approximate the Borda rule, and if there are more than a couple of hundreds voters.

5.2 Approximation Algorithms for Minimax

Similarly to Section 5.1, we empirically tested how well the randomized algorithm (Algorithm 2) and the deterministic algorithm from Section 4.1 approximate the Minimax rule. We refer to the two algorithms as RANDOMIZED and DETERMINISTIC, respectively. We also tested two other natural heuristics. For each two candidates c and c' , let $n(c, c')$ denote the number of voters who (i) rank c and c' among their ℓ most preferred candidates and prefer c over c' or (ii) who rank c but not c' among their top ℓ positions. Then:

1. In our first heuristic, for each two candidates, c and c' , we use a method similar to Minimax, but replace $sc_{MM}(c, c')$ by $n(c, c')$. Then, similarly as in the case of Minimax we compute for each candidate c the score $\min_{c' \neq c} n(c, c')$ and pick the candidate w with the maximal score.
2. In the second heuristic, we replace $sc_{MM}(c, c')$ by

$$n \cdot \frac{n(c, c')}{n(c, c') + n(c', c)}.$$

In our simulation DETERMINISTIC outperformed the two heuristic algorithms we mentioned above, hence we present our results only for DETERMINISTIC and RANDOMIZED. We observed the following:

1. The randomized algorithm for the Minimax rule needs to ask each voter to compare more candidates than in the case of Borda in order to achieve a good approximation ratio. For $m = 50$ candidates, asking each voter to compare $\ell = 8$ of them already gave good results for sufficiently many voters.
2. The deterministic algorithm usually performs better than the randomized one, yet there are distributions (e.g., the one-dimensional Euclidean model) where the quality of winners returned by the deterministic algorithm is much worse than those returned by the randomized algorithm. However, for each distribution that we tested, the randomized algorithm consistently was giving good results when the number n of voters and ℓ (number of candidates to rank) were sufficiently large.

6 Conclusion & Discussion

In this paper we theoretically and experimentally analyzed how well certain election rules can be approximated when we are only given part of the voters' preferences. We investigated how well one can approximate positional scoring rules and the Minimax rule through deterministic and randomized elicitation methods, providing both upper-bounds on the approximation ratio (impossibility results), and providing algorithms (almost) matching these bounds.

We conclude that the randomized approach is usually superior; the exceptions include preference distributions with little or no structure, which rarely appear in practice. For the Borda rule, with hundreds of voters it is usually sufficient to ask each voter to compare two random candidates to achieve a high approximation guarantee. For approximating the Minimax rule, one typically needs more voters and to ask them to compare more candidates—e.g., for 50 candidates, we obtained high approximation guarantees for the Minimax rule only when we set the number of voters to around a thousand and $\ell = 8$.

Finally, let us provide a short discussion on our methodology and possible future research directions that stem from our work. In this project we focused on how well the scores of certain candidates can be approximated while having only partial information. Notice that a good approximation guarantee of an algorithm does not imply that the algorithm will return the winner with high probability. E.g., when there are two comparably good candidates our rules will pick either of them with almost the same probability. Though, we can guarantee that the score of the picked candidate will be close (with high probability) to the score of the true winner. This means that if the winner has a sufficiently high margin of victory, the rules can identify such a winner with high probability. Our model is also more flexible in the following sense: it cannot only be used to find a good candidate as a winner, but also to return a good ranking or to estimate the scores.

There are several ways in which our work can be extended. First, one could look at other natural voting rules that are based on the concept of a score/distance (e.g., Bucklin or Kemeny). Further, instead of using the standard concept of approximation, one could consider other means for evaluating “closeness” of the outcomes. One can for example measure the number of swaps in the preference profile that are required to make c a winner. Finally, it is natural to ask whether one can provide stronger guarantees in certain special cases of the model of preference aggregation (e.g., assuming certain models of ground truth).

Acknowledgments

Piotr Skowron was supported by a postdoctoral fellowship of the Alexander von Humboldt Foundation, Bonn, Germany and by the Foundation for Polish Science within the Homing programme (Project title: “Normative Comparison of Multiwinner Election Rules”).

References

Anshelevich, E., and Postl, J. 2017. Randomized social choice functions under metric preferences. *Journal of Arti-*

ficial Intelligence Research 58:797–827.

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

Ayadi, M.; Ben Amor, N.; and Lang, J. 2019. Approximating voting rules from truncated ballots. Technical report. Working Paper.

Ayadi, M.; Ben Amor, N.; Lang, J.; and Peters, D. 2019. Single transferable vote: Incomplete knowledge and communication issues. In *Proceedings of the 18th International Conference on Autonomous Agents and Multiagent Systems*, 1288–1296.

Benade, G.; Nath, S.; Procaccia, A.; and Shah, N. 2017. Preference elicitation for participatory budgeting. In *Proceedings of the 31st AAAI Conference on Artificial Intelligence*, 376–382.

Benadé, G.; Procaccia, A.; and Qiao, M. 2019. Low-distortion social welfare functions. To appear.

Bentert, M., and Skowron, P. 2019. Comparing election methods where each voter ranks only few candidates. Technical Report arXiv:1901.10848 [cs.DS], arXiv.org.

Boutilier, C., and Rosenschein, J. 2015. 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.

Boutilier, C.; Caragiannis, I.; Haber, S.; Lu, T.; Procaccia, A.; and Sheffet, O. 2012. Optimal social choice functions: A utilitarian view. In *Proceedings of the 13th ACM Conference on Electronic Commerce*, 197–214.

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.

Caragiannis, I.; Chatzigeorgiou, X.; Krimpas, G.; and Voudouris, A. 2019. Optimizing positional scoring rules for rank aggregation. *Artificial Intelligence* 267:58–77.

Caragiannis, I.; Krimpas, G.; and Voudouris, A. 2015. Aggregating partial rankings with applications to peer grading in massive online open courses. In *Proceedings of the 14th International Conference on Autonomous Agents and Multiagent Systems*, 675–683.

Caragiannis, I.; Krimpas, G.; and Voudouris, A. 2016. How effective can simple ordinal peer grading be? In *Proceedings of the 17th ACM Conference on Economics and Computation*, 323–340.

Doucette, J. 2014. Imputation, social choice, and partial preferences. In *Proceedings of AAAI 2014*, 3069–3070. AAAI Press.

Doucette, J. 2015. Resolving elections with partial preferences using imputation. In *Proceedings of Canadian AI 2015*, volume 9091 of *Lecture Notes in Computer Science*, 325–329. Springer.

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

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

Gross, S.; Anshelevich, E.; and Xia, L. 2017. Vote until two of you agree: Mechanisms with small distortion and sample complexity. In *Proceedings of the 31st Conference on Artificial Intelligence*.

Hansen, J. 2016. The random pairs voting rule: Introduction and evaluation with a large dataset. In *Proceedings of COMSOC-16*.

Lu, T., and Boutilier, C. 2011. Robust approximation and incremental elicitation in voting protocols. In *Proceedings of the 22nd International Joint Conference on Artificial Intelligence*, 287–293.

Mallows, C. 1957. Non-null ranking models. *Biometrika* 44:114–130.

Miller, G. 1956. The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review* 63(2):81–97.

Pierczynski, G., and Skowron, P. 2019. Approval-based elections and distortion of voting rules. In *Proceedings of the 28th International Joint Conference on Artificial Intelligence*. To appear.

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.

Raman, K., and Joachims, T. 2014. Methods for ordinal peer grading. In *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 1037–1046.

Xia, L., and Conitzer, V. 2011. A maximum likelihood approach towards aggregating partial orders. In *Proceedings of the 22nd International Joint Conference on Artificial Intelligence*, 446–451.

Zwicker, W. 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.