The Landscape of Communication Complexity Classes*

Mika Göös[1], Toniann Pitassi[2], and Thomas Watson[3]

1 **University of Toronto, Computer Science Department, Toronto, ON, Canada**
mika.goos@mail.utoronto.ca
2 **University of Toronto, Computer Science Department, Toronto, ON, Canada**
toni@cs.toronto.edu
3 **University of Toronto, Computer Science Department, Toronto, ON, Canada**
thomasw@cs.toronto.edu

Abstract

We prove several results which, together with prior work, provide a nearly-complete picture of the relationships among classical communication complexity classes between P and PSPACE, short of proving lower bounds against classes for which no explicit lower bounds were already known. Our article also serves as an up-to-date survey on the state of structural communication complexity.

Among our new results we show that $\mathsf{MA} \not\subseteq \mathsf{ZPP}^{\mathsf{NP}[1]}$, that is, Merlin–Arthur proof systems cannot be simulated by zero-sided error randomized protocols with one NP query. Here the class $\mathsf{ZPP}^{\mathsf{NP}[1]}$ has the property that generalizing it in the slightest ways would make it contain $\mathsf{AM} \cap \mathsf{coAM}$, for which it is notoriously open to prove any explicit lower bounds. We also prove that $\mathsf{US} \not\subseteq \mathsf{ZPP}^{\mathsf{NP}[1]}$, where US is the class whose canonically complete problem is the variant of set-disjointness where yes-instances are uniquely intersecting. We also prove that $\mathsf{US} \not\subseteq \mathsf{coDP}$, where DP is the class of differences of two NP sets. Finally, we explore an intriguing open issue: are rank-1 matrices inherently more powerful than rectangles in communication complexity? We prove a new separation concerning PP that sheds light on this issue and strengthens some previously known separations.

1998 ACM Subject Classification F.1.3 Complexity Measures and Classes

Keywords and phrases Landscape, communication, complexity, classes

Digital Object Identifier 10.4230/LIPIcs.ICALP.2016.86

1 Introduction

Complexity classes form the infrastructure of classical complexity theory. They are used to express the power of models of computation, characterize the complexities of important computational problems, and catalyze proofs of other results. A central project is to ascertain the full, intricate landscape of relationships among complexity classes.

Beginning with [3], there has been a lot of research on the analogues of classical (Turing machine) complexity classes in two-party communication complexity. The analogue of P (the class of decision problems solvable in polynomial time) is the class of functions $F\colon \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$ for which Alice and Bob, given x and y respectively, can evaluate $F(x,y)$ with a protocol that uses polylogarithmically many bits of communication. For other classical complexity classes representing other models of computation, one can generally define,

* The full version is available at http://eccc.hpi-web.de/report/2015/049/. This work was supported by NSERC funding.

43rd International Colloquium on Automata, Languages, and Programming (ICALP 2016).
Editors: Ioannis Chatzigiannakis, Michael Mitzenmacher, Yuval Rabani, and Davide Sangiorgi;
Article No. 86; pp. 86:1–86:15

Leibniz International Proceedings in Informatics
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

in a canonical way, associated communication complexity classes representing associated models of communication. There are many motivations for studying the relationships (inclusions and non-inclusions) between these communication complexity classes.

- A holy grail of classical complexity is to prove separations of classes between P and PSPACE. Separations relative to oracles can often be viewed as class separations in the restricted setting of *query complexity*; see [52] for an excellent survey. Communication complexity can be viewed as a restricted (but generally less restricted than query complexity) setting for which lower bounds are more difficult to obtain. Such separations in restricted settings are sometimes construed as evidence for the classical separations, or at least as barriers to refuting the classical separations. A stronger form of relativization barriers is known as *algebrization* [2], which directly employs communication complexity class separations.
- Proving lower bounds against strong communication complexity classes has applications to other areas of theoretical computer science. One of the most notorious open problems in communication complexity is to prove lower bounds against the analogue of the polynomial hierarchy (PH) for any explicit two-party function. Proving PH lower bounds is a necessary step for obtaining strong rank rigidity lower bounds [44, 36, 37, 53] (as well as margin complexity rigidity lower bounds [35]), which in turn are related to circuit complexity [50]. Lower bounds against PH are also related to graph complexity [42, 25]. It even remains open to prove communication lower bounds against the subclass of PH known as AM (Arthur–Merlin games) for any explicit function (which would be relevant to streaming delegation [8, 31, 19, 7, 9, 32]).
- Communication complexity has a menagerie of techniques for proving lower bounds (among the oldest being discrepancy and corruption). These techniques often provide lower bounds against powerful communication complexity classes, and in some cases turn out to be *equivalent* to the communication measures corresponding to those classes (e.g., discrepancy is equivalent to PP communication [29], and corruption is equivalent to SBP communication [18]). See [17] for more background on this. Thus, by studying complexity classes, as a byproduct we study the relative strength of lower bound techniques.
- The various models of communication corresponding to complexity classes are mathematically interesting because protocols in these models can be viewed as succinct representations of boolean matrices. The study of classes exposes natural questions about the combinatorial power of such succinct representations.

We contribute to the exploration of the communication complexity landscape by filling in many of the remaining gaps in the known relationships among classes, and discovering new techniques and insights along the way. In Section 2 we state our results more precisely and provide some intuition for the proofs. In the full version, we summarize the state of affairs (including our new results) by showing a map of known inclusions and non-inclusions between pairs of traditional communication classes, and we provide a comprehensive survey of these results. This updates previous surveys by Babai, Frankl, and Simon [3] and Halstenberg and Reischuk [21].

We refer to [33, 26] for background on communication complexity. In the full version we provide a catalog of communication complexity class definitions; throughout the text, we provide definitions on a "need-to-know" basis. If $\mathcal{C}$ is the name of a model (e.g., P for deterministic or NP for nondeterministic), we follow the convention of using $\mathcal{C}$ to denote both a complexity class and the corresponding complexity measure: $\mathcal{C}(F)$ denotes the minimum cost of a correct protocol for the (possibly partial) two-party function F in model $\mathcal{C}$, and $\mathcal{C}$ denotes the class of all (families of) partial functions F with $\mathcal{C}(F) \leq \mathrm{poly}(\log n)$.

2 Our Contributions

Several of our results concern two-party composed functions, so we introduce some general notation for this. A *composed function* is of the form $f \circ g^m$ where $f\colon \{0,1\}^m \to \{0,1\}$ is a (possibly partial) *outer function* and $g\colon \{0,1\}^b \times \{0,1\}^b \to \{0,1\}$ is an *inner function* also called a *gadget*. We write $F := f \circ g^m\colon \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$ where $n := m \cdot b$. We view the inputs to Alice and Bob as $x, y \in (\{0,1\}^b)^m$, which are partitioned into *blocks* $x_i, y_i \in \{0,1\}^b$ for $i \in [m]$. The goal is to compute $F(x,y) := f(g(x_1,y_1),\ldots,g(x_m,y_m))$.

2.1 MA $\not\subseteq$ ZPP$^{\mathsf{NP}[1]}$

A Merlin–Arthur (MA) communication protocol is a proof system in which a nondeterministic party called Merlin sends a proof string (depending on the input) to Alice and Bob (collectively constituting Arthur), who then execute a randomized protocol to verify the proof. Merlin–Arthur communication protocols have been studied many times [28, 43, 2, 16, 30, 19, 20], starting with the work of Klauck [28], who gave a $\Omega(\sqrt{n})$ lower bound on the MA communication complexity of set-disjointness. In contrast, for the related (and stronger) model of Arthur–Merlin (AM) communication protocols, in which Merlin's proof string may depend on Alice's and Bob's randomness, no nontrivial lower bound is known for any explicit function, and such lower bounds have become very sought-after in the recent literature [35, 40, 32, 9].

Our first result concerns the relationship between MA and another class, ZPP$^{\mathsf{NP}[1]}$, which is a slightly obscure but intriguing character with many curious properties. A ZPP-type protocol is randomized and may output the correct answer or $\perp$ (representing "don't know"), and must output the correct answer with high probability on every input; granting the protocol access to one query to an NP oracle yields ZPP$^{\mathsf{NP}[1]}$. It is not a priori clear that the model is robust with respect to the choice of threshold for the success probability, since standard amplification by repetition would increase the number of NP oracle queries. However, it was shown in [11] that ZPP$^{\mathsf{NP}[1]}$ does indeed admit efficient amplification as long as the success probability is $> 1/2$ (the proof for time-bounded complexity also works for communication complexity); hence we define the model with success probability some constant $> 1/2$, say $3/4$.

If we allowed ZPP$^{\mathsf{NP}[1]}$ to have success probability $< 1/2$, the class would change drastically: it would contain AM $\cap$ coAM (see the full version), and hence proving explicit lower bounds for the communication version would yield breakthrough AM communication lower bounds. Granting the model access to two nonadaptive NP queries (and requiring success probability $> 1/2$) would also encompass AM $\cap$ coAM. Thus, in a sense, ZPP$^{\mathsf{NP}[1]}$ represents a boundary beyond which AM lower bounds would be the next step. The class ZPP$^{\mathsf{NP}[1]}$ is also sandwiched between BPP and $\mathsf{S}_2\mathsf{P}$ [6]; $\mathsf{S}_2\mathsf{P}$ is a subclass of the polynomial hierarchy that has not been studied before in communication complexity (the definition appears in the full version), and no nontrivial lower bounds against it are known for any explicit function. This is another sense in which ZPP$^{\mathsf{NP}[1]}$ constitutes a new frontier toward the elusive goal of proving explicit PH communication lower bounds. We also mention that ZPP$^{\mathsf{NP}[1]}$ shows up frequently in the literature on the "two queries problem" (e.g., if $\mathsf{P}_{\|}^{\mathsf{NP}[2]} \subseteq$ ZPP$^{\mathsf{NP}[1]}$ then PH $= \mathsf{S}_2\mathsf{P}$ [49]).

We prove that MA $\not\subseteq$ ZPP$^{\mathsf{NP}[1]}$ in the setting of communication complexity. This can be interpreted as saying that one-round non-interactive[1] proof systems cannot be made to have

[1] Here, the term non-interactive means that Alice and Bob cannot interact with Merlin other than receiving the proof string.

zero-sided error, even if the proof is generalized to an NP oracle query that depends on the randomness.

Before officially stating the theorem, we give the relevant formal definitions. An MA communication protocol computing $F\colon \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$ consists of a randomized two-party protocol which takes as input, in addition to the usual inputs x and y, a proof string (witness) $w \in \{0,1\}^k$ that is visible to both Alice and Bob. The completeness criterion is that for every $(x,y) \in F^{-1}(1)$ there exists a w such that the protocol accepts with probability at least $3/4$, and the soundness criterion is that for every $(x,y) \in F^{-1}(0)$ and every w, the protocol rejects with probability at least $3/4$. The cost is the witness length k plus the length of the subsequent transcript between Alice and Bob.

A $\mathsf{ZPP}^{\mathsf{NP}[1]}$ protocol Π computing F is a distribution over $\mathsf{P}^{\mathsf{NP}[1]}$-type protocols, each of which is of the following form: There is a deterministic protocol where for each leaf v having associated rectangle R_v, there is also an associated collection of "witness rectangles" $\left\{S_{v,w} \subseteq R_v \ : \ w \in \{0,1\}^k\right\}$ and an associated "output function" $o_v\colon \{0,1\} \to \{0,1,\perp\}$. The output of the $\mathsf{P}^{\mathsf{NP}[1]}$-type protocol on input (x,y) is obtained by running the deterministic part to reach a leaf v, then applying o_v to the indicator of whether $(x,y) \in \bigcup_w S_{v,w}$. The correctness criterion is that for every $(x,y) \in F^{-1}$, $\mathbb{P}\big[\Pi(x,y) \in \{F(x,y),\perp\}\big] = 1$ and $\mathbb{P}\big[\Pi(x,y) = F(x,y)\big] \geq 3/4$. The cost is the witness length k plus the maximum communication cost of the deterministic part of any of the constituent $\mathsf{P}^{\mathsf{NP}[1]}$-type protocols. The result of [11] shows that changing the success probability from $3/4$ to any other constant strictly between $1/2$ and 1 would only change the measure $\mathsf{ZPP}^{\mathsf{NP}[1]}(F)$ by a constant factor.

We prove a lower bound for the block-equality function Block-Eq, defined as follows:[2] Given $\sqrt{n}$ instances of the equality function Eq of length $\sqrt{n}$, is at least one of them a yes-instance? More formally, we have Block-Eq $\coloneqq$ Or $\circ$ Eq^m where the input to Or is $m \coloneqq \sqrt{n}$ bits, and each input to Eq is $b \coloneqq \sqrt{n}$ bits. In other words, writing $x \coloneqq x_1 \cdots x_{\sqrt{n}} \in (\{0,1\}^{\sqrt{n}})^{\sqrt{n}}$ and $y \coloneqq y_1 \cdots y_{\sqrt{n}} \in (\{0,1\}^{\sqrt{n}})^{\sqrt{n}}$, we have Block-Eq$(x,y) = 1$ iff $x_i = y_i$ for some i. Note that Block-Eq $\in \mathsf{MA}$ since i can be nondeterministically guessed by Merlin, and then $x_i = y_i$ can be verified using a randomized protocol for Eq. (It was first noticed in [34] that Block-Eq $\in \Sigma_2\mathsf{P} \cap \Pi_2\mathsf{P}$, which is a superset of MA.)

▶ **Theorem 1.** $\mathsf{ZPP}^{\mathsf{NP}[1]}(\text{Block-Eq}) = \Theta(\sqrt{n})$, *and hence* $\mathsf{MA} \not\subseteq \mathsf{ZPP}^{\mathsf{NP}[1]}$.

To prove Theorem 1 (Section 3), we apply a new lower bound technique that combines the corruption bound with the 1-monochromatic rectangle size bound and asserts that they hold *simultaneously* (under the same distribution over inputs). We prove that, perhaps surprisingly, this combined technique gives a lower bound for $\mathsf{ZPP}^{\mathsf{NP}[1]}$ (though neither of the individual bounds suffices).

To apply our technique to Block-Eq, we first note that it is straightforward to achieve the two bounds separately: the 1-monochromatic rectangle size bound follows by simple counting, and the corruption bound follows by using Razborov's corruption lemma for the set-intersection function Inter [45] together with a simple reduction from Inter to Block-Eq. However, the latter does *not* result in a distribution satisfying the 1-monochromatic rectangle size bound for Block-Eq. To fix this problem, we argue that if we average Razborov's distribution over all ways of implementing the reduction (of which there are many), then the corruption bound is still satisfied, and now the 1-monochromatic rectangle size bound is also satisfied.

[2] The complement of block-equality is often known as list-non-equality.

2.2 US $\not\subseteq$ ZPP$^{\mathsf{NP}[1]}$

For the set-intersection function Inter, Alice and Bob are each given a subset of $[n]$ (and we identify the subset with its characteristic vector, a length-n bit string), and the goal is to output 1 when the sets are intersecting and 0 when they are disjoint.[3] Phrased as a composed function, Inter $:=$ Or $\circ$ And^n (for single-bit And). This is the canonical NP-complete problem in communication complexity, holding a comparable status to satisfiability, the canonical NP-complete problem in time-bounded complexity.

In the literature, "unique-set-intersection" commonly refers to the partial function version of Inter where the intersection is promised to have size 0 or 1. We propose a change in terminology, in order to be consistent with the following corresponding terminology from time-bounded complexity (see, e.g., [4, 51, 10]): Unique-satisfiability is the problem of determining whether the number of satisfying assignments of a formula is exactly 1, and is complete for the complexity class called US. Unambiguous-satisfiability is the problem of determining whether the number of satisfying assignments of a formula is 0 or 1 under the promise that one of these cases holds, and is complete for the complexity class called UP.

Therefore, we make the following declarations: Unique-set-intersection is the total function Unique-Inter: $\{0,1\}^n \times \{0,1\}^n \to \{0,1\}$ that maps (x,y) to 1 iff $|x \cap y| = 1$, i.e., Unique-Inter $:=$ Unique-Or $\circ$ And^n where Unique-Or$(z) = 1$ iff the Hamming weight of z is 1. Unambiguous-set-intersection is the partial function Unambig-Inter: $\{0,1\}^n \times \{0,1\}^n \to \{0,1\}$ that maps (x,y) to $|x \cap y|$ if the latter is in $\{0,1\}$, i.e., Unambig-Inter $:=$ Unambig-Or $\circ$ And^n where Unambig-Or(z) equals the Hamming weight of z if the latter is in $\{0,1\}$.

Note that Unique-Inter is US-complete, where a cost-k US communication protocol is defined as a collection of rectangles $\left\{R_w \subseteq \{0,1\}^n \times \{0,1\}^n \ : \ w \in \{0,1\}^k\right\}$, where on input (x,y) the output of the protocol is 1 iff (x,y) is in R_w for exactly one w.

▶ **Theorem 2.** ZPP$^{\mathsf{NP}[1]}$(Unique-Inter) $= \Theta(n)$, *and hence* US $\not\subseteq$ ZPP$^{\mathsf{NP}[1]}$.

We give two proofs of Theorem 2. Both proofs show that Theorem 2 holds even under the promise that the input sets intersect in at most two coordinates. Also, in both proofs, handling ZPP$^{\mathsf{NP}[1]}$ instead of P$^{\mathsf{NP}[1]}$ incurs almost no extra complication.

The first proof (Section 3) employs the same lower bound technique as in Theorem 1, but where we use Razborov's corruption lemma [45] directly (and we must do a little analysis to verify the 1-monochromatic rectangle size bound). The optional second proof (relegated to the full version) uses information complexity tools (including an adaptation of the "partial information cost" approach from [24]) and, although longer to write, has some minor advantages over the first proof: It is more self-contained, as it does not rely on the corruption lemma (only on some basic facts that are standard in information complexity). Also, it directly handles success probability $1/2 + \epsilon$ (for any constant $\epsilon > 0$) without relying on the amplification result of [11] (whereas the first proof assumes success probability 0.999).

2.3 US $\not\subseteq$ coDP

The class DP was introduced in [39] to capture the complexity of certain exact versions of optimization problems. A set (of all 1-inputs of a function) is in DP iff it is the difference between two NP sets. The classes P, NP, and DP are the 0^{th}, 1^{st}, and 2^{nd} (respectively) levels of the so-called boolean hierarchy.

[3] We let "set-disjointness" refer to the complementary function where 1-inputs are disjoint.

We have $\mathsf{US} \subseteq \mathsf{DP}$ since to check that there is exactly one witness, we can use an NP computation to check that there is at least one witness, and another to check that there are at least two witnesses, and require that the first computation returns 1 and the second returns 0. However, it is unlikely that $\mathsf{US} \subseteq \mathsf{coDP}$: [10] showed that this inclusion cannot hold in the classical time-bounded setting unless the polynomial hierarchy collapses. This result does not yield a communication separation, since it is unknown whether the polynomial hierarchy collapses in the communication setting. Nevertheless, we show that indeed $\mathsf{US} \not\subseteq \mathsf{coDP}$ in communication complexity.

Formally, a cost-k coDP communication protocol is defined as a pair of collections of rectangles, $\{S_w \subseteq \{0,1\}^n \times \{0,1\}^n \ : \ w \in \{0,1\}^k\}$ and $\{T_w \subseteq \{0,1\}^n \times \{0,1\}^n \ : \ w \in \{0,1\}^k\}$, where on input (x,y) the output is 0 iff $(x,y) \in \bigcup_w S_w \setminus \bigcup_w T_w$.

▶ **Theorem 3.** $\mathsf{coDP}(\textsc{Unique-Inter}) = \Theta(n)$*, and hence* $\mathsf{US} \not\subseteq \mathsf{coDP}$*.*

To prove Theorem 3 (Section 3), we show that the same lower bound technique we introduced for $\mathsf{ZPP}^{\mathsf{NP}[1]}$ (the combination of the corruption bound and the 1-monochromatic rectangle size bound) *also* lower bounds coDP complexity. Thus we can simply reuse the application of the technique to Unique-Inter from Theorem 2. (Reusing the application to Block-Eq from Theorem 1 would show that Block-Eq $\notin \mathsf{coDP}$, but in fact Block-Eq $\notin \mathsf{P}^{\mathsf{NP}} \supseteq \mathsf{coDP}$ was already known [23].)

2.4 $\mathsf{ZPP}^{\mathsf{NP}[1]} \subseteq \mathsf{PostBPP}$

Consider bounded-error randomized computations (like in BPP) but with *postselection*: the output may come from $\{0,1,\perp\}$ and must be correct with high probability *conditioned* on not outputting $\perp$ (and the probability of this conditioning event must be positive). The complexity class corresponding to this model was originally called $\mathsf{BPP}_{\mathsf{path}}$ [22], but the name $\mathsf{PostBPP}$ (inspired by [1]) has gained popularity in the recent literature ([17] is one example) and seems more appropriate, so we use it instead.

According to modern conventions, the standard way to define the cost of a $\mathsf{PostBPP}$ communication protocol for F would be as the communication cost plus $\log(1/\alpha)$, where α is the minimum over all $(x,y) \in F^{-1}$ of the probability of not outputting $\perp$. (Allowing public randomness and not charging for α would enable $\mathsf{PostBPP}$ protocols to compute every function with constant cost.) Similarly, the cost of a PP (i.e., unbounded-error randomized) protocol would be the communication cost plus $\log(1/\epsilon)$ where $1/2+\epsilon$ is the minimum over all $(x,y) \in F^{-1}$ of the probability of outputting the correct answer.

However, for reasons that will become clear in Section 2.5, we choose to revert to the original convention of [3] and define $\mathsf{PostBPP}$ and PP in a slightly different but equivalent way: we do not charge for α or ϵ but we require the public randomness to be uniformly distributed over $\{0,1\}^k$ and we charge for k. For both $\mathsf{PostBPP}$ and PP, this cost measure is equivalent to the above "modern" definition within a constant factor and additive $O(\log n)$ term, by standard sparsification of the public randomness [38].

Formally, we define a $\mathsf{PostBPP}$ communication protocol Π for F in the following succinct way: For each outcome of the public randomness (which is uniformly distributed over $\{0,1\}^k$) there is a deterministic protocol outputting values in $\{0,1,\perp\}$. For each $(x,y) \in F^{-1}$ we must have $\mathbb{P}\big[\Pi(x,y) = F(x,y)\big] > 2 \cdot \mathbb{P}\big[\Pi(x,y) = 1 - F(x,y)\big]$. The cost is the randomness length k plus the maximum communication cost of any of the constituent deterministic protocols.

A priori it is not clear that any explicit lower bounds for $\mathsf{ZPP}^{\mathsf{NP}[1]}$ follow from prior work. The following result shows that in fact they do, since many explicit lower bounds for $\mathsf{PostBPP}$ were known.

▶ **Theorem 4.** $\mathsf{PostBPP}(F) \leq O\big(\mathsf{ZPP}^{\mathsf{NP}[1]}(F) + \log n\big)$ *for all* F*, and hence* $\mathsf{ZPP}^{\mathsf{NP}[1]} \subseteq \mathsf{PostBPP}$*.*

It turns out that Theorem 4 can be derived from the lower bound technique we develop for $\mathsf{ZPP}^{\mathsf{NP}[1]}$ in Section 3; however, that approach is more complicated than necessary and, more importantly, is specific to communication complexity. We give a proof of Theorem 4 (in the full version) using a black-box simulation that also works for time-bounded complexity, without exploiting any special properties of communication.

Intuitively, the worst case for simulating a $\mathsf{ZPP}^{\mathsf{NP}[1]}$ protocol is the following situation: Whenever the NP oracle responds "0" the protocol outputs the right answer, and whenever the NP oracle responds "1" the protocol outputs $\perp$ but would have output the wrong answer if the response were "0". In this situation, pretending the oracle always responds "0" would yield a BPP protocol (this is where we crucially need the success probability to be $> 1/2$). To handle more general situations, we must also randomly guess and verify a witness for the NP query, outputting $\perp$ if the witness is invalid.

2.5 Open issue: Rank-1 vs. rectangles

The classes $\mathsf{PostBPP}$ and PP can be further generalized by allowing the use of private randomness, which does not count toward the cost. This gives rise to the so-called "unrestricted probabilities" classes $\mathsf{UPostBPP}$ (which was defined, but not extensively studied, in [17]) and UPP (which is well-studied [41, 13, 48, 46]). In $\mathsf{UPostBPP}$ and UPP we can dispense with public randomness altogether as the public coins could be tossed privately by Alice and then sent to Bob.

Combinatorially, $\mathsf{PostBPP}$ and PP protocols of cost c induce a distribution over 2^c labeled rectangles (rank-1 matrices with 0-1 entries) each occurring with a "restricted" probability of at least 2^{-c} (see the full version). In the case of $\mathsf{UPostBPP}$ and UPP there is a similar characterization with rectangles replaced by nonnegative rank-1 matrices (see the full version). A natural question arises:

> **Informal question:** *Are rank-1 matrices inherently more powerful than rectangles in communication complexity?*

While it has been shown that, e.g., $\mathsf{PP} \neq \mathsf{UPP}$ [5, 47], the known examples of functions $F \in \mathsf{UPP} \setminus \mathsf{PP}$ can actually be computed without exploiting the full power of private randomness (their rank-1 property): we can use a UPP protocol whose associated rank-1 matrices are still rectangles, but occurring with *unrestricted*, possibly tiny, probability. We conclude that "PP vs. UPP" is not the right way to formalize our informal question (and the existing proofs for $\mathsf{PP} \neq \mathsf{UPP}$ do not incidentally answer our question), since UPP protocols can be more powerful than PP protocols for reasons unrelated to their rank-1 property.[4]

A better formalization is as follows. We define new communication classes, $\mathsf{UPostBPP}_{\square} \subseteq \mathsf{UPostBPP}$ and $\mathsf{UPP}_{\square} \subseteq \mathsf{UPP}$, in the same way as $\mathsf{PostBPP}$ and PP, except allowing the public randomness to be arbitrarily distributed over $\{0,1\}^k$ (still charging for k and not for α or ϵ). Combinatorially, we have a distribution over 2^k labeled rectangles, but with no restrictions on their probabilities. Our informal question can now be formalized as follows:

[4] The Log Rank Conjecture (and its variants) also do not adequately formalize our question, since the definition of a protocol imposes constraints on how its rectangles interrelate, whereas there are no analogous constraints on the rank-1 matrices making up a low-rank decomposition. A fairer formalization along these lines would be to compare the power (in representing boolean matrices) of sums of rank-1 matrices vs. linear combinations of rectangles; nothing seems to be known about this question.

Formal question: *Do we have* $\mathsf{UPostBPP} = \mathsf{UPostBPP}_\square$*? How about* $\mathsf{UPP} = \mathsf{UPP}_\square$*?*

The seemingly minor syntactic generalization introduced in the definitions of the $\square$-classes makes a huge difference: We observe (in the full version) that $\mathsf{P}^{\mathsf{NP}} \subseteq \mathsf{UPostBPP}_\square$,[5] whereas it is known that $\mathsf{PostBPP}$ and P^{NP} are incomparable. Hence $\mathsf{UPostBPP}_\square$ is a strict superset of both $\mathsf{PostBPP}$ and P^{NP}. This leaves us with no known examples of functions to witness a separation for our "rank-1 vs. rectangle" question; currently the best gap is $\mathsf{UPostBPP}(F) \leq O(1)$ vs. $\mathsf{UPostBPP}_\square(F) \geq \Omega(\log n)$ where F is the usual Greater-Than function defined by $F(x, y) = 1$ iff $x > y$ when $x, y \in [2^n]$ are viewed as numbers. There is also no clear analogue of the "rank-1 vs. rectangle" distinction in query complexity, so a separation of the two notions in communication complexity might require interesting techniques. In fact, in the context of SBP (subclass of $\mathsf{PostBPP}$), it can be shown that rank-1 matrices do not add any power over mere rectangles [17].

2.6 $\mathsf{PP} \not\subseteq \mathsf{UPostBPP}_\square$

Our final result is to develop and apply a useful lower bound method for the class $\mathsf{UPostBPP}_\square$ introduced above. $\mathsf{PostBPP}$ already has a tight rectangle-based lower bound technique, which was dubbed "extended discrepancy" in [15] but was used earlier in [28] to show that $\mathsf{PP} \not\subseteq \mathsf{PostBPP}$. We strengthen the latter result to show that $\mathsf{PP} \not\subseteq \mathsf{UPostBPP}_\square$. (Showing $\mathsf{PP} \not\subseteq \mathsf{UPostBPP}$ remains open.) In our proof, we make use of the main theorem from [17], which applies to composed functions where the gadget is as follows.

▶ **Definition 5.** The *confounding* gadget g is defined by $g(x_i, y_i) \coloneqq \langle x_i, y_i \rangle \bmod 2$, where $x_i, y_i \in \{0,1\}^b$ and the block length b is $b(m) \coloneqq 100 \log m$.

We introduce the confounded-majority function, defined as Conf-Maj $\coloneqq f \circ g^m$ where f is the majority function and g is the confounding gadget. Note that Conf-Maj has input length $n \coloneqq m \cdot b = m \cdot 100 \log m$ and is in PP since Alice and Bob can pick $i \in [m]$ uniformly at random and then exchange $b + 1 \leq O(\log n)$ bits to evaluate $g(x_i, y_i)$.

▶ **Theorem 6.** $\mathsf{UPostBPP}_\square(\text{Conf-Maj}) = \Theta(n)$*, and hence* $\mathsf{PP} \not\subseteq \mathsf{UPostBPP}_\square$*.*

To prove Theorem 6 (in the full version) we introduce a lower bound technique for $\mathsf{UPostBPP}_\square$ that strengthens the extended discrepancy bound (for $\mathsf{PostBPP}$) by requiring it to hold under a product distribution over inputs (analogously to how [40] showed that the "monochromatic rectangle size bound under product distributions" gives a lower bound for P^{NP}). However, only a $\Omega(\sqrt{n \log n})$ lower bound for Conf-Maj follows using this technique, so to get the $\Omega(n)$ lower bound in Theorem 6, we generalize the technique further by allowing a rectangle's *size* to be measured with respect to some product distribution while its *error* is measured with respect to some other (arbitrary) distribution. (This is very analogous to the idea of relative discrepancy [14, 12].) To apply our general lower bound technique to Conf-Maj, we employ the communication-to-query machinery from [17] in a new, somewhat indirect way.

Finally, we mention another intriguing property of $\mathsf{UPostBPP}_\square$: By our lower bound technique and the results of [15] it follows immediately that to prove the Log Rank Conjecture,

[5] This inclusion also holds for time-bounded complexity. In defining the time-bounded version of $\mathsf{UPostBPP}_\square$, we would allow the distribution of the random string to depend nonuniformly on the input length n, though for the inclusion of P^{NP}, the distribution is computable in exponential time given the string 1^n.

i.e., that $\mathsf{P}(F) \leq \mathrm{poly}(\log \mathrm{rank}(F))$ for all total boolean matrices F, it suffices to prove the same with $\mathsf{UPostBPP}_{\square}$ instead of P. See the full version for more details.

3 Lower Bounds for Block-Equality and Unique-Set-Intersection

We now describe a technique for lower bounding both $\mathsf{ZPP}^{\mathsf{NP}[1]}$ and coDP communication.

▶ **Lemma 7.** *Suppose μ_0 is a distribution over $F^{-1}(0)$, μ_1 is a distribution over $F^{-1}(1)$, and C is a constant such that for every rectangle $R \subseteq \{0,1\}^n \times \{0,1\}^n$, $\mu_0(R) \leq C \cdot \mu_1(R) + \delta$, and if R is 1-monochromatic (i.e., contains no 0-inputs) then $\mu_1(R) \leq \delta$. Then*

(i) $\mathsf{ZPP}^{\mathsf{NP}[1]}(F) \geq \Omega(\log(1/\delta))$,

(ii) $\mathsf{coDP}(F) \geq \Omega(\log(1/\delta))$.

The first half of the technique ($\mu_0(R) \leq C \cdot \mu_1(R) + \delta$) is the corruption bound (which is a tight lower bound technique for so-called coSBP [18]), and the other half is the 1-monochromatic rectangle size bound (which is a tight lower bound technique for NP [33, §2.4]). The combined technique gives a lower bound for both $\mathsf{ZPP}^{\mathsf{NP}[1]}$ and coDP, even though neither of these classes appears to be a "combination" of coSBP and NP.

We prove parts (i) and (ii) of Lemma 7 in Section 3.1 and Section 3.2. Then we apply the technique to Block-Eq in Section 3.3 (thus proving Theorem 1), and finally we apply the technique to Unique-Inter in Section 3.4 (thus proving Theorem 2 and Theorem 3).

3.1 Proof of Lemma 7(i)

Suppose for contradiction there is a cost-$o(\log(1/\delta))$ $\mathsf{ZPP}^{\mathsf{NP}[1]}$ protocol Π computing F. Then in particular we have $\delta \leq o(1)$. By the amplification result of [11], we may assume $\mathbb{P}\big[\Pi(x,y) = \perp\big] \leq 1/10C$ for all $(x,y) \in F^{-1}$. By Markov's inequality and a union bound, we may fix a $\mathsf{P}^{\mathsf{NP}[1]}$-type protocol Π^* in the support of Π such that $\mathbb{P}_{(x,y)\sim\mu_0}\big[\Pi^*(x,y) = \perp\big] \leq 1/5C$ and $\mathbb{P}_{(x,y)\sim\mu_1}\big[\Pi^*(x,y) = \perp\big] \leq 1/5C$. Let the notation $k, R_v, S_{v,w}, o_v$ be with respect to Π^* (see the definition of $\mathsf{ZPP}^{\mathsf{NP}[1]}$ in Section 2.1), and note that without loss of generality, each o_v is non-constant (otherwise we could redefine $S_{v,w} = \emptyset$ for all w and redefine $o_v(1)$ arbitrarily).

For $b \in \{0,1,\perp\}$, define $W_b := \bigcup_{v,w\,:\,o_v(1)=b} S_{v,w}$ as the set of "witnessed" inputs (the NP oracle responds "1") on which Π^* outputs b, and define $N_b := \bigcup_{v\,:\,o_v(0)=b} \big(R_v \setminus \bigcup_w S_{v,w}\big)$ as the set of "non-witnessed" inputs (the NP oracle responds "0") on which Π^* outputs b. Note that $\{W_0, N_0, W_1, N_1, W_\perp, N_\perp\}$ partitions $\{0,1\}^n \times \{0,1\}^n$. By assumption, $\mu_0(W_\perp \cup N_\perp) \leq 1/5C$ and $\mu_1(W_\perp \cup N_\perp) \leq 1/5C$. By the correctness of Π, for $b \in \{0,1\}$ we have $(W_b \cup N_b) \cap F^{-1}(1-b) = \emptyset$.

▶ **Claim 8.** $\mu_0(W_0) \leq 1/4$.

▶ **Claim 9.** $\mu_0(N_0) \leq 1/4$.

This provides the contradiction since then $\mu_0\big(\{0,1\}^n \times \{0,1\}^n\big) = \mu_0(W_0) + \mu_0(N_0) + \mu_0(W_1 \cup N_1) + \mu_0(W_\perp \cup N_\perp) \leq 1/4 + 1/4 + 0 + 1/5C < 1$.

Proof of Claim 8. For each v, w such that $o_v(1) = 0$, we have $\mu_1(S_{v,w}) = 0$ and hence $\mu_0(S_{v,w}) \leq \delta$. Thus by a union bound, $\mu_0(W_0) \leq \sum_{v,w\,:\,o_v(1)=0} \mu_0(S_{v,w}) \leq 2^{o(\log(1/\delta))} \cdot \delta \leq \delta^{1-o(1)} \leq 1/4$. ◀

Proof of Claim 9. If v is such that $o_v(0) = 0$, then we have

$$\mu_0\big(R_v \setminus \bigcup_w S_{v,w}\big) \;\le\; \mu_0(R_v) \;\le\; C \cdot \mu_1(R_v) + \delta \;=\; C \cdot \mu_1\big(\bigcup_w S_{v,w}\big) + \delta$$

by the fact that $\big(R_v \setminus \bigcup_w S_{v,w}\big) \cap F^{-1}(1) = \emptyset$. Also, since each o_v is non-constant, we have

$$\begin{aligned}
\textstyle\sum_{v\,:\,o_v(0)=0} \mu_1\big(\bigcup_w S_{v,w}\big) \;&=\; \textstyle\sum_{v\,:\,o_v(0)=0,\,o_v(1)=\perp} \mu_1\big(\bigcup_w S_{v,w}\big) \\
&\quad + \textstyle\sum_{v\,:\,o_v(0)=0,\,o_v(1)=1} \mu_1\big(\bigcup_w S_{v,w}\big) \\
&\le\; \mu_1(W_\perp) + \textstyle\sum_{v,w\,:\,o_v(1)=1} \mu_1(S_{v,w}) \\
&\le\; \mu_1(W_\perp \cup N_\perp) + 2^{o(\log(1/\delta))} \cdot \delta \\
&\le\; 1/5C + \delta^{1-o(1)}
\end{aligned}$$

where the third line follows since $S_{v,w}$ is 1-monochromatic if $o_v(1) = 1$. Combining these, we have

$$\begin{aligned}
\mu_0(N_0) \;&=\; \textstyle\sum_{v\,:\,o_v(0)=0} \mu_0\big(R_v \setminus \bigcup_w S_{v,w}\big) \\
&\le\; \textstyle\sum_{v\,:\,o_v(0)=0} \Big(C \cdot \mu_1\big(\bigcup_w S_{v,w}\big) + \delta\Big) \\
&\le\; C \cdot \Big(\textstyle\sum_{v\,:\,o_v(0)=0} \mu_1\big(\bigcup_w S_{v,w}\big)\Big) + 2^{o(\log(1/\delta))} \cdot \delta \\
&\le\; C \cdot \big(1/5C + \delta^{1-o(1)}\big) + \delta^{1-o(1)} \\
&\le\; 1/4.
\end{aligned}$$

◀

3.2 Proof of Lemma 7(ii)

Suppose for contradiction there is a cost-k coDP protocol Π computing F where $k \le o(\log(1/\delta))$. Then in particular we have $\delta \le o(1)$. We have a pair of collections of rectangles, $\{S_w \;:\; w \in \{0,1\}^k\}$ and $\{T_w \;:\; w \in \{0,1\}^k\}$, such that if $F(x,y) = 0$ then $(x,y) \in \bigcup_w S_w$ and $(x,y) \notin \bigcup_w T_w$, and if $F(x,y) = 1$ then $(x,y) \notin \bigcup_w S_w$ or $(x,y) \in \bigcup_w T_w$. Since $\mu_0\big(\bigcup_w S_w\big) = 1$, there exists a w^* such that $\mu_0(S_{w^*}) \ge 2^{-k} \ge \delta^{1/3}$ and hence $\mu_1(S_{w^*}) \ge \frac{1}{C} \cdot (\delta^{1/3} - \delta) \ge \delta^{1/2}$. Since $S_{w^*} \cap F^{-1}(1) \subseteq \bigcup_w T_w$, there exists a w' such that $\mu_1(T_{w'}) \ge \mu_1\big(S_{w^*} \cap F^{-1}(1)\big) \cdot 2^{-k} > \delta^{1/2} \cdot \delta^{1/2} = \delta$. But $T_{w'}$ is 1-monochromatic since $F^{-1}(0) \cap \bigcup_w T_w = \emptyset$, so this is a contradiction.

3.3 Proof of Theorem 1

Let μ_0 be the uniform distribution over $\textsc{Block-Eq}^{-1}(0)$, and let μ_1 be the uniform distribution over the subset of $\textsc{Block-Eq}^{-1}(1)$ consisting of all (x,y) for which $x_i = y_i$ for a unique i.

▶ **Lemma 10.** *$\mu_0(R) \le 45 \cdot \mu_1(R) + 2^{-\Omega(\sqrt{n})}$ holds for every rectangle $R \subseteq \{0,1\}^n \times \{0,1\}^n$.*

▶ **Lemma 11.** *$\mu_1(R) \le 2^{-\Omega(\sqrt{n})}$ holds for every 1-monochromatic rectangle R of* $\textsc{Block-Eq}$.

Together, Lemma 10 and Lemma 11 show that the hypothesis of Lemma 7 holds with $F \coloneqq \textsc{Block-Eq}$, $C \coloneqq 45$, and $\delta \coloneqq 2^{-\Omega(\sqrt{n})}$. The lower bound in Theorem 1 follows. For the upper bound, in fact $\mathsf{ZPP}(\textsc{Block-Eq}) \le O(\sqrt{n})$ holds [33, §4.1.1] (though it is slightly quicker to see that $\mathsf{NP}(\textsc{Block-Eq}) \le O(\sqrt{n})$ holds by guessing i and deterministically verifying that $x_i = y_i$).

For the proofs of the lemmas, we define $m \coloneqq \sqrt{n}$ and $b \coloneqq \sqrt{n}$ (as in the notation for the decomposition $\textsc{Block-Eq} \coloneqq \textsc{Or} \circ \textsc{Eq}^m$ where Eq takes b-bit inputs).

Proof of Lemma 10. For $x^0, x^1, y^0, y^1 \in \{0,1\}^b$, we say the tuple (x^0, x^1, y^0, y^1) is *valid* iff $x^0 \neq y^0$, $x^0 \neq y^1$, $x^1 \neq y^0$, and $x^1 = y^1$. We say $\Xi := \big((x_1^0, x_1^1, y_1^0, y_1^1), \ldots, (x_m^0, x_m^1, y_m^0, y_m^1)\big)$ is valid iff it is a tuple of valid tuples. If Ξ is valid then the injection $\Phi_\Xi \colon \{0,1\}^m \times \{0,1\}^m \to \{0,1\}^n \times \{0,1\}^n$ defined by $\Phi_\Xi(u,v) := \big(x_1^{u_1} \cdots x_m^{u_m}, y_1^{v_1} \cdots y_m^{v_m}\big)$ is a reduction from $\textsc{Inter} := \textsc{Or} \circ \textsc{And}^m$ (for single-bit And) to Block-Eq:

$$\textsc{Inter}(u,v) \;=\; \textsc{Block-Eq}\big(\Phi_\Xi(u,v)\big).$$

(In other words, the image of Φ_Ξ, as a submatrix of the Block-Eq matrix, is a copy of the Inter matrix.)

Define $\textsc{Unambig-Inter} := \textsc{Unambig-Or} \circ \textsc{And}^m$ where the partial function Unambig-Or is Or restricted to the domain of strings of Hamming weight 0 or 1; i.e., $\textsc{Unambig-Inter}^{-1}(0)$ consists of all pairs of disjoint sets, and $\textsc{Unambig-Inter}^{-1}(1)$ consists of all pairs of uniquely intersecting sets.

▶ **Lemma 12** ([45]). *There exists a distribution ν_0 over $\textsc{Unambig-Inter}^{-1}(0)$ and a distribution ν_1 over $\textsc{Unambig-Inter}^{-1}(1)$ such that $\nu_0(R) \leq 45 \cdot \nu_1(R) + 2^{-\Omega(m)}$ holds for every rectangle $R \subseteq \{0,1\}^m \times \{0,1\}^m$. Moreover, the uniquely intersecting coordinate in ν_1 is uniformly distributed.*

We claim that for $a \in \{0,1\}$ we have $\mu_a = \mathbb{E}_\Xi\, \Phi_\Xi(\nu_a)$ where a valid Ξ is chosen uniformly at random independently of ν_a. In other words, μ_a equals the distribution obtained by choosing Ξ, then independently taking a sample from ν_a, then applying Φ_Ξ to the sample (i.e., the uniform mixture of the distributions $\Phi_\Xi(\nu_a)$). We only argue that $\mu_1 = \mathbb{E}_\Xi\, \Phi_\Xi(\nu_1)$ (the argument for $\mu_0 = \mathbb{E}_\Xi\, \Phi_\Xi(\nu_0)$ is essentially the same). In fact, we make the stronger claim that for every $(u,v) \in \textsc{Unambig-Inter}^{-1}(1)$, say with $u_i = v_i = 1$, the distribution $\mathbb{E}_\Xi\, \Phi_\Xi(u,v)$ is uniform over the subset of $\textsc{Block-Eq}^{-1}(1)$ consisting of all (x,y) for which $x_i = y_i$ and $x_j \neq y_j$ for all $j \neq i$. The original claim follows from this since the uniquely intersecting coordinate i is uniformly distributed. The stronger claim follows immediately from the facts that the coordinates of Ξ are independent, that (x_i^1, y_i^1) is uniformly distributed over $\textsc{Eq}^{-1}(1)$, and that for $j \neq i$, (x_j^0, y_j^0), (x_j^0, y_j^1), and (x_j^1, y_j^0) are all marginally uniformly distributed over $\textsc{Eq}^{-1}(0)$. The claim is established.

Now for every rectangle $R \subseteq \{0,1\}^n \times \{0,1\}^n$, if we let $\Phi_\Xi^{-1}(R)$ denote the rectangle of all points in $\{0,1\}^m \times \{0,1\}^m$ that map into R under Φ_Ξ, then we have

$$\begin{aligned}
\mu_0(R) &= \mathbb{E}_\Xi \big(\Phi_\Xi(\nu_0)(R)\big) \\
&= \mathbb{E}_\Xi\, \nu_0\big(\Phi_\Xi^{-1}(R)\big) \\
&\leq \mathbb{E}_\Xi \Big(45 \cdot \nu_1\big(\Phi_\Xi^{-1}(R)\big) + 2^{-\Omega(m)}\Big) \\
&= 45 \cdot \mathbb{E}_\Xi\, \nu_1\big(\Phi_\Xi^{-1}(R)\big) + 2^{-\Omega(m)} \\
&= 45 \cdot \mathbb{E}_\Xi \big(\Phi_\Xi(\nu_1)(R)\big) + 2^{-\Omega(m)} \\
&= 45 \cdot \mu_1(R) + 2^{-\Omega(\sqrt{n})}.
\end{aligned}$$

◀

Proof of Lemma 11. Note that μ_1 is uniform over a set of size

$$m \cdot 2^b \cdot (2^{2b} - 2^b)^{m-1} \;=\; m \cdot 2^b \cdot 2^{2b(m-1)} \cdot (1 - 2^{-b})^{m-1} \;\geq\; \Omega(m \cdot 2^b \cdot 2^{2b(m-1)}).$$

If $R := A \times B$ is 1-monochromatic then $|A| \leq m \cdot 2^{b(m-1)}$ (since for any $y \in B$ there are at most $m \cdot (2^b)^{m-1}$ many x's for which $\textsc{Block-Eq}(x,y) = 1$), and similarly $|B| \leq m \cdot 2^{b(m-1)}$, and hence $|R| \leq m^2 \cdot 2^{2b(m-1)}$. It follows that

$$\mu_1(R) \;\leq\; \frac{m^2 \cdot 2^{2b(m-1)}}{\Omega(m \cdot 2^b \cdot 2^{2b(m-1)})} \;\leq\; O(m \cdot 2^{-b}) \;\leq\; 2^{-\Omega(\sqrt{n})}.$$

◀

3.4 Proof of Theorem 2 and Theorem 3

We again use the corruption lemma from [45], but now we need to take a closer look at the distribution over 1-inputs. Let $n = 4\ell - 1$. Let μ_0 be the distribution over $\textsc{Unique-Inter}^{-1}(0)$ that samples uniformly random disjoint sets of size ℓ, and let μ_1 be the distribution over $\textsc{Unique-Inter}^{-1}(1)$ that samples uniformly random uniquely intersecting sets of size ℓ.

▶ **Lemma 13** ([45]). *$\mu_0(R) \leq 45 \cdot \mu_1(R) + 2^{-\Omega(n)}$ holds for every rectangle $R \subseteq \{0,1\}^n \times \{0,1\}^n$.*

▶ **Lemma 14.** *$\mu_1(R) \leq 2^{-\Omega(n)}$ holds for every 1-monochromatic rectangle R of* Unique-Inter.

Together, Lemma 13 and Lemma 14 show that the hypothesis of Lemma 7 holds with $F \coloneqq \textsc{Unique-Inter}$, $C \coloneqq 45$, and $\delta \coloneqq 2^{-\Omega(n)}$. Theorem 2 and Theorem 3 follow.

Proof of Lemma 14. For each $i \in [n]$ let us define the rectangle $R_i \coloneqq \big\{(x,y) \in R \,:\, x_i = y_i = 1\big\}$, and note that the R_i's partition R. For each i we have $|R_i| \leq 2^{n-1}$ since every $(x,y) \in R_i$ is disjoint on the coordinates $[n] \setminus \{i\}$.[6] Hence $|R| \leq n2^{n-1} \leq 2^{(1+o(1))n}$.

Note that μ_1 can be sampled by the following process.

1. Pick a uniformly random $i \in [n]$.
2. Pick a uniformly random $H \subseteq [n] \setminus \{i\}$ of size $2\ell - 2$. There are $\binom{n-1}{2\ell-2} = \Theta(2^n/\sqrt{n})$ choices.
3. Pick a uniformly random partition $H = H_1 \cup H_2$ into sets of size $\ell - 1$. There are $\binom{2\ell-2}{\ell-1} = \Theta(2^{0.5n}/\sqrt{n})$ choices.
4. Let $x \coloneqq \{i\} \cup H_1$ and $y \coloneqq \{i\} \cup H_2$.

Hence μ_1 is uniform over its support of size $n \cdot \Theta(2^n/\sqrt{n}) \cdot \Theta(2^{0.5n}/\sqrt{n}) = \Theta(2^{1.5n}) \geq 2^{(1.5-o(1))n}$. It follows that $\mu_1(R) \leq 2^{(1+o(1))n}/2^{(1.5-o(1))n} \leq 2^{-\Omega(n)}$. ◀

Acknowledgements. We thank anonymous referees for helpful comments.

References

1 Scott Aaronson. Quantum computing, postselection, and probabilistic polynomial-time. *Proceedings of the Royal Society A*, 461(2063):3473–3482, 2005. doi:10.1098/rspa.2005.1546.

2 Scott Aaronson and Avi Wigderson. Algebrization: A new barrier in complexity theory. *ACM Transactions on Computation Theory*, 1(1), 2009. doi:10.1145/1490270.1490272.

3 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.

4 Andreas Blass and Yuri Gurevich. On the unique satisfiability problem. *Information and Control*, 55(1–3):80–88, 1982. doi:10.1016/S0019-9958(82)90439-9.

5 Harry Buhrman, Nikolai Vereshchagin, and Ronald de Wolf. On computation and communication with small bias. In *Proceedings of the 22nd Conference on Computational Complexity (CCC)*, pages 24–32. IEEE, 2007. doi:10.1109/CCC.2007.18.

[6] By [27], the bound $|R_i| \leq 2^{n-1}$ also holds assuming every input in R has intersection size either 1 or ≥ 3. Using this, it follows that Theorem 2 and Theorem 3 hold under the promise that at most two coordinates intersect.

6 Jin-Yi Cai and Venkatesan Chakaravarthy. On zero error algorithms having oracle access to one query. *Journal of Combinatorial Optimization*, 11(2):189–202, 2006. doi:10.1007/s10878-006-7130-0.

7 Amit Chakrabarti, Graham Cormode, Navin Goyal, and Justin Thaler. Annotations for sparse data streams. In *Proceedings of the 25th Symposium on Discrete Algorithms (SODA)*, pages 687–706. ACM-SIAM, 2014. doi:10.1137/1.9781611973402.52.

8 Amit Chakrabarti, Graham Cormode, Andrew McGregor, and Justin Thaler. Annotations in data streams. *ACM Transactions on Algorithms*, 11(1):7, 2014. doi:10.1145/2636924.

9 Amit Chakrabarti, Graham Cormode, Andrew McGregor, Justin Thaler, and Suresh Venkatasubramanian. Verifiable stream computation and Arthur–Merlin communication. In *Proceedings of the 30th Computational Complexity Conference (CCC)*. Schloss Dagstuhl, 2015. To appear.

10 Richard Chang, Jim Kadin, and Pankaj Rohatgi. On unique satisfiability and the threshold behavior of randomized reductions. *Journal of Computer and System Sciences*, 50(3):359–373, 1995. doi:10.1006/jcss.1995.1028.

11 Richard Chang and Suresh Purini. Amplifying $\mathrm{ZPP}^{\mathrm{SAT}[1]}$ and the two queries problem. In *Proceedings of the 23rd Conference on Computational Complexity (CCC)*, pages 41–52. IEEE, 2008. doi:10.1109/CCC.2008.32.

12 Lila Fontes, Rahul Jain, Iordanis Kerenidis, Sophie Laplante, Mathieu Lauriere, and Jérémie Roland. Relative discrepancy does not separate information and communication complexity. Technical Report TR15-028, Electronic Colloquium on Computational Complexity (ECCC), 2015.

13 Jürgen Forster. A linear lower bound on the unbounded error probabilistic communication complexity. *Journal of Computer and System Sciences*, 65(4):612–625, 2002. doi:10.1016/S0022-0000(02)00019-3.

14 Anat Ganor, Gillat Kol, and Ran Raz. Exponential separation of information and communication for boolean functions. In *Proceedings of the 47th Symposium on Theory of Computing (STOC)*. ACM, 2015. To appear.

15 Dmitry Gavinsky and Shachar Lovett. En route to the log-rank conjecture: New reductions and equivalent formulations. In *Proceedings of the 41st International Colloquium on Automata, Languages, and Programming (ICALP)*, pages 514–524. Springer, 2014. doi:10.1007/978-3-662-43948-7_43.

16 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.

17 Mika Göös, Shachar Lovett, Raghu Meka, Thomas Watson, and David Zuckerman. Rectangles are nonnegative juntas. In *Proceedings of the 47th Symposium on Theory of Computing (STOC)*. ACM, 2015. To appear.

18 Mika Göös and Thomas Watson. Communication complexity of set-disjointness for all probabilities. In *Proceedings of the 18th International Workshop on Randomization and Computation (RANDOM)*, pages 721–736. Schloss Dagstuhl, 2014. doi:10.4230/LIPIcs.APPROX-RANDOM.2014.721.

19 Tom Gur and Ran Raz. Arthur–Merlin streaming complexity. In *Proceedings of the 40th International Colloquium on Automata, Languages, and Programming (ICALP)*, pages 528–539. Springer, 2013. doi:10.1007/978-3-642-39206-1_45.

20 Tom Gur and Ron Rothblum. Non-interactive proofs of proximity. In *Proceedings of the 6th Innovations in Theoretical Computer Science Conference (ITCS)*, pages 133–142. ACM, 2015. doi:10.1145/2688073.2688079.

21 Bernd Halstenberg and Rüdiger Reischuk. Relations between communication complexity classes. *Journal of Computer and System Sciences*, 41(3):402–429, 1990. doi:10.1016/0022-0000(90)90027-I.

22 Yenjo Han, Lane Hemaspaandra, and Thomas Thierauf. Threshold computation and cryptographic security. *SIAM Journal on Computing*, 26(1):59–78, 1997. doi:10.1137/S0097539792240467.

23 Russell Impagliazzo and Ryan Williams. Communication complexity with synchronized clocks. In *Proceedings of the 25th Conference on Computational Complexity (CCC)*, pages 259–269. IEEE, 2010. doi:10.1109/CCC.2010.32.

24 T.S. Jayram, Ravi Kumar, and D. Sivakumar. Two applications of information complexity. In *Proceedings of the 35th Symposium on Theory of Computing (STOC)*, pages 673–682. ACM, 2003. doi:10.1145/780542.780640.

25 Stasys Jukna. On graph complexity. *Combinatorics, Probability, & Computing*, 15(6):855–876, 2006. doi:10.1017/S0963548306007620.

26 Stasys Jukna. *Boolean Function Complexity: Advances and Frontiers*, volume 27 of *Algorithms and Combinatorics*. Springer, 2012.

27 Volker Kaibel and Stefan Weltge. A short proof that the extension complexity of the correlation polytope grows exponentially. *Discrete & Computational Geometry*, 53(2):397–401, 2015. doi:10.1007/s00454-014-9655-9.

28 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.

29 Hartmut Klauck. Lower bounds for quantum communication complexity. *SIAM Journal on Computing*, 37(1):20–46, 2007. doi:10.1137/S0097539702405620.

30 Hartmut Klauck. On Arthur Merlin games in communication complexity. In *Proceedings of the 26th Conference on Computational Complexity (CCC)*, pages 189–199. IEEE, 2011. doi:10.1109/CCC.2011.33.

31 Hartmut Klauck and Ved Prakash. Streaming computations with a loquacious prover. In *Proceedings of the 4th Innovations in Theoretical Computer Science Conference (ITCS)*, pages 305–320. ACM, 2013. doi:10.1145/2422436.2422471.

32 Hartmut Klauck and Ved Prakash. An improved interactive streaming algorithm for the distinct elements problem. In *Proceedings of the 41st International Colloquium on Automata, Languages, and Programming (ICALP)*, pages 919–930. Springer, 2014. doi:10.1007/978-3-662-43948-7_76.

33 Eyal Kushilevitz and Noam Nisan. *Communication Complexity*. Cambridge University Press, 1997.

34 Tak Wah Lam and Walter Ruzzo. Results on communication complexity classes. *Journal of Computer and System Sciences*, 44(2):324–342, 1992. doi:10.1016/0022-0000(92)90025-E.

35 Nathan Linial and Adi Shraibman. Learning complexity vs communication complexity. *Combinatorics, Probability, & Computing*, 18(1–2):227–245, 2009. doi:10.1017/S0963548308009656.

36 Satyanarayana Lokam. Spectral methods for matrix rigidity with applications to size-depth trade-offs and communication complexity. *Journal of Computer and System Sciences*, 63(3):449–473, 2001. doi:10.1006/jcss.2001.1786.

37 Satyanarayana Lokam. Complexity lower bounds using linear algebra. *Foundations and Trends in Theoretical Computer Science*, 4(1–2):1–155, 2009. doi:10.1561/0400000011.

38 Ilan Newman. Private vs. common random bits in communication complexity. *Information Processing Letters*, 39(2):67–71, 1991. doi:10.1016/0020-0190(91)90157-D.

39 Christos Papadimitriou and Mihalis Yannakakis. The complexity of facets (and some facets of complexity). *Journal of Computer and System Sciences*, 28(2):244–259, 1984. doi:10.1016/0022-0000(84)90068-0.

40 Periklis Papakonstantinou, Dominik Scheder, and Hao Song. Overlays and limited memory communication. In *Proceedings of the 29th Conference on Computational Complexity (CCC)*, pages 298–308. IEEE, 2014. doi:10.1109/CCC.2014.37.

41 Ramamohan Paturi and Janos Simon. Probabilistic communication complexity. *Journal of Computer and System Sciences*, 33(1):106–123, 1986. doi:10.1016/0022-0000(86)90046-2.

42 Pavel Pudlák, Vojtech Rödl, and Petr Savický. Graph complexity. *Acta Informatica*, 25(5):515–535, 1988. doi:10.1007/BF00279952.

43 Ran Raz and Amir Shpilka. On the power of quantum proofs. In *Proceedings of the 19th Conference on Computational Complexity (CCC)*, pages 260–274. IEEE, 2004. doi:10.1109/CCC.2004.1313849.

44 Alexander Razborov. On rigid matrices. Technical report, Steklov Mathematical Institute, 1989. In Russian.

45 Alexander Razborov. On the distributional complexity of disjointness. *Theoretical Computer Science*, 106(2):385–390, 1992. doi:10.1016/0304-3975(92)90260-M.

46 Alexander Razborov and Alexander Sherstov. The sign-rank of AC^0. *SIAM Journal on Computing*, 39(5):1833–1855, 2010. doi:10.1137/080744037.

47 Alexander Sherstov. Halfspace matrices. *Computational Complexity*, 17(2):149–178, 2008. doi:10.1007/s00037-008-0242-4.

48 Alexander Sherstov. The unbounded-error communication complexity of symmetric functions. *Combinatorica*, 31(5):583–614, 2011. doi:10.1007/s00493-011-2580-0.

49 Rahul Tripathi. The 1-versus-2 queries problem revisited. *Theory of Computing Systems*, 46(2):193–221, 2010. doi:10.1007/s00224-008-9126-x.

50 Leslie Valiant. Graph-theoretic arguments in low-level complexity. In *Proceedings of the 6th Symposium on Mathematical Foundations of Computer Science (MFCS)*, pages 162–176. Springer, 1977. doi:10.1007/3-540-08353-7_135.

51 Leslie Valiant and Vijay Vazirani. NP is as easy as detecting unique solutions. *Theoretical Computer Science*, 47(3):85–93, 1986. doi:10.1016/0304-3975(86)90135-0.

52 Nikolai Vereshchagin. Relativizability in complexity theory. In *Provability, Complexity, Grammars*, volume 192 of *AMS Translations, Series 2*, pages 87–172. American Mathematical Society, 1999.

53 Henning Wunderlich. On a theorem of Razborov. *Computational Complexity*, 21(3):431–477, 2012. doi:10.1007/s00037-011-0021-5.