

# Optimal Communication-Distortion Tradeoff in Voting

DEBMALYA MANDAL, Columbia University, USA

NISARG SHAH, University of Toronto, Canada

DAVID P. WOODRUFF, Carnegie Mellon University, USA

In recent work, Mandal et al. [32] study a novel framework for the winner selection problem in voting, in which a voting rule is seen as a combination of an elicitation rule and an aggregation rule. The elicitation rule asks voters to respond to a query based on their preferences over a set of alternatives, and the aggregation rule aggregates voter responses to return a winning alternative. They study the tradeoff between the communication complexity of a voting rule, which measures the number of bits of information each voter must send in response to its query, and its distortion, which measures the quality of the winning alternative in terms of utilitarian social welfare. They prove upper and lower bounds on the communication complexity required to achieve a desired level of distortion, but their bounds are not tight. Importantly, they also leave open the question whether the best randomized rule can significantly outperform the best deterministic rule.

We settle this question in the affirmative. For a winner selection rule to achieve distortion  $d$  with  $m$  alternatives, we show that the communication complexity required is  $\tilde{\Theta}\left(\frac{m}{d}\right)$  when using deterministic elicitation, and  $\tilde{\Theta}\left(\frac{m}{d^3}\right)$  when using randomized elicitation; both bounds are tight up to logarithmic factors. Our upper bound leverages recent advances in streaming algorithms. To establish our lower bound, we derive a new lower bound on a multi-party communication complexity problem.

We then study the  $k$ -selection problem in voting, where the goal is to select a set of  $k$  alternatives. For a  $k$ -selection rule that achieves distortion  $d$  with  $m$  alternatives, we show that the best communication complexity is  $\tilde{\Theta}\left(\frac{m}{kd}\right)$  when the rule uses deterministic elicitation and  $\tilde{\Theta}\left(\frac{m}{kd^3}\right)$  when the rule uses randomized elicitation. Our optimal bounds yield the non-trivial implication that the  $k$ -selection problem becomes strictly easier as  $k$  increases.

CCS Concepts: • **Theory of computation** → **Sketching and sampling**; **Algorithmic game theory and mechanism design**; *Communication complexity*.

Additional Key Words and Phrases: voting, sketching, distortion

## ACM Reference Format:

Debmalya Mandal, Nisarg Shah, and David P. Woodruff. 2020. Optimal Communication-Distortion Tradeoff in Voting. In *Proceedings of the 21st ACM Conference on Economics and Computation (EC '20)*, July 13–17, 2020, Virtual Event, Hungary. ACM, New York, NY, USA, 20 pages. <https://doi.org/10.1145/3391403.3399510>

## 1 INTRODUCTION

Making collective decisions through voting has been a subject of interest (at least) since the rise of democracy in ancient Athens. However, formal study of voting theory has more recent roots in the work of de Condorcet [21] in the late 18th Century. In the centuries subsequent to his work, social choice theorists pondered about the following canonical voting problem: If  $n$  voters express ranked preferences over a set of  $m$  alternatives, how should their preferences be aggregated to find the

---

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

EC '20, July 13–17, 2020, Virtual Event, Hungary

© 2020 Association for Computing Machinery.

ACM ISBN 978-1-4503-7975-5/20/07...\$15.00

<https://doi.org/10.1145/3391403.3399510>

most socially desirable alternative? However, the lack of an objective notion of what is “socially desirable” led to a plethora of voting rules being proposed, with no clear consensus, even among experts, as to which voting rule is the best [15].

In the recent decades, the marriage between social choice theory and computer science has given rise to the field of computational social choice [16], and one of the key influences of computer science has been to view voting as an optimization problem. Specifically, Procaccia and Rosenschein [37] proposed the framework of *implicit utilitarian voting*, whereby voters’ expressed ranked preferences over alternatives are seen as proxy for their underlying numerical utility functions. The overall framework consists of two steps. First, we must set an objective that one would want to optimize if voters’ numerical utility functions were known. For example, one goal could be to simply choose an alternative maximizing the sum of voters’ utilities (a.k.a. the *utilitarian social welfare*); this objective function has firm foundations in economic theory [11, 26, 28]. Given that it is impossible to perfectly optimize such an objective given the lack of complete information about voters’ utility functions, the next step is to seek the best *worst-case approximation* of the objective function that can be achieved given available information. This worst-case approximation is referred to as *distortion* in this framework. Arguably, this notion of a worst-case approximation is another key contribution of computer science to economic theory, which has led to successful paradigms such as algorithmic mechanism design [27, 36, 38] and the price of anarchy [31].

One benefit of this distortion framework is that it yields an *optimal* voting rule to aggregate voters’ expressed ranked preferences. Caragiannis et al. [17] and Boutilier et al. [14] identify the optimal deterministic and randomized aggregation rules, and show that their distortion is  $\Theta(m^2)$  and  $\tilde{\Theta}(\sqrt{m})$ , respectively, where  $m$  is the number of alternatives and  $\tilde{\Theta}$  hides logarithmic factors.

Later, Benadè et al. [9] observe that the implicit utilitarian voting framework has another benefit: not only can it be used to derive the optimal method to aggregate ranked votes, or  $k$ -approval votes, or votes expressed in any other *input format* for that matter, it can also be used to compare the efficacy of different input formats. Pushing this idea to the next level, Mandal et al. [32] propose optimizing both the input format and the vote aggregation method simultaneously. Specifically, they view a *winner selection* rule (i.e. a rule which returns a single winning alternative) as a combination of an elicitation rule, which specifies how voters should submit their votes in a certain format given their numerical utility functions, and an aggregation rule, which specifies how voters’ votes should be aggregated to find a single winning alternative. Then, they formally study the elicitation-distortion tradeoff: to achieve a desired distortion of  $d$ , what is the minimum number of bits of information that must be elicited from each voter about her utility function? For deterministic elicitation rules, they give an upper bound of  $\tilde{O}(m/d)$  and a lower bound of  $\Omega(m/d^2)$ . For randomized elicitation rules, they give a lower bound of  $\Omega(m/d^3)$ . We omit a discussion of the deterministic or randomized nature of the aggregation rule for now; a detailed discussion is presented in Section 7. Their work leaves open two key questions:

- What is the optimal communication complexity required to achieve distortion  $d$  in the winner selection problem using deterministic and randomized elicitation?
- Can the optimal winner selection rule with randomized elicitation significantly (i.e. beyond logarithmic factors) outperform that with deterministic elicitation?

In this paper, we answer both these questions by identifying the optimal elicitation-distortion tradeoff in the winner selection problem in voting.

We also examine the  $k$ -selection problem, where the goal is to return a set of  $k$  alternatives. This is a widely studied problem in voting, often known as *committee selection* or *multiwinner voting* [17, 22, 23]. For the classical setting where the elicitation rule is fixed to be the one which elicits ranked preferences, Caragiannis et al. [17] study the optimal distortion which can be achieved

using deterministic and randomized aggregation rules. But no prior work considers optimizing the elicitation rule, along with the aggregation rule, for this problem. We provide optimal bounds in this case as well.

## 1.1 Our Results

Let us briefly introduce our problem a bit more formally (a detailed model is presented in Section 2). For  $k \in \mathbb{N}$ , we define  $[k] \triangleq \{1, \dots, k\}$ . There is a set of voters  $N = [n]$  and a set of  $m$  alternatives  $A$ . Each voter  $i$  has a valuation function  $v_i : A \rightarrow \mathbb{R}_{\geq 0}$ . Following a standard assumption in voting theory [14, 17, 32], we assume normalized valuations, where  $\sum_{a \in A} v_i(a) = 1$  for each voter  $i$ . Given the vector of valuations  $\vec{v} = (v_1, \dots, v_n)$ , we want to maximize a certain objective function. However, eliciting real-valued  $v_i$  precisely requires asking voters to communicate potentially infinitely many bits of information. We are interested in examining how well we can approximate a given objective function in the worst case given a bound on the number of bits of information we are allowed to elicit from each voter. This worst-case approximation ratio is termed *distortion* in the literature, and the (expected) number of bits elicited from each voter is termed the *communication complexity*. We note that following traditional modeling, we assume that the rule asks the same “query” (i.e. how to map valuation function to a discrete response) to all voters. The model is more formally laid out in Section 2.

**1.1.1 Winner Selection.** In the winner selection (i.e. 1-selection) problem, we are interested in finding a single alternative with the goal of maximizing the social welfare:  $\text{sw}(a, \vec{v}) = \sum_{i \in N} v_i(a)$ . Specifically, we are interested in the communication complexity required to achieve distortion at most  $d$ , for a given  $d$ . For this problem, the results of Mandal et al. [32] Pareto dominate prior results in the literature. Hence, we only present a comparison of our results to theirs.

With deterministic elicitation, Mandal et al. propose a voting rule — **PREFTHRESHOLD** — which achieves an upper bound of  $\tilde{O}(m/d)$  communication complexity, and establish a weaker  $\Omega(m/d^2)$  lower bound. We show that their upper bound is tight up to logarithmic factors by proving that every winner selection rule with deterministic elicitation requires  $\Omega(m/d)$  communication complexity. We note that the **PREFTHRESHOLD** voting rule of Mandal et al. uses deterministic aggregation, whereas our lower bound holds even for randomized aggregation, thus establishing that with optimal deterministic elicitation, randomized aggregation does not provide significant benefit over deterministic aggregation for the winner selection problem.

With randomized elicitation, the story is inverted. Mandal et al. do not offer any better upper bounds than the  $\tilde{O}(m/d)$  achieved with deterministic elicitation (except in very restricted cases), but do offer a lower bound of  $\Omega(m/d^3)$ . In this case, we show that their lower bound is tight (up to logarithmic factors) by proposing a new winner selection rule with randomized elicitation which uses  $\tilde{O}(m/d^3)$  communication complexity.

Our optimal results imply that for the winner selection problem, randomized elicitation indeed offers a significant benefit (i.e. beyond logarithmic factors) over deterministic elicitation.

**1.1.2  $k$ -Selection.** In the  $k$ -selection problem, the goal is to select a set of  $k$  alternatives, where  $k$  is given. Trivially,  $k = 1$  is exactly the winner selection problem mentioned above. Hence, we focus on the case of  $k > 1$ . To fully specify this problem, we need to define the objective function we want to maximize. Following Caragiannis et al. [17], we say that the social welfare of a set  $S$  of  $k$  alternatives is  $\text{sw}(S, \vec{v}) = \sum_{i \in N} \max_{a \in S} v_i(a)$ . In other words, each voter  $i$  derives value from only her most favorite alternative in  $S$ . This formulation is applicable in contexts where the  $k$  alternatives act as substitutes. For example, in political elections for choosing a committee of  $k$  representatives, it is often assumed that a voter is “represented” by her most favorite candidate who is elected in the

committee [19, 34]. We follow this framework, albeit note that there are other equally interesting formulations of social welfare of a set of alternatives, which could lead to interesting future work (see Section 7).

For the  $k$ -selection problem, the only known bounds on distortion are those established by Caragiannis et al. [17]. They consider the specific elicitation rule which asks each voter to provide a ranking of the alternatives by value; this requires  $O(m \log m)$  bits of elicitation. As noted by Mandal et al. [32], rankings are not a very efficient form of elicitation in the winner selection problem: they allow achieving only  $\tilde{\Theta}(\sqrt{m})$  distortion with  $\Theta(m \log m)$  bits of communication (even with randomized aggregation), whereas their PREFTHRESHOLD method achieves  $O(1)$  distortion with just  $O(m \log \log m)$  bits of elicitation (and deterministic aggregation!). Our results imply that the same holds for the  $k$ -selection problem. Since our bounds significantly outperform those of Caragiannis et al., we omit a detailed presentation of their (complicated) bounds, and directly present our results.

With deterministic elicitation, we show that the optimal communication complexity required to achieve distortion  $d$  is  $\tilde{\Theta}(\frac{m}{kd})$ . Note that this bound decreases linearly as  $k$  increases. The same holds for randomized elicitation, which leads to optimal communication complexity of  $\tilde{\Theta}(\frac{m}{kd^3})$ .

We remark that apriori it is not even clear that the problem becomes easier as  $k$  increases. Increasing the value of  $k$  allows returning larger sets that achieve higher social welfare, but it also raises the *optimal* social welfare against which a voting rule needs to compete. For example, the bounds established by Caragiannis et al. [17] for ranked elicitation do not monotonically decrease with  $k$  (but they are also loose enough that they do not rule out the possibility of the optimal bounds for ranked elicitation decreasing monotonically with  $k$ ). A more detailed discussion is presented in Section 6.3.

## 1.2 Related Work

We begin by describing the results of Mandal et al. [32] in detail, as their work is the most relevant to ours. For deterministic elicitation, they construct an intuitive voting rule PREFTHRESHOLD, in which each voter is asked to report her approximate utility (with granularity parametrized by  $\ell$ ) for her  $t$  most preferred alternatives. Using a simple deterministic aggregation rule and by setting appropriate values of  $t$  and  $\ell$ , they show that distortion  $d$  can be achieved with  $\tilde{O}(m/d)$  bits of communication. Our result establishes this rule as asymptotically optimal (up to logarithmic factors) for deterministic elicitation. For randomized elicitation, they construct a voting rule RANDSUBSET, which outperforms PREFTHRESHOLD by logarithmic factors, but leave open the possibility of a rule that significantly outperforms PREFTHRESHOLD, only establishing a  $\Omega(m/d^3)$  lower bound. We show that their lower bound is tight (again, up to logarithmic factors) by constructing a voting rule that achieves it. We note that for certain cases, Mandal et al. provide exactly tight bounds; for example, they show that the optimal distortion with  $\log m$  bits of communication is  $\Theta(m^2)$ , achieved by the plurality voting rule. We are only concerned with optimality up to logarithmic factors.

More recently, Amanatidis et al. [2] also consider the elicitation-distortion tradeoff in voting. However, they take a *query complexity* approach to measuring communication. Specifically, they consider a setting where the preference rankings of the agents are known to the center, and the center can perform two types of queries: a *value query* asks a voter to report her precise utility for an alternative, and a *comparison query* asks a voter whether her utility for one alternative is at least  $x$  times her utility for another alternative. Our upper bound results are incomparable to theirs because they work with cardinal queries on top of known ordinal information. We do not assume any upfront knowledge about voters' preferences, and the set of possible queries are restricted to eliciting finitely many bits. At first glance, it may seem that our lower bounds carry over to their

framework. However, they allow adaptively asking different queries to different voters, whereas our lower bounds apply when the queries are common across voters.

When asking different questions to different voters is allowed, Caragiannis and Procaccia [18] showed that one can achieve significantly lower distortion using simple techniques (e.g.  $O(1)$  distortion with only  $\log m$  bits per voter). They show that this is possible via specific limited communication schemes, which they call embeddings. Recently [12] consider a setting where the center can ask a (randomized) threshold query to different voters with possibly different thresholds for different voters, and then the agents approve all the alternatives that they rank higher than this threshold. For this case, Bhaskar et al. [12] show that achieving constant distortion is possible even with *vanishing* number of bits per voter (specifically, with total number of bits independent of the number of voters). We argue that having a common ballot that all voters respond to is a natural assumption and is the most common practice for conducting voting in the real world.

Broadly, our work sits within the framework of implicit utilitarian voting in which no assumptions are made on voters' underlying numerical utility functions [10, 12, 14, 17, 37]. In certain contexts (especially for political elections), it is also common to assume that voters and alternatives lie in an underlying metric space, and voters' utilities (or costs) for alternatives respect the triangle inequality [1, 4, 5, 13, 24, 25, 35].

We also note that our use of sketching voter utility functions closely resembles the line of work on sketching combinatorial valuation functions [7, 8]. Their goal is to compress exponentially many numbers into polynomially many bits, whereas in our case, there are only polynomially many numbers (but infinitely many bits in exact representation) which need to be compressed.

To the best of our knowledge, ours is the first work to use sketching to design optimal voting rules. In particular, we use  $L_p$  samplers [29, 30, 33] which given a sequence of updates to an underlying vector, process the stream and finally output a coordinate proportional to its  $p$ -th power. Space-efficient  $L_p$ -samplers can be used for various applications, like moment estimation, or finding heavy hitters. The reader is referred to the work of Monemizadeh and Woodruff [33] and Andoni et al. [3] for further applications.

## 2 MODEL

For  $k \in \mathbb{N}$ , define  $[k] = \{1, \dots, k\}$ . Let  $x \sim D$  denote that random variable  $x$  has distribution  $D$ . Let  $\log$  denote logarithm to base 2,  $\ln$  denote logarithm to base  $e$ , and  $\text{med}$  denote the median.

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  $v_i : A \rightarrow \mathbb{R}_{\geq 0}$ , where  $v_i(a) \geq 0$  represents the value of voter  $i$  for alternative  $a$ . Equivalently, we view  $v_i \in \mathbb{R}_{\geq 0}^m$  as a vector which contains the voter's value for each alternative. Collectively, voter valuations are denoted by valuation profile  $\vec{v} = (v_1, \dots, v_n)$ . Given a valuation profile  $\vec{v}$ , the (utilitarian) social welfare of an alternative  $a$  is  $\text{sw}(a, \vec{v}) = \sum_{i \in N} v_i(a)$ . In the  $k$ -selection problem, we are interested in social welfare of a set of  $k$  alternatives. For a set  $S \subseteq A$ , define  $v_i(S) = \max_{a \in S} v_i(a)$  and  $\text{sw}(S, \vec{v}) = \sum_{i \in N} v_i(S)$ . Our goal is to elicit information about voter valuations, and use it to find an alternative (in the winner selection problem) or a set of  $k$  alternatives (in the  $k$ -selection problem) with high social welfare.

*Valuations:* We adopt the standard normalization assumption [6] that  $\sum_{a \in A} v_i(a) = 1$  for each  $i \in N$ . This is akin to a "one voter, one vote" principle. Alternatively, one can think of  $v_i(a)$  as the intensity of voter  $i$ 's relative preference for  $a$  as compared to other alternatives. Let  $\Delta^m$  denote the  $m$ -simplex, i.e., the set of all vectors in  $\mathbb{R}_{\geq 0}^m$  whose coordinates sum to 1. Hence, we have that  $v_i \in \Delta^m$  for each  $i \in N$ .

*Query space:* The literature on voting considers different types of responses, e.g., plurality votes,  $k$ -approval votes (which ask voters to report the set of their  $k$  favorite alternatives), threshold

approval votes (which ask voters to approve all alternatives for which their value is at least a given threshold), and ranked votes. Mandal et al. [32] unify these through the framework of communication complexity.

Consider an 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. Following Mandal et al. [32], we say that this interaction elicits  $\log k$  bits of information. It effectively partitions  $\Delta^m$  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  $\Delta^m$  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  $\Delta^m$  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  $\log k(q)$ .

*Voting Rule:* A voting rule consists of two parts: an *elicitation rule*  $\Pi_f$  and an *aggregation rule*  $\Gamma_f$ . The (randomized) elicitation rule  $\Pi_f$  is a distribution over  $\mathcal{Q}$ , according to which a query  $q$  is sampled. Each voter  $i$  provides a response  $\rho_i$  to this query, depending on her valuation  $v_i$ . We say that the elicitation rule is *deterministic* if it has singleton support (i.e., it chooses a query deterministically). The (randomized) aggregation rule  $\Gamma_f$ , takes voter responses  $\vec{\rho} = (\rho_1, \dots, \rho_n)$  as input, and returns an output. For the 1-selection problem, this is a distribution over the alternatives, and for the  $k$ -selection problem, this is a distribution over sets of  $k$  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 / subset returned by  $f$  when voter valuations are  $\vec{v} = (v_1, \dots, v_n)$ . We measure the performance of  $f$  via two metrics.

- (1) The *communication complexity* of  $f$  for  $m$  alternatives, denoted  $C^m(f) = \mathbb{E}_{q \sim \Pi_f}[\log k(q)]$ , 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  $\text{dist}^m(f)$ , 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,

$$\text{dist}(f) = \sup_{\vec{v} \in (\Delta^m)^n} \frac{\max_{a \in A} \text{sw}(a, \vec{v})}{\mathbb{E}_{\hat{a} \sim f(\vec{v})} \text{sw}(\hat{a}, \vec{v})}, \quad [1\text{-selection}]$$

$$\text{dist}(f) = \sup_{\vec{v} \in (\Delta^m)^n} \frac{\max_{S \subseteq A: |S|=k} \text{sw}(S, \vec{v})}{\mathbb{E}_{\hat{S} \sim f(\vec{v})} \text{sw}(\hat{S}, \vec{v})}. \quad [k\text{-selection}]$$

### 3 WINNER SELECTION: RANDOMIZED ELICITATION UPPER BOUND

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. Mandal et al. [32] study this trade-off for the winner selection problem. They propose a voting rule with deterministic elicitation and aggregation which achieves  $O(d)$  distortion with  $\tilde{O}(m/d)$  communication complexity, but their lower bound on communication complexity for achieving  $O(d)$  distortion with deterministic elicitation (and possibly randomized aggregation) is only  $\Omega(m/d^2)$ . For randomized elicitation, they do not propose any voting rule which improves upon their deterministic elicitation bound for general  $d$ , and present a weaker  $\Omega(m/d^3)$  lower bound.

In this and next section, we fill the gaps for both deterministic and randomized elicitation. We start by presenting our new voting rule with randomized elicitation and deterministic aggregation, which achieves distortion  $d$  with communication complexity  $\widetilde{O}(m/d^3)$ . This matches, up to logarithmic factors, the lower bound established by Mandal et al. [32] for randomized elicitation. Note that their lower bound holds even for randomized aggregation, whereas our rule achieves it with deterministic aggregation. The main tool we use in our improved algorithm is the notion of an  $L_p$ -sampler introduced by Monemizadeh and Woodruff [33]. An  $L_p$ -sampler processes a sequence of updates to an underlying vector  $x \in \mathbb{R}^m$  (with less memory than what is required to simply store  $x$ ). After processing the entire stream of updates, its goal is to output a random coordinate of  $x$  such that the  $i$ -th coordinate is sampled with probability approximately proportional to  $|x(i)|^p$ .

**Definition 1.** Let  $x \in \mathbb{R}^m$  and  $\delta > 0$  be a small constant. An  $L_2$ -sampler with relative error  $\nu$  is an algorithm that, with probability at least  $1 - \delta$ , outputs a coordinate  $\widehat{j}$  such that for any  $j \in [m]$ ,

$$\Pr[\widehat{j} = j] \in \left[ (1 - \nu) \frac{|x(j)|^2}{\|x\|_2^2}, (1 + \nu) \frac{|x(j)|^2}{\|x\|_2^2} \right] \pm O(m^{-c}),$$

where  $c \geq 1$  is an arbitrary constant. With the remaining probability (at most  $\delta$ ), the algorithm can output FAIL. When  $\nu = 0$ , this is known as a perfect  $L_2$ -sampler.

Our goal is to use such a sampler to sample according to the social welfare vector  $\text{sw} = \sum_{i \in N} v_i$ . However, each  $v_i$  is held privately by voter  $i$ . Hence, we need a multi-agent version of the  $L_2$ -sampler, which can obtain the required information from each agent  $i$  about her vector  $v_i$ , and then perform  $L_2$ -sampling on  $\text{sw} = \sum_{i \in N} v_i$ . This is where we crucially use the fact that the  $L_2$ -sampler of Jayaram and Woodruff [29] uses a linear “sketch”  $A$ . Therefore, we can obtain the linear sketch  $A(v_i)$  from each voter  $i$ , and combine these to compute  $A(\text{sw}) = \sum_{i \in N} A(v_i)$ . Let us describe the high-level template of the  $L_2$ -sampler of Jayaram and Woodruff [29]:

- (a) Duplicate the input  $x \in \mathbb{R}^m$  by copying each coordinate  $m^{c-1}$  times to obtain  $X \in \mathbb{R}^{m^c}$ , and then scale each coordinate by an i.i.d. random variable to get a vector  $\zeta \in \mathbb{R}^{m^c}$ .
- (b) Run the duplicated input  $X$  and the scaled input  $\zeta$  through the count-sketch algorithm [20] to get a sketch  $A(x)$ .
- (c) Select an index  $\widehat{j}$  using a statistic of  $A(x)$ .
- (d) Use a statistical test to determine whether to output  $\widehat{j}$  or to output FAIL.

---

**ALGORITHM 1:**  $L_2$ -Sampler

---

**Input:**  $\{x_i\}_{i=1}^n, \delta$ .

- (1) Run COMMUNICATE  $(\{x_i\}_{i=1}^n, \delta)$  to get  $A(x_i) = (A_i^1, A_i^2, A_i^3, A_i^4)$  from each agent  $i$ .
  - (2) Compute  $A^k = \sum_{i=1}^n A_i^k$  for  $k \in [4]$ .
  - (3) Run SAMPLE  $(\{A^k\}_{k \in [4]})$  to get  $(\text{STATUS}, (j^*, \widehat{x}(j^*)))$ .
  - (4) IF: STATUS == FAIL:
    - RETURN FAIL.
  - (5) ELSE:
    - RETURN  $(j^*, \widehat{x}(j^*))$ .
- 

In our multi-agent setup, we assume that there is a vector  $x_i$  held by each agent  $i$ , where  $x_i(j) \in \{j/\Delta : j \in [\Delta] \cup \{0\}\}$  for some fixed  $\Delta \in \mathbb{N}$ . We do not allow arbitrary real numbers because we are interested in analyzing the exact number of bits that each agent will need to communicate. We want to perform perfect  $L_2$ -sampling on the vector  $x = \sum_i x_i$ . To do so, we essentially require that each agent  $i$  run steps (a) and (b) to compute  $A(x_i)$  and communicate it to the center. Since

these sketches are linear, the center computes  $A(x) = \sum_i A(x_i)$ , and uses it to select a random index  $\hat{j}$  or output FAIL. The algorithm is presented as Algorithm 1. It first runs COMMUNICATE, which requires each agent to perform Steps (a) and (b) of the  $L_2$ -sampler, and communicate the obtained sketch  $A(x_i)$ . Then, it sums the sketches at the center, and finally uses SAMPLE to perform Steps (c) and (d) of the  $L_2$ -sampler to decide whether to return FAIL or an index  $j^*$  and an estimate of its value  $x_{j^*}$ . The details of the procedures COMMUNICATE and SAMPLE are provided in Algorithms 2 and 3 respectively; for further explanation of these procedures, we refer the reader to the work of Jayaram and Woodruff [29]. The purpose of the next theorem is to re-cast the guarantees provided by Jayaram and Woodruff [29] for their perfect  $L_2$ -sampler into a guarantee of our multi-agent variant, Algorithm 1.<sup>1</sup>

---

**ALGORITHM 2: COMMUNICATE**


---

**Input:**  $\{x_i\}_{i=1}^n, \delta$ .

- (1) Set  $d = \Theta(\log m)$ ,  $\eta = 1/\sqrt{\log m}$ , and  $\mu \sim \text{Unif}[1/2, 3/2]$ . Let  $c$  and  $c'$  be sufficiently large constants.
  - (2) Generate four independent hash tables and  $m^c$  i.i.d. random variables as follows.
    - a. Hash table  $A^1$  of dimension  $d \times 6/\eta^2$  used for selecting an alternative.
    - b. Hash tables  $A^2$  and  $A^3$  of dimensions  $c' \log(1/\delta) \times O(1)$  used for  $L_2$ -norm estimation.
    - c. Hash table  $A^4$  of dimension  $c' \log(1/\delta) \times O(1)$  used for total frequency estimation.
    - d. I.i.d. exponential random variables  $t_j$  for  $j \in [m^c]$  used for scaling valuations.
  - (3) Each agent  $i$  performs the following computation.
    - a. Duplicate  $x_i$  to get  $X_i \in \mathbb{R}^{m^c}$ .
    - b. Scale  $X_i$  as  $\zeta_i(j) = X_i(j)/t_j$  for  $j \in [m^c]$ .
    - c. Run count-sketches using hash tables  $A^1, A^2, A^3$ , and  $A^4$  on duplicated vector  $X_i$ , and the scaled vector  $\zeta_i$  to get  $A_i^1, A_i^2, A_i^3$ , and  $A_i^4$ .
  - (4) Each agent  $i$  sends  $A(x_i) = (A_i^1, A_i^2, A_i^3, A_i^4)$  to the center.
  - (5) RETURN  $A(x_i) = (A_i^1, A_i^2, A_i^3, A_i^4)$  for  $i \in [n]$ .
- 

---

**ALGORITHM 3: SAMPLE**


---

**Input:**  $\{A^k\}_{k=1}^4$ .

- (1) Compute  $A^k = \sum_{i=1}^n A_i^k$  for  $k \in [4]$ .
  - (2) Get an approximation  $y$  of  $\zeta = \sum_i \zeta_i$ , where  $y_j = \text{med}_{r \in [d]} g_r^1(j) A_{r, h_r^1(j)}^1$  for each  $j \in [m^c]$ .
  - (3) Get an approximation of  $\|X\|_2$  (where  $X = \sum_{i \in N} X_i$ ) by  $R = \text{med}_{r \in [c' \log(1/\delta)]} \|A^2[r, \cdot]\|_2$ .
  - (4) Get an approximation of  $\|\zeta\|_2$  by  $R' = \text{med}_{r \in [c' \log(1/\delta)]} \|A^3[r, \cdot]\|_2$ .
  - (5) Let  $y_{(k)}$  denote the  $k^{\text{th}}$  largest value in  $y$ .
  - (6) IF:  $y_{(1)} - y_{(2)} < 100\mu\eta R + \eta R'$  OR  $y_{(2)} < 50\eta\mu R$ :
    - RETURN (FAIL,  $\emptyset$ ).
  - (7) ELSE:
    - Let  $\hat{j} \in \arg \max_{j \in [m^c]} y_j$ . Let  $j^*$  be the corresponding non-duplicated index.
    - Get an approximation  $\tilde{y}$  of  $\zeta$ , where  $\tilde{y}_j = \text{med}_{r \in [d]} g_r^4(j) A_{r, h_r(j)}^4$  for each  $j \in [m^c]$ .
    - Compute an approximation of  $x(j^*)$ :  $\hat{x}(j^*) = t_{\hat{j}} \times \tilde{y}_{\hat{j}}$ .
    - RETURN (SUCCESS,  $(j^*, \hat{x}(j^*))$ ).
- 

<sup>1</sup>All the missing proofs are provided in the full version of the paper available from the authors' web pages.



**THEOREM 1.** Let  $\Delta \in \mathbb{N}$ ,  $c$  be a sufficiently large constant, and  $\delta \geq 1/\text{poly}(m)$ . Suppose each agent  $i$  holds a vector  $x_i$  such that  $x_i(j) \in \{0, 1/\Delta, 2/\Delta, \dots, 1\}$  for each  $j \in [m]$ . Let  $x = \sum_i x_i$ . Then, Algorithm 1 outputs FAIL with probability at most  $\delta$ , and with the remaining probability, its output  $(j^*, \hat{x}(j^*))$  satisfies the following two conditions.

- For each  $j \in [m]$ ,

$$\Pr[j^* = j] = \frac{x(j)^2}{\|x\|_2^2} \pm O(m^{-c}).$$

- Conditioned on the event that  $j^* = j$ ,

$$\hat{x}(j^*) \in \left[ \frac{1}{2} \cdot x(j), 2 \cdot x(j) \right].$$

Moreover, under this algorithm, each agent communicates  $O(\log^2 m \log(\Delta \log m))$  bits.

Next, we want to use this algorithm to design our voting rule. The elicitation rule of our voting rule is simple: each voter  $i$  approximates her valuation function  $v_i$  to  $v_i^\Delta$  by rounding each  $v_i(a)$  to the nearest multiple of  $1/\Delta$ , and then sends the sketch  $A(v_i^\Delta)$  as specified by Algorithm 1. We show that for an appropriately chosen  $\Delta$ , the error introduced in this step does not significantly affect the final result. This is one of the contributions of the next result.

Another contribution is that instead of requesting just one sketch from each voter, we generate  $t$  independent sketch functions (for an appropriately chosen  $t$ ), and request the corresponding sketches from each voter. Recall that our randomized voting rule is allowed to select a random query  $q$  which maps each voter's valuation to a response and ask voters to respond to  $q$ . Communicating query  $q$  to the voters is free of cost. Equivalently, one can imagine that there is a public tape, and the voting rule can write any information required to represent query  $q$  on this tape, free of cost. Hence, to request these sketches, the voting rule generates four random hash functions as well as  $m^c$  i.i.d. exponential random variables for each of  $t$  sketches, and writes them onto the public tape, so voters know exactly which sketches they are supposed to compute.

In the aggregation rule, we run the  $t$  samplers on the combined sketches obtained from the voters. If any of those samplers fail (we show this happens with a low probability), then we simply return an arbitrary alternative. Otherwise, we return the alternative for which the corresponding estimated count returned by our sampler is the highest. The voting rule is formally presented as Algorithm 4. A key contribution of the next result is to show that using  $L_2$ -samplers to sample alternatives according to the square of their social welfare helps achieve the optimal distortion. To the best of our knowledge, this is the first result using  $L_2$ -samplers to design a voting rule. Although the elicitation rule closely follows the  $L_2$ -sampler designed by [29], the novelty lies in designing the aggregation rule and analyzing the distortion achieved.

**THEOREM 2.** For any  $d$ , MAX- $L_2$ -SAMPLER achieves  $\text{dist}(\text{MAX-}L_2\text{-SAMPLER}) = O(d)$  with

$$C(\text{MAX-}L_2\text{-SAMPLER}) = O\left(\frac{m}{d^3} \log^3 m\right).$$

**PROOF.** Recall that our parameter choices are  $\Delta = 128m^3$ , and  $t = 4m/d^3$ . Let  $E$  denote the event that none of  $t$  independent  $L_2$ -samplers fail. Since each  $L_2$ -sampler is run with failure probability  $\delta = 1/(4t) \geq 1/\text{poly}(m)$ , the probability that each  $L_2$ -sampler fails is at most  $1/(4t)$ . Hence, by the union bound, the probability of the event  $E$  is at least  $1 - t \times 1/(4t) = 3/4$ . Since the expected social welfare achieved by the voting rule is at least  $3/4$  times the expected social welfare achieved conditioned on  $E$ , we condition on  $E$  being true for the rest of the proof. The final distortion can only increase by a factor of at most  $4/3$ .

---

**ALGORITHM 4: MAX- $L_2$ -SAMPLER**

---

**Elicitation Rule:**

- Set  $t = 4m/d^3$  and  $\Delta = 128m^3$ .
- Each voter  $i$  computes an approximate valuation  $v_i^\Delta$ , where, for each alternative  $a \in A$ ,  $v_i^\Delta(a)$  is  $v_i(a)$  rounded to the nearest multiple of  $1/\Delta$ .
- Run  $t$  independent copies of the procedure COMMUNICATE  $\left(\{v_i^\Delta\}_{i=1}^n, \frac{1}{4t}\right)$  of Algorithm 1.
- Obtain responses  $A^k(v_i^\Delta)$  for each  $k \in [t]$  and  $i \in [n]$ .

**Aggregation Rule:**

- For each  $k \in [t]$ , run procedure SAMPLE  $\left(\{A^k(v_i^\Delta)\}_{i \in [n]}\right)$  of Algorithm 1 to obtain either FAIL or a pair  $(a^k, \widehat{\text{sw}}(a^k))$ .
- If at least one algorithm returns FAIL, then output an arbitrary alternative.
- Otherwise, return  $\widehat{a} = a^{k^*}$ , where  $k^* \in \arg \max_{k \in [t]} \widehat{\text{sw}}(a^k)$ .

**Communication Complexity:**

$$C(\text{MAX-}L_2\text{-SAMPLER}) = O\left(\frac{m}{d^3} \log^3 m\right).$$

**Distortion:**

$$\text{dist}(\text{MAX-}L_2\text{-SAMPLER}) = O(d).$$

---

Since our algorithm just calls  $t$   $L_2$ -samplers in parallel, using Theorem 1, the communication complexity of the voting rule is clearly bounded by  $O(t \log^2 m \log(m^3 \log m)) = O\left(\frac{m}{d^3} \log^3 m\right)$ . It remains to show that its distortion is  $O(d)$ .

Let  $(a^k, \widehat{\text{sw}}(a^k))$  denote the output of the  $k$ -th  $L_2$  sampler. Throughout the proof, we will use three notions of welfare: (1) the *true welfare*  $\text{sw}(a) = \sum_{i=1}^n v_i(a)$ , (2) the *rounded welfare*  $\text{sw}^\Delta(a) = \sum_{i=1}^n v_i^\Delta(a)$  where  $v_i^\Delta(a)$  is  $v_i(a)$  rounded to the nearest multiple of  $1/\Delta$ , and (3) the *estimated welfare*  $\widehat{\text{sw}}(a^k)$  returned by the  $k$ -th  $L_2$ -sampler, which is an estimate of  $\text{sw}^\Delta(a^k)$ . We will write  $\text{sw}$  and  $\text{sw}^\Delta$  to denote the vectors of true and rounded welfare.

First we notice the following obvious relationship between  $\text{sw}(a)$  and  $\text{sw}^\Delta(a)$ .

$$\forall a \in A : |\text{sw}^\Delta(a) - \text{sw}(a)| = \left| \sum_{i=1}^n v_i^\Delta(a) - \sum_{i=1}^n v_i(a) \right| \leq \sum_{i=1}^n |v_i^\Delta(a) - v_i(a)| \leq \frac{n}{\Delta} \quad (1)$$

Let  $a^* \in \arg \max_{a \in A} \text{sw}(a)$  be an alternative with the highest social welfare. Next, we show a lower bound on the expected social welfare of the alternative  $\widehat{a}$  returned by our voting rule. The proof of the next lemma is given in the appendix.

**Lemma 1.**  $\mathbb{E}[\text{sw}(\widehat{a})] \geq \frac{1}{256} \frac{n}{m^3}$ .

Finally, to derive an upper bound on the distortion (i.e. to derive an upper bound on  $\text{sw}(a^*)/\mathbb{E}[\text{sw}(\widehat{a})]$ ), we consider two cases.

**Case 1:** Suppose  $\text{sw}^\Delta(a^*) \geq \frac{\|\text{sw}^\Delta\|_2}{\sqrt{td/3}}$ . In this case,  $\text{sw}^\Delta(a^*)^2 \geq \frac{3\|\text{sw}^\Delta\|_2^2}{td}$ . Since each  $a^k$  is generated by a perfect  $L_2$  sampler, we have

$$\Pr[a^k = a^*] \geq \frac{\text{sw}^\Delta(a^*)^2}{\|\text{sw}^\Delta\|_2^2} - O(m^{-c}) \geq \frac{3}{td} - O(m^{-c}).$$

Note that  $td = 4m/d^2 = O(m)$ . Hence, for a sufficiently large  $c$ , we can ensure that this probability is at least  $2/td$ . Therefore, the probability that none of  $a^k$ ,  $k \in [t]$ , are equal to  $a^*$  is at most

$(1 - \frac{2}{td})^t \leq 1 - \frac{1}{d}$ . This implies that with probability at least  $1/d$ ,  $a^*$  appears as  $a^k$  for at least one  $k \in [t]$ . Since we select the final alternative  $\hat{a}$  as  $a^k$  with the highest estimated welfare  $\widehat{\text{sw}}$ , we have

$$\begin{aligned} \mathbb{E}[\text{sw}(\hat{a})] &\geq \mathbb{E}[\text{sw}^\Delta(\hat{a})] - \frac{n}{\Delta} \geq \frac{1}{2} \cdot \mathbb{E}[\widehat{\text{sw}}(\hat{a})] - \frac{n}{\Delta} \geq \frac{1}{2d} \cdot \mathbb{E}[\widehat{\text{sw}}(a^*)] - \frac{n}{\Delta} \\ &\geq \frac{1}{4d} \cdot \text{sw}^\Delta(a^*) - \frac{n}{\Delta} \geq \frac{1}{4d} \cdot \text{sw}(a^*) - \frac{n}{\Delta} \left(1 + \frac{1}{4d}\right). \end{aligned}$$

The first inequality follows from Equation (1). The second inequality follows from the guarantee in Theorem 1. The third inequality follows because  $\widehat{\text{sw}}(\hat{a}) \geq \widehat{\text{sw}}(a^k)$  for each  $k \in [t]$ , and  $a^* \in \{a^k : k \in [t]\}$  with probability at least  $1/d$ . The fourth inequality again uses Theorem 1. The final inequality again uses Equation (1). Rearranging, we get the following bound on the distortion.

$$\frac{\text{sw}(a^*)}{\mathbb{E}[\text{sw}(\hat{a})]} \leq 4d + \frac{\frac{n}{\Delta} \left(1 + \frac{1}{4d}\right)}{\mathbb{E}[\text{sw}(\hat{a})]} \leq 4d + 2 \left(1 + \frac{1}{4d}\right) \leq 8d,$$

where the second inequality follows from substituting  $\Delta = 128m^3$  and using Lemma 1.

**Case 2:** Suppose  $\text{sw}^\Delta(a^*) < \frac{\|\text{sw}^\Delta\|_2}{\sqrt{td/3}}$ . Fix any  $k \in [t]$ . We claim that with probability at least  $1/2$ , we have  $(\text{sw}^\Delta(a^k))^2 / \|\text{sw}^\Delta\|_2^2 \geq \frac{1}{2m}$ . This is because every alternative  $a$  with  $(\text{sw}^\Delta(a))^2 / \|\text{sw}^\Delta\|_2^2 \leq \frac{1}{2m}$  is picked with probability at most  $\frac{1}{2m}$  and there are at most  $m$  such alternatives.

Thus, for each  $k \in [t]$ , the following holds.

$$\begin{aligned} \mathbb{E}[\text{sw}(a^k)] &\geq \mathbb{E}[\text{sw}^\Delta(a^k)] - \frac{n}{\Delta} \geq \frac{1}{2} \cdot \frac{\|\text{sw}^\Delta\|_2}{\sqrt{2m}} - \frac{n}{\Delta} \\ &\geq \frac{1}{9} \cdot \sqrt{\frac{td}{m}} \cdot \text{sw}^\Delta(a^*) - \frac{n}{\Delta} \geq \frac{1}{9} \cdot \sqrt{\frac{td}{m}} \cdot \text{sw}(a^*) - \frac{n}{\Delta} \left(1 + \frac{1}{9} \sqrt{\frac{td}{m}}\right). \end{aligned}$$

The first and the final inequalities use Equation (1), and the third inequality uses the fact that we are in the case of  $\|\text{sw}^\Delta\|_2 > \text{sw}^\Delta(a^*) \cdot \sqrt{td/3}$ .

Since the final alternative  $\hat{a} = a^k$  for some  $k \in [t]$ , we get the following bound on the distortion.

$$\frac{\text{sw}(a^*)}{\mathbb{E}[\text{sw}(\hat{a})]} \leq 9 \sqrt{\frac{m}{td}} + \frac{\frac{n}{\Delta} \left(1 + \frac{1}{9} \sqrt{\frac{td}{m}}\right)}{\mathbb{E}[\text{sw}(\hat{a})]} \leq \frac{9d}{2} + 2 \left(1 + \frac{2}{9d}\right) \leq 9d,$$

where the second inequality follows from substituting  $\Delta = 128m^3$  and  $t = 4m/d^3$ , and using the lower bound on  $\mathbb{E}[\text{sw}(\hat{a})]$  from Lemma 1.  $\square$

## 4 WINNER SELECTION: DETERMINISTIC ELICITATION LOWER BOUND

In this section, we derive an  $\Omega(m/d)$  lower bound on the communication complexity of voting rules which achieve distortion at most  $d$  using deterministic elicitation (and possibly randomized aggregation). Mandal et al. [32] provide an upper bound of  $\tilde{\Theta}(m/d)$ , establishing that our lower bound is optimal up to logarithmic factors.

To prove our lower bound, we use a reduction to the multi-party set disjointness problem. To keep the paper self-contained, we begin with a brief background of the multi-party communication complexity literature.

### 4.1 Setup of Multi-Party Communication Complexity

In multi-party communication complexity, there are  $t$  players, denoted 1 through  $t$ . Each player  $i$  holds a private input  $X_i \in \mathcal{X}_i$ . We refer to  $(X_1, \dots, X_t)$  as the *input profile*. The players are assumed

to be computationally omnipotent and limited only in their communication capabilities. The goal is to compute the output of a function  $f : \mathcal{X}_1 \times \mathcal{X}_2 \times \dots \times \mathcal{X}_t \rightarrow \{0, 1\}$  on an input profile. To compute this function, players communicate messages about their private input. We assume a *blackboard* model, in which each player writes her messages on the blackboard, and they are visible to all other players for free.

In this framework, a *deterministic protocol*  $\Gamma$  specifies how players should write messages on the blackboard given their private input and any previous messages they see on the blackboard. We use  $\Gamma(X_1, \dots, X_t)$  to denote the transcript generated on the blackboard from messages written by all players given input profile  $(X_1, \dots, X_t)$ .

**Definition 2** (Deterministic Communication Cost). The *deterministic communication cost* of protocol  $\Gamma$ , denoted by  $D(\Gamma)$ , is the maximum length of the transcript  $\Gamma(X_1, \dots, X_t)$ , where the maximum is taken over all input profiles  $(X_1, \dots, X_t)$ .

We say that  $\Gamma$  is a protocol for  $f$  if there exists a function  $\Pi_{\text{out}}$  mapping the set of possible transcripts to  $\{0, 1\}$  such that  $\Pi_{\text{out}}(\Pi(X_1, \dots, X_t)) = f(X_1, \dots, X_t)$  for every input profile  $(X_1, \dots, X_t)$ .

**Definition 3** (Deterministic Communication Complexity). The *deterministic communication complexity* of  $f$ , denoted  $D(f)$ , is the deterministic communication cost of the best deterministic protocol for  $f$ , i.e.  $D(f) = \min_{\Gamma : \Gamma \text{ is a protocol for } f} D(\Gamma)$ .

We now briefly review randomized protocols and randomized communication complexity. We assume *public randomness*. That is, players have free access to a shared stream of random coin tosses. A *randomized protocol*  $\Pi$  specifies how players should write messages on the blackboard given their private input, random coin tosses, and any previous messages they see on the blackboard. Let  $\Pi(X_1, \dots, X_t)$  be the random variable that denotes the transcript generated when all the players follow the protocol given input profile  $(X_1, \dots, X_t)$ ; here the randomness is due to the public coin tosses.

**Definition 4** (Randomized Communication Cost). The *randomized communication cost* of protocol  $\Pi$ , denoted by  $R(\Pi)$ , is the maximum length of the transcript  $\Pi(X_1, \dots, X_t)$ , where the maximum is taken over all input profiles  $(X_1, \dots, X_t)$ , and all public coin tosses.

Given  $\delta \in [0, 1]$ , we say that  $\Pi$  is a  $\delta$ -*error protocol* for  $f$  if there exists a function  $\Pi_{\text{out}}$  mapping the set of possible transcripts to  $\{0, 1\}$  such that  $\Pr[\Pi_{\text{out}}(\Pi(X_1, \dots, X_t)) = f(X_1, \dots, X_t)] \geq 1 - \delta$  for every input profile  $(X_1, \dots, X_t)$ ; again, the randomness here is over the public coin tosses.

**Definition 5** (Randomized Communication Complexity). The  $\delta$ -*error randomized communication complexity* of  $f$ , denoted  $R_\delta(f)$ , is the randomized communication cost of the best  $\delta$ -error randomized protocol for  $f$ , i.e.  $R_\delta(f) = \min_{\Pi : \Pi \text{ is a } \delta\text{-error protocol for } f} R(\Pi)$ .

## 4.2 Multi-Party Fixed-Size Set Disjointness

Our main tool is to analyze the multi-party fixed-size set disjointness problem [32], which is a refinement of the classic multi-party set disjointness problem.

**Definition 6** (Multi-Party (Fixed-Size) Set Disjointness). In the classic *multi-party set disjointness* problem, denoted  $\text{DISJ}_{m,t}$ , there are  $t$  players and a universe of  $m$  elements. Each player  $i$  holds a subset  $S_i$  of the universe. The goal is to determine whether all sets are pairwise disjoint (i.e.  $S_i \cap S_j = \emptyset$  for all distinct  $i, j \in [t]$ ). If this is the case in an input, it is referred to as a *NO* instance; otherwise, it is referred to as a *YES* instance.

In multi-party fixed-size set disjointness, denoted  $\text{FDISJ}_{m,s,t}$ , there is an additional parameter  $s \in [m]$  such that the set held by each player has size exactly  $s$  (i.e.  $|S_i| = s$  for each  $i \in [t]$ ).

Often, set disjointness is studied under a promise, which allows assuming additional structure of YES instances (equivalently, the protocol is free to return any answer on YES instances without this structure).

**Definition 7** (Unique Intersection Promise). The classical promise in the literature is the *unique intersection* promise, which guarantees that in every YES instance, there exists an element  $x$  such that  $x \in S_i$  for each  $i \in [t]$  and  $(S_i \setminus \{x\}) \cap (S_j \setminus \{x\}) = \emptyset$  for all distinct  $i, j \in [t]$ .

We refer to such an element  $x$  as a *common element*. Mandal et al. [32] show that  $R_\delta(\text{FDISJ}_{m,m/t,t}) = \Omega(m/t)$  under the unique intersection promise, and then they give a reduction from multi-party fixed-size set disjointness to voting. Specifically, they use a voting rule  $f$  with deterministic elicitation to construct a 0-error protocol for  $\text{FDISJ}_{m,\Theta(m/d),\Theta(d)}$  that uses roughly  $\Theta(d) \cdot C(f)$  bits of total elicitation. Then, using  $R_\delta(\text{FDISJ}_{m,\Theta(m/d),\Theta(d)}) = \Omega(m/d)$ , they derive  $C(f) = \Omega(m/d^2)$ .<sup>2</sup> In order to improve the lower bound on the communication complexity of a voting rule, we first show that  $D(\text{FDISJ}_{m,m/t,t}) = \Omega(m)$ . The change from randomized communication complexity to its deterministic counterpart is not an issue because the reduction of Mandal et al. [32] eventually constructs a deterministic protocol anyway. The improvement from  $\Omega(m/t)$  to  $\Omega(m)$  helps shave off a factor of  $t = \Theta(d)$ . However, this requires moving from the strong unique intersection promise to the weaker *substantial intersection promise*, which we define next.

**Definition 8** (Substantial Intersection Promise). We introduce a weaker promise, which we refer to as the *substantial intersection* promise, which guarantees that for a given constant  $\gamma > 0$ , in every YES instance, there exists at least one element  $x$  and a subset of players  $P \subseteq [t]$  with  $|P| \geq \gamma \cdot t$  such that  $x \in S_i$  for each  $i \in P$ .

Note that unique intersection promise provides a stronger guarantee than substantial intersection promise; hence, any communication complexity lower bound under the former promise immediately implies the same lower bound under the latter promise.

### 4.3 Lower Bound on Communication Complexity of Voting Rules

We begin by establishing an improved lower bound on the deterministic communication complexity of  $\text{FDISJ}$  under the weaker substantial intersection promise.

**THEOREM 3.** *Under the substantial intersection promise with  $\gamma \leq 1/76$  and  $t \leq (m/2) \cdot (1 - 1/e)$ , we have  $D(\text{FDISJ}_{m,m/t,t}) \geq m$ .*

**PROOF.** Suppose for contradiction that there exists a deterministic protocol  $\Gamma$  for  $\text{FDISJ}_{m,m/t,t}$  with  $D(\Gamma) = r < m$ . For  $i \in [t]$ , let  $\mathcal{X}_i = \{S_i : S_i \subseteq [m] \wedge |S_i| = m/t\}$  be the collection of possible sets held by player  $i$ , and let  $\mathcal{X} = \{(S_1, \dots, S_t) : S_i \subseteq [m] \wedge |S_i| = m/t, \forall i \in [t]\}$  be the collection of possible input profiles.

A subset  $\mathcal{S} \subseteq \mathcal{X}$  of input profiles is called a *rectangle* if  $\mathcal{S} = \prod_{i \in [t]} \mathcal{S}_i$  where  $\mathcal{S}_i \subseteq \mathcal{X}_i$  for all  $i \in [t]$ . A rectangle  $\mathcal{S}$  is *monochromatic* with respect to deterministic protocol  $\Gamma$  if  $\Gamma$  generates identical transcripts on all input profiles in  $\mathcal{S}$ . Since we are working with the *blackboard* model and at most  $r$  bits are communicated under  $\Gamma$  in the worst case, it is easy to observe that  $\Gamma$  partitions  $\mathcal{X}$  into at most  $2^r$  monochromatic rectangles.<sup>3</sup>

<sup>2</sup>The additional factor of  $d$  appears because  $R_\delta(\text{FDISJ}_{m,s,t})$  measures the *total* communication from all players, whereas  $C(f)$  measures the communication from each player. So a lower bound of  $\Omega(m/d)$  on the *total* communication translates to a lower bound of  $\Omega(m/d^2)$  on the communication from each of  $t = \Theta(d)$  players in their reduction.

<sup>3</sup>This is a standard argument. Note that  $\Gamma$  can equivalently be described by a binary tree, where each internal node represents one of the players writing the next bit on the blackboard. Since at most  $r$  bits are written, the tree has at most  $2^r$  leaves, and each leaf is obtained on a set of input profiles that form a monochromatic rectangle.

Under a NO instance, each of  $t$  players hold a disjoint subset of size  $m/t$ . Hence, the total number of NO instances is  $m!/((m/t)!)^t$ . Thus, by the pigeonhole principle, at least one of the monochromatic rectangles (call it  $\mathcal{S}^*$ ) must contain at least  $m!/(2^r((m/t)!)^t)$  NO instances. Hence,

$$\begin{aligned} |\mathcal{S}^*| &\geq \frac{m!}{2^r ((m/t)!)^t} > \frac{m^{m+1/2} e^{-m}}{2^m (e(m/t)m^{1/2} e^{-m/t})^t} = \frac{m^{m+1/2} e^{-m}}{2^m e^t (m/t)^{m+1/2} e^{-m}} = \frac{m^{1/2} \cdot t^m \cdot t^{t/2}}{2^m \cdot e^t \cdot m^{t/2}} \\ &\geq \frac{t^m}{2^m \cdot e^{m/2}} = \left( \frac{t}{2\sqrt{e}} \right)^m. \end{aligned} \quad (2)$$

Here, the first inequality holds due to Stirling's approximation<sup>4</sup> and the fact that  $r < m$ , and the second inequality holds because for  $t \leq (m/2) \cdot (1 - 1/e)$ , we have

$$\frac{t^{t/2}}{e^t m^{t/2}} = \frac{1}{e^t (m/t)^{t/2}} \geq \frac{1}{e^{m/2 \cdot (1-1/e)} e^{m/(2e)}} = \frac{1}{e^{m/2}}.$$

Here, we used the fact that  $(m/t)^t$  achieves its maximum value at  $t = m/e$ , and hence is bounded from above by  $e^{m/e}$ .

Now, let us write the rectangle  $\mathcal{S}^* = \prod_{i \in [t]} \mathcal{S}_i^*$ , where  $\mathcal{S}_i^* \subseteq \mathcal{X}_i$  for each  $i \in [t]$ . For each  $i \in [t]$ , let us also define  $\text{supp}(\mathcal{S}_i^*) = \bigcup_{S_i: S_i \in \mathcal{S}_i^*} S_i$  to be the set of all elements which appear in at least one set in  $\mathcal{S}_i^*$ . Then, we have the following upper bound on the number of possible instances in  $\mathcal{S}^*$ .

$$|\mathcal{S}^*| = \prod_{i \in [t]} \binom{|\mathcal{S}_i^*|}{m/t} \leq \prod_{i \in [t]} \left( \frac{|\mathcal{S}_i^*| e t}{m} \right)^{m/t}. \quad (3)$$

We now claim that at least  $t/4$  players  $i$  must have  $|\mathcal{S}_i^*| \geq m/19$ . Suppose this is not true. That is, all but at most  $t/4$  players  $i$  have  $|\mathcal{S}_i^*| \leq m/19$ , and the remaining at most  $t/4$  players  $i$  have  $|\mathcal{S}_i^*| \leq m$ . Then, by Equation (3), we would have

$$|\mathcal{S}^*| \leq \left( (et)^{m/t} \right)^{t/4} \cdot \left( \left( \frac{et}{19} \right)^{m/t} \right)^{3t/4} \leq \left( \frac{et}{19^{3/4}} \right)^m < \left( \frac{t}{2\sqrt{e}} \right)^m,$$

which would contradict Equation (2).

Hence, there exist at least  $t/4$  players  $i$  with  $|\text{supp}(\mathcal{S}_i^*)| \geq m/19$ . Thus, by the pigeonhole principle, at least one element  $x^*$  must appear in  $\text{supp}(\mathcal{S}_i^*)$  for at least  $t/76 \geq \gamma t$  players  $i$ . Selecting a corresponding set from  $\mathcal{S}_i^*$  containing  $x^*$  for each such player  $i$  and an arbitrary set from  $\mathcal{S}_i^*$  for each remaining player  $i$  generates a YES instance under the substantial intersection promise, and this YES instance also belongs to  $\mathcal{S}^*$  which also contains at least one NO instance. Hence, protocol  $\Gamma$  would not be able to distinguish these YES and NO instances, establishing the contradiction. This shows that  $D(\text{FDISJ}_{m, m/t, t}) \geq m$ .  $\square$

We now use Theorem 3 to derive a lower bound on the communication complexity of voting rules with deterministic elicitation.

**THEOREM 4.** *For any  $d$ , if voting rule  $f$  uses deterministic elicitation and satisfies  $\text{dist}(f) \leq d$ , then  $C(f) = \Omega(m/d)$ .*

Note that our lower bound of  $\Omega(m/d)$  from Theorem 4 applies even when the rule is allowed to use randomized aggregation, whereas Mandal et al. [32] achieve a matching upper bound of  $\widetilde{O}(m/d)$  using deterministic aggregation. This establishes that, when using deterministic elicitation, there is

<sup>4</sup>For all  $n \in \mathbb{N}$ ,  $\frac{n!}{n^{n+1/2} e^{-n}} \in [1, e]$ .

no significant asymptotic benefit of using randomized aggregation. This matches our observation for the case of randomized elicitation.

However, such observation applies only when we are considering both the optimal elicitation and optimal aggregation rule. It need not be true when considering a fixed (possibly suboptimal) elicitation method. For example, when each voter sends a ranking of the alternatives by their value for her, it is known that randomized aggregation yields optimal distortion of  $\tilde{\Theta}(\sqrt{m})$  [14], which is significantly lower than the optimal distortion of  $\Theta(m^2)$  with the deterministic aggregation [17].

## 5 $k$ -SELECTION: UPPER BOUNDS

We now turn to the  $k$ -selection problem, where the goal is to select a set  $S \subseteq A$  of  $|S| = k$  alternatives. Recall that the value that voter  $i$  derives from set  $S$  is defined as  $v_i(S) = \max_{a \in S} v_i(a)$ . In Section 5.1 below, we present a deterministic  $k$ -selection rule which achieves a distortion of at most  $d$  with communication complexity  $\tilde{O}(m/(kd))$ . Then, in Section 5.2, we present a randomized  $k$ -selection rule with distortion  $d$  and communication complexity  $\tilde{O}(m/(kd^3))$ . Later, in Section 6, we show that these bounds are almost tight. While these results for the  $k$ -selection problem subsume those for the winner selection problem ( $k = 1$ ) from the previous sections, they require much more intricate algorithms and techniques, and have larger logarithmic factors hidden within the  $\tilde{O}$  notation.

In Section 3, we used  $L_2$ -sampling to design a new winner selection with randomized elicitation that achieves the optimal trade-off between communication complexity and distortion. Unfortunately, it seems that  $L_2$ -sampling (or more generally,  $L_p$ -sampling) is not well-suited for the  $k$ -selection problem. The difficulty is that in the  $k$ -selection problem, a voter's value for a set of alternatives is defined as her *maximum* value for any alternative in the set. This inevitably leads to the structure of a coverage problem, where, after choosing one alternative in the desired set, a good choice of the next alternative crucially depends on the alternative just chosen. It is unlikely to be able to obtain a set of alternatives that *collectively* provide "good coverage" by taking *independent*  $L_2$ -samples obtained from a sampler. Hence, we use different techniques to design  $k$ -selection rules.

### 5.1 Deterministic Elicitation

Let us first focus on designing a  $k$ -selection rule with deterministic elicitation. The detailed algorithm is provided in the full version. Here, we discuss the main idea behind the algorithm. We partition the range of possible values that a voter might have for an alternative into  $\log m$  exponentially spaced "value-buckets", and observe that there is one bucket such that if we zero-out valuations that are not in that bucket, the optimal set with respect to the modified valuations is still an  $O(\log m)$  approximation to the original optimal set. Furthermore, for each voter, we can look at the number of alternatives for which the voter's value is in this bucket. This number can be further placed into  $\log m$  exponentially spaced "quantity-buckets". We again show that there exists one quantity bucket such that only considering this bucket further loses only a factor of  $O(\log m)$  social welfare. Once we are restricted to a fixed value-bucket and a fixed quantity-bucket, the problem is simple: if the size of the quantity-bucket is large ( $\geq \frac{m}{kd}$ ), then we show that a uniformly random set of size  $k$  well-approximates the optimal solution. On the other hand, if the size of the quantity-bucket is small ( $\leq \frac{m}{kd}$ ), then every voter can simply send all the alternatives that are in this bucket for her, and we can estimate the optimal solution. Of course, the algorithm is not aware of the *right* pair of value- and quantity-buckets, so it simply chooses one uniformly at random, further losing at most  $O(\log^2 m)$  factor.

**THEOREM 5.** *For  $d \geq 144 \log^4 m$ , there is a voting rule with deterministic elicitation for the  $k$ -selection rule with distortion  $O(d)$  and communication complexity  $O(\frac{m}{kd} \log^6 m)$ .*

## 5.2 Randomized Elicitation

We now discuss our  $k$ -selection rule with randomized elicitation. Once again, the detailed algorithm is presented in the full version, but we discuss the high-level idea below. This algorithm is developed in two stages: first, we design an algorithm that works when the number of voters  $n$  is polynomially bounded by  $m$  (say  $n \leq m^4$ ). Later, we show how the general problem can be reduced to this case.

Suppose  $n \leq m^4$ . As we did in the case of deterministic elicitation, by losing only logarithmic factors, we can reduce to an instance where all voters have either 0 or a non-zero value for each alternative, the non-zero values are approximately equal (as they are in a single “value-bucket”), and the voters have non-zero value for approximately the same number of alternatives (as they are in a single “quantity-bucket”). If this number is large ( $\geq \frac{m}{kd}$ ), then again a uniformly random set of size  $k$  provides a good solution. On the other hand, if this number is small ( $\leq \frac{m}{kd^3}$ ), then voters can just send the alternatives for which they have non-zero values, and we are only using about  $\frac{m}{kd^3}$  bits of communication. So the only interesting case is when this number is in the range  $[\frac{m}{kd^3}, \frac{m}{kd}]$ .

A first attempt would be the following. The voting rule creates a subset of size  $m/d$ , in which each of  $m$  alternatives is included with probability  $1/d$ . Then, every voter reports which of these alternatives she has a non-zero value for, and the voting rule finds the optimal set subject to this information. It can be shown that this leads to distortion  $O(d)$ . However, this protocol requires each voter to send  $\tilde{O}\left(\frac{m}{kd^2}\right)$  bits, as each voter may have non-zero value for up to  $\frac{m}{kd}$  alternatives.

This problem can be easily resolved if we allow voters to further sample a smaller set of alternatives on which they place non-zero value, and then take the intersection with the common set, so each element is retained with probability only  $\text{poly}(\log m)/d$ . However, recall that, while our model allows choosing a randomized query, it does not allow having voters respond randomly to a fixed query (e.g. two voters with the same valuation responding differently to a query). To circumvent this problem, the voting rule can create random seeds and send them to the voters as part of the query for subsampling. Although voters need to subsample independently, the seeds must be anonymous, as the query must be the same for all voters. To circumvent this issue, we “identify” each voter by the set of alternatives on which it places non-zero value. So the voting rule samples a large prime  $\phi$ , and sends it to each voter. Now, voter  $i$  interprets her set (on which she places non-zero value) as a number, and computes her ID modulo the prime number  $\phi$ . When the number of voters is small ( $\text{poly}(m)$ ), we show that we can choose a prime large enough so that voters with distinct sets get distinct IDs. Now, the voters use their IDs to read the random seeds provided by the rule, sample a set where each alternative is retained with probability  $\text{poly}(\log m)/d$ , and send the intersection with the common set back along with the ID they computed for themselves. To keep the communication complexity low, we thus need the ID to be small, which is why we need the prime number  $\phi$ , and in turn, the number of voters  $n$  to be  $\text{poly}(m)$ . This rule has communication complexity  $\tilde{O}\left(\frac{m}{kd^3}\right)$ . Our proof shows that its distortion is  $O(d)$  by carefully arguing that some alternative from the optimal set survives with a non-negligible probability in the intersection, and will be present enough times in the sets sent by the voter, so we can identify it with high probability.

**THEOREM 6.** *When  $n \leq m^4$  and  $d = \Omega(\log^6 m)$ , there is a voting rule with randomized elicitation for the  $k$ -selection rule with distortion  $O(d)$  and communication complexity  $O\left(\frac{m}{kd^3} \log^{21} m\right)$ .*

For  $n \geq m^4$ , we reduce the problem to  $n \leq m^4$  by choosing a random subset of  $m^4$  voters and only considering their responses. Using Chernoff bounds, we show that welfare of this subset of voters provides a good approximation of welfare of all voters, with high probability, and thus the optimal set computed from the responses from this subset of voters provides low distortion.



THEOREM 7. When  $d = \Omega(\log^6 m)$ , there is a voting rule with randomized elicitation for the  $k$ -selection rule with distortion  $O(d)$  and communication complexity  $O\left(\frac{m}{kd^3} \log^{21} m\right)$ .

## 6 $k$ -SELECTION: LOWER BOUNDS

In this section, we derive lower bounds on the communication complexity required to achieve a given distortion  $d$  in the  $k$ -selection problem. For voting rules with deterministic (resp. randomized) elicitation, we establish  $\Omega\left(\frac{m}{kd}\right)$  (resp.  $\Omega\left(\frac{m}{kd^3}\right)$ ) lower bound on the communication complexity; note that both these bounds are tight up to logarithmic factors given our upper bounds from Section 5. To establish the lower bound for deterministic voting rules, we extend the technique introduced in Section 4 for 1-selection to  $k$ -selection. To establish the lower bound for randomized voting rules, we use a new embedding technique in Section 6.2.

Before we introduce our lower bounds, we remark that it is also possible to reduce the 1-selection problem to the  $k$ -selection problem for a fixed  $k > 1$ . That is, we show how to use a  $k$ -selection voting rule to solve the 1-selection problem, and use a lower bound on the latter to derive a lower bound on the former. Since  $k > 1$  is fixed, we cannot trivially set  $k = 1$  and then run the  $k$ -selection rule. Nonetheless, our reduction is simple. However, it produces lower bounds that are weaker by a factor of  $k$  compared to the optimal bounds we present below. We still present this reduction in the full version due to its conceptual novelty; to the best of our knowledge, this is the first reduction from the 1-selection problem to the  $k$ -selection problem for a fixed  $k > 1$ .

### 6.1 Deterministic Elicitation

We first present a lower bound of  $\Omega\left(\frac{m}{kd}\right)$  on the communication complexity of  $k$ -selection rules which achieves distortion at most  $d$  using deterministic elicitation. As noted above, this is achieved by using our new lower bound on the total communication complexity of multi-party fixed-size set disjointness problem from Theorem 3, and using a technique similar to that from Theorem 4 to reduce this problem to the  $k$ -selection problem.

THEOREM 8. Let  $f$  be a  $k$ -selection voting rule which uses deterministic elicitation and achieves  $\text{dist}(f) \leq d$ . Then,  $C(f) = \Omega\left(\frac{m}{kd}\right)$ .

### 6.2 Randomized Elicitation

We now move to presenting a lower bound on the communication complexity  $C(f)$  of a voting rule  $f$  which achieves distortion at most  $d$  using randomized elicitation. For this, the approach of Theorem 8 unfortunately fails. Specifically, the argument goes through to show that, with probability at least  $k/t$ , the voting rule must return a set  $S$  containing an element appearing in more than one set. However, the randomness here is no longer only in aggregation, but in elicitation as well. Since our protocol is communication-restricted, we cannot assume direct access to all sets that are returned with probability at least  $k/t$  by the voting rule on this instance. This can be handled using a trick similar to what Mandal et al. [32] use in their lower bound proof for randomized elicitation with  $k = 1$ . We could run the voting rule  $\frac{t}{k} \cdot \ln(1/\delta)$  times, and record the generated sets. This ensures that with probability at least  $1 - \delta$ , at least one of the sets returned contains at least one element appearing in more than one set. Finding such an element results in a  $\delta$ -error protocol for  $\text{FDISJ}_{m,m/t,t}$ , which must have communication cost  $\Omega(m/t)$ . However, the communication cost of this protocol is roughly  $\frac{t}{k} \cdot t \cdot C(f)$ , which, using  $t = \Theta(kd)$ , gives  $C(f) = \Omega\left(\frac{m}{k^2 d^3}\right)$ . This bound is a factor of  $k$  looser than the bound we want.

Hence, we take a very different approach here. We go back to using  $\text{FDISJ}$  under the unique intersection promise, for which Mandal et al. [32] show that  $R_\delta(\text{FDISJ}_{m,s,t}) = \Omega(s)$  whenever

$m \geq \frac{3}{2}st$ . However, instead of using a big instance with  $m$  elements and  $t = \Theta(kd)$  players like they do, we use a small instance with  $m/k$  elements and  $t = \Theta(d)$  players, and embed it among a set of  $k$  instances where the other  $k - 1$  are randomly generated YES instances also containing  $m/k$  elements and  $t$  players. We use a symmetrization trick to ensure that the voting rule cannot distinguish the real instance from the generated ones, and must return its common element with a sufficiently high probability. We then repeat this setup to find the common element of the real instance with probability at least  $1 - \delta$ , and analyze the communication required. Interestingly, this gives us the tight bound we desire.

**THEOREM 9.** *Let  $f$  be a  $k$ -selection voting rule which uses randomized elicitation and achieves  $\text{dist}(f) \leq d$ . Then,  $C(f) = \Omega\left(\frac{m}{kd^3}\right)$ .*

### 6.3 Discussion of Our Lower Bounds

Given that our matching upper and lower bounds decrease with  $k$ , this shows that the  $k$ -selection problem becomes strictly easier as  $k$  increases. This is not obvious apriori because, while higher values of  $k$  allow a voting rule to achieve higher expected social welfare by selecting more alternatives, they also raise the optimal social welfare against which the expected social welfare of the voting rule is compared.

It is worth noting that Caragiannis et al. [17] also study the  $k$ -selection problem, but under the specific elicitation rule where each voter provides a ranking of the alternatives. They examine the optimal distortion which can be achieved using any aggregation rule. For deterministic aggregation, their bounds imply that the optimal distortion is roughly inversely proportional to  $k$ . For randomized aggregation, their upper bound *increases* as  $\sqrt{k}$  up to  $k \approx \sqrt[3]{m}$ , and then decreases as  $1/k$ , whereas their lower bound stays constant until  $k \approx \sqrt{m}$ , and then decreases roughly as  $1/k$ , leaving open the question of the optimal bound for small  $k$ . In contrast, when the communication complexity is kept fixed, Theorems 8 and 9 show that the optimal distortion for  $k$ -selection decreases as  $1/k$  using deterministic elicitation and as  $1/\sqrt[3]{k}$  using randomized elicitation.

## 7 DISCUSSION & FUTURE WORK

Our work leaves open a number of directions for future research. On a technical level, our upper and lower bounds are tight, but only up to logarithmic factors. When the desired distortion is either very small ( $d = \text{polylog}(m)$ ) or very large ( $d = m/\text{polylog}(m)$ ), improving the logarithmic factors and determining the exact communication complexity becomes important. For example, our lower bounds show that to achieve  $O(1)$  distortion,  $\Omega(m/k)$  communication is necessary in the  $k$ -selection problem even with randomized elicitation and aggregation. Can this be achieved using only  $O(m/k)$  communication, without any logarithmic factors? Does randomized elicitation offer any asymptotic advantage over deterministic elicitation in the  $d = O(1)$  regime?

We also note that for the 1-selection problem with deterministic elicitation, the upper bound of  $\tilde{O}(m/d)$  achieved by Mandal et al. [32] uses deterministic aggregation whereas our matching lower bound of  $\Omega(m/d)$  holds even for randomized aggregation, thus establishing that there is no significant advantage of using randomized aggregation over deterministic aggregation. However, for the 1-selection problem with randomized elicitation, and the  $k$ -selection problem with deterministic or randomized elicitation, our upper bounds use randomized aggregation. We believe that it should be possible to achieve the same bounds using deterministic aggregation (implying that deterministic aggregation is almost as powerful as randomized aggregation), but leave this for future work.

We note that our lower bounds use a reduction from the set-disjointness problem, where lower bounds are known in a very powerful setup which allows a protocol to use multiple rounds of adaptive elicitation. In contrast, our voting model is defined for voting rules which use a single

round of uniform elicitation (where the same query is posed to all voters). Establishing lower bounds for more general forms of elicitation is an interesting direction for future work.

We remark that our results also have implications for other problems studied in voting. One interesting example is *participatory budgeting*, where each alternative has an associated cost, there is a total budget, and the goal is to return the optimal feasible set of alternatives (where feasibility means that the total cost should not exceed the budget). This problem models the real-world process of participatory budgeting, where residents of a city vote over which public projects should be funded. We view designing optimal voting rules for participatory budgeting as a key direction for future work.

Finally, we note that our use of communication complexity (i.e. the number of bits that each voter needs to transmit) can be viewed as an extremely crude measure of the cognitive burden that an elicitation rule imposes on the voters. Quite possibly, humans may find it easier to provide a certain response even if it technically requires transmitting a larger number of bits. The study of more realistic measures of cognitive burden in voting is a challenging direction for future work, which may require interdisciplinary ideas.

## ACKNOWLEDGMENTS

We thank Ariel D. Procaccia for useful discussions. Mandal was supported in part by the Post-Doctoral fellowship from the Columbia Data Science Institute. 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] B. Abramowitz, E. Anshelevich, and W. Zhu. 2019. Awareness of voter passion greatly improves the distortion of metric social choice. In *Proceedings of the 15th Conference on Web and Internet Economics (WINE)*. 3–16.
- [2] G. Amanatidis, G. Birmpas, A. Filos-Ratsikas, and A. A. Voudouris. 2020. Peeking Behind the Ordinal Curtain: Improving Distortion via Cardinal Queries. In *Proceedings of the 34th AAAI Conference on Artificial Intelligence (AAAI)*. Forthcoming.
- [3] A. Andoni, R. Krauthgamer, and K. Onak. 2011. Streaming algorithms via precision sampling. In *Proceedings of the 52nd IEEE Annual Symposium on Foundations of Computer Science*. 363–372.
- [4] E. Anshelevich, O. Bhardwaj, E. Elkind, J. Postl, and P. Skowron. 2018. Approximating optimal social choice under metric preferences. *Artificial Intelligence* 264 (2018), 27–51.
- [5] E. Anshelevich and J. Postl. 2017. Randomized social choice functions under metric preferences. *Journal of Artificial Intelligence Research* 58 (2017), 797–827.
- [6] H. Aziz. 2020. Justifications of Welfare Guarantees under Normalized Utilities. *ACM SIGecom Exchanges* 17, 2 (2020), 71–75.
- [7] A. Badanidiyuru, S. Dobzinski, H. Fu, R. Kleinberg, N. Nisan, and T. Roughgarden. 2012. Sketching valuation functions. In *Proceedings of the 23rd Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*. 1025–1035.
- [8] M.-F. Balcan and N. J. Harvey. 2018. Submodular Functions: Learnability, Structure, and Optimization. *SIAM Journal on Computing* 47, 3 (2018), 703–754.
- [9] G. Benadè, S. Nath, A. D. Procaccia, and N. Shah. 2017. Preference Elicitation for Participatory Budgeting. In *Proceedings of the 31st AAAI Conference on Artificial Intelligence (AAAI)*. 376–382.
- [10] G. Benadè, A. D. Procaccia, and M. Qiao. 2019. Low-Distortion Social Welfare Functions. In *Proceedings of the 33rd AAAI Conference on Artificial Intelligence (AAAI)*. 1788–1795.
- [11] J. Bentham. 1996. *The Collected Works of Jeremy Bentham: An Introduction to the Principles of Morals and Legislation*. Clarendon Press.
- [12] U. Bhaskar, V. Dani, and A. Ghosh. 2018. Truthful and Near-Optimal Mechanisms for Welfare Maximization in Multi-Winner Elections. In *Proceedings of the 32nd AAAI Conference on Artificial Intelligence (AAAI)*. 925–932.
- [13] A. Borodin, O. Lev, N. Shah, and T. Strangway. 2019. Primarily about Primaries. In *Proceedings of the 33rd AAAI Conference on Artificial Intelligence (AAAI)*. 1804–1811.

- [14] C. Boutilier, I. Caragiannis, S. Haber, T. Lu, A. D. Procaccia, and O. Sheffet. 2015. Optimal Social Choice Functions: A Utilitarian View. *Artificial Intelligence* 227 (2015), 190–213.
- [15] S. J. Brams and P. C. Fishburn. 2002. Voting Procedures. In *Handbook of Social Choice and Welfare*. 173–236.
- [16] F. Brandt, V. Conitzer, U. Endress, J. Lang, and A. D. Procaccia (Eds.). 2016. *Handbook of Computational Social Choice*. Cambridge University Press.
- [17] I. Caragiannis, S. Nath, A. D. Procaccia, and N. Shah. 2017. Subset Selection Via Implicit Utilitarian Voting. *Journal of Artificial Intelligence Research* 58 (2017), 123–152.
- [18] I. Caragiannis and A. D. Procaccia. 2011. Voting Almost Maximizes Social Welfare Despite Limited Communication. *Artificial Intelligence* 175, 9–10 (2011), 1655–1671.
- [19] J. R. Chamberlin and P. N. Courant. 1983. Representative Deliberations and Representative Decisions: Proportional Representation and the Borda Rule. *American Political Science Review* 77, 3 (1983), 718–733.
- [20] M. Charikar, K. Chen, and M. Farach-Colton. 2002. Finding frequent items in data streams. In *Proceedings of the 29th International Colloquium on Automata, Languages and Programming (ICALP)*. 693–703.
- [21] M. de Condorcet. 1785. *Essai sur l'application de l'analyse à la probabilité de décisions rendues à la pluralité de voix*. Imprimerie Royal. Facsimile published in 1972 by Chelsea Publishing Company, New York.
- [22] E. Elkind, P. Faliszewski, P. Skowron, and A. Slinko. 2017. Properties of Multiwinner Voting Rules. *Social Choice and Welfare* 48, 3 (2017), 599–632.
- [23] P. Faliszewski, P. Skowron, A. Slinko, and N. Talmon. 2017. Multiwinner Voting: A New Challenge for Social Choice Theory. *Trends in Computational Social Choice* 74 (2017), 27–47.
- [24] V. Gkatzelis, D. Halpern, and N. Shah. 2020. Resolving the Optimal Metric Distortion Conjecture. arXiv:2004.07447.
- [25] A. Goel, A. K. Krishnaswamy, and K. Munagala. 2017. Metric distortion of social choice rules: Lower bounds and fairness properties. In *Proceedings of the 18th ACM Conference on Economics and Computation (EC)*. 287–304.
- [26] J. C. Harsanyi. 1955. Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparisons of Utility. *Journal of Political Economy* 63, 4 (1955), 309–321.
- [27] J. D. Hartline. 2012. Approximation in Economic Design. Book in preparation.
- [28] C. Hillinger. 2005. The Case for Utilitarian Voting. *Homo Oeconomicus* 22, 3 (2005).
- [29] R. Jayaram and D. P. Woodruff. 2018. Perfect  $L_p$  Sampling in a Data Stream. In *2018 IEEE 59th Annual Symposium on Foundations of Computer Science (FOCS)*. IEEE, 544–555.
- [30] H. Jowhari, M. Sağlam, and G. Tardos. 2011. Tight Bounds for  $L_p$  Samplers, Finding Duplicates in Streams, and Related Problems. In *Proceedings of the 30th ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems*. 49–58.
- [31] E. Koutsoupias and C. Papadimitriou. 1999. Worst-Case Equilibria. In *Annual Symposium on Theoretical Aspects of Computer Science*. 404–413.
- [32] D. Mandal, A. D. Procaccia, N. Shah, and D. Woodruff. 2019. Efficient and Thrifty Voting by Any Means Necessary. In *Proceedings of the 33rd Annual Conference on Neural Information Processing Systems (NeurIPS)*. 7178–7189.
- [33] M. Monemizadeh and D. P. Woodruff. 2010. 1-Pass Relative-Error  $L_p$ -Sampling with Applications. In *Proceedings of the 21st Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*. 1143–1160.
- [34] B. L. Monroe. 1995. Fully Proportional Representation. *American Political Science Review* 89, 4 (1995), 925–940.
- [35] K. Munagala and K. Wang. 2019. Improved Metric Distortion for Deterministic Social Choice Rules. In *Proceedings of the 20th ACM Conference on Economics and Computation (EC)*. 245–262.
- [36] N. Nisan and A. Ronen. 2001. Algorithmic Mechanism Design. *Games and Economic Behavior* 35, 1–2 (2001), 166–196.
- [37] A. D. Procaccia and J. S. Rosenschein. 2006. The Distortion of Cardinal Preferences in Voting. In *Proceedings of the 10th International Workshop on Cooperative Information Agents (CIA)*. 317–331.
- [38] A. D. Procaccia and M. Tennenholtz. 2009. Approximate mechanism design without money. In *Proceedings of the 10th ACM Conference on Economics and Computation (EC)*. 177–186.