A Lower Bound for Sampling Disjoint Sets

Mika Göös
Institute for Advanced Study, Princeton, NJ, USA
mika@ias.edu

Thomas Watson
University of Memphis, TN, USA
Thomas.Watson@memphis.edu

Abstract

Suppose Alice and Bob each start with private randomness and no other input, and they wish to engage in a protocol in which Alice ends up with a set $x \subseteq [n]$ and Bob ends up with a set $y \subseteq [n]$, such that (x, y) is uniformly distributed over all pairs of disjoint sets. We prove that for some constant $\beta < 1$, this requires $\Omega(n)$ communication even to get within statistical distance $1 - \beta^n$ of the target distribution. Previously, Ambainis, Schulman, Ta-Shma, Vazirani, and Wigderson (FOCS 1998) proved that $\Omega(\sqrt{n})$ communication is required to get within some constant statistical distance $\varepsilon > 0$ of the uniform distribution over all pairs of disjoint sets of size $\sqrt{n}$.

2012 ACM Subject Classification Theory of computation → Communication complexity

Keywords and phrases Communication complexity, set disjointness, sampling

Digital Object Identifier 10.4230/LIPIcs.APPROX-RANDOM.2019.51

Category RANDOM

Related Version A full version of the paper is available at `https://eccc.weizmann.ac.il/report/2019/066/`.

Funding *Mika Göös*: Supported by NSF grant CCF-1412958.
Thomas Watson: Supported by NSF grant CCF-1657377.

1 Introduction

In most traditional computational problems, the goal is to take an input and produce the "correct" output, or produce one of a set of acceptable outputs. In a *sampling* problem, on the other hand, the goal is to generate a random sample from a specified probability distribution D, or at least from a distribution that is close to D. There has been a surge of interest in studying sampling problems from a complexity theory perspective [7, 36, 73, 1, 58, 32, 74, 13, 72, 47, 77, 15, 78, 75, 79, 76]. Unlike more traditional computational problems, sampling problems do not necessarily need to have any real input, besides the uniformly random bits fed into a sampling algorithm.

One commonly studied type of target distribution is "input–output pairs" of a function f, i.e., $(D, f(D))$ where D is perhaps the uniform distribution over inputs to f. Using an algorithm for computing f, one can sample $(D, f(D))$ by first sampling from D, then evaluating f on that input. However, for some functions f, generating an input jointly with the corresponding output may be computationally easier than evaluating f on an adversarially-chosen input. Thus in general, sampling lower bounds tend to be more challenging to prove than lower bounds for functions.

Many of the above-cited works focus on concrete computational models such as low-depth circuits. We consider the model of 2-party communication complexity, for which comparatively less is known about sampling. Which problem should we study? Well, the single most important function in communication complexity is Set-Disjointness, in which Alice gets a set

© Mika Göös and Thomas Watson;
licensed under Creative Commons License CC-BY

$x \subseteq [n]$, Bob gets a set $y \subseteq [n]$, and the goal is to determine whether $x \cap y = \emptyset$. Identifying the sets with their characteristic bit strings, this can be viewed as $\textsc{Disj}\colon \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$ where $\textsc{Disj}(x,y) = 1$ iff $x \wedge y = 0^n$. The applications of communication bounds for Set-Disjointness are far too numerous to list, but they span areas such as streaming, circuit complexity, proof complexity, data structures, property testing, combinatorial optimization, fine-grained complexity, cryptography, and game theory. Because of its central role, Set-Disjointness has become the de facto testbed for proving new types of communication bounds. This function has been studied in the contexts of randomized [9, 49, 62, 10, 17] and quantum [25, 43, 63, 2, 66, 70] protocols; multi-party number-in-hand [6, 10, 27, 41, 48, 18, 22] and number-on-forehead [40, 71, 12, 66, 28, 57, 11, 69, 68, 61, 60] models; Merlin–Arthur and related models [50, 3, 35, 39, 38, 4, 64, 29]; with a bounded number of rounds of interaction [52, 46, 80, 19, 23]; with bounds on the sizes of the sets [42, 56, 59, 31, 26, 65]; very precise relationships between communication and error probability [20, 21, 39, 33, 30]; when the goal is to find the intersection [24, 34, 79, 8]; in space-bounded, online, and streaming models [53, 16, 5]; and direct product theorems [54, 12, 14, 45, 51, 67, 69, 68]. We contribute one more result to this thorough assault on Set-Disjointness.

Here is the definition of our 2-party sampling model: Let D be a probability distribution over $\{0,1\}^n \times \{0,1\}^n$; we also think of D as a matrix with rows and columns both indexed by $\{0,1\}^n$ where $D_{x,y}$ is the probability of outcome (x,y). We define $\mathsf{Samp}(D)$ as the minimum communication cost of any protocol where Alice and Bob each start with private randomness and no other input, and at the end Alice outputs some $x \in \{0,1\}^n$ and Bob outputs some $y \in \{0,1\}^n$ such that (x,y) is distributed according to D. Note that $\mathsf{Samp}(D) = 0$ iff D is a product distribution (x and y are independent), and $\mathsf{Samp}(D) \le n$ for all D (since Alice can privately sample (x,y) and send y to Bob). Allowing public randomness would not make sense since Alice and Bob could read a properly-distributed (x,y) off of the randomness without communicating. We define $\mathsf{Samp}_\varepsilon(D)$ as the minimum of $\mathsf{Samp}(D')$ over all distributions D' with $\Delta(D,D') \le \varepsilon$, where Δ denotes statistical (total variation) distance, defined as

$$\Delta(D,D') \coloneqq \max_{\text{event } E} \big|\mathbb{P}_D[E] - \mathbb{P}_{D'}[E]\big| = \max_{\text{event } E} \big(\mathbb{P}_D[E] - \mathbb{P}_{D'}[E]\big) = \tfrac{1}{2} \sum_{\text{outcome } o} \big|\mathbb{P}_D[o] - \mathbb{P}_{D'}[o]\big|.$$

1.1 A story

Our story begins with [7], which proved that $\mathsf{Samp}_\varepsilon\big((D, \textsc{Disj}(D))\big) \ge \Omega(\sqrt{n})$ for some constant $\varepsilon > 0$, where D is uniform over the set of all pairs of sets of size $\sqrt{n}$ (note that this D is a product distribution and is approximately balanced between 0-inputs and 1-inputs of $\textsc{Disj}$); here it does not matter which party is responsible for outputting the bit $\textsc{Disj}(D)$. The main tool in the proof was a lemma that was originally employed in [9] to prove an $\Omega(\sqrt{n})$ bound on the randomized communication complexity of *computing* $\textsc{Disj}$. The latter bound was improved to $\Omega(n)$ via several different proofs [49, 62, 10], which leads to a natural question: Can we improve the sampling bound of [7] to $\Omega(n)$ by using the techniques of [49, 62, 10] instead of [9]?

For starters, the answer is "no" for the particular D considered in [7] – there is a trivial exact protocol with $O(\sqrt{n}\log n)$ communication since it only takes that many bits to specify a set of size $\sqrt{n}$. What about other interesting distributions D? The following illuminates the situation.

$\triangleright$ Observation 1. For any D and constants $\varepsilon > \delta > 0$, if $\mathsf{Samp}_\varepsilon\big((D, \textsc{Disj}(D))\big) \ge \omega(\sqrt{n})$ then $\mathsf{Samp}_\delta(D) \ge \Omega\big(\mathsf{Samp}_\varepsilon\big((D, \textsc{Disj}(D))\big)\big)$.

Proof. It suffices to show $\mathsf{Samp}_\varepsilon\big((D, \textsc{Disj}(D))\big) \leq \mathsf{Samp}_\delta(D) + O(\sqrt{n})$. First, note that for any sampling protocol, if we condition on a particular transcript then the output distribution becomes product (Alice and Bob are independent after they stop communicating). Second, [17] proved that for every product distribution and every constant $\gamma > 0$, there exists a deterministic protocol that uses $O(\sqrt{n})$ bits of communication and computes $\textsc{Disj}$ with error probability $\leq \gamma$ on a random input from the distribution. Now to ε-sample $(D, \textsc{Disj}(D))$, Alice and Bob can δ-sample D to obtain (x, y), and then conditioned on that sampler's transcript, they can run the average-case protocol from [17] for the corresponding product distribution with error $\varepsilon - \delta$. A simple calculation shows this indeed gives statistical distance ε. $\lhd$

The upshot is that to get an improved bound, *the hardness of sampling* $(D, \textsc{Disj}(D))$ *would come entirely from the hardness of just sampling* D. Thus such a result would not really be "about" the Set-Disjointness function, it would be about the distribution on inputs. Instead of abandoning this line of inquiry, we realize that if D itself is somehow defined in terms of $\textsc{Disj}$, then a bound for sampling D would still be saying something about the complexity of Set-Disjointness. In fact, the proof in [7] actually shows something stronger than the previously-stated result: If D is instead defined as the uniform distribution over pairs of *disjoint* sets of size $\sqrt{n}$ (which are 1-inputs of $\textsc{Disj}$), then $\mathsf{Samp}_\varepsilon(D) \geq \Omega(\sqrt{n})$. After this pivot, we are now facing a direction in which we can hope for an improvement. We prove that by removing the restriction on the sizes of the sets, the sampling problem becomes maximally hard. Our result holds for error $\varepsilon < 1$ that is exponentially close to 1, but the result is already new and interesting for constant $\varepsilon > 0$.

► **Theorem 1.** *Let U be the uniform distribution over the set of all $(x, y) \in \{0,1\}^n \times \{0,1\}^n$ with $x \wedge y = 0^n$. There exists a constant $\beta < 1$ such that $\mathsf{Samp}_{1-\beta^n}(U) = \Omega(n)$.*

The proof from [7] was a relatively short application of the technique from [9], but for Theorem 1, harnessing known techniques for proving linear communication lower bounds turns out to be more involved.

For calibration, the uniform distribution over *all* (x, y) achieves statistical distance $1 - 0.75^n$ from U since there are 4^n inputs and 3^n disjoint inputs. We can do a little better: Suppose for each coordinate independently, Alice picks 0 with probability $\sqrt{1/3}$ and picks 1 with probability $1 - \sqrt{1/3}$, and Bob does the same. This again involves no communication, and it achieves statistical distance $1 - \big(2\sqrt{1/3} - 1/3\big)^n \leq 1 - 0.82^n$ from U. Theorem 1 shows that the constant 0.82 cannot be improved arbitrarily close to 1 without a lot of communication. (In the setting of lower bounds for circuit samplers, significant effort has gone into handling statistical distances exponentially close to the maximum possible [32, 13, 76].)

1.2 Interpreting the result

We first observe that our sampling model is equivalent to two other models. One of these we call (for lack of a better word) "synthesizing" the distribution D: Alice and Bob get inputs $x, y \in \{0,1\}^n$ respectively, in addition to their private randomness, and their goal is to accept with probability exactly $D_{x,y}$. We let $\mathsf{Synth}(D)$ denote the minimum communication cost of any synthesizing protocol for D, and $\mathsf{Synth}_\varepsilon(D)$ denote the minimum of $\mathsf{Synth}(D')$ over all D' with $\Delta(D, D') \leq \varepsilon$. The other model is the nonnegative rank of a matrix: $\text{rank}_+(D)$ is defined as the minimum k for which D can be written as a sum of k many nonnegative rank-1 matrices.

▷ Observation 2. For every distribution D, the following are all within $\pm O(1)$ of each other:

$$\mathsf{Samp}(D), \quad \mathsf{Synth}(D), \quad \log \operatorname{rank}_+(D).$$

Proof. $\mathsf{Synth}(D) \leq \mathsf{Samp}(D) + 2$ since a synthesizing protocol can just run a sampling protocol and accept iff the result equals the given input (x, y).

$\log \operatorname{rank}_+(D) \leq \mathsf{Synth}(D)$ since for each transcript of a synthesizing protocol, the matrix that records the probability of getting that transcript on each particular input has rank 1; summing these matrices over all accepting transcripts yields a nonnegative rank decomposition of D.

To see that $\mathsf{Samp}(D) \leq \lceil \log \operatorname{rank}_+(D) \rceil$, suppose $D = M^{(1)} + M^{(2)} + \cdots + M^{(k)}$ is a sum of nonnegative rank-1 matrices. For each i, by scaling we can write $M^{(i)}_{x,y} = p_i\, u^{(i)}_x v^{(i)}_y$ for some distributions $u^{(i)}$ and $v^{(i)}$ over $\{0,1\}^n$, where p_i is the sum of all entries of $M^{(i)}$. Since D is a distribution, $p \coloneqq (p_1, \ldots, p_k)$ is a distribution over $[k]$. To sample from D, Alice can privately sample $i \sim p$ and send it to Bob using $\lceil \log k \rceil$ bits, then Alice can sample $x \sim u^{(i)}$ and Bob can independently sample $y \sim v^{(i)}$ with no further communication. ◁

By this characterization, Theorem 1 can be viewed as a lower bound on the approximate nonnegative rank of the DISJ matrix, where the approximation is in ℓ_1 (which has an average-case flavor). In the recent literature, "approximate nonnegative rank" generally refers to approximation in ℓ_∞ (which is a worst-case requirement), and this model is equivalent to the so-called smooth rectangle bound and WAPP communication complexity [44, 55, 37].

2 Proof

2.1 Overview

Our proof of Theorem 1 is by a black-box reduction to the well-known *corruption lemma* for Set-Disjointness due to Razborov [62]. We start with a high-level overview.

For notation: Let $|z|$ denote the Hamming weight of a string $z \in \{0,1\}^n$. For $\ell \in \mathbb{N}$, let U^ℓ be the uniform distribution over all $(x, y) \in \{0,1\}^n \times \{0,1\}^n$ with $|x \wedge y| = \ell$. Note that $U = U^0$. For a distribution D over $\{0,1\}^n \times \{0,1\}^n$ and an event $E \subseteq \{0,1\}^n \times \{0,1\}^n$, let $D_E \coloneqq \sum_{(x,y)\in E} D_{x,y}$. For a randomized protocol Π, let $\operatorname{acc}_\Pi(x, y)$ denote the probability that Π accepts (x, y).

Step I: Uniform corruption

The corruption lemma states that if a rectangle $R \subseteq \{0,1\}^n \times \{0,1\}^n$ contains a noticeable fraction of *disjoint* pairs, then it must contain about as large a fraction of *uniquely intersecting* pairs. More quantitatively, there exist a constant $C > 0$ and two distributions D^ℓ, $\ell = 0, 1$, defined over disjoint ($\ell = 0$) and uniquely intersecting pairs ($\ell = 1$) such that for every rectangle R,

$$\text{if} \quad D^0_R \geq 2^{-o(n)} \quad \text{then} \quad D^1_R \geq C \cdot D^0_R.$$

The original proof [62] defined D^ℓ as the uniform distribution over all pairs (x, y) with fixed sizes $|x| = |y| = \lceil n/4 \rceil$ and $|x \wedge y| = \ell$. For our purpose, we need the corruption lemma to hold relative to the aforementioned distributions U^ℓ, $\ell = 0, 1$, which have no restrictions on set sizes. We derive in Subsection 2.2 a corruption lemma for U^ℓ from the original lemma for D^ℓ. To do this, we exhibit a reduction that uses public randomness and no communication to transform a sample from D^ℓ into a sample from a distribution that is close to U^ℓ in a suitable sense, for $\ell = 0, 1$.

Step II: Truncate and scale

For simplicity, let us think about proving Theorem 1 for a small error $\varepsilon > 0$. Assume for contradiction there is some distribution D, $\Delta(U, D) \leq \varepsilon$, such that $\mathsf{Synth}(D) \leq o(n)$ as witnessed by a private-randomness synthesizing protocol Π' with $\mathrm{acc}_{\Pi'}(x, y) = D_{x,y}$. Note that the total acceptance probability over disjoint inputs is close to 1:

$$\textstyle\sum_{x,y\,:\,|x\wedge y|=0} \mathrm{acc}_{\Pi'}(x,y) \;\geq\; 1-\varepsilon \quad \text{and thus} \quad \mathbb{E}_{(x,y)\sim U^0}[\mathrm{acc}_{\Pi'}(x,y)] \;\geq\; (1-\varepsilon)3^{-n}.$$

Our eventual goal (in Step III) is to apply our corruption lemma to the transcript rectangles, but the above threshold $(1 - \varepsilon)3^{-n}$ is too low for this. To raise the threshold to $2^{-o(n)}$ as needed for corruption, we would like to scale up all the acceptance probabilities accordingly. To "make room" for the scaling, we first carry out a certain truncation step. Specifically, in Subsection 2.3 we transform Π' into a public-randomness protocol Π:

1. First, we **truncate** (using a *truncation lemma* [37]) the values $\mathrm{acc}_{\Pi'}(x, y)$, which has the effect of decreasing some of them, but any $\mathrm{acc}_{\Pi'}(x, y)$ that is under 3^{-n} remains approximately the same. This results in an intermediate protocol Π'' that still satisfies $\mathbb{E}_{(x,y)\sim U^0}[\mathrm{acc}_{\Pi''}(x, y))] \geq \Omega((1 - \varepsilon)3^{-n})$ (using the assumption that $\Delta(U, D) \leq \varepsilon$).
2. Second, we **scale** (using the low cost of Π'') the truncated probabilities up by a large factor $3^n 2^{-o(n)}$. This results in a protocol Π with large typical acceptance probabilities:

$$\mathbb{E}_{(x,y)\sim U^0}[\mathrm{acc}_{\Pi}(x,y)] \;\geq\; 2^{-o(n)}. \tag{1}$$

Step III: Iterate corruption

Because Π has such large acceptance probabilities (Equation 1), our corruption lemma can be applied: there is some constant $C' > 0$ such that

$$\mathbb{E}_{(x,y)\sim U^1}[\mathrm{acc}_{\Pi}(x,y)] \;\geq\; C' \cdot \mathbb{E}_{(x,y)\sim U^0}[\mathrm{acc}_{\Pi}(x,y)]. \tag{2}$$

Since Π is a truncated-and-scaled version of Π', this allows us to infer that

$$\mathbb{E}_{(x,y)\sim U^1}[\mathrm{acc}_{\Pi'}(x,y)] \;\geq\; \Omega((1-\varepsilon)3^{-n}) \quad \text{and thus} \quad \textstyle\sum_{x,y\,:\,|x\wedge y|=1} \mathrm{acc}_{\Pi'}(x,y) \;\geq\; \Omega((1-\varepsilon)n)$$

using the fact that $|\mathrm{supp}(U^1)| = n3^{n-1} = (n/3) \cdot |\mathrm{supp}(U^0)|$. Thus for $\varepsilon = 1 - \omega(1/n)$, this means Π' must have placed a total probability mass > 1 on uniquely intersecting inputs, which is the sought contradiction.

To prove Theorem 1 for very large error $\varepsilon = 1-\beta^n$, in Subsection 2.4 we iterate the above argument for U^ℓ over $0 \leq \ell \leq o(n)$. Namely, analogously to Equation 2, we show that the average acceptance probability of Π over $U^{\ell+1}$ is at least a constant times the average over U^ℓ. Meanwhile, the support sizes increase as $|\mathrm{supp}(U^{\ell+1})| \geq \omega(1) \cdot |\mathrm{supp}(U^\ell)|$ for $\ell \leq o(n)$. These facts together imply a large constant factor increase in the total probability mass that Π' places on $\mathrm{supp}(U^{\ell+1})$ as compared to $\mathrm{supp}(U^\ell)$. Starting with even a tiny probability mass over $\mathrm{supp}(U^0)$, this iteration will eventually lead to a contradiction.

2.2 Step I: Uniform corruption

The goal of this step is to derive Lemma 3 from Lemma 2.

▶ **Lemma 2** (Corruption [62]). *For every rectangle $R \subseteq \{0,1\}^n \times \{0,1\}^n$ we have $D^1_R \geq \frac{1}{45}D^0_R - 2^{-0.017n}$ where, assuming $n = 4k-1$, D^ℓ is the uniform distribution over all (x, y) with $|x| = |y| = k$ and $|x \wedge y| = \ell$.*

▶ **Lemma 3** (Uniform Corruption). *For every rectangle $R \subseteq \{0,1\}^n \times \{0,1\}^n$ we have $U^1_R \geq \frac{1}{765}U^0_R - 2^{-0.008n}$.*

Proof. Assume for convenience that $n/2$ has the form $4k-1$ (otherwise use the nearest such number instead of $n/2$ throughout). We prove that Lemma 2 for $n/2$ implies Lemma 3 for n by the contrapositive. Thus, D^0 and D^1 are distributions over $\{0,1\}^{n/2} \times \{0,1\}^{n/2}$ while U^0 and U^1 are distributions over $\{0,1\}^n \times \{0,1\}^n$. Assume there exists a rectangle $R \subseteq \{0,1\}^n \times \{0,1\}^n$ such that $U^1_R < \frac{1}{765}U^0_R - 2^{-0.008n}$. We exhibit a distribution over rectangles $Q \subseteq \{0,1\}^{n/2} \times \{0,1\}^{n/2}$ such that $\mathbb{E}[D^1_Q] < \frac{1}{45}\mathbb{E}[D^0_Q] - 2^{-0.017n/2}$; by linearity of expectation this implies that there exists such a Q with $D^1_Q < \frac{1}{45}D^0_Q - 2^{-0.017n/2}$.

To this end, we define a distribution F over functions $f\colon \{0,1\}^{n/2} \times \{0,1\}^{n/2} \to \{0,1\}^n \times \{0,1\}^n$ of the form $f(x,y) = (f_1(x), f_2(y))$ and then let Q_f be the rectangle $f^{-1}(R) \coloneqq \{(x,y) : f(x,y) \in R\}$. Let H be the distribution over $\{(v,w) \in \mathbb{N} \times \mathbb{N} : v+w \leq n\}$ obtained by sampling $(x,y) \sim U^0$ and outputting $(|x|,|y|)$; i.e., $H_{v,w} \coloneqq \frac{n!}{v!\,w!\,(n-v-w)!} \cdot 3^{-n}$. To sample $f \sim F$:

1. Sample (v,w) from H conditioned on $v \geq k$, $w \geq k$, and $v+w \leq 2k+n/2$.
2. Sample a uniformly random permutation π of $[n]$.
3. Given $(x,y) \in \{0,1\}^{n/2} \times \{0,1\}^{n/2}$, define $(x',y') \in \{0,1\}^n \times \{0,1\}^n$ by letting

$$x'_i y'_i \coloneqq \begin{cases} x_i y_i & \text{for the first } n/2 \text{ coordinates } i; \\ 10 & \text{for the next } v-k \text{ coordinates } i; \\ 01 & \text{for the next } w-k \text{ coordinates } i; \\ 00 & \text{for the remaining } n/2-(v-k)-(w-k) \geq 0 \text{ coordinates } i. \end{cases}$$

4. Let $f(x,y) \coloneqq (\pi(x'), \pi(y'))$ (i.e., permute the coordinates according to π).

For $\ell \in \{0,1\}$ let $F(D^\ell)$ denote the distribution obtained by sampling $(x,y) \sim D^\ell$ and $f \sim F$ and outputting $f(x,y)$, and note that $F(D^\ell)_R = \mathbb{E}_F[D^\ell_{Q_F}]$. Now we claim that $F(D^\ell)$ and U^ℓ are close, in the following senses:

(1) For every event E, $F(D^0)_E \geq U^0_E - 2^{-0.01n}$.

(2) For every event E, $F(D^1)_E \leq U^1_E \cdot 17$.

Using R as the event E, we have

$$\begin{aligned} F(D^1)_R &\leq U^1_R \cdot 17 \\ &< 17\big(\tfrac{1}{765}U^0_R - 2^{-0.008n}\big) \\ &\leq 17\big(\tfrac{1}{765}(F(D^0)_R + 2^{-0.01n}) - 2^{-0.008n}\big) \\ &\leq \tfrac{1}{45}F(D^0)_R - 2^{-0.017n/2} \end{aligned}$$

as desired. To see *(1)*, note that $F(D^0)$ is precisely U^0 conditioned on $v \geq k$, $w \geq k$, and $v+w \leq 2k+n/2$, and this conditioning event has probability $\geq 1-2^{-0.01n}$ by Chernoff bounds:

$$\begin{aligned} \mathbb{P}[v<k] \;=\; \mathbb{P}[w<k] &= \mathbb{P}[\mathrm{Bin}(n,1/3) < n/8+1/4] \leq 2^{-0.12n} \\ \mathbb{P}[v+w > 2k+n/2] &= \mathbb{P}[\mathrm{Bin}(n,2/3) > 3n/4+1/2] \leq 2^{-0.02n} \end{aligned}$$

Thus letting C be the complement of the conditioning event, we have $F(D^0)_E \geq U^0_{E \setminus C} \geq U^0_E - U^0_C \geq U^0_E - 2^{-0.01n}$. To see *(2)*, consider any outcome $(x,y) \in \{0,1\}^n \times \{0,1\}^n$ with $|x \wedge y| = 1$. We have $U^1_{x,y} = 1/(n3^{n-1})$. Abbreviating $a \coloneqq |x|$ and $b \coloneqq |y|$, assume $a \geq k$, $b \geq k$, and $a+b \leq 2k+n/2$ since otherwise $F(D^1)_{x,y} = 0$ and there would be nothing to prove. Henceforth consider the probability space with the randomness of D^1 and of F. Let I be the event that $F_1(D^1) \wedge F_2(D^1) = x \wedge y$, i.e., that the intersecting coordinate of $F(D^1)$ is the same as for (x,y). We have

$$F(D^1)_{x,y} \;=\; \underbrace{\mathbb{P}[I]}_{(*)} \cdot \underbrace{\mathbb{P}[v=a \text{ and } w=b]}_{(**)} \cdot \underbrace{\mathbb{P}\big[F(D^1)=(x,y) \,\big|\, I \text{ and } v=a \text{ and } w=b\big]}_{(***)}.$$

For the three terms on the right side, we have

$$(*) = \tfrac{1}{n}, \quad (**) \le H_{a,b}/(1-2^{-0.01n}) \le \tfrac{n!}{a!\,b!\,(n-a-b)!}\cdot 3^{-n}\cdot 1.01, \quad (***) = 1/\tfrac{(n-1)!}{(a-1)!\,(b-1)!\,(n-a-b+1)!}.$$

We have

$$\tfrac{n!}{a!\,b!\,(n-a-b)!} \,/\, \tfrac{(n-1)!}{(a-1)!\,(b-1)!\,(n-a-b+1)!} \;=\; \tfrac{n\cdot(n-a-b+1)}{a\cdot b} \;\le\; \tfrac{n\cdot(n-2k+1)}{k\cdot k} \;\le\; \tfrac{n\cdot(n-2n/8+1)}{(n/8)\cdot(n/8)} \;=\; (\tfrac{3}{4}+\tfrac{1}{n})\cdot 64.$$

Combining, we get

$$F(D^1)_{x,y} \,/\, U^1_{x,y} \;=\; (*)\cdot(**)\cdot(***)\cdot n3^{n-1} \;\le\; \tfrac{1.01}{3}\cdot(\tfrac{3}{4}+\tfrac{1}{n})\cdot 64 \;\le\; 17. \qquad ◀$$

2.3 Step II: Truncate and scale

The goal of this step is to construct a truncated-and-scaled protocol Π from any given low-cost Π' that synthesizes a distribution close to U.

For a nonnegative matrix M, we define its *truncation* $\overline{M}$ to be the same matrix but where each entry > 1 is replaced with 1.

▶ **Lemma 4** (Truncation Lemma [37]). *For every $2^n \times 2^n$ nonnegative rank-1 matrix M and every d there exists a $O(d + \log n)$-communication public-randomness protocol Π such that for every (x, y) we have $\mathrm{acc}_\Pi(x, y) \in \overline{M}_{x,y} \pm 2^{-d}$.*

Let $c \ge 1$ be the hidden constant in the big O in Lemma 4, and let $\delta := 0.00005/c$. Toward proving Theorem 1, suppose for contradiction $\mathsf{Samp}(D) \le \delta n$ for some distribution D with $\Delta(U, D) \le 1-2^{-\delta n}$ (so $\beta := 2^{-\delta}$ in Theorem 1) and thus $\sum_{x,y\,:\,|x\wedge y|=0} \min(3^{-n}, D_{x,y}) \ge 2^{-\delta n}$. By Observation 2, $\mathsf{Synth}(D) \le \delta n + 2$, so consider a synthesizing protocol Π' for D with communication cost $\le \delta n + 2$. Let A be the set of all accepting transcripts of Π'. For each $\tau \in A$ let N^τ be the nonnegative rank-1 matrix such that $N^\tau_{x,y}$ is the probability Π' generates τ on input (x, y); thus $D_{x,y} = \sum_{\tau\in A} N^\tau_{x,y}$. Let Π^τ be the public-randomness protocol from Lemma 4 applied to $M^\tau := 3^n N^\tau$ and $d := 15\delta n$. Let Π be the public-randomness protocol that picks a uniformly random $\tau \in A$ and then runs Π^τ. The communication cost of Π is $\le c\cdot(d + \log n) \le 0.001n$.

▷ Claim 5. For every input (x, y) we have $\frac{3^n}{|A|}\min(3^{-n}, D_{x,y}) - 2^{-d} \le \mathrm{acc}_\Pi(x, y) \le \frac{3^n}{|A|} D_{x,y} + 2^{-d}$.

Proof. We have

$$\begin{aligned}
\mathrm{acc}_\Pi(x, y) &= \tfrac{1}{|A|}\textstyle\sum_{\tau\in A} \mathrm{acc}_{\Pi^\tau}(x, y)\\
&\in \tfrac{1}{|A|}\textstyle\sum_{\tau\in A}(\overline{M}^\tau_{x,y} \pm 2^{-d})\\
&\subseteq \tfrac{1}{|A|}\textstyle\sum_{\tau\in A} \min(1, 3^n N^\tau_{x,y}) \pm 2^{-d}\\
&= \tfrac{3^n}{|A|}\textstyle\sum_{\tau\in A} \min(3^{-n}, N^\tau_{x,y}) \pm 2^{-d}.
\end{aligned}$$

From this it follows that:

$$\begin{aligned}
\mathrm{acc}_\Pi(x, y) &\ge \tfrac{3^n}{|A|}\min\big(3^{-n}, \textstyle\sum_{\tau\in A} N^\tau_{x,y}\big) - 2^{-d} &&= \tfrac{3^n}{|A|}\min(3^{-n}, D_{x,y}) - 2^{-d}\\
\mathrm{acc}_\Pi(x, y) &\le \tfrac{3^n}{|A|}\textstyle\sum_{\tau\in A} N^\tau_{x,y} + 2^{-d} &&= \tfrac{3^n}{|A|} D_{x,y} + 2^{-d}. \qquad ◁
\end{aligned}$$

We can now formally state the large typical acceptance probability property (Equation 1 from the overview): writing $U_\Pi := \mathbb{E}_{(x,y)\sim U}[\mathrm{acc}_\Pi(x,y)]$ (and similarly for other input distributions),

$$\begin{aligned} U_\Pi &\geq \tfrac{1}{3^n}\textstyle\sum_{x,y\,:\,|x\wedge y|=0}\left(\tfrac{3^n}{|A|}\min(3^{-n},D_{x,y})-2^{-d}\right) && \text{(by Claim 5)}\\ &= \tfrac{1}{|A|}\textstyle\sum_{x,y\,:\,|x\wedge y|=0}\min(3^{-n},D_{x,y})-2^{-d}\\ &\geq \tfrac{1}{|A|}2^{-\delta n}-2^{-15\delta n}\\ &\geq \tfrac{1}{|A|}2^{-\delta n-1} && (3)\end{aligned}$$

where the last line follows because $|A|\leq 2^{\delta n+2}$ and $2^{-2\delta n-2}$ is at least twice $2^{-15\delta n}$.

2.4 Step III: Iterate corruption

Here we derive the final contradiction: Π' places an acceptance probability mass exceeding 1 on $\mathrm{supp}(U^{\delta n})$. This is achieved by iterating our corruption lemma, starting with Equation 3 as the base case.

For $z\in\{0,1\}^n$ let U^z be the uniform distribution over all $(x,y)\in\{0,1\}^n\times\{0,1\}^n$ with $x\wedge y=z$ (so U^ℓ is the uniform mixture of all U^z with $|z|=\ell$; in particular, $U^0=U^{0^n}$), and if $|z|<n$ then let $\widehat{U}^z$ be the uniform mixture of $U^{z'}$ over all z' that can be obtained from z by flipping a single 0 to 1 (so $U^{\ell+1}$ is the uniform mixture of all $\widehat{U}^z$ with $|z|=\ell$; in particular, $U^1=\widehat{U}^{0^n}$).

▷ Claim 6. For every $z\in\{0,1\}^n$ with $|z|\leq n/2$ we have $\widehat{U}^z_\Pi\geq\frac{1}{765}U^z_\Pi-2^{-0.003n}$.

Proof. Since all relevant inputs (x,y) have $x_iy_i=11$ for all i such that $z_i=1$, we can ignore those coordinates and think of $\widehat{U}^z$ and U^z as U^1 and U^0 respectively, but defined on the remaining $n-|z|\geq n/2$ coordinates (instead of on all n coordinates). Thus by Lemma 3, for every outcome of the public randomness of Π and every accepting transcript, say corresponding to rectangle R, we have $\widehat{U}^z_R\geq\frac{1}{765}U^z_R-2^{-0.008n/2}$. Summing over all the (at most $2^{0.001n}$ many) accepting transcripts, and then taking the expectation over the public randomness, yields the claim since $2^{0.001n}\cdot 2^{-0.008n/2}\leq 2^{-0.003n}$. ◁

▷ Claim 7. For every $\ell=0,\ldots,\delta n$ we have $U^\ell_\Pi\geq\frac{1}{|A|}2^{-\delta n-1-11\ell}$.

Proof. We prove this by induction on ℓ. The base case $\ell=0$ is Equation 3. For the inductive step, assume the claim is true for ℓ. Since $U^{\ell+1}$ and U^ℓ are the uniform mixtures of $\widehat{U}^z$ and U^z respectively over all z with $|z|=\ell$ (so $U^{\ell+1}_\Pi=\mathbb{E}_z[\widehat{U}^z_\Pi]$ and $U^\ell_\Pi=\mathbb{E}_z[U^z_\Pi]$), by linearity of expectation Claim 6 implies

$$U^{\ell+1}_\Pi \;\geq\; \tfrac{1}{765}U^\ell_\Pi-2^{-0.003n} \;\geq\; \tfrac{1}{|A|}2^{-\delta n-1-11\ell-\log_2(765)}-2^{-0.003n} \;\geq\; \tfrac{1}{|A|}2^{-\delta n-1-11(\ell+1)}$$

where the last inequality follows because $|A|\leq 2^{\delta n+2}$ and $2^{-\delta n-2-\delta n-1-11\delta n-\log_2(765)}\geq 2^{-14\delta n}$ is at least twice $2^{-0.003n}$, which gives $U^{\ell+1}_\Pi\geq\frac{1}{|A|}2^{-\delta n-1-11\ell-\log_2(765)-1}$, and $\log_2(765)+1\leq 11$. ◁

Choosing $\ell=\delta n$ we have

$$U^\ell_\Pi-2^{-d} \;\geq\; \tfrac{1}{|A|}2^{-\delta n-1-11\ell}-2^{-15\delta n} \;\geq\; \tfrac{1}{|A|}2^{-\delta n-2-11\ell} \qquad (4)$$

because $|A|\leq 2^{\delta n+2}$ and $2^{-\delta n-2-\delta n-1-11\delta n}\geq 2^{-14\delta n}$ is at least twice $2^{-15\delta n}$. Thus, for $\ell=\delta n$,

$$
\begin{aligned}
\textstyle\sum_{x,y} D_{x,y} &\geq \textstyle\sum_{x,y\,:\,|x\wedge y|=\ell} D_{x,y} \\
&\geq \textstyle\sum_{x,y\,:\,|x\wedge y|=\ell} \frac{|A|}{3^n}(\mathrm{acc}_\Pi(x,y) - 2^{-d}) && \text{(by Claim 5)} \\
&= \tfrac{|A|}{3^n}\tbinom{n}{\ell}3^{n-\ell}(U^\ell_\Pi - 2^{-d}) \\
&\geq \tfrac{|A|}{3^n}(\tfrac{n}{\ell})^\ell 3^{n-\ell}\tfrac{1}{|A|}2^{-\delta n-2-11\ell} && \text{(using Equation 4)} \\
&= (\tfrac{n}{\ell\cdot 3\cdot 2^{11}})^\ell 2^{-\delta n-2} \\
&= (\tfrac{1}{\delta\cdot 3\cdot 2^{11}\cdot 2})^{\delta n}/4 \\
&\geq 1.6^{\delta n} \\
&> 1,
\end{aligned}
$$

contradicting the fact that D is a distribution.

References

1 Scott Aaronson. The Equivalence of Sampling and Searching. *Theory of Computing Systems*, 55(2):281–298, 2014. doi:10.1007/s00224-013-9527-3.

2 Scott Aaronson and Andris Ambainis. Quantum Search of Spatial Regions. *Theory of Computing*, 1(1):47–79, 2005. doi:10.4086/toc.2005.v001a004.

3 Scott Aaronson and Avi Wigderson. Algebrization: A New Barrier in Complexity Theory. *ACM Transactions on Computation Theory*, 1(1):2:1–2:54, 2009. doi:10.1145/1490270.1490272.

4 Amir Abboud, Aviad Rubinstein, and Ryan Williams. Distributed PCP Theorems for Hardness of Approximation in P. In *Proceedings of the 58th Symposium on Foundations of Computer Science (FOCS)*, pages 25–36. IEEE, 2017. doi:10.1109/FOCS.2017.12.

5 Josh Alman, Joshua Wang, and Huacheng Yu. Cell-Probe Lower Bounds from Online Communication Complexity. In *Proceedings of the 50th Symposium on Theory of Computing (STOC)*, pages 1003–1012. ACM, 2018. doi:10.1145/3188745.3188862.

6 Noga Alon, Yossi Matias, and Mario Szegedy. The Space Complexity of Approximating the Frequency Moments. *Journal of Computer and System Sciences*, 58(1):137–147, 1999. doi:10.1006/jcss.1997.1545.

7 Andris Ambainis, Leonard Schulman, Amnon Ta-Shma, Umesh Vazirani, and Avi Wigderson. The Quantum Communication Complexity of Sampling. *SIAM Journal on Computing*, 32(6):1570–1585, 2003. doi:10.1137/S009753979935476.

8 Sepehr Assadi, Yu Chen, and Sanjeev Khanna. Polynomial Pass Lower Bounds for Graph Streaming Algorithms. In *Proceedings of the 51st Symposium on Theory of Computing (STOC)*, pages 265–276. ACM, 2019. doi:10.1145/3313276.3316361.

9 László Babai, Peter Frankl, and Janos Simon. Complexity Classes in Communication Complexity Theory. In *Proceedings of the 27th Symposium on Foundations of Computer Science (FOCS)*, pages 337–347. IEEE, 1986. doi:10.1109/SFCS.1986.15.

10 Ziv Bar-Yossef, T.S. Jayram, Ravi Kumar, and D. Sivakumar. An Information Statistics Approach to Data Stream and Communication Complexity. *Journal of Computer and System Sciences*, 68(4):702–732, 2004. doi:10.1016/j.jcss.2003.11.006.

11 Paul Beame and Dang-Trinh Huynh-Ngoc. Multiparty Communication Complexity and Threshold Circuit Size of AC^0. In *Proceedings of the 50th Symposium on Foundations of Computer Science (FOCS)*, pages 53–62. IEEE, 2009. doi:10.1109/FOCS.2009.12.

12 Paul Beame, Toniann Pitassi, Nathan Segerlind, and Avi Wigderson. A Strong Direct Product Theorem for Corruption and the Multiparty Communication Complexity of Disjointness. *Computational Complexity*, 15(4):391–432, 2006. doi:10.1007/s00037-007-0220-2.

13 Christopher Beck, Russell Impagliazzo, and Shachar Lovett. Large Deviation Bounds for Decision Trees and Sampling Lower Bounds for AC^0-Circuits. In *Proceedings of the 53rd Symposium on Foundations of Computer Science (FOCS)*, pages 101–110. IEEE, 2012. doi:10.1109/FOCS.2012.82.

14 Avraham Ben-Aroya, Oded Regev, and Ronald de Wolf. A Hypercontractive Inequality for Matrix-Valued Functions with Applications to Quantum Computing and LDCs. In *Proceedings of the 49th Symposium on Foundations of Computer Science (FOCS)*, pages 477–486. IEEE, 2008. doi:10.1109/FOCS.2008.45.

15 Itai Benjamini, Gil Cohen, and Igor Shinkar. Bi-Lipschitz Bijection Between the Boolean Cube and the Hamming Ball. In *Proceedings of the 55th Symposium on Foundations of Computer Science (FOCS)*, pages 81–89. IEEE, 2014. doi:10.1109/FOCS.2014.17.

16 Lucas Boczkowski, Iordanis Kerenidis, and Frédéric Magniez. Streaming Communication Protocols. *ACM Transactions on Computation Theory*, 10(4):19:1–19:21, 2018. doi:10.1145/3276748.

17 Ralph Bottesch, Dmitry Gavinsky, and Hartmut Klauck. Correlation in Hard Distributions in Communication Complexity. In *Proceedings of the 19th International Workshop on Randomization and Computation (RANDOM)*, pages 544–572. Schloss Dagstuhl, 2015. doi:10.4230/LIPIcs.APPROX-RANDOM.2015.544.

18 Mark Braverman, Faith Ellen, Rotem Oshman, Toniann Pitassi, and Vinod Vaikuntanathan. A Tight Bound for Set Disjointness in the Message-Passing Model. In *Proceedings of the 54th Symposium on Foundations of Computer Science (FOCS)*, pages 668–677. IEEE, 2013. doi:10.1109/FOCS.2013.77.

19 Mark Braverman, Ankit Garg, Young Kun-Ko, Jieming Mao, and Dave Touchette. Near-Optimal Bounds on the Bounded-Round Quantum Communication Complexity of Disjointness. *SIAM Journal on Computing*, 47(6):2277–2314, 2018. doi:10.1137/16M1061400.

20 Mark Braverman, Ankit Garg, Denis Pankratov, and Omri Weinstein. From Information to Exact Communication. In *Proceedings of the 45th Symposium on Theory of Computing (STOC)*, pages 151–160. ACM, 2013. doi:10.1145/2488608.2488628.

21 Mark Braverman and Ankur Moitra. An Information Complexity Approach to Extended Formulations. In *Proceedings of the 45th Symposium on Theory of Computing (STOC)*, pages 161–170. ACM, 2013. doi:10.1145/2488608.2488629.

22 Mark Braverman and Rotem Oshman. On Information Complexity in the Broadcast Model. In *Proceedings of the 34th Symposium on Principles of Distributed Computing (PODC)*, pages 355–364. ACM, 2015. doi:10.1145/2767386.2767425.

23 Mark Braverman and Rotem Oshman. A Rounds vs. Communication Tradeoff for Multi-Party Set Disjointness. In *Proceedings of the 58th Symposium on Foundations of Computer Science (FOCS)*, pages 144–155. IEEE, 2017. doi:10.1109/FOCS.2017.22.

24 Joshua Brody, Amit Chakrabarti, Ranganath Kondapally, David Woodruff, and Grigory Yaroslavtsev. Beyond Set Disjointness: The Communication Complexity of Finding the Intersection. In *Proceedings of the 33rd Symposium on Principles of Distributed Computing (PODC)*, pages 106–113. ACM, 2014. doi:10.1145/2611462.2611501.

25 Harry Buhrman, Richard Cleve, and Avi Wigderson. Quantum vs. Classical Communication and Computation. In *Proceedings of the 30th Symposium on Theory of Computing (STOC)*, pages 63–68. ACM, 1998. doi:10.1145/276698.276713.

26 Harry Buhrman, David Garcia-Soriano, Arie Matsliah, and Ronald de Wolf. The Non-Adaptive Query Complexity of Testing k-Parities. *Chicago Journal of Theoretical Computer Science*, 2013(6):1–11, 2013. doi:10.4086/cjtcs.2013.006.

27 Amit Chakrabarti, Subhash Khot, and Xiaodong Sun. Near-Optimal Lower Bounds on the Multi-Party Communication Complexity of Set Disjointness. In *Proceedings of the 18th Conference on Computational Complexity*, pages 107–117. IEEE, 2003. doi:10.1109/CCC.2003.1214414.

28 Arkadev Chattopadhyay and Anil Ada. Multiparty Communication Complexity of Disjointness. Technical Report TR08-002, Electronic Colloquium on Computational Complexity (ECCC), 2008. URL: `https://eccc.weizmann.ac.il//eccc-reports/2008/TR08-002/`.

29 Lijie Chen. On The Hardness of Approximate and Exact (Bichromatic) Maximum Inner Product. In *Proceedings of the 33rd Computational Complexity Conference (CCC)*, pages 14:1–14:45. Schloss Dagstuhl, 2018. `doi:10.4230/LIPIcs.CCC.2018.14`.

30 Yuval Dagan, Yuval Filmus, Hamed Hatami, and Yaqiao Li. Trading Information Complexity for Error. *Theory of Computing*, 14(1):1–73, 2018. `doi:10.4086/toc.2018.v014a006`.

31 Anirban Dasgupta, Ravi Kumar, and D. Sivakumar. Sparse and Lopsided Set Disjointness via Information Theory. In *Proceedings of the 16th International Workshop on Randomization and Computation (RANDOM)*, pages 517–528. Springer, 2012. `doi:10.1007/978-3-642-32512-0_44`.

32 Anindya De and Thomas Watson. Extractors and Lower Bounds for Locally Samplable Sources. *ACM Transactions on Computation Theory*, 4(1):3:1–3:21, 2012. `doi:10.1145/2141938.2141941`.

33 Yuval Filmus, Hamed Hatami, Yaqiao Li, and Suzin You. Information Complexity of the AND Function in the Two-Party and Multi-Party Settings. In *Proceedings of the 23rd International Computing and Combinatorics Conference (COCOON)*, pages 200–211. Springer, 2017. `doi:10.1007/978-3-319-62389-4_17`.

34 Dmitry Gavinsky. Communication Complexity of Inevitable Intersection. Technical Report abs/1611.08842, arXiv, 2016. `arXiv:1611.08842`.

35 Dmitry Gavinsky and Alexander Sherstov. A Separation of NP and coNP in Multiparty Communication Complexity. *Theory of Computing*, 6(1):227–245, 2010. `doi:10.4086/toc.2010.v006a010`.

36 Oded Goldreich, Shafi Goldwasser, and Asaf Nussboim. On the Implementation of Huge Random Objects. *SIAM Journal on Computing*, 39(7):2761–2822, 2010. `doi:10.1137/080722771`.

37 Mika Göös, Shachar Lovett, Raghu Meka, Thomas Watson, and David Zuckerman. Rectangles Are Nonnegative Juntas. *SIAM Journal on Computing*, 45(5):1835–1869, 2016. `doi:10.1137/15M103145X`.

38 Mika Göös, Toniann Pitassi, and Thomas Watson. Zero-Information Protocols and Unambiguity in Arthur–Merlin Communication. *Algorithmica*, 76(3):684–719, 2016. `doi:10.1007/s00453-015-0104-9`.

39 Mika Göös and Thomas Watson. Communication Complexity of Set-Disjointness for All Probabilities. *Theory of Computing*, 12(9):1–23, 2016. `doi:10.4086/toc.2016.v012a009`.

40 Vince Grolmusz. The BNS Lower Bound for Multi-Party Protocols Is Nearly Optimal. *Information and Computation*, 112(1):51–54, 1994. `doi:10.1006/inco.1994.1051`.

41 André Gronemeier. Asymptotically Optimal Lower Bounds on the NIH-Multi-Party Information Complexity of the AND-Function and Disjointness. In *Proceedings of the 26th International Symposium on Theoretical Aspects of Computer Science (STACS)*, pages 505–516. Schloss Dagstuhl, 2009. `doi:10.4230/LIPIcs.STACS.2009.1846`.

42 Johan Håstad and Avi Wigderson. The Randomized Communication Complexity of Set Disjointness. *Theory of Computing*, 3(1):211–219, 2007. `doi:10.4086/toc.2007.v003a011`.

43 Peter Høyer and Ronald de Wolf. Improved Quantum Communication Complexity Bounds for Disjointness and Equality. In *Proceedings of the 19th Symposium on Theoretical Aspects of Computer Science (STACS)*, pages 299–310. Springer, 2002. `doi:10.1007/3-540-45841-7_24`.

44 Rahul Jain and Hartmut Klauck. The Partition Bound for Classical Communication Complexity and Query Complexity. In *Proceedings of the 25th Conference on Computational Complexity (CCC)*, pages 247–258. IEEE, 2010. `doi:10.1109/CCC.2010.31`.

45 Rahul Jain, Hartmut Klauck, and Ashwin Nayak. Direct Product Theorems for Classical Communication Complexity via Subdistribution Bounds. In *Proceedings of the 40th Symposium on Theory of Computing (STOC)*, pages 599–608. ACM, 2008. `doi:10.1145/1374376.1374462`.

46 Rahul Jain, Jaikumar Radhakrishnan, and Pranab Sen. A Lower Bound for the Bounded Round Quantum Communication Complexity of Set Disjointness. In *Proceedings of the 44th Symposium on Foundations of Computer Science (FOCS)*, pages 220–229. IEEE, 2003. doi:10.1109/SFCS.2003.1238196.

47 Rahul Jain, Yaoyun Shi, Zhaohui Wei, and Shengyu Zhang. Efficient Protocols for Generating Bipartite Classical Distributions and Quantum States. *IEEE Transactions on Information Theory*, 59(8):5171–5178, 2013. doi:10.1109/TIT.2013.2258372.

48 T.S. Jayram. Hellinger strikes back: A note on the multi-party information complexity of AND. In *Proceedings of the 13th International Workshop on Randomization and Computation (RANDOM)*, pages 562–573. Springer, 2009. doi:10.1007/978-3-642-03685-9_42.

49 Bala Kalyanasundaram and Georg Schnitger. The Probabilistic Communication Complexity of Set Intersection. *SIAM Journal on Discrete Mathematics*, 5(4):545–557, 1992. doi:10.1137/0405044.

50 Hartmut Klauck. Rectangle Size Bounds and Threshold Covers in Communication Complexity. In *Proceedings of the 18th Conference on Computational Complexity (CCC)*, pages 118–134. IEEE, 2003. doi:10.1109/CCC.2003.1214415.

51 Hartmut Klauck. A Strong Direct Product Theorem for Disjointness. In *Proceedings of the 42nd Symposium on Theory of Computing (STOC)*, pages 77–86. ACM, 2010. doi:10.1145/1806689.1806702.

52 Hartmut Klauck, Ashwin Nayak, Amnon Ta-Shma, and David Zuckerman. Interaction in Quantum Communication. *IEEE Transactions on Information Theory*, 53(6):1970–1982, 2007. doi:10.1109/TIT.2007.896888.

53 Hartmut Klauck and Supartha Podder. New Bounds for the Garden-Hose Model. In *Proceedings of the 34th International Conference on Foundation of Software Technology and Theoretical Computer Science (FSTTCS)*, pages 481–492. Schloss Dagstuhl, 2014. doi:10.4230/LIPIcs.FSTTCS.2014.481.

54 Hartmut Klauck, Robert Spalek, and Ronald de Wolf. Quantum and Classical Strong Direct Product Theorems and Optimal Time-Space Tradeoffs. *SIAM Journal on Computing*, 36(5):1472–1493, 2007. doi:10.1137/05063235X.

55 Gillat Kol, Shay Moran, Amir Shpilka, and Amir Yehudayoff. Approximate Nonnegative Rank Is Equivalent to the Smooth Rectangle Bound. *Computational Complexity*, 28(1):1–25, 2019. doi:10.1007/s00037-018-0176-4.

56 Eyal Kushilevitz and Enav Weinreb. The Communication Complexity of Set-Disjointness with Small Sets and 0-1 Intersection. In *Proceedings of the 50th Symposium on Foundations of Computer Science (FOCS)*, pages 63–72. IEEE, 2009. doi:10.1109/FOCS.2009.15.

57 Troy Lee and Adi Shraibman. Disjointness is Hard in the Multiparty Number-on-the-Forehead Model. *Computational Complexity*, 18(2):309–336, 2009. doi:10.1007/s00037-009-0276-2.

58 Shachar Lovett and Emanuele Viola. Bounded-Depth Circuits Cannot Sample Good Codes. *Computational Complexity*, 21(2):245–266, 2012. doi:10.1007/s00037-012-0039-3.

59 Mihai Patrascu. Unifying the Landscape of Cell-Probe Lower Bounds. *SIAM Journal on Computing*, 40(3):827–847, 2011. doi:10.1137/09075336X.

60 Vladimir Podolskii and Alexander Sherstov. Inner Product and Set Disjointness: Beyond Logarithmically Many Parties. Technical Report abs/1711.10661, arXiv, 2017. arXiv:1711.10661.

61 Anup Rao and Amir Yehudayoff. Simplified Lower Bounds on the Multiparty Communication Complexity of Disjointness. In *Proceedings of the 30th Computational Complexity Conference (CCC)*, pages 88–101. Schloss Dagstuhl, 2015. doi:10.4230/LIPIcs.CCC.2015.88.

62 Alexander Razborov. On the Distributional Complexity of Disjointness. *Theoretical Computer Science*, 106(2):385–390, 1992. doi:10.1016/0304-3975(92)90260-M.

63 Alexander Razborov. Quantum Communication Complexity of Symmetric Predicates. *Izvestiya: Mathematics*, 67(1):145–159, 2003. doi:10.1070/IM2003v067n01ABEH000422.

64 Aviad Rubinstein. Hardness of Approximate Nearest Neighbor Search. In *Proceedings of the 50th Symposium on Theory of Computing (STOC)*, pages 1260–1268. ACM, 2018. doi:10.1145/3188745.3188916.

65 Mert Saglam and Gábor Tardos. On the Communication Complexity of Sparse Set Disjointness and Exists-Equal Problems. In *Proceedings of the 54th Symposium on Foundations of Computer Science (FOCS)*, pages 678–687. IEEE, 2013. doi:10.1109/FOCS.2013.78.

66 Alexander Sherstov. The Pattern Matrix Method. *SIAM Journal on Computing*, 40(6):1969–2000, 2011. doi:10.1137/080733644.

67 Alexander Sherstov. Strong Direct Product Theorems for Quantum Communication and Query Complexity. *SIAM Journal on Computing*, 41(5):1122–1165, 2012. doi:10.1137/110842661.

68 Alexander Sherstov. Communication Lower Bounds Using Directional Derivatives. *Journal of the ACM*, 61(6):1–71, 2014. doi:10.1145/2629334.

69 Alexander Sherstov. The Multiparty Communication Complexity of Set Disjointness. *SIAM Journal on Computing*, 45(4):1450–1489, 2016. doi:10.1137/120891587.

70 Yaoyun Shi and Yufan Zhu. Quantum Communication Complexity of Block-Composed Functions. *Quantum Information and Computation*, 9(5–6):444–460, 2009.

71 Pascal Tesson. *Computational Complexity Questions Related to Finite Monoids and Semigroups*. PhD thesis, McGill University, 2003.

72 Emanuele Viola. Extractors for Turing-Machine Sources. In *Proceedings of the 16th International Workshop on Randomization and Computation (RANDOM)*, pages 663–671. Springer, 2012. doi:10.1007/978-3-642-32512-0_56.

73 Emanuele Viola. The Complexity of Distributions. *SIAM Journal on Computing*, 41(1):191–218, 2012. doi:10.1137/100814998.

74 Emanuele Viola. Extractors for Circuit Sources. *SIAM Journal on Computing*, 43(2):655–672, 2014. doi:10.1137/11085983X.

75 Emanuele Viola. Quadratic Maps Are Hard to Sample. *ACM Transactions on Computation Theory*, 8(4):18:1–18:4, 2016. doi:10.1145/2934308.

76 Emanuele Viola. Sampling Lower Bounds: Boolean Average-Case and Permutations. Technical Report TR18-060, Electronic Colloquium on Computational Complexity (ECCC), 2018. URL: https://eccc.weizmann.ac.il/report/2018/060.

77 Thomas Watson. Time Hierarchies for Sampling Distributions. *SIAM Journal on Computing*, 43(5):1709–1727, 2014. doi:10.1137/120898553.

78 Thomas Watson. Nonnegative Rank vs. Binary Rank. *Chicago Journal of Theoretical Computer Science*, 2016(2):1–13, 2016. doi:10.4086/cjtcs.2016.002.

79 Thomas Watson. Communication Complexity with Small Advantage. In *Proceedings of the 33rd Computational Complexity Conference (CCC)*, pages 9:1–9:17. Schloss Dagstuhl, 2018. doi:10.4230/LIPIcs.CCC.2018.9.

80 Omri Weinstein and David Woodruff. The Simultaneous Communication of Disjointness with Applications to Data Streams. In *Proceedings of the 42nd International Colloquium on Automata, Languages, and Programming (ICALP)*, pages 1082–1093. Springer, 2015. doi:10.1007/978-3-662-47672-7_88.