Electronic Colloquium on Computational Complexity, Report No. 54 (2017)

Lifting randomized query complexity to randomized communication complexity

Anurag Anshu† Naresh B. Goud† Rahul Jain * Srijita Kundu†
Priyanka Mukhopadhyay†

March 22, 2017

Abstract

We show that for any (partial) query function $f : \{0,1\}^n \to \{0,1\}$, the randomized communication complexity of f composed with Index_m^n (with $m = \mathrm{poly}(n)$) is at least the randomized query complexity of f times $\log n$. Here $\mathrm{Index}_m : [m] \times \{0,1\}^m \to \{0,1\}$ is defined as $\mathrm{Index}_m(x, y) = y_x$ (the xth bit of y).

Our proof follows on the lines of Raz and Mckenzie [RM99] (and its generalization due to [GPW15]), who showed a lifting theorem for deterministic query complexity to deterministic communication complexity. Our proof deviates from theirs in an important fashion that we consider partitions of rectangles into many sub-rectangles, as opposed to a particular sub-rectangle with desirable properties, as considered by Raz and McKenzie. As a consequence of our main result, some known separations between quantum and classical communication complexities follow from analogous separations in the query world.

1 Introduction

Communication complexity and query complexity are two concrete models of computation which are very well studied. In the communication model there are two parties Alice, with input x and Bob, with input y, and they wish to compute a joint function $f(x, y)$ of their inputs. In the query model one party Alice tries to compute a function $f(x)$ by querying bits of a database string x. There is a natural way in which a query protocol can be viewed as a communication protocol between Alice, with no input, and Bob, with input x, in which the only communication allowed is queries to the bits of x and answers to these queries. Given this, we can (informally) view the query model as a "simpler" sub-model of the communication model. Indeed several results in query complexity are easier to argue and obtain than the corresponding results in communication complexity. One interesting technique that is often employed with great success is to first show a result in the query model and then "lift" it to a result in the communication model via some "lifting theorem".

One of the first such lifting theorems was shown by Raz and McKenzie [RM99] (and its generalization by [GPW15]). For a (partial) query function $f : \{0,1\}^n \to \{0,1\}$ and a communication function $g : \{0,1\}^m \times \{0,1\}^k \to \{0,1\}$ let the composed function $f \circ g^n$ be defined as $f \circ g^n((x_1, y_1), \ldots, (x_n, y_n)) = f(g(x_1, y_1), \ldots, g(x_n, y_n))$. Raz and McKenzie [RM99] (and the generalization due to [GPW15]) showed that for every query function $f : \{0,1\}^n \to \{0,1\}$ the

*Centre for Quantum Technologies, National University of Singapore and MajuLab, UMI 3654, Singapore.
†Center for Quantum Technologies, National University of Singapore, Singapore.

ISSN 1433-8092

deterministic communication complexity of f composed with Index_m (with $m = \mathrm{poly}(n)$) is at least the deterministic query complexity of f times $\log n$. Here $\mathrm{Index}_m : [m] \times \{0,1\}^m \to \{0,1\}$ is defined as $\mathrm{Index}_m(x, y) = y_x$ (the xth bit of y). Subsequently several lifting theorems for different complexity measures have been shown, example lifting *approximate-degree* to *approximate-rank* [She11] and *approximate junta-degree* to *smooth-corruption-bound* [GLM$^+$15] etc.

Our result

In this work we show lifting of (bounded error) randomized query complexity to (bounded error) randomized communication complexity. For a (partial) query function $f : \{0,1\}^n \to \{0,1\}$ let the randomized query complexity with (worst-case) error $\varepsilon > 0$ of f be denoted $\mathrm{R}_\varepsilon(f)$. Similarly for a communication function $g : \{0,1\}^m \times \{0,1\}^k \to \{0,1\}$, let the randomized communication complexity with (worst-case) error $\varepsilon > 0$ of g be denoted $\mathrm{R}_\varepsilon(g)$. We show the following.

Theorem 1. *For all (partial) functions $f : \{0,1\}^n \to \{0,1\}$,*

$$\mathrm{R}_{1/4}(f \circ \mathrm{Index}_m^n) = \Omega(\mathrm{R}_{1/3}(f) \cdot \log n),$$

where $m = \mathrm{poly}(n)$ [1].

On the other hand it is easily seen with a simple simulation of a query protocol using a communication protocol that $\mathrm{R}_{1/3}(f \circ \mathrm{Index}_m^n) = O(\mathrm{R}_{1/3}(f) \cdot \log m)$. This implies $\mathrm{R}_{1/3}(f \circ \mathrm{Index}_m^n) = \Theta(\mathrm{R}_{1/3}(f) \cdot \log n)$ with $m = \mathrm{poly}(n)$.

Our result implies a recent result of [ABBD$^+$16] where they exhibited a power 2.5 separation between classical randomized and quantum communication complexities for a total function. It also implies exponential separation between two-round quantum communication complexity and randomized communication complexity first shown by [Raz99]

Our techniques

Our techniques are largely based on the techniques of Raz and McKenzie [RM99] as presented in [GPW15] with an important modification to deal with distributional error protocols instead of deterministic protocols. Let $\mathcal{T}$ be a deterministic communication protocol tree for $f \circ \mathrm{Index}_m^n$ (with $m = \mathrm{poly}(n)$). We use this to create a randomized query protocol Π (see. Figure 1) for f. Let z be an input for which we are supposed to output $f(z)$. We start with the root of $\mathcal{T}$ and continue to simulate $\mathcal{T}$ (using randomness) till we find a co-ordinate $i \in [n]$ where $\mathcal{T}$ has *worked enough* so that $\mathrm{Index}_m(x_i, y_i)$ is becoming (only slightly) *determined*. Using the properties of Index_m we conclude that $\mathcal{T}$ must have communicated $O(\log n)$ bits by now. We go ahead a query z_i (the ith bit of z) and *synchronize* with z_i, that is go to the appropriate sub-event of the current node in $\mathcal{T}$ consistent with z_i. We then continue to simulate $\mathcal{T}$. This way we do (in expectation) one query in Π for $O(\log n)$ communication bits in $\mathcal{T}$. On reaching a leaf of $\mathcal{T}$ we make the same output as $\mathcal{T}$. This output is correct with high probability since the unqueried bits are sufficiently *un-determined*.

The synchronizing of $\mathcal{T}$ with z_i was done by Raz and McKenzie [RM99] via a *Projection Lemma* by going to a "sub-event" (of small probability) of the current node in $\mathcal{T}$. They could afford to do so since $\mathcal{T}$ was a deterministic protocol and hence it was correct *everywhere*. On the other hand we are forced to work with a "partition" of the node into sub-events where each sub-event is consistent with either z_i being 0 or z_i being 1. This allows us to move according to the "flow" of $\mathcal{T}$ so that we can "capture" correctness of $\mathcal{T}$ wherever it has to offer. This requires us to show a different *Partition Lemma* which works in place of the Projection Lemma of Raz and McKenzie.

[1] We state our result for Boolean functions f, however it holds for general relations.

2 Preliminaries

In this section, we present some notations and basic lemmas needed for the proof of our main result.

Let $f : \{0,1\}^n \to \{0,1\}$ be a (partial) function. Let $\varepsilon > 0$ be an error parameter. Let the randomized query complexity, denoted $\mathrm{R}_\varepsilon(f)$, be the maximum number of queries made by the best randomized query protocol computing f with error at most ε on any input $x \in \{0,1\}^n$. Let θ be a distribution on $\{0,1\}^n$. Let the distributional query complexity, denoted $\mathrm{D}^\theta_\varepsilon(f)$, be the maximum number of queries made by the best deterministic query protocol computing f with average error at most ε under θ. The distributional and randomized query complexities are related by the following Yao's Lemma.

Fact 2 (Yao's Lemma)**.** *Let $\varepsilon > 0$. We have $\mathrm{R}_\varepsilon(f) = \max_\theta \mathrm{D}^\theta_\varepsilon(f)$.*

Similarly, we can define randomized and distributional communication complexities with a similar Yao's Lemma relating them.

Let λ be a *hard* distribution on $\{0,1\}^n$ such that $\mathrm{D}^\lambda_{1/3}(f) = \mathrm{R}_{1/3}(f)$, as guaranteed by Yao's Lemma. Let $m = O(n^{100})$ and let $\mathrm{Bal}_m \subset \{0,1\}^m$ be the set of all strings of length m with equal number of 0's and 1's. Observe that $|\mathrm{Bal}_m| = \binom{m}{m/2} \geq 2^m/\sqrt{m}$. The notation Bal^n_m will refer to the set $\mathrm{Bal}_m \times \mathrm{Bal}_m \times \ldots \mathrm{Bal}_m$. Let $\mathrm{Index}_m : [m] \times \mathrm{Bal}_m \to \{0,1\}$ be defined as $\mathrm{Index}_m(x,y) = y_x$ (the xth bit of y). Consider the following *lifted* distribution for the composed function $f \circ \mathrm{Index}^n_m$: $\mu(x,y) = \lambda(G(x,y))/|G^{-1}(G(x,y))|$, where $G := \mathrm{Index}^n_m$. We observe for this distribution that $\mu(x)$ and $\mu(y)$ are uniform in their support.

Let Alice and Bob's inputs for the composed function be respectively $x = (x_1, \ldots, x_n) \in [m]^n$ and $y = (y_1, \ldots, y_n) \in \mathrm{Bal}^n_m$.

We use the following notation, which coincides largely with the notation used in [GPW15].

- For a node v in a communication protocol tree, let $R^v = X^v \times Y^v$ denote its associated rectangle. If Alice or Bob send the bit b at v, let v_b be the corresponding child of v and $X^{v,b} \subseteq X^v$ and $Y^{v,b} \subseteq Y^v$ be the set of inputs of Alice and Bob respectively on which they do this.
- For a string $x \in [m]^n$ and an interval $I \subset [n]$, let x_I be the restriction of x to the interval I. We use shorthand x_i for $x_{\{i\}}$. We use similar notation for string $y \in \mathrm{Bal}^n_m$. For a set $A \subset [m]^n$, let $A_I := \{x_I : x \in A\}$ be the restriction of A to the interval I. For a set $U \subseteq [m]$, let $A|_{U,i} := \{x \in A : x_i \in U\}$ (we will omit the substript 'i' from $A|_{U,i}$ when it is clear from context).
- For an interval I and $A \subseteq [m]^{|I|}$and $x_{-i} \in [m]^{|I|-1}$, let $A_{x_{-i}} = \{x \in A : x_{I\setminus\{i\}} = x_{-i}\}$ and $A_{-i} = \cup_{x_{-i} \in [m]^{|I|-1}} A_{x_{-i}}$. Let $\mathrm{MinDeg}_i(A) = \min\limits_{x_{-i}:|A_{x_{-i}}|>0} |A_{x_{-i}}|$ and $\mathrm{AvgDeg}_i(A) = \frac{|A|}{|A_{-i}|}$.
- $A \subseteq [m]^n$ is called *thick* for $I \subseteq [n]$ if $\mathrm{MinDeg}_i(A_I) \geq m^{17/20}$ for all $i \in I$.
- For $i \in [n]$, a set $B \subset \mathrm{Bal}^n_m$ and a string $\eta \in \mathrm{Bal}_m$, let $B_{\eta,i} := \{y \in B : y_i = \eta\}$.
- For $y \in \mathrm{Bal}_m$ and $z \in \{0,1\}$, define $U(y,z) := \{x \in [m] : \mathrm{Index}_m(x,y) = z\}$.
- We say that a set $A \subseteq [m]^n$ is uniform for an interval I if it holds that $|(A_{x_I})_{[n]\setminus I}|$ is the same for all $x_I \in A_I$.

We will need the following claims.

Claim 3. *Let $A \subset [m]^n$ be such that for a given interval I, $\mathrm{AvgDeg}_i(A_I) \geq d$ for all $i \in I$. Let $A' \subset A$ be such that $\frac{|A'_I|}{|A_I|} \geq \frac{1}{n^2}$. Then $\mathrm{AvgDeg}_i(A'_I) \geq \frac{d}{n^2}$ for all $i \in I$.*

Proof. Fix an $i \in I$. We have that $\frac{|A'_I|}{|A'_{I\setminus\{i\}}|} \geq \frac{|A'_I|}{|A_{I\setminus\{i\}}|}$ (as $A'_{I\setminus\{i\}} \subset A_{I\setminus\{i\}}$). Thus, by definition of average degree,

$$\mathrm{AvgDeg}_i(A'_I) = \frac{|A'_I|}{|A'_{I\setminus\{i\}}|} \geq \frac{|A'_I|}{|A_{I\setminus\{i\}}|} \geq \frac{1}{n^2}\frac{|A_I|}{|A_{I\setminus\{i\}}|} \geq \frac{d}{n^2}.$$

□

Claim 4. *Let $A \subset [m]^n$ be a set such that A is thick on an interval I. Then for every $i \in I$ and every x_i, A_{x_i} is thick on the interval $I/\{i\}$.*

Proof. Fix an i. For every $j \in I \setminus \{i\}$, we have that $\mathrm{MinDeg}_j(A_I) \geq m^{17/20}$, that is, for every $x_{-j} \in A_{I\setminus\{j\}}$, $|A_{I,x_{-j}}| \geq m^{17/20}$. But, writing $x'_{-j} = x_i x'_{-\{j,i\}}$ we conclude that for every $x'_{-j} \in A_{I\setminus\{i,j\}}$, $|A_{I\setminus\{i\},x'_{-j}}| \geq m^{17/20}$. □

Claim 5. *Consider a tree with nodes and weighted directed edges such that for every node, the sum of weights of edges going to its children sum to 1. Call a non-leaf node aborted if it has no children. For any node, let the sum of weights of the edges going to aborted children be at most δ. Let the depth of the tree be d. Consider a random walk that starts from the root and goes to the children according to the weights of the edges. Then the overall probability of abort is at most $\delta \cdot d$.*

Proof. We construct a new tree in which nodes at a particular level which do not abort are coarse-grained into a single node and the aborting nodes are coarse grained into another node (which we again call abort node). For this tree, the probability of a node having an aborted child is still at most δ and the overall probability of abort is at least as large as in the original tree, which is

$$\delta + (1-\delta)\cdot\delta + (1-\delta)^2\cdot\delta \ldots + (1-\delta)^{d-1}\delta \leq d\delta.$$

This completes the proof. □

The following Thickness Lemma was shown in [GPW15] with slightly different parameters. We reproduce a proof for completeness.

Lemma 6 (Thickness Lemma, [GPW15])**.** *If $n \geq 2$ and $A \subseteq [m]^n$ is such that $\mathrm{AvgDeg}_i(A) \geq d$ for all $i \in I$ for some $I \subseteq [n]$, then there exists $A' \subseteq A$ such that*

(i) $|A'| \geq \left(1 - \frac{1}{n^2}\right)|A|$,

(ii) $\mathrm{MinDeg}_i(A') \geq \frac{d}{n^3}$ *for all $i \in I$.*

Proof. We give the following algorithm that produces a series of $A^0, A^1, A^2, \ldots$ of subsets of A. We claim that the algorithm terminates and produces a subset that satisfies the properties required of A'.

1. Set $j = 0, A^0 = A$.
2. If $\mathrm{MinDeg}_i(A^j) \geq d/n^3$ for all $i \in I$, then output $A' = A^j$.
3. Find $i \in I$ such that $\mathrm{MinDeg}_i(A^j) < d/n^3$ and $x_{-i} \in [m]^{n-1}$ such that $d/n^3 > |A_{x_{-i}}| > 0$.
4. Set $A^{j+1} = A^j \setminus \{x \in A^j : x_{[n]\setminus\{i\}} = x_{-i}\}$ and $j = j+1$.

Clearly the output of the algorithm satisfies condition (ii). We need to prove that the algorithm terminates and the size of A^j at termination is as required. In an A^j which is not final output, if x_{-i} and i are picked, then since all the x_{-i} strings present in A^j are removed in A^{j+1}, we have $|A^{j+1}_{-i}| \leq |A^j_{-i}| - 1$ and for every $i \neq i'$, $|A^{j+1}_{-i'}| \leq |A^j_{-i'}|$. Since all the $|A^j_{-i}|$-s for $i \in I$ can decrease to zero at the most, the algorithm runs for at most

$$\sum_{i \in I} |A_{-i}| = \sum_{i \in I} \frac{|A|}{\mathrm{AvgDeg}_i(A)} \leq \frac{|I||A|}{d}$$

iterations. Moreover, the number of strings being removed from A^j at each step is at most d/n^3, since we only remove strings x_{-i} for which $d/n^3 > |A_{x_{-i}}| > 0$. That is, $|A^{j+1}| > |A^j| - d/n^3$. So the total number of element of A that can be removed is at most $(|I||A|/d) \cdot (d/n^3) \leq |A|/n^2$. Hence the final A' satisfies $|A'| \geq \left(1 - \frac{1}{n^2}\right)|A|$. □

Now we prove a Partition Lemma which helps to partition a current node when the algorithm Π performs a query.

Lemma 7 (Partition Lemma)**.** *For $I \subseteq [n]$, let $B \subseteq \mathrm{Bal}_m^{|I|}$ be such that $\frac{|B|}{|\mathrm{Bal}_m^{|I|}|} \geq 2^{-n^2}$. If $A \subseteq [m]^{|I|}$ is thick, then for all $i \in I$ and for all $z_i \in \{0,1\}$ with probability at least $1 - 2^{-m^{7/20}}$ when $y_i = \eta$ is drawn randomly from the distribution $\frac{|B_{\eta,i}|}{|B|}$ it holds that*

$$\frac{1}{2} + n^{-5} \geq \frac{|(A|_{U(y_i,z_i)})_{x_{-i}}|}{|A_{x_{-i}}|} \geq \frac{1}{2} - n^{-5} \quad \forall x_{-i} \in [m]^{|I|-1}.$$

We say a $y_i \in \mathrm{Bad}(A,B,i)$ if it does not satisfy above property or $\frac{|B_{y_i}|}{|\mathrm{Bal}_m^{|I|-1}|} < 2^{-n^2}$. Then, for all $y_i \notin \mathrm{Bad}(A,B,i)$, we have that $(A|_{U(y_i,z_i)})_{-i}$ is the same set as A_{-i}.

Proof. Fix an index $i \in I$ and $z_i \in \{0,1\}$. Fix an ε to be chosen later. For arguments below y is fixed outside I. Let a string y_i satisfy the property $P(x_{-i})$ if it holds that

$$\frac{1}{2} + 2\varepsilon \geq \frac{|(A|_{U(y_i,z_i)})_{x_{-i}}|}{|A_{x_{-i}}|} \geq \frac{1}{2} - 2\varepsilon.$$

Since A is thick, it holds that for all x_{-i}, we have $|A_{x_{-i}}| \geq m^{17/20}$. The number of assignments of $\{0,1\}$ to the strings in $A_{x_{-i}}$ such that the fraction of 0-assignments and fraction of 1 assignments differ by 2ε is at most $2^{m^{17/20}(1-\varepsilon^2)}$ by Chernoff bound. Now for a fixed $y_{-i} \in \mathrm{Bal}_m^{|I|-1}$ and x_{-i} consider $B_{y_{-i}}$ at the i-th block. The total number of strings in the i-th block of $B_{y_{-i}}$ (even after an over counting to include strings not in Bal_m), such that the indices in $A_{x_{-i}}$ do not give balanced output, is at most

$$2^{m - m^{17/20} + m^{17/20}(1-\varepsilon^2)} \leq 2^{m - m^{17/20} \cdot \varepsilon^2}.$$

This implies, for a y drawn from $B_{y_{-i}}$, the probability that y_i does not satisfy the property $P(x_{-i})$ is at most $2^{m - m^{17/20} \cdot \varepsilon^2} \frac{1}{|B_{y_{-i}}|}$. Thus the probability that y_i does not satisfy any of the properties $P(x_{-i})$ for all x_{-i} is at most $m^n \cdot 2^{m - m^{17/20} \cdot \varepsilon^2} \frac{1}{|B_{y_{-i}}|}$. Hence, the probability that y drawn according to the distribution given in the statement of the lemma satisfies $P(x_{-i})$ for all x_{-i} is at least

$$\begin{aligned} 1 - m^n \cdot 2^{m - m^{17/20} \cdot \varepsilon^2} \sum_{y_{-i}} \frac{\Pr[y_{-i}]}{|B_{y_{-i}}|} &= 1 - 2^{n \log m + m - m^{17/20} \cdot \varepsilon^2} \sum_{y_{-i}} \frac{|B_{y_{-i}}|}{|B| \cdot |B_{y_{-i}}|} \\ &= 1 - 2^{n \log m + m - m^{17/20} \cdot \varepsilon^2} \sum_{y_{-i}} \frac{1}{|B|}. \end{aligned}$$

Since number of strings $y_{-i} \leq |\mathrm{Bal}_m|^{|I|-1}$ and $|B| \geq |\mathrm{Bal}_m|^{|I|} \cdot 2^{-n^2}$, we conclude that above probability is at least

$$1 - 2^{n \log m + m - m^{17/20} \cdot \varepsilon^2} \cdot |\mathrm{Bal}_m|^{-1} \cdot 2^{n^2} \geq 1 - \sqrt{m} \cdot 2^{2n^2 - m^{17/20} \cdot \varepsilon^2}.$$

Choosing $2\varepsilon = m^{-1/4}$, we conclude the proof. □

Following is a Uniformity Lemma, which ensures that as long as B is a large set, the distribution μ on Alice's side, conditioned on Bob's side being in set B, behaves like a uniform distribution. The proof of the lemma is deferred to Appendix A.

Lemma 8 (Uniformity lemma)**.** *Fix an interval $I \subset [n]$, a subset $B \subseteq \mathrm{Bal}_m^n$ and a string $z \in \{0,1\}^n$. Let B have the additional property that $B_{[n]\setminus I}$ is a fixed string and $\frac{|B|}{|\mathrm{Bal}_m^{|I|}|} \geq 2^{-n^2}$. Define the distribution*

$$\sigma^{I,B,z}(x'_I) := \frac{\sum_{x,y: x_I = x'_I, G(x,y)=z, y \in B} \mu(x,y)}{\sum_{x,y: G(x,y)=z, y\in B} \mu(x,y)},$$

which is the distribution μ conditioned on z, B. Then following properties hold:

1. *For any $A \subset [m]^n$ such that A is thick and uniform in I, we have that*

$$\sum_{x_1 x_2 \dots x_{|I|} \in A_I} \sigma^{I,B,z}(x_1 x_2 \dots x_{|I|}) \geq \frac{1 + 4n^{-5}}{m^{|I|}} \cdot (1 - n^{-67})|A_I|.$$

2. *For any $A' \subset A$, where A is thick and uniform in I, we have*

$$\sum_{x_1 x_2 \dots x_{|I|} \in A'_I} \sigma^{I,B,z}(x_1 x_2 \dots x_{|I|}) \leq \frac{1 + 4n^{-5}}{m^{|I|}} (|A'_I| + n^{-67}|A_I|).$$

Following is an immediate corollary of the lemma.

Corollary 9. *Let I be an interval and $1 \notin I$ be an index. Fix a string z and a set $B \in \mathrm{Bal}_m^n$ such that $B_{[n]\setminus I}$ is fixed string, with $y_1 \in \mathrm{Bal}_m$ being the string at index $\{1\}$. Let $A \subseteq [m]^n$ be a subset such that it is uniform and thick in the interval $I \cup \{1\}$ and the support of $A_{\{1\}}$ is on the set $\{x_1 : \mathrm{Index}_m(x_1, y_1) = z_1\}$. Let $A' \subset A$ be a subset of A. Consider the distribution $\sigma^{I,B,z}$ as defined in Lemma 8. Let θ_1 be uniform distribution on $\{x_1 : \mathrm{Index}_m(x_1, y_1) = z_1\}$. Then it holds that*

$$\frac{\sum_{x_1 x_I \in A'} \sigma^{I,B,z}(x_I) \cdot \theta(x_1)}{\sum_{x_1 x_I \in A} \sigma^{I,B,z}(x_I) \cdot \theta(x_1)} \leq \frac{|A'|}{|A|} + n^{-67}.$$

Proof. Consider a partition of A into sets $A_1, A_2, \dots$ obtained by fixing the strings in $\{1\}$. This ensures that for all j, A_j are uniform and thick in the interval I. This also induces a partition of A' into corresponding sets $A'_1, A'_2, \dots$. Then we have

$$\frac{\sum_{x_1 x_I \in A'} \sigma^{I,B,z}(x_I) \cdot \theta(x_1)}{\sum_{x_1 x_I \in A} \sigma^{I,B,z}(x_I) \cdot \theta(x_1)} = \frac{\sum_{x_1} \theta(x_1) \sum_{x_I \in A'_{x_1}} \sigma^{I,B,z}(x_I)}{\sum_{x_1} \theta(x_1) \sum_{x_I \in A_{x_1}} \sigma^{I,B,z}(x_I)}.$$

Lemma 8 allows us to conclude

$$\frac{\sum_{x_1 x_I \in A'} \sigma^{I,B,z}(x_I) \cdot \theta(x_1)}{\sum_{x_1 x_I \in A} \sigma^{I,B,z}(x_I) \cdot \theta(x_1)} \leq (1 + n^{-67}) \frac{\sum_{x_1} \theta(x_1)(|A'_{x_1}| + n^{-67}|A_{x_1}|)}{\sum_{x_1} \theta(x_1)|A_{x_1}|}$$

$$= (1+n^{-67})\frac{\sum_{x_1}|A'_{x_1}| + n^{-67}\sum_{x_1}|A_{x_1}|}{\sum_{x_1}|A_{x_1}|} = (1+n^{-67})\frac{|A'| + n^{-67}|A|}{|A|}.$$

This completes the proof.

□

Finally, we prove a Flattening Lemma, which decomposes a set into uniform sets. This is helpful during the communication steps of the protocol Π.

Lemma 10 (Flattening Lemma)**.** *Given a set $A \subset [m]^n$ that is uniform for interval I and a subset $A' \subset A$, there exists a partition of A' into disjoint sets $A_1, A_2, \ldots$ such that for all k, $|(A_k)_I| = \frac{|A_I|\cdot|A'|}{|A|}$ and A_k are uniform sets in the interval I.*

Proof. For any string $x \in [m]^n$, let $p_A(x) = \frac{1}{|A|}$ if $x \in A$ and 0 otherwise, and let $p_{A'}(x) = \frac{1}{|A'|}$ if $x \in A'$ and 0 otherwise. Let

$$p'_A(x_I) := \sum_{x_{-I}\in A_{-I}} p_A(x_{-I}x_I) \quad \text{and} \quad p'_{A'}(x_I) := \sum_{x_{-I}\in A_{-I}} p_{A'}(x_{-I}x_I).$$

Since $A' \subset A$, we have that $p_{A'}(x) \le \frac{|A|}{|A'|}p_A(x)$, which implies that

$$p'_{A'}(x_I) \le \frac{|A|}{|A'|}p'_A(x_I) = \frac{|A|}{|A'|\cdot|A_I|}$$

for all x_I, as A is uniform in I. Thus, the min-entropy of $p'_{A'}$ is at least $\log\frac{|A'|\cdot|A_I|}{|A|}$. Hence, $p'_{A'}$ can be written as $p'_{A'} = \sum_k q_k U_k$, where U_k is uniform in I and has size equal to $\frac{|A'|\cdot|A_I|}{|A|}$. From this, we can construct sets $A_1, A_2, \ldots$ for each of U_1, U_2 respectively as follows. For any string x_I, consider the set of strings x_{-I} such that $x_I x_{-I} \in A'$. Let this set be $\mathrm{Ext}(x_I)$. Let $\mathrm{Pos}(x_I)$ be the set of all indices k such that $U_k(x_I) > 0$. Append some arbitrary $\frac{q_k}{\sum_{k'\in\mathrm{Pos}(x_I)} q_k}$ fraction of strings[2] from $\mathrm{Ext}(x_I)$ to x_I and put it in A_k. Continue this way for all x_I. Note that $(A_k)_I = U_k$ and hence the lemma follows. □

3 Proof of main result

We restate Theorem 1 and provide its proof below.

Theorem 11. *For all (partial) functions f, it holds that*

$$\mathrm{R}^{\mu}_{1/4}(f \circ \mathrm{Index}^n_m) \ge \Omega(\mathrm{R}_{1/3}(f)\cdot\log n),$$

where $m = \mathrm{poly}(n)$.

Proof. For a given function f, recall the definition of λ (hard distribution for f) and μ (lifted distribution for $f \circ \mathrm{Index}^n_m$) from Section 2. Let $\mathcal{T}$ be a deterministic communication tree for f achieving $\mathrm{D}^{\mu}_{1/4}(f\circ\mathrm{Index}^n_m)$. Let $k := \mathrm{D}^{\mu}_{1/4}(f\circ\mathrm{Index}^n_m)$ be the depth of $\mathcal{T}$. Using our algorithm Π given in Figure 1, we get a randomized query protocol for f which makes an error of at most $\frac{1.1}{4}$ under λ (as implied by Lemma 13) and makes at most $O(k/\log n)$ expected number of queries (as implied by Lemma 17). This can be converted into an algorithm with $O(k/\log n)$

[2]We assume here that this is a natural number. The addition in our final error bound due this assumption is $\frac{1}{\mathrm{poly}(n)}$ since the set A' is thick whenever the Flattenig Lemma is invoked.

number of queries (in the worst case) and distributional error $\frac{1}{3}$, using standard application of Markov's inequality. This shows that

$$\mathrm{R}_{\frac{1}{3}}(f) = \mathrm{D}^{\lambda}_{\frac{1}{3}}(f) \leq O\left(\frac{k}{\log n}\right).$$

$\square$

We construct a tree which represents the evolution of the algorithm Π and is helpful in our analysis. The steps have been depicted in Figures 2 and 3. The first tree we construct is $\mathcal{T}_{\mathcal{AB}}$ which represents the evolution of the sets A, B as they partition into smaller sets, but with extra care for query steps of the algorithm. There will be two types of nodes in the tree, the *regular nodes* and the *intermediate nodes*. The intermediate nodes shall not be counted in the tree depth, but they shall help with the analysis. All the nodes of the tree are labelled by subsets of $[m]^j, \mathrm{Bal}^j_m$ (where $j \in \{1, 2, \dots n\}$) and the current interval. The root node is $([m]^n, \mathrm{Bal}^n_m, [n])$ and the rest of the tree is constructed as follows. Consider the step 2(a.i.A) of the algorithm Π when A is going to be partitioned into A_1, A_2. Set the intermediate children (children which are intermediate nodes) of (A, B, I) to be $(A_1, B, I), (A_2, B, I)$ and assign the weights of the edges going from parent to children as $\Pr_\mu[A_1 \times B | A \times B], \Pr_\mu[A_2 \times B | A \times B]$. Now,consider the step 2(a.i.C). Let the children of intermediate nodes be the subsets $A_1, A_2, \dots$ obtained by the Flattening Lemma. Set these to be regular nodes and assign them the labels $(A_1, B, I), (A_2, B, I), \dots$. Same construction holds for step 3. In step 2(a.i.D), let the children be regular nodes obtained by Thickness Lemma. Consider the step 2(b.iii) and let the query at this step be done at i-th index. We shall consider a pair of regular children for every non aborting η and an aborted intermediate child for the aborting η. Fixing an η, let the partitions of A be A_1, A_2 (corresponding to $U(\eta, 0), U(\eta, 1)$). Set the children to be $(A_1, B_\eta, I \backslash \{i\}), (A_2, B_\eta, I \backslash \{i\})$ and label the edges with η, probability of obtaining η conditioned on $A \times B$, and probability of the child conditioned on fixing η. For the step 2(a.ii) that involves partition of B, the children are regular nodes and weights of edges are associated probabilities. This finishes the construction of the tree.

For a node $(A, B, I) \in \mathcal{T}_{\mathcal{AB}}$, let $\mathrm{Par}((A, B, I))$ represent the regular parent node of (A, B, I) and $\mathrm{Chil}((A, B, I))$ represent the set of regular children nodes of (A, B, I) (ignoring the intermediate nodes, if they existed). With some abuse of notation (as it shall be clear from the context), $\mathrm{Par}_{\mathcal{A}}(A)$ shall represent the set A associated to the parent node, $\mathrm{Par}_{\mathcal{B}}(B)$ shall represent the set B associated to the parent node and $\mathrm{Chil}_{\mathcal{I}}(I)$ represent the set of sets in the children node. Similarly, we shall consider $\mathrm{Par}_{\mathcal{A}}(I)$ and $\mathrm{Chil}_{\mathcal{A}}(I)$. We shall also use the notation $\mathrm{IntPar}((A, B, I))$ as intermediate parent of a node (if they exist) and $\mathrm{IntChil}((A, B, I))$ as the set of intermediate children (if they exist). A level in tree shall be represented as $t > 0$, with $t = 1$ representing the root node. In any level t, let $\mathcal{R}(t)$ represent set of all regular nodes which are at distance t from the root, ignoring all the intermediate nodes as we count. Let $\mathcal{Q}(t)$ (or query nodes) represent the set of all nodes at level t which were obtained from their parent through the step 2(b). Let $\mathcal{C}(t)$ (or non-query nodes) represent the set of remaining nodes at level t. Let the regular nodes that did not abort for set A at level t be $\mathcal{N}_{abort}(t)$.

Note that the depth of the tree $\mathcal{T}_{\mathcal{AB}}$ is at most $O(n \log n)$, as (without loss of generality) there are at most $O(n \log n)$ communication steps in $\mathcal{T}$ and at most n query steps and constant number of operations for each of these steps in the algorithm Π.

Error analysis of algorithm Π

We first show the following lemma, which states some conditions that remain invariant under our algorithm.

Lemma 12 (Invariance Lemma)**.** *Throughout the execution of the algorithm* Π*, we show the following invariant:*

1. Initialize v as root of the protocol tree $\mathcal{T}$, initialize the interval $I = [n]$, Alice's part of rectangle $A = [m]^n$ and Bob's part of rectangle $B = \mathrm{Bal}_m^n$.
2. While v is not a leaf do:
   (a) If $\mathrm{AvgDeg}_i(A_I) \geq m^{19/20}$ for all $i \in I$:
      i. If Alice sends a bit at v:
         A. Pick $b \in \{0, 1\}$ with probability $\Pr_\mu[(A \cap X^{v,b}) \times B | A \times B]$. If $\frac{|(A \cap X^{v,b})|}{|A|} \leq \frac{1}{n^2}$, for the picked b, then Abort.
         B. Set $v \leftarrow v_b$
         C. Apply the process as in Flattening Lemma to partition $A \cap X^{v,b}$ into disjoint sets $A_1, A_2, \ldots$ which are uniform in interval I. Set $A \leftarrow A_k$ with probability $\Pr_\mu[A_k \times B | (A \cap X^{v,b}) \times B]$.
         D. Apply the Thickness Lemma to get A' satisfying $|A'| > \left(1 - \frac{1}{n^2}\right)|A|$ and A' is thick for I. Set $A \leftarrow A'$ with probability $\Pr_\mu[A' \times B | A \times B]$. Set $A \leftarrow A \setminus A'$ with probability $1 - \Pr_\mu[A' \times B | A \times B]$ and Abort.
      ii. If Bob sends a bit at v:
         A. Pick b with probability $\Pr_\mu[A \times (B \cap Y^{v,b}) | A \times B]$. If $\frac{|(B \cap Y^{v,b})_I|}{2^{m|I|}} \leq 2^{-n^2}$ for the picked b, then Abort.
         B. Set $v \leftarrow v_b$ and $B \leftarrow B \cap Y^{v,b}$.
   (b) If $\mathrm{AvgDeg}_i(A_I) < m^{19/20}$ for some $i \in I$:
      i. Query z_i.
      ii. For an $\eta \in \mathrm{Bal}_m$, select η with probability $\Pr_\mu[A \times B_\eta | A \times B]$. If the picked $\eta \in \mathrm{Bad}(A, B, i)$, then Abort.
      iii. Set $B \leftarrow B_{\eta,i}$, $A \leftarrow A|_{U(\eta, z_i), i}$.
      iv. Set $I \leftarrow I \backslash \{i\}$.
3. Apply the process as in Flattening Lemma to partition A into disjoint sets $A_1, A_2, \ldots$. Set $A \leftarrow A_k$ with probability $\Pr_\mu[A_k \times B | A \times B]$. Apply Thickness Lemma to get A' satisfying $|A'| > \left(1 - \frac{1}{n^2}\right)|A|$ and A' is thick for I. Set $A \leftarrow A'$ with probability $\Pr_\mu[A' \times B | A \times B]$ and Abort with probability $1 - \Pr_\mu[A' \times B | A \times B]$ by setting $A \leftarrow A \setminus A'$.
4. Assume $I = \{1, 2, \ldots |I|\}$ (without loss of generality for the procedure described here). Set $i = 1$.
5. While $i \leq |I|$ do:
   (a) For an $\eta \in \mathrm{Bal}_m$, select η with probability $\Pr_\mu[A \times B_\eta | A \times B]$. If the picked $\eta \in \mathrm{Bad}(A, B, i)$, then Abort.
   (b) Set $B \leftarrow B_{\eta,i}$, $i \leftarrow i + 1$.
6. Output as $\mathcal{T}$ does on the leaf v.

Figure 1: Randomized query algorithm Π for f.

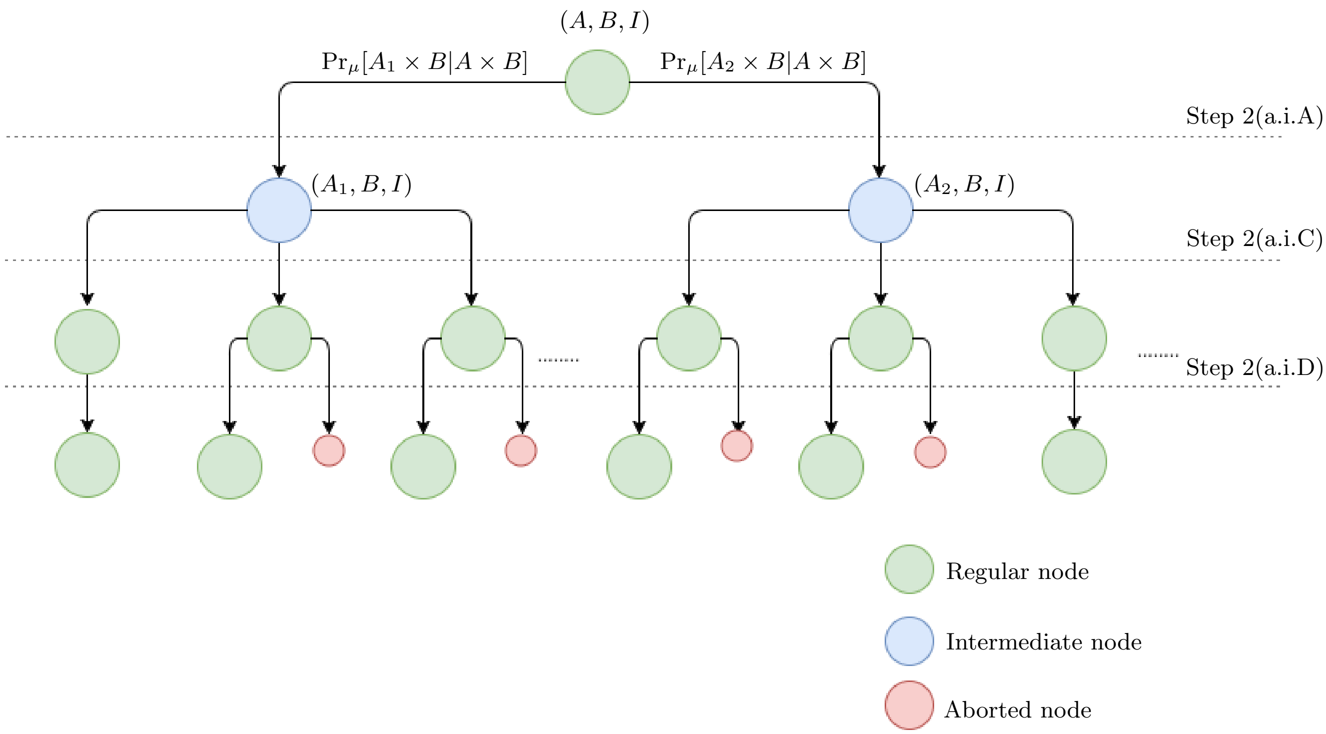

Figure 2: Structure of tree after Alice's communication step of the algorithm. Regular nodes after Step 2(a.i.C) are uniform in interval I. Regular nodes after Step 2(a.i.D) are uniform and thick in interval I. Aborted nodes after Step 2(a.i.D) are small in size.

1. *A is uniform in the current interval I, after execution of step 2(a.i.D). Same holds after the execution of step 3.*
2. *A is thick in the current interval I, after execution of step 2(a.i.D). Same holds after the execution of step 3.*
3. *$|B_I| \geq 2^{m|I|-n^2}$ for the current interval I for nodes of $\mathcal{T}_{\mathcal{AB}}$.*

Proof. 1. After the partition of rectangle A in the step 2(a.i.A), step 2(a.i.C) ensures uniformity. After step 2(a.i.D), the uniformity is maintained since the strings are removed from the current interval. Similar arguments apply after the execution of step 3.

2. On reaching step 2.(a.i.A), it holds that $\mathrm{AvgDeg}_i(A_I) \geq m^{19/20}$ for all $i \in I$. If Step 2(a.i.A) did not abort, then $|A \cap X^{v,b}|/|A| \geq 1/n^2$. The application of Flattening Lemma in step 2(a.i.C) gives new sets such that for each $A_k \subset A$, $|(A_k)_I| \geq |A_I|/n^2$. Thus, $\mathrm{AvgDeg}_i((A_k)_I) \geq m^{19/20}/n^2$ for all $i \in I$, from Claim 3. Then applying Thickness Lemma in step 2(a.i.D), we obtain a subset A'_k of A_k satisfying $\mathrm{MinDeg}_i((A'_k)_I) \geq m^{19/20}/n^4 \geq m^{17/20}$ for all $i \in I$. Thus A'_k is thick for I after execution of step 2(a.i.D). Similar arguments apply after the execution of step 3.

3. The item follows since a B which does not satisfy this condition is aborted.

$\square$

Now we are in a position to do the error analysis.

Lemma 13. *The algorithm Π makes an error of at most $1/4 + O(\log n/n)$.*

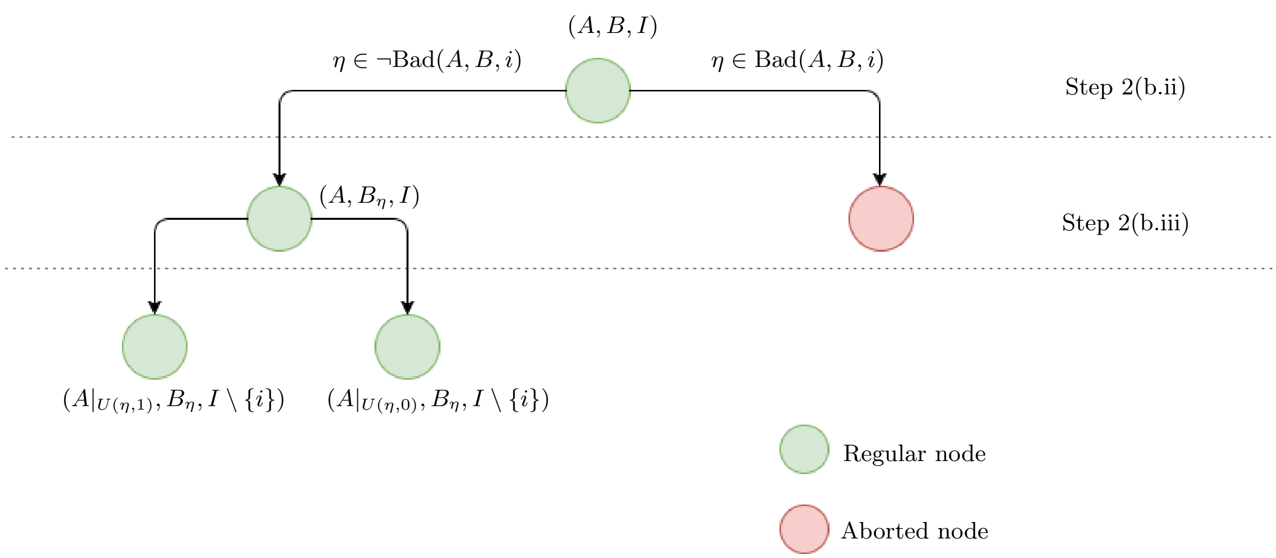

Figure 3: Structure of tree after query step of the algorithm. Queried index is i.

Proof. We begin with computing the overall probability of abort. We will first compute the probability of abort associated to A subsets. We start with step 2(a.i,A). For a node $(A, B, I) \in \mathcal{T}_{\mathcal{AB}}$, consider the quantity $|(A \cap X^{v,b})|/|A|$. It is upper bounded by $\frac{1}{n^2}$ if there is abort. Thus, child of (A, B, I) aborts only if size of $\mathrm{Chil}_{\mathcal{A}}(A)$ is smaller than $\frac{1}{n^2}$ times the size of A. We appeal to Uniformity Lemma 8 and its corollary 9 (it can be verified that the conditioned required for the lemma and corollary are satisfied for the considered sets) to conclude that the probability of transition from parent to child is upper bounded by

$$\Pr_{\mu}[(A \cap X^{v,b}) \times B | A \times B] \leq |(A \cap X^{v,b})|/|A| + n^{-67} \leq n^{-2}.$$

Similar argument holds for steps 2.(a.i.D) and 3. As noted before, the tree $\mathcal{T}_{\mathcal{AB}}$ has depth at most $O(n \log n)$. Hence, we obtain that the overall probability to abort is at most $O(\log n/n)$, appealing to Claim 5.

Marginalizing over Alice, we now compute the probability of abort associated to B subsets. At the steps 2(b.ii) and 5(a), the sampled η is from $\mathrm{Bad}(A, B, i)$ with probability at most $2^{-n^{35}} + 2^{-n^2}$. At the steps 2(a.ii.A), the abort occurs if $\frac{|B \cap Y^{v,b}|}{2^{m|I|}} \leq 2^{-n^2}$. Without loss of generality, the depth of $\mathcal{T}$ is at most $O(n \log n)$, resulting in at most $2^{O(n \log n)}$ subsets of Bal^n_m. Thus the overall probability of abort of this form is at most $2^{O(n \log n)} \cdot 2^{-n^2} \leq 2^{-n}$.

Thus by union bound, the overall probability of aborting in the algorithm is at most $O(\log n/n)$.

Conditioned on non abort, we proceed as follows. We consider the algorithm when it reaches a leaf L in $\mathcal{T}$ and does not abort. Let I be the set of corresponding un-queried bits (dropping the index L) and let the queried string be $z^*_{n/\{I\}}$. Without loss of generality, let $I = \{1, 2, \ldots |I|\}$. Let the rectangle corresponding to this leaf be $A \times B$. We have the following claim.

Claim 14. *For all $r < |I|$ and the sequence $y_1, y_2, \ldots y_r$, the y_{r+1} drawn from $B_{y_1, y_2, \ldots y_r}$ belongs to $\neg\mathrm{Bad}(A, B, i)$ and $|B_{y_1, y_2, \ldots y_r}| \geq 2^{m(|I|-r)-n^2}$.*

Proof. The property that y_{r+1} drawn from $B_{y_1, y_2, \ldots y_r}$ belongs to $\neg\mathrm{Bad}(A, B, i)$ is guaranteed by Step 5. To lower bound the size, we consider the base case $r = 0$. Then $|B| \geq 2^{m|I|-n^2}$ from the non-abort condition. Moreover, y_1 belongs to $\neg\mathrm{Bad}(A, B, 1)$, which, by definition, implies that $|B_{y_1}| \geq 2^{m|I|-n^2-m} = 2^{m(|I|-1)-n^2}$. Continuing this way, the claim follows. $\square$

We start with the following distribution over the strings $z \in \{0,1\}^n$:

$$\rho^L(z) = \sum_{(x,y)\in A\times B:G(x,y)=z} \frac{1}{|A|\cdot|B|}.$$

Note that $\rho^L(z)$ is only supported on those strings z for which $z_{[n]/I} = z^*_{[n]/I}$. We shall establish the following claim.

Claim 15. *For the strings z such that $z_{[n]/I} = z^*_{[n]/\{I\}}$, it holds that*

$$\rho^L(z) \in \frac{1}{2^{|I|}}\left[1-2n^{-4}, 1+2n^{-4}\right].$$

Proof. We shall keep track of three invariant properties: A is thick in the interval I, A is uniform in the interval I and $|B| \geq 2^{m|I|-n^2}$. Note that these conditions are true after the execution of step 3, as argued in Invariance Lemma 12. We start with computing

$$\rho^L(z_1) = \frac{1}{|B|}\sum_{y\in B}\sum_{x_{-1}} \frac{|A_{x_{-1}}|}{|A|}\frac{|(A|_{U(y_1,z_1)})_{x_{-1}}|}{|A_{x_{-1}}|}.$$

From Partition Lemma (which we can apply due to our invariant), it holds that

$$\frac{|(A|_{U(y_1,z_1)})_{x_{-1}}|}{|A_{x_{-1}}|} \in \left[\frac{1}{2}-n^{-5}, \frac{1}{2}+n^{-5}\right]$$

for every x_{-1}, y. Thus, we conclude that $\rho^L(z_1) \in [\frac{1}{2}-n^{-5}, \frac{1}{2}+n^{-5}]$.

Now we proceed to compute $\rho^L_{z_2|z_1}$. Fix a x_1, y_1 such that $\mathrm{Index}_m(x_1, y_1) = z_1$. Since A is thick in the interval I, A_{x_1} is thick as well in the interval $I/\{1\}$ (using Claim 4). On the other hand, A_{x_1} is uniform in the interval $I/\{1\}$ (as otherwise A would not have been uniform in interval I). From Claim 14, we have that $|B_{y_1}| \geq 2^{m(|I|-1)-n^2}$. Thus, we have maintained the invariant properties that we started with. Hence, we can apply the Partition Lemma again to obtain that $\rho^L_{z_2|z_1} \in [\frac{1}{2}-n^{-5}, \frac{1}{2}+n^{-5}]$.

Continuing in the same fashion, we see that the invariant properties continue to be maintained. Hence by recursive application of the Partition Lemma, we obtain the claim. $\square$

A corollary of this claim is the following.

Corollary 16. *Consider the distribution $\tau^L(z)$ over strings z conditioned on reaching the leaf. It holds that $\tau^L(z) \in \lambda(z|z_{n/\{I\}}) \cdot [1-4\cdot n^{-4}, 1+4\cdot n^{-4}] \cdot \delta(z^*_{n/\{I\}} = z_{n/\{I\}})$.*

Proof. We have that

$$\tau^L(z) = \frac{\sum_{x,y\in A\times B:G(x,y)=z}\mu(x,y)}{\sum_{x,y\in A\times B}\mu(x,y)} = \frac{\lambda(z)\sum_{x,y\in A\times B:G(x,y)=z}\frac{1}{|A||B|}}{\sum_z \lambda(z)\sum_{x,y\in A\times B:G(x,y)=z}\frac{1}{|A||B|}}.$$

From above claim, we have that

$$\sum_{x,y\in A\times B:G(x,y)=z}\frac{1}{|A||B|} \in \frac{1}{2^{|I|}}[1-2\cdot n^{-4}, 1+2\cdot n^{-4}]$$

as long as $z_{n/\{I\}} = z^*_{n/\{I\}}$. Thus, we find that

$$\tau^L(z) \in \lambda(z|z_{n/\{I\}}) \cdot [1 - 4 \cdot n^{-4}, 1 + 4 \cdot n^{-4}] \cdot \delta(z^*_{n/\{I\}} = z_{n/\{I\}}).$$

This proves the corollary □

Thus we conclude that conditioned on being on a non-aborting leaf L, τ^L differs by at most n^{-4} (in trace distance) with λ conditioned on $z_{n/\{I\}} = z^*_{n/\{I\}}$.

Furthermore, probability of our algorithm reaching a non-abort leaf is as according to the probability in the tree $\mathcal{T}$. This can be seen as follows: by construction of the algorithm, our probabilities of transition into sub-rectangles are as directed by $\mathcal{T}$ during all the non-query steps of the algorithm. During the query steps, the probability of transition of our algorithm to the event corresponding to query outcome z_i is as directed by λ. However, it can be argued identically along the lines of Claim 15 (using Invariance Lemma 12) that this transition probability, up to a multiplicative error of $1 \pm n^{-5}$ for each query step, is as directed by $\mathcal{T}$. Thus, overall error due to this discrepancy is at most n^{-4}.

Since $\mathcal{T}$ made an error of at most 1/4, the query algorithm makes an error of at most $1/4 + 2n^{-4} + O(\log n/n) \leq 1/4 + O(\log n/n)$. This proves the lemma. □

Expected number of queries of Π

We prove the following lemma.

Lemma 17. *The algorithm Π makes at most $\frac{2k}{5 \log n}$ expected number of queries, where the depth of the tree $\mathcal{T}$ is k.*

Proof. We will consider a potential function at the level t as follows, where the quantities are computed with respect to the tree $\mathcal{T}_{\mathcal{AB}}$. For every node $(A, B, I) \in \mathcal{T}_{\mathcal{AB}}$, let the probability of this node, computed by summing over all the conditional probabilities from root to the node be $\Pr(A, B)$. Define

$$P(t) = \sum_{(A,I) \in \mathcal{N}_{abort}(t)} \Pr(A, B) \log\left(\frac{m^{|I|}}{|A_I|}\right). \tag{1}$$

Consider the change in the potential function as the level t progresses.

$$\begin{aligned}
P(t+1)-P(t) &= \sum_{(A,B,I)\in\mathcal{N}_{abort}(t)} \Pr(A,B)\log\left(\frac{m^{|I|}}{|A_I|}\right) - \sum_{(A,B,I)\in\mathcal{N}_{abort}(t)} \Pr(A,B)\log\left(\frac{m^{|I|}}{|A_I|}\right)\\
&\overset{a}{=} \sum_{(A,B,I)\in\mathcal{C}(t+1)\cap\mathcal{N}_{abort}(t)} \Pr(A,B)\log\left(\frac{m^{|I|}}{|A_I|}\right) - \sum_{(A,B,I)\in\mathrm{Par}(\mathcal{C}(t+1))\cap\mathcal{N}_{abort}(t)} \Pr(A,B)\log\left(\frac{m^{|I|}}{|A_I|}\right)\\
&+ \sum_{(A,B,I)\in\mathcal{Q}(t+1)\cap\mathcal{N}_{abort}(t)} \Pr(A,B)\log\left(\frac{m^{|I|}}{|A_I|}\right) - \sum_{(A,B,I)\in\mathrm{Par}(\mathcal{Q}(t+1))\cap\mathcal{N}_{abort}(t)} \Pr(A,B)\log\left(\frac{m^{|I|}}{|A_I|}\right)\\
&\overset{b}{\leq} \sum_{(A,B,I)\in\mathcal{C}(t+1)\cap\mathcal{N}_{abort}(t)} \Pr(A,B)\log\left(\frac{|\mathrm{Par}_{\mathcal{A}}(A)_{\mathrm{Par}_{\mathcal{A}}(I)}|}{|A_I|}\right)\\
&+ \sum_{(A,B,I)\in\mathcal{Q}(t+1)} \Pr(A,B)\log\left(\frac{|\mathrm{Par}_{\mathcal{A}}(A)_{\mathrm{Par}_{\mathcal{A}}(I)}|m^{|I|}}{|A_I|m^{|\mathrm{Par}_{\mathcal{I}}(I)|}}\right)\\
&\overset{c}{=} \sum_{(A,B,I)\in\mathcal{C}(t+1)\cap\mathcal{N}_{abort}(t)} \Pr(A,B)\log\left(\frac{|\mathrm{Par}_{\mathcal{A}}(A)_I|}{|A_I|}\right)\\
&+ \sum_{(A,B,I)\in\mathcal{Q}(t+1)} \Pr(A,B)\log\left(\frac{|\mathrm{Par}_{\mathcal{A}}(A)_{\mathrm{Par}_{\mathcal{A}}(I)}|}{|A_I|m}\right), \qquad (2)
\end{aligned}$$

where in (a), we note that parents of non-aborted nodes cannot be aborted nodes. The inequality (b) holds by partitioning the probability of a parent node as a sum of probabilities of its child nodes. An inequality comes instead of an equality since we use the fact that a parent of a non aborted node may have an aborted intermediate child. In (b), we have also used the fact that in query step, every child of a node belongs to $\mathcal{N}_{abort}$ and in any non-query step, the size of I does not change. Equality (c) follows since in a query step, the size of I decreases by 1.

Lets consider the second to last expression above. In a step that involves partitioning of a B set, there is no change in the expression in logarithm. Thus, we consider only the steps involving partitioning of A sets. We begin with steps 2(a.i.A)-2.a(i.C). Consider a node (A^*, B) and its intermediate children $(\mathrm{IntChil}(A^*)_1, B), (\mathrm{IntChil}(A^*)_2, B)$ in a non-query step. Let the collection of A's obtained from $\mathrm{IntChil}(A^*)_1$ and $\mathrm{IntChil}(A^*)_2$ by the Flattening Lemma be $\mathcal{C}_1$ and $\mathcal{C}_2$ respectively. Define $\frac{1}{p} := \frac{|A^*|}{|\mathrm{IntChil}(A^*)_1|}$. For an $A \in \mathcal{C}_1$, define $\frac{1}{q_1} := \frac{|A^*_I|}{|A_I|}$, which is the same for all $A \in \mathcal{C}_1$ as argued in the Flattening Lemma. Similarly define q_2. This allows us to conclude that for $j \in \{1, 2\}$,

$$\sum_{A:A\in\mathcal{C}_j} \Pr(A,B)\log\left(\frac{|A^*_I|}{|A_I|}\right) = \Pr(\mathrm{IntChil}(A^*)_j, B)\log\left(\frac{|A^*_I|}{|A_I|}\right),$$

where the A appearing on the right hand side belongs to $\mathcal{C}_j$. From the statement of the Flattening Lemma, we have that $\frac{1}{q_1} \leq \frac{1}{p}$ and $\frac{1}{q_2} \leq \frac{1}{1-p}$. Thus, we conclude that for $j \in \{1, 2\}$,

$$\sum_{A:A\in\mathcal{C}_j} \Pr(A,B)\log\left(\frac{|A^*_I|}{|A_I|}\right) \leq \Pr(\mathrm{IntChil}(A^*)_j, B)\log\left(\frac{|A^*|}{|\mathrm{IntChil}(A^*)_j|}\right).$$

Now using Uniformity Lemma 8, we have the upper bound $\Pr[A \times B | A^* \times B] \leq \frac{|A|}{|A^*|} + n^{-67} \leq (1+n^{-60})\frac{|A|}{|A^*|}$, due to the fact that $\frac{|A|}{|A^*|} \geq \frac{1}{n^2}$ (else the intermediate child was an aborted child). Thus, the contribution of A^* to second to last expression in Equation 2 is upper bounded by

$$(1+\frac{1}{n^{60}}) \cdot (p\log\frac{1}{p} + (1-p)\log\frac{1}{1-p}) \leq 1.1$$

With this we conclude,

$$\begin{aligned}\sum_{(A,B,I)\in\mathcal{C}(t+1)\cap\mathcal{N}_{abort}(t+1)} \Pr(A)\log\left(\frac{|\mathrm{Par}_{\mathcal{A}}(A)_{\mathrm{Par}_{\mathcal{I}}(I)}|}{|A_I|}\right) &\leq 1.1 \sum_{(A,B,I)\in\mathcal{C}(t+1)} \Pr(A,B)\\ &= 1.1 \sum_{(A^*,B,I)\in\mathcal{C}^*(t)} \Pr(A^*,B),\end{aligned}$$

where by $\mathcal{C}^*(t)$, we represent the set of parent nodes at level t that are partitioned via a non-query step. Same analysis holds for the step 2.(a.i.D).

Now we evaluate the last expression in Equation 2. Let the queried index be i. Since AvgDeg_i has reduced below $m^{19/20}$ and $\mathrm{Par}_{\mathcal{I}}(I) = I \cup \{i\}$, we conclude that $\frac{|\mathrm{Par}_{\mathcal{A}}(A)_{\mathrm{Par}_{\mathcal{I}}(I)}|}{|A_I|} \leq m^{19/20}$. This implies

$$\sum_{(A,B,I)\in\mathcal{Q}(t+1)} \Pr(A,B)\log\left(\frac{|\mathrm{Par}_{\mathcal{A}}(A)_{\mathrm{Par}_{\mathcal{I}}(I)}|}{m|A_I|}\right) \leq \left(-\frac{1}{20}\log m\right) \sum_{(A^*,B,I)\in\mathcal{Q}^*(t)} \Pr(A^*,B),$$

where the set of parent nodes that are partitioned via step 2(b) at level t are represented by $\mathcal{Q}^*_{\mathcal{A}}(t)$.

Finally, we observe that $\sum_{(A^*,B,I)\in\mathcal{C}^*(t)} \Pr(A^*,B)$ is the expected number of communication steps taken by Alice at step t (which is upper bounded by k) and $\sum_{(A^*,B,I)\in\mathcal{Q}^*(t)} \Pr(A^*,B)$ is the expected number of query steps to be taken at step t. Let the last step of the algorithm be last. Using telescopic sum, we have that

$$P(\mathrm{last}) - P(1) = \sum_{t=1}^{\mathsf{last}-1} P(t+1) - P(t).$$

Thus, setting q to be the expected number of queries made by the algorithm, we have that

$$P(\mathsf{last}) - P(1) \leq 2.2k - \frac{q}{20}\cdot\log m.$$

But $P(1) = 0$ and $P(\mathsf{last}) > 0$. Thus setting the value of m, we have that $q \leq \frac{2.2}{5\log n}\cdot k$. This proves the lemma.

□

Acknowledgement

This work is supported by the Singapore Ministry of Education and the National Research Foundation, also through the Tier 3 Grant Random numbers from quantum processes MOE2012-T3-1-009.

References

[ABBD+16] A. Anshu, A. Belovs, S. Ben-David, M. Gs, R. Jain, R. Kothari, T. Lee, and M. Santha. Separations in communication complexity using cheat sheets and information complexity. In *2016 IEEE 57th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 555–564, Oct 2016.

[GLM+15] Mika Göös, Shachar Lovett, Raghu Meka, Thomas Watson, and David Zuckerman. Rectangles are nonnegative juntas. In *Proceedings of the Forty-seventh Annual ACM Symposium on Theory of Computing*, STOC '15, pages 257–266, New York, NY, USA, 2015. ACM.

[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, Oct 2015.

[Raz99] Ran Raz. Exponential separation of quantum and classical communication complexity. In *Proceedings of the Thirty-first Annual ACM Symposium on Theory of Computing*, STOC '99, pages 358–367, New York, NY, USA, 1999. ACM.

[RM99] Ran Raz and Pierre McKenzie. Separation of the monotone nc hierarchy. *Combinatorica*, 19(3):403–435, 1999.

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

A Proof of Uniformity Lemma 8

We shall use the following claim for the proof.

Claim 18. *Let $I \subset [n]$ be an interval and $B \in \mathrm{Bal}_m^{|I|}$ be a set such that $\frac{|B|}{|\mathrm{Bal}_m|^{|I|}} \geq 2^{-k}$ for some $k \geq n^2$. Then there exists a set of indices J with $|J| \geq m|I| - 2n^4 \cdot k$, such that for all $j \in J$, $\frac{|B_{y_j}|}{|B|} \in [\frac{1}{2} \pm \frac{1}{n^2}]$, for $y_j \in \{0,1\}$.*

Proof. Let $Y := Y_1Y_2 \dots Y_{m|I|}$ be the random variable with associated probability distribution $\Pr_Y(y) = \frac{1}{|B|}$ if $y \in B$ and 0 otherwise. We have that

$$H(Y) = \log|B| \geq |I| \log|\mathrm{Bal}_m| - k \geq m|I| - k - n\log m \geq m|I| - 2k.$$

Hence, $\sum_j H(Y_j) \geq H(Y) \geq m|I| - 2k$. Since $H(Y_j) \leq 1$, this can be re-written as $\sum_j (1 - H(Y_j)) \leq 2k$. By Markov's ineqaulity, the number of indices for which $H(Y_j) < 1 - \frac{1}{n^4}$ is at most $n^4 \cdot k$. Let the rest of indices form the set J. By our construction, for every $j \in J$, we have $H(Y_j) \geq 1 - \frac{1}{n^4}$. Thus, $\Pr[Y_j = 0] \in [\frac{1}{2} \pm \frac{1}{n^2}]$. Since $\Pr[Y_j = 0] = \frac{|\{y : y_j = 0, y \in B\}|}{|B|}$, and same argument applies for $Y_j = 1$, the claim follows. ∎

Proof of Lemma 8. Below, we shall drop the subscript from A_I, for brevity.

We start with the observation that

$$\sigma^{I,B,z}(x'_I) = \frac{\sum_{x,y : x_I = x'_I, G(x,y) = z, y \in B} \mu(x,y)}{\sum_{x,y : G(x,y) = z, y \in B} \mu(x,y)} = \frac{|\{y_I : G(x'_I, y_I) = z_I, y_I \in B_I\}|}{\sum_{x_I} |\{y_I : G(x_I, y_I) = z_I, y_I \in B_I\}|},$$

which follows since B is fixed in the interval $[n] \setminus I$. Without loss of generality, let $I = \{1, 2, \dots |I|\}$.

We will compute the distribution $\sigma^{I,B,z}(x'_1)$. For this, consider

$$\sum_{x'_{-1}} |\{y_I : G(x'_I, y_I) = z_I, y_I \in B_I\}| = \sum_{y_I \in B} |\{x'_{I \setminus \{1\}} : G(x'_1 x'_{I \setminus \{1\}}, y_I) = z_I\}|$$

$$= (\frac{m}{2})^{|I|-1} |\{y_I \in B : \mathrm{Index}_m(x'_1, y_1) = z_1\}|,$$

where second equality holds since for every y_I which is balanced, the number of $x_{I \setminus \{1\}}$ giving a particular $z_{I \setminus \{1\}}$ is $(\frac{m}{2})^{|I|-1}$. Thus, we find that

$$\sum_{x'_{-1}} |\{y_I : G(x'_I, y_I) = z_I, y_I \in B_I\}| = (\frac{m}{2})^{|I|-1} |B_{(z_1, x'_1)}|,$$

where $B_{(z_1,x_1')}$ is the set $\{y \in B : y_{x_1'} = z_1\}$. This allows us to conclude:

$$\sigma^{I,B,z}(x_1') = \frac{|B_{(z_1,x_1')}|}{\sum_{x_1} |B_{(z_1,x_1)}|}.$$

Similarly, we conclude that

$$\sigma^{I,B,z}(x_2'|x_1) = \frac{|(B_{(z_1,x_1)})_{(z_2,x_2')}|}{\sum_{x_2} |(B_{(z_1,x_1)})_{(z_2,x_2)}|},$$

and so on.

Now we introduce some notations. We let $x_1 \in \mathrm{Good}_1$ if it holds that $\frac{|B_{(z_1,x_1)}|}{|B|} \leq \frac{1}{2} + n^{-2}$. Otherwise we let $x_1 \in \mathrm{Bad}_1$. Similarly, for a fixed x_1, let $x_2 \in \mathrm{Good}_2(x_1)$ if it holds that $\frac{|(B_{(z_1,x_1)})_{(z_2,x_2)}|}{|B_{(z_1,x_1)}|} \leq \frac{1}{2} + n^{-2}$. Otherwise let $x_2 \in \mathrm{Bad}_2(x_1)$. Continuing in this fashion, we define the sets $\mathrm{Good}_r(x_1, x_2, \ldots x_{r-1})$ and $\mathrm{Bad}_r(x_1, x_2, \ldots x_{r-1})$ for any $r \leq |I|$.

Furthermore, we set $x_1 \in \mathrm{Large}_1$ if it holds that $\frac{|B_{(z_1,x_1)}|}{|B|} \geq m^{-|I|+1}/3$. We define the set $\mathrm{Large}_2(x_1)$ as the set of all x_2 for which $\frac{|(B_{z_1,x_1})_{(z_2,x_2)}|}{|B_{(z_1,x_1)}|} \geq m^{-|I|+1}/3$. Continuing in this fashion, we define the sets $\mathrm{Large}_r(x_1, x_2, \ldots x_{r-1})$ for any $r \leq |I|$.

Finally, we say that $x_1 x_2 \ldots x_{|I|} \in \mathrm{Good}$ if it holds that $x_1 \in \mathrm{Good}_1$ and $x_2 \in \mathrm{Good}_2(x_1)$ and so on or if it holds that there exists an r such that $x_r \in \neg\mathrm{Large}_r(x_1 x_2 \ldots x_{r-1})$. We say that $x_1 x_2 \ldots x_{|I|} \in \mathrm{Bad}$ if either '$x_1 \in \mathrm{Bad}_1$' or '$x_2 \in \mathrm{Bad}_2(x_1)$ and $x_1 \in \mathrm{Large}_1 \cap \mathrm{Good}_1$', or '$x_3 \in \mathrm{Bad}_3(x_1 x_2)$ and $x_2 \in \mathrm{Large}_2(x_1) \cap \mathrm{Good}_2(x_1)$ and $x_1 \in \mathrm{Large}_1 \cap \mathrm{Good}_1$' and so on. We say that $x_1 x_2 \ldots x_{|I|} \in \mathrm{VeryGood}$ if it holds that $x_1 \in \mathrm{Good}_1 \cap \mathrm{Large}_1$ and $x_2 \in \mathrm{Good}_2(x_1) \cap \mathrm{Large}_2(x_1)$ and so on.

We are now in a position to prove the following claim.

Claim 19. *Fix a $x_I := x_1 x_2 \ldots x_{|I|}$. Following properties hold.*

1. $\sigma^{I,B,z}(x_I) \leq \frac{2^{|I|}(1+2n^{-5})}{m^{|I|}}$.
2. *If $x_I \in \mathrm{Good}$, then $\sigma^{I,B,z}(x_I) \leq \frac{1+4n^{-5}}{m^{|I|}}$. If $x_I \in \mathrm{VeryGood}$, then in addition we have $\sigma^{I,B,z}(x_I) \geq \frac{1-4n^{-5}}{m^{|I|}}$.*
3. *If $x_I \in \mathrm{Bad}$, then we have $|\mathrm{Bad}_1| \leq 2n^6$, $|\mathrm{Bad}_2(x_1)| \leq 2n^6 + 2n^5 \log m$, $|\mathrm{Bad}_3(x_1, x_2)| \leq 2n^6 + 3n^5 \log m, \ldots \mathrm{Bad}_{|I|}(x_1 x_2 \ldots x_{|I|-1}) \leq 2n^6 + |I| n^5 \log m \leq 2n^6 \log m$.*
4. *If $x_1 \in \mathrm{VeryGood}$, then we have $|\neg\mathrm{Good}_1| \leq 2n^6$, $|\neg\mathrm{Good}_2(x_1)| \leq 2n^6 + 2n^5 \log m$, $\neg\mathrm{Good}_3(x_1 x_2)| \leq 2n^6 + 3n^5 \log m, \ldots |\neg\mathrm{Good}_{|I|}(x_1 x_2 \ldots x_{|I|-1})| \leq 2n^6 \log m$.*

Proof. The claim is established by an inductive argument. Below, we show the first two steps of the argument.

Claim 18 ensures that number of $x_1 \in \mathrm{Good}_1$ such that $\frac{|B_{(z_1,x_1)}|}{|B|} \geq \frac{1}{2} - n^{-2}$ is at least $m - 2n^6$. Thus,

$$(m + 2n^6)|B|/2 = (m - 2n^6)|B|/2 + 2n^6|B| \geq \sum_{x_1} |B_{(z_1,x_1)}| \geq (m - 2n^6)|B|/2$$

and $|\mathrm{Bad}_1| \leq 2n^6$. This gives,

$$\frac{|B_{(z_1,x_1)}|}{m|B|} \cdot (2 - 3n^{-6}) \leq \sigma^{I,B,z}(x_1) \leq \frac{|B_{(z_1,x_1)}|}{m|B|} \cdot (2 + 3n^{-6})$$

for all x_1.

If x_1 is such that $\frac{|B_{(z_1,x_1)}|}{|B|} \leq m^{-|I|+1}/3$ (that is, $x_1 \in \neg\mathrm{Large}_1$), then we automatically have $\sigma^{I,B,z}(x_1 x_{I\setminus\{1\}}) \leq \sigma^{I,B,z}(x_1) \leq m^{-|I|}$ for all $x_{I\setminus\{1\}}$. So we restrict ourselves to $x_1 \in \mathrm{Large}_1$, for which $|B_{(z_1,x_1)}| \geq |\mathrm{Bal}_m|^{|I|} 2^{-n^2-n\log m}$.

Claim 18 once again ensures that the number of $x_2 \in \mathrm{Good}_2(x_1)$ is at least $m - 2(n^6 + n^5 \log m)$. Thus, $\sum_{x_2} |(B_{(z_2,x_2)})_{(z_1,x_1)}| \geq (m - 2n^6 - 2n^5\log m)|B_{(z_1,x_1)}|/2$. This gives,

$$\frac{|(B_{(z_2,x_2)})_{(z_1,x_1)}|}{m|B_{(z_1,x_1)}|} \cdot (2-4n^{-6}) \leq \sigma^{I,B,z}(x_2|x_1) \leq \frac{|(B_{(z_2,x_2)})_{(z_1,x_1)}|}{m|B_{(z_1,x_1)}|} \cdot (2+4n^{-6}).$$

We also conclude that $|\mathrm{Bad}_2(x_1)| \leq 2n^6 + 2n^5 \log m$. Moreover, we find the following expression for $\sigma^{I,B,z}(x_1x_2)$:

$$\frac{1}{m^2}(2-4n^{-6})^2 \cdot \frac{|B_{(z_1,x_1)}|}{|B|} \cdot \frac{|(B_{(z_2,x_2)})_{(z_1,x_1)}|}{|B_{(z_2,x_2)}|} \leq \sigma^{I,B,z}(x_1x_2) \leq \frac{1}{m^2}(2+4n^{-6})^2 \cdot \frac{|B_{(z_1,x_1)}|}{|B|} \cdot \frac{|(B_{(z_2,x_2)})_{(z_1,x_1)}|}{|B_{(z_2,x_2)}|},$$

for any $x_1 \in \mathrm{Large}_1$. In particular, if $x_1 \in \mathrm{Good}_1$ and $x_2 \in \mathrm{Good}_2(x_1)$, then $\frac{1}{2} - n^{-2} \leq \frac{|B_{(z_1,x_1)}|}{|B|} \leq \frac{1}{2} + n^{-2}$ and $\frac{1}{2} - n^{-2} \leq \frac{|(B_{(z_2,x_2)})_{(z_1,x_1)}|}{|B_{(z_1,x_1)}|} \leq \frac{1}{2} + n^{-2}$. Thus, we obtain the bounds $\frac{1}{m^2} \cdot (1 + 6 \cdot n^{-6}) \leq \sigma^{I,B,z}(x_1x_2) \leq \frac{1}{m^2} \cdot (1 + 6 \cdot n^{-6})$.

The argument proceeds similarly in an inductive fashion. We finally conclude that if $x_1 \in \mathrm{Large}_1$ and $x_2 \in \mathrm{Large}_2(x_1)$ and $x_3 \in \mathrm{Large}_3(x_1x_2)$,..., then

$$\sigma^{I,B,z}(x_1x_2\ldots x_{|I|}) \in \frac{1}{m^{|I|}}(2 \pm 4n^{-6})^{|I|} \cdot \frac{|B_{(z_1,x_1)}|}{|B|} \cdot \frac{|(B_{(z_2,x_2)})_{(z_1,x_1)}|}{|B_{(z_1,x_1)}|} \cdots.$$

This gives a general upper bound on $\sigma^{I,B,z}(x_1x_2\ldots x_{|I|}) \leq \frac{1}{m^{|I|}}(2+4n^{-6})^{|I|}$, proving item 1.

In addition, if $x_1 \in \mathrm{Good}_1, x_2 \in \mathrm{Good}_2(x_1), \ldots$, then the bound is improved to $\sigma^{I,B,z}(x_1x_2\ldots x_{|I|}) \in \frac{1 \pm 2n^{-5}}{m^{|I|}}$. If $x_1 \in \neg\mathrm{Large}_1$ or $x_2 \in \neg\mathrm{Large}_2(x_1)$ or so on, then by definition $\sigma^{I,B,z}(x_1x_2\ldots x_{|I|}) \leq m^{-|I|}$. This proves the item 2. Items 3 and 4 have been shown above by upper bounding the sizes of 'Bad' sets. This proves the claim.

□

Now, we are in a position to prove the items of the lemma. We are automatically using the uniformity of A in I to restrict the computation to the interval I.

1. Consider

$$\sum_{x_1x_2\ldots x_{|I|} \in A} \sigma^{I,B,z}(x_1x_2\ldots x_{|I|}) \geq \sum_{x_1x_2\ldots x_{|I|} \in A\cap\mathrm{VeryGood}} \frac{1-4n^{-5}}{m^{|I|}} = \frac{1-4n^{-5}}{m^{|I|}}|A\cap\mathrm{VeryGood}|.$$

For any $x_1x_2\ldots x_{|I|-1}$, the number of strings $x_{|I|}$ such that $x_1x_2\ldots x_{|I|-1}x_{|I|} \in A$ is at least $|A_{x_1x_2\ldots x_{|I|-1}}|$. By Claim 19, we further have that number of strings $x_{|I|}$ such that $x_1x_2\ldots x_{|I|-1}x_{|I|} \in A \cap \mathrm{VeryGood}$ is at least $|A_{x_1x_2\ldots x_{|I|-1}}| - 2n^6\log m$. Since $|A_{x_1x_2\ldots x_{|I|-1}}| \geq m^{17/20}$, this allows us to conclude that

$$\sum_{x_{|I|}} \mathbb{I}(x_1x_2\ldots x_{|I|} \in A \cap \mathrm{VeryGood}) \geq (1 - \frac{2n^6\log m}{m^{17/20}}) \sum_{x_{|I|}} \mathbb{I}(x_1x_2\ldots x_{|I|} \in A),$$

where $\mathbb{I}(.)$ is the indicator function. Thus, we have removed the constraint on $x_{|I|}$ at the cost of a multiplicative factor of $(1 - \frac{2n^6\log m}{m^{17/20}})$. We can continue this way for $x_{|I|-1}$ (constraints on which depend only on $x_1, x_2 \ldots x_{|I|-2}$) and so on, to obtain

$$|A \cap \mathrm{VeryGood}| \geq (1 - \frac{2n^6\log m}{m^{17/20}})^n |A| \geq (1 - \frac{2n^7\log m}{m^{17/20}})|A|.$$

This proves the item by substituting the value of m.

2. Consider

$$\begin{aligned}\sum_{x_1x_2\ldots x_{|I|}\in A'} \sigma^{I,B,z}(x_1x_2\ldots x_{|I|}) &= \sum_{x_1x_2\ldots x_{|I|}\in A'\cap \mathrm{Good}} \sigma^{I,B,z}(x_1x_2\ldots x_{|I|}) \\ &\quad + \sum_{x_1x_2\ldots x_{|I|}\in A'\cap \mathrm{Bad}} \sigma^{I,B,z}(x_1x_2\ldots x_{|I|}) \\ &\le \sum_{x_1x_2\ldots x_{|I|}\in A'\cap \mathrm{Good}} \frac{1+4n^{-5}}{m^{|I|}} + \sum_{x_1x_2\ldots x_{|I|}\in A'\cap \mathrm{Bad}} \frac{2^{|I|}(1+2n^{-5})}{m^{|I|}} \\ &\le \frac{1+4n^{-5}}{m^{|I|}}(|A'| + 2^{|I|}|A\cap \mathrm{Bad}|)\end{aligned}$$

Now, we upper bound $|A\cap \mathrm{Bad}|$. We recall the definition of the set Bad and that A is thick and uniform in I. As in the definition of the set Bad, suppose $x_1x_2\ldots x_{|I|}$ is such that $x_1 \in \mathrm{Bad}_1$. Using the relation $|\mathrm{Bad}_1| \le 2n^6$ from Claim 19, consider

$$\sum_{x_1} \mathbb{I}(x_1x_2\ldots x_{|I|} \in A, x_1 \in \mathrm{Bad}_1) \le \frac{2n^6}{m^{17/20}} \sum_{x_1} \mathbb{I}(x_1x_2\ldots x_{|I|} \in A),$$

which holds since A is thick. Similarly, suppose $x_1 \in \mathrm{Good}_1 \cap \mathrm{Large}_1$ but $x_2 \in \mathrm{Bad}_2(x_1)$. In such a scenario, Claim 19 ensures that $|\mathrm{Bad}_2(x_1)| \le 2n^6 \log m$. Thus we conclude in a similar fashion that

$$\sum_{x_1} \mathbb{I}(x_1x_2\ldots x_{|I|} \in A, x_1 \in \mathrm{Good}_1\cap\mathrm{Large}_1, x_2 \in \mathrm{Bad}_2(x_1)) \le \frac{2n^6\log m}{m^{17/20}} \sum_{x_1} \mathbb{I}(x_1x_2\ldots x_{|I|} \in A).$$

We can continue this way for all other conditions present in the definition of Bad. Since there are at most n such conditions, we conclude that $|A\cap \mathrm{Bad}| \le \frac{2n^7 \log m}{m^{17/20}}|A|$. Thus, we obtain the upper bound

$$\sum_{x_1x_2\ldots x_{|I|}\in A'} \sigma^{I,B,z}(x_1x_2\ldots x_{|I|}) \le \frac{1+4n^{-5}}{m^{|I|}}(|A'| + \frac{n^8}{m^{17/20}}|A|).$$

This establishes the item, by substituting the value of m.

Thus, the lemma concludes.

□

ECCC ISSN 1433-8092

https://eccc.weizmann.ac.il