Voting Rules Require Communication

Jakob de Raaij **Itai Shapira**

Abstract

Voting rules are fundamental tools in collective decision-making, yet their practical implementation often faces significant challenges due to the cognitive and logistical burdens of collecting and processing complete preference profiles. This work investigates the communication complexity of approximating the Borda voting rule, focusing on reducing the information burden required to compute approximate scores for all candidates simultaneously within a small additive margin of error, ε. We establish tight bounds on the deterministic communication complexity of approximating the Borda scores, proving an upper bound of $\mathcal{O}(nm \log \frac{1}{\varepsilon})$ and lower bounds of $\Omega(nm)$ for general ε and $\Omega(nm \log m)$ for sufficiently small ε, where n is the number of voters and m is the number of candidates. Furthermore, we demonstrate that randomized protocols can significantly reduce communication complexity. We give a randomized protocol using at most $\mathcal{O}(\frac{1}{\varepsilon^2} m(\log m + \log \frac{1}{\delta}))$ bits of communication that approximates the Borda scores up to ε for all candidates simultaneously with probability at least $1 - \delta$. Our results highlight the trade-offs between communication efficiency and outcome accuracy, offering insights into the practical implementation of voting rules under informational constraints.

1 Introduction

In collective decision-making, voting rules serve to aggregate individual preferences into a societal outcome. These rules dictate how votes are transformed into decisions or rankings, striving to balance fairness, efficiency, and representativeness [12]. Examples range from simple plurality voting, where each individual has one vote and the alternative with the most votes wins, to more sophisticated methods like the Borda count, which assigns scores to alternatives based on individuals' rankings over them [6], and the Bradley-Terry model, which ranks candidates using pairwise comparisons between them [4]. These methods provide a foundation for aggregating preferences in diverse contexts, from elections to recommendation systems.

Despite the theoretical appeal of more sophisticated voting rules, their practical application faces substantial hurdles. As the number of voters and candidates increases, collecting and processing complete preference information becomes unfeasible [5]. Expecting participants to rank all alternatives imposes significant cognitive and logistical burdens, particularly in scenarios with a large number of options. Moreover, logistical constraints make it challenging to efficiently manage and process such extensive data [7]. These difficulties often hinder the adoption of voting rules that rely on detailed preference aggregation in real-world settings.

To tackle these challenges, communication complexity [10] offers a natural lens through which to evaluate voting rules. At its core, communication complexity formalizes the information burden on voters through a system of queries. Each query addresses a specific preference question, such as "Do you prefer A over B?" or "Is A your favorite option?" [7]. Summing the total number of queries required across all voters quantifies the communication burden of a voting rule, providing a principled basis for systematic comparisons. High communication complexity demonstrates that a voting rule is impractical in real-world settings due to the excessive burden it places on voters. By identifying voting rules that are infeasible under informational constraints, communication complexity helps focus attention on more efficient alternatives.[1] This framework highlights the trade-off between practical constraints and the quality of outcomes, offering valuable insights into the efficiency of voting rules.

In this work, we study the communication complexity of *approximating* the Borda voting rule (see Definition 2.3). By approximation, we mean computing scores for all alternatives that deviate from the scores found in the exact rule by a small additive margin ε, hopefully reducing the communication required. This approach is particularly compelling in scenarios where participants are willing to accept rules that produce outcomes slightly less representative of their preferences in exchange for a significant reduction in communication burden [5]. For example, in cases where Borda is desirable but impractical due to high communication requirements, approximating Borda may offer a better balance – reducing the burden significantly while retaining much of its quality. This approach is often preferable to both exact Borda, which is communication-intensive, and simpler rules like plurality, which sacrifice outcome quality for minimal communication.

We address the following key questions:

> *What are the theoretical limits on the amount of information required to compute an ε-approximation of scores for Borda count?*

and:

> *Can randomized methods surpass the non-random lower bounds, reducing communication complexity further while providing accurate approximations with high probability?*

By studying the communication complexity of approximate voting rules, we aim to bridge the gap between theoretical social choice models and their practical implementation.

Our Results. In Section 2, we formalize the problem by introducing voting rules, the Borda rule with scores normalized to $[0, 1]$, and communication complexity. We also define the notion of approximate voting rules, highlighting the trade-off between accuracy and communication efficiency, and extend the relevant communication complexity results to our approximation setting. In Section 3, we analyze deterministic protocols and establish both upper and lower bounds. We prove that approximating the Borda scores of all m candidates within an error of ϵ across n voters requires at most $\mathcal{O}\left(nm \log \frac{1}{\epsilon}\right)$ bits of communication (Theorem 3.3). Complementing this, we show that any deterministic protocol must use at least $\Omega(nm \log m)$ bits for small ϵ (Theorem 3.5), matching the complexity of exact computation. For larger ε, we show a lower bound of $\Omega(nm)$ bit for any deterministic protocol (Theorem 3.8), translating a result from [15] to our notion of approximation. In Section 4, we demonstrate that allowing randomness significantly reduces communication complexity. Specifically, we show that approximating Borda scores within an error of ϵ with high probability requires only $\mathcal{O}\left(\frac{1}{\epsilon^2} m\big(\log m + \log \frac{1}{\delta}\big)\right)$ bits (Theorem 4.2).

Related Research. Our project contributes to an ongoing line of research on the outcomes of voting rules given only partial information about voters' preferences. Conitzer and Sandholm [5] studied the communication complexity of voting rules without approximation. For many well-known voting rules, they proved matching upper and lower bounds on the communication complexity of computing the winner deterministically or non-deterministically. In particular, they showed that any deterministic or non-deterministic voting protocol that determines the winner of the Borda voting rule requires $\Omega(nm \log m)$ bits of communication—enough for each voter to communicate their entire ranking. In several ways, our project extends their results by introducing the notion of approximation to their (mostly dismal) results. Service and Adams [15] studied the problem of approximating scoring rules using a different notion of approximation than ours, focusing on a multiplicative form. Specifically, they considered the problem of finding an alternative a such that its actual score $s(a)$ is in $[(1-\epsilon)s(w), s(w)]$, where w is the winner, i.e., the alternative with the highest score. For all $\epsilon \in (0, 1)$, they presented a protocol with $\mathcal{O}(nm \log \frac{1}{\epsilon})$ communication that finds such an alternative. They also showed that for $\epsilon \leq \frac{1}{4}$, any deterministic communication protocol in their notion of approximation requires $\mathcal{O}(nm)$ bits of communication. Halpern et al. [8] characterized

[1] While low communication complexity suggests feasibility, it does not guarantee it, as voters may still face significant cognitive or logistical challenges in generating their responses. For example, a single query to a voter, such as "Is A your 42nd favorite choice?" might still require essentially fully ranking all alternatives to compute a response.

positional scoring rules computable with limited queries and established tight bounds for deterministic protocols. Their work is relevant to our study of approximating the Borda rule, as it highlights the challenges of balancing communication efficiency with computation under limited information. Heckel et al. [9] studied the randomized communication complexity of determining the k alternatives with the highest scores, allowing for a small number of wrongly included alternatives, with high probability. Their results critically depend on the size of the gaps between scores of alternatives.

2 Model and Preliminaries

In this section, we present the formal framework for our analysis, including definitions of voting rules, the Borda rule, approximation concepts, and the communication complexity framework used to evaluate the information requirements of computing and approximating these rules.

2.1 Voting Rules

Let V be a set of n voters and let $A = \{a_i\}_{i\in[m]}$ be a set of m alternatives, where each alternative corresponds to a candidate in a voting setting. Each voter $i \in V$ holds a private ranking σ_i over the alternatives, specifying their preferences. A ranking is a bijection $\sigma : [m] \to A$, where $\sigma(j)$ denotes the candidate ranked in the j-th position. For two candidates $a, b \in A$ and a ranking σ, we write $a \succ_\sigma b$ if $\sigma^{-1}(a) < \sigma^{-1}(b)$, indicating that a is preferred to b under σ. Let $L(A)$ denote the set of all $m!$ possible rankings over the candidates in A. A preference profile is a tuple $\sigma = (\sigma_1, \ldots, \sigma_n) \in L(A)^n$, where $\sigma_i \in L(A)$ is the ranking of voter i. A voting rule $\mathcal{V}$ is a function $\mathcal{V} : L(A)^n \to L(A)$, mapping a profile of preferences $\sigma \in L(A)^n$ over a set of alternatives A to a ranking $\tau \in L(A)$.

A scoring rule is a specific type of voting rule that assigns numerical scores to candidates based on their positions in the rankings of the voters. Formally:

Definition 2.1 (Scoring Rule)**.** *Given a vector $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_m)$, where $\alpha_1 \geq \alpha_2 \geq \cdots \geq \alpha_m \geq 0$, the* scoring rule *induced by α, denoted s^α, is defined as follows: for a profile of preferences $\sigma = (\sigma_1, \ldots, \sigma_n) \in L(A)^n$, the* score *of a candidate $a \in A$ is:*

$$s^\alpha(\sigma)(a) = \frac{1}{n}\sum_{i=1}^{n} \alpha_{\sigma_i^{-1}(a)},$$

The rule determines a final ranking $\tau \in L(A)$ by sorting the candidates in descending order of their scores $s^\alpha(a)$, breaking ties arbitrarily if necessary.

Remark 2.2. *With slight abuse of notation, we will identify s^α both as a scoring function that assigns a score to each candidate and as a voting function that outputs a ranking of candidates given a profile. When clear from the context, we will omit explicitly mentioning α or σ and simply write $s(a)$ to denote the score of a candidate a.*

An important example of a scoring rule is **plurality**, which is induced by the vector $\alpha = (1, 0, 0, \ldots, 0)$. Under this rule, each voter assigns a score of 1 to their top-ranked candidate and 0 to all others.

Once we associate the output ranking with the richer scoring vector, which assigns a numerical value to each rank, we obtain a natural notion of ”approximation.” This allows us to compute scores that deviate slightly from their exact values while preserving the overall structure and essence of the scoring rule.

A key scoring rule, which we will focus on in this project, is the Borda rule:

Definition 2.3 (Borda Rule)**.** *The (normalized)* ***Borda scoring rule*** *s^* is the scoring rule induced by*

$$\alpha = \frac{1}{n(m-1)}(m-1, m-2, \ldots, 0)$$

The Borda rule assigns a normalized score $s^*(a) \in [0, 1]$ to each candidate $a \in A$. Each voter awards their most preferred candidate $\frac{1}{n}$ point, the second-most preferred $\frac{1}{n} - \frac{1}{n(m-1)}$ points, and so on, until the least preferred gets 0 points. We note that this differs from the most common definition of Borda by dividing all points by $\frac{1}{n(m-1)}$ to make sure that as the number of voters or alternatives increases, the Borda score stays less than one. Moreover, this normalization allows us to interpret $s(a)$ as the probability that a voter, chosen uniformly at random, prefers a over another alternative chosen uniformly at random (see Lemma 4.1).

The Borda scoring rule, as we will show below, is one of the most communication-intensive scoring rules (details to follow). However, its theoretical properties make it a highly attractive choice in practice [12]. Simplified versions of the Borda rule are widely adopted in real-world applications. For instance, it is used to select the Most Valuable Player (MVP) in major American sports leagues [14], determine the winners of the Ballon d'Or (FIFA Player of the Year) awards [1], and in the Eurovision Song Contest [2]. A generalization of the Borda rule, the Bradley-Terry model (see [16] for the connection), is extensively used to align AI systems with human feedback [4, 11].

2.2 Approximation

Below we define our notion of additive approximation.

Definition 2.4 (ϵ-Approximation)**.** *For a given voting profile σ, a scoring vector s' is said to be an ϵ-approximate to the outcome $s(\sigma)$ of the scoring rule if*

$$\|s' - s(\sigma)\|_\infty \leq \epsilon,$$

where $\|\cdot\|_\infty$ denotes the maximum absolute difference across all candidates. We denote the set of all vectors that are a ε-approximation to $s(\sigma)$ as

$$B_\varepsilon(s(\sigma)) = \{s' \in \mathbb{R}^n : \|s' - s(\sigma)\|_\infty \leq \epsilon\}.$$

Notice that in this definition, we require the approximate scoring vector to approximate all alternatives simultaneously.

For scoring rules like Borda, the discrete nature of scores necessitates careful consideration.

Remark 2.5. *The Borda score of any candidate changes in discrete units of $u = \frac{1}{n(m-1)}$, corresponding to a single rank swap in one voter's ranking.*

Consequently, throughout the presentation of our results, it will at times be more natural to express the approximation error ϵ in terms of these units u.

2.3 Communication Complexity

We briefly review the model of communication complexity and the concept of a fooling set, which is a fundamental tool for proving lower bound. For a more throughout review, we refer to [3],[10], and [13]. We extend these definitions to an approximation setting and prove Theorem 2.7 which will be essential for all lower bounds in this paper.

In our setting, n voters each hold private rankings $\sigma_i \in L(A)$ over a set of candidates A. Together, the voters aim to approximate a function $s : L(A)^n \to \mathbb{R}^m$. A *protocol* specifies the communication process by which voters share information about their rankings. We consider the blackboard model, where in each step a voter communicates a single bit to all other voters (i.e., writes it on a public blackboard), based on their ranking and the information exchanged so far. Upon termination, the protocol returns a score vector based on the entire communication. We say a protocol is ε-approximates a scoring function s if for all $\sigma = (\sigma_1, ..., \sigma_n)$, the returned score vector s' is an ε-approximation to $s(\sigma)$, i.e. $s' \in B_\varepsilon(s(\sigma))$.

In a *deterministic* protocol, the next voter to communicate is a function of the communication history thus far, and this voter's communication is a function of their private input (their ranking) and the communication history thus far. The deterministic communication complexity of ε-approximating a scoring rule s is the worst-case number of bits exchanged over all possible input profiles in the most efficient ε-approximate protocol.

A common technique for proving lower bounds on deterministic communication complexity is giving a *fooling set*, a collection of input profiles that exhibit specific combinatorial properties. Intuitively, a fooling set demonstrates that any protocol attempting to compute s must distinguish between a large number of distinct inputs, inherently requiring significant communication.

Definition 2.6 (t-mixing fooling set for ϵ-approximation)**.** *A t-mixing* fooling set *for ε-approximating a scoring rule s is a collection of input profiles*

$$\mathcal{F} = \left\{\sigma^j\right\}_j, \sigma^j = (\sigma_i^j)_{i\in[n]}$$

such that for any $M \subseteq [|\mathcal{F}|]$, $|M| = t$ there exists a mixed profile $\sigma^M = (\sigma_i^{r_i})_{i\in[n]}$, where $r_i \in M$, such that $B_\varepsilon(s(\sigma^M)) \cap B_\varepsilon(s(\sigma^j)) = \emptyset$ for some $j \in M$.

The essence of the fooling set argument is that we can *mix* components of any t different profiles in $\mathcal{F}$ to change the resulting scoring vector such that there is no vector that is both an ε-approximating to this new scoring vector and any of the scoring vectors of a profile in the fooling set. This intuition is formalized in the following theorem. The proof essentially follows the structure of all fooling set lower bound arguments (see f.e. [10] or [15]), adapted to our approximation setting.

Theorem 2.7. *Assume there exists a t-mixing fooling set $\mathcal{F}$ for ε-approximation for a scoring rule s. Then the deterministic communication complexity of s is at least $\log_2 \frac{|\mathcal{F}|}{t-1}$.*

Proof. Let $\mathcal{F} = \{\sigma^1, ..., \sigma^{|\mathcal{F}|}\}$. Assume towards a contradiction that there exists a deterministic communication protocol that ε-approximates s and communicates less than $\log_2 \frac{|\mathcal{F}|}{t-1}$ bits in the worst case. Thus, there exist less than $\frac{|\mathcal{F}|}{t-1}$ possible strings of bits communicated. By the pigeonhole principle, there exists a set $M \subseteq [|\mathcal{F}|]$ of size at least t such that the same communication occurs in the protocol for all voting profiles σ^j with $j \in M$. Thus, for all $\sigma^j, j \in M$, the protocol returns the same scoring vector s'. Since the protocol ε-approximates s, we know that $s' \in B_\varepsilon(s(\sigma^j))$ for all $j \in M$.

However, by the definition of $\mathcal{F}$, we know that there exists a voting profile $\sigma^M = (\sigma_i^{r_i})_{i\in[n]}$, where $r_i \in M$, such that $B_\varepsilon(s(\sigma^M)) \cap B_\varepsilon(s(\sigma^j)) = \emptyset$ for some $\sigma^j \in \mathcal{F}$. Since each voter's communication at a given time only depends only on their own ranking and the prior communication, the same communication as for all $\sigma^j, j \in M$ occurs also on σ^M. Consequently, the protocol returns s' on voting profile σ^M. We know that for some $j \in M$, there does not exist a vector that ε-approximates both σ^j and σ^M, i.e., $B_\varepsilon(s(\sigma^M)) \cap B_\varepsilon(s(\sigma^j)) = \emptyset$. Since s' is an ε-approximation to $s(\sigma^j)$ for all $j \in M$, i.e., $s' \in B_\varepsilon(s(\sigma^j))$, it follows that $s' \notin B_\varepsilon(s(\sigma^M))$, a contradiction to the protocol being an ε-approximation to s. □

Two other commonly used notions of communication complexity are randomized and nondeterministic communication complexity.

In a *randomized* protocol, both the function choosing the next voter to communicate and each voter's function choosing their communication have access to random bits. Whether these random bits are private or public does not change the bound we obtain in this paper [10]. The randomized communication complexity of ε-approximating a scoring rule s is the worst-case number of bits exchanged over all possible input profiles in the most efficient protocol that for every input profile is ε-approximate with probability at least $2/3$ over the protocol's randomness.

There are multiple (mostly equivalent) ways to define nondeterministic communication complexity. One way to think of *nondeterministic* protocols is as allowing for the function determining each voter's communication to be nondeterministic. However, it seems to be more common and convenient to define them using an omniscient prover [13] that writes a proof, i.e., bits $c \in \{0, 1\}^*$, on the commonly seen blackboard before the voters start communicating. Out of space constraints, we do not formalize nondeterministic communication complexity further but note at this point that fooling set lower bounds generally also apply to nondeterministic protocols and that Theorem 2.7 can be extended to this case with a formal definition of nondeterminism, thus extending the lower bound results in this paper for deterministic protocols to nondeterministic protocols as well.

3 Non-Random Communication Complexity

3.1 Upper Bounds

A trivial upper bound on the communication complexity can be obtained by having each voter communicate their entire ranking:

Theorem 3.1. *The deterministic communication complexity of any voting rule is* $\mathcal{O}(nm \log m)$.

Proof. Consider the protocol where each voter communicates their ranking by sending the position $\sigma_i^{-1}(a)$ of each candidate $a \in A$, requiring $m \cdot \log m$ bits per voter and thus $n \cdot m \log m = \mathcal{O}(nm \log m)$ for n voters in total. □

This result is tight, as shown by the following matching lower bounds on even the nondeterministic communication complexity:

Theorem 3.2. *[Special case of Theorem 3.5 for* $\varepsilon = 0$*] The nondeterministic communication complexity of computing the Borda rule is* $\Omega(nm \log m)$

This result was shown by Conitzer and Sandholm [5] for deciding the winner under the Borda rule. Together, these two results demonstrate that the Borda rule is asymptotically as communication-intensive as a scoring rule can be, as its communication complexity is $\Theta(nm \log m)$. This reflects the "down to the wire" nature of the rule, where already a single rank swap by one voter can change the outcome. Consequently, any deterministic protocol must asymptotically extract as much information as the full rankings from every voter, as even minor details are critical.

Allowing an ϵ-margin simplifies the problem. Voters no longer need to communicate precise rankings but can instead communicate coarser score intervals. By grouping candidates into bins of size proportional to ϵ, voters only need to indicate the bin index for each candidate, reducing communication requirements:

Theorem 3.3 (Non-Random Upper Bound)**.** *The deterministic communication complexity of simultaneously approximating the Borda score* $s^*(a)$ *for all alternatives* $a \in A$ *within an error of* ϵ *is* $\mathcal{O}\left(nm \log \frac{1}{\epsilon}\right)$.

Proof. Let $k \in [m]$ be a parameter to be chosen later. Consider the following protocol: Each voter $i \in V$ partitions their ranking into $\left\lceil \frac{m}{k} \right\rceil$ bins, with each bin covering up to k positions. Specifically, the first $\left\lfloor \frac{m}{k} \right\rfloor$ bins each contain exactly k positions, and the final bin contains the remaining positions if m is not divisible by k.[2]

Each voter i then communicates only the bin index $\ell_i(a)$ for each candidate a, so that $\sigma_i^{-1}(a)$ (the exact position of a in i's ranking) satisfies:

$$(\ell_i(a) - 1)k + 1 \leq \sigma_i^{-1}(a) \leq \ell_i(a)k.$$

The communication cost is $m \log \left(\frac{m}{k}\right)$ bits per voter, as each bin index $\ell_i(a)$ requires $\log \left(\frac{m}{k}\right)$ bits. The true Borda score $s^*(a)$ for candidate a is given by: $s^*(a) = (n(m-1))^{-1} \sum_{i=1}^{n} (m - \sigma_i^{-1}(a))$. The highest position of a within bin $\ell_i(a)$ is $(\ell_i(a) - 1)k + 1$, giving the upper bound:

$$s^*(a) \leq \frac{1}{n(m-1)} \sum_{i=1}^{n} \left(m - ((\ell_i(a) - 1)k + 1)\right). \tag{1}$$

The lowest position of a within bin $\ell_i(a)$ is $\ell_i(a)k$, giving the lower bound:

$$s^*(a) \geq \frac{1}{n(m-1)} \sum_{i=1}^{n} \left(m - \ell_i(a)k\right). \tag{2}$$

The difference between the upper and lower bounds on $s^*(a)$ gives the error in our approximation:

$$\Delta(s^*(a)) = \frac{1}{n(m-1)} \sum_{i=1}^{n} (m - (\ell_i(a) - 1)k - 1 - (m - \ell_i(a)k)) = \frac{k-1}{m-1} < \frac{k}{m-1}.$$

With $k < \epsilon(m-1)$, each voter communicates $m \log \left(\frac{m}{k}\right) = m \log \frac{1}{\epsilon}$ bits, resulting in a total communication complexity of $\mathcal{O}\left(nm \log \frac{1}{\epsilon}\right)$. □

[2]For example, if $m = 5$ and $k = 2$: Bin 1 covers positions 1–2, Bin 2 covers positions 3–4, and Bin 3 covers position 5.

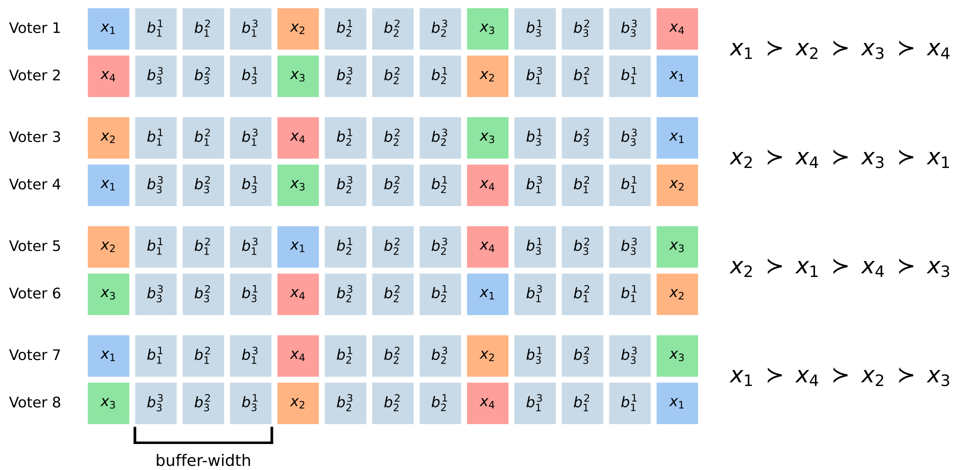

Figure 1: Example for illustration purposes of one profile in the proof of Theorem 3.5 with $m = 13$, $n = 8$ and 3 slack candidates. The slack candidates, denoted as b_i^j, are used as buffers inserted between primary candidates x_i. Each voter pair corresponds to a permutation of the set of primary candidates X.

Two remarks are in order. First, this upper bound helps us quantify the trade-off between approximation error and communication complexity. In the extreme case where $\epsilon = \mathcal{O}(\frac{1}{mn})$, the approximation is indistinguishable from exact computation, as the margin of error is smaller than the smallest possible difference in candidate scores. In this case, the communication complexity of the algorithm is $\Theta(nm \log nm)$. As ϵ increases, broader score intervals enable communication gains. For instance, with $\epsilon = 1/2$, the protocol effectively classifies candidates as "up" (top half) or "down" (bottom half), yielding a communication complexity of $\mathcal{O}(nm)$. This is intuitively optimal due to the "down to the wire" argument discussed earlier.

Second, the key property of the Borda rule leveraged in the proof is that under- or over-estimating a candidate's position in a single voter's ranking by k positions results in a limited score deviation. This distinguishes scoring rules like Borda from rules such as plurality, where small misclassifications can cause significant score increases (e.g., a large $1/n$ increment in the normalized setting). More generally, for a scoring rule induced by weights $\alpha = (\alpha_1, \ldots, \alpha_m)$, a similar argument applies: the best score for a candidate a (Equation (1)) is $n^{-1} \sum_{i=1}^{n} \alpha_{(\ell_i - 1)k}$, and the worst score (Equation (2)) is $n^{-1} \sum_{i=1}^{n} \alpha_{\ell_i k}$. Their difference is $n^{-1} \sum_{i=1}^{n} (\alpha_{(\ell_i - 1)k} - \alpha_{\ell_i k})$, thus:

Corollary 3.4. *For a scoring rule induced by $\alpha = (\alpha_1, \ldots, \alpha_m)$, for any $\epsilon > 0$, if there exists some k such that*

$$\max_{i \leq m-k} \alpha_i - \alpha_{i+k} < \epsilon,$$

then the communication complexity of ϵ-approximating the scoring rule is at most $\mathcal{O}\left(nm \log \frac{m}{k}\right)$.

Having established an upper bound on the communication complexity of approximating Borda scores, we now turn to proving a lower bound.

3.2 Lower Bound

We provide two different lower-bound arguments which are tight for different values of ε. First, we essentially show that for $\varepsilon = \mathcal{O}(\frac{1}{nm})$, no reduction in communication is possible. I.e., to compute the Borda scores up to a constant number of units (u, see Remark 2.5), asymptotically as much information as the full rankings are required. Our fooling set construction is loosely based on the fooling set that Conitzer and Sandholm used in [5] to prove Theorem 3.2. We then show that even as ϵ grows, as long as $\epsilon \leq \frac{1}{4} - c$ for any arbitrarily small constant $c > 0$, $\Omega(nm)$ communication is required. The fooling set construction used in this second proof is very similar to a fooling set used in [15]. We note that their notion of approximation significantly varies from our notion and that it thus is non-trivial that their arguments translate.

Theorem 3.5 (Non-Random Lower Bound)**.** *Let $n = 2n'$ voters, m candidates, and $\epsilon = ku = \frac{k}{n(m-1)}$ for some $k \in \{0, \dots, m-1\}$. The nondeterministic communication complexity of approximating the score of all candidates within an error of ϵ is $\Omega\left(n\frac{m}{k}\log\left(\frac{m}{k}\right)\right)$.*

Proof. We construct a large 2-mixing fooling set of profiles $\mathcal{F}$, where in each profile $\sigma \in \mathcal{F}$, each candidate has a balanced Borda score of $\frac{1}{2}$. However, for any two profiles $\sigma, \tau \in \mathcal{F}$, we can mix the rankings of voters from σ and τ into a ranking υ such that the score of at least one candidate in υ exceeds $\frac{1}{2}$ by at least 2ϵ, so that no scoring vector can ε-approximate all of $s^*(\sigma), s^*(\tau)$, and $s^*(\upsilon)$ simultaneously. That is, $B_\varepsilon(s^*(\upsilon)) \cap B_\varepsilon(\frac{1}{2}\mathbf{1}) = \emptyset$, where $\mathbf{1}$ is the m-dimensional 1-vector. By Theorem 2.7, the existence of a sufficientlly large such $\mathcal{F}$ proves this theorem.

We partition the set of candidates into two types: the "primary" candidates X and the "slack" candidates S. Intuitively, slack candidates are inserted between each pair of primary candidates to act as buffers. These buffers ensure that mixing creates a change in score proportional to the buffer size B, amplifying the difference between the previous score and the new mixed score.

Let $A = X \cup S$, where $X = \{x_1, \dots, x_{|X|}\}$ and $S = \{b_i^j\}_{i \in [B], j \in [|X|-1]}$. We refer to B as the buffer width, which is the number of slack candidates placed between each pair of primary candidates in the rankings. The total number of candidates satisfies: $(B+1)|X| - B = m$. We will determine $|X|$ and B later to optimize the size of $\mathcal{F}$.

Next, we define a (large) class of profiles $\mathcal{F}$ and prove it is a fooling set. Let $\boldsymbol{\pi} = (\pi_1, \dots, \pi_{n'})$ denote a sequence of n' permutations over X. The profile associated with $\boldsymbol{\pi}$, denoted by $P^{\boldsymbol{\pi}} = (\sigma_1, \dots, \sigma_n)$, is defined as follows, for $i = 1, ..., n'$ (see Figure 1):

$$\sigma_{2i} : x_{\pi_i(1)} \succ b_1^1 \succ \cdots \succ b_B^1 \succ x_{\pi_i(2)} \succ b_1^2 \succ \cdots \succ b_B^2 \succ \cdots \succ x_{\pi_i(|X|)},$$

$$\sigma_{2i-1} : x_{\pi_i(|X|)} \succ b_B^{|X|} \succ \cdots \succ b_1^{|X|} \succ x_{\pi_i(|X|-1)} \succ \cdots \succ b_1^1 \succ x_{\pi_i(1)}.$$

The set $\mathcal{F}$ consists of all profiles $P^{\boldsymbol{\pi}}$ for all sequences $\boldsymbol{\pi}$ of n' permutations over X. Thus, the size of $\mathcal{F}$ is $(|X|!)^{n'}$. For any $\boldsymbol{\pi}$, the Borda score of each $x \in X$ under $P^{\boldsymbol{\pi}}$ is $\frac{1}{2}$, as the contributions from voters $2i$ and $2i-1$ even out.

Next, consider two profiles $P^{\boldsymbol{\pi}}, P^{\boldsymbol{\tau}} \in \mathcal{F}$, where $\boldsymbol{\pi} \neq \boldsymbol{\tau}$. Since $\boldsymbol{\pi} \neq \boldsymbol{\tau}$, there exists at least one index $i \in [|X|]$ such that $\pi_i \neq \tau_i$. This implies that for at least one candidate $x \in X$, $\pi_i^{-1}(x) \neq \tau_i^{-1}(x)$. Without loss of generality, assume $\pi_i^{-1}(x) < \tau_i^{-1}(x)$. Construct a mixed profile $P^{(\boldsymbol{\pi},\boldsymbol{\tau})} = (\sigma_1, \dots, \sigma_{2n'})$ as follows (see Figure 2):

$$\sigma_j = \begin{cases} \sigma_j^{\boldsymbol{\pi}}, & \text{if } j \neq 2i-1, \\ \sigma_{2i-1}^{\boldsymbol{\tau}}, & \text{if } j = 2i-1. \end{cases}$$

That is, for voters $2i-1$ and $2i$, we take the voter from the ranking in which x has a higher Borda contribution, ensuring that x's position improves in the mixed profile. In $P^{(\boldsymbol{\pi},\boldsymbol{\tau})}$, the score of x increases because v_{2i-1} now ranks x higher than in $P^{\boldsymbol{\pi}}$. The increase in x's score is given by $\Delta s(x) n(m-1) \geq B+1$. To ensure $s(x) \geq \frac{1}{2} + 2\epsilon$, it suffices that $\frac{B+1}{n(m-1)} \geq 2\epsilon$, which implies $B \geq 2k-1$, where $\epsilon = \frac{k}{n(m-1)}$. Finally, the size of the fooling set is determined by $|X|$: $|X| = \mathcal{O}(m/k)$. The size of $\mathcal{F}$ is thus: $|\mathcal{F}| = (|X|!)^{n'} = \left(\frac{m}{k}!\right)^{n'}$. Therefore, the communication complexity lower bound becomes:

$$\log(|\mathcal{F}|) = O\left(n'|X| \cdot \log(|X|)\right) = \Omega\left(n\frac{m}{k}\log\left(\frac{m}{k}\right)\right).$$

□

Remark 3.6. *For $\varepsilon = \mathcal{O}(\frac{1}{mn})$ we get that $k = \mathcal{O}(1)$, in which case the lower bound becomes $\Omega(nm\log m)$, which matches the upper bound.*

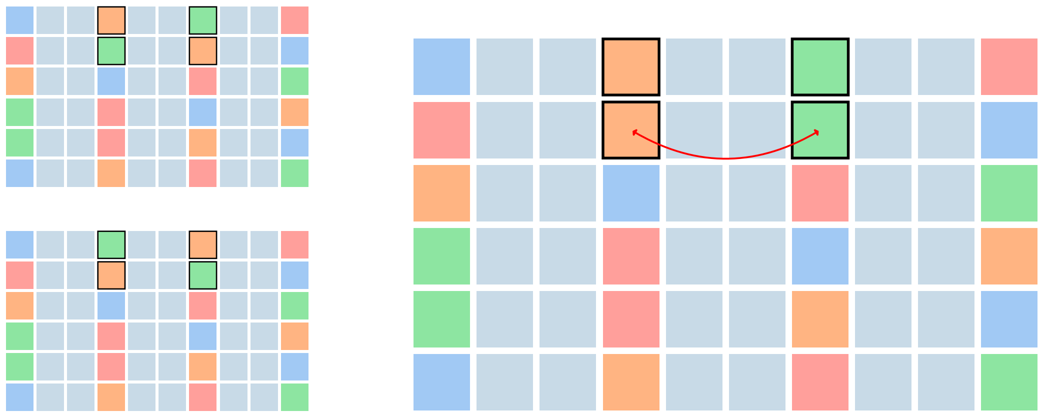

Figure 2: Mixing profiles strategy: The left column shows two profiles that differ in at least one pair of voters, while the right column illustrates the mixing strategy, where the second-row voter in the first pair changes compared to the first profile.

Remark 3.7. *Notice that the above constructions could be applied to general scoring rules. The only property we used about Borda is the fact that moving $B+1$ scores increases the candidate's score "enough".*

We now move on to the second lower bound.

Theorem 3.8. *Let there be $n = 2n'$ voters, m candidates, and $\varepsilon \leq 1/4 - c$ for an arbitrarily small constant $c > 0$. The nondeterministic communication complexity of approximating the score of all candidates within an error of ε is $\Omega(nm)$.*

Similarly to the proof of Theorem 3.5, we construct a large fooling set. The main differences are that the fooling set is t-mixing and that "buffer" candidates are not used between every pair of candidates but only between and upper and lower section of candidates. Because of this change, we no longer consider permutations over the primary candidates but pair those candidates up and consider the ways in which each pair can be split between the upper and lower half. This different use of buffer candidates makes us loose the additional $\log m$ in the bound while allowing us to create fooling sets for larger ε. Moreover, the proof is non-constructive but uses an argument using the probabilistic method to argue the existence of the fooling set. The proof is relying heavily on a fooling set idea and two lemmas proven in [15] for a different notion of approximation.

For space reasons, the full proof can be found in Appendix A.

4 Randomized Communication Complexity

The non-randomized upper and lower bounds established in Section 3 demonstrate that deterministic protocols require substantial communication to compute or approximate the Borda scores, especially for small ϵ. However, introducing randomness can significantly reduce the communication burden while maintaining high accuracy.

4.1 Key Insight: Pairwise Comparisons and Sampling

The key insight enabling this reduction comes from the observation that the Borda score of a candidate a can be computed by only looking at the fraction of voters who prefer a over b for any pair (a, b) of candidates. This transformation reduces the need to access full rankings and shifts the computational focus to pairwise preferences. For a given profile σ, denote by $P^\sigma \in \mathbb{R}^{m \times m}$ the *pairwise majority matrix*, where the entry P^σ_{ab} represents the number of voters who prefer a to b:

$$P^\sigma_{ab} = |\{i \in V : a \succ_{\sigma_i} b\}| .$$

The following Lemma 4.1 shows that the Borda scores can be computed using only information from P^σ. Notice that P^σ is agnostic to n and requires storing only $m \times m$ real numbers, making it significantly more compact than full rankings.

Lemma 4.1. *The Borda score $s^*(a)$ of a candidate $a \in A$ can be computed as:*

$$s^*(a) = \frac{1}{n(m-1)} \sum_{b \in A \setminus \{a\}} P^\sigma_{ab},$$

where P^σ_{ab} is the pairwise majority matrix for the profile σ.

Lemma 4.1 shows that the Borda score $s^*(a)$ is equivalent to the average fraction of voters who prefer a over other candidates. Formally,

$$s^*(a) = \frac{1}{m-1}\left[\frac{\sum_{b \in A \setminus \{a\}} P_{ab}}{n}\right] = \mathbb{E}_{b \sim U(A \setminus \{a\})}\left[\mathbb{E}_{i \sim U(V)}\left[\mathbf{1}_{a \succ_i b}\right]\right],$$

where $U(X)$ denotes the uniform distribution over the set X. This characterization motivates the use of randomized protocols, as we can approximate the pairwise comparisons by sampling voters and applying a simple concentration argument.

Proof of Lemma 4.1.

$$\begin{aligned} n(m-1)s^*(a) &= \sum_{i=1}^{n} \left(m - \sigma_i^{-1}(a)\right) = \sum_{i=1}^{n} \sum_{b \in A, b \neq a} \mathbf{1}[\sigma_i^{-1}(a) < \sigma_i^{-1}(b)] \\ &= \sum_{b \in A, b \neq a} \sum_{i=1}^{n} \mathbf{1}[\sigma_i^{-1}(a) < \sigma_i^{-1}(b)] = \sum_{b \in A, b \neq a} P^\sigma_{ab}. \end{aligned}$$

□

4.2 Randomized Protocol

Building on this insight, we now present a randomized protocol for approximating Borda scores:

Theorem 4.2. *There exists an algorithm with communication complexity*

$$\mathcal{O}\left(\frac{1}{\varepsilon^2}\, m \left(\log m + \log \frac{1}{\delta}\right)\right)$$

that returns approximate scores $\hat{s}(a)$ for all $a \in A$ such that with probability at least $1 - \delta$ for all $a \in A$, $|\hat{s}(a) - s^*(a)| < \varepsilon$.

Proof. We aim to approximate the Borda scores $s(a)$ for all candidates $a \in A$ using a randomized protocol. Let n' be a parameter to be chosen later. For each pair of distinct alternatives $(a, b) \in A^2$, we sample n' voters uniformly and independently at random from V. For each voter i in the sample, we define a random variable: $X^i_{ab} = \mathbf{1}[a \succ_{\sigma_i} b]$, indicating whether voter i prefers a over b. Define the approximate Borda score for a candidate a as:

$$\hat{s}(a) = \frac{1}{(m-1)} \sum_{b \in A \setminus \{a\}} \frac{1}{n'} \sum_{i=1}^{n'} X^i_{ab}.$$

By Lemma 4.1, $\mathbb{E}[\hat{s}(a)] = s^*(a)$. By Hoeffding's inequality,

$$\Pr\left[|\hat{s}(a) - s^*(a)| \geq \epsilon\right] \leq 2(m-1)\exp\left(-2n'\epsilon^2\right).$$

To ensure that all candidates are approximated within ϵ simultaneously, we apply a union bound:

$$\Pr\left[\forall a \in A : |\hat{s}(a) - s(a)| \leq \epsilon\right] \geq 1 - 2m(m-1)\exp\left(-2n'\epsilon^2\right).$$

To achieve a confidence level of $1 - \delta$, we choose n' such that:

$$2m(m-1)\exp\left(-2n'\epsilon^2\right) \leq \delta.$$

This simplifies to:

$$n' = \frac{1}{2\epsilon^2} \log\left(\frac{2m(m-1)}{\delta}\right).$$

Each sampled voter communicates 1 bit per pair (a, b), and there are $m(m-1)/2$ pairs. Thus, the total communication complexity is:

$$\frac{m(m-1)}{2} \cdot n' = \mathcal{O}\left(\frac{m}{\epsilon^2} \log\left(\frac{m}{\delta}\right)\right).$$

□

Remark 4.3. *More generally, for any scoring rule that can be computed directly from the pairwise majority matrix P (also known as a* C2 voting rule*), this randomized protocol can be extended.*

5 Conclusion

Reducing the communication burden in voting is crucial for enabling preference aggregation in cases with many voters and candidates. This work analyzed the communication complexity of approximating the Borda rule, offering tight bounds for deterministic protocols and demonstrating the potential of randomized methods to significantly reduce communication costs while maintaining accuracy. These results highlight the practicality of approximate voting rules, particularly in low-stakes settings where voters are willing to trade some accuracy for efficiency. Future work could explore communication-efficient protocols for other voting rules, examine worst-case versus average-case scenarios, and investigate methods that preserve privacy while minimizing communication. By advancing understanding of these trade-offs, we take a step toward making voting rules more scalable and adaptable for modern applications.

Our randomized protocol is *non-adaptive*, meaning that the protocol runs independently of the realization of sampled voter preferences. However, introducing adaptivity into the sampling process leads to an interesting connection with multi-armed bandit problems. Specifically, for each of the $\binom{m}{2}$ pairs of candidates (a, b), we can identify the pair as an arm, with the mean reward of the arm corresponding to P_{ab}/n, the fraction of voters who prefer a over b. For example, to estimate the Borda score of a candidate a, one could iteratively sample pairwise comparisons between a and other candidates $b \in A \setminus \{a\}$, adapting the number of samples based on the uncertainty in P_{ab}.

References

[1] Explained: how the ballon d'or winner is selected. *The Times*, 2024. URL https://www.thetimes.co.uk/article/explained-ballon-dor-winner-selection-process-nominees-kphx07k2d.

[2] Time for a key change? how we would reform eurovision voting. *Electoral Reform Society*, 2024. URL https://www.electoral-reform.org.uk/time-for-a-key-change-how-we-would-reform-eurovision-voting/.

[3] YAO AC-C. Some complexity questions related to distributed computing. In *Proc. 11th Annual ACM Symposium on Theory of Computing, 1979*, pages 209–213, 1979.

[4] Ralph Allan Bradley and Milton E Terry. Rank analysis of incomplete block designs: I. the method of paired comparisons. *Biometrika*, 39(3/4):324–345, 1952.

[5] Vincent Conitzer and Tuomas Sandholm. Communication complexity of common voting rules. In *Proceedings of the 6th ACM Conference on Electronic Commerce*, EC '05, page 78–87, 2005. doi: 10.1145/1064009.1064018.

[6] Michael Dummett. The borda count and agenda manipulation. *Social Choice and Welfare*, 15 (2):289–296, 1998.

[7] Yannai A Gonczarowski, Noam Nisan, Rafail Ostrovsky, and Will Rosenbaum. A stable marriage requires communication. *Games and Economic Behavior*, 118:626–647, 2019.

[8] Daniel Halpern, Safwan Hossain, and Jamie Tucker-Foltz. Computing voting rules with elicited incomplete votes. *arXiv preprint arXiv:2402.11104*, 2024.

[9] Reinhard Heckel, Max Simchowitz, Kannan Ramchandran, and Martin J. Wainwright. Approximate ranking from pairwise comparisons, 2018. URL https://arxiv.org/abs/1801.01253.

[10] Eyal Kushilevitz and Noam Nisan. *Communication Complexity*. Cambridge University Press, 1996.

[11] Long Ouyang, Jeffrey Wu, Xu Jiang, Diogo Almeida, Carroll Wainwright, Pamela Mishkin, Chong Zhang, Sandhini Agarwal, Katarina Slama, Alex Ray, John Schulman, Jacob Hilton, Fraser Kelton, Luke Miller, Maddie Simens, Amanda Askell, Peter Welinder, Paul Christiano, Jan Leike, and Ryan Lowe. Training language models to follow instructions with human feedback. *Advances in Neural Information Processing Systems*, 35:27730–27744, 2022.

[12] A. D. Procaccia. In F. Brandt, V. Conitzer, U. Endress, J. Lang, and A. D. Procaccia, editors, *Handbook of Computational Social Choice*, chapter 13. Cambridge University Press, 2016.

[13] Tim Roughgarden. Communication complexity (for algorithm designers), 2015. URL https://arxiv.org/abs/1509.06257.

[14] Quantim Schmitz. Why the right player usually wins the mvp, and why voter fatigue might actually be good. *Quantim Schmitz*, 2023.

[15] Travis C. Service and Julie A. Adams. Communication complexity of approximating voting rules. In *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems - Volume 2*, AAMAS '12, page 593–602, Richland, SC, 2012. International Foundation for Autonomous Agents and Multiagent Systems. ISBN 0981738125.

[16] Anand Siththaranjan, Cassidy Laidlaw, and Dylan Hadfield-Menell. Distributional preference learning: Understanding and accounting for hidden context in rlhf. *arXiv preprint arXiv:2312.08358*, 2023.

A Proof of Theorem 3.8

Similar to the proof of Theorem 3.5, we construct a large t-mixing fooling set of profiles $\mathcal{F}$, where in each profile $\sigma \in \mathcal{F}$, each candidate has a balanced Borda score of $\frac{1}{2}$. For any t profiles, we can mix the rankings of voters from these t profiles into a ranking υ such that the score of at least one candidate in υ exceeds $\frac{1}{2}$ by at least 2ϵ, so that no scoring vector can ε-approximate $s^*(\upsilon)$ and $s^*(\frac{1}{2}\mathbf{1})$ simultaneously. That is, $B_\varepsilon(s^*(\upsilon)) \cap B_\varepsilon(\frac{1}{2}\mathbf{1}) = \emptyset$.

To construct $\mathcal{F}$, we again partition the set of candidates into two types: the "primary" candidates X and the "slack" candidates S. However, differently to above, we don't insert slack candidates between any two candidates, but instead split the candidates in an upper and a lower half, separated by the slack candidates.

Let $A = X^0 \cup X^1 \cup S$, where $X^0 = \{x_1^0, \dots, x_\ell^0\}$, $X^1 = \{x_1^1, \dots, x_\ell^1\}$, and $S = \{b_1, ..., b_{m-2\ell}\}$. We will determine ℓ later to optimize the size of $\mathcal{F}$. Let $R \in \{0,1\}^{n' \times \ell}$ be a 01-matrix. We define the profile associated with R as $P^R = (\sigma_1^R, ..., \sigma_n^R)$, where for $i = 1, ..., n'$

$$\sigma_{2i}^R : x_1^{R_{i,1}} \succ x_2^{R_{i,2}} \succ \cdots \succ x_\ell^{R_{i,\ell}} \succ b_1 \succ \cdots \succ b_{2m-\ell} \succ x_\ell^{1-R_{i,\ell}} \cdots \succ x_1^{1-R_{i,1}},$$

$$\sigma_{2i-1}^R : x_1^{1-R_{i,1}} \succ x_2^{1-R_{i,2}} \succ \cdots \succ x_\ell^{1-R_{i,\ell}} \succ b_{2m-\ell} \succ \cdots \succ b_1 \succ x_\ell^{R_{i,\ell}} \cdots \succ x_1^{R_{i,1}}.$$

It turns out to be difficult to construct a set of matrices R such that the fooling set consisting of the corresponding P^R is both large and has the mixing property we want. However, we can argue using the probabilistic method that such a fooling set exists. This idea is summarized in the next two Lemmas, proven in slightly altered versions in [15]. The first roughly shows that t randomly selected voting profiles with high probability fulfill the condition we expect any t voting profiles in a fooling set to have. The second aggregates this result over all subsets of size t of a large fooling set.

Lemma A.1 ([15]). *Let $R^1, ..., R^t \in \{0,1\}^{n' \times \ell}$ be t random matrices where each element is independently 0 and 1 with equal probability. For $j = 1, ..., t$, let $P^j = (\sigma_i^j)_{i \in [n]}$ be the voting profile corresponding to R^j. Let $\mu = (1 - \left(\frac{1}{2}\right)^{t-1})n'$. Then, with probability at least*

$$\exp\left(\frac{-\ell\delta^2\mu}{2}\right),$$

there exists an alternative $x \in X^0$ and $r = (r_1, ..., r_n) \in [t]^n$ such that in the mixed voting profile $P^r = (\sigma_i^{r_i})_{i \in [n]}$, x is ranked in the top ℓ by at least $\frac{n}{2} + (1-\delta)\mu$ voters, where $\delta \in (0,1)$.

Proof ([15]). For any $x \in X^0$ and $i \in [n']$, we define the random variable as $Z_{i,x} = 1$ if and only if there exist $j, j' \in [t]$ such that x is ranked in the top ℓ by both σ_{2i}^j and $\sigma_{2i-1}^{j'}$, else $Z_{i,x} = 0$. The only case in which $Z_{i,x} = 0$ is if x is in the top ℓ of σ_{2i}^j for all $j \in [t]$ or if x is in the top ℓ of σ_{2i-1}^j for all $j \in [t]$, i.e., if all of $\{R_{i,x}^j\}_{j \in [t]}$ are 0 or all are 1. This happens with probability $1 - \left(\frac{1}{2}\right)^{t-1}$.

For a fixed x, we can find a mixed voting profile P^r as define above by picking the two voters σ_{2i}^j and $\sigma_{2i-1}^{j'}$ that rank z in the top ℓ for all $i \in [n']$ where $Z_{x,i} = 1$ and picking all other voters from P^1. Thus, x is ranked in the top ℓ by $\frac{n}{2} + \sum_{i=1}^{n'} Z_{i,x}$ voters.

Note that $\mathbb{E}\left[\sum_{i=1}^{n'} Z_{i,x}\right] = (1 - \left(\frac{1}{2}\right)^{t-1})n' = \mu$. We can use a Chernoff inequality to bound the probability that this sum is far away from its expectation:

$$\Pr\left[\sum_{i=1}^{n'} Z_{i,x} \le (1-\delta)\mu\right] \le \exp\left(\frac{-\delta^2\mu}{2}\right).$$

Since all elements in the matrices $R^1, ..., R^t$ are chosen independently, we can conclude that the probability that there does not exist a $x \in X^0$ and mixed voting profile P^r in which x is ranked in the top ℓ by at least $\frac{n}{2} + (1-\delta)\mu$ voters is at most $\exp\left(\frac{-\ell\delta^2\mu}{2}\right)$. □

Lemma A.2 ([15]). *Let*

$$k = \frac{2}{t}\exp\left(\frac{\ell\delta^2\mu}{2t}\right).$$

There exists a set $\mathcal{F} = \{P^j\}_{j\in[k]}$ of voting profiles $P^j = (\sigma_i^j)_{i\in[n]}$ such that

1. $s^*(P^j) = \frac{1}{2}\mathbf{1}$ *for $j = 1, ..., k$, and*
2. *for any subset $M \subseteq [k]$, $|M| = t$, there exists an alternative $x \in X^0$ and $r = (r_1, ..., r_n) \in M^n$ such that in the mixed voting profile $P^r = (\sigma_i^{r_i})_{i\in[n]}$, x is ranked in the top ℓ by at least $\frac{n}{2} + (1-\delta)\mu$ voters, where $\mu = (1 - \left(\frac{1}{2}\right)^{t-1})n'$.*

Proof ([15]). Let $R^1, ..., R^{tk} \in \{0,1\}^{n'\times\ell}$ be tk random matrices where each element is independently 0 and 1 with equal probability. For $j = 1, ..., \ell$, let $P^j = (\sigma_i^j)_{i\in[n]}$ be the voting profile corresponding to R^j, and let $\mathcal{G} = \{P^j\}_{j\in[tk]}$. By definition, $s^*(P^j) = \frac{1}{2}\mathbf{1}$ for $j = 1, ..., k$.

Using a union bound and Lemma A.1, we can bound the probability that $\mathcal{G}$ does not fulfill property 2 as

$$\Pr\left[\mathcal{G} \text{ doesn't satisfy } 2\right] \le \binom{tk}{t}\Pr[P_1, ..., P_t \text{ doesn't satisfy } 2] < (tk)^t \exp\left(\frac{-\ell\delta^2\mu}{2}\right) = 1.$$

Since the probability that $\mathcal{G}$ does not fulfill property 2 is strictly less than 1, it follows that there exists such a $\mathcal{G}$. Note that by this property, $\mathcal{G}$ can contain the same voting profile at most $t-1$ times, so we have to remove at most $\frac{t-2}{t-1}tk$ voting profiles from $\mathcal{G}$ to ensure no voting profile appears more than once. We obtain a set $\mathcal{F}$ of size $|\mathcal{F}| \ge |\mathcal{G}| - \frac{t-2}{t-1}tk = \frac{t}{t-1}k \ge k$ that fulfills both properties in the lemma since it is a subset of $\mathcal{G}$. □

Equipped with these two lemmas, all that is left is to define the values of the placeholder variables.

Proof of Theorem 3.8. Let $\mathcal{F} = \{P_j\}_{j\in[k]}, P_j = (\sigma_i^j)_{i\in[n]}$ be the set defined in Lemma A.2. By property 2, for any $M \subseteq [k]$, $|M| = t$ there exists an alternative $a \in A$ and a mixed profile $P^M = (\sigma_i^{r_i})_{i\in[n]}$, where $r_i \in M$, such that

$$s^*(P^M)(a) \ge \frac{m-\ell}{(m-1)n}\left(\frac{n}{2} + (1-\delta)\mu\right).$$

We define $\ell = (1-c_\ell)m$, $\delta = 1-c_d$ and $t = \log_2(c_t)+1$ for constants (i.e., independent of m, n, ε) $c_\ell, c_d \in (0,1)$ and $c_t \ge 2$, to get $s^*(P^M)(a) \ge \frac{1}{2} + \frac{c_\ell c_d}{2}(1 - \frac{1}{c_t})$.

If $s^*(P^M)(a) \ge \frac{1}{2} + 2\varepsilon$, we get that $B_\varepsilon(s^*(P^M)) \cap B_\varepsilon(s^*(P^j)) = \emptyset$ for all $j \in M$ and thus that $\mathcal{F}$ is a t-mixing fooling set for ε-approximating s^*, the Borda voting rule. By Theorem 2.7, this implies that the communication complexity of ε-approximating s^* is at least

$$\log_2\left(\frac{|F|}{t}\right) = \log_2\left(\frac{2}{t^2}\exp\left(\frac{\ell\delta^2\mu}{2t}\right)\right) = \Theta(mn)$$

for all $\varepsilon \le \frac{c_\ell c_d}{4}(1 - \frac{1}{c_t})$. This holds for all $\varepsilon \le \frac{1}{4} - c$ for any arbitrary constant $c > 0$ by picking c_ℓ and c_d sufficiently close to 1 and c_t sufficiently large. □

Remark A.3. *Using $t = 2$, we can get with the above proof technique that the communication complexity of Borda is $\Omega(mn)$ for all $\varepsilon \le \frac{1}{8} - c$ for an arbitrary constant $c > 0$. For $t = 2$, it seems feasible to give a construction for such a fooling set $\mathcal{F}$ by making the matrices R contain codewords of a binary code with constant rate and constant relative distance. However, we also note that making this proof constructive for $t = 2$ only has aesthetic value and cannot improve the bound using this class of fooling sets.*