Electronic Colloquium on Computational Complexity, Report No. 10 (2017)

Raz-McKenzie simulation with the inner product gadget

Xiaodi Wu* Penghui Yao† Henry Yuen‡

Abstract

In this note we show that the Raz-McKenzie simulation algorithm which lifts deterministic query lower bounds to deterministic communication lower bounds can be implemented for functions f composed with the Inner Product gadget $g_{\text{IP}}(x, y) = \sum_i x_i y_i \mod 2$ of logarithmic size. In other words, given a function $f : \{0,1\}^n \to \{0,1\}$ with deterministic query complexity $\mathsf{D}(f)$, we show that the deterministic communication complexity of the composed function $f \circ g_{\text{IP}}^n$ is $\Theta(\mathsf{D}(f) \log n)$, where

$$f \circ g_{\text{IP}}^n(x, y) = f(g_{\text{IP}}(x^1, y^1), \ldots, g_{\text{IP}}(x^n, y^n))$$

where $x = (x^1, \ldots, x^n)$, $y = (y^1, \ldots, y^n)$ and each x^i and y^i are $O(\log n)$ bit strings. In [RM97] and [GPW15], the simulation algorithm is implemented for functions composed with the Indexing gadget, where the size of the gadget is *polynomial* in the input length of the outer function f.

1 Introduction

Let $f : \{0,1\}^n \to \{0,1\}$ be a function. Its deterministic query complexity $\mathsf{D}(f)$ (also known as its decision tree complexity) is the minimum number of queries to an unknown input $z \in \{0,1\}^n$ that an algorithm must make in order to compute $f(z)$. In [RM97] and [GPW15], it is shown that there is a *gadget* $g : \{0,1\}^{b_1} \times \{0,1\}^{b_2} \to \{0,1\}$ such that the composed function (now of two inputs) $f \circ g^n$ has *deterministic communication complexity* at least $\Omega(\mathsf{D}(f)) \cdot \min\{b_1, b_2\}$. This is tight, because there is a simple communication protocol that simulates the query algorithm for f in order to compute $f \circ g^n$. This was proved via a *simulation argument*, which converts any deterministic communication protocol $\mathcal{P}$ for $f \circ g$ with communication complexity C into a query algorithm $\mathcal{A}$ for f with query complexity $O(C/\min\{b_1, b_2\})$.

The gadget used in these works is the Indexing gadget, in which $g(x, y) = y_x$, where we think of x as an index to a bit of y. Here, $b_1 = O(\log n)$ and $b_2 = \text{poly}(n)$. Thus the blow-up of the problem size going from the query function f to the communication function $f \circ g$ is polynomial. In this note, we adapt the proof of [GPW15] to show that the Raz-McKenzie simulation argument also works also when the gadget is the inner product function of logarithmic size.

*Computer and Information Science Department, University of Oregon.
†Joint Center for Quantum Information and Computer Science (QuICS), University of Maryland.
‡Department of Computer Science, UC Berkeley.

ISSN 1433-8092

We make a few remarks about our adaptation. The analysis of the Raz-McKenzie simulation in [GPW15] relies on two main lemmas, the *Thickness Lemma* and the *Projection Lemma*. The Thickness Lemma is invoked whenever the simulation algorithm simulates the communication protocol, and the Projection Lemma is invoked whenever the algorithm performs a query. Our proof only differs in the analysis of the Projection Lemma. We believe our proof is slightly simpler than the one given in [GPW15] (which itself is a simplification of [RM97]).

2 Notation and setup

We largely adhere to the same notation that was used in [GPW15].

Define $\Sigma = \{0,1\}^b \setminus \{0\}$. By 0, we mean the all 0's string. Let $g_{\text{IP}} : \Sigma \times \Sigma \to \{0,1\}$ be the inner product function (i.e. $g(x,y) = \sum_i x_i y_i \mod 2$).

Fix a deterministic communication protocol $\mathcal{P}$ for the composed function $f \circ g^n$. The simulation algorithm maintains a current node v in the communication protocol tree. Let $R_v \subseteq \Sigma^n \times \Sigma^n$ denote the rectangle associated with v. The simulation also maintains a rectangle $S \subseteq R_v$ as well. We write $S = A \times B$.

1. **Projections**. Given a subset $I \subseteq [n]$ and a set $A \subseteq \Sigma^n$, we write $A_I = \{(x_i)_{i \in I} : (x_1, \ldots, x_n) \in A\} \subseteq \Sigma^I$.
2. **Pruning**. Let $U \subseteq \Sigma$, $i \in [n]$, and $A \subseteq \Sigma^n$. Then $A^{i,U} = \{x \in A : x_i \in U\}$. Given a rectangle $S = A \times B \subseteq \Sigma^n \times \Sigma^n$ and a rectangle $R = U \times V \subseteq \Sigma \times \Sigma$, then $S^{i,R}$ denotes $A^{i,U} \times B^{i,V}$.
3. **Neighbors**. Let $I \subseteq [n]$, $J \subseteq I$, and $A \subseteq \Sigma^n$. Let $\alpha \in A_J$. Then we write $\mathsf{Neighbor}(\alpha, A_I)$ to denote the set of $\beta \in A_{I \setminus J}$ such that $(\alpha, \beta) \in A_I$. We say that β is a *neighbor* of α if $\beta \in \mathsf{Neighbor}(\alpha, A_I)$.
4. **Min Entropy**. Let $I \subseteq [n]$, $i \in I$, and $A \subseteq \Sigma^n$. Then $\mathsf{MinEntropy}_i(A_I)$ denotes
$$\log \min_{\alpha \in A_{I \setminus \{i\}}} \left|\mathsf{Neighbor}(\alpha, A_I)\right|.$$
5. **Average Entropy**. Let $I \subseteq [n]$, $i \in I$, and $A \subseteq \Sigma^n$. Then $\mathsf{AvgEntropy}_i(A_I)$ denotes
$$\log \frac{|A_I|}{|A_{I \setminus \{i\}}|}.$$
6. **Thickness**. Given $I \subseteq [n]$, and $A \subseteq \Sigma^n$. Then A is κ-*thick* in I if for all $i \in I$,
$$\mathsf{MinEntropy}_i(A_I) \geq \kappa.$$
Given a rectangle $S = A \times B$, we say that S is thick in I if both A and B are thick in I.
7. **Potential function**. Define the potential function to be $\Phi(S, I) := 2b|I| - \log |A_I \times B_I|$.
8. **Average Entropy threshold**. Define $\kappa_{avg} = 0.9b$.
9. **Min Entropy threshold**. Define $\kappa_{min} = 0.6b$.
10. **Size of gadget**. Define $b = 1000 \log n$.

3 The query simulation and analysis

We describe the Raz-McKenzie simulation algorithm for the communication protocol $\mathcal{P}$. Its presentation is essentially the same as in [GPW15].

---

1. Initialize $S \leftarrow \Sigma^n \times \Sigma^n$, $Q \leftarrow \emptyset$.
2. **While** v is not a leaf, do.
3. Assume that at v, Alice sends a bit to Bob. Otherwise, execute Line 4 to line 17 with A and B exchanged.
4. **If** $\mathsf{AvgEntropy}_i(A_I) \geq \kappa_{avg}$ for all $i \in I$, then
5. Let v_0, v_1 be the children of v, R_{v_0} and R_{v_1} be the rectangle associated with v_0 and v_1, respectively. Set $S^0 := S \cap R_{v_0}$ and $S^1 := S \cap R_{v_1}$.
6. Let $b^* = \text{argmax}_b \{|S_I^b|\}$.
7. Let $\tilde{S} = \tilde{A} \times \tilde{B} \subseteq S^{b^*}$ satisfy the Thickness Lemma.
8. Update $S \leftarrow \tilde{S}$.
9. **else if** $\mathsf{AvgEntropy}_j(A_I) < \kappa_{avg}$ for some $j \in I$, then
10. Query z_j.
11. Let R a z_j-monochromatic rectangle satisfying the Projection Lemma.
12. Update $S \leftarrow S^{j,R}$, $I \leftarrow I \setminus \{j\}$.
13. **End.**
14. **End.**
15. Output the value associated with v.

Figure 1: The simulation algorithm

3.1 Thickness and Projection Lemmas

The analysis of the simulation algorithm, like in [RM97, GPW15], depends on two main lemmas (called the Thickness Lemma and the Projection Lemma) which we now present.

Lemma 1 (Thickness Lemma)**.** *Let $I \subseteq [n]$ and let $S = A \times B \subseteq \Sigma^n \times \Sigma^n$ be a rectangle satisfying $\mathsf{AvgEntropy}_i(A_I) \geq d$ and $\mathsf{AvgEntropy}_i(B_I) \geq d$. Then there exists a subrectangle $S' = A' \times B' \subseteq S$ such that*

1. *For all $i \in I$, $\mathsf{MinEntropy}_i(A'_I) \geq d - \log 4n$, $\mathsf{MinEntropy}_i(B'_I) \geq d - \log 4n$.*

2. $|S'| \geq |S|/2$.

Proof. The proof of this is the same as in [GPW15]. □

Lemma 2 (Sampling properties of random subspaces)**.** *There exist distributions $\mathcal{R}_0, \mathcal{R}_1$ over rectangles $R_A \times R_B \subseteq \Sigma \times \Sigma$ with the following properties. For all $t \in \{0,1\}$, the rectangles sampled in $\mathcal{R}_t$ are t-monochromatic. Furthermore, for all $v \in \Sigma$, let A_v be the indicator random variable for the event $v \in R_A$. Then under both $\mathcal{R}_0$ and $\mathcal{R}_1$, the variables $\{A_v\}$ are such that*

A. $\Pr(A_v = 1) \geq \Omega(2^{-b/2})$ *for all $v \in \Sigma$, and*

B. For all distinct $v, w \in \Sigma$, $\Pr(A_v A_w = 1) \leq \Pr(A_v = 1)\Pr(A_w = 1)$.

Define variables B_v analogously, and the analogous properties hold.

We defer the proof of this sampling Lemma to the end, and show how it can be used for the Projection Lemma.

Lemma 3 (Projection Lemma)**.** *Let $I \subseteq [n]$, and let $S = A \times B \subseteq \Sigma^n \times \Sigma^n$ be a rectangle that is κ-thick in I, $j \in I$ and $t \in \{0,1\}$. Then there exists a rectangle $R \subseteq \Sigma \times \Sigma$ that is t-monochromatic such that*

1. *$S^{j,R}$ is thick in $I \setminus \{j\}$.*
2. *$\Phi(S^{j,R}, I \setminus \{j\}) \leq \Phi(S, I) - 2b + \mathsf{AvgEntropy}_j(A_I) + \mathsf{AvgEntropy}_j(B_I) + 1$.*

Proof. We use the probabilistic method. Imagine sampling a rectangle $R \sim \mathcal{R}_t$ as described in Lemma 2. Let $R = U \times V$. Assume for now that

$$\frac{\left|S^{j,R}_{I\setminus\{j\}}\right|}{\left|S_{I\setminus\{j\}}\right|} \geq 1/2 \tag{1}$$

with non-zero probability. In particular, this implies that $A^{j,U}_{I\setminus\{j\}}$ and $B^{j,V}_{I\setminus\{j\}}$ are non-empty with non-zero probability. Let R be a rectangle satisfying this.

The first item follows from: for all $i \neq j$,

$$\mathsf{MinEntropy}_i(A^{j,U}_{I\setminus\{j\}}) \geq \mathsf{MinEntropy}_i(A_I),$$

and similarly for B. This is because as long as $A^{j,U}_{I\setminus\{j\}}$ is non-empty, the min-entropy of under discarding a coordinate does not decrease.

For the second item, observe that

$$\begin{aligned}\left|S^{j,R}_{I\setminus\{j\}}\right| &= \frac{\left|S^{j,R}_{I\setminus\{j\}}\right|}{\left|S_{I\setminus\{j\}}\right|} \cdot \frac{\left|S_{I\setminus\{j\}}\right|}{|S_I|} \cdot |S_I| \\ &= \frac{\left|S^{j,R}_{I\setminus\{j\}}\right|}{\left|S_{I\setminus\{j\}}\right|} \cdot 2^{-(\mathsf{AvgEntropy}_j(A_I)+\mathsf{AvgEntropy}_j(B_I))} \cdot |S_I|\end{aligned}$$

Let $S' = S^{j,R}$ and $I' = I \setminus \{j\}$. By (1), we have that

$$\begin{aligned}
\Phi(S', I') &= 2b|I'| - \log\left|S'_{I'}\right| \\
&= 2b(|I| - 1) - \log\left|S'_{I'}\right| \\
&\leq 2b|I| - \log|S_I| - 2b + \mathsf{AvgEntropy}_j(A_I) + \mathsf{AvgEntropy}_j(B_I) + 1 \\
&= \Phi(S, I) - 2b + \mathsf{AvgEntropy}_j(A_I) + \mathsf{AvgEntropy}_j(B_I) + 1.
\end{aligned}$$

Thus the proof will be complete once we establish (1). We focus first on Alice's side of the rectangle. For all $x_j \in A_{\{j\}}$ and neighbors $x_{-j} \in \mathsf{Neighbor}(x_j, A_I)$, let $A_{x_{-j}}$ and A_{x_j} denote the indicator variables for the events that $x_{-j} \in A_{I'}^{j,U}$ and $x_j \in U$, respectively. We have that $A_{x_{-j}} = 1$ iff

$$T_{x_{-j}} := \sum_{x_j \in \mathsf{Neighbor}(x_{-j}, A_I)} A_{x_j} \geq 1.$$

Fix a $x_{-j} \in A_{I'}$. We bound the probability of the event $A_{x_{-j}} = 0$:

$$\Pr(A_{x_{-j}} = 0) \leq \Pr\left(\left|T_{x_{-j}} - \mu\right| \geq \mu - 1\right) \leq \frac{\mathrm{Var}\left(T_{x_{-j}}\right)}{(\mu - 1)^2}$$

where $\mu = \mathbb{E}_{R \sim \mathcal{R}_t} T_{x_{-j}}$ and we used Chebyshev's inequality. First, we calculate the variance:

$$\begin{aligned}
\mathrm{Var}\left(T_{x_{-j}}\right) &= \sum_{x_j \in \mathsf{Neighbor}(x_{-j}, A_I)} \mathrm{Var}(A_{x_j}) + \sum_{x_j \neq x'_j \in \mathsf{Neighbor}(x_{-j}, A_I)} \mathrm{Cov}(A_{x_j}, A_{x'_j}) \\
&\leq \sum_{x_j \in \mathsf{Neighbor}(x_{-j}, A_I)} \Pr(A_{x_j} = 1) + \sum_{x_j \neq x'_j \in \mathsf{Neighbor}(x_{-j}, A_I)} \left(\Pr(A_{x_j} A_{x'_j} = 1) - \Pr(A_{x_j} = 1)\Pr(A_{x'_j} = 1)\right) \\
&\leq \mu
\end{aligned}$$

where in the first inequality we used the fact that the A_{x_j} are indicator variables, and that $\Pr(A_v A_w = 1) \leq \Pr(A_v = 1)\Pr(A_w = 1)$ from Lemma 2.

Next, we calculate μ. By Lemma 2 again, we have that

$$\begin{aligned}
\mu &= \mathop{\mathbb{E}}_{R \sim \mathcal{R}_t} \sum_{x_j \in \mathsf{Neighbor}(x_{-j}, A_I)} A_{x_j} \\
&= \sum_{x_j \in \mathsf{Neighbor}(x_{-j}, A_I)} \Pr(A_{x_j} = 1) \\
&\geq 2^{\kappa_{min}} \cdot \Omega(2^{-b/2}).
\end{aligned}$$

In the inequality, we used the fact that the number of neighbors of x_{-j} is at least $2^{\kappa_{min}}$, by the thickness property of S. Thus, we have that

$$\Pr(A_{x_{-j}} = 0) \leq \frac{2}{\mu} \leq O\left(2^{b/2 - \kappa_{min}}\right).$$

Set $\delta = 2^{b/2-\kappa_{min}}$. Now note that $|A^{j,R}_{I'}| = \sum_{x_{-j} \in A_{I'}} A_{x_{-j}}$. By Markov's inequality, we have that the probability more than $O(\sqrt{\delta})$ fraction of $x_{-j} \in A_{I'}$ are missing from $A^{j,R}_{I'}$ is at most $O(\sqrt{\delta})$. Thus with probability at least $1 - O(\sqrt{\delta})$, we have

$$|A^{j,R}_{I'}| \geq \left(1 - O(\sqrt{\delta})\right)|A_{I'}|.$$

By similar reasoning we obtain the same bound for $B^{j,R}_{I'}$. By the union bound we have

$$|S^{j,R}_{I'}| \geq \left(1 - O(\sqrt{\delta})\right)^2 |S_{I'}| = \left(1 - O(\sqrt{\delta})\right)|S_{I'}|$$

with probability $1 - O(\sqrt{\delta})$.

Since this probability is greater than 0, this implies there exists a rectangle R satisfying both conclusions of the Projection Lemma. □

3.1.1 Proof of the sampling Lemma

Proof. We first describe a process to sample $R_A \times R_B \sim \mathcal{R}_0$. Sample a pair (V, W) of orthogonal subspaces of dimension $d = b/2$. Set $R_A = V \setminus \{0\}$ and $R_B = W \setminus \{0\}$.

It is easy to verify that rectangles sampled in this way are 0-monochromatic. We now verify (A). We first observe that the marginal distribution of R_A is a uniformly random d-dimensional subspace. Let $\binom{b}{d}_2$ denote the number of d-dimensional subspaces in $\mathbb{F}_2^b$; it is well known that

$$\binom{b}{d}_2 = \frac{(2^b - 1)\cdots(2^b - 2^{d-1})}{(2^d - 1)\cdots(2^d - 2^{d-1})}.$$

Fix a $v \in \Sigma$. The probability that a random d-dimensional subspace contains v is exactly

$$\frac{\binom{b-1}{d-1}_2}{\binom{b}{d}_2} = \frac{2^d - 1}{2^b - 1}.$$

This establishes both (A) and its analogue for B_v. Now fix distinct $v, w \in \Sigma$. Note that since v, w are distinct and are non-zero, they are linearly independent, and thus span a 2-dimensional space. The probability that a random d-dimensional subspace contains both v and w is exactly

$$\frac{\binom{b-2}{d-2}_2}{\binom{b}{d}_2} = \frac{(2^d - 1)(2^d - 2)}{(2^b - 1)(2^b - 2)}.$$

It is easy to verify that

$$\frac{\binom{b-2}{d-2}_2}{\binom{b}{d}_2} \leq \left(\frac{\binom{b-1}{d-1}_2}{\binom{b}{d}_2}\right)^2,$$

which establishes (B).

Now we describe how to sample from $\mathcal{R}_1$. First, sample a vector u with odd Hamming weight. Let $G(u)$ denote the $(b-1)$-dimensional orthogonal complement of span$\{u\}$. Sample a pair (V, W) of orthogonal subspaces of dimension $d' = b/2 - 1$ from within $G(u)$. Set $R_A = V + u$ and $R_B = W + u$.

Again it is easy to verify that rectangles sampled in this way are 1-monochromatic. We now verify (A). Conditioned on u, V is a uniformly random d'-dimensional subspace within $G(u)$. Therefore we have

$$\Pr(A_v = 1) = \sum_{\substack{u:|u| \text{ is odd} \\ \langle u,v\rangle=1}} \Pr(u) \cdot \Pr(u+v \in V|u) = \sum_{\substack{u:|u| \text{ is odd} \\ \langle u,v\rangle=1}} \Pr(u) \cdot \frac{\binom{b-2}{d'-1}_2}{\binom{b-1}{d'}_2} = \sum_{\substack{u:|u| \text{ is odd} \\ \langle u,v\rangle=1}} \Pr(u) \cdot \frac{2^{d'}-1}{2^{b-1}-1}$$

The vector u is uniformly random over all odd Hamming weight vectors, which is an affine subspace of dimension $b-1$. Thus $\Pr(u) = 2^{-(b-1)}$. The number of u's such that $|u|$ is odd and $\langle u, v\rangle = 1$ (which implies that $\langle u+v, v\rangle = 0$ and thus $u+v$ is in $G(u)$) is the number of solutions to this system of equations:

$$\langle u, \vec{1}\rangle = 1$$
$$\langle u, v\rangle = 1$$

where $\vec{1}$ denotes the all one's vector. There are 2^{b-2} solutions to this system when $v \neq \vec{1}$, and 2^{b-1} solutions when $v = \vec{1}$. Therefore

$$\Pr(A_v = 1) = \begin{cases} (2^{d'}-1)/(2^b-2) & \text{if } v \neq \vec{1} \\ (2^{d'}-1)/(2^{b-1}-1) & \text{if } v = \vec{1}. \end{cases}$$

Now fix distinct $v, w \in \Sigma$. Then we have

$$\begin{aligned} \Pr(A_v A_w = 1) &= \sum_{\substack{u:|u| \text{ is odd} \\ \langle u,v\rangle=\langle u,w\rangle=1}} \Pr(u) \cdot \Pr(A_v A_w = 1|u) \\ &= \sum_{\substack{u:|u| \text{ is odd} \\ \langle u,v\rangle=\langle u,w\rangle=1}} \Pr(u) \cdot \frac{\binom{b-3}{d'-2}_2}{\binom{b-1}{d'}_2}. \end{aligned}$$

Note that unless the vectors v, w, and the all ones vector $\vec{1}$ are all linearly independent, the number of u such that $|u|$ is odd and $\langle u, v\rangle = \langle u, w\rangle = 1$ is zero. This is because if $v + w = \vec{1}$, then $\langle u, v\rangle + \langle u, w\rangle = \langle u, \vec{1}\rangle = |u| = 0 \mod 2$, a contradiction.

Thus in the case they are all linearly independent, the number of solutions u is at most 2^{b-3}, so therefore

$$\Pr(A_v A_w = 1) \leq \frac{1}{4} \frac{\binom{b-3}{d'-2}_2}{\binom{b-1}{d'}_2}.$$

But the above calculation implies that $\Pr(A_v A_w = 1) \leq \Pr(A_v = 1)\Pr(A_w = 1)$, which establishes (B).

□

3.2 Correctness of the simulation algorithm

Claim 4. *The simulation algorithm maintains the following loop invariants:*

1. *$S \subseteq R_v$ where v is the current node in the protocol tree.*
2. *S is κ_{min}-thick in I.*
3. *$g(x_i, y_i) = z_i$ for all $(x, y) \in S$ and for all $i \in [n] \setminus I$.*

Proof. The proof of this is similar in [GPW15]. The invariants are trivially satisfied initially. Assume that they hold at the beginning of an iteration. We can also assume without loss of generality that it is Alice's turn to communicate in this iteration.

Suppose line 4 holds (i.e. the simulation is in the communication phase), and assume without loss of generality that $b^* = 0$. Let $S = A \times B$ and $R_0 = X_0 \times Y_0$. Let $\widetilde{A} = A \cap X_0$ and $\widetilde{B} = B$. Since $S \subseteq \widetilde{A} \times \widetilde{B} \subseteq R_0$, Invariants 1 and 3 hold. For all $i \in I$ we have that

$$\mathsf{AvgEntropy}_i(\widetilde{A}_I) = \frac{\left|(A \cap X_0)_I\right|}{(A \cap X_0)_{I \setminus \{i\}}} \geq \frac{\left|(A \cap X_0)_I\right|/2}{\left|A_{I \setminus \{i\}}\right|} \geq 2^{\kappa_{avg}-1}.$$

$\mathsf{AvgEntropy}_i(\widetilde{B}_I) = \mathsf{AvgEntropy}_i(B_I)$. We apply the Thickness Lemma by setting d to $\kappa_{avg} - 1$. Invariant 2 holds because $\kappa_{avg} - 1 - \log 4n \geq \kappa_{min}$.

Suppose line 9 holds (i.e. the simulation is in the query phase). By Invariant 2, S is κ_{min} thick in I. We may apply the Projection Lemma to S and get a rectangle R. Invariant 1 is unchanged. Invariants 2 and 3 hold because of the Projection Lemma and the definition of R.

□

The correctness of the algorithm is argued similarly to the one in [GPW15]. Let v be the leaf reached at termination. It suffices to show that there exists an $(x, y) \in R_v$ such that $g^n(x, y) = z$. Assuming that the algorithm continues by executing line 10-12, repeatedly, once for each remaining coordinate $i \in I$ in arbitrary order until only one coordinate remains unqueried (ignoring the condition 2 in the Projection Lemma). Let $(v, I, R = A \times B)$ be the state at the end of this extended execution, where $I = \{i\}$ is a singleton. Then $\mathsf{MinEntropy}_i(A_{\{i\}}) \geq 2^{\kappa_{min} b}$ and $\mathsf{MinEntropy}_i(B_{\{i\}}) \geq 2^{\kappa_{min} b}$. Hence $A_{\{i\}} \times B_{\{i\}}$ is not monochromatic, because of the properties of the inner product gadget. Pick an (x_i, y_i) such that $g(x_i, y_i) = z_i$ and pick an $(x, y) \in A \times B$ with this value of (x_1, y_1). By Invariant 1, correctness is established.

3.3 Query complexity of the algorithm

Claim 5. *The number of queries made by the simulation algorithm is at most $O(C/b)$, where C is the communication complexity of the protocol for $f \circ g^n$.*

Proof. The proof of this is the same as in [GPW15]. Note that our potential function $\Phi(S, I) = 0$ initially, and is always nonnegative. There are most C communication rounds in the algorithm. In every communication round, the size of S decreases by at most 4 (because we first divide S based on the bit being communicated, and then we cut it further due to the Thickness Lemma), so the potential function increases by at most 2. Since the potential function only increases in the communication rounds, the potential function has a maximum of $2C$. On the other hand, in each query round of the algorithm, the potential function decreases by $\Omega(b)$. Thus there are at most $O(C/b)$ query rounds in the algorithm. □

3.4 Concluding remarks

We end with two observations about our adaptation of [GPW15]. First, we note that the only place where we used the fact that the inner product gadgets are logarithmic size are in the Thickness Lemma (the Projection Lemma does not rely on this). In contrast, both the Thickness and Projection Lemmas in [GPW15] require that the size of the gadget has some dependence on n. We believe our simplifications are a step towards obtaining a simulation argument with *constant-sized* gadgets.

Next, readers might observe that we've renamed MinDeg and AvgDeg to MinEntropy and AvgEntropy, respectively. The reason for this is because we believe it is useful to understand the simulation argument in terms of entropies, rather than in terms of the combinatorial counting measures of minimum degree and average degree.

To illustrate this, imagine that instead of just maintaining the *set* S, the algorithm keeps track of a *distribution* $\mathcal{S}$ over $\Sigma^n \times \Sigma^n$. Instead of maintaining that the minimum degree of various blocks of S are high, we ensure that the conditional min-entropy of the blocks of $\mathcal{S}$ are large, i.e., that $H_{min}(\mathcal{S}_i^I | \mathcal{S}_{\neq i}^I) \geq \kappa_{min}$. To determine when to make a query or not, the algorithm checks whether the conditional Shannon entropy of the various blocks of $\mathcal{S}$ are large or not, i.e., that $H(\mathcal{S}_i^I | \mathcal{S}_{\neq i}^I) \geq \kappa_{avg}$. Correspondingly, the potential function $\Phi(\mathcal{S}, I)$ should be nothing but the *KL divergence* between the marginal distribution $\mathcal{S}^I$ and the uniform distribution.

Using smoother measures such as KL divergence and Shannon entropy makes certain things nicer. For example, the proof of the Projection Lemma can be expressed as an application of the chain rule for KL divergence. Suppose that z_1 is being queried, because $H(\mathcal{A}_1 | \mathcal{A}_{\neq 1}) \leq \kappa_{avg}$ (where $\mathcal{S} = \mathcal{A} \times \mathcal{B}$). Furthermore assume that $I = [n]$. Then observe that

$$\Phi(\mathcal{S}, [n]) = D_{KL}(\mathcal{S} \,\|\, U_{[n]}) = D_{KL}(\mathcal{S}_{\neq 1} \,\|\, U_{\neq 1}) + \mathop{\mathbb{E}}_{s \sim \mathcal{S}_{\neq 1}} D_{KL}\left(\frac{\mathcal{S}_1}{s} \,\middle\|\, U_1\right)$$

where $U_{[n]}$ denotes the uniform distribution on $\Sigma^n \times \Sigma^n$, $\mathcal{S}_{\neq 1}$ denotes the marginal distribution of $\mathcal{S}$ on blocks 2 through n, $\frac{\mathcal{S}_1}{s}$ denotes the distribution of $\mathcal{S}_1$ *conditioned* on a sample s drawn from $\mathcal{S}_{\neq 1}$. Notice that

$$\mathop{\mathbb{E}}_{s \sim \mathcal{S}_{\neq 1}} D_{KL}\left(\frac{\mathcal{S}_1}{s} \,\middle\|\, U_1\right) = 2b - H(\mathcal{S}_1 | \mathcal{S}_{\neq 1}) \geq b - H(\mathcal{A}_1 | \mathcal{A}_{\neq 1}) \geq 0.1b.$$

Thus $D_{KL}(\mathcal{S}_{\neq 1} \,\|\, U_{\neq 1}) \leq \Phi(\mathcal{S}, [n]) - 0.1b$. We would almost be done in showing that the potential function decreases, except we have to show that $D_{KL}(\frac{\mathcal{S}_{\neq 1}}{1,\mathcal{R}} \,\|\, U_{\neq 1}) \approx D_{KL}(\mathcal{S}_{\neq 1} \,\|\, U_{\neq 1})$, where $\frac{\mathcal{S}_{\neq 1}}{1,\mathcal{R}}$ denotes the distribution $\mathcal{S}_{\neq 1}$ conditioned on the first coordinate being projected to the random rectangle sampled according to $\mathcal{R}$. Though showing this will not be as clean a calculation as the one performed above, we believe that this calculation sketch reveals more intuition about why the Projection Lemma should be true.

We hope that this perspective may be useful in extending the Raz-McKenzie simulation to the *randomized* setting — that is, giving a simulation method to lift randomized query lower bounds to randomized communication lower bounds.

Acknowledgments. We wish to thank Govind Ramnarayan and Aviad Rubinstein for useful comments.

References

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

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

ECCC ISSN 1433-8092

https://eccc.weizmann.ac.il