Electronic Colloquium on Computational Complexity, Revision 2 of Report No. 54 (2017)

Lifting randomized query complexity to randomized communication complexity

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

July 1, 2017

Abstract

We show that for a relation $f \subseteq \{0,1\}^n \times \mathcal{O}$ and a function $g : \{0,1\}^m \times \{0,1\}^m \to \{0,1\}$ (with $m = O(\log n)$),

$$\mathrm{R}_{1/3}(f \circ g^n) = \Omega\left(\mathrm{R}_{1/3}(f) \cdot \left(\log \frac{1}{\mathrm{disc}(M_g)} - O(\log n)\right)\right),$$

where $f \circ g^n$ represents the composition of f and g^n, M_g is the sign matrix for g, $\mathrm{disc}(M_g)$ is the *discrepancy* of M_g under the uniform distribution and $\mathrm{R}_{1/3}(f)$ ($\mathrm{R}_{1/3}(f \circ g^n)$) denotes the randomized query complexity of f (randomized communication complexity of $f \circ g^n$) with worst case error $\frac{1}{3}$.

In particular, this implies that for a relation $f \subseteq \{0,1\}^n \times \mathcal{O}$,

$$\mathrm{R}_{1/3}(f \circ \mathrm{IP}_m^n) = \Omega\left(\mathrm{R}_{1/3}(f) \cdot m\right),$$

where $\mathrm{IP}_m : \{0,1\}^m \times \{0,1\}^m \to \{0,1\}$ is the Inner Product (modulo 2) function and $m = O(\log(n))$.

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

*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

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]). Raz and McKenzie [RM99] (and the generalization due to [GPW15]) showed that for a relation $f \subseteq \{0,1\}^n \times \mathcal{O}$, the deterministic communication complexity of f composed with Index_m (with $m = \mathrm{poly}(n)$) is at least the deterministic query complexity of f times $\Omega(\log n)$. 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, for example lifting *approximate-degree* to *approximate-rank* [She11], *approximate junta-degree* to *smooth-corruption-bound* [GLM+15] and, more recently, randomized query complexity to randomized communication complexity using the Index_m function [GPW17] (for $m = \mathrm{poly}(n)$).

Our result

In this work we show lifting of (bounded error) randomized query complexity to (bounded error) randomized communication complexity. Let $\mathrm{R}_\varepsilon(\cdot)$ denote the randomized query/communication complexity, as is clear from the context, with worst case error ε. We show the following:

Theorem 1. *Let $f \subseteq \{0,1\}^n \times \mathcal{O}$ be a relation and $g : \{0,1\}^m \times \{0,1\}^m \to \{0,1\}$ be a function (with $m = O(\log n)$). It holds that*

$$\mathrm{R}_{1/3}(f \circ g^n) = \Omega\left(\mathrm{R}_{1/3}(f) \cdot \left(\log \frac{1}{\mathrm{disc}(M_g)} - O(\log n)\right)\right),$$

where M_g is the sign matrix for g and $\mathrm{disc}(M_g)$ is the discrepancy of M_g under the uniform distribution.

In particular, this implies that for a relation $f \subseteq \{0,1\}^n \times \mathcal{O}$,

$$\mathrm{R}_{1/3}(f \circ \mathrm{IP}_m^n) = \Omega\left(\mathrm{R}_{1/3}(f) \cdot m\right),$$

where $\mathrm{IP}_m : \{0,1\}^m \times \{0,1\}^m \to \{0,1\}$ is the Inner Product (modulo 2) function and $m = O(\log(n))$. In comparison, the Index_m function used by [GPW17] required $m = \mathrm{poly}(n)$.

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

Our techniques

Our techniques are partly 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 T be a deterministic communication protocol tree for $f \circ g^n$ (with $\log \frac{1}{\mathrm{disc}(M_g)} = \Omega(\log n)$). We use this to create a randomized query protocol Π (see Algorithm 2) for f. Let z be an input for which we are supposed to output a b such that $(z,b) \in f$. We start with the root of T and continue to simulate T (using randomness) till we find a co-ordinate $i \in [n]$ where T has *worked enough* so that $g(x_i, y_i)$ is becoming (only slightly) *determined*. Using the properties of g we conclude that T must have communicated $O(\log \frac{1}{\mathrm{disc}(M_g)})$ bits by now. We go ahead and 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 T consistent with z_i.

To keep the unqueried bits z_i (with i belonging to the unqueried interval I) sufficiently *undetermined*, we keep track of how much T has worked to determine x_i (or y_i), using the conditional probability of x_i (or y_i), given all possible $x_{I\setminus\{i\}}$ (or $y_{I\setminus\{i\}}$) at other unqueried locations. When either of the conditional probabilities $p_A(x_i|x_{I\setminus\{i\}})$ or $p_B(y_i|y_{I\setminus\{i\}})$ (where

A, B are current rectangles) becomes too high for a sufficiently large number of strings, we conclude that a query must be made.

Lets suppose that the conditional probability at some location becomes high in A. We only want to make a query in that part of A where the conditional probability violation takes place. This eventually lets us compare the number of queries we make with the number of bits communicated. So within a query subroutine, we first probabilistically split A into the strings HIGH where the conditional probability becomes too high, and $A \setminus$ HIGH, where this does not happen. A query is then made in the HIGH part, and only the x_i and y_i that are consistent with the z_i (that we learn from the query) are retained and partitioned into a collection of rectangles. After this, the conditional probability can be restored to a low enough value for the rest of the indices, and we can move on with communication steps.

As long as we have a bound on the conditional probabilities in the unqueried locations, the unqueried z_i are sufficiently undetermined and we can move from node to node of T according to the "flow" of T, for every input z. We prove this in Lemma 12 in Section 3. This lets our algorithm to sample the leaves of T close to the desired probabilities, and thus the correctness of T on (x, y) in expectation ensures the correctness of our algorithm on z in expectation.

During the course of our simulation, we may end up at some "bad" subevents, where we will not be able to maintain a sufficiently large number of (x, y) consistent with z. We need to abort the algorithm on such subevents. When we have a bound on the conditional probability and we are going with the "flow" of T, we can ensure that the probability of going to such bad subevents is small. But, if we need to do a series of queries in one go, we will not be able to maintain the requisite bound on the conditional probabilities in between queries. So it may be possible that when we do a query and try to synchronize x_i and y_i with z_i, we do not find any (or sufficiently many) x_i and y_i that are consistent with z_i. In the technical Lemma 7 we show that there is a way around this: if we do some "preprocessing" on A before carrying out queries, the probability of this bad subevent happening is still small. The arguments here are similar to showing that g is a good *strong extractor* for *blockwise sources* with appropriate conditional min-entropy in each block. Thus the algorithm aborts with small probability. Similar arguments hold for B.

2 Preliminaries

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

Let $f \subseteq \{0,1\}^n \times \mathcal{O}$ be a relation. 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 be an integer. Let $g : \{0,1\}^m \times \{0,1\}^m \to \{0,1\}$ be a function, interpreted as a *gadget*. Let $G := g^n$. Define the following distributions:

$$\mu^0(x, y) := \frac{\mathbb{1}_{g(x,y)=0}}{|g^{-1}(0)|}, \quad \mu^1(x, y) := \frac{\mathbb{1}_{g(x,y)=1}}{|g^{-1}(1)|}.$$

For every $z \in \{0,1\}^n$, define $\mu^z := \mu^{z_1} \times \mu^{z_2} \times \ldots \times \mu^{z_n}$. The *lifted* distribution for the composed relation $f \circ g^n$ is $\mu := \mathbb{E}_{z\sim\lambda}\mu^z$.

Let Alice and Bob's inputs for the composed function be respectively $x = (x_1, \ldots, x_n) \in (\{0,1\}^m)^n$ and $y = (y_1, \ldots, y_n) \in (\{0,1\}^m)^n$.

We use the following notation, some of which is adapted from notation used in [GPW15].

- For a node v in a communication protocol tree, let $X^v \times Y^v$ 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 send b.
- For a string $x \in (\{0,1\}^m)^n$ 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 a string $y \in (\{0,1\}^m)^n$.
- For a set $A \subset (\{0,1\}^m)^n$, let $A_I := \{x_I : x \in A\}$ be the restriction of A to the interval I and $A_{x_I} := \{x' \in A : x'_I = x_I\}$. Our convention is for an $x_I \notin A_I$, A_{x_I} is the null set. We use similar notation for B.
- For $A \subseteq (\{0,1\}^m)^n$, we represent the uniform probability distribution on strings x in A with $p_A(x)$. We use similar notation for B.
- For $A \subseteq (\{0,1\}^m)^n$, an index $i \in I$ and $x_{I\setminus\{i\}} \in \{0,1\}^{m(|I|-1)}$, let

  $$p_{\max}(A, x_{I\setminus\{i\}}) := \max_{x_i \in \{0,1\}^m} p_A(x_i | x_{I\setminus\{i\}}).$$

  We say a $p_{\max}$ bound of α holds for A with respect to I, if $p_{\max}(A, x_{I\setminus\{i\}}) \leq \alpha$ for all $i \in I$ and all such $x_{I\setminus\{i\}}$. Similar terminology holds for B.
- For $A \subseteq (\{0,1\}^m)^n, I \subseteq [n]$, let

  $$\textsc{High}(A, \alpha, I) := \{x \in A : \exists i \in I, p_A(x_i | x_{I\setminus\{i\}}) > \alpha\}.$$

  Similarly, we define $\textsc{High}(B, \alpha, I)$.
- For $y_i \in \{0,1\}^m$ and $z_i \in \{0,1\}$, let $U(y_i, z_i) := \{x_i \in \{0,1\}^m : g(x_i, y_i) = z_i\}$. We use $A|_{U(y_i,z_i)}$ to denote $\cup_{x_i \in U(y_i,z_i)} A_{x_i}$. Similarly, for $x_i \in \{0,1\}^m$ and $z_i \in \{0,1\}$, let $V(x_i, z_i) := \{y_i \in \{0,1\}^m : g(x_i, y_i) = z_i\}$. We use $B|_{V(x_i,z_i)}$ to denote $\cup_{y_i \in V(x_i,z_i)} B_{y_i}$.
- Let M_g represent the sign matrix for the function g. That is, $M_g(x,y) = (-1)^{g(x,y)}$. For an interval I and $x_I, y_I \in \{0,1\}^{m|I|}$, we define $M_g^{\otimes|I|}(x_I, y_I) := \Pi_{i\in I} M_g(x_i, y_i)$. Observe that $M_g^{\otimes|I|}$ is also a sign matrix.

Discrepancy of M_g, defined below, gives a lower bound on the communication complexity of g. We consider it with respect to the uniform distribution, although it extends to any general distribution (see for example, [LSS08]).

Definition 3. *Discrepancy of M_g with respect to the uniform distribution is defined as*

$$\mathrm{disc}(M_g) := \frac{1}{2^{2m}} \max_{A\subseteq\{0,1\}^m, B\subseteq\{0,1\}^m} |\sum_{x\in A, y\in B} M(x,y)|.$$

This definition extends to the matrix $M_g^{\otimes r}$, for any integer $r > 1$, in a natural fashion. Following lemma follows from [LSS08, Theorems 16,17].

Lemma 4. *Let M_g be a sign matrix and $r > 1$ be an integer. Then it holds that*

$$\mathrm{disc}(M_g^{\otimes r}) \leq (8 \cdot \mathrm{disc}(M_g))^r.$$

For the rest of the proof, we define a parameter β as

$$\beta := \frac{1}{2} \log \frac{1}{\mathrm{disc}(M_g)}.$$

Our working assumption is that $\beta \geq 100 \log n$.

We shall use the sets $\textsc{Small}(A, A', I), \textsc{Small}(B, B', I), \textsc{UnBal}_{\mathcal{X}}(A, B, I)$ and $\textsc{UnBal}_{\mathcal{Y}}(A, B, I)$ defined in Lemmas 5 and 7 respectively in our algorithm and analysis.

Following lemma is similar to the Thickness Lemma in [GPW15].

Lemma 5. *For $A \subseteq (\{0,1\}^m)^n, I \subseteq [n]$ and $A' \subseteq A$, there exists $A'' \subseteq A'$ such that for all $i \in I$ and $x_{I \setminus \{i\}} \in A''_{I\setminus\{i\}}$, $|A''_{x_{I\setminus\{i\}}}| \geq \frac{1}{n^3}|A_{x_{I\setminus\{i\}}}|$, and*

$$p_A(A' \setminus A'') < \frac{1}{n^2}.$$

Define $\textsc{Small}(A, A', I) := A' \setminus A''$.

Similarly, for $B \subseteq (\{0,1\}^m)^n, I \subseteq [n]$ and $B' \subseteq B$, there exists $B'' \subseteq B'$ such that for all $i \in I$ and $y_{I \setminus \{i\}} \in B''_{I\setminus\{i\}}$, $|B''_{y_{I\setminus\{i\}}}| \geq \frac{1}{n^3}|B_{y_{I\setminus\{i\}}}|$, and

$$p_B(B' \setminus B'') < \frac{1}{n^2}.$$

Define $\textsc{Small}(B, B', I) := B' \setminus B''$.

Proof. We prove the statement for the sets A, A'. Similar argument holds for B, B'. The set A'' is obtained by running the following algorithm on A'. It is easy to see that the A'' obtained satisfies the property required.

Algorithm 1: Decomposing $A' = A'' \cup \textsc{Small}(A, A', I)$

Initialize $A^0 = A', j = 0$
while $|A^j_{x_{I\setminus\{i\}}}| < \frac{1}{n^3}|A_{x_{I\setminus\{i\}}}|$ *for some* $i \in I$ *and* $x_{I\setminus\{i\}} \in A^j_{I\setminus\{i\}}$ **do**
  Pick such an $i \in I$ and $x_{I\setminus\{i\}}$
  Set $A^{j+1} = A^j \setminus \{x' \in A^j : x'_{I\setminus\{i\}} = x_{I\setminus\{i\}}\}$
end
Output $A'' = A^j$

To bound the size of $\textsc{Small}(A, A', I)$, let $(i_j, x^j_{I\setminus\{i_j\}})$ be the pair picked by the algorithm in the j-th iteration. Then we have, $|A^j_{x_{I\setminus\{i_j\}}}| < \frac{1}{n^3}|A_{x_{I\setminus\{i_j\}}}|$. Note that a particular $x^j_{I\setminus\{i_j\}}$ can only be removed once in the algorithm, so at most all $x_{I\setminus\{i\}}$ for all $i \in I$ can be removed. So the total strings removed is at most

$$\sum_{(i_j, x^j_{I\setminus\{i_j\}})} |A^j_{x^j_{I\setminus\{i_j\}}}| < \frac{1}{n^3} \sum_{(i_j, x^j_{I\setminus\{i_j\}})} |A_{x^j_{I\setminus\{i_j\}}}| \leq \frac{1}{n^3} \sum_{i\in I} \sum_{x_{I\setminus\{i\}}} |A_{x_{I\setminus\{i\}}}| = \frac{|I|}{n^3}|A| \leq \frac{1}{n^2}|A|.$$

This proves the lemma. ∎

Lemma 6. *Fix a real number $k \in (0,1)$. Let $A, B \subseteq (\{0,1\}^m)^n$ and $I \subseteq [n]$ be such that $p_A(x_I) \leq 2^{-|I|(m-k\beta)}$ for all $x_I \in \{0,1\}^{m|I|}$ and $p_B(y_I) \leq 2^{-|I|(m-k\beta)}$ for all $y_I \in \{0,1\}^{m|I|}$. Then it holds that*

$$\left|\sum_{x_I, y_I} p_A(x_I) p_B(y_I) M_g^{\otimes|I|}(x_I, y_I)\right| \leq (8 \cdot 2^{-(2-2k)\beta})^{|I|}.$$

Proof. First we show that without loss of generality, we can assume that $p_A(x_I)$ and $p_B(y_I)$ are uniform in their support. The probability distributions $p_A(x_I), p_B(y_I)$ have min-entropy at least $|I|(m-k\beta)$. Thus, they can be decomposed as a convex combination of probability distributions having min-entropy at least $|I|(m-k\beta)$ which are uniform in their support. Thus, we write

$$p_A(x_I) = \sum_i \lambda_i p_A^i(x_I), \quad p_B(y_I) = \sum_j \mu_j p_B^j(y_I), \quad \text{with } \sum_i \lambda_i = \sum_j \mu_j = 1, \lambda_i \geq 0, \mu_j \geq 0.$$

We obtain by triangle inequality that

$$|\sum_{x_I,y_I} p_A(x_I)p_B(y_I)M_g^{\otimes|I|}(x_I,y_I)| \leq \sum_{i,j}\lambda_i\mu_j|\sum_{x_I,y_I} p_A^i(x_I)p_B^j(y_I)M_g^{\otimes|I|}(x_I,y_I)|.$$

Thus, the maximum value is attained at distributions that are uniform in their support.

Thus, assuming that $p_A(x_I)$ and $p_B(y_I)$ are uniform in their support, let A', B' be their respective supports. Since the min-entropy of $p_A(x_I)$ and $p_B(y_I)$ is at least $|I|(m-k\beta)$, we have that $|A'| \geq 2^{|I|(m-k\beta)}$ and $|B'| \geq 2^{|I|(m-k\beta)}$. Thus,

$$\begin{aligned}
|\sum_{x_I,y_I} p_A(x_I)p_B(y_I)M_g^{\otimes|I|}(x_I,y_I)| &= \frac{1}{|A'||B'|}|\sum_{x_I\in A',y_I\in B'} M_g^{\otimes|I|}(x_I,y_I)| \\
&= \frac{2^{2m|I|}}{|A'||B'|}\cdot\frac{1}{2^{2m|I|}}|\sum_{x_I\in A',y_I\in B'} M_g^{\otimes|I|}(x_I,y_I)| \\
&\leq \frac{2^{2m|I|}}{2^{2|I|(m-k\beta)}}\mathrm{disc}(M_g^{\otimes|I|}) \\
&\leq 2^{2|I|k\beta}(8\cdot\mathrm{disc}(M_g))^{|I|} \quad (\text{Lemma } 4) \\
&= (8\cdot 2^{-(2-2k)\beta})^{|I|}
\end{aligned}$$

This completes the proof. ∎

Lemma 7. *For $A, B \subseteq (\{0,1\}^m)^n$ suppose a $p_{\max}$ bound of $2^{-m+0.8\beta}$ holds with respect to $I \subseteq [n]$. We say $x \in \textsc{UnBal}_{\mathcal{X}}(A,B,I)$ if it does not satisfy the property*

$$\Pr_{y_I\sim p_B}[g^{|I|}(x_I,y_I)=z_I] \in \frac{1}{2^{|I|}}\left[1-2^{-0.05\beta}, 1+2^{-0.05\beta}\right] \qquad \forall z\in\{0,1\}^{|I|}.$$

We have,

$$p_A(\textsc{UnBal}_{\mathcal{X}}(A,B,I)) \leq 2^{-0.05\beta}.$$

Similarly, we say $y \in \textsc{UnBal}_{\mathcal{Y}}(A,B,I)$ if it does not satisfy the property

$$\Pr_{x_I\sim p_A}[g^{|I|}(x_I,y_I)=z_I] \in \frac{1}{2^{|I|}}\left[1-2^{-0.05\beta}, 1+2^{-0.05\beta}\right] \qquad \forall z\in\{0,1\}^{|I|}.$$

We have,

$$p_B(\textsc{UnBal}_{\mathcal{Y}}(A,B,I)) \leq 2^{-0.05\beta}.$$

Proof. We prove the first part. The second part follows similarly. Fix an interval $J \subseteq I$. Since a $p_{\max}$ bound of $2^{-m+0.8\beta}$ holds for the given A, we have that for all x_J, $p_A(x_J) \leq 2^{-|J|(m-0.8\beta)}$. Consider any subset $A'_J \subseteq A_J$ such that $p_A(A'_J) \geq 2^{-0.1|J|\beta}$. It holds that

$$p_A(x_J|A'_J) \leq 2^{0.1|J|\beta}2^{-|J|(m-0.8\beta)} = 2^{-|J|(m-0.9\beta)}.$$

Thus, invoking Lemma 6, we obtain

$$|\sum_{x_J,y_J} p_A(x_J|A'_J)p_B(y_J)M_g^{\otimes|J|}(x_J,y_J)| \le (8\cdot 2^{-0.2\beta})^{|J|}. \tag{1}$$

Let $\textsc{Bad}_J^{(1)}$ be the set of all $x_J \in A_J$ for which $\sum_{y_J} p_B(y_J)M_g^{\otimes|J|}(x_J,y_J) \ge (8\cdot 2^{-0.2\beta})^{|J|}$ and $\textsc{Bad}_J^{(0)}$ be the set of all $x_J \in A_J$ for which $\sum_{y_J} p_B(y_J)M_g^{\otimes|J|}(x_J,y_J) \le -(8\cdot 2^{-0.2\beta})^{|J|}$. Let $\textsc{Bad}_J := \textsc{Bad}_J^{(1)} \cup \textsc{Bad}_J^{(0)}$. From Equation 1, we conclude that $p_A(\textsc{Bad}_J^{(i)}) < 2^{-0.1\beta|J|}$ for $i \in \{0,1\}$. Thus,

$$p_A(\textsc{Bad}_J) \le p_A(\textsc{Bad}_J^{(1)}) + p_A(\textsc{Bad}_J^{(0)}) \le 2\cdot 2^{-0.1\beta|J|}.$$

Using $\beta \ge 100\log n$, we obtain

$$p_A(\cup_{J\subseteq I}\textsc{Bad}_J) \le \sum_{J\subseteq I} p_A(\textsc{Bad}_J) \le 2\sum_{r=1}^{|I|}\binom{|I|}{r}2^{-0.1r\beta} \le 2\sum_{r=1}^{n}2^{r\log n - 0.1r\beta} \le 2^{-0.05\beta}.$$

Now we show that $\textsc{UnBal}(A,B,I) \subseteq \cup_{J\subseteq I}\textsc{Bad}_J$. For this, we consider an x such that $x_J \in \neg\textsc{Bad}_J$ for all $J \subseteq I$. That is,

$$|\sum_{y_J} p_B(y_J)M^{\otimes|J|}(x_J,y_J)| \le (8\cdot 2^{-0.2\beta})^{|J|} \quad \text{for all } J\subseteq I.$$

Following claim shows that $x \notin \textsc{UnBal}(A,B,I)$, which completes the proof. This claim is a restatement of [GLM+15, Lemma 13].

Claim 8. *Consider an x satisfying*

$$|\sum_{y_J} p_B(y_J)M^{\otimes|J|}(x_J,y_J)| \le (8\cdot 2^{-0.2\beta})^{|J|} \quad \textit{for all } J\subseteq I. \tag{2}$$

It holds that

$$\Pr_{y_I\sim p_B}[g^{|I|}(x_I,y_I)=z_I] \in \frac{1}{2^{|I|}}\left[1-2^{-0.05\beta}, 1+2^{-0.05\beta}\right] \qquad \forall z\in\{0,1\}^{|I|}.$$

Proof. Fix an x satisfying Equation 2. Let $\chi_J(z) := (-1)^{\oplus_{j\in J}z_j}$ be the parity function. The fact that M_g is the sign matrix for g implies

$$|\sum_{y_J} p_B(y_J)\chi_J\left(g^{|J|}(x_J,y_J)\right)| \le (8\cdot 2^{-0.2\beta})^{|J|} \quad \text{for all } J\subseteq I. \tag{3}$$

Let $p(z_I)$ be the distribution of z_I for the given x and averaged over $y \sim p_B$, that is $p(z_I) = \Pr_{y_I\sim p_B}[g^{|I|}(x_I,y_I)=z_I]$. We fourier expand $p(z_I) := \sum_{J\subseteq I}\chi_J(z_I)\hat{p}(J)$, where

$$\hat{p}(J) = \frac{1}{2^{|I|}}\sum_{z_I} p(z_I)\chi_J(z_I) = \frac{1}{2^{|I|}}\sum_{y_J} p_B(y_J)\chi_J\left(g^{|J|}(x_J,y_J)\right).$$

From Equation 3, we have that $2^{|I|}|\hat{p}(J)| \le (8\cdot 2^{-0.2\beta})^{|J|}$. Furthermore, $\hat{p}(\phi) = \frac{1}{2^{|I|}}$, where ϕ is the empty set. Thus we conclude (using $\beta \ge 100\log n$)

$$\begin{aligned}
|p(z_I) - \frac{1}{2^{|I|}}| &= |\sum_{J\subseteq I, J\neq\phi}\chi_J(z_I)\hat{p}(J)| \\
&\le \frac{1}{2^{|I|}}\sum_{J\subseteq I, J\neq\phi}(8\cdot 2^{-0.2\beta})^{|J|} \\
&= \frac{(1+8\cdot 2^{-0.2\beta})^{|I|}-1}{2^{|I|}} \\
&\le \frac{2^{\log n - 0.1\beta}}{2^{|I|}} \le \frac{2^{-0.05\beta}}{2^{|I|}}.
\end{aligned}$$

This establishes the claim. □

□

3 Proof of main result

We show the following which implies Theorem 1.

Theorem 9. *Let $f \subseteq \{0,1\}^n \times \mathcal{O}$ be a relation and $g : \{0,1\}^m \times \{0,1\}^m \to \{0,1\}$ be a function (with $m = O(\log n)$). It holds that*

$$\mathrm{D}^{\mu}_{1/4}(f \circ g^n) = \Omega(\mathrm{R}_{1/3}(f) \cdot (\beta - 100 \log n)),$$

where $\beta = \frac{1}{2} \log \frac{1}{\mathrm{disc}(g)}$.

Proof. If $\beta \leq 100 \log n$, then the statement is trivially true. Thus, we assume $\beta > 100 \log n$. For a given relation f, recall the definition of λ (hard distribution for f) and μ (lifted distribution for $f \circ g^n$) from Section 2. Let T be a deterministic communication tree for f achieving $\mathrm{D}^{\mu}_{1/4}(f \circ g^n)$. Let $c := \mathrm{D}^{\mu}_{1/4}(f \circ g^n)$ be the depth of T. Using our algorithm Π given in Algorithm 2 and described in the form of a flowchart 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 10) and makes at most $O(c/(\beta - 3 \log n))$ expected number of queries (as implied by Lemma 16). This can be converted into an algorithm with $O(c/(\beta - 3 \log n))$ number of queries (in the worst case) and distributional error $\frac{1}{3}$, using standard application of Markov's inequality. This shows that

$$\mathrm{R}_{\frac{1}{3}}(f) = \mathrm{D}^{\lambda}_{\frac{1}{3}}(f) \leq O\left(\frac{c}{\beta - 3 \log n}\right),$$

which shows the desired. □

For an input z, we construct a tree $\mathcal{T}$ which represents the evolution of the algorithm Π, depending on the random choices made by it in steps 4, 24, 11, 31, 14, 34 and the Filter steps of Algorithm 2. All the nodes of the tree are labeled by unique triplets $(A \times B, I, v)$ where $I \subseteq [n]$ is the current interval, $A \subseteq (\{0,1\}^m)^n, B \subseteq (\{0,1\}^m)^n$ are the current parts of the rectangle held by Alice and Bob respectively, and v is the current node of T. The root node is $((\{0,1\}^m)^n \times (\{0,1\}^m)^n, [n], r)$ where r is the root of T, and the children of any node are all the nodes that can be reached from it depending on random choices made. Each edge is labeled by the conditional probability of the algorithm reaching the child node, conditioned on it reaching the parent node for that z. The overall probability of the algorithm reaching a node $(A \times B, I, v)$ on input z, denoted by $\Pr_{\mathcal{T},z}[(A \times B, I, v)]$ is obtained by multiplying all the conditional probabilities along the path from the root to $(A \times B, I, v)$.

Note that there are at most $O(n \log n)$ communication steps in T and at most n query steps in Π (along with a constant number of additional operations for each of these steps).

Error analysis of algorithm Π

In this section we shall prove the following main lemma.

Lemma 10. *The algorithm Π makes an error of at most $1/4 + O(\log n/n)$ when input z is drawn from λ.*

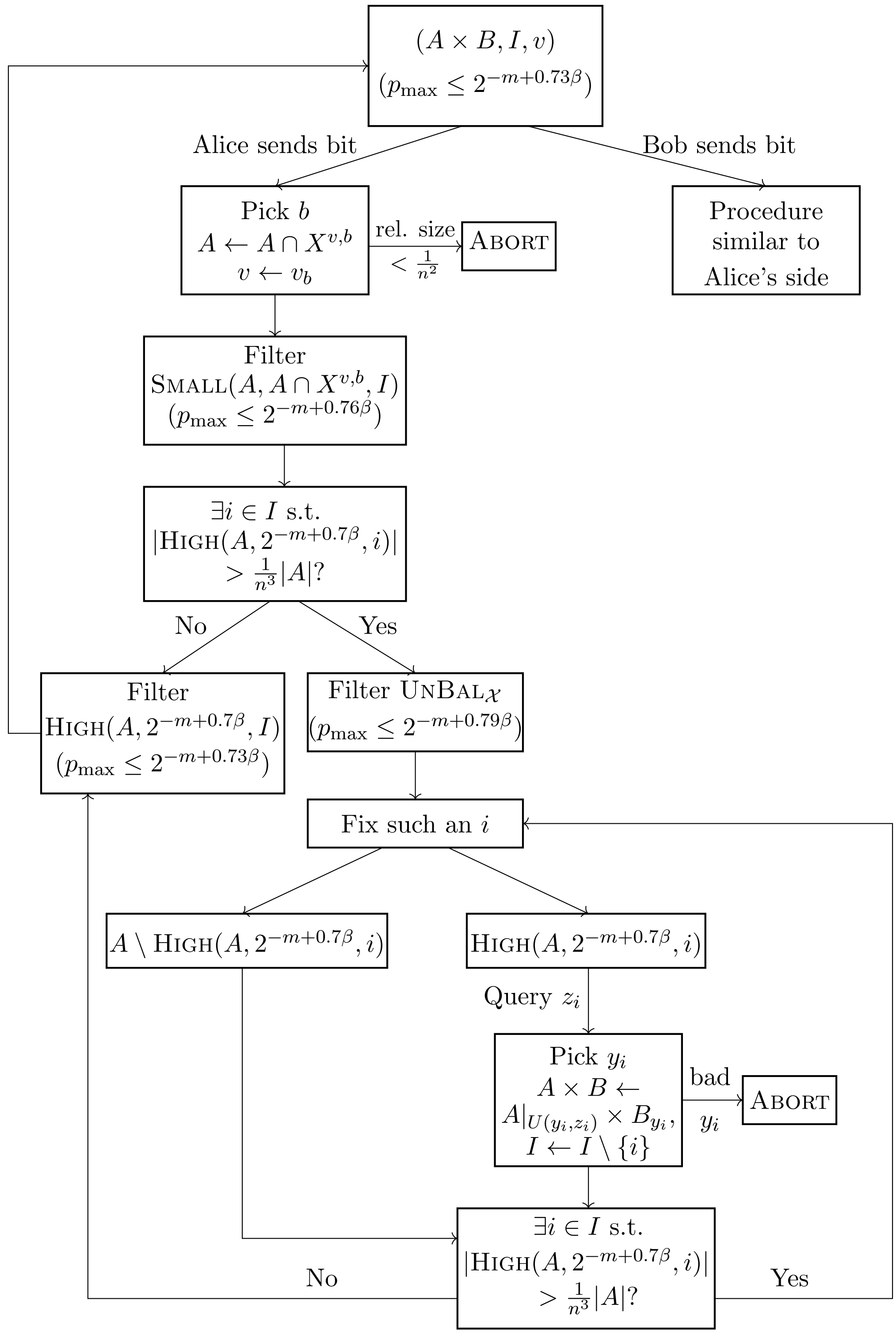

Figure 1: A flowchart description of the algorithm

The proof of this lemma has the following components. First, we prove an Invariance Lemma which will show that an appropriate $p_{\max}$ bound holds at right steps in the algorithm. Second, we show a lemma related to the size of $|(A \times B) \cap G(z)|$ relative to $|A \times B|$ at any node of $\mathcal{T}$ for any input z to Π. This lets us do two things: first we can show that the transition probabilities of Π are almost the same for any z that is consistent with the bits queried so far, and second, we can show that the probability of the algorithm going to an aborted node is small. This allows us to argue that the probability of Π reaching a leaf is close to the desired probability. Thus, the expected value of the error of Π over the leaves is close to the expected value of the error of T over the leaves, which is small.

Lemma 11 (Invariance Lemma)**.** *Throughout the execution of* Π*, we show the following invariants:*

1. *$p_{\max}$ bounds of $2^{-m+0.76\beta}, 2^{-m+0.79\beta}, 2^{-m+0.73\beta}$ and $2^{-m+0.73\beta}$ for the current A with respect to the current interval I hold after steps 27, 29, 41 and 16 respectively;*
2. *$p_{\max}$ bounds of $2^{-m+0.76\beta}, 2^{-m+0.79\beta}, 2^{-m+0.73\beta}$ and $2^{-m+0.73\beta}$ for the current B with respect to the current interval I hold after steps 7, 9, 21 and 36 respectively;*

Proof. We prove the statement for A. A similar argument holds for B.

1. **After step 27:** If Abort does not happen here, A is set to $A' = A^b \setminus A^b|_{\textsc{Small}(A,A^b,I)}$ (where we use A^b to denote $A \cap X^{v,b}$). For all $x \in A'$ and for all $i \in I$, Lemma 5 implies that $|A'_{x_{I\setminus\{i\}}}| \geq \frac{1}{n^3} \cdot |A_{x_{I\setminus\{i\}}}|$. Moreover, $|A'_{x_{I\setminus\{i\}}\circ x_i}| \leq |A_{x_{I\setminus\{i\}}\circ x_i}|$ for every $x_i, x_{I\setminus\{i\}}$. Since a $p_{\max}$ bound of $2^{-m+0.73\beta}$ holds for A (refer to the topmost node in Flowchart 1), we have,
$$\max_{x_i} \frac{|A'_{x_{I\setminus\{i\}}\circ x_i}|}{|A'_{x_{I\setminus\{i\}}}|} \leq \max_{x_i} \frac{n^3 \cdot |A_{x_{I\setminus\{i\}}\circ x_i}|}{|A_{x_{I\setminus\{i\}}}|} \leq 2^{-m+0.76\beta}.$$
Hence, the $p_{\max}$ bound of $2^{-m+0.76\beta}$ holds for A'.

   After step 29: Here an A for which a $p_{\max}$ bound of $2^{-m+0.76\beta}$ holds is set to A'. We have for every $x \in A'$ and $i \in I$, $|A'_{x_{I\setminus\{i\}}}| \geq \frac{1}{n^3}|A_{x_{I\setminus\{i\}}}|$. Moreover, $|A'_{x_{I\setminus\{i\}}\circ x_i}| \leq |A_{x_{I\setminus\{i\}}\circ x_i}|$ for all $x_i, x_{I\setminus\{i\}}$. So for any $x_{I\setminus\{i\}}$,
$$\max_{x_i} \frac{|A'_{x_{I\setminus\{i\}}\circ x_i}|}{|A'_{x_{I\setminus\{i\}}}|} \leq \max_{x_i} \frac{n^3 \cdot |A_{x_{I\setminus\{i\}}\circ x_i}|}{|A_{x_{I\setminus\{i\}}}|} \leq 2^{-m+0.79\beta}$$
due to the $p_{\max}$ bound on A.

   After step 41: A similar argument holds here. Since the strings in both $\textsc{High}(A, 2^{-m+0.7\beta}, I)$ and $\textsc{Small}(A, A \setminus \textsc{High}(A, 2^{-m+0.7\beta}, I), I)$ are removed, we have for the remaining strings in the set A',
$$\max_{x_i} \frac{|A'_{x_{I\setminus\{i\}}\circ x_i}|}{|A'_{x_{I\setminus\{i\}}}|} \leq \max_{x_i} \frac{n^3 \cdot |(A \setminus \textsc{High}(A, 2^{-m+0.7\beta}, I))_{x_{I\setminus\{i\}}\circ x_i}|}{|A_{x_{I\setminus\{i\}}}|} \leq 2^{-m+0.73\beta}$$
by the definition of $\textsc{High}(A, 2^{-m+0.7\beta}, I)$.

   After step 16: The set A is fixed to A_{x_i}, for some x_i and $i \in I$, after this step. Since a $p_{\max}$ bound of $2^{-m+0.73\beta}$ held before this step in interval I, and the interval $I \setminus \{i\}$ is a subset of I, the same $p_{\max}$ bound continues to hold.

$\square$

Thus, we conclude that at every step, the $p_{\max}$ bound is at most $2^{-m+0.8\beta}$. We will use this upper bound below unless more precise upper bound is required.

Lemma 12 (Uniformity lemma)**.** *Let $(A \times B, I, v)$ be a node of $\mathcal{T}$ at which a $p_{\max}$ bound of $2^{-m+0.8\beta}$ holds for A, B with respect to I. Then the number of inputs (x, y) in $A \times B$ consistent with z, denoted by $\rho_{(A\times B,I)}(z) = |(A \times B) \cap G^{-1}(z)|$, satisfies*

$$\rho_{(A\times B,I)}(z) \in \frac{1}{2^{|I|}} \left[1 - 2^{-0.04\beta}, 1 + 2^{-0.04\beta}\right] \cdot |A| \cdot |B|.$$

Proof. Without loss of generality, we assume $I = \{1, \dots, l\}$, which means that the bits of z that have been queried till now are $l+1, \dots n$. Since Π reaches $(A \times B, I, v)$ on z, we must have that $g^{n-l}(x_{[n]\setminus[l]}, y_{[n]\setminus[l]}) = z_{[n]\setminus[l]}$. We view $\frac{1}{|A||B|}\rho_{(A\times B,I)}(z)$ as a probability distribution over z_I, which we denote by $p(z_I)$. Our bound shall follow by computing $p(z_1)p(z_2|z_1)\dots p(z_l|z_1, \dots z_{l-1})$. Setting $I = \{1\}$ in Lemma 6, we conclude that

$$\begin{aligned} p(z_1) &= \sum_{x_1,y_1} p_A(x_1)p_B(y_1)\mathbb{1}_{g(x_1,y_1)=z_1} \\ &= \frac{1}{2} + \frac{1-2z_1}{2} \sum_{x_1,y_1} p_A(x_1)p_B(y_1)M_g(x_1,y_1) \\ &\in \frac{1}{2}\left[1 - 2^{-0.05\beta}, 1 + 2^{-0.05\beta}\right]. \end{aligned}$$

Now, we consider $p(z_2|z_1)$. For this, it is sufficient to consider $p(z_2|x_1, y_1)$ for any $x_1 \in A_{\{1\}}, y_1 \in B_{\{1\}}$ satisfying $g(x_1, y_1) = z_1$. Since a $p_{\max}$ bound of $2^{-m+0.8\beta}$ also holds for the sets A_{x_1}, B_{y_1} with respect to $I \setminus \{1\}$, we can appeal to Lemma 6 to conclude that

$$p(z_2|x_1, y_1) \in \frac{1}{2}\left[1 - 2^{-0.05\beta}, 1 + 2^{-0.05\beta}\right] \implies p(z_2|z_1) \in \frac{1}{2}\left[1 - 2^{-0.05\beta}, 1 + 2^{-0.05\beta}\right].$$

Proceeding in similar fashion, we conclude that

$$p(z) \in \frac{1}{2^{|I|}}\left[(1 - 2^{-0.05\beta})^n, (1 + 2^{-0.05\beta})^n\right] \in \frac{1}{2^{|I|}}\left[1 - 2^{\log n - 0.05\beta}, 1 + 2^{\log n - 0.05\beta}\right].$$

This completes the proof. ☐

The lemma has the following immediate corollary.

Corollary 13. *For any node $(A_1 \times B_1, I_1, v_1)$ and its successor node $(A_2 \times B_2, I_2, v_2)$ such that a $p_{\max}$ bound of $2^{-m+0.8\beta}$ holds for A_1, B_1 and A_2, B_2 with respect to I_1 and I_2 respectively, the probability that the algorithm Π reaches $(A_2 \times B_2, I_2, v_2)$, conditioned on it reaching $(A_1 \times B_1, I_1, v_1)$ on any input z, lies in the range*

$$\left[1 - 2^{-0.039\beta}, 1 + 2^{-0.039\beta}\right] \cdot 2^{-|I_1\setminus I_2|} \cdot \frac{|A_2| \cdot |B_2|}{|A_1| \cdot |B_1|}.$$

Proof. Note that the transitions in $\mathcal{T}$ from $(A_1 \times B_1, I_1, v_1)$ to $(A_1' \times B_1, I_1', v_1')$ and so on until to $(A_2 \times B_2, I_2, v_2)$ happen according to the relative sizes of the rectangles. Hence the probability of these transitions on z are given by $\rho_{(A_1'\times B_1',I_1)}(z)/\rho_{(A_1\times B_1,I_1)}(z)$ and so on until $\rho_{(A_2\times B_2,I_2)}(z)/\rho_{(A_2'\times B_2',I_2')}(z)$. So,

$$\Pr_{\mathcal{T},z}[(A_2 \times B_2, I_2, v_2)|(A_1 \times B_1, I_1, v_1)] = \frac{\rho_{(A_1'\times B_1',I_1')}(z)}{\rho_{(A_1\times B_1,I_1)}(z)} \cdot \dots \cdot \frac{\rho_{(A_2\times B_2,I_2)}(z)}{\rho_{(A_2'\times B_2',I_2')}(z)} = \frac{\rho_{(A_2\times B_2,I_2)}(z)}{\rho_{(A_1\times B_1,I_1)}(z)}$$

Appealing to the p_{max} bound, we can apply Lemma 12 to obtain that for every z,

$$\rho_{(A_1\times B_1, I_1)}(z) \in \frac{1}{2^{|I_1|}}[1-2^{-0.04\beta}, 1+2^{-0.04\beta}]\cdot|A_1|\cdot|B_1|$$

and a similar result holds for $(A_2 \times B_2, I_2, v_2)$. Plugging in the upper and lower bounds for both these quantities, we get the desired result.

□

For bounding the probability of the algorithm aborting, we need the following claim.

Claim 14. *Consider a tree τ representing a random process with directed edges weighed by the conditional probability of going to a child node conditioned on being in a parent node. Some of the nodes are marked as aborted nodes, and we have that for any node, the sum of weights of the edges going to aborted children be at most δ. If the depth of τ is d, then the overall probability of the random process reaching an aborted node is at most $\delta \cdot d$.*

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

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

which gives us the required bound for the probability of reaching an aborted node in τ. □

Lemma 15. *The overall probability of the algorithm Π aborting on any input $z \in \{0,1\}^n$ is $O(\log n/n)$.*

Proof. We will compute the abort probability for steps corresponding to A. A similar argument holds for B. The algorithm aborts on steps 25, 35 and the Filter steps. We shall consider each of these separately, and further subdivide the argument into query and communication parts.

First we consider the communication steps. For this, we upper bound the abort probability for a fixed z, from which the actual abort probability can be upper bounded by averaging over $\lambda(z)$.

On steps 25 and 26: We first consider the conditional abort probability on a communication sub-routine of Alice starting from step 23 and ending at step 26, conditioned on being at a node $(A \times B, I)$ at the beginning of this subroutine at step 23. This gives us the conditional probability of not aborting at either step 25 or in the Filter procedure in step 26. Note that a $p_{\max}$ bound of $2^{-m+0.8\beta}$ holds at A and a $p_{\max}$ bound of $2^{-m+0.8\beta}$ holds for all possible non-aborting A_j obtained from it at step 26. A $p_{\max}$ bound of $2^{-m+0.8\beta}$ holds for B at the beginning and does not change in these steps. Hence by Corollary 13,

$$\sum_{\text{non-aborting } j} \frac{\rho_{(A_j\times B,I)}(z)}{\rho_{(A\times B,I)}(z)} \geq (1-2^{-0.039\beta}) \sum_{\text{non-aborting } j} \frac{|A_j|}{|A|}.$$

Note that at first A is partitioned into two subsets A^0 and A^1 according to the picked b in step 24. At most one of A^0 and A^1 could have lead to an abort and because of our aborting condition we have

$$\sum_{\text{non-aborting } b\in\{0,1\}} \frac{|A^b|}{|A|} \geq 1 - \frac{1}{n^2}.$$

Moreover, from Lemma 5,

$$p_A(\textsc{Small}(A, A^b, I)) \leq \frac{1}{n^2}$$

which gives us

$$\begin{aligned}\sum_{\text{non-aborting } j} \frac{|A_j|}{|A|} &= \sum_{\text{non-aborting } b\in\{0,1\}} \frac{|A^b \setminus \textsc{Small}(A, A^b, I)|}{|A|} \\ &\geq \sum_{\text{non-aborting } b\in\{0,1\}} \frac{1}{|A|}\left(|A^b| - |A^b \cap \textsc{Small}(A, A^b, I)|\right) \\ &\geq \left(1 - \frac{1}{n^2}\right) - \frac{2}{n^2} \geq 1 - \frac{3}{n^2}.\end{aligned}$$

So finally we get,

$$\sum_{\text{non-aborting } j} \frac{\rho_{(A_j\times B,I)}(z)}{\rho_{(A\times B,I)}(z)} \geq (1 - 2^{-0.039\beta})(1 - 3n^{-2}) \geq 1 - 3n^{-2} - 2^{-0.039\beta}.$$

Hence, the probability of abort conditioned on reaching this node is at most $4n^{-2}$, by the choice of β.

On steps 29 and 41 (no queries): We can do very similar calculations for the probability of abort on steps 29, conditioned on the A after step 26. Note that a $p_{\max}$ bound of $2^{-m+0.8\beta}$ holds for both the parent node A in step 26 and the non-aborting child node A' in step 29. Hence

$$\frac{\rho_{(A'\times B,I)}(z)}{\rho_{(A\times B,I)}(z)} \geq (1 - 2^{-0.039\beta}) \cdot \frac{|A'|}{|A|}.$$

Now in A' the strings $\textsc{UnBal}(A, B, I)$ and the strings $\textsc{Small}(A, A \setminus \textsc{UnBal}(A, B, I), I)$ are removed. By Lemma 7, the total probability loss due to removal of the strings in $\textsc{UnBal}(A, B, I)$ is $2^{-0.05\beta}$ and the total probability loss due to removal of the strings in $\textsc{Small}(A, A\setminus\textsc{UnBal}(A, B, I), I)$ is n^{-2} by Lemma 5. Hence the total conditional probability of aborting is again upper bounded by $O(n^{-2})$ by the choice of β. A similar argument holds for the abort probability in step 41 if there are no queries carried out.

Note that each of the aborts whose conditional probabilities we have calculated so far happen once after Alice communicates a bit. Since there are at most $O(n \log n)$ bits communicated, by Claim 14, the overall probability of abort in these steps is at most $O(\log n/n)$. Now we proceed to the steps where a query has taken place.

On steps 35 and 41 (at least one query): Now assuming at least one query happens, we calculate the probability of abort on steps 35 and 41, conditioned on being at a node $(A\times B, I, v)$ before the while loop began. A $p_{\max}$ bound of $2^{-m+0.8\beta}$ holds for A, B with respect to I by the Invariance Lemma. By Lemma 12 we can say,

$$\rho_{(A\times B,I)}(z) \leq \frac{1 + 2^{-0.04\beta}}{2^l}|A| \cdot |B| \tag{4}$$

Warm up, one query: Consider the simplest case where the while loop has only one iteration, querying say z_1 (where we assume, without loss of generality, that $\{1\} \in I$). In the while loop, first A is split into $A'(= \textsc{High}(A, 2^{-m+0.7\beta}, \{1\}))$ and $A \setminus A'$. $A \setminus A'$ exits the while loop without any queries being done, and then a $\textsc{Filter}$ step is carried out on it, after which a $p_{\max}$ bound of $2^{-m+0.8\beta}$ holds by the Invariance Lemma. Suppose the part that is not aborted in the $\textsc{Filter}$ step is A'', then $|A''| \geq (1 - 2n^{-2})|A \setminus A'|$, since at most n^{-2} fraction is removed in $\textsc{High}$ and

SMALL parts each (for this, notice that at this stage, $|\textsc{High}(A \setminus A', 2^{-m+0.7\beta}, \{i\})| \leq \frac{|A \setminus A'|}{n^3}$ for all $i \in I$). By Lemma 12 we have,

$$\rho_{(A'' \times B, I)}(z) \geq \frac{1 - 2^{-0.04\beta}}{2^{|I|}} |A''| \cdot |B| \geq \frac{(1 - 2^{-0.04\beta})(1 - 2n^{-2})}{2^{|I|}} |A \setminus A'| \cdot |B|.$$

On A', z_1 is queried and A' is set to $A'|_{U(y_1,z_1)}$ depending on the choice of y_1 in step 34. Some of these y_1 lead to abort in step 35. Let Ab (representing abort) denote this set of y_1, that is, $|A'|_{U(y_1,z_1)}| \leq \frac{1}{n^3}|A'|$. The non-aborting part then goes through another FILTER step, after which at most $2n^{-2}$ fraction of $A'|_{U(y_1,z_1)}$ is removed, and it has a $p_{\max}$ bound of $2^{-m+0.8\beta}$ with respect to $I \setminus \{1\}$. So if we let $\{A_k \times B_k\}_k$ denote the rectangles on which a query happens and which are not aborted on steps 35 or 41, then by Lemma 12,

$$\begin{aligned}
\sum_k \rho_{(A_k \times B_k, I \setminus \{1\})}(z) &\geq \frac{(1 - 2^{-0.04\beta})(1 - 2n^{-2})}{2^{|I|-1}} \sum_{y_1 \notin \mathrm{Ab}} |A'|_{U(y_1,z_1)}| \cdot |B_{y_1}| \\
&= \frac{(1 - 2^{-0.04\beta})(1 - 2n^{-2})}{2^{|I|-1}} \left(\sum_{y_1} |A'|_{U(y_1,z_1)}| \cdot |B_{y_1}| - \sum_{y_1 \in \mathrm{Ab}} |A'|_{U(y_1,z_1)}| \cdot |B_{y_1}| \right).
\end{aligned}$$

We bound each of the summations in the above expression separately. For the first term, note that

$$\sum_{y_1} |A'|_{U(y_1,z_1)}| \cdot |B_{y_1}| = |B| \sum_{x \in A'} \Pr_{y_1 \sim p_B}[g(x_1, y_1) = z_1] \geq \frac{1}{2}(1 - 2^{-0.05\beta})|A'| \cdot |B| \tag{5}$$

where the inequality holds due to Lemma 7. For the second term we have,

$$\sum_{y_1 \in \mathrm{Ab}} |A'|_{U(y_1,z_1)}| \cdot |B_{y_1}| \leq \frac{1}{n^3}|A'| \cdot |B| \sum_{y_1 \in \mathrm{Ab}} p_B(y_1) \leq |A'| \cdot |B| \cdot n^{-3}. \tag{6}$$

This gives us

$$\sum_k \rho_{(A_k \times B_k, I \setminus \{1\})}(z) \geq \frac{(1 - 2^{-0.04\beta})(1 - 2n^{-2})}{2^{|I|}} (1 - 2^{-0.05\beta} - \frac{2}{n^3})|A'| \cdot |B|.$$

So the total probability of not aborting is at least

$$\begin{aligned}
&\frac{\sum_k \rho_{(A_k \times B_k, I \setminus \{1\})}(z) + \rho_{(A'' \times B, I)}(z)}{\rho_{(A \times B, I)}(z)} \\
&\geq (1 - 3n^{-2}) \cdot \frac{1 - 2^{-0.04\beta}}{1 + 2^{-0.04\beta}} \cdot (1 - 2^{-0.05\beta} - \frac{2}{n^3}) \left(\frac{|A'| + |A \setminus A'|}{|A|} \right) \\
&\geq 1 - 6n^{-2},
\end{aligned}$$

for the choice of β. So the conditional probability of aborting in this step is at most $6n^{-2}$.

More than one query: For a larger number of queries, there are more possible divisions of A, but the calculations are similar, applying different cases of Lemma 7. There are different sequences of queries for the different partitions of the rectangle $(A \times B)$ at the beginning of the while loop. Recall that the unqueried interval for $(A \times B)$ is I. To capture the branching sequence, we consider the subtree $\mathcal{T}^q$ of $\mathcal{T}$, with root at $(A \times B, I)$ (we shall drop the v label) and the leaves at all the nodes that reach (but do not exit) step 41. For every non-aborting leaf node

$(A_L \times B_L, I_L)$ of $\mathcal{T}^q$, there is a child node $(A'_L \times B_L, I_L)$ in $\mathcal{T}$ that goes through step 41 and does not abort. We have the following relation for all leaves $L \in \mathcal{T}^q$, using Lemma 12,

$$\rho_{(A'_L \times B_L, I_L)}(z) \geq (1 - 2^{-0.04\beta})(1 - 2n^{-2})\frac{2^{|I \setminus I_L|}}{2^{|I|}}|A_L| \cdot |B_L|. \tag{7}$$

Let $\mathrm{Leaf}(\mathcal{T}^q)$ represent the non-aborting leaves of $\mathcal{T}^q$. We shall argue that

$$\sum_{L \in \mathrm{Leaf}(\mathcal{T}^q)} \rho_{(A'_L \times B_L, I_L)}(z) \geq (1 - 4n^{-2})\frac{1}{2^{|I|}}|A||B|.$$

Combined with Equation 4, this shall allow us to upper bound the probability of abort. To show the desired inequality, it suffices to lower bound

$$\sum_{L \in \mathrm{Leaf}(\mathcal{T}^q)} 2^{|I \setminus I_L|}|A_L| \cdot |B_L|, \tag{8}$$

To achieve this, we shall evaluate the expression starting from the leaves and going up to the roots.

We call a node *penultimate* if it is a parent of a leaf node. Consider a penultimate node $L = (A^* \times B^*, I^*)$, which was partitioned into children $\{(A_k \times B_k, I')\}_k$. Suppose the partition happened through a query step (observe that I' is same for each child, as they are all queried at the same location). Let $i = I^* \setminus I'$ be the queried location and Ab_i be the set of aborting y_i's. Following relation holds using Lemma 7, where the argument is similar to that used in Equations 5 and 6:

$$\begin{aligned} \sum_{(A_k \times B_k, I')} |A_k||B_k| &= \sum_{y_i \notin \mathrm{Ab}_i} |A^*|_{U(y_i, z_i)}||B^*_{y_i}| \\ &\geq \frac{1}{2}(1 - 2^{-0.05\beta} - \frac{2}{n^3})|A^*||B^*|. \end{aligned}$$

If the children of penultimate node were not formed due to any query step, then none of them were aborted (abort only occurs at Step 35 within the While loop) and I_L did not change. Then it trivially holds that

$$\sum_{(A_k \times B_k, I')} |A_k||B_k| = |A^*||B^*|.$$

Now, consider the tree $\mathcal{T}_1^q$ formed by removing all the leafs from $\mathcal{T}^q$. Let $\mathrm{Leaf}(\mathcal{T}_1^q)$ be the leaves of $\mathcal{T}_1^q$. Above argument allows us to conclude that the summation in Equation 8 is lower bounded by the following:

$$(1 - 2^{-0.05\beta} - \frac{2}{n^3}) \sum_{L \in \mathrm{Leaf}(\mathcal{T}_1^q)} 2^{|I \setminus I_L|}|A_L| \cdot |B_L|,$$

Continuing the same process, we can reduce the tree till it is just the node $(A \times B, I)$. Then Equation 8 is lower bounded as

$$\sum_{L \in \mathrm{Leaf}(\mathcal{T}^q)} 2^{|I \setminus I_L|}|A_L| \cdot |B_L| \geq (1 - 2^{-0.05\beta} - \frac{2}{n^3})^n |A||B| \geq (1 - 4n^{-2})|A||B|,$$

for the choice of β. This shows that the abort probability is at most $6n^{-2}$.

This gives the total probability of the algorithm aborting to be at most $O(\log n/n)$. □

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

Proof of Lemma 10. The probability that Π makes an error is at most the sum of the probability that Π aborts, given by Lemma 15, and the probability that it makes an error on a leaf. We know by Lemma 15 that the overall probability of abort on any z is at most $O(\log n/n)$, hence the overall probability of abort when z is drawn from λ is also at most $O(\log n/n)$.

To bound the error at a leaf, let us denote the output of a leaf L of T by b^L and probability that T on input (x, y) drawn uniformly from $G^{-1}(z)$ for a fixed z, reaches leaf L by q_z^L. By correctness of T on the distribution μ we have,

$$\Pr_{(x,y)\sim\mu}[((x, y), T(x, y)) \in f \circ g^n] = \mathbb{E}_{z\sim\lambda}\left[\sum_{L:(z,b^L)\in f} q_z^L\right] \geq \frac{3}{4}. \tag{9}$$

Let us further denote the probability of Π reaching a leaf on a fixed input z by $q_z'^L$. We will lower bound

$$\Pr_{z\sim\lambda}[(z, \Pi(z)) \in f] = \mathbb{E}_{z\sim\lambda}\left[\sum_{L:(z,b^L)\in f} q_z'^L\right].$$

Due to (9), it is enough to show that q_z^L and $q_z'^L$ are close. Let the rectangle associated with the leaf L of T be denoted as $A^L \times B^L$. Since T has no internal randomness and conditioned on a particular z the underlying distribution is uniform in its support, the probability that an input drawn uniformly from $G^{-1}(z)$ reaches L is given only by the relative number of $(x, y) \in A^L \times B^L$ that are consistent with $G^{-1}(z)$. That is,

$$q_z^L = \frac{|(A^L \times B^L) \cap G^{-1}(z)|}{|G^{-1}(z)|}.$$

Now there are many nodes $(A_k \times B_k, I_k, L)$ in $\mathcal{T}$ corresponding to the node L and we have that $A^L = (\cup_k A_k) \cup A^L_{\text{ABORT}}$, $B^L = (\cup_k B_k) \cup B^L_{\text{ABORT}}$. Moreover, we know from Corollary 13 that the probability of Π going to node $(A_k \times B_k, I_k, L)$ is proportional to $|(A_k \times B_k) \cap G^{-1}(z)| \cdot [1 - 2^{-0.039\beta}, 1 + 2^{-0.039\beta}]$ (the non-aborting leaf nodes all have a $p_{\max}$ bound of $2^{-m+0.8\beta}$ for A, B). So,

$$\begin{aligned}
q_z'^L &\geq (1 - 2^{-0.039\beta}) \frac{\sum_{(A_k \times B_k, I_k, L)\in\mathcal{T}} |(A_k \times B_k) \cap G^{-1}(z)|}{|G^{-1}(z)|} \\
&= (1 - 2^{-0.039\beta}) \frac{|(A^L \times B^L) \cap G^{-1}(z)| - |(A^L_{\text{ABORT}} \times B^L_{\text{ABORT}}) \cap G^{-1}(z)|}{|G^{-1}(z)|} \\
&= (1 - 2^{-0.039\beta}) q_z^L - (1 - 2^{-0.039\beta}) \frac{|(A^L_{\text{ABORT}} \times B^L_{\text{ABORT}}) \cap G^{-1}(z)|}{|G^{-1}(z)|}
\end{aligned}$$

We now appeal to Lemma 15 to conclude that overall probability of abort is at most $O(\frac{\log n}{n})$. This gives us the probability of the algorithm making an error on a leaf to be

$$1 - \mathbb{E}_{z\sim\lambda}\left[\sum_{L:(z,b^L)\in f} q_z'^L\right] \leq 1 - (1 - 2^{-0.039\beta}) \mathbb{E}_{z\sim\lambda}\left[\sum_{L:(z,b^L)\in f} q_z^L\right] + O\left(\frac{\log n}{n}\right) \leq \frac{1}{4} + O\left(\frac{\log n}{n}\right).$$

□

Expected number of queries of Π

Lemma 16. *The algorithm Π makes at most $\frac{2c}{0.7\beta - \log 4n}$ expected number of queries, where c is the number of bits communicated in T in the worst case.*

Proof. Consider the non-aborting leaf nodes $\{(A_k \times B_k, I_k, L_k)\}_k$ in $\mathcal{T}$, where L_k is a leaf node of T. In each of A_k, B_k, some strings are fixed in intervals J_A, J_B respectively (we drop the label k from these intervals, as it will be clear from context), where J_A and J_B are disjoint. Moreover $J_A \cup J_B = [n] \setminus I_k$.

Assume for simplicity that $I_k = \{1, 2, \ldots |I_k|\}$, $J_A = \{|I_k|+1, |I_k|+2, \ldots |I_k|+|J_A|\}$ and $J_B = \{|I_k|+|J_A|+1, |I_k|+|J_A|+2, \ldots n\}$. For any pair of strings $(x, y) \in A_k \times B_k$, we have that,

$$\frac{1}{|A_k||B_k|} = p_{A_k}(x)p_{B_k}(y) = p_{A_k}(x_{I_k \cup J_A}) \cdot p_{A_k}(x_{J_B}|x_{I_k \cup J_A}) \times p_{B_k}(y_{I_k \cup J_B}) \cdot p_{B_k}(y_{J_A}|y_{I_k \cup J_B}) \quad (10)$$

We evaluate the term $p_{A_k}(x_{J_B}|x_{I_k \cup J_A})$ in the following way. Suppose the queries in J_B happened in the sequence $\{|I_k|+|J_A|+1, |I_k|+|J_A|+2, \ldots n\}$. Let A'_k be the ancestor of A_k when index $\{n\}$ was queried. Since $A_k \subseteq A'_k$, we have

$$p_{A_k}(x_{J_B}|x_{I_k \cup J_A}) = \frac{1}{|(A_k)_{x_{I_k \cup J_A}}|} \geq \frac{1}{|(A'_k)_{x_{I_k \cup J_A}}|} = p_{A'_k}(x_{J_B}|x_{I_k \cup J_A}).$$

Now, the fact that the query happened at index $\{n\}$ implies $p_{A'_k}(x_n|x_{I_k \cup J_A}) \geq 2^{-m+0.7\beta}$. Thus, $p_{A'_k}(x_{J_B}|x_{I_k \cup J_A}) \geq 2^{-m+0.7\beta} \cdot p_{A'_k}(x_{J_B \setminus \{n\}}|x_{I_k \cup J_A \cup \{n\}})$. Now, we can consider the ancestor A''_k of A'_k at which x_{n-1} was queried and further lower bound this quantity. Continuing this way, we obtain

$$p_{A_k}(x_{J_B}|x_{I_k \cup J_A}) \geq 2^{-(m-0.7\beta)|J_B|}.$$

Similar argument for B_k gives us

$$p_{B_k}(y_{J_A}|y_{I_k \cup J_B}) \geq 2^{-(m-0.7\beta)|J_A|}.$$

Combining, we obtain

$$p_{A_k}(x_{J_B}|x_{I_k \cup J_A})p_{B_k}(y_{J_A}|y_{I_k \cup J_B}) \geq 2^{-(m-0.7\beta)(|J_B|+|J_A|)}.$$

Now, there is at least one x_{I_k} such that $p_{A_k}(x_{I_k \cup J_A}) \geq 2^{-m \cdot |I_k|}$ (recall that x_{J_A} is fixed). Similarly, there is at least one y_{I_k} such that $p_{B_k}(y_{I_k \cup J_B}) \geq 2^{-m \cdot |I_k|}$ (recall that y_{J_B} is fixed). Thus, collectively, we find from Equation 10 that

$$\frac{1}{|A_k||B_k|} \geq 2^{-2m \cdot |I_k|} \cdot 2^{-(|J_A|+|J_B|)(m-0.7\beta)}.$$

This implies that

$$\frac{2^{m(2n-|J_A|-|J_B|)}}{|A_k||B_k|} \geq 2^{m(2n-|J_A|-|J_B|)} \cdot 2^{-2m \cdot |I_k|} \cdot 2^{-(|J_A|+|J_B|)(m-0.7\beta)} = 2^{0.7\beta(|J_A|+|J_B|)}.$$

Taking logarithm on both sides and taking expectation over all $(A_k \times B_k, I_k)$, we get

$$0.7\beta \cdot \mathbb{E}_{((A_k \times B_k), I_k)}(|J_A|+|J_B|) \leq \mathbb{E}_{((A_k \times B_k), I_k)}\left(\log\left(\frac{2^{2m \cdot n}}{|A_k||B_k|}\right) - m \cdot (|J_A|+|J_B|)\right). \quad (11)$$

Now, recall that $|J_A|+|J_B|$ is the number of queries in the rectangle $(A_k \times B_k)$. Let

$$E_q := \mathbb{E}_{((A_k \times B_k), I_k)}(|J_A|+|J_B|)$$

be the expected number of queries in the algorithm Π. We upper bound the right hand side of Equation 11. For this, we view the protocol Π as a communication protocol Π^{cc} in the following sense. Let Alice and Bob publicly share z distributed according to λ. Alice and Bob simulate all the steps in Π, and communicate abort if required. Suppose the players reach steps 10 or 30 during the simulation. If Alice (Bob) finds that there exists an $i \in I$ that satisfies the 'while' condition, then she (he) communicates i to Bob (Alice) with $\log n$ bits. For every query done by Alice or Bob at some index i, the respective player samples z_i from the shared randomness. If either of Alice and Bob need to fix an x_i or y_i (respectively), they communicate the fixed string to the other party with m bits of communication. In similar manner, rest of the steps are simulated. We find that for each query done in Π, Π^{cc} communicates at most $m + \log n + 2$ bits. Thus, expected number of bits communicated in Π^{cc} is upper bounded by $2c + (m + \log 4n)E_q$, where the term $2c$ arise due to the possibility of abort in non-query steps (note that there is at most one additional abort per communication step). Using Claim 17 below, we conclude

$$\mathbb{E}_{((A_k \times B_k), I_k)} \log \left(\frac{2^{2m \cdot n}}{|A_k||B_k|} \right) \leq 2c + (m + \log 4n)E_q.$$

Using this in Equation 11, we obtain

$$0.7\beta \cdot E_q \leq 2c + (m + \log 4n)E_q - m \cdot E_q \implies (0.7\beta - \log 4n)E_q \leq 2c.$$

This completes the proof.

Claim 17. *The expected number of bits communicated in Π^{cc} is lower bounded by*

$$\mathbb{E}_{((A_k \times B_k), I_k)} \log \left(\frac{2^{2m \cdot n}}{|A_k||B_k|} \right).$$

Proof. Given the protocol Π^{cc}, we construct a tree $\mathcal{T}^{cc}$ as follows. The nodes of tree $\mathcal{T}^{cc}$ are labeled by the rectangles $(A \times B)$ that appear in the protocol Π^{cc}. The root node is $((\{0,1\}^m)^n \times (\{0,1\}^m)^n)$. For each node $(A \times B)$, its children are the set of all nodes that are obtained by partitioning $(A \times B)$ in the given randomized fashion. We include the aborted nodes in $\mathcal{T}^{cc}$. Observe that $\mathcal{T}^{cc}$ is different from the tree $\mathcal{T}$ constructed earlier, in the sense that the latter is constructed for a fixed z. For a node $(A \times B) \in \mathcal{T}^{cc}$, let $N_{\text{Child}}(A \times B)$ be the number of children of $(A \times B)$ (we set this number to 1 for a leaf node). The expected number of bits communicated in Π^{cc} is equal to

$$\mathbb{E}_{(A \times B) \in \mathcal{T}^{cc}} \log N_{\text{Child}}(A \times B).$$

For a node $(A \times B)$, let $\{(A_k \times B_k)\}_k$ be the set of its children. Then it holds that

$$\sum_{(A_k \times B_k)} \frac{|A_k||B_k|}{|A||B|} \log \frac{|A||B|}{|A_k||B_k|} \leq \log N_{\text{Child}}(A \times B).$$

Since the transition probability in Π^{cc} from $(A \times B)$ to $(A_k \times B_k)$ is $\frac{|A_k||B_k|}{|A||B|}$, the claim concludes. ∎

∎

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

[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.

[GPW17] Mika Göös, Toniann Pitassi, and Thomas Watson. Query-to-communication lifting for BPP. 2017. https://arxiv.org/abs/1703.07666.

[LSS08] T. Lee, A. Shraibman, and R. Spalek. A direct product theorem for discrepancy. In *2008 23rd Annual IEEE Conference on Computational Complexity*, pages 71–80, June 2008.

[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.

Algorithm 2: Randomized query algorithm Π for f

Input: $z \in \{0,1\}^n$

Initialize v as root of the protocol tree T, $I = [n]$, $A = (\{0,1\}^m)^n$ and $B = (\{0,1\}^m)^n$

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

  if *Bob sends a bit at v* **then**

    For $b \in \{0,1\}$ pick $B' = B \cap Y^{v,b}$ with probability $|B \cap Y^{v,b}|/|B|$

    if $|B'| < \frac{1}{n^2}|B|$ *for the picked b* **then** Abort

    $\textsc{Filter}(B', \textsc{Small}(B, B', I))$

    Set $v \leftarrow v_b$ and $B \leftarrow B'$

    if *there is an $i \in I$ such that* $|\textsc{High}(B, 2^{-m+0.7\beta}, i)| > \frac{1}{n^3}|B|$ **then**

      $\textsc{Filter}\left(B, \textsc{UnBal}_{\mathcal{Y}}(A, B, I) \cup \textsc{Small}(B, B \setminus \textsc{UnBal}_{\mathcal{Y}}(A, B, I), I)\right)$

      while $|\textsc{High}(B, 2^{-m+0.7\beta}, i)| > \frac{1}{n^3}|B|$ *for some $i \in I$* **do**

        Pick $B' = \textsc{High}(B, 2^{-m+0.7\beta}, i)$ or $B \setminus \textsc{High}(B, 2^{-m+0.7\beta}, i)$ with probability $|\textsc{High}(B, 2^{-m+0.7\beta}, i)|/|B|$ or $1 - |\textsc{High}(B, 2^{-m+0.7\beta}, i)|/|B|$ respectively

        if $B' = \textsc{High}(B, 2^{-m+0.7\beta}, i)$ *is picked* **then**

          Query z_i

          Pick $x_i \in \{0,1\}^m$ with probability $|A_{x_i}|/|A|$

          if $|B'|_{V(x_i,z_i)}| < \frac{1}{n^3}|B'|$ **then** Abort

          Set $B \leftarrow B'|_{V(x_i,z_i)}$, $A \leftarrow A_{x_i}$ and $I \leftarrow I \setminus \{i\}$

        end

        Set $B \leftarrow B'$

      end

    end

    $\textsc{Filter}\left(B, \textsc{High}(B, 2^{-m+0.7\beta}, I) \cup \textsc{Small}(B, B \setminus \textsc{High}(B, 2^{-m+0.7\beta}, I), I)\right)$

  end

  else if *Alice sends a bit at v* **then**

    For $b \in \{0,1\}$ pick $A' = A \cap X^{v,b}$ with probability $|A \cap X^{v,b}|/|A|$

    if $|A'| < \frac{1}{n^2}|A|$ *for the picked b* **then** Abort

    $\textsc{Filter}(A', \textsc{Small}(A, A', I))$

    Set $v \leftarrow v_b$ and $A \leftarrow A'$

    if *there is an $i \in I$ such that* $|\textsc{High}(A, 2^{-m+0.7\beta}, i)| > \frac{1}{n^3}|A|$ **then**

      $\textsc{Filter}\left(A, \textsc{UnBal}_{\mathcal{X}}(A, B, I) \cup \textsc{Small}(A, A \setminus \textsc{UnBal}_{\mathcal{X}}(A, B, I), I)\right)$

      while $|\textsc{High}(A, 2^{-m+0.7\beta}, i)| > \frac{1}{n^3}|A|$ *for some $i \in I$* **do**

        Pick $A' = \textsc{High}(A, 2^{-m+0.7\beta}, i)$ or $A \setminus \textsc{High}(A, 2^{-m+0.7\beta}, i)$ with probability $|\textsc{High}(A, 2^{-m+0.7\beta}, i)|/|A|$ or $1 - |\textsc{High}(A, 2^{-m+0.7\beta}, i)|/|A|$ respectively

        if $A' = \textsc{High}(A, 2^{-m+0.7\beta}, i)$ *is picked* **then**

          Query z_i

          Pick $y_i \in \{0,1\}^m$ with probability $|B_{y_i}|/|B|$

          if $|A'|_{U(y_i,z_i)}| < \frac{1}{n^3}|A'|$ **then** Abort

          Set $A \leftarrow A'|_{U(y_i,z_i)}$, $B \leftarrow B_{y_i}$ and $I \leftarrow I \setminus \{i\}$

        end

        Set $A \leftarrow A'$

      end

    end

    $\textsc{Filter}\left(A, \textsc{High}(A, 2^{-m+0.7\beta}, I) \cup \textsc{Small}(A, A \setminus \textsc{High}(A, 2^{-m+0.7\beta}, I), I)\right)$

  end

end

Output as T does on the leaf v.

Procedure 3: Filter(T, S)

Input: $T \subseteq (\{0,1\}^m)^n$ and $S \subseteq T$

Pick $T' = S$ or $T \setminus S$ with probability $|S|/|T|$ or $1 - |S|/|T|$ respectively

if $T' = S$ *is picked* **then** Abort

Set $T \leftarrow T'$

ECCC ISSN 1433-8092

https://eccc.weizmann.ac.il