Approximate Degree, Secret Sharing, and Concentration Phenomena

Andrej Bogdanov
Department of Computer Science and Engineering, Chinese University of Hong Kong
Institute for Theoretical Computer Science and Communications, Hong Kong
andrejb@cse.cuhk.edu.hk

Nikhil S. Mande
Department of Computer Science, Georgetown University, USA
nikhil.mande@georgetown.edu

Justin Thaler
Department of Computer Science, Georgetown University, USA
justin.thaler@georgetown.edu

Christopher Williamson
Department of Computer Science and Engineering, Chinese University of Hong Kong, Hong Kong
chris@cse.cuhk.edu.hk

Abstract

The ε-approximate degree $\widetilde{\deg}_\varepsilon(f)$ of a Boolean function f is the least degree of a real-valued polynomial that approximates f pointwise to within ε. A sound and complete certificate for approximate degree being at least k is a pair of probability distributions, also known as a *dual polynomial*, that are perfectly k-wise indistinguishable, but are distinguishable by f with advantage $1-\varepsilon$. Our contributions are:

- We give a simple, explicit new construction of a dual polynomial for the AND function on n bits, certifying that its ε-approximate degree is $\Omega\left(\sqrt{n \log 1/\varepsilon}\right)$. This construction is the first to extend to the notion of weighted degree, and yields the first explicit certificate that the 1/3-approximate degree of any (possibly unbalanced) read-once DNF is $\Omega(\sqrt{n})$. It draws a novel connection between the approximate degree of AND and anti-concentration of the Binomial distribution.
- We show that any pair of *symmetric* distributions on n-bit strings that are perfectly k-wise indistinguishable are also statistically K-wise indistinguishable with at most $K^{3/2} \cdot \exp\left(-\Omega\left(k^2/K\right)\right)$ error for all $k < K \leq n/64$. This bound is essentially tight, and implies that any symmetric function f is a reconstruction function with constant advantage for a ramp secret sharing scheme that is secure against size-K coalitions with statistical error $K^{3/2} \cdot \exp\left(-\Omega\left(\widetilde{\deg}_{1/3}(f)^2/K\right)\right)$ for all values of K up to $n/64$ simultaneously. Previous secret sharing schemes required that K be determined in advance, and only worked for $f = \mathsf{AND}$. Our analysis draws another new connection between approximate degree and concentration phenomena.

  As a corollary of this result, we show that for any $d \leq n/64$, any degree d polynomial approximating a symmetric function f to error 1/3 must have coefficients of ℓ_1-norm at least $K^{-3/2} \cdot \exp\left(\Omega\left(\widetilde{\deg}_{1/3}(f)^2/d\right)\right)$. We also show this bound is essentially tight for any $d > \widetilde{\deg}_{1/3}(f)$. These upper and lower bounds were also previously only known in the case $f = \mathsf{AND}$.

2012 ACM Subject Classification Theory of computation → Pseudorandomness and derandomization

Keywords and phrases approximate degree, dual polynomial, pseudorandomness, polynomial approximation, secret sharing

Digital Object Identifier 10.4230/LIPIcs.APPROX-RANDOM.2019.71

Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques (APPROX/RANDOM 2019).
Editors: Dimitris Achlioptas and László A. Végh; Article No. 71; pp. 71:1–71:21
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

Category RANDOM

Related Version A full version of the paper is available at `https://eccc.weizmann.ac.il/report/2019/082/`.

Funding *Andrej Bogdanov*: Supported by Hong Kong RGC GRF CUHK14207618.
Nikhil S. Mande: Supported by NSF Grant CCF-1845125.
Justin Thaler: Supported by NSF Grant CCF-1845125.
Christopher Williamson: Supported by the Hong Kong PhD Fellowship Scheme.

Acknowledgements We thank Mark Bun for telling us about the work of Sachdeva and Vishnoi [22], and Mert Sağlam, Pritish Kamath, Robin Kothari, and Prashant Nalini Vasudevan for helpful comments on a previous version of the manuscript. We are also grateful to Xuangui Huang and Emanuele Viola for sharing the manuscript [14].

1 Introduction

The ε-approximate degree of a function $f\colon \{-1,1\}^n \to \{0,1\}$, denoted $\widetilde{\deg}_\varepsilon(f)$, is the least degree of a multivariate real-valued polynomial p such that $|p(x) - f(x)| \leq \varepsilon$ for all inputs $x \in \{-1,1\}^n$.[1] Such a p is said to be an approximating polynomial for f. This is a central object of study in computational complexity, owing to its polynomial equivalence to many other complexity measures including sensitivity, exact degree, deterministic and randomized query complexity [20], and quantum query complexity [6].

By linear programming duality, f has ε-approximate degree more than k if and only if there exist a pair of probability distributions μ and ν over the domain of f such that μ and ν are perfectly k-wise indistinguishable (i.e., all k-wise projections of μ and ν are identical), but are $(1-\varepsilon)$-distinguishable by f, namely $\mathbb{E}_{X\sim\mu}[f(X)] - \mathbb{E}_{Y\sim\nu}[f(Y)] \geq 1-\varepsilon$. Said equivalently, a sound and complete certificate for ε-approximate degree being more than k is a *dual polynomial* $q = (\mu-\nu)/2$ that contains no monomials of degree k or less, and such that $\sum_x |q(x)| = 1$ and $\sum_x q(x)f(x) \geq \varepsilon$.

Dual polynomials have immediate applications to cryptographic secret sharing: a dual polynomial $q = (\mu-\nu)/2$ for f is a description of a cryptographic scheme for sharing a 1-bit secret amongst n parties, where the secret can be reconstructed by applying f to the shares, and the scheme is secure against coalitions of size k (see [4] for details).

Motivation for explicit constructions of dual polynomials. Recent years have seen significant progress in proving new approximate degree lower bounds by explicitly constructing dual polynomials exhibiting the lower bound [8, 24, 9, 25, 10, 7, 11, 27]. These new lower bounds have in turn resolved significant open questions in quantum query complexity and communication complexity. At the technical core of these results are techniques for constructing a dual polynomial for composed functions $f \circ g := f(g,\ldots,g)$, given dual polynomials for f and g individually.

Often, an explicitly constructed dual polynomial showing that $\widetilde{\deg}_\varepsilon(g) \geq d$ exhibits additional metric properties, beyond what is required simply to witness $\widetilde{\deg}_\varepsilon(g) \geq d$. Much of the major recent progress in proving approximate degree lower bounds has exploited these additional metric properties [10, 7, 11, 27]. Accordingly, even if cases where an approximate degree lower bound for a function g is known, it can often be useful to construct an explicit dual polynomial witnessing the lower bound. Hence, we are optimistic that the new constructions of dual polynomials given in this work will find future applications.

[1] In this work, for convenience we also consider functions mapping $\{0,1\}^n$ to $\{0,1\}$.

Explicit constructions of dual polynomials are also necessary to implement the corresponding secret-sharing scheme, and to analyze the complexity of the algorithm that samples the shares of the secret.

Our results in a nutshell. Our results fall into two categories. In the first category, we reprove several known approximate degree lower bounds by giving the first explicit constructions of dual polynomials witnessing the lower bounds. Specifically, our dual polynomial certifies that the ε-approximate degree of the n-bit AND function is $\Theta(\sqrt{n \log 1/\varepsilon})$. This construction is the first to extend to the notion of weighted degree, and yields the first explicit certificate that the 1/3-approximate degree of any (possibly unbalanced) read-once DNF is $\Omega(\sqrt{n})$. Interestingly, our dual polynomial construction draws a novel and clean connection between the approximate degree of AND and anti-concentration of the Binomial distribution.

In the second category, we prove new and tight results about the size of the coefficients of polynomials that approximate symmetric functions. Specifically, we show that for any $d \leq n/64$, any degree d polynomial approximating f to error 1/3 must have coefficients of weight (ℓ_1-norm) at least $d^{3/2} \cdot \exp\left(\Omega\left(\widetilde{\deg}_{1/3}\left(f\right)^2/d\right)\right)$. We show this bound is tight (up to logarithmic factors in the exponent) for any $d > \widetilde{\deg}_{1/3}(f)$. These bounds were previously only known in the case $f = \mathsf{AND}$ [23, 5]. Our analysis actually establishes a considerably more general result, and as a consequence we obtain new cryptographic secret sharing schemes with symmetric reconstruction procedures (see Section 1.2 for details).

1.1 A New Dual Polynomial for AND

To describe our dual polynomial for AND, it will be convenient to consider the AND function to have domain $\{-1,1\}^n$ and range $\{0,1\}$, with $\mathsf{AND}(x) = 1$ if and only if $x = 1^n$. In their seminal work, Nisan and Szegedy [20] proved that the 1/3-approximate degree of the AND function on n inputs is $\Theta(\sqrt{n})$. More generally, it is now well-known that the ε-approximate degree of AND is $\Theta\left(\sqrt{n \log(1/\varepsilon)}\right)$ [15, 6]. These works do not construct explicit dual polynomials witnessing the lower bounds; this was achieved later in works of Špalek [28] and Bun and Thaler [8].

Our first contribution is the construction of a new dual polynomial ϕ for AND, which is simple enough to describe in a single equation:

$$\phi(x) = \frac{(-1)^n}{Z}\left(\prod_{i\in[n]} x_i\right)\left(\mathbb{E}_S \prod_{i\in S} x_i\right)^2. \tag{1}$$

Here, S is a random subset of $\{1,\ldots,n\}$ of size at most $\frac{1}{2}(n-d)$ (where d determines the degree of the polynomials against which the exhibited lower bound holds), and Z is an (explicit) normalization constant.

In the language of secret sharing, to share a secret $s \in \{-1,1\}$, the dealer samples shares $x \in \{-1,1\}^n$ with probability proportional to $(\mathbb{E}_S \prod_{i\in S} x_i)^2$, conditioned on the parity of the shares $\prod x_i$ being equal to s.

In Corollary 8 we show that ϕ certifies that every degree-d polynomial must differ from the AND function by $2^{-n}\sum_{k=0}^{(n-d)/2}\binom{n}{k}$ at some input. In other words, the approximation error of a degree-d polynomial is lower bounded by the probability that a sum of unbiased independent bits deviates from its mean by $d/2$.

Our function ϕ given in (1), unlike previous dual polynomials [15, 28, 9, 26], also certifies that the *weighted* 1/3-approximate degree of AND with weights $w \in \mathbb{R}^n_{\geq 0}$ is $\Omega(\|w\|_2)$ (see Corollary 9).[2] This lower bound is tight for all w, matching an upper bound of Ambainis [1]. The only difference in our dual polynomial construction for the weighted case is in the distribution over sets S, and the lower bound in the weighted case is derived from anti-concentration of *weighted* sums of Bernoulli random variables.

Both statements are corollaries of the following theorem.

▶ **Theorem 1.** *Define* $\mathsf{AND}\colon \{-1,1\}^n \to \{0,1\}$ *as* $\mathsf{AND}(x) = 1$ *if and only if* $x = 1^n$. *The function* ϕ *defined in Equation* (1) *is a dual witness for* $\widetilde{\deg}_{w,\varepsilon}(\mathsf{AND}) \geq d$ *for* $\varepsilon = \Pr_{X\sim\{-1,1\}^n}[\langle w, X\rangle \geq d]$.

By combining, in a black-box manner, the dual polynomial for the weighted-approximate degree of AND with prior work (e.g., [16, Proof of Theorem 7]), one obtains, for any read-once DNF f, an explicit dual polynomial for the fact that $\widetilde{\deg}_{1/3}(f) \geq \Omega(n^{1/2})$. Very recent work of Ben-David et al. [2] established this result for the first time, shaving logarithmic factors off of prior work [9, 16]. In fact, Ben-David et al. [2] prove more generally that any depth-d read-once AND-OR formula has approximate degree $2^{-O(d)}\sqrt{n}$. Their method, however, does not appear to yield an explicit dual polynomial, even in the case $d = 2$.

Discussion. It has been well known that the ε-approximate degree of the AND function on n variables is $\Theta\left(\sqrt{n\log(1/\varepsilon)}\right)$ [20, 6], a fact which has many applications in theoretical computer science. This is superficially reminiscent of Chernoff bounds, which state that the middle $\Theta\left(\sqrt{n\log(1/\varepsilon)}\right)$ layers of the Hamming cube contain a $1-\varepsilon$ fraction of all inputs (i.e., "most" n-bit strings have Hamming weight close to $n/2$). However, these two phenomena have not previously been connected, and it is not a priori clear why approximate degree should be related to concentration of measure. An approximating polynomial p for f must approximate f at *all* inputs in $\{-1,1\}^n$. Why should it matter that *most* (but very far from all) inputs have Hamming weight close to $n/2$?

The new dual witness for AND constructed in Equation (1) above provides a surprising answer to this question. The connection between (anti-)concentration and approximate degree of AND arises not because of the number of *inputs* to f that have Hamming weight close to $n/2$, but because of the number of *parity functions* on n bits that have *degree* close to $n/2$. This connection appears to be rather deep, as evidenced by our construction's ability to yield a tight lower bound in the case of weighted approximate degree.

1.2 Indistinguishability for Symmetric Distributions

In this section, for convenience we consider functions mapping $\{0,1\}^n$ to $\{0,1\}$. Two distributions μ and ν over $\{0,1\}^n$ are *(statistically)* (k,δ)*-wise indistinguishable* if for all subsets $S \subseteq \{1,\ldots,n\}$ of size k, the induced marginal distributions $\mu|_S$ and $\nu|_S$ are within statistical distance δ. When $\delta = 0$, we say they are *(perfectly)* k*-wise indistinguishable.*

[2] For a polynomial $p(x_1,\ldots,x_n)$, a weight vector $w \in \mathbb{R}^n_{\geq 0}$ assigns weight w_i to variable x_i. The weighted degree of p is the maximum weight over all monomials appearing in p, where the weight of a monomial is the sum of the weights of the variables appearing within it. The weighted ε-approximate degree of f, denoted $\widetilde{\deg}_{w,\varepsilon}(f)$, is the least weighted degree of any polynomial that approximates f pointwise to error ε.

For general pairs of distributions, perfect k-wise indistinguishability does not imply any sort of security against distinguishers of size $k+1$. Any binary linear error-correcting code of distance $k+1$ and block length n induces a pair of distributions (the uniform distribution over codewords and one of its affine shifts) that are perfectly k-wise indistinguishable, yet perfectly $(k+1)$-wise distinguishable.

In contrast, we prove that perfect k-wise indistinguishability for *symmetric* distributions implies strong statistical security against larger adversaries:

▶ **Theorem 2.** *If μ and ν are symmetric over $\{0,1\}^n$ and perfectly k-wise indistinguishable, then they are statistically $(K, O(K^{3/2}) \cdot e^{-k^2/1156K})$-wise indistinguishable for all $1 \leq k < K \leq n/64$.*

Theorem 2 has the following direct consequence for secret sharing schemes over bits with symmetric reconstruction. We say (μ, ν) are α-reconstructible by f if $\mathbb{E}_{X\sim\mu}[f(X)] - \mathbb{E}_{Y\sim\nu}[f(Y)] \geq \alpha$.

▶ **Corollary 3.** *Let f be a symmetric Boolean function. There exists a pair of distributions μ and ν that are $\left(K, K^{3/2} \cdot e^{-\Omega(\widetilde{\deg}_{1/3}(f)^2/K)}\right)$-indistinguishable for all $K \leq n/64$, but are $\Omega(1)$-reconstructible by f.*

Corollary 3 is an immediate consequence of our Theorem 2, and the fact that any symmetric function has an optimal dual polynomial that is itself symmetric. In the special case $f = \mathsf{AND}$ (or equivalently $f = \mathsf{OR}$), Corollary 3 implies the existence of a *visual secret sharing scheme* (see, for example [19]) that is $\left(K, K^{3/2} \cdot e^{-\Omega(n/K)}\right)$-statistically secure against all coalitions of size K, simultaneously for all K up to size $n/64$. This property, where security guarantees are in place for many coalition sizes at the same time, is in contrast to an earlier result of Bogdanov and Williamson [5] where they proved that for any fixed coalition size K, there is a visual secret sharing scheme that is $(K, e^{-\Omega(n/K)})$-statistically secure. In their construction, the distribution of shares μ and ν depend on the value of K.

We remark that the bound of Corollary 3 cannot hold in general for $K = n$, since there exists distributions that are perfectly $\Omega(n)$-wise indistinguishable but are reconstructible by the majority function on all n inputs. We do not however know if a bound of the form $K \leq (1-\Omega(1))n$ is tight in this context.

Tight weight-degree tradeoffs for polynomials approximating symmetric functions

Let $f\colon \{0,1\}^n \to \{0,1\}$ be any function. For any integer $d \geq 0$, denote by $W_\varepsilon(f,d)$ the minimum *weight* of any degree-d polynomial that approximates f pointwise to error ε. By the weight of a polynomial, we mean the ℓ_1-norm of its coefficients over the parity (Fourier) basis[3]. In Section B, we observe that Corollary 3 implies weight-degree trade-off lower bounds for symmetric functions.

▶ **Corollary 4.** *For any symmetric function $f\colon \{0,1\}^n \to \{0,1\}$, any constant $\varepsilon \in (0,1/2)$, and any integer K where $n/64 \geq K \geq \widetilde{\deg}_\varepsilon(f)$, we have $W_\varepsilon(f,K) \geq K^{-3/2} \cdot 2^{\Omega\left(\widetilde{\deg}_{1/3}(f)^2/K\right)}$.*

The following theorem shows that the lower bound obtained in Corollary 4 is tight (up to polylogarithmic factors in the exponent) for all symmetric functions.

[3] In fact, our main weight lower bound (Corollary 4) holds over any set of functions (not just parities) that each depend on at most d variables.

▶ **Theorem 5.** *For any symmetric function $f\colon \{0,1\}^n \to \{0,1\}$, any constant $\varepsilon \in (0,1/2)$ and $K > \widetilde{\deg}_\varepsilon(f) \cdot \sqrt{\log n}$, $W_\varepsilon(f,K) \le 2^{\tilde{O}(\widetilde{\deg}_{1/3}(f)^2/K)}$.*[4]

Theorem 5 also implies that Corollary 3 is tight (up to polylogarithmic factors in the exponent) for all symmetric f and for all $K \ge \widetilde{\deg}_{1/3}(f)\sqrt{\log n}$. This is because any improvement to Corollary 3 would yield an improvement to Corollary 4, contradicting Theorem 5.

Essentially Optimal Ramp Visual Secret Sharing Schemes. The following result shows that in the case $f = \mathsf{AND}$, Corollary 3 is essentially tight for *all* $K \ge 2$, and Theorem 2 is tight as a reduction from perfect to approximate indistinguishability for symmetric distributions. It does so by constructing essentially optimal ramp visual secret sharing schemes.[5]

▶ **Theorem 6.** *For all $2 \le k < K \le n$ there exist symmetric k-wise indistinguishable distributions μ and ν over n-bit strings that are $\sqrt{2^{-4K+3} \cdot \sum_{d>k} \binom{2K}{K+d}^2}$-reconstructible by AND_K, where $\mathsf{AND}_K(x)$ is the AND of the first K bits of x.*

Discussion of Theorem 6. This theorem gives the existence of a ramp visual secret sharing scheme that is perfectly secure against any k parties, but in which any $K > k$ parties can reconstruct the secret with the above advantage. This generalizes the schemes in [5] where only reconstruction by all n parties was considered.

Let us express the reconstruction advantage appearing in Theorem 6 in a manner more easily comparable to other results in this manuscript. Standard results on anti-concentration of the Binomial distribution state that $2^{-2K} \cdot \sum_{d>k} \binom{2K}{K+d} = e^{-\Theta(k^2/K)}$ (see, e.g., [17]). The Cauchy-Schwarz inequality then implies that the reconstruction advantage appearing in Theorem 6 is at least $K^{-1/2} \cdot e^{-O(k^2/K)}$.[6]

Hence, the visual secret sharing schemes given in Theorem 6 are nearly optimal; if the reconstruction advantage could be improved by more than the leading poly(K) factor (or the constant factor in the exponent), then this would contradict Theorem 2 which upper bounds the distinguishing advantage of any statistical test over K bits against symmetric, perfectly k-wise indistinguishable distributions. Theorem 6 also shows that the indistinguishability parameter in Theorem 2 cannot be significantly improved, even in the restricted case where the only statistical test is AND_K.

In Section 4 we describe another application of Theorem 2 to security against share consolidation and "downward self-reducibility" of visual secret shares.

[4] Here and throughout, the $\tilde{O}$ notation hides polylogarithmic factors in n.

[5] A visual secret sharing scheme is a scheme where the reconstruction function is the AND of some subset of the shares. A ramp scheme is one where there is not necessarily a sharp threshold between the perfect secrecy and reconstruction thresholds; in particular, we allow for $K > k + 1$.

[6] Theorem 6 is closely related to Theorem 1, in that Theorem 6 gives *another* anti-concentration-based proof that $\widetilde{\deg}_\varepsilon(\mathsf{AND}_K) \ge k$ for $\varepsilon = K^{-1/2} \cdot e^{-\Theta(k^2/K)}$. However, the two results are incomparable. Theorem 6 does not yield an explicit dual polynomial for AND_K, and the ε-approximate degree lower bound for AND_K implied by Theorem 6 is loose by the $K^{-1/2}$ factor appearing in the expression for ε. On the other hand, Theorem 1 only yields a visual secret sharing scheme with reconstruction by all n parties, while Theorem 6 yields a ramp scheme with non-trivial reconstruction advantage by the AND of the first K (out of n) parties.

1.3 Related Works

Prior Work. Servedio, Tan, and Thaler [23] established Corollary 4 and Theorem 5 in the special case $f = \mathsf{OR}$, showing that degree d polynomials that approximate the OR function require weight $2^{\tilde{\Theta}(n/d)} = 2^{\tilde{\Theta}(\widetilde{\deg}_{1/3}(\mathsf{OR})^2/d)}$.[7] They used this result to establish tight weight-degree tradeoffs for polynomial threshold functions computing decision lists. As previously mentioned, Bogdanov and Willamson [5] generalized the weight-vs-degree lower bound from [23] beyond polynomials, thereby obtaining a visual secret-sharing scheme for any fixed K that is $(K, e^{-\Omega(n/K)})$-statistically secure.

Elkies [13] and Sachdeva and Vishnoi [22] exploit concentration of measure to prove a tight upper bound on the degree of univariate polynomials that approximate the function $t \mapsto t^n$ over the domain $[-1, 1]$. Their techniques inspired our (much more technical) proof of Theorem 2.

Other Related Work. This work subsumes Bogdanov's manuscript [3], which shows a slightly weaker lower bound on the weighted approximate degree of AND, and does not derive an explicit dual polynomial. In independent work, Huang and Viola [14] prove a weaker form of our Corollary 3: their distributions μ, ν depend on the value of K. They also prove (a slightly tighter version of) Theorem 5, thereby establishing that the statistical distance in Corollary 3 is tight.

1.4 Techniques and Organization

The proof of Theorem 1 (Section 2) is an elementary verification that the function ϕ given in (1) is a dual polynomial. The only property that is not immediate is correlation with AND. Verifying this property amounts to upper bounding the normalization constant Z, which follows from orthogonality of the Fourier characters.

In the proof of Theorem 2 (Section 3), a K-bit statistical distinguisher for symmetric distribution is first decomposed into a sum of at most $K + 1$ tests Q_w that evaluate to 1 only when the input has Hamming weight exactly w. Lemma 13 shows that the univariate symmetrizations p_w of these distinguishers can be pointwise approximated by a degree-k polynomial with error at most $O(K^{1/2}) \cdot e^{-\Omega(k^2/K)}$.

To construct the desired approximation, we derive an identity relating the moment generating function of the squared Chebyshev coefficients of p_w (interpreted as relative probabilities) to the average magnitude of a polynomial g related to p_w on the unit complex circle (Claims 16 and 17). We bound these magnitudes analytically (Claim 18) and derive tail inequalities for the Chebyshev coefficients from bounds on the moment generating function as in standard proofs of Chernoff-Hoeffding bounds.

In the special case when the secrecy parameters k and K are fixed and the number of parties n approaches infinity, $p_w(t)$ turns out to equal $C_w(t-1)^w(t+1)^{K-w}$, where C_w is some quantity independent of t. In this case, the Chebyshev coefficients are the regular coefficients of the polynomial $g^\infty(s) = 2^{-w}C_w(s-1)^{2w}(s+1)^{2(K-w)}$.[8] When $w = 0$, $K/2$, or 1, the coefficients of g^∞ are exponentially concentrated around the middle as they follow

[7] These bounds for OR were implicit in [23], but not explicitly highlighted. The upper bound was explicitly stated in [12, Lemma 4.1], which gave applications to differential privacy, and the lower bound in [9, Lemma 32], which used it to establish tight weight-degree tradeoffs for polynomial threshold functions computing read-once DNFs.

[8] The i-th coefficient of g^∞ is the value of the i-th Kravchuk polynomial with parameter $2K$ evaluated at $2w$.

the binomial distribution. We prove that this exponential decay in magnitudes happens for all values of w, which requires understanding complicated cancellations in the algebraic expansion of $g^{\infty}(s)$.

We generalize this analysis to the finitary setting $n \geq 64K$.

We prove Theorem 5 (Section B) by writing any symmetric function f as a sum of at most $\ell := \min\{|f^{-1}(0)|, |f^{-1}(1)|\}$ many conjunctions, and approximating each conjunction to such low error (namely error $\ll \ell$) that the sum of all approximations is an approximation for f itself. Theorem 5 then follows by constructing low-weight, low-degree polynomial approximations for each conjunction in the sum.

Theorem 6 (Section C) is proved by lower bounding the error of degree k polynomial approximations to the symmetrization f of the function $\mathsf{AND}_K\left(x|_{\{1,\ldots,K\}}\right)$. By duality, a lower bound on approximation error translates into a secret sharing scheme with the same reconstruction advantage. To lower bound the error, we estimate the values of the coefficients in the Chebyshev expansion of f with indices larger than k. Owing to orthogonality, the largest of these coefficients lower bounds the approximation error of any degree-k polynomial.

In Section 4 we formulate a security of secret sharing against consolidation and downward self-reducibility of visual schemes, and derive these properties from the main results.

2 Dual Polynomial For the Weighted Approximate Degree of AND

In this section we prove Theorem 1 and derive its two corollaries about the unweighted and weighted approximate degree of AND.

Notation and Definitions. Let $[n] = \{1, \ldots, n\}$. Given a vector $w \in \mathbb{R}^n_{\geq 0}$, define the weight of a monomial $\chi_S(x) = \prod_{i \in S} x_i, x_i \in \{-1, 1\}$ to equal $\sum_{i \in S} w_i$. Define the w-weighted degree of a polynomial to be the maximum weight of a monomial in it. That is, if $p = \sum_{S \subseteq [n]} c_S \chi_S$, then define

$$\deg_w(p) = \max_{S : c_S \neq 0} w(S).$$

Define the w-weighted ε-approximate degree $\widetilde{\deg}_{w,\varepsilon}(f)$ to be the minimum w-weighted degree of a polynomial p that satisfies $|p(x) - f(x)| \leq \varepsilon$ for all x in the domain of f. Given two real-valued functions f, g over domain $\{-1, 1\}^n$, define $\langle f, g \rangle := \frac{1}{2^n} \sum_{x \in \{-1,1\}^n} f(x) \cdot g(x)$.

▶ **Lemma 7.** *For any finite set X and any function $f : X \to \mathbb{R}$, $\widetilde{\deg}_{w,\varepsilon}(f) \geq d$ iff there exists a function $\phi : X \to \mathbb{R}$ satisfying the following conditions.*

- Pure high degree*: For any real polynomial p of weighted degree is at most d, $\langle \phi, p \rangle = 0$.*
- Normalization*: $\sum_{x \in X} |\phi(x)| = 1$,*
- Correlation*: $\langle \phi, f \rangle \geq \varepsilon$,*

We call ϕ a dual witness for $\widetilde{\deg}_{w,\varepsilon}(f) \geq d$. The lemma follows by linear programming duality and is a straightforward generalization of previous results (see e.g. [28, 9]). We prove the "if" direction, which is sufficient for our purposes.

Proof. For any p of weighted degree at most d,

$$\|f - p\|_\infty = \|f - p\|_\infty \|\phi\|_1 \geq \langle \phi, f - p \rangle = \langle \phi, f \rangle - \langle \phi, p \rangle \geq \varepsilon. \qquad ◀$$

The dual polynomial of interest is

$$\phi(x) = \frac{(-1)^n}{Z} \chi_{[n]}(x) \cdot \mathbb{E}_{S\sim\mathcal{H}}[\chi_S(x)]^2,$$

where $x \in \{-1,1\}^n$, $\mathcal{H}$ is the uniform distribution over the sets $\{S \subseteq [n] : w(S) \leq (\|w\|_1 - d)/2\}$, and Z is the normalization constant

$$Z = \sum_{x\in\{-1,1\}^n} \mathbb{E}_{S\sim\mathcal{H}}[\chi_S(x)]^2.$$

Proof of Theorem 1. We prove the theorem by showing that ϕ satisfies the three conditions of Lemma 7. The expression $\mathbb{E}_{S\sim\mathcal{H}}[\chi_S(x)]^2$ can be written as a sum of products of pairs of monomials of weight at most $(\|w\|_1 - d)/2$, so its weighted degree is at most $\|w\|_1 - d$. Thus every monomial that occurs in the expansion of $\chi_{[n]}(x)\,\mathbb{E}_{S\sim\mathcal{H}}[\chi_S(x)]^2$ must have weighted degree *at least* d, and so ϕ has pure high weighted degree at least d as desired.

The scaling by Z in the definition of ϕ ensures that ϕ has L_1 norm 1. The correlation of ϕ and AND is given by $\langle \phi, \mathsf{AND}\rangle = \phi(1^n) = \frac{1}{Z}$. Finally, the normalization constant Z evaluates to

$$\begin{aligned} Z &= \sum_{x\in\{-1,1\}^n} \mathbb{E}_{S\sim\mathcal{H}}[\chi_S(x)]^2 = \sum_{x\in\{-1,1\}^n} \mathbb{E}_{S\sim\mathcal{H}}[\chi_S(x)]\,\mathbb{E}_{T\sim\mathcal{H}}[\chi_T(x)] \\ &= \sum_{x\in\{-1,1\}^n} \mathbb{E}_{S,T\sim\mathcal{H}}[\chi_{S\Delta T}(x)] = \mathbb{E}_{S,T\sim\mathcal{H}} \sum_{x\in\{-1,1\}^n} \chi_{S\Delta T}(x) \\ &= 2^n \Pr[S = T] = \frac{2^n}{|\mathcal{H}|}, \end{aligned}$$

since the inner summation over x evaluates to 2^n when $S = T$, and zero otherwise.

It remains to show that $1/Z = |\mathcal{H}|/2^n$ equals the desired expression for ε. For a set $S \subseteq [n]$, let $X(S) \in \{-1,1\}^n$ be the string that assigns values 1 and -1 to elements inside and outside S, respectively. Then $w(S) = \|w\|_1/2 + \langle w, X(S)\rangle/2$, so

$$\frac{|\mathcal{H}|}{2^n} = \Pr_{S\subseteq[n]}[w(S) \geq \|w\|_1/2 + d/2] = \Pr_{X\sim\{-1,1\}^n}[\langle w, X\rangle \geq d]. \qquad ◀$$

▶ **Corollary 8** (Approximate degree of AND). *Recall that* AND: $\{-1,1\}^n \to \{0,1\}$ *denotes the function satisfying* $\mathsf{AND}(x) = 1$ *if and only if* $x = 1^n$. *If* p *has degree at most* d*, then* $|p(x) - \mathsf{AND}(x)| \geq \Pr[X \leq (n-d)/2]$ *for some* x*, where* X *is a* $\mathrm{Binomial}(n, 1/2)$ *random variable.*

The expression on the right is lower bounded by the larger of $1/2 - O(d/\sqrt{n})$ and $2^{-O(d^2/n)}$. In the large d regime ($d \geq \sqrt{n}$), this bound is tight [15, 6].

Proof. Apply Theorem 1 to the weight vector $w = (1, 1, \dots, 1)$. ◀

Earlier constructions of dual polynomials for AND are quite different from our Corollary 8 [15, 28, 9, 26] and are based on real-valued polynomial interpolation. Specifically, for a carefully chosen set $T \subseteq \{0, 1, \dots, n\}$ of size $|T| = 2d$, the prior constructions consider a *univariate* polynomial $p(t) = \prod_{i\in[n]\setminus T}(t - i)$, and they define $\psi(x) = p(|x|)$, where $|x|$ denotes the Hamming weight of x. Clearly ψ has degree at most $n - |T|$. A fairly complicated calculation is required to show that, for an appropriate choice of T, defining ψ in this way ensures that $|\psi(1^n)|$ captures an ε-fraction of the L_1-mass of ψ.

▶ **Corollary 9** (Weighted approximate degree of AND). $\widetilde{deg}_{w,3/32}(\mathsf{AND}) \geq \|w\|_2/2$.

The proof uses the Paley-Zygmund inequality:

▶ **Lemma 10** (Paley-Zygmund inequality). *Let $Z \geq 0$ be any random variable with finite variance. Then, for any $0 < \theta < 1$,*

$$\Pr[Z \geq \theta\,\mathbb{E}(Z)] \geq (1-\theta)^2 \frac{(\mathbb{E}[Z])^2}{\mathbb{E}[Z^2]}.$$

Proof of Corollary 9. We apply the Paley-Zygmund inequality to $\langle w, X\rangle^2$. First, $\mathbb{E}[\langle w, X\rangle]^2 = \|w\|_2^2$ and $\mathbb{E}[\langle w, X\rangle^4] = \sum w_i^4 + 3\sum w_i^2 w_j^2 \leq 3\|w\|_2^2$. Then

$$\Pr\left[\langle w, X\rangle \geq \frac{\|w\|_2}{2}\right] = \frac{1}{2}\Pr\left[|\langle w, X\rangle| \geq \frac{\|w\|_2}{2}\right] = \frac{1}{2}\Pr\left[\langle w, X\rangle^2 \geq \frac{\|w\|_2^2}{4}\right] \geq \frac{1}{2}\cdot\frac{9}{16}\cdot\frac{1}{3} = \frac{3}{32},$$

where the first equality follows from the sign-symmetry of X. Applying Theorem 1 with $d = \|w\|_2/2$ yields the claim. ◀

3 Approximate Indistinguishability from Perfect Indistinguishability

In this section, we prove Theorem 2, which states that any pair of symmetric and perfectly k-wise indistinguishable distributions over $\{0,1\}^n$ are also approximately indistinguishable against statistical tests that observe $K > k$ of the bits. We may and will assume without loss of generality that the statistical test is a symmetric function,[9] meaning that it depends only on the Hamming weight of the observed bits of its input.

Let X and Y denote an arbitrary pair of symmetric $(k, 0)$-wise indistinguishable distributions over $\{0,1\}^n$. We will be interested in obtaining an upper bound on the statistical distance of their projections to any K indices of $[n]$, namely the advantage $\mathbb{E}_X[T(X|_S) - \mathbb{E}_Y[T(Y|_S)]$ where $T : \{0,1\}^K \to \{0,1\}$ is a symmetric function and $S \subseteq [n]$ is any set of size K. We can decompose T into a sum of tests $Q_w : \{0,1\}^K \to \{0,1\}$, where Q_w outputs 1 if and only if the Hamming weight of its input is exactly w. Specifically, we decompose T as

$$T = \sum_{w=0}^{K} b_w Q_w, \tag{2}$$

where each b_w is either zero or one. We will bound the distinguishing advantage of each Q_w in the sum individually. This advantage is captured by a univariate function p_w that expresses Q_w in terms of the Hamming weight of its input, after shifting and scaling the Hamming weight to reside in the interval $[-1, 1]$.

▶ **Fact 11.** *Let $S \subseteq [n]$ be any set of size K. There exists a univariate polynomial p_w of degree at most K such that the following holds. For all $t \in \{-1, -1 + 2/n, \ldots, 1 - 2/n, 1\}$, $p_w(t) = \mathbb{E}_Z[Q_w(Z|_S)]$ where Z is a random string of Hamming weight $\phi^{-1}(t) = (1-t)n/2 \in \{0, 1, \ldots, n\}$.*

Proof. This statement is a simple extension of Minsky and Papert's classic symmetrization technique [18]. Specifically, Minsky and Papert showed that for any polynomial $p_n : \{0,1\}^n \to \mathbb{R}$, there exists a univariate polynomial P of degree at most the total degree of p_n, such that for all $i \in \{0, \ldots, n\}$, $P(i) = \mathbb{E}_{|x|=i}[p_n(x)]$. Apply this result to $p_n(x) = Q_w(x|_S)$ and let $p_w(t) = P(\phi^{-1}(t)) = P((1-t)n/2)$. The fact then follows from the observation that the total degree of $Q_w(x|_S)$ is at most K, since this function is a K-junta. ◀

[9] In the full version, we include simple proofs that (1) the marginal distributions of a symmetric distribution are symmetric and that (2) the best distinguisher between a pair of symmetric distributions is a symmetric function.

In particular, the value $p_w(t)$ is a probability for every $t \in \{-1, -1+2/n, \dots, 1-2/n, 1\}$. Moreover, this probability must equal zero when the Hamming weight of Z is less than w or greater than $n - K + w$. Therefore p_w has K distinct zeros at the points $Z_w = Z_- \cup Z_+$, where

$$Z_- = \{-1 + 2h/n : h = 0, ..., K - w - 1\}, \qquad Z_+ = \{1 - 2h/n : h = 0, ..., w - 1\}. \tag{3}$$

and so p_w must have the form

$$p_w(t) = C_w \cdot \prod_{z \in Z_w} (t - z) \tag{4}$$

for some C_w that does not depend on t.[10] As $p_w(t)$ is probability when $t \in \{-1, -1 + 2/n, \dots, 1 - 2/n, 1\}$, the function p_w is 1-bounded at those inputs. In fact, p_w is uniformly bounded on the interval $[-1, 1]$:

▷ Claim 12. Assuming $n \geq 64K$, $|p_w(t)| \leq 2$ for all $t \in [-1, 1]$.

The proof is omitted due to space limitations but follows a similar structure as the proof of Claim 18 which appears in Section A. Formula (4) and Claim 12 will be applied to show that p_w has a good uniform polynomial approximation on the interval $[-1, 1]$.

▶ **Lemma 13.** *Assuming $n \geq 64K$, there exists a degree-k polynomial q_w such that $|p_w(t) - q_w(t)| \leq 4\sqrt{K}\exp(-k^2/1156K)$ for all $t \in [-1, 1]$.*

Lemma 13 is the main technical result of this section. It is proved in Section 3.1.

Proof of Theorem 2. Now let T be a general distinguisher on K inputs, which we may and will assume to be a symmetric Boolean-valued function. We bound the distinguishing advantage as follows. Recalling that X and Y are $(k, 0)$-indistinguishable symmetric distributions over $\{0, 1\}^n$, for any set $S \subseteq [n]$ of size K we have:

$$\begin{aligned}
&\mathbb{E}[T(X|_S)] - \mathbb{E}[T(Y|_S)] \\
&\quad = \sum_{w=0}^{K} b_w \left(\mathbb{E}[Q_w(X|_S)] - \mathbb{E}[Q_w(Y|_S)]\right) \qquad \text{(by (2))} \\
&\quad \leq \sum_{w=0}^{K} \left|\mathbb{E}[Q_w(X|_S)] - \mathbb{E}[Q_w(Y|_S)]\right| \qquad \text{(by boundedness of } b_w\text{)} \\
&\quad = \sum_{w=0}^{K} \left|\mathbb{E}[p_w(\phi(|X|)] - \mathbb{E}[p_w(\phi(|Y|))]\right| \qquad \text{(by symmetry of } X, Y\text{, and Fact 11)} \\
&\quad \leq \sum_{w=0}^{K} \left|\mathbb{E}[q_w(\phi(|X|))] - \mathbb{E}[q_w(\phi(|Y|))]\right| + 8\sqrt{K}\exp(-k^2/1156K) \qquad \text{(by Lemma 13)} \\
&\quad = O(K^{3/2}) \cdot e^{-k^2/1156K} \qquad \text{(by } k\text{-wise indistinguishability of } X, Y\text{)}
\end{aligned}$$

Therefore, X and Y are $(K, O(K^{3/2}) \cdot e^{-k^2/1156K})$-wise indistinguishable for $2 \leq K \leq n/64$. ◀

[10] p_w, C_w, and Z_w also depend on K and n but we omit those arguments from the notation as they will be fixed in the proof.

3.1 Proof of Lemma 13

We will prove Lemma 13 by studying the Chebyshev expansion of p_w. To this end we take a brief detour into Chebyshev polynomials and an even briefer one into Fourier analysis.

Chebyshev polynomials

The Chebyshev polynomials are a family of real polynomials $\{T_d\}$, 1-bounded on $[-1,1]$, with T_d having degree d. We extend the definition to negative indices by setting $T_{-d} = T_d$. The Chebyshev polynomials are orthogonal with respect to the measure $d\sigma(t) = (1-t^2)^{-1/2}dt$ supported on $[-1,1]$. Therefore every degree-K polynomial $p\colon \mathbb{R} \to \mathbb{R}$ has a unique (symmetrized) Chebyshev expansion

$$p(t) = \sum_{d=-K}^{K} c_d T_d(t), \qquad c_{-d} = c_d$$

where $c_{-K}, \dots, c_K$ are the *Chebyshev coefficients* of p.

The Chebyshev polynomials satisfy the following identity, which plays an important role in our analysis:

▶ **Fact 14.** $t \cdot T_d(t) = \frac{1}{2}T_{d-1}(t) + \frac{1}{2}T_{d+1}(t)$.

This formula, together with the "base cases" $T_0(t) = 1$ and $T_1(t) = t$, specifies all Chebyshev polynomials.

We will also need the following form of Parseval's identity for univariate polynomials.

▷ Claim 15 (Parseval's identity). For every complex polynomial h, the sum of the squares of the magnitudes of the coefficients of h equals $\mathbb{E}_z[|h(z)|^2]$, where z is a random complex number of magnitude 1.

Proof outline

We will argue that the Chebyshev expansion $\sum_{d=-K}^{K} c_d T_d(t)$ of $p_w(t)$ has small weight on the coefficients c_d when $|d| > k$. Zeroing out those coefficients then yields a good degree-k approximation of p_w as desired.

The upper bound on the Chebyshev coefficients of p_w is derived in two steps. The first step, which is of an algebraic nature, expresses the Chebyshev coefficients of p_w as regular coefficients of a related polynomial g.[11] We are interested in the coefficients of the derived polynomial $g_\varepsilon(s) = g((1+\varepsilon)s)$, which represent the Chebyshev coefficients c_d of p_w amplified by the exponential scaling factor $(1+\varepsilon)^d$.

The second step, which is analytic, upper bounds the magnitude of the coefficients of $g_\varepsilon(s)$. The main tool is Parseval's identity, which identifies the sum of the squares of these coefficients by the average magnitude of g_ε over the complex unit circle $\mathbb{E}_\theta\, |g((1+\varepsilon)e^{i\theta})|^2$. We bound the *maximum* magnitude $\max_\theta |g((1+\varepsilon)e^{i\theta})|^2$ by explicitly analyzing the function g. This step comprises the bulk of our proof.

The third step translates the bound on the squared 2-norm $\sum_{d=-K}^{K}(1+\varepsilon)^{2d}c_d^2$ of the amplified coefficients into a tail bound on c_d by optimizing over a suitable value of ε. This is analogous to the standard derivation of Chernoff-Hoeffding bounds by analysis of the moment generating function of the relevant random variable.

[11] We omit the dependence on w as this parameter remains constant throughout the proof.

We now sketch how this outline is executed for the special case where n tends to infinity while k and K remain fixed. Although this setting is technically much easier, it allows us to highlight the main conceptual points of our argument. The analysis for finite n can be viewed as an approximation of this proof strategy.

Sketch of the limiting case $n \to \infty$

By the expansion (4) of p_w, as n tends to infinity p_w converges uniformly to the function

$$p_w^\infty(t) = C_w \cdot (t-1)^w (t+1)^{K-w},$$

as this corresponds to Fact 11 when the bits of the string Z are independent and $(1-t)/2$-biased. As $p_w^\infty(t)$ is a probability for every $t \in [-1,1]$, Claim 12 follows immediately.

Step 1. Our algebraic treatment of the Chebyshev transform yields that the Chebyshev coefficient c_d of p_w^∞ is the $(K+d)$-th regular coefficient of the polynomial

$$g^\infty(s) = C_w \left(\frac{s-1}{\sqrt{2}}\right)^{2w} \left(\frac{s+1}{\sqrt{2}}\right)^{2(K-w)}. \tag{5}$$

Step 2. The evaluation of the polynomial $g_\varepsilon^\infty(s) = g^\infty((1+\varepsilon)s)$ at $s = e^{i\theta}$ satisfies the identity

$$\left|g^\infty\left((1+\varepsilon)e^{i\theta}\right)\right| = (1+\varepsilon)^K \cdot (1+\delta)^K \cdot C_w \cdot \left(1 - \frac{\cos\theta}{1+\delta}\right)^w \left(1 + \frac{\cos\theta}{1+\delta}\right)^{K-w}, \tag{6}$$

where $\delta = \varepsilon^2/2(1+\varepsilon)$. This happens to equal

$$(1+\varepsilon)^K(1+\delta)^K p_w(\cos\theta/(1+\delta)), \tag{7}$$

and is in particular uniformly bounded by $(1+\varepsilon)^K(1+\delta)^K$ for all θ. This similarity between p^∞ and g_ε^∞ is the crux of our analysis.

Step 3. By Parseval's identity, after suitable shifting and cancellation, the amplified sum of Chebyshev coefficients $\sum_{d=-K}^{K}(1+\varepsilon)^{2d}c_d^2$ is upper bounded by $(1+\delta)^{2K}$. Therefore the tail $\sum_{k\geq d} c_d^2$ can have value at most $(1+\delta)^{2K}/(1+\varepsilon)^{2k} \leq \exp(2K\varepsilon^2 - 2(\varepsilon - \varepsilon^2/2)k)$. This upper bound holds for all $\varepsilon \in [0,1]$, and plugging in the approximate minimizer $\varepsilon = k/2K$ yields a bound of the desired form $\exp(-\Omega(k^2/K))$.

Outline of the general case

We now give the outline of our full proof for the general case and relevant technical statements that we use to prove our main upper bound. Identity (5) generalizes to the following statement:

▷ Claim 16. The Chebyshev coefficient c_d of p_w is the $(K+d)$-th regular coefficient of the polynomial

$$g(s) = C_w \prod_{z \in Z_w} \left(\frac{s^2 - 2sz + 1}{2}\right),$$

where C_w is as in Equation (4).

The general form of identity (6) is:

▷ Claim 17. For $\varepsilon > 0$, $\delta = \varepsilon^2/2(1+\varepsilon)$, and $\theta \in [-\pi, \pi]$,

$$\left|g((1+\varepsilon)e^{i\theta})\right|^2 = (1+\varepsilon)^{2K}(1+\delta)^{2K} \cdot C_w^2 \prod_{z \in Z_w} h_{\delta(1+1/(1+\delta))}\left(\frac{\cos\theta}{1+\delta}, z\right)$$

where $h_\delta(s, z) = (s-z)^2 + \delta(1-z^2)$.

Owing to the second term in h_δ, there is no identity analogous to (7) when n is finite and p_w has zeros inside $(-1, 1)$. Nevertheless, $\prod_{z \in Z_w} h_\delta(s, z)$ can be uniformly bounded either by a sufficiently small multiple of $p_w(s)^2$, or a fixed quantity that is constant in the parameter range of interest.

▷ Claim 18. Assume $n \geq 64K$ and $w \leq K/2$. Then

$$C_w^2 \cdot \prod_{z \in Z_w} h_\delta(s, z) \leq \begin{cases} e^{65\delta K} \cdot p_w(s)^2 & \text{if } |s| \leq 1 - w/16K \\ e^{65\delta K} & \text{if } 1 - w/16K \leq |s| \leq 1. \end{cases}$$

We now prove Lemma 13. Due to space limitations, we omit the proof of Claim 16, which follows via induction and is an application of Fact 14, and the proof of Claim 17, which consists of a lengthy but relatively straightforward calculation. Claim 18 is proved in Section A.

▶ **Fact 19.** $p_w(t) = p_{K-w}(1-t)$.

Proof. By Fact 11, both sides are degree-K polynomials that agree on $n+1 > K$ points so they are identical. ◀

Proof of Lemma 13. By Fact 19 we may and will assume that $w \leq K/2$. Let $p_w = \sum_{d=-K}^{K} c_d T_d$. The approximating polynomial q_w is $\sum_{|d|<k} c_d T_d$. It remains to prove a tail upper bound on the Chebyshev coefficients. By Claim 16, the $(K+d)$-th coefficient of $g(s)$ is c_d. Therefore the polynomial $g_\varepsilon(s) = g((1+\varepsilon)s)$ has coefficients $(1+\varepsilon)^{K+d} c_d$ as d ranges from $-K$ to K. We apply Parseval's identity (Claim 15) to g_ε.

It follows that

$$\begin{aligned}\sum_{d=-K}^{K} (1+\varepsilon)^{2(K+d)} c_d^2 &= \mathbb{E}_\theta \left|g((1+\varepsilon)e^{i\theta})\right|^2 \\ &\leq \max_{\theta \in [-\pi,\pi]} \left|g((1+\varepsilon)e^{i\theta})\right|^2 \\ &= \max_{s \in [-1,1]} (1+\varepsilon)^{2K}(1+\delta)^{2K} \cdot C_w^2 \prod_{z \in Z_w} h_{\delta(1+1/(1+\delta))}(s/(1+\delta), z),\end{aligned}$$

by Claim 17. Since $0 \leq \delta = \varepsilon^2/2(1+\varepsilon) \leq 1/2$, for simplicity we may replace $h_{\delta(1+1/(1+\delta))}(s/(1+\delta), z)$ by $h_{2\delta}(s, z)$ in the above inequality. This gives the following approximation bound for $\alpha = \max_{t \in [-1,1]} |p_w(t) - q_w(t)|$:

$$\begin{aligned}\alpha &= \max_{t \in [-1,1]} \left|\sum\nolimits_{|d| \geq k} c_d T_d(t)\right| \\ &\leq \sum\nolimits_{|d| \geq k} |c_d| \max_{t \in [-1,1]} |T_d(t)| \\ &\leq 2\sum_{d \geq k} |c_d| \qquad \text{(by symmetry and boundedness of } T_d\text{)}\end{aligned}$$

$$\begin{aligned}\ldots &\le 2\sqrt{K}\cdot\sqrt{\sum\nolimits_{d\ge k} c_d^2} \qquad \text{(by Cauchy-Schwarz)}\\ &\le 2\sqrt{K}\cdot\sqrt{(1+\varepsilon)^{-2(K+k)}\sum\nolimits_{d\ge k}(1+\varepsilon)^{2(K+d)}c_d^2}\\ &\le 2\sqrt{K}\sqrt{(1+\varepsilon)^{-2k}\cdot(1+\delta)^{2K}\cdot\max_{s\in[-1,1]} C_w^2\prod_{z\in Z_w} h_{2\delta}(s,z)}.\end{aligned}$$

By the boundedness of p_w (Claim 12), the upper bounds in Claim 18 can be unified by the inequality $C_w^2\prod_{z\in Z_w} h_{2\delta}(s,z)\le 4e^{130\delta K}$ that is valid for all $s\in[-1,1]$. Since $1+\delta\le e^{\delta}$ and $1+\varepsilon\ge e^{\varepsilon-\varepsilon^2/2}$ for $0\le\varepsilon\le 1$,

$$\alpha\le 2\sqrt{K}\cdot\sqrt{\frac{(1+\delta)^{2K}}{(1+\varepsilon)^{2k}}\cdot 4e^{130\delta K}}\le 4\sqrt{K}\cdot\sqrt{e^{132\delta K-2\varepsilon k+\varepsilon^2 k}}\le 4\sqrt{K}\cdot\sqrt{e^{67\varepsilon^2 K-2\varepsilon k}},$$

where the last inequality follows from the definition $\delta=\varepsilon^2/2(1+\varepsilon)$. Setting $\varepsilon=k/34K$ we obtain that $\alpha\le 4\sqrt{K}\cdot e^{-k^2/1156K}$. ◀

4 Robustness of Symmetric Secret Sharing Against Consolidation

Consider a secret sharing scheme with tn parties, divided in n blocks of size t, that is perfectly secure against size-k coalitions. If all parties in each block come together and consolidate their information even into a single bit, the number of *blocks* against which the scheme remains secure drops to k/t. In general this is the best possible, with linear schemes providing tight examples.

The following corollary, proven in the full version, shows that if the distribution over shares is symmetric then much better security against this type of attack can be obtained.

▶ **Corollary 20.** *Let $f_1,\ldots,f_n\colon\{0,1\}^t\to\{0,1\}$. Assume X,Y are k-wise indistinguishable symmetrically distributed random variables over tn-bit strings. Write $X=X_1\ldots X_n$, $Y=Y_1\ldots Y_n$, where all blocks X_i, Y_i have size t. For every K, the n-bit random variables $X'=f_1(X_1)\ldots f_n(X_n)$ and $Y'=f_1(Y_1)\ldots f_n(Y_n)$ are $O((tK)^{3/2}n^K e^{-k^2/1156tK})$-close to being perfectly K-wise indistinguishable, assuming $K\le n/64$.*

References

1 Andris Ambainis. Quantum Search with Variable Times. *Theory Comput. Syst.*, 47(3):786–807, 2010.

2 Shalev Ben-David, Adam Bouland, Ankit Garg, and Robin Kothari. Classical Lower Bounds from Quantum Upper Bounds. In *59th IEEE Annual Symposium on Foundations of Computer Science, FOCS 2018, Paris, France, October 7-9, 2018*, pages 339–349, 2018.

3 Andrej Bogdanov. Approximate degree of AND via Fourier analysis. *Electronic Colloquium on Computational Complexity (ECCC)*, 25:197, 2018.

4 Andrej Bogdanov, Yuval Ishai, Emanuele Viola, and Christopher Williamson. Bounded Indistinguishability and the Complexity of Recovering Secrets. In *Advances in Cryptology - CRYPTO 2016 - 36th Annual International Cryptology Conference, Santa Barbara, CA, USA, August 14-18, 2016, Proceedings, Part III*, pages 593–618, 2016.

5 Andrej Bogdanov and Christopher Williamson. Approximate Bounded Indistinguishability. In *44th International Colloquium on Automata, Languages, and Programming, ICALP 2017, July 10-14, 2017, Warsaw, Poland*, pages 53:1–53:11, 2017.

6 Harry Buhrman, Richard Cleve, Ronald de Wolf, and Christof Zalka. Bounds for Small-Error and Zero-Error Quantum Algorithms. In *40th Annual Symposium on Foundations of Computer Science, FOCS '99, 17-18 October, 1999, New York, NY, USA*, pages 358–368, 1999.

7 Mark Bun, Robin Kothari, and Justin Thaler. The polynomial method strikes back: tight quantum query bounds via dual polynomials. In *Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing, STOC 2018, Los Angeles, CA, USA, June 25-29, 2018*, pages 297–310, 2018.

8 Mark Bun and Justin Thaler. Dual Lower Bounds for Approximate Degree and Markov-Bernstein Inequalities. In *Automata, Languages, and Programming - 40th International Colloquium, ICALP 2013, Riga, Latvia, July 8-12, 2013, Proceedings, Part I*, pages 303–314, 2013.

9 Mark Bun and Justin Thaler. Hardness Amplification and the Approximate Degree of Constant-Depth Circuits. In *Automata, Languages, and Programming - 42nd International Colloquium, ICALP 2015, Kyoto, Japan, July 6-10, 2015, Proceedings, Part I*, pages 268–280, 2015.

10 Mark Bun and Justin Thaler. A Nearly Optimal Lower Bound on the Approximate Degree of AC^0. In *58th IEEE Annual Symposium on Foundations of Computer Science, FOCS 2017, Berkeley, CA, USA, October 15-17, 2017*, pages 1–12, 2017.

11 Mark Bun and Justin Thaler. The Large-Error Approximate Degree of AC^0. *Electronic Colloquium on Computational Complexity (ECCC)*, 25:143, 2018.

12 Karthekeyan Chandrasekaran, Justin Thaler, Jonathan Ullman, and Andrew Wan. Faster private release of marginals on small databases. In *Innovations in Theoretical Computer Science, ITCS'14, Princeton, NJ, USA, January 12-14, 2014*, pages 387–402, 2014.

13 Noam D. Elkies (https://mathoverflow.net/users/14830/noam-d elkies). Uniform approximation of x^n by a degree d polynomial: estimating the error. MathOverflow. URL: `https://mathoverflow.net/q/70527`.

14 Xuangui Huang and Emanuele Viola. Almost Bounded Indistinguishability and Degree-Weight Tradeoffs, 2019. Manuscript.

15 Jeff Kahn, Nathan Linial, and Alex Samorodnitsky. Inclusion-exclusion: Exact and approximate. *Combinatorica*, 16(4):465–477, 1996.

16 Pritish Kamath and Prashant Vasudevan. Approximate Degree of AND-OR trees, 2014. Manuscript available at `https://www.scottaaronson.com/showcase3/kamath-pritish-vasudevan-prashant.pdf`.

17 Philip N. Klein and Neal E. Young. On the Number of Iterations for Dantzig-Wolfe Optimization and Packing-Covering Approximation Algorithms. *SIAM J. Comput.*, 44(4):1154–1172, 2015.

18 Marvin Minsky and Seymour Papert. *Perceptrons*. MIT Press, Cambridge, MA, 1969.

19 Moni Naor and Adi Shamir. Visual Cryptography. In *Advances in Cryptology - EUROCRYPT '94, Workshop on the Theory and Application of Cryptographic Techniques, Perugia, Italy, May 9-12, 1994, Proceedings*, pages 1–12, 1994.

20 Noam Nisan and Mario Szegedy. On the Degree of Boolean Functions as Real Polynomials. *Computational Complexity*, 4:301–313, 1994.

21 Ramamohan Paturi. On the Degree of Polynomials that Approximate Symmetric Boolean Functions (Preliminary Version). In *Proceedings of the 24th Annual ACM Symposium on Theory of Computing, May 4-6, 1992, Victoria, British Columbia, Canada*, pages 468–474, 1992.

22 Sushant Sachdeva and Nisheeth K. Vishnoi. Faster Algorithms via Approximation Theory. *Foundations and Trends in Theoretical Computer Science*, 9(2):125–210, 2014.

23 Rocco A. Servedio, Li-Yang Tan, and Justin Thaler. Attribute-Efficient Learning and Weight-Degree Tradeoffs for Polynomial Threshold Functions. In *COLT 2012 - The 25th Annual Conference on Learning Theory, June 25-27, 2012, Edinburgh, Scotland*, pages 14.1–14.19, 2012.

24 Alexander A. Sherstov. Approximating the AND-OR Tree. *Theory of Computing*, 9:653–663, 2013.

25 Alexander A Sherstov. Breaking the Minsky–Papert Barrier for Constant-Depth Circuits. *SIAM Journal on Computing*, 47(5):1809–1857, 2018.

26 Alexander A. Sherstov. The Power of Asymmetry in Constant-Depth Circuits. *SIAM J. Comput.*, 47(6):2362–2434, 2018.

27 Alexander A Sherstov and Pei Wu. Near-Optimal Lower Bounds on the Threshold Degree and Sign-Rank of AC^0. *arXiv preprint arXiv:1901.00988*, 2019. To appear in STOC 2019.

28 Robert Špalek. A Dual Polynomial for OR. *CoRR*, abs/0803.4516, 2008.

A Proof of Claim 18

The objective is to uniformly bound the value of the function

$$h_\delta(s) = C_w^2 \cdot \prod_{z \in Z_w} h_\delta(s,z), \qquad \text{where} \qquad h_\delta(s,z) = (s-z)^2 + \delta(1-z^2)$$

for $s \in [-1,1]$. When k, K are fixed and n becomes large, all zeros in Z_w approach -1 or $+1$, $h_\delta(s,z)$ uniformly approaches $h_0(s,z) = (s-z)^2$, $h_w(s)$ approaches $h_0(s) = p_w^\infty(s)$ and is therefore uniformly bounded.

The main difficulty in extending this argument to finite n is that $h_\delta(s,z)$ can no longer be uniformly bounded by a multiple of $(s-z)^2$ since when s equals z, the latter function vanishes but the former one doesn't. For this reason, we divide the analysis into two parameter regimes. When s is bounded away from the set of zeros Z_w, an approximation of the infinitary term-by-term argument can be carried out. When s is near the zeroes, we argue that $h_\delta(s)$ cannot be much larger than $h_\delta(s_0)$ for an s_0 that is even farther away from Z_w, and then argue that $h_0(s_0) = p_w(s_0)^2$ must be small because it represents the square of a probability of a rare event.

▶ **Fact 21.** *$h_\delta(s,z)h_\delta(s,-z) = h_\delta(-s,z)h_\delta(-s,-z)$.*

▶ **Fact 22.** *$h_\delta(s,z) \le h_\delta(|s|,z)$ when $z \le 0$ and $s \ge 0$.*

▶ **Fact 23.** *$h_\delta(s,z) \le h_\delta(s_0,z)$ when $s_0 \le s \le 1$, $s_0 \le 2z-1$, and $|z| \le 1$.*

Proof. The fact is equivalent to checking that $(s_0-z)^2 - (s-z)^2 \ge 0$ when $s_0 \le s \le 1$ and $s_0 \le 2z-1$. If $s \le z$ then we have that $s_0 \le s \le z$ from which it immediately follows that $(s_0-z)^2 \ge (s-z)^2$. If $s > z$ then $(s-z)^2$ is at most $(1-z)^2$. However, since $|z| \le 1$, we have that $s_0 \le 2z-1 \le z$ and thus $(s_0-z)^2$ is always at least $(z-(2z-1))^2 = (1-z)^2$. Again we have that $(s_0-z)^2 \ge (s-z)^2$. ◀

We begin by reducing to the case of non-negative inputs $s \in [0,1]$.

▷ Claim 24. Assuming $w \le K/2$, $h_\delta(s) \le h_\delta(|s|)$.

Proof. When $w \le K/2$ then elements of Z_w (3) can be split into w pairs of the form $A = \{(-1+2h/n, 1-2h/n) \colon 0 \le h < w\}$, and $K-2w$ remaining elements $B = \{-1 + 2h/n \colon w \le h < K-w\}$ are all non-positive. By Fact 21, $\prod_{(-z,z)\in A} h_\delta(s,z)h_\delta(s,-z) = \prod_{(-z,z)\in A} h_\delta(|s|,z)h_\delta(|s|,-z)$. By Fact 22, $\prod_{z\in B} h_\delta(s,z) \le \prod_{z\in B} h_\delta(|s|,z)$. Therefore the product $\prod_{z\in Z_w} h_\delta(s,z) \le \prod_{z\in Z_w} h_\delta(|s|,z)$. ◁

The following claim handles values of s in the range $[0, 1-w/16K]$.

▷ Claim 25. Assuming $0 \leq s \leq 1 - w/16K$,

$$h_\delta(s,z) \leq \begin{cases} (1+\delta)(s-z)^2, & \text{if } z \leq -1/\sqrt{2}. \\ (1+(64K/w)\delta)(s-z)^2, & \text{if } z \geq 1 - w/32K \end{cases}$$

Proof. The ratio $h_\delta(s,z)/(s-z)^2$ equals $1+((1-z^2)/(s-z)^2)\delta$. The number $(1-z^2)/(s-z)^2$ is at most 1 when $s \geq 0$ and $z \leq -1/\sqrt{2}$ and at most the following when $z \geq 1 - w/32K$.

$$\frac{1-(1-w/32K)^2}{((1-w/16K)-(1-w/32K))^2} \leq \frac{2w/32K}{(w/32K)^2} = 64K/w. \qquad \lhd$$

▶ **Corollary 26.** *Assuming* $0 \leq s \leq 1 - w/16K$ *and* $n \geq 64K$, $h_\delta(s) \leq e^{65\delta K} h_0(s)$.

Proof. By the choice of parameters, all zeros in Z_- meet the criterion for the first inequality in Claim 25, while all zeros in Z_+ meet the criterion for the second one. Therefore

$$\begin{aligned} h_\delta(s) &= C_w^2 \prod_{z\in Z_-} h_\delta(s,z) \prod_{z\in Z_+} h_\delta(s,z) \\ &\leq C_w^2 \prod_{z\in Z_-} (1+\delta)(s-z)^2 \prod_{z\in Z_+} (1+(64K/w)\delta)(s-z)^2 \\ &\leq (1+\delta)^{K-w}(1+(64K/w)\delta)^w \cdot C_w^2 \prod_{z\in Z_-} h_0(s,z) \prod_{z\in Z_+} h_0(s,z) \\ &\leq e^{\delta K} \cdot e^{64\delta K} \cdot h_0(s). \end{aligned}$$

◀

The following two claims handle values of s in the range $[1 - w/16K, 1]$.

▷ Claim 27. Assuming $w \leq K$ and $1 - w/8K \leq s_0 \leq 1 - w/16K \leq s \leq 1$,

$$h_\delta(s,z) \leq \begin{cases} h_\delta(s_0,z), & \text{if } z \geq 1 - w/32K \\ (1+w/8K)^2 \cdot h_\delta(s_0,z), & \text{if } z \leq -w/8K. \end{cases}$$

Proof. By the choice of parameters the first inequality follows from Fact 23. For the second one, we upper bound the ratio

$$\frac{(s-z)^2}{(s_0-z)^2} \leq \frac{(1-z)^2}{(1-z-w/8K)^2} = \left(1 + \frac{w/8K}{1-z-w/8K}\right)^2 \leq \left(1+\frac{w}{8K}\right)^2.$$

This is greater than one, so $(s-z)^2 + \delta(1-z^2) \leq (1+w/8K)^2((s_0-z)^2 + \delta(1-z^2))$ as desired. ◁

▶ **Corollary 28.** *Assuming* $1 - w/8K \leq s_0 \leq 1 - w/16K \leq s \leq 1$ *and* $n \geq 2K$, $h_\delta(s) \leq e^{w/4} h_\delta(s_0)$.

Proof. By the choice of parameters, all zeros in Z_- meet the criterion for the first inequality in Claim 27, while all zeros in Z_+ meet the criterion for the second one. Therefore

$$\begin{aligned} h_\delta(s) &= C_w^2 \prod_{z\in Z_-} h_\delta(s,z) \prod_{z\in Z_+} h_\delta(s,z) \\ &\leq C_w^2 \prod_{z\in Z_-} (1+w/8K)^2 \cdot h_\delta(s_0,z) \prod_{z\in Z_+} h_\delta(s_0,z) \\ &= (1+w/8K)^{2|Z_-|} \cdot h_\delta(s_0) \\ &\leq (1+w/8K)^{2K} \cdot h_\delta(s_0) \leq e^{w/4} h_\delta(s_0). \end{aligned}$$

◀

▷ Claim 29. If s_0 is of the form $1 - 2h/n$ for some integer $0 \leq h \leq wn/e^2K$ then $0 \leq p_w(s_0) \leq e^{-w}$.

Proof. By Fact 11, $p_w(s_0)$ is the probability that a random string of Hamming weight h and length n has exactly w ones in its first K positions. The probability that it has at least w ones in its first K positions is at most

$$\binom{K}{w} \cdot \frac{h}{n} \cdot \frac{h-1}{n-1} \cdots \frac{h-w+1}{n-w+1} \leq \left(\frac{eK}{w}\right)^w \left(\frac{h}{n}\right)^w \leq e^{-w}. \qquad \lhd$$

Proof of Claim 18. By Claim 24 we may assume $s \in [0,1]$. When $0 \leq s \leq 1 - w/16K$ the result follows from Corollary 26. When $1 - w/16K \leq |s| \leq 1$, by the assumption $n \geq 64K$ there must exist a value s_0 between $1 - w/8K$ and $1 - w/16K$ that is of the form $1 - 2h/n$. In particular $h \leq wn/e^2K$. Then

$$h_\delta(s) \leq e^{w/4} h_\delta(s_0) \leq e^{w/4} e^{65\delta K} p_w(s_0)^2 \leq e^{65\delta K - 7w/4},$$

where the inequalities follow from Corollary 28, Corollary 26, and Claim 29, respectively. ◁

B Proofs of Corollary 4 and Theorem 5

B.1 Proof of Corollary 4

Proof of Corollary 4. Corollary 3 implies the existence of a $\phi \left(= \frac{\mu - \nu}{2}\right)$ satisfying $\|\phi\|_1 = 1$, $\langle f, \phi \rangle = \varepsilon$ for some $\varepsilon = \Omega(1)$ and $\langle \phi, q \rangle \leq K^{3/2} \cdot 2^{-\Omega\left(\widetilde{\deg}_{1/3}(f)^2/K\right)}$ for any parity of degree at most K.

For any p of degree K and weight at most w,

$$\|f - p\|_\infty = \|f - p\|_\infty \|\phi\|_1 \geq \langle \phi, f - p \rangle = \langle \phi, f \rangle - \langle \phi, p \rangle \geq \varepsilon - w \cdot K^{3/2} \cdot 2^{-\Omega\left(\widetilde{\deg}_{1/3}(f)^2/K\right)}.$$

Thus, we conclude that $W_{\varepsilon/2}(f, K) = K^{-3/2} \cdot 2^{\Omega\left(\widetilde{\deg}_{1/3}(f)^2/K\right)}$. Corollary 4 now follows using standard error reduction techniques that show that $\widetilde{\deg}_\varepsilon(f) = \Theta(\widetilde{\deg}_{1/3}(f))$ for all constants $0 < \varepsilon < 1/2$. ◀

B.2 Proof of Theorem 5

We first require the following lemma. This lemma builds on ideas in [23, Claim 2], which showed a similar result for $t = \Theta(1)$.

▶ **Lemma 30.** *For any $y \in \{0,1\}^n$, denote by EQ_y the function on $\{0,1\}^n$ that outputs 1 on input y, and 0 otherwise. Then for any $t > 0$ and $d > \sqrt{nt \log n}$, we have $W_{n^{-O(t)}}(\mathsf{EQ}_y, d) \leq 2^{O(nt \log^2(n)/d)}$.*

Proof. Note that for any $y \in \{-1,1\}^n$, the function EQ_y is just the AND function on n input bits (with 0-1 valued output), with possibly negated input variables. Thus it suffices to give an approximating polynomial for the AND function on n bits. We now express AND_n as $\mathsf{AND}_\ell \circ \mathsf{AND}_{n/\ell}$, where ℓ is a parameter we will set later. We compute the inner $\mathsf{AND}_{n/\ell}$ exactly and approximate the outer AND_ℓ to error $n^{-\Omega(t)}$. This can be done with a polynomial p of degree $O\left(\sqrt{\ell \log(n^t)}\right)$ [15, 6]. Combining the fact that p is bounded by $1 + n^{-\Omega(t)} \leq 2$ at all Boolean inputs with Parseval's identity and the Cauchy-Schwarz inequality, it can be seen that the weight of p is at most $\ell^{O\left(\sqrt{\ell\log(n^t)}\right)}$.[12] It is well known that the exact multilinear polynomial representation of $\mathsf{AND}_{n/\ell}$ has constant weight. Hence, by composing p with the multilinear polynomial that exactly computes $\mathsf{AND}_{n/\ell}$, we obtain an approximation q for AND_n of degree $O\left(n\sqrt{\frac{t\log n}{\ell}}\right)$, error $n^{-\Omega(t)}$, and weight $2^{O\left(\sqrt{\ell t\log^3 n}\right)}$. We now fix the value of ℓ to $\ell := \frac{n^2 t\log n}{d^2} < n$, thereby ensuring that the degree of q is at most d. With this setting of ℓ, the weight of q is at most $2^{O(nt\log^2(n)/d)}$, proving the lemma. ◀

Proof of Theorem 5. Let $f : \{0,1\}^n \to \{0,1\}$ be any symmetric function, corresponding to the univariate predicate $D_f : \{0\} \cup [n] \to \{0,1\}^n$. For the purpose of this proof, let us denote by k_f the smallest i for which f is constant on inputs of Hamming weight in the interval $[i+1, n-i-1]$. Without loss of generality, $f(x) = 0$ for strings of x Hamming weight between $k_f + 1$ and $n - k_f - 1$. The case where $f = 1$ on input strings of Hamming weight between $k_f + 1$ and $n - k_f - 1$ can be proved using a similar argument. Define $\mathrm{supp}(f) := \{x \in \{0,1\}^n : f(x) = 1\}$. Note that $|\mathrm{supp}(f)| \le 2 \cdot n^{k_f}$.

Observe that $f(x) = \sum_{y\in\mathrm{supp}(f)} \mathsf{EQ}_y(x)$. Lemma 30 implies, for each $y \in \mathrm{supp}(f)$, the existence of polynomials p_y of degree K and weight $2^{O(nk_f\log^2(n)/K)}$, which approximate EQ_y to error $\frac{1}{6}\cdot n^{-k_f}$. Define a polynomial $p\colon \{0,1\}^n \to \mathbb{R}$ by $p(x) = \sum_{y\in\mathrm{supp}(f)} p_y(x)$. Clearly p has degree K, weight at most $n^{O(k_f)} \cdot 2^{O(nk_f\log^2(n)/K)} = 2^{\tilde{O}(nk_f/K)}$, and error at most $|\mathrm{supp}(f)| \cdot n^{-k_f}/6 \le 1/3$, where the upper bounds on the weight and error follow from the triangle inequality.

The theorem now follows standard error reduction techniques and Paturi's theorem [21], which states that for symmetric functions, $\widetilde{\deg}(f) = \Theta\left(\sqrt{n\cdot k_f}\right)$. ◀

▶ Remark 31. The upper bound obtained in Theorem 5 is more general than as stated, and the only property of symmetric functions it exploits is that symmetric functions of low approximate degree are highly biased. More specifically, the proof of Theorem 5 shows that any function $f\colon \{0,1\}^n \to \{0,1\}$ with $\min\{|f^{-1}(0)|, |f^{-1}(1)|\} \le n^t$ satisfies $W_\varepsilon(f,K) \le 2^{\tilde{O}(nt/K)}$ for any $K \ge \sqrt{nt\log n}$.

C Proof of Theorem 6

Proof outline. As we explain in more detail in the proof itself, it is sufficient to establish the theorem for fixed k and K and infinitely many n because the statement is downward reducible in n.

Using the Chebyshev approximation formulas from Section 3 we derive explicit lower bounds on the large Chebyshev coefficients on the polynomial p_0 representing the distinguishing advantage of the AND function on K inputs. Owing to orthogonality and boundedness of the Chebyshev polynomials, this is a lower bound on the approximate degree of AND_K. By strong duality as given in the following Claim (see [4]) we obtain Theorem 6.

▷ Claim 32. If $\widetilde{\deg}_{\varepsilon/2}(F_n) \ge k$ then there exists a pair of perfectly k-wise indistinguishable distributions μ, ν over $\{0,1\}^n$ such that $\mathbb{E}_{X\sim\mu}[F_n(X)] - \mathbb{E}_{Y\sim\nu}[F_n(Y)] \ge \varepsilon$.

[12] Building on [6], It is possible to derive explicit ε-approximating polynomials for AND where the degree is $O\left(\sqrt{\ell\log(1/\varepsilon)}\right)$ and the weight is $2^{O\left(\sqrt{\ell\log(1/\varepsilon)}\right)}$ rather than $\ell^{O\left(\sqrt{\ell\log(1/\varepsilon)}\right)}$. Using this tighter weight bound would improve our final result by a factor of $\log n$ in the exponent. We omit this tighter result for brevity.

Recall that the Chebyshev polynomials are orthogonal under the measure $d\sigma(t) = (1-t^2)^{-1/2}dt$ supported on $[-1,1]$. We will need the following identity for their average square magnitude under this measure:

$$\mathbb{E}_{t\sim\sigma}[T_d(t)^2] = 1/2 \qquad \text{when } d > 0. \tag{8}$$

Proof of Theorem 6. By symmetry of the distinguishers, μ and ν can be assumed symmetric. Let F_n denote the function on $\{0,1\}^n$ that outputs $\mathsf{AND}_K\left(x|_{\{1,\dots,K\}}\right)$, i.e., F_n outputs the AND of the first $K < n$ bits of the input. We prove the theorem for $G_n(x_1,\dots,x_n) = \mathsf{NOR}(x|_{\{1,\dots,K\}})$. By the symmetry of 0 and 1 inputs the theorem also holds for F_n.

First, we claim that the statement of Theorem 6 is stronger as n becomes larger, so it is sufficient to prove it in the limiting case when n approaches infinity and k, K are fixed. Suppose that μ and ν are distributions over n bit strings that are k-wise indistinguishable yet are ε-reconstructable by G_n. We must show that there are distributions μ' and ν' over $\{0,1\}^{n-1}$ are k-wise indistinguishable yet are ε-reconstructable by G_{n-1}. But this holds for μ' (respectively ν') that generate a random sample from μ (respectively, ν) and then throw away the last bit.

If the statement was false then by Claim 32 there would exist degree-k polynomials $\tilde{G}_n$ that approximate G_n pointwise on $\{0,1\}^n$ to within error $\varepsilon = \sqrt{2^{-4K+1}\sum_{d>K}\binom{2K}{K+d}^2}$ for almost all n. Applying the construction from the proof of Fact 11 to $\tilde{G}_n$, there exist univariate degree-k polynomials $\tilde{p}_0^n$ approximating p_0^n on the set of points $W_n = \{-1+2h/n \colon 0 \le h \le n\}$ to within error ε. We emphasize the dependence on n as it will play a role in the proof.

By Formula (3) the polynomial p_0^n has the form

$$p_0^n(t) = C_0^n \prod_{z\in Z_0^n} (t-z),$$

where $Z_0^n = \{-1+2h/n \colon 0 \le h < K\}$ (the set Z_+ is empty). The value $p_n^0(1)$ is the probability that G_n accepts the all-zero string, so it must equal one. The constant C_0^n must therefore equal $\prod_{z\in Z_0^n}(1-z)^{-1}$. As n tends to infinity, the set Z_0 converges to a single zero at -1 of multiplicity K, so the sequence p_0^n converges uniformly to the polynomial

$$p_0^\infty(t) = 2^{-K}(t+1)^K.$$

By the triangle inequality, for every $\delta > 0$ and all sufficiently large n, $\tilde{p}_0^n$ is within $\varepsilon+\delta$ of p_0^∞ on the set W_n. A degree-k polynomial is determined by its values on W_{k+1} and the set of degree-k polynomials that are within $\varepsilon+\delta$ of p_0^∞ on W_{k+1} is compact. Therefore the sequence of approximating polynomials $\tilde{p}_0^n$ must contain a subsequence (for values of n that are multiples of $k+1$) that converges (uniformly) to a limiting degree-k polynomial $\tilde{p}_0^\infty$. Since $\tilde{p}_0^n$ is within $\varepsilon+\delta$ of p_0^n on W_n for infinitely many n, $\tilde{p}_0^\infty$ must be within $\varepsilon+2\delta$ of p_0^∞ on W_n for infinitely many n. The union of these sets W_n is dense in $[-1,1]$, and by continuity p_0^∞ can be $\varepsilon+\delta$-approximated by the degree-k polynomial $\tilde{p}_0^\infty$ everywhere on $[-1,1]$. As δ was arbitrary it follows that the ε-approximate degree of p_0^∞ can be at most k.

All that remains to prove that this is not true, i.e., to show a lower bound of k on the ε-approximate degree of p_0^∞. This lower bound is known (see, e.g., [13]); we provide the details in the full version. ◀