THEORY OF COMPUTING, Volume 12 (9), 2016, pp. 1–23
www.theoryofcomputing.org

SPECIAL ISSUE: APPROX-RANDOM 2014

Communication Complexity of Set-Disjointness for All Probabilities

Mika Göös Thomas Watson

Received December 21, 2014; Revised June 8, 2015; Published August 15, 2016

Abstract: We study SET-DISJOINTNESS in a generalized model of randomized two-party communication where the probability of acceptance must be at least $\alpha(n)$ on yes-inputs and at most $\beta(n)$ on no-inputs, for some functions $\alpha(n) > \beta(n)$. Our main result is a complete characterization of the private-coin communication complexity of SET-DISJOINTNESS for all functions α and β, and a near-complete characterization for public-coin protocols. In particular, we obtain a simple proof of a theorem of Braverman and Moitra (STOC 2013), who studied the case where $\alpha = 1/2 + \varepsilon(n)$ and $\beta = 1/2 - \varepsilon(n)$. The following contributions play a crucial role in our characterization and are interesting in their own right.

(1) We introduce two communication analogues of the classical complexity class that captures *small bounded-error* computations: we define a "restricted" class SBP (which lies between MA and AM) and an "unrestricted" class USBP. The distinction between them is analogous to the distinction between the well-known communication classes PP and UPP.

(2) We show that the SBP communication complexity is precisely captured by the classical *corruption* lower bound method. This sharpens a theorem of Klauck (CCC 2003).

A preliminary version of this paper appeared in the Proceedings of the 18th International Workshop on Randomization and Computation (RANDOM), 2014 [22].

ACM Classification: F.1.3

AMS Classification: 68Q17, 68Q15

Key words and phrases: communication complexity, set-disjointness, complexity classes

 DOI: 10.4086/toc.2016.v012a009

(3) We use information complexity arguments to prove a linear lower bound on the USBP complexity of SET-DISJOINTNESS.

1 Introduction

In the SET-DISJOINTNESS problem, Alice is given an $x \subseteq [n]$, Bob is given a $y \subseteq [n]$, and their task is to decide whether $x \cap y = \emptyset$. Equivalently, viewing x and y as binary strings, we define

$$\textsc{Disj}(x,y) := \neg \bigvee_{i \in [n]} (x_i \wedge y_i)\,.$$

Set-disjointness is the preeminent coNP-complete problem in communication complexity [2, 15]. A fundamental result of Kalyanasundaram and Schnitger [27] (with alternative proofs given by [35, 4]) states that every randomized protocol for SET-DISJOINTNESS requires $\Omega(n)$ bits of communication to achieve a constant error probability that is bounded away from $1/2$. These lower bounds have been extremely useful in applications of communication complexity to other areas of theoretical computer science, including circuit complexity, distributed computing, streaming, data structures, combinatorial optimization, and more; see [32, 26, 15].

In this work, we study SET-DISJOINTNESS in a generalized setting where the probability of acceptance must be at least $\alpha(n)$ on yes-inputs and at most $\beta(n)$ on no-inputs, for any prescribed functions $\alpha(n) > \beta(n)$.

1.1 Main result

Our main result is a complete characterization of the private-coin communication complexity of SET-DISJOINTNESS for all functions α and β, and a near-complete characterization for public-coin protocols. Roughly speaking, we prove that the randomized complexity is

$$\Theta(n \cdot (1 - \beta/\alpha))$$

for typical functions α and β; see Section 1.4 for the statement of the exact bounds.

As a special case, we obtain a simple proof of a result of Braverman and Moitra [8]. They showed that the communication complexity of SET-DISJOINTNESS is $\Theta(\varepsilon n)$ in case $\alpha = 1/2 + \varepsilon(n)$ and $\beta = 1/2 - \varepsilon(n)$. While this special case might suggest that the complexity is determined by the additive gap $\alpha - \beta$, our characterization reveals that, in fact:

> **Central tenet:** *It is not the additive gap between α and β that determines the complexity of* SET-DISJOINTNESS*; what matters is the multiplicative gap.*

Our proof follows this ideology: we show that in order to understand the communication complexity for all α and β it suffices to understand the *small bounded-error* case where α is tiny (e. g., exponentially small in n) and $\beta = \alpha/2$. The basic reason is because a protocol's multiplicative gap between α and β can be efficiently amplified at the expense of decreasing α, and thus upper bounds for the general case yield upper bounds for the small bounded-error case.

1.2 SBP: Small bounded-error probabilities

In classical time-bounded (i. e., poly-time Turing machine) complexity theory, small bounded-error acceptance probabilities are captured by a counting class called SBP, which was introduced by Böhler, Glaßer, and Meister [5] and has also been studied in [41]. In particular, [5] observed that SBP is sandwiched between the Arthur–Merlin classes MA and AM [3].

In this work, we introduce two communication complexity analogues of SBP: a *restricted* class called SBP, and an *unrestricted* class called USBP. These classes are natural and interesting in their own right. Most importantly, they serve to structure our argument.

Randomized communication complexity. In what follows, we assume familiarity with basic definitions of communication complexity [32, 26]. Fix a two-party function $f\colon \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$ where on input (x,y) Alice is given x and Bob is given y. We say (x,y) is a b-input if $(x,y) \in f^{-1}(b)$. We let $\mathsf{R}^{\mathsf{pub}}_{\alpha,\beta}(f)$, respectively $\mathsf{R}^{\mathsf{priv}}_{\alpha,\beta}(f)$, denote the minimum communication complexity (as a function of n) of a public-coin, respectively private-coin, protocol for f such that the probability of acceptance is at least $\alpha(n)$ on all 1-inputs and at most $\beta(n)$ on all 0-inputs. As is customary [2], for any communication measure $\mathsf{C}(f)$ we often let C stand for the class of functions f with $\mathsf{C}(f) = \mathrm{polylog}(n)$.

PP and UPP. To motivate our upcoming definitions for SBP, we take a little detour and recall the communication classes associated with the standard complexity class PP. There are in fact two distinct measures—*restricted* and *unrestricted*—as introduced in [2, 34]:

$$\mathsf{PP}(f) \coloneqq \min_{\varepsilon(n)>0} \mathsf{R}^{\mathsf{pub}}_{1/2+\varepsilon,1/2-\varepsilon}(f) + \log(1/\varepsilon)\,,$$
$$\mathsf{UPP}(f) \coloneqq \min_{\varepsilon(n)>0} \mathsf{R}^{\mathsf{priv}}_{1/2+\varepsilon,1/2-\varepsilon}(f)\,.$$

In the (restricted) public-coin model, one needs to charge the additional $\log(1/\varepsilon)$ term in order for the measure to be well-behaved when ε is tiny. (For example, note that $\mathsf{R}^{\mathsf{pub}}_{1/2+\varepsilon,1/2-\varepsilon}(f) \le 2$ for $\varepsilon = 2^{-n-1}$ since public randomness can be used to guess Alice's input.) The original definition of $\mathsf{PP}(f)$ given in [2] actually charged for the number of public coin flips *instead* of the $+\log(1/\varepsilon)$; however, by standard sparsification techniques (see [33] and [32, Theorem 3.14]) the two versions are essentially equivalent—they are within a constant factor plus $O(\log n)$—and the definition we have stated is much more prevalent in recent literature. It also follows from standard sparsification that we may convert any PP protocol into a UPP protocol of comparable cost: $\mathsf{UPP}(f) \le O(\mathsf{PP}(f) + \log n)$. In the converse direction, an exponential separation between UPP and PP is known [10, 36, 37].

SBP and USBP. Analogously to the above, we define

$$\mathsf{SBP}(f) \coloneqq \min_{\alpha(n)>0} \mathsf{R}^{\mathsf{pub}}_{\alpha,\alpha/2}(f) + \log(1/\alpha)\,,$$
$$\mathsf{USBP}(f) \coloneqq \min_{\alpha(n)>0} \mathsf{R}^{\mathsf{priv}}_{\alpha,\alpha/2}(f)\,.$$

Here the constant factor $1/2 = \beta/\alpha$ can be replaced by any positive constant less than 1 while affecting the complexity measures by only a constant factor: if we run a protocol ℓ times and accept iff all iterations accept, then β/α gets raised to the power ℓ while the communication and the $\log(1/\alpha)$ term each get multiplied by ℓ. We call this procedure *and-amplification* (in contrast to the usual *majority-amplification*). We also note that by standard sparsification, $\mathsf{USBP}(f) \le O(\mathsf{SBP}(f) + \log n)$ holds for all f. In the converse direction, at the time of this writing we did not know whether USBP is significantly more powerful than SBP (this question has subsequently been resolved in the negative—see Section 6), though a small separation is witnessed by the greater-than function, which has constant USBP complexity but $\Theta(\log n)$ SBP and PP complexity [9].

Relationship to Arthur–Merlin classes. Klauck [29, 31] and Aaronson and Wigderson [1] took up the study of communication complexity analogues of Arthur–Merlin games. Their results have already found applications in data streaming [11, 12, 23]. We do not define the communication models MA and AM here, but we note that the classical inclusions continue to hold in the communication setting (for the same reasons):

$$\mathsf{MA} \subseteq \mathsf{SBP} \subseteq \mathsf{AM}.$$

Indeed, if $\mathsf{MA}(f) = m$ then by majority-amplification and by absorbing Merlin's nondeterminism into the randomness we obtain

$$\mathsf{R}^{\mathsf{pub}}_{2^{-m-1},2^{-m-2}}(f) \le O(m^2).$$

Thus $\mathsf{SBP}(f) \le O(\mathsf{MA}(f)^2)$ (and the quadratic blow-up is necessary for "black-box" simulations [17]). On the other hand, $\mathsf{AM}(f) \le O(\mathsf{SBP}(f) + \log n)$ holds by sparsifying the randomness and using the Goldwasser–Sipser protocol [20].

1.3 Results for SBP and USBP

We prove that SBP communication complexity is exactly characterized by the well-known *corruption* lower bound method (also known as the *rectangle bound* or *one-sided discrepancy*). The definition of the corruption bound $\mathsf{Corr}(f)$ is given in Section 3, but for now, we note that $\mathsf{Corr}(f)$ essentially depends on the size of the largest approximately 1-monochromatic rectangle in the communication matrix of f. (For an extensive discussion of the different lower bound methods in communication complexity, see [24].) Previously, Klauck [29] showed that $\mathsf{Corr}(f)$ lies somewhere between the MA and AM communication complexities of f; namely $\Omega(\mathsf{AM}(f)) \le \mathsf{Corr}(f) \le O(\mathsf{MA}(f)^2)$. Klauck also gave a combinatorial near-characterization of $\mathsf{Corr}(f)$ (tight up to logarithmic factors) using so-called *one-sided uniform threshold covers*. The following theorem (proved in Section 3) sharpens these results by pinpointing precisely the class between MA and AM that is characterized by corruption.

Theorem 1.1. $\mathsf{SBP}(f) = \Theta(\mathsf{Corr}(f))$ *for all* f.

One way to frame Theorem 1.1 is as follows. A lot of effort (e. g., [29, 24, 28, 25, 19]) has been spent on comparing the relative strengths of different lower bound methods in communication complexity with the goal of finding a natural method that captures the bounded-error randomized communication complexity of every function. Theorem 1.1 can be viewed as achieving a diametrically opposite goal: we start with a historically important lower bound method (i. e., corruption) and find a natural communication measure that it captures. Theorem 1.1 is also somewhat analogous, in content and proof, to another result of Klauck [30] showing that the discrepancy bound captures PP.

Razborov [35] famously proved that $\mathsf{Corr}(\text{DISJ}) = \Theta(n)$. (The first linear lower bound for SET-DISJOINTNESS [27] did not use corruption.) By the results of [29], this implies that $\mathsf{MA}(\text{DISJ}) \geq \Omega(\sqrt{n})$. We immediately have a stronger corollary.

Corollary 1.2. $\mathsf{SBP}(\text{DISJ}) = \Theta(n)$.

To obtain lower bounds for USBP we show that the *information complexity* framework, as formulated by Bar-Yossef, Jayram, Kumar, and Sivakumar [4] (see also [14]), can be adapted to suit our purposes. The main technical result of this work is the following, proved in Section 4.

Theorem 1.3. $\mathsf{USBP}(\text{DISJ}) = \Theta(n)$.

We note that the statement of Theorem 1.3 is similar in spirit to Forster's theorem [18] stating that the UPP complexity of the INNER-PRODUCT function is $\Theta(n)$. Note also that Corollary 1.2 is of course a corollary of Theorem 1.3, too, but the corruption-based proof via Theorem 1.1 is arguably more elementary than the proof of Theorem 1.3. Finally, we note that the well-studied GAP-HAMMING-DISTANCE promise problem [13, 40, 39] (where 1-inputs have distance $\geq n/2 + \sqrt{n}$ and 0-inputs have distance $\leq n/2 - \sqrt{n}$) has SBP and USBP complexities $\Theta(\sqrt{n})$, where the lower bound follows by Theorem 1.3 and a standard reduction from DISJ, and the upper bound follows by and-amplification of the trivial protocol that checks inequality at a random bit position.

1.4 Characterization for all α and β

Using our results for SBP and USBP in a black-box manner we derive the following (near) complete characterization for the randomized communication complexity of SET-DISJOINTNESS in Section 2.

Theorem 1.4 (Private-coin)**.** *For all* $\alpha(n) > \beta(n)$*,*

$$\mathsf{R}^{\mathsf{priv}}_{\alpha,\beta}(\text{DISJ}) = \Theta(n \cdot (1 - \beta/\alpha) + \log n)\,.$$

Theorem 1.5 (Public-coin)**.** *There is a universal constant* $C > 0$ *such that for all* $\alpha(n) > \beta(n)$*,*

$$\mathsf{R}^{\mathsf{pub}}_{\alpha,\beta}(\text{DISJ}) = \begin{cases} \Theta(n \cdot (1-\beta/\alpha)) & \textit{when } \log(1/\alpha) \leq C \cdot n \cdot (1-\beta/\alpha)\,, \\ 2 & \textit{when } \log(1/\alpha) \geq \lceil n \cdot (1-\beta/\alpha) \rceil\,. \end{cases}$$

We stress that for the public-coin characterization (and in particular, the result of [8] as a corollary), it suffices to rely only on Razborov's corruption lemma (via Corollary 1.2), and not on any information complexity techniques. Braverman and Moitra [8] observed that

$$\mathsf{R}^{\mathsf{pub}}_{1/2+\varepsilon,\,1/2-\varepsilon}(\text{DISJ}) \geq \Omega(\varepsilon^2 n)$$

follows from the standard bounded-error lower bound by majority-amplification, and they obtained the tight $\Omega(\varepsilon n)$ bound by developing information complexity techniques tailored to this setting. Our idea is that and-amplification imposes only an ε factor loss (rather than the ε^2 factor loss imposed by majority-amplification) while still reducing to a case where the corruption method applies.

We also note that for public-coin protocols there remains a small gap in the parameters around the threshold $\log(1/\alpha) = \Theta(n \cdot (1-\beta/\alpha))$ that is not covered by our theorem. As we discuss in Section 2, the power of the public coins kicks in at this threshold.

Finally, we mention that all the SET-DISJOINTNESS lower bounds in this paper continue to hold under the *unique-intersection* promise where the inputs are either disjoint or intersect in exactly one coordinate: for Corollary 1.2 this property is inherited from Razborov's proof; for Theorem 1.3 this property is implicit in our proof.

1.5 Extended formulations for maximum-clique

The authors of [6] proved that for every positive constant $\delta < 1/2$, so-called $n^{1/2-\delta}$-approximate extended formulations for a certain natural linear programming encoding of the maximum-clique problem for graphs have complexity $2^{\Omega(n^{2\delta})}$. Their proof involves developing a generalization of the corruption lemma from [35]. In [8], this result was improved using information-theoretic methods to show a tight lower bound for maximum-clique: for every positive constant $\delta < 1$, such $n^{1-\delta}$-approximate extended formulations have complexity $2^{\Omega(n^{\delta})}$. In [7], a simplified proof of the latter result was given, also using information-theoretic methods.

The reason the proof due to [6] only works up to $n^{1/2-\delta}$-approximation (rather than $n^{1-\delta}$-approximation) is similar to the reason why majority-amplification yields

$$\mathsf{R}^{\mathsf{pub}}_{1/2+\varepsilon,\,1/2-\varepsilon}(\text{DISJ}) \geq \Omega(\varepsilon^2 n)$$

(rather than $\Omega(\varepsilon n)$) from the standard bounded-error lower bound for DISJ. Although communication complexity lower bounds do not seem to imply extended formulation lower bounds in a black-box way, the tools for the former have generally been useful for the latter. This suggests that perhaps an idea analogous to our and-amplification-based communication lower bound proof could be used to "bootstrap" the result of [6] to get the tight lower bound for maximum-clique. In Section 5 we confirm that this is indeed the case, thus obtaining a proof of the tight bound that is simpler than the ones given in [8, 7] and avoids their use of information-theoretic methods (instead relying only on the corruption-based argument of [6]).

2 The complexity of Set-Disjointness

We now prove Theorem 1.4 and Theorem 1.5 using Theorem 1.3 and Corollary 1.2.

2.1 Upper bounds

Public-coin protocols. We start with a simple $\mathsf{R}^{\mathsf{pub}}_{1,\beta/\alpha}$ protocol for DISJ of cost $\Theta(n \cdot (1-\beta/\alpha))$.

Basic public-coin protocol Π.

1. Use public randomness to pick a uniformly random $S \subseteq [n]$ of size $\lceil n \cdot (1-\beta/\alpha) \rceil$.
2. Alice sends the substring $x|_S$ to Bob.
3. Bob outputs $\textsc{Disj}(x|_S, y|_S)$.

It is straightforward to check that Π is indeed an $\mathsf{R}^{\mathsf{pub}}_{1,\beta/\alpha}$ protocol. To obtain an $\mathsf{R}^{\mathsf{pub}}_{\alpha,\beta}$ protocol for the first part of Theorem 1.5 (without needing any restriction on the parameters), we can reject with probability $1-\alpha$ at the beginning and otherwise run Π. To obtain a protocol of cost 2 for the second part of Theorem 1.5, we need to better exploit the power of public coins. If we modify Π so that additional public coins are used to guess $x|_S$, then Alice can just send one bit indicating whether the guess is correct, and Bob can send the output bit (rejecting if the guess was wrong). This yields an

$$\mathsf{R}^{\mathsf{pub}}_{1/2^{|S|},\,\beta/\alpha 2^{|S|}}$$

protocol which, by the restriction that $\alpha \le 1/2^{|S|}$, can be adapted into an $\mathsf{R}^{\mathsf{pub}}_{\alpha,\beta}$ protocol by automatically rejecting with probability $1-\alpha 2^{|S|}$.

In fact, the above protocols can be seen as special cases of the following general protocol, which interpolates between them. For simplicity of presentation, let us assume that $\log(1/\alpha)$ is an integer and $\log(1/\alpha) \le |S|$. In step 2 of the basic protocol Π, Alice can expedite the sending of her message to Bob as follows: Alice and Bob interpret additional public coins as guessing the first $\log(1/\alpha)$ bits of Alice's message. Alice can use one bit of communication to indicate whether this guess is correct, and if so she can send the other $|S| - \log(1/\alpha)$ bits of her message normally. The probability that the public guess is correct is $2^{-\log(1/\alpha)} = \alpha$. Thus, this new protocol ends up working in a familiar way: with probability $1-\alpha$ the public guess fails (in which case we reject), but otherwise we are able to run Π successfully. This results in an $\mathsf{R}^{\mathsf{pub}}_{\alpha,\beta}$ protocol of cost $|S| - \log(1/\alpha) + 2$. Here the $+2$ comes from Alice indicating whether the public guess is correct and Bob sending the final answer.

Private-coin protocols. By sparsification, we may assume the basic protocol Π uses only $O(\log n)$ bits of public randomness. Thus we have

$$\mathsf{R}^{\mathsf{priv}}_{1,\beta/\alpha}(\textsc{Disj}) \le O(n \cdot (1-\beta/\alpha) + \log n)$$

since Alice can pick S privately and send it to Bob along with $x|_S$. An $\mathsf{R}^{\mathsf{priv}}_{\alpha,\beta}$ protocol for Theorem 1.4 can be obtained as previously: automatically reject with probability $1-\alpha$ and otherwise run the $\mathsf{R}^{\mathsf{priv}}_{1,\beta/\alpha}$ protocol.

2.2 Lower bounds

Private-coin lower bounds. Let Π be an $\mathsf{R}^{\mathsf{priv}}_{\alpha,\beta}$ protocol for $\textsc{Disj}$. We prove that the cost of Π is both $\Omega(n \cdot (1-\beta/\alpha))$ and $\Omega(\log n)$, as required for Theorem 1.4.

First, if we do and-amplification by iterating the protocol $\lceil 1/(1-\beta/\alpha) \rceil$ times and accepting iff all runs accept, we get an $\mathsf{R}^{\mathsf{priv}}_{\alpha',\alpha'/2}$ protocol for $\textsc{Disj}$ with $\alpha' := \alpha^{\lceil 1/(1-\beta/\alpha) \rceil}$ (since $(\beta/\alpha)^{\lceil 1/(1-\beta/\alpha) \rceil} < 1/2$).

By Theorem 1.3 the amplified protocol must use $\Omega(n)$ communication and hence Π must have used $\Omega(n\cdot(1-\beta/\alpha))$ communication.

Second, Forster's result [18] that the UPP complexity of INNER-PRODUCT is $\Omega(n)$ gives us the $\Omega(\log n)$ lower bound for Π. Indeed, the INNER-PRODUCT function reduces to DISJ with exponential blow-up (see [38, Proposition 6.5]) and we may convert Π into a UPP protocol by shifting the acceptance threshold near $1/2$.

Public-coin lower bounds. Let Π be an $\mathsf{R}^{\mathsf{pub}}_{\alpha,\beta}$ protocol for DISJ. We consider the two parts of Theorem 1.5 separately.

For the first part, suppose $\log(1/\alpha)\leq C\cdot n\cdot(1-\beta/\alpha)$ for a to-be-specified constant C. We proceed exactly as above: We first and-amplify Π into an $\mathsf{R}^{\mathsf{priv}}_{\alpha',\alpha'/2}$ protocol. The parameters satisfy

$$\log(1/\alpha')=\log(1/\alpha)\cdot\lceil 1/(1-\beta/\alpha)\rceil\leq C\cdot n\cdot(1-\beta/\alpha)\cdot\lceil 1/(1-\beta/\alpha)\rceil\leq 2C\cdot n\,.$$

Hence if C is a sufficiently small universal constant then the $\Omega(n)$ lower bound for the amplified protocol (provided now by Corollary 1.2) must be coming from the communication cost and not from the $\log(1/\alpha')$ term. We conclude that the original protocol Π must have used $\Omega(n\cdot(1-\alpha/\beta))$ communication.

For the second part, we do not need any restriction on the parameters. We claim that since DISJ has a 2×2 identity submatrix, we cannot have $\mathsf{R}^{\mathsf{pub}}_{\alpha,\beta}(\text{DISJ})\leq 1$.[1] Suppose for contradiction there is a 1-bit protocol and yet $\text{DISJ}(x,y)=\text{DISJ}(x',y')=1$ and $\text{DISJ}(x,y')=\text{DISJ}(x',y)=0$. Say r is the probability Alice declares the output and $1-r$ is the probability Bob declares the output. Conditioned on Alice declaring the output let $p_x,p_{x'}$ be the acceptance probability for the x and x' rows, and conditioned on Bob declaring the output let $q_y,q_{y'}$ be the acceptance probability for the y and y' columns. Letting $\pi_{xy}:=rp_x+(1-r)q_y$ be the overall acceptance probability on input (x,y), we have

$$\alpha-\beta\leq\pi_{xy}-\pi_{x'y}=r(p_x-p_{x'})\quad\text{and}\quad\alpha-\beta\leq\pi_{x'y'}-\pi_{xy'}=r(p_{x'}-p_x)\,,$$

a contradiction.

3 SBP is characterized by corruption

In this section we prove Theorem 1.1, which states that $\mathsf{SBP}(f)=\Theta(\mathsf{Corr}(f))$ for all f. We start by defining the corruption bound. We say a distribution μ over inputs is *balanced* (with respect to f) if $\mu(f^{-1}(1))=\mu(f^{-1}(0))=1/2$. We say a rectangle R is 1-*biased* (with respect to f and μ) if $\mu(R\cap f^{-1}(0))\leq\mu(R)/8$. The corruption bound is defined as

$$\mathsf{Corr}(f):=\max_{\text{balanced }\mu}\ \min_{\text{1-biased }R}\ \log\left(\frac{1}{\mu(R)}\right).$$

[1] For the public-coin version of UPP, an equivalence actually holds. For all f, we have $\mathsf{UPP}^{\mathsf{pub}}(f)\leq 2$, and it is not difficult to show that the following are equivalent: (i) $\mathsf{UPP}^{\mathsf{pub}}(f)\leq 1$, (ii) there exist row and column values $p_x,q_y\in[0,1]$ and $r\in[0,1]$ such that $\left|rp_x+(1-r)q_y-f(x,y)\right|<1/2$, (iii) the rows and columns can be permuted so each row and each column is monotonically nondecreasing (0's then 1's), (iv) f does not contain as a submatrix the 2×2 identity (or its complement). To see that (iii) $\Rightarrow$ (ii), take $r=1/2$, and $p_x=$ fraction of 1's in the x row, and $q_y=(y-1/2)/(\text{number of columns})$ where y is viewed as a positive integer.

If the corruption bound of f is high, then for some μ all 1-biased rectangles are small (and hence if many inputs must be covered by such rectangles, then many such rectangles will be needed). It was proved in [29] that the constant factor of $1/8$ (in the definition of 1-biased) can be replaced by any positive constant at most $1/8$ while affecting the corruption bound by only a constant factor. It was also proved in [29] that the bound is robust with respect to the balance condition on μ.

3.1 SBP is lower bounded by corruption

Here we show the lower bound $\mathsf{SBP}(f) \geq \Omega(\mathsf{Corr}(f))$. The intuition is as follows. The first step is to fix the public randomness of an SBP protocol in such a way that the average-case behavior of the resulting deterministic protocol mimics the worst-case behavior of the original protocol. Typically, this sort of thing is done by invoking the distributive law (linearity of expectation), but here we need a more elaborate calculation due to the asymmetric nature of SBP. Then, the rest of the argument follows along similar lines as the proof in [29] (that $\mathsf{Corr}(f) \leq O(\mathsf{MA}(f)^2)$), showing that the 1-inputs are mostly covered by "small" transcript rectangles (of our deterministic protocol), hence many such rectangles are needed.

We proceed with the formal proof. Let Π be an $\mathsf{R}^{\mathsf{pub}}_{\alpha,\alpha/32}$ protocol for f; recall that by and-amplification we may assume $\beta = \alpha/32$ rather than $\beta = \alpha/2$ in the definition of SBP. Assuming $\log(1/\alpha) < \mathsf{Corr}(f)/2$, we show that Π uses $\Omega(\mathsf{Corr}(f))$ bits of communication. To this end, fix a balanced distribution μ such that for all 1-biased rectangles R, $\mu(R) \leq 2^{-\mathsf{Corr}(f)}$.

Identify the possible outcomes of public randomness with $\{1,\ldots,m\}$, and let Π_i denote Π running with public randomness i. Let p_i be the probability the public randomness is i (so $p_i = 1/m$ if the public randomness is uniformly distributed). Let q_i be the probability over μ that Π_i accepts, conditioned on the input being a 1-input. Let r_i be the same but conditioned on a 0-input. Now

$$\sum_i p_i q_i = \Pr_{i,(x,y)\sim\mu}\big[\Pi_i(x,y) \text{ accepts} \bigm| f(x,y)=1\big] \geq \alpha\,, \tag{3.1}$$

$$\sum_i p_i r_i = \Pr_{i,(x,y)\sim\mu}\big[\Pi_i(x,y) \text{ accepts} \bigm| f(x,y)=0\big] \leq \alpha/32\,. \tag{3.2}$$

Claim 3.1. *There exists an i^* such that $q_{i^*} \geq \alpha/2$ and $r_{i^*} \leq q_{i^*}/16$.*

Proof of Claim. Suppose for contradiction that for all i either $q_i < \alpha/2$ or $r_i > q_i/16$. Let $S \subseteq \{1,\ldots,m\}$ be such that for all $i \in S$, $q_i < \alpha/2$, and for all $i \in \overline{S}$, $r_i > q_i/16$. Then

$$\sum_i p_i r_i \geq \sum_{i\in\overline{S}} p_i r_i \geq \sum_{i\in\overline{S}} p_i q_i/16 = \frac{1}{16}\Big(\sum_i p_i q_i - \sum_{i\in S} p_i q_i\Big) \geq \frac{1}{16}\Big(\alpha - \Big(\sum_{i\in S} p_i\Big)\alpha/2\Big) \geq \alpha/32\,.$$

Furthermore, at least one of the inequalities must be strict, contradicting (3.2). □

Fix an i^* guaranteed by Claim 3.1. Using i^* as the public randomness in Π, we can now apply the usual corruption argument. Consider the 1-rectangles that correspond to accepting transcripts of Π_{i^*}. Call a 1-rectangle R *large* if $\mu(R) > 2^{-\mathsf{Corr}(f)}$ and *small* otherwise. Recall that by our assumption on μ, no large 1-rectangle is 1-biased: for every large 1-rectangle R we have $\mu(R \cap f^{-1}(0)) > \mu(R)/8$. Under μ, the total measure of large 1-rectangles is at most half the total measure of all 1-rectangles, since otherwise

$$
\begin{aligned}
r_{i^*} &= 2 \Pr_{(x,y)\sim\mu}\big[\Pi_{i^*}(x,y) \text{ accepts and } f(x,y)=0\big] \\
&= 2 \sum_{\text{1-rectangles } R} \mu(R\cap f^{-1}(0)) \\
&\geq 2 \sum_{\text{large 1-rectangles } R} \mu(R\cap f^{-1}(0)) \\
&> \frac{1}{4} \sum_{\text{large 1-rectangles } R} \mu(R) \\
&> \frac{1}{4}\cdot\frac{1}{2} \sum_{\text{1-rectangles } R} \mu(R) \\
&\geq \frac{1}{8} \sum_{\text{1-rectangles } R} \mu(R\cap f^{-1}(1)) \\
&= \frac{1}{8} \Pr_{(x,y)\sim\mu}\big[\Pi_{i^*}(x,y) \text{ accepts and } f(x,y)=1\big] \\
&= \frac{1}{8}\cdot q_{i^*}/2 \\
&= q_{i^*}/16\,.
\end{aligned}
$$

Therefore

$$\sum_{\text{small 1-rectangles } R} \mu(R) \;\geq\; \frac{1}{2} \sum_{\text{1-rectangles } R} \mu(R) \;\geq\; \frac{1}{2}\cdot q_{i^*}/2 \;\geq\; \alpha/8 \;>\; 2^{-\mathsf{Corr}(f)/2-3}\,.$$

Thus there are at least $2^{-\mathsf{Corr}(f)/2-3}/2^{-\mathsf{Corr}(f)} = 2^{\mathsf{Corr}(f)/2-3}$ small 1-rectangles, which implies that Π uses at least $\mathsf{Corr}(f)/2-3$ bits of communication.

3.2 SBP is upper bounded by corruption

Here we show the upper bound $\mathsf{SBP}(f) \leq O(\mathsf{Corr}(f))$. The intuition is as follows. If the corruption bound is small, that means for every balanced distribution over inputs there exists a rectangle (which can be viewed as a 2-bit protocol) that exhibits average-case SBP-like behavior—accepting a random 1-input with not-too-small probability, and accepting a random 0-input with constant-factor-smaller probability. We use the minimax theorem to convert this property into a distribution over rectangles, with a worst-case SBP guarantee. Several technical issues arise with this argument. One is the asymmetry between 1-inputs and 0-inputs, but this can be massaged away using a linear transformation of probabilities before invoking minimax. Another is that the corruption bound can yield an average-case SBP rectangle with a different "α" for different balanced distributions, whereas the minimax application requires a single α to work uniformly for all balanced distributions. This issue is fixed by passing to an appropriate subrectangle to decrease the α if necessary, for any given balanced distribution.

We proceed with the formal proof. For notational convenience we let $\mathbf{0}$ and $\mathbf{1}$ stand for the events $f^{-1}(0)$ and $f^{-1}(1)$, respectively. For example,

$$\mu(\mathbf{0}\,|\,R) = \mu(R \cap f^{-1}(0))/\mu(R) \quad \text{and} \quad \mu(R\,|\,\mathbf{0}) = \mu(R \cap f^{-1}(0))/\mu(f^{-1}(0)).$$

Define $\alpha \coloneqq 2^{-\mathsf{Corr}(f)}$.

Claim 3.2. *For every balanced μ there exists a rectangle R with $\mu(R\,|\,\mathbf{1}) \geq \alpha$ and $\mu(R\,|\,\mathbf{0}) \leq \alpha/2$.*

Proof of Claim. Fix a balanced distribution μ. By definition of corruption, there exists a rectangle S such that $\mu(S) \geq \alpha$ and $\mu(\mathbf{0}\,|\,S) \leq 1/8$. Decompose S as the disjoint union $S_1 \cup S_2 \cup \cdots \cup S_m$ where the S_i's are the individual rows of S, sorted in nondecreasing order of $\mu(\mathbf{0}\,|\,S_i)$. Let $S_{\leq i} \coloneqq S_1 \cup S_2 \cup \cdots \cup S_i$. For every i we know that $\mu(\mathbf{0}\,|\,S_{\leq i}) \leq \mu(\mathbf{0}\,|\,S) \leq 1/8$. If there exists an i such that $\alpha \leq \mu(S_{\leq i}\,|\,\mathbf{1}) \leq 2\alpha$ then $R \coloneqq S_{\leq i}$ witnesses the claim since

$$\mu(S_{\leq i}\,|\,\mathbf{0}) = \frac{\mu(\mathbf{0}\,|\,S_{\leq i}) \cdot \mu(S_{\leq i}\,|\,\mathbf{1}) \cdot \mu(\mathbf{1})}{\mu(\mathbf{0}) \cdot \mu(\mathbf{1}\,|\,S_{\leq i})} \leq \frac{(1/8) \cdot 2\alpha \cdot (1/2)}{(1/2) \cdot (7/8)} = 2\alpha/7 \leq \alpha/2.$$

Otherwise, since $\mu(S_{\leq m}\,|\,\mathbf{1}) = \mu(S\,|\,\mathbf{1}) = \mu(\mathbf{1}\,|\,S) \cdot \mu(S)/\mu(\mathbf{1}) \geq (7/8) \cdot \alpha/(1/2) > \alpha$ and $\mu(S_{\leq 0}\,|\,\mathbf{1}) = 0 < \alpha$, there must exist an i such that $\mu(S_{\leq i}\,|\,\mathbf{1}) > 2\alpha$ and $\mu(S_{\leq i-1}\,|\,\mathbf{1}) < \alpha$ and thus $\mu(S_i\,|\,\mathbf{1}) > 2\alpha - \alpha = \alpha$. In this case, the rectangle $R \coloneqq S_i \cap \mathbf{1}$ witnesses the claim since $\mu(S_i \cap \mathbf{1}\,|\,\mathbf{1}) = \mu(S_i\,|\,\mathbf{1}) > \alpha$ and $\mu(S_i \cap \mathbf{1}\,|\,\mathbf{0}) = 0 \leq \alpha/2$. $\square$

Let M be the matrix with rows indexed by inputs $(x,y) \in \{0,1\}^n \times \{0,1\}^n$ and columns indexed by rectangles $R \subseteq \{0,1\}^n \times \{0,1\}^n$ such that

$$M_{(x,y),R} \coloneqq \begin{cases} 1 & \text{if } f(x,y) = 1 \text{ and } (x,y) \in R, \\ 0 & \text{if } f(x,y) = 1 \text{ and } (x,y) \notin R, \\ 0 & \text{if } f(x,y) = 0 \text{ and } (x,y) \in R, \\ \frac{\alpha}{1-\alpha/2} & \text{if } f(x,y) = 0 \text{ and } (x,y) \notin R. \end{cases}$$

We claim that for every distribution μ over inputs, there exists a rectangle R such that $\mathbb{E}(M_{\mu,R}) \geq \alpha$ (where $\mathbb{E}$ denotes expectation). If $\mu(\mathbf{0}) = 0$ then take $R \coloneqq \{0,1\}^n \times \{0,1\}^n$, and if $\mu(\mathbf{1}) = 0$ then take $R \coloneqq \emptyset$. Otherwise, let μ' be the balanced version of μ and invoke Claim 3.2 to find an R such that $\mu'(R\,|\,\mathbf{1}) \geq \alpha$ and $\mu'(R\,|\,\mathbf{0}) \leq \alpha/2$. Then we have

$$\begin{aligned} \mathbb{E}(M_{\mu,R}) &= \mu(R\,|\,\mathbf{1}) \cdot \mu(\mathbf{1}) + \frac{\alpha}{1-\alpha/2} \cdot \mu(\overline{R}\,|\,\mathbf{0}) \cdot \mu(\mathbf{0}) \\ &= \mu'(R\,|\,\mathbf{1}) \cdot \mu(\mathbf{1}) + \frac{\alpha}{1-\alpha/2} \cdot \mu'(\overline{R}\,|\,\mathbf{0}) \cdot \mu(\mathbf{0}) \\ &\geq \alpha \cdot \mu(\mathbf{1}) + \frac{\alpha}{1-\alpha/2} \cdot (1-\alpha/2) \cdot \mu(\mathbf{0}) \\ &= \alpha. \end{aligned}$$

Now by the minimax theorem (we can use the most basic version since the matrix M is finite), there exists a distribution D over rectangles such that for every input (x,y), $\mathbb{E}(M_{(x,y),D}) \geq \alpha$. If $f(x,y) = 1$

this means the probability a random rectangle from D contains (x,y) is at least α. If $f(x,y)=0$ this means $\alpha/(1-\alpha/2)$ times the probability a random rectangle from D does not contain (x,y) is at least α, in other words the probability a random rectangle from D contains (x,y) is at most $\alpha/2$. Thus the protocol that picks a random rectangle from D and accepts iff the input is in the rectangle shows that $\mathsf{R}^{\mathsf{pub}}_{\alpha,\alpha/2}(f) \le 2$ and hence $\mathsf{SBP}(f) \le 2+\log(1/\alpha) = 2+\mathsf{Corr}(f)$.

4 USBP lower bound

In this section we prove Theorem 1.3, which states that $\mathsf{USBP}(\textsc{Disj}) = \Theta(n)$. We first give an informal overview.

Our proof uses the by-now standard information complexity approach [4, 14]. In this approach, one considers some suitably distributed random input (X,Y) and measures the amount of information that the protocol transcript $\Pi(X,Y)$ (i. e., the concatenation of all messages sent) "leaks" about the input as quantified by the *mutual information* $\mathbb{I}(\Pi(X,Y)\,;X,Y)$. Lower bounding the mutual information has the side effect of lower bounding the *entropy* $\mathbb{H}(\Pi(X,Y))$ of the transcript, which in turn lower bounds the length of the transcript and thereby the communication complexity. It is often useful to involve the *conditional* versions of these information measures, defined by

$$\mathbb{H}(\Pi\,|\,Z) := \mathbb{E}_{z\sim Z}\mathbb{H}(\Pi\,|\,Z=z) \quad \text{and} \quad \mathbb{I}(\Pi\,;X,Y\,|\,Z) := \mathbb{E}_{z\sim Z}\mathbb{I}(\Pi\,;X,Y\,|\,Z=z)$$

where Z is some random variable (jointly distributed with X and Y). We refer the reader to [16] for discussions of these basic information theory concepts.

A key benefit of studying mutual information is that one automatically obtains for it a *direct sum* property (as in [14, 4]), as long as the coordinates (X_i,Y_i), $i\in[n]$, are mutually independent. This way, the task of proving an $\Omega(n)$ lower bound for the original problem reduces to the task of proving an $\Omega(1)$ information lower bound for some constant-size "gadget." For the SET-DISJOINTNESS function $\textsc{Disj} := \textsc{And}_n \circ \textsc{Nand}^n$ this gadget is typically NAND.

Our proof follows this outline. The reduction to the single-gadget case will be packaged into Lemma 4.1 and is standard. By contrast, in proving the $\Omega(1)$ information lower bound for the single gadget, we need to overcome the following two new technical issues.

(1) Small acceptance probabilities. Since the protocol is only required to succeed with a tiny probability $\alpha(n)$ on 1-inputs, the transcript of Π may be useless most of the time: Imagine a protocol that rejects with probability $1-\alpha$ at the start (and otherwise does something useful). The entropy of the transcript of such protocols can be as low as $O(\alpha)$.

To address this issue, we do not work with the transcript distribution of Π directly, but rather with the *conditional distribution* given that the protocol accepts. That is, for 1-inputs (x,y), we consider the random variable

$$T(x,y) := \Pi(x,y)\,|\,\Pi(x,y) \text{ is an accepting transcript}$$

and proceed to lower bound $\mathbb{I}(T(X,Y)\,;X,Y)$ instead. One subtlety is that conditioning on acceptance does not "commute" with the reduction to the single-gadget case. We must consider the distribution of T that arises from first conditioning on acceptance and then doing the reduction, which is generally not the

Protocol Π^*. On input $(x,y) \in \{0,1\}^2$:

1. If $x=0$ Alice sends a "1." If $x=1$ Alice sends a "1" with probability α and rejects otherwise (by sending a "0").

2. Suppose Alice sent a "1." Then if $y=0$ Bob accepts (by sending a "1"). If $y=1$ then Bob accepts with probability α and rejects otherwise.

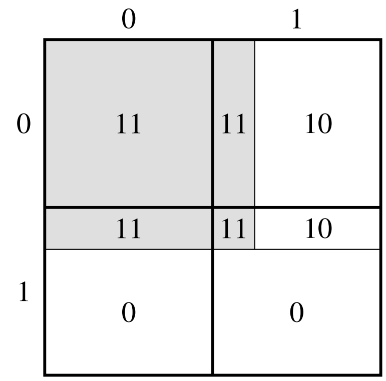

Figure 1: Protocol Π^* for NAND. In the illustration on the right, each of the input blocks is further subdivided into rectangles according to the outcomes of the private coins. The rectangles are labeled with the associated transcripts.

same distribution as if we did the reduction and then conditioned on acceptance. However, this is not a significant technical obstacle.

(2) Large acceptance probabilities. The acceptance probability of a protocol Π can vary between α and 1 when run on different 1-inputs. This, together with our conditioning idea above, introduces a new problem: there are USBP protocols for NAND such that the associated T *leaks no information about the input!*

Indeed, consider the protocol Π^* for NAND given in Figure 1. This protocol accepts the 1-input $(0,0)$ with probability 1, the 1-inputs $(0,1)$ and $(1,0)$ with probability α, and the 0-input $(1,1)$ with probability α^2. Choosing α such that $\alpha^2 \le \alpha/2$ we obtain a USBP protocol for NAND where the associated conditioned-on-acceptance variable T^* is constant (the protocol Π^* has only one accepting transcript, namely "11").

To avoid this problem, we use a more complicated gadget than NAND; see Figure 2a. The new gadget G contains two instances of NAND: in Figure 2b one instance of NAND corresponds to the pair of edges AB and another one to AC. We show that the bad situation described above—i. e., the 1-input $(0,0)$ of NAND having much higher acceptance probability than the other two 1-inputs $(0,1)$ and $(1,0)$ of NAND—cannot happen simultaneously for both instances. One subtlety (arising from the conditioning not "commuting" with the reduction, as described above) is that the bad situation actually depends on the transcript, with some transcripts being bad for AB and some being bad for AC, but none being bad for both. We prove an information lower bound (conditioned on acceptance) for whichever instance of NAND behaves better for "most" transcripts.

We note that a similar technical issue arose in the proof of Braverman and Moitra [8] when analyzing the case $\alpha = 1/2+\varepsilon$, $\beta = 1/2-\varepsilon$. Their solution involved applying a certain type of *random-self-reduction* (they called it *smoothing*) to the inputs before invoking the protocol. This approach is highly tailored to their setting and does not seem to be directly helpful to us. Nevertheless, our gadget G was inspired by their analysis.

(a) **(b)**

Figure 2: (a) Truth table of the gadget G. (b) Distributions $Q_{\rm A}$, $Q_{\rm B}$, and $Q_{\rm C}$ are uniform over the endpoints of the edges A, B, and C, respectively.

4.1 Proof of Theorem 1.3

Define the gadget $G\colon \{0,1,2\}\times\{1,2\}\to\{0,1\}$ as the indicator for non-equality; see Figure 2a. Define the function $F\colon \{0,1,2\}^n\times\{1,2\}^n\to\{0,1\}$ by $F := \textsc{And}_n\circ G^n$, i. e., $F(x,y)=1$ iff $G(x_i,y_i)=1$ for all $i\in[n]$. Since G reduces to $\textsc{Disj}$ on 2 bits by the map $(0\mapsto 00,\ 1\mapsto 01,\ 2\mapsto 10)$, we find that F reduces to $\textsc{Disj}$ on $2n$ bits. Hence it suffices to prove that $\mathsf{USBP}(F)\geq\Omega(n)$.

Input distribution. Define $Q_{\rm A}, Q_{\rm B}, Q_{\rm C}$ to be the following three distributions over $\{0,1,2\}\times\{1,2\}$: $Q_{\rm A}$ is uniform over $\{(0,1),(0,2)\}$, $Q_{\rm B}$ is uniform over $\{(0,1),(2,1)\}$, and $Q_{\rm C}$ is uniform over $\{(0,2),(1,2)\}$; see Figure 2b. For $i\in[n]$, $u\in\{0,1,2\}$, $v\in\{1,2\}$, and z a length-$(n-1)$ string over the alphabet $\{\rm A,B,C\}$ indexed by $[n]\setminus\{i\}$, define $D_{i,u,v,z}$ to be the distribution over pairs $(x,y)\in\{0,1,2\}^n\times\{1,2\}^n$ obtained by setting $x_i=u$, $y_i=v$, and for each $j\neq i$ (independently) sampling (x_j,y_j) from Q_{z_j}. Note that $\mathrm{support}(D_{i,u,v,z})\subseteq F^{-1}(G(u,v))$ and that x and y are independent when sampled from $D_{i,u,v,z}$.

Reduction to the single-gadget case. Let Π be an $\mathsf{R}^{\mathsf{priv}}_{\alpha,\alpha/4}$ protocol for F (recall that by and-amplification we may assume $\beta=\alpha/4$ in the definition of USBP). Let $\Pi(x,y)$ denote the transcript of Π on input (x,y). Thus $\Pi(x,y)$ is a random variable whose outcome depends on the private coins of the protocol. For $(x,y)\in F^{-1}(1)$ define $T(x,y)$ as the random variable whose distribution is that of $\Pi(x,y)$ conditioned on $\Pi(x,y)$ being an accepting transcript.

Suppose for contradiction that the transcripts have length less than $n/2400$. Using the direct sum methodology we will next find a coordinate i (and a string z) such that the protocol leaks very little information about the i-th input (conditioned on the data z). This is formalized in the following lemma whose proof we defer to Section 4.2 as it is essentially identical to the corresponding argument in [4]. Below, $\|X-Y\|$ denotes the statistical distance between the distributions of the random variables X and Y.

Lemma 4.1. *If $\gamma>0$ is such that for all i and z either*

$$\|T(D_{i,0,1,z})-T(D_{i,0,2,z})\|\geq\gamma\quad\text{or}\quad\|T(D_{i,0,1,z})-T(D_{i,2,1,z})\|\geq\gamma\quad\text{or}\quad\|T(D_{i,0,2,z})-T(D_{i,1,2,z})\|\geq\gamma,$$

then some transcript has length at least $n\gamma^2/6$.

Contrapositively, letting $\gamma = 1/20$, Lemma 4.1 implies that there exists an i and z (which we fix henceforth) such that $\|T(D_{i,0,1,z}) - T(D_{i,0,2,z})\|$, $\|T(D_{i,0,1,z}) - T(D_{i,2,1,z})\|$, and $\|T(D_{i,0,2,z}) - T(D_{i,1,2,z})\|$ are all less than γ.

The single-gadget case. Let (u,v) be an input to G and let τ be a transcript. We define

$$\begin{aligned} \pi_{uv}(\tau) &:= \Pr\big[\Pi(D_{i,u,v,z}) = \tau\big] \quad \text{for any } (u,v)\,, \\ t_{uv}(\tau) &:= \Pr\big[T(D_{i,u,v,z}) = \tau\big] \quad \text{for } (u,v) \in G^{-1}(1)\,. \end{aligned}$$

Note that $\pi_{uv}(\tau)/t_{uv}(\tau)$ is generally *not* (as might appear at a glance) the acceptance probability of Π on a random input from $D_{i,u,v,z}$ (this is related to the conditioning on acceptance not "commuting" with the reduction to the single-gadget case, as mentioned in the intuition).

We henceforth adopt the convention that $0/0 := 0$. Let

$$S := \left\{ \text{accepting } \tau \,:\, \frac{\pi_{01}(\tau)}{t_{01}(\tau)} \le \frac{\pi_{02}(\tau)}{t_{02}(\tau)} \right\}$$

and let $\overline{S} := \{\text{accepting } \tau\} \backslash S$. Connecting to the intuition, S consists of transcripts that are bad for AC, and $\overline{S}$ consists of transcripts that are bad for AB. Since $\|T(D_{i,0,1,z}) - T(D_{i,0,2,z})\| < \gamma$, we have either

$$\Pr\big[T(D_{i,0,1,z}) \in S\big] \ge \frac{1-\gamma}{2} \quad \text{or} \quad \Pr\big[T(D_{i,0,2,z}) \in \overline{S}\big] \ge \frac{1-\gamma}{2}\,.$$

Henceforth assume the former case; a completely analogous argument handles the latter case. We now only need to consider $(u,v) \in \big\{(0,1),(2,1),(0,2),(2,2)\big\}$, forming the AB instance of NAND in G. (In the latter case, we would only need to consider $(u,v) \in \big\{(0,1),(1,1),(0,2),(1,2)\big\}$, forming the AC instance of NAND in G.)

Note that $\pi_{22}(\tau)\cdot\pi_{01}(\tau) = \pi_{21}(\tau)\cdot\pi_{02}(\tau)$ by the basic rectangular structure of τ. Also note that if $G(u,v) = 1$ and τ is accepting, then the following both hold.

- We have $\pi_{uv}(\tau) = 0$ iff $t_{uv}(\tau) = 0$, and hence $\pi_{uv}(\tau) = \frac{\pi_{uv}(\tau)}{t_{uv}(\tau)} \cdot t_{uv}(\tau)$.
- Assuming $\pi_{uv}(\tau)$ and $t_{uv}(\tau)$ are nonzero, we have $\frac{\pi_{uv}(\tau)}{t_{uv}(\tau)} \ge \alpha$. This is because

$$\begin{aligned} \pi_{uv}(\tau) &= \mathbb{E}_{(x,y)\sim D_{i,u,v,z}}\Big(\Pr[\Pi(x,y) \text{ accepts}] \cdot \Pr\big[\Pi(x,y) = \tau \;\big|\; \Pi(x,y) \text{ accepts}\big]\Big) \\ &\ge \alpha \cdot \mathbb{E}_{(x,y)\sim D_{i,u,v,z}} \Pr\big[\Pi(x,y) = \tau \;\big|\; \Pi(x,y) \text{ accepts}\big] \\ &= \alpha \cdot t_{uv}(\tau) \end{aligned}$$

  since support$(D_{i,u,v,z}) \subseteq F^{-1}(1)$ and Π is correct.

For $(u,v) \in \{(2,1),(0,2)\}$ define $\gamma_{uv}(\tau) := |t_{uv}(\tau) - t_{01}(\tau)|$ and note that

$$\sum_{\text{accepting } \tau} \gamma_{uv}(\tau) = 2\|T(D_{i,u,v,z}) - T(D_{i,0,1,z})\| < 2\gamma. \tag{4.1}$$

Claim 4.2. *For all accepting τ,*

$$\frac{t_{21}(\tau)\cdot t_{02}(\tau)}{t_{01}(\tau)} \geq t_{01}(\tau)-\gamma_{21}(\tau)-\gamma_{02}(\tau)\,. \tag{4.2}$$

Proof of Claim. It suffices to show that

$$t_{21}(\tau)\cdot t_{02}(\tau) \geq t_{01}(\tau)^2 - t_{01}(\tau)\big(\gamma_{21}(\tau)+\gamma_{02}(\tau)\big)\,. \tag{4.3}$$

(If $t_{01}(\tau)\neq 0$ then (4.2) follows by dividing (4.3) by $t_{01}(\tau)$, and if $t_{01}(\tau)=0$ then (4.2) follows since its right side is nonpositive and its left side is 0; recall our convention that $0/0 := 0$.) For some signs $\sigma_{uv}(\tau)\in\{1,-1\}$, the left side of (4.3) equals

$$\big(t_{01}(\tau)+\sigma_{21}(\tau)\gamma_{21}(\tau)\big)\cdot\big(t_{01}(\tau)+\sigma_{02}(\tau)\gamma_{02}(\tau)\big)\,,$$

which expands to

$$t_{01}(\tau)^2+\sigma_{21}(\tau)t_{01}(\tau)\gamma_{21}(\tau)+\sigma_{02}(\tau)t_{01}(\tau)\gamma_{02}(\tau)+\sigma_{21}(\tau)\sigma_{02}(\tau)\gamma_{21}(\tau)\gamma_{02}(\tau)\,. \tag{4.4}$$

If $\sigma_{21}(\tau)=\sigma_{02}(\tau)$ then (4.4) is at least the right side of (4.3) since the last term of (4.4) is nonnegative. If $\sigma_{21}(\tau)\neq\sigma_{02}(\tau)$, say $\sigma_{21}(\tau)=-1$ and $\sigma_{02}(\tau)=1$, then (4.4) is at least the right side of (4.3) since the sum of the last two terms in (4.4) is $t_{01}(\tau)\gamma_{02}(\tau)-\gamma_{21}(\tau)\gamma_{02}(\tau)=t_{21}(\tau)\gamma_{02}(\tau)\geq 0$. $\square$

We have

$$\begin{aligned}
\Pr\big[\Pi(D_{i,2,2,z})\text{ accepts}\big] &= \sum_{\text{accepting }\tau}\pi_{22}(\tau)\\
&\geq \sum_{\tau\in S}\pi_{22}(\tau)\\
&\geq \sum_{\tau\in S}\frac{\pi_{21}(\tau)\cdot\pi_{02}(\tau)}{\pi_{01}(\tau)}\\
&= \sum_{\tau\in S}\frac{\frac{\pi_{21}(\tau)}{t_{21}(\tau)}\cdot\frac{\pi_{02}(\tau)}{t_{02}(\tau)}}{\frac{\pi_{01}(\tau)}{t_{01}(\tau)}}\cdot\frac{t_{21}(\tau)\cdot t_{02}(\tau)}{t_{01}(\tau)}\\
&\geq \sum_{\tau\in S}\alpha\cdot\big(t_{01}(\tau)-\gamma_{21}(\tau)-\gamma_{02}(\tau)\big)\\
&> \alpha\cdot\Big(\frac{1-\gamma}{2}-2\gamma-2\gamma\Big)\\
&> \alpha/4\,.
\end{aligned}$$

To see that the fifth line follows from the fourth, consider each $\tau\in S$: If $t_{21}(\tau)\neq 0$ and $t_{02}(\tau)\neq 0$ then it follows by

$$\frac{\pi_{21}(\tau)}{t_{21}(\tau)}\geq\alpha \quad\text{and}\quad \frac{\pi_{02}(\tau)}{t_{02}(\tau)}\Big/\frac{\pi_{01}(\tau)}{t_{01}(\tau)}\geq 1$$

(since $\tau \in S$) and Claim 4.2. On the other hand, if $t_{21}(\tau) = 0$ or $t_{02}(\tau) = 0$, say, $t_{21}(\tau) = 0$, then it follows since the summand on the fourth line is 0, and $t_{01}(\tau) - \gamma_{21}(\tau) = 0$ so the summand on the fifth line is nonpositive. The sixth line follows from the fifth by

$$\sum_{\tau \in S} t_{01}(\tau) = \Pr\big[T(D_{i,0,1,z}) \in S\big] \geq \frac{1-\gamma}{2} \quad \text{and} \quad \sum_{\tau \in S} \gamma_{21}(\tau) \leq \sum_{\text{accepting } \tau} \gamma_{21}(\tau)$$

and (4.1), and similarly for γ_{02}.

We conclude that $\Pr\big[\Pi(x,y) \text{ accepts}\big] > \alpha/4$ for some $(x,y) \in \operatorname{support}(D_{i,2,2,z}) \subseteq F^{-1}(0)$, contradicting the correctness of Π. This finishes the proof of Theorem 1.3.

4.2 Proof of Lemma 4.1

Define jointly distributed random variables $X = X_1 \cdots X_n \in \{0,1,2\}^n$, $Y = Y_1 \cdots Y_n \in \{1,2\}^n$, and $Z = Z_1 \cdots Z_n \in \{\text{A},\text{B},\text{C}\}^n$ as follows: Z is uniform, and given a particular choice of Z, for each $i \in [n]$ (independently) (X_i, Y_i) is sampled from Q_{Z_i}. Thus the marginal distribution of (X,Y) is that for each i (independently), (X_i,Y_i) has probability $1/3$ for each of $(0,1),(0,2)$, and probability $1/6$ for each of $(2,1),(1,2)$. Since the support of (X,Y) is in $F^{-1}(1)$, we may also view T as a random variable distributed jointly with (X,Y,Z). Let Z_{-i} denote $Z_1 \cdots Z_{i-1} Z_{i+1} \cdots Z_n$.

For any $i \in [n]$, $z_i \in \{\text{A},\text{B},\text{C}\}$, and $z_{-i} \in \{\text{A},\text{B},\text{C}\}^{n-1}$ indexed by $[n] \backslash \{i\}$, we can view

$$\big(T(D_{i,Q_{z_i},z_{-i}}), Q_{z_i}\big)$$

as a pair of jointly distributed random variables that is distributed identically to

$$\big(T, (X_i,Y_i) \;\big|\; Z_i = z_i, Z_{-i} = z_{-i}\big)\,.$$

For all i and z_{-i}, by a standard lemma (see [4, Lemma 6.2 and Proposition 6.10]) we have

$$\mathbb{I}\big(T(D_{i,Q_\text{A},z_{-i}}); Q_\text{A}\big) \geq \|T(D_{i,0,1,z_{-i}}) - T(D_{i,0,2,z_{-i}})\|^2/2$$

and similarly for B and C. Therefore

$$\begin{aligned}
\text{Maximum length of transcript} &\geq \mathbb{H}(T \mid Z) \\
&\geq \mathbb{I}(T; X,Y \mid Z) \\
&\geq \sum_i \mathbb{I}(T; X_i, Y_i \mid Z) \\
&= \sum_i \mathbb{E}_{z_{-i}} \frac{1}{3} \sum_{z_i} \mathbb{I}\big(T; X_i, Y_i \;\big|\; Z_i = z_i, Z_{-i} = z_{-i}\big) \\
&= \sum_i \mathbb{E}_{z_{-i}} \frac{1}{3} \sum_{z_i} \mathbb{I}\big(T(D_{i,Q_{z_i},z_{-i}}); Q_{z_i}\big) \\
&\geq \sum_i \mathbb{E}_{z_{-i}} \frac{1}{3} \cdot \frac{1}{2} \Big(\|T(D_{i,0,1,z_{-i}}) - T(D_{i,0,2,z_{-i}})\|^2 \\
&\qquad + \|T(D_{i,0,1,z_{-i}}) - T(D_{i,2,1,z_{-i}})\|^2 \\
&\qquad + \|T(D_{i,0,2,z_{-i}}) - T(D_{i,1,2,z_{-i}})\|^2 \Big)
\end{aligned}$$

$$\begin{aligned} &\geq \sum_i \mathbb{E}_{z_{-i}}(\gamma^2/6) \\ &= n\gamma^2/6\,. \end{aligned}$$

where the third line follows by a standard direct sum property for conditional information cost [4].

5 Extended formulations for Maximum-Clique

Recall that the *nonnegative rank* of a nonnegative $\ell \times m$ matrix M, denoted $\mathrm{rank}_+(M)$, is the least r such that $M = UV$ for some nonnegative matrices U and V of dimensions $\ell \times r$ and $r \times m$, respectively. We say M is an ε-UDISJ matrix (where ε is to be thought of as a function of n, and UDISJ stands for UNIQUE-SET-DISJOINTNESS) iff it is $2^n \times 2^n$ with rows and columns indexed by subsets $x, y \subseteq [n]$, each entry with $|x \cap y| = 0$ is 1, each entry with $|x \cap y| = 1$ is $1 - \varepsilon$, and all other entries are nonnegative. The authors of [6] proved the following two things.

(i) For every ε-UDISJ matrix M, $\mathrm{rank}_+(M) \geq 2^{\Omega(\varepsilon^2 n)}$.
(ii) For some ε-UDISJ matrix M, $\mathrm{rank}_+(M)$ lower bounds the complexity of $1/\varepsilon$-approximate extended formulations for a certain natural linear programming encoding of the maximum-clique problem for graphs.

The lower bound in (i) was improved to $2^{\Omega(\varepsilon n)}$ in [8], and a simpler alternative proof of this stronger bound was given in [7]. Combining such nonnegative rank lower bounds with (ii) yields lower bounds on the approximate extended formulation complexity for maximum-clique. It is not known how to exploit the additional structure of the matrix M from the proof of (ii) when lower bounding the nonnegative rank.

The paper [6] proved (i) by developing a generalization of the corruption lemma from [35]. Both of the subsequent papers [8, 7] employed information-theoretic methods. We now show that an idea analogous to our and-amplification-based communication lower bound proof (Theorem 1.4 and Theorem 1.5) can be used to “bootstrap” (i) to get the stronger $2^{\Omega(\varepsilon n)}$ lower bound, thus bypassing the information-theoretic methods of [8, 7].

Given a nonnegative matrix M, let M^k denote the *entrywise* k-th power of M. It is an elementary fact that $\mathrm{rank}_+(M^k) \leq \mathrm{rank}_+(M)^k$. If M is an ε-UDISJ matrix, then choosing $k = \Theta(1/\varepsilon)$, we find that M^k is a c-UDISJ matrix for some $c \geq 1/2$. Then by (i) we have $\mathrm{rank}_+(M^k) \geq 2^{\Omega(n)}$ and thus

$$\mathrm{rank}_+(M) \geq \mathrm{rank}_+(M^k)^{1/k} \geq \left(2^{\Omega(n)}\right)^{1/k} \geq 2^{\Omega(\varepsilon n)}\,.$$

6 Subsequent developments

It has been proven in [21] that $\mathsf{SBP}(f) \leq O(\mathsf{USBP}(f) + \log n)$ for all f (even partial f) using a new and elementary “Truncation Lemma” which enables rectangle-based techniques to be applied to low nonnegative rank matrices in certain situations. Combining this result with Corollary 1.2 yields an alternative proof of Theorem 1.3 and shows the class equality $\mathsf{SBP} = \mathsf{USBP}$.

Furthermore, this result implies that the tight extended formulation lower bound for maximum-clique (discussed in Section 1.5 and Section 5) can be derived from Razborov’s corruption lemma [35] in a black-box way, without needing any of the machinery in [6, 8, 7]: Specifically, any ε-UDISJ matrix having nonnegative rank r can be interpreted as witnessing

$$\mathsf{R}^{\mathsf{priv}}_{\alpha,(1-\varepsilon)\alpha}(\text{UDISJ}) \leq \log r + O(1)$$

(for some $\alpha > 0$). This protocol can be and-amplified to witness $\mathsf{USBP}(\text{UDISJ}) \leq O((\log r)/\varepsilon)$, and thus $\mathsf{SBP}(\text{UDISJ}) \leq O((\log r)/\varepsilon + \log n)$. Finally, Corollary 1.2, which is an elementary consequence of Razborov's corruption lemma, implies that $\log r \geq \Omega(\varepsilon n)$.

It has also been proven in [21] that $\mathsf{MA} \neq \mathsf{SBP}$, solving another open problem posed in the original version of this paper [22].

Acknowledgements

We thank Mark Braverman, Tom Gur, Raghu Meka, Toniann Pitassi, and anonymous reviewers for comments and discussions.

References

[1] SCOTT AARONSON AND AVI WIGDERSON: Algebrization: A new barrier in complexity theory. *ACM Trans. Comput. Theory*, 1(1):2:1–2:54, 2009. Prelimianry version in STOC'08. [doi:10.1145/1490270.1490272] 4

[2] LÁSZLÓ BABAI, PETER FRANKL, AND JANOS SIMON: Complexity classes in communication complexity theory. In *Proc. 27th FOCS*, pp. 337–347. IEEE Comp. Soc. Press, 1986. [doi:10.1109/SFCS.1986.15] 2, 3

[3] LÁSZLÓ BABAI AND SHLOMO MORAN: Arthur–Merlin games: A randomized proof system, and a hierarchy of complexity classes. *J. Comput. System Sci.*, 36(2):254–276, 1988. [doi:10.1016/0022-0000(88)90028-1] 3

[4] ZIV BAR-YOSSEF, T. S. JAYRAM, RAVI KUMAR, AND D. SIVAKUMAR: An information statistics approach to data stream and communication complexity. *J. Comput. System Sci.*, 68(4):702–732, 2004. Preliminary version in FOCS'02. [doi:10.1016/j.jcss.2003.11.006] 2, 5, 12, 14, 17, 18

[5] ELMAR BÖHLER, CHRISTIAN GLAßER, AND DANIEL MEISTER: Error-bounded probabilistic computations between MA and AM. *J. Comput. System Sci.*, 72(6):1043–1076, 2006. Preliminary version in MFCS'03. [doi:10.1016/j.jcss.2006.05.001] 3

[6] GÁBOR BRAUN, SAMUEL FIORINI, SEBASTIAN POKUTTA, AND DAVID STEURER: Approximation limits of linear programs (beyond hierarchies). *Math. Oper. Res.*, 40(3):756–772, 2015. Preliminary version in FOCS'12. [doi:10.1287/moor.2014.0694, arXiv:1204.0957] 6, 18, 19

[7] GÁBOR BRAUN AND SEBASTIAN POKUTTA: Common information and unique disjointness. *Algorithmica*, pp. 1–33, 2016. Preliminary versions in FOCS'13 and ECCC. [doi:10.1007/s00453-016-0132-0] 6, 18, 19

[8] MARK BRAVERMAN AND ANKUR MOITRA: An information complexity approach to extended formulations. In *Proc. 45th STOC*, pp. 161–170. ACM Press, 2013. Preliminary version in ECCC. [doi:10.1145/2488608.2488629] 2, 5, 6, 13, 18, 19

[9] MARK BRAVERMAN AND OMRI WEINSTEIN: A discrepancy lower bound for information complexity. In *Proc. 16th Internat. Workshop on Randomization and Computation (RANDOM'12)*, pp. 459–470. Springer, 2012. Preliminary version in ECCC. [doi:10.1007/978-3-642-32512-0_39, arXiv:1112.2000] 4

[10] HARRY BUHRMAN, NIKOLAI VERESHCHAGIN, AND RONALD DE WOLF: On computation and communication with small bias. In *Proc. 22nd IEEE Conf. on Computational Complexity (CCC'07)*, pp. 24–32. IEEE Comp. Soc. Press, 2007. [doi:10.1109/CCC.2007.18] 3

[11] AMIT CHAKRABARTI, GRAHAM CORMODE, ANDREW MCGREGOR, AND JUSTIN THALER: Annotations in data streams. *ACM Trans. Algorithms*, 11(1):7, 2014. Preliminary version in ECCC. [doi:10.1145/2636924] 4

[12] AMIT CHAKRABARTI, GRAHAM CORMODE, ANDREW MCGREGOR, JUSTIN THALER, AND SURESH VENKATASUBRAMANIAN: Verifiable stream computation and Arthur–Merlin communication. In *Proc. 30th IEEE Conf. on Computational Complexity (CCC'15)*, pp. 217–243. Schloss Dagstuhl, 2015. Preliminary version in ECCC. [doi:10.4230/LIPIcs.CCC.2015.217] 4

[13] AMIT CHAKRABARTI AND ODED REGEV: An optimal lower bound on the communication complexity of Gap-Hamming-Distance. *SIAM J. Comput.*, 41(5):1299–1317, 2012. Preliminary version in STOC'11. [doi:10.1137/120861072, arXiv:1009.3460] 5

[14] AMIT CHAKRABARTI, YAOYUN SHI, ANTHONY WIRTH, AND ANDREW YAO: Informational complexity and the direct sum problem for simultaneous message complexity. In *Proc. 42nd FOCS*, pp. 270–278. IEEE Comp. Soc. Press, 2001. [doi:10.1109/SFCS.2001.959901] 5, 12

[15] ARKADEV CHATTOPADHYAY AND TONIANN PITASSI: The story of set disjointness. *ACM SIGACT News*, 41(3):59–85, 2010. [doi:10.1145/1855118.1855133] 2

[16] THOMAS COVER AND JOY THOMAS: *Elements of Information Theory*. Wiley, 2006. 12

[17] SCOTT DIEHL: Lower bounds for swapping Arthur and Merlin. In *Proc. 11th Internat. Workshop on Randomization and Computation (RANDOM'07)*, pp. 449–463. Springer, 2007. [doi:10.1007/978-3-540-74208-1_33] 4

[18] JÜRGEN FORSTER: A linear lower bound on the unbounded error probabilistic communication complexity. *J. Comput. System Sci.*, 65(4):612–625, 2002. Preliminary version in CCC'01. [doi:10.1016/S0022-0000(02)00019-3] 5, 8

[19] ANAT GANOR, GILLAT KOL, AND RAN RAZ: Exponential separation of information and communication for boolean functions. In *Proc. 47th STOC*, pp. 557–566. ACM Press, 2015. Preliminary version in ECCC. [doi:10.1145/2746539.2746572] 4

[20] SHAFI GOLDWASSER AND MICHAEL SIPSER: Private coins versus public coins in interactive proof systems. In *Proc. 18th STOC*, pp. 59–68. ACM Press, 1986. [doi:10.1145/12130.12137] 4

[21] MIKA GÖÖS, SHACHAR LOVETT, RAGHU MEKA, THOMAS WATSON, AND DAVID ZUCKERMAN: Rectangles are nonnegative juntas. In *Proc. 47th STOC*, pp. 257–266. ACM Press, 2015. Preliminary version in ECCC. [doi:10.1145/2746539.2746596] 18, 19

[22] MIKA GÖÖS AND THOMAS WATSON: Communication complexity of set-disjointness for all probabilities. In *Proc. 18th Internat. Workshop on Randomization and Computation (RANDOM'14)*, pp. 721–736. Schloss Dagstuhl, 2014. Preliminary versin in ECCC. [doi:10.4230/LIPIcs.APPROX-RANDOM.2014.721] 1, 19

[23] TOM GUR AND RAN RAZ: Arthur–Merlin streaming complexity. *Inform. and Comput.*, 243:145–165, 2015. Preliminary version in ICALP'13. [doi:10.1016/j.ic.2014.12.011, arXiv:1302.0418] 4

[24] RAHUL JAIN AND HARTMUT KLAUCK: The partition bound for classical communication complexity and query complexity. In *Proc. 25th IEEE Conf. on Computational Complexity (CCC'10)*, pp. 247–258. IEEE Comp. Soc. Press, 2010. [doi:10.1109/CCC.2010.31, arXiv:0910.4266] 4

[25] RAHUL JAIN, TROY LEE, AND NISHEETH VISHNOI: A quadratically tight partition bound for classical communication complexity and query complexity, 2014. [arXiv:1401.4512] 4

[26] STASYS JUKNA: *Boolean Function Complexity: Advances and Frontiers*. Volume 27 of *Algorithms and Combinatorics*. Springer, 2012. [doi:10.1007/978-3-642-24508-4] 2, 3

[27] BALA KALYANASUNDARAM AND GEORG SCHNITGER: The probabilistic communication complexity of set intersection. *SIAM J. Discrete Math.*, 5(4):545–557, 1992. Preliminary version in SCT'87. [doi:10.1137/0405044] 2, 5

[28] IORDANIS KERENIDIS, SOPHIE LAPLANTE, VIRGINIE LERAYS, JÉRÉMIE ROLAND, AND DAVID XIAO: Lower bounds on information complexity via zero-communication protocols and applications. *SIAM J. Comput.*, 44(5):1550–1572, 2015. Preliminary versions in FOCS'12 and ECCC. [doi:10.1137/130928273, arXiv:1204.1505] 4

[29] HARTMUT KLAUCK: Rectangle size bounds and threshold covers in communication complexity. In *Proc. 18th IEEE Conf. on Computational Complexity (CCC'03)*, pp. 118–134. IEEE Comp. Soc. Press, 2003. [doi:10.1109/CCC.2003.1214415, arXiv:cs/0208006] 4, 5, 9

[30] HARTMUT KLAUCK: Lower bounds for quantum communication complexity. *SIAM J. Comput.*, 37(1):20–46, 2007. Preliminary version in FOCS'01. [doi:10.1137/S0097539702405620, arXiv:quant-ph/0106160] 5

[31] HARTMUT KLAUCK: On Arthur Merlin games in communication complexity. In *Proc. 26th IEEE Conf. on Computational Complexity (CCC'11)*, pp. 189–199. IEEE Comp. Soc. Press, 2011. [doi:10.1109/CCC.2011.33, arXiv:1101.0523] 4

[32] EYAL KUSHILEVITZ AND NOAM NISAN: *Communication Complexity*. Cambridge Univ. Press, 1997. 2, 3

[33] ILAN NEWMAN: Private vs. common random bits in communication complexity. *Inform. Process. Lett.*, 39(2):67–71, 1991. [doi:10.1016/0020-0190(91)90157-D] 3

[34] RAMAMOHAN PATURI AND JANOS SIMON: Probabilistic communication complexity. *J. Comput. System Sci.*, 33(1):106–123, 1986. Preliminary version in FOCS'84. [doi:10.1016/0022-0000(86)90046-2] 3

[35] ALEXANDER A. RAZBOROV: On the distributional complexity of disjointness. *Theoret. Comput. Sci.*, 106(2):385–390, 1992. Preliminary version in ICALP'90. [doi:10.1016/0304-3975(92)90260-M] 2, 5, 6, 18

[36] ALEXANDER A. SHERSTOV: Halfspace matrices. *Comput. Complexity*, 17(2):149–178, 2008. Preliminary version in CCC'07. [doi:10.1007/s00037-008-0242-4] 3

[37] ALEXANDER A. SHERSTOV: The pattern matrix method. *SIAM J. Comput.*, 40(6):1969–2000, 2011. Preliminary version in STOC'08. [doi:10.1137/080733644, arXiv:0906.4291] 3

[38] ALEXANDER A. SHERSTOV: The unbounded-error communication complexity of symmetric functions. *Combinatorica*, 31(5):583–614, 2011. Preliminary version in FOCS'08. [doi:10.1007/s00493-011-2580-0] 8

[39] ALEXANDER A. SHERSTOV: The communication complexity of Gap Hamming Distance. *Theory of Computing*, 8(8):197–208, 2012. Preliminary version in ECCC. [doi:10.4086/toc.2012.v008a008] 5

[40] THOMAS VIDICK: A concentration inequality for the overlap of a vector on a large set, with application to the communication complexity of the Gap-Hamming-Distance problem. *Chicago J. Theoret. Comput. Sci.*, 2012(1):1–12, 2012. [doi:10.4086/cjtcs.2012.001] 5

[41] THOMAS WATSON: The complexity of estimating min-entropy. *Comput. Complexity*, 25(1):153–175, 2016. Preliminary version in ECCC. [doi:10.1007/s00037-014-0091-2] 3

AUTHORS

Mika Göös
Graduate Student
University of Toronto, Toronto, ON
mika.goos@mail.utoronto.edu
http://www.cs.utoronto.ca/~mgoos/

Thomas Watson
Assistant Professor
University of Memphis, Memphis, TN
Thomas.Watson@memphis.edu
http://www.cs.toronto.edu/~thomasw/

ABOUT THE AUTHORS

MIKA GÖÖS is a Ph. D. student in the Theory Group at the Department of Computer Science at the University of Toronto, advised by Toniann Pitassi. His research interests lie in computational complexity with a particular fondness for distributed graph algorithms and communication complexity.

THOMAS WATSON is an assistant professor of computer science at the University of Memphis. Previously, he was a postdoc at the University of Toronto, supervised by Toniann Pitassi. He received his Ph. D. in computer science from the University of California, Berkeley, supervised by Luca Trevisan. He received his undergraduate degree in computer science, math, and computer engineering from the University of Wisconsin-Madison, where his interest in theoretical computer science research was sparked and fostered by Dieter van Melkebeek. His research interests include most of complexity theory, but recently he has been working on communication complexity.