Efficient and Thrifty Voting by Any Means Necessary*

Debmalya Mandal
Columbia University
dm3557@columbia.edu

Ariel D. Procaccia
Carnegie Mellon University
arielpro@cs.cmu.edu

Nisarg Shah
University of Toronto
nisarg@cs.toronto.edu

David P. Woodruff
Carnegie Mellon University
dwoodruf@cs.cmu.edu

Abstract

We take an unorthodox view of voting by expanding the design space to include both the *elicitation rule*, whereby voters map their (cardinal) preferences to votes, and the *aggregation rule*, which transforms the reported votes into collective decisions. Intuitively, there is a tradeoff between the communication requirements of the elicitation rule (i.e., the number of bits of information that voters need to provide about their preferences) and the efficiency of the outcome of the aggregation rule, which we measure through *distortion* (i.e., how well the utilitarian social welfare of the outcome approximates the maximum social welfare in the worst case). Our results chart the Pareto frontier of the communication-distortion tradeoff.

1 Introduction

AI systems are increasingly being used for making various decisions in society, including criminal justice [1], participatory budgeting [2] etc. Most of these decisions are high-stake and we need to ensure that such systems reflect appropriate societal values. However, in many situations, it is not a priori clear what the right choices are [3], and a promising direction is to design systems that aggregate people's opinions. The field of social choice theory studies the aggregation of individual preferences into collective decisions. While its origins can be traced back to the contributions of Condorcet [4] and others in the 18th Century, the field was founded in its modern form in the 20th Century. With his famous impossibility result, Arrow [5] pioneered the axiomatic approach to voting, in which voting rules that aggregate ranked preferences of individuals are compared qualitatively based on the axiomatic desiderata they satisfy or violate. This approach underlies most of the work on voting in social choice theory [see, e.g., 6, 7].

By contrast, research in computational social choice [8] has put more emphasis on *quantitative* evaluation of voting rules. In particular, Procaccia and Rosenschein [9] introduced the *implicit utilitarian voting* framework, in which it is assumed that individuals (a.k.a. voters) have underlying cardinal utilities for the different alternatives, and express ranked preferences that are consistent with their utilities. The goal is to choose an alternative that maximizes (utilitarian) *social welfare* — the sum of utilities — by relying on the reported rankings as a proxy for the latent utilities. Specifically, voting rules are compared by their *distortion*, which is the worst-case ratio of the maximum social welfare to the social welfare of the alternative they choose. The implicit utilitarian voting approach has received significant attention in the past decade [2, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23], and voting rules based on it have been deployed on the online voting platform robovote.org.

Benadè et al. [2] observe that implicit utilitarian voting has another advantage: it allows comparing not only voting rules that aggregate ranked preferences, but also voting rules that aggregate other

*All the missing proofs are included in the supplementary material.

33rd Conference on Neural Information Processing Systems (NeurIPS 2019), Vancouver, Canada.

types of ballots, which they refer to as *input formats*. They further argue that we can associate each input format with the best rule for aggregating votes in that format, and ultimately compare the input formats themselves based on the lowest distortion they make possible. They also introduce a new input format, *threshold approval*, whereby each voter is asked to report whether her utility for each alternative is above or below a given threshold; this input format allows achieving logarithmic distortion. The results of Benadè et al. [2] beg the question: why should we set only a single threshold? What if we set two thresholds and ask each voter to report whether her utility for each alternative is below the lower threshold, between the two thresholds, or above the higher threshold? What if we set five thresholds? Or a million for that matter? Intuitively there is a tradeoff between the number of thresholds and the distortion that can be achieved. However, perhaps adding thresholds is not the most efficient way to drive down distortion; there may be other input formats that encapsulate more useful information. (Spoiler alert: this is indeed the case.)

Our goal in this paper is to characterize the *optimal* tradeoff between elicitation and distortion. Ranking always asks voters to rank their alternatives and always asks for the same amount of information from the votrs. On the other hand, consider the input format threshold approval. The larger the number of thresholds, the finer the information we elicit about voter utilities, and the lower the distortion. This is an example of a tradeoff between the amount of information voters are required to report, and the distortion that can be achieved. As we elicit more information from voters about their utilities, we should be able to achieve lower distortion. But exactly how low? To answer this question, we need a precise way to reason about the complexity of vote elicitation. We use the nomenclature of communication complexity [24], and, in particular, examine the number of bits needed to report a vote. Note that this is simply the logarithm of the number of possible votes that a voter can provide in a given input format. Hence, plurality votes that ask a voter to report which of the m alternatives is her top choice contain $\log m$ bits of information, while ranked preferences that ask a voter to rank all m alternatives contain $\log m! = \Theta(m \log m)$ bits of information. Our main research question is this:

For any k, given a budget of at most k bits per vote, what is the minimum distortion any voting rule can achieve?

1.1 Our Results

Before outlining our results, we describe our framework in a bit more detail (a formal model is presented in Section 2). A voting rule f is composed of two parts. Its *elicitation rule* Π_f elicits information from voters about their utilities. Essentially, it chooses a (possibly randomized) mapping from utility functions to finitely many (say k) possible responses, and each voter uses this mapping to cast her vote. The communication complexity of f, denoted $\mathrm{C}(f)$, is then $\mathbb{E}[\log k]$, where the expectation is due to random choices made by Π_f. The *aggregation rule* Γ_f aggregates the votes cast by voters to choose a single alternative (possibly in a randomized way). The distortion of f, denoted $\mathrm{dist}(f)$, is the worst-case ratio of the maximum social welfare to the (expected) social welfare of this chosen alternative. The distortion is typically a function of the number of alternatives m. Our goal is to study the tradeoff between $\mathrm{C}(f)$ and $\mathrm{dist}(f)$.

Figure 1 shows our results and positions them in the context of prior work. We note that any upper bound with deterministic elicitation (resp. aggregation) also serves as an upper bound with randomized elicitation (resp. aggregation), and the converse holds for lower bounds. For deterministic elicitation, it is known that plurality voting rule achieves $\Theta(m^2)$ distortion with deterministic aggregation and $\log m$ communication complexity, and that it is trivial to achieve $\Theta(m)$ distortion with randomized aggregation and zero communication complexity [18]. Our lower bounds from Section 4 establish that these are the best possible asymptotic bounds with communication complexity at most $\log m$. We show that these bounds do not hold for randomized elicitation by constructing a new voting rule in Section 3, RANDSUBSET, which uses randomized elicitation and achieves $o(m)$ distortion with $o(\log m)$ communication complexity.

We also propose a family of voting rules, PREFTHRESHOLD, which use deterministic elicitation and aggregation, and can achieve d distortion with $O(m \log(d \log m)/d)$ communication complexity, shown as the solid line in Figure 1. This is an improvement over bounds achievable by existing deterministic elicitation methods (even with optimal randomized aggregation): threshold approval voting has distortion $\Omega(\sqrt{m})$ with communication complexity m [2], and ranked voting has distortion $\Omega(\sqrt{m})$ with communication complexity $m \log m$ [11]. In fact, this is also an improvement

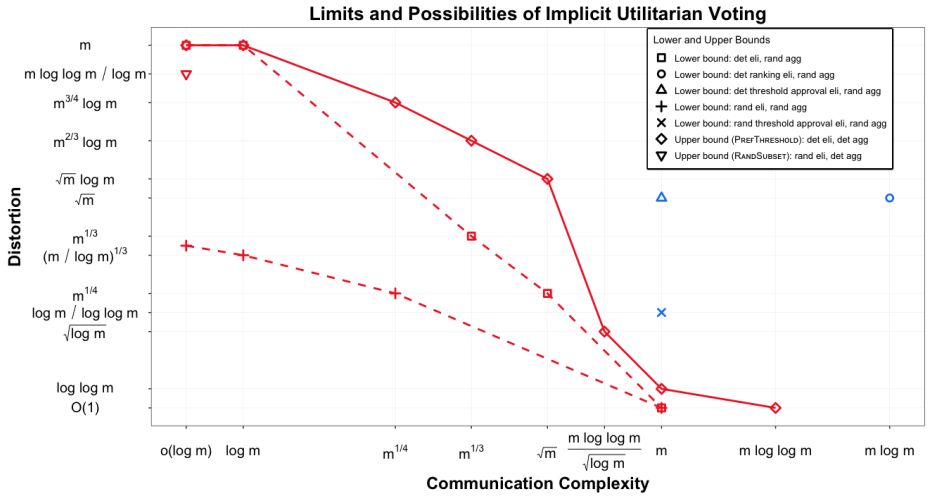

Figure 1: The figure depicts lower and upper bounds on distortion which can be achieved as a function of communication complexity. We use the following abbreviation: det = deterministic, rand = randomized, eli = elicitation, agg = aggregation. Red dashed lines show our lower bounds (Theorems 4 and 6), which apply to all voting rules using deterministic or randomized elicitation. Red diamonds (solid line) and inverted triangle show some of the upper bounds achieved by two families of rules we introduce — PREFTHRESHOLD (Theorem 1) and RANDSUBSET (Theorem 2) — which use deterministic and randomized elicitation, respectively, and deterministic aggregation. We obtain tradeoffs which Pareto-dominate the best possible tradeoffs that existing elicitation methods (shown in blue) — such as threshold approval voting [2] and ranked voting [11] — allow even with randomized aggregation.

over the randomized elicitation version of threshold approval voting, which still has distortion $\Omega(\log m/ \log logm)$ with communication complexity m [2].

In Section 5, we leverage tools from multi-party communication complexity to show that the bounds achieved by PREFTHRESHOLD are nearly optimal: any voting rule with d distortion must have $\Omega(m/d^2)$ communication complexity with deterministic elicitation and $\Omega(m/d^3)$ communication complexity with randomized elicitation. These are presented as dashed lines in Figure 1. Note that our upper and lower bounds differ by a factor that is almost linear or almost quadratic in d, and sublogarithmic in m. This implies a surprising fact: when our goal is to achieve near-constant distortion, randomization cannot significantly help.

1.2 Related Work

There are two threads of research on implicit utilitarian voting. The first thread does not make any assumptions on utilities, other than that they are normalized [2, 9, 10, 11, 17, 18, 22, 23]. The second thread assumes that utilities are induced by a metric [12, 13, 14, 15, 16, 19, 20, 21]; this structure generally enables lower distortion. Our approach is consistent with the former thread.

In addition to the work of Benadè et al. [2], discussed above, an especially relevant paper is that of Caragiannis and Procaccia [10]. Their goal is also to achieve low distortion while keeping the communication requirements low. To that end, they employ specific voting techniques such as approving a single alternative (like plurality) or approving a subset of alternatives (like approval voting) — these require $\log m$ and m bits per voter, respectively — but use what they call an *embedding* to describe how voters translate their cardinal preferences into votes. However, the key difference between the work of Caragiannis and Procaccia and our work is that our design space is much larger: we simultaneously optimize both the embedding and the voting technique (together, these form our elicitation rule), as well as the aggregation rule.[2]

[2]That said, in this work we focus only on deterministic embeddings. That is, we study elicitation rules in which voters deterministically translate their cardinal preferences into votes, and show that the foregoing result is impossible to achieve in this case. We discuss implications of randomized embeddings in Section 6.

Further afield, Conitzer and Sandholm [25] study how much information about the voters' ranked preferences has to be elicited in order to compute the outcome under a given voting rule. By contrast, we are interested in designing the voting rule, and the very way in which preferences are represented, in order to minimize distortion. In addition, the voting rules we design ask voters to report their approximate utility for their top few choices or for a randomly chosen subset of alternatives. Related ideas have been explored previously [26] or in parallel [27] in the computational social choice literature, albeit in fundamentally different models.

Another loosely related line of work was initiated by Balcan and Harvey [28] and Badanidiyuru et al. [29]. Their goal is to sketch *combinatorial* valuation functions, that is, to encode such functions using a polynomial number of bits in a way that the value of each subset can be recovered approximately. We deal with much simpler valuation functions, but, on the other hand, are looking to achieve much lower communication complexity. We also note that in several query models it is standard to directly query a real number [28, 29, 30, 31, 32]; by contrast, in our framework, asking for even a single real number leads to infinite communication complexity.

2 Model

For $k \in \mathbb{N}$, define $[k] = \{1, \ldots, k\}$. Let $x \sim D$ denote that random variable x has distribution D. Let $\log$ denote the logarithm to base 2, and $\ln$ denote the logarithm to base e. There is a set of alternatives A with $|A| = m$, and a set of voters $N = [n]$. Each voter $i \in N$ is endowed with a valuation $v_i : A \to \mathbb{R}_+$, where $v_i(a) \geq 0$ represents the value of voter i for alternative a. Equivalently, we view $v_i \in \mathbb{R}_+^m$ as a vector which contains the voter's value for each alternative. We slightly abuse notation and let $v_i(S) = \sum_{a \in S} v_i(a)$ for $S \subseteq A$. Collectively, voter valuations are denoted $\vec{v} = (v_1, \ldots, v_n)$. Given $\vec{v}$, the (utilitarian) social welfare of an alternative a is $\mathrm{sw}(a, \vec{v}) = \sum_{i \in N} v_i(a)$. Our goal is to elicit information about voter valuations and use it to find an alternative with high social welfare.

Valuations: We adopt the standard normalization assumption that $\sum_{a \in A} v_i(a) = 1$ for each $i \in N$. This can be thought of as a "one voter, one vote" principle for cardinal valuations as it prevents voters from overshadowing other voters by expressing very high values. [3] Let Δ^m denote the m-simplex, i.e., the set of all vectors in $\mathbb{R}_+^m$ whose coordinates sum to 1. Hence, we have that $v_i \in \Delta^m$ for each $i \in N$. Given such a vector $v_i \in \Delta^m$, let $\mathrm{supp}(v_i) \subseteq A$ denote the *support* of v_i, i.e., the set of alternatives a for which $v_i(a) > 0$.

Query space: Consider any interaction with voter i which elicits finitely many bits of information and in which the voter responds deterministically. In this interaction, the voter must provide one of finitely many (say k) possible responses. We say that this interaction elicits $\log k$ bits of information.[4] It effectively partitions Δ^m into k *compartments*, where the compartment corresponding to each response is the set of all valuations which would result in the voter choosing that response. In other words, any interaction which elicits $\log k$ bits of information is equivalent to a *query* which partitions Δ^m into k compartments and asks the voter to pick the compartment in which her valuation belongs.

Let $\mathcal{Q}$ denote the set of all queries which partition Δ^m into finitely many compartments. For a query $q \in \mathcal{Q}$, let $k(q)$ denote the number of compartments created by q; the number of bits elicited is $\log k(q)$.[5] This query space incorporates traditional elicitation methods studied in the social choice literature. For instance, plurality votes (which ask voters to report their favorite alternative) use m compartments, k-approval votes (which ask voters to report the set of their k favorite alternatives) use $\binom{m}{k}$ compartments, threshold approval votes (which ask voters to approve alternatives for which their value is at least a given threshold) use 2^m compartments, and ranked votes (which ask voters to rank all alternatives) use $m!$ compartments.

Voting Rule: A voting rule (or simply, a *rule*) f consists of two parts: an *elicitation rule* Π_f and an *aggregation rule* Γ_f. The (randomized) elicitation rule Π_f is a distribution over $\mathcal{Q}$, according to which a query q is sampled. Each voter i provides a response ρ_i to this query, depending on her valuation

[3] Effectively, voters can only report the intensity of their relative preference for one alternative over another.

[4] For a multi-round interaction, we can concatenate the voter's responses in different rounds. This is equivalent to a single-round interaction in which the voter is asked to provide the entire string upfront.

[5] Note that the number of bits elicited may not be an integer, but 2 raised to the power of the number of bits must be an integer. We could take the ceiling to enforce an integral number of bits, and this would only minimally increase elicitation, but some of our lower bounds are sensitive to this non-integral formulation.

v_i. We say that the elicitation rule is *deterministic* if it has singleton support (i.e., it chooses a query deterministically). The (randomized) aggregation rule Γ_f takes voter responses $\vec{\rho} = (\rho_1, \ldots, \rho_n)$ as input, and returns a distribution over alternatives. We say that the aggregation rule is *deterministic* if it always returns a distribution with singleton support. Slightly abusing notation, we denote by $f(\vec{v})$ the (randomized) alternative returned by f when voter valuations are $\vec{v} = (v_1, \ldots, v_n)$. We measure the performance of f via two metrics.

1. The *communication complexity* of f for m alternatives, denoted $\mathrm{C}^m(f) = \mathbb{E}_{q \sim \Pi_f} \log k(q)$, is the expected number of bits of information elicited by f from each voter. We drop m from the superscript when its value is clear from the context.
2. The *distortion* of f for m alternatives, denoted $\mathrm{dist}^m(f)$, is the worst-case ratio of the optimal social welfare to the expected social welfare achieved by f. Again, we drop m from the superscript when its value is clear from the context. Formally,

$$\mathrm{dist}(f) = \sup_{\vec{v} \in (\Delta^m)^n} \frac{\max_{a \in A} \mathrm{sw}(a, \vec{v})}{\mathbb{E}_{\widehat{a} \sim f(\vec{v})} \mathrm{sw}(\widehat{a}, \vec{v})}.$$

While it is desirable for a voting rule to have low communication complexity and low distortion, typically eliciting more information from voters enables achieving low distortion. Our goal is to understand the Pareto frontier of the tradeoff between communication complexity and distortion.

3 Upper Bounds

In this section, we derive upper bounds on the best distortion a voting rule can achieve given an upper bound on its communication complexity (equivalently, this gives an upper bound on the communication complexity required to achieve a given level of distortion). We construct two families of voting rules: PREFTHRESHOLD, which use deterministic elicitation and aggregation, and RANDSUBSET, which convert a given voting rule into one which uses randomized elicitation.

3.1 Deterministic Elicitation, Deterministic Aggregation

We begin by designing voting rules which use deterministic elicitation and deterministic aggregation — the most practical combination. Caragiannis et al. [18] show that plurality achieves $\Theta(m^2)$ distortion with $\log m$ communication complexity, and even voting rules that elicit ranked preferences, and thus have $\Theta(m \log m)$ communication complexity, cannot achieve asymptotically better distortion.

We propose a novel voting rule $\textsc{PrefThreshold}_{t,\ell}$, parametrized by $t \in [m]$ and $\ell \in \mathbb{N}$. It is presented as Algorithm 1. Its elicitation rule asks each voter to report the set of her t most preferred alternatives, and for each alternative in this set, report her approximate value for it by picking one of $\ell + 1$ subintervals of $[0, 1]$. Note that for $t = 1$, we use ℓ subintervals of $[1/m, 1]$; this is valid because a voter's value for her most favorite alternative must be at least $1/m$. The aggregation rule is also intuitive: it uses the approximate values to compute an estimated social welfare of each alternative, and picks an alternative with the highest estimated social welfare. The next theorem provides bounds on the communication and distortion of $\textsc{PrefThreshold}_{t,\ell}$.

Theorem 1. *prefth For $t \in [m] \setminus \{1\}$ and $\ell \in \mathbb{N}$, we have*

$$C(\textsc{PrefThreshold}_{t,\ell}) = \Theta\left(t \log \frac{m(\ell+1)}{t}\right), \mathit{dist}(\textsc{PrefThreshold}_{t,\ell}) = O\left(m^{1+2/\ell}/t\right).$$

For $t = 1$ and $\ell \in \mathbb{N}$, we have

$$C(\textsc{PrefThreshold}_{1,\ell}) = \log(m\ell),\ \mathit{dist}(\textsc{PrefThreshold}_{t,\ell}) = O\left(m^{1+1/\ell}\right).$$

$\textsc{PrefThreshold}_{t,\ell}$ offers a tradeoff between two parameters, t and ℓ. Increasing either parameter increases the communication complexity but reduces the distortion. It is easy to see that there is no (asymptotic) benefit of choosing $\ell > \log m$. We make several observations.

- $\boldsymbol{t = 1, \ell = 2}$ gives us subquadratic distortion of $O(m\sqrt{m})$ with just one more bit of elicitation than plurality (i.e. $\log m + 1$ bits).
- $\boldsymbol{t = m^{1-\gamma}, \ell = \log m}$ gives us sublinear distortion of $O(m^\gamma)$ (for $\gamma \in (0, 1)$) with polynomial communication complexity of $O(m^{1-\gamma} \log m)$.

ALGORITHM 1: PREFTHRESHOLD$_{t,\ell}$, where $t \in [m]$ and $\ell \in \mathbb{N}$.

Elicitation Rule:

- If $t > 1$, create $\ell + 1$ buckets: $B_0 = [0, 1/m^2]$ and $B_p = (1/m^{2-2(p-1)/\ell}, 1/m^{2-2p/\ell}]$ for $p \in [\ell]$.
- If $t = 1$, create ℓ buckets: $B_1 = [m^{-1}, m^{-1+1/\ell}]$ and $B_p = (m^{-1+(p-1)/\ell}, m^{-1+p/\ell}]$ for $p \in [\ell] \setminus \{1\}$.
- The query asks each voter i to identify set S_i^t of the t alternatives for which she has the highest value (breaking ties arbitrarily), and for each $a \in S_i^t$, identify bucket index $p_{i,a}$ such that $v_i(a) \in B_{p_{i,a}}$.

Aggregation Rule:

- For each p, let U_p denote the upper endpoint of bucket B_p.
- For each voter $i \in N$ and alternative $a \in A$, define $\widehat{v_i}(a) = U_{p_{i,a}}$ if $a \in S_i^t$ and $\widehat{v_i}(a) = 0$ o.w.
- For an alternative $a \in A$, define the *estimated social welfare* as $\widehat{\mathrm{sw}}(a) = \sum_{i \in N} \widehat{v_i}(a)$.
- Return an alternative with the highest estimated social welfare, i.e., $\widehat{a} \in \arg\max_{a \in A} \widehat{\mathrm{sw}}(a)$.

- $\boldsymbol{t = m/\sqrt{\log m}, \ell = \log m}$ has distortion $O\left(\sqrt{\log m}\right)$ with communication $o(m)$, and Pareto-dominates threshold approval voting, which has higher communication complexity of m and higher distortion of $\Omega(\log m / \log\log m)$, even with randomized aggregation [2].
- $\boldsymbol{t = m, \ell = \log m}$ leads to constant distortion with communication $O(m \log\log m)$. By contrast, eliciting ranking leads to higher communication complexity of $\Theta(m \log m)$, and also significantly higher distortion of $\Omega(\sqrt{m})$, even with randomized aggregation [11].

3.2 Randomized Elicitation, Randomized Aggregation

We now present a generic approach to designing voting rules with randomized elicitation. Given a voting rule f and an integer $s \leq m$, instead of using f to select one alternative from A directly, we sample $S \subseteq A$ with $|S| = s$ at random and use f to select one alternative from S. Recall that for $p \in \mathbb{N}$, $\mathrm{C}^p(f)$ and $\mathrm{dist}^p(f)$ denote the communication complexity and distortion of f for p alternatives, respectively.

Clearly, this approach reduces the communication complexity from $\mathrm{C}^m(f)$ to $\mathrm{C}^s(f)$. Its effect on distortion, however, is more subtle. On the one hand, selecting an alternative from S instead of A results in an inevitable loss of welfare because we can only hope to do as well as the best alternative in S. On the other hand, the welfare we achieve is related to the welfare of the best alternative in S via the factor $\mathrm{dist}^s(f)$, which can be significantly lower than $\mathrm{dist}^m(f)$. We show that in some cases, this approach reduces distortion in addition to reducing communication complexity. The key challenge in making this approach work is that we cannot apply f directly to select one alternative from S, as the total value of the alternatives in S need not be 1. We circumvent this obstacle by eliciting an approximate value of $v_i(S)$ from each voter i, making a number of copies of voter i that is approximately proportional to $v_i(S)$ with each copy now having a total value of 1 for alternatives in S, and running f on the resulting instance.

ALGORITHM 2: RANDSUBSET(f, s), where f is a voting rule and $s \in [m]$

Elicitation Rule:

- Pick $S \subseteq A$ with $|S| = s$ uniformly at random from among all subsets of A of size s.
- Partition $[0, 1]$ into $\lceil \log(4m) \rceil$ buckets as follows: $B_0 = \left[0, \frac{1}{4m}\right]$, $B_j = \left(\frac{2^{j-1}}{4m}, \frac{2^j}{4m}\right]$ for $j \in \lceil \log(4m) \rceil$.
- Ask two reports from each voter i:
  1. The bucket index p_i such that $v_i(S) = \sum_{a \in S} v_i(a) \in B_{p_i}$;
  2. A response ρ_i to the elicitation rule of f for the set of alternatives S according to the renormalized valuation $\widehat{v_i}$ defined as $\widehat{v_i}(a) = v_i(a)/v_i(S)$ for each $a \in S$.

Aggregation Rule:

- Let L_p denote the lower endpoint of bucket B_p for $p \in \lceil \log(4m) \rceil \cup \{0\}$.
- Run the aggregation rule of f on an input which consists of $4m \cdot L_{p_i}$ copies of ρ_i for each $i \in N$.

Theorem 2. *For every voting rule f and $s \in [m]$, we have $\mathrm{C}^m(\text{RANDSUBSET}(f, s)) = \mathrm{C}^s(f) + \log\lceil \log(4m) \rceil$ and $\mathrm{dist}^m(\text{RANDSUBSET}(f, s)) \leq \frac{4m}{s} \cdot \mathrm{dist}^s(f)$.*

Using $f = \text{PREFTHRESHOLD}_{t,\ell}$ and Theorem 1, we obtain that for $s \in [m]$, $t \in [s]$, and $\ell \in \mathbb{N}$, there is a new voting rule $g = \text{RANDSUBSET}\left(\text{PREFTHRESHOLD}_{t,\ell}, s\right)$ with

$$\mathrm{C}^m(g) = O\left(t \log(s(\ell+1)/t) + \log\log m\right) \text{ and } \mathrm{dist}^m(g) = O\left(m \cdot s^{2/\ell}/t\right).$$

Setting $\ell = \log s$, we get $O(m/t)$ distortion. Then, we set $s = t$ to minimize communication complexity to $O(t \log\log t + \log\log m)$. This is slightly better than using $\text{PREFTHRESHOLD}_{t,\log m}$, which achieves $O(m/t)$ distortion with $O(t \log \frac{m \log m}{t})$ communication complexity. In particular, for $t = O(1)$ this reduces communication complexity by a factor of $\log m / \log\log m$.

An interesting choice is $t = \frac{\log m}{\log\log m}$, which leads to distortion $O\left(m \log\log m / \log m\right) = o(m)$ and communication complexity $O\left(t \log\log t + \log\log m\right) = o(\log m)$. Note that this rule has randomized elicitation but deterministic aggregation. By contrast, we later show that with deterministic elicitation, no voting rule can achieve $o(m)$ distortion with communication complexity at most $\log m$, even when randomized aggregation is allowed (Theorem 4).

4 Direct Lower Bounds For Deterministic Elicitation

We now turn to deriving lower bounds on the distortion of a voting rule given an upper bound on its communication complexity (equivalently, this gives a lower bound on the communication complexity required to achieve a given level of distortion). In this section, we focus on deterministic elicitation.

Consider a voting rule f which uses deterministic elicitation and has communication complexity at most $\log k$. Hence, the (deterministic) query of f must partition Δ^m into at most k compartments. Wthout loss of generality, we can assume that f uses exactly k compartments. This is because if f uses k' compartments where $k' < k$, then we can partition some of its compartments into smaller compartments and derive a new voting rule g which uses exactly k compartments, receives at least the information that f receives from the voters, and simulates the aggregation rule of f to achieve the same distortion. Now, establishing a lower bound on the distortion of f requires analyzing the following game between two players, the voting rule f and the adversary.

1. The voting rule f decides the partition of Δ^m into k compartments.
2. The adversary decides the response of each voter.
3. The voting rule f picks a winning alternative (or a distribution over winning alternatives, if its aggregation rule is randomized).
4. The adversary picks valuations of voters consistent with their responses in the second step.

We use this framework to derive lower bounds on the distortion of voting rules that use deterministic elicitation. We first focus on deterministic aggregation. Perhaps the simplest such voting rule is plurality, which has $\log m$ communication complexity and achieves $\Theta(m^2)$ distortion. This raises an important question: *What distortion can we achieve with deterministic elicitation, deterministic aggregation, and communication complexity less than* $\log m$*?* The next lemma shows that the answer turns out to be disappointing.

Theorem 3. *Every voting rule that has deterministic elicitation, deterministic aggregation, and communication complexity strictly less than* $\log m$ *has unbounded distortion.*

Now, plurality has communication complexity $\log m$ and achieves $\Theta(m^2)$ distortion. *Can a different voting rule achieve better distortion using only* $\log m$ *communication complexity?* Perhaps unsurprisingly, we answer this in the negative. But the proof of this intuitive result is quite intricate.

Further, using randomized aggregation we can trivially achieve $O(m)$ distortion with zero communication complexity (by returning the uniform distribution over alternatives). One may wonder: *How much information do we need from the voters to achieve sublinear distortion?* It is easy to show that eliciting plurality votes is not sufficient. Surprisingly, we show that this holds for *every* $\log m$-bit elicitation. That is, even with randomized aggregation, eliciting $\log m$ bits per voter is asymptotically no better than blindly selecting an alternative uniformly at random!

Theorem 4. *Let f be a voting rule with deterministic elicitation and $C(f) \leq \log m$. If f uses deterministic (resp. randomized) aggregation, then dist$(f) = \Omega(m^2)$ (resp. $\Omega(m)$).*

For deterministic aggregation, Theorem 4 shows that eliciting $\log m$ bits per voter is not sufficient to achieve $o(m^2)$ distortion. By contrast, we know from Theorem 1 that we can achieve $O(m)$ distortion by eliciting $O(\log m)$ bits per voter. Similarly, for randomized aggregation, Theorem 4 shows that eliciting $\log m$ bits per voter is not sufficient to achieve $o(m)$ distortion. However, we can achieve $o(m)$ distortion if we are willing to elicit $\omega(\log m)$ bits per voter (Theorem 1),[6] or if we are willing to use randomized elicitation (Theorem 2).

5 Lower Bounds Through Multi-Party Communication Complexity

In this section, we leverage tools from the literature on multi-party communication complexity to derive lower bounds for both deterministic and randomized elicitation. Specifically, we derive lower bounds on the communication complexity of voting rules that achieve a given level of distortion. We begin by reviewing existing results on multi-party communication complexity, and then derive new results, which help us prove the desired lower bounds in our voting context.

5.1 Setup

In multi-party communication complexity, there are t computationally omnipotent players. Each player i holds a private input $X_i \in \mathcal{X}_i$. The *input profile* is the vector $(X_1, \ldots, X_t)$. The goal is to compute the output of a function $f : \mathcal{X}_1 \times \mathcal{X}_2 \times \ldots \times \mathcal{X}_t \to \{0, 1\}$ on the input profile.

A shared *protocol* Π specifies how the players exchange information among themselves and with the center. We use the *blackboard model*, in which messages written by one player are visible to all other players. Let $\Pi(X_1, \ldots, X_t)$ be the random variable denoting the message transcript generated when all players follow the protocol on input profile $(X_1, \ldots, X_t)$; here, the randomness is due to coin tosses by the players or the protocol. The *communication cost* of Π, denoted $|\Pi|$, is the maximum length of $\Pi(X_1, \ldots, X_t)$ over all input profiles $(X_1, \ldots, X_t)$ and all coin tosses. Given $\delta \geq 0$, we say that Π is a *δ-error protocol* for f if there exists a function Π_{out} such that for every input profile $(X_1, \ldots, X_t)$, $\Pr\left[\Pi_{\text{out}}(\Pi(X_1, \ldots, X_t)) = f(X_1, \ldots, X_t)\right] \geq 1 - \delta$. The *$\delta$-error communication complexity* of f, denoted $R_\delta(f)$, is the communication cost of the best δ-error protocol for f.

5.2 Multi-Party Fixed-Size Set-Disjointness

The main ingredient of our proof is a standard problem in multi-party communication complexity called the *multi-party set-disjointness* problem, denoted $\text{DISJ}_{m,t}$. Here, each player i holds an arbitrary set S_i from a universe of size m. The goal is to distinguish between two types of inputs.

- NO inputs: The sets are pairwise disjoint, i.e., $S_i \cap S_j = \emptyset$ for all $i \neq j$.
- YES inputs: The sets have a unique element in common, but are otherwise pairwise disjoint, i.e., there exists x such that $S_i \cap S_j = \{x\}$ for all $i \neq j$.

It is promised that the input will be one of these two types (in other words, the protocol is free to choose any output on an input that does not satisfy this promise). Following a series of results [33, 34, 35], Gronemeier [36] and Jayram [37] finally established the optimal lower bound of $\Omega(m/t)$.

We introduce a variant of this problem, which we call *multi-party fixed-size set-disjointness* and denote $\text{FDISJ}_{m,s,t}$. It is almost identical to $\text{DISJ}_{m,t}$, except that we know each player i holds a set S_i of a given size s. Our goal is to still determine whether the sets are pairwise disjoint ($S_i \cap S_j = \emptyset$ for all $i \neq j$) or pairwise uniquely intersecting (there exists x such that $S_i \cap S_j = \{x\}$ for all $i \neq j$). We use the lower bound on $R_\delta(\text{DISJ}_{m,t})$ to derive the following lower bound on $R_\delta(\text{FDISJ}_{m,s,t})$.

Theorem 5. *For a sufficiently small constant $\delta > 0$ and $m \geq (3/2)st$, $R_\delta(\text{FDISJ}_{m,s,t}) = \Omega(s)$.*

5.3 Lower Bounds on the Communication Complexity of Voting Rules

We now use our lower bound on the δ-error communication complexity of $\text{FDISJ}_{m,s,t}$ to derive a lower bound on the communication complexity of a voting rule in terms of its distortion. We derive different bounds depending on whether the elicitation rule of f is deterministic or randomized. For randomized elicitation, our bound is weaker.

[6] For $t = \omega(1)$, $\text{PREFTHRESHOLD}_{t,\log m}$ has distortion $O(m/t) = o(m)$ and communication $O(t \log m)$.

Table 1: Comparison between our lower bounds (Theorem 6) and upper bounds (Theorem 1)

Distortion	Lower Bounds		Upper Bound
	Deterministic Elicitation	Randomized Elicitation	
$O(m^{\gamma})$	$\Omega(m^{1-2\gamma})$	$\Omega(m^{1-3\gamma})$	$O(m^{1-\gamma}\log m)$
$O(\log m)$	$\Omega\left(m/\log^2 m\right)$	$\Omega\left(m/\log^3 m\right)$	$O(m\log\log m/\log m)$
$O(1)$	$\Omega(m)$	$\Omega(m)$	$O(m\log\log m)$

The key insight in the proof is that we can use a voting rule f with $\text{dist}(f) \leq t/2$ to construct a δ-error protocol for solving $\text{FDISJ}_{m,s,t}$, and hence we can use the lower bound on $R_\delta(\text{FDISJ}_{m,s,t})$ from Theorem 5 to derive a lower bound on $\text{C}(f)$. At a high level, consider an instance $(S_1, \ldots, S_t)$ of $\text{FDISJ}_{m,s,t}$. We ask each player i to respond to the query of f according to an artificial valuation function constructed using S_i. We then use these responses to create an input for the aggregation rule of f. We show that by asking each player an additional question about the alternative returned by the aggregation rule, and possibly running this process a number of times, we can solve $\text{FDISJ}_{m,s,t}$.

Theorem 6. *Consider a voting rule f with elicitation rule Π_f and dist$(f) = d$. If Π_f is deterministic, then $C(f) \geq \Omega\left(m/d^2\right)$, and if Π_f is randomized, then $C(f) \geq \Omega\left(m/d^3\right)$.*

Finally, Table 1 summarizes our upper and lower bounds for some special cases. Achieving sublinear distortion makes polynomial communication complexity both necessary (even with randomized aggregation) and sufficient (even with deterministic aggregation). If $\text{dist}(f) = O(\log m)$, our upper and lower bounds differ by only polylogarithmic factors. And for constant distortion, they differ by only a sublogarithmic factor.

6 Discussion

We initiated a formal study of the communication-distortion tradeoff in voting, but our work leaves many open questions. The most immediate direction is to close the gap between our upper and lower bounds. Regarding our upper bounds, both families of voting rules that we introduce — PREFTHRESHOLD and RANDSUBSET — use deterministic aggregation, and we do not have better upper bounds using randomized aggregation. Our lower bounds from Theorem 6 are also identical for deterministic and randomized aggregation. This raises an elegant question: *Can randomized aggregation help?* Also, using randomized elicitation in RANDSUBSET, we can achieve sublinear distortion with communication complexity at most $\log m$; Theorem 4 shows that this is not possible with deterministic elicitation. This raises another elegant question: *What is the best possible distortion with randomized elicitation and communication complexity at most* $\log m$*?* It would also be interesting to improve upon our lower bounds in Section 5, potentially by using a different problem from the multi-party communication complexity literature.

Taking a broader viewpoint, we can consider more general forms of elicitation, like *non-uniform* questions across voters, and questions *adaptive* to past responses. One can also explore the effect of imposing other restrictions on the voting rule such as truthfulness [23, 38]. On a conceptual level, perhaps the main take-away message of our paper is that it pays off to elicit and aggregate preferences "by any means necessary," that is, potentially through highly nonstandard aggregation and, especially, elicitation rules. In the setting of Caragiannis and Procaccia [10], voters are software agents, and this is natural. But when voters are people, it is crucial to understand the implications of such unconventional approaches, both in terms of how communication complexity corresponds to cognitive burden, and in terms of the interpretability and transparency of aggregation rules.

Acknowledgments

Mandal was partially supported by the Post-Doctoral fellowship from the Columbia Data Science Institute, and part of this work was done while he was a graduate student at Harvard University. Procaccia was partially supported by the National Science Foundation under grants IIS-1350598, IIS-1714140, CCF-1525932, and CCF-1733556; by the Office of Naval Research under grants N00014-16-1-3075 and N00014-17-1-2428; by a J.P. Morgan AI Research Award; and by a Guggenheim

Fellowship. Shah was partially supported by the Natural Sciences and Engineering Research Council under a Discovery grant. Woodruff was partially supported by the National Science Foundation under grant CCF-1815840, and part of this work was done while he was visiting the Simons Institute for the Theory of Computing.

References

[1] Jon Kleinberg, Himabindu Lakkaraju, Jure Leskovec, Jens Ludwig, and Sendhil Mullainathan. Human Decisions and Machine Predictions. *The quarterly journal of economics*, 133(1): 237–293, 2017.

[2] G. Benadè, S. Nath, A. D. Procaccia, and N. Shah. Preference elicitation for participatory budgeting. In *Proceedings of the 31st AAAI Conference on Artificial Intelligence (AAAI)*, pages 376–382, 2017.

[3] Edmond Awad, Sohan Dsouza, Richard Kim, Jonathan Schulz, Joseph Henrich, Azim Shariff, Jean-François Bonnefon, and Iyad Rahwan. The moral machine experiment. *Nature*, 563(7729): 59, 2018.

[4] Marquis de Condorcet. Essai sur l'application de l'analyse à la probabilité de décisions rendues à la pluralité de voix. Imprimerie Royal, 1785. Facsimile published in 1972 by Chelsea Publishing Company, New York.

[5] K. Arrow. *Social Choice and Individual Values*. Wiley, 1951.

[6] K. J. Arrow, A. Sen, and K. Suzumura. *Handbook of Social Choice and Welfare*, volume 2. Elsevier, 2010.

[7] A. Sen. Social choice theory. *Handbook of Mathematical Economics*, 3:1073–1181, 1986.

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

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

[10] I. Caragiannis and A. D. Procaccia. Voting almost maximizes social welfare despite limited communication. *Artificial Intelligence*, 175(9–10):1655–1671, 2011.

[11] C. Boutilier, I. Caragiannis, S. Haber, T. Lu, A. D. Procaccia, and O. Sheffet. Optimal social choice functions: A utilitarian view. *Artificial Intelligence*, 227:190–213, 2015.

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

[13] E. Anshelevich and J. Postl. Randomized social choice functions under metric preferences. *Journal of Artificial Intelligence Research*, 58:797–827, 2017.

[14] S. Gross, E. Anshelevich, and L. Xia. Vote until two of you agree: Mechanisms with small distortion and sample complexity. In *Proceedings of the 31st AAAI Conference on Artificial Intelligence (AAAI)*, pages 544–550, 2017.

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

[16] Y. Cheng, S. Dughmi, and D. Kempe. On the distortion of voting with multiple representative candidates. In *Proceedings of the 32nd AAAI Conference on Artificial Intelligence (AAAI)*, pages 973–980, 2018.

[17] G. Benadé, A. D. Procaccia, and M. Qiao. Low-distortion social welfare functions. In *Proceedings of the 33rd AAAI Conference on Artificial Intelligence (AAAI)*, 2019. Forthcoming.

[18] I. Caragiannis, S. Nath, A. D. Procaccia, and N. Shah. Subset selection via implicit utilitarian voting. *Journal of Artificial Intelligence Research*, 58:123–152, 2017.

[19] A. Borodin, O. Lev, N. Shah, and T. Strangway. Primarily about primaries. In *Proceedings of the 33rd AAAI Conference on Artificial Intelligence (AAAI)*, 2019. Forthcoming.

[20] M. Ghodsi, M. Latifian, and M. Seddighin. On the distortion value of the elections with abstention. In *Proceedings of the 33rd AAAI Conference on Artificial Intelligence (AAAI)*, 2019. Forthcoming.

[21] M. Feldman, A. Fiat, and I. Golumb. On voting and facility location. In *Proceedings of the 17th ACM Conference on Economics and Computation (EC)*, pages 269–286, 2016.

[22] U. Bhaskar and A. Ghosh. On the welfare of cardinal voting mechanisms. In *Proceedings of the 38th IARCS Annual Conference on Foundations of Software Technology and Theoretical Computer Science (FSTTCS)*, pages 27:1–27:22, 2018.

[23] U. Bhaskar, V. Dani, and A. Ghosh. Truthful and near-optimal mechanisms for welfare maximization in multi-winner elections. In *Proceedings of the 32nd AAAI Conference on Artificial Intelligence (AAAI)*, pages 925–932, 2018.

[24] E. Kushilevitz and N. Nisan. *Communication Complexity*. Cambridge University Press, 1996.

[25] V. Conitzer and T. Sandholm. Communication complexity of common voting rules. In *Proceedings of the 6th ACM Conference on Economics and Computation (EC)*, pages 78–87, 2005.

[26] Y. Filmus and J. Oren. Efficient voting via the top-k elicitation scheme: A probabilistic approach. In *Proceedings of the 15th ACM Conference on Economics and Computation (EC)*, pages 295–312, 2014.

[27] M. Bentert and P. Skowron. Comparing election methods where each voter ranks only few candidates. arXiv:1901.10848, 2019.

[28] M.-F. Balcan and N. J. Harvey. Submodular functions: Learnability, structure, and optimization. *SIAM Journal on Computing*, 47(3):703–754, 2018.

[29] A. Badanidiyuru, S. Dobzinski, H. Fu, R. Kleinberg, N. Nisan, and T. Roughgarden. Sketching valuation functions. In *Proceedings of the 23rd Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1025–1035, 2012.

[30] E. Balkanski, A. Rubinstein, and Y. Singer. The limitations of optimization from samples. In *Proceedings of the 49th Annual ACM Symposium on Theory of Computing (STOC)*, pages 1016–1027, 2017.

[31] U. Feige, V. S. Mirrokni, and J. Vondrák. Maximizing non-monotone submodular functions. *SIAM Journal on Computing*, 40(4):1133–1153, 2011.

[32] V. S. Mirrokni, M. Schapira, and J. Vondrák. Tight information-theoretic lower bounds for welfare maximization in combinatorial auctions. In *Proceedings of the 9th ACM Conference on Electronic Commerce (EC)*, pages 70–77, 2008.

[33] Z. Bar-Yossef, T. S. Jayram, R. Kumar, and D. Sivakumar. An information statistics approach to data stream and communication complexity. *Journal of Computer and System Sciences*, 68 (4):702–732, 2004.

[34] N. Alon, Y. Matias, and M. Szegedy. The space complexity of approximating the frequency moments. *Journal of Computer and System Sciences*, 58(1):137–147, 1999.

[35] A. Chakrabarti, S. Khot, and X. Sun. Near-optimal lower bounds on the multi-party communication complexity of set disjointness. In *Proceedings of the 18th IEEE Annual Conference on Computational Complexity (CCC)*, pages 107–117, 2003.

[36] A. Gronemeier. Asymptotically optimal lower bounds on the NIH-multi-party information complexity of the AND-function and disjointness. In *Proceedings of the 26th International Symposium on Theoretical Aspects of Computer Science (STACS)*, pages 505–516, 2009.

[37] T. S. Jayram. Hellinger strikes back: A note on the multi-party information complexity of AND. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques*, pages 562–573. 2009.

[38] A. Gibbard. Manipulation of schemes that mix voting with chance. *Econometrica*, 45:665–681, 1977.

Appendix

A Upper Bounds

A.1 Deterministic Elicitation, Deterministic Aggregation

Theorem 1. *For $t \in [m] \setminus \{1\}$ and $\ell \in \mathbb{N}$, we have*

$$C(\textsc{PrefThreshold}_{t,\ell}) = \log\left[\binom{m}{t} \cdot (\ell+1)^t\right] = \Theta\left(t \log \frac{m(\ell+1)}{t}\right),$$
$$\textit{dist}(\textsc{PrefThreshold}_{t,\ell}) = O\left(m^{1+2/\ell}/t\right).$$

For $t = 1$ and $\ell \in \mathbb{N}$, we have

$$C(\textsc{PrefThreshold}_{1,\ell}) = \log(m\ell),\ \textit{dist}(\textsc{PrefThreshold}_{t,\ell}) = O\left(m^{1+1/\ell}\right).$$

Proof. It is evident that the number of possible responses that a voter can provide under $\textsc{PrefThreshold}_{t,\ell}$ is $\binom{m}{t} \cdot (\ell+1)^t$ if $t > 1$, and $m\ell$ if $t = 1$. Taking the logarithm of this gives us the desired communication complexity.

We now establish the distortion of $\textsc{PrefThreshold}_{t,\ell}$. Let $\vec{v} = (v_1, \ldots, v_n)$ be the underlying valuations of voters. For alternative $a \in A$, recall that $\mathrm{sw}(a, \vec{v}) = \sum_{i \in N} v_i(a)$, and

$$\widehat{\mathrm{sw}}(a) = \sum_{i \in N} \widehat{v_i}(a) = \sum_{i \in N : a \in S_i^t} \widehat{v_i}(a) = \sum_{i \in N : a \in S_i^t} U_{p_{i,a}}.$$

Let $\widehat{a} \in \arg\max_{a \in A} \widehat{\mathrm{sw}}(a)$ be the alternative chosen by the rule, and let $a^* \in \arg\max_{a \in A} \mathrm{sw}(a, \vec{v})$ be an alternative maximizing social welfare.

We begin by finding an upper bound on $\mathrm{sw}(a^*, \vec{v})$ in terms of $\widehat{\mathrm{sw}}(\widehat{a})$.

$$\begin{aligned}
\mathrm{sw}(a^*, \vec{v}) &= \sum_{i \in N} v_i(a^*) = \sum_{i \in N : a^* \in S_i^t} v_i(a^*) + \sum_{i \in N : a^* \notin S_i^t} v_i(a^*) \\
&\leq \sum_{i \in N : a^* \in S_i^t} v_i(a^*) + \sum_{i \in N : a^* \notin S_i^t} \left(\frac{\sum_{a \in S_i^t} v_i(a)}{t}\right) \\
&\leq \sum_{i \in N : a^* \in S_i^t} \widehat{v_i}(a^*) + \frac{\sum_{a \in A \setminus \{a^*\}} \sum_{i \in N : a^* \notin S_i^t \wedge a \in S_i^t} \widehat{v_i}(a)}{t} \\
&\leq \widehat{\mathrm{sw}}(a^*) + \frac{\sum_{a \in A \setminus \{a^*\}} \widehat{\mathrm{sw}}(a)}{t} \leq \widehat{\mathrm{sw}}(\widehat{a}) + \frac{(m-1) \cdot \widehat{\mathrm{sw}}(\widehat{a})}{t} = \frac{m+t-1}{t} \cdot \widehat{\mathrm{sw}}(\widehat{a}),
\end{aligned} \tag{1}$$

where the third transition holds because for every $i \in N$ with $a^* \notin S_i^t$ and every $a \in S_i^t$, we have $v_i(a^*) \leq v_i(a)$; the fourth transition holds because for every $i \in N$ and $a \in S_i^t$, $v_i(a) \leq \widehat{v_i}(a)$; the fifth transition follows from the definition of $\widehat{\mathrm{sw}}$; and the sixth transition holds because $\widehat{a}$ is a maximizer of $\widehat{\mathrm{sw}}$.

We now establish the distortion for $t > 1$. The first step is to derive an upper bound on $\widehat{\text{sw}}(\widehat{a})$ in terms of $\text{sw}(\widehat{a}, \vec{v})$. Our bucketing implies that for all $i \in N$ and $a \in S_i^t$, we have $v_i(a) \leq \widehat{v_i}(a) \leq m^{2/\ell} v_i(a) + \frac{1}{m^2}$. Using this, we can derive the following.

$$\widehat{\text{sw}}(\widehat{a}) = \sum_{i \in N : \widehat{a} \in S_i^t} \widehat{v_i}(\widehat{a}) \leq \sum_{i \in N : \widehat{a} \in S_i^t} \left(m^{2/\ell} v_i(\widehat{a}) + \frac{1}{m^2} \right) \leq m^{2/\ell} \text{sw}(\widehat{a}, \vec{v}) + \frac{n}{m^2}. \tag{2}$$

Next, we derive a lower bound on $\widehat{\text{sw}}(\widehat{a})$, which helps establish a lower bound on $\text{sw}(\widehat{a}, \vec{v})$. Note that for each voter $i \in N$, $\sum_{a \in S_i^t} v_i(a) \geq t/m$. Hence,

$$\sum_{a \in A} \widehat{\text{sw}}(a) = \sum_{i \in N} \sum_{a \in S_i^t} \widehat{v_i}(a) \geq \sum_{i \in N} \sum_{a \in S_i^t} v_i(a) \geq \frac{n \cdot t}{m}.$$

Because $\widehat{a}$ is a maximizer of $\widehat{\text{sw}}$, this yields $\widehat{\text{sw}}(\widehat{a}) \geq n \cdot t/m^2$. Substituting this into Equation (2), we get

$$\frac{n}{m^2} + \text{sw}(\widehat{a}, \vec{v}) \cdot m^{2/\ell} \geq \widehat{\text{sw}}(\widehat{a}) \geq \frac{n \cdot t}{m^2} \Rightarrow \text{sw}(\widehat{a}, \vec{v}) \geq \frac{n \cdot (t-1)}{m^2} \cdot m^{-2/\ell} \geq \frac{n}{m^2} \cdot m^{-2/\ell}. \tag{3}$$

Applying Equations (1), (2), and (3) in this order, we have

$$\begin{aligned} \frac{\text{sw}(a^*, \vec{v})}{\text{sw}(\widehat{a}, \vec{v})} &\leq \frac{m+t-1}{t} \cdot \frac{\widehat{\text{sw}}(\widehat{a})}{\text{sw}(\widehat{a}, \vec{v})} \leq \frac{m+t-1}{t} \cdot \left(m^{2/\ell} + \frac{n}{m^2 \cdot \text{sw}(\widehat{a}, \vec{v})} \right) \\ &\leq \frac{m+t-1}{t} \cdot \left(m^{2/\ell} + m^{2/\ell} \right) \in O(m^{1+2/\ell}/t). \end{aligned}$$

For $t = 1$, we have that for every $i \in N$ and $a \in S_i^t$, $v_i(a) \leq \widehat{v_i}(a) \leq m^{1/\ell} v_i(a)$. Hence, in Equation (2), the additive factor of n/m^2 disappears and the multiplicative factor of $m^{2/\ell}$ becomes $m^{1/\ell}$, yielding $\widehat{\text{sw}}(\widehat{a}) \leq \text{sw}(\widehat{a}, \vec{v}) \cdot m^{1/\ell}$. Similarly, Equation (3) becomes $\text{sw}(\widehat{a}, \vec{v}) \geq \frac{n}{m^2} \cdot m^{-1/\ell}$. Following the same line of proof as for the case of $t > 1$, we obtain

$$\frac{\text{sw}(a^*, \vec{v})}{\text{sw}(\widehat{a}, \vec{v})} \leq m \cdot \frac{\widehat{\text{sw}}(\widehat{a})}{\text{sw}(\widehat{a}, \vec{v})} \leq m \cdot m^{1/\ell},$$

which is the desired bound on distortion. □

A.2 Randomized Elicitation, Randomized Aggregation

Theorem 2. *For every voting rule f and $s \in [m]$, we have $C^m(\textsc{RandSubset}(f, s)) = C^s(f) + \log\lceil \log(4m) \rceil$ and $dist^m(\textsc{RandSubset}(f, s)) \leq \frac{4m}{s} \cdot dist^s(f)$.*

Proof. Let $\vec{v} = (v_1, \ldots, v_n)$ denote the underlying valuations of voters. First, let us consider a fixed choice of $S \subseteq A$ with $|S| = s$. Due to our bucketing, we have that for every $i \in N$,

$$\frac{v_i(S)}{2} - \frac{1}{4m} \leq L_{p_i} \leq v_i(S). \tag{4}$$

Recall that in the input to the aggregation rule of f, we have $4m \cdot L_{p_i}$ copies of the response ρ_i of voter i. Hence, the social welfare function approximated by the aggregation rule of f is given by

$$\forall a \in S,\ \widehat{\text{sw}}(a, \vec{v}) = \sum_{i \in N} 4m \cdot L_{p_i} \cdot \frac{v_i(a)}{v_i(S)} = 4m \sum_{i \in N} v_i(a) \cdot \frac{L_{p_i}}{v_i(S)}.$$

Combining this with Equation (4), we have that for each $a \in S$,

$$\widehat{\text{sw}}(a, \vec{v}) \geq 4m \sum_{i \in N} v_i(a) \cdot \left(\frac{1}{2} - \frac{1}{4m \cdot v_i(S)} \right) = 2m \cdot \text{sw}(a, \vec{v}) - \sum_{i \in N} \frac{v_i(a)}{v_i(S)} \geq 2m \cdot \text{sw}(a, \vec{v}) - n, \tag{5}$$

as well as

$$\widehat{\text{sw}}(a, \vec{v}) \leq 4m \sum_{i \in N} v_i(a) \cdot 1 = 4m \cdot \text{sw}(a, \vec{v}). \tag{6}$$

Let $\widehat{a}$ denote the alternative chosen by our rule. Because the distortion of f for choosing an alternative from S is $\text{dist}^s(f)$, we have that $\mathbb{E}[\widehat{\text{sw}}(\widehat{a}, \vec{v})] \geq \max_{a \in S} \widehat{\text{sw}}(a, \vec{v})/\text{dist}^s(f)$. Note that so far, we have fixed S. The expectation on the left hand side is due to the fact that even for fixed S, $\widehat{a}$ can be randomized if f is randomized.

Next, we take expectation over the choice of S, and use the fact that the optimal alternative $a^* \in \arg\max_{a \in A} \text{sw}(a, \vec{v})$ belongs to S with probability s/m. We obtain

$$\mathbb{E}[\widehat{\text{sw}}(\widehat{a}, \vec{v})] \geq \frac{\mathbb{E}[\max_{a \in S} \widehat{\text{sw}}(a, \vec{v})]}{\text{dist}^s(f)} \geq \frac{\frac{s}{m} \cdot \widehat{\text{sw}}(a^*, \vec{v})}{\text{dist}^s(f)} \geq \frac{\frac{s}{m}\left(2m \cdot \text{sw}(a^*, \vec{v}) - n\right)}{\text{dist}^s(f)}, \tag{7}$$

where the final transition follows from Equation (5). On the other hand, from Equation (6), we have

$$\mathbb{E}[\widehat{\text{sw}}(\widehat{a}, \vec{v})] \leq 4m\, \mathbb{E}[\text{sw}(\widehat{a}, \vec{v})]. \tag{8}$$

Combining Equations (7) and (8), we have that

$$\text{dist}^m(\textsc{RandSubset}(f, s)) = \frac{\text{sw}(a^*, \vec{v})}{\mathbb{E}[\text{sw}(\widehat{a}, \vec{v})]} \leq \frac{\text{sw}(a^*, \vec{v})}{\frac{\text{sw}(a^*, \vec{v})}{2} - \frac{n}{4m}} \cdot \frac{m}{s} \cdot \text{dist}^s(f) \leq \frac{4m}{s} \cdot \text{dist}^s(f),$$

where the final transition uses the fact that $\text{sw}(a^*, \vec{v}) \geq (1/m) \cdot \sum_{a \in A} \text{sw}(a, \vec{v}) = n/m$. This establishes the desired distortion bound. Since each voter answers the query of f for s alternatives and chooses one of $\lceil \log(4m) \rceil$ buckets, we get $\text{C}^m(\textsc{RandSubset}(f, s)) = \text{C}^s(f) + \log\lceil \log(4m) \rceil$, as desired. □

B Lower Bounds

B.1 Direct Lower Bounds for Deterministic Elicitation

We start by establishing a straightforward lemma. Recall that for a valuation $v \in \Delta^m$, $\text{supp}(v)$ denotes the support of v.

Lemma 1. *Let f be a voting rule which uses deterministic elicitation and deterministic aggregation. Let q^* be the query used by f. If some compartment of q^* contains two valuations v^1 and v^2 such that $\text{supp}(v^1) \cap \text{supp}(v^2) = \emptyset$, then the distortion of f is unbounded.*

Proof. Suppose compartment P contains valuations v^1 and v^2 such that $\text{supp}(v^1) \cap \text{supp}(v^2) = \emptyset$. Let $\widehat{a}$ be the alternative returned by f when all voters pick compartment P. Pick $t \in \{1, 2\}$ such that $\widehat{a} \notin \text{supp}(v^t)$. Note that $v^t(\widehat{a}) = 0$, but there exists $a^* \in \text{supp}(v^t)$ such that $v^t(a^*) > 0$.

Define voter valuations $\vec{v} = (v_1, \ldots, v_n)$ such that $v_i = v^t$ for each $i \in N$. This yields $\text{sw}(\widehat{a}, \vec{v}) = 0$ and $\text{sw}(a^*, \vec{v}) > 0$, which implies that f must have infinite distortion. □

Theorem 3. *Every voting rule that has deterministic elicitation, deterministic aggregation, and communication complexity strictly less than $\log m$ has unbounded distortion.*

Proof. We need the following definition. For $a \in A$, we say that the *unit valuation* corresponding to a is the valuation $v^a \in \Delta^m$ for which $v^a(a) = 1$. Let f be a voting rule that has deterministic elicitation and deterministic aggregation, and let $\text{C}(f) < \log m$. Hence, the query used by f must partition Δ^m into less than m compartments.

Because there are m unit valuations, by the pigeonhole principle there must exist distinct $a, b \in A$ such that v^a and v^b belong to the same compartment. Because $\text{supp}(v^a) \cap \text{supp}(v^b) = \emptyset$, Lemma 1 implies that the distortion of f must be infinite. □

Theorem 4. *Let f be a voting rule which uses deterministic elicitation and has $\text{C}(f) \leq \log m$. If f uses deterministic aggregation, then $\text{dist}(f) = \Omega(m^2)$. If f uses randomized aggregation, then $\text{dist}(f) = \Omega(m)$.*

Proof. Let f be a voting rule which has deterministic elicitation and $\mathrm{C}(f) \leq \log m$. As argued above, we can assume $\mathrm{C}(f) = \log m$ without loss of generality. Hence, the query q^* used by f partitions Δ^m into m compartments. Let $\mathcal{P} = (P_1, \ldots, P_m)$ denote the set of compartments. If f has unbounded distortion, we are done. Suppose f has bounded distortion.

Due to Lemma 1, each of m unit vectors must belong to a different compartment. Since there are m compartments, we identify each compartment by the unit valuation it contains. For $a \in A$, let P^a denote the compartment containing unit valuation v^a. Before we construct adversarial valuations, we need to define *low valuations* and *high valuations*.

Low valuations: We say that a valuation $v \in \Delta^m$ is a *low valuation* if $|\mathrm{supp}(v)| = m/5$ and $v(a) = 5/m$ for every $a \in \mathrm{supp}(v)$. Let $\Delta^{m,\text{low}}$ denote the set of all low valuations. Due to Lemma 1, we have

$$v \in \Delta^{m,\text{low}} \cap P^a \Rightarrow a \in \mathrm{supp}(v) \wedge v(a) = \frac{5}{m}. \tag{9}$$

Let $\mathcal{L} = \{P \in \mathcal{P} : P \cap \Delta^{m,\text{low}} \neq \emptyset\}$ be the set of compartments containing at least one low valuation, and $A^{\mathcal{L}} = \{a \in A : P^a \in \mathcal{L}\}$ be the set of alternatives corresponding to these compartments.

We claim that $|A^{\mathcal{L}}| = |\mathcal{L}| \geq 4m/5 + 1$. Suppose for contradiction that $|A^{\mathcal{L}}| \leq 4m/5$. Then, $|A \setminus A^{\mathcal{L}}| \geq m/5$. Hence, there exists a low valuation $v \in \Delta^{m,\text{low}}$ such that $\mathrm{supp}(v) \subseteq A \setminus A^{\mathcal{L}}$. Let $a \in A$ be the alternative for which $v \in P^a$. Because P^a contains a low valuation, $a \in A^{\mathcal{L}}$ by definition. Thus, the construction of v ensures $v(a) = 0$. We have $v \in \Delta^{m,\text{low}} \cap P^a$ with $v(a) = 0$, which contradicts Equation (9). Hence, $|A^{\mathcal{L}}| \geq 4m/5 + 1$.

High valuations: We say that a valuation $v \in \Delta^m$ is a *high valuation* if $|\mathrm{supp}(v)| = 2$ and $v(a) = 1/2$ for each $a \in \mathrm{supp}(v)$. Let $\Delta^{m,\text{high}}$ denote the set of high valuations. Note that $|\Delta^{m,\text{high}}| = \binom{m}{2}$. Similarly to the case of low valuations, we can apply Lemma 1, and obtain that

$$v \in \Delta^{m,\text{high}} \cap P^a \Rightarrow a \in \mathrm{supp}(v) \wedge v(a) = \frac{1}{2}. \tag{10}$$

For $a \in A$, let $\mathcal{H}^a = \{P \in \mathcal{L} : \exists v \in \Delta^{m,\text{high}} \cap P \text{ s.t. } a \in \mathrm{supp}(v)\}$. In words, $\mathcal{H}^a$ is the set of compartments from $\mathcal{L}$ which contain at least one high valuation v for which $v(a) = 1/2$. Let $A^{\text{high}} = \{a \in A : |\mathcal{H}^a| \geq m/5\}$. We claim that $|A^{\text{high}}| \geq m/6$.

Suppose this is not true. Let $B = |A \setminus A^{\text{high}}|$. Then, $|B| \geq 5m/6$. Consider $a \in B$. Each of the $m-1$ high valuations which contain a in their support must belong to some compartments in $\mathcal{H}^a \cup (\mathcal{P} \setminus \mathcal{L})$. Since $|\mathcal{H}^a| \leq m/5 - 1$ for $a \in B$ and $|\mathcal{P} \setminus \mathcal{L}| \leq m/5 - 1$, the $m-1$ high valuations containing a in their support are distributed across at most $2m/5 - 2$ compartments. However, due to Lemma 1, a compartment other than P^a can contain at most one high valuation with a in its support. Hence, P^a must contain at least $m - 1 - (2m/5 - 3) = 3m/5 + 2$ high valuations. Thus, we have established that $|B| \geq 5m/6$ and for each $a \in B$, P^a contains at least $3m/5 + 2$ high valuations. Thus, the number of high valuations is at least $(5m/6) \cdot (3m/5 + 2) > m^2/2 > \binom{m}{2}$, which is a contradiction. Thus, we have $|A^{\text{high}}| \geq m/6$.

We are now ready to prove the desired result for both deterministic and randomized aggregation.

Voter responses: When responding to the query q^*, suppose each compartment $P \in \mathcal{L}$ is picked by a set N_P of $n/|\mathcal{L}|$ voters.

Deterministic aggregation: Let $\widehat{a}$ denote the alternative picked by f. We claim that $\widehat{a} \in A^{\mathcal{L}}$. If $\widehat{a} \notin A^{\mathcal{L}}$, consider voter valuations $\vec{v}$ such that every voter i picking compartment $P^a \in \mathcal{L}$ has valuation $v_i = v^a$. Since $\widehat{a} \notin A^{\mathcal{L}}$, we have $v_i(\widehat{a}) = 0$ for each $i \in N$, i.e., $\mathrm{sw}(\widehat{a}, \vec{v}) = 0$. Since $\mathrm{sw}(a, \vec{v}) > 0$ for some $a \in A$, f has infinite distortion, which is a contradiction. Thus, we must have $\widehat{a} \in A^{\mathcal{L}}$.

Now, let us construct the voter valuations as follows. Pick a low valuation $\widehat{v} \in P^{\widehat{a}} \cap \Delta^{m,\text{low}}$, which exists because we have established $\widehat{a} \in A^{\mathcal{L}}$. Note that $\widehat{v}(\widehat{a}) = 5/m$. For each $i \in N_{P^{\widehat{a}}}$, let $v_i = \widehat{v}$. Pick $a^* \in A^{\text{high}} \setminus \{\widehat{a}\}$. Let $\bar{P}$ be the compartment containing the high valuation under which both $\widehat{a}$ and a^* have utility $1/2$. For each $P \in \mathcal{H}^{a^*} \setminus \{P^{\widehat{a}}, \bar{P}\}$, and for each $i \in N_P$, let v_i be the high valuation in P such that $v_i(a^*) = 1/2$ and $v_i(\widehat{a}) = 0$. For every other $P^a \in \mathcal{L}$ and every $i \in N_{P^a}$, let $v_i = v^a$.

Observe that under these valuations, $\mathrm{sw}(\widehat{a}, \vec{v}) = \Theta(n/m^2)$, whereas, since $|\mathcal{H}^{a^*}| \geq m/5$ and $|\mathcal{L}| \leq |\mathcal{P}| = m$, $\mathrm{sw}(a^*, \vec{v}) = \Theta(n)$. We conclude that $\mathrm{dist}(f) = \Omega(m^2)$.

Randomized aggregation: Note that f must select at least one alternative $a^* \in A^{\text{high}}$ with probability at most $1/|A^{\text{high}}| \leq 6/m$. Construct voter valuations such that for every $P \in \mathcal{H}^{a^*}$ and every $i \in N_P$, v_i is the high valuation under which $v_i(a^*) = 1/2$. For every $P^a \in \mathcal{L} \setminus \mathcal{H}^{a^*}$, and for every $i \in N_{P^a}$, let $v_i = v^a$. It holds that $\mathrm{sw}(a^*, \vec{v}) = \Theta(n)$ (as before), whereas $\mathrm{sw}(a, \vec{v}) = O(n/m)$ for every $a \in A \setminus \{a^*\}$. Because f selects a^* with probability at most $6/m$, we have $\mathbb{E}_{\widehat{a} \sim f(\vec{v})}[\mathrm{sw}(\widehat{a}, \vec{v})] = O(n/m)$, implying $\mathrm{dist}(f) = \Omega(m)$, as required. □

B.2 Lower Bound for Plurality Votes

In this section, we show that eliciting plurality votes (whereby each voter picks her most favorite alternative) results in $\Omega(m)$ distortion, even with randomized aggregation. This is implied by Theorem 4, which proves this for any elicitation that has at most $\log m$ communication complexity. However, for the special case of plurality votes, we can provide a much simpler proof.

Proposition 1. *Every voting rule which elicits plurality votes incurs $\Omega(m)$ distortion.*

Proof. For simplicity, let the number of voters n be divisible by the number of alternatives m. Consider an input profile in which the set of voters N is partitioned into equal-size sets $\{N_a\}_{a \in A}$ such that for each $a \in A$, a is the most favorite alternative of every voter in N_a.

Take any voting rule f. It must return some alternative $a^* \in A$ with probability at most $1/m$. Now, construct adversarial valuations of voters $\vec{v}$ as follows.

- For all $i \in N_{a^*}$, $v_i(a^*) = 1$ and $v_i(a) = 0$ for all $a \in A \setminus \{a^*\}$.
- For all $\widehat{a} \in A \setminus \{a^*\}$ and $i \in N_{\widehat{a}}$, $v_i(\widehat{a}) = v_i(a^*) = 1/2$ and $v_i(a) = 0$ for all $a \in A \setminus \{a^*, \widehat{a}\}$.

Under these valuations, we have $\mathrm{sw}(a^*, \vec{v}) \geq n/2$, while $\mathrm{sw}(a, \vec{v}) = (n/m) \cdot (1/2)$ for every $a \in A \setminus \{a^*\}$. Hence, the distortion of f is

$$\mathrm{dist}(f) \geq \frac{\mathrm{sw}(a^*, \vec{v})}{\frac{1}{m}\mathrm{sw}(a^*, \vec{v}) + \frac{m-1}{m}\frac{n}{2m}} = \Omega(m),$$

where the final transition holds when substituting $\mathrm{sw}(a^*, \vec{v}) \geq n/2$. □

C Lower Bounds Through Multi-Party Communication Complexity

C.1 Lower Bound on the Communication Complexity of $\mathrm{FDISJ}_{m,s,t}$

In this section, we prove a lower bound on the communication complexity of multi-party fixed-size set-disjointness. Let us recall Theorem 5.

Theorem 5. *For a sufficiently small constant $\delta > 0$ and $m \geq (3/2)st$, $R_\delta(\mathrm{FDISJ}_{m,s,t}) = \Omega(s)$.*

Proof. Suppose there is a δ-error protocol Π for $\mathrm{FDISJ}_{m,s,t}$. We use it to construct a 2δ-error protocol Π' for $\mathrm{DISJ}_{m',t'}$, where $m' = st/2$ and $t' = 2t$.

Consider an instance $(S'_1, \ldots, S'_{t'})$ of $\mathrm{DISJ}_{m',t'}$. Due to the promise that the sets are either pairwise disjoint or pairwise uniquely intersecting, we have that at most one of the m' elements can appear in multiple sets. Hence, $\sum_{i=1}^{t'} |S'_i| \leq m' - 1 + t'$. Due to the pigeonhole principle, there must exist at least $t'/2 = t$ sets of size at most $2(m' + t' - 1)/t'$. Note that

$$\frac{2(m' + t' - 1)}{t'} = \frac{st/2 + 2t - 1}{t} = \frac{s}{2} + 2 - \frac{1}{t} \leq s.$$

The final transition holds when $s \geq 4$. When $s < 4$, the lower bound of $\Omega(s)$ is trivial.

Consider a set of t players $\{i_1, \ldots, i_t\}$ such that $|S'_{i_k}| \leq s$ for each $k \in [t]$. Suppose that each such player i_k adds $s - |S'_{i_k}|$ unique elements to S'_{i_k} and creates a set S_{i_k} with $|S_{i_k}| = s$. The number of

unique elements required is at most st. Hence, the total number of elements used in sets $S_{i_1}, \ldots, S_{i_t}$ is at most $m' + st = (3/2)st \leq m$. In other words, these sets can be created using the m-element universe of $\text{FDISJ}_{m,s,t}$. Further, it is easy to check that sets $S_{i_1}, \ldots, S_{i_t}$ are pairwise disjoint (resp. pairwise uniquely intersecting) if and only if sets $S'_1, \ldots, S'_{t'}$ are pairwise disjoint (resp. pairwise uniquely intersecting). Thus, $(S_{i_1}, \ldots, S_{i_t})$ is a valid instance of $\text{FDISJ}_{m,s,t}$ and has the same solution as the instance $(S'_1, \ldots, S'_{t'})$ of $\text{DISJ}_{m',t'}$.

Our goal is to construct a 2δ-error protocol Π' for $\text{DISJ}_{m',t'}$ that solves $(S'_1, \ldots, S'_{t'})$ by effectively running the given δ-error protocol Π for $\text{FDISJ}_{m,s,t}$ on $(S'_{i_1}, \ldots, S'_{i_t})$. We could ask each player i to report a single bit indicating whether $|S'_i| \leq s$, determine t players for which this holds, and then run Π on them. However, this would add a t'-bit overhead. Instead, we would like to bound the overhead in terms of the communication cost of Π, denoted $|\Pi|$, which could be significantly smaller.

This is achieved as follows. We first order the players according to a uniformly random permutation σ. Then, we simulate Π. Every time Π wants to interact with a new player, we ask players that we have not interacted with so far, in the order in which they appear in σ, whether their sets have size at most s, until we find one such player. Then, we let Π interact with this player. Protocol Π' terminates naturally when protocol Π terminates (and returns the same answer), but terminates abruptly if, at any point, it has interacted with more than $2|\Pi|/\delta$ players (and returns an arbitrary answer).

Note that $|\Pi|$ is also an upper bound with the number of players that Π needs to interact with. Let X be the smallest index such that there are at least $|\Pi|$ players having sets of size at most s among the first X players in σ. Then, because at least half of the players have sets of size at most s, we have $\mathbb{E}[X] \leq 2 \cdot |\Pi|$. Due to Markov's inequality, we have that $\Pr[X > 2|\Pi|/\delta] \leq \delta$. Hence, the probability that Π' terminates abruptly is at most δ. When it does not terminate abruptly, it returns the wrong answer with probability at most δ (as Π is a δ-error protocol). Hence, due to the union bound, we conclude that Π' is a 2δ-error protocol for $\text{DISJ}_{m',t'}$.

Finally, we have that $|\Pi'| \leq 2|\Pi|/\delta + |\Pi| = |\Pi|(1 + 2/\delta)$. When δ is sufficiently small, Gronemeier [36] showed that $|\Pi'| \geq R_{2\delta}(\text{DISJ}_{m',t'}) = \Omega(m'/t') = \Omega(s)$. Hence, we have that $|\Pi| = \Omega(s)$. Since this holds for every δ-error protocol Π for $\text{FDISJ}_{m,s,t}$, we have $R_\delta(\text{FDISJ}_{m,s,t}) = \Omega(s)$. □

C.2 Lower Bounds on the Communication Complexity of Voting Rules

Theorem 6. *For a voting rule f with elicitation rule Π_f and dist$(f) = d$, the following hold.*

- *If Π_f is deterministic, then $C(f) \geq \Omega\left(m/d^2\right)$.*
- *If Π_f is randomized, then $C(f) \geq \Omega\left(m/d^3\right)$.*

Proof. Let $t = 2 \cdot \text{dist}(f)$ and $s = 2m/(3t)$. Note that for these parameters, we have $R_\delta(\text{FDISJ}_{m,s,t}) = \Omega(s)$ from Theorem 5.

Consider an input $(S_1, \ldots, S_t)$ to $\text{FDISJ}_{m,s,t}$ with a universe U of size m. Let us create an instance of the voting problem with a set of n voters N and a set of m alternatives A. Each alternative in A corresponds to a unique element of U. Partition the set of voters N into t equal-size buckets $\{N_1, \ldots, N_t\}$. Here, bucket N_i corresponds to player i, and consists of n/t voters that each have valuation v^{S_i} given by $v^{S_i}(a) = 1/s$ for each $a \in S_i$ and $v^{S_i}(a) = 0$ for each $a \notin S_i$. Let $\vec{v}$ denote the resulting profile of voter valuations. Note that under these valuations, $\text{sw}(a, \vec{v}) = \frac{n}{ts} \sum_{i=1}^{t} \mathbb{1}[a \in S_i]$, where $\mathbb{1}$ is the indicator variable. Due to the promise that an element either belongs to at most one set or belongs to every set, we have $\text{sw}(a, \vec{v}) \in \{0, n/(ts), n/s\}$. We say that a is a "good" alternative if $\text{sw}(a, \vec{v}) = n/s$ and a "bad" alternative otherwise.

We define two processes that will help covert our voting rule f into a protocol for $\text{FDISJ}_{m,s,t}$.

Process E: In this process, we ask each player i to respond to the query posed by voting rule f (possibly selected in a randomized manner) according to valuation v^{S_i}. We note that this requires a total of $t \cdot \text{C}(f)$ bits of communication from the players.

Process A: We take players' responses from process E, create n/t copies of the response of each player, and pass the resulting profile as input to the aggregation rule Γ_f to obtain the returned alternative $\widehat{a}$ (possibly selected in a randomized manner). We end the process by determining if $\widehat{a}$ is a

good alternative or a bad alternative. This requires eliciting 2 extra bits of information: we can ask any two players i and j whether their sets contain $\widehat{a}$, and due to the promise of $\mathrm{FDISJ}_{m,s,t}$, we know that $\widehat{a}$ is good if and only if it belongs to both S_i and S_j.

Knowing whether $\widehat{a}$ is good or bad is useful for solving the given instance of $\mathrm{FDISJ}_{m,s,t}$ due to the following reason.

1. If $(S_1, \ldots, S_t)$ is a "NO input", then we know that every alternative is a bad alternative. Hence, $\mathrm{sw}(a, \vec{v}) \leq (n/t) \cdot (1/s) = n/(ts)$ for each $a \in A$. In particular, this implies $\mathrm{sw}(\widehat{a}, \vec{v}) \leq n/(ts)$ with probability 1.
2. If $(S_1, \ldots, S_t)$ is a "YES input", then there exists a unique good alternative $a^* \in A$ with $\mathrm{sw}(a^*, \vec{v}) = n/s$, and every other alternative a is a bad alternative with $\mathrm{sw}(a, \vec{v}) \leq n/(ts)$. Because $\mathrm{dist}(f) = t/2$, we have that $\mathbb{E}[\mathrm{sw}(\widehat{a}, \vec{v})] \geq \frac{n/s}{t/2} = \frac{2n}{ts}$. This implies that $\Pr[\mathrm{sw}(\widehat{a}, \vec{v}) = n/s] = \Pr[\widehat{a} = a^*] \geq 1/t$ because if $\Pr[\widehat{a} = a^*] < 1/t$, then $\mathbb{E}[\mathrm{sw}(\widehat{a}, \vec{v})] < (1/t) \cdot (n/s) + 1 \cdot n/(ts) = 2n/(ts)$, which is a contradiction.

We are now ready to use f to construct a protocol for $\mathrm{FDISJ}_{m,s,t}$, and use Theorem 5 to derive a lower bound on $\mathrm{C}(f)$. We consider two cases depending on whether the elicitation rule Π_f is deterministic or randomized.

1. *Deterministic elicitation*: In this case, we run process E once and then run process A $t \ln(1/\delta)$ times. In a NO input, we always get a bad alternative. In a YES input, each run of process A returns a good alternative with probability at least $1/t$. Hence, the probability that we get a good alternative at least once is at least $1 - (1 - 1/t)^{t \ln(1/\delta)} \geq 1 - \delta$. Hence, this is a δ-error protocol for $\mathrm{FDISJ}_{m,s,t}$ which requires $t \cdot \mathrm{C}(f) + t \ln(1/\delta) \cdot 2$ bits of total communication from the players. Using Theorem 5, we have that $t \cdot (\mathrm{C}(f) + 2 \ln(1/\delta)) = \Omega(s)$. Using $s = 2m/(3t)$ and $t = 2d$, we have $\mathrm{C}(f) = \Omega(m/d^2)$.
2. *Randomized elicitation*: In this case, we run E once followed by running A once. And we repeat this entire process $t \ln(1/\delta)$ times. Note that we need to repeat process E because the elicitation is also randomized. Like in the previous case, we always get a bad alternative in a NO input, and get a good alternative with probability at least $1/t$ in each run in a YES input. Hence, in a YES input, we get a good alternative in at least one run with probability at least $1 - (1 - 1/t)^{t \ln(1/\delta)} \geq 1 - \delta$. This results in a δ-error protocol for $\mathrm{FDISJ}_{m,s,t}$ which requires $t \ln(1/\delta) \cdot (t \cdot \mathrm{C}(f) + 2)$ bits of total communication from the players. Using Theorem 5, we have $t \ln(1/\delta) \cdot (t \cdot \mathrm{C}(f) + 2) = \Omega(s)$. Using $s = 2m/(3t)$ and $t = 2d$, we have $\mathrm{C}(f) = \Omega(m/d^3)$.

These are the desired lower bounds on $\mathrm{C}(f)$. □