58th Annual IEEE Symposium on Foundations of Computer Science

On the Power of Statistical Zero Knowledge

Adam Bouland*, Lijie Chen†, Dhiraj Holden*, Justin Thaler‡ and Prashant Nalini Vasudevan†
*Computer Science and Artificial Intelligence Laboratory, MIT, Cambridge, MA, USA
†Institute for Interdisciplinary Information Sciences, Tsinghua University, Beijing, China
‡Department of Computer Science, Georgetown University, Washington, DC USA

***Abstract*—We examine the power of statistical zero knowledge proofs (captured by the complexity class SZK) and their variants. First, we give the strongest known relativized evidence that SZK contains hard problems, by exhibiting an oracle relative to which SZK (indeed, even NISZK) is not contained in the class UPP, containing those problems solvable by randomized algorithms with unbounded error. This answers an open question of Watrous from 2002. Second, we "lift" this oracle separation to the setting of communication complexity, thereby answering a question of Göös et al. (ICALP 2016). Third, we give relativized evidence that *perfect* zero knowledge proofs (captured by the class PZK) are weaker than general zero knowledge proofs. Specifically, we exhibit oracles which separate SZK from PZK, NISZK from NIPZK and PZK from coPZK. The first of these results answers a question raised in 1991 by Aiello and Håstad (Information and Computation), and the second answers a question of Lovett and Zhang (2016). We also describe additional applications of these results outside of structural complexity.**

The technical core of our results is a stronger hardness amplification theorem for approximate degree, which roughly says that composing the gapped-majority function with any function of high approximate degree yields a function with high threshold degree.

***Keywords*-Oracle Separation, Statistical Zero Knowledge proof, Perfect Zero Knowledge Proof, Hardness Amplification**

I. Introduction

Zero knowledge proof systems, first introduced by Goldwasser, Micali and Rackoff [28], have proven central to the study of complexity theory and cryptography. Abstractly, a zero knowledge proof is a form of interactive proof in which the verifier can efficiently simulate the honest prover on "yes" instances. Therefore, the verifier learns nothing other than whether its input is a "yes" or "no" instance.

In this work, we study *statistical* zero knowledge proofs systems. Here, "efficiently simulate" means that the verifier can, by itself, sample from a distribution which is statistically close to the distribution of the transcript of its interaction with the honest prover[1]. The resulting class of decision problems that have statistical zero knowledge proofs is denoted SZK. One can similarly define variants of this class, such as non-interactive statistical zero knowledge (where the proof system is non-interactive, denoted NISZK), or perfect zero knowledge (where the verifier can exactly simulate the honest prover, denoted PZK).

Many problems, some of which are not necessarily in NP, have been shown to admit SZK protocols. These include Graph Non-isomorphism, as well as problems believed to be hard on average, such as Quadratic Residuosity (as well as the closely related discrete logarithm problem), and the Approximate Shortest Vector and Closest Vector problems in lattices [23]–[25], [28], [39]. Although SZK contains problems believed to be hard, it lies very low in the polynomial hierarchy (below $\mathsf{AM} \cap \mathsf{coAM}$), and cannot contain NP-complete problems unless the polynomial hierarchy collapses [8], [10], [22]. Owing in part to its unusual property of containing problems believed to be hard but not NP-complete, SZK has been the subject of intense interest among complexity theorists and cryptographers.

Despite its importance, many basic questions about the hardness of SZK and its variants remain open. Our results in this work can be understood as grouped into three classes, detailed in each of the next three subsections. However, we prove these results via a unified set of techniques.

A. Group 1: Evidence for the Hardness of SZK

Motivation. Several cryptosystems have been based on the believed hardness of problems in SZK, most notably Quadratic Residuosity and the Approximate Shortest Vector and Closest Vector problems mentioned above. If one could solve SZK-hard problems efficiently, it would break these cryptosystems. Hence, a natural task is to show lower bounds demonstrating that problems in SZK cannot be solved easily. For example, one might want to show that quantum computers or other, more powerful models of computation cannot solve SZK-hard problems efficiently.

Of course, proving such results unconditionally is very difficult, because SZK is contained in $\mathsf{AM} \cap \mathsf{coAM}$ [8], [22], so even proving lower bounds against classical algorithms solving SZK-hard problems would require separating P from NP.[2] Therefore, a more reasonable goal has been to create oracles relative to which SZK is not contained in other complexity classes; one can then unconditionally prove

[1] *Computational* zero-knowledge, in which the zero-knowledge condition is that the verifier can sample from a distribution that is *computationally indistinguishable* from the transcript, has also been the subject of intense study. In this work we focus exclusively on statistical zero knowledge.

[2] Since $\mathsf{SZK} \subseteq \mathsf{AM} \cap \mathsf{coAM} \subseteq \mathsf{PH}$, if $\mathsf{P} \neq \mathsf{SZK}$, then $\mathsf{P} \neq \mathsf{PH}$, which in particular implies $\mathsf{P} \neq \mathsf{NP}$.

0272-5428/17 $31.00 © 2017 IEEE
DOI 10.1109/FOCS.2017.71

that "black-box" algorithms from other complexity classes cannot break SZK.

Additional Context. While much progress has been made in this direction (see Section I-F for details), the problem of giving an oracle separation between SZK and PP has been open since it was posed by Watrous in 2002 [1] and additionally mentioned as an open problem in [4]. Here, PP is the set of decision problems decidable in polynomial time by randomized algorithms with unbounded error. Since a PP algorithm can flip polynomially many coins in its decision process, the gap between the acceptance probabilities of yes and no instances can be exponentially small. PP is a very powerful complexity class – it contains NP and coNP (since it is trivially closed under complement) as well as $\mathsf{BPP}_{\mathsf{path}}$. Furthermore, by Toda's theorem [45], P^{PP} contains the entire polynomial hierarchy. Additionally, Aaronson showed PP = PostBQP, the set of problems decidable by quantum algorithms equipped with postselection (the ability to discard all runs of an experiment which do not achieve an exponentially unlikely outcome). As a result, it is difficult to prove lower bounds against PP.

Our Results. We answer Watrous' question by giving an oracle separating SZK from PP. In fact, we prove something significantly stronger: our oracle construction separates NISZK from UPP.[3]

Theorem I.1. *There exists an oracle $\mathcal{O}$ such that* $\mathsf{NISZK}^{\mathcal{O}} \not\subset \mathsf{UPP}^{\mathcal{O}}$.

B. Group 2: Limitations on the Power of Perfect Zero Knowledge

Motivation. Much progress has been made on understanding the relationship between natural variants of SZK [21], [26], [34]–[36]. For example, it is known that SZK = coSZK [36], and if NISZK = coNISZK then SZK = NISZK = coNISZK [26]. Additionally Lovett and Zhang [34] recently gave an oracle separation between NISZK and coNISZK as well as SZK and NISZK. However, many questions remain open, especially regarding the power of *perfect* zero-knowledge proof systems.

Many important SZK protocols, such as the ones for Graph Non-Isomorphism and Quadratic Nonresiduosity, are in fact PZK protocols. This illustrates the power of perfect zero knowledge. In this work, we are primarily concerned with studying the *limitations* of perfect zero knowledge. We are particularly interested in four questions: Does SZK = PZK? What about their non-interactive variants, NISZK and NIPZK? Is PZK closed under complement, the way that SZK is? What about NIPZK? Answering any of these questions in the negative would require showing P ≠ NP,[4] so it is natural to try to exhibit oracles relative to which SZK ≠ PZK, NISZK ≠ NIPZK, PZK ≠ coPZK, and NIPZK ≠ coNIPZK.

Additional Context. In 1991, Aiello and Håstad [7] gave evidence that PZK contains hard problems by creating an oracle relative to which PZK is not contained in BPP. On the other hand, they also gave an oracle that they *conjectured* separates SZK from PZK (but were unable to prove this). Exhibiting such an oracle requires a technique that can tell the difference between zero simulation error (PZK) and simulation to inverse exponential error (SZK), and prior to our work, no such technique was known. The question of whether SZK = PZK has been asked by Goldwasser [27] as well. The analogous question for the non-interactive classes NISZK and NIPZK is also well motivated, and was explicitly asked in recent work of Lovett and Zhang [34].

Determining whether variants of SZK satisfy the same closure properties as SZK is natural as well: indeed, a main result of Lovett and Zhang [34] is an oracle relative to which NISZK ≠ coNISZK.

Our Results. We give oracles separating SZK from PZK, NISZK from NIPZK, PZK from coPZK, and NIPZK from coNIPZK. The first two results answer the aforementioned questions raised by Aiello and Håstad [7] (though our oracle is different from the candidate proposed by Aiello and Håstad), and Lovett and Zhang [34]. Along the way, we show that PZK is contained in PP in a relativizing manner – this is in sharp contrast to SZK (see Theorem I.1).

Theorem I.2. *For any oracle $\mathcal{O}$,* $\mathsf{PZK}^{\mathcal{O}} \subseteq \mathsf{PP}^{\mathcal{O}}$. *In addition, there exist oracles $\mathcal{O}_1$ and $\mathcal{O}_2$ such that* $\mathsf{SZK}^{\mathcal{O}_1} \not\subseteq \mathsf{PZK}^{\mathcal{O}_1}$, $\mathsf{NISZK}^{\mathcal{O}_1} \not\subseteq \mathsf{NIPZK}^{\mathcal{O}_1}$, $\mathsf{PZK}^{\mathcal{O}_2} \not\subseteq \mathsf{coPZK}^{\mathcal{O}_2}$, *and* $\mathsf{NIPZK}^{\mathcal{O}_2} \not\subseteq \mathsf{coNIPZK}^{\mathcal{O}_2}$.

A summary of known relationships between complexity classes in the vicinity of SZK, including the new results established in this work, is provided in Figure 1.

C. Group 3: Communication Complexity

Motivation and Context. Paturi and Simon [38] introduced the model of *unbounded error communication complexity*, captured by the communication complexity class $\mathsf{UPP}^{\mathsf{cc}}$.[5] In this model, two parties with inputs (x, y) execute a randomized communication protocol, and are only required to output $f(x, y)$ with probability strictly better than random guessing. Unbounded error communication protocols are extremely powerful, owing to this weak success criterion.

[3] UPP is traditionally defined as an oracle complexity class, in which machines must output the correct answer with probability strictly greater than $1/2$, and are charged for oracle queries but not for computation time. In this model, the gap between $1/2$ and the probability of outputting the correct answer can be *arbitrarily* (in particular, superexponentially) small.

[4] P = NP implies P = PH, and therefore SZK = P.

[5] As is standard, given a query model C^{dt} (or a communication model C^{cc}), we define a corresponding complexity class, also denoted C^{dt} (or C^{cc}), consisting of all problems that have polylogarithmic cost protocols in the model.

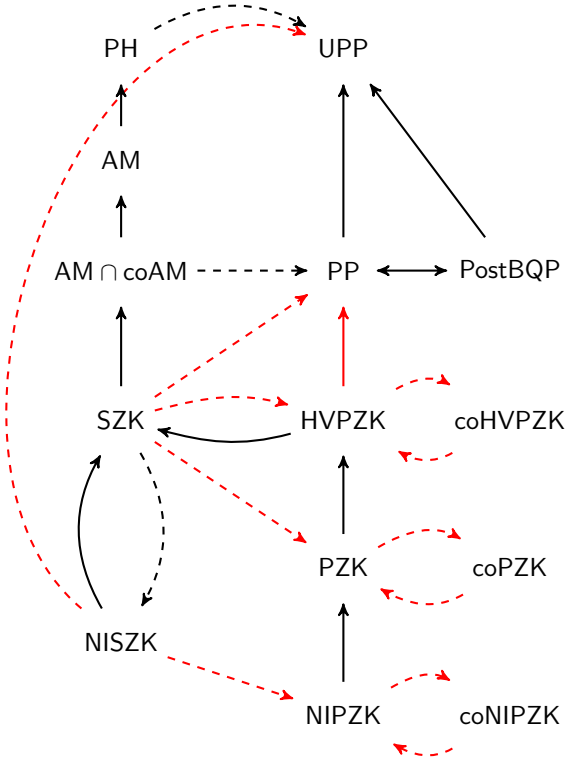

Figure 1. $\mathcal{C}_1 \to \mathcal{C}_2$ indicates $\mathcal{C}_1$ is contained in $\mathcal{C}_2$ with respect to *every* oracle, and $\mathcal{C}_1 \dashrightarrow \mathcal{C}_2$ denotes that there is an oracle $\mathcal{O}$ such that $\mathcal{C}_1^{\mathcal{O}} \not\subseteq \mathcal{C}_2^{\mathcal{O}}$. Red indicates new results. Certain non-inclusions that are depicted are subsumed by other non-inclusions (e.g., NISZK not in UPP subsumes SZK not in PP). We include some redundant arrows to facilitate comparison of our results to prior work.

In fact, $\mathsf{UPP}^{\mathsf{cc}}$ represents the frontier of our understanding of communication complexity: it is the most powerful communication model against which we know how to prove lower bounds. We direct the interested reader to [30] for a thorough overview of communication complexity classes and their known relationships.

What Lies Beyond the Frontier? In an Arthur-Merlin game, a computationally-unbounded prover (Merlin) attempts to convince a computationally-bounded verifier (Arthur) of the value of a given Boolean function on a given input. The communication analogue of Arthur-Merlin games is captured by the communication complexity class $\mathsf{AM}^{\mathsf{cc}}$.

Many works have pointed to $\mathsf{AM}^{\mathsf{cc}}$ as one of the simplest communication models against which we do not know how to prove superlogarithmic lower bounds. Works attempting to address this goal include [16], [29]–[33], [37]. In fact, there are even simpler communication models against which we do not know how to prove lower bounds: it is known that $\mathsf{NISZK}^{\mathsf{cc}} \subseteq \mathsf{SZK}^{\mathsf{cc}} \subseteq \mathsf{AM}^{\mathsf{cc}} \cap \mathsf{coAM}^{\mathsf{cc}} \subseteq \Sigma_2^{\mathsf{cc}}$, and we currently cannot prove lower bounds even against $\mathsf{NISZK}^{\mathsf{cc}}$.

Despite our inability to prove lower bounds against these classes, prior to our work it was possible that $\mathsf{AM}^{\mathsf{cc}}$ is actually contained in $\mathsf{UPP}^{\mathsf{cc}}$ (which, as described above, is a class against which we *can* prove lower bounds). The prior works that had come closest to ruling this out were as follows.

- $\mathsf{AM}^{\mathsf{cc}} \cap \mathsf{coAM}^{\mathsf{cc}} \not\subseteq \mathsf{PP}^{\mathsf{cc}}$. This was established (using a partial function) by Klauck [31], who proved it by combining Vereshchagin's analogous query complexity separation with Sherstov's pattern matrix method [42].
- $\Sigma_2^{\mathsf{cc}} \not\subseteq \mathsf{UPP}^{\mathsf{cc}}$. This result was proved (using a total function) by Razborov and Sherstov [40].

Based on this state of affairs, Göös et al. [30] explicitly posed the problem of showing that $\mathsf{AM}^{\mathsf{cc}} \cap \mathsf{coAM}^{\mathsf{cc}} \not\subseteq \mathsf{UPP}^{\mathsf{cc}}$.

Our Results. In this work, we do even better than showing that $\mathsf{AM}^{\mathsf{cc}} \not\subseteq \mathsf{UPP}^{\mathsf{cc}}$. By "lifting" our oracle separation of NISZK and UPP to the communication setting, we show (using a partial function) that $\mathsf{NISZK}^{\mathsf{cc}} \not\subseteq \mathsf{UPP}^{\mathsf{cc}}$. Hence, if $\mathsf{UPP}^{\mathsf{cc}}$ is taken to represent the frontier of our understanding of communication complexity, our result implies that $\mathsf{NISZK}^{\mathsf{cc}}$ (and hence $\mathsf{AM}^{\mathsf{cc}}$) is truly beyond the frontier. This also answers the question of Göös et al. [30].

Theorem I.3. *There is a (promise) problem in* $\mathsf{NISZK}^{\mathsf{cc}}$ *that is not in* $\mathsf{UPP}^{\mathsf{cc}}$.

D. Other Consequences of Our Results

In addition to the above oracle and communication separations, our results have a number of applications in other areas of theoretical computer science. For example, our results have implications regarding the power of complexity classes capturing the power of quantum computing with "more powerful" modified versions of quantum mechanics [3], [5], imply limitations on the Polarization Lemma of Sahai and Vadhan [41], yield novel lower bounds for certain forms of property testing algorithms, and imply upper bounds for *streaming interactive proofs* [16], [20]. These results are described in detail in the full version of the paper.

E. Overview of Our Techniques

1) Oracle Separation of NISZK *and* UPP *(Proof Overview for Theorem I.1):* To describe our methods, it is helpful to introduce the notions of approximate degree and threshold degree, both of which are measures of Boolean function complexity that capture the difficulty of point-wise approximation by low-degree polynomials. The ε-approximate degree of a function $f\colon \{0,1\}^n \to \{0,1\}$, denoted $\widetilde{\deg}_{\varepsilon}(f)$, is the least degree of a real polynomial that point-wise approximates f to error ε. The threshold degree of f, denoted $\deg_{\pm}(f)$, is the least degree of a real polynomial that agrees in sign[6] with f at all points. It is easy to see that threshold degree is equivalent to the limit of the

[6]By a polynomial p "agreeing in sign" with f, we mean that $p(x) > 0$ whenever $f(x) = 1$, and $p(x) < 0$ when $f(x) = 0$.

approximate degree as the error parameter ε approaches 1/2 from below.

A recent and growing line of work has addressed a variety of open problems in complexity theory by establishing various forms of hardness amplification for approximate degree. Roughly speaking, these results show how to take a function f which is hard to approximate by degree d polynomials to error $\varepsilon = 1/3$, and turn f into a related function F that is hard to approximate by degree d polynomials even when ε is very close to 1/2. In most of these works, F is obtained from f by block-composing f with a "hardness-amplifying function" g. We denote such a block-composition by $g(f)$.

The technical core of our result lies in establishing a new form of hardness amplification for approximate degree. Specifically, let g be the partial function $\mathsf{GapMaj}_n\colon \{0,1\}^n \to \{0,1\}$ (throughout this introduction, whenever necessary, we use subscripts after function names to clarify the number of variables on which the function is defined). Here GapMaj is the gapped majority function, defined, for some $1 \geq \delta > 0.5$, to be 1 if $\geq \delta$ fraction of its inputs are 1, to be 0 if $\geq \delta$ fraction of its inputs are 0, and to be undefined otherwise (in this introduction, we will ignore the precise choice of δ that we use in our formal results).[7]

Theorem I.4. *(Informal) Let $f\colon \{0,1\}^M \to \{0,1\}$. Suppose that $\widetilde{\deg}_{1/3}(f) \geq d$. Define $F\colon \{0,1\}^{n\cdot M} \to \{0,1\}$ via $F = \mathsf{GapMaj}_n(f)$. Then $\deg_{\pm}(F) = \Omega(\min(d,n))$.*

In our main application of Theorem I.4, we apply the theorem to a well-known (partial) function $f = \mathsf{Col}_M$ called the Collision problem. This function is known to have approximate degree $\tilde{\Omega}(M^{1/3})$, so Theorem I.4 implies that $F := \mathsf{GapMaj}_{M^{1/3}}(\mathsf{Col}_M)$ has threshold degree $\tilde{\Omega}(M^{1/3})$. Standard results then imply that the UPP *query complexity* of F is $\tilde{\Omega}(M^{1/3})$ as well. That is, $F \notin \mathsf{UPP}^{\mathsf{dt}}$.

Corollary I.5 (Informal)**.** *Let $m = M^{4/3}$, and define $F\colon \{0,1\}^m \to \{0,1\}$ via $F := \mathsf{GapMaj}_{M^{1/3}}(\mathsf{Col}_M)$. Then $\mathsf{UPP}^{\mathsf{dt}}(F) = \tilde{\Omega}(m^{1/4})$.*

We then show that $\mathsf{GapMaj}_{M^{1/3}}(\mathsf{Col}_M)$ is in $\mathsf{NISZK}^{\mathsf{dt}}$. Hence, we obtain a separation between $\mathsf{NISZK}^{\mathsf{dt}}$ and $\mathsf{UPP}^{\mathsf{dt}}$. The desired oracle separating NISZK from UPP follows via standard methods.

Comparison of Theorem I.4 to Prior Work. The hardness amplification result from prior work that is most closely related to Theorem I.4 is due to Sherstov [43]. Sherstov's result makes use of a notion known as (positive) one-sided approximate degree [12], [43]. Positive one-sided approximate degree is a measure that is intermediate between approximate degree and threshold degree—the positive one-sided approximate degree of f, denoted $\deg^+_\varepsilon(f)$, is always at most as large as the approximate degree of f but can be much smaller, and it is always at least as large as the threshold degree of f but can be much larger (see Section II-A for a formal definition of positive one-sided approximate degree).

Theorem I.6 (Sherstov)**.** *Let $f\colon \{0,1\}^M \to \{0,1\}$. Suppose that $\deg^+_{1/3}(f) \geq d$. Define $F\colon \{0,1\}^{n\cdot M} \to \{0,1\}$ via $F = \mathsf{AND}_n(f)$. Then $\deg_{\pm}(F) = \Omega(\min(d,n))$.*[8]

There are two differences between Theorems I.4 and I.6. The first is that the hardness-amplifier in Theorem I.4 is GapMaj, while in Theorem I.6 it is AND. GapMaj is a "simpler" function than AND in the following sense: block-composing f with GapMaj preserves membership in complexity classes such as $\mathsf{NISZK}^{\mathsf{dt}}$ and $\mathsf{SZK}^{\mathsf{dt}}$; this is not the case for AND, as AND itself is not in $\mathsf{SZK}^{\mathsf{dt}}$. This property is essential for us to obtain threshold degree lower bounds even for functions that are in $\mathsf{NISZK}^{\mathsf{dt}}$.

The second difference is that Theorem I.4 holds under the assumption that $\widetilde{\deg}_{1/3}(f) \geq d$, while Theorem I.6 makes the stronger assumption that $\deg^+_{1/3}(f) \geq d$. While we do not exploit this second difference in our applications, ours is the first form of hardness amplification that works for approximate degree rather than one-sided approximate degree. This has been exploited in subsequent work [15].

Proof Sketch for Theorem I.4. A *dual polynomial* is a dual solution to an appropriate linear program capturing the threshold degree of any function. Specifically, for a (partial) function f defined on a subset of $\{0,1\}^n$, a dual polynomial witnessing the fact that $\widetilde{\deg}_\varepsilon(f) \geq d$ is a function $\psi\colon \{0,1\}^n \to \mathbb{R}$ that satisfies the following three properties.

(a) ψ is uncorrelated with all polynomials p of total degree at most d. That is, for any $p\colon \{0,1\}^n \to \mathbb{R}$ such that $\deg(p) \leq d$, it holds that $\sum_{x\in\{0,1\}^n} \psi(x)\cdot p(x) = 0$. We refer to this property by saying that ψ has *pure high degree* d.

(b) ψ has ℓ_1 norm equal to 1, i.e., $\sum_{x\in\{0,1\}^n} |\psi(x)| = 1$.

(c) ψ has correlation at least ε with f. That is, if D denotes the domain on which f is defined, then $\sum_{x\in D} \psi(x)\cdot f(x) - \sum_{x\in\{0,1\}^n\setminus D} |\psi(x)| > \varepsilon$.

It is not hard to see that a dual witness for the fact that $\deg_{\pm}(f) \geq d$ is a function ψ satisfying Properties (a) and

[7] We clarify that if f is a partial function then $\mathsf{GapMaj}_n(f)$ is technically not a composition of functions, since for some inputs $x = (x_1,\ldots,x_n)$ on which $\mathsf{GapMaj}_n(f)$ is defined, there may be values of i for which x_i is outside of the domain of f. See Section II-C for further discussion of this point.

[8] Sherstov stated his result for $\mathsf{OR}_n(f)$ under the assumption that f has large *negative* one-sided approximate degree. Our statement of Theorem I.6 is the equivalent result under the assumption that f has large positive one-sided approximate degree.

(b) above, that additionally is *perfectly* correlated with f. That is, ψ additionally satisfies

$$\sum_{x \in D} \psi(x) \cdot f(x) - \sum_{x \in \{0,1\}^n \setminus D} |\psi(x)| = 1. \qquad (1)$$

In this case, $\psi \cdot f$ is non-negative, and is referred to as an *orthogonalizing distribution* for f.

We prove Theorem I.4 by constructing an explicit orthogonalizing distribution for $\mathsf{GapMaj}_n(f)$. Specifically, we show how to take a dual polynomial witnessing the fact that $\widetilde{\deg}_{1/3}(f) \geq d$, and turn it into an orthogonalizing distribution witnessing the fact that $\deg_{\pm}(F) = \Omega(\min(d, n))$.

Our construction of an orthogonalizing distribution for $\mathsf{GapMaj}_n(f)$ is inspired by and reminiscent of Sherstov's construction of an orthogonalizing distribution for $\mathsf{AND}_n(f)$ [43], which in turn builds on a dual polynomial for $\mathsf{AND}_n(f)$ constructed by Bun and Thaler [12]. In more detail, Bun and Thaler constructed a dual polynomial ψ_{BT} of pure high degree d that had correlation $1 - 2^{-n}$ with $\mathsf{AND}_n(f)$. Sherstov's dual witness was defined as $\psi_{BT} + \psi_{corr}$, where ψ_{corr} is an *error-correction term* that also has pure high degree $\Omega(d)$. The purpose of ψ_{corr} is to "zero-out" ψ_{BT} at all points where ψ_{BT} differs in sign from f, without affecting the sign of ψ_{BT} on any other inputs.

Naively, one might hope that $\psi_{BT} + \psi_{corr}$ is also a dual witness to the fact that $\deg_{\pm}(\mathsf{GapMaj}_n(f))$ is large. Unfortunately, this is not the case, as it does not satisfy Equation (1) with respect to $\mathsf{GapMaj}_n(f)$. It is helpful to think of this failure as stemming from two issues. First, $\psi_{BT} + \psi_{corr}$ places non-zero weight on many inputs on which $\mathsf{GapMaj}_n(f)$ is undefined (i.e., on inputs for which fewer than δn copies of f evaluate to 1 and fewer than δn copies of f evaluate to 0). Second, there are inputs on which $\mathsf{GapMaj}_n(f)$ is defined, yet $\mathsf{AND}_d(f)$ does not agree with $\mathsf{GapMaj}_n(f)$.

To address both of these issues, we add a *different* error-correction term ψ'_{corr} of pure high degree $\tilde{\Omega}(\min(n, d))$ to ψ_{BT}. Our correction term does not just zero out the value of ψ_{BT} on inputs on which it disagrees in sign with $\mathsf{AND}_n(f)$, but also zeros it out on inputs for which $\mathsf{GapMaj}_n(f)$ is undefined, and on inputs on which $\mathsf{AND}_n(f)$ does not agree with $\mathsf{GapMaj}_n(f)$.

Moreover, we show that adding ψ'_{corr} does not affect the sign of ψ_{BT} on other inputs – achieving this requires some new ideas in both the definition ψ'_{corr} and its analysis. Putting everything together, we obtain a dual witness $\psi_{BT} + \psi'_{corr}$ showing that $\deg_{\pm}(\mathsf{GapMaj}_n(f)) = \Omega(\min(n, d))$.

2) Limitations on the Power of Perfect Zero Knowledge (Proof Overview For Theorem I.2): We begin the proof of Theorem I.2 by showing that HVPZK (*honest verifier* perfect zero knowledge) is contained in PP in a relativizing manner. Since the inclusions PP $\subseteq$ UPP, NIPZK $\subseteq$ HVPZK, PZK $\subseteq$ HVPZK, and NISZK $\subseteq$ SZK hold with respect to any oracle, this means that our oracle separating NISZK from UPP (Theorem I.1) also separates SZK from PZK and NISZK from NIPZK.

We then turn to showing that PZK and NIPZK are not closed under complement with respect to some oracle. Since the proofs are similar, we focus on the case of PZK in this overview.

Since both PZK and coPZK are contained in PP with respect to any oracle, our oracle separation of NISZK from PP (Theorem I.1) does not imply an oracle relative to which PZK $\neq$ coPZK. Instead, to obtain this result we prove a new amplification theorem for one-sided approximate degree. Using similar techniques as Theorem I.4, we show that if f has high positive one-sided approximate degree, then block-composing f with the gapped AND function yields a function with high threshold degree. Here GapAND is partial function that outputs 1 if all inputs are 1, outputs 0 if at least a δ fraction of inputs are 0, and is undefined otherwise.

Theorem I.7. *(Informal) Let* $f\colon \{0,1\}^M \to \{0,1\}$. *Suppose that* $\deg^+_{1/3}(f) \geq d$. *Then* $\deg_{\pm}(\mathsf{GapAND}_n(f)) = \Omega(\min(d, n))$.

We then show that (a) $\mathsf{PZK}^{\mathsf{dt}}$ is closed under composition with GapAND and (b) there is a function f in $\mathsf{PZK}^{\mathsf{dt}}$ whose *complement* $\bar{f}$ has high positive one-sided approximate degree. If $\mathsf{PZK}^{\mathsf{dt}}$ were closed under complement, then $\bar{f}$ would be in $\mathsf{PZK}^{\mathsf{dt}}$. By amplifying the hardness of $\bar{f}$ using Theorem I.7, we obtain a problem that is still in $\mathsf{PZK}^{\mathsf{dt}}$ (this holds by property (a)) yet outside of $\mathsf{PP}^{\mathsf{dt}}$ (this holds by property (b), together with Theorem I.7). This is easily seen to contradict the fact PZK is in PP relative to all oracles. Hence, $\bar{f}$ is a function in $\mathsf{coPZK}^{\mathsf{dt}}$ that is not in $\mathsf{PZK}^{\mathsf{dt}}$, and standard techniques translate this fact into an oracle separating coPZK from PZK. We provide details of these results in the full version of the paper.

3) Lifting to Communication Complexity: Proof Overview For Theorem I.3: To extend our separation between NISZK and UPP to the world of communication complexity, we build on recently developed methods of Bun and Thaler [14], who themselves used and generalized the breakthrough work of Razborov and Sherstov [40]. Razborov and Sherstov showed that if F has high threshold degree and this is witnessed by an orthogonalizing distribution that satisfies an additional smoothness condition, then F can be transformed into a related function F' that has high $\mathsf{UPP}^{\mathsf{cc}}$ complexity (specifically, F' is obtained from F via the *pattern matrix method* introduced in [42]). So in order to turn $\mathsf{GapMaj}(\mathsf{Col})$ into a function with high $\mathsf{UPP}^{\mathsf{cc}}$ complexity, it is enough to give a *smooth* orthogonalizing distribution for F.

Bun and Thaler [14] showed how to take the dual witness Sherstov constructed for $\mathsf{OR}(f)$ in the proof of Theorem I.6 and smooth it out, assuming the inner function f satisfies

some modest additional conditions. Fortunately, a variant of Col called the Permutation Testing Problem (PTP for short) satisfies these additional conditions, and since our construction of an orthogonalizing distribution for GapMaj(PTP) is reminiscent of Sherstov's orthogonalizing distribution for OR(f), we are able to modify the methods of Bun and Thaler to smooth out our dual witness for GapMaj(PTP). Although there are many technical details to work through, adopting the methodology of Bun and Thaler to our setting does not require substantially new ideas, and we do not consider it to be a major technical contribution of this work. Nonetheless, it does require the careful management of various subtleties arising from our use of promise problems as opposed to total Boolean functions, and our final communication lower bound inherits many of the advantages of our Theorem I.4 relative to prior work (such as applying to functions with high approximate degree rather than high one-sided approximate degree).

F. Other Works Giving Evidence for the Hardness of SZK

As mentioned in Section I-B, Aiello and Håstad showed that PZK (and also SZK) is not contained in BPP relative to some oracle [7]. Agrawal et al. later used similar techniques to show that SZK is not contained in the class SRE (which can be viewed as a natural generalization of BPP) relative to some oracle [6]. Aaronson [2] gave an oracle relative to which SZK is not contained in BQP – and therefore quantum computers cannot break SZK-hard cryptosystems in a black-box manner. Building on that work, Aaronson [4] later gave oracle separations against the class QMA (a quantum analogue of NP) and the class $\mathsf{A_0PP}$ (a class intermediate between QMA and PP). Therefore even quantum proofs cannot certify SZK in a black-box manner.[9]

Until recently, the lower bound most closely related to our oracle separation of NISZK and UPP (cf. Theorem I.1) was Vereshchagin's result from 1995, which gave an oracle relative to which $\mathsf{AM} \cap \mathsf{coAM}$ is not contained in PP [46]. Our result is an improvement on Vereshchagin's because the inclusions $\mathsf{NISZK} \subseteq \mathsf{SZK} \subseteq \mathsf{AM} \cap \mathsf{coAM}$ can be proved in a relativizing manner (cf. Figure 1). It also generalizes Aaronson's oracle separation between SZK and $\mathsf{A_0PP}$ [4].

Vereshchagin [46] also reports that Beigel claimed a simple proof of the existence of a function f that is in the query complexity class $\mathsf{AM}^{\mathsf{dt}}$, but is not in the query complexity class $\mathsf{UPP}^{\mathsf{dt}}$. Our result improves on Beigel's in two regards. First, since $\mathsf{NISZK}^{\mathsf{dt}} \subseteq \mathsf{AM}^{\mathsf{dt}}$, separating $\mathsf{NISZK}^{\mathsf{dt}}$ from $\mathsf{UPP}^{\mathsf{dt}}$ is more difficult than separating $\mathsf{AM}^{\mathsf{dt}}$ from $\mathsf{UPP}^{\mathsf{dt}}$. Second, Beigel only claimed a superpolylogarithmic lower bound on the $\mathsf{UPP}^{\mathsf{dt}}$ query complexity of f, while we give a polynomial lower bound.

[9]Note, however, that oracle separations do not necessarily imply the analogous separations in the "real world" – see [9] and [17] for instances in which the situation in the presence of oracles is far from the situation in the real world.

Theorem I.1 also improves on very recent work of Chen [18], [19], which gave a query separation between the classes $\mathsf{P}^{\mathsf{SZK}}$ and PP.

Outline for the Rest of the Paper: In the interest of space, we shall only formally state and prove our hardness amplification results for approximate degree (Theorems I.4 and I.7), as we consider this to be our primary technical contribution. Formal statements and proofs of the other results mentioned in this section may be found in the full version of this paper [11].

In Section II, we define various objects we shall be using in our discussion and state a few relevant facts about them. In Section III, we state and prove our hardness amplification results.

II. Technical Preliminaries

A. Approximate Degree, Threshold Degree, and Their Dual Characterizations

We first recall the definitions of approximate degree, positive one-sided approximate degree, and threshold degree for partial functions.

Definition II.1. Let $D \subseteq \{0,1\}^M$, and let f be a function mapping D to $\{0,1\}$.

- The *approximate degree* of f with approximation constant $0 \leq \varepsilon < 1/2$, denoted $\widetilde{\deg}_\varepsilon(f)$, is the least degree of a real polynomial $p\colon \{0,1\}^M \to \mathbb{R}$ such that $|p(x) - f(x)| \leq \varepsilon$ when $x \in D$, and $-\varepsilon \leq p(x) \leq 1+\varepsilon$ for all $x \notin D$. We refer to such a p as an *approximating polynomial* for f. We use $\widetilde{\deg}(f)$ to denote $\widetilde{\deg}_{1/3}(f)$.
- The *threshold degree* of f, denoted $\deg_\pm(f)$, is the least degree of a real polynomial p such that $p(x) > 0$ when $f(x) = 1$, and $p(x) < 0$ when $f(x) = 0$.
- The *postive one-sided approximate degree* of f with approximation constant $0 \leq \varepsilon < 1/2$, denoted $\deg^+_\varepsilon(f)$, is the least degree of a real polynomial p such that $|p(x) - 1| \leq \varepsilon$ for all $x \in f^{-1}(1)$, and $p(x) \leq \varepsilon$ when $x \in f^{-1}(0)$. We refer to such a p as a *positive one-sided approximating polynomial* for f. We use $\deg^+(f)$ to denote $\deg^+_{1/3}(f)$.

There are clean dual characterizations for each of the three quantities defined in Definition II.1. We state these characterizations without proof, and direct the interested reader to [13], [43], [44] for details.

For a function $\psi\colon \{0,1\}^M \to \mathbb{R}$, define the ℓ_1 norm of ψ by $\|\psi\|_1 = \sum_{x \in \{0,1\}^M} |\psi(x)|$. If the support of a function $\psi\colon \{0,1\}^M \to \mathbb{R}$ is (a subset of) a set $D \subseteq \{0,1\}^M$, we will write $\psi\colon D \to \mathbb{R}$. For functions $f, \psi\colon D \to \mathbb{R}$, denote their inner product by $\langle f, \psi \rangle := \sum_{x \in D} f(x)\psi(x)$. We say that a function $\psi\colon \{0,1\}^M \to \mathbb{R}$ has *pure high degree* d if ψ is uncorrelated with any polynomial $p\colon \{0,1\}^M \to \mathbb{R}$ of total degree at most d, i.e., if $\langle \psi, p \rangle = 0$.

Theorem II.2. *Let $f : D \to \{0,1\}$ with $D \subseteq \{0,1\}^M$ be a partial function and ε be a real number in $[0,1/2)$. $\widetilde{\deg}_\varepsilon(f) > d$ if and only if there is a real function $\psi : \{0,1\}^M \to \mathbb{R}$ such that:*

1) *(Pure high degree): ψ has pure high degree of d.*
2) *(Unit ℓ_1-norm): $\|\psi\|_1 = 1$.*
3) *(Correlation): $\sum_{x\in D} \psi(x)f(x) - \sum_{x\notin D} |\psi(x)| > \varepsilon$.*

Theorem II.3. *Let $f : D \to \{0,1\}$ with $D \subseteq \{0,1\}^M$ be a partial function. $\deg_\pm(f) > d$ if and only if there is a real function $\psi : D \to \mathbb{R}$ such that:*

1) *(Pure high degree): ψ has pure high degree of d.*
2) *(Sign Agreement): $\psi(x) \geq 0$ when $f(x) = 1$, and $\psi(x) \leq 0$ when $f(x) = 0$.*
3) *(Non-triviality): $\|\psi\|_1 > 0$.*

Theorem II.4. *Let $f : D \to \{0,1\}$ with $D \subseteq \{0,1\}^M$ be a partial function and ε be a constant in $[0,1/2)$. $\deg^+_\varepsilon(f) > d$ if and only if there is a real function $\psi : D \to \mathbb{R}$ such that:*

1) *(Pure high degree): ψ has pure high degree of d.*
2) *(Unit ℓ_1-norm): $\|\psi\|_1 = 1$.*
3) *(Correlation): $\langle \psi, f\rangle > \varepsilon$.*
4) *(Negative Sign Agreement): $\psi(x) \leq 0$ whenever $f(x) = 0$.*

B. $\mathsf{PP}^{\mathsf{dt}}$ and $\mathsf{UPP}^{\mathsf{dt}}$

Now we define the two natural analogues of PP complexity in the query model.

Definition II.5. Let $f : D \to \{0,1\}$ with $D \subseteq \{0,1\}^M$ be a partial function. Let $\mathcal{T}$ be a randomized decision tree which computes f with a probability better than $1/2$. Let α be the maximum real number such that

$$\min_{x\in D} \Pr[\mathcal{T} \text{ outputs } f(x) \text{ on input } x] \geq \frac{1}{2} + \alpha.$$

Then we define the PP query cost of $\mathcal{T}$ for f to be $\mathsf{PP}^{\mathsf{dt}}(\mathcal{T}; f) = C(\mathcal{T}; f) + \log_2(1/\alpha)$, where $C(\mathcal{T}; f)$ denotes the maximum number of queries $\mathcal{T}$ incurs on an input in the worst case. We define $\mathsf{UPP}^{\mathsf{dt}}(\mathcal{T}; f) = C(\mathcal{T}; f)$. We define $\mathsf{PP}^{\mathsf{dt}}(f)$ (respectively, $\mathsf{UPP}^{\mathsf{dt}}(f)$) as the minimum of $\mathsf{PP}^{\mathsf{dt}}(\mathcal{T}; f)$ (respectively, $\mathsf{UPP}^{\mathsf{dt}}(\mathcal{T}; f)$) over all $\mathcal{T}$ that computes f with a probability better than $1/2$.

$\mathsf{PP}^{\mathsf{dt}}$ is closely related to approximate degree with error very close to $1/2$. We have the following well-known relationship between them.

Lemma II.6. *Let $f : D \to \{0,1\}$ with $D \subseteq \{0,1\}^M$ be a partial function. Suppose $\widetilde{\deg}_{1/2-2^{-d}}(f) > d$ for some positive integer d. Then $\mathsf{PP}^{\mathsf{dt}}(f) > d/2$.*

Meanwhile, $\mathsf{UPP}^{\mathsf{dt}}$ is exactly characterized by threshold degree.

Lemma II.7. *Let $f : D \to \{0,1\}$ with $D \subseteq \{0,1\}^M$ be a partial function. Then $\mathsf{UPP}^{\mathsf{dt}}(f) = \deg_\pm(f)$.*

C. Gap Majority and Gap AND

In this subsection we introduce transformations of partial functions which will be used in this paper.

Definition II.8. Let $f : D \to \{0,1\}$ with $D \subseteq \{0,1\}^M$ be a partial function and n be a positive integer, $0.5 < \varepsilon \leq 1$ be a real number. We define the gap majority version of f, denoted by $\mathsf{GapMaj}_{n,\varepsilon}(f)$, as follows:

Given an input $x = (x_1, x_2, \ldots, x_n) \in \{0,1\}^{M\cdot n}$, we define $n_{\mathsf{Yes}}(x) := \sum_{i=1}^{n} \mathbb{1}_{x_i\in D \wedge f(x_i)=1}$ and $n_{\mathsf{No}}(x) := \sum_{i=1}^{n} \mathbb{1}_{x_i\in D\wedge f(x_i)=0}$.[10] Then

$$\mathsf{GapMaj}_{n,\varepsilon}(f)(x) = \begin{cases} 1 & \text{when } n_{\mathsf{Yes}}(x) \geq \varepsilon\cdot n \\ 0 & \text{when } n_{\mathsf{No}}(x) \geq \varepsilon\cdot n \\ \text{undefined} & \text{otherwise} \end{cases}$$

For brevity, we will occasionally write $\mathsf{GapMaj}(f)$ when n and ε are clear from context.

We also define the GapAND function. This is a partial function that agrees with the total function AND wherever it is defined.

Definition II.9. Let n be a positive integer, $0 < \varepsilon < 1$ be a constant. We define the Gapped AND function, $\mathsf{GapAND}_{n,\varepsilon} : D \to \{0,1\}$ with $D \subseteq \{0,1\}^n$, as the function that outputs 1 if all inputs are 1; outputs 0 if at least $\varepsilon \cdot n$ inputs are 0; and is undefined otherwise.

For a partial function $f : D \to \{0,1\}$ with $D \subseteq \{0,1\}^M$, we define $\mathsf{GapAND}_{n,\varepsilon}(f)$ to be a true block-composition of partial functions, i.e., $\mathsf{GapAND}_{n,\varepsilon}(f)(x_1, \ldots, x_n) = \mathsf{GapAND}_{n,\varepsilon}(f(x_1), \ldots, f(x_n))$ whenever the right hand side of the equality is defined, and $\mathsf{GapAND}_{n,\varepsilon}(f)$ is undefined otherwise.

Remark II.10. *Note that $\mathsf{GapMaj}_{n,\varepsilon}(f)$ is not technically a block-composition of partial functions, since $\mathsf{GapMaj}_{n,\varepsilon}(f)(x_1, \ldots, x_n)$ is defined even on some inputs for which some $f(x_i)$ is not defined.*

III. Hardness Amplification For Approximate Degree

In this section we prove a novel hardness amplification theorem. Specifically, we show that for any function f with high approximate degree, composing f with GapMaj yields a function with high threshold degree, and hence the resulting function is hard for any UPP algorithm in the query

[10] Here, $\mathbb{1}$ is the indicator function which takes 1 when the boolean expression is true and 0 otherwise.

model. Similarly, we show that if f has high positive one-sided approximate degree, then composing f with GapAND yields a function with high threshold degree.

Note that this hardness amplification theorem is tight, in the sense that if f has low approximate degree, then composing f with GapMaj yields a function that has low UPP query complexity, and the same holds for composing f with GapAND if f has low positive one-sided approximate degree. See the full version for details.

A. Notation

For a partial function f, an integer n and a real $\varepsilon \in (1/2, 1]$, we denote $\mathsf{GapMaj}_{n,\varepsilon}(f)$ by F for convenience, where n and ε will always be clear in the context. We also use $x = (x_1, x_2, \ldots, x_n)$ to denote an input to F, where x_i represents the input to the ith copy of f.

The following simple lemma establishes some basic properties of dual witnesses exhibiting the fact that $\widetilde{\deg}_\varepsilon(f) > d$ or $\deg^+_\varepsilon(f) > d$.

Lemma III.1. *Let $f\colon D \to \{0,1\}$ with $D \subseteq \{0,1\}^M$ be a partial function, ε be a real in $[0, 1/2)$, and d be an integer such that $\widetilde{\deg}_\varepsilon(f) > d$.*

Let $\mu\colon \{0,1\}^M \to \mathbb{R}$ be a dual witness to the fact $\widetilde{\deg}_\varepsilon(f) > d$ as per Theorem II.2. If f satisfies the stronger condition that $\deg^+_\varepsilon(f) > d$, let μ to be a dual witness to the fact that $\deg^+_\varepsilon(f) > d$ as per Theorem II.4.

We further define $\mu_+(x) := \max\{0, \mu(x)\}$ and $\mu_-(x) := -\min\{0, \mu(x)\}$ to be two non-negative real functions on $\{0,1\}^M$, and μ^i_- and μ^i_+ be the restrictions of μ_- and μ_+ on $f^{-1}(i)$ respectively for $i \in \{0,1\}$. Then the following holds:

- *μ_+ and μ_- have disjoint supports.* (2)
- *$\langle \mu_+, p\rangle = \langle \mu_-, p\rangle$ for any polynomial p of degree at most d. Hence, $\|\mu_+\|_1 = \|\mu_-\|_1 = \frac{1}{2}$.* (3)
- *$\|\mu^1_+\|_1 > \varepsilon$ and $\|\mu^0_-\|_1 > \varepsilon$. If $\deg^+_\varepsilon(f) > d$, then $\|\mu^1_+\|_1 = 1/2$.* (4)

The lemma follows directly from Theorem II.2. We provide a proof in the full version of the paper for completeness.

B. Warm Up : A PP Lower Bound

As a warmup, we establish a simpler hardness amplification theorem for $\mathsf{PP}^{\mathsf{dt}}$.

Theorem III.2. *Let $f : D \to \{0,1\}$ with $D \subseteq \{0,1\}^M$ be a partial function, n, d be two positive integers, and $1/2 < \varepsilon < 1$ and $0 < \varepsilon_2 < 1/2$ be two constants such that $2\varepsilon_2 > \varepsilon$. Suppose $\widetilde{\deg}_{\varepsilon_2}(f) > d$. Then*

$$\mathsf{PP}^{\mathsf{dt}}(\mathsf{GapMaj}_{n,\varepsilon}(f)) > \Omega\left\{\min\left(d, (2\varepsilon_2 - \varepsilon)^2 \cdot n\right)\right\}.$$

Proof: For $i \in \{0,1\}$ let $\mu_+, \mu_-, \mu^i_+, \mu^i_-$ be functions whose existence is guaranteed by Lemma III.1, combined with the assumption that $\widetilde{\deg}_{\varepsilon_2}(f) > d$.

In light of Lemma II.6, it suffices to show that $\widetilde{\deg}_{1/2-2^{-T}}(\mathsf{GapMaj}_{n,\varepsilon}(f)) > T$, for $T = \Omega\left\{\min\left(d, (2\varepsilon_2 - \varepsilon)^2 \cdot n\right)\right\}$. We prove this by constructing a dual witness to this fact, as per Theorem II.2.

We first define the following two non-negative functions on $\{0,1\}^{n\cdot M}$:

$$\psi^+(x) := \prod_{i=1}^n \mu_+(x_i) \quad \text{and} \quad \psi^-(x) := \prod_{i=1}^n \mu_-(x_i).$$

Our dual witness ψ is simply their linear combination: $\psi := 2^{n-1} \cdot (\psi^+ - \psi^-)$. We remark that ψ is precisely the function denoted by ψ_{BT} alluded to in Section I-E1. Now we verify that ψ is the dual witness we want.

Proving the ψ has unit ℓ_1-norm. Since μ_+ and μ_- have disjoint supports by Condition (2) of Lemma III.1, so does ψ^+ and ψ^-. Therefore $\|\psi\|_1 = 2^{n-1} \cdot (2^{-n} + 2^{-n}) = 1$.

Proving the ψ has pure high degree d. Let $p\colon \{0,1\}^{n\cdot M} \to \mathbb{R}$ be any monomial of degree at most d, and let $p_i\colon \{0,1\}^M \to \mathbb{R}$ be such that $p(x_1, \ldots, x_n) = \prod_{i=1}^n p_i(x_i)$. Then it holds that $\langle \psi^+, p\rangle = \prod_{i=1}^n \langle \mu_+, p_i\rangle = \prod_{i=1}^n \langle \mu_-, p_i\rangle = \langle \psi^-, p\rangle$, where the second equality holds by Condition (3) of Lemma III.1.

As a polynomial is a sum of monomials, by linearity, it follows that $\langle \psi, p\rangle = \langle \psi^+, p\rangle - \langle \psi^-, p\rangle = 0$ for any polynomial p with degree at most d.

Proving that ψ has high correlation with F. Define $\mathcal{D}_0 := 2 \cdot \mu_-$ and $\mathcal{D}_1 := 2 \cdot \mu_+$. Note μ_+ and μ_- are non-negative functions with norm $1/2$, so $\mathcal{D}_0$ and $\mathcal{D}_1$ can be thought as distributions on $\{0,1\}^M$. We further define distributions $\mathcal{U}_i$ on $\{0,1\}^{n\cdot M}$ for $i \in \{0,1\}$ as $\mathcal{U}_i := \mathcal{D}_i^{\otimes n}$. Observe that $\mathcal{U}_0 = 2^n \cdot \psi^-$ and $\mathcal{U}_1 = 2^n \cdot \psi^+$ as functions.

Then by Condition (4) of Lemma III.1, we have $\Pr_{x\sim\mathcal{D}_1}[f(x) = 1] = 2 \cdot \|\mu^1_+\|_1 > 2\varepsilon_2 > \varepsilon$, and $\Pr_{x\sim\mathcal{D}_0}[f(x) = 0] = 2 \cdot \|\mu^0_-\|_1 > 2\varepsilon_2 > \varepsilon$.

Let D_F denote the domain of F. By the definition of $F = \mathsf{GapMaj}_{n,\varepsilon}(f)$ and a simple Chernoff bound, we have

$$2^n \cdot \sum_{x\in D_F} \psi^+(x) \cdot F(x) = \Pr_{x\sim\mathcal{U}_1}[F(x) = 1] \geq 1 - 2^{-c_1\Delta^2\cdot n}, \tag{5}$$

where c_1 is a universal constant and $\Delta := 2\varepsilon_2 - \varepsilon$. For brevity, let k denote $c_1\Delta^2 \cdot n$.

Since $2^n \cdot \|\psi^+\|_1 = 1$, inequality (5) further implies that $2^n \cdot \sum_{x\notin D_F} \psi^+(x) \leq 2^{-k}$. Similarly, we have $\Pr_{x\sim\mathcal{U}_0}[F(x) = 0] \geq 1 - 2^{-k}$, which implies that $2^n \cdot \sum_{x\notin D_F} \psi^-(x) \leq 2^{-k}$.

Putting everything together, we can calculate the correlation between F and ψ as follows:

$$\begin{aligned}
&\sum_{x\in D_F} F(x)\psi(x) - \sum_{x\notin D_F} |\psi(x)| \\
\geq& 2^{n-1}\cdot \sum_{x\in D_F} \psi^+(x)F(x) - \\
& 2^{n-1}\cdot\left(\sum_{x\notin D_F}\psi^-(x) + \sum_{x\notin D_F}\psi^+(x)\right) \\
\geq& 1/2 - 2^{-k-1} - 2^{-k} \\
>& 1/2 - 2^{-k+1}.
\end{aligned}$$

Setting $T = \min(d, k-1)$, then we can see that ψ is a dual witness for $\widetilde{\deg}_{1-2^{-T}}(\mathsf{GapMaj}_{n,\varepsilon}(f)) > T$. Clearly $T = \Omega\left\{\min\left(d, (2\varepsilon_2-\varepsilon)^2\cdot n\right)\right\}$. Invoking Lemma II.6 completes the proof. ∎

C. The UPP Lower Bound

The dual witness $\psi \sim \psi^+ - \psi^-$ constructed in the previous subsection is not a dual witness for the high threshold degree of $F = \mathsf{GapMaj}_n(f)$ for two reasons: it puts weight on some points outside of the domain of F, and it does not satisfy the sign-agreement condition of Theorem II.3.

In order to obtain a valid dual witness for threshold degree, we add two error correction terms ψ^+_{corr} and ψ^-_{corr} to ψ. The purpose of the error correction terms is to zero out the erroneous values, while simultaneously maintaining the high pure degree property and avoiding changing the sign of ψ on inputs at which it does not agree in sign with F. We achieve this through an error correction lemma that may be of independent interest.

Lemma III.3 (Error Correction Lemma)**.** *Let A be a subset of $\{0,1\}^M$, and φ be a function on $\{0,1\}^M$. Let $\varphi_\circ$ and $\varphi_\times$ be the restrictions of φ on A and $\{0,1\}^M \setminus A$ respectively. That is, $\varphi_\circ(x_i) = \varphi(x_i)$ if $x_i \in A$ and $\varphi_\circ(x_i) = 0$ otherwise, and similarly $\varphi_\times(x_i) = \varphi(x_i)$ if $x_i \notin A$ and $\varphi_\times(x_i) = 0$ otherwise. Define $\psi : \{0,1\}^{n\cdot M} \to \{0,1\}$ as $\psi(x_1, x_2, \ldots, x_n) := \prod_{i=1}^n \varphi(x_i)$, and $n_A(x) := \sum_{i=1}^n \mathbb{1}_{x_i\in A}$.*

Suppose $\alpha = \|\varphi_\times\|_1/\|\varphi_\circ\|_1 < 1/40$, and let $0.5 < \varepsilon < 1$ be a real number and n be a sufficient large integer. Then there exists a function $\psi_{corr}\colon \{0,1\}^{n\cdot M} \to \mathbb{R}$ such that:

- $\psi_{corr}(x) = \psi(x)$, *when* $n_A(x) \leq \varepsilon\cdot n$. (6)
- $|\psi_{corr}(x)| \leq \psi(x)/2$, *when* $n_A(x) > \varepsilon\cdot n$. (7)
- ψ_{corr} *has pure high degree of at least* $(1-(1+10\alpha)\cdot\varepsilon)\cdot n - 4$. (8)

The proof of Lemma III.3 is deferred to the full version of the paper. Here, we show that it implies the desired hardness amplification results.

Theorem III.4 (Formal version of Theorems I.4 and I.7)**.** *Let $f : D \to \{0,1\}$ with $D \subseteq \{0,1\}^M$ be a partial function, n be a sufficiently large integer, d be an integer, and $1/2 < \varepsilon < 1$ and $0.49 < \varepsilon_2 < 1/2$ be two constants. Let $a = \frac{2\varepsilon_2}{1-2\varepsilon_2}$. Then the following holds.*

If $\widetilde{\deg}_{\varepsilon_2}(f) > d$, *then* $\deg_\pm(\mathsf{GapMaj}_{n,\varepsilon}(f)) > \min\left(d, \left(1-\left(1+\frac{10}{a}\right)\cdot\varepsilon\right)\cdot n - 4\right)$.

If $\deg^+_{\varepsilon_2}(f) > d$, *then* $\deg_\pm(\mathsf{GapAND}_{n,\varepsilon}(f)) > \min\left(d, \left(1-\left(1+\frac{10}{a}\right)\cdot\varepsilon\right)\cdot n - 4\right)$.

Proof: We prove both claims in the theorem by exhibiting a single dual solution that witnesses both.

As in the proof of Theorem III.2, for $i \in \{0,1\}$, let $\mu_+, \mu_-, \mu^i_+, \mu^i_-$ denote the functions whose existence is guaranteed by Lemma III.1, combined with the assumption that either $\widetilde{\deg}_\varepsilon(f) > d$ or $\deg^+_\varepsilon(f) > d$. Also as in the proof of Theorem III.2, define the following two non-negative functions on $\{0,1\}^{n\cdot M}$:

$$\psi^+(x) := \prod_{i=1}^n \mu_+(x_i) \quad \text{and} \quad \psi^-(x) := \prod_{i=1}^n \mu_-(x_i).$$

Given an input $x = (x_1, x_2, \ldots, x_n)$, let $n_{\mathsf{Yes}}(x) := \sum_{i=1}^n \mathbb{1}_{f(x_i)=1}$ and $n_{\mathsf{No}}(x) := \sum_{i=1}^n \mathbb{1}_{f(x_i)=0}$ as in Definition II.8. Now apply Lemma III.3 with the following parameters.

- Set $A = f^{-1}(1)$, $\varphi = \mu_+$. Then for α as defined in Lemma III.3, we have $\alpha = \frac{\|\mu_+\|_1 - \|\mu^1_+\|_1}{\|\mu^1_+\|_1} \leq \frac{1-2\varepsilon_2}{2\varepsilon_2} = a^{-1}$ by Conditions (3) and (4) of Lemma III.1. Note that $a^{-1} < 1/40$ by the assumption that $0.49 < \varepsilon_2$. Hence, by Lemma III.3, there exists a function $\psi^+_{corr} : \{0,1\}^{n\cdot M} \to \mathbb{R}$ such that:
  - $\psi^+_{corr}(x) = \psi^+(x)$, for all x such that $n_{\mathsf{Yes}}(x) \leq \varepsilon\cdot n$ (9)
  - $|\psi^+_{corr}(x)| \leq \psi^+(x)/2$, for all x such that $n_{\mathsf{Yes}}(x) > \varepsilon\cdot n$ (10)
  - ψ^+_{corr} has pure high degree at least $\left(1-\left(1+\frac{10}{a}\right)\cdot\varepsilon\right)\cdot n - 4$ (11)
- Similarly, set $A = f^{-1}(0)$, $\varphi = \mu_-$. Again by Lemma III.3, there exists a function $\psi^-_{corr} : \{0,1\}^{n\cdot M} \to \mathbb{R}$ such that:
  - $\psi^-_{corr}(x) = \psi^-(x)$, for all x such that $n_{\mathsf{No}}(x) \leq \varepsilon\cdot n$ (12)

- $|\psi^-_{corr}(x)| \le \psi^-(x)/2$, for all x such that $n_{\mathsf{No}}(x) > \varepsilon \cdot n$ (13)
- ψ^-_{corr} has pure high degree of at least $\left(1 - \left(1 + \frac{10}{a}\right) \cdot \varepsilon\right) \cdot n - 4$ (14)

For convenience, let $N = \left(1 - \left(1 + \frac{10}{a}\right) \cdot \varepsilon\right) \cdot n - 4$. We are ready to construct the dual witness ψ that establishes the claimed threshold degree lower bounds. Define $\psi\colon \{0,1\}^{n \cdot M} \to \mathbb{R}$ by

$$\psi := (\psi^+ - \psi^+_{corr}) - (\psi^- - \psi^-_{corr}).$$

We first establish two properties of ψ.

- When $n_{\mathsf{Yes}}(x) \ge \varepsilon \cdot n$, $\psi(x) = \psi^+(x) - \psi^+_{corr}(x) \ge \psi^+(x)/2 \ge 0$ (15)
- When $n_{\mathsf{No}}(x) \ge \varepsilon \cdot n$, $\psi(x) = -(\psi^-(x) - \psi^-_{corr}(x)) \le -\psi^-(x)/2 \le 0$ (16)

Verifying Condition (15) and (16). To establish that Condition (15) holds, observe that since $n_{\mathsf{Yes}}(x) \ge \varepsilon \cdot n$, and $\varepsilon > 1/2$ by assumption, it follows that $n_{\mathsf{No}}(x) \le (1-\varepsilon)\cdot n \le \varepsilon \cdot n$. This implies that $\psi^-(x) = \psi^-_{corr}(x)$ by Condition (12) and $|\psi^+_{corr}(x)| \le \psi^+(x)/2$ by Condition (10). Then $\psi(x) = \psi^+(x) - \psi^+_{corr}(x) \ge \psi^+(x)/2 \ge 0$, where the last inequality follows from the fact that ψ^+ is non-negative.

Similarly, for Condition (16), as $n_{\mathsf{No}}(x) \ge \varepsilon \cdot n$, it follows that $n_{\mathsf{Yes}}(x) \le (1-\varepsilon) \cdot n \le \varepsilon \cdot n$. This implies that $\psi^+(x) = \psi^+_{corr}(x)$ by Condition (9) and $|\psi^-_{corr}(x)| \le \psi^-(x)/2$ by Condition (13). Note ψ^- is also non-negative. Hence $\psi(x) = -(\psi^-(x) - \psi^-_{corr}(x)) \le -(\psi^-(x)/2) \le 0$.

We now verify that ψ is a dual witness for $\deg_\pm(F) > \min(d, N)$ (recall that F denotes $\mathsf{GapMaj}(f)$).

Analyzing the pure high degree of ψ. Write $\psi := \psi^+ - \psi^- - \psi^+_{corr} + \psi^-_{corr}$. We already established that $\psi^+ - \psi^-$ has pure high degree d in the proof of Theorem III.2, and both ψ^+_{corr} and ψ^+_{corr} have pure high degree at least N (cf. Conditions (11) and (14)). By linearity, ψ itself has pure high degree at least $\min(d, N)$.

Showing that the support of ψ is a subset of the inputs on which F is defined. Let x be an input outside of the domain of F. Then by the definition of GapMaj, it must be the case that both $n_{\mathsf{Yes}}(x)$ and $n_{\mathsf{No}}(x)$ are strictly less than $\varepsilon \cdot n$. This means that $\psi^+(x) = \psi^+_{corr}(x)$ and $\psi^-(x) = \psi^-_{corr}(x)$ by Conditions (9) and (12), and hence $\psi(x) = 0$. Therefore, the support of ψ is a subset of the domain of F.

Showing that ψ agrees in sign with F. When $F(x) = 1$, by the definition of GapMaj, we have $n_{\mathsf{Yes}}(x) \ge \varepsilon \cdot n$. Then $\psi(x) \ge 0$ follows directly from Condition (15). Similarly, when $F(x) = 0$, we have $n_{\mathsf{No}}(x) \ge \varepsilon \cdot n$ and $\psi(x) \le 0$ by Condition (16). Therefore, ψ agrees in sign with F.

Showing that ψ is non-trivial. Pick an input x_0 to f such that $\mu^1_+(x_0) > 0$, and let $x = (x_0, x_0, \ldots, x_0)$. Then we have $f(x_0) = 1$ and $n_{\mathsf{Yes}}(x) = n \ge \varepsilon \cdot n$. Therefore, $\psi(x) = \psi^+(x) - \psi^+_{corr}(x) \ge \psi^+(x)/2 = (\mu^1_+(x_0))^n/2 > 0$ by Condition (15). So ψ is non-trivial.

Putting everything together and invoking Theorem II.3 proves the first claim of Theorem III.4.

Showing ψ is also a dual witness for $\mathsf{GapAND}_{n,\varepsilon}(f)$. Now we show that, when $\deg^+_{\varepsilon_2}(f) > d$, the same function ψ is also a dual witness for $\deg_\pm(\mathsf{GapAND}_{n,\varepsilon}(f)) > \min(d, N)$.

We already proved that the pure high degree of ψ is as claimed, and that it is non-trivial. So it remains to verify ψ only puts weight in the domain of $\mathsf{GapAND}_{n,\varepsilon}(f)$, and that ψ agrees in sign with $\mathsf{GapAND}_{n,\varepsilon}(f)$.

By Condition (4) of Lemma III.1, we have $|\mu^1_+| = |\mu_+| = \frac{1}{2}$, which means μ_+ only puts weight inputs in $f^{-1}(1)$. So ψ^+ only takes non-zero values when $n_{\mathsf{Yes}}(x) = n$. Also, note that when $n_{\mathsf{No}}(x) \le \varepsilon \cdot n$, we have $\psi^-(x) = \psi^-_{corr}(x)$ by Condition (12). Therefore, ψ only puts weight on inputs when $n_{\mathsf{Yes}}(x) = n$ or $n_{\mathsf{No}}(x) > \varepsilon \cdot n$. All such inputs are in the domain of $\mathsf{GapAND}_{n,\varepsilon}(f)$.

Finally, we verify that ψ agrees in sign with $\mathsf{GapAND}_{n,\varepsilon}(f)$. When $\mathsf{GapAND}_{n,\varepsilon}(f)(x) = 1$, we have $n_{\mathsf{Yes}}(x) = n \ge \varepsilon \cdot n$, hence $\psi(x) \ge 0$ by Condition (15). When $\mathsf{GapAND}_{n,\varepsilon}(f)(x) = 0$, we have $n_{\mathsf{No}}(x) \ge \varepsilon \cdot n$, so $\psi(x) \le 0$ follows immediately from Condition (16). Applying Theorem II.3 again, this completes the proof for the second claim of Theorem III.4. ∎

IV. Open Problems

Our works leaves a number of open related problems. As one example, we have shown that the function $\mathsf{GapMaj}(f)$ is hard for $\mathsf{UPP}^{\mathsf{dt}}$, for any function f of high approximate degree, and that $\mathsf{GapAND}(f)$ is hard for $\mathsf{UPP}^{\mathsf{dt}}$, for any function of high positive one-sided approximate degree. Can one extend this work to characterize when $f \circ g$ is hard for $\mathsf{UPP}^{\mathsf{dt}}$, based on some properties of f and g? We conjecture that the $\mathsf{UPP}^{\mathsf{dt}}$ complexity of $\mathsf{GapMaj}(f)$ (respectively, $\mathsf{GapAND}(f)$) is characterized by the *rational approximate degree* of f (respectively, positive one-sided approximate degree of f). Such a result would complement the characterization of the threshold degree of $\mathsf{AND}(f)$ in terms of positive one-sided rational approximate degree given in [43].

However, the main open question highlighted by our work is to break through the UPP frontier in communication complexity. We formalize this question via the following challenge: prove any superlogarithmic lower bound for an explicit problem in a natural communication model that cannot be efficiently simulated by $\mathsf{UPP}^{\mathsf{cc}}$. Our work shows that any communication model capable of efficiently computing the pattern matrix of $\mathsf{GapMaj}(\mathsf{PTP})$ is a candidate

for achieving this goal. Thomas Watson has suggested the following as perhaps the simplest candidate: consider the $\mathsf{NISZK}^{\mathsf{cc}}$ model, but restricted to be one-way, in the sense that neither Merlin nor Bob can talk to Alice. This model effectively combines the key features of the $\mathsf{NISZK}^{\mathsf{cc}}$ and $\mathbf{OIP}_{+}^{[2]}$ (cf. [16]) communication models. There is a logarithmic cost "one-way NISZK" protocol for the pattern matrix of $\mathsf{GapMaj}(\mathsf{PTP})$, so this model cannot be efficiently simulated by $\mathsf{UPP}^{\mathsf{cc}}$. Curiously, despite the ability of this model to compute functions outside of $\mathsf{UPP}^{\mathsf{cc}}$, to the best of our knowledge it is possible that even the INDEX function requires polynomial cost in this model. Note that while Chakrabarti et al. [16] gave an efficient $\mathbf{OIP}_{+}^{[2]}$ communication protocol for INDEX, their protocol is not zero-knowledge.

Acknowledgments

We thank Scott Aaronson, Jayadev Acharya, Shalev Ben-David, Clément Canonne, Oded Goldreich, Mika Göös, Gautam Kamath, Robin Kothari, Tomoyuki Morimae, Harumichi Nishimura, Ron Rothblum, Mike Saks, Salil Vadhan and Thomas Watson for helpful discussions. Adam Bouland was supported in part by the NSF Graduate Research Fellowship under grant no. 1122374 and by the NSF Alan T. Waterman award under grant no. 1249349. Lijie Chen was supported in part by the National Basic Research Program of China Grant 2011CBA00300, 2011CBA00301, the National Natural Science Foundation of China Grant 61361136003. Dhiraj Holden was supported in part by an Akamai Presidential fellowship, by NSF MACS - CNS-1413920, and by a SIMONS Investigator award Agreement Dated 6-5-12. Prashant Vasudevan was supported in part by the Qatar Computing Research Institute under the QCRI-CSAIL partnership, and by the National Science Foundation Frontier grant CNS 1413920.

References

[1] S. Aaronson, "Personal communication."

[2] ——, "Quantum lower bound for the collision problem," in *Proceedings of STOC*, 2002, pp. 635–642.

[3] ——, "Quantum computing and hidden variables," *Physical Review A*, vol. 71, no. 3, p. 032325, 2005.

[4] ——, "Impossibility of succinct quantum proofs for collision-freeness," *Quantum Information & Computation*, vol. 12, no. 1-2, pp. 21–28, 2012.

[5] S. Aaronson, A. Bouland, J. Fitzsimons, and M. Lee, "The space "just above" BQP," in *Proceedings of ITCS*, 2016, pp. 271–280.

[6] S. Agrawal, Y. Ishai, D. Khurana, and A. Paskin-Cherniavsky, "Statistical randomized encodings: A complexity theoretic view," in *Proceedings of ICALP (Part I)*, 2015, pp. 1–13.

[7] W. Aiello and J. Håstad, "Relativized perfect zero knowledge is not BPP," *Information and Computation*, vol. 93, pp. 223–240, 1991.

[8] ——, "Statistical zero-knowledge languages can be recognized in two rounds," *J. Comput. Syst. Sci.*, vol. 42, no. 3, pp. 327–345, 1991.

[9] B. Barak, "How to go beyond the black-box simulation barrier," in *Proceedings of FOCS*, 2001, pp. 106–115.

[10] R. B. Boppana, J. Håstad, and S. Zachos, "Does co-NP have short interactive proofs?" *Inf. Process. Lett.*, vol. 25, no. 2, pp. 127–132, 1987.

[11] A. Bouland, L. Chen, D. Holden, J. Thaler, and P. N. Vasudevan, "On the power of statistical zero knowledge," *Electronic Colloquium on Computational Complexity (ECCC)*, vol. 23, p. 140, 2016. [Online]. Available: http://eccc.hpi-web.de/report/2016/140

[12] M. Bun and J. Thaler, "Hardness amplification and the approximate degree of constant-depth circuits," in *Proceedings of ICALP (Part I)*, 2015, pp. 268–280.

[13] ——, "Dual polynomials for collision and element distinctness," *Theory of Computing*, vol. 12, no. 1, pp. 1–34, 2016.

[14] ——, "Improved bounds on the sign-rank of AC^0," *Electronic Colloquium on Computational Complexity (ECCC)*, vol. 23, p. 75, 2016.

[15] ——, "A nearly optimal lower bound on the approximate degree of AC^0," *Electronic Colloquium on Computational Complexity (ECCC)*, vol. 24, p. 51, 2017, to appear in *FOCS* 2017.

[16] A. Chakrabarti, G. Cormode, A. McGregor, J. Thaler, and S. Venkatasubramanian, "Verifiable stream computation and Arthur-Merlin communication," in *Proceedings of CCC*, 2015, pp. 217–243.

[17] R. Chang, B. Chor, O. Goldreich, J. Hartmanis, J. Håstad, D. Ranjan, and P. Rohatgi, "The random oracle hypothesis is false," *J. Comput. Syst. Sci.*, vol. 49, no. 1, pp. 24–39, 1994.

[18] L. Chen, "Adaptivity vs postselection," *arXiv:1606.04016*, 2016.

[19] ——, "A note on oracle separations for BQP," *arXiv:1605.00619*, 2016.

[20] G. Cormode, J. Thaler, and K. Yi, "Verifying computations with streaming interactive proofs," *PVLDB*, vol. 5, no. 1, pp. 25–36, 2011.

[21] M. Fischlin, "On the impossibility of constructing non-interactive statistically-secret protocols from any trapdoor one-way function," in *Proceedings of the The Cryptographer's Track at the RSA Conference on Topics in Cryptology*, ser. CT-RSA '02, 2002, pp. 79–95.

[22] L. Fortnow, "The complexity of perfect zero-knowledge (extended abstract)," in *Proceedings of STOC*, 1987, pp. 204–209.

[23] O. Goldreich and S. Goldwasser, "On the limits of non-approximability of lattice problems," in *Proceedings of STOC*, 1998, pp. 1–9.

[24] O. Goldreich and E. Kushilevitz, "A perfect zero-knowledge proof for a problem equivalent to discrete logarithm," in *Proceedings of CRYPTO*, 1988, pp. 57–70.

[25] O. Goldreich, S. Micali, and A. Wigderson, "Proofs that yield nothing but their validity or all languages in NP have zero-knowledge proof systems," *Journal of the ACM (JACM)*, vol. 38, no. 3, pp. 690–728, 1991.

[26] O. Goldreich, A. Sahai, and S. Vadhan, "Can statistical zero knowledge be made non-interactive? Or on the relationship of SZK and NISZK," in *Proceedings of CRYPTO*, 1999, pp. 467–484.

[27] S. Goldwasser, "Zero knowledge probabilistic proof systems," https://www.youtube.com/watch?v=J4TkHuTmHsg#t=1h15m20s, Simons Institute for the Theory of Computing, 2015.

[28] S. Goldwasser, S. Micali, and C. Rackoff, "The knowledge

complexity of interactive proof systems," *SIAM Journal on computing*, vol. 18, no. 1, pp. 186–208, 1989.

[29] M. Göös, T. Pitassi, and T. Watson, "Zero-information protocols and unambiguity in arthur-merlin communication," in *Proceedings of ITCS*, 2015, pp. 113–122.

[30] ——, "The landscape of communication complexity classes," in *Proceedings of ICALP (Part I)*, 2016, pp. 86:1–86:15.

[31] H. Klauck, "On Arthur Merlin games in communication complexity," in *Proceedings of CCC*, 2011, pp. 189–199.

[32] N. Linial and A. Shraibman, "Learning complexity vs communication complexity," *Combinatorics, Probability & Computing*, vol. 18, no. 1-2, pp. 227–245, 2009.

[33] S. V. Lokam, "Spectral methods for matrix rigidity with applications to size-depth trade-offs and communication complexity," *J. Comput. Syst. Sci.*, vol. 63, no. 3, pp. 449–473, 2001.

[34] S. Lovett and J. Zhang, "On the impossibility of entropy reversal, and its application to zero-knowledge proofs," *ECCC TR16-118*, 2016.

[35] L. Malka, "How to achieve perfect simulation and a complete problem for non-interactive perfect zero-knowledge," *Journal of Cryptology*, vol. 28, no. 3, pp. 533–550, 2015.

[36] T. Okamoto, "On relationships between statistical zero-knowledge proofs," in *Proceedings of STOC*, 1996, pp. 649–658.

[37] P. A. Papakonstantinou, D. Scheder, and H. Song, "Overlays and limited memory communication," in *Proceedings of CCC*, 2014, pp. 298–308.

[38] R. Paturi and J. Simon, "Probabilistic communication complexity," *J. Comput. Syst. Sci.*, vol. 33, no. 1, pp. 106–123, 1986.

[39] C. Peikert and V. Vaikuntanathan, "Noninteractive statistical zero-knowledge proofs for lattice problems," in *Proceedings of CRYPTO*, 2008, pp. 536–553.

[40] A. A. Razborov and A. A. Sherstov, "The sign-rank of AC^0," *SIAM J. Comput.*, vol. 39, no. 5, pp. 1833–1855, 2010.

[41] A. Sahai and S. Vadhan, "A complete problem for statistical zero knowledge," *Journal of the ACM (JACM)*, vol. 50, no. 2, pp. 196–249, 2003.

[42] A. A. Sherstov, "The pattern matrix method," *SIAM Journal on Computing*, vol. 40, no. 6, pp. 1969–2000, 2011.

[43] ——, "Breaking the Minsky-Papert barrier for constant-depth circuits," in *Proceedings of STOC*. ACM, 2014, pp. 223–232.

[44] ——, "The power of asymmetry in constant-depth circuits," in *Proceedings of FOCS*. IEEE, 2015, pp. 431–450.

[45] S. Toda, "PP is as hard as the polynomial-time hierarchy," *SIAM J. Comput.*, vol. 20, no. 5, pp. 865–877, 1991.

[46] N. K. Vereshchagin, "Lower bounds for perceptrons solving some separation problems and oracle separation of AM from PP," in *Proceedings of ISTCS*, 1995, pp. 46–51.