Electronic Colloquium on Computational Complexity, Report No. 19 (2022)

Simulation Methods in Communication Lower Bounds, Revisited

Guangxu Yang
University of Electronic Science and Technology of China
yanggx187@gmail.com

Jiapeng Zhang*
Department of Computer Science
University of Southern California
jiapengz@usc.edu

February 17, 2022

Abstract

The notion of lifting theorems is a generic method to lift hardness of one-party functions to two-party lower bounds in communication model. It has many applications in different areas such as proof complexity, game theory, combinatorial optimization. Among many lifting results, a central idea is called Raz-McKenize simulation (FOCS 1997). This simulation provides a systematic way to convert a communication protocol into a corresponding decision tree. Though it is very convenient in many applications, there are still some challenges in this framework. A major problem is that Raz-McKenize simulation requires a very large gadget.

In this paper, we revise Raz-McKenzie simulation. We introduce a white-box simulation, proving lifting theorems for block sensitivity with constant-size gadgets. Concretely, we show there is a constant-size gadget g such that for any Boolean function f, the corruption bound of $f \circ g^n$ is lower bounded by $\Omega(\mathrm{bs}(f))$. Combined with a result of Beame et al. (CCC 2005), this implies the randomized communication complexity of $f \circ g^n$ is lower bounded by $\Omega(\mathrm{bs}(f))$. Besides the result itself, we believe our simulation technique may have more applications in diverse areas. We also discuss why our simulation method has a potential to avoid the large-size gadget bottleneck in Raz-McKenzie simulation.

1 Introduction

The methodology of lifting theorems is a reductive lower bound technique that converts lower bounds of f in query model (simpler-to-understand) to lower bounds of lifted functions $f \circ g^n$ in communication model. For a function $f : \{0,1\}^n \to R$ and a function $g : X \times Y \to \{0,1\}$ (we call it *gadget*), their composition $f \circ g^n : X^n \times Y^n \to R$ is defined by

$$(f \circ g^n)(x, y) := f(g(x_1, y_1), \ldots, g(x_n, y_n)).$$

In the communication model, Alice holds $x \in X^n$ and Bob holds $y \in Y^n$ respectively, and their goal is to compute $(f \circ g^n)(x, y)$ through a communication channel.

There is a substantial number of works studying lifting theorems for a variety of *query-to-communication* models including: deterministic [RM97, GPW18, CFK$^+$19, LMM$^+$20], randomized [GPW20], non-deterministic [GLM$^+$16, PSW21] and degree-to-rank [She11, PR17, PR18, RPRC16,

*Research supported by NSF CAREER award 2141536.

ISSN 1433-8092

CLRS16, KMR21, LRS15]. In these papers and others, lifting theorems have been applied to simplify and resolve longstanding open problems, including new separations in communication complexity [GPW18, GPW20], resolution of the clique vs independent set problem [Göö15] and refutation of the log approximate-rank conjecture [CMS20]. Lifting theorems also have applications in other literature, such as proof complexity [GLM$^+$16, HN12, GP18, DRNV16, dR19], monotone circuit complexity [GGKS18], date structure lower bounds [CKLM18], communication complexity of Nash equilibrium [GR21, BR20], monotone span programs, linear secret sharing schemes [RPRC16, PR17, PR18] and lower bounds on the extension complexity of linear and semi-definite programs [GJW18, CLRS16, KMR21, LRS15].

Among many of these proofs, a very important idea is the *simulation technique*. It converts any communication protocol for $f \circ g^n$ (where f is arbitrary but the gadget g is chosen carefully) into a decision tree for f. Simulation methods are generic and flexible in that it translates to different communication models [GKPW19, GJPW18, GLM$^+$16], and can be specialized to various gadgets [LM19, CKLM18] as well. Applications of simulation-based lifting theorems often depend on the size of the gadget where reducing the gadget size to a constant would be a fundamental breakthrough with many interesting applications [GP18, GJW18, GR21]. An ideal lifting theorem with constant gadget size would give a unified way to prove tight lower bounds for most known functions. However, existing lifting theorems [RM97, WYY17, CKLM19, GPW18, CFK$^+$19, LMM$^+$20] based on simulation methods need gadgets of very large size ($q = \mathsf{poly}(n)$). In Raz-McKenzie simulation [RM97, GPW18], there is a crucial average-to-worst reduction (known as the "thickness lemma"), which requires $q = \Omega(n^2)$ to maintain the pseudo-random property. Based on robust sunflowers ([ALWZ21]), Lovett et al. [LMM$^+$20] gives a simulation with $q = \Theta(n \log q)$, which is the best known result. Thus, the major open problem of the lifting literature is to prove a query-to-communication lifting theorem that uses gadgets of constant size.

In this paper, we revise the simulation method and adapt it to lifting theorems with constant gadgets. Concretely, we show there are constants $\epsilon > 0$ and $q > 1$ such that $\mathrm{Corr}_\epsilon(f \circ g^n) = \Omega(\log q \cdot \mathrm{bs}(f))$, where $g : [q] \times [q] \to \{0,1\}$ is any balanced gadget with small discrepancy. As a byproduct, we show that $\mathbf{BPP}^{\mathrm{cc}}(f \circ g^n) = \Omega(\log q \cdot \mathrm{bs}(f))$. Even though similar results were known before [Zha09, GP18, HN12, ABDK20], our result still gives some novel insights which we believe has a potential to give more applications.

1.1 Our Results

In this work, we study lifting theorems with low-discrepancy gadgets of constant size. In what follows, we denote by $\mathbf{P}^{\mathrm{cc}}$ the deterministic communication complexity and $\mathbf{BPP}^{\mathrm{cc}}$ the randomized (public-coin) communication complexity with constant error probability. We denote $\mathrm{bs}(f)$ as the block sensitivity of a Boolean function f and $\mathrm{disc}(g)$ as the discrepancy of a gadget g (Please see formal definition in Section 2). We call a gadget g balanced if exactly half of the inputs satisfy $g(x, y) = 0$.

As a warm-up, we first present a deterministic lifting theorem for block sensitivity.

Theorem 1.1. *There is a constant $q > 0$. For any Boolean function $f : \{0,1\}^n \to \{0,1\}$ and any balanced gadget $g : [q] \times [q] \to \{0,1\}$ with $\mathrm{disc}(g) \le 2^{-\log q}$, we have that*

$$\mathbf{P}^{\mathrm{cc}}(f \circ g^n) = \Omega(\log q \cdot \mathrm{bs}(f)).$$

A large family of gadgets are balanced with $disc(g) \le 2^{-\log q}$. For example, the inner product gadget (previously known in [CKLM19, WYY17]) has this property. We make two remarks: 1).

we believe the balanced assumption can be removed; 2) we can relax the discrepancy condition a little bit in which requires only $\mathrm{disc}(g) = 2^{-\gamma \cdot \log q}$ for a constant such as $\gamma = 0.1$. We do not make these efforts here since we would like to keep our presentation as simple as possible.

By refining our ideas further, we are also able to prove a stronger statement. Let $\mathrm{Corr}_\epsilon(f)$ be the corruption bound of f with error ϵ.

Theorem 1.2. *There exist constants $\epsilon > 0$ and $q > 1$. For any Boolean function $f : \{0,1\}^n \to \{0,1\}$ and any balanced gadget $g : [q] \times [q] \to \{0,1\}$ with $disc(g) \le 2^{-\log q}$, we have that*

$$\mathrm{Corr}_\epsilon(f \circ g^n) = \Omega(\log q \cdot \mathrm{bs}(f)).$$

Since corruption bound is a lower bound of randomized communication complexity [BPSW05], we have the following corollary.

Corollary 1.3. *There exist constants $\epsilon > 0$ and $q > 1$. For any Boolean function $f : \{0,1\}^n \to \{0,1\}$ and any balanced gadget $g : [q] \times [q] \to \{0,1\}$ with $disc(g) \le 2^{-\log q}$, we have that*

$$\mathbf{BPP}^{\mathrm{cc}}(f \circ g^n) = \Omega(\log q \cdot \mathrm{bs}(f)).$$

Based on known connections between corruption bound and information complexity [KLL+12], we also have the following corollary.

Corollary 1.4. *There exist constants $\epsilon > 0$ and $q > 1$. For any Boolean function $f : \{0,1\}^n \to \{0,1\}$ and any balanced gadget $g : [q] \times [q] \to \{0,1\}$ with $disc(g) \le 2^{-\log q}$, we have that*

$$\mathsf{IC}^{\mathrm{ext}}(f \circ g^n) = \Omega(\log q \cdot \mathrm{bs}(f)).$$

Applications. Lifting theorems for (critical) block sensitivity gave several known applications in diverse areas, such as the polynomial relation between classical and quantum communication complexities of certain block-composed functions [Zha09], length–space lower bounds for semi-algebraic proof systems[HN12, GP18], monotone circuit depth lower bound [GP18] and randomized communication complexity of approximate Nash equilibrium [GR21]. We believe our method can be used to prove lifting theorems for critical block sensitivity, but it is not the main focus of this paper.

Comparisons with Previous Results: Similar lifting results for (critical) block sensitivity were obtained in [Zha09, HN12, GP18, ABDK20].

Zhang [Zha09] and Göös and Pitassi [GP18] proved $\mathbf{BPP}^{\mathrm{cc}}(f \circ g^n) = \Omega(\log q \cdot \mathrm{bs}(f))$ for some constant-size gadgets g. Their proof is built on a reduction from *Unique-Disjointness*, a well-known hard problem in communication model. Huynh and Jakob [HN12] and Anshu et al. [ABDK20] proved this statement by analyzing information complexity. Both proofs work only on versatile gadgets: A versatile gadget must have *flippability* and *random self-reducibility* which makes it a very special type of gadgets. We note that a flippable gadget is also balanced.

Our method is very different in that we prove it by the simulation framework. As a benefit, we extend their results in two ways: 1) we only require the balanced gadget g to have small discrepancy; 2) we also obtain lower bounds for the corruption bound.

1.2 Overview of Our Technique

In this part, we explain our simulation with a comparison to Raz-McKenzie framework. For simplicity, we use some informal notations here that may have small difference with formal version in Section 3. For now, we assume that $\mathrm{bs}(f) = \mathrm{bs}(f, 0^n) = \mathrm{s}(f, 0^n) = n$. That is, $f(0^n) \neq f(e_i)$ for all indicator vector $e_i \in \{0,1\}^n$. For a rectangle $R \subseteq [q]^n \times [q]^n$, we denote $D(R) = \{(x,y) \in R : g^n(x,y) = 0^n\}$ and $g^n(R) = \{g^n(x,y) : (x,y) \in R\}$.

The idea of our simulation is to (explicitly) find a long path in any communication protocol tree that computes $f \circ g^n$. To achieve this, we define a potential function for rectangles. For each rectangle R, its potential is $E(R) = \log |D(R)|$.

We recall that every node in a communication tree has an associated rectangle. Starting from the root, we find a path as follows: for each intermediate node, the path always visits the child (either left or right) with its associated rectangle maximizes the potential function. We observe two important properties about the potential function.

- For every rectangle R, $E(R) \leq E([q]^n \times [q]^n)$.
- Each step in the path decreases the potential function by at most 1.

Let ℓ be the length of the path, and let R_ℓ be the leaf rectangle reached by the path. We then have that

$$E([q]^n \times [q]^n) - \ell \leq E(R_\ell).$$

On the other hand, we introduce a projection operation. We show that for each rectangle R with $0^n \in g^n(R)$ and $e_i \notin g^n(R)$, a projection on the coordinate i outputs a "sub-rectangle"[1] R' of R such that it has potential function $E(R') \geq E(R) + \Omega(\log q)$. Furthermore, we prove that for every monochromatic rectangle R with $0^n \in g^n(R)$, there is a sub-rectangle R^* of R such that

$$E(R^*) \geq E(R) + \Omega(n \log q),$$

where R^* is obtained by applying the projection on R for each coordinate e_i. We note that R_ℓ is indeed a monochromatic function, hence

$$E([q]^n \times [q]^n) - \ell \leq E(R_\ell) \leq E(R^*_\ell) - \Omega(n \cdot \log q) \leq E([q]^n \times [q]^n) - \Omega(n \cdot \log q).$$

Thus we have that $\ell = \Omega(n \cdot \log q)$.

Our simulation v.s. Raz-McKenzie simulations. Firstly, we point out that our simulation is a method to find a long path in a communication protocol tree. By contrast, Raz-McKenzie simulation is a method that fully converts a communication protocol into a decision tree. Under this strong transformation, any query lower bound on the Boolean function f can be automatically transformed into a communication lower bound of $f \circ g^n$. However, the price of full simulation is expensive. A crucial step in Raz-McKenzie simulation is called the *full rectangle lemma*. Many structures such as Fourier analysis and robust sunflowers were used to prove this full rectangle lemma. But all of them require a gadget g of size at least $\Omega(n)$, which is far away from constants. However this may not be necessary to prove communication lower bounds. It is actually sufficient

[1] Not exactly sub-rectangle, but have a similarity

to find a long path in the communication protocol, which can potentially save the cost from the full rectangle lemma.

Secondly, our simulation is a white-box simulation rather than black-box simulations such as Raz-McKenzie. We call it white-box because the potential function depends on the function f itself. In comparison, the potential function used in Raz-McKenzie is oblivious to the function f. To prove a communication lower bound, we observe that it is sufficient to lift the "hard part" of the function f. For example, if $\mathrm{s}(f) = \mathrm{s}(f, 0^n) = n$, we only need to lift the set $\{0^n, e_1, \ldots, e_n\}$. However, a black-box simulation always lifts the whole domain $\{0,1\}^n$, which also requires the full rectangle lemma.

In summary, our simulation avoids the full rectangle lemma, which is the main bottleneck in Raz-McKenzie simulation. We believe our simulation can be also used in many other lifting applications. For example, it has a potential to attack the main open question in lifting theorems: to prove that $\mathbf{P}^{\mathrm{cc}}(f \circ g^n) = \Omega(\mathbf{P}^{\mathrm{dt}}(f))$ for some constant-size gadget g. Even it may be challenging to achieve this ambitious goal, we can first study those functions f with a good expression on its hard part. In summary, our new simulation has a good potential to obtain many new results.

1.3 Paper Organization.

We recall some useful definitions in Section 2. In Section 3, we explain the proof of deterministic lifting (Theorem 1.1). In Section 4, we give a proof of the randomized lifting (Theorem 1.2). Finally, we discuss open problems in Section 5. Missing proofs are put in Appendix A.

Acknowledgements. Authors thank Jack DePascale for helpful discussions. We are grateful to Kewen Wu for reading early versions of this paper and providing useful suggestions.

2 Preliminary

We assume the reader is familiar with the basic definitions of communication complexity (see, e.g., [Kus97]). For any $n \in \mathbb{N}$, we denote $[n] = \{1, ..., n\}$. Unless otherwise specified, the log in this paper will be base 2.

Let $I \subseteq [n]$ be a set of coordinates. For any set $X \subseteq [q]^n$, we define $X_I = \{(x_i)_{i \in I} \in [q]^I : (x_1, ..., x_n) \in X\}$ be the projection of X onto the coordinates in I and $X\,|_{X_i=u} = \{x \in X : x_i = u\}$ be the restriction of X on $X_i = u$.

Given a boolean function $g : X \times Y \to \{0,1\}$, we denote by $g^I : X^I \times Y^I \to \{0,1\}^I$ the function that takes as inputs $|I|$ pairs from $X \times Y$ that are indexed by I, and outputs the string in $\{0,1\}^I$ whose i-th bit is the output of g on the i-th pair. In particular, we denote $g^n := g^{[n]}$.

Definition 2.1 (Discrepancy)**.** *Let $g : [q] \times [q] \to \{0,1\}$ be a function. Given a rectangle $R \subseteq [q] \times [q]$, the discrepancy of g with respect to R, denoted as $\mathrm{disc}_R(g)$, is*

$$\mathrm{disc}_R(g) = \frac{\big||\{(x,y) \in R : g(x,y) = 0\}| - |\{(x,y) \in R : g(x,y) = 1\}|\big|}{q^2}.$$

The discrepancy of g, denoted as $\mathrm{disc}(g)$, is the maximum of $\mathrm{disc}_R(g)$ over all rectangle $R \subseteq [q] \times [q]$.

Definition 2.2 (Entropy)**.** *Let D be a probability distribution. The entropy of D is*

$$\mathcal{H}(D) := \sum_a \Pr[D = a] \cdot \log(1/\Pr[D = a]).$$

For two (discrete) random variables A and B, the conditional entropy of B given A is defined as

$$\mathcal{H}(B|A) = \sum_{a,b} \Pr[ab] \cdot \log \frac{1}{\Pr[b|a]} = \sum_a \Pr[a] \cdot \mathcal{H}(B|A = a).$$

The following useful inequality will be used in our paper.

Fact 2.3. *For any random variables A and B, we have $\mathcal{H}(B|A) \leq \mathcal{H}(B)$.*

For a string $x \in \{0,1\}^n$ and a set $I \subseteq [n]$, the string x^I is obtained from x by flipping all coordinates in I. We use x^i as a shorthand for $x^{\{i\}}$.

Definition 2.4 (Sensitivity)**.** *Let $f : \{0,1\}^n \to \{0,1\}$ be a Boolean function. For an input $x \in \{0,1\}^n$, the sensitivity $\mathrm{s}(f,x)$ of f at x is the number of coordinates i such that $f(x) \neq f(x^i)$. The sensitivity of f is defined by*

$$\mathrm{s}(f) = \max_{x \in \{0,1\}^n} \mathrm{s}(f,x).$$

Definition 2.5 (Block sensitivity)**.** *Let $f : \{0,1\}^n \to \{0,1\}$ be a Boolean function. For an input $x \in \{0,1\}^n$, the block sensitivity $\mathrm{bs}(f,x)$ of f at x is the maximum number k such that there are disjoint blocks $B_1, \ldots, B_k \subseteq [n]$ satisfying $f(x) \neq f(x^{B_i})$ for every $i \in [k]$. The block-sensitivity of f is defined by*

$$\mathrm{bs}(f) = \max_{x \in \{0,1\}^n} \mathrm{bs}(f,x).$$

Let $F : X \times Y \to \{0,1\}$. we use $\mathcal{R}$ denote the set of rectangles of $X \times Y$. For each $z \in \{0,1\}$, we denote $F^{-1}(z) := \{(x,y) \in X \times Y : F(x,y) = z\}$. For a (joint) distribution μ on (X,Y) and a set $S \subseteq X \times Y$, we define $\mu(S) := \Pr_{(x,y)\sim\mu}[(x,y) \in S]$.

Definition 2.6 (Corruption Bound)**.** *Let $F : X \times Y \to \{0,1\}$ and $\epsilon > 0$. The corruption bound of F with error ϵ, denoted by $\mathrm{Corr}_\epsilon(F)$, is*

$$\max_{\substack{z \in \{0,1\} \\ \mu \text{ on } X \times Y}} \min \left\{ \log \frac{1}{\mu(F^{-1}(z) \cap R)} : R \in \mathcal{R} \text{ with } \mu(F^{-1}(1-z) \cap R) \leq \epsilon \cdot \mu(F^{-1}(z) \cap R) \right\}.$$

By [BPSW05], the corruption bound is a lower bound of the randomized communication complexity.

3 Deterministic Lifting for Block Sensitivity

Now we discuss Theorem 1.1. In this section, we may skip some proofs since they are implied by similar lemmas in Section 4. The main purpose of this section is to give a clear explanation of our simulation framework. For simplicity in presentation, we first prove the following theorem.

Theorem 3.1. *There exists a constant $q > 0$. For any Boolean function $f : \{0,1\}^n \to \{0,1\}$ and balanced gadget $g : [q] \times [q] \to \{0,1\}$ with $disc(g) \leq 2^{-\log q}$, we have that*

$$\mathbf{P}^{\mathrm{cc}}(f \circ g^n) = \Omega(\log q \cdot \mathrm{s}(f)).$$

Comparing with Theorem 1.1, we replace the block sensitivity by sensitivity. Without loss of generality, we assume that $\mathrm{s}(f) = \mathrm{s}(f, 0^n) = n$. That is, $f(0^n) \neq f(e_i)$ for all $i \in [n]$. We now introduce some notations to simplify our presentation. For any $I \subseteq [n]$, we define

$$D_0^I := \{(x, y) \in [q]^I \times [q]^I : g^I(x, y) = 0^I\},$$

and define for each $i \in I$

$$D_i^I := \{(x, y) \in [q]^I \times [q]^I : g^I(x, y) = e_i\},$$

where $e_i \in \{0,1\}^I$ is the indicator vector.

Definition 3.2 (Potential Function)**.** *For each $I \subseteq [n]$ and $X, Y \subseteq [q]^I$, let $R = X \times Y \subseteq [q]^I \times [q]^I$ and define its potential function of R as*

$$E_I(R) := \log\left(\frac{|R \cap D_0^I|}{|D_0^I|}\right).$$

We use $E(R)$ as a shorthand for $E_I(R)$ when I is clear in the context.

We note that $E(R)$ is always non-positive. During the simulation along the communication tree, each communication round will decrease the potential function and each projection round will increase the potential function. Now we define our projection process.

Definition 3.3 (Projection)**.** *Let $R = X \times Y \subseteq [q]^I \times [q]^I$ be a rectangle. For $i \in I$ and $u \in [q]$, the projection by Alice at (i, u) is a rectangle $\Pi_{i,u}^A(R) = X' \times Y' \subseteq [q]^{I\setminus\{i\}} \times [q]^{I\setminus\{i\}}$ where*

- $X' := \{x' \in [q]^{I\setminus\{i\}} : (x', u) \in X\}$,
- $Y' := \{y' \in [q]^{I\setminus\{i\}} : \exists v \in [q], g(u, v) = 0 \text{ and } (y', v) \in Y\}$.

Similarly, we define the projection by Bob as $\Pi_{i,u}^B(R) = X'' \times Y''$ where

- $X'' := \{x' \in [q]^{I\setminus\{i\}} : \exists v \in [q], g(v, u) = 0 \text{ and } (x', v) \in X\}$,
- $Y'' := \{y' \in [q]^{I\setminus\{i\}} : (y', u) \in Y\}$.

Our projection borrows ideas from Raz-McKenzie but with some difference. In Raz-McKenzie, a projection on the coordinate i corresponds to a query of the i-th coordinate in the decision tree. In our projection, we fix the coordinate i as 0. Since we aim to find a long path in the communication tree, we do projections towards the direction that maximizes our potential function. Now we are ready to describe our simulation process in Algorithm 1.

Algorithm 1 Simulation algorithm for deterministic lifting

Input: Communication protocol tree Π

Output: A communication path with length $\Omega(\mathrm{bs}(f) \cdot \log q)$.

Initialize $v =$ root of Π, $R_v = X_v \times Y_v = [q]^n \times [q]^n$, and $I = [n]$.
Let P be an empty string.
while R_v is not a monochromatic rectangle **do**
    Let v_0, v_1 be the children of v in Π.
    if Alice sends a bit at v **then**
        Let X_{v_0}, X_{v_1} be the partition of X_v according to Alice's partition.
        Let $R_{v_0} = X_{v_0} \times Y_v$ and $R_{v_1} = X_{v_1} \times Y_v$.
        Let $b \in \{0, 1\}$ be the bit maximizes $E(R_{v_b})$.
        Let $P = P \circ b$.
        Update $R_v = R_{v_b}$ and $v = v_b$.
    end if
    if Bob sends a bit at v **then**
        Let Y_{v_0}, Y_{v_1} be the partition of Y_v according to Bob's partition.
        Let $R_{v_0} = X_v \times Y_{v_0}$ and $R_{v_1} = X_v \times Y_{v_1}$.
        Let $b \in \{0, 1\}$ be the bit maximizes $E(R_{v_b})$.
        Let $P = P \circ b$.
        Update $R_v = R_{v_b}$ and $v = v_b$.
    end if
    while $R_v \cap D_i^I = \emptyset$ for some $i \in I$ **do**
        Do projection by a player $C \in \{\text{Alice}, \text{Bob}\}$ at (i, u) that maximizes $E(\Pi_{i,u}^C(R_v))$.
        Update $R_v = \Pi_{i,u}^C(R_v)$ and $I = I \setminus i$.
    end while
end while
Output P.

3.1 Proof of Theorem 3.1

To prove Theorem 3.1, we first observe some important properties of our projection.

Lemma 3.4 (Partition Lemma)**.** *Let $R = X \times Y \subseteq [q]^I \times [q]^I$ be a rectangle, then for any partition $R = R_0 \cup R_1$, there is some $b \in \{0, 1\}$ such that $E(R_b) \geq E(R) - 1$.*

The proof of this Lemma 3.4 is straightforward: Since R_0, R_1 is a partition of R, either $|R_0 \cap D_0^I| \geq |R \cap D_0^I|/2$ or $|R_1 \cap D_0^I| \geq |R \cap D_0^I|/2$. If $|R_0 \cap D_0^I| \geq |R \cap D_0^I|/2$, then $E(R_0) \geq E(R) - 1$; otherwise $E(R_1) \geq E(R) - 1$.

Lemma 3.5 (Projection Lemma)**.** *Let $R = X \times Y \subseteq [q]^I \times [q]^I$ be a rectangle. If there is a coordinate $i \in I$ such that $R \cap D_i^I = \emptyset$, then there is a value $u \in [q]$ such that,*

- *either $E(\Pi_{i,u}^A(R)) \geq E(R) + \Omega(\log q)$,*
- *or $E(\Pi_{i,u}^B(R)) \geq E(R) + \Omega(\log q)$.*

Lemma 3.5 is the crucial step in our proof. Since the purpose of this section is merely introduce the proof outline of our main results, we omit the proof of Lemma 3.5 here. We direct readers

interested in its proof to Lemma 4.9 in Section 4, which is the randomized (thus stronger) version. Now we are ready to prove Theorem 3.1.

Proof. Let $(v_0, v_1, \ldots, v_t)$ be the path found by Algorithm 1. We aim to prove that $t = \Omega(n \cdot \log q)$. Let $R_0, R_1, \ldots, R_t$ be rectangles in each step respectively. We first note several important facts:

- $E(R_0) = 0$ and $E(R_j) \leq 0$ for every $j \in [t]$.
- $E(R_j) \geq E(R_{j-1}) - 1$ for every $j \in [t]$.

On the other hand, each projection increases the potential function by $\Omega(\log q)$. Hence we have the following statement:

$$t \geq (\text{number of projections happened in the algorithm}) \cdot \Omega(\log q).$$

Notice that a projection on e_i happens whenever e_i is not in current rectangle, and the projection fixes the i-th coordinate of f to 0. So the projection will be executed by n times to result in a monochromatic rectangle, which shows $t = \Omega(n \cdot \log q)$ as desired. □

3.2 Proof of Theorem 1.1

Now we generalize our proof for block-sensitivity. Let f be a Boolean function such that $\mathrm{bs}(f) = \mathrm{bs}(f, 0^n) = k$. Let $B_1, \ldots, B_k$ be disjoint sets that flips $f(0^n)$. We define a set $X \in [q]^n$ as

$$X = \{x \in [q]^n : \forall \ell \in [k], i_1, i_2 \in B_\ell, x_{i_1} = x_{i_2}\}.$$

Similarly, we define

$$Y = \{y \in [q]^n : \forall \ell \in [k], i_1, i_2 \in B_\ell, y_{i_1} = y_{i_2}\}.$$

Let $D = X \times Y$ be a rectangle. The main idea here is to enforce all coordinates in B_ℓ identical. Let $D_0^n = \{(x, y) \in D : g^n(x, y) = 0^n\}$ and $D_\ell^n := \{(x, y) \in D : \forall i, g(x_i, y_i) = 1 \text{ iff } i \in B_\ell\}$. We can then apply the simulation process in the sub-rectangle D.

4 Randomized Lifting for Block Sensitivity

Similar as Section 3, we assume that $f(0^n) \neq f(e_i)$, for any $i \in [n]$. As we discussed in Section 3.2, this assumption does not lose generality. Let $D_0^I = \{(x, y) \in \{0,1\}^I \times \{0,1\}^I : g^I(x, y) = 0^I\}$ and $D_i^I = \{(x, y) \in \{0,1\}^I \times \{0,1\}^I : g^I(x, y) = e_i\}$ be the same notations used in Section 3. We use D_0 as a shorthand for $D_0^{[n]}$ and D_i a shorthand for $D_i^{[n]}$.

Our simulation for randomized lifting is slightly different from the one used above. To explain connections better, we give a different algorithm (Algorithm 2) to find a long path. Comparing to the previous one, this new algorithm defers all projections unto leaves. By analyzing deferred projections in the leaf found by Algorithm 2, we can prove the following statement.

Theorem 4.1. *Let R be a monochromatic rectangle such that $R \cap D_0 \neq \emptyset$, then*

$$|R \cap D_0| \leq 2^{-\Omega(\mathrm{bs}(f) \cdot \log q)} \cdot |D_0|.$$

Algorithm 2 Simulation with Deferred Projections

Input: Communication protocol tree Π

Output: A communication path with length $\Omega(\mathrm{bs}(f) \cdot \log q)$.

Initialize $v =$ root of Π, $R_v = X_v \times Y_v = [q]^n \times [q]^n$, $I = [n]$.
Let P be an empty string.
while R_v is not a monochromatic rectangle **do**
    Let v_0, v_1 be the children of v in Π.
    if Alice sends a bit at v **then**
        Let X_{v_0}, X_{v_1} be the partition of X_v according to Alice's partition.
        Let $R_{v_0} = X_{v_0} \times Y_v$ and $R_{v_1} = X_{v_1} \times Y_v$.
        Let $b \in \{0, 1\}$ be the bit maximizes $E(R_{v_b})$.
        Let $P = P \circ b$.
        Update $R_v = R_{v_b}$ and $v = v_b$.
    end if
    if Bob sends a bit at v **then**
        Let Y_{v_0}, Y_{v_1} be the partition of Y_v according to Bob's partition.
        Let $R_{v_0} = X_v \times Y_{v_0}$ and $R_{v_1} = X_v \times Y_{v_1}$.
        Let $b \in \{0, 1\}$ be the bit maximizes $E(R_{v_b})$.
        Let $P = P \circ b$.
        Update $R_v = R_{v_b}$ and $v = v_b$.
    end if
end while
while $R_v \cap D_i^I = \emptyset$ for some $i \in I$ **do**
    Do projection by a player $C \in \{\text{Alice}, \text{Bob}\}$ at (i, u) that maximizes $E(\Pi_{i,u}^C(R_v))$.
    Update $R_v = \Pi_{i,u}^C(R_v)$ and $I = I \setminus i$.
end while
Output P.

Now we focus on proving the randomized lifting. We first define a hard distribution μ that will be used in our proof.

1. Sample a bit $b \in \{0, 1\}$ uniformly at random.
2. If $b = 0$, output a uniformly random $(x, y) \in D_0$.
3. If $b = 1$, output a uniformly random $(x, y) \in \bigcup_{i \in [n]} D_i$.

In the rest of the content, the notation μ is always used to refer this hard distribution. To prove the main theorem (Theorem 1.2), it is sufficient to prove the following theorem.

Theorem 4.2. *There is a constant $\epsilon \in (0, 9 \cdot 10^{-6})$ such that the following holds. For any rectangle*

$R \subseteq [q]^n \times [q]^n$, *if*

$$\mu\left(\bigcup_{i=1}^{n} D_i \cap R\right) \leq \epsilon \cdot \mu\left(D_0 \cap R\right),$$

then

$$\mu\left(D_0 \cap R\right) \leq 2^{-\Omega(\mathrm{bs}(f)\cdot \log q)}.$$

This theorem is a randomized version of Theorem 4.1. The main difference is that we replace the monochromatic assumption (i.e., $\bigcup_{i=1}^{n} D_i \cap R = \emptyset$) by a bias assumption (i.e., $\mu(\bigcup_{i=1}^{n} D_i \cap R) \leq \epsilon \cdot \mu(D_0 \cap R)$). We then give a quick proof of Theorem 1.2 by assuming Theorem 4.2.

Proof of Theorem 1.2. Fix any R such that $\epsilon \cdot \mu((f \circ g^n)^{-1}(0) \cap R) \geq \mu((f \circ g^n)^{-1}(1) \cap R)$, by Theorem 4.2, we have

$$\mathrm{Corr}_\epsilon(f \circ g^n) \geq \log \frac{1}{\mu((f \circ g^n)^{-1}(0) \cap R)} = \log \frac{1}{\mu(D_0 \cap R)} = \Omega(\mathrm{bs}(f) \cdot \log q).$$

□

4.1 Proof of Theorem 4.2

To prove Theorem 4.2, we define a new potential function for randomized lifting. For a distribution $\mathcal{D}$ on $X \times Y$, we call it a product distribution if it samples $x \in X$ and $y \in Y$ independently.

Definition 4.3. *For any set $I \subseteq [n]$, let $R = X \times Y$ be a rectangle on $[q]^I \times [q]^I$. For any product distribution $\mathcal{D}$ on R, let $\mathcal{D}' = \mathcal{D}|_{g^I(X,Y)=0^I}$, we define the potential function of $\mathcal{D}$ as*

$$E_I(\mathcal{D}) := \mathcal{H}(\mathcal{D}') - 2 \cdot |I| \log q + |I|.$$

We use $E(\mathcal{D})$ as a shorthand for $E_I(\mathcal{D})$ when I is clear in the context.

Since g is a balanced gadget, we note that $E(\mathcal{D}) = E(R)$ when $\mathcal{D}$ is uniform on R, where $E(R)$ is the potential function used in deterministic lifting. Hence, this definition is a natural generalization of the deterministic case. Similarly, we also have $E(\mathcal{D}) \leq 0$ for any product distribution $\mathcal{D}$.

In deterministic lifting, we apply a projection on a coordinate i whenever $R \cap D_i = \emptyset$. In the randomized case, however, the condition $R \cap D_i = \emptyset$ may not hold. Instead, we use a bias condition defined below.

Definition 4.4. *Let $\mathcal{D}$ be a distribution on $[q]^I \times [q]^I$. For any $(x, y) \in D_0^I$ and $i \in I$, the bias of (x, y) at coordinate i is defined by,*

$$\gamma_i^I(x, y) = \Pr_{(x', y') \sim \mathcal{D}} \left[g(x'_i, y'_i) = 1 \mid \forall \ell \neq i, x'_\ell = x_\ell \text{ and } y'_\ell = x_\ell \right].$$

If $\mathcal{D}$ is a product distribution, the bias of $\mathcal{D}$ at coordinate i is defined by

$$\gamma_i^I(\mathcal{D}) = \mathop{\mathbb{E}}_{(x,y) \sim \mathcal{D}'} [\gamma_i(x, y)].$$

For mixed distribution $\mathcal{D} = \sum_j p_j \cdot \mathcal{D}_j$, we define its bias as

$$\gamma_i^I(\mathcal{D}) = \sum_j p_j \cdot \gamma_i(\mathcal{D}_j).$$

We use $\gamma_i(x, y)$ as a shorthand for $\gamma_i^I(x, y)$ and $\gamma_i(\mathcal{D})$ as a shorthand for $\gamma_i^I(\mathcal{D})$ when I is clear in the context.

In deterministic case, we have that $\mu(D_i \cap R) = 0$ (bias of $\gamma_i(\mathcal{D})$ is 0) for any monochromatic rectangle R with $R \cap D_0 \neq \emptyset$. We have a similar lemma in the randomized case.

Lemma 4.5. *For any fixed $\epsilon \in (0, 1/2)$, let $R \subseteq [q]^n \times [q]^n$ be a rectangle such that*

$$\mu\left(\bigcup_{i=1}^{n} D_i \cap R\right) \leq \epsilon \cdot \mu(D_0 \cap R).$$

Let $\mathcal{D}$ be the uniform distribution on R. We have that,

- $E(\mathcal{D}) = -\log \frac{|D_0|}{|D_0 \cap R|} = -\log \frac{\mu(D_0)}{\mu(R \cap D_0)} \geq -\log \frac{1}{\mu(D_0 \cap R)}$.
- $\sum_i \gamma_i(\mathcal{D}) \leq \epsilon \cdot \mathrm{bs}(f)$.

The proof of this lemma follows from definitions. Now we introduce projections for randomized lifting. We first need to generalize the notion of product distribution a little bit.

Definition 4.6. *Let R be a rectangle. A distribution $\mathcal{D}$ on R is called a mixed distribution if there exists product distributions $\mathcal{D}_1, \ldots, \mathcal{D}_\ell$ and $p_1, \ldots, p_\ell > 0$ with $\sum_j p_j = 1$, such that $\mathcal{D} = \sum_j p_j \cdot \mathcal{D}_j$.*

Definition 4.7. *For a mixed distribution $\mathcal{D} = \sum p_j \mathcal{D}_j$, its potential function is defined by*

$$E(\mathcal{D}) = \sum_j p_j \cdot E(\mathcal{D}_j).$$

We define projections as follows.

Definition 4.8. *Let $\mathcal{D} = X \times Y$ be a product distribution on $[q]^I \times [q]^I$, and let $\mathcal{D}' = \mathcal{D}\,|_{g^I(X,Y)=0^I}$. For $i \in I$, the projection of $\mathcal{D}$ by Alice at coordinate i, denoted by $\Pi_i^A(\mathcal{D})$, is a mixed distribution on $[q]^{I\setminus\{i\}} \times [q]^{I\setminus\{i\}}$ defined as follows:*

- *For each $u \in [q]$, let $p_u = \Pr_{(x,y)\sim\mathcal{D}'}[x_i = u]$ and let $B_u := \{v \in [q] : g(u, v) = 0\}$.*
- *For each $u \in [q]$, let $\mathcal{D}_u = X_u \times Y_u$ be the product distribution on $[q]^{I\setminus\{i\}} \times [q]^{I\setminus\{i\}}$ defined by $X_u = X_{I\setminus\{i\}}\,|_{X_i=u}$ and $Y_u = Y_{I\setminus\{i\}}\,|_{Y_i\in B_u}$.*
- *Output $\Pi_i^A(\mathcal{D}) = \sum_u p_u \cdot \mathcal{D}_u$.*

Similarly, we can define projections $\Pi_i^B(\mathcal{D})$ by Bob. If $\mathcal{D} = \sum_j p_j \cdot \mathcal{D}_j$ is a mixed distribution, projections is naturally defined by $\Pi_i^A(\mathcal{D}) = \sum_j p_j \cdot \Pi_i^A(\mathcal{D}_j)$.

We give a comparison to the deterministic case. In the deterministic case, since we only care about the disperser property [LMM+20], we choose an element $u \in [q]$ that maximizes the potential function. By contrast, in randomized case, projections have to inherit bias, so we decompose all elements $u \in [q]$, and combine them as a mixed distribution. Given definitions of potential functions and projections, we now prove a projection lemma for randomized lifting.

Lemma 4.9. *There is a constant $\epsilon > 0$ and a constant $q > 0$. For any mixed distribution $\mathcal{D} = \sum_j p_j \cdot \mathcal{D}_j$ on $[q]^I \times [q]^I$ with $\gamma_i(\mathcal{D}) \leq \epsilon$ for some $i \in I$, one the following statements must be true:*

- $E(\Pi_i^A(\mathcal{D})) \geq E(\mathcal{D}) + c \cdot \log q$,

- $E(\Pi_i^B(\mathcal{D})) \geq E(\mathcal{D}) + c \cdot \log q$,

where c is a constant depending on ϵ and q.

We defer the proof of this lemma to Section 4.2. Now we present the following lemma indicating that randomized projections keep bias.

Lemma 4.10. *Let $\mathcal{D}$ be a mixed distribution on $[q]^I \times [q]^I$. For any $i \neq \ell$, we have that,*

$$\gamma_\ell\left(\Pi_i^A(\mathcal{D})\right) \leq \gamma_\ell(\mathcal{D}).$$

Similarly, it also holds that $\gamma_\ell\left(\Pi_i^B(\mathcal{D})\right) \leq \gamma_\ell(\mathcal{D})$.

This lemma is proven by a standard convexity inequality. We defer the proof to the appendix.

Proof of Theorem 4.2. Our aim is to prove $\log \frac{1}{\mu(D_0 \cap R)} \geq \Omega(\mathrm{bs}(f) \cdot \log q)$. Let q and ϵ be constants from Lemma 4.9. Let $\epsilon' = \epsilon/10$. We set ϵ' as the constant for Theorem 4.2. We recall several facts:

- By Lemma 4.5, for the uniform distribution $\mathcal{D}$ on R, $-E(\mathcal{D}) \leq \log \frac{1}{\mu(D_0 \cap R)}$.
- By Lemma 4.5, $\sum_i \gamma_i(\mathcal{D}) \leq \epsilon' \cdot \mathrm{bs}(f)$.
- For any mixed distribution $\mathcal{D}$, $E(\mathcal{D}) \leq 0$.

Let $P = \{i \in [n] : \gamma_i(\mathcal{D}) \leq 10 \cdot \epsilon' = \epsilon\}$. Then by an average argument, we have that $|P| \geq n/2$ (recall that we assumed $\mathrm{bs}(f) = n$). Now we apply projections on all coordinates in P. By combining Lemma 4.10 and Lemma 4.10, each round of the projection increases the potential function by $c \cdot \log q = \Omega(\log q)$. Hence we have that $-E(\mathcal{D}) \geq -\Omega(\mathrm{bs}(f) \cdot \log q)$. We then conclude the proof since $-E(\mathcal{D}) \leq \log \frac{1}{\mu(D_0 \cap R)}$. □

4.2 Proof of Projection Lemma

We first prove the projection lemma for production distributions. In this section, we reuse some notations to make our writing clean. We use X, Y to denote distributions on $[q]^I$. We fix parameters as $\epsilon = 9 \cdot 10^{-6}$,$\epsilon_1 = 10^{-5}$, $\lambda = 1 - 10^{-4}$ and $q = 2^{10^{10}}$ in the rest of this section. We believe these constants can be improved significantly. However, our main focus in this paper is to explain a clean proof with constant-size gadgets.

There are two steps in our proof. Firstly, we show that bias implies entropy loss. Concretely, we prove the following lemma.

Lemma 4.11. *Let $\mathcal{D} = X \times Y$ be a product distribution on $[q]^I \times [q]^I$ and denote $(X'_I, Y'_I) = \mathcal{D}|_{g(X,Y)=0^I}$. If $\gamma_i(\mathcal{D}) \leq \epsilon_1$ for some $i \in I$, one of the following statements must be true:*

- $\mathcal{H}\left(X'_i \mid \left(X'_{I\setminus i}, Y'_{I\setminus i}\right)\right) \leq \lambda \cdot \log q$,
- $\mathcal{H}\left(Y'_i \mid \left(X'_{I\setminus i}, Y'_{I\setminus i}\right)\right) \leq \lambda \cdot \log q$.

We prove this lemma by contradiction. If both X'_i and Y'_i have extremely high conditional entropy, they also have high min-entropy. By an extractor-like property, $g(X'_i, Y'_i)$ should close to uniform (even after fixing other coordinates), which contradicts to the fact that $\gamma_i(\mathcal{D}) \leq \epsilon_1$. Here we use the fact that a low-discrepancy gadget g has an extractor-like property. A similar extractor-like property was used in many previous papers such as [GLM$^+$16, CFK$^+$19]. The difference is that they need a block-wise extractor for distributions with high block-wise min-entropy. It is unlikely to construct a constant-size gadget with block-wise extractor property. By contrast, we study one-block extractors for distributions with high conditional entropy. Hence we avoided the large-size gadget barrier. We defer the proof of Lemma 4.11 to the appendix.

Now we prove projection lemma for those coordinates with entropy loss.

Lemma 4.12. *Let $\mathcal{D} = X \times Y$ be a product distribution and let $(X'_I, Y'_I) = \mathcal{D}|_{g^I(X,Y)=0^I}$. For any $\lambda > 0$ and $q > 2$, if $\mathcal{H}(Y'_i \mid (X'_{I\setminus i}, Y'_{I\setminus i})) \leq \lambda \cdot \log q$, the projection by Alice at coordinate i increases potential function, i.e.,*

$$E\left(\Pi_i^A(\mathcal{D})\right) \geq E(\mathcal{D}) + (1-\lambda)\cdot \log q - 1.$$

Similarly, if $\mathcal{H}(X'_i \mid (X'_{I\setminus i}, Y'_{I\setminus i})) \leq \lambda \cdot \log q$, the projection by Bob at coordinate i increases potential function,

$$E(\Pi_i^B(\mathcal{D})) \geq E(\mathcal{D}) + (1-\lambda)\cdot \log q - 1.$$

Proof. From the definition, we have that $\Pi_i^A(\mathcal{D}) = \sum_u p_u \cdot \mathcal{D}_u$, where $\mathcal{D}_u = X_u \times Y_u$ is a product distribution on $[q]^{I\setminus\{i\}} \times [q]^{I\setminus\{i\}}$. By chain rule,

$$\mathcal{H}(X'_I, Y'_I) = \mathcal{H}(X'_i) + \mathcal{H}(X'_{I\setminus i}, Y'_{I\setminus i} | X'_i) + \mathcal{H}(Y'_i | X'_I, Y'_{I\setminus i}).$$

Denote $(X'_u, Y'_u) = (X_u, Y_u)\ |_{g^{I\setminus i}(X_u,Y_u)=0^{I\setminus\{i\}}}$. We observe that $(X'_{I\setminus i}, Y'_{I\setminus i} | X'_i = u) = (X'_u, Y'_u)$. It implies that,

$$\mathcal{H}(X'_I, Y'_I) = \mathcal{H}(X'_i) + \sum p_u \cdot \mathcal{H}(X'_u, Y'_u) + \mathcal{H}(Y'_i | X'_I, Y'_{I\setminus i}).$$

By using Fact 2.3, we have that

$$\mathcal{H}(X'_I, Y'_I) \leq \mathcal{H}(X'_i) + \sum p_u \cdot \mathcal{H}(X'_u, Y'_u) + \mathcal{H}(Y'_i | X'_{I\setminus i}, Y'_{I\setminus i}).$$

From the assumption of $\mathcal{H}(Y'_i | X'_{I\setminus i}, Y'_{I\setminus i}) \leq \lambda \cdot \log q$, it concludes that,

$$\mathcal{H}(X'_I, Y'_I) \leq \log q + \sum p_u \cdot \mathcal{H}(X'_u, Y'_u) + \lambda \cdot \log q.$$

Recall the definition of potential function, then we have

$$E(\Pi_i^A(\mathcal{D})) = \sum p_u \cdot E(\mathcal{D}_u) \geq E(\mathcal{D}) + (1-\lambda)\cdot \log q - 1.$$

The claim then follows. □

Build on Lemma 4.11 and Lemma 4.12, we are ready to prove the projection lemma.

Proof of Lemma 4.9. Let $\mathcal{D} = \sum_{j\in[m]} p_j \cdot \mathcal{D}_j$ be a mixed distribution with $\gamma_i(\mathcal{D}) \leq \epsilon$. Let

$$J = \{j \in [m] : \gamma_i(\mathcal{D}_j) \leq \epsilon/0.9 = \epsilon_1\}.$$

By the average argument, we have that $\sum_{j\in J} p_j \geq 0.1$.

For each $j \in J$, $\mathcal{D}_j$ is a product distribution with $\gamma_i(\mathcal{D}_j) \leq \epsilon_1$. By Lemma 4.11 and Lemma 4.12, one of the following statements must hold:

$$E(\Pi_i^A(\mathcal{D}_j)) \geq E(\mathcal{D}_j) + (1-\lambda) \cdot \log q - 1$$

or

$$E(\Pi_i^B(\mathcal{D}_j)) \geq E(\mathcal{D}_j) + (1-\lambda) \cdot \log q - 1.$$

Let $J_A \subseteq J$ be the set of distributions $\mathcal{D}_j$ with $E(\Pi_i^A(\mathcal{D}_j)) \geq E(\mathcal{D}_j) + (1-\lambda) \cdot \log q - 1$.. Without loss of generality, we assume that $\sum_{j \in J_A} p_j \geq (\sum_{j \in J} p_j)/2 \geq 0.05$.
On the other hand, for any $j \notin J_A$, we apply Lemma 4.12 with $\lambda = 1$, then

$$E(\Pi_i^A(\mathcal{D}_j)) \geq E(\mathcal{D}_j) - 1.$$

In sum,

$$E(\Pi_i^A(\mathcal{D})) \geq \sum_{j \in J_A} p_j \left(E(\mathcal{D}_j) + (1-\lambda)\log q - 2)\right) + \sum_{j \notin J_A} p_j \left(E(\mathcal{D}_j) - 1)\right).$$

Now we set $c = (1-\lambda) \cdot 0.05 - \frac{1}{\log q}$ and derive

$$E(\Pi_i^A(\mathcal{D})) \geq E(\mathcal{D}) + c \cdot \log q.$$

□

5 Discussion and Open problems

5.1 New direct sum theorems

Lifting theorems can be viewed as a generalization of direct sum theorems [CFK+19]. In the setting of randomized communication complexity, it is known that the "ability of g to admit a direct sum theorem" is characterized exactly by the information cost of g (denoted $\mathsf{IC}(g)$) [BYJKS04]. This leads to the natural conjecture that a lifting theorem should hold for every gadget g that has sufficiently high information cost. [CFK+19] proposed the following conjecture for large gadget size,

Conjecture 5.1 ([CFK+19])**.** *Let $g : [q] \times [q] \to \{0,1\}$ be a two-party function, where $q = n^C$ for some constant $C > 0$. Then for any Boolean function $f : \{0,1\}^n \to \{0,1\}$,*

$$\mathbf{BPP}^{\mathrm{cc}}(f \circ g) = \Omega\left(\mathbf{BPP}^{\mathrm{dt}}(f) \cdot \mathsf{IC}(g)\right).$$

Similar to above conjecture, we propose the following conjecture for constant gadget size:

Conjecture 5.2. *There is a constant $q > 0$. For any Boolean function $f : \{0,1\}^n \to \{0,1\}$ and gadget $g : [q] \times [q] \to \{0,1\}$, we have that*

$$\mathsf{IC}^{\mathrm{ext}}(f \circ g^n) = \Omega(\mathrm{bs}(f) \cdot \mathsf{IC}(g)).$$

We note that above conjecture is a generalization of the direct sum argument in [BYJKS04].

5.2 The corruption bound and log-rank conjecture

In order to prove log rank conjecture [Lov14], Adi Shraibman [Shr19] showed that for any function f, there is a constant ϵ such that

$$\mathbf{P}^{\mathrm{cc}}(f) \leq \mathrm{Corr}_{\epsilon}(f) \cdot \log^2 \mathrm{rank}(f).$$

This result is similar to the result $\mathbf{P}^{\mathrm{dt}}(f) \leq \deg^2(f)\mathrm{bs}(f)$ [NS94] in query complexity.

However, for certain block-composed functions $f \circ g^n$ (where f is arbitrary but the gadget g is chosen carefully), in degree-to-rank lifting [She11], which shows that

$$\log \mathrm{rank}(f \circ g^n) = \Omega(\deg(f))$$

Remembering our result, that is

$$\mathrm{Corr}_{\epsilon}(f \circ g^n) = \Omega(\mathrm{bs}(f))$$

We can chose the gadget g carefully to lift the result $\mathbf{P}^{\mathrm{dt}}(f) \leq \deg(f)\mathrm{bs}(f)$ [Mid04] in query complexity to communication complexity. Thus, there is a constant gadget g such that

$$\mathbf{P}^{\mathrm{cc}}(f \circ g^n) \leq \mathbf{P}^{\mathrm{dt}}(f) \leq \deg(f) \cdot \mathrm{bs}(f) \leq \mathrm{Corr}_{\epsilon}(f \circ g^n) \cdot \log \mathrm{rank}(f \circ g^n).$$

Therefore, it's natural to ask the question: Can we mimic the query protocol in [Mid04] to get a communication protocol and prove the following conjecture?

Conjecture 5.3. *For any Boolean function $f : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$, there is a constant ϵ such that*

$$\mathbf{P}^{\mathrm{cc}}(f) \leq \mathrm{Corr}_{\epsilon}(f) \cdot \log \mathit{rank}(f).$$

We note this conjecture, if true, will improve the results in [Shr19].

References

[ABDK20] Anurag Anshu, Shalev Ben-David, and Srijita Kundu. On query-to-communication lifting for adversary bounds. *arXiv preprint arXiv:2012.03415*, 2020.

[ALWZ21] Ryan Alweiss, Shachar Lovett, Kewen Wu, and Jiapeng Zhang. Improved bounds for the sunflower lemma. *Annals of Mathematics*, 194(3):795–815, 2021.

[BPSW05] Paul Beame, Toniann Pitassi, Nathan Segerlind, and Avi Wigderson. A direct sum theorem for corruption and the multiparty nof communication complexity of set disjointness. In *20th Annual IEEE Conference on Computational Complexity (CCC'05)*, pages 52–66. IEEE, 2005.

[BR20] Yakov Babichenko and Aviad Rubinstein. Communication complexity of nash equilibrium in potential games. In *2020 IEEE 61st Annual Symposium on Foundations of Computer Science (FOCS)*, pages 1439–1445. IEEE, 2020.

[BYJKS04] 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.

[CFK+19] Arkadev Chattopadhyay, Yuval Filmus, Sajin Koroth, Or Meir, and Toniann Pitassi. Query-to-communication lifting using low-discrepancy gadgets. *arXiv preprint arXiv:1904.13056*, 2019.

[CKLM18] Arkadev Chattopadhyay, Michal Kouckỳ, Bruno Loff, and Sagnik Mukhopadhyay. Simulation beats richness: New data-structure lower bounds. In *Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing*, pages 1013–1020, 2018.

[CKLM19] Arkadev Chattopadhyay, Michal Kouckỳ, Bruno Loff, and Sagnik Mukhopadhyay. Simulation theorems via pseudo-random properties. *computational complexity*, 28(4):617–659, 2019.

[CLRS16] Siu On Chan, James R Lee, Prasad Raghavendra, and David Steurer. Approximate constraint satisfaction requires large lp relaxations. *Journal of the ACM (JACM)*, 63(4):1–22, 2016.

[CMS20] Arkadev Chattopadhyay, Nikhil S Mande, and Suhail Sherif. The log-approximate-rank conjecture is false. *Journal of the ACM (JACM)*, 67(4):1–28, 2020.

[dR19] Susanna F de Rezende. *Lower Bounds and Trade-offs in Proof Complexity*. PhD thesis, KTH Royal Institute of Technology, 2019.

[DRNV16] Susanna F De Rezende, Jakob Nordström, and Marc Vinyals. How limited interaction hinders real communication (and what it means for proof and circuit complexity). In *2016 IEEE 57th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 295–304. IEEE, 2016.

[GGKS18] Ankit Garg, Mika Göös, Pritish Kamath, and Dmitry Sokolov. Monotone circuit lower bounds from resolution. In *Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing*, pages 902–911, 2018.

[GJPW18] Mika Göös, TS Jayram, Toniann Pitassi, and Thomas Watson. Randomized communication versus partition number. *ACM Transactions on Computation Theory (TOCT)*, 10(1):1–20, 2018.

[GJW18] Mika Göös, Rahul Jain, and Thomas Watson. Extension complexity of independent set polytopes. *SIAM Journal on Computing*, 47(1):241–269, 2018.

[GKPW19] Mika Göös, Pritish Kamath, Toniann Pitassi, and Thomas Watson. Query-to-communication lifting for p np. *computational complexity*, 28(1):113–144, 2019.

[GLM+16] Mika Göös, Shachar Lovett, Raghu Meka, Thomas Watson, and David Zuckerman. Rectangles are nonnegative juntas. *SIAM J. Comput.*, 45(5):1835–1869, 2016.

[Göö15] Mika Göös. Lower bounds for clique vs. independent set. In *2015 IEEE 56th Annual Symposium on Foundations of Computer Science*, pages 1066–1076. IEEE, 2015.

[GP18] Mika Göös and Toniann Pitassi. Communication lower bounds via critical block sensitivity. *SIAM Journal on Computing*, 47(5):1778–1806, 2018.

[GPW18] Mika Göös, Toniann Pitassi, and Thomas Watson. Deterministic communication vs. partition number. *SIAM Journal on Computing*, 47(6):2435–2450, 2018.

[GPW20] Mika Göös, Toniann Pitassi, and Thomas Watson. Query-to-communication lifting for bpp. *SIAM Journal on Computing*, 49(4):FOCS17–441, 2020.

[GR21] Mika Göös and Aviad Rubinstein. Near-optimal communication lower bounds for approximate nash equilibria. *SIAM Journal on Computing*, pages FOCS18–316, 2021.

[HN12] Trinh Huynh and Jakob Nordström. On the virtue of succinct proofs: Amplifying communication complexity hardness to time-space trade-offs in proof complexity. In *Proceedings of the forty-fourth annual ACM symposium on Theory of computing*, pages 233–248, 2012.

[KLL+12] Iordanis Kerenidis, Sophie Laplante, Virginie Lerays, Jeremie Roland, and David Xiao. Lower bounds on information complexity via zero-communication protocols and applications. *2012 IEEE 53rd Annual Symposium on Foundations of Computer Science*, Oct 2012.

[KMR21] Pravesh K Kothari, Raghu Meka, and Prasad Raghavendra. Approximating rectangles by juntas and weakly exponential lower bounds for lp relaxations of csps. *SIAM Journal on Computing*, pages STOC17–305, 2021.

[Kus97] Eyal Kushilevitz. Communication complexity. In *Advances in Computers*, volume 44, pages 331–360. Elsevier, 1997.

[LM19] Bruno Loff and Sagnik Mukhopadhyay. Lifting theorems for equality. In *36th International Symposium on Theoretical Aspects of Computer Science (STACS 2019)*. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, 2019.

[LMM+20] Shachar Lovett, Raghu Meka, Ian Mertz, Toniann Pitassi, and Jiapeng Zhang. Lifting with sunflowers. In *Electron. Colloquium Comput. Complex*, page 111, 2020.

[Lov14] Shachar Lovett. Recent advances on the log-rank conjecture in communication complexity. *arXiv preprint arXiv:1403.8106*, 2014.

[LRS15] James R Lee, Prasad Raghavendra, and David Steurer. Lower bounds on the size of semidefinite programming relaxations. In *Proceedings of the forty-seventh annual ACM symposium on Theory of computing*, pages 567–576, 2015.

[Mid04] Gatis Midrijanis. Exact quantum query complexity for total boolean functions. *arXiv preprint quant-ph/0403168*, 2004.

[NS94] Noam Nisan and Mario Szegedy. On the degree of boolean functions as real polynomials. *Computational complexity*, 4(4):301–313, 1994.

[PR17] Toniann Pitassi and Robert Robere. Strongly exponential lower bounds for monotone computation. In *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing, STOC 2017, Montreal, QC, Canada, June 19-23, 2017*, pages 1246–1255. ACM, 2017.

[PR18] Toniann Pitassi and Robert Robere. Lifting nullstellensatz to monotone span programs over any field. In *Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing, STOC 2018, Los Angeles, CA, USA, June 25-29, 2018*, pages 1207–1219. ACM, 2018.

[PSW21] Toniann Pitassi, Morgan Shirley, and Thomas Watson. Nondeterministic and randomized boolean hierarchies in communication complexity. *computational complexity*, 30(2):1–48, 2021.

[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.

[RPRC16] Robert Robere, Toniann Pitassi, Benjamin Rossman, and Stephen A. Cook. Exponential lower bounds for monotone span programs. In *Proceedings of the 57th Symposium on Foundations of Computer Science (FOCS)*, pages 406–415. IEEE Computer Society, 2016.

[She11] Alexander A Sherstov. The pattern matrix method. *SIAM Journal on Computing*, 40(6):1969–2000, 2011.

[Shr19] Adi Shraibman. The corruption bound, log-rank, and communication complexity. *Information Processing Letters*, 141:16–21, 2019.

[WYY17] Xiaodi Wu, Penghui Yao, and Henry S Yuen. Raz-mckenzie simulation with the inner product gadget. In *Electron. Colloquium Comput. Complex.*, volume 24, page 10, 2017.

[Zha09] Shengyu Zhang. On the tightness of the buhrman-cleve-wigderson simulation. In *International Symposium on Algorithms and Computation*, pages 434–440. Springer, 2009.

A Missing proofs

A.1 Proof of Lemma 4.10

We quickly recall the statement of this lemma. Let $\mathcal{D}$ be a mixed distribution on $[q]^I \times [q]^I$. For any $i \neq \ell$, $\gamma_\ell(\mathcal{D}) \geq \gamma_\ell\left(\Pi_i^A(\mathcal{D})\right)$. Similarly, it also holds that $\gamma_\ell(\mathcal{D}) \geq \gamma_\ell\left(\Pi_i^B(\mathcal{D})\right)$.

We first prove the following lemma:

Lemma A.1. *For any $x_1, ..., x_n \geq 0$ and $y_1, ..., y_n \geq 0$.*

$$\frac{1}{\sum_{j=1}^n y_j} \cdot \sum_{j=1}^n \frac{y_j \cdot x_j}{x_j + y_j} \leq \frac{\sum_{j=1}^n x_j}{\sum_{j=1}^n (x_j + y_j)}.$$

Proof. We prove this lemma by induction. We first prove the base case ($n = 2$),

$$\frac{y_1}{y_1 + y_2} \cdot \frac{x_1}{x_1 + y_1} + \frac{y_2}{y_1 + y_2} \cdot \frac{x_2}{x_2 + y_2} \leq \frac{x_1 + x_2}{x_1 + x_2 + y_1 + y_2}.$$

Let f be a convex function defined as $f(x) = \frac{x}{x+1}$. By Jensen inequality, we have

$$\frac{y_1}{y_1+y_2} \cdot f\left(\frac{x_1}{y_1}\right) + \frac{y_2}{y_1+y_2} \cdot f\left(\frac{x_2}{y_2}\right) \le f\left(\frac{y_1}{y_1+y_2} \cdot \frac{x_1}{y_1} + \frac{y_2}{y_1+y_2} \cdot \frac{x_2}{y_2}\right) \le f\left(\frac{x_1+x_2}{y_1+y_2}\right).$$

We then finish the proof by observing that $f(x_1/y_1) = x_1/(x_1+y_1)$, $f(x_2/y_2) = x_2/(x_2+y_2)$ and $f((x_1+x_2)/(y_1+y_2)) = (x_1+x_2)/(x_1+x_2+y_1+y_2)$.

Now we prove the induction part. Assume this statement holds for $n = k-1$, we prove the case of $n = k$. Let $x_1, \ldots, x_k$ and $y_1, \ldots, y_n$ be the numbers. We first apply the base case on x_{k-1}, x_k and y_{k-1}, y_k,

$$\frac{y_{k-1}}{y_{k-1}+y_k} \cdot \frac{x_{k-1}}{x_{k-1}+y_{k-1}} + \frac{y_k}{y_{k-1}+y_k} \cdot \frac{x_k}{x_k+y_k} \le \frac{x_{k-1}+x_k}{x_{k-1}+y_{k-1}+x_k+y_k}.$$

By multiplying each term by $(y_{k-1}+y_k)$, we have

$$\frac{x_{k-1} \cdot y_{k-1}}{x_{k-1}+y_{k-1}} + \frac{x_k \cdot y_k}{x_k+y_k} \le \frac{(x_{k-1}+x_k)(y_{k-1}+y_k)}{x_{k-1}+y_{k-1}+x_k+y_k}.$$

Set $y'_{k-1} = y_{k-1}+y_k$, $x'_{k-1} = x_{k-1}+x_k$ and $x'_i = x_i$, $y'_i = y_i$ for $i < k-1$, we have that

$$\frac{1}{\sum_{j=1}^k y_j} \cdot \sum_{j=1}^k \frac{y_j \cdot x_j}{x_j+y_j} = \frac{1}{\sum_{j=1}^k y_j} \cdot \left(\sum_{j=1}^{k-2} \frac{y_j \cdot x_j}{x_j+y_j} + \frac{y_{k-1} \cdot x_{k-1}}{x_{k-1}+y_{k-1}} + \frac{y_k \cdot x_k}{x_k+y_k}\right) \le \frac{1}{\sum_{j=1}^{k-1} y'_j} \cdot \sum_{j=1}^{k-1} \frac{y'_j \cdot x'_j}{x'_j+y'_j}.$$

By applying hypothesis on $x'_1, ..., x'_{k-1}$ and $y'_1, ..., y'_{k-1}$, we have that

$$\frac{1}{\sum_{j=1}^{k-1} y'_j} \cdot \sum_{j=1}^{k-1} \frac{y'_j \cdot x'_j}{x'_j+y'_j} \le \frac{\sum_{j=1}^{k-1} x'_j}{\sum_{j=1}^{k-1}(x'_j+y'_j)} = \frac{\sum_{j=1}^k x_j}{\sum_{j=1}^k (x_j+y_j)}.$$

Thus, the case $n = k$ is holds. □

Now we are ready to prove Lemma 4.10.

Proof. We first prove it for product distributions. Recall that $\Pi_i^A(\mathcal{D}) = \sum_{u \in [q]} p_u \cdot \mathcal{D}_u$. For each $u \in [q]$, let

$$r_u := \Pr_{(x,y) \sim \mathcal{D}_u}\left[(x,y) \in D_0^{I \setminus \{i\}}\right].$$

For $\ell \neq i$, let

$$s_u := \Pr_{(x,y) \sim \mathcal{D}_u}\left[(x,y) \in D_\ell^{I \setminus \{i\}}\right].$$

We recall connections between r_u, s_u and p_u,

$$p_u = \frac{r_u}{\sum r_u}, \qquad \gamma_\ell(\mathcal{D}_u) = \frac{s_u}{s_u+r_u}, \qquad \gamma_\ell(\mathcal{D}) = \frac{\sum s_u}{\sum (s_u+r_u)}.$$

By Lemma A.1, we have

$$\gamma_\ell(\Pi_i^A(\mathcal{D})) = \sum_u p_u \cdot \gamma_\ell(\mathcal{D}_u) = \frac{1}{\sum r_u} \cdot \sum_u \frac{r_u \cdot s_u}{s_u+r_u} \le \frac{\sum s_u}{\sum s_u + r_u} = \gamma_\ell(\mathcal{D}).$$

The claim then follows for production distributions. For a mixed distribution $\mathcal{D} = \sum_u p_j \cdot \mathcal{D}_j$, we apply the above inequality for each $\mathcal{D}_j$,

$$\gamma_\ell(\Pi_i^A(\mathcal{D})) = \gamma_\ell\left(\sum_j p_j \cdot \Pi_i^A(\mathcal{D}_j)\right) = \sum_j p_j \cdot \gamma_\ell(\Pi_i^A(\mathcal{D}_j)) \le \sum_j p_j \cdot \gamma_\ell(\mathcal{D}_j) = \gamma_\ell(\mathcal{D}).$$

□

A.2 Proof of Lemma 4.11

We recall parameters $\epsilon = 9 \cdot 10^{-6}$, $\epsilon_1 = 10^{-5}$, $\delta = 0.2$, $\lambda = 1 - 10^{-4}$, $\lambda' = 1 - 10^{-3}$, $\eta = 10^{-4}$ and $q = 2^{10^{10}}$. We restate Lemma 4.11.

Lemma A.2 (Restated)**.** *Let $\mathcal{D} = X \times Y$ be a product distribution on $[q]^I \times [q]^I$ and denote $(X'_I, Y'_I) = \mathcal{D}|_{g^I(X,Y)=0^I}$. If $\gamma_i(\mathcal{D}) \le \epsilon_1$ for some $i \in I$, one of the following statements must be true:*

- $\mathcal{H}\left(X'_i \mid \left(X'_{I\setminus i}, Y'_{I\setminus i}\right)\right) \le \lambda \cdot \log q$,
- $\mathcal{H}\left(Y'_i \mid \left(X'_{I\setminus i}, Y'_{I\setminus i}\right)\right) \le \lambda \cdot \log q$.

In order to show entropy loss on the distribution $\mathcal{D}$ is large when $\gamma_i(\mathcal{D})$ is small (Lemma 4.11), we first show entropy loss for gadget function. Let $g : [q] \times [q] \to \{0,1\}$ be the gadget function with discrepancy at most $2^{-\log q}$. Such gadget function g satisfies the following "extractor-like" property.

Lemma A.3. *Let $\mathcal{D} = X \times Y$ be a product distribution on $[q]^I \times [q]^I$ and denote $(X'_I, Y'_I) = \mathcal{D}|_{g(X,Y)=0^I}$. For a fixed $(x', y') \in D_0^{I\setminus\{i\}}$, we define the X_i, Y_i, X'_i, Y'_i:*

$$X_i = X\,|_{X_{I\setminus\{i\}}=x'} \quad \text{and } Y_i = Y\,|_{Y_{I\setminus\{i\}}=y'}$$

and let

$$X'_i = X'_I\,|_{(X'_{I\setminus\{i\}}, Y'_{I\setminus\{i\}})=(x',y')} \quad \text{and } Y'_i = Y'_I\,|_{(X'_{I\setminus\{i\}}, Y'_{I\setminus\{i\}})=(x',y')}\,.$$

If $\mathcal{H}(X'_i) \ge \lambda' \cdot \log q$, $\mathcal{H}(Y'_i) \ge \lambda' \cdot \log q$, then

$$\Pr[g(X_i, Y_i) = 1] \ge \eta \cdot \Pr[g(X_i, Y_i) = 0].$$

The proof of Lemma A.3 is a standard technique in previous lifting theorems [CFK+19]. Note that X'_i and Y'_i are independent because $\mathcal{D}$ is a product distribution. There is only a slight difference between our lemma and the standard techniques. In previous proofs, they require that both X'_i and Y'_i have large min-entropy, and then apply an extractor property. Here, we only requires that X'_i and Y'_i have large Shannon entropy. The cost of using Shannon entropy is that, we have to choose the entropy-loss parameter $\lambda' = 1 - \Omega(1)$, which also keeps the extractor property.

Given a Boolean random variable D,we denote the bias of D by $\mathrm{bias}(D) = |\Pr[D = 0] - \Pr[D = 1]|$. We use the following lemma in [CFK+19].

Lemma A.4 ([CFK+19])**.** *Let $g : [q] \times [q] \to \{0,1\}$ be a function with discrepancy at most $2^{-\log q}$, Let X, Y be independent random variables taking values in $[q]$ such that $H_\infty(X)+H_\infty(Y) \geq 1.2 \cdot \log q$ then*

$$bias(g(X,Y)) \leq q^{-0.2}.$$

Claim A.5. *Let X be a distribution on $[q]$. If $\mathcal{H}(X) \geq 0.998 \cdot \log q$, then there is a distribution X' such that*

- $||X - X'||_{TV} \leq 0.01$,
- $H_\infty(X') \geq 0.6 \cdot \log q$,

Proof. Since $\mathcal{H}(X) \geq \lambda_0 \cdot \log q$, By Markov inequality, we have that

$$\Pr_x \left[\log \frac{1}{\Pr[X = x]} \leq 0.61 \cdot \log q \right] \leq 2.5 \cdot 0.002 \leq 0.01.$$

Let $T = \{x \in [q] : \log \frac{1}{\Pr[X=x]} \leq 0.61 \log q\}$, we get X' by the following way:

1. For each $x \in T$, $\Pr[X' = x] = 0$.
2. For each $x \in [q] \setminus T$, $\Pr[X' = x] = \frac{\Pr[X=x]}{1-t} \leq \frac{1}{1-0.005} \cdot q^{-0.61} \leq q^{-0.6}$.

We note that $||X - X'||_{\mathrm{TV}} \leq 0.01$ and $H_\infty(X') \geq 0.6 \cdot \log q$. □

Lemma A.6. *Let X, Y be independent random variable on $[q]$. If $\mathcal{H}(X) \geq 0.998 \cdot \log q$ and $\mathcal{H}(Y) \geq 0.998 \cdot \log q$, then there is a constant $\eta' \geq 0.1$ such that*

$$\Pr[g(X,Y) = 1] \geq \eta' \cdot \Pr[g(X,Y) = 0],$$

Proof. By Claim A.5, there is a independent variable X' and Y' such that

$$H_\infty(X') \geq 0.6 \cdot \log q \text{ and } H_\infty(Y') \geq 0.6 \cdot \log q$$

and

$$||X - X'||_{\mathrm{TV}} \leq 0.01 \text{ and } ||Y - Y'||_{\mathrm{TV}} \leq 0.01.$$

By Lemma A.4, we have

$$\mathrm{bias}(g(X,Y)) \leq \mathrm{bias}(g(X',Y')) + 0.04 \leq q^{-0.2} + 0.04 \leq 0.05.$$

Setting $\eta' = 0.1$, we have

$$\Pr[g(X,Y) = 1] \geq \eta' \cdot \Pr[g(X,Y) = 0].$$

□

Now we are ready to prove Lemma A.3.

Proof of Lemma A.3. Assume $\Pr[g(X_i, Y_i) = 0] \geq 0.9999$. By the definition of $\mathcal{D}$ and $\mathcal{D}'$, we have

$$\mathcal{H}(X_i') = \mathcal{H}(X_i | g(X_i, Y_i) = 0) \leq \frac{\mathcal{H}(X_i | g(X_i, Y_i))}{\Pr[g(X_i, Y_i) = 0]} \leq \frac{\mathcal{H}(X_i)}{\Pr[g(X_i, Y_i) = 0]}.$$

Since $\Pr[g(X_i, Y_i) = 0] \geq 0.9999$, we have

$$\mathcal{H}(X_i) \geq 0.9999 \cdot \mathcal{H}(X_i') \geq 0.998 \cdot \log q.$$

Similarly,

$$\mathcal{H}(Y_i) \geq 0.9999 \cdot \mathcal{H}(Y_i') \geq 0.998 \cdot \log q.$$

Since $\mathcal{D}$ is product distribution, X_i and Y_i are independent random variables, By Lemma A.6, we have

$$\Pr[g(X_i, Y_i) = 1] \geq 0.1 \cdot \Pr[g(X_i, Y_i) = 0].$$

If $\Pr[g(X_i, Y_i) = 0] \leq 0.9999$, then

$$\Pr[g(X_i, Y_i) = 1] \geq 10^{-4} \cdot \Pr[g(X_i, Y_i) = 0].$$

□

We note that if we fix coordinates in $I \backslash i$, we only need to deal with gadget function in coordinate i. Since $\gamma_i(\mathcal{D})$ is small, by Lemma A.3, we have entropy loss either in Alice's side or in Bob's side. Now we are ready to prove Lemma 4.11.

Proof of Lemma 4.11. We prove the statement by contradiction. We assume that $\mathcal{H}(X_i' | (X_{I\backslash i}', Y_{I\backslash i}')) \geq \lambda \cdot \log q$ and $\mathcal{H}(Y_i' | (X_{I\backslash i}', Y_{I\backslash i}')) \geq \lambda \cdot \log q$. Fix $(x', y') \in D_0^{I\backslash i}$, we use X_i' as a shorthand for $X_i' \mid_{(X_{I\backslash\{i\}}', Y_{I\backslash i}') = (x_{I\backslash i}', y_{I\backslash i}')}$ and Y_i' as a shorthand for $Y_i' \mid_{(X_{I\backslash\{i\}}', Y_{I\backslash i}') = (x_{I\backslash i}', y_{I\backslash i}')}$. Recall that $\lambda' = 1 - 10^{-3}$, $\lambda = 1 - 10^{-4}$ and $\delta = 2$.

If $\mathcal{H}(X_i' | (X_{I\backslash i}', Y_{I\backslash i}')) \geq \lambda \cdot \log q$, then by an average argument,

$$\Pr_{(x', y') \sim (X_{I\backslash i}', Y_{I\backslash i}')} [\mathcal{H}(X_i')) \geq \lambda' \cdot \log q] \geq 1 - \delta.$$

Similarly, if $\mathcal{H}(Y_i' | (X_{I\backslash i}', Y_{I\backslash i}')) \geq \lambda \cdot \log q$, we have

$$\Pr_{(x', y') \sim (X_{I\backslash i}', Y_{I\backslash i}')} [\mathcal{H}(Y_i') \geq \lambda' \cdot \log q] \geq 1 - \delta.$$

By union bound, we have

$$\Pr_{(x', y') \sim (X_{I\backslash i}', Y_{I\backslash i}')} [\mathcal{H}(X_i') \geq \lambda' \cdot \log q \text{ and } \mathcal{H}(Y_i') \geq \lambda' \cdot \log q] \geq 1 - 2\delta.$$

By Lemma A.3, if $(x', y') \sim (X_{I\backslash i}', Y_{I\backslash i}')$, let $X_i = X \mid_{X_{I\backslash\{i\}} = x'}$ and $Y_i = Y \mid_{Y_{I\backslash\{i\}} = y'}$, then with probability $1 - 2\delta$, we have

$$\Pr[g(X_i, Y_i) = 1] \geq \eta \cdot \Pr[g(X_i, Y_i) = 0].$$

Recall the definition of $\gamma_i(\mathcal{D})$, then we have

$$\gamma_i(\mathcal{D}) \geq \frac{\eta}{2} \cdot (1 - 2\delta).$$

Setting $\epsilon < \frac{\eta}{2} \cdot (1 - 2\delta) \leq 3 \cdot 10^{-4}$ contradicts the fact that $\gamma_i(\mathcal{D}) \leq \epsilon$. □

ECCC ISSN 1433-8092

https://eccc.weizmann.ac.il