Partition Bound Is Quadratically Tight for Product Distributions*

Prahladh Harsha†[1], Rahul Jain‡[2], and Jaikumar Radhakrishnan[3]

1 **Tata Institute of Fundamental Research, Mumbai, India**
`prahladh@tifr.res.in`
2 **Center for Quantum Technologies, MajuLab and NUS, Singapore**
`rahul@comp.nus.edu.sg`
3 **Tata Institute of Fundamental Research, Mumbai, India**
`jaikumar@tifr.res.in`

Abstract

Let $f : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$ be a 2-party function. For every product distribution μ on $\{0,1\}^n \times \{0,1\}^n$, we show that

$$\mathsf{CC}^{\mu}_{0.49}(f) = O\left(\left(\log \mathsf{prt}_{1/8}(f) \cdot \log\log \mathsf{prt}_{1/8}(f)\right)^2\right),$$

where $\mathsf{CC}^{\mu}_{\varepsilon}(f)$ is the distributional communication complexity of f with error at most ε under the distribution μ and $\mathsf{prt}_{1/8}(f)$ is the *partition bound* of f, as defined by Jain and Klauck [*Proc. 25th CCC*, 2010]. We also prove a similar bound in terms of $\mathsf{IC}_{1/8}(f)$, the *information complexity* of f, namely,

$$\mathsf{CC}^{\mu}_{0.49}(f) = O\left(\left(\mathsf{IC}_{1/8}(f) \cdot \log \mathsf{IC}_{1/8}(f)\right)^2\right).$$

The latter bound was recently and independently established by Kol [*Proc. 48th STOC*, 2016] using a different technique.

We show a similar result for query complexity under product distributions. Let $g : \{0,1\}^n \to \{0,1\}$ be a function. For every bit-wise product distribution μ on $\{0,1\}^n$, we show that

$$\mathsf{QC}^{\mu}_{0.49}(g) = O\left(\left(\log \mathsf{qprt}_{1/8}(g) \cdot \log\log \mathsf{qprt}_{1/8}(g)\right)^2\right),$$

where $\mathsf{QC}^{\mu}_{\varepsilon}(g)$ is the distributional query complexity of f with error at most ε under the distribution μ and $\mathsf{qprt}_{1/8}(g))$ is the *query partition bound* of the function g.

Partition bounds were introduced (in both communication complexity and query complexity models) to provide LP-based lower bounds for randomized communication complexity and randomized query complexity. Our results demonstrate that these lower bounds are polynomially tight for *product* distributions.

1998 ACM Subject Classification F.1.3 Complexity Measures and Classes, G.2.1 Combinatorics

Keywords and phrases partition bound, product distribution, communication complexity, query complexity

Digital Object Identifier 10.4230/LIPIcs.ICALP.2016.135

* A full version of the paper is available at `http://arxiv.org/abs/1512.01968` [8]. This work was done when the three authors were visiting the Simons Institute for the Theory of Computing, Berkeley, USA.

† Prahladh Harsha's research supported in part by ISF-UGC grant 1399/4 and Google India Fellowship.
‡ Rahul Jain's research partly supported by the Singapore Ministry of Education via Academic Research Fund Tier 3 MOE2012-T3-1-009 and Young Researcher Award, National University of Singapore.

1 Introduction

Over the last decade, several lower bound techniques using linear programming formulations and information complexity methods have been developed for problems in communication complexity and query complexity. One of the central questions in communication complexity is to understand the tightness of these lower bound techniques. For instance, over the last few years, considerable effort has gone into understanding the *information complexity* measure. Informally speaking, (internal) information complexity is the amount of information the two parties reveal to each other about their respective inputs while computing the joint function. It is known that for product distributions, the internal information complexity not only lower bounds but also upper bounds the distributional communication complexity (up to logarithmic multiplicative factors in the communication complexity) [1]. On the other hand, recent works due to Ganor, Kol and Raz [3, 4, 5] show that there exist non-product distributions which exhibit exponential separation between internal information complexity and distributional communication complexity[1]. However, it is still open if internal information complexity (or a polynomial of it) upper bounds the public-coin randomized communication complexity (up to logarithmic multiplicative factors in the input size) [2].

Jain and Klauck [9], using tools from linear programming, gave a uniform treatment of several of the existing lower bound techniques and proposed the *partition bound.* This leads to following related (but incomparable) conjecture: does a polynomial of the partition bound yield an upper bound on the communication complexity? We are not aware of any counterexample to this conjecture[2].

We consider these questions when the inputs to Alice and Bob are drawn from a product distribution and show the following.

▶ **Theorem 1.** *Let $f : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$, and let $\mathsf{IC}_\varepsilon(f)$ and $\mathsf{prt}_\varepsilon(f)$ be the information complexity and partition bound respectively of f with error at most ε. For a product distribution μ on $\{0,1\}^n \times \{0,1\}^n$, the distributional communication complexity of f under distribution μ with error at most 0.49, denoted by $\mathsf{CC}^\mu_{0.49}(f)$, can be bounded above as follows:*

$$\mathsf{CC}^\mu_{0.49}(f) = O\left(\left(\mathsf{IC}_{1/8}(f) \cdot \log \mathsf{IC}_{1/8}(f)\right)^2\right), \tag{1.1}$$

$$\mathsf{CC}^\mu_{0.49}(f) = O\left(\left(\log \mathsf{prt}_{1/8}(f) \cdot \log\log \mathsf{prt}_{1/8}(f)\right)^2\right). \tag{1.2}$$

Our technique yields bounds more general than those stated above (see discussion after Proposition 7 for this generalization). We remark that recently (and independently of this work) Kol [11] obtained the bound (1.1) using very different techniques. Kol's result is stronger in the sense that her bound is in terms of the information complexity $\mathsf{IC}^\mu(f)$ for the product distribution μ, while our result is in terms of the worst case information complexity $\mathsf{IC}(f)$ (note, $\mathsf{IC}_\varepsilon(f) = \max_\mu \mathsf{IC}^\mu_\varepsilon(f)$). In fact, Kol showed that

$$\mathsf{CC}^\mu_{\delta+\varepsilon}(f) = O\left(\mathsf{IC}^\mu_\delta(f)^2 \cdot \operatorname{poly}\log \mathsf{IC}^\mu_\delta(f)/\varepsilon^5\right), \tag{1.3}$$

and concluded that

$$\mathsf{CC}^\mu_{0.49}(f) = O\left(\mathsf{IC}_{1/8}(f)^2 \cdot \operatorname{poly}\log \mathsf{IC}_{1/8}(f)\right). \tag{1.4}$$

[1] The third result of Ganor, Kol and Raz [5] actually demonstrates an exponential separation between external information and communication complexity, albeit not for computing a Boolean function.

[2] The recent work of Göös et al. [6] demonstrates the existence of a total function for which the partition bound is strictly sublinear in the randomized communication complexity. This still does not rule out communication complexity being bound by a polynomial of the partition bound.

Kol's result (1.3) is incomparable to our second result in terms of partition bound (1.2).

We consider a similar question in query complexity and show the following.

▶ **Theorem 2.** *Let $g : \{0,1\}^n \to \{0,1\}$ be a function and μ be a bit-wise product distribution on $\{0,1\}^n$. Let $\mathsf{qprt}_\varepsilon(g)$ be the query partition bound for g with error ε. Then, the distributional query complexity with error at most 0.49 under the distribution μ, denoted by $\mathsf{QC}^\mu_{0.49}(f)$, can be bounded above as follows:*

$$\mathsf{QC}^\mu_{0.49}(g) = O\left(\left(\log \mathsf{qprt}_{1/8}(g) \cdot \log\log \mathsf{qprt}_{1/8}(g)\right)^2\right) .$$

A similar quadratic upper bound for query complexity for product distributions in terms of approximate certificate complexity was obtained by Smyth [14]. His proof uses Reimer's inequality while our proof technique is based on Nisan and Wigderson's [13] more elementary approach.

Organization. The communication complexity result is proven in §2 while the query complexity result is deferred to the full version [8] for lack of space.

2 Communication Complexity

2.1 Preliminaries

We work in Yao's two-party communication model [15] (see Kushilevitz and Nisan [12] for an excellent introduction to the area). Let $\mathcal{X}$, $\mathcal{Y}$ and $\mathcal{Z}$ be finite non-empty sets, and let $f : \mathcal{X} \times \mathcal{Y} \to \mathcal{Z}$ be a function. A two-party protocol for computing f consists of two parties, Alice and Bob, who get inputs $x \in \mathcal{X}$ and $y \in \mathcal{Y}$ respectively, and exchange messages in order to compute $f(x,y) \in \mathcal{Z}$ (using shared randomness).

For a distribution μ on $\mathcal{X} \times \mathcal{Y}$, let the ε-error distributional communication complexity of f under μ (denoted by $\mathsf{CC}^\mu_\varepsilon(f)$), be the number of bits communicated (for the worst-case input) by the best deterministic protocol for f with average error at most ε under μ. Let $\mathsf{CC}^{\text{pub}}_\varepsilon(f)$, the public-coin randomized communication complexity of f with worst case error ε, be the number of bits communicated (for the worst-case input) by the best public-coin randomized protocol that for each input (x,y) computes $f(x,y)$ correctly with probability at least $1-\varepsilon$. Randomized and distributional complexity are related by the following special case of von Neumann's minmax principle.

▶ **Theorem 3** (Yao's minmax principle [16])**.** $\mathsf{CC}^{\text{pub}}_\varepsilon(f) = \max_\mu \mathsf{CC}^\mu_\varepsilon(f)$.

We will prove Theorem 1 by first showing an upper bound on communication complexity in terms of the smooth rectangle bound and then observing that the smooth rectangle bound is bounded above by the partition bound.

Smooth rectangle bound

The smooth rectangle bound was introduced by Jain and Klauck [9] as a generalization of the rectangle bound. Just like the rectangle bound, the smooth rectangle bound also provides a lower bound for randomized communication complexity. Informally, the smooth rectangle bound for a function f under a distribution μ, is the maximum over all functions g , which are close to f under the distribution μ, of the rectangle bound of g. However, it will be more convenient for us to work with the following linear programming formulation. (See [9, Lemma 2] and [10, Lemma 6] for the relations between the LP formulation and the more "natural" formulation in terms of rectangle bound.) It is evident from the LP formulation that the smooth rectangle bound is a further relaxation of the partition bound (defined in the appendix). We will formulate our results in terms of a distributional version of the above smooth rectangle bound. For $\mu : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}$ and any $z \in \mathcal{Z}$ and rectangle R, let $\mu_z(R) := \mu(R \cap f^{-1}(z))$ and $\mu_{\bar{z}}(R) := \mu(R) - \mu_z(R)$. Furthermore, let $\mu_z := \mu_z(\mathcal{X} \times \mathcal{Y})$ and $\mu_{\bar{z}} := \mu_{\bar{z}}(\mathcal{X} \times \mathcal{Y})$. The smooth rectangle and its distributional version are defined below.

▶ **Definition 4** (Smooth rectangle bound)**.**

- For a function $f : \mathcal{X} \times \mathcal{Y} \to \mathcal{Z}$ and $\varepsilon \in (0,1)$, the (ε, δ)-smooth rectangle bound of f denoted $\mathsf{srec}_{\varepsilon,\delta}(f)$ is defined to be $\max\{\mathsf{srec}^z_{\varepsilon,\delta}(f) : z \in \mathcal{Z}\}$, where $\mathsf{srec}^z_{\varepsilon,\delta}(f)$ is the optimal value of the following linear program.
- For a distribution μ on $\mathcal{X} \times \mathcal{Y}$ and function $f : \mathcal{X} \times \mathcal{Y} \to \mathcal{Z}$, the (ε, δ)-smooth rectangle bound of f with respect to μ denoted $\mathsf{srec}^\mu_{\varepsilon,\delta}(f)$ is defined to be $\max\{\mathsf{srec}^{z,\mu}_{\varepsilon,\delta}(f) : z \in \mathcal{Z}\}$, where $\mathsf{srec}^{z,\mu}_{\varepsilon,\delta}(f)$ is the optimal value of the following linear program.

$\underline{\mathsf{srec}^z_{\varepsilon,\delta}(f)}$

$$\begin{aligned} \min \quad & \sum_R w_R \\ & \sum_{R \ni (x,y)} w_R \geq 1 - \varepsilon, \quad \forall (x,y) \in f^{-1}(z) \\ & \sum_{R \ni (x,y)} w_R \leq \delta, \quad \forall (x,y) \notin f^{-1}(z) \\ & \sum_{R \ni (x,y)} w_R \leq 1, \quad \forall (x,y) \\ & w_R \geq 0, \quad \forall R\ . \end{aligned}$$

$\underline{\mathsf{srec}^{z,\mu}_{\varepsilon,\delta}(f)}$

$$\begin{aligned} \min \quad & \sum_R w_R \\ \sum_{(x,y) \in f^{-1}(z)} \mu_{x,y} & \sum_{R \ni (x,y)} w_R \geq (1 - \varepsilon) \cdot \mu_z && (2.1) \\ & \sum_{R \ni (x,y)} w_R \leq \delta, \quad \forall (x,y) \notin f^{-1}(z) && (2.2) \\ & \sum_{R \ni (x,y)} w_R \leq 1, \quad \forall (x,y) && (2.3) \\ & w_R \geq 0, \quad \forall R\ . \end{aligned}$$

We will refer to the constraint in (2.1) as the covering constraint and the ones in (2.2) as the packing constraints. Note that while there is a single covering constraint (averaged over all the inputs (x, y) that satisfy $f(x, y) = z$) there are packing constraints corresponding to each $(x, y) \notin f^{-1}(z)$.

Similar to Yao's minmax principle Theorem 3, we have the following proposition relating the distributional version of the smooth rectangle bound to the smooth rectangle bound.

▶ **Proposition 5.** $\mathsf{srec}_{\varepsilon,\delta}(f) = \max_\mu \mathsf{srec}^\mu_{\varepsilon,\delta}(f)$.

The main result of this section is the following

▶ **Theorem 6.** *For any Boolean function* $f : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$ *and any* product *distribution* μ *on* $\{0,1\}^n \times \{0,1\}^n$*, we have the following.*

1. $\mathsf{CC}^\mu_{0.49}(f) = O\left((\log \mathsf{srec}^\mu_{1/n^2,1/n^2}(f))^2 \cdot \log n\right)$.
2. *Furthermore, if there exists* $k \geq 20$ *such that*

$$\lceil 100 \log \mathsf{srec}^\mu_{\delta,\delta}(f) \rceil \leq k,$$

for $\delta \leq 1/(30 \cdot 100(k+1)^4)$*, then*

$$\mathsf{CC}^\mu_{0.49}(f) = O(k^2).$$

The above theorem is useful only when we have a upper bound on the smooth rectangle bound for very small δ. The following proposition shows that such upper bounds for smooth rectangle bound for such small δ can be obtained in terms of either the information complexity or the partition bound.

▶ **Proposition 7.** *For any Boolean function* $f : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$ *and any* $\delta \in (0,1)$*, we have the following bounds on* $\mathsf{srec}_{\delta,\delta}(f)$*:*

$$\log \mathsf{srec}_{\delta,\delta}(f) \le O\left(\log \frac{1}{\delta}\right) \cdot \mathsf{IC}_{1/8}(f),$$
$$\log \mathsf{srec}_{\delta,\delta}(f) \le O\left(\log \frac{1}{\delta}\right) \cdot \log \mathsf{prt}_{1/8}(f).$$

(This proposition depends on the error-reduction properties of information complexity and partition bound; a proof appears in the full version [8].) Using this proposition, we can reduce the error (i.e., δ) to $1/n^2$ and show that $\mathsf{CC}^{\mu}_{0.49}(f) = O\left(\left(\log \mathsf{prt}_{1/8}(f)\right)^2 \cdot (\log n)^3\right)$. However, we can also reduce the error to $1/\mathrm{poly}(\log \mathsf{prt}_{1/8}(f))$ and show that there exists a $k = O\left(\log \mathsf{prt}_{1/8}(f) \cdot \log\log \mathsf{prt}_{1/8}(f)\right)$ that satisfies the hypothesis for the second part of Theorem 6. The bound (1.2) in Theorem 1 now follows by combining Propositions 7 and 5 and Theorem 6. A similar argument yields the bound (1.1).

In particular, the above discussion shows that our techniques apply to any complexity measure (not necessarily partition bound and information complexity) which can be used to bound the smooth rectangle bound for very small δ. An interesting question that arises in this context is if we could bound smooth rectangle bound for small δ in terms of smooth rectangle bound for large δ, say $\delta = 1/3$ (i.e., is error-reduction for srec feasible?). This question was answered in the negative for partial functions by Göös et al. [7] who show that there exists a partial function f that has $\mathsf{srec}_{1/3}(f) = O(\log n)$ and yet $\mathsf{srec}_{1/4}(f) = \Omega(n)$.

2.2 Proof of Theorem 6

In this section, we construct a communication protocol tree with a small number of leaves from the optimal solutions to the LPs corresponding to $\mathsf{srec}^{0,\mu}_{\varepsilon,\delta}$ and $\mathsf{srec}^{1,\mu}_{\varepsilon,\delta}$. The construction of the protocol tree with a small number of leaves is inspired by a construction due to Nisan and Wigderson, in the context of log-rank conjecture [13, Theorem 2] (see also [12, Combinatorial proof of Theorem 2.11]). Unlike the earlier constructions, our protocol works for a distribution and allows for error. As a result, the decomposition into sub-problems needs to be performed more carefully. This step critically uses the product nature of the distribution μ.

The decomposition is accomplished using an inductive argument. We will work with the quantity $\mathsf{srec}^0 + \mathsf{srec}^1$. That is, we will show that if this sum is small, then there is a protocol with few leaves. Suppose $\mathsf{srec}^0 \le \mathsf{srec}^1$. Since srec^0 is small, we will conclude that there is a large rectangle biased towards 0 (see Lemma 8). Based on this large rectangle, the entire communication matrix is partitioned into three parts: (1) the large biased rectangle itself, (2) a rectangle whose corresponding sub-problem admits an LP solution leading to a smaller srec^1 value (the underlying product nature of the distribution μ is used here) and (3) a rectangle where the total measure with respect to μ drops significantly (see Lemma 9).

We say that a rectangle R is $(1-\alpha)$-biased towards to 0 if $\mu_1(R) \le \alpha\mu_0(R)$.

▶ **Lemma 8** (large biased rectangle). *Let μ be a product distribution. If $\mathsf{srec}^{0,\mu}_{\varepsilon,\delta}(f) \le D$, then for every $\rho \in (0,1)$ there exists a rectangle S such that S is $(1-\rho)$-biased towards 0 and*

$$\mu(S) \ge \mu_0(S) \ge \frac{1}{D} \cdot \left((1-\varepsilon) \cdot \mu_0 - \left(\frac{\delta}{\rho}\right) \cdot \mu_1 \right).$$

(The proof appears in §2.3.) We will apply the above lemma with $\rho = \sqrt{\delta}$ and conclude that there exists a large rectangle $S = X_0 \times Y_0$ that is $(1-\sqrt{\delta})$-biased towards 0. Let $X_1 = \mathcal{X} \setminus X_0$ and $Y_1 = \mathcal{Y} \setminus Y_0$. For $i, j \in \{0,1\}$, define rectangles $R^{(ij)} := X_i \times Y_j$, $R^{(1*)} := X_1 \times \mathcal{Y}$, and $R^{(*1)} := \mathcal{X} \times Y_1$. (Note, $S = R^{(00)}$.) For $i, j \in \{0, 1, *\}$, let $\mu^{(ij)}$ be the restriction of μ to the rectangle $R^{(ij)}$. We show in the lemma below that the function f when restricted to either $R^{(10)}$ or $R^{(01)}$ has the property that the corresponding srec^1 drops by a constant factor. Define

$$\varepsilon(f) := 1 - \frac{\left(\sum_{(x,y)\in f^{-1}(1)} \mu_{x,y} \sum_{R:(x,y)\in R} w_R\right)}{\mu_1},$$

$$\varepsilon^{(ij)}(f) := 1 - \frac{\left(\sum_{(x,y)\in f^{-1}(1)\cap R^{(ij)}} \mu_{x,y} \sum_{R:(x,y)\in R} w_R\right)}{\mu_1(R^{(ij)})}; \qquad \text{for } i,j \in \{0,1\}.$$

It follows from the covering constraint that $\varepsilon(f) \le \varepsilon$. Furthermore, $\varepsilon(f)$ is an average of the $\varepsilon^{(ij)}$'s in the sense that $\varepsilon(f) = \left(\sum_{i,j\in\{0,1\}} \mu_1(R^{(ij)})\varepsilon^{(ij)}\right)/\mu_1$.

▶ **Lemma 9.** *Suppose the product distribution μ and rectangles $R^{(ij)}$ are as above; in particular, $R^{(00)}$ is $(1-\sqrt{\delta})$-biased towards 0. There exists an $(ij) \in \{(01),(10)\}$ such that one of the following holds: (a) $2\mu^{(ij)}(f^{-1}(1)) \le \mu^{(ij)}(f^{-1}(0))$ or (b) $\mathsf{srec}^{1,\mu^{(ij)}}_{\varepsilon^{(ij)}+30\sqrt[4]{\delta},\delta}(f) \le 0.9D$ where $\varepsilon^{(ij)}$ is as defined above.*

We will prove this lemma in §2.3. Let us assume the above lemmas and obtain the low cost communication protocol claimed in Theorem 6.

Suppose $\mu^{(01)}$ satisfies $\mathsf{srec}^{1,\mu^{(01)}}_{\varepsilon^{(01)}+30\sqrt[4]{\delta},\delta}(f) \le 0.9D$ as given by the above lemma. Consider the decomposition of the space $\mathcal{X} \times \mathcal{Y}$ given by $(R^{(00)}, R^{(01)}, R^{(1*)} = R^{(10)} \cup R^{(11)})$. We note that $R^{(00)}$ is a large biased rectangle, $R^{(01)}$ has lower srec^1 value while $R^{(1*)}$ has lower μ value (since $R^{(00)}$ is large) and its srec values are no larger than that of the entire space. In the case when $\mu^{(10)}$ satisfies $\mathsf{srec}^{1,\mu^{(10)}}_{\varepsilon^{(10)}+30\sqrt[4]{\delta},\delta}(f) \le 0.9D$, we similarly have the decomposition $(R^{(00)}, R^{(10)}, R^{(*1)} = R^{(01)} \cup R^{(11)})$.

This suggests a natural inductive protocol Π for f that we formalize in the lemma below.

For our induction it will be convenient to work with μ that are not necessarily normalized. So, we will only assume $\mu : \mathcal{X} \times \mathcal{Y} \to [0,1]$ but not that $|\mu| := \mu(\mathcal{X} \times \mathcal{Y}) = \sum_{(x,y)\in\mathcal{X}\times\mathcal{Y}} \mu(x,y) = 1$. For a protocol Π, let the advantage of Π be defined by

$$\mathsf{adv}_\mu(\Pi) = \sum_{(x,y):f(x,y)=\Pi(x,y)} \mu(x,y) - \sum_{(x,y):f(x,y)\ne\Pi(x,y)} \mu(x,y).$$

Let $L(\Pi)$ be the number of leaves in Π.

We now formulate the induction hypothesis as follows.

▶ **Lemma 10.** *Fix a function $f : \mathcal{X} \times \mathcal{Y} \to \{0,1\}$ and a product distribution (not necessarily normalized) $\mu : \mathcal{X} \times \mathcal{Y} \to [0,1]$ such that $|\mu| \ge 0$. Let $\varepsilon, \delta \in (0,1)$ and $\Delta \in (0, |\mu|)$. Let s, t be* *non-negative integers such that*

$$s \geq s(\mu,\varepsilon,\delta) := \left\lceil 100 \cdot \log 2(\mathsf{srec}^{0,\mu}_{\varepsilon,\delta}(f) + \mathsf{srec}^{1,\mu}_{\varepsilon,\delta}(f)) \right\rceil ;$$
$$t \geq t(\mu,\varepsilon,\delta) := \lceil 100 \cdot 2^s \log(|\mu|/\Delta) \rceil .$$

Then, there is a protocol Π *such that*

$$L(\Pi) \leq 4\binom{s+t}{t} - 1; \tag{2.4}$$
$$\mathsf{adv}_\mu(\Pi) \geq \left(\frac{1}{10} - \varepsilon - 30(s+1)\sqrt[4]{\delta}\right)|\mu| - \Delta \cdot L(\Pi). \tag{2.5}$$

▶ Remark. Since $\varepsilon \leq \frac{1}{2}$, our definitions imply that $\mathsf{srec}^{1,\mu}_{\varepsilon,\delta}(f) + \mathsf{srec}^{1,\mu}_{\varepsilon,\delta}(f)) \geq \frac{1}{2}$; thus $s \geq 0$. Similarly, since $\Delta \leq |\mu|$, we have $t \geq 0$.

Proof. First, we observe that if $\max\{\mu_0,\mu_1\} \geq 2\min\{\mu_0,\mu_1\}$, then the protocol Π consisting of just one leaf, with the most popular value as label, meets the requirements: for, $\mathsf{adv}_\mu(\Pi) \geq \frac{1}{3}|\mu|$ and $L(\Pi) = 1$, and our claim holds. Also, we may assume that $\varepsilon - 30(s+1)\sqrt[4]{\delta} < \frac{1}{10}$, for otherwise the claim is trivially true.

We now proceed by induction on $s+t$, assuming that μ is balanced: $\max\{\mu_0,\mu_1\} \leq 2\min\{\mu_0,\mu_1\}$.

Base case ($s = 0$)

Since $s = 0$, we have $\log \mathsf{srec}^{1,\mu}_{\varepsilon,\delta}(f) \leq \frac{1}{100}$. We will show a protocol Π where Alice sends one bit after which Bob announces the answer. Consider the optimal solution $\langle w_R : R \text{ a rectangle}\rangle$ to the LP corresponding to $\mathsf{srec}^{1,\mu}_{\varepsilon,\delta}(f)$; thus, $\mathsf{OPT} := \sum_R w_R = \mathsf{srec}^{1,\mu}_{\varepsilon,\delta}(f) \leq 2^{1/100} \leq 2$. Let $R = R_X \times R_Y$ be a random rectangle picked with probability proportional to w_R (using public coins). In the protocol Π, Alice tells Bob if $x \in R_X$, and Bob returns the answer 1 if $(x,y) \in R_Y$ and returns 0 otherwise. Let $p_{xy} := \Pr_R[(x,y) \in R]$. Then, by (2.1) we have $\sum_{(x,y)\in f^{-1}(1)} \mu(x,y)p_{xy} \geq (1-\varepsilon)\mu_1/\mathsf{OPT}$, and by (2.2), we have $\sum_{(x,y)\in f^{-1}(0)} \mu(x,y)p_{xy} \leq \delta\mu_0/\mathsf{OPT}$. Thus,

$$\begin{aligned}
\mathbb{E}_R\left[\sum_{(x,y):\Pi(x,y)\neq f(x,y)} \mu(x,y)\right] &= \sum_{(x,y)\in f^{-1}(1)} \mu(x,y)(1-p_{xy}) + \sum_{(x,y)\in f^{-1}(0)} \mu(x,y)p_{x,y} \\
&\leq \mu_1 - (1-\varepsilon)\mu_1/\mathsf{OPT} + \delta\mu_0/\mathsf{OPT} \\
&\leq \mu_1 - ((1-\varepsilon)\mu_1 - \delta\mu_0)/\mathsf{OPT} \\
&\leq \frac{1}{2}(\mu_1 + \varepsilon\mu_1 + \delta\mu_0) \qquad (\text{since } \mathsf{OPT} \leq 2).
\end{aligned} \tag{2.6}$$

Fix a choice R for which the quantity under the expectation is at most $\frac{1}{2}(\mu_1 + \varepsilon\mu_1 + \delta\mu_0)$. Then,

$$\begin{aligned}
\mathsf{adv}(\Pi) &= |\mu| - 2\sum_{(x,y):\Pi(x,y)\neq f(x,y)} \mu(x,y) \\
&\geq |\mu| - (\mu_1 + \varepsilon\mu_1 + \delta\mu_0) \\
&\geq \left(\frac{1}{3} - \varepsilon - \delta\right)|\mu| \qquad (\text{since } \mu_1 \leq 2\mu_0).
\end{aligned}$$

Base case ($t = 0$)

In this case, $|\mu| = \Delta$, and the protocol Π with a single leaf that gives the most probable answer achieves $\mathsf{adv}(\Pi) \geq 0 \geq |\mu| - \Delta$.

Induction step

We will use Lemma 8 to decompose the communication matrix into a small number of rectangles. After an exchange of a few bits to determine in which rectangle the input lies, Alice and Bob will be left with a problem for which s or t is significantly smaller. Assume $\mathsf{srec}^{1,\mu}_{\varepsilon,\delta}(f) \geq \mathsf{srec}^{1,\mu}_{\varepsilon,\delta}(f)$; in particular, $\mathsf{srec}^{1,\mu}_{\varepsilon,\delta}(f) \leq 2^{s/100}$.

Formally, from Lemma 8 (taking $\rho = \sqrt{\delta}$), we obtain a rectangle $R^{(00)} = X_0 \times Y_0$ such that (a) $R^{(00)}$ is $(1-\sqrt{\delta})$-biased towards 0, and (b) $\mu(R^{(00)}) \geq \frac{1}{2^{s/100}}(1-\varepsilon-2\sqrt{\delta})|\mu_0| \geq \frac{1}{3\cdot 2^{s/100}}(1-\varepsilon-2\sqrt{\delta})|\mu|$. Recall the definitions of the rectangles $R^{(10)}, R^{(01)}, R^{(11)}, R^{(1*)}, R^{(*1)}$ and the corresponding restrictions of μ, namely, $\mu^{(01)}, \mu^{(10)}, \mu^{(11)}, \mu^{(1*)}, \mu^{(*1)}$. Suppose the choice of ij in Lemma 9 for which one of the alternatives holds is $ij = 01$ (the other case $ij = 10$ is symmetric). The protocol Π proceeds as follows. Alice starts by telling Bob if $x \in X_0$.

Alice says $x \in X_0$. Now, Bob tells Alice if $y \in Y_0$.

Bob says $y \in Y_0$. The protocol $\Pi^{(00)}$ in this case has one leaf with answer 0; thus $\mathsf{adv}(\Pi^{(00)}) \geq |\mu^{(00)}| \cdot (1-\sqrt{\delta})$.

Bob says $y \notin Y_0$. Alice and Bob follow the protocol $\Pi^{(01)}$ promised by induction for $R^{(01)}$ under $\mu^{(01)}$. To bound the number of leaves in $\Pi^{(01)}$, we will consider the two alternatives ((a) and (b)) specified in Lemma 9 separately. First (alternative (a)) suppose $2\mu^{(01)}(f^{-1}(1)) \leq \mu^{(01)}(f^{-1}(0))$; then we immediately declare 0 as the response, so that $L(\Pi^{(01)}) = 1$ and $\mathsf{adv}(\Pi^{(01)}) \geq |\mu^{(01)}|/3$. If alternative (b) holds, then we have

$$\mathsf{srec}^{1,\mu^{(01)}}_{\varepsilon^{(01)}+30\sqrt[4]{\delta},\delta}(f) \leq 0.9\mathsf{srec}^{1,\mu}_{\varepsilon,\delta}(f). \tag{2.7}$$

Then, we obtain $\Pi^{(01)}$ by induction. We take $\varepsilon^{(01)} + 30\sqrt[4]{\delta}$ as ε (if this quantity is greater than 1, then we use a trivial protocol with one leaf and zero advantage). With the reduction promised in (2.7), we may use a value of s that is the old s minus 1. Thus, we have

$$L(\Pi^{(01)}) \leq 4\binom{(s-1)+t}{t} - 1;$$
$$\mathsf{adv}(\Pi^{(01)}) \geq |\mu^{(01)}| \cdot \left(\frac{1}{10} - (\varepsilon^{(01)} + 30\sqrt[4]{\delta}) - 30s\sqrt[4]{\delta}\right) - \Delta \cdot L(\Pi^{(01)}).$$

Alice says $x \notin X_0$. Alice and Bob follow the protocol $\Pi^{(1*)}$ obtained by applying the induction hypothesis to the rectangle $R^{(1*)}$ and the associated distribution $\mu^{(1*)}$. Observe that

$$|\mu^{(1*)}| \leq |\mu| - \mu(R^{(00)}) \leq |\mu|\left(1 - \frac{1}{3\cdot 2^{s/100}}(1-\varepsilon-2\sqrt{\delta})\right) \leq |\mu|\left(1 - \frac{1}{4\cdot 2^{s}}\right). \tag{2.8}$$

For the last inequality we used $\varepsilon + 2\sqrt{\delta} \leq \frac{1}{10}$, for otherwise (2.5) holds trivially. Now, (2.8) implies that $\log|\mu^{(1*)}| \leq \log|\mu| - \frac{1}{1002^s}$; so, for our induction we may take $t \leftarrow t-1$. The parameters ε, δ and Δ remain the same. The original LP solutions are still valid for the subproblem, so we use the same s. The protocol $\Pi^{(1*)}$ obtained by induction satisfies the following inequalities.

$$L(\Pi^{(1*)}) \leq 4\binom{s+(t-1)}{t-1} - 1;$$
$$\mathsf{adv}(\Pi^{(1*)}) \geq |\mu^{(1*)}| \cdot \left(\frac{1}{10} - \varepsilon^{(1*)} - 30(s+1)\sqrt[4]{\delta}\right) - \Delta \cdot L(\Pi^{(1*)}).$$

Putting all the contributions together, we obtain

$$\begin{aligned} L(\Pi) &= 1 + L(\Pi^{(01)}) + L(\Pi^{(1*)}) \\ &\leq 1 + \left(4\binom{(s-1)+t}{t} - 1\right) + \left(4\binom{s+(t-1)}{t-1} - 1\right) \\ &= 4\binom{s+t}{t} - 1; \end{aligned}$$
$$\begin{aligned} \mathsf{adv}(\Pi) &\geq |\mu^{(00)}| \cdot (1 - \sqrt{\delta}) \\ &\quad + |\mu^{(01)}| \cdot \left(\frac{1}{10} - (\varepsilon^{(01)} + 30\sqrt[4]{\delta}) - 30s\sqrt[4]{\delta}\right) - \Delta \cdot L(\Pi^{(01)}) \\ &\quad + |\mu^{(1*)}| \cdot \left(\frac{1}{10} - \varepsilon^{(1*)} - 30(s+1)\sqrt[4]{\delta}\right) - \Delta \cdot L(\Pi^{(1*)}) \\ &\geq \left(\frac{1}{10} - \varepsilon - 30(s+1)\sqrt[4]{\delta}\right)|\mu| - \Delta \cdot L(\Pi). \end{aligned}$$ ◀

The above lemma yields a protocol whose protocol tree has a small number of leaves, but not necessarily small depth. We can balance the protocol tree using the following proposition.

▶ **Proposition 11** ([12, Lemma 2.8]). *If f has a deterministic communication protocol tree with ℓ leaves, then f has a protocol tree with depth at most $O(\log \ell)$.*

We are now in a position to complete the proof of the main theorem of this section.

Proof of Theorem 6. To prove the first part of Theorem 6, we invoke Lemma 10 with $\Delta = 1/2^{4n}$ and $\varepsilon = \delta = 1/n^2$ to derive a protocol tree Π with at most

$$L(\Pi) = n^{O\left(\log \mathsf{srec}^{1,\mu}_{1/n^2,1/n^2}(f)\right)^2}$$

leaves and advantage at least 1/20. The first part now follows from Proposition 11.

To prove the second part of Theorem 6, we invoke Lemma 10 with $s = k$, $\Delta = 1/2^{5k^2}$ and $\varepsilon = \delta = 1/(30 \cdot 100(k+1)^4)$ where k satisfies the hypothesis. With this setting of parameters $t = \lceil 500 \cdot 2^k k^2 \rceil \leq 2^{2k}$ (for $k \geq 20$). Lemma 10 implies a protocol tree Π with at most

$$L(\Pi) \leq (t+s)^s \leq t^{2s} \leq 2^{4k^2}$$

leaves and advantage at most 1/20. The second claim then follows from Proposition 11. ◀

2.3 Proofs of Lemmas 8–9

Proof of Lemma 8. Fix $z \in \{0,1\}$. In the following we say that a rectangle R is *biased* (towards 0) if $\mu_1(R) \leq \rho \cdot \mu_0(R)$; otherwise, we say it is unbiased. Fix a solution $\langle w_R :$

R is a rectangle$\rangle$ that achieves the optimum $\mathsf{srec}^{0,\mu}_{\varepsilon,\delta}(f) \leq D$. It follows

$$\begin{aligned}\sum_{R:\text{unbiased}} w_R \cdot \mu_0(R) &\leq \sum_{R:\text{unbiased}} w_R \cdot \frac{\mu_1(R)}{\rho} \\ &\leq \frac{1}{\rho} \cdot \sum_R w_R \cdot \mu_1(R) \\ &= \frac{1}{\rho} \sum_{(x,y)\in f^{-1}(1)} \mu(x,y) \sum_{R:(x,y)\in R} w_R \\ &\leq \frac{\delta}{\rho} \cdot \mu_1,\end{aligned}$$

where the last inequality follows from the packing constraints (2.2). We now use the covering constraints (2.1) to conclude that

$$\sum_{R:\text{biased}} w_R \cdot \mu_0(R) = \sum_R w_R \cdot \mu_0(R) - \sum_{R:\text{unbiased}} w_r \cdot \mu_0(R) \geq (1-\varepsilon)\cdot \mu_0 - \frac{\delta}{\rho} \cdot \mu_1. \quad (2.9)$$

Define a probability distribution on the rectangles R as follows $p(R) := w_R/\mathsf{srec}^{0,\mu}_{\varepsilon,\delta}(f)$. Then (2.9) can be rewritten as

$$\mathbb{E}_R\left[\mathbb{I}_{\text{biased}}(R) \cdot \mu_0(R)\right] \geq \frac{1}{D} \cdot \left((1-\varepsilon)\cdot \mu_0 - \frac{\delta}{\rho}\cdot \mu_1\right).$$

Hence, there exists a large biased rectangle $S = X_0 \times Y_0$ as claimed. ◀

Proof of Lemma 9. Since $R^{(00)}$ is $(1-\sqrt{\delta})$-biased towards 0, we have from the packing and covering constraints (2.2) and (2.3) that

$$\begin{aligned}&\sum_{(x,y)\in R^{(00)}} \mu_{x,y} \sum_{R\ni(x,y)} w_R \\ &= \sum_{(x,y)\in R^{(00)}\cap f^{-1}(1)} \mu_{x,y} \sum_{R\ni(x,y)} w_R + \sum_{(x,y)\in R^{(00)}\cap f^{-1}(0)} \mu_{x,y} \sum_{R\ni(x,y)} w_R \\ &\leq \mu_1(R^{(00)}) + \delta\mu_0(R^{(00)}) \leq (\sqrt{\delta}+\delta)\mu_0(R^{(00)}) \leq 2\sqrt{\delta}\mu(R^{(00)}).\end{aligned}$$

Hence,

$$\sum_R w_R \cdot \left(\frac{\mu(R^{(00)}\cap R)}{\mu(R^{(00)})}\right) \leq 2\sqrt{\delta}. \quad (2.10)$$

Group the rectangles in to subsets as follows:

$$B^{(01)} := \left\{R : \frac{\mu(R^{(01)}\cap R)}{\mu(R^{(01)})} \geq \frac{10\sqrt[4]{\delta}}{D}\right\}, \qquad B^{(10)} := \left\{R : \frac{\mu(R^{(10)}\cap R)}{\mu(R^{(10)})} \geq \frac{10\sqrt[4]{\delta}}{D}\right\},$$

$$B := \left\{R : \frac{\mu(R^{(11)}\cap R)}{\mu(R^{(11)})} \geq \frac{10}{D}\right\}.$$

By (2.3), we have

$$\sum_{(x,y)\in R^{(11)}} \mu_{x,y} \sum_{R\ni(x,y)} w_R \leq \sum_{(x,y)\in R^{(11)}} \mu_{x,y} = \mu(R^{(11)}).$$

Or equivalently,

$$\sum_R \frac{w_R}{D} \cdot \frac{\mu(R^{(11)} \cap R)}{\mu(R^{(11)})} \le \frac{1}{D}.$$

Hence,

$$\sum_{R \in B} w_R \le 0.1D. \tag{2.11}$$

We will now argue that either $\sum_{R \in B^{(01)}} w_R \le 0.9D$ or $\sum_{R \in B^{(10)}} w_R \le 0.9D$. Suppose, for contradiction, that neither is true. Then, by (2.11) we have

$$\sum_{R \in \left(B^{(01)} \cap B^{(10)}\right) \setminus B} w_R \ge 0.7D. \tag{2.12}$$

Since μ is a product distribution we have

$$\frac{\mu(R^{(01)} \cap R)}{\mu(R^{(01)})} \cdot \frac{\mu(R^{(10)} \cap R)}{\mu(R^{(10)})} = \frac{\mu(R^{(00)} \cap R)}{\mu(R^{(00)})} \cdot \frac{\mu(R^{(11)} \cap R)}{\mu(R^{(11)})}.$$

Using the above we have

$$\begin{aligned}
&\sum_R w_R \cdot \left(\frac{\mu(R^{(00)} \cap R)}{\mu(R^{(00)})}\right) \\
&\ge \sum_{R \in \left(B^{(01)} \cap B^{(10)}\right) \setminus B} w_R \cdot \left(\frac{\mu(R^{(00)} \cap R)}{\mu(R^{(00)})}\right) \\
&\ge \sum_{R \in \left(B^{(01)} \cap B^{(10)}\right) \setminus B} w_R \cdot \left(\frac{\mu(R^{(01)} \cap R)}{\mu(R^{(01)})}\right) \cdot \left(\frac{\mu(R^{(10)} \cap R)}{\mu(R^{(10)})}\right) \Big/ \left(\frac{\mu(R^{(11)} \cap R)}{\mu(R^{(11)})}\right) \\
&\ge \sum_{R \in \left(B^{(01)} \cap B^{(10)}\right) \setminus B} w_R \cdot \left(\frac{10\sqrt[4]{\delta}}{D}\right) \cdot \left(\frac{10\sqrt[4]{\delta}}{D}\right) \Big/ \left(\frac{10}{D}\right) \\
&\ge \frac{10\sqrt{\delta}}{D} \cdot (0.7D) \\
&= 7\sqrt{\delta}.
\end{aligned}$$

This contradicts (2.10). Hence, either $\sum_{R \in B^{(01)}} w_R \le 0.9D$ or $\sum_{R \in B^{(10)}} w_R \le 0.9D$. Assume, wlog that $\sum_{R \in B^{(01)}} w_R \le 0.9D$. If f is 1/2-biased towards 0 with respect to the distribution $\mu^{(01)}$, then the alternative (a) of the lemma holds, and we are done. Otherwise, that is $\mu_0(R^{(01)}) \le 2\mu_1(R^{(01)})$ or equivalently $\mu(R^{(01)}) \le 3\mu_1^{(01)}(R^{(01)})$. We will infer from this that $\mathsf{srec}^{1,\mu^{(01)}}_{\varepsilon^{(01)}+30\sqrt[4]{\delta},\delta}(f) \le 0.9D$. Consider the primal solution given by

$$w'_R = \begin{cases} w_R, & \text{if } R \in B^{(01)} \\ 0, & \text{if } R \notin B^{(01)}. \end{cases}$$

Clearly, w'_R, being a part of the original solution, satisfies (2.2) and (2.3), and has objective value at most $0.9D$. All we need to show is that it satisfies the covering constraint (2.1). For this, we first consider

$$\sum_{R \notin B^{(01)}} w_R \cdot \left(\frac{\mu_1(R^{(01)} \cap R)}{\mu(R^{(01)})}\right) \le \sum_{R \notin B^{(01)}} w_R \cdot \left(\frac{\mu(R^{(01)} \cap R)}{\mu(R^{(01)})}\right) \le \frac{10\sqrt[4]{\delta}}{D} \cdot D \le 10\sqrt[4]{\delta}. \tag{2.13}$$

Now,

$$\begin{aligned}
&\sum_{(x,y)\in f^{-1}(1)\cap R^{(01)}} \mu_{x,y} \sum_{R\in(x,y)} w'_R \\
&= \sum_{(x,y)\in f^{-1}(1)\cap R^{(01)}} \mu_{x,y} \sum_{R\in(x,y),R\in B^{(01)}} w_R \\
&= \sum_{(x,y)\in f^{-1}(1)\cap R^{(01)}} \mu_{x,y} \left(\sum_{R\in(x,y)} w_R - \sum_{R\in(x,y),R\notin B^{(01)}} w_R \right) \\
&= (1-\varepsilon^{(01)})\mu_1(R^{(01)}) - \sum_{(x,y)\in f^{-1}(1)\cap R^{(01)}} \mu_{x,y} \sum_{R\in(x,y),R\notin B^{(01)}} w_R \\
&= (1-\varepsilon^{(01)})\mu_1(R^{(01)}) - \sum_{R\notin B^{(01)}} w_R \mu_1(R^{(01)}\cap R) \\
&\geq (1-\varepsilon^{(01)})\mu_1(R^{(01)}) - 10\sqrt[4]{\delta}\mu(R^{(01)}) \qquad \text{[From (2.13)]} \\
&\geq (1-\varepsilon^{(01)})\mu_1(R^{(01)}) - 30\sqrt[4]{\delta}\mu_1(R^{(01)}) \qquad [\text{Since } \mu(R^{(01)}) \leq 3\mu_1(R^{(01)})] \\
&= (1-\varepsilon^{(01)} - 30\sqrt[4]{\delta})\mu_1(R^{(01)})
\end{aligned}$$

Thus, (2.1) holds for $R^{(01)}$ with ε replaced by $\varepsilon^{(01)} + 30\sqrt[4]{\delta}$. ◀

Acknowledgements. We thank Mika Göös for pointing out an error in an earlier version of the paper and for referring us to the literature.

References

1. Boaz Barak, Mark Braverman, Xi Chen, and Anup Rao. How to compress interactive communication. *SIAM J. Comput.*, 42(3):1327–1363, 2013. (Preliminary Version in *42nd STOC*, 2010). doi:10.1137/100811969.
2. Lila Fontes, Rahul Jain, Iordanis Kerenidis, Sophie Laplante, Mathieu Laurière, and Jérémie Roland. Relative discrepancy does not separate information and communication complexity. In Magnús M. Halldórsson, Kazuo Iwama, Naoki Kobayashi, and Bettina Speckmann, editors, *Proc. 42nd International Colloq. of Automata, Languages and Programming (ICALP), Part I*, volume 9134 of *LNCS*, pages 506–516. Springer, 2015. doi:10.1007/978-3-662-47672-7_41.
3. Anat Ganor, Gillat Kol, and Ran Raz. Exponential separation of information and communication. In *Proc. 55th IEEE Symp. on Foundations of Comp. Science (FOCS)*, pages 176–185, 2014. doi:10.1109/FOCS.2014.27.
4. Anat Ganor, Gillat Kol, and Ran Raz. Exponential separation of information and communication for boolean functions. In *Proc. 47th ACM Symp. on Theory of Computing (STOC)*, pages 557–566, 2015. doi:10.1145/2746539.2746572.
5. Anat Ganor, Gillat Kol, and Ran Raz. Exponential separation of communication and external information. In *Proc. 48th ACM Symp. on Theory of Computing (STOC)*, 2016. (to appear).
6. Mika Göös, T. S. Jayram, Toniann Pitassi, and Thomas Watson. Randomized communication vs. partition number. Technical Report TR15-169, Elect. Colloq. on Comput. Complexity (ECCC), 2015. URL: http://eccc.hpi-web.de/report/2015/169.
7. Mika Göös, Shachar Lovett, Raghu Meka, Thomas Watson, and David Zuckerman. Rectangles are nonnegative juntas. In *Proc. 47th ACM Symp. on Theory of Computing (STOC)*, pages 257–266, 2015. doi:10.1145/2746539.2746596.

8 Prahladh Harsha, Rahul Jain, and Jaikumar Radhakrishnan. Partition bound is quadratically tight for product distributions, 2016. `arXiv:1512.01968`.

9 Rahul Jain and Hartmut Klauck. The partition bound for classical communication complexity and query complexity. In *Proc. 25th IEEE Conf. on Computational Complexity*, pages 247–258, 2010. `arXiv:0910.4266`, `doi:10.1109/CCC.2010.31`.

10 Rahul Jain and Penghui Yao. A strong direct product theorem in terms of the smooth rectangle bound, 2012. `arXiv:1209.0263`.

11 Gillat Kol. Interactive compression for product distributions. In *Proc. 48th ACM Symp. on Theory of Computing (STOC)*, 2016. (to appear).

12 Eyal Kushilevitz and Noam Nisan. *Communication Complexity*. Cambridge University Press, 1997. `doi:10.1017/CBO9780511574948`.

13 Noam Nisan and Avi Wigderson. On rank vs. communication complexity. *Combinatorica*, 15(4):557–565, 1995. (Preliminary version in *35th FOCS*, 1994). `doi:10.1007/BF01192527`.

14 Clifford D. Smyth. Reimer's inequality and Tardos' conjecture. In *Proc. 34th ACM Symp. on Theory of Computing (STOC)*, pages 218–221, 2002. `doi:10.1145/509907.509942`.

15 Andrew Chi-Chih Yao. Some complexity questions related to distributive computing (preliminary report). In *Proc. 11th ACM Symp. on Theory of Computing (STOC)*, pages 209–213, 1979. `doi:10.1145/800135.804414`.

16 Andrew Chi-Chih Yao. Lower bounds by probabilistic arguments (extended abstract). In *Proc. 24th IEEE Symp. on Foundations of Comp. Science (FOCS)*, pages 420–428, 1983. `doi:10.1109/SFCS.1983.30`.