Electronic Colloquium on Computational Complexity, Report No. 70 (2024)

Gadgetless Lifting Beats Round Elimination: Improved Lower Bounds for Pointer Chasing

Xinyu Mao * Guangxu Yang * Jiapeng Zhang *

April 10, 2024

Abstract

The notion of *query-to-communication lifting theorems* is a generic framework to convert query lower bounds into two-party communication lower bounds. Though this framework is very generic and beautiful, it has some inherent limitations such as it only applies to lifted functions. In order to address this issue, we propose *gadgetless lifting*, a new framework to prove communication lower bounds. The crux of the new approach is to lift lower bounds for *a family of restricted protocols* into lower bounds for *general protocols*.

To demonstrate the power of gadgetless lifting, we prove an $\Omega(n/k+k)$ communication lower bound on $(k-1)$-*round distributional complexity* of *the k-step pointer chasing problem* under the *uniform input distribution*, improving the $\Omega(n/k - k\log n)$ lower bound due to Yehudayoff (Combinatorics Probability and Computing, 2020). Our lower bound almost matches the upper bound of $\widetilde{O}(n/k+k)$ communication by Nisan and Wigderson (STOC 91).

A key step in gadgetless lifting is how to choose the definition of restricted protocols. We start with proving communication lower bounds for restricted protocols, and then lift it into general settings. In this paper, our definition of restricted protocols is inspired by the structure-vs-pseudorandomness decomposition by Göös, Pitassi, and Watson (FOCS 17) and Yang and Zhang (STOC 24).

Previously, round-communication trade-offs were mainly obtained by round elimination and information complexity. However, both methods have some obstacles in some settings. In general, we believe gadgetless lifting provides a new solution to address previous barriers by round elimination and information complexity.

1 Introduction

Pointer chasing is a well-known problem [RY20] that demonstrates the power of interaction in communication and has broad applications in different areas. It was used for proving monotone constant-depth hierarchy theorem [NW91, KPPY84], lower bounds on the time complexity of distributed computation [NDSP11], lower bounds on the space complexity of streaming algorithms [FKM+09, GO16, ACK19], adaptivity hierarchy theorem for property testing [CG18], exponential

*Research supported by NSF CAREER award 2141536.
Thomas Lord Department of Computer Science, University of Southern California.
Email: {xinyumao, guangxuy, jiapengz}@usc.edu

ISSN 1433-8092

separations in local differential privacy [JMR20], memory bounds for continual learning [CPP22] and limitations of the transformer architecture [PNP24]. It is a two-party function defined below.

Definition 1.1 (k-step pointer chasing function)**.** For $k \geq 1$, the k-step pointer chasing function $\mathrm{PC}_k : [n]^n \times [n]^n \to \{0,1\}$ is defined as follows. Given input $f_A, f_B \in [n]^n$, for $r = 0, 1, \ldots, k$ we recursively define pointers via

$$\mathsf{pt}_r(f_A, f_B) \stackrel{\text{def}}{=} \begin{cases} 1 & \text{if } r = 0; \\ f_A(\mathsf{pt}_{r-1}(f_A, f_B)) & \text{if } r > 0 \text{ is odd;} \\ f_B(\mathsf{pt}_{r-1}(f_A, f_B)) & \text{if } r > 0 \text{ is even.} \end{cases}$$

The output of PC_k is the parity of the last pointer, namely, $\mathrm{PC}_k(f_A, f_B) \stackrel{\text{def}}{=} \mathsf{pt}_k(f_A, f_B) \bmod 2$.

Upper bounds. If Alice and Bob could communicate for k rounds, a simple protocol could be the following: Alice and Bob alternatively send $f_A(\mathsf{pt}_{r-1}(f_A, f_B))$ or $f_A(\mathsf{pt}_{r-1}(f_A, f_B))$. The total communication cost for this simple protocol is $O(k \cdot \log n)$. However, if Alice and Bob can only communicate $(k-1)$ rounds, the upper bound then becomes non-trivial. In this direction, Nisan and Wigderson [NW91] proposed a randomized $(k-1)$-round protocol with $O((n/k + k)\log n)$ communication bits.

- In the beginning, Alice and Bob use public randomness to pick a set of coordinates $I \subseteq [n]$ of size $10n/k$, and then send $f_A(I)$ and $f_B(I)$ to the other party.
- On the other hand, Alice and Bob also simulate (r rounds) deterministic protocol but skip one round if one party finds that the pointer is located in I.
- If the skip round never happens, Alice and Bob simply abort at the last round. A simple calculation shows the probability of this event is low.

This randomized protocol is indeed very simple. Alice and Bob only share coordinate-wise information. In fact, this is a restricted rectangle in our setting.

Lower Bounds. Consider $(k-1)$ round protocols where Alice speaks first. For deterministic protocols, Nisan and Wigderson [NW91] proved an $\Omega(n - k\log n)$ communication lower bound. In the same paper, they also proved an $\Omega(n/k^2 - k\log n)$ communication lower bound for protocols that achieve $2/3$ accuracy *under uniform input distribution*.

Since then, lower bounds for pointer chasing and its close variants have been substantially studied by a good amount of papers [DGS84, DJS98, PRV01, Kla00, KNTSZ07, FKM+09, GO16, ACK19]. Finally, Yehudayoff [Yeh20] proved an $\Omega(n/k - k\log n)$ lower bound for protocols achieving constant advantage under uniform input distribution.

Now the main gap between the upper bound [NW91] and the lower bound [Yeh20] is the extra $k\log n$ term. This gap becomes significant if $k \geq \sqrt{n}$. In this paper, we further improve the lower bound and close the gap.

1.1 Our results

We prove that any protocol that achieves constant advantage under uniform input distribution must communicate $\Omega(n/k + k)$ bits.

Theorem 1.2. *Let Π be a $(k-1)$-round protocol for PC_k where Alice speaks first such that*

$$\Pr_{f_A, f_B \leftarrow [n]^n} [\Pi(f_A, f_B) = PC_k(f_A, f_B)] \geq 2/3.$$

Then the communication complexity of Π is $\Omega(n/k + k)$.

By Yao's minimax principle, it implies a lower bound for the $(k-1)$ round randomized communication complexity.

Corollary 1.3. *Every $(k-1)$-round randomized protocol for PC_k with error at most $1/3$ (where Alice speaks first) has communication complexity $\Omega(n/k + k)$.*

We observe there is still a $(\log n)$ gap between our lower bound and the protocol by [NW91]. We conjecture that our lower bound is tight and there is a chance to remove the $\log n$ factor in the upper bound side. A simple deterministic protocol with $(k-1)$ rounds and $O(n)$ communication bits could be the following: Alice and Bob send the parity of $f_A(x)$ and $f_B(x)$ for all $x \in [n]$ in the beginning. Hence they can skip the last round as they already know the parity. This simple protocol shows that [NW91]'s protocol is not tight when $k = o(\log n)$. We believe similar ideas could be extended for large k.

Applications. Given the connections between PC_k and diverse applications [FKM+09, NDSP11, CG18, JMR20, CPP22, PNP24], our improved lower bounds automatically lead to several applications. We list two applications below.

Corollary 1.4 (Direct sum extension of pointer chasing)**.** *The $(k-1)$-round randomized communication complexity of PC_k with d pairs of functions is $\Omega(d \cdot n/k^2 + d)$*

This corollary improves the previous $\Omega(d \cdot n/k^3 - dk \log n - 2d)$ lower bound presented in [FKM+09], which has applications in BFS trees streaming lower bound.

Corollary 1.5 (Exponential separations in local differential privacy)**.** *Let A be a $(k-1)$-round sequentially interactive ε-locally private protocol solving PC_k with error probability $\gamma \leq 1/3$. Then the sample complexity of A is $\Omega\left(\frac{1}{e^\varepsilon} \cdot (n/k + k)\right)$ and there is a k round protocol with sample complexity $\widetilde{O}\left(\frac{k \log n}{\varepsilon^2}\right)$.*

This corollary improves the previous $\Omega\left(\frac{n}{e^\varepsilon k^2}\right)$ lower bound for $k < \sqrt{n/\log n}$ given by [JMR20].

1.2 Gadgetless Lifting: A New Framework to Prove Communication Lower Bounds

Query-to-communication lifting theorem [RM97, GPW15, GPW17, CFK+19, LMM+22] is a remarkable framework for proving communication complexity lower bounds with a wide variety of applications. Let $f : Z^n \to \{0, 1\}$ be a boolean function and let $g : X \times Y \to Z$ be a two-party gadget function. The framework proceeds in two steps:

1. Given any communication protocol for $f \circ g^n$, we construct a decision tree for f by simulating the communication protocol.

2. Prove a lower bound for the decision tree complexity of f.

Combined with these two steps, a communication lower bound for $f \circ g^n$ is then generically achieved. Though this framework is beautiful, however, it requires a gadget as f is a one-party function. As a consequence, this framework only applies to lifted functions.

Many important problems, such as pointer chasing, do not have a lifted form. Therefore, lifting theorems do not apply to these questions. To address this limitation, we propose a new framework called *gadgetless lifting*. Unlike the query-to-communication lifting theorems that lift *decision tree lower bounds* into *communication lower bounds*, the gadgetless lifting lifts lower bounds for *restricted protocols* into lower bounds for *general protocols.* This new framework proceeds in two steps.

1. Identify a family of restricted protocols.

2. Simulate general protocols by restricted protocols and prove communication lower bounds for restricted protocols.

As the basic components in gadgetless lifting are the restricted protocols, the gadgets are no longer needed. A crux of gadgetless lifting is how to decide the restricted protocols. In this paper, we capture it as those protocols that "all shared useful information are local information". For example, the protocol by [NW91] only share local information such as $f_A(x)$ or $f_B(x)$ for some $x \in [n]$. We formalize this idea by lemma 3.1, which is inspired by the structure-vs-pseudorandomness decomposition by Göös, Pitassi, and Watson [GPW17] and Yang and Zhang [YZ24].

In the study of lifted functions, it has been shown that query-to-communication lifting theorems bypassed some fundamental barriers from previous methods. Similarly, gadgetless lifting can also bypass obstacles from existing methods. We discuss two of them below.

Avoiding the loss in round elimination method. Previously, the only method to prove round-communication trade-offs is the *round elimination method* [NW91]. In these two papers, [NW91] and [Yeh20], the authors studied the pointer chasing problem via the round elimination method. Denote by $\boldsymbol{M}_1, \ldots, \boldsymbol{M}_t$ the messages sent in the first t rounds, and let $\boldsymbol{Z}_0, \boldsymbol{Z}_1, \ldots, \boldsymbol{Z}_k$ be the pointers. In round elimination-based proofs, [NW91, Yeh20] analyzed the random variable

$$\boldsymbol{R}_t = (\boldsymbol{M}_1, \ldots, \boldsymbol{M}_t, \boldsymbol{Z}_1, \ldots, \boldsymbol{Z}_{t-1}) \text{ for } t \le k$$

Particularly, they proved that $\mathbf{H}(\boldsymbol{R}_k) \ge \Omega(n/k)$. Together with the fact that $\mathbf{H}(\boldsymbol{Z}_1, \ldots, \boldsymbol{Z}_k) = k \log n$, it implies that $\mathbf{H}(\boldsymbol{M}) \ge \Omega(n/k - k \log n)$. The $(k \log n)$ loss is associated with many round elimination-based analysis [NW91, Kla00, KNTSZ07, FKM+09, Yeh20]. In this paper, we avoid the $k \log n$ loss via the gadgetless lifting.

Breaking square-root loss barrier in information complexity. Another popular method in proving communication lower bounds is the information complexity. However, as mentioned by Yahudayoff [Yeh20], entropy-based analyses are likely to meet a square-root loss barrier. This barrier usually comes from applying Pinsker's inequality (or its variant) to bound statistical distance from a small entropy gap. As a consequence, many results such as [NW91] can only prove an $\Omega(n/k^2 - k \log n)$ lower bound.

As mentioned in [Yeh20], the square-root loss also appears in many works when using the entropy-based method to prove lower bounds. For example, it appears in the parallel repetition theorem and is related to the 'strong parallel repetition' conjecture which is motivated by Khot's unique games conjecture [Kho02]. This loss also appears in direct-sum theorems [BBCR10] and direct-product theorems [BRWY13] in communication complexity.

[Yeh20] overcomes this square-root loss barrier via a non-standard measurement called triangular discrimination. By contrast, our approach overcomes the barrier more naturally without using entropy.

Potentail applications. We noticed that our method can also be naturally extended to multiparty settings such as the numbers in hand model and numbers on forehead model. Moreover, some important open problems, such as round-communication tradeoff of bipartite matching problem [DNO14, ANRW15, BVDBE+22] and set pointer chasing problem [FKM+09, GO16], are difficult to solve using the round elimination method due to its inherent limitations. Our method offers the potential to solve these challenging problems.

2 Preliminaries

Notations. We use capital letters X to denote a set and use bold symbols like $\boldsymbol{R}$ to denote random variables. Particularly, for a set X, we use $\boldsymbol{X}$ to denote the random variable uniformly distributed over the set X. We use $\leftarrow$ to denote sampling from a distribution or choosing an element from a set uniformly at random.

2.1 Density-Restoring Partition

Min-entropy and dense distribution. For a random variable $\boldsymbol{X}$, we use $\mathrm{supp}(\boldsymbol{X})$ to denote the support of $\boldsymbol{X}$.

Definition 2.1 (Min-entropy and deficiency)**.** The min-entropy of a random variable $\boldsymbol{X}$ is defined by

$$\mathbf{H}_\infty(\boldsymbol{X}) := \min_{x \in \mathrm{supp}(\boldsymbol{X})} \log\left(\frac{1}{\mathbf{Pr}[\boldsymbol{X} = x]}\right).$$

Suppose that $\boldsymbol{X}$ is supported on $[n]^J$. We define the *deficiency* of $\boldsymbol{X}$ as

$$\mathbf{D}_\infty(\boldsymbol{X}) := |J| \log n - \mathbf{H}_\infty(\boldsymbol{X}).$$

For $I \subseteq J$, $x \in [n]^J$, let $x(I) \stackrel{\mathsf{def}}{=} (x(i))_{i \in I} \in [n]^I$ be the projection of x on coordinates in I.

Definition 2.2 (Dense distribution)**.** Let $\gamma \in (0, 1)$. A random variable $\boldsymbol{X}$ supported on $[n]^J$ is said to be *γ-dense* if for all nonempty $I \subseteq J$, $\mathbf{H}_\infty(\boldsymbol{x}(I)) \geq \gamma |I| \log n$.

The following lemma is the crux of the structure-vs-pseudorandomness method by [GPW17]. It essentially says that a flat random variable could be decomposed into a convex combination of flat random variables with disjoint support and dense properties.

Lemma 2.3 (Density-restoring partition)**.** *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 $\alpha_i \in [n]^{I_i}$ 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. $\mathbf{D}_\infty\left(\boldsymbol{X}^i(J \setminus I_i)\right) \le \mathbf{D}_\infty\left(\boldsymbol{X}(J)\right) - (1-\gamma)\log n \cdot |I_i| + \delta_i$, where $\delta_i \stackrel{\text{def}}{=} \log(|X|/|\cup_{j\ge i} X^j|)$.

The proof of this lemma, simple and elegant, is included in the appendix for completeness.

2.2 Communication Protocols

We recall basic definitions and facts about communication protocols.

Protocol Tree. Let X and Y be the input space of Alice and Bob respectively. A deterministic 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 either Alice or Bob — we denote the owner by $\mathsf{owner}(v)$;
- 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.

Fact 2.4. *The set of all inputs that leads to an internal vertex v is a rectangle, denoted by $\Pi_v = X_v \times Y_v \subseteq X \times Y$.*

The *communication complexity* of Π, denoted by $\mathsf{CC}(\Pi)$, is the depth of the tree. The *round complexity* of Π, is the minimum number k such that in every path from the root to some leaf, the owner switches at most $(k-1)$ times. Clearly, if a protocol has k round, then its communication complexity is at least k. We can safely make the following assumptions for any protocol Π:

- Π has k round on every input; and
- Π communicates $\mathsf{CC}(\Pi)$ bits on every input.

Indeed, for any protocol, we can add empty messages and rounds in the end, which boosts the communication complexity by a factor of 2.

3 Proof of Main Theorem

We use a *decomposition and sampling process* DS, as shown in Algorithm 1, 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 three steps:

1. First, section 3.1 analyzes crucial invariants during the running of DS.
2. Next, section 3.2 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.
3. Finally, section 3.3 proves that the average fixed size can be bounded from above by $O(\mathsf{CC}(\Pi))$. Consequently, if Π enjoys high accuracy, we get a lower bound of $\mathsf{CC}(\Pi)$.

3.1 The Decomposition and Sampling Process

During the sampling process, we maintain a useful structure of R mainly by a partitioning-then-sampling mechanism: At the beginning, R is set to be the set of all inputs. Walking down the protocol tree, we decompose the rectangle into structured sub-rectangles; then we sample a decomposed rectangle with respect to its size. In the end, we arrive at a leaf node v and a subrectangle of Π_v.

Lemma 3.1 (Loop invariant)**.** *Set $\gamma \stackrel{\text{def}}{=} 1 - \frac{0.1}{\log n}$. Then in the running of $\mathsf{DS}(\Pi)$, we have the following loop invariants: After each iteration,*

- *($\diamond$) $X \times Y \subseteq \Pi_v$;*
- *($\clubsuit$) $\boldsymbol{X}(J_A), \boldsymbol{Y}(J_B)$ are γ-dense;*
- *($\heartsuit$) there exists some $\alpha_A \in [n]^{\overline{J_A}}, \alpha_B \in [n]^{\overline{J_B}}$ such that $x(\overline{J_A}) = \alpha_A, y(\overline{J_B}) = \alpha_B$ for all $x \in X, y \in Y$;*
- *($\spadesuit$) there exists some $z_r \in [n]$ such that $\mathsf{pt}_r(f_A, f_B) = z_r$ for all $(f_A, f_B) \in X \times Y$.*

Proof. Item ($\diamond$) is true because every time v is updated, $X \times Y$ is updated accordingly to a sub-rectangle of Π_v and updating $X \times Y$ into its sub-rectangles does not violate this condition.

Since we applied density restoring partition at the end of each iteration, Item ($\clubsuit$) and ($\heartsuit$) is guaranteed by lemma 2.3 and the way that X, Y, J_A, J_B are updated.

We prove the last item ($\spadesuit$) by induction. Assume that the statement holds after the first $(t-1)$ iterations. WLOG, assume that at the beginning of the t-th iteration, v is owned by Alice. Consider the following two cases.

- <u>Case 1. Not a new round:</u> Line 13 is not executed in the t-th iteration. Since r remains unchanged and we only update R to be a sub-rectangle of itself, the statement still holds.
- <u>Case 2. A new round begins:</u> Line 13 is executed and r is increased by 1. Let ρ denote the value of r before Line 13, then after this iteration, we have $r = \rho + 1$. The induction hypothesis guarantees that there exists some $z_{\rho-1} \in [n]$ such that

  $$\mathsf{pt}_{\rho-1}(f_A, f_B) = z_{\rho-1} \text{ for all} (f_A, f_B) \in X \times Y.$$

  Due to the partition and the update in Line 11 and Line 12, $|\mathrm{supp}(\boldsymbol{X}(z_{\rho-1}))| \leq n/2$. Hence, $\boldsymbol{X}(z_{\rho-1})$ cannot be γ-dense as we set $\gamma = 1 - \frac{0.1}{\log n}$. Observe that after the update in Line 17, $\boldsymbol{X}(J_A)$ is γ-dense. Consequently, we must have $z_{\rho-1} \in \overline{J_A}$, and by item ($\heartsuit$), there exists some $z_\rho \in [n]$ such that $f_A(z_{\rho-1}) = z_\rho \ \forall f_A \in X$. By definition, for all $(f_A, f_B) \in X \times Y$,

  $$\mathsf{pt}_\rho(f_A, f_B) = f_A(\mathsf{pt}_{\rho-1}(f_A, f_B)) = f_A(z_{\rho-1}) = z_\rho.$$

  This is exactly the same statement after the t-th iteration (as we have $r = \rho + 1$).

$\square$

The restricted rectangles in this loop invariant are inspired by the protocols of Nisan and Wigderson [NW91]. This lemma aims to capture the fact that Alice and Bob cannot get any additional useful information other than coordinate-wise information during their communication.

Algorithm 1: The decomposition and sampling process DS

Input: A protocol Π for the problem PC_k.

Output: A rectangle $R = X \times Y$, and $J_A, J_B \subseteq [n]$.

Initialize $v :=$ root of $\Pi, r := 1, X := Y := [n]^n, J_A := J_B := [n]$, bad := FALSE.

while *v is not a leaf node* **do**

  //Invariant: (1) $X \times Y \subseteq \Pi_v$; (2) there exists some $z_{r-1} \in [n]$ such that $\mathsf{pt}_{r-1}(f_A, f_B) = z_{r-1}$ $\forall (f_A, f_B) \in X \times Y$.

  Let $u_0 := \Pi(v, 0), u_1 := \Pi(v, 1)$ be the two children of v.

  if owner(v) = Alice **then**

    Partition X into $X = X^0 \cup X^1$ such that $X^b \times Y \subseteq \Pi_{u_b}$ for $b \in \{0, 1\}$.

    Sample $\boldsymbol{b} \in \{0, 1\}$ such that $\mathbf{Pr}[\boldsymbol{b} = b] = |X^b|/|X|$ for $b \in \{0, 1\}$.

    Update $X := X^{\boldsymbol{b}}, v := u_{\boldsymbol{b}}$.

    if owner$(u_{\boldsymbol{b}})$ = Bob **then**

      // A new round.

      Further Partition X into $X = X^0 \cup X^1$ where $X^b := \{f_A \in X : f_A(z_{r-1}) \bmod 2 = b\}$.

      Sample $\boldsymbol{b}' \in \{0, 1\}$ such that $\mathbf{Pr}[\boldsymbol{b}' = b] = |X^b|/|X|$ for $b \in \{0, 1\}$.

      Update $X := X^{\boldsymbol{b}'}, r := r + 1$.

    Let $X = X^1 \cup \cdots \cup X^m$ be the decomposition of X promised by lemma 2.3 with associated sets $I_1, \ldots, I_m \subseteq J_A$.

    // Invoking lemma 2.3 with $J = J_A, M = n, \gamma = 1 - \frac{0.1}{\log n}$.

    Sample a random element $\boldsymbol{j} \in [m]$ such that $\mathbf{Pr}[\boldsymbol{j} = j] = |X^j|/|X|$ for $j \in [m]$.

    Update $X := X^{\boldsymbol{j}}, J_A := J_A \setminus I_{\boldsymbol{j}}$.

    if owner$(u_{\boldsymbol{b}})$ = Bob $\wedge$ $z_{r-1} \notin J_B$ **then**

      bad := TRUE.

  if owner(v) = Bob **then**

    Partition Y into $Y = Y^0 \cup Y^1$ such that $X \times Y^b \subseteq \Pi_{u_b}$ for $b \in \{0, 1\}$.

    Sample $\boldsymbol{b} \in \{0, 1\}$ such that $\mathbf{Pr}[\boldsymbol{b} = b] = |Y^b|/|Y|$ for $b \in \{0, 1\}$.

    Update $Y := Y^{\boldsymbol{b}}, v := u_{\boldsymbol{b}}$.

    if owner$(u_{\boldsymbol{b}})$ = Alice **then**

      Further Partition Y into $Y = Y^0 \cup Y^1$ where $Y^b := \{f_B \in Y : f_B(z_{r-1}) \bmod 2 = b\}$.

      Sample $\boldsymbol{b}' \in \{0, 1\}$ such that $\mathbf{Pr}[\boldsymbol{b}' = b] = |Y^b|/|Y|$ for $b \in \{0, 1\}$.

      Update $Y := Y^{\boldsymbol{b}'}, r := r + 1$.

    Let $Y = Y^1 \cup \cdots \cup Y^m$ be the decomposition of Y promised by lemma 2.3 with associated sets $I_1, \ldots, I_m \subseteq J_B$.

    Sample a random element $\boldsymbol{j} \in [m]$ such that $\mathbf{Pr}[\boldsymbol{j} = j] = |Y^j|/|Y|$ for $j \in [m]$.

    Update $Y := Y^{\boldsymbol{j}}, J_B := J_B \setminus I_{\boldsymbol{j}}$.

    if owner$(u_{\boldsymbol{b}})$ = Alice $\wedge$ $z_{r-1} \notin J_A$ **then**

      bad := TRUE.

3.2 Relating Accuracy and Average Fixed Size

From lemma 3.1 we know that the coordinates in $\overline{J_A}$ and $\overline{J_B}$ are fixed if we only look at the inputs in $X \times Y$. Intuitively, the advantage of the protocol comes from such fixed coordinates, since the 'alive' coordinates J_A, J_B are dense in the sense that $\boldsymbol{X}(J_A), \boldsymbol{Y}(J_B)$ is γ-dense. This intuition is formalized in the following lemma.

Lemma 3.2 (Relating accuracy and avarage fixed size)**.** *Let* Π *be a* $(k-1)$*-round deterministic protocol where Alice speaks first. Then*

$$\Pr_{f_A, f_B \leftarrow [n]} [\Pi(f_A, f_B) = PC_k(f_A, f_B)] \leq \frac{n^{1-\gamma}}{2} + n^{-\gamma} \cdot (k-1) \cdot \mathop{\mathbf{E}}_{(R, J_A, J_B) \leftarrow \mathsf{DS}(\Pi)} \left[|\overline{J_A}| + |\overline{J_B}| \right].$$

The proof of the lemma is by the following two claims. The first claim 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 claim shows the probability that the flag is raised is bounded in terms of the average fixed size.

Claim 3.3. *If* $\mathsf{DS}(\Pi)$ *outputs* $(R = X \times Y, J_A, J_B)$ *and* $\mathsf{bad} = \textsc{False}$ *in the end, then*

$$\Pr_{(f_A, f_B) \leftarrow R} [\Pi(f_A, f_B) = PC_k(f_A, f_B)] \leq \frac{n^{1-\gamma}}{2}.$$

Claim 3.4. $\mathbf{Pr}_{\mathsf{DS}(\Pi)} [\mathsf{bad} = \textit{True}] \leq n^{-\gamma} \cdot (k-1) \cdot \mathbf{E}_{(R, J_A, J_B) \leftarrow \mathsf{DS}(\Pi)} \left[|\overline{J_A}| + |\overline{J_B}| \right]$.

Next, we first prove lemma 3.2 using the above two claims, and the proof of the claims is followed.

Proof of lemma 3.2. Note that in the running of $\mathsf{DS}(\Pi)$, we always update R to a randomly chosen rectangle and the probability of each rectangle being chosen is proportional to its size. Consequently,

$$\begin{aligned}
&\Pr_{f_A, f_B \leftarrow [n]^n} [\Pi(f_A, f_B) = \mathrm{PC}_k(f_A, f_B)] \\
=& \Pr_{(R, J_A, J_B) \leftarrow \mathsf{DS}(\Pi), (f_A, f_B) \leftarrow R} [\Pi(f_A, f_B) = \mathrm{PC}_k(f_A, f_B)] \\
\leq& \Pr_{\mathsf{DS}(\Pi)} [\mathsf{bad} = \textsc{True}] + \Pr_{(R, J_A, J_B) \leftarrow \mathsf{DS}(\Pi), (f_A, f_B) \leftarrow R} [\Pi(f_A, f_B) = \mathrm{PC}_k(f_A, f_B) \wedge \mathsf{bad} = \textsc{False}] \\
\leq& \frac{n^{1-\gamma}}{2} + n^{-\gamma} \cdot (k-1) \cdot \mathop{\mathbf{E}}_{(R, J_A, J_B) \leftarrow \mathsf{DS}(\Pi)} \left[|\overline{J_A}| + |\overline{J_B}| \right].
\end{aligned}$$

where the last step is by claim 3.3 and claim 3.4. ◻

It remains to prove the two claims.

Proof of claim 3.3. WLOG, assume $k-1$ is odd and the protocol always has k round. Let z_{k-1} be the pointer guaranteed by the loop invariant (lemma 3.1), i.e., $\mathsf{pt}_{k-1}(f_A, f_B) = z_{k-1}$ for all $(f_A, f_B) \in R$. Since $\mathsf{bad} = \textsc{False}$, we have $z_{k-1} \in J_A$. Again by the loop invariant, $\mathbf{H}_\infty(\boldsymbol{X}(z_{k-1})) \geq \gamma$. Moreover, since R is contained in some leaf node of Π, Π output the same answer in R, say $b^* \in \{0,1\}$. Consequently,

$$\begin{aligned}\Pr_{(f_A,f_B)\leftarrow R}[\Pi(f_A,f_B) = \mathrm{PC}_k(f_A,f_B)] &= \Pr_{f_A\leftarrow X}[f_A(z_{k-1}) \bmod 2 = b^*] \\ &\le \sum_{\sigma\in[n]:\sigma \bmod 2 = b^*} \Pr_{f_A\leftarrow X}[f_A(z_{k-1}) = \sigma] \\ &\le \frac{n}{2}\cdot n^{-\gamma}.\end{aligned}$$

□

Proof of claim 3.4. Let $\mathcal{E}_\ell$ denote the event that the flag bad is raised when $r = \ell+1$ (i.e., when the ℓ-th round ends) for the first time. Clearly, $\Pr[\mathsf{bad} = \textsc{True}] = \sum_{\ell=1}^{k-1}\Pr[\mathcal{E}_\ell]$. It suffices to bound each $\Pr[\mathcal{E}_\ell]$.

Assume ℓ is odd, meaning that Alice speaks in the ℓ-th round. Let coin denote the randomness used for the first $(\ell-1)$ rounds. Let $X^{(\ell-1)}, J_A^{(\ell-1)}, J_B^{(\ell-1)}$ be the sets X, J_A, J_B when executing $\mathsf{DS}(\Pi)$ using coin until the ℓ-th round begins. Let $z_{\ell-1}$ be the pointer promised by the invariant. For $\mathcal{E}_\ell$ to happen, we must have $\mathsf{bad} = \textsc{False}$ until the ℓ-th round begins, meaning that $z_{\ell-1} \in J_A^{(\ell-1)}$.

Note that the random variable $\boldsymbol{z}_\ell$ exactly has the same distribution as $\boldsymbol{X}^{(\ell-1)}(z_{\ell-1})$. This is because, in the ℓ-th round (i.e., until r steps to $\ell+1$), we decompose $X^{(\ell-1)}$ into finer sets and update X to be one of them with probability proportional to their size. Therefore,

$$\begin{aligned}\Pr_{\mathsf{coin}'}[\mathcal{E}_t] = \Pr_{\mathsf{coin}'}\left[z_\ell \notin J_B^{(\ell-1)}\right] = \Pr_{f_A\leftarrow \boldsymbol{X}^{(\ell-1)}}\left[f_A(z_{\ell-1}) \notin J_B^{(\ell-1)}\right] &= \sum_{\sigma\in \overline{J_B^{(\ell-1)}}} \Pr_{f_A\leftarrow \boldsymbol{X}^{(\ell-1)}}[f_A(z_{\ell-1}) = \sigma] \\ &\le \left|\overline{J_B^{(\ell-1)}}\right| \cdot n^{-\gamma},\end{aligned}$$

where we fix coin and the probability runs over coin′, the randomness used afterward; the last inequality holds because $z_{\ell-1} \in J_A^{(\ell-1)}$ and $\boldsymbol{X}^{(\ell-1)}\left(J_A^{(\ell-1)}\right)$ is γ-dense (by Item (♣) in lemma 3.1). Averaging over coin, we get

$$\Pr_{\mathsf{DS}(\Pi)}[\mathcal{E}_\ell] \le \mathop{\mathbf{E}}_{\mathsf{coin}}\left[\left|\overline{J_B^{(\ell-1)}}\right|\right]\cdot n^{-\gamma} \le \mathop{\mathbf{E}}_{(R,J_A,J_B)\leftarrow\mathsf{DS}(\Pi)}\left[|\overline{J_B}|\right]\cdot n^{-\gamma},$$

where the second inequality holds because J_B becomes smaller and smaller during the execution.

For even ℓ's, we analogously have $\Pr[\mathcal{E}_\ell] \le \mathbf{E}\left[|\overline{J_A}|\right]\cdot n^{-\gamma}$, and hence the claim follows from union bound. □

3.3 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., $\mathbf{E}_{(R,J_A,J_B)\leftarrow\mathsf{DS}(\Pi)}\left[|\overline{J_A}| + |\overline{J_B}|\right]$), in what follows we show that the average fixed size is at most $O(\mathrm{CC}(\Pi))$. Formally, we prove that

Lemma 3.5. *Let Π be a $(k-1)$-round deterministic protocol where Alice speaks first. Then*

$$\mathop{\mathbf{E}}_{(R,J_A,J_B)\leftarrow\mathsf{DS}(\Pi)}\left[|\overline{J_A}| + |\overline{J_B}|\right] \le \frac{3\mathrm{CC}(\Pi)}{(1-\gamma)\log n}.$$

Remark 3.6. We shall set $\gamma := 1 - \frac{0.1}{\log n}$ and hence the right-handed side equals $30\mathrm{CC}(\Pi)$.

Proof. We shall prove this lemma by density increment argument. That is, we study the change of the density function

$$\mathbf{D}_\infty(R) \stackrel{\text{def}}{=} \mathbf{D}_\infty(\boldsymbol{X}(J_A)) + \mathbf{D}_\infty(Y(J_B)). \tag{1}$$

in each iteration. Let $\boldsymbol{\phi_t}$ be the value of $\mathbf{D}_\infty(R)$ at the end of the t-th iteration. Assume without loss of generality Alice speaks (i.e., $\mathsf{owner}(v) = \mathsf{Alice}$) in the t-th iteration.

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

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

$$\sum_{b\in\{0,1\}} \frac{|X^b|}{|X|} \log\left(\frac{|X|}{|X^b|}\right) \le 1. \tag{2}$$

2. Next, suppose that updating v leads to the switch of the owner, i.e., Line 13 is triggered. Since we also partition X into two parts and update X with probability proportional to the size of each part, the same argument applies. That is, taking expectation over the random choice of $\boldsymbol{b'}$, $\mathbf{D}_\infty(\boldsymbol{X}(J_A))$ increases by at most 1 in expectation.

3. Finally, we further partition X according to lemma 2.3. Say X is partitioned into $X = X^1 \cup \cdots \cup X^m$ and let $I_1, \ldots, I_m$ be the index sets promised by lemma 2.3; and for all $j \in [m]$ we have

$$\mathbf{D}_\infty(\boldsymbol{X}^j(J_A \setminus I_j)) \le \mathbf{D}_\infty(\boldsymbol{X}(J_A)) - (1-\gamma)\log n |I_j| + \delta_j,$$

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

$$\mathbf{D}_\infty(\boldsymbol{X}(J_A)) - \mathop{\mathbf{E}}_{j\leftarrow \boldsymbol{j}}\left[\mathbf{D}_\infty(\boldsymbol{X}^j(J_A \setminus I_j))\right] \ge \mathop{\mathbf{E}}_{j\leftarrow \boldsymbol{j}}\left[(1-\gamma)\log n \cdot |I_j| - \delta_j\right]. \tag{3}$$

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

$$\mathop{\mathbf{E}}_{j\leftarrow \boldsymbol{j}}\left[\delta_j\right] = \sum_{j=1}^m p_j \log \frac{1}{\sum_{v\ge j} p_j} \le \int_0^1 \frac{1}{1-x} \mathrm{d}x \le 1. \tag{4}$$

Let $\mathcal{F}_{t-1}$ be the σ-algebra generated by the random coins used for the first $(t-1)$ iterations. Let $\boldsymbol{\beta}_t$ be the increment of $|\overline{J_A}|$ and $|\overline{J_B}|$ in the t-th iteration. Observe that $\boldsymbol{\beta}_t = |I_{\boldsymbol{j}}|$ by definition. By eq. (3) and eq. (4), taking expectation over random choice of $\boldsymbol{j}$, $\mathbf{D}_\infty(\boldsymbol{X}(J_A))$ decrease by at least $(1-\gamma)\log n \cdot \mathbf{E}\left[\boldsymbol{\beta}_t \mid \mathcal{F}_{t-1}\right] - 1$ due to the density restoring partition. Then

$$\mathbf{E}\left[\boldsymbol{\phi}_t - \boldsymbol{\phi}_{t-1}\right] = \mathbf{E}\left[\mathbf{E}\left[\boldsymbol{\phi}_t - \boldsymbol{\phi}_{t-1} \mid \mathcal{F}_{t-1}\right]\right] \le \mathbf{E}\left[1 + \boldsymbol{\eta_t} - ((1-\gamma)\log n \cdot \boldsymbol{\beta}_t - 1)\right], \tag{5}$$

where $\boldsymbol{\eta_t} \stackrel{\text{def}}{=} \mathbb{1}[\text{owner switches in the } t\text{-th iteration}]$.

Write $c \overset{\text{def}}{=} \mathrm{CC}(\Pi)$ and assume we always have c iterations. [1] In the beginning, $\boldsymbol{\phi}_0 = \mathbf{D}_\infty([n]^n \times [n]^n) = 0$. Since the density function is always non-negative by definition, we have $\boldsymbol{\phi}_c \geq 0$ and thus $\mathbf{E}[\boldsymbol{\phi}_c - \boldsymbol{\phi}_0] \geq 0$. On the other hand, by telescoping,

$$\mathbf{E}[\boldsymbol{\phi}_c - \boldsymbol{\phi}_0] = \sum_{t=1}^{c} \mathbf{E}[\boldsymbol{\phi}_t - \boldsymbol{\phi}_{t-1}] \leq 2c + \sum_{t=1}^{c} \mathbf{E}[\boldsymbol{\eta}_t - (1-\gamma)\log n \cdot \boldsymbol{\beta}_t],$$

where the inequality follows from eq. (5). Observe that $\sum_{t=1}^{c} \boldsymbol{\eta}_t$ is at most k and $\sum_{t=1}^{c} \boldsymbol{\beta}_t = |\overline{J_A}| + |\overline{J_B}|$ by definition. We conclude that

$$\mathbf{E}\left[|\overline{J_A}| + |\overline{J_B}|\right] = \mathbf{E}\left[\sum_{t=1}^{c} \boldsymbol{\beta}_t\right] \leq \frac{2c+k}{(1-\gamma)\log n} \leq \frac{3c}{(1-\gamma)\log n},$$

as desired. $\square$

Proving the main theorem. Now our main theorem easily follows from the two lemmas.

Proof of theorem 1.2. Set $\gamma \overset{\text{def}}{=} 1 - \frac{0.1}{\log n}$. By lemma 3.5 and lemma 3.2, we get

$$\begin{aligned}\texttt{Accuracy}(\Pi) \overset{\text{def}}{=} \Pr_{f_A, f_B \leftarrow [n]^n}[\Pi(f_A, f_B) = \mathrm{PC}_k(f_A, f_B)] &\leq \frac{n^{1-\gamma}}{2} + n^{-\gamma} \cdot (k-1) \cdot \frac{3\mathrm{CC}(\Pi)}{(1-\gamma)\log n} \\ &\leq 0.54 + \frac{1.08(k-1)}{n} \cdot 30\mathrm{CC}(\Pi),\end{aligned}$$

where we use $\frac{n^{1-\gamma}}{2} \leq 0.54, n^{-\gamma} \leq \frac{1.08}{n}$. Since we assumed $\texttt{Accuracy}(\Pi) \geq 2/3$, we conclude that

$$\mathrm{CC}(\Pi) \geq \frac{2/3 - 0.54}{1.08 \cdot 30} \cdot \frac{n}{k-1} > 0.0039 \cdot \frac{n}{k-1} = \Omega(n/k).$$

We also trivially have $\mathrm{CC}(\Pi) \geq k-1$ as Π has $(k-1)$ rounds; putting it together we conclude that $\mathrm{CC}(\Pi) = \Omega(n/k + k)$. $\square$

Acknowledgements

We thank Sepehr Assadi for his helpful comments and suggestions.

References

[ACK19] 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. 1, 2

[ANRW15] Noga Alon, Noam Nisan, Ran Raz, and Omri Weinstein. Welfare maximization with limited interaction. In *2015 IEEE 56th Annual Symposium on Foundations of Computer Science*, pages 1499–1512. IEEE, 2015. 5

[1]Namely, Π communicates c bits on all inputs.

[BBCR10] 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. 5

[BRWY13] Mark Braverman, Anup Rao, Omri Weinstein, and Amir Yehudayoff. Direct products in communication complexity. In *2013 IEEE 54th Annual Symposium on Foundations of Computer Science*, pages 746–755. IEEE, 2013. 5

[BVDBE+22] Joakim Blikstad, Jan Van Den Brand, Yuval Efron, Sagnik Mukhopadhyay, and Danupon Nanongkai. Nearly optimal communication and query complexity of bipartite matching. In *2022 IEEE 63rd Annual Symposium on Foundations of Computer Science (FOCS)*, pages 1174–1185. IEEE, 2022. 5

[CFK+19] Arkadev Chattopadhyay, Yuval Filmus, Sajin Koroth, Or Meir, and Toniann Pitassi. Query-to-communication lifting for bpp using inner product. In *46th International Colloquium on Automata, Languages, and Programming (ICALP 2019)*. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, 2019. 3

[CG18] Clément L Canonne and Tom Gur. An adaptivity hierarchy theorem for property testing. *computational complexity*, 27:671–716, 2018. 1, 3

[CPP22] Xi Chen, Christos Papadimitriou, and Binghui Peng. Memory bounds for continual learning. In *2022 IEEE 63rd Annual Symposium on Foundations of Computer Science (FOCS)*, pages 519–530. IEEE, 2022. 2, 3

[DGS84] Pavol Duris, Zvi Galil, and Georg Schnitger. Lower bounds on communication complexity. In *Proceedings of the sixteenth annual ACM symposium on Theory of computing*, pages 81–91, 1984. 2

[DJS98] Carsten Damm, Stasys Jukna, and Jiří Sgall. Some bounds on multiparty communication complexity of pointer jumping. *Computational Complexity*, 7:109–127, 1998. 2

[DNO14] Shahar Dobzinski, Noam Nisan, and Sigal Oren. Economic efficiency requires interaction. In *Proceedings of the forty-sixth annual ACM symposium on Theory of computing*, pages 233–242, 2014. 5

[FKM+09] Joan Feigenbaum, Sampath Kannan, Andrew McGregor, Siddharth Suri, and Jian Zhang. Graph distances in the data-stream model. *SIAM Journal on Computing*, 38(5):1709–1727, 2009. 1, 2, 3, 4, 5

[GO16] Venkatesan Guruswami and Krzysztof Onak. Superlinear lower bounds for multipass graph processing. *Algorithmica*, 76:654–683, 2016. 1, 2, 5

[GPW15] Mika Göös, Toniann Pitassi, and Thomas Watson. Deterministic communication vs. partition number. In *2015 IEEE 56th Annual Symposium on Foundations of Computer Science*, pages 1077–1088. IEEE, 2015. 3

[GPW17] 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. 3, 4, 5, 15

[JMR20] Matthew Joseph, Jieming Mao, and Aaron Roth. Exponential separations in local differential privacy. In *Proceedings of the Fourteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 515–527. SIAM, 2020. 2, 3

[Kho02] Subhash Khot. On the power of unique 2-prover 1-round games. In *Proceedings of the thiry-fourth annual ACM symposium on Theory of computing*, pages 767–775, 2002. 5

[Kla00] Hartmut Klauck. On quantum and probabilistic communication: Las vegas and one-way protocols. In *Proceedings of the thirty-second annual ACM symposium on Theory of computing*, pages 644–651, 2000. 2, 4

[KNTSZ07] Hartmut Klauck, Ashwin Nayak, Amnon Ta-Shma, and David Zuckerman. Interaction in quantum communication. *IEEE Transactions on Information Theory*, 53(6):1970–1982, 2007. 2, 4

[KPPY84] Maria Klawe, Wolfgang J Paul, Nicholas Pippenger, and Mihalis Yannakakis. On monotone formulae with restricted depth. In *Proceedings of the sixteenth annual ACM symposium on Theory of computing*, pages 480–487, 1984. 1

[LMM+22] Shachar Lovett, Raghu Meka, Ian Mertz, Toniann Pitassi, and Jiapeng Zhang. Lifting with sunflowers. *Leibniz international proceedings in informatics*, 215, 2022. 3

[NDSP11] Danupon Nanongkai, Atish Das Sarma, and Gopal Pandurangan. A tight unconditional lower bound on distributed randomwalk computation. In *Proceedings of the 30th annual ACM SIGACT-SIGOPS symposium on Principles of distributed computing*, pages 257–266, 2011. 1, 3

[NW91] 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. 1, 2, 3, 4, 7

[PNP24] Binghui Peng, Srini Narayanan, and Christos Papadimitriou. On limitations of the transformer architecture. *arXiv preprint arXiv:2402.08164*, 2024. 2, 3

[PRV01] Stephen J Ponzio, Jaikumar Radhakrishnan, and Srinivasan Venkatesh. The communication complexity of pointer chasing. *Journal of Computer and System Sciences*, 62(2):323–355, 2001. 2

[RM97] Ran Raz and Pierre McKenzie. Separation of the monotone nc hierarchy. In *Proceedings 38th Annual Symposium on Foundations of Computer Science*, pages 234–243. IEEE, 1997. 3

[RY20] Anup Rao and Amir Yehudayoff. *Communication Complexity: and Applications*. Cambridge University Press, 2020. 1

[Yeh20] Amir Yehudayoff. Pointer chasing via triangular discrimination. *Combinatorics, Probability and Computing*, 29(4):485–494, 2020. 2, 4, 5

[YZ24] Guangxu Yang and Jiapeng Zhang. Communication lower bounds for collision problems via density increment arguments. *STOC 2024 (to appear)*, 2024. 4

Appendix

The following lemma and proof are from Lemma 5 in [GPW17].

Lemma 3.7 (lemma 2.3 restated)**.** *Let $\gamma \in (0,1)$. Let X be a subset of $[n]^M$ and $J \subseteq [M]$. Suppose that there exists an $\beta \in [n]^{\bar{J}}$ such that $\forall x \in X, x(\bar{J}) = \beta$. 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. *$\mathbf{D}_\infty\left(\boldsymbol{X}^i(J \setminus I_i)\right) \leq \mathbf{D}_\infty\left(\boldsymbol{X}(J)\right) - (1-\gamma)\log n \cdot |I_i| + \delta_i$, where $\delta_i \stackrel{\text{def}}{=} \log(|X|/|\cup_{j\geq i} X^j|)$.*

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

Algorithm 2: 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 $\mathbf{H}_\infty(\boldsymbol{X}(I)) < \gamma|I|\log n$ 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 $i := i + 1$.

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 $X^{\geq i} \stackrel{\text{def}}{=} \cup_{j\geq i} X^i$. Then

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

where the first equality holds as $(\boldsymbol{X}^{\geq i}|\boldsymbol{X}^{\geq i}(I_i) = \alpha_i) = \boldsymbol{X}^i$. 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.

Finally, Item 3 is proved by straightforward calculation:

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

□

ECCC ISSN 1433-8092

https://eccc.weizmann.ac.il