A Quadratically Tight Partition Bound for Classical Communication and Query Complexity

Rahul Jain[1], Troy Lee[2], and Nisheeth K. Vishnoi[3]

1 **Centre for Quantum Technologies and Department of Computer Science, National University of Singapore. Email:** `rahul@comp.nus.edu.sg`

2 **Nanyang Technological University and Center for Quantum Technologies. Email:** `troyjlee@gmail.com`

3 **EPFL Switzerland. Email:** `nisheeth.vishnoi@gmail.com`

Abstract

The *partition bound* introduced in [4] is a way to prove lower bounds in classical communication and query complexity. While the partition bound provides a strong and general way to prove lower bounds, it remains open how tight the bounds obtained from this method are. In this work we give quadratically tight lower bounds via a strengthened version of the partition bound, which we call the *public-coin partition bound*. Formally, we show that, for all relations, the logarithms of the communication and query complexity versions of our public-coin partition bounds are within a quadratic factor of the public-coin randomized communication and randomized query complexity respectively.

Keywords: Partition bound, communication complexity, lower bounds, linear programs.

1 Introduction

Proving communication and query complexity bounds has been a challenging and active research direction in complexity theory with far reaching applications to VLSI design, streaming algorithms and combinatorial optimization, see [6]. Many different methods have been proposed which aim to capture communication and query complexity in different settings, see [7]. Understanding the tightness of these lower bound methods is a central question. Perhaps the most important and famous among these is understanding the relationship between *log-rank* of the communication matrix and deterministic communication complexity (for boolean functions). This question has wide-spread consequences for example in understanding *non-negative rank* of matrices and also to deep questions in combinatorics. Another important question is understanding relationship between *approximate* γ_2 and classical randomized communication complexity which can help relate classical and quantum communication complexity (for total functions).

Among the strongest known lower bound methods, both in classical randomized (public-coin) communication complexity and randomized query complexity, is the *partition bound* introduced in [4]. To the best of our knowledge, there is no function or relation where the lower bound obtained from the partition bound method is asymptotically weaker either for randomized (public-coin) communication complexity or for randomized query complexity. However, it has not been established that the partition bound captures these quantities tightly. Again understanding relationship between partition bound and randomized communication complexity will enhance our understanding of non-negative rank of matrices.

The other widely studied lower bound method is (internal) *information complexity* which very recently has been shown to be exponentially smaller than (distributional) communication complexity (under a specific distribution) for a particular relation [3].

Our Contribution

In this work we introduce, both for communication complexity and query complexity, a strengthening of the partition bound which we call the *public-coin partition bound*. Analogous to the partition bound, our new bound is also a linear-programming based lower bound method and is stronger than the partition bound for all relations, both in communication complexity and query complexity. Formally, in one direction we show that (the base two logarithm of) its communication and query complexity versions continue to form a lower bound on the public-coin communication complexity and randomized query complexity respectively. Importantly, we are also able to show a near-converse: The square of (the base two logarithm of) the communication and query complexity versions of our public-coin partition bound form an upper bound on the public-coin communication complexity and randomized query complexity respectively.

In the first look the definition of our lower bound may look quite similar to the definition of communication complexity itself and hence its tightness may not appear surprising. However the fact that our lower bound is expressed as a linear program may have potential advantages. It may be possible to provide new lower bounds for specific functions and relations by providing good feasible solutions to the dual. It may be possible to relate it (by comparing feasible solutions) to other linear programming based lower bound methods e.g. the partition bound and the *smooth-rectangle bound* [4] and hence helpful in understanding the tightness of these bounds. This may also shed light into the *direct-product* question in communication complexity since a direct-product result is known in terms of smooth-rectangle bound for all relations [5]. A direct-product result states that if less than k times the communication required to compute a single instance of a relation f (with constant error) is provided for computing k simultaneous instances of f, then the overall success is exponentially small in k. It may also be possible to attack the direct-product question directly by analyzing if the linear program corresponding to public-coin partition bound exhibits a product structure.

We now proceed to formally stating our results and proofs.

2 Our Result in the Communication Complexity Setting

In this section we introduce our new bound in the communication complexity setting. Let us first recall the partition bound of [4].

▶ Definition 1 (Partition bound [4]). Let $f \subseteq \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$ be a relation. Let $\varepsilon > 0$. The ε-partition bound of f, denoted $\mathsf{prt}_\varepsilon(f)$, is given by the optimal value of the following linear program. Below R represents a rectangle in $\mathcal{X} \times \mathcal{Y}$ and $(x, y, z) \in \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$.

Primal

$$\min \quad \sum_z \sum_R w_{z,R}$$

s.t.

$$\forall (x,y): \quad \sum_{z:(x,y,z)\in f} \sum_{R:(x,y)\in R} w_{z,R} \geq 1-\varepsilon$$

$$\forall (x,y): \quad \sum_{R:(x,y)\in R} \sum_z w_{z,R} = 1$$

$$\forall (z,R): \quad w_{z,R} \geq 0$$

Dual

$$\max \quad (1-\varepsilon)\sum_{(x,y)} \mu_{x,y} + \sum_{(x,y)} \phi_{x,y}$$

s.t.

$$\forall (z,R): \quad \sum_{(x,y)\in R:(x,y,z)\in f} \mu_{x,y} + \sum_{(x,y)\in R} \phi_{x,y} \leq 1$$

$$\forall (x,y): \quad \mu_{x,y} \geq 0, \phi_{x,y} \in \mathbb{R}$$

Our bound is obtained by refining the above linear program by adding a new set of variables; one for each setting of the public-coins in the protocol. These extra variables allow us to construct a protocol starting from a solution to the linear program.

▶ Definition 2 (Public-coin partition bound). Let $f \subseteq \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$ be a relation. Let $\varepsilon > 0$. The ε-public-coin partition bound of f, denoted $\mathsf{pprt}_\varepsilon(f)$, is given by the optimal value of the following linear program. Below R represents a rectangle in $\mathcal{X} \times \mathcal{Y}$ and P represents a partition along with outputs in $\mathcal{Z}$; that is $P = \{(z_1, R_1), (z_2, R_2), \cdots, (z_m, R_m)\}$, such that $\{R_1, \cdots, R_m\}$ form a partition of $\mathcal{X} \times \mathcal{Y}$ into rectangles and $\forall i \in [m], z_i \in \mathcal{Z}$.

Primal

$$\min \quad \sum_z \sum_R w_{z,R}$$

s.t.

$$\begin{aligned}
\forall (x,y):& \quad \sum_{z:(x,y,z)\in f} \sum_{R:(x,y)\in R} w_{z,R} \geq 1-\varepsilon \\
\forall (x,y):& \quad \sum_{R:(x,y)\in R} \sum_{z} w_{z,R} = 1 \\
\forall (z,R):& \quad w_{z,R} = \sum_{P:(z,R)\in P} a_P \\
& \quad \sum_P a_P = 1 \\
\forall (z,R):& \quad w_{z,R} \geq 0; \quad \forall P : a_P \geq 0
\end{aligned}$$

Dual

$$\max \quad (1-\varepsilon) \sum_{(x,y)} \mu_{x,y} + \sum_{(x,y)} \phi_{x,y} + \lambda$$

s.t.

$$\begin{aligned}
\forall (z,R):& \quad \sum_{(x,y)\in R:(x,y,z)\in f} \mu_{x,y} + \sum_{(x,y)\in R} \phi_{x,y} + v_{z,R} \leq 1 \\
\forall P:& \quad \sum_{(z,R)\in P} v_{z,R} \geq \lambda \\
\forall (x,y):& \quad \mu_{x,y} \geq 0, \phi_{x,y} \in \mathbb{R} \\
\forall (z,R):& \quad v_{z,R} \in \mathbb{R} \\
& \quad \lambda \in \mathbb{R}
\end{aligned}$$

It is possible to get rid of the variables $w_{z,R}$ in this program, however we keep them since explicit comparison with partition bound is easier in this form. We present a simplified linear program in Section C.

Notice that the dual of the linear program used to define the partition bound can be obtained from the dual of the linear program for the public-coin partition bound by setting the variables λ and $v_{z,A}$ to 0. Thus, any lower bound on $\mathsf{prt}_\varepsilon(f)$ obtained by demonstrating a feasible solution to the corresponding dual extends to a feasible dual solution of the public-coin partition bound dual; resulting in the same lower bound on $\mathsf{pprt}_\varepsilon(f)$. In particular it is always true that

$$\mathsf{prt}_\varepsilon(f) \leq \mathsf{pprt}_\varepsilon(f).$$

The following is our main theorem in the communication complexity setting.

▶ Theorem 1. Let $f \subseteq \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$ be a relation. Let $\varepsilon > 0$. Let $\mathsf{R}^{\mathsf{pub}}_\varepsilon(f)$ represents the public-coin communication complexity of f with worst-case error ε. Then,

$$\log_2 \mathsf{pprt}_{2\varepsilon}(f) \leq \mathsf{R}^{\mathsf{pub}}_{2\varepsilon}(f) \leq \left(\log_2 \mathsf{pprt}_\varepsilon(f) + \log_2 \frac{1}{\varepsilon} + 1\right)^2.$$

Note that such a result is not known to be true for the partition bound. We prove the lower bound and the upper bound separately. We start by showing that our bound is indeed a lower bound (please refer to [6] for standard definitions in communication complexity).

▶ Lemma 1. Let $f \subseteq \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$ be a relation. Let $\varepsilon > 0$. Let $\mathsf{R}^{\mathsf{pub}}_\varepsilon(f)$ represents the public-coin communication complexity of f with worst-case error ε. Then,

$$\log_2 \mathsf{pprt}_\varepsilon(f) \leq \mathsf{R}^{\mathsf{pub}}_\varepsilon(f).$$

This proof proceeds along similar lines as the proof of [4] for an analogous result for the partition bound.

Proof. Let $\mathcal{P}$ be a public coin randomized protocol for f with communication $c \stackrel{\mathsf{def}}{=} \mathsf{R}^{\mathsf{pub}}_{\varepsilon}(f)$ and worst case error ε. For a binary string r, let $\mathcal{P}_r$ represent the deterministic communication protocol obtained from $\mathcal{P}$ by fixing the public coins to r. Every deterministic communication protocol amounts to partitioning the inputs in $\mathcal{X} \times \mathcal{Y}$ into rectangles and outputting an element in $\mathcal{Z}$ corresponding to each rectangle in the partition. Let $P_r = \{(z_1^r, R_1^r), (z_2^r, R_2^r), \cdots, (z_m^r, R_m^r)\}$, be the corresponding partition along with the outputs, i.e., $\{R_1^r, \cdots, R_m^r\}$ partition $\mathcal{X} \times \mathcal{Y}$ into rectangles and $\forall i \in [m], z_i^r \in \mathcal{Z}$. Let q_r represent the probability that the string r is chosen in $\mathcal{P}$. For P_r define $a'_{P_r} \stackrel{\mathsf{def}}{=} q_r$. For the partitions P that do not correspond to a random string r in $\mathcal{P}$, define $a'_P = 0$. For all (z, R) define

$$w'_{z,R} \stackrel{\mathsf{def}}{=} \sum_{P:(z,R)\in P} a'_P.$$

With these definitions, it can be seen that for all $(x, y, z) \in \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$:

$$\Pr[\mathcal{P} \text{ outputs } z \text{ on input } (x, y)] = \sum_{R:(x,y)\in R} w'_{z,R}.$$

Since the protocol has error at most ε on all inputs we get the constraint:

$$\forall (x, y) : \sum_{z:(x,y,z)\in f} \sum_{R:(x,y)\in R} w'_{z,R} \geq 1 - \varepsilon.$$

Also, since the $\Pr[\mathcal{P}$ outputs some $z \in \mathcal{Z}$ on input $(x, y)] = 1$, we get the constraint:

$$\forall (x, y) : \sum_{z} \sum_{R:(x,y)\in R} w'_{z,R} = 1.$$

We also have by construction:

$$\sum_P a'_P = 1; \quad \forall (z, R) : w'_{z,R} \geq 0; \quad \forall P : a'_P \geq 0.$$

Therefore $\{w'_{z,R}\} \cup \{a'_P\}$ is feasible for the primal of $\mathsf{pprt}_{\varepsilon}(f)$. We know that for each r, $|P_r| \leq 2^c$, since the communication in $\mathcal{P}_r$ is at most c bits. Hence,

$$\mathsf{pprt}_{\varepsilon}(f) \leq \sum_z \sum_R w'_{z,R} = \sum_r a'_{P_r} \cdot |P_r| \leq 2^c \sum_r a'_{P_r} = 2^c.$$

◀

Next we show the other half of Theorem 1; that the square of the logarithm of our new bound forms an upper bound on the public-coin communication complexity.

▶ Lemma 2. Let $f \subseteq \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$ be a relation. Let $\varepsilon > 0$. We have,

$$\mathsf{R}^{\mathsf{pub}}_{2\varepsilon}(f) \leq \left(\log_2 \mathsf{pprt}_{\varepsilon}(f) + \log_2 \frac{1}{\varepsilon} + 1\right)^2.$$

Proof. Let $\mathsf{pprt}_{\varepsilon}(f) = 2^c$. Let $\{w_{z,R}\} \cup \{a_P\}$ be an optimal solution for the primal. Let n_P be the number of rectangles in P. We have,

$$\sum_P a_P \cdot n_P = \sum_{z,R} w_{z,R} = 2^c.$$

Next we will identify the partitions with many more rectangles than the expected value 2^c, assign 0 probability to them and re-normalize the probability among other partitions. Define $B \stackrel{\mathsf{def}}{=} \{P \mid n_P \geq \frac{1}{\varepsilon}2^c\}$. Then $\delta \stackrel{\mathsf{def}}{=} \sum_{P \in B} a_P \leq \varepsilon$. Define $a'_P \stackrel{\mathsf{def}}{=} \frac{1}{1-\delta} a_P$ for $P \notin B$ and $a'_P \stackrel{\mathsf{def}}{=} 0$ for $P \in B$. Define $w'_{z,R} \stackrel{\mathsf{def}}{=} \sum_{P:(z,R)\in P} a'_P$. Then we have,

$$
\begin{aligned}
\forall (x,y): \quad & \sum_{z:(x,y,z)\in f} \sum_{R:(x,y)\in R} w'_{z,R} \geq 1 - 2\varepsilon, \\
\forall (x,y): \quad & \sum_{R:(x,y)\in R} \sum_{z} w'_{z,R} = 1, \\
\forall (z,R): \quad & w'_{z,R} = \sum_{P:(z,R)\in P} a'_p, \\
& \sum_{P} a'_P = 1, \\
\forall (z,R): \quad & w'_{z,R} \geq 0 \\
\forall P: \quad & a'_P \geq 0.
\end{aligned}
$$

Next we show that a partition with m rectangles can be realized by a communication protocol with communication $(\lceil \log_2 m \rceil)^2$. This argument proceeds as in the proof of Theorem 2.11 in [6], which relates non-deterministic communication complexity to deterministic communication complexity. We provide a proof in Section A. Given this, consider a public-coin communication protocol Π as follows:

1. Alice and Bob (using public coins) choose a $P = \{(z_1, R_1), (z_2, R_2), \cdots, (z_m, R_m)\}$ with probability a'_P.
2. They communicate to realize the partition $\{R_1, R_2, \cdots, R_m\}$ with communication bounded by $(c + \log_2 \frac{1}{\varepsilon} + 1)^2$.
3. If they end up with rectangle R_i, they output z_i.

It is clear that, in the worst case, the amount of communication of the protocol is bounded by $(c + \log_2 \frac{1}{\varepsilon} + 1)^2$. The condition

$$
\forall (x,y): \sum_{z:(x,y,z)\in f} \sum_{R:(x,y)\in R} w'_{z,R} \geq 1 - 2\varepsilon
$$

implies that the protocol has worst case error at most 2ε. Therefore,

$$
\mathsf{R}^{\mathsf{pub}}_{2\varepsilon}(f) \leq \left(\log_2 \mathsf{pprt}_{\varepsilon}(f) + \log_2 \frac{1}{\varepsilon} + 1 \right)^2.
$$

◀

Thus, we complete the proof of Theorem 1.

3 Our Result in the Query Complexity Setting

In this section we introduce our new bound in the query complexity setting. Let $f \subseteq \{0,1\}^n \times \mathcal{Z}$ be a relation. An *assignment* $A : S \to \{0,1\}^l$ is an assignment of values to some subset S of n variables (with $|S| = l$). We say that A is *consistent* with $x \in \{0,1\}^n$ if $x_i = A(i)$ for all $i \in S$. We write $x \in A$ as shorthand for 'A is consistent with x'. We write $|A|$ to represent the size of A which is the cardinality of S (not to be confused with the number of consistent inputs). Furthermore we say that an index i *appears* in A, iff $i \in S$

where S is the subset of $[n]$ corresponding to A. Let $\mathcal{A}$ denote the set of all assignments. Below we assume $x \in \{0,1\}^n$, $A \in \mathcal{A}$ and $z \in \mathcal{Z}$. Below P represents a partition along with outputs in $\mathcal{Z}$; that is $P = \{(z_1, A_1), (z_2, A_2), \cdots, (z_m, A_m)\}$, such that $\{A_1, \cdots, A_m\}$ form a partition of $\{0,1\}^n$ into assignments (that is for each $x \in \{0,1\}^n$, there is a unique $i \in [m]$ such that $x \in A_i$) and $\forall i \in [m], z_i \in \mathcal{Z}$. Let us first recall the partition bound of [4].

▶ Definition 3 (Partition bound [4]). Let $f \subseteq \{0,1\}^n \times \mathcal{Z}$ be a relation. Let $\varepsilon > 0$. The ε-partition bound of f, denoted $\mathsf{prt}_\varepsilon(f)$, is given by the optimal value of the following linear program.

Primal

$$\begin{aligned}
&\min \quad \sum_z \sum_A w_{z,A} \cdot 2^{|A|} \\
&\text{s.t.} \\
&\forall x: \quad \sum_{z:(x,z)\in f} \sum_{A: x\in A} w_{z,A} \geq 1-\varepsilon \\
&\forall x: \quad \sum_{A: x\in A} \sum_z w_{z,A} = 1 \\
&\forall (z,A): \quad w_{z,A} \geq 0
\end{aligned}$$

Dual

$$\begin{aligned}
&\max \quad (1-\varepsilon)\sum_x \mu_x + \sum_x \phi_x \\
&\text{s.t.} \\
&\forall (z,A): \quad \sum_{x\in A:(x,z)\in f} \mu_x + \sum_{x\in A} \phi_x \leq 2^{|A|} \\
&\forall x: \quad \mu_x \geq 0, \phi_x \in \mathbb{R}
\end{aligned}$$

Our strengthened bound is defined as follows.

▶ Definition 4 (Public-coin partition bound). Let $f \subseteq \{0,1\}^n \times \mathcal{Z}$ be a relation. Let $\varepsilon > 0$. The ε-public-coin partition bound of f, denoted $\mathsf{pprt}_\varepsilon(f)$, is given by the optimal value of the following linear program.

Primal

$$\begin{aligned}
&\min \quad \sum_z \sum_A w_{z,A} \cdot 2^{|A|} \\
&\text{s.t.} \\
&\forall x: \quad \sum_{z:(x,z)\in f} \sum_{A: x\in A} w_{z,A} \geq 1-\varepsilon \\
&\forall x: \quad \sum_{A: x\in A} \sum_z w_{z,A} = 1 \\
&\forall (z,A): \quad w_{z,A} = \sum_{P:(z,A)\in P} a_P \\
&\qquad \sum_P a_P = 1 \\
&\forall (z,A): \quad w_{z,A} \geq 0; \quad \forall P: a_P \geq 0
\end{aligned}$$

Dual

$$\begin{aligned}
&\max \quad (1-\varepsilon)\sum_x \mu_x + \sum_x \phi_x + \lambda \\
&\text{s.t.} \\
&\forall (z,A): \quad \sum_{x\in A:(x,z)\in f} \mu_x + \sum_{x\in A} \phi_x + v_{z,A} \leq 2^{|A|} \\
&\forall P: \quad \sum_{(z,A)\in P} v_{z,A} \geq \lambda \\
&\forall x: \quad \mu_x \geq 0, \phi_x \in \mathbb{R} \\
&\forall (z,A): \quad v_{z,A} \in \mathbb{R} \\
&\qquad \lambda \in \mathbb{R}
\end{aligned}$$

Again it is possible to get rid of the variables $w_{z,A}$ in this program, however we keep them since explicit comparison with partition bound is easier in this form. We present a simplified linear program in Section D.

In a manner similar to the communication complexity setting, it can be seen that, for all f,

$$\mathsf{prt}_\varepsilon(f) \leq \mathsf{pprt}_\varepsilon(f).$$

Our main theorem is an analog of Theorem 1 in the query complexity setting and is not known to be true for the original partition bound.

▶ Theorem 2. Let $f \subseteq \{0,1\}^n \times \mathcal{Z}$ be a relation. Let $\varepsilon > 0$. Let $\mathsf{R}_\varepsilon(f)$ represent the randomized query complexity of f with worst case error ε. Then,

$$\frac{1}{2}\log_2 \mathsf{pprt}_{2\varepsilon}(f) \leq \mathsf{R}_{2\varepsilon}(f) \leq \left(\log \mathsf{pprt}_\varepsilon(f) + \log_2 \frac{1}{\varepsilon}\right)^2.$$

We start by proving the easy direction of Theorem 2.

▶ Lemma 3. Let $f \subseteq \{0,1\}^n \times \mathcal{Z}$ be a relation. Let $\varepsilon > 0$. Let $\mathsf{R}_\varepsilon(f)$ represent the randomized query complexity of f with worst case error ε. Then,

$$\frac{1}{2}\log_2 \mathsf{pprt}_\varepsilon(f) \leq \mathsf{R}_\varepsilon(f).$$

Our proof follows arguments similar to [4] for an analogous result for the partition bound.

Proof. Let $\mathcal{P}$ be a randomized query algorithm which achieves $c \overset{\mathsf{def}}{=} \mathsf{R}_\varepsilon(f)$. Let $\mathcal{P}_r$ be the deterministic query algorithm, arising from $\mathcal{P}$, corresponding to random string r. We know that each such deterministic query algorithm is a binary decision tree of depth at most c (please refer to [2] for standard definitions related to query complexity). We note that the leaves of a decision tree represent a partition of the inputs into assignments along with outputs in Z. Let $P_r = \{(z_1^r, A_1^r), (z_2^r, A_2^r), \cdots, (z_m^r, A_m^r)\}$ represent the partition and outputs corresponding to random string r, where $\{A_1^r, \cdots, A_m^r\}$ form a partition of $\{0,1\}^n$ into assignments and $\forall i \in [m], z_i^r \in \mathcal{Z}$. Let q_r represent the probability of string r in $\mathcal{P}$. For P_r define $a'_{P_r} \overset{\mathsf{def}}{=} q_r$. For the partitions P that do not correspond to any string r in $\mathcal{P}$, define $a'_P = 0$. For any (z, A) define,

$$w'_{z,A} \overset{\mathsf{def}}{=} \sum_{P:(z,A)\in P} a'_P.$$

As in the proof of Lemma 1, we can argue that $\{w'_{z,A}\} \cup \{a'_P\}$ is feasible for the primal of $\mathsf{pprt}_\varepsilon(f)$. Note that for each (z, A) with $w'_{z,A} > 0$, we have $|A| \leq c$. Moreover, $|P_r| \leq 2^c$ since the depth of the corresponding binary decision tree is at most c. Now,

$$\begin{aligned}\mathsf{pprt}_\varepsilon(f) &= \sum_z \sum_A w'_{z,A} 2^{|A|} \leq 2^c \left(\sum_z \sum_A w'_{z,A}\right) \\ &\leq 2^c \left(\sum_r a'_{P_r} \cdot |P_r|\right) \leq 2^{2c} \sum_r a'_{P_r} = 2^{2c}.\end{aligned}$$

Thus, the result follows. ◀

Next we show that the square of the logarithm of our new bound forms an upper bound on randomized query complexity, thus, completing the proof of Theorem 2.

▶ Lemma 4. Let $f \subseteq \{0,1\}^n \times \mathcal{Z}$ be a relation. Let $\varepsilon > 0$. Then,

$$\mathsf{R}_{2\varepsilon}(f) \leq \left(\log \mathsf{pprt}_\varepsilon(f) + \log_2 \frac{1}{\varepsilon}\right)^2.$$

Proof. Let $\mathsf{pprt}_\varepsilon(f) = 2^c$. Let $\{w_{z,A}\} \cup \{a_P\}$ be an optimal solution for the primal. We have,

$$\sum_P \sum_{A:(z,A)\in P} a_P \cdot 2^{|A|} = \sum_{z,A} w_{z,A} \cdot 2^{|A|} = 2^c.$$

Next we will identify the partitions containing an assignment of size much larger than c, assign 0 probability to them and re-normalize the probability among other partitions. Define $B \stackrel{\mathsf{def}}{=} \{P \mid \exists (z,A) \in P \text{ with } |A| > c + \log_2 \frac{1}{\varepsilon}\}$. Then $\delta \stackrel{\mathsf{def}}{=} \sum_{P \in B} a_P \leq \varepsilon$. Define $a'_P \stackrel{\mathsf{def}}{=} \frac{1}{1-\delta} a_P$ for $P \notin B$ and $a'_P \stackrel{\mathsf{def}}{=} 0$ for $P \in B$. Define $w'_{z,A} \stackrel{\mathsf{def}}{=} \sum_{P:(z,A)\in P} a'_P$. Then we have,

$$\begin{aligned}
\forall x : &\quad \sum_{z:(x,z)\in f} \sum_{A: x\in A} w'_{z,A} \geq 1 - 2\varepsilon, \\
\forall x : &\quad \sum_{A: x\in A} \sum_{z} w'_{z,A} = 1, \\
\forall (z,A) : &\quad w'_{z,A} = \sum_{P:(z,A)\in P} a'_P, \\
&\quad \sum_P a'_P = 1, \\
\forall (z,A) : &\quad w'_{z,A} \geq 0, \\
\forall P : &\quad a'_P \geq 0.
\end{aligned}$$

Next we show that a partition with assignments each of length at most m can be realized by a query protocol with m^2 queries. This argument proceeds as in the proof of Theorem 11 in [2], relating certificate complexity to deterministic query complexity (for total functions). We present the proof in Section B. Given this, consider a randomized query protocol Π as follows:

1. Alice (randomly) chooses a $P = \{(z_1, A_1), (z_2, A_2), \cdots, (z_s, A_s)\}$ with probability a'_P.
2. She queries to realize the partition $\{A_1, A_2, \cdots, A_s\}$ with $(c + \log_2 \frac{1}{\varepsilon})^2$ queries.
3. If she ends up with assignment A_i, she outputs z_i.

It is clear that, in the worst case, the number of queries made by the protocol is $(c+\log_2 \frac{1}{\varepsilon})^2$. The condition

$$\forall x : \sum_{z:(x,z)\in f} \sum_{A: x\in A} w'_{z,A} \geq 1 - 2\varepsilon$$

implies that the protocol has worst case error at most 2ε. Therefore,

$$\mathsf{R}_{2\varepsilon}(f) \leq \left(\log_2 \mathsf{prt}_{\varepsilon}(f) + \log_2 \frac{1}{\varepsilon}\right)^2 .$$

◀

4 Conclusion and open questions

In this work we present linear programming based lower bound methods for public-coin communication complexity and randomized query complexity and show that they are quadratically tight for all relations. This is the first time any lower bound method has been shown to be quadratically tight (for all relations) either in communication complexity or query complexity. Some interesting open questions related to this work are as follows:

1. What is the relationship between the public-coin partition bound, the partition bound and the smooth-rectangle bound introduced in [4] (all of which are a linear program based lower bound methods)? What is the relationship between the public-coin partition bound and the information complexity lower bound method?

2. A strong direct product theorem is shown for all relations in terms of the smooth-rectangle bound by [5] and recently in terms of information-complexity by [1]. Can a similar result be shown in terms of the public-coin partition bound or the partition bound?
3. Is the public-coin partition bound linearly tight for communication complexity and query complexity? This basically boils down to the following: Can a partition of the communication matrix with 2^k rectangles always be realized (with small error) using $O(k)$ communication public-coin protocol? Can a partition of $\{0,1\}^n$ with 2^k partial assignments, each of size at most k, always be realized (with small error) using $O(k)$ query randomized query protocol?
4. Can explicit lower bounds for interesting functions and relations be shown using the public-coin partition bound?

Acknowledgments

We thank anonymous referees for many suggestions helpful in improving the presentation of the paper.

This work is supported by the Singapore Ministry of Education Tier 3 Grant and the Core Grants of the Centre for Quantum Technologies, Singapore. Part of the work done when NKV was visiting CQT.

References

1 Mark Braverman and Omri Weinstein. An Interactive Information Odometer with Applications. ECCC-TR14-047, 2014.
2 Harry Buhrman and Ronald de Wolf. Complexity measures and decision tree complexity: A survey. *Theoretical Computer Science*, 288(1):21–43, 2002.
3 Anat Ganor, Gillat Kol and Ran Raz. Exponential Separation of Information and Communication. ECCC-TR14-049, 2014.
4 Rahul Jain, Hartmut Klauck. The partition bound for classical communication complexity and query complexity. In *Proceeding of the 25th IEEE Conference on Computational Complexity*, (CCC), pp.247-258, 2010.
5 Rahul Jain, Penghui Yao. A strong direct product theorem in terms of the smooth rectangle bound. Arxiv:1209.0263, 2012.
6 Eyal Kushilevitz and Noam Nisan. *Communication Complexity.* Cambridge University Press, 1997.
7 Troy Lee, Adi Shraibman. *Lower Bounds in Communication Complexity.* Foundations and Trends in Theoretical Computer Science, Vol. 3(4), 2009.

A Communication protocol to realize a partition

Let $\{R_1, R_2, \ldots, R_m\}$ be a partition of $S_A \times S_B$ into rectangles. Initially $S_A = \mathcal{X}, S_B = \mathcal{Y}$. Let $\forall i \in [m] : \ R_i = A_i \times B_i$, where $A_i \subseteq \mathcal{X}, B_i \subseteq \mathcal{Y}$. Let x and y be the inputs to Alice and Bob respectively. The protocol is as follows:

1. Alice determines if there exists an $i \in [m]$ such that
   a. $x \in A_i$ and
   b. the number of rectangles in $\{R_1, R_2, \ldots, R_m\}$ that row intersect with R_i are at most $m/2$. We say that R_{i_1} and R_{i_2} row intersect if $A_{i_1} \cap A_{i_2}$ is non-empty.

   If such an i exists she communicates i to Bob using $\lceil \log_2 m \rceil$ bits. They both now consider $\{R_1 \cap A_i, R_2 \cap A_i, \ldots, R_m \cap A_i\}$ as a partition of $(A_i \times S_B)$ and repeat (and set $S_A = A_i$). If Alice cannot find any such i she indicates this to Bob by sending 0.
2. On receiving 0 from Alice, Bob determines if there exists a $j \in [m]$ such that
   a. $y \in B_j$ and
   b. the number of rectangles in $\{R_1, R_2, \ldots, R_m\}$ that column intersect with R_j are at most $m/2$. We say that R_{j_1} and R_{j_2} column intersect if $B_{j_1} \cap B_{j_2}$ is non-empty.

   If such a j exists he communicates it to Alice using $\lceil \log_2 m \rceil$ bits. They both now consider $\{R_1 \cap B_j, R_2 \cap B_j, \ldots, R_m \cap B_j\}$ as a partition of $(S_A \times B_j)$ and repeat (and set $S_B = B_j$).

We can note that either Alice or Bob must succeed in finding a desired i, j respectively since the rectangle that contains (x, y) satisfies the requirements in either 1 or 2 above (since $\{R_1, R_2, \ldots, R_m\}$ is a partition of $S_A \times S_B$). Moreover, the communication in each round is at most $\lceil \log_2 m \rceil$ and the number of (non-empty) rectangles surviving after each round reduce by a factor of 2. Hence, the process ends after at most $\lceil \log_2 m \rceil$ rounds. Thus, the total communication is bounded by $(\lceil \log_2 m \rceil)^2$.

B Query protocol to realize a partition

Let $\{A_1, A_2, \ldots, A_s\}$ be a partition of $\{0,1\}^n$ such that $|A_i| \leq m$ for each $i \in [s]$. Let x be the string in the database.

1. Alice queries the bits of x corresponding to A_1. If the bits revealed are consistent with A_1 then she considers A_1 to be the desired assignment and stops.
2. If the bits revealed are not consistent with A_1 then note that one bit is revealed for all A_i, $i \in [s]$ since $\{A_1, A_2, \ldots, A_s\}$ is a partition of $\{0,1\}^n$. Hence, the size of each A_i (consistent with the bits revealed so far) reduces by at least 1. Alice considers now the new set of modified A_is and repeats.

We note that the number of such rounds is at most m and in each round at most m bits are revealed. Hence, the total number of queries is at most m^2.

C Simplified public-coin partition bound for communication complexity

Here we present the simplified linear program.

Primal

$$\min \quad \sum_{P} \sum_{(z,R)\in P} a_p$$

s.t.

$$\forall (x,y): \quad \sum_{P} \sum_{(z,R)\in P:(x,y)\in R,(x,y,z)\in f} a_p \geq 1-\varepsilon$$

$$\sum_{P} a_P \leq 1$$

$$\forall P: \quad a_P \geq 0$$

Dual

$$\max \quad (1-\varepsilon)\sum_{(x,y)} \mu_{x,y} - \lambda$$

s.t.

$$\forall P: \quad \sum_{(z,R)\in P} \sum_{(x,y)\in R:(x,y,z)\in f} \mu_{x,y} \leq \lambda + \sum_{(z,R)\in P} 1$$

$$\forall (x,y): \quad \mu_{x,y} \geq 0$$

$$\lambda \geq 0$$

D Simplified public-coin partition bound for query complexity

Here we present the simplified linear program.

Primal

$$\min \quad \sum_{P} \sum_{A:(z,A)\in P} a_P \cdot 2^{|A|}$$

s.t.

$$\forall x: \quad \sum_{P} \sum_{(z,A):x\in A,(x,z)\in f} a_P \geq 1-\varepsilon$$

$$\sum_{P} a_P \leq 1$$

$$\forall P: \quad a_P \geq 0$$

Dual

$$\max \quad (1-\varepsilon)\sum_{x} \mu_x - \lambda$$

s.t.

$$\forall P: \quad \sum_{(z,A)\in P} \sum_{x\in A:(x,z)\in f} \mu_x \leq \lambda + \sum_{(z,A)\in P} 2^{|A|}$$

$$\forall x: \quad \mu_x \geq 0$$

$$\lambda \geq 0$$