A $\mathsf{ZPP}^{\mathsf{NP}[1]}$ Lifting Theorem

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

Abstract

The complexity class $\mathsf{ZPP}^{\mathsf{NP}[1]}$ (corresponding to zero-error randomized algorithms with access to one NP oracle query) is known to have a number of curious properties. We further explore this class in the settings of time complexity, query complexity, and communication complexity.

- For starters, we provide a new characterization: $\mathsf{ZPP}^{\mathsf{NP}[1]}$ *equals* the restriction of $\mathsf{BPP}^{\mathsf{NP}[1]}$ where the algorithm is only allowed to err when it forgoes the opportunity to make an NP oracle query.
- Using the above characterization, we prove a *query-to-communication lifting theorem*, which translates any $\mathsf{ZPP}^{\mathsf{NP}[1]}$ decision tree lower bound for a function f into a $\mathsf{ZPP}^{\mathsf{NP}[1]}$ communication lower bound for a two-party version of f.
- As an application, we use the above lifting theorem to prove that the $\mathsf{ZPP}^{\mathsf{NP}[1]}$ communication lower bound technique introduced by Göös, Pitassi, and Watson (ICALP 2016) is not tight. We also provide a "primal" characterization of this lower bound technique as a complexity class.

2012 ACM Subject Classification Theory of computation → Computational complexity and cryptography

Keywords and phrases Query complexity, communication complexity, lifting

Digital Object Identifier 10.4230/LIPIcs.STACS.2019.59

Related Version A full version of the paper is available at `https://eccc.weizmann.ac.il/report/2017/139/`.

Funding Supported by NSF grant CCF-1657377.

Acknowledgements I thank Mika Göös and Toniann Pitassi for discussions, and anonymous reviewers for thoughtful comments.

1 Introduction

Query-to-communication lifting is a paradigm for proving lower bounds in communication complexity [30, 26, 34] using lower bounds in query complexity (a.k.a. decision tree complexity) [40, 9, 26]. This technique has yielded a wide array of applications, including lower bounds for the Clique vs. Independent Set communication game and the related Alon–Saks–Seymour conjecture in graph theory [15, 6], separations between communication complexity and partition number [20, 3, 16, 4, 22], lower bounds for monotone circuits, monotone span programs, and proof complexity [35, 7, 25, 19, 14, 36, 33], new and unified proofs of quantum communication lower bounds [38] and of separations between randomized and quantum communication complexity [22, 1, 2], lower bounds for LP and SDP relaxations of CSPs [12, 31, 29], separations between communication complexity classes [10, 28, 18, 21, 17, 8], and lower bounds for finding Nash equilibria [37, 5].

The basic format of the technique is a two-step approach in which a relatively simple problem-specific argument is combined with fairly heavy-duty general-purpose machinery for handling communication protocols. More specifically:

36th International Symposium on Theoretical Aspects of Computer Science (STACS 2019).
Editors: Rolf Niedermeier and Christophe Paul; Article No. 59; pp. 59:1–59:16
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

Class	Reference
P	[35, 20]
NP	[18, 15]
BPP	[22]
P^{NP}	[17]
$\mathsf{ZPP}^{\mathsf{NP}[1]}$	Theorem 2
SBP	[18]
AWPP	[38]
$\mathsf{PostBPP}$	[18]
PP	[38]

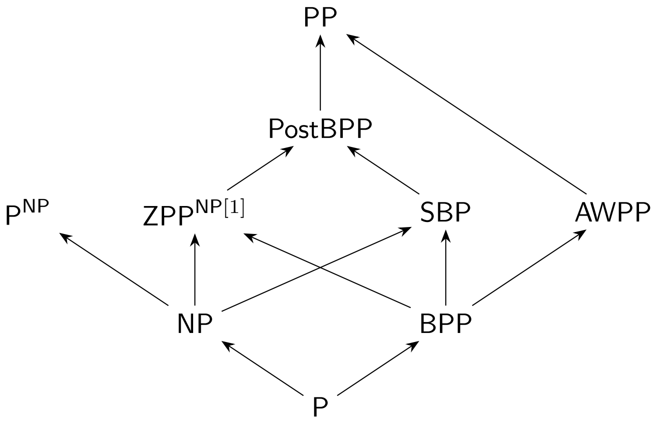

Figure 1 Classes with a known query-to-communication lifting theorem. $\mathcal{C}_1 \to \mathcal{C}_2$ denotes $\mathcal{C}_1 \subseteq \mathcal{C}_2$.

(1) Capture the combinatorial core of the desired communication complexity lower bound by proving an analogous query complexity lower bound.

(2) Apply a *lifting theorem* showing that the query complexity of any boolean function f is essentially the same as the communication complexity of a two-party version of f.

The availability of a lifting theorem greatly eases the burden on the lower bound prover, since query lower bounds are generally much easier to prove than communication lower bounds.

The lifting theorem is with respect to a particular model of computation: deterministic, randomized, nondeterministic, and so on; it is convenient to associate these models with their corresponding classical time-bounded complexity classes: P, BPP, NP, and so on. This idea has led to an ongoing project: prove lifting theorems for the query/communication analogues of various classical complexity classes. Figure 1 shows the main classes for which a lifting theorem is known, along with primary references. Even the less well-known classes sometimes correspond to standard measures in the query/communication settings; e.g., AWPP corresponds to approximate polynomial degree in query complexity and to log of approximate rank in communication complexity. Some notable classes for which a lifting theorem is not known include BQP, UP, and MA. Proving a lifting theorem for AM would be a breakthrough, as it is notoriously open to prove any strong AM-type communication lower bound for an explicit function, but is trivial to do so in the query complexity setting.

Our central contribution is a lifting theorem for the slightly exotic class $\mathsf{ZPP}^{\mathsf{NP}[1]}$, which corresponds to randomized algorithms that can make one call to an NP oracle, output the correct answer with probability $\geq 3/4$, and output $\perp$ with the remaining probability. This model is interesting partly because it has so many curious properties, one of which is that it is robust with respect to the success probability threshold: by [13], the success probability can be efficiently amplified as long as it is $> 1/2$ (which is nontrivial since the standard method for amplification would use multiple independent trials, resulting in multiple NP oracle queries). In terms of relations to other classes, $\mathsf{ZPP}^{\mathsf{NP}[1]}$ contains BPP [11] and is contained in $\mathsf{S}_2\mathsf{P}$ [11] and in $\mathsf{PostBPP}$ (a.k.a. $\mathsf{BPP}_{\mathsf{path}}$) [21]. If we generalized $\mathsf{ZPP}^{\mathsf{NP}[1]}$ to allow success probability slightly $< 1/2$, or to allow two nonadaptive NP oracle calls, either way the class would contain $\mathsf{AM} \cap \mathsf{coAM}$, and hence proving explicit lower bounds for the communication version would yield breakthrough AM communication lower bounds; in this sense, $\mathsf{ZPP}^{\mathsf{NP}[1]}$ is just shy of the communication lower bound frontier. $\mathsf{ZPP}^{\mathsf{NP}[1]}$ also shows up frequently in the literature on the "two queries problem" [39].

Our starting point is to uncover another curious property of $\mathsf{ZPP}^{\mathsf{NP}[1]}$: we prove it is equivalent (in time, query, and communication complexities) to a new model we dub $\mathsf{CautiousBPP}^{\mathsf{NP}[1]}$, which corresponds to randomized algorithms that can make one call to an NP oracle, output the correct answer with probability $\geq 3/4$, and are only allowed to err when they choose not to call the NP oracle. This equivalence plays a crucial role in our proof of the lifting theorem for $\mathsf{ZPP}^{\mathsf{NP}[1]}$.

Once we have the lifting theorem, the natural application domain is to prove new $\mathsf{ZPP}^{\mathsf{NP}[1]}$-type communication lower bounds. [21] developed a technique for proving such lower bounds, and we use our lifting theorem to derive new separations, which imply that the technique from [21] is not tight. This is analogous to the main application from [17], in which a P^{NP} lifting theorem was used to show that the P^{NP}-type communication lower bound technique from [24, 32] is not tight. For context, we note that certain other communication complexity classes have similar lower bound techniques that *are* tight; e.g., the discrepancy bound captures PP communication [27], and the corruption bound captures SBP communication [23]. So for what class is the lower bound technique from [21] tight, if not $\mathsf{ZPP}^{\mathsf{NP}[1]}$? We also answer this question in the full version of this paper. The class did not have a standard name, but it turns out to have a reasonably simple definition.

1.1 Statement of results

We formally define $\mathsf{ZPP}^{\mathsf{NP}[1]}$ and $\mathsf{CautiousBPP}^{\mathsf{NP}[1]}$ and their query/communication analogues in Section 2. For any model $\mathcal{C}$ (such as $\mathsf{ZPP}^{\mathsf{NP}[1]}$ or $\mathsf{CautiousBPP}^{\mathsf{NP}[1]}$) we use $\mathcal{C}$ for the polynomial time complexity class, $\mathcal{C}^{\mathsf{dt}}$ and $\mathcal{C}^{\mathsf{cc}}$ for the polylog query and communication complexity classes, and $\mathcal{C}^{\mathsf{dt}}(f)$ and $\mathcal{C}^{\mathsf{cc}}(F)$ for the corresponding query and communication complexities of a partial function $f\colon \{0,1\}^n \to \{0,1\}$ and a partial two-party function $F\colon \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$ (we also consider F's where Alice and Bob have unequal but polynomially-related input lengths). We use $\tilde{\Theta}$ to hide polylog(n) factors. We prove the following characterization in Section 3.

▶ **Theorem 1.**

(i) $\mathsf{ZPP}^{\mathsf{NP}[1]} = \mathsf{CautiousBPP}^{\mathsf{NP}[1]}$.

(ii) $\mathsf{ZPP}^{\mathsf{NP}[1]\mathsf{dt}}(f) = \tilde{\Theta}(\mathsf{CautiousBPP}^{\mathsf{NP}[1]\mathsf{dt}}(f))$ *for all* f.

(iii) $\mathsf{ZPP}^{\mathsf{NP}[1]\mathsf{cc}}(F) = \tilde{\Theta}(\mathsf{CautiousBPP}^{\mathsf{NP}[1]\mathsf{cc}}(F))$ *for all* F.

We now prepare to state the lifting theorem. For $f\colon \{0,1\}^n \to \{0,1\}$ (called the *outer function*) and $g\colon \mathcal{X} \times \mathcal{Y} \to \{0,1\}$ (called the *gadget*), their composition $f \circ g^n\colon \mathcal{X}^n \times \mathcal{Y}^n \to \{0,1\}$ is the two-party function where Alice gets $x = (x_1, \ldots, x_n) \in \mathcal{X}^n$, Bob gets $y = (y_1, \ldots, y_n) \in \mathcal{Y}^n$, and the goal is to evaluate $(f \circ g^n)(x, y) \coloneqq f(g(x_1, y_1), \ldots, g(x_n, y_n))$. Note that any deterministic (P-type) decision tree for f can be turned into a deterministic protocol for $f \circ g^n$ where Alice and Bob communicate to evaluate $g(x_i, y_i)$ whenever the decision tree queries the i^{th} input bit of f. A similar thing can be done in other models besides deterministic. The essence of a lifting theorem is to go in the other direction: convert a protocol for $f \circ g^n$ into a comparable-cost decision tree for f. In other words, if g is sufficiently complicated, then it hides the input bits to f so well that a communication protocol cannot do any better than just running a decision tree for f.

We use the *index* gadget $\textsc{Ind}_m\colon [m] \times \{0,1\}^m \to \{0,1\}$ mapping $(x, y) \mapsto y_x$, where m is a sufficiently large polynomial in n. This gadget has previously been used for the P, BPP, and P^{NP} lifting theorems. (In some cases, lifting theorems with simpler gadgets are known, but for many applications the index gadget is fine.)

▶ **Theorem 2.** *Let $m = m(n) := n^C$ for a large enough constant C. For every $f\colon \{0,1\}^n \to \{0,1\}$,*

(i) $\mathsf{ZPP}^{\mathsf{NP}[1]\mathsf{cc}}(f \circ \textsc{Ind}_m^n) = \tilde{\Theta}(\mathsf{ZPP}^{\mathsf{NP}[1]\mathsf{dt}}(f))$,

(ii) $\mathsf{CautiousBPP}^{\mathsf{NP}[1]\mathsf{cc}}(f \circ \textsc{Ind}_m^n) = \Theta(\mathsf{CautiousBPP}^{\mathsf{NP}[1]\mathsf{dt}}(f) \cdot \log n)$.

Note that part (i) of Theorem 2 is a corollary of part (ii), since by Theorem 1,

$$\begin{aligned}\mathsf{ZPP}^{\mathsf{NP}[1]\mathsf{cc}}(f \circ \textsc{Ind}_m^n) &= \tilde{\Theta}(\mathsf{CautiousBPP}^{\mathsf{NP}[1]\mathsf{cc}}(f \circ \textsc{Ind}_m^n)) \\ &= \tilde{\Theta}(\mathsf{CautiousBPP}^{\mathsf{NP}[1]\mathsf{dt}}(f)) = \tilde{\Theta}(\mathsf{ZPP}^{\mathsf{NP}[1]\mathsf{dt}}(f)).\end{aligned}$$

We are not aware of a way to prove part (i) directly, without going through Theorem 1. To prove part (ii) (in Section 4), we combine tools and techniques from the proofs of lifting theorems for BPP [22], NP [18, 15], and P^{NP} [17], along with some new technical contributions.

Two of the main results in [21] are $\mathsf{MA}^{\mathsf{cc}} \not\subseteq \mathsf{ZPP}^{\mathsf{NP}[1]\mathsf{cc}}$ and $\mathsf{US}^{\mathsf{cc}} \not\subseteq \mathsf{ZPP}^{\mathsf{NP}[1]\mathsf{cc}}$, where MA and US are the classes associated with "Merlin–Arthur games" and "unique witnesses" respectively (more precise definitions are deferred to Section 5). The proofs introduced a certain lower bound technique – let us use $\mathcal{B}^{\mathsf{cc}}(F)$ for the largest bound attainable for F using this technique, and $\mathcal{B}^{\mathsf{cc}}$ for the class of all F's with $\mathcal{B}^{\mathsf{cc}}(F) \leq \text{polylog}(n)$ – and showed that $\mathsf{MA}^{\mathsf{cc}} \not\subseteq \mathcal{B}^{\mathsf{cc}}$, $\mathsf{US}^{\mathsf{cc}} \not\subseteq \mathcal{B}^{\mathsf{cc}}$, and $\mathsf{ZPP}^{\mathsf{NP}[1]\mathsf{cc}} \subseteq \mathcal{B}^{\mathsf{cc}}$. The definition of $\mathcal{B}^{\mathsf{cc}}$ is not important for now, but we provide it in the full version of this paper, where we show that it can be characterized as a more natural complexity class.

Since $\mathsf{ZPP}^{\mathsf{NP}[1]\mathsf{cc}}$ is closed under complement (whereas $\mathcal{B}^{\mathsf{cc}}$ is not), we have $\mathsf{ZPP}^{\mathsf{NP}[1]\mathsf{cc}} \subseteq \mathcal{B}^{\mathsf{cc}} \cap \mathsf{co}\mathcal{B}^{\mathsf{cc}}$. A natural question is whether the latter is actually an equality, i.e., whether the lower bound technique of [21] for $\mathsf{ZPP}^{\mathsf{NP}[1]\mathsf{cc}}$ is tight. Since [21] observed that $\mathsf{MA}^{\mathsf{cc}}, \mathsf{US}^{\mathsf{cc}} \subseteq \mathsf{co}\mathcal{B}^{\mathsf{cc}}$, we have $\mathsf{MA}^{\mathsf{cc}} \cap \mathsf{coMA}^{\mathsf{cc}}$, $\mathsf{US}^{\mathsf{cc}} \cap \mathsf{coUS}^{\mathsf{cc}} \subseteq \mathcal{B}^{\mathsf{cc}} \cap \mathsf{co}\mathcal{B}^{\mathsf{cc}}$, and thus the following result (proven in Section 5 using Theorem 2) answers this question in the negative (in two different ways).

▶ **Theorem 3.**

(i) $\mathsf{MA}^{\mathsf{cc}} \cap \mathsf{coMA}^{\mathsf{cc}} \not\subseteq \mathsf{ZPP}^{\mathsf{NP}[1]\mathsf{cc}}$.

(ii) $\mathsf{US}^{\mathsf{cc}} \cap \mathsf{coUS}^{\mathsf{cc}} \not\subseteq \mathsf{ZPP}^{\mathsf{NP}[1]\mathsf{cc}}$.

2 Definitions

We set up notation and provide the formal definitions of $\mathsf{ZPP}^{\mathsf{NP}[1]}$ and $\mathsf{CautiousBPP}^{\mathsf{NP}[1]}$. For the query and communication complexity versions, we follow the convention of using the complexity class names as complexity measures. That is, $\mathcal{C}^{\mathsf{dt}}(f)$ denotes the minimum cost of any correct $\mathcal{C}$-type decision tree for f, and $\mathcal{C}^{\mathsf{dt}}$ also denotes the class of families of partial f's with $\mathcal{C}^{\mathsf{dt}}(f) \leq \text{polylog}(n)$; similarly, $\mathcal{C}^{\mathsf{cc}}(F)$ denotes the minimum cost of any correct $\mathcal{C}$-type communication protocol for F, and $\mathcal{C}^{\mathsf{cc}}$ also denotes the class of families of partial F's with $\mathcal{C}^{\mathsf{cc}}(F) \leq \text{polylog}(n)$ (assuming Alice and Bob have polynomially-related input lengths).

In the query complexity setting, "query" actually has two meanings for us: a decision tree makes queries to individual input bits, then it forms an NP-type (DNF) oracle query.

We think of a randomized algorithm M as taking a uniformly random string $s \in \{0,1\}^r$ (for some number of coins r that depends on the input length); we let $M_s(x)$ denote M running on input x with outcome s. Similarly, we think of a randomized (in our case, $\mathsf{ZPP}^{\mathsf{NP}[1]}$-type or $\mathsf{CautiousBPP}^{\mathsf{NP}[1]}$-type) decision tree T or communication protocol Π as the uniform distribution over a multiset of corresponding non-randomized T_s's or Π_s's indexed by $s \in \{0,1\}^r$; we denote this as $T \sim \{T_s : s \in \{0,1\}^r\}$ or $\Pi \sim \{\Pi_s : s \in \{0,1\}^r\}$.

2.1 $\mathsf{ZPP}^{\mathsf{NP}[1]}$

$\mathsf{ZPP}^{\mathsf{NP}[1]}$ consists of all languages L for which there is a polynomial-time randomized algorithm M (taking input x and coin tosses $s \in \{0,1\}^r$) and a language $L' \in \mathsf{NP}$ such that the following hold.

Syntax: The computation of $M_s(x)$ produces an oracle query q and a function $out\colon \{0,1\} \to \{0,1,\bot\}$; the output is then $out(L'(q))$.

Correctness: The output is always $L(x)$ or $\bot$, and is $L(x)$ with probability $\geq 3/4$.

We define a $\mathsf{ZPP}^{\mathsf{NP}[1]}$-type decision tree T for f on input x as follows.

Syntax: $T \sim \{T_s : s \in \{0,1\}^r\}$ where each T_s makes queries to the bits of x until it reaches a leaf, which is labeled with a DNF D and a function $out\colon \{0,1\} \to \{0,1,\bot\}$; the output is then $out(D(x))$.

Correctness: The output is always $f(x)$ or $\bot$, and is $f(x)$ with probability $\geq 3/4$.

Cost: The maximum height of any T_s, plus the maximum width of any DNF appearing at a leaf.

We define a $\mathsf{ZPP}^{\mathsf{NP}[1]}$-type communication protocol Π for F on input (x, y) as follows.

Syntax: $\Pi \sim \{\Pi_s : s \in \{0,1\}^r\}$ where each Π_s communicates until it reaches a leaf, which is labeled with a multiset of rectangles $\{R^w : w \in \{0,1\}^k\}$ (for some k) and a function $out\colon \{0,1\} \to \{0,1,\bot\}$; the output is then out applied to the indicator of whether $(x, y) \in \bigcup_w R^w$.

Correctness: The output is always $F(x, y)$ or $\bot$, and is $F(x, y)$ with probability $\geq 3/4$.

Cost: The maximum communication cost of any Π_s, plus the maximum k at any leaf.

A priori, the value $3/4$ seems arbitrary since it is not clear whether $\mathsf{ZPP}^{\mathsf{NP}[1]}$ is amenable to amplification of the success probability (naively doing repeated trials would increase the number of NP queries). However, [13] showed that amplification is actually possible, so we may use any constant $> 1/2$ for the success probability in the definition of $\mathsf{ZPP}^{\mathsf{NP}[1]}$ (while affecting the measures $\mathsf{ZPP}^{\mathsf{NP}[1]\mathsf{dt}}(f)$ and $\mathsf{ZPP}^{\mathsf{NP}[1]\mathsf{cc}}(F)$ by only constant factors).

2.2 $\mathsf{CautiousBPP}^{\mathsf{NP}[1]}$

$\mathsf{CautiousBPP}^{\mathsf{NP}[1]}$ consists of all languages L for which there is a polynomial-time randomized algorithm M (taking input x and coin tosses $s \in \{0,1\}^r$) and a language $L' \in \mathsf{NP}$ such that the following hold.

Syntax: The computation of $M_s(x)$ either directly outputs a bit (without invoking the oracle) or produces an oracle query q and a nonconstant function $out\colon \{0,1\} \to \{0,1\}$; in the latter case the output is then $out(L'(q))$.

Correctness: The output is $L(x)$ with probability $\geq 3/4$, and is $L(x)$ for all s such that $M_s(x)$ makes an oracle query.

We define a $\mathsf{CautiousBPP}^{\mathsf{NP}[1]}$-type decision tree T for f on input x as follows.

Syntax: $T \sim \{T_s : s \in \{0,1\}^r\}$ where each T_s makes queries to the bits of x until it reaches a leaf, which is labeled with either an output bit, or a DNF D and a nonconstant function $out\colon \{0,1\} \to \{0,1\}$; in the latter case the output is then $out(D(x))$.

Correctness: The output is $f(x)$ with probability $\geq 3/4$, and is $f(x)$ for all s such that $T_s(x)$ makes a DNF query.

Cost: The maximum height of any T_s, plus the maximum width of any DNF appearing at a leaf.

We define a $\mathsf{CautiousBPP}^{\mathsf{NP}[1]}$-type communication protocol Π for F on input (x, y) as follows.

Syntax: $\Pi \sim \{\Pi_s : s \in \{0,1\}^r\}$ where each Π_s communicates until it reaches a leaf, which is labeled with either an output bit, or a multiset of rectangles $\{R^w : w \in \{0,1\}^k\}$ (for some k) and a nonconstant function out: $\{0,1\} \to \{0,1\}$; in the latter case the output is then out applied to the indicator of whether $(x,y) \in \bigcup_w R^w$.
Correctness: The output is $F(x,y)$ with probability $\geq 3/4$, and is $F(x,y)$ for all s such that $\Pi_s(x,y)$ makes a "union of rectangles" query.
Cost: The maximum communication cost of any Π_s, plus the maximum k at any leaf.

The success probability of any $\mathsf{CautiousBPP}^{\mathsf{NP}[1]}$-type computation can be amplified by taking the majority vote of multiple independent trials – except if at least one trial results in an NP-type oracle query then (to avoid multiple oracle queries) we just use the output of one such trial since we know it will be correct. Thus just like for BPP-type computations, success probability $1/2 + \varepsilon$ can be amplified to $1 - \delta$ with a $O(\frac{1}{\varepsilon^2} \log \frac{1}{\delta})$ factor overhead in cost.

3 $\mathsf{ZPP}^{\mathsf{NP}[1]} = \mathsf{CautiousBPP}^{\mathsf{NP}[1]}$

We now prove Theorem 1, starting with part (i). First assume $L \in \mathsf{ZPP}^{\mathsf{NP}[1]}$, witnessed by a randomized algorithm M (taking input x and coin tosses $s \in \{0,1\}^r$) and $L' \in \mathsf{NP}$. To see that $L \in \mathsf{CautiousBPP}^{\mathsf{NP}[1]}$, consider this randomized algorithm with oracle access to L':

1. Sample $s \in \{0,1\}^r$ and run $M_s(x)$ until it produces q and out.
2. If $out(0) = out(1)$ then output this common bit, or an arbitrary bit if $out(0) = out(1) = \perp$.
3. Else if one of $out(0), out(1)$ is $\perp$ then output whichever is not $\perp$.
4. Else invoke the oracle on q and output $out(L'(q))$.

Consider any s for which this algorithm outputs the wrong bit: then it did not make an oracle query (since M never outputs the wrong bit), and $M_s(x)$ would have output $\perp$ (because of either line 2, or line 3 with $out(L'(q)) = \perp$ and $out(1 - L'(q)) \neq L(x)$). Hence this algorithm correctly solves L, with error probability at most that of M.

For the converse direction, we generalize the argument from [11] that $\mathsf{BPP} \subseteq \mathsf{ZPP}^{\mathsf{NP}[1]}$. Assume $L \in \mathsf{CautiousBPP}^{\mathsf{NP}[1]}$, witnessed by a randomized algorithm M (taking input x and coin tosses $s \in \{0,1\}^r$) and $L' \in \mathsf{NP}$. Assume that this has already been amplified so the error probability is $< 1/4r$ (by the remark at the end of Section 2.2). For a fixed input x and $b \in \{0,1\}$, let

$$S_b \;\coloneqq\; \{s \in \{0,1\}^r : M_s(x) \text{ outputs } b \text{ without invoking the oracle}\}.$$

To see that $L \in \mathsf{ZPP}^{\mathsf{NP}[1]}$, consider the following randomized algorithm:

1. Sample $s \in \{0,1\}^r$ and run $M_s(x)$ until it produces either an output b (so $s \in S_b$) or q and out.
2. If it produced q and out then ask the NP oracle for the value of $L'(q)$ and output $out(L'(q))$.
3. Else sample independent strings $s^1, \ldots, s^{4r} \in \{0,1\}^r$ and ask the NP oracle whether $\bigcup_i (S_b \oplus s^i) \neq \{0,1\}^r$ (i.e., whether there exists an s' such that for every i, $s' \oplus s^i \notin S_b$); output $\perp$ if so and b if not.

Note that this algorithm never outputs the wrong bit: if $s \in S_b$ for $b = 1 - L(x)$, then $|S_b| < 2^r/4r$ so by a union bound, $\left|\bigcup_i (S_b \oplus s^i)\right| < 4r \cdot (2^r/4r) = 2^r$ and hence the NP oracle returns 1 on line 3 and the algorithm outputs $\perp$. For the success probability, consider two cases. If $|S_0 \cup S_1| \leq 2^r/4$, then line 2 executes (guaranteeing correct output) with probability $\geq 3/4$. Otherwise, since $|S_{1-L(x)}| < 2^r/4r$, we must have $|S_{L(x)}| > 2^r/4 - 2^r/4r > 2^r/5$ (we may assume r is at least a large enough constant), so by a union bound over all $s' \in \{0,1\}^r$, the probability over $s^1, \ldots, s^{4r}$ that $\bigcup_i (S_{L(x)} \oplus s^i) \neq \{0,1\}^r$ is $< 2^r \cdot (4/5)^{4r} \leq (5/6)^r \leq 1/5$. In this latter case, the probability of outputting $\perp$ is

$$\begin{aligned}&\mathbb{P}[b = 1 - L(x)] + \mathbb{P}\big[\textstyle\bigcup_i (S_b \oplus s^i) \neq \{0,1\}^r \,\big|\, b = L(x)\big] \cdot \mathbb{P}[b = L(x)]\\ &\leq\ 1/4r + (1/5) \cdot |S_{L(x)}|/2^r \ \leq\ 1/4.\end{aligned}$$

In both cases the success probability is $\geq 3/4$.

Parts (ii) and (iii) are proved in the same way as part (i), but we must carefully analyze the cost. We summarize the differences. For the $\mathsf{ZPP}^{\mathsf{NP}[1]} \subseteq \mathsf{CautiousBPP}^{\mathsf{NP}[1]}$ direction, exactly the same argument works but using T_s or Π_s in place of M_s, and making the same DNF query or "union of rectangles" query rather than the same NP oracle query on line 4. This shows $\mathsf{CautiousBPP}^{\mathsf{NP}[1]\mathsf{dt}}(f) \leq \mathsf{ZPP}^{\mathsf{NP}[1]\mathsf{dt}}(f)$ and $\mathsf{CautiousBPP}^{\mathsf{NP}[1]\mathsf{cc}}(F) \leq \mathsf{ZPP}^{\mathsf{NP}[1]\mathsf{cc}}(F)$.

Now consider the $\mathsf{CautiousBPP}^{\mathsf{NP}[1]} \subseteq \mathsf{ZPP}^{\mathsf{NP}[1]}$ direction for parts (ii) and (iii). By standard sparsification of the randomness, we may assume T or Π uses only $O(\log n)$ coin tosses (while affecting the success probability by only $\pm o(1)$). Then as noted at the end of Section 2.2, we may amplify with $O(\log\log n)$ repetitions so r becomes $O(\log n \cdot \log\log n)$ and the error probability becomes $\leq 1/\log^2 n < 1/4r$. As above, we use T_s or Π_s in place of M_s, and make the same DNF query or "union of rectangles" query rather than the same NP oracle query on line 2. For line 3, we note that the predicate $\bigcup_i (S_b \oplus s^i) \neq \{0,1\}^r$, as a function of the input x or (x,y), can be computed by nondeterministically guessing s' and running $T_{s' \oplus s^i}(x)$ or $\Pi_{s' \oplus s^i}(x,y)$ for each $i \in [4r]$; this can be expressed as a DNF of width $4r\cdot$(cost of amplified T), or as a union of 2^k rectangles with $k = r + 4r \cdot$ (cost of amplified Π). Thus, the overall cost is $O(r \cdot (\text{cost of amplified } T \text{ or } \Pi)) \leq O((\text{cost of original } T \text{ or } \Pi) \cdot \log n \cdot \log^2 \log n)$. This shows the following, finishing the proof of Theorem 1:

$$\begin{aligned}\mathsf{ZPP}^{\mathsf{NP}[1]\mathsf{dt}}(f) &\leq O\big(\mathsf{CautiousBPP}^{\mathsf{NP}[1]\mathsf{dt}}(f) \cdot \log n \cdot \log^2 \log n\big),\\ \mathsf{ZPP}^{\mathsf{NP}[1]\mathsf{cc}}(F) &\leq O\big(\mathsf{CautiousBPP}^{\mathsf{NP}[1]\mathsf{cc}}(F) \cdot \log n \cdot \log^2 \log n\big).\end{aligned}$$

4 Proof of the Lifting Theorem

We now prove Theorem 2. As noted in Section 1.1, we just need to show part (ii). It is straightforward to see that for all f,

$$\mathsf{CautiousBPP}^{\mathsf{NP}[1]\mathsf{cc}}(f \circ \textsc{Ind}_m^n) \ \leq\ O(\mathsf{CautiousBPP}^{\mathsf{NP}[1]\mathsf{dt}}(f) \cdot \log n)$$

since we can have the communication protocol run the optimal decision tree for f, communicating $O(\log n)$ bits to evaluate $\textsc{Ind}_m(x_i, y_i)$ whenever this bit is queried, and if a width-w DNF oracle query is formed then we can convert each of its $\leq n^w$ conjunctions into $\leq m^w$ rectangles, resulting in a "union of rectangles" oracle query that contributes $k = O(w \log n)$ to the cost. Thus, the bulk of the proof is to show that for all f,

$$\mathsf{CautiousBPP}^{\mathsf{NP}[1]\mathsf{dt}}(f) \ \leq\ O(\mathsf{CautiousBPP}^{\mathsf{NP}[1]\mathsf{cc}}(f \circ \textsc{Ind}_m^n)/\log n). \tag{1}$$

In Section 4.1 we provide relevant technical background from the proofs of earlier lifting theorems (mainly the one for BPP [22]). In Section 4.2 we describe how to simulate the communication protocol with a decision tree. In Section 4.3 we prove a key technical lemma.

4.1 Background

Abbreviate $G \coloneqq \textsc{Ind}_m^n$. We consider deterministic communication protocols on G's input domain $[m]^n \times (\{0,1\}^m)^n$, which we view as partitioned into *slices* $G^{-1}(z) = \{(x,y) : G(x,y) = z\}$, one for each $z \in \{0,1\}^n$. We let $|\Pi|$ denote the worst-case number of bits communicated by a deterministic protocol Π. We use boldface letters for random variables.

Let $\mathbb{H}_\infty(\boldsymbol{x}) \coloneqq \min_x \log(1/\mathbb{P}[\boldsymbol{x} = x])$ denote the usual min-entropy of a random variable $\boldsymbol{x}$. Supposing $\boldsymbol{x}$ is distributed over a set X, we define the *deficiency* of $\boldsymbol{x}$ as the nonnegative quantity $\mathbb{D}_\infty(\boldsymbol{x}) \coloneqq \log|X| - \mathbb{H}_\infty(\boldsymbol{x})$. A basic property is that if X is a Cartesian product then marginalizing $\boldsymbol{x}$ to some coordinates cannot increase the deficiency. For a set X we let $\boldsymbol{X}$ denote a random variable uniformly distributed on X.

The following definition and claim originate in the proof of the lifting theorems for NP, SBP, and PostBPP [18, 15]. They describe an invariant that Alice maintains throughout the simulation, and how to restore it (by fixing some coordinates, which will correspond to querying those input bits of f) when it gets violated.

▶ **Definition 4.** *A random variable $\boldsymbol{x} \in [m]^J$ (where $J \subseteq [n]$ is some index set) is called* δ-dense *if for every nonempty $I \subseteq J$, the coordinates $\boldsymbol{x}_I$ (marginally distributed over $[m]^I$) have min-entropy rate at least δ, i.e., $\mathbb{H}_\infty(\boldsymbol{x}_I) \geq \delta \cdot |I| \log m$.*

▷ Claim 5. If $A \subseteq [m]^J$ then there exist an $I \subseteq J$ of size $|I| \leq O(\mathbb{D}_\infty(\boldsymbol{A})/\log n)$ and a nonempty $A' \subseteq A$ such that $\boldsymbol{A}'$ is fixed on I and 0.9-dense on $J \smallsetminus I$.

It is simple to check that all 2^n slices of G's input domain have the same size, and that the uniform distribution over any slice is marginally nearly-uniform on both Alice's input and Bob's input. The following lemma from [22] provides a sufficient condition for similar properties to hold even after we have queried some of the input bits of f.

▶ **Definition 6.** *For a partial assignment $\rho \in \{0,1,*\}^n$, define its* free *positions as* $\operatorname{free}\rho \coloneqq \rho^{-1}(*) \subseteq [n]$*, and its* fixed *positions as* $\operatorname{fix}\rho \coloneqq [n] \smallsetminus \operatorname{free}\rho$*. A rectangle $X \times Y$ is called* ρ-structured *if $\boldsymbol{X}_{\operatorname{free}\rho}$ is 0.9-dense, $\boldsymbol{X}_{\operatorname{fix}\rho}$ is fixed, and each element of $G(X \times Y) \subseteq \{0,1\}^n$ is consistent with ρ.*

▶ **Definition 7.** *A distribution $\mathcal{D}_1$ is said to be* ε-pointwise-close *to a distribution $\mathcal{D}_2$ if for every outcome, the probability under $\mathcal{D}_1$ is within a factor $1 \pm \varepsilon$ of the probability under $\mathcal{D}_2$. The distributions are said to be* ε-close *if the statistical (total variation) distance is $\leq \varepsilon$.*

▶ **Lemma 8** ([22])**.** *Suppose $X \times Y$ is ρ-structured and $\mathbb{D}_\infty(\boldsymbol{Y}) \leq n^3$. Then:*

(i) *For any $z \in \{0,1\}^n$ consistent with ρ, the uniform distribution on $G^{-1}(z) \cap X \times Y$ (which is nonempty) has both of its marginal distributions $o(1)$-close to $\boldsymbol{X}$ and $\boldsymbol{Y}$, respectively.*

(ii) *$G(\boldsymbol{X}, \boldsymbol{Y})$ is $o(1)$-pointwise-close to the uniform distribution over the set of all z consistent with ρ.*

Now we come to the main part of the proof of the BPP lifting theorem from [22]. It shows that, given query access to z, we can approximately sample the transcript that would be generated by a communication protocol on a random input from z's slice. In fact, this simulation maintains some invariants, which we need to expose (in the "furthermore" part of the lemma) for use in the subsequent "NP oracle query" phase of our simulation.

▶ **Definition 9.** *A deterministic protocol $\overline{\Pi}$ is said to be a* refinement *of a deterministic protocol Π if they have the same input domain and for every transcript rectangle $X \times Y$ of $\overline{\Pi}$, there exists a transcript rectangle of Π that contains $X \times Y$.*

▶ **Lemma 10** ([22]). *For every deterministic protocol Π on G's input domain with $|\Pi| \leq n \log m$, there exist a refinement $\overline{\Pi}$ and a randomized decision tree T of cost $O(|\Pi|/\log n)$ that on input $z \in \{0,1\}^n$ outputs a transcript of $\overline{\Pi}$ or $\perp$, such that the following two distributions are $o(1)$-close:*

$$\begin{aligned} \boldsymbol{t} &:= \text{output distribution of } T \text{ on input } z, \\ \boldsymbol{t}' &:= \text{transcript generated by } \overline{\Pi} \text{ when run on a random input } (\boldsymbol{x}, \boldsymbol{y}) \sim G^{-1}(z). \end{aligned}$$

Furthermore, for every (non-$\perp$) transcript output by T on input z with positive probability, the associated rectangle $X \times Y$ satisfies:

(i) *$X \times Y$ is ρ-structured, where ρ corresponds to the results of the queries made by T (and is hence consistent with z),*

(ii) $\mathbb{D}_\infty(\boldsymbol{Y}) \leq n^{2.5}$,

(iii) $\mathbb{D}_\infty(\boldsymbol{X}_{\text{free}\,\rho}) \leq O(|\Pi|)$.

4.2 Simulation

▶ **Lemma 11.** *Let $X \times Y$ be a ρ-structured rectangle in G's input domain such that $\mathbb{D}_\infty(\boldsymbol{Y}) \leq n^{2.5}$. Suppose $\{R^w \subseteq X \times Y : w \in \{0,1\}^k\}$ is a collection of rectangles whose union covers exactly $G^{-1}(f^{-1}(1)) \cap X \times Y$. Then f can be computed by a width-$O((\mathbb{D}_\infty(\boldsymbol{X}_{\text{free}\,\rho}) + k)/\log n)$ DNF on the domain of inputs consistent with ρ.*

Lemma 11 is our key tool for converting the $\mathsf{NP}^{\mathsf{cc}}$ oracle query to an $\mathsf{NP}^{\mathsf{dt}}$ oracle query. The proof, which we give in Section 4.3, combines insights from the lifting theorem proofs for NP [18, 15] and P^{NP} [17] with new calculations. For now we use Lemma 11 to argue (1), thus finishing the proof of Theorem 2.

Let $\Pi \sim \{\Pi_s : s \in \{0,1\}^r\}$ be a $\mathsf{CautiousBPP}^{\mathsf{NP}[1]}$-type communication protocol for $f \circ G$ (and note WLOG the cost is $\leq n \log m$). Here is a $\mathsf{CautiousBPP}^{\mathsf{NP}[1]}$-type decision tree for f on input z:

1. Sample $s \in \{0,1\}^r$ and (eliding the dependence on s) let $\overline{\Pi}$ and T be the refinement and randomized decision tree from Lemma 10 applied to Π_s.
2. Sample T's coin tosses s' and run $T_{s'}$ on input z until it either outputs $\perp$ (in which case we halt and output an arbitrary bit) or produces a transcript t of $\overline{\Pi}$.
3. Let $X \times Y$ be the rectangle associated with t, and let t^* be the transcript of Π_s whose rectangle contains $X \times Y$.
4. If t^* outputs a bit, then we halt and output the same bit; otherwise let $\{R^w : w \in \{0,1\}^k\}$ and *out*: $\{0,1\} \to \{0,1\}$ be the rectangles and nonconstant function associated with t^*.
5. Since $X \times Y$ satisfies properties (i), (ii), (iii) from Lemma 10, we may apply Lemma 11 to the collection $\{R^w \cap X \times Y : w \in \{0,1\}^k\}$ (whose union covers exactly $G^{-1}(f^{-1}(\mathit{out}(1))) \cap X \times Y$ by the correctness of Π), using f if $\mathit{out}(1) = 1$ or $\neg f$ if $\mathit{out}(1) = 0$, to obtain a width-$O((|\Pi_s| + k)/\log n)$ DNF D that computes f or $\neg f$ (respectively) on all inputs consistent with ρ.
6. Output $\mathit{out}(D(z))$.

Since T makes $O(|\Pi_s|/\log n)$ queries and the DNF on line 5 has width $O((|\Pi_s| + k)/\log n)$, the above decision tree indeed has cost $O((\text{cost of } \Pi)/\log n)$. If it reaches line 5 and makes a DNF query, then the output is correct since z is consistent with ρ and hence $\mathit{out}(D(z)) = f(z)$. For the success probability, call t *good* if the corresponding t^* either outputs $f(z)$ directly or makes a "union of rectangles" query, and note that if the above decision tree generates a good t then the output is correct (by the previous sentence). Hence, letting $\boldsymbol{t}, \boldsymbol{t}'$ be the $o(1)$-close random variables from Lemma 10 applied to $\Pi_{\boldsymbol{s}}$ (with $(\boldsymbol{x}, \boldsymbol{y}) \sim G^{-1}(z)$), we have

$$\begin{aligned}
\mathbb{P}[\text{output is correct}] &\geq \mathbb{E}_{\boldsymbol{s}}\big[\mathbb{P}_{\boldsymbol{s}'}[\boldsymbol{t} \text{ is a good transcript}]\big] \\
&\geq \mathbb{E}_{\boldsymbol{s}}\big[\mathbb{P}_{\boldsymbol{x},\boldsymbol{y}}[\boldsymbol{t}' \text{ is good}] - o(1)\big] \\
&= \mathbb{E}_{\boldsymbol{x},\boldsymbol{y}}\big[\mathbb{P}_{\boldsymbol{s}}[\boldsymbol{t}' \text{ is good}]\big] - o(1) \\
&= \mathbb{E}_{\boldsymbol{x},\boldsymbol{y}}\big[\mathbb{P}_{\boldsymbol{s}}[\Pi_{\boldsymbol{s}}(\boldsymbol{x},\boldsymbol{y}) \text{ outputs } f(z)]\big] - o(1) \\
&\geq \mathbb{E}_{\boldsymbol{x},\boldsymbol{y}}[3/4] - o(1) \\
&= 3/4 - o(1).
\end{aligned}$$

We conclude that $\mathsf{CautiousBPP}^{\mathsf{NP}[1]\mathsf{dt}}(f) \leq O(\mathsf{CautiousBPP}^{\mathsf{NP}[1]\mathsf{cc}}(f \circ G)/\log n)$.[1]

4.3 Forming a DNF

We now prove Lemma 11. Fix any $z \in f^{-1}(1)$ consistent with ρ, and define $J := \operatorname{free}\rho$. We need to show that there exists a width-$O((\mathbb{D}_\infty(\boldsymbol{X}_J) + k)/\log n)$ conjunction that accepts z but does not accept any input in $f^{-1}(0)$ consistent with ρ.

For each rectangle $R^w = X^w \times Y^w$ define the set of *weighty rows* as

$$A^w := \big\{x \in X^w : |Y^w_x| \geq 2^{nm-n^3}\big\} \quad \text{where} \quad Y^w_x := \big\{y \in Y^w : G(x,y) = z\big\}.$$

▷ Claim 12. There exists a $w \in \{0,1\}^k$ such that $|A^w| \geq |X|/2^{k+1}$.

Proof. Suppose for contradiction this is not the case. Then by Lemma 8.(i) we have

$$\frac{\big|G^{-1}(z) \cap \big(\bigcup_w A^w\big) \times Y\big|}{|G^{-1}(z) \cap X \times Y|} \leq \frac{|\bigcup_w A^w|}{|X|} + o(1) \leq \frac{2^k \cdot |X|/2^{k+1}}{|X|} + o(1) < 3/4. \qquad (2)$$

On the other hand, since the R^w's cover $G^{-1}(z) \cap X \times Y$ and since $k \leq n \log m$ WLOG,

$$\big|G^{-1}(z) \cap \big(X \smallsetminus \textstyle\bigcup_w A^w\big) \times Y\big| \leq \big|\textstyle\bigcup_{w,\, x \notin A^w} Y^w_x\big| \leq 2^k \cdot |X| \cdot 2^{nm-n^3} \leq |X| \cdot 2^{nm-n^{2.9}},$$

and by Lemma 8.(ii) and $\mathbb{D}_\infty(\boldsymbol{Y}) \leq n^{2.5} \leq n^3$ we have

$$|G^{-1}(z) \cap X \times Y| \geq |X| \cdot |Y| \cdot (1 - o(1))/2^{|J|} \geq |X| \cdot 2^{nm-n^{2.5}} \cdot (1 - o(1))/2^n \geq |X| \cdot 2^{nm-n^{2.6}},$$

and thus

$$\frac{\big|G^{-1}(z) \cap \big(X \smallsetminus \bigcup_w A^w\big) \times Y\big|}{|G^{-1}(z) \cap X \times Y|} \leq \frac{|X| \cdot 2^{nm-n^{2.9}}}{|X| \cdot 2^{nm-n^{2.6}}} = 2^{n^{2.6}-n^{2.9}} < 1/4. \qquad (3)$$

Now (2) and (3) form a contradiction. This proves the claim. ◁

[1] Let us summarize the fundamental reason we are unable to make this proof work directly for $\mathsf{ZPP}^{\mathsf{NP}[1]}$ (instead of $\mathsf{CautiousBPP}^{\mathsf{NP}[1]}$) without going through Theorem 1. Suppose we reach line 5 with $\mathit{out}(1) = \perp$ and $\mathit{out}(0) \neq \perp$. We would like to form a DNF that accepts those z's consistent with ρ where $G^{-1}(z) \cap X \times Y$ is covered by the union of $\big\{R^w \cap X \times Y : w \in \{0,1\}^k\big\}$ – and then output $\perp$ if the DNF accepts and output $\mathit{out}(0)$ if it rejects. The issue is that there may be some z's consistent with ρ such that $f(z) = \mathit{out}(0)$ but $G^{-1}(z) \cap X \times Y$ is partially covered by the union – even a fairly small coverage might result in the DNF accepting z. This could cause the overall probability of outputting $\perp$ on z to be much higher in the decision tree than in the communication protocol.

Now fix a $w \in \{0,1\}^k$ such that $|A^w| \geq |X|/2^{k+1}$ and hence $\mathbb{D}_\infty(\boldsymbol{A}^w_J) \leq \mathbb{D}_\infty(\boldsymbol{X}_J) + k + 1$. Applying Claim 5 to A^w_J, we can obtain an $I \subseteq J$ of size $|I| \leq O((\mathbb{D}_\infty(\boldsymbol{X}_J) + k)/\log n)$ and a nonempty $A' \subseteq A^w$ such that $\boldsymbol{A}'$ is fixed on $I \cup \operatorname{fix}\rho$ and 0.9-dense on $J \smallsetminus I$. Consider the conjunction that accepts iff the I coordinates of the input equal z_I; we now argue that this conjunction satisfies the desired properties. It certainly has the right width and accepts z.

Define $\sigma \in \{0,1,*\}^n$ as the partial assignment that extends ρ by fixing the coordinates in I to z_I. Pick any $x' \in A'$ and let $B := Y^w_{x'}$. Then $A' \times B$ is σ-structured (note that for all $(x,y) \in A' \times B$, $G(x,y)_I = G(x',y)_I = z_I$ since $x_I = x'_I$) and $\mathbb{D}_\infty(\boldsymbol{B}) \leq n^3$, and thus by Lemma 8.(ii), $G(\boldsymbol{A}', \boldsymbol{B})$ is $o(1)$-pointwise-close to the uniform distribution over all strings consistent with σ. In particular, for every z' consistent with σ (i.e., for every z' consistent with ρ that is accepted by the conjunction) there exists an $(x,y) \in A' \times B$ such that $G(x,y) = z'$; since $A' \times B \subseteq R^w \subseteq G^{-1}(f^{-1}(1))$, this implies that $f(z') = 1$. In summary, the conjunction does not accept any input in $f^{-1}(0)$ consistent with ρ. This finishes the proof of Lemma 11.

5 Applications

We prove Theorem 3 in this section. Since $\mathsf{MA}^{\mathsf{cc}} \cap \mathsf{coMA}^{\mathsf{cc}}$, $\mathsf{US}^{\mathsf{cc}} \cap \mathsf{coUS}^{\mathsf{cc}} \subseteq \mathcal{B}^{\mathsf{cc}} \cap \mathsf{co}\mathcal{B}^{\mathsf{cc}}$, Theorem 3 cannot be shown using the lower bound technique from [21], so we instead prove the analogous separations in query complexity and apply our lifting theorem. We start by defining the query/communication versions of MA and US.

Merlin–Arthur games (MA) are the model where Merlin nondeterministically sends a message to Arthur (comprised of Alice and Bob in the communication setting), who is randomized and decides whether to accept. On a 1-input, there should exist a witness Merlin can send that makes Arthur accept with probability 1, and on a 0-input, Arthur should reject with probability $\geq 1/2$ no matter what Merlin sends. In the query/communication settings, the cost is Merlin's message length plus Arthur's query/communication cost.

The US model is like ordinary nondeterminism, except that an input is accepted iff there is *exactly one* witness that leads to acceptance (so, rejection means there are either 0 or ≥ 2 accepted witnesses). In query complexity, the cost is the maximum width of any of the witness conjunctions. In communication complexity, the cost is the log of the number of witness rectangles.

5.1 MA ∩ coMA

We now prove Theorem 3.(i). We start with a general technique for proving $\mathsf{CautiousBPP}^{\mathsf{NP}[1]\mathsf{dt}}$ lower bounds. For a bit b, we say a conjunction is b-monochromatic for a partial function f if it rejects all $(1-b)$-inputs.

▶ **Lemma 13.** *Suppose f has no monochromatic conjunction of width $< k$. Then*

$$\mathsf{CautiousBPP}^{\mathsf{NP}[1]\mathsf{dt}}(f) \;\geq\; \min(k, \mathsf{BPP}^{\mathsf{dt}}(f)).$$

Proof. If f has a $\mathsf{CautiousBPP}^{\mathsf{NP}[1]}$-type decision tree of cost $< k$, then this decision tree must never make a DNF query (in which case it is just a BPP-type decision tree, showing that $\mathsf{BPP}^{\mathsf{dt}}(f) \leq \mathsf{CautiousBPP}^{\mathsf{NP}[1]\mathsf{dt}}(f)$). To see this, suppose for contradiction some leaf is labeled with a DNF query D and a function out, and consider the conjunction that accepts the inputs that lead to that leaf and are accepted by an arbitrarily chosen term of D (which WLOG is consistent with the partial assignment leading to the leaf). Then this conjunction has width $< k$ and is $out(1)$-monochromatic (as any input accepted by it would make the $\mathsf{CautiousBPP}^{\mathsf{NP}[1]}$-type decision tree output $out(1)$ after making a DNF query, for some outcome of the coin tosses, and hence could not be an $out(0)$-input). ◀

Let $n = 2\ell^2$, and define the partial function $f\colon \{0,1\}^n \to \{0,1\}$ that interprets its input as a pair of $\ell \times \ell$ boolean matrices (A, B), such that $f(A, B) = 1$ iff A has an all-1 row and every row of B is at most half 1's, and $f(A, B) = 0$ iff B has an all-1 row and every row of A is at most half 1's. Note that $f \in \mathsf{MA}^{\mathsf{dt}} \cap \mathsf{coMA}^{\mathsf{dt}}$ since an MA-type decision tree can guess a row in A and check that a random bit from that row is 1, and a coMA-type decision tree can guess a row in B and check that a random bit from that row is 1. This upper bound lifts to $f \circ \textsc{Ind}_m^n \in \mathsf{MA}^{\mathsf{cc}} \cap \mathsf{coMA}^{\mathsf{cc}}$. We now show that $f \notin \mathsf{CautiousBPP}^{\mathsf{NP[1]dt}}$ which, by Theorems 1 and 2, implies that $f \circ \textsc{Ind}_m^n \notin \mathsf{ZPP}^{\mathsf{NP[1]cc}}$. This will yield Theorem 3.(i).

By Lemma 13, it suffices to show that (1) f has no monochromatic conjunction of width $\leq \ell/2$, and (2) $\mathsf{BPP}^{\mathsf{dt}}(f) \geq \Omega(\ell)$.

To see (1), consider any conjunction C of width $\leq \ell/2$: Since it does not touch every row of A, and it touches at most half the bits in each row of B, we can construct a 1-input accepted by C by putting all 1's in an untouched row of A, and filling the rest of the matrix entries with 0's (except those whose value is determined by C accepting). Similarly, there must exist a 0-input accepted by C. Thus C is not monochromatic.

For (2), it suffices to show that every cost-$o(\ell)$ deterministic decision tree succeeds with probability $< 3/4$ over the input distribution obtained by filling a uniformly random one of the 2ℓ rows with 1's (and letting all other entries of (A, B) be 0's). If the decision tree accepts after seeing only 0's, then conditioned on a random 0-input it continues to accept (and hence err) with probability $\geq 1 - o(1)$ (since the all-0's path of the decision tree only touches a $o(1)$ fraction of the rows). Similarly, if it rejects after seeing only 0's, then conditioned on a random 1-input it continues to reject (and hence err) with probability $\geq 1 - o(1)$. In either case, it errs with probability $\geq 1/2 - o(1)$ over an unconditioned random input.

5.2 $\mathsf{US} \cap \mathsf{coUS}$

We now prove Theorem 3.(ii). Let weight$(\cdot)$ refer to Hamming weight. For even n, define the partial function $f\colon \{0,1\}^n \to \{0,1\}$ that interprets its input as $(a, b) \in \{0,1\}^{n/2} \times \{0,1\}^{n/2}$, such that $f(a, b) = 1$ iff weight$(a) = 1$ and weight$(b) \in \{0, 2\}$, and $f(a, b) = 0$ iff weight$(b) = 1$ and weight$(a) \in \{0, 2\}$. Note that $f \in \mathsf{US}^{\mathsf{dt}} \cap \mathsf{coUS}^{\mathsf{dt}}$ since a US-type decision tree can guess the location of a 1 in a, and a coUS-type decision tree can guess the location of a 1 in b. This upper bound lifts to $f \circ \textsc{Ind}_m^n \in \mathsf{US}^{\mathsf{cc}} \cap \mathsf{coUS}^{\mathsf{cc}}$. We now show that $f \notin \mathsf{CautiousBPP}^{\mathsf{NP[1]dt}}$ which, by Theorems 1 and 2, implies that $f \circ \textsc{Ind}_m^n \notin \mathsf{ZPP}^{\mathsf{NP[1]cc}}$. This will yield Theorem 3.(ii).

Note that Lemma 13 cannot help us here, since this f does have small monochromatic conjunctions (e.g., a conjunction with two positive literals from a is 0-monochromatic), so we devise a different technique. In fact, we show something stronger than $f \notin \mathsf{CautiousBPP}^{\mathsf{NP[1]dt}}$. Define $\mathsf{BPP}^{\mathsf{NP[1]}}$ in the natural way (two-sided error, and allowed to err after an NP oracle query is made), and notice that the class may depend on the exact choice of success probability (since the standard method of amplification involves multiple independent trials, which would increase the number of NP oracle queries). Let us use $\mathsf{BPP}_p^{\mathsf{NP[1]}}$ to indicate that the success probability must be $\geq p$ on each input. As $\mathsf{CautiousBPP}^{\mathsf{NP[1]}}$ can be efficiently amplified (see the end of Section 2.2), the following lemma implies that $f \notin \mathsf{CautiousBPP}^{\mathsf{NP[1]dt}}$.

▶ **Lemma 14.** *For every constant $\varepsilon > 0$, $\mathsf{BPP}_{3/4+\varepsilon}^{\mathsf{NP[1]dt}}(f) \geq \Omega(n)$.*

Proof. It suffices to show that every cost-$o(n)$ $\mathsf{P}^{\mathsf{NP[1]}}$-type decision tree succeeds with probability $\leq 3/4 + o(1)$ over the uniform distribution on valid inputs to f. Let v be the leaf reached after seeing only 0's, and say v is labeled with DNF D and function *out*: $\{0,1\} \to \{0,1\}$. Assume *out*$(1) = 1$ (the case *out*$(1) = 0$ is argued similarly). Consider the joint random variables $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{a}', \boldsymbol{b}'$ where $\boldsymbol{a}$ has a unique 1 in a random position, $\boldsymbol{b}$ is all 0's, $\boldsymbol{a}'$ is obtained from $\boldsymbol{a}$ by flipping a random 0 to 1, and $\boldsymbol{b}'$ is obtained from $\boldsymbol{b}$ by flipping a random 0 to 1. Note that $(\boldsymbol{a}, \boldsymbol{b})$ is the input distribution conditioned on $\text{weight}(a) = 1$ and $\text{weight}(b) = 0$, and $(\boldsymbol{a}', \boldsymbol{b}')$ is the input distribution conditioned on $\text{weight}(a) = 2$ and $\text{weight}(b) = 1$. We have $\mathbb{P}[(\boldsymbol{a}, \boldsymbol{b}) \text{ reaches } v] \geq 1 - o(1)$ and thus $\mathbb{P}[(\boldsymbol{a}, \boldsymbol{b}) \text{ reaches } v \text{ and is accepted}] \geq \mathbb{P}[(\boldsymbol{a}, \boldsymbol{b}) \text{ is accepted}] - o(1)$. Also, conditioned on any outcome of $(\boldsymbol{a}, \boldsymbol{b})$ that reaches v and is accepted, with probability $\geq 1 - o(1)$ the two flipped bits are not among those read along the path to v and not among those read by an arbitrarily chosen term of D that accepts $(\boldsymbol{a}, \boldsymbol{b})$, in which case $(\boldsymbol{a}', \boldsymbol{b}')$ also reaches v and is accepted. Thus, $\mathbb{P}\big[(\boldsymbol{a}', \boldsymbol{b}') \text{ reaches } v \text{ and is accepted} \,\big|\, (\boldsymbol{a}, \boldsymbol{b}) \text{ reaches } v \text{ and is accepted}\big] \geq 1 - o(1)$. Combining these, we get

$$\begin{aligned}
& \mathbb{P}[(\boldsymbol{a}', \boldsymbol{b}') \text{ is accepted}] \\
\geq\ & \mathbb{P}\big[(\boldsymbol{a}', \boldsymbol{b}') \text{ and } (\boldsymbol{a}, \boldsymbol{b}) \text{ both reach } v \text{ and are accepted}\big] \\
=\ & \mathbb{P}\big[(\boldsymbol{a}', \boldsymbol{b}') \text{ reaches } v \text{ and is accepted} \,\big|\, (\boldsymbol{a}, \boldsymbol{b}) \text{ reaches } v \text{ and is accepted}\big] \\
& \cdot \mathbb{P}[(\boldsymbol{a}, \boldsymbol{b}) \text{ reaches } v \text{ and is accepted}] \\
\geq\ & (1 - o(1)) \cdot (\mathbb{P}[(\boldsymbol{a}, \boldsymbol{b}) \text{ is accepted}] - o(1)) \\
=\ & \mathbb{P}[(\boldsymbol{a}, \boldsymbol{b}) \text{ is accepted}] - o(1).
\end{aligned}$$

Thus, under the uniform distribution on valid inputs to f,

$$\begin{aligned}
\mathbb{P}[\text{err}] \ &\geq\ \mathbb{P}\big[\text{err} \,\big|\, \text{weight}(a) = 1 \text{ and weight}(b) = 0\big]/4 \\
&\quad + \mathbb{P}\big[\text{err} \,\big|\, \text{weight}(a) = 2 \text{ and weight}(b) = 1\big]/4 \\
&=\ \big(\mathbb{P}[(\boldsymbol{a}, \boldsymbol{b}) \text{ is rejected}] + \mathbb{P}[(\boldsymbol{a}', \boldsymbol{b}') \text{ is accepted}]\big)/4 \\
&=\ \big(1 - (\mathbb{P}[(\boldsymbol{a}, \boldsymbol{b}) \text{ is accepted}] - \mathbb{P}[(\boldsymbol{a}', \boldsymbol{b}') \text{ is accepted}])\big)/4 \ \geq\ (1 - o(1))/4. \qquad \blacktriangleleft
\end{aligned}$$

We complement Lemma 14 by noting that $\mathsf{BPP}^{\mathsf{NP}[1]\mathsf{dt}}_{3/4}(f) \leq 2$: With probability 1/4 each:

- accept iff $\text{weight}(a) \leq 1$,
- accept iff $\text{weight}(a) \geq 1$,
- reject iff $\text{weight}(b) \leq 1$,
- reject iff $\text{weight}(b) \geq 1$.

Hence $\mathsf{BPP}^{\mathsf{NP}[1]\mathsf{dt}}_{3/4} \not\subseteq \mathsf{BPP}^{\mathsf{NP}[1]\mathsf{dt}}_{3/4+\varepsilon}$, which implies that $\mathsf{BPP}^{\mathsf{NP}[1]}_{3/4} \not\subseteq \mathsf{BPP}^{\mathsf{NP}[1]}_{3/4+\varepsilon}$ in a relativized world. Thus, unlike $\mathsf{ZPP}^{\mathsf{NP}[1]}$, $\mathsf{BPP}^{\mathsf{NP}[1]}$ is not generally amenable to efficient amplification; this phenomenon has subsequently been fully explored in [41].

References

1 Scott Aaronson and Andris Ambainis. Forrelation: A Problem That Optimally Separates Quantum from Classical Computing. In *Proceedings of the 47th Symposium on Theory of Computing (STOC)*, pages 307–316. ACM, 2015. doi:10.1145/2746539.2746547.

2 Scott Aaronson, Shalev Ben-David, and Robin Kothari. Separations in Query Complexity Using Cheat Sheets. In *Proceedings of the 48th Symposium on Theory of Computing (STOC)*, pages 863–876. ACM, 2016. doi:10.1145/2897518.2897644.

3 Andris Ambainis, Martins Kokainis, and Robin Kothari. Nearly Optimal Separations Between Communication (Or Query) Complexity And Partitions. In *Proceedings of the 31st Computational Complexity Conference (CCC)*, pages 4:1–4:14. Schloss Dagstuhl, 2016. doi:10.4230/LIPIcs.CCC.2016.4.

4 Anurag Anshu, Aleksandrs Belovs, Shalev Ben-David, Mika Göös, Rahul Jain, Robin Kothari, Troy Lee, and Miklos Santha. Separations in Communication Complexity Using Cheat Sheets And Information Complexity. In *Proceedings of the 57th Symposium on Foundations of Computer Science (FOCS)*, pages 555–564. IEEE, 2016. doi:10.1109/FOCS.2016.66.

5 Yakov Babichenko and Aviad Rubinstein. Communication Complexity of Approximate Nash Equilibria. In *Proceedings of the 49th Symposium on Theory of Computing (STOC)*, pages 878–889. ACM, 2017. doi:10.1145/3055399.3055407.

6 Shalev Ben-David, Pooya Hatami, and Avishay Tal. Low-Sensitivity Functions from Unambiguous Certificates. In *Proceedings of the 8th Innovations in Theoretical Computer Science Conference (ITCS)*, pages 28:1–28:23. Schloss Dagstuhl, 2017. doi:10.4230/LIPIcs.ITCS.2017.28.

7 Maria Luisa Bonet, Juan Luis Esteban, Nicola Galesi, and Jan Johannsen. On the Relative Complexity of Resolution Refinements and Cutting Planes Proof Systems. *SIAM Journal on Computing*, 30(5):1462–1484, 2000. doi:10.1137/S0097539799352474.

8 Adam Bouland, Lijie Chen, Dhiraj Holden, Justin Thaler, and Prashant Vasudevan. On The Power of Statistical Zero Knowledge. In *Proceedings of the 58th Symposium on Foundations of Computer Science (FOCS)*, pages 708–719. IEEE, 2017. doi:10.1109/FOCS.2017.71.

9 Harry Buhrman and Ronald de Wolf. Complexity Measures and Decision Tree Complexity: A Survey. *Theoretical Computer Science*, 288(1):21–43, 2002. doi:10.1016/S0304-3975(01)00144-X.

10 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.

11 Jin-Yi Cai and Venkatesan Chakaravarthy. On Zero Error Algorithms Having Oracle Access to One Query. *Journal of Combinatorial Optimization*, 11(2):189–202, 2006. doi:10.1007/s10878-006-7130-0.

12 Siu On Chan, James Lee, Prasad Raghavendra, and David Steurer. Approximate Constraint Satisfaction Requires Large LP Relaxations. *Journal of the ACM*, 63(4):34:1–34:22, 2016. doi:10.1145/2811255.

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

14 Susanna de Rezende, Jakob Nordström, and Marc Vinyals. How Limited Interaction Hinders Real Communication (and What It Means for Proof and Circuit Complexity). In *Proceedings of the 57th Symposium on Foundations of Computer Science (FOCS)*, pages 295–304. IEEE, 2016. doi:10.1109/FOCS.2016.40.

15 Mika Göös. Lower Bounds for Clique vs. Independent Set. In *Proceedings of the 56th Symposium on Foundations of Computer Science (FOCS)*, pages 1066–1076. IEEE, 2015. doi:10.1109/FOCS.2015.69.

16 Mika Göös, T.S. Jayram, Toniann Pitassi, and Thomas Watson. Randomized Communication vs. Partition Number. In *Proceedings of the 44th International Colloquium on Automata, Languages, and Programming (ICALP)*, pages 52:1–52:15. Schloss Dagstuhl, 2017. doi:10.4230/LIPIcs.ICALP.2017.52.

17 Mika Göös, Pritish Kamath, Toniann Pitassi, and Thomas Watson. Query-to-Communication Lifting for P$^{\mathrm{NP}}$. In *Proceedings of the 32nd Computational Complexity Conference (CCC)*, pages 12:1–12:16. Schloss Dagstuhl, 2017. doi:10.4230/LIPIcs.CCC.2017.12.

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

19 Mika Göös and Toniann Pitassi. Communication Lower Bounds via Critical Block Sensitivity. In *Proceedings of the 46th Symposium on Theory of Computing (STOC)*, pages 847–856. ACM, 2014. doi:10.1145/2591796.2591838.

20 Mika Göös, Toniann Pitassi, and Thomas Watson. Deterministic Communication vs. Partition Number. In *Proceedings of the 56th Symposium on Foundations of Computer Science (FOCS)*, pages 1077–1088. IEEE, 2015. doi:10.1109/FOCS.2015.70.

21 Mika Göös, Toniann Pitassi, and Thomas Watson. The Landscape of Communication Complexity Classes. In *Proceedings of the 43rd International Colloquium on Automata, Languages, and Programming (ICALP)*, pages 86:1–86:15. Schloss Dagstuhl, 2016. doi:10.4230/LIPIcs.ICALP.2016.86.

22 Mika Göös, Toniann Pitassi, and Thomas Watson. Query-to-Communication Lifting for BPP. In *Proceedings of the 58th Symposium on Foundations of Computer Science (FOCS)*, pages 132–143. IEEE, 2017. doi:10.1109/FOCS.2017.21.

23 Mika Göös and Thomas Watson. Communication Complexity of Set-Disjointness for All Probabilities. *Theory of Computing*, 12(1):1–23, 2016. Special issue for selected papers from APPROX–RANDOM 2014. doi:10.4086/toc.2016.v012a009.

24 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.

25 Jan Johannsen. Depth Lower Bounds for Monotone Semi-Unbounded Fan-In Circuits. *RAIRO - Theoretical Informatics and Applications*, 35:277–286, 2001. doi:10.1051/ita:2001120.

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

27 Hartmut Klauck. Lower Bounds for Quantum Communication Complexity. *SIAM Journal on Computing*, 37(1):20–46, 2007. doi:10.1137/S0097539702405620.

28 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.

29 Pravesh Kothari, Raghu Meka, and Prasad Raghavendra. Approximating Rectangles by Juntas and Weakly-Exponential Lower Bounds for LP Relaxations of CSPs. In *Proceedings of the 49th Symposium on Theory of Computing (STOC)*, pages 590–603. ACM, 2017. doi:10.1145/3055399.3055438.

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

31 James Lee, Prasad Raghavendra, and David Steurer. Lower Bounds on the Size of Semidefinite Programming Relaxations. In *Proceedings of the 47th Symposium on Theory of Computing (STOC)*, pages 567–576. ACM, 2015. doi:10.1145/2746539.2746599.

32 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.

33 Toniann Pitassi and Robert Robere. Strongly Exponential Lower Bounds for Monotone Computation. In *Proceedings of the 49th Symposium on Theory of Computing (STOC)*, pages 1246–1255. ACM, 2017. doi:10.1145/3055399.3055478.

34 Anup Rao and Amir Yehudayoff. *Communication Complexity*. In preparation, 2017.

35 Ran Raz and Pierre McKenzie. Separation of the Monotone NC Hierarchy. *Combinatorica*, 19(3):403–435, 1999. doi:10.1007/s004930050062.

36 Robert Robere, Toniann Pitassi, Benjamin Rossman, and Stephen Cook. Exponential Lower Bounds for Monotone Span Programs. In *Proceedings of the 57th Symposium on Foundations of Computer Science (FOCS)*, pages 406–415. IEEE, 2016. doi:10.1109/FOCS.2016.51.

37 Tim Roughgarden and Omri Weinstein. On the Communication Complexity of Approximate Fixed Points. In *Proceedings of the 57th Symposium on Foundations of Computer Science (FOCS)*, pages 229–238. IEEE, 2016. doi:10.1109/FOCS.2016.32.

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

39 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.

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

41 Thomas Watson. Amplification with One NP Oracle Query. Technical Report TR18-058, Electronic Colloquium on Computational Complexity (ECCC), 2018. URL: https://eccc.weizmann.ac.il/report/2018/058.