



# The Landscape of Communication Complexity Classes

Mika Göös      Toniann Pitassi      Thomas Watson

Department of Computer Science, University of Toronto

April 1, 2015

## 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  $MA \not\subseteq ZPP^{NP[1]}$ , that is, Merlin–Arthur proof systems cannot be simulated by zero-sided error randomized protocols with one  $NP$  query. Here the class  $ZPP^{NP[1]}$  has the property that generalizing it in the slightest ways would make it contain  $AM \cap coAM$ , for which it is notoriously open to prove any explicit lower bounds. We also prove that  $US \not\subseteq ZPP^{NP[1]}$ , 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  $US \not\subseteq coDP$ , 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.

## 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 [BFS86], 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  $F: \{0, 1\}^n \times \{0, 1\}^n \rightarrow \{0, 1\}$  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, 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.

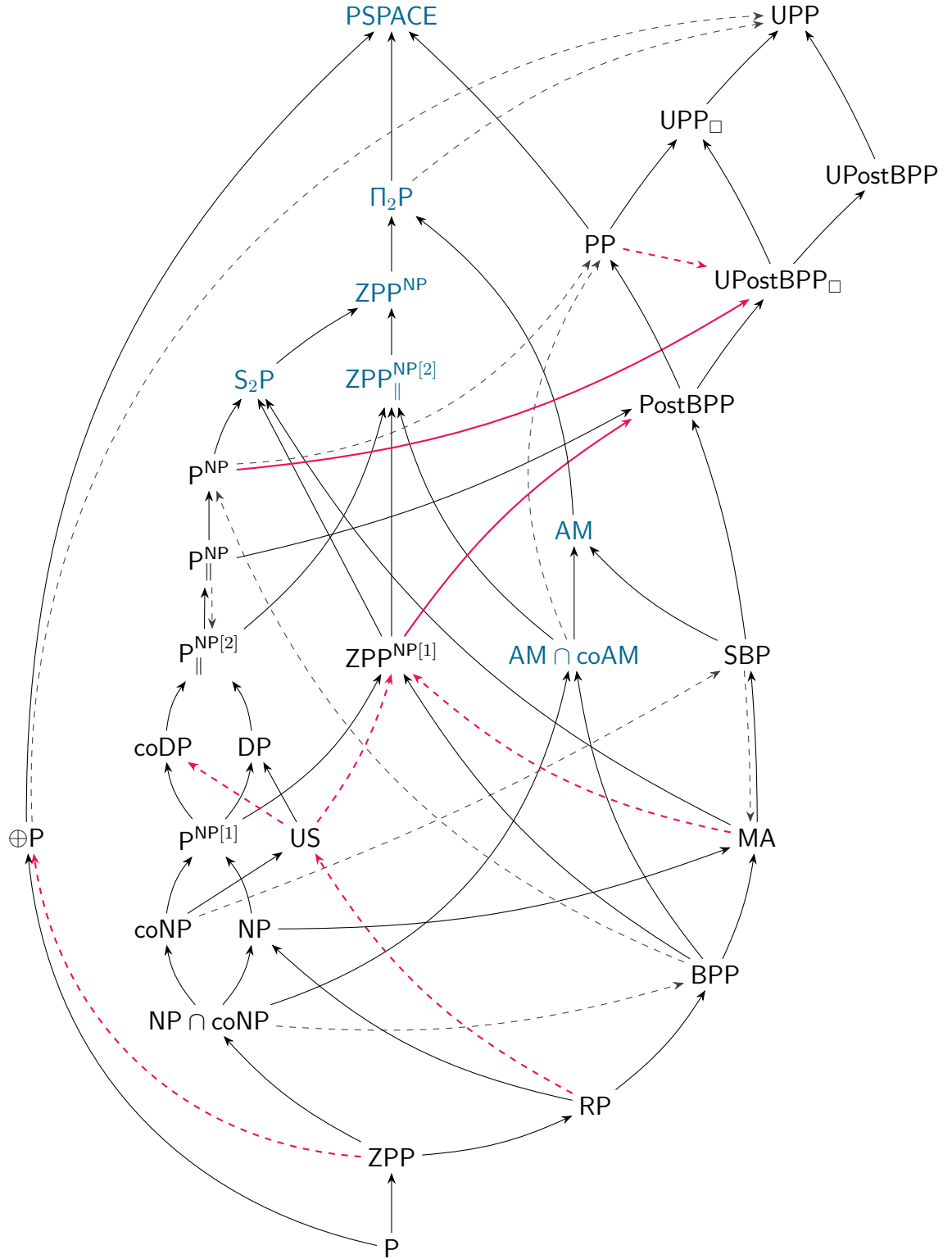
- A holy grail of classical complexity is to prove separations of classes between  $P$  and  $PSPACE$ . Separations relative to oracles can often be viewed as class separations in the restricted

setting of *query complexity*; see [Ver99] 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* [AW09], 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 [Raz89, Lok01, Lok09, Wun12] (as well as margin complexity rigidity lower bounds [LS09]), which in turn are related to circuit complexity [Val77]. Lower bounds against PH are also related to graph complexity [PRS88, Juk06]. 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 [CCMT14, KP13, GR13, CCGT14, CCM<sup>+</sup>15, KP14]).
- 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 [Kla07], and corruption is equivalent to SBP communication [GW14]). See [GLM<sup>+</sup>15] 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. At a glance, the state of affairs (including our new results) is summarized in [Figure 1](#), which shows a map of known inclusions and non-inclusions between pairs of communication classes. In [Section 2](#) we state our results more precisely and provide some intuition for the proofs. In [Section 3](#), we provide a comprehensive survey of all the nontrivial (non-)inclusions among the traditional classes depicted in [Figure 1](#). This updates previous surveys by Babai, Frankl, and Simon [BFS86] and Halstenberg and Reischuk [HR90].

We refer to [KN97, Juk12] for background on communication complexity. In [Appendix B](#) 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  $\mathcal{C}(F) \leq \text{poly}(\log n)$ .



**Figure 1:**  $\mathcal{C}_1 \rightarrow \mathcal{C}_2$  denotes  $\mathcal{C}_1 \subseteq \mathcal{C}_2$ , and  $\mathcal{C}_1 \dashrightarrow \mathcal{C}_2$  denotes  $\mathcal{C}_1 \not\subseteq \mathcal{C}_2$ . **Red** indicates new results. **Blue** indicates classes for which no explicit lower bounds are known.

## 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  $f: \{0, 1\}^m \rightarrow \{0, 1\}$  is a (possibly partial) *outer function* and  $g: \{0, 1\}^b \times \{0, 1\}^b \rightarrow \{0, 1\}$  is an *inner function* also called a *gadget*. We write  $F := f \circ g^m: \{0, 1\}^n \times \{0, 1\}^n \rightarrow \{0, 1\}$  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  $F(x, y) := f(g(x_1, y_1), \dots, g(x_m, y_m))$ .

### 2.1 MA $\not\subseteq$ ZPP<sup>NP[1]</sup>

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 [Kla03, RS04, AW09, GS10, Kla11, GR13, GR15], starting with the work of Klauck [Kla03], 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 [LS09, PSS14, KP14, CCM<sup>+</sup>15].

Our first result concerns the relationship between MA and another class, ZPP<sup>NP[1]</sup>, 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<sup>NP[1]</sup>. 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 [CP08] that ZPP<sup>NP[1]</sup> 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<sup>NP[1]</sup> to have success probability  $< 1/2$ , the class would change drastically: it would contain  $\text{AM} \cap \text{coAM}$  (see Section 3), 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  $\text{AM} \cap \text{coAM}$ . Thus, in a sense, ZPP<sup>NP[1]</sup> represents a boundary beyond which AM lower bounds would be the next step. The class ZPP<sup>NP[1]</sup> is also sandwiched between BPP and S<sub>2</sub>P [CC06]; S<sub>2</sub>P is a subclass of the polynomial hierarchy that has not been studied before in communication complexity (the definition appears in Appendix B), and no nontrivial lower bounds against it are known for any explicit function. This is another sense in which ZPP<sup>NP[1]</sup> constitutes a new frontier toward the elusive goal of proving explicit PH communication lower bounds. We also mention that ZPP<sup>NP[1]</sup> shows up frequently in the literature on the “two queries problem” (e.g., if  $\text{P}_{\parallel}^{\text{NP}[2]} \subseteq \text{ZPP}^{\text{NP}[1]}$  then  $\text{PH} = \text{S}_2\text{P}$  [Tri10]).

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

<sup>1</sup>Here, the term non-interactive means that Alice and Bob cannot interact with Merlin other than receiving the

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  $F: \{0, 1\}^n \times \{0, 1\}^n \rightarrow \{0, 1\}$  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  $\text{ZPP}^{\text{NP}[1]}$  protocol  $\Pi$  computing  $F$  is a distribution over  $\text{P}^{\text{NP}[1]}$ -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”  $\{S_{v,w} \subseteq R_v : w \in \{0, 1\}^k\}$  and an associated “output function”  $o_v: \{0, 1\} \rightarrow \{0, 1, \perp\}$ . The output of the  $\text{P}^{\text{NP}[1]}$ -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  $(x, y) \in F^{-1}$ ,  $\mathbb{P}[\Pi(x, y) \in \{F(x, y), \perp\}] = 1$  and  $\mathbb{P}[\Pi(x, y) = F(x, y)] \geq 3/4$ . The cost is the witness length  $k$  plus the maximum communication cost of the deterministic part of any of the constituent  $\text{P}^{\text{NP}[1]}$ -type protocols. The result of [CP08] shows that changing the success probability from  $3/4$  to any other constant strictly between  $1/2$  and  $1$  would only change the measure  $\text{ZPP}^{\text{NP}[1]}(F)$  by a constant factor.

We prove a lower bound for the block-equality function BLOCK-EQ, defined as follows:<sup>2</sup> 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  $\text{BLOCK-EQ} := \text{OR} \circ \text{EQ}^m$  where the input to OR is  $m := \sqrt{n}$  bits, and each input to EQ is  $b := \sqrt{n}$  bits. In other words, writing  $x := x_1 \cdots x_{\sqrt{n}} \in (\{0, 1\}^{\sqrt{n}})^{\sqrt{n}}$  and  $y := y_1 \cdots y_{\sqrt{n}} \in (\{0, 1\}^{\sqrt{n}})^{\sqrt{n}}$ , we have  $\text{BLOCK-EQ}(x, y) = 1$  iff  $x_i = y_i$  for some  $i$ . Note that  $\text{BLOCK-EQ} \in \text{MA}$  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 [LR92] that  $\text{BLOCK-EQ} \in \Sigma_2\text{P} \cap \Pi_2\text{P}$ , which is a superset of MA.)

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

To prove Theorem 1 (Section 4), 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  $\text{ZPP}^{\text{NP}[1]}$  (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 [Raz92] 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.

---

proof string.

<sup>2</sup>The complement of block-equality is often known as list-non-equality.

## 2.2 US $\not\subseteq$ ZPP<sup>NP[1]</sup>

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.<sup>3</sup> Phrased as a composed function,  $\text{INTER} := \text{OR} \circ \text{AND}^n$  (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., [BG82, VV86, CKR95]): 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:  $\{0, 1\}^n \times \{0, 1\}^n \rightarrow \{0, 1\}$  that maps  $(x, y)$  to 1 iff  $|x \cap y| = 1$ , i.e.,  $\text{UNIQUE-INTER} := \text{UNIQUE-OR} \circ \text{AND}^n$  where  $\text{UNIQUE-OR}(z) = 1$  iff the Hamming weight of  $z$  is 1. Unambiguous-set-intersection is the partial function UNAMBIG-INTER:  $\{0, 1\}^n \times \{0, 1\}^n \rightarrow \{0, 1\}$  that maps  $(x, y)$  to  $|x \cap y|$  if the latter is in  $\{0, 1\}$ , i.e.,  $\text{UNAMBIG-INTER} := \text{UNAMBIG-OR} \circ \text{AND}^n$  where  $\text{UNAMBIG-OR}(z)$  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  $\{R_w \subseteq \{0, 1\}^n \times \{0, 1\}^n : w \in \{0, 1\}^k\}$ , 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.**  $\text{ZPP}^{\text{NP}[1]}(\text{UNIQUE-INTER}) = \Theta(n)$ , and hence  $\text{US} \not\subseteq \text{ZPP}^{\text{NP}[1]}$ .

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  $\text{ZPP}^{\text{NP}[1]}$  instead of  $\text{P}^{\text{NP}[1]}$  incurs almost no extra complication.

The first proof ([Section 4](#)) employs the same lower bound technique as in [Theorem 1](#), but where we use Razborov’s corruption lemma [[Raz92](#)] directly (and we must do a little analysis to verify the 1-monochromatic rectangle size bound). The optional second proof (relegated to [Appendix A](#)) uses information complexity tools (including an adaptation of the “partial information cost” approach from [[JKS03](#)]) 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 [[CP08](#)] (whereas the first proof assumes success probability 0.999).

## 2.3 US $\not\subseteq$ coDP

The class DP was introduced in [[PY84](#)] 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

<sup>3</sup>We let “set-disjointness” refer to the complementary function where 1-inputs are disjoint.

two NP sets. The classes P, NP, and DP are the 0<sup>th</sup>, 1<sup>st</sup>, and 2<sup>nd</sup> (respectively) levels of the so-called boolean hierarchy.

We have  $US \subseteq DP$  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  $US \subseteq \text{coDP}$ : [CKR95] 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  $US \not\subseteq \text{coDP}$  in communication complexity.

Formally, a cost- $k$   $\text{coDP}$  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.**  $\text{coDP}(\text{UNIQUE-INTER}) = \Theta(n)$ , and hence  $US \not\subseteq \text{coDP}$ .

To prove [Theorem 3 \(Section 4\)](#), we show that the same lower bound technique we introduced for  $\text{ZPP}^{\text{NP}[1]}$  (the combination of the corruption bound and the 1-monochromatic rectangle size bound) *also* lower bounds  $\text{coDP}$  complexity. Thus we can simply reuse the application of the technique to  $\text{UNIQUE-INTER}$  from [Theorem 2](#). (Reusing the application to  $\text{BLOCK-EQ}$  from [Theorem 1](#) would show that  $\text{BLOCK-EQ} \notin \text{coDP}$ , but in fact  $\text{BLOCK-EQ} \notin \text{P}^{\text{NP}} \supseteq \text{coDP}$  was already known [IW10].)

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

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  $\text{BPP}_{\text{path}}$  [HHT97], but the name  $\text{PostBPP}$  (inspired by [Aar05]) has gained popularity in the recent literature ([GLM<sup>+</sup>15] 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  $\text{PostBPP}$  communication protocol for  $F$  would be as the communication cost plus  $\log(1/\alpha)$ , where  $\alpha$  is the minimum over all  $(x, y) \in F^{-1}$  of the probability of not outputting  $\perp$ . (Allowing public randomness and not charging for  $\alpha$  would enable  $\text{PostBPP}$  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 [BFS86] and define  $\text{PostBPP}$  and PP in a slightly different but equivalent way: we do not charge for  $\alpha$  or  $\epsilon$  but we require the public randomness to be uniformly distributed over  $\{0, 1\}^k$  and we charge for  $k$ . For both  $\text{PostBPP}$  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 [New91].

Formally, we define a  $\text{PostBPP}$  communication protocol  $\Pi$  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  $\mathbb{P}[\Pi(x, y) = F(x, y)] > 2 \cdot \mathbb{P}[\Pi(x, y) = 1 - F(x, y)]$ . 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  $\text{ZPP}^{\text{NP}[1]}$  follow from prior work. The following result shows that in fact they do, since many explicit lower bounds for  $\text{PostBPP}$  were known (see [Section 3](#)).

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

The proof of [Theorem 4 \(Section 5\)](#) also works for time-bounded complexity, and does not exploit any special properties of communication. Intuitively, the worst case for simulating a  $\text{ZPP}^{\text{NP}[1]}$  protocol is the following situation: Whenever the  $\text{NP}$  oracle responds “0” the protocol outputs the right answer, and whenever the  $\text{NP}$  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  $\text{BPP}$  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  $\text{NP}$  query, outputting  $\perp$  if the witness is invalid.

## 2.5 Open issue: Rank-1 vs. rectangles

The classes  $\text{PostBPP}$  and  $\text{PP}$  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  $\text{UPostBPP}$  (which was defined, but not extensively studied, in [\[GLM<sup>+</sup>15\]](#)) and  $\text{UPP}$  (which is well-studied [\[PS86, For02, She11b, RS10\]](#)). In  $\text{UPostBPP}$  and  $\text{UPP}$  we can dispense with public randomness altogether as the public coins could be tossed privately by Alice and then sent to Bob.

Combinatorially,  $\text{PostBPP}$  and  $\text{PP}$  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  $2^{-c}$  ([Observations 51](#) and [56](#)). In the case of  $\text{UPostBPP}$  and  $\text{UPP}$  there is a similar characterization with rectangles replaced by nonnegative rank-1 matrices ([Observations 52](#) and [57](#)). 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.,  $\text{PP} \neq \text{UPP}$  [\[BVdW07, She08\]](#), the known examples of functions  $F \in \text{UPP} \setminus \text{PP}$  can actually be computed without exploiting the full power of private randomness (their rank-1 property): we can use a  $\text{UPP}$  protocol whose associated rank-1 matrices are still rectangles, but occurring with *unrestricted*, possibly tiny, probability. We conclude that “ $\text{PP}$  vs.  $\text{UPP}$ ” is not the right way to formalize our informal question (and the existing proofs for  $\text{PP} \neq \text{UPP}$  do not incidentally answer our question).

A better formalization is as follows. We define new communication classes,  $\text{UPostBPP}_{\square} \subseteq \text{UPostBPP}$  and  $\text{UPP}_{\square} \subseteq \text{UPP}$ , in the same way as  $\text{PostBPP}$  and  $\text{PP}$ , except allowing the public randomness to be arbitrarily distributed over  $\{0, 1\}^k$  (still charging for  $k$  and not for  $\alpha$  or  $\epsilon$ ). 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:

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

The seemingly minor syntactic generalization introduced in the definitions of the  $\square$ -classes makes a huge difference: We observe ([Section 7](#)) that  $\text{P}^{\text{NP}} \subseteq \text{UPostBPP}_{\square}$ ,<sup>4</sup> whereas it is known that  $\text{PostBPP}$

<sup>4</sup>This inclusion also holds for time-bounded complexity. In defining the time-bounded version of  $\text{UPostBPP}_{\square}$ , we would allow the distribution of the random string to depend nonuniformly on the input length  $n$ , though for the inclusion of  $\text{P}^{\text{NP}}$ , the distribution is computable in exponential time given the string  $1^n$ .



and  $\mathsf{P}^{\mathsf{NP}}$  are incomparable (see Section 3). Hence  $\mathsf{UPostBPP}_{\square}$  is a strict superset of both  $\mathsf{PostBPP}$  and  $\mathsf{P}^{\mathsf{NP}}$ . 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 [GLM<sup>+</sup>15].

## 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 [GL14] but was used earlier in [Kla03] 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 [GLM<sup>+</sup>15], which applies to composed functions where the gadget is as follows.

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

We introduce the confounded-majority function, defined as  $\mathsf{CONF-MAJ} := f \circ g^m$  where  $f$  is the majority function and  $g$  is the confounding gadget. Note that  $\mathsf{CONF-MAJ}$  has input length  $n := m \cdot b = m \cdot 100 \log m$  and is in  $\mathsf{PP}$  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.**  $\mathsf{UPostBPP}_{\square}(\mathsf{CONF-MAJ}) = \Theta(n)$ , and hence  $\mathsf{PP} \not\subseteq \mathsf{UPostBPP}_{\square}$ .

To prove Theorem 6 (Section 6) 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 [PSS14] showed that the “monochromatic rectangle size bound under product distributions” gives a lower bound for  $\mathsf{P}^{\mathsf{NP}}$ ). However, only a  $\Omega(\sqrt{n \log n})$  lower bound for  $\mathsf{CONF-MAJ}$  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 [GKR15, FJK<sup>+</sup>15].) To apply our general lower bound technique to  $\mathsf{CONF-MAJ}$ , we employ the communication-to-query machinery from [GLM<sup>+</sup>15] 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 [GL14] it follows immediately that to prove the Log Rank Conjecture, i.e., that  $\mathsf{P}(F) \leq \text{poly}(\log \text{rank}(F))$  for all total boolean matrices  $F$ , it suffices to prove the same with  $\mathsf{UPostBPP}_{\square}$  instead of  $\mathsf{P}$ . See Section 6 for more details.

## 3 Cartography

In this section we explore in detail the known (non-)inclusions shown on the map in Figure 1. We have not drawn any redundant arrows in the map: other relationships can be inferred from those shown; e.g., if  $\mathcal{C}_1 \not\subseteq \mathcal{C}_2$  and  $\mathcal{C}_1 \subseteq \mathcal{C}_3$  and  $\mathcal{C}_4 \subseteq \mathcal{C}_2$ , then  $\mathcal{C}_3 \not\subseteq \mathcal{C}_4$ .

**Inclusions.** The following inclusions also hold for time-bounded (Turing machine) complexity, as they do not exploit any special properties of communication. Also recall that all our classes consist of partial functions (promise problems); in particular, none of these inclusions exploit special properties of total functions.

$BPP \subseteq ZPP^{NP[1]}$ : This was first shown implicitly in [NW94, GZ11]. A particularly clean and elegant argument was given in [CC06].

$ZPP^{NP[1]} \subseteq \text{PostBPP}$ : This is our [Theorem 4](#).

$P_{\parallel}^{NP} \subseteq \text{PostBPP}$ : This was shown in [HHT97].

$P^{NP} \subseteq \text{UPostBPP}_{\square}$ : We sketch the proof of this in [Section 7 \(Observation 25\)](#).

$SBP \subseteq AM$ : This was shown in [GS86].

$AM \cap \text{coAM} \subseteq ZPP_{\parallel}^{NP[2]}$ : This follows from the well-known facts that  $AM = \text{coR} \cdot NP$  and that  $ZPP = RP \cap \text{coRP}$  relativizes. (We do not know whether  $ZPP_{\parallel}^{NP[2]}$  admits efficient amplification, but for concreteness we define it with success probability  $3/4$ .) The same argument shows that  $AM \cap \text{coAM}$  would be in  $ZPP^{NP[1]}$  if we allowed the latter to have success probability a constant less than  $1/2$ .

$P^{NP} \subseteq S_2P$ : This was shown in [RS98, Can96].

$MA \subseteq S_2P$ : This was shown in [RS98]. (It was shown in [Can96] that  $BPP \subseteq S_2P$ .)

$ZPP^{NP[1]} \subseteq S_2P$ : This was shown in [CC06].

$S_2P \subseteq ZPP^{NP}$ : This was shown in [Cai07]. See also [FIKU08].

**Non-inclusions.** For a non-inclusion  $\mathcal{C}_1 \not\subseteq \mathcal{C}_2$ , the result is strengthened if we show that some total function is in  $\mathcal{C}_1$  but not in  $\mathcal{C}_2$ . All the following non-inclusions are known to hold for a total function, except in cases where we say otherwise.

$MA \not\subseteq ZPP^{NP[1]}$ : This is our [Theorem 1](#).

$US \not\subseteq ZPP^{NP[1]}$ : This is our [Theorem 2](#).

$US \not\subseteq \text{coDP}$ : This is our [Theorem 3](#).

$RP \not\subseteq US$ : This is fairly simple to show, but was not recorded in the literature before, so we take the opportunity to do so in [Section 7 \(Observation 26\)](#).

$NP \cap \text{coNP} \not\subseteq BPP$ : This was shown in [Kla03]. Of course, it is known that  $NP \cap \text{coNP} = P$  for total functions, so the function witnessing this is necessarily partial.

$\text{coNP} \not\subseteq SBP$ : This was shown in [GW14] using the corruption lemma of [Raz92].

$P_{\parallel}^{NP[q+1]} \not\subseteq P_{\parallel}^{NP[q]}$ : This holds for all constants  $q \geq 0$  [HR90]. It is also known that  $P^{NP[q]} = P_{\parallel}^{NP[2^q-1]}$  for all constants  $q \geq 0$  [Bei91], and hence  $P^{NP[q+1]} \not\subseteq P^{NP[q]}$ .

$BPP \not\subseteq P^{NP}$ : This was implicitly shown in [PSS14], though only for a partial function (the variant of gap-Hamming-distance with a constant relative gap). Progress toward witnessing  $BPP \not\subseteq P^{NP}$  by a total function can be made in two directions: finding a total function not in  $P^{NP}$  that is in a small superclass of  $BPP$ , and

finding a total function in BPP that is not in a large subclass of  $P^{NP}$ . For the former,  $MA \not\subseteq P^{NP}$  is witnessed by BLOCK-EQ [IW10]. For the latter,  $BPP \not\subseteq P_{\parallel}^{NP}$  is witnessed by GREATER-THAN [HR90].<sup>5</sup>

SBP  $\not\subseteq$  MA: This was shown in [GLM<sup>+</sup>15], though only for a partial function.

$AM \cap coAM \not\subseteq PP$ : This was shown in [Kla11] (by combining the results of [Ver95, She11a]), though only for a partial function.

$P^{NP} \not\subseteq PP$ : This was shown in [BVdW07]; there it was only stated that  $UPP \not\subseteq PP$ , but the function witnessing this is, in fact, in  $P^{NP}$ .

$PP \not\subseteq UPostBPP_{\square}$ : This is our [Theorem 6](#). Previously,  $PP \not\subseteq PostBPP$  was shown in [Kla03], and  $PP \not\subseteq P^{NP}$  was known since the negation would imply  $\oplus P \subseteq P^{NP}$  by binary search, and a fairly simple proof that  $\oplus P \not\subseteq P^{NP}$  was given in [PSS14].

ZPP  $\not\subseteq \oplus P$ : This is fairly simple to show, but was not recorded in the literature before, so we take the opportunity to do so in [Section 7 \(Observation 27\)](#). Of course, it is known that  $ZPP = P$  for total functions, so the function witnessing this is necessarily partial. The non-equality total function NEQ witnesses  $RP \not\subseteq \oplus P$ .

$\oplus P \not\subseteq UPP$ : This was shown in [For02].

$\Pi_2 P \not\subseteq UPP$ : This was shown in [RS10].

**Open issues.** In summary, everything is now known about the relations between pairs of classes in [Figure 1](#), except for the following conjectured non-inclusions:

- $PP \not\subseteq UPostBPP$  (or even  $UPP \not\subseteq UPostBPP$ ),
- $UPostBPP \not\subseteq UPP_{\square}$  (or even  $UPP \not\subseteq UPP_{\square}$  or  $UPostBPP \not\subseteq UPostBPP_{\square}$ ),
- $AM \cap coAM \not\subseteq UPP$  (or even  $\Sigma_2 P \cap \Pi_2 P \not\subseteq UPostBPP_{\square}$  or  $AM \not\subseteq UPostBPP_{\square}$ ),
- $S_2 P \not\subseteq UPP$ ,

and except for conjectured non-inclusions that would entail explicit AM lower bounds<sup>6</sup> or explicit  $S_2 P$  lower bounds:

- $coNP \not\subseteq AM$  (or even  $PSPACE \not\subseteq AM \cap coAM$  or  $UPP \not\subseteq AM \cap coAM$ ),
- $SBP \not\subseteq \Sigma_2 P$  (or even  $PSPACE \not\subseteq S_2 P$  or  $UPP \not\subseteq S_2 P$ ),
- $\oplus P \not\subseteq \Pi_2 P$ ,
- $MA \not\subseteq ZPP_{\parallel}^{NP[2]}$ ,
- $P_{\parallel}^{NP} \not\subseteq ZPP_{\parallel}^{NP[2]}$ ,
- $AM \cap coAM \not\subseteq S_2 P$ ,
- $UPostBPP_{\square} \not\subseteq PSPACE$ ,

and except for showing the following non-inclusions for total functions:

---

<sup>5</sup>It was only claimed in [HR90] that  $GREATER-THAN \not\subseteq P_{\parallel}^{NP[q]}$  for any constant  $q$ , but in fact their proof shows that  $GREATER-THAN \not\subseteq P_{\parallel}^{NP}$ .

<sup>6</sup>Note that if we had an  $AM \cap coAM$  lower bound for an explicit function  $F$ , then we would also have an AM lower bound for the explicit function that maps  $((b, x), y) \mapsto F(x, y) \oplus b$  where  $b \in \{0, 1\}$ .

- $\text{BPP} \not\subseteq \text{P}^{\text{NP}}$  (or even  $\text{ZPP}_{\parallel}^{\text{NP}[2]} \not\subseteq \text{P}^{\text{NP}}$ ),
- $\text{SBP} \not\subseteq \text{MA}$  (or even  $\text{AM} \not\subseteq \text{MA}$ ),
- $\text{AM} \cap \text{coAM} \not\subseteq \text{PP}$  (or even  $\text{ZPP}_{\parallel}^{\text{NP}[2]} \not\subseteq \text{PostBPP}$ ).

## 4 Lower Bounds for Block-Equality and Unique-Set-Intersection

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

**Lemma 7.** *Suppose  $\mu_0$  is a distribution over  $F^{-1}(0)$ ,  $\mu_1$  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)  $\text{ZPP}^{\text{NP}[1]}(F) \geq \Omega(\log(1/\delta))$ ,
- (ii)  $\text{coDP}(F) \geq \Omega(\log(1/\delta))$ .

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  $\text{coSBP}$  [GW14]), and the other half is the 1-monochromatic rectangle size bound (which is a tight lower bound technique for  $\text{NP}$  [KN97, §2.4]). The combined technique gives a lower bound for both  $\text{ZPP}^{\text{NP}[1]}$  and  $\text{coDP}$ , even though neither of these classes appears to be a “combination” of  $\text{coSBP}$  and  $\text{NP}$ .

We prove parts (i) and (ii) of Lemma 7 in Section 4.1 and Section 4.2. Then we apply the technique to  $\text{BLOCK-EQ}$  in Section 4.3 (thus proving Theorem 1), and finally we apply the technique to  $\text{UNIQUE-INTER}$  in Section 4.4 (thus proving Theorem 2 and Theorem 3).

### 4.1 Proof of Lemma 7.(i)

Suppose for contradiction there is a cost- $o(\log(1/\delta))$   $\text{ZPP}^{\text{NP}[1]}$  protocol  $\Pi$  computing  $F$ . Then in particular we have  $\delta \leq o(1)$ . By the amplification result of [CP08], we may assume  $\mathbb{P}[\Pi(x, y) = \perp] \leq 1/10C$  for all  $(x, y) \in F^{-1}$ . By Markov’s inequality and a union bound, we may fix a  $\text{P}^{\text{NP}[1]}$ -type protocol  $\Pi^*$  in the support of  $\Pi$  such that  $\mathbb{P}_{(x,y) \sim \mu_0}[\Pi^*(x, y) = \perp] \leq 1/5C$  and  $\mathbb{P}_{(x,y) \sim \mu_1}[\Pi^*(x, y) = \perp] \leq 1/5C$ . Let the notation  $k, R_v, S_{v,w}, o_v$  be with respect to  $\Pi^*$  (see the definition of  $\text{ZPP}^{\text{NP}[1]}$  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  $\text{NP}$  oracle responds “1”) on which  $\Pi^*$  outputs  $b$ , and define  $N_b := \bigcup_{v: o_v(0)=b} (R_v \setminus \bigcup_w S_{v,w})$  as the set of “non-witnessed” inputs (the  $\text{NP}$  oracle responds “0”) on which  $\Pi^*$  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  $\Pi$ , 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  $\mu_0(\{0, 1\}^n \times \{0, 1\}^n) = \mu_0(W_0) + \mu_0(N_0) + \mu_0(W_1 \cup N_1) + \mu_0(W_{\perp} \cup N_{\perp}) \leq 1/4 + 1/4 + 0 + 1/5C < 1$ .

*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,  $\mu_0(W_0) \leq \sum_{v,w: o_v(1)=0} \mu_0(S_{v,w}) \leq 2^{o(\log(1/\delta))} \cdot \delta \leq \delta^{1-o(1)} \leq 1/4$ .  $\square$

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

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

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

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

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

$$\begin{aligned} \mu_0(N_0) &= \sum_{v: o_v(0)=0} \mu_0(R_v \setminus \bigcup_w S_{v,w}) \\ &\leq \sum_{v: o_v(0)=0} (C \cdot \mu_1(\bigcup_w S_{v,w}) + \delta) \\ &\leq C \cdot \left( \sum_{v: o_v(0)=0} \mu_1(\bigcup_w S_{v,w}) \right) + 2^{o(\log(1/\delta))} \cdot \delta \\ &\leq C \cdot (1/5C + \delta^{1-o(1)}) + \delta^{1-o(1)} \\ &\leq 1/4. \end{aligned} \quad \square$$

## 4.2 Proof of Lemma 7.(ii)

Suppose for contradiction there is a cost- $k$  coDP protocol  $\Pi$  computing  $F$  where  $k \leq o(\log(1/\delta))$ . Then in particular we have  $\delta \leq o(1)$ . 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  $\mu_0(\bigcup_w S_w) = 1$ , there exists a  $w^*$  such that  $\mu_0(S_{w^*}) \geq 2^{-k} \geq \delta^{1/3}$  and hence  $\mu_1(S_{w^*}) \geq \frac{1}{C} \cdot (\delta^{1/3} - \delta) \geq \delta^{1/2}$ . Since  $S_{w^*} \cap F^{-1}(1) \subseteq \bigcup_w T_w$ , there exists a  $w'$  such that  $\mu_1(T_{w'}) \geq \mu_1(S_{w^*} \cap F^{-1}(1)) \cdot 2^{-k} > \delta^{1/2} \cdot \delta^{1/2} = \delta$ . But  $T_{w'}$  is 1-monochromatic since  $F^{-1}(0) \cap \bigcup_w T_w = \emptyset$ , so this is a contradiction.

## 4.3 Proof of Theorem 1

Let  $\mu_0$  be the uniform distribution over  $\text{BLOCK-EQ}^{-1}(0)$ , and let  $\mu_1$  be the uniform distribution over the subset of  $\text{BLOCK-EQ}^{-1}(1)$  consisting of all  $(x,y)$  for which  $x_i = y_i$  for a unique  $i$ .

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

**Lemma 11.**  $\mu_1(R) \leq 2^{-\Omega(\sqrt{n})}$  holds for every 1-monochromatic rectangle  $R$  of  $\text{BLOCK-EQ}$ .

Together, Lemma 10 and Lemma 11 show that the hypothesis of Lemma 7 holds with  $F := \text{BLOCK-EQ}$ ,  $C := 45$ , and  $\delta := 2^{-\Omega(\sqrt{n})}$ . The lower bound in Theorem 1 follows. For the upper bound, in fact  $\text{ZPP}(\text{BLOCK-EQ}) \leq O(\sqrt{n})$  holds [KN97, §4.1.1] (though it is slightly quicker to see that  $\text{NP}(\text{BLOCK-EQ}) \leq O(\sqrt{n})$  holds by guessing  $i$  and deterministically verifying that  $x_i = y_i$ ).

For the proofs of the lemmas, we define  $m := \sqrt{n}$  and  $b := \sqrt{n}$  (as in the notation for the decomposition  $\text{BLOCK-EQ} := \text{OR} \circ \text{EQ}^m$  where  $\text{EQ}$  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

$$\Xi := ((x_1^0, x_1^1, y_1^0, y_1^1), \dots, (x_m^0, x_m^1, y_m^0, y_m^1))$$

is valid iff it is a tuple of valid tuples. If  $\Xi$  is valid then the injection  $\Phi_\Xi: \{0, 1\}^m \times \{0, 1\}^m \rightarrow \{0, 1\}^n \times \{0, 1\}^n$  defined by

$$\Phi_\Xi(u, v) := (x_1^{u_1} \cdots x_m^{u_m}, y_1^{v_1} \cdots y_m^{v_m})$$

is a reduction from  $\text{INTER} := \text{OR} \circ \text{AND}^m$  (for single-bit AND) to  $\text{BLOCK-EQ}$ :

$$\text{INTER}(u, v) = \text{BLOCK-EQ}(\Phi_\Xi(u, v)).$$

(In other words, the image of  $\Phi_\Xi$ , as a submatrix of the  $\text{BLOCK-EQ}$  matrix, is a copy of the  $\text{INTER}$  matrix.)

Define  $\text{UNAMBIG-INTER} := \text{UNAMBIG-OR} \circ \text{AND}^m$  where the partial function  $\text{UNAMBIG-OR}$  is  $\text{OR}$  restricted to the domain of strings of Hamming weight 0 or 1. That is,  $\text{UNAMBIG-INTER}^{-1}(0)$  consists of all pairs of disjoint sets, and  $\text{UNAMBIG-INTER}^{-1}(1)$  consists of all pairs of uniquely intersecting sets.

**Lemma 12 ([Raz92]).** *There exists a distribution  $\nu_0$  over  $\text{UNAMBIG-INTER}^{-1}(0)$  and a distribution  $\nu_1$  over  $\text{UNAMBIG-INTER}^{-1}(1)$  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  $\nu_1$  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  $\Xi$  is chosen uniformly at random independently of  $\nu_a$ . In other words,  $\mu_a$  equals the distribution obtained by choosing  $\Xi$ , then independently taking a sample from  $\nu_a$ , then applying  $\Phi_\Xi$  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  $(u, v) \in \text{UNAMBIG-INTER}^{-1}(1)$ , say with  $u_i = v_i = 1$ , the distribution  $\mathbb{E}_\Xi \Phi_\Xi(u, v)$  is uniform over the subset of  $\text{BLOCK-EQ}^{-1}(1)$  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  $\Xi$  are independent, that  $(x_i^1, y_i^1)$  is uniformly distributed over  $\text{EQ}^{-1}(1)$ , 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  $\text{EQ}^{-1}(0)$ . 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  $\Phi_\Xi$ , then we have

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

*Proof of Lemma 11.* Note that  $\mu_1$  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  $\text{BLOCK-EQ}(x, y) = 1$ ), and similarly  $|B| \leq m \cdot 2^{b(m-1)}$ , and hence  $|R| \leq m^2 \cdot 2^{2b(m-1)}$ . It follows that

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

#### 4.4 Proof of Theorem 2 and Theorem 3

We again use the corruption lemma from [Raz92], but now we need to take a closer look at the distribution over 1-inputs. Let  $n = 4\ell - 1$ . Let  $\mu_0$  be the distribution over  $\text{UNIQUE-INTER}^{-1}(0)$  that samples uniformly random disjoint sets of size  $\ell$ , and let  $\mu_1$  be the distribution over  $\text{UNIQUE-INTER}^{-1}(1)$  that samples uniformly random uniquely intersecting sets of size  $\ell$ .

**Lemma 13** ([Raz92]).  $\mu_0(R) \leq 45 \cdot \mu_1(R) + 2^{-\Omega(n)}$  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  $\text{UNIQUE-INTER}$ .

Together, Lemma 13 and Lemma 14 show that the hypothesis of Lemma 7 holds with  $F := \text{UNIQUE-INTER}$ ,  $C := 45$ , and  $\delta := 2^{-\Omega(n)}$ . Theorem 2 and Theorem 3 follow.

*Proof of Lemma 14.* For each  $i \in [n]$  let us define the rectangle  $R_i := \{(x, y) \in R : x_i = y_i = 1\}$ , 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  $[n] \setminus \{i\}$ .<sup>7</sup> Hence  $|R| \leq n2^{n-1} \leq 2^{(1+o(1))n}$ .

Note that  $\mu_1$  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  $x := \{i\} \cup H_1$  and  $y := \{i\} \cup H_2$ .

Hence  $\mu_1$  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  $\mu_1(R) \leq 2^{(1+o(1))n} / 2^{(1.5-o(1))n} \leq 2^{-\Omega(n)}$ .  $\square$

## 5 Proof of Theorem 4

Consider an optimal  $\text{ZPP}^{\text{NP}[1]}$  protocol  $\Pi$  for  $F$  with deterministic communication cost  $c$  and witness length  $k$ . By standard sparsification, we may assume the public randomness is uniformly distributed over  $\{0, 1\}^{O(\log n)}$ . We use  $r$  to denote an outcome of this randomness, we let  $\Pi^r$  denote the  $\text{P}^{\text{NP}[1]}$ -type protocol induced by  $r$ , and we let the notation  $R_v^r, S_{v,w}^r, o_v^r$  be with respect to  $\Pi^r$  (see the

<sup>7</sup>By [KW15], the bound  $|R_i| \leq 2^{n-1}$  also holds assuming every input in  $R$  has intersection size either 1 or  $\geq 3$ . Using this, it follows that Theorem 2 and Theorem 3 hold under the promise that at most two coordinates intersect.

definition of  $\text{ZPP}^{\text{NP}[1]}$  in Section 2.1). We claim that the following protocol  $\tilde{\Pi}$  is a PostBPP protocol for  $F$  of cost  $O(\text{ZPP}^{\text{NP}[1]}(F) + \log n)$ .

**Input:**  $(x, y)$   
**Output:**  $\in \{0, 1, \perp\}$

- 1 pick  $r$  uniformly at random
- 2 run the deterministic part of  $\Pi^r(x, y)$  to a leaf  $v^r$
- 3 pick  $a \in \{0, 1\}$  uniformly at random
- 4 **if**  $a = 1$  **then**
- 5 pick  $w \in \{0, 1\}^k$  uniformly at random
- 6 **if**  $(x, y) \in S_{v^r, w}^r$  **then** output  $o_{v^r}^r(1)$
- 7 **else** output  $\perp$
- 8 **else if**  $a = 0$  **then**
- 9 output  $o_{v^r}^r(0)$  with probability  $2^{-(k+2)}$
- 10 output  $\perp$  with the remaining probability
- 11 **end**

$\tilde{\Pi}$  has communication cost  $c + O(1)$  and randomness cost  $O(\log n) + k + O(1)$  and hence cost  $O(c + k + \log n)$ . We now argue the correctness. Let  $\tilde{\Pi}^r$  denote  $\tilde{\Pi}$  with a particular  $r$  chosen on line 1. Fix an input  $(x, y)$ , and let  $\chi^r \in \{0, 1\}$  indicate whether  $(x, y) \in \bigcup_w S_{v^r, w}^r$  (i.e., the NP oracle's response). Let  $A := \{r : \chi^r = 1\}$ ,  $\bar{A} := \{r : \chi^r = 0\}$ ,  $B := \{r : o_{v^r}^r(\chi^r) = F(x, y)\}$ , and  $\bar{B} := \{r : o_{v^r}^r(\chi^r) = \perp\}$ . We have

$$\mathbb{P}[\tilde{\Pi}^r(x, y) = F(x, y)] \geq \begin{cases} 2^{-(k+1)} & \text{if } r \in A \cap B \\ 0 & \text{if } r \in A \cap \bar{B} \\ 2^{-(k+3)} & \text{if } r \in \bar{A} \cap B \\ 0 & \text{if } r \in \bar{A} \cap \bar{B} \end{cases}$$

and

$$\mathbb{P}[\tilde{\Pi}^r(x, y) = 1 - F(x, y)] \leq \begin{cases} 2^{-(k+3)} & \text{if } r \in A \\ 0 & \text{if } r \in \bar{A} \end{cases}$$

Thus since  $\mathbb{P}[r \in B] \geq 3/4$  we have

$$\begin{aligned} \mathbb{P}[\tilde{\Pi}(x, y) = F(x, y)] &\geq 2^{-(k+3)} \cdot \mathbb{P}[r \in B] + (2^{-(k+1)} - 2^{-(k+3)}) \cdot \mathbb{P}[r \in A \cap B] \\ &= 2^{-(k+3)} \cdot (\mathbb{P}[r \in B] + 3 \cdot \mathbb{P}[r \in A \cap B]) \\ &\geq 2^{-(k+3)} \cdot (3 \cdot \mathbb{P}[r \in \bar{B}] + 3 \cdot \mathbb{P}[r \in A \cap B]) \\ &\geq 3 \cdot 2^{-(k+3)} \cdot \mathbb{P}[r \in A] \\ &\geq 3 \cdot \mathbb{P}[\tilde{\Pi}(x, y) = 1 - F(x, y)] \end{aligned}$$

so  $\tilde{\Pi}$  is a correct PostBPP protocol for  $F$ .



## 6 Lower Bound for Majority

In this section we prove [Theorem 6](#). We first give a general lower bound technique for  $\text{UPostBPP}_\square$  in [Section 6.1](#), then we describe the machinery we borrow from [\[GLM<sup>+</sup>15\]](#) in [Section 6.2](#), and finally we give the proof of [Theorem 6](#) in [Section 6.3](#).

### 6.1 Lower bound technique

**Definition 15.** For  $F: \{0, 1\}^n \times \{0, 1\}^n \rightarrow \{0, 1\}$ ,  $R \subseteq \{0, 1\}^n \times \{0, 1\}^n$ , and  $\mu$  a distribution over  $\{0, 1\}^n \times \{0, 1\}^n$ , we say  $R$  is  $\mu$ -unbiased (with respect to  $F$ ) if  $\frac{1}{2} \cdot \mu(R \cap F^{-1}(0)) \leq \mu(R \cap F^{-1}(1)) \leq 2 \cdot \mu(R \cap F^{-1}(0))$ , and is  $\mu$ -biased otherwise.

**Lemma 16.** Suppose  $\mu$  is a distribution over  $F^{-1}$  and  $\rho$  is a product distribution over  $\{0, 1\}^n \times \{0, 1\}^n$  such that for every rectangle  $R \subseteq \{0, 1\}^n \times \{0, 1\}^n$ , if  $\rho(R) \geq \delta$  then  $R$  is  $\mu$ -unbiased (with respect to  $F$ ), and if  $\rho(R) \geq 1/2$  then  $\mu(R) > 0$ . Then  $\text{UPostBPP}_\square(F) \geq \Omega(\log(1/\delta))$ .

The case where  $\mu = \rho$  is equivalent to extended discrepancy [\[GL14\]](#) under product distributions, and leads to the lower bound  $\text{UPostBPP}_\square(\text{CONF-MAJ}) \geq \Omega(\sqrt{n \log n})$  (details omitted). The more general form is needed to get the  $\Omega(n)$  lower bound. The results of [\[GL14\]](#) show that for total  $F$ ,  $\text{P}(F) \leq \text{poly}(\log n)$  follows from the assumptions that  $F$ 's matrix has  $\text{poly}(\log n)$  rank (over the reals) and that every rectangle  $S$  has a subrectangle that has measure  $\geq 2^{-\text{poly}(\log n)}$  and is biased (both with respect to the uniform distribution over  $S$ ). By letting  $\rho = \mu$  be uniform over an arbitrary  $S$  in [Lemma 16](#), the latter property follows from the existence of a  $\text{poly}(\log n)$ -cost  $\text{UPostBPP}_\square$  protocol. Hence to prove the Log Rank Conjecture, it suffices to prove the same with  $\text{UPostBPP}_\square$  instead of  $\text{P}$ .

*Proof of Lemma 16.* Suppose  $\Pi$  is a cost- $k$   $\text{UPostBPP}_\square$  protocol for  $F$ . By [Observation 51](#), we may assume  $\Pi$  is just a distribution over  $2^k$  many  $\{0, 1\}$ -labeled rectangles. For  $R$  a rectangle and  $o \in \{0, 1\}$ , we let  $\pi_{R,o}$  denote the probability of getting  $R$  with label  $o$  under  $\Pi$ .

We start by recording the following observation: For every distribution  $\nu$  over  $F^{-1}$  there exists an  $(R, o)$  such that  $\pi_{R,o} > 0$  and  $R$  is  $\nu$ -biased. This follows by considering the 2-player 0-sum game where the row strategies are inputs  $(x, y) \in F^{-1}$ , the column strategies are  $(R, o)$  pairs with  $\pi_{R,o} > 0$ , and the payoff to the column player is 1 if  $(x, y) \in R$  and  $F(x, y) = o$ , is  $-2$  if  $(x, y) \in R$  and  $F(x, y) = 1 - o$ , and is 0 if  $(x, y) \notin R$ . The mixed column strategy  $\pi$  demonstrates that the game has positive value, and hence for every mixed row strategy there exists a pure column strategy for which the expected payoff to the column player is positive. This implies the observation. Only the straightforward direction of the Minimax Theorem is used.

Consider the following procedure.

```

1 let  $Q_0 := \{0, 1\}^n \times \{0, 1\}^n$ 
2 for  $i = 1, 2, \dots$  do
3   let  $R_i$  be a  $(\mu \upharpoonright Q_{i-1})$ -biased rectangle such that  $\pi_{R_i, o_i} > 0$  for some  $o_i$ 
4   let  $A_i \times B_i := R_i \cap Q_{i-1}$ 
5   let  $Q_i := Q_{i-1} \setminus (\text{either } A_i \times \{0, 1\}^n \text{ or } \{0, 1\}^n \times B_i, \text{ whichever is smaller under } \rho)$ 
6 until  $\rho(Q_i) < 1/2$ 

```

We show by induction on  $i$  that lines 3 and 4 always succeed,  $Q_i$  is a rectangle,  $\rho(Q_i) \geq 1 - i \cdot \sqrt{\delta}$ , and the  $R_j$ 's for  $j \in \{1, \dots, i\}$  are all distinct from each other and disjoint from  $Q_i$ . The base case

$i = 0$  is trivial, so assume this holds for  $i - 1$ . Since  $Q_{i-1}$  is a rectangle and  $\rho(Q_{i-1}) \geq 1/2$  by line 6, we have  $\mu(Q_{i-1}) > 0$  by assumption (with  $R := Q_{i-1}$ ) and hence the conditioning on line 3 is valid. By the above observation (with  $\nu := (\mu \mid Q_{i-1})$ ), line 3 succeeds. Since  $Q_{i-1}$  is a rectangle, so are  $R_i \cap Q_{i-1}$  (hence line 4 succeeds) and  $Q_i$ . Since  $R_i$  is  $(\mu \mid Q_{i-1})$ -biased, we have that  $R_i \cap Q_{i-1}$  is  $\mu$ -biased and hence  $\rho(R_i \cap Q_{i-1}) < \delta$  by assumption (with  $R := R_i \cap Q_{i-1}$ ). Since  $\rho$  is a product distribution, either  $\rho(A_i \times \{0, 1\}^n) < \sqrt{\delta}$  or  $\rho(\{0, 1\}^n \times B_i) < \sqrt{\delta}$ . Hence  $\rho(Q_i) \geq \rho(Q_{i-1}) - \sqrt{\delta} \geq 1 - i \cdot \sqrt{\delta}$ . Since  $R_i \cap Q_{i-1}$  is  $\mu$ -biased,  $R_i$  is not disjoint from  $Q_{i-1}$  and hence  $R_i$  is distinct from the  $R_j$ 's for  $j \in \{1, \dots, i-1\}$  (since the latter are all disjoint from  $Q_{i-1}$ ). Since  $Q_i \subseteq Q_{i-1}$ , the  $R_j$ 's for  $j \in \{1, \dots, i-1\}$  are also disjoint from  $Q_i$ . Finally, line 5 ensures that  $R_i$  is disjoint from  $Q_i$ , since  $Q_i \subseteq Q_{i-1} \setminus (A_i \times B_i) = Q_{i-1} \setminus R_i$ . This completes the induction step.

Since the  $R_i$ 's are all distinct and  $\pi_{R_i, o_i} > 0$ , there are at most  $2^k$  iterations. Let  $i^*$  be the final value of  $i$ . By line 6, we have  $1/2 > \rho(Q_{i^*}) \geq 1 - i^* \cdot \sqrt{\delta}$  and hence  $2^k \geq i^* > 1/2\sqrt{\delta}$ . Thus  $k > \frac{1}{2} \cdot \log(1/\delta) - 1$ .  $\square$

## 6.2 Conjunction rectangles

We now state the ‘‘Packing with Conjunctions Theorem’’ from [GLM<sup>+</sup>15], which is the technical heart of the main ‘‘Junta Theorem’’ from that paper. The theorem makes no reference to the outer function  $f$ ; it is simply a statement about the function  $G := g^m$  where  $g$  is the confounding gadget with block length  $b$ .

**Definition 17.** *Two distributions over  $\{0, 1\}^m$  are  $\epsilon$ -close if for every  $z \in \{0, 1\}^m$ , the probabilities of  $z$  under the two distributions are within a factor  $(1 \pm \epsilon)$  of each other.*

**Definition 18.** *A rectangle  $S$  is a  $(d, \epsilon)$ -conjunction rectangle if there exists a width- $d$  conjunction  $h: \{0, 1\}^m \rightarrow \{0, 1\}$  (i.e.,  $h$  can be written as  $(\ell_1 \wedge \dots \wedge \ell_w)$  where  $w \leq d$  and each  $\ell_i$  is an input variable or its negation) such that the distributions over  $\{0, 1\}^m$  obtained in the following two ways are  $\epsilon$ -close:*

- picking a uniformly random  $z \in \{0, 1\}^m$  and a uniformly random  $(x, y) \in G^{-1}(z)$  and conditioning on  $(x, y) \in S$ ,
- picking a uniformly random  $z \in h^{-1}(1)$ .

**Definition 19.** *A distribution  $\nu$  over  $\{0, 1\}^n \times \{0, 1\}^n$  is the lift of a distribution  $\xi$  over  $\{0, 1\}^m$  if  $\nu(x, y) = \xi(z)/|G^{-1}(z)|$  where  $z := G(x, y)$ . Note that a lifted distribution is a convex combination of distributions that are uniform over a set  $G^{-1}(z)$ .*

**Theorem 20.** *For  $\epsilon := 1/100$ , and for every  $d \geq 0$ , every lifted distribution  $\nu$ , and every rectangle  $R$  with  $\nu(R) \geq 2^{-db/20}$ , there exist disjoint  $(d, \epsilon)$ -conjunction subrectangles  $S_1, S_2, \dots \subseteq R$  such that  $\nu(\bigcup_i S_i \mid R) \geq 1 - \epsilon$ .*

The proof in [GLM<sup>+</sup>15] actually gives  $\epsilon := 2^{-\Theta(b)}$ , but we only need  $\epsilon := 1/100$ .

## 6.3 Proof of Theorem 6

For convenience, assume  $m$  is odd. We have CONF-MAJ  $:= f \circ G$  where  $G := g^m$ . Let  $M := \{z \in \{0, 1\}^m : |z| \in \{\lfloor m/2 \rfloor, \lceil m/2 \rceil\}\}$  (the ‘‘middle layers’’ of the Hamming cube), and let  $L := G^{-1}(M)$

(the “lifted version” of the set  $M$ ). Let the distribution  $\mu$  be the lift of the uniform distribution over  $M$  (so  $\mu$  is supported on  $L$ ), and let  $\rho$  be the uniform distribution over  $\{0, 1\}^n \times \{0, 1\}^n$  (which is a product distribution). We argue the following two claims, both of which exploit [Theorem 20](#). Recall that  $b := 100 \log m$  and  $n := m \cdot b$ .

**Claim 21.** *For every rectangle  $R$ , if  $\rho(R) \geq 0.99999^n$  then  $\mu(R) \geq 0.997^n$ .*

**Claim 22.** *For every rectangle  $R$ , if  $\mu(R) \geq 0.997^n$  then  $R$  is  $\mu$ -unbiased.*

[Theorem 6](#) follows because the assumptions of [Lemma 16](#) hold with  $\delta := 0.99999^n$ : The first part (if  $\rho(R) \geq \delta$  then  $R$  is  $\mu$ -unbiased) holds by [Claim 21](#) and [Claim 22](#). The second part (if  $\rho(R) \geq 1/2$  then  $\mu(R) > 0$ ) holds by [Claim 21](#) alone.

**Observation 23.** *Let  $\rho'$  be the lift of the uniform distribution over  $\{0, 1\}^m$ , and note the following.*

- (i)  $\rho$  and  $\rho'$  are  $(1/2)$ -close. (This is straightforward to verify using  $b := 100 \log m$  and the fact that  $|g^{-1}(0)|, |g^{-1}(1)| \in 2^{2b-1} \pm 2^{b+1}$ .)
- (ii)  $\mu(\cdot) = \rho'(\cdot | L)$ .
- (iii) The first distribution in [Definition 18](#) picks  $z$  with probability  $\rho'(G^{-1}(z) | S)$ ; hence this value is in  $(1 \pm \epsilon) \cdot \mathbb{P}_{z' \in h^{-1}(1)}[z' = z]$  where the notation  $\mathbb{P}_{z' \in h^{-1}(1)}$  means a uniformly random choice.

In the proof of [Claim 21](#) we use the following fact, which holds by Stirling approximations.

**Fact 24.** *For all  $s \geq t$  we have  $\binom{s}{t} = \Theta\left(\frac{1}{\sqrt{s}} \cdot \left(\frac{s}{t}\right)^{t+1/2} \cdot \left(\frac{s}{s-t}\right)^{s-t+1/2}\right) \geq \Omega\left(\frac{1}{\sqrt{s}} \cdot \left(\frac{s}{\max(t, s-t)}\right)^s\right)$ .*

*Proof of Claim 21.* Assuming  $\rho(R) \geq 0.99999^n$ , we have  $\rho'(R) \geq 0.99998^n$ . Apply [Theorem 20](#) with  $\nu := \rho'$  and  $d := m/1000$  (noting that  $0.99998^n \geq 2^{-db/20}$ ) to get disjoint  $(d, \epsilon)$ -conjunction subrectangles  $S_1, S_2, \dots \subseteq R$  with associated conjunctions  $h_1, h_2, \dots$ , such that  $\rho'(\bigcup_i S_i | R) \geq 1 - \epsilon$  (where  $\epsilon := 1/100$ ). For each  $i$ , assuming for convenience that  $h_i$  depends on exactly  $d$  variables, exactly  $j$  of which are positive literals, we have

$$\begin{aligned} |h_i^{-1}(1) \cap M| &= \binom{m-d}{\lfloor m/2 \rfloor - j} + \binom{m-d}{\lceil m/2 \rceil - j} \\ &\geq \binom{m-d}{\lfloor m/2 \rfloor - j} \\ &\geq \Omega\left(\frac{1}{\sqrt{m}} \cdot \left(\frac{m-d}{\lfloor m/2 \rfloor}\right)^{m-d}\right) \\ &\geq 2^{m-d} \cdot \Omega\left(\frac{1}{\sqrt{m}} \cdot 0.9987^{m-d}\right) \\ &\geq 2^{m-d} \cdot 0.9985^m \end{aligned}$$

where the third line follows by [Fact 24](#) and the fourth line follows by  $d := m/1000$ . Thus we have

$$\mathbb{P}_{z \in h_i^{-1}(1)}[z \in M] = \frac{|h_i^{-1}(1) \cap M|}{2^{m-d}} \geq 0.9985^m, \text{ and hence}$$

$$\begin{aligned} \rho'(L | S_i) &= \sum_{z \in M} \rho'(G^{-1}(z) | S_i) \\ &\in \sum_{z \in M} (1 \pm \epsilon) \cdot \mathbb{P}_{z' \in h_i^{-1}(1)}[z' = z] \\ &= (1 \pm \epsilon) \cdot \mathbb{P}_{z \in h_i^{-1}(1)}[z \in M] \\ &\geq 0.9983^m \end{aligned}$$

where the second line follows since  $S_i$  is a  $(d, \epsilon)$ -conjunction rectangle. Then we have

$$\begin{aligned}
\rho'(L | R) &\geq \sum_i \rho'(L \cap S_i | R) \\
&= \sum_i \rho'(L | S_i) \cdot \rho'(S_i | R) \\
&\geq \sum_i 0.9983^m \cdot \rho'(S_i | R) \\
&= 0.9983^m \cdot \rho'(\bigcup_i S_i | R) \\
&\geq 0.998^m
\end{aligned}$$

where the last line follows by  $\rho'(\bigcup_i S_i | R) \geq 1 - \epsilon$ , and finally

$$\mu(R) = \rho'(R | L) \geq \rho'(L | R) \cdot \rho'(R) \geq 0.998^m \cdot 0.99998^n \geq 0.997^n. \quad \square$$

*Proof of Claim 22.* Apply [Theorem 20](#) with  $\nu := \mu$  and  $d := m/10$  (noting that  $0.997^n \geq 2^{-db/20}$ ) to get disjoint  $(d, \epsilon)$ -conjunction subrectangles  $S_1, S_2, \dots \subseteq R$  with associated conjunctions  $h_1, h_2, \dots$ , such that  $\mu(\bigcup_i S_i | R) \geq 1 - \epsilon$  (where  $\epsilon := 1/100$ ). Recall that  $f: \{0, 1\}^m \rightarrow \{0, 1\}$  is the majority function. For each  $i$ , assuming for convenience that  $h_i$  depends on exactly  $d$  variables, exactly  $j$  of which are positive literals, we have

$$\frac{\mathbb{P}_{z \in h_i^{-1}(1)}[z \in f^{-1}(0) \cap M]}{\mathbb{P}_{z \in h_i^{-1}(1)}[z \in f^{-1}(1) \cap M]} = \frac{\binom{m-d}{\lfloor m/2 \rfloor - j}}{\binom{m-d}{\lceil m/2 \rceil - j}} = \frac{\lceil m/2 \rceil - j}{\lceil m/2 \rceil - d + j} = 1 + \frac{d-2j}{\lceil m/2 \rceil - d + j} \in \left[\frac{3}{4}, \frac{4}{3}\right]$$

since  $d := m/10$ . Now fix any output  $o \in \{0, 1\}$ , and let  $E_o := \text{CONF-MAJ}^{-1}(o) = G^{-1}(f^{-1}(o))$ . We have

$$\begin{aligned}
\rho'(E_o \cap L | S_i) &= \sum_{z \in f^{-1}(o) \cap M} \rho'(G^{-1}(z) | S_i) \\
&\in \sum_{z \in f^{-1}(o) \cap M} (1 \pm \epsilon) \cdot \mathbb{P}_{z' \in h_i^{-1}(1)}[z' = z] \\
&= (1 \pm \epsilon) \cdot \mathbb{P}_{z \in h_i^{-1}(1)}[z \in f^{-1}(o) \cap M] \\
&\geq (1 - \epsilon) \cdot \frac{3}{4} \cdot \mathbb{P}_{z \in h_i^{-1}(1)}[z \in f^{-1}(1-o) \cap M] \\
&\geq (1 - \epsilon) \cdot \frac{3}{4} \cdot (1 - \epsilon) \cdot \rho'(E_{1-o} \cap L | S_i) \\
&\geq \frac{2}{3} \cdot \rho'(E_{1-o} \cap L | S_i).
\end{aligned}$$

If  $\mu(S_i) > 0$  (equivalently,  $\rho'(L | S_i) > 0$ ) then

$$\begin{aligned}
\mu(E_o | S_i) &= \rho'(E_o | L \cap S_i) \\
&= \rho'(E_o \cap L | S_i) / \rho'(L | S_i) \\
&\geq \frac{2}{3} \cdot \rho'(E_{1-o} \cap L | S_i) / \rho'(L | S_i) \\
&= \frac{2}{3} \cdot \mu(E_{1-o} | S_i)
\end{aligned}$$

and hence  $\mu(E_o | S_i) \geq \frac{2}{5}$ . Now we have

$$\begin{aligned}
\mu(E_o | R) &\geq \sum_i \mu(E_o \cap S_i | R) \\
&= \sum_{i: \mu(S_i) > 0} \mu(E_o | S_i) \cdot \mu(S_i | R) \\
&\geq \sum_{i: \mu(S_i) > 0} \frac{2}{5} \cdot \mu(S_i | R)
\end{aligned}$$

$$\begin{aligned}
&= \frac{2}{5} \cdot \mu(\bigcup_i S_i \mid R) \\
&\geq \frac{1}{3}.
\end{aligned}$$

where the last line follows by  $\mu(\bigcup_i S_i \mid R) \geq 1 - \epsilon$ . Thus  $\mu(E_o \mid R) \geq \frac{1}{2} \cdot \mu(E_{1-o} \mid R)$  and hence  $\mu(R \cap E_o) \geq \frac{1}{2} \cdot \mu(R \cap E_{1-o})$ . Since this holds for either  $o \in \{0, 1\}$ ,  $R$  is  $\mu$ -unbiased with respect to CONF-MAJ.  $\square$

## 7 Additional Observations

**Observation 25.**  $\text{UPostBPP}_{\square}(F) \leq O(\text{P}^{\text{NP}}(F))$  for all  $F$ , and hence  $\text{P}^{\text{NP}} \subseteq \text{UPostBPP}_{\square}$ .

*Proof sketch.* It is a classical fact that if we consider “super-witnesses” that consist of a string of purported responses to the NP oracle queries along with purported witnesses for all the queries for which the purported response is “1”, and if we order the super-witnesses reverse-lexicographically by the string of oracle responses (the witnesses do not matter for the ordering), then the output of a  $\text{P}^{\text{NP}}$  computation is determined by the first super-witness for which all the purported oracle query witnesses check out. (This fact was phrased as an “overlay” characterization in [PSS14] and was also used in the proof that  $\text{P}^{\text{NP}} \subseteq \text{S}_2\text{P}$  [RS98, Can96].) To get a  $\text{UPostBPP}_{\square}$  protocol, we can pick a random super-witness with probabilities geometrically decreasing according to the order, output  $\perp$  if one of the purported witnesses does not check out, and otherwise produce the same output as the computation path given the purported oracle responses.  $\square$

Let NEQ be the non-equality function, which is in RP.

**Observation 26.**  $\text{US}(\text{NEQ}) = \Theta(n)$ , and hence  $\text{RP} \not\subseteq \text{US}$ .

*Proof.* The matrix of NEQ is the complement of the identity matrix. Consider a collection of rectangles that touches each off-diagonal entry exactly once, and touches each diagonal entry either zero times or at least twice. If we sum these rectangles as 0-1 matrices over the reals, the resulting matrix  $M$  has all off-diagonal entries = 1 and all diagonal entries  $\neq 1$ . Subtracting the all-1’s matrix from  $M$  results in a diagonal matrix with all nonzero diagonal entries, which has full rank. Thus  $M$  has rank at least  $2^n - 1$  since the all-1’s matrix has rank 1. However, the number of rectangles upper bounds the rank since each rectangle has rank 1.  $\square$

Consider the partial function WHICH-EQ:  $\{0, 1\}^n \times \{0, 1\}^n \rightarrow \{0, 1\}$  that partitions each of its two inputs into two strings of length  $n/2$ ,  $x := x_0x_1$  and  $y := y_0y_1$ , where  $\text{WHICH-EQ}(x, y) = 0$  iff  $x_0 = y_0$  and  $x_1 \neq y_1$ , and  $\text{WHICH-EQ}(x, y) = 1$  iff  $x_0 \neq y_0$  and  $x_1 = y_1$ . Note that  $\text{WHICH-EQ} \in \text{ZPP}$ .

**Observation 27.**  $\oplus\text{P}(\text{WHICH-EQ}) = \Theta(n)$ , and hence  $\text{ZPP} \not\subseteq \oplus\text{P}$ .

*Proof.* Consider any total boolean matrix  $M$  that agrees with WHICH-EQ on the latter’s domain. We claim that  $M$  contains a  $2^{n/2-1} \times 2^{n/2-1}$  identity or complement-of-identity submatrix; hence  $M$  has rank at least  $2^{n/2-1} - 1$  over  $GF(2)$ . If  $M(z, z) = 1$  for at least half of all  $z \in \{0, 1\}^n$ , fix  $z_0$  so that  $M(z_0z_1, z_0z_1) = 1$  for all  $z_1$  in some  $Z_1 \subseteq \{0, 1\}^{n/2}$  of size  $2^{n/2-1}$ , and note that for  $x_1, y_1 \in Z_1$ ,  $M(z_0x_1, z_0y_1)$  indicates whether  $x_1 = y_1$ . On the other hand, if  $M(z, z) = 0$  for at least half of all  $z \in \{0, 1\}^n$ , fix  $z_1$  so that  $M(z_0z_1, z_0z_1) = 0$  for all  $z_0$  in some  $Z_0 \subseteq \{0, 1\}^{n/2}$  of size  $2^{n/2-1}$ , and note that for  $x_0, y_0 \in Z_0$ ,  $M(x_0z_1, y_0z_1)$  indicates whether  $x_0 \neq y_0$ .  $\square$

## 8 Conclusion and Open Questions

It is open to prove that any explicit function is not in  $S_2P$ ; we wish to highlight this as a new frontier (presumably incomparable to the  $AM \cap coAM$  frontier) toward proving explicit lower bounds for the communication polynomial hierarchy.

Is [Lemma 16](#) a tight lower bound technique for  $UPostBPP_{\square}$ ? (This is related to the open question of whether the lower bound technique given in [\[PSS14\]](#) for  $P^{NP}$  is tight.) It is also open to prove a  $UPostBPP_{\square}$  lower bound for the majority function lifted with a constant-size gadget (which, without loss of generality, would be AND or XOR). Finally, we mention that for some known results, there is room for quantitative improvement; e.g., is there an  $F \in MA$  such that  $ZPP^{NP[1]}(F) \geq \omega(\sqrt{n})$ ?

Our survey in [Section 3](#) lists all the open problems that fall directly within the scope of this paper. Although we aimed for our survey to be fairly comprehensive, there are some further topics concerning communication complexity classes that we have not addressed. Our discussion has excluded classes involving limited ambiguity (such as UP, FewP [\[KNSW94, Kla10\]](#), and UAM [\[GPW15\]](#)), more exotic counting classes (such as Few, APP,<sup>8</sup> WPP, AWPP, WAPP, LWPP, SPP,  $C=P$ , and  $Mod_mP$  [\[DKMW04\]](#) for integers  $m > 2$ ), classes defined using the dot operator (such as  $BP \cdot UP$  [\[Kla10\]](#),  $U \cdot BPP$ , and  $N \cdot BPP$  which may differ from MA), and classes with oracles other than NP. One can also ask about more complicated relationships among the classes (e.g., concerning intersections of different classes, although we have mentioned  $NP \cap coNP$ ,  $AM \cap coAM$ , and  $\Sigma_2P \cap \Pi_2P$ ), and about closure properties (e.g., it is open whether UPP is closed under intersection). Finally, we have not considered average-case models, quantum models, multi-party models, variable partition models, round-restricted models, asymmetric models, search problems, or functions with non-boolean codomains.

## A Appendix: Information Complexity Proof of [Theorem 2](#)

In this appendix we provide an alternate proof of [Theorem 2](#) using information complexity tools.

### A.1 Preliminaries

In this proof it is more convenient to consider the private-randomness version of  $ZPP^{NP[1]}$ , in which a protocol consists of a single  $P^{NP[1]}$ -type protocol over the domain  $(\{0, 1\}^n \times \{0, 1\}^q) \times (\{0, 1\}^n \times \{0, 1\}^q)$  (for some  $q$ ), and on input  $(x, y)$  the protocol is applied to  $((x, r_x), (y, r_y))$  where  $r_x, r_y \in \{0, 1\}^q$  are chosen independently uniformly at random. This model is equivalent to the public-randomness version, within a constant factor and additive  $O(\log n)$  term in the cost, by standard sparsification of randomness and the fact that the success probability can be amplified as long as it is a constant greater than  $1/2$  [\[CP08\]](#).

Throughout this appendix, we use bold letters for random variables,  $\mathbb{P}$  for probability,  $\mathbb{E}$  for expectation,  $\mathbb{H}$  for Shannon entropy,  $\mathbb{I}$  for mutual information,  $H$  for Hellinger distance, and  $\Delta$  for statistical (total variation) distance. Recall that if  $\Psi_1, \Psi_2$  are distributions over a finite set  $\Omega$ , then  $H^2(\Psi_1, \Psi_2) := 1 - \sum_{\omega \in \Omega} \sqrt{\Psi_1(\omega)\Psi_2(\omega)}$  and  $\Delta(\Psi_1, \Psi_2) := \frac{1}{2} \sum_{\omega \in \Omega} |\Psi_1(\omega) - \Psi_2(\omega)|$ . We use the following (by-now standard) lemmas [\[Lin91, BYJKS04\]](#).

---

<sup>8</sup>Not to be confused with the measure APP from [\[Kla03\]](#), which is equivalent to PostBPP.

**Lemma 28.** *Suppose  $\Psi, \Lambda$  are jointly distributed random variables and  $\Lambda$  is uniform over two outcomes, say  $\{1, 2\}$ . Then  $H^2(\Psi_1, \Psi_2) \leq \mathbb{I}(\Psi ; \Lambda)$  where  $\Psi_\Lambda := (\Psi \mid \Lambda = \Lambda)$  for  $\Lambda \in \{1, 2\}$ .*

**Lemma 29.** *If  $\Psi_1, \Psi_2$  are distributions, then  $H^2(\Psi_1, \Psi_2) \leq \Delta(\Psi_1, \Psi_2) \leq \sqrt{2}H(\Psi_1, \Psi_2)$ .*

## A.2 Proof of Theorem 2

Suppose for contradiction there is a cost- $o(n)$  private-randomness  $\text{ZPP}^{\text{NP}[1]}$  protocol  $\Pi$  computing UNIQUE-INTER with success probability  $1/2 + \epsilon$  (for any constant  $\epsilon > 0$ ). Let the notation  $k, R_v, S_{v,w}, o_v$  be with respect to  $\Pi$  (similarly to the definition of  $\text{ZPP}^{\text{NP}[1]}$  in Section 2.1), and let  $c$  be the communication cost of the deterministic part of  $\Pi$ . Consider the following jointly distributed random variables.

- Let  $\mathbf{i}$  be uniform over  $[n]$ .
- Let  $\mathbf{z} := z_1 \cdots z_n$  be distributed as follows. For each  $j \in [n]$  (independently), if  $j = \mathbf{i}$  then  $z_j$  is uniform over the two outcomes  $\{\{00\}, \{11\}\}$ , and if  $j \neq \mathbf{i}$  then  $z_j$  is uniform over the two outcomes  $\{\{00, 10\}, \{00, 01\}\}$ .
- Let  $\mathbf{x} := x_1 \cdots x_n$  and  $\mathbf{y} := y_1 \cdots y_n$  be distributed as follows. For each  $j \in [n]$  (independently),  $x_j y_j$  is uniform over the elements of the outcome of  $z_j$ .
- Let  $\mathbf{r}_x, \mathbf{r}_y$  be the private random strings, which are independent of  $\mathbf{x}, \mathbf{y}$ .
- Let  $\mathbf{v} \in \{0, 1\}^c$  be the leaf reached (i.e., the deterministic transcript) of  $\Pi((\mathbf{x}, \mathbf{r}_x), (\mathbf{y}, \mathbf{r}_y))$ .
- Let  $\chi \in \{0, 1\}$  indicate whether  $((\mathbf{x}, \mathbf{r}_x), (\mathbf{y}, \mathbf{r}_y)) \in \bigcup_w S_{v,w}$  (i.e., the NP oracle's response).
- Let  $\mathbf{w} \in \{\varepsilon\} \cup \{0, 1\}^k$  (where  $\varepsilon$  is the empty string) be distributed as follows. If  $\chi = 0$  then  $\mathbf{w} := \varepsilon$ . If  $\chi = 1$  then let  $\mathbf{w} \in \{0, 1\}^k$  be chosen arbitrarily such that  $((\mathbf{x}, \mathbf{r}_x), (\mathbf{y}, \mathbf{r}_y)) \in S_{v,w}$ .
- Let  $\mathbf{\Pi} := \mathbf{v}\mathbf{w}$ , which is distributed over  $\{0, 1\}^c \cup \{0, 1\}^{c+k}$ . Note that  $\mathbf{\Pi}$  is a deterministic function of  $((\mathbf{x}, \mathbf{r}_x), (\mathbf{y}, \mathbf{r}_y))$ .

Let  $\mathbf{x}_{-i}, \mathbf{y}_{-i}, \mathbf{z}_{-i}$  denote the restrictions of  $\mathbf{x}, \mathbf{y}, \mathbf{z}$  to coordinates in  $[n] \setminus \{\mathbf{i}\}$ . We have

$$\mathbb{I}(\mathbf{\Pi} ; \mathbf{x}_{-i}, \mathbf{y}_{-i} \mid \mathbf{i}, \mathbf{z}) \leq \mathbb{H}(\mathbf{\Pi} \mid \mathbf{i}, \mathbf{z}) \leq c + k \leq o(n).$$

By the standard direct sum property for mutual information [BYJKS04, JKS03], if  $\mathbf{j}$  is uniform over  $[n] \setminus \{\mathbf{i}\}$  (and independent of the other random variables, conditioned on  $\mathbf{i}$ ) then

$$\mathbb{I}(\mathbf{\Pi} ; \mathbf{x}_j, \mathbf{y}_j \mid \mathbf{i}, \mathbf{j}, \mathbf{z}) \leq \frac{1}{n-1} \cdot \mathbb{I}(\mathbf{\Pi} ; \mathbf{x}_{-i}, \mathbf{y}_{-i} \mid \mathbf{i}, \mathbf{z}) \leq o(1).$$

Define two more random variables (which are deterministic functions of  $(\mathbf{i}, \mathbf{j}, \mathbf{z})$ ) as follows:  $\mathbf{h} := \{\mathbf{i}, \mathbf{j}\}$  and

$$\mathbf{g} := \begin{cases} \text{HEADS} & \text{if } \mathbf{i} < \mathbf{j} \text{ and } \mathbf{z}_i = \{11\}, \text{ or if } \mathbf{i} > \mathbf{j} \text{ and } \mathbf{z}_i = \{00\} \\ \text{TAILS} & \text{if } \mathbf{i} < \mathbf{j} \text{ and } \mathbf{z}_i = \{00\}, \text{ or if } \mathbf{i} > \mathbf{j} \text{ and } \mathbf{z}_i = \{11\} \end{cases}.$$

Let  $\mathbf{x}_{-h}, \mathbf{y}_{-h}, \mathbf{z}_{-h}$  denote the restrictions of  $\mathbf{x}, \mathbf{y}, \mathbf{z}$  to coordinates in  $[n] \setminus \mathbf{h}$ . Since we have  $\mathbb{I}(\mathbf{\Pi} ; \mathbf{x}_j, \mathbf{y}_j \mid \mathbf{g}, \mathbf{h}, \mathbf{i}, \mathbf{j}, \mathbf{z}) \leq o(1)$ , there must exist outcomes  $g^* \in \{\text{HEADS}, \text{TAILS}\}$ ,  $h^* \in \binom{[n]}{2}$ , and  $z_{-h}^*$  such that if we let  $E$  denote the event  $(\mathbf{g} = g^*, \mathbf{h} = h^*, \mathbf{z}_{-h} = z_{-h}^*)$  then

$$\mathbb{I}(\mathbf{\Pi} ; \mathbf{x}_j, \mathbf{y}_j \mid E, \mathbf{i}, \mathbf{j}, \mathbf{z}_h) \leq o(1). \tag{1}$$

		$\mathbf{y}_1\mathbf{y}_2$			
		00	01	10	11
$\mathbf{x}_1\mathbf{x}_2$	00	A		C	
	01				
	10	B		D	F
	11			E	

**Figure 2:** The shaded rectangles are all 1-inputs, and the unshaded rectangles are all 0-inputs.

Assume  $g^* = \text{HEADS}$  and  $h^* = \{1, 2\}$  (the other cases are analogous). As illustrated in **Figure 2**, define  $A := (00, 00)$ ,  $B := (10, 00)$ ,  $C := (00, 10)$ ,  $D := (10, 10)$ ,  $E := (11, 10)$ ,  $F := (10, 11)$ . Conditioned on  $E$ , there are the following four equally likely outcomes of  $(\mathbf{i}, \mathbf{j}, \mathbf{z}_h)$ .

- $(\mathbf{i} = 2, \mathbf{j} = 1, \mathbf{z}_i = \{00\}, \mathbf{z}_j = \{00, 10\})$  so  $(\mathbf{x}_1\mathbf{x}_2, \mathbf{y}_1\mathbf{y}_2)$  is uniform over  $\{A, B\}$ .
- $(\mathbf{i} = 2, \mathbf{j} = 1, \mathbf{z}_i = \{00\}, \mathbf{z}_j = \{00, 01\})$  so  $(\mathbf{x}_1\mathbf{x}_2, \mathbf{y}_1\mathbf{y}_2)$  is uniform over  $\{A, C\}$ .
- $(\mathbf{i} = 1, \mathbf{j} = 2, \mathbf{z}_i = \{11\}, \mathbf{z}_j = \{00, 10\})$  so  $(\mathbf{x}_1\mathbf{x}_2, \mathbf{y}_1\mathbf{y}_2)$  is uniform over  $\{D, E\}$ .
- $(\mathbf{i} = 1, \mathbf{j} = 2, \mathbf{z}_i = \{11\}, \mathbf{z}_j = \{00, 01\})$  so  $(\mathbf{x}_1\mathbf{x}_2, \mathbf{y}_1\mathbf{y}_2)$  is uniform over  $\{D, F\}$ .

Letting  $e \in \{AB, AC, DE, DF\}$  be the random variable indicating which of these cases holds, (1) says that  $\mathbb{I}(\Pi ; \mathbf{x}_j, \mathbf{y}_j \mid E, e) \leq o(1)$ . Thus for each outcome  $e$  we have

$$\mathbb{I}(\Pi ; \mathbf{x}_j, \mathbf{y}_j \mid E, e = e) \leq o(1).$$

Conditioned on  $(E, e = e)$ ,  $(\mathbf{x}_j, \mathbf{y}_j)$  is uniform over two outcomes, so we can apply **Lemma 28** with  $\Psi := (\Pi \mid E, e = e)$  and  $\Lambda := (\mathbf{x}_j, \mathbf{y}_j \mid E, e = e)$ .

Hence, if for  $s \in \{A, B, C, D, E, F\}$  we let  $E_s$  denote the event  $((\mathbf{x}_1\mathbf{x}_2, \mathbf{y}_1\mathbf{y}_2) = s, \mathbf{z}_{-h^*} = \mathbf{z}_{-h^}^*)$  and we define the distribution  $\Pi_s := (\Pi \mid E_s)$ , then (noting that  $\Pi_A$  is distributed identically to  $(\Pi \mid E, e = AB, \mathbf{x}_1\mathbf{y}_1 = 00)$  and similarly for the other possibilities) we have  $H(\Pi_A, \Pi_B), H(\Pi_A, \Pi_C), H(\Pi_D, \Pi_E), H(\Pi_D, \Pi_F) \leq o(1)$ . Assuming that “close” means “within Hellinger distance  $o(1)$ ” or equivalently “within statistical distance  $o(1)$ ” (by **Lemma 29**), by the triangle inequality,  $\Pi_A, \Pi_B, \Pi_C$  are all close and  $\Pi_D, \Pi_E, \Pi_F$  are all close. In particular, the same holds for the distributions  $\mathbf{v}_s := (\mathbf{v} \mid E_s)$  (equivalently,  $\mathbf{v}_s$  is the marginal of the first  $c$  bits of  $\Pi_s$ ):  $\mathbf{v}_A, \mathbf{v}_B, \mathbf{v}_C$  are all close and  $\mathbf{v}_D, \mathbf{v}_E, \mathbf{v}_F$  are all close.

Note that  $(\mathbf{x}, \mathbf{y} \mid E_s)$  is uniform over a rectangle consisting only of 0-inputs if  $s \in \{A, B, C\}$  and only of 1-inputs if  $s \in \{D, E, F\}$ . Since for every leaf  $v$ , the event  $\mathbf{v} = v$  consists of a rectangle in the domain of  $((\mathbf{x}, \mathbf{r}_x), (\mathbf{y}, \mathbf{r}_y))$ , we have  $\mathbb{P}[\mathbf{v}_A = v] \cdot \mathbb{P}[\mathbf{v}_D = v] = \mathbb{P}[\mathbf{v}_B = v] \cdot \mathbb{P}[\mathbf{v}_C = v]$ . This implies that  $H(\mathbf{v}_A, \mathbf{v}_D) = H(\mathbf{v}_B, \mathbf{v}_C) \leq o(1)$ , and hence  $\mathbf{v}_A, \mathbf{v}_B, \mathbf{v}_C, \mathbf{v}_D, \mathbf{v}_E, \mathbf{v}_F$  are all close.

Let  $E_{s,v}$  denote the intersection of the event  $\mathbf{v} = v$  with  $E_s$ . Let  $\mathbf{w}_{s,v}$  denote the distribution  $(\mathbf{w} \mid E_{s,v})$  assuming  $\mathbb{P}[E_{s,v}] > 0$  (equivalently, assuming  $\mathbb{P}[\mathbf{v}_s = v] > 0$ ).

**Claim 30.** *There exists a leaf  $v^*$  such that the following all hold.*



- $\mathbb{P}[E_{s,v^*}] > 0$  for all  $s \in \{A, B, C, D, E, F\}$ .
- $\mathbb{P}[o_{v^*}(\boldsymbol{\chi}) = \perp \mid E_{D,v^*}] < 1$ .
- $\mathbb{P}[o_{v^*}(\boldsymbol{\chi}) = \perp \mid E_{A,v^*}] < 1$ .
- $\Delta(\mathbf{w}_{B,v^*}, \mathbf{w}_{C,v^*}) < 1$ .
- $\Delta(\mathbf{w}_{E,v^*}, \mathbf{w}_{F,v^*}) < 1$ .

*Proof.* Since  $\mathbb{E}_{v \sim \mathbf{v}_D} \mathbb{P}[o_v(\boldsymbol{\chi}) = \perp \mid E_{D,v}] = \mathbb{P}[o_v(\boldsymbol{\chi}) = \perp \mid E_D] \leq 1/2 - \epsilon$ , we have

$$\mathbb{P}_{v \sim \mathbf{v}_D} \left[ \mathbb{P}[o_v(\boldsymbol{\chi}) = \perp \mid E_{D,v}] < 1 \text{ and } \mathbb{P}[E_{D,v}] > 0 \right] \geq 1/2 + \epsilon \quad (2)$$

by Markov's inequality. Similarly, (2) holds with A in place of D, and thus

$$\mathbb{P}_{v \sim \mathbf{v}_D} \left[ \mathbb{P}[o_v(\boldsymbol{\chi}) = \perp \mid E_{A,v}] < 1 \text{ and } \mathbb{P}[E_{A,v}] > 0 \right] \geq 1/2 + \epsilon - o(1) \quad (3)$$

since  $\Delta(\mathbf{v}_A, \mathbf{v}_D) \leq o(1)$ . Next, we show that

$$\mathbb{P}_{v \sim \mathbf{v}_D} \left[ \Delta(\mathbf{w}_{B,v}, \mathbf{w}_{C,v}) < 1 \text{ and } \mathbb{P}[E_{B,v}] > 0 \text{ and } \mathbb{P}[E_{C,v}] > 0 \right] \geq 1 - o(1) \quad (4)$$

holds. Similarly,

$$\mathbb{P}_{v \sim \mathbf{v}_D} \left[ \Delta(\mathbf{w}_{E,v}, \mathbf{w}_{F,v}) < 1 \text{ and } \mathbb{P}[E_{E,v}] > 0 \text{ and } \mathbb{P}[E_{F,v}] > 0 \right] \geq 1 - o(1) \quad (5)$$

will hold. The claim then follows from (2), (3), (4), and (5) by a union bound over  $v \sim \mathbf{v}_D$ . It remains to show (4). Let

$$V := \left\{ v : \Delta(\mathbf{w}_{B,v}, \mathbf{w}_{C,v}) = 1 \text{ and } \mathbb{P}[E_{B,v}] > 0 \text{ and } \mathbb{P}[E_{C,v}] > 0 \right\}$$

and let  $T := \{vw : v \in V \text{ and } w \in \text{supp}(\mathbf{w}_{B,v})\}$ . Note that  $\mathbb{P}[\boldsymbol{\Pi}_C \in T] = 0$  since  $\text{supp}(\mathbf{w}_{B,v}) \cap \text{supp}(\mathbf{w}_{C,v}) = \emptyset$  for each  $v \in V$ . Thus  $\mathbb{P}[\mathbf{v}_B \in V] = \mathbb{P}[\boldsymbol{\Pi}_B \in T] \leq 0 + \Delta(\boldsymbol{\Pi}_B, \boldsymbol{\Pi}_C) \leq o(1)$ . It follows that  $\mathbb{P}[\mathbf{v}_D \in V] \leq o(1) + \Delta(\mathbf{v}_D, \mathbf{v}_B) \leq o(1)$ . We also have  $\mathbb{P}_{v \sim \mathbf{v}_D} [\mathbb{P}[E_{B,v}] = 0] \leq \Delta(\mathbf{v}_B, \mathbf{v}_D) \leq o(1)$  and  $\mathbb{P}_{v \sim \mathbf{v}_D} [\mathbb{P}[E_{C,v}] = 0] \leq \Delta(\mathbf{v}_C, \mathbf{v}_D) \leq o(1)$ . Hence the left side of (4) is at least

$$1 - \mathbb{P}_{v \sim \mathbf{v}_D} [v \in V] - \mathbb{P}_{v \sim \mathbf{v}_D} [\mathbb{P}[E_{B,v}] = 0] - \mathbb{P}_{v \sim \mathbf{v}_D} [\mathbb{P}[E_{C,v}] = 0] \geq 1 - o(1). \quad \square$$

By the correctness of  $\Pi$ , we have  $\mathbb{P}[o_{v^*}(\boldsymbol{\chi}) = 0 \mid E_{A,v^*}] > 0$  and  $\mathbb{P}[o_{v^*}(\boldsymbol{\chi}) = 1 \mid E_{D,v^*}] > 0$ . Thus 0 and 1 are both possible outputs of  $o_{v^*}$ , and hence  $\perp$  is not a possible output of  $o_{v^*}$ . In what follows, note that  $E_s$  can be viewed as a subset of the domain of  $((\mathbf{x}, \mathbf{r}_x), (\mathbf{y}, \mathbf{r}_y))$ .

First suppose  $o_{v^*}(1) = 0$  and  $o_{v^*}(0) = 1$ . For all  $s \in \{A, B, C\}$ , we actually have  $\mathbb{P}[o_{v^*}(\boldsymbol{\chi}) = 0 \mid E_{s,v^*}] = 1$  and hence  $\mathbb{P}[\boldsymbol{\chi} = 1 \mid E_{s,v^*}] = 1$  and hence  $\mathbb{P}[\mathbf{w}_{s,v^*} \neq \varepsilon] = 1$ . Since  $\Delta(\mathbf{w}_{B,v^*}, \mathbf{w}_{C,v^*}) < 1$ , this implies that there exists a  $w^* \in \{0, 1\}^k$  such that  $\mathbb{P}[\mathbf{w}_{B,v^*} = w^*] > 0$  and  $\mathbb{P}[\mathbf{w}_{C,v^*} = w^*] > 0$ . Hence there exists a  $((x, r_x), (y, r_y)) \in S_{v^*, w^*} \cap E_B$  and a  $((x', r'_x), (y', r'_y)) \in S_{v^*, w^*} \cap E_C$ . Since  $S_{v^*, w^*}$  is a rectangle,  $((x, r_x), (y', r'_y)) \in S_{v^*, w^*}$  and hence  $\Pi$  outputs  $o_{v^*}(1) = 0$ . This contradicts the correctness since  $(x, y')$  is a 1-input (having  $x \cap y' = \{1\}$  and lying in the D cell).

On the other hand, suppose  $o_{v^*}(1) = 1$  and  $o_{v^*}(0) = 0$ . The argument is very similar: For all  $s \in \{D, E, F\}$ , we actually have  $\mathbb{P}[o_{v^*}(\boldsymbol{\chi}) = 1 \mid E_{s,v^*}] = 1$  and hence  $\mathbb{P}[\boldsymbol{\chi} = 1 \mid E_{s,v^*}] = 1$  and hence  $\mathbb{P}[\mathbf{w}_{s,v^*} \neq \varepsilon] = 1$ . Since  $\Delta(\mathbf{w}_{E,v^*}, \mathbf{w}_{F,v^*}) < 1$ , this implies that there exists a  $w^* \in \{0, 1\}^k$  such that  $\mathbb{P}[\mathbf{w}_{E,v^*} = w^*] > 0$  and  $\mathbb{P}[\mathbf{w}_{F,v^*} = w^*] > 0$ . Hence there exists a  $((x, r_x), (y, r_y)) \in S_{v^*, w^*} \cap E_E$  and a  $((x', r'_x), (y', r'_y)) \in S_{v^*, w^*} \cap E_F$ . Since  $S_{v^*, w^*}$  is a rectangle,  $((x, r_x), (y', r'_y)) \in S_{v^*, w^*}$  and hence  $\Pi$  outputs  $o_{v^*}(1) = 1$ . This contradicts the correctness since  $(x, y')$  is a 0-input (having  $x \cap y' = \{1, 2\}$  and lying in the bottom-right cell in [Figure 2](#)).

## B Appendix: Catalog of Communication Complexity Classes

We now provide formal definitions of all the communication complexity classes considered in [Section 3](#). If  $\mathcal{C}$  is the name of a model and  $F: \{0, 1\}^n \times \{0, 1\}^n \rightarrow \{0, 1\}$  is a partial function, then we let  $\mathcal{C}(F)$  denote the minimum cost of a correct protocol for  $F$  in model  $\mathcal{C}$ , and we also let  $\mathcal{C}$  denote the class of all (families of) partial functions  $F$  with  $\mathcal{C}(F) \leq \text{poly}(\log n)$ . We let  $\text{co}\mathcal{C}(F) := \mathcal{C}(\neg F)$ .

For example,  $\text{P}(F)$  is the minimum cost of a deterministic protocol for  $F$ , and  $\text{P}$  is the set of partial functions with  $\text{poly}(\log n)$ -cost deterministic protocols. We group the remaining models into four categories (corresponding to the four subsections): the NP query hierarchy, bounded-error randomized models, models with postselection or unbounded error, and models with alternation.

In the definitions that follow, we always use  $\Pi$  to denote a protocol,  $F$  to denote an arbitrary partial function, and  $(x, y)$  to denote an arbitrary input in the domain of  $F$  (the models are worst-case, so the correctness criteria always hold for all such  $(x, y)$ ). All randomized models are assumed to have public randomness except when noted otherwise.

### B.1 The NP query hierarchy

#### Definition 31. (NP)

*Syntax:*  $\Pi$  is a collection of rectangles  $\{R_w : w \in \{0, 1\}^k\}$ , and  $\Pi$  outputs 1 or 0 indicating whether  $(x, y) \in \bigcup_w R_w$ .

*Correctness:*  $\Pi(x, y) = F(x, y)$ .

*Cost:*  $k$ .

#### Definition 32. (US)

*Syntax:*  $\Pi$  is a collection of rectangles  $\{R_w : w \in \{0, 1\}^k\}$ , and  $\Pi$  outputs 1 or 0 indicating whether the number of  $w$ 's such that  $(x, y) \in R_w$  is exactly one.

*Correctness:*  $\Pi(x, y) = F(x, y)$ .

*Cost:*  $k$ .

#### Definition 33. (DP)

*Syntax:*  $\Pi$  is a pair of collections of rectangles,  $\{S_w : w \in \{0, 1\}^k\}$  and  $\{T_w : w \in \{0, 1\}^k\}$ , and  $\Pi$  outputs 1 or 0 indicating whether  $(x, y) \in \bigcup_w S_w \setminus \bigcup_w T_w$ .

*Correctness:*  $\Pi(x, y) = F(x, y)$ .

*Cost:*  $k$ .

#### Definition 34. ( $\text{P}_{\parallel}^{\text{NP}[q]}$ for constant $q$ )

*Syntax:*  $\Pi$  is a deterministic protocol where for each leaf  $v$  with associated rectangle  $R_v$ , there are  $q$  associated collections of subrectangles  $\{S_{v,i,w} \subseteq R_v : w \in \{0, 1\}^k\}$  ( $i \in [q]$ ) and an associated output function  $o_v: \{0, 1\}^q \rightarrow \{0, 1\}$  that is applied to the indicators of whether  $(x, y) \in \bigcup_w S_{v,i,w}$  for each  $i$ .

*Correctness:*  $\Pi(x, y) = F(x, y)$ .

*Cost:*  $k +$  the communication cost of the deterministic part.

**Definition 35.** ( $\mathsf{P}^{\mathsf{NP}}$ )

*Syntax:*  $\Pi$  is a protocol tree where each internal node  $v$  is labeled with either (i) a 1-bit function of Alice's or of Bob's input in the usual way, or (ii) an "NP query" consisting of a collection of rectangles  $\{S_{v,w} : w \in \{0,1\}^{k_v}\}$ , where the indicator of whether  $(x,y) \in \bigcup_w S_{v,w}$  determines which child to descend to in the protocol tree. The output of  $\Pi$  is determined by the leaf reached.

*Correctness:*  $\Pi(x,y) = F(x,y)$ .

*Cost:* The maximum over all root-to-leaf paths of the following: the length of the path plus the sum of  $k_v$  over all type-(ii) nodes  $v$  on the path.

**Definition 36.** ( $\mathsf{P}^{\mathsf{NP}[q]}$  for constant  $q$ )

*Syntax:*  $\Pi$  is a  $\mathsf{P}^{\mathsf{NP}}$ -type protocol where there are at most  $q$  NP queries on each root-to-leaf path.

*Correctness:*  $\Pi(x,y) = F(x,y)$ .

*Cost:* Same as [Definition 35](#). Affecting the cost only by a constant factor, it can be assumed that all NP queries happen at the end and all have the same witness length  $k_v$ .

**Definition 37.** ( $\mathsf{P}_{\parallel}^{\mathsf{NP}}$ )

*Syntax:*  $\Pi$  is a  $\mathsf{P}^{\mathsf{NP}}$ -type protocol where the result of each NP query is not revealed until the last query on any path down the tree. Thus, each type-(ii) node has 1 child if it has a type-(ii) descendant, and has  $2^q$  children if it has no type-(ii) descendants (where  $q$  is the number of type-(ii) nodes on that path). Hence without loss of generality, all the NP queries are consecutive.

*Correctness:*  $\Pi(x,y) = F(x,y)$ .

*Cost:* Same as [Definition 35](#).

## B.2 Bounded-error randomized models

**Definition 38.** (ZPP)

*Syntax:*  $\Pi$  is a distribution over deterministic protocols outputting values in  $\{0,1,\perp\}$ .

*Correctness:*  $\mathbb{P}[\Pi(x,y) \in \{F(x,y), \perp\}] = 1$  and  $\mathbb{P}[\Pi(x,y) = F(x,y)] \geq 3/4$ .

*Cost:* The maximum communication cost of any constituent deterministic protocol.

**Definition 39.** (RP)

*Syntax:*  $\Pi$  is a distribution over deterministic protocols outputting values in  $\{0,1\}$ .

*Correctness:* If  $F(x,y) = 1$  then  $\mathbb{P}[\Pi(x,y) = 1] \geq 1/2$ .

If  $F(x,y) = 0$  then  $\mathbb{P}[\Pi(x,y) = 0] = 1$ .

*Cost:* The maximum communication cost of any constituent deterministic protocol.

**Definition 40.** (BPP)

*Syntax:*  $\Pi$  is a distribution over deterministic protocols outputting values in  $\{0,1\}$ .

*Correctness:*  $\mathbb{P}[\Pi(x,y) = F(x,y)] \geq 3/4$ .

*Cost:* The maximum communication cost of any constituent deterministic protocol.

**Definition 41.** (MA)

*Syntax:*  $\Pi$  is a distribution over deterministic protocols that take an additional input  $w \in \{0, 1\}^k$ , which is visible to both Alice and Bob.

*Correctness: Completeness:* if  $F(x, y) = 1$  then  $\exists w : \mathbb{P}[\Pi(x, y, w) = 1] \geq 3/4$ .

*Soundness:* if  $F(x, y) = 0$  then  $\forall w : \mathbb{P}[\Pi(x, y, w) = 0] \geq 3/4$ .

*Cost:*  $k$  + the maximum communication cost of any constituent deterministic protocol.

**Definition 42.** (AM)

*Syntax:*  $\Pi$  is a distribution over nondeterministic (NP-type) protocols.

*Correctness:*  $\mathbb{P}[\Pi(x, y) = F(x, y)] \geq 3/4$ .

*Cost:* The maximum cost of any constituent nondeterministic protocol.

**Definition 43.** ( $\text{ZPP}_{\parallel}^{\text{NP}[q]}$  for constant  $q$ )

*Syntax:*  $\Pi$  is a distribution over  $\text{P}_{\parallel}^{\text{NP}[q]}$ -type protocols outputting values in  $\{0, 1, \perp\}$ .

*Correctness:* Same as [Definition 38](#).

*Cost:* The maximum cost of any constituent  $\text{P}_{\parallel}^{\text{NP}[q]}$ -type protocol.

**Definition 44.** ( $\text{ZPP}_{\parallel}^{\text{NP}}$ )

*Syntax:*  $\Pi$  is a distribution over  $\text{P}_{\parallel}^{\text{NP}}$ -type protocols outputting values in  $\{0, 1, \perp\}$ .

*Correctness:* Same as [Definition 38](#).

*Cost:* The maximum cost of any constituent  $\text{P}_{\parallel}^{\text{NP}}$ -type protocol.

**Definition 45.** ( $\text{ZPP}^{\text{NP}[q]}$  for constant  $q$ )

*Syntax:*  $\Pi$  is a distribution over  $\text{P}^{\text{NP}[q]}$ -type protocols outputting values in  $\{0, 1, \perp\}$ .

*Correctness:* Same as [Definition 38](#).

*Cost:* The maximum cost of any constituent  $\text{P}^{\text{NP}[q]}$ -type protocol.

**Definition 46.** ( $\text{ZPP}^{\text{NP}}$ )

*Syntax:*  $\Pi$  is a distribution over  $\text{P}^{\text{NP}}$ -type protocols outputting values in  $\{0, 1, \perp\}$ .

*Correctness:* Same as [Definition 38](#).

*Cost:* The maximum cost of any constituent  $\text{P}^{\text{NP}}$ -type protocol.

**B.3 Models with postselection or unbounded error**

Although SBP is not defined in terms of postselection or unbounded error, we include the definition here since it provides a nice segue.

**Definition 47.** (SBP)

*Syntax:*  $\Pi$  has public randomness uniformly distributed over  $\{0, 1\}^k$ , with each outcome having an associated deterministic protocol outputting values in  $\{0, 1\}$ .

*Correctness:*  $\min_{(x,y) \in F^{-1}(1)} \mathbb{P}[\Pi(x, y) = 1] > 2 \cdot \max_{(x,y) \in F^{-1}(0)} \mathbb{P}[\Pi(x, y) = 1]$ .

*Cost:*  $k$  + the maximum communication cost of any constituent deterministic protocol.

**Definition 48.** (PostBPP)

*Syntax:*  $\Pi$  has public randomness uniformly distributed over  $\{0, 1\}^k$ , with each outcome having an associated deterministic protocol outputting values in  $\{0, 1, \perp\}$ .

*Correctness:*  $\mathbb{P}[\Pi(x, y) = F(x, y)] > 2 \cdot \mathbb{P}[\Pi(x, y) = 1 - F(x, y)]$ .

*Cost:*  $k$  + the maximum communication cost of any constituent deterministic protocol.

**Definition 49.** (UPostBPP $_{\square}$ )

*Syntax:* Same as [Definition 48](#), except the public randomness is arbitrarily distributed over  $\{0, 1\}^k$ .

*Correctness:* Same as [Definition 48](#).

*Cost:* Same as [Definition 48](#).

**Definition 50.** (UPostBPP)

*Syntax:*  $\Pi$  is a private-randomness protocol outputting values in  $\{0, 1, \perp\}$ .

*Correctness:* Same as [Definition 48](#).

*Cost:* The communication cost of the underlying deterministic protocol.

We have  $\text{UPostBPP}(F) \leq \text{UPostBPP}_{\square}(F) \leq \text{PostBPP}(F)$  for all  $F$ , and hence  $\text{PostBPP} \subseteq \text{UPostBPP}_{\square} \subseteq \text{UPostBPP}$ .

**Observation 51.** *Without loss of generality, in a PostBPP or UPostBPP $_{\square}$  protocol, each of the constituent deterministic protocols consists of a single rectangle (with fixed output 0 or 1 on inputs in the rectangle, and output  $\perp$  on inputs outside the rectangle).*

*Proof.* We may modify a PostBPP or UPostBPP $_{\square}$  protocol so that after choosing the original public randomness, it then picks a uniformly random leaf rectangle (of which we assume there are exactly  $2^c$ ) from the associated deterministic protocol, outputs the same value on inputs in the rectangle, and outputs  $\perp$  on all inputs outside the rectangle. The correctness is unaffected. The number of random bits becomes  $k + c$ , and the communication cost becomes 2, so the overall cost becomes  $k + c + 2$ . If after the transformation, any rectangle has label  $\perp$ , we can instead assume it is an empty rectangle with non- $\perp$  label.  $\square$

**Observation 52.**  $\text{UPostBPP}(F) \in \min(\log \text{rank}_+(M^0) + \log \text{rank}_+(M^1)) \pm O(1)$  where  $\text{rank}_+$  denotes nonnegative rank, and the minimum is over nonnegative real matrices  $M^0, M^1$  (indexed by inputs) such that for each  $(x, y) \in F^{-1}$ ,  $M_{x,y}^{F(x,y)} > 2 \cdot M_{x,y}^{1-F(x,y)}$ .

**Definition 53.** (PP)

*Syntax:*  $\Pi$  has public randomness uniformly distributed over  $\{0, 1\}^k$ , with each outcome having an associated deterministic protocol outputting values in  $\{0, 1\}$ .

*Correctness:*  $\mathbb{P}[\Pi(x, y) = F(x, y)] > 1/2$ .

*Cost:*  $k$  + the maximum communication cost of any constituent deterministic protocol.

**Definition 54.** (UPP $_{\square}$ )

*Syntax:* Same as [Definition 53](#), except the public randomness is arbitrarily distributed over  $\{0, 1\}^k$ .

*Correctness:* Same as [Definition 53](#).

*Cost:* Same as [Definition 53](#).

**Definition 55.** (UPP)

*Syntax:*  $\Pi$  is a private-randomness protocol outputting values in  $\{0, 1\}$ .

*Correctness:* Same as [Definition 53](#).

*Cost:* The communication cost of the underlying deterministic protocol.

We have  $\text{UPP}(F) \leq \text{UPP}_{\square}(F) \leq \text{PP}(F)$  for all  $F$ , and hence  $\text{PP} \subseteq \text{UPP}_{\square} \subseteq \text{UPP}$ .

**Observation 56.** *Without loss of generality, in a PP or  $\text{UPP}_{\square}$  protocol, each of the constituent deterministic protocols consists of a single rectangle (with output only depending on whether the input is in the rectangle).*

*Proof.* We may modify a PP or  $\text{UPP}_{\square}$  protocol so that after choosing the original public randomness, it then picks a uniformly random leaf rectangle (of which we assume there are exactly  $2^c$ ) from the associated deterministic protocol, outputs the same value on inputs in the rectangle, and flips a coin to determine the output on all inputs outside the rectangle. The correctness is unaffected. The number of random bits becomes  $k + c + 1$ , and the communication cost becomes 2, so the overall cost becomes  $k + c + 3$ .  $\square$

**Observation 57.**  $\text{UPP}(F) \in \min(\log \text{rank}_+(M^0) + \log \text{rank}_+(M^1)) \pm O(1)$  where  $\text{rank}_+$  denotes nonnegative rank, and the minimum is over nonnegative real matrices  $M^0, M^1$  (indexed by inputs) such that for each  $(x, y) \in F^{-1}$ ,  $M_{x,y}^{F(x,y)} > M_{x,y}^{1-F(x,y)}$ .

**B.4 Models with alternation****Definition 58.** ( $\text{S}_2\text{P}$ )

*Syntax:*  $\Pi$  is a matrix with rows indexed by  $w^0 \in \{0, 1\}^k$  and columns indexed by  $w^1 \in \{0, 1\}^k$ , with each entry  $(w^0, w^1)$  having an associated deterministic protocol  $\Pi_{w^0, w^1}$  outputting values in  $\{0, 1\}$ .

*Correctness:* If  $F(x, y) = 1$  then  $\exists w^1 \forall w^0 : \Pi_{w^0, w^1}(x, y) = 1$ .

If  $F(x, y) = 0$  then  $\exists w^0 \forall w^1 : \Pi_{w^0, w^1}(x, y) = 0$ .

*Cost:*  $k$  + the maximum communication cost of any constituent deterministic protocol.

**Definition 59.** ( $\Sigma_{\ell}\text{P}$  for constant  $\ell$ )

*Syntax:*  $\Pi$  is a complete  $2^k$ -ary tree of depth  $\ell$  (root-to-leaf paths have  $\ell$  edges) representing a formula with alternating layers of OR and AND gates, with an OR gate at the root, and where each leaf is the indicator for a rectangle (if  $\ell$  is odd) or the complement of a rectangle (if  $\ell$  is even).

*Correctness:*  $\Pi(x, y) = F(x, y)$ .

*Cost:*  $k$ .

**Definition 60.** ( $\Pi_{\ell}\text{P}$  for constant  $\ell$ )

*Syntax:*  $\Pi$  is a complete  $2^k$ -ary tree of depth  $\ell$  (root-to-leaf paths have  $\ell$  edges) representing a formula with alternating layers of AND and OR gates, with an AND gate at the root, and where each leaf is the indicator for a rectangle (if  $\ell$  is even) or the complement of a rectangle (if  $\ell$  is odd).

*Correctness:*  $\Pi(x, y) = F(x, y)$ .

*Cost:*  $k$ .

The class PH is defined as  $\bigcup_{\ell} \Sigma_{\ell}P = \bigcup_{\ell} \Pi_{\ell}P$  (where the union is over constants  $\ell$ ).

**Definition 61.** (PSPACE)

*Syntax:*  $\Pi$  is a formula where each leaf is the indicator for a rectangle.

*Correctness:*  $\Pi(x, y) = F(x, y)$ .

*Cost:* The log of the size of the formula.

Although  $\oplus P$  is not defined in terms of alternation, we include the definition in this subsection since in a sense, it is at least as powerful as alternation:  $\text{PH} \subseteq \text{BP} \cdot \oplus P$  [Tod91].

**Definition 62.** ( $\oplus P$ )

*Syntax:*  $\Pi$  is a collection of rectangles  $\{R_w : w \in \{0, 1\}^k\}$ , and  $\Pi$  outputs 1 or 0 indicating whether the number of  $w$ 's such that  $(x, y) \in R_w$  is odd.

*Correctness:*  $\Pi(x, y) = F(x, y)$ .

*Cost:*  $k$ .

**Observation 63.**  $\oplus P(F) \in \log \text{rank}(F) \pm O(1)$  where the rank is over  $GF(2)$ .

## References

- [Aar05] 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.
- [AW09] 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.
- [Bei91] Richard Beigel. Bounded queries to SAT and the boolean hierarchy. *Theoretical Computer Science*, 84(2):199–223, 1991. doi:10.1016/0304-3975(91)90160-4.
- [BFS86] 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.
- [BG82] 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.
- [BVdW07] 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.
- [BYJKS04] Ziv Bar-Yossef, T.S. Jayram, Ravi Kumar, and D. Sivakumar. An information statistics approach to data stream and communication complexity. *Journal of Computer and System Sciences*, 68(4):702–732, 2004. doi:10.1016/j.jcss.2003.11.006.
- [Cai07] Jin-Yi Cai.  $S_2P \subseteq ZPP^{NP}$ . *Journal of Computer and System Sciences*, 73(1):25–35, 2007. doi:10.1016/j.jcss.2003.07.015.

- [Can96] Ran Canetti. More on BPP and the polynomial-time hierarchy. *Information Processing Letters*, 57(5):237–241, 1996. doi:10.1016/0020-0190(96)00016-6.
- [CC06] Jin-Yi Cai and Venkatesan Chakaravathy. 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.
- [CCGT14] 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.
- [CCM<sup>+</sup>15] 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.
- [CCMT14] 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.
- [CKR95] 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.
- [CP08] Richard Chang and Suresh Purini. Amplifying  $ZPP^{\text{SAT}[1]}$  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.
- [DKMW04] Carsten Damm, Matthias Krause, Christoph Meinel, and Stephan Waack. On relations between counting communication complexity classes. *Journal of Computer and System Sciences*, 69(2):259–280, 2004. doi:10.1016/j.jcss.2004.03.002.
- [FIKU08] Lance Fortnow, Russell Impagliazzo, Valentine Kabanets, and Christopher Umans. On the complexity of succinct zero-sum games. *Computational Complexity*, 17(3):353–376, 2008. doi:10.1007/s00037-008-0252-2.
- [FJK<sup>+</sup>15] 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.
- [For02] 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.
- [GKR15] 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.



- [GL14] 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.
- [GLM<sup>+</sup>15] 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.
- [GPW15] Mika Göös, Toniann Pitassi, and Thomas Watson. Zero-information protocols and unambiguity in Arthur–Merlin communication. In *Proceedings of the 6th Innovations in Theoretical Computer Science Conference (ITCS)*, pages 113–122. ACM, 2015. doi:10.1145/2688073.2688074.
- [GR13] 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.
- [GR15] 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.
- [GS86] Shafi Goldwasser and Michael Sipser. Private coins versus public coins in interactive proof systems. In *Proceedings of the 18th Symposium on Theory of Computing (STOC)*, pages 59–68. ACM, 1986. doi:10.1145/12130.12137.
- [GS10] 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.
- [GW14] 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.
- [GZ11] Oded Goldreich and David Zuckerman. Another proof that  $BPP \subseteq PH$  (and more). In *Studies in Complexity and Cryptography*, pages 40–53. Springer, 2011. doi:10.1007/978-3-642-22670-0\_6.
- [HHT97] 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.
- [HR90] 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)90027-1.
- [IW10] 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.

- [JKS03] 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.
- [Juk06] Stasys Jukna. On graph complexity. *Combinatorics, Probability, & Computing*, 15(6):855–876, 2006. doi:10.1017/S0963548306007620.
- [Juk12] Stasys Jukna. *Boolean Function Complexity: Advances and Frontiers*, volume 27 of *Algorithms and Combinatorics*. Springer, 2012.
- [Kla03] 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.
- [Kla07] Hartmut Klauck. Lower bounds for quantum communication complexity. *SIAM Journal on Computing*, 37(1):20–46, 2007. doi:10.1137/S0097539702405620.
- [Kla10] Hartmut Klauck. A strong direct product theorem for disjointness. In *Proceedings of the 42nd Symposium on Theory of Computing (STOC)*, pages 77–86. ACM, 2010. doi:10.1145/1806689.1806702.
- [Kla11] 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.
- [KN97] Eyal Kushilevitz and Noam Nisan. *Communication Complexity*. Cambridge University Press, 1997.
- [KNSW94] Mauricio Karchmer, Ilan Newman, Michael Saks, and Avi Wigderson. Non-deterministic communication complexity with few witnesses. *Journal of Computer and System Sciences*, 49(2):247–257, 1994. doi:10.1016/S0022-0000(05)80049-2.
- [KP13] 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.
- [KP14] 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.
- [KW15] 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.
- [Lin91] Jianhua Lin. Divergence measures based on the Shannon entropy. *IEEE Transactions on Information Theory*, 37(1):145–151, 1991. doi:10.1109/18.61115.
- [Lok01] 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.

- [Lok09] 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.
- [LR92] 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.
- [LS09] Nathan Linial and Adi Shraibman. Learning complexity vs communication complexity. *Combinatorics, Probability, & Computing*, 18(1–2):227–245, 2009. doi:10.1017/S0963548308009656.
- [New91] 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.
- [NW94] Noam Nisan and Avi Wigderson. Hardness vs. randomness. *Journal of Computer and System Sciences*, 49(2):149–167, 1994. doi:10.1016/S0022-0000(05)80043-1.
- [PRS88] Pavel Pudlák, Vojtech Rödl, and Petr Savický. Graph complexity. *Acta Informatica*, 25(5):515–535, 1988. doi:10.1007/BF00279952.
- [PS86] 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.
- [PSS14] 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.
- [PY84] 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.
- [Raz89] Alexander Razborov. On rigid matrices. Technical report, Steklov Mathematical Institute, 1989. In Russian.
- [Raz92] Alexander Razborov. On the distributional complexity of disjointness. *Theoretical Computer Science*, 106(2):385–390, 1992. doi:10.1016/0304-3975(92)90260-M.
- [RS98] Alexander Russell and Ravi Sundaram. Symmetric alternation captures BPP. *Computational Complexity*, 7(2):152–162, 1998. doi:10.1007/s000370050007.
- [RS04] 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.
- [RS10] Alexander Razborov and Alexander Sherstov. The sign-rank of  $AC^0$ . *SIAM Journal on Computing*, 39(5):1833–1855, 2010. doi:10.1137/080744037.
- [She08] Alexander Sherstov. Halfspace matrices. *Computational Complexity*, 17(2):149–178, 2008. doi:10.1007/s00037-008-0242-4.

- [She11a] Alexander Sherstov. The pattern matrix method. *SIAM Journal on Computing*, 40(6):1969–2000, 2011. doi:10.1137/080733644.
- [She11b] Alexander Sherstov. The unbounded-error communication complexity of symmetric functions. *Combinatorica*, 31(5):583–614, 2011. doi:10.1007/s00493-011-2580-0.
- [Tod91] Seinosuke Toda. PP is as hard as the polynomial-time hierarchy. *SIAM Journal on Computing*, 20(5):865–877, 1991. doi:10.1137/0220053.
- [Tri10] 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.
- [Val77] 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.
- [Ver95] Nikolai Vereshchagin. Lower bounds for perceptrons solving some separation problems and oracle separation of AM from PP. In *Proceedings of the 3rd Israel Symposium on Theory of Computing and Systems (ISTCS)*, pages 46–51. IEEE, 1995. doi:10.1109/ISTCS.1995.377047.
- [Ver99] Nikolai Vereshchagin. Relativizability in complexity theory. In *Provability, Complexity, Grammars*, volume 192 of *AMS Translations, Series 2*, pages 87–172. American Mathematical Society, 1999.
- [VV86] 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.
- [Wun12] Henning Wunderlich. On a theorem of Razborov. *Computational Complexity*, 21(3):431–477, 2012. doi:10.1007/s00037-011-0021-5.