A Min-Entropy Approach to Multi-Party Communication Lower Bounds

Mi-Ying (Miryam) Huang
Thomas Lord Department of Computer Science, University of Southern California, Los Angeles, CA, USA

Xinyu Mao
Thomas Lord Department of Computer Science, University of Southern California, Los Angeles, CA, USA

Shuo Wang
Department of Mathematics, Massachusetts Institute of Technology, Cambridge, MA, USA

Guangxu Yang
Thomas Lord Department of Computer Science, University of Southern California, Los Angeles, CA, USA

Jiapeng Zhang
Thomas Lord Department of Computer Science, University of Southern California, Los Angeles, CA, USA

Abstract

Information complexity is one of the most powerful techniques to prove information-theoretical lower bounds, in which Shannon entropy plays a central role. Though Shannon entropy has some convenient properties, such as the chain rule, it still has inherent limitations. One of the most notable barriers is the square-root loss, which appears in the square-root gap between entropy gaps and statistical distances, e.g., Pinsker's inequality. To bypass this barrier, we introduce a new method based on min-entropy analysis. Building on this new method, we prove the following results.

- An $\Omega(N^{\sum_i \alpha_i - \max_i\{\alpha_i\}}/k)$ randomized communication lower bound of the k-party *set-intersection* problem where the i-th party holds a random set of size $\approx N^{1-\alpha_i}$.
- A tight $\Omega(n/k)$ randomized lower bound of the k-party *Tree Pointer Jumping* problems, improving an $\Omega(n/k^2)$ lower bound by Chakrabarti, Cormode, and McGregor (STOC 08).
- An $\Omega(n/k + \sqrt{n})$ lower bound of the *Chained Index* problem, improving an $\Omega(n/k^2)$ lower bound by Cormode, Dark, and Konrad (ICALP 19).

Since these problems served as hard problems for numerous applications in streaming lower bounds and cryptography, our new lower bounds directly improve these streaming lower bounds and cryptography lower bounds.

On the technical side, min-entropy does not have nice properties such as the chain rule. To address this issue, we enhance the structure-vs-pseudorandomness decomposition used by Göös, Pitassi, and Watson (FOCS 17) and Yang and Zhang (STOC 24); both papers used this decomposition to prove communication lower bounds. In this paper, we give a new breath to this method in the multi-party setting, presenting a new toolkit for proving multi-party communication lower bounds.

2012 ACM Subject Classification Theory of computation → Communication complexity

Keywords and phrases communication complexity, lifting theorems, set intersection, chained index

Digital Object Identifier 10.4230/LIPIcs.CCC.2025.33

Related Version *Previous Version*: https://eccc.weizmann.ac.il/report/2024/067/
Previous Version: https://eccc.weizmann.ac.il/report/2023/164/

Funding *Mi-Ying (Miryam) Huang*: Supported by NSF CAREER award 2141536.
Xinyu Mao: Supported by NSF CAREER award 2141536.
Shuo Wang: Supported by NSF CCF award 2227876.

COMPUTATIONAL COMPLEXITY CONFERENCE

40th Computational Complexity Conference (CCC 2025).
Editor: Srikanth Srinivasan; Article No. 33; pp. 33:1–33:29

Leibniz International Proceedings in Informatics
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

Guangxu Yang: Supported by NSF CAREER award 2141536.
Jiapeng Zhang: Supported by NSF CAREER award 2141536.

Acknowledgements We thank anonymous reviewers for their helpful comments.

1 Introduction

Information complexity is one of the most powerful tools in proving communication complexity lower bounds [17, 5, 6, 23, 40] and streaming lower bounds [5, 16, 2, 31, 3, 13, 37, 12]. The idea of information complexity is to analyze the mutual information between the inputs held by the communication parties and the communication transcript. The definition of information complexity is similar to communication complexity, with information cost replacing communication cost. For a protocol Π, a popular notion of information cost is defined by $\mathrm{IC}(\Pi) \stackrel{\mathsf{def}}{=} \mathrm{I}(\boldsymbol{X};\Pi(\boldsymbol{X},\boldsymbol{Y})|\boldsymbol{Y}) + \mathrm{I}(\boldsymbol{Y};\Pi(\boldsymbol{X},\boldsymbol{Y})|\boldsymbol{X})$, where $\boldsymbol{X}$ and $\boldsymbol{Y}$ are the input distribution of Alice and Bob respectively and I is the mutual information. Intuitively, $\mathrm{IC}(\Pi)$ captures the mutual information of the inputs and the communication transcript, which is a lower bound of the communication cost. Besides this specific definition, there are many different variants that are smartly designed for diverse applications. However, they all share a similar idea: capture the information cost (usually by Shannon entropy) between the input distribution and the transcript.

Despite a vast number of applications successfully given by the information complexity-based approaches, this framework still has some inherent limitations. Indeed, some significant barriers are *not only* associated with some specific variants of information cost notions, but further deeply caused by the *entropy* itself. In this direction, one notable limitation is the square-root loss barrier.

Square-root loss barrier

We first use a simple example to illustrate this phenomenon. Let $\boldsymbol{I}$ be a random variable that outputs 1 with probability $1/2+\varepsilon$ and 0 with probability $1/2-\varepsilon$. This is a biased coin with a $\Theta(\varepsilon)$ statistical distance to the uniform distribution. However, on the other hand, the entropy gap between them has only $\Theta(\varepsilon^2)$. This square gap is not significant if ε is a constant. However, the loss would become very large when it becomes very small. Beyond this simple example, this is indeed a *general gap between entropy loss and statistical distance*. For example, any result that applies Pinsker's inequality has a good chance of creating this gap.

▶ **Lemma 1** (Pinsker's inequality). *If P and Q are two distributions, then*

$$D_{TV}(P,Q) \le \sqrt{\frac{1}{2}D_{KL}(P\|Q)}$$

Here $D_{TV}(P,Q)$ is the total variation distance of P and Q and $D_{KL}(P\|Q)$ is the KL-divergence.

This quadratic gap makes it difficult to get good bounds via entropy-based analysis in many applications. For instance, proofs of *multiparty unique-set disjointness* [5], *set disjointness under product distribution* [23, 40], the *chained index problem* [21], *multi-party pointer jumping problem* [15], *tree pointer jumping problem* [16], *pointer chasing problem* [39], among others, all meet the square-root loss comparing with the upper bounds.

Despite resolving the square-root loss for some specific problems, these efforts are ad-hoc and use non-standard variants of Shannon entropy. Hence, it is hard to extend them for broader applications. A natural question arises: Could we use any measurement other than the Shannon entropy (or its close variants)?

Now, we revisit the example above. For a random variable $\boldsymbol{X}$ supported on $\{0,1\}^n$ with entropy $\mathrm{H}(\boldsymbol{X}) \geq n-\varepsilon$, we know the statistical distance between $\boldsymbol{X}$ and the uniform distribution is $\Theta(\sqrt{\varepsilon})$ by Pinsker's inequality. Furthermore, improving Pinsker's inequality is hard as it is tight in general. However, on the other hand, for a random variable $\boldsymbol{X}$ with *min-entropy* $n-\varepsilon$, a simple calculation shows that the statistical distance between $\boldsymbol{X}$ and the uniform distribution is $\Theta(\varepsilon)$. In this paper, min-entropy is a good candidate for avoiding square root loss in general settings.

Analysis of min-entropy via structure-vs-pseudorandomness

Though the min-entropy itself does not meet the square-root loss, there are other challenges in analyzing it. One of the most significant challenges is that, unlike the Shannon entropy, there is *no chain rule* for min-entropy, where a chain rule is an essential tool in entropy.

In order to overcome this issue, we adopt the structure-vs-pseudorandomness decomposition to serve as the "chain rule" in min-entropy analysis. This approach has been successfully applied in sunflower lemmas [36, 1] and query-to-communication lifting theorems [29, 30, 35, 46, 38]. Though this approach has been successfully applied in several areas, it has not been studied in *multi-party settings*. In this paper, we extend this approach to the multi-party setting. Beyond the three problems studied in this paper, we believe the min-entropy-based analysis could provide more applications to multi-party problems.

1.1 Our Results

Building on min-entropy analysis, we improve the lower bounds for three communication problems: (1) *Set Intersection* [7], (2) *Tree Pointer Jumping* [16], and (3) *Chained Index* [21].

1.1.1 Set Intersection Problem

To show the advantages of our min-entropy approach, we consider the *search version* of the *set-disjointness* problem, which is called the *set-intersection* problem. There are two versions of the set-intersection problems. The first one requires the players to find the whole intersection; the second one only asks the players to find one element from the intersection. Together, the set-intersection problems have been studied in many papers [34, 42, 11, 14, 41, 7, 45, 27, 26, 32, 9, 40]. In this paper, we focus on the second version. The setting is: each player i is assigned a subset S_i of $[N]$, and the goal changes to finding an element $a \in \bigcap_{i=1}^{k} S_i$.

We consider the communication complexity under product distribution here, and there are two typical product distributions that have been widely studied before. One is the *fixed-size product distribution*, where each player i receives a uniformly random subset $S_i \subseteq [N]$ with $|S_i| = n_i$. The other one is the *Bernoulli product distribution*, where each player i receives a random set S_i sampled as follows: for each element $a \in [N]$, $a \in S_i$ independently with probability m_i. Babai, Frankl, and Simon [4] first proposed the communication complexity of the *set-disjointness* problem under fixed-size product distribution where $n_i = \sqrt{N}$, and gave an $\Omega(\sqrt{N})$ lower bound. Their proof could also be adapted to the setting of Bernoulli product distribution with $m_i = N^{-1/2}$. Recently, this bound was extended to the k-party setting by a recent paper by Dershowitz, Oshman, and Roth [23]. They showed that when $k \leq \log N/6$, the communication complexity under the Bernoulli product distribution, where $m_i = N^{-1/k}$, is $\Omega(N^{1-1/k}/k^2)$. Both of these decision-version lower bounds gave various applications.

In the context of the *set-intersection*, lower bounds are less known, though it also provides many applications. Bauer, Farshim, and Mazaheri [7] first gave a lower bound under Bernoulli product distribution with applications to cryptography. To be more specific, they proved:

▶ **Theorem 2** ([7]). *For the 2-party set-intersection problem under Bernoulli product distribution, where $m_i = N^{-\alpha_i}$, $\alpha_1 + \alpha_2 \le 1$, its communication complexity is $\Omega\left(N^{\alpha_1+\alpha_2+\min\{\alpha_1,\alpha_2\}-1}\right)$.*

Note that this problem is exactly the search version of the set-disjointness problem considered in [4, 23]. Compared to the set-disjointness problem, set-intersection could be studied in a larger range of parameters, i.e., $\alpha_1 + \alpha_2 < 1 - \Omega(1)$, where the intersection could be very large, i.e., as large as $N^{\Omega(1)}$ with high probability.

However, the theorem by [7] is far from tight and does not provide a non-trivial bound when α_1, α_2 are small. One of the main obstacles here is that the size of the intersections is large, but players only need to find one common element from many valid answers.

We consider the communication problem in [7], and extend it to the k-party setting. Concretely, we assume that each player holds a (random) set S_i of size $|S_i| \approx N^{1-\alpha_i}$ with $\sum_i \alpha_i \le 1$, and prove the following results for *set-intersection*.

▶ **Theorem 3.** *For the k-party set-intersection problem under Bernoulli product distribution, where $m_i = N^{-\alpha_i}$, $\sum_i \alpha_i \le 1$ and $k \le 0.1 \cdot \min\{N^{\min_i\{\alpha_i\}/2}, N^{(1-\max_i\{\alpha_i\})/3}\}$[1]:*

1. *the communication complexity is $\Omega\left(N^{\sum_i \alpha_i - \max_i\{\alpha_i\}}/k\right)$ to achieve a constant accuracy;*
2. *there exists a protocol that solves this problem under the distribution mentioned above with a constant accuracy and uses $O\left(k \log N \cdot N^{\sum_i \alpha_i - \max_i\{\alpha_i\}}\right)$ communication cost.*

Note that this theorem establishes the first non-trivial lower bound when $2\alpha_1 + \alpha_2 \le 1$ (we assume $\alpha_1 \le \alpha_2$ here). Actually, it implies that $\widetilde{\Theta}\left(N^{\sum_i \alpha_i - \max_i\{\alpha_i\}}\right)$ is tight (up to logarithmic factors when $k < \log N$) for the distributional-version of set-intersection considered in [7]. Also, our bound is similar to [23] when $\alpha_1 = \cdots = \alpha_k = 1/k$, where they proved a lower bound of $\Omega(N^{1-1/k}/k^2)$ for $k \le \log N/6$. Similar to the two-party setting, our result significantly strengthens theirs when the size of the intersection is large.

1.1.2 Tree Pointer Jumping Problem

The Tree Pointer Jumping problem is a communication problem introduced by Chakrabarti, Cormode, and McGregor [16] with applications in streaming lower bounds. For $t, k \ge 2$, we consider a complete k-level t-ary tree T rooted at v_1. The k-party Tree Pointer Jumping problem, denoted by $TPJ_{k,t}(\phi)$, takes as an input a function $\phi : V(T) \to [t]$ with $\phi(v) \in \{0, 1\}$ if v is a leaf of T , where $V(T)$ is the set of nodes of T. We use $\mathcal{F}$ to denote the set of all valid functions ϕ here. For each input ϕ, we define the functions g_ϕ by,

$$g_\phi(v) = \begin{cases} \text{the } \phi(v)\text{-th child of } v & \text{if } v \text{ is not an internal node;} \\ \phi(v) & \text{if } v \text{ is a leaf.} \end{cases}$$

The output of $TPJ_{k,t}(\phi)$ is defined by $TPJ_{k,t}(\phi) := g_\phi(g_\phi(\cdots g_\phi(v_1)\cdots))$. In the communication setting, the input ϕ is distributed to k players. The problem is described as follows:

[1] We assume all the distributions considered in this paper satisfy this constraint.

- Player i receives the labels of the i-th level nodes, i.e., the first player receives $\phi(v_1)$,..., and the last player receives the labels of the leaves.
- In each round, players send messages in reverse order: *from the last player to the first player*. The cost of this round is the total number of bits sent by all players.
- Players could communicate $(k-1)$ rounds, and the *first player* outputs the answer.

The goal of the players is to compute $TPJ_{k,t}(\phi)$ while minimizing *the maximum cost of each round*. For any $(r-1)$-round protocol Π, we use $R_{\max}(\Pi)$ to denote the maximum communication cost in all rounds. In this direction, [16] first proved the following lower bound.

▶ **Theorem 4** ([16]). *For any $(k-1)$-round protocol Π with $\mathbf{Pr}_{\phi\sim\mathrm{Unif}(\mathcal{F})}[\Pi(\phi) = TPJ_{k,t}(\phi)] \geq 2/3$, we have that $R_{\max}(\Pi) = \Omega(t/k^2)$.*

Chakrabarti, Cormode, and McGregor [16] first used Theorem 4 to improve multi-pass streaming lower bound for median finding. Later on, Chakrabarti and Wirth [18] used this theorem to show a pass/approximation trade-off for the SET-COVER in the semi-streaming setting. In this paper, we improve the lower bound from Theorem 4 based on min-entropy analysis.

▶ **Theorem 5.** *For any $(k-1)$-round protocol Π with $\mathbf{Pr}_{\phi\sim\mathrm{Unif}(\mathcal{F})}[\Pi(\phi) = TPJ_{k,t}(\phi)] \geq 2/3$, we have that $R_{\max}(\Pi) = \Omega(t/k)$.*

As corollaries, our improved lower bounds can be directly used to improve the applications given by [16] and [18]. Since this paper is a merged version of [44] and [33], we omit the proof of Theorem 5. Readers can refer to [33] for the complete proof and further applications.

1.1.3 Chained Index Problem

The Chained Index problem, introduced by Cormode, Dark, and Konrad [21], is another useful tool with many applications in streaming lower bounds [21, 24, 25, 10, 22]. For this problem, we consider the following communication setting.

- There are k players. Each player i receives an input $z_i = (\sigma_i, x_i) \in [n] \times \{0,1\}^n$
- It is promised that $x_1(\sigma_2) = \cdots = x_{k-1}(\sigma_k)$. Here $x_i(\sigma_{i+1})$ is the σ_{i+1}-th coordinate of x_i.
- Their goal is to compute $x_i(\sigma_{i+1})$ through a one-way communication from the first player to the last player, where the last player should output the answer.

We say that a one-way protocol solves the Chained Index problem if for every input $(z_1, \ldots, z_k)$, the last player always outputs the correct answer with probability $2/3$. The communication cost of this protocol is the total communication bits of all players. Using tools from information complexity, Cormode, Dark, and Konrad proved the following lower bounds for the Chained Index problem.

▶ **Theorem 6** ([21]). *Any one-way communication protocol that solves the Chained Index problem has randomized communication complexity $\Omega(n/k^2)$.*

Since it has been introduced, many streaming lower bounds [21, 24, 25, 10, 22] were built on Theorem 6. Interested readers can find detailed discussions in the papers mentioned above. Theorem 6 was obtained by direct entropy-based analysis. In this paper, we improve this lower bound.

▶ **Theorem 7.** *Any one-way communication protocol that solves the Chained Index problem has randomized communication complexity* $\Omega(n/k + \sqrt{n})$.

1.2 Proof Outline

In this section, we give a brief overview to our proof technique and we use the set-intersection problem for illustration. The proofs for chained index problem is similar in spirit.

Instead of considering Bernoulli distributions, we consider the following product distribution to simplify our presentation:

- Each player i independently and uniformly samples $cN^{1-\alpha_i}$ elements from $[N]$ (may have duplicates), where c equals $(1 + 2/k)$ here.

Thus, each player i receives a vector in $[N]^{cN^{1-\alpha_i}}$ and gets its set $S_i \subseteq [N]$ by removing the duplicate elements in the vector. In general, for any $I \subseteq [cN^{1-\alpha_i}]$ and $\beta \in [N]^I$, we consider β as a subset of $[N]$ in a similar way. We prove the lower bound under this distribution, and then reductions are established in Section 3.3 to prove our main theorem.

It is well known that a deterministic protocol Π partitions the input domain into $2^{|\Pi|}$ rectangles by step-by-step communication. The crucial idea of our proof is to further partition these leaf rectangles in the protocol tree into many structured rectangles. A structured rectangle $R = X_1 \times X_2 \times \cdots \times X_k$ satisfies that: for each i, X_i is fixed on some coordinates and pseudorandom on the remaining coordinates. The formal definition is given below.

▶ **Definition 8** (Structured rectangles). *Assuming $R = X_1 \times X_2 \times \cdots \times X_k$, where each X_i is a subset of $[N]^{cN^{1-\alpha_i}}$, is a rectangle. We say R is a structured rectangle if there exist subsets of coordinates $J_1, J_2, \cdots, J_k$ with $J_i \subseteq [cN^{1-\alpha_i}]$ satisfying that*

- *For each i, there exists a $\beta_i \in [N]^{J_i^c}$ such that $\forall x_i \in X_i, x_i(J_i^c) = \beta_i$. Here, J_i^c is the complement of J_i defined by $J_i^c := [cN^{1-\alpha_i}] - J_i$ and $x_i(J_i^c) \in [N]^{J_i^c}$ is the values of x_i on J_i^c.*
- *For each i, X_i has a high block-wise min-entropy (see definitions in Section 2) on the coordinates J_i.*

The notion of structured rectangle has also been widely used in query-to-communication lifting theorems [30, 19, 35].

In the decomposition, we recursively (starting from the root to the leaves) decompose all rectangles in the protocol tree, i.e., for a node (which is also a rectangle), we decompose it based on the decomposition of its ancestors. This is the key step compared to existing decomposition (pre-sampling techniques) in cryptography, which may lead to new applications. The formal process of this decomposition is referred to Section 3.

After the decomposition process, each leaf has been partitioned into many structured rectangles. For a structured rectangle $R = X_1 \times X_2 \times \cdots \times X_k$ associated with $J_1, \ldots, J_k$ and $\beta_i \in [N]^{J_i^c}$ for $i \in [k]$, we say that:

1. R is *bad* if $\cap_i \beta_i \neq \emptyset$.
2. R is *good* if $\cap_i \beta_i = \emptyset$. We also call good structured rectangles as pseudorandom rectangles.

Then, our proof consists of the following two parts.

- If the communication complexity of Π is small, the total size of bad structured rectangles is small compared to the size of the input domain (formalized by Lemma 17);
- On the other hand, we show that players can not find a common intersection from pseudorandom rectangles (formalized by Lemma 18).

Combining the two parts, we are able to prove the main theorem. We defer the detailed proofs to Section 3.

Comparison with existing methods

Similar questions have been widely studied in several recent papers [7, 32, 23, 40]. All of these papers used standard known techniques in communication complexity such as information complexity.

These papers achieved tight bounds for *set-disjointness* (decision version), or set-intersection enumeration (finding whole intersections). However, all of their bounds for search problems are sub-optimal whenever the size of the set intersection (the solution space for the search problem) is large. By contrast, our method follows the structure-vs-pseudorandomness approach by Yang and Zhang [46], which was inspired by lifting theorems [30, 19, 35]. In [46], authors first introduced the techniques in proving query-to-communication lifting theorems directly to a communication setting without gadgets and proved the communication complexity lower bound for the collision problem in a two-party setting.

Compared to [46], we further extend their approach in two aspects: 1) we generalize this method into the multi-party setting; 2) we adopt it to prove communication lower bounds for search problems with many solutions. To the best of our knowledge, existing lower-bound methods could not address communication problems in these two settings. We believe these two settings could provide many applications.

1.3 Subsequent works and future directions

A late work by Sundaresan [43] further improved the lower bound of the Chained Index problem to $\Omega(n - k \log n)$ via a reduction to a variant called the biased index problem.

Göös et al. [28] investigated quantum-classical separations in the communication model. To be more specific, they exhibit a total search problem whose communication complexity in the quantum simultaneous message passing model is exponentially smaller than in the classical two-way randomized model, which they call the *Bipartite NullCodeword* problem. To establish classical lower bounds, they employed a structure-vs-randomness approach, akin to the techniques used in [46] and this paper.

Another notable result was demonstrated by Mao, Yang, and Zhang in [38], where they improved the lower bound for a classical communication problem known as the *k-step pointer chasing problem* by the structure-vs-randomness approach.

In [8], Beame and Whitmeyer established near-optimal lower bounds for the k-party collision-finding problem of the strong bit pigeonhole principle which implies the tree-like semantic cutting-planes refutation lower bounds in proof complexity. However, their lower bounds only hold for the strong bit pigeonhole principle, they left lower bounds for the weak bit pigeonhole principle as an open question.

Paper organization

In Section 2, we give preliminaries. Section 3 shows an almost tight bound for the Set Intersection problem. In Section 4, we prove an improved lower bound for the Chained Index problem.

2 Preliminary

Definitions for set intersection problem

To begin with, we formally define the product distributions for set intersection problem adopted in this paper. For fixed parameters: k is the number of parties, N is the size of the domain, and $\alpha_i \in (0,1)$ are parameters indicating the size of each player's set. We consider the following three types of hardness distributions in this paper (two of them have appeared in Section 1):

1. Each player i independently and uniformly samples $cN^{1-\alpha_i}$ elements from $[N]$ (may have duplicates), where c equals $(1+2/k)$.
2. Each player i independently and uniformly samples $c_i N^{1-\alpha_i}$ distinct elements from $[N]$, where $1-1/k \le c_i \le 1+1/k$.
3. Each player i independently samples its set S_i with that every element $a \in [N]$ is contained in S_i with probability $N^{-\alpha_i}$.

We assume that $\sum_i \alpha_i \le 1$, otherwise the existence of intersections can be not guaranteed. Furthermore, if $\sum_i \alpha_i \le 1-C$ holds for some constant $C>0$, the common intersection of all players could be very large ($\approx N^C$).

The hardness distribution 3 is the Bernoulli product distribution with wide applications. Previous papers have mainly focused on this distribution. We prove the lower bound under distribution 1, and use two simple reductions to get the lower bound results for the hardness distributions 2 and 3. We refer to the two reductions to Section 3.3. In what follows, our discussion mainly focuses on distribution 1.

For a distribution D and a communication protocol Π, we define the *accuracy* of Π on D by:

$$\mathrm{Acc}_\Pi(D) := \mathop{\mathbf{Pr}}_{S_1,\cdots,S_k\sim D}\left[\Pi(S_1,\cdots,S_k) \in \bigcap_{i=1}^{k} S_i\right].$$

For simplicity, we define this accuracy notion, which does not take the cases when sets are disjoint into consideration, differently from [7] in which they also consider the accuracy of distinguishing disjoint cases, namely, they define

$$\mathrm{Acc}'_\Pi(D) := \mathop{\mathbf{Pr}}_{S_1,\cdots,S_k\sim D}\left[\Pi(S_1,\cdots,S_k) \in \bigcap_{i=1}^{k} S_i \text{ or } \Pi(S_1,\cdots,S_k) = \bigcap_{i=1}^{k} S_i = \emptyset\right].$$

Since we aim to establish lower bounds for those protocols achieving $\mathrm{Acc}_\Pi(D) = \Omega(1)$, we only consider the range of $\alpha_1,\ldots,\alpha_k$ with[2]

$$\mathop{\mathbf{Pr}}_{S_1,\cdots,S_k\sim D}\left[\bigcap_{i=1}^{k} S_i \neq \emptyset\right] > 1/2.$$

In this paper, our lower bound result shows that achieving $\mathrm{Acc}_\Pi(D) > \epsilon$, where epsilon is a constant less than $1/2$, requires large amounts of communication. This also implies a non-trivial hardness result to achieve $\mathrm{Acc}'_\Pi(D) > \epsilon + 1/2$ since the disjoint cases could contribute at most $1/2$ to $\mathrm{Acc}'_\Pi(D)$ when $\mathbf{Pr}_{S_1,\cdots,S_k\sim D}\left[\bigcap_{i=1}^{k} S_i \neq \emptyset\right] > 1/2$ holds. Hence, our results also imply hardness results under the [7] setting.

[2] $\mathbf{Pr}_{S_1,\cdots,S_k\sim D}\left[\bigcap_{i=1}^{k} S_i \neq \emptyset\right] > 1/2$ is guaranteed by the definitions of hardness distribution 3 when $\sum_i \alpha_i \le 1$.

Next, we introduce some useful notions in communication complexity. In a k-party communication problem, where each party holds an input x_i from a domain Δ_i, a rectangle is defined by $R := X_1 \times X_2 \times \cdots \times X_k$ ($X_i \subseteq \Delta_i$).

For a set $X_i \subseteq \Delta_i$, we denote $\boldsymbol{X}_i$ as the uniform distribution on X_i. In the set-intersection problem (particularly hard distribution 1), we consider the cases that each input is in $\Delta_i = [N]^{M_i}$ where $M_i = cN^{1-\alpha_i}$, and an instance $x_i \in [N]^{M_i}$ can be transformed into a subset of $[N]$ by removing duplicate elements. Also, for two instances $x_i \in [N]^{M_i}, x_j \in [N]^{M_j}$, we define $x_i \cap x_j$ by the intersection of the two subsets of $[N]$ deduced from x_i and x_j.

For a set of coordinates $J_i \subseteq [M_i]$, we use $\boldsymbol{X}_i(J_i)$ to denote marginal distribution of $\boldsymbol{X}_i$ on J_i. For an instance $x_i \in [N]^{M_i}$ and a set of coordinates $J_i \subseteq [M_i]$, define $x_i(J_i)$ to be an instance in $[N]^{J_i}$ by projecting x_i on J_i.

Structure-vs-pseudorandomness decomposition

We use capital letters X to denote a set and bold symbols like $\boldsymbol{R}$ to denote random variables. For a set X, we use $\boldsymbol{X}$ to denote the random variable uniformly distributed over the set X. We introduce the formal definition of *min-entropy*.

▶ **Definition 9** (Min-entropy). *For a random variable $\boldsymbol{X}$ taking value on Δ, its min-entropy is defined as follows:*

$$H_\infty(\boldsymbol{X}) = \min_{x \in \Delta} \left(\log \frac{1}{\mathbf{Pr}[\boldsymbol{X} = x]} \right).$$

A useful concept adopted in this paper is the dense notion used in lifting theorems [30, 35].

▶ **Definition 10** (Density function). *We define the one-side density function for a random variable $\boldsymbol{X}$ on its support $[N]^J$ as:*

$$\mathcal{D}(\boldsymbol{X}) := |J| \log N - H_\infty(\boldsymbol{X}).$$

Note that $\mathcal{D}(\boldsymbol{X}) \geq 0$ always holds by definitions and $\mathcal{D}(\boldsymbol{X}) = 0$ when $\boldsymbol{X}$ is a uniform distribution.

The density function is also known as the entropy deficiency in lifting theorem papers, and we design the k-side density function in order to extend the two-party results to the k-party setting.

▶ **Definition 11** (k-side density function). *For a structured rectangle $R = X_1 \times X_2 \times \cdots \times X_k$, where each X_i is subset of $[N]^{M_i}$ and associated with a set $J_i \subseteq [M_i]$, we define its k-side density function as:*

$$\mathcal{D}(R) = \mathcal{D}\big(\boldsymbol{X}_1(J_1)\big) + \mathcal{D}\big(\boldsymbol{X}_2(J_2)\big) + \cdots + \mathcal{D}\big(\boldsymbol{X}_k(J_k)\big).$$

In structure-vs-pseudorandomness decomposition, one of the most important notions, which captures the pseudorandomness, is the block-wise density.

▶ **Definition 12** (Block-wise density [29]). *For $\gamma > 0$. A random variable $\boldsymbol{X}$ supported on $[N]^n$ is said to be γ-dense if for all nonempty $I \subseteq [n]$, we have that $H_\infty(\boldsymbol{X}(I)) \geq \gamma \cdot |I| \cdot \log N$, here $\boldsymbol{X}(I)$ is the marginal distribution of $\boldsymbol{X}$ on the set I.*

The definition of γ-dense measures the pseudorandomness of a random variable. In our proof, a typical choice of $\gamma = 1 - \frac{1}{10k \log N}$ for set intersection and $\gamma = 1 - \frac{2\epsilon}{k}$ for the Chain Index problem.[3]

The following lemma tells us that a random variable could be decomposed by a combination of random variables with dense properties by fixing some positions:

▶ **Lemma 13** (Density-restoring partition [30]). *Let X be a subset of $[N]^M$ and J be a subset of $[M]$, and there exists an $\beta \in N^{J^c}$ such that $\forall x \in X, x(J^c) = \beta$. Then, there exists a partition of X:*

$$X := X^1 \cup X^2 \cup \cdots \cup X^r$$

such that every X^i is associated with a set $I_i \subseteq J$ and a value $\tau_i \in [N]^{I_i}$. Then, they satisfy the following properties:

1. $\forall x \in X^i, x(I_i) = \tau_i$;
2. $\boldsymbol{X}^i(J - I_i)$ *is γ-dense;*
3. $D\Big(\boldsymbol{X}^i(J - I_i)\Big) \le D\Big(\boldsymbol{X}(J)\Big) - (1-\gamma)|I_i| \log N + \delta_i$.

Here, we define $\delta_i := \log(|X|/|\cup_{j \ge i} X^j|)$.

We also use the following simple version of Lemma 13 for some proofs.

▶ **Proposition 14.** *Let $Z_1, \ldots, Z_T$ be a partition of the set Z. Then*

$$\sum_{i=1}^{T} \frac{|Z_i|}{|Z|} \cdot \log |Z_i| \ge \log |Z| - \log T.$$

For dense random variables, we also have the following useful lemma.

▶ **Lemma 15.** *If $\boldsymbol{X}_1, \boldsymbol{X}_2, \cdots, \boldsymbol{X}_\ell$ are $\ell < k$ independent $\big(1 - \frac{1}{10k \log N}\big)$-dense random variables and each $\boldsymbol{X}_i$ takes value from $[N]^{J_i}$ with $|J_i| \le c \cdot N^{1-\alpha_i}$, where c is a constant and $N^{\alpha_i} = \omega(k)$, then for any element $a \in [N]$, it holds*

$$\mathbf{Pr}\left[a \in \bigcap_{i=1}^{\ell} \boldsymbol{X}_i\right] \le \frac{e c^\ell}{N^{\sum_i \alpha_i}},$$

here $e \approx 2.7$ denotes the Euler's number.

Proof. We know that all $\boldsymbol{X}_i$'s are independent. Thus, we first bound the probability that $\mathbf{Pr}[a \in \boldsymbol{X}_i]$. Assuming that $J_i = (j_1, j_2, \cdots, j_{|J_i|})$, we have the following argument

$$\mathbf{Pr}[a \in \boldsymbol{X}_i] = \mathbf{Pr}[a \in \bigcup_{q=1}^{|J_i|} \boldsymbol{X}_i(j_q)] \le \sum_{q=1}^{|J_i|} \mathbf{Pr}[a \in \boldsymbol{X}_i(j_q)] \le \frac{c(1+1/k)}{N^{\alpha_i}},$$

where the last inequality comes from the definition of $(1 - \frac{1}{10k \log N})$-dense. Hence, we know that

$$\mathbf{Pr}[a \in \cap_i \boldsymbol{X}_i] = \prod_i \mathbf{Pr}[a \in \boldsymbol{X}_i] \le c^\ell (1+1/k)^\ell \cdot \frac{1}{N^{\sum_i \alpha_i}} \le c^\ell \frac{e}{N^{\sum_i \alpha_i}}.$$ ◀

[3] $\gamma = 0.9$ in previous structure-vs-pseudorandomness decomposition [30, 20, 19, 35].

3 Lower Bounds for Set Intersection

In this section, we prove the communication lower bound for the hardness distribution 1 (where each player i gets cN^{α_i} independent and uniform samples from $[N]$). Then, in Section 3.3, we use reductions to obtain lower bounds for hardness distributions 2 and 3. Formally, we prove that:

▶ **Theorem 16.** *If a communication protocol* Π *solves* k*-party set-intersection problem under the hardness distribution 1 with accuracy bigger than* 0.1*, the communication complexity* $CC(\Pi)$ *is* $\Omega\left(\frac{N^{\sum_i \alpha_i - \max_i\{\alpha_i\}}}{k}\right)$.

3.1 The Decomposition and Sampling Process

The key idea of this proof, as we introduce in Section 1, is to decompose rectangles (nodes[4]) of the protocol tree into structured rectangles and analyze the accuracy of the protocol could achieve in those decomposed structured rectangles. We design a *decomposition and sampling process* in this section to

- decompose the rectangles of the protocol tree into structured rectangles;
- sample a decomposed rectangle with respect to its size.

We define the root rectangle of the protocol tree to be R^{root}, which contains all valid inputs. R^{root} is also a structured rectangle by definitions. We start from R^{root} and begin our decomposition and sampling process, which uses a random walk on the protocol tree from the root R^{root} to a leaf, and do the decomposition along the path. See Algorithm 1 for the formal decomposition process.

We use R^{cur} to denote the current rectangle of the decomposition and sampling process. It begins with $R^{\text{cur}} = R^{\text{root}}$, and at each step R^{cur} is partitioned into two subrectangles R^0, R^1 by the protocol. Then, we replace R^{cur} with R^0 or R^1 with probability $|R^0|/|R^{\text{cur}}|$ or $|R^1|/|R^{\text{cur}}|$ (which also equals to $|X^0|/|X_i^{\text{cur}}|$ or $|X^1|/|X_i^{\text{cur}}|$ as we defined in Algorithm 1), and reach a new rectangle. After reaching the new rectangle, the structured property of R^{cur} may get destroyed, and our decomposition works here to maintain the structured property. We use the density-restoring partition (Lemma 13) to further decompose the current rectangle R^{cur} into r subrectangles $R^{\text{cur}} = R^1 \cup R^2 \cup \cdots \cup R^r$, and each R^j is a structured rectangle. Again, we choose R^j to be our next rectangle with probability $|R^j|/|R^{\text{cur}}|$, and do the process above recursively until reaching a leaf rectangle. As shown in the decomposition and sampling process, we eventually sample a structured rectangle in the leaf level with respect to its size.

Note that at some point of the random walk, the current rectangle R^{cur} *may not exist on the protocol tree since we do the density-restoring partition to further decompose the rectangles. However, every* R^{cur} *that potentially appears in the random walk must be fully contained in a rectangle of the protocol tree. Thus, the protocol* Π *also partitions* R^{cur} *into two sub-rectangles if* R^{cur} *is not in the leaf level of the protocol tree.*

Note that the output R^{cur} of the process above is a random variable over rectangles. We define $\boldsymbol{R}^{\text{leaf}}$ to be the random variables over decomposed structured rectangles in the leaf level (not leaf rectangles of the protocol tree, but sub-rectangles of those leaves after decomposition) sampled by the process above, and $\boldsymbol{R}^{\text{leaf}}$ is associated with random sets $\boldsymbol{J}_i^{\text{leaf}}$s. For convenience, we define the support of $\boldsymbol{R}^{\text{leaf}}$ to be $\mathcal{R}^{\text{leaf}}$. One may see the two important properties of the decomposition and sampling process:

[4] Note that a node of the protocol tree is a rectangle.

Algorithm 1 The decomposition and sampling process.

Input: A rectangle $R^{\text{root}} = X_1 \times X_2 \times \cdots \times X_k$, where each X_i equals $[N]^{cN^{1-\alpha_i}}$.
Output: A rectangle $R^{\text{cur}} = X_1^{\text{cur}} \times X_2^{\text{cur}} \times \cdots \times X_k^{\text{cur}}$, and k sets $J_1, J_2, \cdots, J_k$.

for each i, $J_i \leftarrow [cN^{1-\alpha_i}]$;
$R^{\text{cur}} \leftarrow R^{\text{root}}$;
while R^{cur} *is not in a leaf level* [5] **do**
  without loss of generality, we assume it is player i's turn to speak;
  X_i^{cur} is partitioned by: $X_i^{\text{cur}} = X^0 \cup X^1$, and R^{cur} is thus partitioned by: $R^{\text{cur}} = R^0 \cup R^1$;
  toss a $(\frac{|X^0|}{|X_i^{\text{cur}}|}, \frac{|X^1|}{|X_i^{\text{cur}}|})$ biased coin c;
  if $c = 0$:
    $X_i^{\text{cur}} \leftarrow X^0$;
    $R^{\text{cur}} \leftarrow R^0$;
  if $c = 1$:
    $X_i^{\text{cur}} \leftarrow X^1$;
    $R^{\text{cur}} \leftarrow R^1$;
  if $\boldsymbol{X}_i^{\text{cur}}(J_i)$ is $(1 - \frac{1}{10k \log N})$-dense:
    continue;
  else:
    decompose X_i^{cur} by Lemma 13 with $J = J_i$, get $X^1, \cdots, X^r, I_1, \cdots, I_r$;
    R^{cur} is thus decomposed by $R^{\text{cur}} = R^1 \cup \cdots \cup R^r$;
    sample a random element $\boldsymbol{j} \in [r]$: $\boldsymbol{j}$ w.p. $|X^j|/|X_i^{\text{cur}}|$ equals j for each j;
    $X_i^{\text{cur}} \leftarrow X^j$, $J_i \leftarrow J_i \backslash I_j$;
    $R^{\text{cur}} \leftarrow X_1^{\text{cur}} \times X_2^{\text{cur}} \times \cdots \times X_k^{\text{cur}}$;

- Every rectangle $R \in \mathcal{R}^{\text{leaf}}$ is a structured rectangle;
- For a rectangle $R = X_1 \times X_2 \times \cdots \times X_k \in \mathcal{R}^{\text{leaf}}$, we have that

$$\mathbf{Pr}[\boldsymbol{R}^{\text{leaf}} = R] = \prod_i \frac{|X_i|}{N^{cN^{1-\alpha_i}}} = \frac{|R|}{N^{c\sum_i N^{1-\alpha_i}}}.$$

The verification of the two properties is straightforward from the definition of our decomposition and sampling process. The first statement offers a structured property that makes it easier to analyze the rectangles. The second statement tells us that the probability that $\boldsymbol{R}^{\text{leaf}} = R$ equals the probability that the input lies in R. This is crucial in later bounding the accuracy of Π.

Next, we bound the accuracy of Π. For every structured rectangle $R = X_1 \times X_2 \times \cdots \times X_k \in \mathcal{R}^{\text{leaf}}$ associated with $J_1, J_2, \cdots, J_k$, we define J_i^c as $[cN^{1-\alpha_i}] - J_i$, namely the fixed parts of X_i. Hence, for each X_i, it holds $\forall x \in X_i, x(J_i^c) = \beta_i$ since R is a structured rectangle. We can then divide all the rectangles in $\mathcal{R}^{\text{leaf}}$ into two types:

1. R is a *bad* structured rectangle if $\cap_i \beta_i \neq \emptyset$;
2. R is a *good* structured rectangle if $\cap_i \beta_i = \emptyset$.

Assume R is a bad structured rectangle. Then, there exists a universal common element a[6] such that $a \in \cap_i x_i$ for any possible instance $(x_1, x_2, \cdots, x_k)$ in R. The protocol is thus able to

[6] We can choose any element that lies in $\cap_i \beta_i$ here.

achieve perfect correctness by outputting a when the input lies in R. Hence, we need to show with a low probability that $\boldsymbol{R}^{\text{leaf}}$ is a bad rectangle, namely the probability that the input lies in bad rectangles is small. To be more specific, we prove the following lemma:

▶ **Lemma 17.** *If* $CC(\Pi) \le 0.0001N^{\sum_i \alpha_i - \max_i\{\alpha_i\}}/k$, *it holds that* $\mathbf{Pr}_{R\sim\boldsymbol{R}^{leaf}}[R \text{ is bad}] \le 0.05$.

For those good structured rectangles, we show the following facts: For a good structured rectangle, the protocol Π cannot achieve high accuracy since there is no intersection on the fixed parts, while the other parts are dense. Formally, we prove the following lemma:

▶ **Lemma 18.** *For a good structured rectangle* $R = X_1 \times X_2 \times \cdots \times X_k$, *it holds that for any* $a \in [N]$,

$$\mathbf{Pr}[a \in \cap_i \boldsymbol{X}_i] \le 0.05.$$

Combining the three lemmas above, we can easily prove Theorem 16.

Proof of Theorem 16. We prove the theorem by showing that communication protocol Π with $\mathrm{CC}(\Pi) \le 0.0001N^{\sum_i \alpha_i - \max_i\{\alpha_i\}}/k$ can achieve at most 0.1 accuracy.

It is well known that a communication protocol Π partitions the whole input domain into several leaf rectangles and assigns an answer to each leaf rectangle. With our decomposition and sampling process, original leaf rectangles are further decomposed into the two types of structured rectangles mentioned above. The accuracy of Π comes from the following two parts:

1. The probability $\mathbf{Pr}[\boldsymbol{R}^{\text{leaf}} \text{ is bad}] = p_1$.
2. The probability that the protocol outputs the correct answer in a good structured rectangle is p_2.

From Lemma 17 and 18, we know that $p_1 \le 0.05, p_2 \le 0.05$. By a union bound, the total accuracy is thus no more than $p_1 + p_2 \le 0.1$ as desired. ◀

It suffices to prove the two important lemmas above.

3.2 Proofs of Technical Lemmas

We first prove Lemma 17 by the following round-by-round analysis.

Proof of Lemma 17. Recall the decomposition process from line 4 to line 12. In each communication round, player i sends one bit, and partitions X_i^{cur} into two parts X^0, X^1. Then, X_i^{cur} is replaced by X^0 (or X^1) with probability $\frac{|X^0|}{|X_i^{\text{cur}}|}$ (or $\frac{|X^1|}{|X_i^{\text{cur}}|}$). In this process, the density function $\mathcal{D}(\boldsymbol{X}_i^{\text{cur}}(J_i))$ would increase since the size of $|X_i^{\text{cur}}|$ decreases. This contributes to the density function with an increment of:

- $\log(\frac{|X_i^{\text{cur}}|}{|X^0|})$ with probability $|X^0|/|X_i^{\text{cur}}|$;
- $\log(\frac{|X_i^{\text{cur}}|}{|X^1|})$ with probability $|X^1|/|X_i^{\text{cur}}|$.

Thus, in expectation, the density function of $R^{\text{cur}} = X_1^{\text{cur}} \times X_2^{\text{cur}} \times \cdots \times X_k^{\text{cur}}$ after partitioning will increase

$$\frac{|X^0|}{|X_i^{\text{cur}}|} \log\left(\frac{|X_i^{\text{cur}}|}{|X^0|}\right) + \frac{|X^1|}{|X_i^{\text{cur}}|} \log\left(\frac{|X_i^{\text{cur}}|}{|X^1|}\right) \le 1, \tag{1}$$

where $|X_i^{\text{cur}}|$ denotes the size of X_i^{cur} before partitioning. Furthermore, if $\boldsymbol{X}_i^{\text{cur}}(J_i)$ is no longer $(1 - \frac{1}{10k\log n})$-dense, we partition X_i^{cur} by Lemma 13 and get $X_i^{\text{cur}} = X^1 \cup X^2 \cup \cdots \cup X^r$ and

$I_1 \cup I_2 \cup \cdots \cup I_r$ with $X^j(I_j) = \tau_j$ for all j. We use Lemma 13, where we take $\gamma = 1 - 1/(10k \log N)$, and get:

$$\mathcal{D}(\boldsymbol{X}^j(J_i - I_i)) \le \mathcal{D}(\boldsymbol{X}_i^{\mathrm{cur}}(J_i)) - (1-\gamma)|I_j| \log N + \gamma_j = \mathcal{D}(\boldsymbol{X}_i^{\mathrm{cur}}(J_i)) - \frac{|I_j|}{10k} + \delta_j. \tag{2}$$

Recall that $\delta_j := \log(|X_i^{\mathrm{cur}}|/|\cup_{p \ge j} X^p|)$ here. In the decomposition process, X_i^{cur} is replaced with X^j with probability $|X^j|/|X_i^{\mathrm{cur}}|$. Hence, taking the expectation in one communication round, we have

$$\mathbb{E}[\delta_j] = \sum_j \frac{|X^j|}{|X_i^{\mathrm{cur}}|} \log(|X_i^{\mathrm{cur}}|/|\cup_{p \ge j} X^p|) \le \int_0^1 \log \frac{1}{1-x} dx = 1. \tag{3}$$

Thus, combining (1), (2) and (3) and taking expectations, we know that after $\mathrm{CC}(\Pi)$ rounds of communication (where each round communicates exact one bit message), it holds:

$$\mathbb{E}_{R \sim \boldsymbol{R}^{\mathrm{leaf}}}[\mathcal{D}(R)] \le 2 \cdot \mathrm{CC}(\Pi) - \frac{\mathbb{E}_{J_1 \sim \boldsymbol{J}_1^{\mathrm{leaf}}, \cdots, J_k \sim \boldsymbol{J}_k^{\mathrm{leaf}}} \left[\sum_{j=1}^k |J_j^c| \right]}{10k}.$$

Here, the $2 \cdot \mathrm{CC}(\Pi)$ comes from (1) and (3). We know that $\mathbb{E}_{R \sim \boldsymbol{R}^{\mathrm{leaf}}}[\mathcal{D}(R)] \ge 0$ from definitions. Hence, we have

$$\sum_{j=1}^k \mathbb{E}_{J_j \sim \boldsymbol{J}_j^{\mathrm{leaf}}}[|J_j^c|] \le 20k \cdot \mathrm{CC}(\Pi). \tag{4}$$

We can bound the probability that the bad structured rectangle appears round by round. At each round of communication, if we choose X^j to replace X_i^{cur}, then we will fix $|I_j|$ more positions for X_i^{cur}. We then consider the probability that this new fixed part contributes to forming a bad structured rectangle with future fixed positions.

Let $R^j = X_1^{\mathrm{cur}} \times X_2^{\mathrm{cur}} \times \cdots X^j \cdots \times X_k^{\mathrm{cur}}$, for any $x = (x_1^{\mathrm{cur}}, x_2^{\mathrm{cur}}, \cdots, x^j, \cdots, x_k^{\mathrm{cur}}) \in R^j$, we label it as a error term if $\exists a \in \tau_j$, $a \in \bigcap_{p \ne i} x_p^{\mathrm{cur}}(J_p)$ [7]. By Lemma 15, for any $a \in \tau_j$,

$$\mathbf{Pr}[a \in \bigcap_{p \ne i} \boldsymbol{X}_p^{\mathrm{cur}}(J_p)] \le \frac{ec^{k-1}}{N^{(\sum_{p=1}^k \alpha_p) - \alpha_i}}$$

By a union bound, the probability that error terms appear in R^j is

$$\mathbf{Pr}[\exists a \in \tau_j, a \in \bigcap_{p \ne i} \boldsymbol{X}_p^{\mathrm{cur}}(J_p)] \le \frac{|I_j| \cdot ec^{k-1}}{N^{(\sum_{p=1}^k \alpha_p) - \alpha_i}}$$

Also, we know that the total number of fixed elements equals $\sum_{i=1}^k |J_i^c|$, which is identical to the summation of $|I_j|$ of every step, thus, the average probability of error terms at the end of the decomposition process is at most

$$\frac{ec^{k-1}}{N^{(\sum_i \alpha_i) - \max_i\{\alpha_i\}}} \cdot \sum_{i=1}^k \mathbb{E}_{J_i \sim \boldsymbol{J}_i^{\mathrm{leaf}}} \left[|J_i^c| \right].$$

[7] τ_j is a fixed subset of $[N]$ with size at most $|I_j|$ since X^j is fixed on I_j. Input x may be labeled many times during the decomposition process.

We note that for any $R \in \mathcal{R}^{\text{leaf}}$, if R is bad, then all instances $x \in R$ have been labeled as an error term in the decomposition process, together with (4), we have

$$\Pr_{R\sim \boldsymbol{R}^{\text{leaf}}}[R \text{ is bad}] \le \frac{ec^{k-1}}{N^{(\sum_i \alpha_i)-\max_i\{\alpha_i\}}} \cdot \sum_{i=1}^{k} \mathbb{E}_{J_i\sim \boldsymbol{J}_i^{\text{leaf}}}\left[|J_i^c|\right] \le 0.05.$$

The last inequality holds since $c = (1 + 2/k)$ and $\text{CC}(\Pi) \le 0.0001N^{\sum_i \alpha_i - \max_i\{\alpha_i\}}/k$. ◀

Next, we show that in the good structured rectangles, the protocol Π cannot achieve large accuracy in finding the common element. This also comes from the structured properties of the rectangles:

Proof of Lemma 18. Notice that we consider the rectangle $R = X_1 \times X_2 \times \cdots \times X_k$ associated with $J_1, J_2, \cdots J_k$ that has no common elements on fixed parts J_i^c. Thus, for any element $a \in [N]$, there exists at least a party i which does not contain a on its fixed part. Thus, we use Lemma 15 for $\boldsymbol{X}_i(J_i)$ with $\ell = 1$, and get

$$\mathbf{Pr}[a \in \boldsymbol{X}_i] = \mathbf{Pr}[a \in \boldsymbol{X}_i(J_i)] \le ce/N^{\alpha_i} = o(1).$$ ◀

3.3 Lower Bounds for Other Hardness Distributions

In this section, we first establish a reduction from Bernoulli hardness distribution (hardness distribution 3) to hardness distribution 2 by the following lemma:

▶ **Lemma 19.** *If a communication protocol Π that solves set-intersection under hardness distribution 3 with accuracy ϵ, there exists parameters $c_1, \cdots, c_k$ with each $1-1/k \le c_i \le 1+1/k$ for hardness distribution 2 so that Π can find set intersection under this distribution with accuracy $\epsilon - 2k\exp(-\frac{N^{1-\max_i\{\alpha_i\}}}{3k^2})$, which is bigger than $\epsilon - 0.01$ when $N^{1-\max_i\{\alpha_i\}} \ge 100k^2\log k$.*

Proof. We first use Chernoff bound to bound the probability of the size of set S_i of each player i exceeding $(1 + 1/k) \cdot N^{1-\alpha_i}$ or less than $(1 - 1/k) \cdot N^{1-\alpha_i}$ under the hardness distribution 3:

$$\mathbf{Pr}[||S_i| - N^{1-\alpha_i}| > 1/k \cdot N^{1-\alpha_i}] \le 2\exp\left(-\frac{N^{1-\alpha_i}}{3k^2}\right).$$

We use A to denote the event that $\exists i, ||X_i| - N^{1-\alpha_i}| > 1/k \cdot N^{1-\alpha_i}$. Then, by a union bound, we know that:

$$\mathbf{Pr}[A] \le 2k \cdot \exp\left(-\frac{N^{1-\max_i\{\alpha_i\}}}{3k^2}\right).$$

Then, condition on $\neg A$, we have the success probability of Π in finding set intersection under hardness distribution 3 is bigger than $\epsilon - 2k \cdot \exp\left(-\frac{N^{1-\max_i\{\alpha_i\}}}{3k^2}\right)$. Furthermore, condition on $\neg A$, the hardness distribution 3 can be represented by a combination of product distributions:

$$\sum_{c_1,c_2,\cdots,c_k} \sigma(c_1, c_2, \cdots, c_k) D_{c_1,c_2,\cdots,c_k},$$

where $D_{c_1,c_2,\cdots,c_k}$ denotes the hardness distribution 2 with parameters $c_1, c_2, \cdots, c_k$. Then, the lemma follows from an averaging argument. ◀

It suffices to construct a reduction from hardness distribution 2 to hardness distribution 1.

▶ **Lemma 20.** *If there exists a communication protocol* Π *with communication complexity* C *which solves set-intersection under hardness distribution 2 with accuracy* ϵ*, there exists a communication protocol* Π' *with communication complexity* C *which solves set-intersection under hardness distribution 1 with accuracy* $\epsilon - 0.05$ *when* $k^2 N^{-\min_i\{\alpha_i\}} \leq \frac{1}{100}$ *holds.*

Proof. We construct the communication protocol Π' as follows:
1. For each player i, remove the duplicate elements of its input and get a $S_i \subseteq [N]$.
2. Randomly sample $c_i N^{1-\alpha_i}$ elements from S_i, Π' fail if $|Y_i| < c_i N^{1-\alpha_i}$.
3. Run the communication protocol Π on Y_is to find intersection.

We know that the successful probability of Π' under hardness distribution 1 is bigger than

$$\epsilon - \mathbf{Pr}[\Pi' \text{ fail at step 2}].$$

It suffices to bound $\mathbf{Pr}[\Pi'$ fail at step 2$]$. From the union bound, we have:

$$\begin{aligned}\mathbf{Pr}[\Pi' \text{ fail at step 2}] &\leq k \cdot \mathbf{Pr}[|S_i| < c_i N^{1-\alpha_i}] \\ &\leq k \cdot \mathbf{Pr}[\#\text{repeated elements in } S_i > (c - c_i) N^{1-\alpha_i}].\end{aligned}$$

We know that

$$\mathbb{E}[\#\text{repeated elements}] = cN^{1-\alpha_i}\left(1 - (1 - 1/N)^{cN^{1-\alpha_i}-1}\right) \leq c^2 N^{1-2\alpha_i}.$$

From Markov's Inequality, we have

$$\mathbf{Pr}[\#\text{repeated elements} > (c-c_i)N^{1-\alpha_i}] \leq \mathbb{E}[\#\text{repeated elements}]/(c-c_i)N^{1-\alpha_i} \leq kc^2 N^{-\alpha_i}.$$

If $kN^{-\alpha_i} \leq \frac{1}{100k}$ holds, which is guaranteed by the constraints, $\mathbf{Pr}[\Pi'$ fail at step 2$] \leq 0.05$ also holds. This concludes the lemma. ◀

3.4 Efficient Protocols for the Hardness Distribution

In this section, we first explain an efficient protocol for the hardness distribution 3, where we use D_3 to denote the distribution, showing that our lower bound result is almost tight for this distribution. Also, this protocol can be easily extended to some more general product distributions sharing *"similarities"* with the Bernoulli product distribution. Formally, we prove:

▶ **Theorem 21.** *There is a protocol* Π*, which solves the hardness distribution 3, with* $Acc_\Pi(D_3) \geq 0.1$ *and*

$$CC(\Pi) = O(N^{\sum_i \alpha_i - \max_i\{\alpha_i\}} \log N).$$

Furthermore, this protocol can be extended to more general distributions. Let D *be any distribution that satisfies the following properties:*
1. *each party holds a set of size* $\Theta(N^{1-\alpha_i})$*;*
2. *the size of intersecting part of all parties is* $\Omega(N^{1-\sum_i \alpha_i})$*;*

there exists a protocol Π' *with* $O(k \log N \cdot N^{\sum_i \alpha_i - \max_i\{\alpha_i\}})$ *communication cost that achieves* $\Omega(1)$ *accuracy under* D*.*

Proof. To begin with, we first propose an efficient protocol to solve D_3. Without loss of generality, we assume $\alpha_1 \geq \alpha_2 \geq \cdots \geq \alpha_k$ and each party i gets a subset $S_i \subseteq [N]$. Then, the communication protocol Π proceeds as follows:

1. The first party uniformly and randomly picks $\min\{|S_1|, N^{\sum_i \alpha_i - \max_i\{\alpha_i\}}\}$ elements from S_1 and sends them, denoted by M_1, to the second party.
2. The second party receives the message M_1 from the first one, and sends $M_2 := M_1 \bigcap S_2$ to the third party.
3. The process goes on, and the last party computes $M_{k-1} \bigcap S_k$. If it is not empty, the last party outputs any element in it. Otherwise, the protocol fails.

Then, we bound $\mathrm{Acc}_\Pi(D_3)$ and its communication complexity to show Π is highly efficient. From the definitions, we know that

$$\mathrm{Acc}_\Pi(D_3) = \mathbf{Pr}[M_1 \cap S_2 \cap \cdots \cap S_k \neq \emptyset].$$

Also, we have that

$$\mathbf{Pr}[M_1 \cap S_2 \cap \cdots \cap S_k \neq \emptyset | |M_1| = m] = 1 - \left(1 - \frac{1}{N^{\sum_i \alpha_i - \max_i\{\alpha_i\}}}\right)^m \geq \frac{m}{e \cdot N^{\sum_i \alpha_i - \max_i\{\alpha_i\}}}.$$

The last inequality holds since $m \leq N^{\sum_i \alpha_i - \max_i\{\alpha_i\}}$. From Chernoff bound, we know that the probability that $\mathbf{Pr}\left[|M_1| \leq N^{\sum_i \alpha_i - \max_i\{\alpha_i\}}/2\right] \leq e^{-N^{1-\alpha_1}/12} \leq e^{-10k^3}$. The last inequality is from the constraint of $k \leq 0.1 \cdot \min\{N^{\min_i\{\alpha_i\}/2}, N^{(1-\max_i\{\alpha_i\})/3}\}$. Furthermore, when

$$|M_1| \geq N^{\sum_i \alpha_i - \max_i\{\alpha_i\}}/2,$$

it holds that

$$\mathbf{Pr}[M_1 \cap S_2 \cap \cdots \cap S_k \neq \emptyset | |M_1| = m] \geq \frac{1}{2e}.$$

Combining the facts above, we have $\mathrm{Acc}_\Pi(D_3) \geq \frac{1}{2e}\left(1 - e^{-10k^3}\right) \geq 0.1$.

On the other hand, we bound the communication complexity by bounding the expected size of $|M_i|$. $\mathbb{E}[|M_1|] \leq N^{\sum_i \alpha_i - \max_i\{\alpha_i\}} \log N$ holds from definitions. Furthermore, we have

$$\mathbb{E}[|M_i|] \leq \mathbb{E}[|M_{i-1}|] \cdot N^{-\alpha_i}.$$

Then, $\mathbb{E}[\sum_i M_i] \leq O(N^{\sum_i \alpha_i - \max_i\{\alpha_i\}} \log n)$ follows by $N^{-\alpha_i} \leq N^{\min_i\{\alpha_i\}} \leq 1/2$. This concludes the first statement.

Next, we slightly change the protocol above to match the second statement. The protocol Π' proceeds as follows:

1. The first party uniformly and randomly picks $\Theta(N^{\sum_i \alpha_i - \max_i\{\alpha_i\}})$ elements from S_1 and sends them, denoted by M_1, to the second party.
2. The second party receives the message M_1 from the first one, and sends $M_2 := M_1 \bigcap S_2$ to the third party.
3. The process goes on, and the last party computes $M_{k-1} \bigcap S_k$. If it is not empty, the last party outputs any element in it.

Obviously, the communication complexity of this protocol Π' is $O(k \log n \cdot N^{\sum_i \alpha_i - \max_i\{\alpha_i\}})$. Also, we know the accuracy is bigger than

$$\Omega\left(1 - (1 - \frac{\Omega(N^{1-\sum_i \alpha_i})}{|S_1|})^{\Theta(N^{\sum_i \alpha_i - \max_i\{\alpha_i\}})}\right) = \Omega(1). \quad \blacktriangleleft$$

Thus, our lower bounds show that those trivial protocols are nearly optimal.

4 Lower Bounds for Chained Index

Recall that in the Chained Index problem, the player i receives an input $z_i = (\sigma_i, x_i) \in [n] \times \{0,1\}^n$. The players aim to compute $x_{k-1}(\sigma_k)$ through a one-way communication. In this section, we show an improved lower bound for the Chained Index problem. In light of Yao's principle, we consider the following hard distribution.

The distribution χ_k.

1. Uniformly sample $\sigma_1, ..., \sigma_k \in [n]$.
2. Sample $(x_1, \ldots, x_k) \sim (\{0,1\}^n)^k$ conditioned on $x_1(\sigma_2) = \cdots = x_{k-1}(\sigma_k)$.
3. Output $z = (z_1, \ldots, z_k)$ where $z_i = (\sigma_i, x_i)$ for every $i \in [k]$.

For a subset $R \subseteq ([n] \times \{0,1\}^n)^k$, define the weight of R under χ_k as

$$\chi_k(R) \stackrel{\text{def}}{=} \Pr_{z \sim \chi_k}[z \in R] = \frac{\#\{((\sigma_1, x_1), \ldots, (\sigma_k, x_k)) \in R : x_1(\sigma_2) = \cdots = x_{k-1}(\sigma_k)\}}{n^k \cdot 2^{(k-1)(n-1)+n+1}}.$$

We prove the following lower bound. We say a one-way k-party protocol has ***signature*** $(C_1, \ldots, C_k)$ if, for each $i \in [k]$, the i-th party sends at most C_i bits (on all inputs).

▶ **Theorem 22.** *Let $\varepsilon \in (0, 1/4]$ be a constant. Let Π be a protocol for the k-party chained index problem with signature $(C_1, \ldots, C_k)$. If Π has 2ε advantage, i.e.,*

$$\Pr_{z=(\sigma_1,x_1,\ldots,\sigma_k,x_k)\sim\chi_k}[\Pi(z) = x_{k-1}(\sigma_k)] \geq \frac{1}{2} + 2\varepsilon.$$

Then $\sum_{t=1}^{k} C_t \geq \max\left\{\frac{1}{32}\varepsilon^2 n/k, \frac{1}{8}\varepsilon\sqrt{n}\right\} = \Omega(n/k + \sqrt{n})$.

We use a decomposition and sampling process DS, as shown in Algorithm 2, in our analysis. DS takes as input a protocol Π, and samples a rectangle R that is contained in Π_v for some leaf node v. Our proof proceeds in two steps:

1. Section 4.1 shows that the accuracy of Π is captured by a quantity called *average fixed size*, which is a natural quantity that arises in the running of DS.
2. Section 4.2 proves that the average fixed size can be bounded from above by $O(k \cdot \mathrm{CC}(\Pi))$. Consequently, if Π enjoys high accuracy, we get a lower bound of $\mathrm{CC}(\Pi)$.

We first recall some basic definitions.

k-party one-way protocols

A deterministic k-party one-way communication protocol Π is specified by a rooted binary tree. For every internal vertex v,

- it has 2 children, denoted by $\Pi(v, 0)$ and $\Pi(v, 1)$;
- v is owned by some party – we denote the owner by $\mathsf{owner}(v) \in [k]$;
- every leaf node specifies an output.

Starting from the root, the owner of the current node cur partitions its input space into two parts X_0 and X_1, and sets the current node to $\Pi(\mathsf{cur}, b)$ if its input belongs to X_b.

The *communication complexity* of Π, denoted by $\mathrm{CC}(\Pi)$, is the depth of the tree. On a path from root to some leaf, each time the owner switches, we call it a new *round*; in a one-way protocol, the label of the owner is non-decreasing.

▶ **Fact 23.** *The set of all inputs that lead to an internal vertex v is a rectangle, denoted by $\Pi_v = X_1 \times \cdots \times X_k$.*

Normalized protocols

We normalized a protocol Π as follows so as to make it defined on all inputs, including those not in $\mathrm{supp}(\chi_k)$. For the i-th party, given input $(\sigma_i, x_i) \in [n] \times \{0,1\}^n$ and previous transcripts trans, output 0 if the input is *invalid*, i.e., given trans, there is no input in $\mathrm{supp}(\chi_k)$ matches x_i. Otherwise, the i-th party outputs 1 and proceeds as Π. Clearly, by normalizing, we communicate k more bits.

Algorithm 2 Decomposition and Sampling Process DS.

Input: A protocol Π

Output: A rectangle $R = (\{\sigma_1\} \times X_1) \times \cdots \times (\{\sigma_k\} \times X_k)$ and k sets $J_1, \ldots, J_k \subseteq [n]$.

for $i \in [k]$ **do**
  $X_i := \{0,1\}^n, J_i := [n]$. `// Initialization`
Sample $\sigma_1 \sim [n]$.
$v :=$ root of $\Pi, R := (\{\sigma_1\} \times \{0,1\}^n) \times ([n] \times \{0,1\}^n)^{k-1}, \mathsf{bad} :=$ FALSE.
while *v is not a leaf node* **do**
  $i := \mathsf{owner}(v), u_0 := \Pi(v,0), u_1 := \Pi(v,1)$.
  `//Loop invariant: (1)` $R \subseteq \Pi_v$`; (2)` $\boldsymbol{X}_i(J_i)$ `is` γ`-dense.`
  X_i is partitioned into $X_i = X^0 \cup X^1$ according to Π.
  Sample $\boldsymbol{b}$ such that $\mathbf{Pr}[\boldsymbol{b} = b] = \frac{\chi_k(R^b)}{\chi_k(R)}$ where $R^b = (\{\sigma_1\} \times X_1) \times \cdots (\{\sigma_i\} \times X^b) \times ([n] \times \{0,1\}^n)^{k-i}$ for $b \in \{0,1\}$.
  `//`R `is always a shorthand for` $(\{\sigma_1\} \times X_1) \times \cdots \times (\{\sigma_i\} \times X_i) \times ([n] \times \{0,1\}^n)^{k-i}$
  Update $X_i := X^{\boldsymbol{b}}$.
  Let $X_i = X^1 \cup \cdots \cup X^m$ be the decomposition of X_i promised by Lemma 13 with associated sets $I_1, \ldots, I_m \subseteq J_i$. `// Invoking Lemma 13 with` $\gamma = 1 - \frac{2\varepsilon}{k}, J = J_i, N = 2$.
  Sample $\boldsymbol{j} \in [m]$ such that $\mathbf{Pr}[\boldsymbol{j}' = j] = \frac{\chi_k(R^j)}{\chi_k(R)}$ where $R^j = (\{\sigma_1\} \times X_1) \times \cdots \times (\{\sigma_i\} \times X^j) \times ([n] \times \{0,1\}^n)^{k-i}$ for $j \in [m]$.
  Update $X_i := X^j, J_i := J_i \setminus I_j$.
  if $\mathsf{owner}(u_{\boldsymbol{b}}) \neq i$ **then**
    Sample $\sigma_{i+1} \in [n]$ such that $\mathbf{Pr}[\sigma_{i+1} = \rho] = \frac{\chi_k(R^\rho)}{\chi_k(R)}$ where $R^\rho = (\{\sigma_1\} \times X_1) \times \cdots \times (\{\sigma_i\} \times X_i) \times (\{\rho\} \times \{0,1\}^n) \times ([n] \times \{0,1\}^n)^{k-i-1}$ for $\rho \in [n]$.
    if $\sigma_{i+1} \notin J_i$ **then** $\mathsf{bad} :=$ TRUE;
  Output $R = (\{\sigma_1\} \times X_1) \times \cdots \times (\{\sigma_k\} \times X_k), J_1, \ldots, J_k$.

▶ **Lemma 24** (Loop invariant). *After each iteration in algorithm 2,*

- $R \subseteq \Pi_v$;
- *for all* $i \in [k]$, $\boldsymbol{X}_i(J_i)$ *is* γ*-dense;*
- *for all* $i \in [k]$, *there exists* $\alpha_i \in \{0,1\}^{\overline{J_i}}$ *such that* $x(\overline{J_i}) = \alpha_i$ *for all* $x \in X_i$.

Proof. The first item is true because every time v is updated, R is updated accordingly to a sub-rectangle of Π_v and updating R into its sub-rectangles does not violate this condition.

Since we applied density restoring partition at the end of each iteration, the second and the third items are guaranteed by Lemma 13 and the way that X_i, J_i are updated. ◀

4.1 Relating Accuracy and Average Fixed Size

As shown in Lemma 24, during the execution of $\mathsf{DS}(\Pi)$, for every $i \in [k]$, the set X_i is "fixed" on $\overline{J_i}$ in the sense that all strings in X_i share the same value on coordinates in $\overline{J_i}$. So we call the expected size of $|\overline{J_i}|$ *average fixed size*. However, in order to relate the accuracy of Π to average fixed size, we need to consider the expectation of $|\overline{J_i}|$ in a slightly different distribution.

▶ **Definition 25** (Average fixed size)**.** *Let $\mathcal{U}_k$ denote the uniform distribution over the input space $([n] \times \{0,1\}^n)^k$. Let $t \in [k]$ and consider the following process, denoted by $\mathsf{Unif}_t(\Pi)$:*

1. *run $\mathsf{DS}(\Pi)$ until the t-th round;*
2. *continue running DS with χ_k replaced by $\mathcal{U}_k$ in the execution of Line 9, Line 13, and Line 16;*
3. *upon entering the $(t+1)$-th round (i.e., until Line 17 is reached with $i = t$), return J_t.*

The average fixed size of the t-th party is defined as $\mathbf{E}_{J_t \sim \mathsf{Unif}_t(\Pi)} \left[|\overline{J_t}| \right]$.

▶ **Lemma 26** (Relating accuracy and average fixed size)**.** *Assume that $\gamma \geq \log \left[1 + \left(\frac{1-2\varepsilon}{1+2\varepsilon} \right)^{1/k} \right]$. Then*

$$\Pr_{z=(\sigma_1,x_1,\ldots,\sigma_k,x_k)\sim\chi_k} [\Pi(z) = x_{k-1}(\sigma_k)] \leq \frac{1}{2} + \varepsilon + \frac{2}{n} \cdot \sum_{t=1}^{k} \mathop{\mathbf{E}}_{J_t \sim \mathsf{Unif}_t(\Pi)} \left[|\overline{J_t}| \right].$$

▶ Remark 27. $\gamma \stackrel{\mathsf{def}}{=} 1 - \frac{2\varepsilon}{k}$ satisfies the condition. Indeed,

$$\log \left[1 + \left(\frac{1-2\varepsilon}{1+2\varepsilon} \right)^{1/k} \right] \leq \left(\frac{1-2\varepsilon}{1+2\varepsilon} \right)^{1/k} \leq 1 - \frac{1}{k} \cdot \frac{4\varepsilon}{1+2\varepsilon} \leq 1 - \frac{2\varepsilon}{k},$$

where the first inequality is by $\log(1+x) \leq x$, and the second is by $(1-x)^r \leq 1 - rx$ for $x \in (-1, 0)$ and $r \in (0, 1)$.

The proof of the lemma is obtained through the following two lemmas. The first lemma readily says that conditioned on the flag bad is not raised, Π has little advantage in the rectangle R output by $\mathsf{DS}(\Pi)$. The second lemma shows the probability that the flag is raised is bounded in terms of the average fixed size.

▶ **Lemma 28.** *If $\mathsf{DS}(\Pi)$ outputs $(R, J_1, \ldots, J_k)$ and $\mathsf{bad} = \textsc{False}$ in the end, then*

$$\Pr_{z=(\rho_1,x_1,\ldots,\rho_k,x_k)\sim R} [\Pi(z) = x_{k-1}(\rho_k) | x_1(\rho_2) = \cdots = x_{k-1}(\rho_k)] \leq \frac{1}{2} + \varepsilon.$$

▶ **Lemma 29.** $\Pr_{\mathsf{DS}(\Pi)} [\mathsf{bad} = \textsc{True}] \leq \frac{2}{n} \cdot \sum_{t=1}^{k} \mathbf{E}_{J_t \sim \mathsf{Unif}_t(\Pi)} \left[|\overline{J_t}| \right]$.

Next, we first prove Lemma 26 using the above two lemmas.

Proof of Lemma 26. Note that in the running of $\mathsf{DS}(\Pi)$, we first sample $\sigma_1, \ldots, \sigma_k \sim [n]$ and then always update R to a randomly chosen rectangle; the probability of each rectangle being chosen is proportional to its weight under χ_k. Consequently,

$$\begin{aligned}
&\Pr_{z=(\sigma_1,x_1,\ldots,\sigma_k,x_k)\sim\chi_k} [\Pi(z) = x_{k-1}(\sigma_k)] \\
&= \Pr_{\substack{(R,J_1,\ldots,J_k)\sim\mathsf{DS}(\Pi) \\ (\sigma_1,x_1,\ldots,\sigma_k,x_k)\sim R}} [\Pi(z) = x_{k-1}(\sigma_k) \mid x_1(\sigma_2) = \cdots = x_{k-1}(\sigma_k)] \\
&\leq \Pr_{\mathsf{DS}(\Pi)} [\mathsf{bad} = \textsc{True}] + \Pr_{\substack{(R,J_1,\ldots,J_k)\sim\mathsf{DS}(\Pi) \\ z=(\sigma_1,x_1,\ldots,\sigma_k,x_k)\sim R}} [\Pi(z) = x_{k-1}(\sigma_k) | x_1(\sigma_2) = \cdots = x_{k-1}(\sigma_k) \wedge \mathsf{bad} = \textsc{False}] \\
&\leq \frac{1}{2} + \varepsilon + \frac{2}{n} \cdot \sum_{t=1}^{k} \mathop{\mathbf{E}}_{J_t \sim \mathsf{Unif}_t(\Pi)} \left[|\overline{J_t}| \right].
\end{aligned}$$

where the last step is by Lemma 28 and Lemma 29. ◀

It remains to prove the two lemmas.

Proof of Lemma 28. Say $R = (\{\sigma_1\} \times X_1) \times \cdots \times (\{\sigma_k\} \times X_k)$. Since $\mathsf{bad} = \textsc{False}$ in the end, we have $\sigma_{i+1} \in J_i$ for all $i \in [k-1]$. By Lemma 24, we have $\mathrm{H}_\infty(\boldsymbol{X}_i(\sigma_{i+1})) \geq \gamma$ for all i. Since R is contained in some leaf node of Π, Π output the same answer in R, say $b^* \in \{0,1\}$. Note that for $b \in \{0,1\}$,

$$\Pr_{z=(\rho_1,x_1,\ldots,\rho_k,x_k)\sim R}[x_1(\rho_2) = \cdots = x_{k-1}(\rho_k) = b] = \prod_{i\in[k-1]} \Pr_{x_i\sim X_i}[x_i(\sigma_{i+1}) = b],$$

since we must have $\rho_i = \sigma_i$. Write $p_i \stackrel{\mathsf{def}}{=} \mathbf{Pr}_{x^i\sim X^i}\left[x^i(\sigma_i) = b^*\right]$. Then we have

$$\begin{aligned}
&\Pr_{z=(\rho_1,x_1,\ldots,\rho_k,x_k)\sim R}[\Pi(z) = x_{k-1}(\rho_k) | x_1(\rho_2) = \cdots = x_{k-1}(\rho_k)] \\
=&\Pr_{z=(\rho_1,x_1,\ldots,\rho_k,x_k)\sim R}[x_1(\rho_2) = \cdots = x_{k-1}(\rho_k) = b^* | x_1(\rho_2) = \cdots = x_{k-1}(\rho_k)] \\
=&\Pr_{z=(\rho_1,x_1,\ldots,\rho_k,x_k)\sim R}[x_1(\sigma_2) = \cdots = x_{k-1}(\sigma_k) = b^*] \Big/ \Pr_{z=(\rho_1,x_1,\ldots,\rho_k,x_k)\sim R}[x_1(\sigma_2) = \cdots = x_{k-1}(\sigma_k)] \\
=&\frac{\prod_{i\in[k]} p_i}{\prod_{i\in[k]} p_i + \prod_{i\in[k]}(1-p_i)} = \frac{1}{1 + \prod_{i\in[k]}(1/p_i - 1)}.
\end{aligned}$$

Since $\mathrm{H}_\infty(\boldsymbol{X}_i(\sigma_{i+1})) \geq \gamma$ for all i, we have $p_i \in [1 - 2^{-\gamma}, 2^{-\gamma}]$, which implies

$$\frac{1}{1 + \prod_{i\in[k]}(1/p_i - 1)} \leq \frac{1}{1 + (2^\gamma - 1)^k}.$$

Since we assumed $\gamma \geq \log\left[1 + \left(\frac{1-2\varepsilon}{1+2\varepsilon}\right)^{1/k}\right]$, it holds that $\frac{1}{1+(2^\gamma-1)^k} \leq \frac{1}{2} + \varepsilon$, concluding the proof. ◀

Proof of Lemma 29. Let $\mathcal{B}_t$ denote the event that the flag bad is raised when $i = t$ (i.e., when the i-th round ends) for the first time. Clearly, $\mathbf{Pr}[\mathsf{bad} = \textsc{True}] = \sum_{t=1}^{k-1} \mathbf{Pr}[\mathcal{B}_t]$. It suffices to show $\mathbf{Pr}[\mathcal{B}_t] \leq \mathbf{E}_{J_t\sim\mathsf{Unif}_t(\Pi)}\left[|\overline{J_t}|\right]$. for each t.

Fix $t \in [k-1]$ and the random coins coin used for the first $(t-1)$ rounds, i.e., until Line 17 is reached with $i = t-1$). Let $R_{t-1} = (\{\sigma_1\} \times X_1 \times \cdots \times (\{\sigma_t\} \times \{0,1\}^n) \times ([n] \times \{0,1\}^n)^{k-t}$ be the value of rectangle R when running $\mathsf{DS}(\Pi)$ using coin until the t-th round begins. The core of our proof is to compare the process with one that runs under uniform weight instead of the weight under χ_k; this is why we can deal with the promise.

- Let Real_t be the process that runs $\mathsf{DS}(\Pi)$ until the t-th round begins with coin, then run the t-th round with fresh random coins.
- Let $\mathsf{Unif}_t(\Pi; \mathsf{coin})$ be the process that runs $\mathsf{DS}(\Pi)$ until the t-th round begins with coin, then runs the t-th round with χ_k replaced by $\mathcal{U}_k$.

Note that during the execution of Real_t and Unif_t, the partitions are the same, and the only difference is that when choosing $\boldsymbol{b}, \boldsymbol{j}, \sigma_{t+1}$, the probabilities are different. Let $\widehat{X}_t, \widehat{J}_t, \widehat{\sigma}_{t+1}$ be a possible value of X_t, J_t, σ_{t+1} at the end of the t-th round. In Real_t we update R according to χ_k, and thus the probability that $X_t = \widehat{X}_t, \sigma_{t+1} = \widehat{\sigma}_{t+1}$ in the end of Real_t equals

$$p(\widehat{X}_t, \widehat{\sigma}_{t+1}) = \frac{\chi_k((\{\sigma_1\} \times X_1) \times \cdots \times (\{\sigma_t\} \times \widehat{X}_t) \times (\{\widehat{\sigma}_{t+1}\} \times \{0,1\}^n) \times ([n] \times \{0,1\}^n)^{k-t-1})}{\chi_k(R_{t-1})}.$$

Similarly, the probability that $X_t = \widehat{X}_t, \sigma_{t+1} = \widehat{\sigma}_{t+1}$ in the end of $\mathsf{Unif}_t(\Pi, \mathsf{coin})$ equals

$$\begin{aligned}
q(\widehat{X}_t, \widehat{\sigma}_{t+1}) =& \frac{\mathcal{U}_k((\{\sigma_1\} \times X_1) \times \cdots \times (\{\sigma_t\} \times \widehat{X}_t) \times (\{\widehat{\sigma}_{t+1}\} \times \{0,1\}^n) \times ([n] \times \{0,1\}^n)^{k-t-1})}{\mathcal{U}_k(R_{t-1})} \\
=& \frac{|\widehat{X}_t|}{n2^n}.
\end{aligned}$$

The next claim reveals a connection between the two probabilities, whose proof is by direct calculation and is deferred to the appendix.

▷ Claim 30. For all possible value $\widehat{X}_t, \widehat{\sigma}_{t+1}$, $p(\widehat{X}_t, \widehat{\sigma}_{t+1}) \le 2q(\widehat{X}_t, \widehat{\sigma}_{t+1})$.

Since J_t is determined by the value of X_t and the event $\mathcal{B}_t$ is determined by X_t and σ_{t+1}, the above claim implies that $\mathbf{Pr}_{\mathsf{Real}_t}[\mathcal{B}_t] \le 2\,\mathbf{Pr}_{\mathsf{Unif}_t(\Pi;\mathsf{coin})}[\mathcal{B}_t]$. Note that in $\mathsf{Unif}_t(\Pi;\mathsf{coin})$, σ_{t+1} is chosen uniformly at random, and thus

$$\Pr_{\mathsf{Unif}_t(\Pi;\mathsf{coin})}[\mathcal{B}_t] \le \mathop{\mathbf{E}}_{J_t\sim\mathsf{Unif}_t(\Pi;\mathsf{coin})}\left[|\overline{J}_t|\right]/n.$$

Taking expectation over coin we get $\mathbf{Pr}[\mathcal{B}_t] \le \frac{2}{n}\cdot\mathbf{E}_{J_t\sim\mathsf{Unif}_t(\Pi)}\left[|\overline{J}_t|\right]$, as desired. ◀

4.2 Average Fixed Size is Bounded by Communication

Now that the accuracy of a protocol Π is bounded from above by the average fixed size (i.e., $\sum_{t=1}^{k}\mathbf{E}_{J_t\sim\mathsf{Unif}_t(\Pi)}\left[|\overline{J}_t|\right]$), in what follows we show that the average fixed size is at most $O(k\cdot\mathrm{CC}(\Pi))$. Formally, we prove that

▶ **Lemma 31.** *Assume that Π is a normalized protocol with signature $(C_1,\ldots,C_k)$. Then*

$$\sum_{t=1}^{k}\mathop{\mathbf{E}}_{J_t\sim\mathsf{Unif}_t(\Pi)}\left[|\overline{J}_t|\right] \le \frac{2}{1-\gamma}\cdot\sum_{t\in[k]}C_t.$$

Proof. The proof strategy is similar to the proof of Lemma 29. Fix $t\in[k-1]$ and consider $\mathbf{E}_{J_t\sim\mathsf{Unif}_t(\Pi)}\left[|\overline{J}_t|\right]$. Fix the random coins coin used for the first $(t-1)$ rounds (i.e., until Line 17 is reached with $i=t-1$). Let Real_t and $\mathsf{Unif}_t(\Pi;\mathsf{coin})$ be defined as in the proof of Lemma 29. Moreover, let c_t denote the number of bits sent by the t-th party, i.e., the number of iterations in the t-th round. By a standard density increment argument, we have

▷ Claim 32. $\mathbf{E}_{J_t\sim\mathsf{Unif}_t(\Pi;\mathsf{coin})}\left[|\overline{J}_t|\right] \le \frac{2}{1-\gamma}\mathbf{E}_{\mathsf{Unif}_t(\Pi;\mathsf{coin})}[c_t] \le \frac{2}{1-\gamma}C_t$.

Averaging over coin, we have

$$\mathop{\mathbf{E}}_{J_t\sim\mathsf{Unif}_t(\Pi)}\left[|\overline{J}_t|\right] = \mathop{\mathbf{E}}_{\mathsf{coin},J_t\sim\mathsf{Unif}_t(\Pi;\mathsf{coin})}\left[|\overline{J}_t|\right] \le \frac{2}{1-\gamma}C_t,$$

where the second inequality follows from Claim 32. By summing up all t's, we get the desired result. ◀

It remains to prove Claim 32.

Proof of Claim 32. We shall prove this lemma by a density increment argument. That is, we study the change of the density function $D_\infty(\boldsymbol{X}_t(J_t))$. in each iteration. Let $\boldsymbol{\phi_t}$ be the value of $D_\infty(\boldsymbol{X}_t(J_t))$ at the end of the ℓ-th iteration.

We fix the random coins used for the first $(\ell-1)$ iterations and consider the updates in the current iteration.

1. First, X_t is partitioned into $X_t = X^0\cup X^1$ according to Π. Then, X_t is updated to X^b with probability $\frac{|X^b|}{|X|}$. Consequently, $D_\infty(\boldsymbol{X_t}(J_t))$ will increase as $|X_t|$ shrinks, and in expectation (over the random choice of $\boldsymbol{b}$) the increment is

$$\sum_{b\in\{0,1\}}\frac{|X^b|}{|X_t|}\log\left(\frac{|X_t|}{|X^b|}\right) \le 1. \tag{5}$$

2. Next, we further partition X_t according to Lemma 13. Say X is partitioned into $X_t = X^1 \cup \cdots \cup X^m$ and let $I_1, \ldots, I_m$ be the index sets promised by Lemma 13; and for all $j \in [m]$ we have

$$D_\infty(\boldsymbol{X}^j(J_t \setminus I_j)) \le D_\infty(\boldsymbol{X}_t(J_t)) - (1-\gamma)|I_j| + \delta_j,$$

where $\delta_j = \log(|X_t|/|\cup_{v \ge j} X^v|)$. With probability $p_j \stackrel{\text{def}}{=} |X^j|/|X_t|$, we update $X_t := X^j$ and $J_t := J_t \setminus I_j$. Therefore, taking expectation over the random choice of $\boldsymbol{j}$, the density function will decrease by

$$D_\infty(\boldsymbol{X_t}(J_t)) - \mathop{\mathbf{E}}_{j \sim \boldsymbol{j}}\left[D_\infty(\boldsymbol{X}^j_{\boldsymbol{t}}(J_t \setminus I_j))\right] \ge \mathop{\mathbf{E}}_{j \sim \boldsymbol{j}}\left[(1-\gamma)\cdot|I_j| - \delta_j\right]. \tag{6}$$

Note that $\delta_j \stackrel{\text{def}}{=} \log \frac{1}{\sum_{v \ge j} p_v}$ and thus

$$\mathop{\mathbf{E}}_{j \sim \boldsymbol{j}}\left[\delta_j\right] = \sum_{j=1}^{m} p_j \log \frac{1}{\sum_{v \ge j} p_v} \le \int_0^1 \log \frac{1}{1-x} \mathrm{d}x \le 1. \tag{7}$$

Let $\mathcal{F}_{\ell-1}$ be the σ-algebra generated by the random coins used for the first $(\ell-1)$ iterations. Let $\boldsymbol{\beta}_\ell$ be the increment of $|\overline{J_t}|$ in the ℓ-th iteration. Observe that $\boldsymbol{\beta}_\ell = |I_j|$ by definition. By Equation (6) and Equation (7), taking expectation over random choice of $\boldsymbol{j}$, $D_\infty(\boldsymbol{X}_t(J_t))$ decrease by at least $(1-\gamma)\cdot\mathbf{E}\left[\boldsymbol{\beta}_\ell \mid \mathcal{F}_{\ell-1}\right] - 1$ due to the density restoring partition. Then

$$\mathbf{E}\left[\boldsymbol{\phi}_\ell - \boldsymbol{\phi}_{\ell-1}\right] = \mathbf{E}\left[\mathbf{E}\left[\boldsymbol{\phi}_\ell - \boldsymbol{\phi}_{\ell-1} \mid \mathcal{F}_{\ell-1}\right]\right] \le \mathbf{E}\left[1 - ((1-\gamma)\cdot\boldsymbol{\beta}_\ell - 1)\right]. \tag{8}$$

In the beginning, $\boldsymbol{\phi}_0 = D_\infty(\{0,1\}^n) = 0$. Since the density function is always non-negative by definition, we have $\boldsymbol{\phi}_{\boldsymbol{c}_t} \ge 0$ and thus $\mathbf{E}\left[\boldsymbol{\phi}_{\boldsymbol{c}_t} - \boldsymbol{\phi}_0\right] \ge 0$. On the other hand, by telescoping,

$$\mathbf{E}\left[\boldsymbol{\phi}_{\boldsymbol{c}_t} - \boldsymbol{\phi}_0\right] = \mathbf{E}\left[\sum_{\ell=1}^{\boldsymbol{c}_t}(\boldsymbol{\phi}_\ell - \boldsymbol{\phi}_{\ell-1})\right] \le \mathbf{E}\left[\sum_{\ell=1}^{\boldsymbol{c}_t}(\boldsymbol{\beta}_\ell + 2)\right],$$

where the inequality follows from Equation (8). Observe that $\sum_{t=1}^{\boldsymbol{c}_t} \boldsymbol{\beta}_t = |\overline{J_t}|$ by definition. We conclude that

$$\mathbf{E}\left[|\overline{J_t}|\right] = \mathbf{E}\left[\sum_{\ell=1}^{\boldsymbol{c}_t}\boldsymbol{\beta}_\ell\right] \le \frac{2\,\mathbf{E}\left[\boldsymbol{c}_t\right]}{1-\gamma} \le \frac{2C_t}{1-\gamma},$$

as desired. ◁

4.3 Putting Things Together

Now we are prepared to prove Theorem 22.

Proof of Theorem 22. We first normalize Π so as to make it accept all inputs in $([n] \times \{0,1\}^n)^k$. Denoted by Π' the normalized protocol, then Π' has signature $(C_1+1, \ldots, C_k+1)$.

Set $\gamma = 1 - \frac{2\varepsilon}{k}$. One can check that γ satisfies the requirement in Lemma 26. By Lemma 26 and Lemma 31, we have

$$\text{Accuracy}(\Pi') \stackrel{\text{def}}{=} \mathop{\mathbf{Pr}}_{z=(\sigma_1,x_1,\ldots,\sigma_k,x_k)\sim\chi_k}\left[\Pi'(z) = x_{k-1}(\sigma_k)\right] \le \frac{1}{2} + \varepsilon + \frac{4}{n}\cdot\frac{k}{2\varepsilon}\cdot 2\sum_{t=1}^{k}(C_t+1). \tag{9}$$

Since Π', Π have the same output on valid inputs and we assumed $\text{Accuracy}(\Pi) \geq \frac{1}{2} + 2\varepsilon$, we get $\text{Accuracy}(\Pi') \geq \frac{1}{2} + 2\varepsilon$. Combining with Equation (9) and rearranging, we have

$$\sum_{t=1}^{k} C_t \geq \frac{\varepsilon^2 n}{4k} - k. \tag{10}$$

The above lower bound is vacuous when $k = \omega(\sqrt{n})$. Next, we strengthen the lower bound to $\Omega(n/k + \sqrt{n})$ via simple reductions. Consider two cases.

- Case 1: $k \leq \varepsilon\sqrt{n}/4$. Equation (10) implies that

$$\sum_{t=1}^{k} C_t \geq \frac{\varepsilon^2 n}{4k} - k \geq \frac{\varepsilon^2 n}{4k} - \frac{\varepsilon^2 n}{16k} > \frac{\varepsilon^2 n}{8k} \geq \frac{\varepsilon\sqrt{n}}{2}.$$

- Case 2: $k > \varepsilon\sqrt{n}/4$. Let $\mathcal{T} \stackrel{\text{def}}{=} \{t \in [k] : C_t \geq 1\}$ be the set of talking parties. If $|\mathcal{T}| \geq \varepsilon\sqrt{n}/8 \geq \frac{\varepsilon^2 n}{32k}$, we are done. Otherwise, we construct a protocol $\widehat{\Pi}$ for $\textsc{ChainInd}_{k'}$ where $k' \stackrel{\text{def}}{=} 2|\mathcal{T}| \leq \varepsilon\sqrt{n}/4$, described below.
  - Say the talking parties are $P_{i_1}, \ldots, P_{i_{k'/2}}$. Let P'_{2j-1} emulate P_{i_j}, and P_{2j} emulate $P_{i_j+1}, \ldots P_{i_{j+1}-1}$ by doing nothing. Note that on receiving an input sampled from $\mu_{k'}$, the parties P'_{2j} can imagine they are given inputs for all $P_{i_j+1}, \ldots P_{i_{j+1}-1}$ from μ_k. Since $P_{i_j+1}, \ldots P_{i_{j+1}-1}$ never talks, P'_{2j} perfect emulate $P_{i_j+1}, \ldots P_{i_{j+1}-1}$ them by doing nothing. In sum,

$$\Pr_{z \sim \chi_{k'}}\left[\widehat{\Pi}(z) = \textsc{ChainInd}_{k'}(z)\right] = \Pr_{z \sim \chi_k}\left[\Pi(z) = \textsc{ChainInd}_k(z)\right] \geq \frac{1}{2} + 2\varepsilon.$$

    Observe that $\widehat{\Pi}$ has signature $(C_{i_1}, 0, C_{i_2}, 0, \ldots, C_{i_{k'/2}}, 0)$. Applying Equation (10) to $\widehat{\Pi}$, we get

$$\sum_{i=1}^{k} C_i = \sum_{t=1}^{k'} C_{i_t} \geq \frac{\varepsilon^2 n}{4k'} - k' \geq \frac{\varepsilon\sqrt{n}}{2} \geq \frac{\varepsilon^2 n}{8k}.$$

Therefore, we conclude that $\sum_{i=1}^{k} C_i \geq \max\left\{\frac{1}{32}\varepsilon^2 n/k, \frac{1}{8}\varepsilon\sqrt{n}\right\}$, regardless of the value of k. ◀

References

1. Ryan Alweiss, Shachar Lovett, Kewen Wu, and Jiapeng Zhang. Improved bounds for the sunflower lemma. In *Proceedings of the 52nd Annual ACM SIGACT Symposium on Theory of Computing*, pages 624–630, 2020. doi:10.1145/3357713.3384234.
2. Alexandr Andoni, Andrew McGregor, Krzysztof Onak, and Rina Panigrahy. Better bounds for frequency moments in random-order streams. *arXiv preprint*, 2008. arXiv:0808.2222.
3. Sepehr Assadi, Yu Chen, and Sanjeev Khanna. Polynomial pass lower bounds for graph streaming algorithms. In *Proceedings of the 51st Annual ACM SIGACT Symposium on theory of computing*, pages 265–276, 2019. doi:10.1145/3313276.3316361.
4. Laszlo Babai, Peter Frankl, and Janos Simon. Complexity classes in communication complexity theory. In *27th Annual Symposium on Foundations of Computer Science (sfcs 1986)*, pages 337–347, 1986. doi:10.1109/SFCS.1986.15.
5. Ziv Bar-Yossef, Thathachar S Jayram, Ravi 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. doi:10.1016/J.JCSS.2003.11.006.

6 Boaz Barak, Mark Braverman, Xi Chen, and Anup Rao. How to compress interactive communication. In *Proceedings of the forty-second ACM symposium on Theory of computing*, pages 67–76, 2010. doi:10.1145/1806689.1806701.

7 Balthazar Bauer, Pooya Farshim, and Sogol Mazaheri. Combiners for backdoored random oracles. In *Advances in Cryptology–CRYPTO 2018: 38th Annual International Cryptology Conference, Santa Barbara, CA, USA, August 19–23, 2018, Proceedings, Part II 38*, pages 272–302. Springer, 2018. doi:10.1007/978-3-319-96881-0_10.

8 Paul Beame and Michael Whitmeyer. Multiparty communication complexity of collision finding, 2024. doi:10.48550/arXiv.2411.07400.

9 Anup Bhattacharya, Sourav Chakraborty, Arijit Ghosh, Gopinath Mishra, and Manaswi Paraashar. Disjointness through the lens of vapnik–chervonenkis dimension: Sparsity and beyond. *Comput. Complex.*, 31(2), December 2022. doi:10.1007/s00037-022-00225-6.

10 Sujoy Bhore, Fabian Klute, and Jelle J Oostveen. On streaming algorithms for geometric independent set and clique. In *International Workshop on Approximation and Online Algorithms*, pages 211–224. Springer, 2022. doi:10.1007/978-3-031-18367-6_11.

11 Mark Braverman, Ankit Garg, Denis Pankratov, and Omri Weinstein. From information to exact communication. In *Proceedings of the Forty-Fifth Annual ACM Symposium on Theory of Computing*, STOC '13, pages 151–160, New York, NY, USA, 2013. Association for Computing Machinery. doi:10.1145/2488608.2488628.

12 Mark Braverman, Sumegha Garg, Qian Li, Shuo Wang, David P Woodruff, and Jiapeng Zhang. A new information complexity measure for multi-pass streaming with applications. In *Proceedings of the 56th Annual ACM Symposium on Theory of Computing*, pages 1781–1792, 2024. doi:10.1145/3618260.3649672.

13 Mark Braverman, Sumegha Garg, and David P Woodruff. The coin problem with applications to data streams. In *2020 ieee 61st annual symposium on foundations of computer science (focs)*, pages 318–329. IEEE, 2020. doi:10.1109/FOCS46700.2020.00038.

14 Joshua Brody, Amit Chakrabarti, Ranganath Kondapally, David P. Woodruff, and Grigory Yaroslavtsev. Beyond set disjointness: The communication complexity of finding the intersection. In *Proceedings of the 2014 ACM Symposium on Principles of Distributed Computing*, PODC '14, pages 106–113, New York, NY, USA, 2014. Association for Computing Machinery. doi:10.1145/2611462.2611501.

15 Amit Chakrabarti. Lower bounds for multi-player pointer jumping. In *Twenty-Second Annual IEEE Conference on Computational Complexity (CCC'07)*, pages 33–45. IEEE, 2007. doi:10.1109/CCC.2007.14.

16 Amit Chakrabarti, Graham Cormode, and Andrew McGregor. Robust lower bounds for communication and stream computation. In *Proceedings of the fortieth annual ACM symposium on Theory of computing*, pages 641–650, 2008. doi:10.1145/1374376.1374470.

17 Amit Chakrabarti, Yaoyun Shi, Anthony Wirth, and Andrew Yao. Informational complexity and the direct sum problem for simultaneous message complexity. In *Proceedings 42nd IEEE Symposium on Foundations of Computer Science*, pages 270–278. IEEE, 2001.

18 Amit Chakrabarti and Anthony Wirth. Incidence geometries and the pass complexity of semi-streaming set cover. In *Proceedings of the twenty-seventh annual ACM-SIAM symposium on Discrete algorithms*, pages 1365–1373. SIAM, 2016. doi:10.1137/1.9781611974331.CH94.

19 Arkadev Chattopadhyay, Yuval Filmus, Sajin Koroth, Or Meir, and Toniann Pitassi. Query-To-Communication Lifting for BPP Using Inner Product. In Christel Baier, Ioannis Chatzigiannakis, Paola Flocchini, and Stefano Leonardi, editors, *46th International Colloquium on Automata, Languages, and Programming (ICALP 2019)*, volume 132 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 35:1–35:15, Dagstuhl, Germany, 2019. Schloss Dagstuhl – Leibniz-Zentrum für Informatik. doi:10.4230/LIPIcs.ICALP.2019.35.

20 Sandro Coretti, Yevgeniy Dodis, Siyao Guo, and John Steinberger. Random oracles and non-uniformity. In *Annual International Conference on the Theory and Applications of Cryptographic Techniques*, pages 227–258. Springer, 2018. doi:10.1007/978-3-319-78381-9_9.

21 Graham Cormode, Jacques Dark, and Christian Konrad. Independent sets in vertex-arrival streams. *arXiv preprint*, 2018. arXiv:1807.08331.

22 Jacques Dark, Adithya Diddapur, and Christian Konrad. Interval selection in data streams: Weighted intervals and the insertion-deletion setting. In *43rd IARCS Annual Conference on Foundations of Software Technology and Theoretical Computer Science (FSTTCS 2023)*, pages 24:1–24:17. Schloss-Dagstuhl-Leibniz Zentrum für Informatik, 2023. doi:10.4230/LIPIcs.FSTTCS.2023.24.

23 Nachum Dershowitz, Rotem Oshman, and Tal Roth. The communication complexity of multiparty set disjointness under product distributions. In *Proceedings of the 53rd Annual ACM SIGACT Symposium on Theory of Computing*, pages 1194–1207, 2021. doi:10.1145/3406325.3451106.

24 Moran Feldman, Ashkan Norouzi-Fard, Ola Svensson, and Rico Zenklusen. The one-way communication complexity of submodular maximization with applications to streaming and robustness. In *Proceedings of the 52nd Annual ACM SIGACT Symposium on Theory of Computing*, pages 1363–1374, 2020. doi:10.1145/3357713.3384286.

25 Moran Feldman, Ashkan Norouzi-Fard, Ola Svensson, and Rico Zenklusen. Submodular maximization subject to matroid intersection on the fly. In *30th Annual European Symposium on Algorithms (ESA 2022)*, pages 52:1–52:14. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2022. doi:10.4230/LIPIcs.ESA.2022.52.

26 Dmitry Gavinsky. The communication complexity of the inevitable intersection problem. *Chicago Journal of Theoretical Computer Science*, 2020(3), August 2020. URL: http://cjtcs.cs.uchicago.edu/articles/2020/3/contents.html.

27 Satrajit Ghosh and Mark Simkin. The communication complexity of threshold private set intersection. In Alexandra Boldyreva and Daniele Micciancio, editors, *Advances in Cryptology – CRYPTO 2019*, pages 3–29, Cham, 2019. Springer International Publishing. doi:10.1007/978-3-030-26951-7_1.

28 Mika Göös, Tom Gur, Siddhartha Jain, and Jiawei Li. Quantum communication advantage in tfnp, 2024. doi:10.48550/arXiv.2411.03296.

29 Mika Göös, Shachar Lovett, Raghu Meka, Thomas Watson, and David Zuckerman. Rectangles are nonnegative juntas. *SIAM Journal on Computing*, 45(5):1835–1869, 2016. doi:10.1137/15M103145X.

30 Mika Göös, Toniann Pitassi, and Thomas Watson. Query-to-communication lifting for bpp. In *2017 IEEE 58th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 132–143, 2017. doi:10.1109/FOCS.2017.21.

31 Venkatesan Guruswami and Krzysztof Onak. Superlinear lower bounds for multipass graph processing. *Algorithmica*, 76:654–683, 2016. doi:10.1007/S00453-016-0138-7.

32 Dawei Huang, Seth Pettie, Yixiang Zhang, and Zhijun Zhang. The communication complexity of set intersection and multiple equality testing. *SIAM Journal on Computing*, 50(2):674–717, 2021. doi:10.1137/20M1326040.

33 Mi-Ying(Miryam) Huang, Xinyu Mao, Guangxu Yang, and Jiapeng Zhang. Breaking square-root loss barriers via min-entropy. In *In Electronic Colloquium on Computational Complexity (ECCC) (TR24-067)*, 2024. URL: https://eccc.weizmann.ac.il/report/2024/067/.

34 Bala Kalyanasundaram and Georg Schintger. The probabilistic communication complexity of set intersection. *SIAM Journal on Discrete Mathematics*, 5(4):545–557, 1992. doi:10.1137/0405044.

35 Shachar Lovett, Raghu Meka, Ian Mertz, Toniann Pitassi, and Jiapeng Zhang. Lifting with sunflowers. *Leibniz international proceedings in informatics*, 215, 2022. doi:10.4230/LIPICS.ITCS.2022.104.

36 Shachar Lovett, Noam Solomon, and Jiapeng Zhang. From dnf compression to sunflower theorems via regularity. In *Proceedings of the 34th Computational Complexity Conference*, pages 1–14, 2019. doi:10.4230/LIPICS.CCC.2019.5.

37 Shachar Lovett and Jiapeng Zhang. Streaming lower bounds and asymmetric set-disjointness. In *2023 IEEE 64th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 871–882. IEEE, 2023. doi:10.1109/FOCS57990.2023.00056.

38 Xinyu Mao, Guangxu Yang, and Jiapeng Zhang. Gadgetless lifting beats round elimination: Improved lower bounds for pointer chasing. *arXiv preprint*, 2024. doi:10.48550/arXiv.2411.10996.

39 Noam Nisan and Avi Widgerson. Rounds in communication complexity revisited. In *Proceedings of the twenty-third annual ACM symposium on Theory of computing*, pages 419–429, 1991. doi:10.1145/103418.103463.

40 Rotem Oshman and Tal Roth. The Communication Complexity of Set Intersection Under Product Distributions. In Kousha Etessami, Uriel Feige, and Gabriele Puppis, editors, *50th International Colloquium on Automata, Languages, and Programming (ICALP 2023)*, volume 261 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 95:1–95:20, Dagstuhl, Germany, 2023. Schloss Dagstuhl – Leibniz-Zentrum für Informatik. doi:10.4230/LIPIcs.ICALP.2023.95.

41 Jeff M. Phillips, Elad Verbin, and Qin Zhang. Lower bounds for number-in-hand multiparty communication complexity, made easy. *SIAM Journal on Computing*, 45(1):174–196, 2016. doi:10.1137/15M1007525.

42 M. Saglam and G. Tardos. On the communication complexity of sparse set disjointness and exists-equal problems. In *2013 IEEE 54th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 678–687, Los Alamitos, CA, USA, October 2013. IEEE Computer Society. doi:10.1109/FOCS.2013.78.

43 Janani Sundaresan. Optimal communication complexity of chained index. *ITCS*, 2025.

44 Shuo Wang, Guangxu Yang, and Jiapeng Zhang. Communication complexity of set-intersection problems and its applications. In *In Electronic Colloquium on Computational Complexity (ECCC) (TR23-164)*, 2023.

45 Thomas Watson. Communication Complexity with Small Advantage. In Rocco A. Servedio, editor, *33rd Computational Complexity Conference (CCC 2018)*, volume 102 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 9:1–9:17, Dagstuhl, Germany, 2018. Schloss Dagstuhl – Leibniz-Zentrum für Informatik. doi:10.4230/LIPIcs.CCC.2018.9.

46 Guangxu Yang and Jiapeng Zhang. Communication lower bounds for collision problems via density increment arguments. In *Proceedings of the 56th Annual ACM Symposium on Theory of Computing*, pages 630–639, 2024. doi:10.1145/3618260.3649607.

A Proof of Claim 30

▷ Claim 33 (Claim 30 restated). Let $t \le k$. Let $\sigma_1, \dots, \sigma_{t+1} \in [n], X_1, \dots, X_t \subseteq \{0,1\}^n$. For $\ell \in \{t, t+1\}$, define

$$R_\ell \stackrel{\text{def}}{=} (\{\sigma_1\} \times X_1) \times \cdots \times (\{\sigma_{\ell-1}\} \times X_{\ell-1}) \times (\{\sigma_\ell\} \times \{0,1\}^n) \times ([n] \times \{0,1\}^n)^{k-\ell}.$$

Then

$$\frac{\chi_k(R_{t+1})}{\chi_k(R_t)} \le 2\frac{\mathcal{U}_k(R_{t+1})}{\mathcal{U}_k(R_t)}.$$

Proof of Claim 30. To start with, observe that

$$\frac{\mathcal{U}_k(R_{t+1})}{\mathcal{U}_k(R_t)} = \frac{|R_{t+1}|}{|R_t|} = \frac{|X_t|}{n2^n}. \tag{11}$$

We claim that for $\ell \in \{t, t+1\}$,

$$\chi_k(R_\ell) = \frac{\#\{(x_1, \dots, x_{\ell-1}) \in X_1 \times \cdots \times X_{\ell-1} : x_1(\sigma_2) = \cdots = x_{\ell-1}(\sigma_\ell)\}}{n^\ell \cdot 2^{(n-1)\ell+1}}. \tag{12}$$

Then we have

$$\begin{aligned}\chi_k(R_{t+1}) &= \frac{\#\{(x_1,\dots,x_t)\in X_1\times\cdots\times X_t : x_1(\sigma_2)=\cdots=x_t(\sigma_{t+1})\}}{n^{t+1}\cdot 2^{(n-1)(t+1)+1}}\\ &\le \frac{\#\{(x_1,\dots,x_{t-1})\in X_1\times\cdots\times X_{t-1} : x_1(\sigma_2)=\cdots=x_{t-1}(\sigma_t)\}\cdot|X_t|}{n^{t+1}\cdot 2^{(n-1)(t+1)+1}}\\ &= \chi_k(R_t)\cdot\frac{|X_t|}{n2^{n-1}}.\end{aligned}$$

where the first and the third equality is from Equation (12). Combining with Equation (11) we have the desired result.

It remains to show Equation (12). Suppose that $((\rho_1,x_1),\dots,(\rho_k,x_k))\in R_\ell$ satisfies $x_1(\rho_2)=\cdots=x_{k-1}(\rho_k)$. Then we $\rho_1=\sigma_1,\dots,\rho_\ell=\sigma_\ell$ and

$$x_1(\sigma_2)=\cdots=x_{\ell-1}(\sigma_\ell)=b \text{ for some } b\in\{0,1\}.$$

For every $\rho_{\ell+1},\dots,\rho_k\in[n]$, there exists exactly $2^{(n-1)}$ possible values for each x_j with $\ell\le j\le k-1$ (with one bit fixed to be b) and 2^n possible values for x_k (which is not used at all). Therefore,

$$\begin{aligned}\chi_k(R_\ell) &= \frac{\#\{((\rho_1,x_1),\dots,(\rho_k,x_k))\in R_{t+1} : x_1(\rho_2)=\cdots=x_{k-1}(\rho_k)\}}{n^k\cdot 2^{(n-1)(k-1)+n+1}}\\ &= \frac{n^{k-\ell}\cdot 2^{(n-1)\cdot(k-1-\ell)+n}\cdot\#\{(x_1,\dots,x_{\ell-1})\in X_1\times\cdots\times X_{\ell-1} : x_1(\sigma_2)=\cdots=x_{\ell-1}(\sigma_\ell)\}}{n^k\cdot 2^{(n-1)(k-1)+n+1}}\\ &= \frac{\#\{(x_1,\dots,x_{\ell-1})\in X_1\times\cdots\times X_{\ell-1} : x_1(\sigma_2)=\cdots=x_{\ell-1}(\sigma_\ell)\}}{n^\ell\cdot 2^{(n-1)\ell+1}},\end{aligned}$$

which is exactly what we wanted. ◁

B Proof of Lemma 13

The following lemma and proof are from Lemma 5 in [30].

▶ **Lemma 34** (Lemma 13 restated)**.** *Let $\gamma\in(0,1)$. Let X be a subset of $[n]^M$ and $J\subseteq[M]$. Suppose that there exists an $\beta\in[n]^{\overline{J}}$ such that $\forall x\in X, x(\overline{J})=\beta$. Then, there exists a partition $X=X^1\cup X^2\cup\cdots\cup X^r$ and every X^i is associated with a set $I_i\subseteq J$ and a value $\alpha_i\in\{0,1\}^{I_i}$ that satisfy the following properties.*

1. *$\forall x\in X^i, x(I_i)=\alpha_i$;*
2. *$\boldsymbol{X}^i(J\setminus I_i)$ is γ-dense;*
3. *$D_\infty\left(\boldsymbol{X}^i(J\setminus I_i)\right)\le D_\infty\left(\boldsymbol{X}(J)\right)-(1-\gamma)\log n\cdot|I_i|+\delta_i$, where $\delta_i\stackrel{\text{def}}{=}\log(|X|/|\cup_{j\ge i}X^j|)$.*

Proof. We prove it by a greedy algorithm as follows.

Item 1 is guaranteed by the construction of X^i and I_i.

We prove Item 2 by contradiction. Assume towards contradiction that $\boldsymbol{X}^i(J\setminus I_i)$ is not γ-dense for some i. By definition, there is a nonempty set $K\subseteq J\setminus I_i$ and $\beta\in[n]^K$ violating the min-entropy condition, namely, $\mathbf{Pr}\left[\boldsymbol{X}(K)=\beta\right]>n^{-\gamma|K|}$. Write $X^{\ge i}\stackrel{\text{def}}{=}\cup_{j\ge i}X^i$. Then

$$\mathbf{Pr}\left[\boldsymbol{X}^{\ge i}(I_i\cup K)=(\alpha_i,\beta)\right]=\mathbf{Pr}\left[\boldsymbol{X}^{\ge i}(I_i)=\alpha_i\right]\cdot\mathbf{Pr}\left[\boldsymbol{X}^i(K)=\beta\right]>n^{-\gamma|I_i|}\cdot n^{-\gamma|K|}=n^{-\gamma|I_i\cup K|},$$

where the first equality holds as $(\boldsymbol{X}^{\ge i}|\boldsymbol{X}^{\ge i}(I_i)=\alpha_i)=\boldsymbol{X}^i$. However, this means at moment that I_i is chosen, the set $I_i\cup K\subseteq J$ also violates the min-entropy condition (witnessed by (α_i,β)), contradicting the maximality of I_i.

Algorithm 3 Greedy Algorithm.

Input: $X \subseteq [n]^M$
Output: A partition $X = X^1 \cup X^2 \cup \cdots \cup X^m$

Initialize $i := 1$.
while $X \neq \emptyset$ **do**
  Let $I \subseteq J$ be a maximal subset (possibly $I = \emptyset$) such that $\mathrm{H}_\infty(\boldsymbol{X}(I)) < \gamma|I| \log n$ and let $\alpha_i \in [n]^I$ be a witness of this fact, i.e., $\mathbf{Pr}[\boldsymbol{X}(I) = \alpha_i] > n^{-\gamma|I|}$.
  $X^i := \{x \in X : x(I) = \alpha_i\}$ and $I_i := I$.
  Update $X := X \setminus X^i$, $J := J \setminus I_i$, and $i := i + 1$.

Finally, Item 3 is proved by straightforward calculation:

$$\begin{aligned} D_\infty(\boldsymbol{X}^i(J \setminus I_i)) &= |J \setminus I_i| \log n - \log |X^i| \\ &\le (|J| \log n - |I_i| \log n) - \log \left(\left|X^{\ge i}\right| \cdot n^{-\gamma|I_i|} \right) \\ &= (|J| \log n - \log |X|) - (1-\gamma)|I_i| \cdot \log n + \log \left(\frac{|X|}{|X^{\ge i}|} \right) \\ &= D_\infty\Big(\boldsymbol{X}(J)\Big) - (1-\gamma)|I_i| \log n + \delta_i. \end{aligned}$$

◀