# The Partition Bound for Classical Communication Complexity and Query Complexity

Rahul Jain* Hartmut Klauck†

**Abstract**

We describe new lower bounds for randomized communication complexity and query complexity which we call the *partition bounds*. They are expressed as the optimum value of linear programs. For communication complexity we show that the partition bound is stronger than both the *rectangle/corruption bound* and the $\gamma_2$/*generalized discrepancy bounds*. In the model of query complexity we show that the partition bound is stronger than the *approximate polynomial degree* and *classical adversary bounds*. We also exhibit an example where the partition bound is quadratically larger than the approximate polynomial degree and adversary bounds.

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

†Centre for Quantum Technologies, National University of Singapore. Email: hklauck@gmail.com

# 1 Introduction

The computational models investigated in communication complexity and query complexity, i.e., Yao's communication model [Yao79] and the decision tree model, are simple enough to allow us to prove interesting lower bounds, yet they are rich enough to have numerous applications to other models as well as exhibit nontrivial structure. Research in both these models is concentrated on lower bounds and a recurring theme is methods to prove such bounds. In this paper we present a new method for proving lower bounds in both of these models.

## 1.1 Communication Complexity

In the model of communication complexity there are several general methods to prove lower bounds in the settings of randomized communication and quantum communication. Linial and Shraibman [LS09] identified a quantity called $\gamma_2$, which not only yields lower bounds for quantum protocols, but also subsumes a good number of previously known bounds. There is another quantity called the generalized discrepancy, the name being coined in [CA08], which also coincides with $\gamma_2$ as is implicit in [LS09]. The generalized discrepancy can be derived from the standard discrepancy bound in a way originally suggested by Klauck [K07]. The standard discrepancy bound was first shown to be applicable in the quantum case by Kremer and Yao [Kre95]. Razborov [Raz03] (implicitly) and Sherstov [S08] (explicitly), showed that the $\gamma_2$ method yields a tight $\Omega(\sqrt{n})$ bound for the quantum communication complexity of the Disjointness problem, arguably the most important single function considered in the area (for a matching upper bound see [AA05]). This leaves our knowledge of lower bound methods in the world of quantum communication complexity in a neat form, where there is one "master method" that seems to do better than everything else; the only potential competition coming from *information theoretic techniques*, for example as in Jain, Radhakrishnan and Sen [JRS03], which are not applicable to all problems, and not known to beat $\gamma_2$ either.

In the world of randomized communication things appear to be much less organized. Besides simply applying $\gamma_2$, the main competitors are the rectangle (aka corruption) bound (compare [Y83, BFS86, Raz92, K03, BPSW06]), as well as again information theoretic techniques [BKKS04, JKS03]. Both of these approaches are able to beat $\gamma_2$, by allowing $\Omega(n)$ bounds for the Disjointness problem [Raz03, BKKS04, KS92]. There is an information theoretic proof of a tight $\Omega(n)$ lower bound for the Tribes function; an AND of $\sqrt{n}$ ORs of $\sqrt{n}$ ANDs of distributed pairs of variables [JKS03]. With the rectangle bound one cannot prove a lower bound larger than $\sqrt{n}$ for this problem, and neither with $\gamma_2$. So the two general techniques, rectangle bound and $\gamma_2$, are known to be quadratically smaller than the randomized communication complexity for some problems, and the information theoretic approach seems to be only applicable to problems of a "direct sum" type.

In this paper we propose a new lower bound method for randomized communication complexity which we call the partition bound [1]. We derive this bound using a linear program, which captures a relaxation of the fact that a randomized protocol is a convex combination of deterministic protocols and hence a convex combination of partitions of the communication matrix into rectangles. Linear programs have been previously used to describe lower bounds in communication complexity. Lovasz [L90] gives a program which, as we show, turns out to capture the rectangle bound. Our program for the partition bound however uses stricter constraints to overcome the one-sidedness of Lovasz's

[1] In this paper we are only concerned with the two-party, two-way model and the partition bound for other models can be defined analogously. For example for the *Number on the Forehead Model* it can be defined by replacing rectangles by *cylinder intersections*. For the two-party, one-way model it can be defined analogously.

program. Karchmer, Kushilevitz and Nisan [KKN95] give a linear program for *fractional covers*, as well as a linear program which can be seen to be equivalent to our zero-error partition bound for relations, where it was introduced as a lower bound for deterministic complexity.

We also describe a weaker bound to the partition bound which we call the *smooth rectangle bound*. It is inherently a one-sided bound and is derived by relaxing constraints in the linear program for the partition bound. This bound has recently been used to prove a *strong direct product theorem* for Disjointness in [K09]. Another way to derive the smooth rectangle bound is as follows. Suppose we want to prove a lower bound for a function $f$. Then we could apply the rectangle bound, but sometimes this might not yield a large enough lower bound. Instead we apply the rectangle bound to a function $g$ that is sufficiently close to $f$, under a suitable probability distribution, so that lower bounds for $g$ imply lower bounds for $f$. Maximizing this over all $g$, close to $f$, gives us the smooth rectangle bound. This is the same approach that turns the discrepancy bound into the generalized discrepancy. We will use the term smooth discrepancy in the following, because it better captures the underlying approach.

After defining the partition bound and the smooth rectangle bound we proceed to show that the smooth rectangle bound subsumes both the standard rectangle bound and $\gamma_2$/smooth discrepancy. We also show that the LP formulation of the smooth rectangle bound coincides with its natural definition as described above. This leaves us with one unified general lower bound method for randomized communication complexity, the partition bound.

We also define the *Las-Vegas partition bound* via a linear program and exhibit it to be a lower bound on the *Las-Vegas communication complexity* (refer [KN97]). We compare the zero-error partition bound/Las-Vegas partition bound to other standard lower bound methods on deterministic/Las-Vegas communication complexity.

## 1.2 Query Complexity

We then turn to randomized query complexity. Again there are several prominent lower bound methods in this area. Some of the main methods are the classical version of Ambainis' adversary method, the quantum version is from [A02] and classical versions are by Laplante/Magniez [LM08] and Aaronson [A08]; the approximate polynomial degree [NS94, BBC$^+$01]; the randomized certificate bound defined by Aaronson [A06], this being the query complexity analogue of the rectangle bound in communication complexity, as well as older methods like block-sensitivity [Nis91].

We again propose a new lower bound, the partition bound, defined via a linear program, this time based on the fact that a decision tree partitions the Boolean cube into subcubes. We then proceed to show that our lower bound method subsumes all the other bounds mentioned above. In particular the partition bound is always larger than the classical adversary bound, the approximate degree, and block-sensitivity.

To further illustrate the power of our approach we describe a Boolean function, AND of ORs, which we continue to call Tribes, for which the partition bound yields a tight linear lower bound, while both the adversary bound and the approximate degree are at least quadratically smaller.

## Organization

In Section 2 we define the communication complexity partition bound, smooth rectangle bound and mention other previously known lower bound methods. We then present some of the key comparisions between these bounds as mentioned. In Section 3 we perform the same excercise for query complexity. We defer all proofs to Section A. In Section B we present the definitions of partition bound for Las-Vegas communication and query complexity and show that they serve as corresponding lower bounds respectively. In Section C we define the partition bound for relations for communication complexity and query complexity and show that they serve as correponding lower bounds respectively.

# 2 Communication Complexity Bounds

Let $f : \mathcal{X} \times \mathcal{Y} \to \mathcal{Z}$ be a partial function. All the functions considered in this paper are partial functions unless otherwise specified, hence we will drop the term partial henceforth. It is easily verified that strong duality holds for the programs that appear below and hence optima for the primal and dual are same. Let $\mathcal{R}$ be the set of all rectangles in $\mathcal{X} \times \mathcal{Y}$. We refer the reader to [KN97] for introduction to basic terms in communication complexity. Below we assume $(x, y) \in \mathcal{X} \times \mathcal{Y}, R \in \mathcal{R}, z \in \mathcal{Z}$, unless otherwise specified. Let $f^{-1} \subseteq \mathcal{X} \times \mathcal{Y}$ denote the subset where $f(\cdot)$ is defined. For sets $A, B$ we denote $A - B \stackrel{\mathsf{def}}{=} \{a : a \in A, a \notin B\}$. We assume $\epsilon \geq 0$ unless otherwise specified.

## 2.1 New Bounds: Definitions

**Definition 1 (Partition Bound)** *The $\epsilon$-partition bound of $f$, denoted $\mathsf{prt}_\epsilon(f)$, is given by the optimal value of the following linear program.*

*Primal*

$$\begin{aligned} \textit{min:} \quad & \sum_z \sum_R w_{z,R} \\ & \forall (x,y) \in f^{-1} : \sum_{R:(x,y)\in R} w_{f(x,y),R} \geq 1 - \epsilon, \\ & \forall (x,y) : \sum_{R:(x,y)\in R} \sum_z w_{z,R} = 1, \\ & \forall z, \forall R : w_{z,R} \geq 0 \ . \end{aligned}$$

*Dual*

$$\begin{aligned} \textit{max:} \quad & \sum_{(x,y)\in f^{-1}} (1-\epsilon)\mu_{x,y} + \sum_{(x,y)} \phi_{x,y} \\ & \forall z, \forall R : \sum_{(x,y)\in f^{-1}(z)\cap R} \mu_{x,y} + \sum_{(x,y)\in R} \phi_{x,y} \leq 1, \\ & \forall (x,y) : \mu_{x,y} \geq 0, \phi_{x,y} \in \mathbb{R} \ . \end{aligned}$$

Below we present the definition of smooth-rectangle bound as a one-sided relaxation of the partition bound. As we show in the next subsection, it is upper bounded by the partition bound.

**Definition 2 (Smooth-Rectangle bound)** *The $\epsilon$- smooth rectangle bound of $f$ denoted $\mathsf{srec}_\epsilon(f)$ is defined to be $\max\{\mathsf{srec}^z_\epsilon(f) : z \in \mathcal{Z}\}$, where $\mathsf{srec}^z_\epsilon(f)$ is given by the optimal value of the following linear program.*

*Primal*

$$\begin{aligned} \textit{min:} \quad & \sum_{R\in\mathcal{R}} w_R \\ & \forall (x,y) \in f^{-1}(z) : \sum_{R:(x,y)\in R} w_R \geq 1 - \epsilon, \\ & \forall (x,y) \in f^{-1}(z) : \sum_{R:(x,y)\in R} w_R \leq 1, \\ & \forall (x,y) \in f^{-1} - f^{-1}(z) : \sum_{R:(x,y)\in R} w_R \leq \epsilon, \\ & \forall R : w_R \geq 0 \ . \end{aligned}$$

*Dual*

$$\begin{aligned} \textit{max:} \quad & \sum_{(x,y)\in f^{-1}(z)} ((1-\epsilon)\mu_{x,y} - \phi_{x,y}) - \sum_{(x,y)\in f^{-1}-f^{-1}(z)} \epsilon \cdot \mu_{x,y} \\ & \forall R : \sum_{(x,y)\in f^{-1}(z)\cap R} (\mu_{x,y} - \phi_{x,y}) - \sum_{(x,y)\in (R\cap f^{-1})-f^{-1}(z)} \mu_{x,y} \leq 1, \\ & \forall (x,y) : \mu_{x,y} \geq 0; \phi_{x,y} \geq 0 \ . \end{aligned}$$

Below we present an alternate and "natural" definition of smooth-rectangle bound, which justifies its name. In the next subsection we show that the two definitions are equivalent.

**Definition 3 (Smooth-Rectangle bound : Natural definition)** *In the natural definition, $(\epsilon, \delta)$-smooth-rectangle bound of $f$, denoted $\widetilde{\mathsf{srec}}_{\epsilon,\delta}(f)$, is defined as follows (refer to the definition of $\widetilde{\mathsf{rec}}_{\epsilon}^{z,\lambda}(g)$ in the next subsection):*

$$\widetilde{\mathsf{srec}}_{\epsilon,\delta}(f) \stackrel{\mathsf{def}}{=} \max\{\widetilde{\mathsf{srec}}_{\epsilon,\delta}^{z}(f) : z \in \mathcal{Z}\}.$$

$$\widetilde{\mathsf{srec}}_{\epsilon,\delta}^{z}(f) \stackrel{\mathsf{def}}{=} \max\{\widetilde{\mathsf{srec}}_{\epsilon,\delta}^{z,\lambda}(f) : \lambda \textit{ a (probability) distribution on } \mathcal{X} \times \mathcal{Y} \cap f^{-1}\}.$$

$$\widetilde{\mathsf{srec}}_{\epsilon,\delta}^{z,\lambda}(f) \stackrel{\mathsf{def}}{=} \max\{\widetilde{\mathsf{rec}}_{\epsilon}^{z,\lambda}(g) : g : \mathcal{X} \times \mathcal{Y} \to \mathcal{Z}; \Pr_{(x,y)\leftarrow\lambda}[f(x,y) \neq g(x,y)] < \delta; \lambda(g^{-1}(z)) \geq 0.5\}.$$

Below we define smooth-discrepancy via a linear program. Later we present the natural definition of smooth-discrepancy and in the next subsection we show that the two definitions are equivalent. As we also show in the next subsection smooth-discrepancy is upper bounded by smooth-rectangle bound which in turn is upper bounded by the partition bound.

**Definition 4 (Smooth-Discrepancy)** *Let $f : \mathcal{X} \times \mathcal{Y} \to \{0,1\}$ be a Boolean function. The smooth-discrepancy of $f$, denoted $\mathsf{sdisc}_\epsilon(f)$, is given by the optimal value of the following linear program.*

*Primal*

$$\begin{aligned}
&\textit{min:} \quad \sum_{R\in\mathcal{R}} w_R + v_R \\
&\forall (x,y) \in f^{-1}(1): \quad 1+\epsilon \geq \sum_{R:(x,y)\in R} w_R - v_R \geq 1, \\
&\forall (x,y) \in f^{-1}(0): \quad 1+\epsilon \geq \sum_{R:(x,y)\in R} v_R - w_R \geq 1, \\
&\forall R : w_R, v_R \geq 0 \ .
\end{aligned}$$

*Dual*

$$\begin{aligned}
&\textit{max:} \quad \sum_{(x,y)\in f^{-1}} \mu_{x,y} - (1+\epsilon)\phi_{x,y} \\
&\forall R : \sum_{(x,y)\in f^{-1}(1)\cap R} (\mu_{x,y} - \phi_{x,y}) - \sum_{(x,y)\in R\cap f^{-1}(0)} (\mu_{x,y} - \phi_{x,y}) \leq 1, \\
&\forall R : \sum_{(x,y)\in f^{-1}(0)\cap R} (\mu_{x,y} - \phi_{x,y}) - \sum_{(x,y)\in R\cap f^{-1}(1)} (\mu_{x,y} - \phi_{x,y}) \leq 1, \\
&\forall (x,y) : \mu_{x,y} \geq 0; \phi_{x,y} \geq 0 \ .
\end{aligned}$$

### 2.1.1 Known Bounds: Definitions

Below we present the definition of the rectangle bound via a linear program. This program was first described by Lovasz [L90] although he did not make the connection to the rectangle bound.

**Definition 5 (Rectangle-Bound)** *The $\epsilon$-rectangle bound of $f$, denoted $\mathsf{rec}_\epsilon(f)$, is defined to be $\max\{\mathsf{rec}_\epsilon^z(f) : z \in \mathcal{Z}\}$, where $\mathsf{rec}_\epsilon^z(f)$ is given by the optimal value of the following linear program.*

*Primal*

$$\begin{aligned}
&\textit{min:} \quad \sum_{R} w_R \\
&\quad \forall (x,y) \in f^{-1}(z) : \sum_{R:(x,y)\in R} w_R \geq 1-\epsilon, \\
&\quad \forall (x,y) \in f^{-1} - f^{-1}(z) : \sum_{R:(x,y)\in R} w_R \leq \epsilon, \\
&\quad \forall R : w_R \geq 0 \ .
\end{aligned}$$

*Dual*

$$\begin{aligned}
&\textit{max:} \quad \sum_{(x,y)\in f^{-1}(z)} (1-\epsilon)\cdot\mu_{x,y} - \sum_{(x,y)\in f^{-1}-f^{-1}(z)} \epsilon\cdot\mu_{x,y} \\
&\quad \forall R : \sum_{(x,y)\in f^{-1}(z)\cap R} \mu_{x,y} - \sum_{(x,y)\in (R\cap f^{-1})-f^{-1}(z)} \mu_{x,y} \leq 1, \\
&\quad \forall (x,y) : \mu_{x,y} \geq 0 \ .
\end{aligned}$$

Below we present the alternate, natural and conventional definition of rectangle bound as used in several previous works [Y83, BFS86, Raz92, K03, BPSW06]. In the next subsection we show that the two definitions are equivalent.

**Definition 6 (Rectangle-Bound: Conventional definition)** *In the conventional definition, $\epsilon$-rectangle bound of $f$, denoted $\widetilde{\mathsf{rec}}_\epsilon(f)$ is defined as follows:*

$$\widetilde{\mathsf{rec}}_\epsilon(f) \stackrel{\mathsf{def}}{=} \max\{\widetilde{\mathsf{rec}}^z_\epsilon(f) : z \in \mathcal{Z}\}$$

$$\widetilde{\mathsf{rec}}^z_\epsilon(f) \stackrel{\mathsf{def}}{=} \max\{\widetilde{\mathsf{rec}}^{z,\lambda}_\epsilon(f) : \lambda \text{ a distribution on } \mathcal{X} \times \mathcal{Y} \cap f^{-1} \text{ with } \lambda(f^{-1}(z)) \geq 0.5\}.$$

$$\widetilde{\mathsf{rec}}^{z,\lambda}_\epsilon(f) \stackrel{\mathsf{def}}{=} \min\{\frac{1}{\lambda(f^{-1}(z) \cap R)} : R \in \mathcal{R} \text{ with } \epsilon \cdot \lambda(f^{-1}(z) \cap R) > \lambda(R - f^{-1}(z))\} \ .$$

Below we present the definition of discrepancy via a linear program followed by the conventional definition of discrepancy. It is easily seen that the two are exactly the same.

**Definition 7 (Discrepancy)** *Let $f : \mathcal{X} \times \mathcal{Y} \to \{0, 1\}$ be a Boolean function. The discrepancy of $f$, denoted $\mathsf{disc}(f)$, is given by the optimal value of the following linear program.*

*Primal*

$$\begin{aligned}
\textit{min:} \quad & \sum_R w_R + v_R \\
& \forall (x, y) \in f^{-1}(1) : \sum_{R:(x,y)\in R} w_R - v_R \geq 1, \\
& \forall (x, y) \in f^{-1}(0) : \sum_{R:(x,y)\in R} v_R - w_R \geq 1, \\
& \forall R : w_R, v_R \geq 0 \ .
\end{aligned}$$

*Dual*

$$\begin{aligned}
\textit{max:} \quad & \sum_{(x,y)\in f^{-1}} \mu_{x,y} \\
& \forall R : \sum_{(x,y)\in f^{-1}(1)\cap R} \mu_{x,y} - \sum_{(x,y)\in R\cap f^{-1}(0)} \mu_{x,y} \leq 1, \\
& \forall R : \sum_{(x,y)\in f^{-1}(0)\cap R} \mu_{x,y} - \sum_{(x,y)\in R\cap f^{-1}(1)} \mu_{x,y} \leq 1, \\
& \forall (x, y) : \mu_{x,y} \geq 0 \ .
\end{aligned}$$

**Definition 8 (Discrepancy: Conventional definition)** *Let $f : \mathcal{X} \times \mathcal{Y} \to \{0, 1\}$ be a Boolean function. The discrepancy of $f$, denoted $\mathsf{disc}(f)$ is defined as follows:*

$$\mathsf{disc}(f) \stackrel{\mathsf{def}}{=} \max\{\mathsf{disc}^\lambda(f) : \lambda \text{ a distribution on } \mathcal{X} \times \mathcal{Y} \cap f^{-1}\}.$$

$$\mathsf{disc}^\lambda(f) \stackrel{\mathsf{def}}{=} \min\{\frac{1}{|\sum_{(x,y)\in R}(-1)^{f(x,y)} \cdot \lambda_{x,y}|} : R \in \mathcal{R}\} \ .$$

Below we present the natural definition of smooth-discrepancy which has found shape in previous works [K07, S08]. It is defined in analogous fashion from discrepancy as smooth-rectangle bound is defined from rectangle bound.

**Definition 9 (Smooth-Discrepancy: Natural Definition)** *Let $f : \mathcal{X}\times\mathcal{Y} \to \{0, 1\}$ be a Boolean function. The $\delta$- smooth-discrepancy of $f$, denoted $\widetilde{\mathsf{sdisc}}_\delta(f)$, is defined as follows:*

$$\widetilde{\mathsf{sdisc}}_\delta(f) \stackrel{\mathsf{def}}{=} \max\{\widetilde{\mathsf{sdisc}}^\lambda_\delta(f) : \lambda \text{ a distribution on } \mathcal{X} \times \mathcal{Y} \cap f^{-1}\}.$$

$$\widetilde{\mathsf{sdisc}}^\lambda_\delta(f) \stackrel{\mathsf{def}}{=} \max\{\mathsf{disc}^\lambda(g) : g : \mathcal{X} \times \mathcal{Y} \to \mathcal{Z}; \Pr_{(x,y)\leftarrow\lambda}[f(x, y) \neq g(x, y)] < \delta\}.$$

Below we define the $\gamma_2$ bound of Linial and Shraibman [LS09] where it is also implicitly shown that it is equivalent to smooth-discrepancy. We present a proof of this equivalence later for completeness.

**Definition 10 ($\gamma_2$ bound [LS09])** *Let $A$ be a sign matrix and $\alpha \geq 1$. Then,*

$$\gamma_2(A) \stackrel{\mathsf{def}}{=} \min_{X,Y:XY=A} r(X)c(Y) \quad ; \quad \gamma_2^\alpha(A) \stackrel{\mathsf{def}}{=} \min_{B:\forall(i,j)\ 1\leq A(i,j)B(i,j)\leq\alpha} \gamma_2(B).$$

*Above $r(X)$ represents the largest $\ell_2$ norm of the rows of $X$ and $c(X)$ represents the largest $\ell_2$ norm of the columns of $Y$.*

Below we present two well-known lower bound methods for deterministic communication complexity. We present a comparison of these with the 0-error partition bound in the next subsection.

**Definition 11** (log**-rank bound)** *Let $f : \mathcal{X} \times \mathcal{Y} \to \mathcal{Z}$ be a total function. Let $M_f$ denote the communication matrix associated with $f$; $\mathsf{D}(f)$ denote the deterministic communication complexity of $f$ and $\mathsf{rank}(\cdot)$ represents the rank over the reals. Then it is well known [KN97] that $\mathsf{D}(f) \geq \log_2 \mathsf{rank}(f)$.*

**Definition 12 (Fooling Set)** *Let $f : \mathcal{X} \times \mathcal{Y} \to \mathcal{Z}$ be a total function. A set $S \subseteq \mathcal{X} \times \mathcal{Y}$ is called a* fooling set *(for $f$) if there exists a value $z \in \mathcal{Z}$, such that*

- *For every $(x, y) \in S$, $f(x, y) = z$.*
- *For every two distinct pairs $(x_1, y_1)$ and $(x_2, y_2)$ in $S$, either $f(x_1, y_2) \neq z$ or $f(x_2, y_1) \neq z$.*

*It is easily argued that $\mathsf{D}(f) \geq \log_2 |S|$ [KN97].*

## 2.2 Comparison between bounds

The following theorem captures some key relationships between the bounds defined in the previous section. Below $\mathsf{R}^{\mathsf{pub}}_\epsilon(f)$ denotes the public-coin, $\epsilon$-error communication complexity of $f$. All logs in this paper are taken to base 2.

**Theorem 1** *Let $f : \mathcal{X} \times \mathcal{Y} \to \mathcal{Z}$ be a function and let $\epsilon > 0$.*

1. $\mathsf{R}^{\mathsf{pub}}_\epsilon(f) \geq \log \mathsf{prt}_\epsilon(f)$.
2. $\mathsf{prt}_\epsilon(f) \geq \mathsf{srec}_\epsilon(f)$.
3. $\mathsf{srec}_\epsilon(f) \geq \mathsf{rec}_\epsilon(f)$.
4. *Let $f : \mathcal{X} \times \mathcal{Y} \to \mathcal{Z}$ be a total function, then $\mathsf{D}(f) = O((\log \mathsf{prt}_0(f) + \log n)^2)$. Later we exhibit that the quadratic gap between $\mathsf{D}$ and $\log \mathsf{prt}_0$ is tight. For relations however there could be an exponential gap between $\log \mathsf{prt}_0$ and $\mathsf{D}$ as shown in [KKN95].*
5. *Let $f : \mathcal{X} \times \mathcal{Y} \to \mathcal{Z}$ be a total function, and let $S \subseteq \mathcal{X} \times \mathcal{Y}$ be a fooling set. Then $\mathsf{prt}_0(f) \geq |S|$.*

The following lemma shows the equivalence of the two definitions of the rectangle bound.

**Lemma 1** *Let $f : \mathcal{X} \times \mathcal{Y} \to \mathcal{Z}$ be a function and let $\epsilon > 0$. Then for all $z \in \mathcal{Z}$,*

1. $\mathsf{rec}^z_\epsilon(f) \leq \widetilde{\mathsf{rec}}^z_{\frac{\epsilon}{2}}(f)$.
2. $\mathsf{rec}^z_\epsilon(f) \geq \frac{1}{2} \cdot (\frac{1}{2} - \epsilon) \cdot \widetilde{\mathsf{rec}}^z_{2\epsilon}(f)$.

The following lemma shows the equivalence of the two definitions of the smooth-rectangle bound.

**Lemma 2** *Let $f : \mathcal{X} \times \mathcal{Y} \to \mathcal{Z}$ be a function and let $\epsilon > 0$. Then for all $z \in \mathcal{Z}$,*

1. $\mathsf{srec}^z_\epsilon(f) \leq \widetilde{\mathsf{srec}}^z_{\frac{\epsilon}{2}, \frac{1-\epsilon}{2}}(f)$.

2. $\mathsf{srec}^z_\epsilon(f) \geq \frac{1}{2} \cdot (\frac{1}{4} - \epsilon) \cdot \widetilde{\mathsf{srec}}^z_{2\epsilon, \frac{\epsilon}{2}}(f)$.

The following lemma shows the equivalence of the two definitions of smooth-discrepancy.

**Lemma 3** *Let $f : \mathcal{X} \times \mathcal{Y} \rightarrow \{0,1\}$ be a function and let $\epsilon > 0$. Then*

1. $\widetilde{\mathsf{sdisc}}_{\frac{1}{2} - \frac{\epsilon}{8}}(f) \geq \mathsf{sdisc}_\epsilon(f)$.

2. $\frac{1}{2} \cdot \widetilde{\mathsf{sdisc}}_{\frac{1}{4+2\epsilon}}(f) \leq \mathsf{sdisc}_\epsilon(f)$.

The following lemma states the rectangle bound dominates the discrepancy bound for Boolean functions and hence the smooth-rectangle bound dominates the smooth-discrepancy bound.

**Lemma 4** *Let $f : \mathcal{X} \times \mathcal{Y} \rightarrow \{0,1\}$ be a function; let $z \in \{0,1\}$ and let $\lambda$ be a distribution on $\mathcal{X} \times \mathcal{Y} \cap f^{-1}$. Let $\epsilon, \delta > 0$, then*

$$\mathsf{rec}^z_\epsilon(f) \geq (\frac{1}{2} - \epsilon)\mathsf{disc}^\lambda(f) - \frac{1}{2} \ .$$

*This implies by definition and Lemma 1,*

$$\widetilde{\mathsf{rec}}^z_{\frac{\epsilon}{2}}(f) \geq \mathsf{rec}^z_\epsilon(f) \geq (\frac{1}{2} - \epsilon)\mathsf{disc}(f) - \frac{1}{2} \ ,$$

$$\Rightarrow \widetilde{\mathsf{srec}}^z_{\frac{\epsilon}{2},\delta}(f) \geq (\frac{1}{2} - \epsilon)\widetilde{\mathsf{sdisc}}_\delta(f) - \frac{1}{2} \ .$$

For a function $g : \mathcal{X} \times \mathcal{Y} \rightarrow \{0,1\}$, let $A_g$ be the sign matrix corresponding to $g$, that is $A_g(x,y) \stackrel{\mathsf{def}}{=} (-1)^{g(x,y)}$. Following lemma states the equivalence between smooth-discrepancy and the $\gamma_2$ bound. This fact is implicit in Linial and Shraibman [LS09]. We present a proof of this in Section A for completeness.

**Lemma 5 ([LS09])** *Let $f : \mathcal{X} \times \mathcal{Y} \rightarrow \{0,1\}$ be a Boolean function and let $\alpha > 1$. Then*

$$\frac{1}{2} \cdot \widetilde{\mathsf{sdisc}}_{\frac{1}{2(\alpha+1)}}(f) \leq \gamma_2^\alpha(A_f) \leq 8 \cdot \widetilde{\mathsf{sdisc}}_{\frac{1}{\alpha+1}}(f) \ .$$

From Lemma 4 and Lemma 5 we have the following corollary.

**Corollary 1** *Let $f : \mathcal{X} \times \mathcal{Y} \rightarrow \{0,1\}$ be a Boolean function; let $z \in \{0,1\}$; let $\alpha > 1, \epsilon > 0$. Then,*

$$\widetilde{\mathsf{srec}}^z_{\frac{\epsilon}{2}, \frac{1}{\alpha+1}}(f) \geq (\frac{1}{2} - \epsilon)\frac{1}{8}\gamma_2^\alpha(A_f) - \frac{1}{2} \ .$$

The following theorem captures some separations between some of the bounds we mentioned.

**Theorem 2**
1. $\log \mathsf{prt}_\epsilon(\mathsf{Disj}) \geq \log \mathsf{rec}_\epsilon(\mathsf{Disj}) = \Omega(n)$, *while* $\log \gamma_2^\alpha(\mathsf{Disj}) = O(\sqrt{n})$ *for all* $\epsilon < 1/2$ *and* $\alpha > 1$.

2. *There is a function* $f : \{0,1\}^n \times \{0,1\}^n \rightarrow \{0,1\}$ *such that* $\log \mathsf{prt}_\epsilon(f) \geq \log \mathsf{rec}_\epsilon(f) = \Omega(n)$, *while* $\log \mathsf{rank}(f) = O(n^{0.62})$ *for all* $\epsilon < 1/2$.

3. *Let the function* $\mathsf{LNE} : \{0,1\}^{n^2} \times \{0,1\}^{n^2} \rightarrow \{0,1\}$ *be defined as*

$$\mathsf{LNE}(x_1, \ldots, x_n; y_1, \ldots, y_n) = 1 \iff \forall i : x_i \neq y_i,$$

*where all* $x_i, y_j$ *are strings of length* $n$. *Then* $\mathsf{D}(\mathsf{LNE}) = \log \mathsf{rank}(\mathsf{LNE}) = n^2$, *however* $\mathsf{R}_0(\mathsf{LNE}) = O(n)$ *and* $\log \mathsf{prt}_0(\mathsf{LNE}) = O(n)$.

# 3 Query Complexity Bounds

## 3.1 New Bounds: Definitions

Let $f : \{0,1\}^n \to \{0,1\}^m$ be a partial function. Henceforth all functions considered are partial unless otherwise specified. An *assignment* $A : S \to \{0,1\}^m$ is an assignment of values to some subset $S$ of $n$ variables. 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 $z \in \{0,1\}^m$, unless otherwise specified.

**Definition 13 (Partition Bound)** *Let $f : \{0,1\}^n \to \{0,1\}^m$ be a function and let $\epsilon \geq 0$. The $\epsilon$-partition bound of $f$, denoted $\mathsf{prt}_\epsilon(f)$, is given by the optimal value of the following linear program.*

*Primal*

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

*Dual*

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

We define the query complexity version of the smooth discrepancy bound as follows. We show in the next subsection that it is equivalent to approximate degree (up to log factors).

**Definition 14 (Smooth-Discrepancy)** *Let $f : \{0,1\}^n \to \{0,1\}$ be a function. The smooth-discrepancy of $f$, denoted $\mathsf{sdisc}_\epsilon(f)$, is given by the optimal value of the following linear program.*

*Primal*

$$\begin{aligned}
\text{min:}\quad & \sum_A (w_A + v_A) \cdot 2^{|A|} \\
& \forall x \in f^{-1}(1) : \quad 1 \geq \sum_{A : x \in A} w_A - v_A \geq 1-\epsilon, \\
& \forall x \in f^{-1}(0) : \quad 1 \geq \sum_{A : x \in A} v_A - w_A \geq 1-\epsilon, \\
& \forall x : \quad 1 \geq \sum_{A : x \in A} v_A - w_A \geq -1, \\
& \forall A : w_A, v_A \geq 0 \ .
\end{aligned}$$

*Dual*

$$\begin{aligned}
\text{max:}\quad & \Big(\sum_{x \in f^{-1}} (1-\epsilon)\mu_x - \phi_x\Big) - \Big(\sum_{x \notin f^{-1}} \mu_x + \phi_x\Big) \\
& \forall A : \sum_{x \in f^{-1}(1) \cap A} (\mu_x - \phi_x) - \sum_{x \in A, x \notin f^{-1}(1)} (\mu_x - \phi_x) \leq 2^{|A|}, \\
& \forall A : \sum_{x \in A, x \notin f^{-1}(1)} (\mu_x - \phi_x) - \sum_{x \in A \cap f^{-1}(1)} (\mu_x - \phi_x) \leq 2^{|A|}, \\
& \forall x : \mu_x \geq 0; \phi_x \geq 0 \ .
\end{aligned}$$

### 3.1.1 Known Bounds

In this section we define some known complexity measures of functions. All of these except the (error less) certificate complexity are lower bounds for randomized query complexity. See the survey by Buhrman and de Wolf [BW02] for further information.

**Definition 15 (Certificate Complexity)** *For $z \in \{0,1\}^m$, a $z$-certificate for $f$ is an assignment $A$ such that $x \in A \Rightarrow f(x) = z$. The certificate complexity $\mathsf{C}_x(f)$ of $f$ on $x$ is the size of the smallest*

*$f(x)$-certificate that is consistent with $x$. The certificate complexity of $f$ is $\mathsf{C}(f) \stackrel{\text{def}}{=} \max_{x\in f^{-1}} \mathsf{C}_x(f)$. The $z$-certificate complexity of $f$ is $\mathsf{C}^z(f) \stackrel{\text{def}}{=} \max_{x:f(x)=z} \mathsf{C}_x(f)$.*

**Definition 16 (Sensitivity and Block Sensitivity)** *For $x \in \{0,1\}^n$ and $S \subseteq [n]$, let $x^S$ be $x$ flipped on locations in $S$. The sensitivity $\mathsf{s}_x(f)$ of $f$ on $x$ is the number of different $i \in [n]$ for which $f(x) \neq f(x^{\{i\}})$. The sensitivity of $f$ is $\mathsf{s}(f) \stackrel{\text{def}}{=} \max_{x\in f^{-1}} \mathsf{s}_x(f)$.*

*The block sensitivity $\mathsf{bs}_x(f)$ of $f$ on $x$ is the maximum number $b$ such that there are disjoint sets $B_1, \ldots, B_b$ for which $f(x) \neq f(x^{B_i})$. The block sensitivity of $f$ is $\mathsf{bs}(f) \stackrel{\text{def}}{=} \max_{x\in f^{-1}} \mathsf{bs}_x(f)$. If $f$ is constant, we define $\mathsf{s}(f) = \mathsf{bs}(f) = 0$. It is clear from definitions that $\mathsf{s}(f) \leq \mathsf{bs}(f)$.*

**Definition 17 (Randomized Certificate Complexity [A06])** *A $\epsilon$-error randomized verifier for $x \in \{0,1\}^n$ is a randomized algorithm that, on input $y \in \{0,1\}^n$, queries $y$ and (i) accepts with probability 1 if $y = x$, and (ii) rejects with probability at least $1-\epsilon$ if $f(y) \neq f(x)$. If $y \neq x$ but $f(y) = f(x)$, the acceptance probability can be arbitrary. Then $\mathsf{RC}^x_\epsilon(f)$ is the maximum number of queries used by the best $\epsilon$-error randomized verifier for $x$, and $\mathsf{RC}_\epsilon(f) \stackrel{\text{def}}{=} \max_{x\in f^{-1}} \mathsf{RC}^x_\epsilon(f)$.*

The above definition is stronger than the one in [A06].

**Definition 18 (Approximate Degree)** *Let $f : \{0,1\}^n \to \{0,1\}$ be a Boolean function and let $\epsilon > 0$. A polynomial $\mathbb{R}^n \to \mathbb{R}$ is said to $\epsilon$-approximate $f$, if $|p(x) - f(x)| < \epsilon$ for all $x \in f^{-1}$ and $0 \leq p(x) \leq 1$ for all $x \in \{0,1\}^n$. The $\epsilon$-approximate degree $\widetilde{\mathsf{deg}}_\epsilon(f)$ of $f$ is the minimum degree among all multi linear polynomials that $\epsilon$-approximate $f$. If $\epsilon = 0$ we write $\mathsf{deg}(f)$.*

**Definition 19 (Classical Adversary Bound)** *Let $f : \{0,1\}^n \to \{0,1\}^m$ be a function. Let $p = \{p_x : x \in \{0,1\}^n, p_x$ is a probability distribution on $[n]\}$. The classical adversary bound for $f$ denoted $\mathsf{cadv}(f)$, is defined as*

$$\mathsf{cadv}(f) \stackrel{\text{def}}{=} \min_p \max_{x,y:f(x)\neq f(y)} \frac{1}{\sum_{i:x_i\neq y_i} \min\{p_x(i), p_y(i)\}} \ .$$

The classical adversary bound is defined in an equivalent but slightly different way by Laplante and Magniez [LM08]; the above formulation appears in their proof and is made explicit in [SS06]. Aaronson [A08] defines a slightly weaker version as observed in [LM08]. Laplante and Magniez do not show an general upper bound for the classical adversary bound, but it is easy to see that $\mathsf{cadv}(f) = O(\mathsf{C}(f))$ for all total functions.

**Definition 20 (Quantum Adversary Bound)** *Let $f : \{0,1\}^n \to \mathcal{Z}$ be a function. Let $\Gamma$ be a Hermitian matrix whose rows and columns are labeled by elements in $\{0,1\}^n$, such that $\Gamma(x,y) = 0$ whenever $f(x) \neq f(y)$. For $i \in [n]$, let $D_i$ be a Boolean matrix whose rows and columns are labeled by elements in $\{0,1\}^n$, such that $D_i(x,y) = 1$ if $x_i \neq y_i$ and $D_i(x,y) = 0$ otherwise. The* quantum adversary bound *for $f$, denoted $\mathsf{adv}(f)$ is defined as*

$$\mathsf{adv} \stackrel{\text{def}}{=} \max_{\Gamma\neq 0} \frac{\|\Gamma\|}{\max_i \|\Gamma \circ D_i\|} \ .$$

## 3.2 Comparison between bounds

The following theorem captures the key relations between the above bounds. Below $\mathsf{R}_\epsilon(f)$ denotes the $\epsilon$-error randomized query complexity of $f$.

**Theorem 3** *Let $f : \{0,1\}^n \to \{0,1\}^m$ be a function and $\epsilon > 0$, then*

1. $\mathsf{R}_\epsilon(f) \geq \frac{1}{2} \log \mathsf{prt}_\epsilon(f)$.
2. $\log \mathsf{prt}_0(f) \geq \mathsf{C}(f)$.
3. *Let $\epsilon < 1/2$, then* $\log \mathsf{prt}_{\frac{\epsilon}{4}}(f) \geq \epsilon \cdot \mathsf{bs}(f) + \log \epsilon - 2$.
4. $\log \mathsf{prt}_\epsilon(f) \geq \mathsf{RC}_{\frac{2\epsilon}{1-2\epsilon}}(f) + \log \epsilon$.
5. $\log \mathsf{prt}_\epsilon(f) \geq (1 - 4\epsilon) \cdot \mathsf{cadv}(f) + \log \epsilon$.
6. *Let $f : \{0,1\}^n \to \{0,1\}$ be a Boolean function. Then,* $\log \mathsf{prt}_\epsilon(f) \geq \log \mathsf{sdisc}_{2\epsilon}(f) \geq \widetilde{\mathsf{deg}}_{2\epsilon}(f) + \log \epsilon$.
7. *Let $f : \{0,1\}^n \to \{0,1\}$ be a Boolean function. Then,* $\log \mathsf{sdisc}_\epsilon(f) \leq O(\widetilde{\mathsf{deg}}_\epsilon(f) \cdot \log n)$.
8. *Let $f : \{0,1\}^n \to \{0,1\}$ be a Boolean function. Then,* $\mathsf{D}(f) = O(\log \mathsf{prt}_0(f) \cdot \log \mathsf{prt}_{1/3}(f))$ *and* $\mathsf{D}(f) = O(\log \mathsf{prt}_{1/3}(f)^3)$, *where $\mathsf{D}(f)$ represents the deterministic query complexity of $f$.*

Consider the $\mathsf{Tribes}$ function $\mathsf{Tribes} : \{0,1\}^n \to \{0,1\}$, which is defined by an AND of $\sqrt{n}$ ORs of $\sqrt{n}$ variables. We show that the partition bound gives a tight lower bound for this function while no other general lower bounds methods as mentioned above gives tight lower bound.

**Theorem 4** *Let $\epsilon \in (0, 1/16)$, then*

$$\mathsf{R}_\epsilon(\mathsf{Tribes}) \geq \frac{1}{2} \log \mathsf{prt}_\epsilon(\mathsf{Tribes}) \geq \Omega(n),$$

*while* $\mathsf{C}(\mathsf{Tribes}), \mathsf{cadv}(\mathsf{Tribes}), \mathsf{adv}(\mathsf{Tribes}), \widetilde{\mathsf{deg}}(\mathsf{Tribes}) = O(\sqrt{n})$.

We give an example of a function $f$ such that the $\log \mathsf{prt}_0(f)$ is asymptotically larger than $\mathsf{R}_0(f)$, the Las-Vegas communication complexity. Let $T^h$ represent the complete binary $\mathsf{NAND}$ tree of height $h$. Let $f^h$ be the corresponding function evaluated by $T^h$ with its leaves serving as input variables to $f^h$. It is well known that $\mathsf{R}_0(f^h) = \Theta((\frac{1+\sqrt{33}}{4})^h)$ [SW86]. We show the following:

**Theorem 5** $\log \mathsf{prt}_0(f^h) = \Omega(2^h)$.

The following lemma shows that even the non-zero error partition bound sometimes can be larger than the degree.

**Lemma 6** *There is a function $f : \{0,1\}^n \to \{0,1\}$ such that* $\log \mathsf{prt}_\epsilon(f) \geq \Omega(\mathsf{bs}(f)) = \Omega(n)$, *while* $\mathsf{deg}(f) = O(n^{0.62})$ *for all $\epsilon < 1/2$.*

By composing $\mathsf{Tribes}$ with the function $f$ above we can also get a function for which $\log \mathsf{prt}_\epsilon$ is polynomially larger than $\mathsf{C}$ and $\mathsf{deg}$ simultaneously.

**Remark:** We remark, without proof, that the error in the partition bound (both communication and query) and its relatives can in general be boosted down in the same way as the error for randomized protocols, for example we have: For all relations $f$: $\log \mathsf{prt}_{2^{-k}}(f) = O(k \cdot \log \mathsf{prt}_{1/3}(f))$.

## 4 Open Questions

### Communication Complexity

1. Is $\mathsf{R}_{1/3}(f) \leq \mathsf{poly}(\log \mathsf{prt}_{1/3}(f))$ for all functions/relations $f$?
2. Is $\mathsf{prt}_{1/3}(\mathsf{Tribes}) = \Omega(n)$ ?

### Query Complexity

1. Is $\mathsf{R}_{1/3}(f) = O(\log^2 \mathsf{prt}_{1/3}(f))$ or better still is $\mathsf{R}_{1/3}(f) = O(\log \mathsf{prt}_{1/3}(f))$ ?
2. Is $\mathsf{adv}(f) = O(\log \mathsf{prt}_{1/3}(f))$ ?
3. Is $\mathsf{deg}(f) = \widetilde{O}(\mathsf{prt}_0(f))$ ?

**Acknowledgment:** The work done is supported by the internal grants of the Centre for Quantum Technologies, Singapore.

## References

[A06] S. Aaronson. Quantum certificate complexity. In *SIAM J. Comput.*, 35(4), pages 804–824, 2006.

[A08] S. Aaronson. Lower Bounds for Local Search by Quantum Arguments. In *J. Comput. Syst. Sci.*, 74(3), pages 313–322, 2008.

[AA05] S. Aaronson and A. Ambainis. Quantum search of spatial regions. In *Theory of Computing*, 1(1), pages 47–79, 2005.

[A02] A. Ambainis. Quantum Lower Bounds by Quantum Arguments. In *J. Comput. Syst. Sci.*, 64(4), pages 750–767, 2002.

[BFS86] L. Babai, P. Frankl, and J. Simon. Complexity classes in communication complexity theory. In *Proceedings of 27th IEEE FOCS*, pages 337–347, 1986.

[BKKS04] Z. Bar-Yossef, T. S. Jayram, R. Kumar, D. Sivakumar. An information statistics approach to data stream and communication complexity. In *J. Comput. Syst. Sci.*, 68(4): 702-772, 2004.

[BPSW06] P. Beame, T. Pitassi, N. Segerlind, A. Wigderson. A Strong Direct Product Theorem for Corruption and the Multiparty Communication Complexity of Disjointness. In *Computational Complexity*, 15(4): 391-432, 2006.

[BBC+01] R. Beals, H. Buhrman, R. Cleve, M. Mosca, and R. de Wolf. Quantum lower bounds by polynomials. *Journal of the ACM*, 48(4):778–797, 2001. Earlier version in FOCS'98. quant-ph/9802049.

[BW02] H. Buhrman and R. de Wolf. Complexity measures and decision tree complexity: A survey. *Theoretical Computer Science*, 288(1):21–43, 2002.

[CA08] A. Chattopadhyay, A. Ada. Multiparty Communication Complexity of Disjointness. ECCC Technical Report 15(002), 2008.

[JRS03] R. Jain, J. Radhakrishnan, P. Sen. A Lower Bound for the Bounded Round Quantum Communication Complexity of Set Disjointness. In *FOCS 2003*, pages 220-229, 2003.

[JKS03] T. S. Jayram, R. Kumar, D. Sivakumar. Two applications of information complexity. In *STOC 2003*, pages 673-682, 2003.

[KS92] B. Kalyanasundaram and G. Schnitger. The probabilistic communication complexity of set intersection. *SIAM Journal on Discrete Mathematics*, 5(4):545–557, 1992.

[KKN95] M. Karchmer, E. Kushilevitz, N. Nisan. Fractional Covers and Communication Complexity. In *SIAM J. Discrete Math.*, 8(1): 76-92, 1995.

[K03] H. Klauck. Rectangle Size Bounds and Threshold Covers in Communication Complexity. In: *IEEE Conference on Computational Complexity 2003*, pages 118-134, 2003.

[K07] H. Klauck. Lower Bounds for Quantum Communication Complexity. *SIAM J. Comput.*, 37(1): 20-46, 2007.

[K09] H. Klauck. A Strong Direct Product Theorem for Disjointness. arXiv:0908.2940, 2009.

[Kre95] I. Kremer. Quantum communication. Master's thesis, Hebrew University, Computer Science Department, 1995.

[KN97] E. Kushilevitz and N. Nisan. *Communication Complexity*. Cambridge University Press, 1997.

[LM08] S. Laplante, F. Magniez. Lower Bounds for Randomized and Quantum Query Complexity Using Kolmogorov Arguments. *SIAM Journal on Computing*, 38(1), pages 46–62, 2008.

[LS09] N. Linial, A. Shraibman. Lower bounds in communication complexity based on factorization norms. *Random Struct. Algorithms*, 34(3), pages 368–394, 2009.

[L90] L. Lovasz. Communication Complexity: A Survey. In *Paths, Flows, and VLSI Layout, edited by B. H. Korte*, Springer, 1990.

[Nis91] N. Nisan. CREW PRAMs and decision trees. *SIAM Journal on Computing*, 20(6):999–1007, 1991.

[NS94] N. Nisan and M. Szegedy. On the degree of Boolean functions as real polynomials. *Computational Complexity*, 4(4):301–313, 1994.

[NW95] N. Nisan and A. Wigderson. On Rank vs. Communication Complexity. *Combinatorica* , 15(4):557–565, 1995.

[Raz92] A. Razborov. On the distributional complexity of disjointness. *Theoretical Computer Science*, 106(2):385–390, 1992.

[Raz03] A. Razborov. Quantum communication complexity of symmetric predicates. *Izvestiya of the Russian Academy of Science, mathematics*, 67(1):159–176, 2003. quant-ph/0204025.

[SW86] M. Saks, A. Wigderson. Probabilistic boolean decision trees and the complexity of evaluating game trees. *FOCS 1986*, pages 29-38-94, 1986.

[S08] A. A. Sherstov. The pattern matrix method for lower bounds on quantum communication. *STOC 2008*, pages 85-94, 2008.

[SS06] R. Spalek, M. Szegedy. All Quantum Adversary Methods are Equivalent. *Theory of Computing*, 2(1):1–18, 2006.

[Yao79] A. C-C. Yao. Some Complexity Questions Related to Distributive Computing . In *Proceedings of STOC 1979*, pages 209–213, 1979.

[Y83] A.C.C. Yao. Lower Bounds by Probabilistic Arguments. *24th IEEE Symp. Foundations of Computer Science*, pp. 420–428, 1983.

# A Proofs

**Proof of Theorem 1:**

1. Let $\mathcal{P}$ be a public coin randomized protocol for $f$ with communication $c \stackrel{\mathsf{def}}{=} \mathsf{R}^{\mathsf{pub}}_{\epsilon}(f)$ and worst case error $\epsilon$. For binary string $r$, let $\mathcal{P}_r$ represent the deterministic protocol obtained from $\mathcal{P}$ on fixing the public coins to $r$. Let $r$ occur with probability $q(r)$ in $\mathcal{P}$. Every deterministic protocol amounts to partitioning the inputs in $\mathcal{X} \times \mathcal{Y}$ into rectangles. Let $\mathcal{R}_r$ be the set of rectangles corresponding to different communication strings between Alice and Bob in $\mathcal{P}_r$. We know that $|\mathcal{R}_r| \leq 2^c$, since the communication in $\mathcal{P}_r$ is at most $c$ bits. Let $z^r_R \in \mathcal{Z}$ be the output corresponding to rectangle $R$ in $\mathcal{P}_r$. Let

$$w'_{z,R} \stackrel{\mathsf{def}}{=} \sum_{r: R \in \mathcal{R}_r \text{ and } z^r_R = z} q(r) \ .$$

It is easily 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 $\epsilon$ on all inputs in $f^{-1}$ we get the constraints:

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

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

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

Of course we also have by construction : $\forall z, \forall R : w'_{z,R} \geq 0$. Therefore $\{w'_{z,R} : z \in \mathcal{Z}, R \in \mathcal{R}\}$ is feasible for the primal of $\mathsf{prt}_\epsilon(f)$. Hence,

$$\mathsf{prt}_\epsilon(f) \leq \sum_z \sum_R w'_{z,R} = \sum_r q(r) \cdot |\mathcal{R}_r| \leq 2^c \sum_r q(r) = 2^c \ .$$

2. Fix $z' \in \mathcal{Z}$. We will show that $\mathsf{srec}_\epsilon^{z'}(f) \leq \mathsf{prt}_\epsilon(f)$; this will imply $\mathsf{srec}_\epsilon(f) \leq \mathsf{prt}_\epsilon(f)$. Let $\{w_{z,R} : z \in \mathcal{Z}, R \in \mathcal{R}\}$ be an optimal solution of the primal for $\mathsf{prt}_\epsilon(f)$. Let us define $\forall R \in \mathcal{R} : w_R \stackrel{\mathsf{def}}{=} w_{z',R}$, hence $\forall R \in \mathcal{R}, w_R \geq 0$. Now,

$$\begin{aligned} \forall (x,y) \in f^{-1}(z') : \sum_{R:(x,y)\in R} w_{z',R} \geq 1-\epsilon \quad &\Rightarrow \quad \sum_{R:(x,y)\in R} w_R \geq 1-\epsilon, \\ \forall (x,y) \in f^{-1} - f^{-1}(z') : \sum_{R:(x,y)\in R} w_{f(x,y),R} \geq 1-\epsilon \quad &\Rightarrow \quad \sum_{R:(x,y)\in R} w_R \leq \epsilon, \\ \forall (x,y) : \sum_{R:(x,y)\in R} \sum_z w_{z,R} = 1 \quad &\Rightarrow \quad \sum_{R:(x,y)\in R} w_R \leq 1 \ . \end{aligned}$$

Hence $\{w_R : R \in \mathcal{R}\}$ forms a feasible solution to the primal for $\mathsf{srec}_\epsilon^z(f)$ which implies

$$\mathsf{srec}_\epsilon^z(f) \leq \sum_R w_R \leq \sum_z \sum_R w_{z,R} = \mathsf{prt}_\epsilon(f) \ .$$

3. Fix $z \in \mathcal{Z}$. Since the primal program for $\mathsf{srec}_\epsilon^z(f)$ has extra constraints over the primal program for $\mathsf{rec}_\epsilon^z(f)$, it implies that $\mathsf{rec}_\epsilon^z(f) \leq \mathsf{srec}_\epsilon^z(f)$. Hence $\mathsf{rec}_\epsilon(f) \leq \mathsf{srec}_\epsilon(f)$.

4. (Sketch) Let $W \stackrel{\mathsf{def}}{=} \{w_{z,R}\}$ be an optimal solution to the primal for $\mathsf{prt}_0 f$. It is easily seen that

$$w_{z,R} > 0 \Rightarrow ((x,y) \in R \Rightarrow f(x,y) = z) \ .$$

Using standard Chernoff type arguments we can argue that there exists subset $W' \subseteq W$ with $|W'| = O(n\mathsf{prt}_0 f)$ such that :

$$\forall (x,y) \in f^{-1} : \sum_{R:(x,y)\in R, w_{f(x,y),R} \in W'} w_{f(x,y),R} > 0 \ .$$

Hence $W'$ is a cover of $\mathcal{X} \times \mathcal{Y}$ using monochromatic rectangles. Now using arguments as in Theorem 2.11 of [KN97] it follows that $\mathsf{D}(f) = O((\log \mathsf{prt}_0 f + \log n)^2)$.

5. Define $\mu_{x,y} \stackrel{\mathsf{def}}{=} 1; \phi_{x,y} \stackrel{\mathsf{def}}{=} 0$ iff $(x,y) \in S$ and $\mu_{x,y} = \phi_{x,y} \stackrel{\mathsf{def}}{=} 0$ otherwise. Since no two elements of $S$ can appear in the same rectangle, it is easily seen that the constraints for the dual of $\mathsf{prt}_0(f)$ are satisfied by $\{\mu_{x,y}, \phi_{x,y}\}$ . Hence $\mathsf{prt}_0(f) \geq \sum_{(x,y)} (\mu_{x,y} - \phi_{x,y}) = |S|$.

$\square$

**Proof of Lemma 1:**

1. Fix $z \in \mathcal{Z}$. Let $k \stackrel{\text{def}}{=} \mathsf{rec}^z_\epsilon(f)$. Let $\{\mu_{x,y} : (x,y) \in \mathcal{X} \times \mathcal{Y}\}$ be an optimal solution to the dual for $\mathsf{rec}^z_\epsilon(f)$. We can assume without loss of generality that $(x,y) \notin f^{-1} \Rightarrow \mu_{x,y} = 0$. Let $k_1 \stackrel{\text{def}}{=} \sum_{(x,y)\in f^{-1}(z)} \mu_{x,y}$ and $k_2 \stackrel{\text{def}}{=} \sum_{(x,y)\in f^{-1}-f^{-1}(z)} \mu_{x,y}$. Then,

$$\begin{aligned} k &= (1-\epsilon) \sum_{(x,y)\in f^{-1}(z)} \mu_{x,y} - \epsilon \sum_{(x,y)\in f^{-1}-f^{-1}(z)} \mu_{x,y} \\ \Rightarrow \quad k &= (1-\epsilon)k_1 - \epsilon k_2 \\ \Rightarrow \quad k_1 &\geq k \text{ and } k_1 \geq \epsilon k_2 \quad (\text{since } k, k_2 \geq 0) \ . \end{aligned} \tag{1}$$

Let us define $\lambda_{x,y} \stackrel{\text{def}}{=} \frac{\mu_{x,y}}{2k_1}$ iff $f(x,y) = z$ and $\lambda_{x,y} \stackrel{\text{def}}{=} \frac{\mu_{x,y}}{2k_2}$, otherwise. It is easily seen that $\lambda$ is a distribution on $\mathcal{X} \times \mathcal{Y} \cap f^{-1}$ and $\lambda(f^{-1}(z)) = 0.5$. For all $R \in \mathcal{R}$,

$$\begin{aligned} \sum_{(x,y)\in f^{-1}(z)\cap R} \mu_{x,y} - \sum_{(x,y)\in (R\cap f^{-1})-f^{-1}(z)} \mu_{x,y} &\leq 1 \\ \Rightarrow \sum_{(x,y)\in f^{-1}(z)\cap R} 2k_1\lambda_{x,y} - \sum_{(x,y)\in R-f^{-1}(z)} 2k_2\lambda_{x,y} &\leq 1 \\ \Rightarrow \sum_{(x,y)\in f^{-1}(z)\cap R} 2k_1\lambda_{x,y} - \sum_{(x,y)\in R-f^{-1}(z)} \frac{2k_1}{\epsilon}\lambda_{x,y} &\leq 1 \quad (\text{from (1)}) \\ \Rightarrow \epsilon\left(\sum_{(x,y)\in f^{-1}(z)\cap R} \lambda_{x,y} - \frac{1}{2k_1}\right) &\leq \sum_{(x,y)\in R-f^{-1}(z)} \lambda_{x,y} \\ \Rightarrow \epsilon\left(\sum_{(x,y)\in f^{-1}(z)\cap R} \lambda_{x,y} - \frac{1}{2k}\right) &\leq \sum_{(x,y)\in R-f^{-1}(z)} \lambda_{x,y} \quad (\text{from (1)}) \end{aligned}$$

Let $R \in \mathcal{R}$ be such that $\sum_{(x,y)\in f^{-1}(z)\cap R} \lambda_{x,y} \geq \frac{1}{k}$. Then we have from above

$$\frac{\epsilon}{2}\left(\sum_{(x,y)\in f^{-1}(z)\cap R} \lambda_{x,y}\right) \leq \sum_{(x,y)\in R-f^{-1}(z)} \lambda_{x,y} \ . \tag{2}$$

Therefore by definition $\widetilde{\mathsf{rec}}^{z,\lambda}_{\frac{\epsilon}{2}}(f) \geq k$ which implies $\widetilde{\mathsf{rec}}^{z}_{\frac{\epsilon}{2}}(f) \geq k$.

2. Fix $z \in \mathcal{Z}$. Let $k = \widetilde{\mathsf{rec}}^z_{2\epsilon}(f)$. Let $\lambda$ be a distribution on $\mathcal{X} \times \mathcal{Y} \cap f^{-1}$ such that $\widetilde{\mathsf{rec}}^z_{2\epsilon}(f) = \widetilde{\mathsf{rec}}^{z,\lambda}_{2\epsilon}(f)$ and $\lambda(f^{-1}(z)) \geq 0.5$. Let us define $\mu_{x,y} \stackrel{\text{def}}{=} k \cdot \lambda_{x,y}$ iff $f(x,y) = z$; $\mu_{x,y} \stackrel{\text{def}}{=} k \cdot \frac{\lambda_{x,y}}{2\epsilon}$ iff $(x,y) \in f^{-1} - f^{-1}(z)$ and $\mu_{x,y} = 0$ otherwise. Now let $R \in \mathcal{R}$ be such that $\lambda(f^{-1}(z)\cap R) \leq \frac{1}{k}$, then

$$\sum_{(x,y)\in f^{-1}(z)\cap R} \lambda_{x,y} \leq \frac{1}{k} \quad \Rightarrow \quad \sum_{(x,y)\in f^{-1}(z)\cap R} \mu_{x,y} \leq 1 \ .$$

Let $\lambda(f^{-1}(z)\cap R) > \frac{1}{k}$, then

$$\begin{aligned} 2\epsilon \sum_{(x,y)\in f^{-1}(z)\cap R} \lambda_{x,y} &\leq \sum_{(x,y)\in R-f^{-1}(z)} \lambda_{x,y} \\ \Rightarrow \sum_{(x,y)\in f^{-1}(z)\cap R} \mu_{x,y} &\leq \sum_{(x,y)\in (R\cap f^{-1})-f^{-1}(z)} \mu_{x,y} \ . \end{aligned}$$

Hence the constraints of the dual program for $\mathsf{rec}^z_\epsilon(f)$ are satisfied by $\{\mu_{x,y} : (x,y) \in \mathcal{X} \times \mathcal{Y}\}$. Now,

$$\begin{aligned}
\mathsf{rec}^z_\epsilon(f) &\geq \sum_{(x,y)\in f^{-1}(z)} (1-\epsilon)\cdot \mu_{x,y} - \sum_{(x,y)\in f^{-1}-f^{-1}(z)} \epsilon \cdot \mu_{x,y} \\
&= k \cdot \left( \sum_{(x,y)\in f^{-1}(z)} (1-\epsilon)\cdot \lambda_{x,y} - \sum_{(x,y)\in f^{-1}-f^{-1}(z)} \frac{\lambda_{x,y}}{2} \right) \\
&\geq \frac{k}{2} \cdot (\frac{1}{2} - \epsilon) \quad \text{(since } \lambda(f^{-1}(z)) \geq 0.5\text{)} \ .
\end{aligned}$$

□

**Proof of Lemma 2:**

1. Fix $z \in \mathcal{Z}$. Let $\{\mu_{x,y}, \phi_{x,y} : (x,y) \in \mathcal{X} \times \mathcal{Y}\}$ be an optimal solution to the dual for $\mathsf{srec}^z_\epsilon(f)$. We can assume w.l.o.g. that $(x,y) \notin f^{-1} \Rightarrow \mu_{x,y} = \phi_{x,y} = 0$; also that $(x,y) \notin f^{-1}(z) \Rightarrow \phi_{x,y} = 0$. Let us observe that we can assume w.l.o.g. that $\forall (x,y) \in f^{-1}(z)$, either $\mu_{x,y} = 0$ or $\phi_{x,y} = 0$. Otherwise let us say that for some $(x,y) \in f^{-1}(z) : \mu_{x,y} \geq \phi_{x,y} > 0$. Then using $\mu'_{x,y} \overset{\mathsf{def}}{=} \mu_{x,y} - \phi_{x,y}$ and $\phi'_{x,y} \overset{\mathsf{def}}{=} 0$ instead of $(\mu_{x,y}, \phi_{x,y})$, and the rest the same, is a strictly better solution; that is the objective function is strictly larger in the new case. A similar argument can be made if for some $(x,y) \in f^{-1}(z) : \phi_{x,y} \geq \mu_{x,y} > 0$.

   Let $g : \mathcal{X} \times \mathcal{Y} \to \mathcal{Z}$ be such that $g(x,y) = f(x,y)$ iff $\phi_{x,y} = 0$ and $g(x,y) \neq f(x,y)$ otherwise ($g$ remains undefined wherever $f$ is undefined). For all $(x,y)$ let $\mu'_{x,y} \overset{\mathsf{def}}{=} \mu_{x,y}$ iff $\phi_{x,y} = 0$ and $\mu'_{x,y} = \phi_{x,y}$ otherwise. Then $\forall (x,y), \mu'_{x,y} \geq 0$ and

$$\begin{aligned}
&\forall R \in \mathcal{R} : \sum_{(x,y)\in f^{-1}(z)\cap R} (\mu_{x,y} - \phi_{x,y}) - \sum_{(x,y)\in (R\cap f^{-1})-f^{-1}(z)} \mu_{x,y} \leq 1 \\
&\Rightarrow \forall R \in \mathcal{R} : \sum_{(x,y)\in g^{-1}(z)\cap R} \mu'_{x,y} - \sum_{(x,y)\in (R\cap g^{-1})-g^{-1}(z)} \mu'_{x,y} \leq 1 \ . \qquad (3)
\end{aligned}$$

   Hence $\{\mu'_{x,y} : (x,y) \in \mathcal{X} \times \mathcal{Y}\}$ is a feasible solution to the dual of $\mathsf{rec}^z_\epsilon(g)$. Now,

$$\begin{aligned}
k &\overset{\mathsf{def}}{=} \sum_{(x,y)\in g^{-1}(z)} (1-\epsilon)\cdot \mu'_{x,y} - \sum_{(x,y)\in g^{-1}-g^{-1}(z)} \epsilon \cdot \mu'_{x,y} \qquad (4) \\
&= \sum_{(x,y)\in f^{-1}(z)} (1-\epsilon)\cdot \mu_{x,y} - \sum_{(x,y)\in f^{-1}(z)} \epsilon \cdot \phi_{x,y} - \sum_{(x,y)\in f^{-1}-f^{-1}(z)} \epsilon \cdot \mu_{x,y} \\
&\geq \sum_{(x,y)\in f^{-1}(z)} (1-\epsilon)\cdot \mu_{x,y} - \sum_{(x,y)\in f^{-1}(z)} \phi_{x,y} - \sum_{(x,y)\in f^{-1}-f^{-1}(z)} \epsilon \cdot \mu_{x,y} \\
&= \mathsf{srec}^z_\epsilon(f) \ . \qquad (5)
\end{aligned}$$

   Let $k_1 \overset{\mathsf{def}}{=} \sum_{(x,y)\in g^{-1}(z)} \mu'_{x,y}$ and $k_2 \overset{\mathsf{def}}{=} \sum_{(x,y)\in g^{-1}-g^{-1}(z)} \mu'_{x,y}$. Let $\lambda_{x,y} \overset{\mathsf{def}}{=} \frac{\mu'_{x,y}}{2k_1}$ iff $g(x,y) = z$ and $\lambda_{x,y} \overset{\mathsf{def}}{=} \frac{\mu'_{x,y}}{2k_2}$, otherwise. It is clear that $\lambda$ is a distribution on $\mathcal{X} \times \mathcal{Y} \cap g^{-1}$ and $\lambda(g^{-1}(z)) = 0.5$. As in the proof of Part 1. of Lemma 1, using (3) and (4), we can argue that

$\widetilde{\mathsf{rec}}^{z,\lambda}_{\frac{\epsilon}{2}}(g) \geq \mathsf{rec}^{z,\lambda}_{\epsilon}(g) \geq k$. Also since $\sum_{(x,y)\in f^{-1}}((1-\epsilon)\mu_{x,y}-\phi_{x,y}) \geq 0$ and $\sum_{(x,y)\in f^{-1}(z)}(\mu_{x,y}-\phi_{x,y})-\sum_{(x,y)\in(f^{-1})-f^{-1}(z)}\mu_{x,y} \leq 1$ we can argue that $\sum_{(x,y)\in f^{-1}(z)}\phi_{x,y} \leq (1-\epsilon)k_2$ (we assume $\mathsf{srec}^z_\epsilon(f)$ is at least a large constant) . Therefore,

$$\Pr_{(x,y)\leftarrow\lambda}[g(x,y) \neq f(x,y)] = \sum_{(x,y)\in f^{-1}(z)} \frac{\phi_{x,y}}{2k_2} \leq \frac{1-\epsilon}{2} \ .$$

Hence by definition, $\widetilde{\mathsf{srec}}^z_{\frac{\epsilon}{2},\frac{1-\epsilon}{2}}(f) \geq \widetilde{\mathsf{srec}}^{z,\lambda}_{\frac{\epsilon}{2},\frac{1-\epsilon}{2}}(f) \geq \widetilde{\mathsf{rec}}^{z,\lambda}_{\frac{\epsilon}{2}}(g) \geq k \geq \mathsf{srec}^z_\epsilon(f)$. The last inequality follows from (5).

2. Fix $z \in \mathcal{Z}$. Let $k \stackrel{\mathsf{def}}{=} \widetilde{\mathsf{srec}}^z_{2\epsilon,\frac{\epsilon}{2}}(f)$. Let $\lambda$ be distribution on $\mathcal{X}\times\mathcal{Y}\cap f^{-1}$ such that $\widetilde{\mathsf{srec}}^z_{2\epsilon,\frac{\epsilon}{2}}(f) = \widetilde{\mathsf{srec}}^{z,\lambda}_{2\epsilon,\frac{\epsilon}{2}}(f)$. Let $g : \mathcal{X}\times\mathcal{Y} \to \mathcal{Z}$ be a function such that $\widetilde{\mathsf{srec}}^{z,\lambda}_{2\epsilon,\frac{\epsilon}{2}}(f) = \mathsf{rec}^{z,\lambda}_{2\epsilon}(g)$ and $\lambda(g^{-1}(z)) \geq 0.5$ and $\lambda(f \neq g) \leq \epsilon/2$. Note that we can assume w.l.o.g. that $g(x,y) \neq f(x,y) \Rightarrow f(x,y) = z$.

   For $(x,y) \in f^{-1}$, let us define $\mu_{x,y} \stackrel{\mathsf{def}}{=} k\cdot\lambda_{x,y}$ iff $g(x,y) = f(x,y) = z$ and $\mu_{x,y} \stackrel{\mathsf{def}}{=} k\cdot\frac{\lambda_{x,y}}{2\epsilon}$ iff $f(x,y) \neq z$. Let $\phi_{x,y} \stackrel{\mathsf{def}}{=} k\cdot\frac{\lambda_{x,y}}{2\epsilon}$ iff $z = f(x,y) \neq g(x,y)$. For $(x,y) \notin f^{-1}$, let $\mu_{x,y} = \phi_{x,y} = 0$. Now let $R \in \mathcal{R}$ be such that $\lambda(g^{-1}(z)\cap R) \leq \frac{1}{k}$, then

$$\begin{aligned} \sum_{(x,y)\in g^{-1}(z)\cap R} \lambda_{x,y} \leq \frac{1}{k} \quad \Rightarrow \quad \sum_{(x,y)\in g^{-1}(z)\cap R} \mu_{x,y} &\leq 1 \\ \Rightarrow \sum_{(x,y)\in f^{-1}(z)\cap R} \mu_{x,y} - \phi_{x,y} &\leq 1 \ . \end{aligned}$$

   Let $\lambda(g^{-1}(z)\cap R) > \frac{1}{k}$, then

$$\begin{aligned} 2\epsilon \sum_{(x,y)\in g^{-1}(z)\cap R} \lambda_{x,y} &\leq \sum_{(x,y)\in R-g^{-1}(z)} \lambda_{x,y} \\ \Rightarrow \sum_{(x,y)\in g^{-1}(z)\cap R} \mu_{x,y} &\leq \sum_{(x,y)\in R-g^{-1}(z)} \mu_{x,y} + \phi_{x,y} \\ \Rightarrow \sum_{(x,y)\in f^{-1}(z)\cap R} \mu_{x,y} - \phi_{x,y} &\leq \sum_{(x,y)\in(R\cap f^{-1})-f^{-1}(z)} \mu_{x,y} \ . \end{aligned}$$

   Hence the constraints of the dual program for $\mathsf{srec}^z_\epsilon(f)$ are satisfied by $\{\mu_{x,y},\phi_{x,y} : (x,y) \in \mathcal{X}\times\mathcal{Y}\}$. Now,

$$\begin{aligned} \mathsf{srec}^z_\epsilon(f) &\geq \sum_{(x,y)\in f^{-1}(z)} ((1-\epsilon)\cdot\mu_{x,y} - \phi_{x,y}) - \sum_{(x,y)\in f^{-1}-f^{-1}(z)} \epsilon\cdot\mu_{x,y} \\ &\geq \sum_{(x,y)\in g^{-1}(z)} (1-\epsilon)\cdot\mu_{x,y} - \sum_{(x,y)\in f^{-1}(z)} \phi_{x,y} - \sum_{(x,y)\notin g^{-1}(z)} \epsilon\cdot\mu_{x,y} \\ &= k\cdot\left(\sum_{(x,y)\in g^{-1}(z)} (1-\epsilon)\cdot\lambda_{x,y} - \frac{1}{2\epsilon}\sum_{(x,y):f(x,y)\neq g(x,y)} \lambda_{x,y} - \sum_{(x,y)\notin g^{-1}(z)} \frac{\lambda_{x,y}}{2}\right) \\ &\geq \frac{k}{2}\cdot(\frac{1}{4}-\epsilon) \ . \end{aligned}$$

   The last inequality follows since $\lambda(g^{-1}(z)) \geq 0.5$ and $\lambda(f \neq g) \leq \epsilon/2$.

$\square$

**Proof of Lemma 3:**

1. Let $k \stackrel{\mathsf{def}}{=} \mathsf{sdisc}_\epsilon(f)$. Let $\{\mu_{x,y}, \phi_{x,y}\}$ be an optimal solution to the dual for $\mathsf{sdisc}_\epsilon(f)$. As in the proof of Lemma 2, we can argue that for all $(x,y) \in f^{-1}$, either $\mu_{x,y} = 0$ or $\phi_{x,y} = 0$. For $(x,y) \in f^{-1}$, let us define $\lambda'_{x,y} \stackrel{\mathsf{def}}{=} \max\{\mu_{x,y}, \phi_{x,y}\}$ and let $\lambda_{x,y} \stackrel{\mathsf{def}}{=} \frac{\lambda'_{x,y}}{\sum_{(x,y)\in f^{-1}} \lambda'_{x,y}}$. It is clear that $\lambda$ is a distribution on $f^{-1}$. Let us define $g : \mathcal{X} \times \mathcal{Y} \to \{0,1\}$ such that $g^{-1} = f^{-1}$. For $(x,y) \in f^{-1}$, let $g(x,y) = f(x,y)$ iff $\phi_{x,y} = 0$ and let $g(x,y) \neq f(x,y)$ iff $\phi_{x,y} \neq 0$. Now
$$\begin{aligned} & \forall R : \Big|\sum_{(x,y)\in f^{-1}(1)\cap R} (\mu_{x,y} - \phi_{x,y}) - \sum_{(x,y)\in R\cap f^{-1}(0)} (\mu_{x,y} - \phi_{x,y})\Big| \leq 1 \\ \Rightarrow \quad & \forall R : \Big|\sum_{(x,y)\in g^{-1}(1)\cap R} \lambda'_{x,y} - \sum_{(x,y)\in R\cap g^{-1}(0)} \lambda'_{x,y}\Big| \leq 1 \\ \Rightarrow \quad & \forall R : \Big|\sum_{(x,y)\in g^{-1}(1)\cap R} \lambda_{x,y} - \sum_{(x,y)\in R\cap g^{-1}(0)} \lambda_{x,y}\Big| \leq \frac{1}{\sum_{x,y} \mu_{x,y} + \phi_{x,y}} \leq \frac{1}{k} \ . \end{aligned}$$
Hence $\mathsf{disc}^\lambda(g) \geq k$. Also since $\sum_{(x,y)} \mu_{x,y} - (1+\epsilon)\phi_{x,y} \geq 0$,
$$\Pr_{(x,y)\leftarrow\lambda}[g(x,y) \neq f(x,y)] = \frac{1}{\sum_{x,y} \mu_{x,y} + \phi_{x,y}} \sum_{(x,y)} \phi_{x,y} < \frac{1}{2+\epsilon} \leq \frac{1}{2} - \frac{\epsilon}{8} \ .$$
Hence our result.

2. Let $\delta \stackrel{\mathsf{def}}{=} \frac{1}{4+2\epsilon}$. Let $\lambda$ be a distribution on $f^{-1}$ such that $k \stackrel{\mathsf{def}}{=} \widetilde{\mathsf{sdisc}}_\delta(f) = \widetilde{\mathsf{sdisc}}_\delta^\lambda(f)$ and $\Pr_{(x,y)\leftarrow\lambda}[g(x,y) \neq f(x,y)] < \delta$. For $(x,y) \in f^{-1}$, let $\mu_{x,y} \stackrel{\mathsf{def}}{=} k \cdot \lambda_{x,y}; \phi_{x,y} = 0$ iff $f(x,y) = g(x,y)$ and $\phi_{x,y} \stackrel{\mathsf{def}}{=} k \cdot \lambda_{x,y}; \mu_{x,y} = 0$ iff $f(x,y) \neq g(x,y)$. Then,
$$\begin{aligned} & \forall R : \Big|\sum_{(x,y)\in g^{-1}(1)\cap R} \lambda_{x,y} - \sum_{(x,y)\in R\cap g^{-1}(0)} \lambda_{x,y}\Big| \leq \frac{1}{k} \\ \Rightarrow \quad & \forall R : \Big|\sum_{(x,y)\in f^{-1}(1)\cap R} (\mu_{x,y} - \phi_{x,y}) - \sum_{(x,y)\in R\cap f^{-1}(0)} (\mu_{x,y} - \phi_{x,y})\Big| \leq 1 \ . \end{aligned}$$
Hence $\{\mu_{x,y}, \phi_{x,y}\}$ form a feasible solution to the dual for $\mathsf{sdisc}_\epsilon(f)$. Now,
$$\mathsf{sdisc}_\epsilon(f) \geq \sum_{(x,y)} \mu_{x,y} - (1+\epsilon)\phi_{x,y} > k((1-\delta) - (1+\epsilon)\delta) = k(1 - (2+\epsilon)\delta) = \frac{k}{2} \ .$$

$\square$

**Proof of Lemma 4:** Let $k \stackrel{\mathsf{def}}{=} \mathsf{disc}^\lambda(f)$. Let $\forall (x,y) \in f^{-1} : \mu_{x,y} \stackrel{\mathsf{def}}{=} k \cdot \lambda_{x,y}$ and $\mu_{x,y} = 0$ otherwise. Then we have:
$$\begin{aligned} & \forall R : \sum_{(x,y)\in R\cap f^{-1}(z)} \lambda_{x,y} - \sum_{(x,y)\in R - f^{-1}(z)} \lambda_{x,y} \leq \frac{1}{k} \\ \Rightarrow & \forall R : \sum_{(x,y)\in R\cap f^{-1}(z)} \mu_{x,y} - \sum_{(x,y)\in (R\cap f^{-1}) - f^{-1}(z)} \mu_{x,y} \leq 1 \ . \end{aligned}$$

Hence the constraints for the dual of the linear program for $\mathsf{rec}^z_\epsilon(f)$ are satisfied by $\{\mu_{x,y} : (x,y) \in \mathcal{X} \times \mathcal{Y}\}$. Now,

$$\begin{aligned}
\mathsf{rec}^z_\epsilon(f) &\geq \sum_{(x,y)\in f^{-1}(z)} (1-\epsilon)\cdot \mu_{x,y} - \sum_{(x,y)\in f^{-1}-f^{-1}(z)} \epsilon \cdot \mu_{x,y} \\
&= k \cdot \left( \sum_{(x,y)\in f^{-1}(z)} (1-\epsilon)\cdot \lambda_{x,y} - \sum_{(x,y)\notin f^{-1}(z)} \epsilon \cdot \lambda_{x,y} \right) \\
&= k \cdot \left( \sum_{(x,y)\in f^{-1}(z)} \lambda_{x,y} - \epsilon \right) \\
&\geq k \cdot \left( \frac{1}{2} - \frac{1}{2k} - \epsilon \right) = (\frac{1}{2} - \epsilon)k - \frac{1}{2} \enspace .
\end{aligned}$$

The last inequality follows since $\mathsf{disc}^\lambda(f) = k$. $\square$

**Proof of Lemma 5:** For a sign matrix $A$, let $g_A$ be the corresponding function given by $g_A(x,y) \stackrel{\mathsf{def}}{=} (1 - A(x,y))/2$. For distribution $\lambda$ on $\mathcal{X} \times \mathcal{Y}$, let $P_\lambda$ be the matrix defined by $P_\lambda(x,y) \stackrel{\mathsf{def}}{=} \lambda(x,y)$. For matrix $B$, define $||B||_\Sigma \stackrel{\mathsf{def}}{=} \sum_{i,j} |B(i,j)|$. For matrices $C, D$, let $C \circ D$ denote the entry wise *Hadamard product* of $C, D$. We have the following facts:

**Fact 1 ([LS09])** *For every sign matrix $A$,*

$$\gamma_2^\alpha(A) = \max_B \frac{1}{2\gamma_2^*(B)} \left( (\alpha+1) \langle A, B\rangle - (\alpha - 1)||B||_\Sigma \right) \enspace .$$

*Above, $\gamma_2^*(\cdot)$ is the dual norm of $\gamma_2(\cdot)$.*

**Fact 2 ([LS09])** *Let $A$ be a sign matrix and let $\lambda$ be a distribution. Then,*

$$\frac{1}{8\gamma_2^*(A\circ P_\lambda)} \leq \mathsf{disc}^\lambda(g_A) \leq \frac{1}{\gamma_2^*(A \circ P_\lambda)} \enspace .$$

Therefore we have,

$$\begin{aligned}
\gamma_2^\alpha(A_f) &= \max_B \frac{1}{2\gamma_2^*(B)} \left( (\alpha+1) \langle A_f, B\rangle - (\alpha-1)||B||_\Sigma \right) \\
&= \max_{B:||B||_\Sigma = 1} \frac{1}{2\gamma_2^*(B)} \left( (\alpha+1) \langle A_f, B\rangle - (\alpha-1) \right) \\
&= \max_{g,\lambda} \frac{1}{\gamma_2^*(A_g \circ P_\lambda)} \left( 1 - (\alpha+1)\lambda(f \neq g) \right) \\
&\leq \max_{g,\lambda} 8 \cdot \mathsf{disc}^\lambda(g) \left( 1 - (\alpha+1)\lambda(f\neq g) \right) \\
&\leq \max\{8\cdot \mathsf{disc}^\lambda(g) : g, \lambda \text{ such that } \lambda(f \neq g) < \frac{1}{\alpha+1}\} \\
&= 8 \cdot \widetilde{\mathsf{sdisc}}_{\frac{1}{\alpha+1}}(f) \enspace .
\end{aligned}$$

Similarly,

$$
\begin{aligned}
\gamma_2^\alpha(A_f) &= \max_{g,\lambda} \frac{1}{\gamma_2^*(A_g \circ P_\lambda)} \left(1-(\alpha+1)\lambda(f\neq g)\right) \\
&\geq \max_{g,\lambda} \mathsf{disc}^\lambda(g)\left(1-(\alpha+1)\lambda(f\neq g)\right) \\
&\geq \max\{\frac{1}{2}\cdot \mathsf{disc}^\lambda(g) : g,\lambda \text{ such that } \lambda(f\neq g) < \frac{1}{2(\alpha+1)}\} \\
&= \frac{1}{2}\cdot \widetilde{\mathsf{sdisc}}_{\frac{1}{2(\alpha+1)}}(f) \ .
\end{aligned}
$$

□

**Proof of Theorem 2:**

1. The lower bound is from [Raz92], the upper bound follows from [AA05].
2. The function is described in [NW95].
3. The lower bound $\mathsf{D}(\mathsf{LNE}) = n^2$ is shown in [KN97] where it was shown that $\log \mathsf{rank}(\mathsf{LNE}) = n^2$. It is not hard to see that the Las-Vegas complexity of $\mathsf{LNE}$ is $O(n)$ which is also shown in [KN97].

   In order to show $\log \mathsf{prt}_0(\mathsf{LNE}) = O(n)$, we describe a solution to the primal program for the partition bound for $\mathsf{LNE}$. We will assign a positive weight $w_R$, to every monochromatic rectangle $R$ such that the sum of weights is small. In this case one can set $w_{z,R} \stackrel{\mathsf{def}}{=} w_R$ where $z$ is the color of the monochromatic rectangle $R$ (all other $w_{z',R}$ are 0).

   We present the analysis below assuming that none of $x_1 \dots x_n, y_1 \dots y_n$ is $0^n$. The analysis can be extended easily if such is the case.

   First we consider the 1-inputs of $\mathsf{LNE}$. Let $R_{z_1,\dots,z_n,s_1,\dots,s_n}$ be the rectangle that contains all inputs with $\sum_j x_i(j)\cdot z_i(j) = s_i \mod 2$ and $\sum_j y_i(j)\cdot z_i(j) \neq s_i \mod 2$ for all $i$. Note that these are 1-chromatic rectangles. We give weight $2^n/2^{n^2}$ to each such rectangle. For every 1-input $x_1,\dots,x_n; y_1,\dots,y_n$ and all $s_1,\dots,s_n$

$$
\Pr_{z_1\dots z_n}\left(\sum_j x_i(j)\cdot z_i(j) = s_i \mod 2 \wedge \sum_j y_i(j)\cdot z_i(j) \neq s_i \mod 2 \text{ for all } i\right) = 1/4^n
$$

   for uniform $z_1,\dots,z_n$. Hence

$$
\sum w_{R_{z_1,\dots,z_n,s_1,\dots,s_n}} = 2^n \cdot \frac{2^{n^2}}{4^n}\cdot\frac{2^n}{2^{n^2}} = 1,
$$

   when the sum is over all $R_{z_1,\dots,z_n,s_1,\dots,s_n}$ consistent with $x_1,\dots,x_n; y_1,\dots,y_n$ . The sum of the weights $w_{R_{z_1,\dots,z_n,s_1,\dots,s_n}}$ of all such rectangles is exactly $2^{2n}$.

   Now we turn to the 0-inputs. For each of them there is a position $k+1$, where $x_{k+1} = y_{k+1}$ but $x_i \neq y_i$ for all $i \leq k$. Let $R_{z_1,\dots,z_k,s_1,\dots,s_k,u}$ denote the rectangle that contains all inputs with $\sum_j x_i(j)\cdot z_i(j) = s_i \mod 2$ and $\sum_j y_i(j)\cdot z_i(j) \neq s_i \mod 2$ for all $i \leq k$ and $x_{k+1} = y_{k+1} = u$.

The rectangle $R_{z_1,\ldots,z_k,s_1,\ldots,s_k,u}$ receives weight $2^k/2^{nk}$. As before it can be argued that every 0-input lies in $2^{nk}/2^k$ such rectangles, so the constraints are satisfied. The overall sum of rectangle weights is at most

$$\sum_{k=0}^{n-1} 2^{kn} \cdot 2^k \cdot 2^n \cdot \frac{2^k}{2^{kn}} \leq 2 \cdot 2^{3n}.$$

Hence $\log \mathsf{prt}_0(\mathsf{LNE}) \leq \log \sum_{R \in \mathcal{R}_{\mathsf{LNE}}} w_R = O(n)$.

□

**Proof of Theorem 3:**

1. Let $\{w_{z,A}\}$ be an optimal solution to the primal of $\mathsf{prt}_\epsilon(f)$. Let $\mathcal{P}$ be a randomized algorithm which achieves $\mathsf{R}_\epsilon(f)$. Then $\mathcal{P}$ is a convex combination of deterministic algorithms where each deterministic algorithm is a decision tree of depth at most $\mathsf{R}_\epsilon(f)$. As in the proof of Part 1. of Theorem 1, we can argue that $\sum_z \sum_A w_{z,A} \leq 2^{\mathsf{R}_\epsilon(f)}$. Now since for each $A$ above $|A| \leq \mathsf{R}_\epsilon(f)$,

$$\mathsf{prt}_\epsilon(f) = \sum_z \sum_A w_{z,A} 2^{|A|} \leq 2^{\mathsf{R}_\epsilon(f)} \left( \sum_{z \in \{0,1\}^m} \sum_{A \in \mathcal{A}} w_{z,A} \right) \leq 2^{2\mathsf{R}_\epsilon(f)} \quad .$$

   Hence our result.

2. Let $\{w_{z,A}\}$ be an optimal solution to the primal of $\mathsf{prt}_0(f)$. It is easily observed that $w_{z,A} > 0$ implies that $A$ is a $z$-certificate. Fix $x \in f^{-1}$, now

$$\begin{aligned}\mathsf{prt}_0(f) = \sum_z \sum_A w_{z,A} \cdot 2^{|A|} \geq \sum_{A : x \in A} w_{f(x),A} \cdot 2^{|A|} &\geq 2^{\mathsf{C}_x(f)} \cdot \left( \sum_{A : x \in A} w_{f(x),A} \right) \\ &= 2^{\mathsf{C}_x(f)} \quad .\end{aligned}$$

   Hence $\log \mathsf{prt}_0(f) \geq \max_{x \in f^{-1}} \{\mathsf{C}_x(f)\} = \mathsf{C}(f)$.

3. Fix $x \in f^{-1}$. Let $b \stackrel{\mathsf{def}}{=} \mathsf{bs}_x(f)$ and let $B_1, \ldots, B_b$ be the blocks for which $f(x) \neq f(x^{B_i})$. Let $\mu_x \stackrel{\mathsf{def}}{=} 2^{\epsilon b - 1}; \phi_x \stackrel{\mathsf{def}}{=} -(1-\epsilon)\mu_x$ and for each $i \in [b]$, let $-\phi_{x^{B_i}} = \mu_{x^{B_i}} \stackrel{\mathsf{def}}{=} \frac{2^{\epsilon b-1}}{b}$;. Let $\phi_y = \mu_y \stackrel{\mathsf{def}}{=} 0$ for $y \notin \{x, x^{B_1}, \ldots, x^{B_b}\}$.

   (a) Let $|A| \geq \epsilon b$. It is clear that $\forall z \in \{0,1\}^m : \sum_{x' \in f^{-1}(z) \cap A} \mu_{x'} + \sum_{x' \in A} \phi_{x'} \leq 2^{\epsilon b} \leq 2^{|A|}$.

   (b) Let $|A| < \epsilon b$. Let $z \neq f(x)$ or $x \notin A$. It is clear that $\sum_{x' \in f^{-1}(z) \cap A} \mu_{x'} + \sum_{x' \in A} \phi_{x'} \leq 0 \leq 2^{|A|}$.

   (c) Let $|A| < \epsilon b$ and $z = f(x)$ and $x \in A$. Since at most $\epsilon b$ blocks among $B_1, \ldots, B_b$ can have non-empty intersection with the subset $S \subseteq [n]$ corresponding to $A$, at least $(1-\epsilon)b$ among $\{x^{B_1}, \ldots, x^{B_b}\}$ belong to $A$; therefore (since $\epsilon < 0.5$)

$$\sum_{x' \in f^{-1}(z) \cap A} \mu_{x'} + \sum_{x' \in A} \phi_{x'} \leq \epsilon \cdot 2^{\epsilon b - 1} - (1-\epsilon) b \frac{2^{\epsilon b-1}}{b} < 0 \leq 2^{|A|}.$$

Therefore the constraints for $\mathsf{prt}_{\frac{\epsilon}{4}}(f)$ are satisfied. Now,

$$\mathsf{prt}_{\frac{\epsilon}{4}}(f) \geq \sum_x (1-\frac{\epsilon}{4})\mu_x + \phi_x = (1-\frac{\epsilon}{4})2^{\epsilon b} - (2-\epsilon)2^{\epsilon b-1} = \epsilon 2^{\epsilon b-2} \ .$$

Hence our result.

4. Let $\{w_{z,A}\}$ be an optimal solution to the primal of $\mathsf{prt}_\epsilon(f)$. Let $\alpha \stackrel{\mathsf{def}}{=} \sum_z \sum_A w_{z,A} \cdot 2^{|A|}$. Let $\mathcal{A}' \stackrel{\mathsf{def}}{=} \{A : |A| \leq \log \frac{\alpha}{\epsilon}\}$. Then $\sum_z \sum_{A \notin \mathcal{A}'} w_{z,A} \leq \epsilon$. Fix $x \in f^{-1}$. Let $\mathcal{A}'_x \stackrel{\mathsf{def}}{=} \{A \in \mathcal{A}' : x \in A\}$. We know that

$$\alpha_x \stackrel{\mathsf{def}}{=} \sum_{A \in \mathcal{A}'_x} w_{f(x),A} \geq \sum_{A : x \in A} w_{f(x),A} - \epsilon \geq 1 - 2\epsilon \ .$$

The verifier $V_x$ for $x$ acts as follows:

(a) Choose $A \in \mathcal{A}'_x$ with probability $\frac{w_{f(x),A}}{\alpha_x}$.
(b) Query locations in $A$.
(c) Accept iff locations queried are consistent with $A$. Reject otherwise.

Now it is clear that if the input is $x$ then $V_x$ accepts with probability 1. Also the number of queries of $V_x$ are at most $\log \frac{\alpha}{\epsilon}$ on any input $y$. Let $y$ be such that $f(y) \neq f(x)$. Let $\mathcal{A}'_{x,y} \stackrel{\mathsf{def}}{=} \{A \in \mathcal{A}'_x : y \in A\}$. Then,

$$\sum_{A \in \mathcal{A}'_{x,y}} w_{f(x),A} \leq \sum_{A \in \mathcal{A}' : y \in A} \ \sum_{z \neq f(y)} w_{z,A} \leq \sum_{A \in \mathcal{A} : y \in A} \ \sum_{z \neq f(y)} w_{z,A} + \epsilon \leq 2\epsilon \ .$$

Hence $y$ would be accepted with probability at most $\frac{2\epsilon}{\alpha_x} \leq \frac{2\epsilon}{1-2\epsilon}$. Hence our result.

5. Let $\{w_{z,A}\}$ be an optimal solution to the primal of $\mathsf{prt}_\epsilon(f)$. Let $\alpha \stackrel{\mathsf{def}}{=} \sum_z \sum_A w_{z,A} \cdot 2^{|A|}$ and $k \stackrel{\mathsf{def}}{=} \log \frac{\alpha}{\epsilon}$. Let $\mathcal{A}' \stackrel{\mathsf{def}}{=} \{A : |A| \leq k\}$; then $\sum_z \sum_{A \notin \mathcal{A}'} w_{z,A} \leq \epsilon$. We set $p$ as in the definition of $\mathsf{cadv}$ as follows. For all $x \in f^{-1}$, let $\mathcal{A}'_x \stackrel{\mathsf{def}}{=} \{A \in \mathcal{A}' : x \in A\}$. Define distributions $p_x$ on $[n]$ as follows:

(a) Choose $A \in \mathcal{A}'_x$ with probability $q(x, A) \stackrel{\mathsf{def}}{=} \frac{w_{f(x),A}}{\sum_{A' \in \mathcal{A}'_x} w_{f(x),A'}}$.
(b) Choose $i$ uniformly from the set $\{i : i \text{ appears in } A\}$.

It is easily seen that $p_x$ is a distribution on $[n]$. We will show that

$$\max_{x,y: f(x) \neq f(y)} \frac{1}{\sum_{i : x_i \neq y_i} \min\{p_x(i), p_y(i)\}} \leq \frac{k}{1-4\epsilon}, \tag{6}$$

which proves our main claim.

Take any $x, y$ such that $f(x) \neq f(y)$. Let's define $\forall i \in [n], q_x(i) \stackrel{\mathsf{def}}{=} \sum_{A \in \mathcal{A}'_x : i \text{ appears in } A} q(x, A)$; similarly define $q_y(i)$. It is clear that $\forall i \in [n] : p_x(i) \geq \frac{q_x(i)}{k}$ and $p_y(i) \geq \frac{q_y(i)}{k}$. We will show:

$$\sum_{i : x_i \neq y_i} \min\{q_x(i), q_y(i)\} \geq 1 - 4\epsilon,$$

which implies (6).

Now assume for a contradiction that $\sum_{i:x_i\neq y_i}\min\{q_x(i),q_y(i)\}<1-4\epsilon$. Consider a hybrid input $r\in\{0,1\}^n$ constructed in the following way: if $q_x(i)\geq q_y(i)$ then $r_i\stackrel{\mathsf{def}}{=}x_i$, otherwise $r_i\stackrel{\mathsf{def}}{=}y_i$. Now,

$$\begin{aligned}
\sum_{A:r\in A}\sum_z w_{z,A} &\geq \sum_{A\in\mathcal{A}'_r}\sum_z w_{z,A}\\
&\geq \sum_{A\in\mathcal{A}'_x} w_{f(x),A}-\sum_{i:q_x(i)<q_y(i)} q_x(i)+\sum_{A\in\mathcal{A}'_y} w_{f(y),A}-\sum_{i:q_y(i)\leq q_x(i)} q_y(i)\\
&\geq \sum_{A\in\mathcal{A}'_x} w_{f(x),A}+\sum_{A\in\mathcal{A}'_y} w_{f(y),A}-\sum_{i:x_i\neq y_i}\min\{q_x(i),q_y(i)\}\\
&\geq \sum_{A:x\in A} w_{f(x),A}+\sum_{A:y\in A} w_{f(y),A}-\sum_{i:x_i\neq y_i}\min\{q_x(i),q_y(i)\}-2\epsilon\\
&\geq 2(1-\epsilon)-(1-4\epsilon)-2\epsilon \quad > \quad 1\ .
\end{aligned}$$

This contradicts the assumption that $\{w_{z,A}\}$ is a feasible solution to the primal of $\mathsf{prt}_\epsilon(f)$.

6. $\log\mathsf{prt}_{2\epsilon}(f)\geq\log\mathsf{sdisc}_{2\epsilon}(f)$ follows using similar arguments as before and hence proof skipped.

   We turn to the second part. Let $\{w_A,v_A\}$ be an optimal solution to the primal of $\mathsf{sdisc}_{2\epsilon}(f)$. Let $\alpha\stackrel{\mathsf{def}}{=}\sum_A(w_A+v_A)\cdot 2^{|A|}$. Let $\mathcal{A}'\stackrel{\mathsf{def}}{=}\{A:|A|\leq\log\frac{\alpha}{\epsilon}\}$; then $\sum_{A\notin\mathcal{A}'}w_A+v_A\leq\epsilon$. For $A\in\mathcal{A}'$, let $m_A(x)$ be the multilinear polynomial which is 1 iff $x\in A$ (over the Boolean inputs $x$). Note that the degree of $m_A$ is at most $|A|$. Let $p(x)\stackrel{\mathsf{def}}{=}\sum_{A\in\mathcal{A}'}(w_A-v_A)\cdot m_A(x)$. Then the degree of $p(x)$ is at most $\log\frac{\alpha}{\epsilon}$. Now since the constraints of the primal of $\mathsf{sdisc}_{2\epsilon}(f)$ are satisfied by $\{w_A,v_A\}$, we get,

$$\forall x\in f^{-1}(1): 1+\epsilon\geq p(x)=\sum_{A\in\mathcal{A}':x\in A} w_A-v_A\geq 1-3\epsilon,$$

   and

$$\forall x\in f^{-1}(0): -1-\epsilon\leq p(x)=\sum_{A\in\mathcal{A}':x\in A} w_A-v_A\leq -1+3\epsilon,$$

   and

$$\forall x: -1-\epsilon\leq p(x)\leq 1+\epsilon\ .$$

   Therefore $(p(x)/(1+\epsilon)+1)/2$, $2\epsilon$-approximates $f$ and hence our result.

7. Let $p(x)$ be a polynomial that $\epsilon$-approximates $f$ and has degree $k=\widetilde{\mathsf{deg}}_\epsilon(f)$. Then $q(x)=2p(x)-1$ has the same degree. Write $q(x)$ in the standard form $q(x)=\sum_{S\subseteq[n]}c_S\cdot m_S$, where $m_S$ is the monomial $\prod_{i\in S}x_i$ and the $c_S$ are the coefficients. Now we define the solutions for the program for $\mathsf{sdisc}_{2\epsilon}$. We identify $S\subseteq[n]$ with the partial assignment $A(S)$ that sets the variables in $S$ to 1, and will set $w_A=v_A=0$ for all other partial assignments $A$ associated with $S$. Otherwise if $c_S\leq 0$, we set $w_{A(S)}=0$ and $v_{A(S)}=-c_S$, if $c_S\geq 0$, we set $v_{A(S)}=0$ and $w_{A(S)}=c_S$. It is now easy to see that this yields a feasible solution to the primal program of $\mathsf{sdisc}_\epsilon(f)$.

It remains to show that the cost of the solution is no more than exponential in $O(k \log n)$. Note that $q(x)$ contains at most $\sum_{i \leq k} \binom{n}{i}$ monomials of degree at most $k$, so it suffices to show that every $|c_S|$ is bounded. We will show that any $|c_S|$ in $q(x)$ is indeed bounded by $(2k+1)!$.

Consider $S \subseteq [n]$. Denote $t_S = \sum_{S': S' \subseteq S} |c_{S'}|$ and $t_j = \max_{S:|S|=j} t_S$. Then $t_0 \leq 1$, because $c_\emptyset$ is the constant coefficient of $q$ and equals $q(0^n)$. Similarly we can see that $t_j \leq 2j \cdot t_{j-1} + 1$: For every $S$ of size $j$ we can write

$$t_S \leq \sum_{S' \subset S, |S'|=j-1} t_{S'} + |c_S|,$$

and

$$|c_S| \leq \sum_{S' \subset S, |S'|=j-1} t_{S'} + 1,$$

because $q(e_S) = \sum_{S' \subseteq S} c_{S'} \in [-1, 1]$ for the string $e_S$ containing 1's in $S$.

By induction this proves that $|c_S| \leq (2j+1)!$, and hence the cost of our solution is bounded by $\sum_{j \leq k} \binom{n}{j} \cdot (2j+1)! \cdot 2^j \leq \exp(O(k \log n))$.

8. For a Boolean function $f$, it is known that $\mathsf{D}(f) = O(\mathsf{C}(f)\mathsf{bs}(f))$ and $\mathsf{D}(f) = O(\mathsf{bs}(f)^3)$ (refer to [BW02]). The desired result is implied now using earlier parts of this theorem.

□

**Proof of Theorem 4:** We denote the $\mathsf{Tribes}$ function by $f$. We exhibit a solution to the dual of the linear program for $\mathsf{prt}_\epsilon(f)$. In fact we use a one-sided relaxation of the LP for $\mathsf{prt}_\epsilon(f)$, similar to the smooth rectangle bound. It is easily observed that the optimum of the LP below, denoted $\mathsf{opt}_\epsilon(f)$ is at most $\mathsf{prt}_\epsilon(f)$.

Primal

$$\begin{aligned}
\text{min:} \quad & \sum_A w_A \cdot 2^{|A|} \\
& \forall x \text{ with } f(x)=1: \sum_{A: x \in A} w_A \geq 1-\epsilon, \\
& \forall x \text{ with } f(x)=1: \sum_{A: x \in A} w_A \leq 1, \\
& \forall x \text{ with } f(x)=0: \sum_{A: x \in A} w_A \leq \epsilon, \\
& \forall A: w_A \geq 0 \ .
\end{aligned}$$

Dual

$$\begin{aligned}
\text{max:} \quad & \sum_{x: f(x)=1} (1-\epsilon)\mu_x - \sum_{x: f(x)=0} \epsilon \mu_x + \sum_x \phi_x \\
& \forall A: \sum_{x \in f^{-1}(1) \cap A} \mu_x - \sum_{x \in f^{-1}(0) \cap A} \mu_x + \sum_{x \in A} \phi_x \leq 2^{|A|}, \\
& \forall x: \mu_x \geq 0, \phi_x \leq 0 \ .
\end{aligned}$$

We will work with the dual program and will assign nonzero values for $(\mu_x, \phi_x)$ on three types of inputs. Denote the set $\{(i,j) : j = 1, \ldots, \sqrt{n}\}$ by $B_i$. This is a block of inputs that feeds into a single OR. The first set of inputs has exactly one $x_{i,j} = 1$ per block $B_i$. Clearly these are inputs with $f(x) = 1$, and there are exactly $\sqrt{n}^{\sqrt{n}}$ such inputs. Denote the set of these inputs by $T_1$. Then we consider a set of inputs with $f(x, y) = 0$. Denote by $T_0$ the set of inputs in which all but one block $B_i$ have exactly one 1, and one block $B_i$ has no $x_{i,j} = 1$. Again, there are $\sqrt{n}^{\sqrt{n}}$ such inputs. Finally, $T_2$ contains the set of inputs, in which all $B_i$ except one have exactly one 1, and one block has two 1's. There are $(\sqrt{n})^{\sqrt{n}}(n-\sqrt{n})/2$ such inputs. Let $\delta \stackrel{\mathsf{def}}{=} \frac{1}{4} - 4\epsilon$ and,

$$\begin{aligned}
\text{For all } x \in T_1 \quad &: \quad \mu_x = \frac{2^{\delta n}}{\sqrt{n}^{\sqrt{n}}}; \quad \phi_x = 0, \\
\text{For all } x \in T_0 \quad &: \quad \mu_x = \frac{2^{\delta n}}{4\epsilon \cdot \sqrt{n}^{\sqrt{n}}}; \quad \phi_x = 0, \\
\text{For all } x \in T_2 \quad &: \quad \phi_x = \frac{-4 \cdot 2^{\delta n}}{3(n-\sqrt{n})\sqrt{n}^{\sqrt{n}}}; \quad \mu_x = 0, \\
\text{For all } x \notin T_0 \cup T_1 \cup T_2 \quad &: \quad \mu_x = \phi_x = 0 \ .
\end{aligned}$$

**Claim 1** $\{\mu_x, \phi_x\}$ *as defined is feasible for the dual for* $\mathsf{opt}_\epsilon(f)$.

**Proof** Clearly $\forall x : \mu_x \geq 0, \phi_x \leq 0$. Let $A$ be an assignment with $|A| \geq \delta n$; in this case,

$$\sum_{x \in f^{-1}(1) \cap A} \mu_x - \sum_{x \in f^{-1}(0) \cap A} \mu_x + \sum_{x \in A} \phi_x \leq \sum_{x \in f^{-1}(1)} \mu_x \leq 2^{\delta n} \leq 2^{|A|}.$$

From now on $|A| < \delta n$. Let $A$ fix at least two input positions to 1 in a single block $B_i$. In this case clearly,

$$\sum_{x \in f^{-1}(1) \cap A} \mu_x - \sum_{x \in f^{-1}(0) \cap A} \mu_x + \sum_{x \in A} \phi_x \leq 0 \leq 2^{|A|}.$$

Hence from now on consider $A$ which fixes at most a single input position to 1 in each block $B_i$. For block $i$ let $\alpha_i$ denote the number of positions fixed to 0 in $B_i$; let $\beta_i \in \{0, 1\}$ denote the number of positions fixed to 1 and let $\gamma_i$ denote the number of free positions, i.e., $\sqrt{n} - \alpha_i - \beta_i$.

First consider the case when $k \stackrel{\mathsf{def}}{=} \sum_i \beta_i \leq (1-4\epsilon)\sqrt{n}$ and w.l.o.g. assume that the last $k$ blocks contain a 1. The number of inputs in $T_1$ consistent with $A$ is exactly $\prod_{i=1}^{\sqrt{n}-k} \gamma_i$. The number of inputs in $T_0$ consistent with $A$ is

$$\sum_{i=1}^{\sqrt{n}-k} \prod_{j=1,\ldots,\sqrt{n}-k; j \neq i} \gamma_j \geq \frac{\sqrt{n}-k}{\sqrt{n}} \cdot \prod_{i=1}^{\sqrt{n}-k} \gamma_i \geq 4\epsilon \prod_{i=1}^{\sqrt{n}-k} \gamma_i.$$

Hence,

$$\sum_{x \in f^{-1}(1) \cap A} \mu_x - \sum_{x \in f^{-1}(0) \cap A} \mu_x + \sum_{x \in A} \phi_x \leq \frac{2^{\delta n}}{\sqrt{n}^{\sqrt{n}}}\left(1 - \frac{4\epsilon}{4\epsilon}\right) \prod_{i=1}^{\sqrt{n}-k} \gamma_i \leq 0.$$

Now assume that $k = \sum_i \beta_i \geq (1-4\epsilon)\sqrt{n}$. Again w.l.o.g. the last $k$ blocks have $\beta_i = 1$. There are $\prod_{i=1}^{\sqrt{n}-k} \gamma_i$ inputs in $T_1 \cap A$. The number of inputs in $T_2 \cap A$ is at least

$$\left(\prod_{i=1}^{\sqrt{n}-k} \gamma_i\right) \cdot \left(\sum_{i=\sqrt{n}-k+1}^{\sqrt{n}} \gamma_i\right) \geq \left(\prod_{i=1}^{\sqrt{n}-k} \gamma_i\right) \cdot n(1-\delta-4\epsilon),$$

because we can choose a single 1 for the first $\sqrt{n}-k$ blocks, and a second 1 in any of the last $k$ blocks. Hence

$$\begin{aligned}
\sum_{x\in A\cap T_2}\phi_x &\leq -\left(\prod_{i=1}^{\sqrt{n}-k}\gamma_i\right)\cdot n(1-\delta-4\epsilon)\cdot\frac{4\cdot 2^{\delta n}}{3(n-\sqrt{n})\sqrt{n}^{\sqrt{n}}} \\
&= -\left(\sum_{x\in A\cap T_1}\mu_x\right)\cdot n(1-\delta-4\epsilon)\cdot\frac{4}{3(n-\sqrt{n})} \\
&\leq -\left(\sum_{x\in A\cap T_1}\mu_x\right)\cdot(1-\delta-4\epsilon)\cdot\frac{4}{3} = -\left(\sum_{x\in A\cap T_1}\mu_x\right) \quad .
\end{aligned}$$

Hence the constraints for dual of $\mathsf{opt}_\epsilon(f)$ are satisfied by all $A$. $\square$

Finally we have,

$$\begin{aligned}
\mathsf{prt}_\epsilon(f) \geq \mathsf{opt}_\epsilon(f) &\geq \sum_{x:f(x)=1}(1-\epsilon)\mu_x - \sum_{x:f(x)=0}\epsilon\mu_x + \sum_x \phi_x \\
&= 2^{\delta n}\left(1-\epsilon-\frac{\epsilon}{4\epsilon}-\frac{2}{3}\right) = 2^{\Omega(n)} \quad .
\end{aligned}$$

Hence our result.

The upper bound on $\mathsf{C}(\mathsf{Tribes})$ is obvious, and implies the bound on $\mathsf{cadv}$. The remaining bounds follow from the existence of efficient quantum query algorithms for the problem. $\square$

**Proof of Theorem 5:** Let us define the weights for the dual of $\mathsf{prt}_0(f^h)$ in a recursive fashion. We first define the weights for the inputs for $f^1$ as follows $\mu_{00} \overset{\mathsf{def}}{=} -1, \mu_{01} = \mu_{10} \overset{\mathsf{def}}{=} 1.6, \mu_{11} \overset{\mathsf{def}}{=} 1.2$. Let $x_1, x_2$ be two inputs to $f^{h-1}$, then define the weights for $f^h$ as follows ($x_1x_2$ represents concatenation):

$$\mu_{x_1x_2} \overset{\mathsf{def}}{=} \mu_{x_1}\cdot\mu_{x_2} \text{ iff } f^{h-1}(x_1)=1 \text{ or } f^{h-1}(x_2)=1; \text{ otherwise } \mu_{x_1x_2} \overset{\mathsf{def}}{=} -\mu_{x_1}\cdot\mu_{x_2} \quad .$$

We have the following lemma:

**Lemma 7** *Let $\mathcal{A}_{f^h}$ represent the set of all monochromatic assignments of $f^h$. For all $h$, we have the following invariants.*

1. $\forall A \in \mathcal{A}_{f^h} : |\sum_{x\in A}\mu_x| \leq 2^{|A|}$.
2. $\forall A : |\sum_{x\in A, f^h(x)=1}\mu_x - \sum_{x\in A, f^h(x)=0}\mu_x| \leq 2^{|A|}$.

**Proof** We prove the invariants using induction. For the base case ($h=1$) they can be checked by direct calculation. Assume that they are true for $h-1$ and we need to show for $h$.

1. Let $A = A_1A_2$ be a 0-monochromatic assignment of $f^h$, then $A_1, A_2$ need to be each a 1-monochromatic assignment of $f^{h-1}$. Hence by induction,

$$|\sum_{x\in A}\mu_x| = |(\sum_{x_1\in A_1}\mu_{x_1})\cdot(\sum_{x_2\in A_2}\mu_{x_2})| = |\sum_{x_1\in A_1}\mu_{x_1}|\cdot|\sum_{x_2\in A_2}\mu_{x_2}| \leq 2^{|A_1|}2^{|A_2|} = 2^{|A|} \quad .$$

2. Let $A = A_1A_2$ be a 1-monochromatic assignment of $f^h$, then either $A_1$ or $A_2$ needs to be a 0-monochromatic assignment of $f^{h-1}$. Let w.l.o.g $A_1$ be a 0-monochromatic assignment. Then we have by induction,

$$\begin{aligned}
|\sum_{x\in A}\mu_x| &= |\sum_{x_1\in A_1, x_2\in A_2}\mu_{x_1x_2}| \\
&= |(\sum_{x_1\in A_1}\mu_{x_1})\cdot(\sum_{x_2\in A_2, f^{h-1}(x_2)=1}\mu_{x_2} - \sum_{x_2\in A_2, f^{h-1}(x_2)=0}\mu_{x_2})| \\
&= |(\sum_{x_1\in A_1}\mu_{x_1})|\cdot|(\sum_{x_2\in A_2, f^{h-1}(x_2)=1}\mu_{x_2} - \sum_{x_2\in A_2, f^{h-1}(x_2)=0}\mu_{x_2})| \\
&\le 2^{|A_1|}\cdot 2^{|A_2|} = 2^{|A|} \quad .
\end{aligned}$$

3. Let $A = A_1, A_2$ be some assignment of $f^h$. Then by induction we have,

$$\begin{aligned}
&|\sum_{x\in A, f^h(x)=1}\mu_x - \sum_{x\in A, f^h(x)=0}\mu_x| \\
&= |-\sum_{x_1\in A_1, f^{h-1}(x_1)=1, x_2\in A_2, f^{h-1}(x_2)=1}\mu_{x_1}\mu_{x_2} - \sum_{x_1\in A_1, f^{h-1}(x_1)=0, x_2\in A_2, f^{h-1}(x_2)=0}\mu_{x_1}\mu_{x_2} \\
&\quad + \sum_{x_1\in A_1, f^{h-1}(x_1)=1, x_2\in A_2, f^{h-1}(x_2)=0}\mu_{x_1}\mu_{x_2} + \sum_{x_1\in A_1, f^{h-1}(x_1)=0, x_2\in A_2, f^{h-1}(x_2)=1}\mu_{x_1}\mu_{x_2}| \\
&= |(\sum_{x_1\in A_1, f^{h-1}(x_1)=1}\mu_{x_1} - \sum_{x_1\in A_1, f^{h-1}(x_1)=0}\mu_{x_1})(\sum_{x_2\in A_2, f^{h-1}(x_2)=1}\mu_{x_2} - \sum_{x_2\in A_2, f^{h-1}(x_2)=0}\mu_{x_2})| \\
&\le 2^{|A_1|}2^{|A_2|} = 2^{|A|} \quad .
\end{aligned}$$

□

Therefore $\{\mu_x : x \text{ input of } f^h\}$ satisfy the constraints for the dual of $\mathsf{prt}_0(f^h)$. Now define

$$\alpha_1^h \stackrel{\mathsf{def}}{=} \sum_{x, f^h(x)=1}\mu_x \quad ; \quad \alpha_0^h \stackrel{\mathsf{def}}{=} \sum_{x, f^h(x)=0}\mu_x \quad .$$

Then we see

$$\alpha_0^h = (\alpha_1^{h-1})^2 \quad ; \quad \alpha_1^h = 2\alpha_1^{h-1}\alpha_0^{h-1} - (\alpha_0^{h-1})^2 \quad .$$

This implies $\alpha_0^h - \alpha_1^h = (\alpha_0^{h-1} - \alpha_1^{h-1})^2$. Now since $\alpha_1^1 = 2.2, \alpha_0^1 = 1.2$ we have that $\alpha_0^h - \alpha_1^h = 1$ for all $h \ge 2$. Hence for $h \ge 2$,

$$\alpha_1^h = \alpha_0^h - 1 = (\alpha_1^{h-1})^2 - 1 \ge \frac{(\alpha_1^{h-1})^2}{2} \quad .$$

Therefore $\alpha_1^h \ge \frac{(\alpha_1^1)^{2^{h-1}}}{2^h}$. Since $\alpha_1^1 = 2.2$, we conclude $\log \mathsf{prt}_0(f^h) = \Omega(2^h)$. □

**Proof of Lemma 6:**(Sketch) Examples of such functions are given in [NS94, NW95] with the best construction attributed to Kushilevitz in the latter paper. □

## B Las-Vegas Partition Bound

### Communication Complexity

In this section we consider the Las-Vegas communication complexity. Las-Vegas protocols use randomness and for each input they are allowed to output "don't know" with probability 1/2, however when they do give an answer then it is required to be correct. An equivalent way to view is that these protocols are never allowed to err, but for each input we only count the expected communication (over the coins), instead of the worst case communication (as in deterministic protocols). Below we present a lower bound for Las-Vegas protocols via a linear program, which we call the Las-Vegas partition bound.

**Definition 21 (Las-Vegas Partition Bound)** *Let $f : \mathcal{X} \times \mathcal{Y} \to \mathcal{Z}$ be a partial function. The Las-Vegas partition bound of $f$, denoted $\mathsf{prt}_{LV}(f)$, is given by the optimal value of the following linear program. Let $\mathcal{R}_f$ denote the set of monochromatic rectangles for $f$.*

<u>*Primal*</u>

$$\begin{aligned}
\textit{min:} \quad & \sum_{R \in \mathcal{R}_f} w_R + \sum_{R \in \mathcal{R}} v_R \\
& \forall (x,y) \in f^{-1} : \sum_{R \in \mathcal{R}_f : (x,y) \in R} w_R \geq \frac{1}{2}, \\
& \forall (x,y) : \sum_{R \in \mathcal{R}_f : (x,y) \in R} w_R + \sum_{R : (x,y) \in R} v_R = 1, \\
& \forall R : w_R, v_R \geq 0 \ .
\end{aligned}$$

<u>*Dual*</u>

$$\begin{aligned}
\textit{max:} \quad & \sum_{(x,y) \in f^{-1}} \frac{1}{2} \cdot \mu_{x,y} + \sum_{(x,y)} \phi_{x,y} \\
& \forall R \in \mathcal{R}_f : \sum_{(x,y) \in f^{-1} \cap R} \mu_{x,y} + \sum_{(x,y) \in R} \phi_{x,y} \leq 1, \\
& \forall R \in \mathcal{R} : \sum_{(x,y) \in R} \phi_{x,y} \leq 1, \\
& \forall (x,y) : \mu_{x,y} \geq 0, \phi_{x,y} \in \mathbb{R} \ .
\end{aligned}$$

The constant above is arbitrary and can be any constant in $(0,1)$ and will give asymptotically similar value for the bound. The following lemma follows easily using arguments as before. Below $\mathsf{R}_0(f)$ represents the Las-Vegas communication complexity of $f$; please refer to [KN97] for explicit definition of $\mathsf{R}_0(f)$.

**Lemma 8** *Let $f : \mathcal{X} \times \mathcal{Y} \to \mathcal{Z}$ be a partial function. Then, $\mathsf{R}_0(f) \geq \log \mathsf{prt}_{LV}$ .*

### Query Complexity

The Las-Vegas partition bound for query complexity is defined as follows.

**Definition 22 (Las-Vegas Partition Bound)** *Let $f : \mathcal{X} \to \mathcal{Z}$ be a partial function. The Las-Vegas partition bound of $f$, denoted $\mathsf{prt}_{LV}(f)$, is given by the optimal value of the following linear program. Let $\mathcal{A}_f$ denote the set of monochromatic assignments for $f$.*

<u>*Primal*</u>

$$\begin{aligned}
\textit{min:} \quad & \sum_{A \in \mathcal{A}_f} w_A \cdot 2^{|A|} + \sum_{A \in \mathcal{A}} v_A \cdot 2^{|A|} \\
& \forall x \in f^{-1} : \sum_{A \in \mathcal{A}_f : x \in A} w_A \geq \frac{1}{2}, \\
& \forall x : \sum_{A \in \mathcal{A}_f : x \in A} w_A + \sum_{A : x \in A} v_A = 1, \\
& \forall A : w_A, v_A \geq 0 \ .
\end{aligned}$$

<u>*Dual*</u>

$$\begin{aligned}
\textit{max:} \quad & \sum_{x \in f^{-1}} \frac{1}{2} \cdot \mu_x + \sum_{x} \phi_x \\
& \forall A \in \mathcal{A}_f : \sum_{x \in f^{-1} \cap A} \mu_x + \sum_{x \in A} \phi_x \leq 2^{|A|}, \\
& \forall A \in \mathcal{A} : \sum_{x \in A} \phi_x \leq 2^{|A|}, \\
& \forall x : \mu_x \geq 0, \phi_x \in \mathbb{R} \ .
\end{aligned}$$

As before the constant above is arbitrary and can be any constant in $(0,1)$ and will give asymptotically similar value for the bound. The following lemma follows easily using arguments as before. Below $\mathsf{R}_0(f)$ represents the Las-Vegas query complexity of $f$.

**Lemma 9** *Let $f : \mathcal{X} \to \mathcal{Z}$ be a partial function. Then* $\mathsf{R}_0(f) \geq \log \mathsf{prt}_{LV}$ .

**Remark:** For communication complexity let $\mathsf{prt}^*_{LV}(f)$ be defined similarly to $\mathsf{prt}_{LV}(f)$, except that the constraints

$$\forall (x,y) \in f^{-1} : \sum_{R \in \mathcal{R}_f : (x,y) \in R} w_R \geq 1/2$$

are replaced by

$$\forall (x,y) \in f^{-1} : \sum_{R \in \mathcal{R}_f : (x,y) \in R} w_R = 1/2 \ .$$

Then we can observe $\mathsf{prt}_0(f) \geq \mathsf{prt}^*_{LV}(f) \geq \frac{1}{2}\mathsf{prt}_0(f)$. Note that $\log \mathsf{prt}^*_{LV}(f)$ forms a lower bound for $\mathsf{R}_0(f)$ if there is a Las-Vegas protocol for $f$ that has the probability of output 'don't know' for all inputs. Similarly for query complexity.

# C Partition bound for relations

## Communication Complexity

Here we define the partition bound for relations.

**Definition 23 (Partition Bound for relation)** *Let $f \subseteq \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$ be a relation. The $\epsilon$-partition bound of $f$, denoted $\mathsf{prt}_\epsilon(f)$, is given by the optimal value of the following linear program.*

<u>*Primal*</u>

$$\begin{aligned}
\textit{min:} \quad & \sum_z \sum_R w_{z,R} \\
& \forall (x,y) : \sum_{R:(x,y)\in R} \ \sum_{z:(x,y,z)\in f} w_{z,R} \geq 1-\epsilon, \\
& \forall (x,y) : \sum_{R:(x,y)\in R} \ \sum_{z} w_{z,R} = 1, \\
& \forall z, \forall R : w_{z,R} \geq 0 \ .
\end{aligned}$$

<u>*Dual*</u>

$$\begin{aligned}
\textit{max:} \quad & \sum_{(x,y)} (1-\epsilon)\mu_{x,y} + \phi_{x,y} \\
& \forall z, \forall R : \sum_{(x,y):(x,y)\in R; (x,y,z)\in f} \mu_{x,y} + \sum_{(x,y)\in R} \phi_{x,y} \leq 1, \\
& \forall (x,y) : \mu_{x,y} \geq 0, \phi_{x,y} \in \mathbb{R} \ .
\end{aligned}$$

As in Theorem 1, we can show that partition bound is a lower bound on the communication complexity. Its proof is skipped since it is very similar.

**Lemma 10** *Let $f \subseteq \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$ be a relation. Then,* $\mathsf{R}^{\mathsf{pub}}_\epsilon(f) \geq \log \mathsf{prt}_\epsilon(f)$ .

## Query Complexity

Here we define the partition bound for query complexity for relations.

**Definition 24 (Partition Bound for relations)** *Let $f \subseteq \mathcal{X} \times \mathcal{Z}$ be a relation, let $\epsilon \geq 0$. The $\epsilon$-partition bound of $f$, denoted $\mathsf{prt}_\epsilon(f)$, is given by the optimal value of the following linear program.*

*Primal*

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

*Dual*

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

As in Theorem 3, we can show that partition bound is a lower bound on the randomized query complexity of $f$. Its proof is skipped since it is very similar.

**Theorem 6** *Let $f \subseteq \mathcal{X} \times \mathcal{Z}$ be a relation, let $\epsilon > 0$. Then,* $\mathsf{R}_\epsilon(f) \geq \frac{1}{2}\log \mathsf{prt}_\epsilon(f)$ .