Electronic Colloquium on Computational Complexity, Report No. 46 (2020)

A Robust Version of Hegedűs's Lemma, with Applications

Srikanth Srinivasan*

Abstract

Hegedűs's lemma is the following combinatorial statement regarding polynomials over finite fields. Over a field $\mathbb{F}$ of characteristic $p > 0$ and for q a power of p, the lemma says that any multilinear polynomial $P \in \mathbb{F}[x_1, \ldots, x_n]$ of degree less than q that vanishes at all points in $\{0,1\}^n$ of Hamming weight $k \in [q, n-q]$ must also vanish at all points in $\{0,1\}^n$ of weight $k+q$. This lemma was used by Hegedűs (2009) to give a solution to *Galvin's problem*, an extremal problem about set systems; by Alon, Kumar and Volk (2018) to improve the best-known multilinear circuit lower bounds; and by Hrubeš, Ramamoorthy, Rao and Yehudayoff (2019) to prove optimal lower bounds against depth-2 threshold circuits for computing some symmetric functions.

In this paper, we formulate a robust version of Hegedűs's lemma. Informally, this version says that if a polynomial of degree $o(q)$ vanishes at most points of weight k, then it vanishes at many points of weight $k+q$. We prove this lemma and give the following three different applications.

- Degree lower bounds for the coin problem: The δ-*Coin Problem* is the problem of distinguishing between a coin that is heads with probability $((1/2)+\delta)$ and a coin that is heads with probability $1/2$. We show that over a field of positive (fixed) characteristic, any polynomial that solves the δ-coin problem with error ε must have degree $\Omega(\frac{1}{\delta}\log(1/\varepsilon))$, which is tight up to constant factors.
- Probabilistic degree lower bounds: The *Probabilistic degree* of a Boolean function is the minimum d such that there is a random polynomial of degree d that agrees with the function at each point with high probability. We give tight lower bounds on the probabilistic degree of *every* symmetric Boolean function over positive (fixed) characteristic. As far as we know, this was not known even for some very simple functions such as unweighted Exact Threshold functions, and constant error.
- A robust version of the combinatorial result of Hegedűs (2009) mentioned above.

1 Introduction

The Polynomial Method is a technique of great utility in both Theoretical Computer Science and Combinatorics. The idea of associating polynomials with various combinatorial objects and then using algebraic or geometric techniques to analyze them has proven useful in many settings including, but not limited to, Computational Complexity (Circuit lower bounds [Raz87, Smo87a, Bei93, Wil14c], Pseudorandom generators [Bra10]), Algorithm design (Learning Algorithms [LMN93, KS04, KOS04], Satisfiability algorithms [Wil14c, Wil14b], Combinatorial algorithms [Wil18, AWY15, AW15]), and Extremal Combinatorics [Gut16, CLP17, EG17].

The engine that drives the proofs of many of these results is our understanding of combinatorial and algebraic properties of polynomials. In this paper, we investigate another such naturally stated property of polynomials defined over the Boolean cube $\{0,1\}^n$ and strengthen known results in this direction. We then apply this result to sharpen known results in theoretical computer science and combinatorics.

*Department of Mathematics, Indian Institute of Technology Bombay, Mumbai, India. Email: srikanth@math.iitb.ac.in. Supported by MATRICS grant MTR/2017/000958 awarded by SERB, Government of India.

ISSN 1433-8092

The question we address is related to how well low-degree polynomials can 'distinguish' between different layers of the Boolean cube $\{0,1\}^n$. For $m \in \{0,\ldots,n\}$, let $\{0,1\}^n_m$ be the elements of $\{0,1\}^n$ of Hamming weight exactly m. As a first approximation, let us say that a polynomial $P \in \mathbb{F}[x_1,\ldots,x_n]$ (here $\mathbb{F}$ is some field) distinguishes between level sets $\{0,1\}^n_k$ and $\{0,1\}^n_K$ if it vanishes at all points in the former set and at no point of the latter. Note that the ability of low-degree polynomials to do this depends on the properties of the underlying field $\mathbb{F}$: when $\mathbb{F} = \mathbb{Q}$ (or any field of characteristic 0), the simple polynomial $(\sum_{i=1}^n x_i) - k$ does the job. However, if the field $\mathbb{F}$ has positive characteristic p and more specifically if $K - k$ is divisible by p, then this simple polynomial no longer works and the answer is not so clear.

In this setting, a classical theorem of Lucas tells us that if q is the largest power of p dividing $K - k$, then there is a polynomial of degree q that distinguishes between $\{0,1\}^n_k$ and $\{0,1\}^n_K$. A very interesting lemma of Hegedűs [Heg09] shows that this is tight even if we only require P to be non-zero at *some* point of $\{0,1\}^n_K$. More precisely, Hegedűs's lemma shows the following.[1]

Lemma 1 (Hegedűs's lemma)**.** *Let $\mathbb{F}$ be a field of characteristic $p > 0$. Fix any positive integers n, k, q such that $k \in [q, n-q]$, and q a power of p. If $P \in \mathbb{F}[x_1,\ldots,x_n]$ is any multilinear polynomial that vanishes at* all *$a \in \{0,1\}^n_k$ but does not vanish at* some *$b \in \{0,1\}^n_{k+q}$, then* $\deg(P) \geq q$.

This lemma was first proved in [Heg09] using Gröbner basis techniques. An elementary proof of this was recently given by the author and independently by Alon (see [HRRY19]) using the Combinatorial Nullstellensatz.

Hegedűs's lemma has been used to resolve various questions in both combinatorics and theoretical computer science.

- Hegedűs used this lemma to give an alternate solution to a problem of Galvin, which is stated as follows. Given a positive integer n divisible by 4, what is the smallest size $m = m(n)$ of a family $\mathcal{F}$ of $(n/2)$-sized subsets of $[n]$ such that for any $S \subseteq [n]$ of size $n/2$, there is a $T \in \mathcal{F}$ with $|T \cap S| = n/4$? It is easy to see that $m(n) \leq n/2$ for any n. A matching lower bound was given by Enomoto, Frankl, Ito and Nomora [EFIN87] in the case that $t := (n/4)$ is odd. Hegedűs used the above lemma to give an alternate proof of a lower bound of n in the case that t is an odd prime. His proof was subsequently strengthened to a linear lower bound for all t by Alon et al. [AKV18] and more recently to a near-tight lower bound of $(n/2) - o(n)$ for all t by Hrubeš et al. [HRRY19]. Both these results used the lemma above.

- Alon et al. [AKV18] also used Hegedűs's lemma to prove bounds for generalizations of Galvin's problem. Using this, they were able to prove improved lower bounds against *syntatically multilinear algebraic circuits.* These are algebraic circuits that compute multilinear polynomials in a "transparently multilinear" way (see e.g. [SY10] for more). Alon et al. used Hegedűs's lemma to prove near-quadratic lower bounds against syntactically multilinear algebraic circuits computing certain explicitly defined multilinear polynomials, improving on an earlier $\tilde{\Omega}(n^{4/3})$ lower bound of Raz, Shpilka and Yehudayoff [RSY08].

- Hrubeš et al. [HRRY19] also used Hegedűs's lemma to answer the following question of Kulikov and Podolskii [KP17] on depth-2 threshold circuits. What is the smallest $k = k(n)$ such that there is a depth-2 circuit made up of Majority[2] gates of fan-in at most k that computes the Majority function on n bits? Using Hegedűs's lemma, Hrubeš et al. showed an asymptotically tight lower bound of $n/2 - o(n)$ on $k(n)$.

[1] The lemma is usually stated [Heg09, AKV18, HRRY19] for a more restricted choice of parameters. However, the known proofs extend to yield the stronger statement given here.

[2] The Majority function is the Boolean function f which accepts exactly those inputs that have more 1s than 0s.

Main Result. Our main result in this paper is a 'robust' strengthening of Hegedűs's lemma. Proving 'robust' or 'stability' versions of known results is standard research direction in combinatorics. Such questions are usually drawn from the following template. Given the fact that objects that satisfy a certain property have some fixed structure, we ask if a similar structure is shared by objects that 'almost' or 'somewhat' satisfy the property.

In our setting, we ask if we can recover the degree lower bound in Hegedűs's lemma even if we have a polynomial P that 'approximately' distinguishes between $\{0,1\}^n_k$ and $\{0,1\}^n_{k+q}$: this means that the polynomial P vanishes at 'most' points of weight k but is non-zero at 'many' points of weight $k+q$. Our main lemma is that under suitable definitions of 'most' and 'many', we can recover (up to constant factors) the same degree lower bound as in Lemma 1 above.

Lemma 2 (Main Result (Informal))**.** *Assume that $\mathbb{F}$ is a field of characteristic p. Let n be a growing parameter and assume we have positive integer parameters k, q such that $100q < k < n - 100q$ and q is a power of p. For $\varepsilon = \varepsilon(n,k,q)$, if $P \in \mathbb{F}[x_1, \ldots, x_n]$ that vanishes at a $(1-\varepsilon)$-fraction of points of $\{0,1\}^n_k$ but does not vanish at an $\varepsilon^{0.0001}$ fraction of points of $\{0,1\}^n_{k+q}$, then* $\deg(P) = \Omega(q)$.

Remark 3. *1. To keep the exposition informal, we have not specified exactly what ε is in the above lemma. However, we note below that the ε chosen is nearly the best possible in the sense that if ε is appreciably increased, then there is a sampling-based construction of a polynomial P of degree $o(q)$ satisfying the hypothesis of the above lemma (see Section 3.3).*

2. The reader might wonder why the lemma above is a strengthening of Hegedűs's lemma, given that we require the polynomial P to be non-zero at many points of weight $k+q$, which is a seemingly stronger condition than required in Lemma 1. However, this is in fact a weaker condition. This is because of the following simple algebraic fact: if there is a polynomial P of degree at most d satisfying the hypothesis of Lemma 1 (i.e. vanishing at all points of weight k but not at some point of weight $k+q$), then there is also a polynomial Q of degree at most d that vanishes at all points of weight k but does not vanish at a significant fraction *(at least a $(1-1/p)$ fraction) of points of weight $k+q$. We give a short proof of this in Appendix A. Hence, the above lemma is indeed a generalization of Lemma 1 (up to the constant-factor losses in the degree lower bound).*

Applications. Our investigations into robust versions of Hegedűs's lemma were motivated by questions in computational complexity theory. Using our main result, we are able to sharpen and strengthen known results in complexity as well as combinatorics.

1. **Degree bounds for the Coin Problem:** For a parameter $\delta \in [0, 1/2]$, we define the *δ-coin problem* as follows. We are given N independent tosses of a coin, which is promised to either be of bias $1/2$ (i.e. unbiased) or $(1/2) - \delta$, and we are required to guess which of these is the case with a high degree of accuracy, say with error probability at most ε.

   The coin problem has been studied in a variety of settings in complexity theory (see, e.g. [ABO84, Val84, Vio09, SV10, BV10, CGR14]) and for various reasons such as understanding the power of randomness in bounded-depth circuits, the limitations of blackbox hardness amplification, and devising pseudorandom generators for bounded-width branching programs. More recently, Limaye et al. [LSS+18] proved optimal lower bounds on the size of $\mathrm{AC}^0[\oplus]$ [3] circuits solving the δ-coin problem with constant error, strengthening an earlier lower bound of Shaltiel and Viola [SV10]. This led to the first class of explicit functions for which we have tight (up to polynomial factors) $\mathrm{AC}^0[\oplus]$ lower bounds. These bounds were in turn used by Golovnev, Ilango, Impagliazzo, Kabanets,

[3] Recall that these are bounded-depth circuits made up of AND, OR and $\oplus$ gates.

Kolokolova and Tal [GII+19] to resolve a long-standing open problem regarding the complexity of MCSP in the $AC^0[\oplus]$ model, and by Potukuchi [Pot19] to prove lower bounds for Andreev's problem.

A key result in the lower bound of Limaye et al. [LSS+18] was a tight lower bound on the degree of any polynomial $P \in \mathbb{F}[x_1, \ldots, x_N]$ that solves the δ-coin problem with constant error: they showed that any such polynomial P must have degree at least $\Omega(1/\delta)$. As noted by Agrawal [Agr19], this is essentially equivalent to a recent result of Chattopadhyay, Hatami, Lovett and Tal [CHLT19] on the level-1 Fourier coefficients of low-degree polynomials over finite fields, which in turn is connected to an intriguing new approach [CHLT19] toward constructing pseudorandom generators secure against $AC^0[\oplus]$.

Using the robust Hegedűs lemma, we are able to strengthen the degree lower bound of [LSS+18] to a tight degree lower bound for *all errors*. Specifically, we show that over any field $\mathbb{F}$ of fixed positive characteristic p, any polynomial P that solves the δ-coin problem with error ε must have degree $\Omega(\frac{1}{\delta}\log(1/\varepsilon))$, which is tight for all δ and ε.

2. **Probabilistic degrees of symmetric functions:** In a landmark paper [Raz87], Razborov showed how to use polynomial approximations to prove lower bounds against $AC^0[\oplus]$. The notion of polynomial approximation introduced (implicitly) in his result goes by the name of *probabilistic polynomials,* and is defined as follows. An ε-error probabilistic polynomial of degree d for a Boolean function $f : \{0,1\}^n \to \{0,1\}$ is a random multilinear polynomial $\boldsymbol{P}$ of degree at most d that agrees with f at each point with probability at least $1 - \varepsilon$. The ε-error probabilistic degree of f is the least d for which this holds. (Roughly speaking, a low-degree probabilistic polynomial for f is an efficient randomized algorithm for f, where we think of polynomials as algorithms and degree as a measure of efficiency.)

   Many applications of polynomial approximation in complexity theory [Bei93] and algorithm design [Wil14a] use probabilistic polynomials and specifically bounds on the probabilistic degrees of various *symmetric* Boolean functions.[4] Motivated by this, in a recent result with Tripathi and Venkitesh [STV19], we gave a near-tight characterization on the probabilistic degree of every symmetric Boolean function. Unfortunately, however, our upper and lower bounds were separated by logarithmic factors. This can be crucial: in certain algorithmic applications (see, e.g., [AW15, Footnote, Page 138]), the appearance or non-appearance of an additional logarithmic factor in the degree can be the difference between (say) a truly subquadratic running time of $N^{2-\varepsilon}$ and a running time of $N^{2-o(1)}$, which might be less interesting.

   In the case of characteristic 0 (or growing with n), such gaps look hard to close since we don't even understand completely the probabilistic degree of simple functions like the OR function [MNV16, HS16, BHMS18]. However, in positive (fixed) characteristic, there are no obvious barrriers. Yet, even in this case, the probabilistic degree of very simple symmetric Boolean functions like the *Exact Threshold functions* (functions that accept inputs of exactly one Hamming weight) remained unresolved until this paper.

   In this paper, we resolve this question and more. We are able to give a tight (up to constants) lower bound (matching the upper bounds in [STV19]) on the probabilistic degree of *every* symmetric function over fields of positive (fixed) characteristic.

3. **Robust version of Galvin's problem:** Given that Hegedűs's lemma was used to solve Galvin's problem, it is only natural that we consider the question of using the robust version to solve a robust version of Galvin's problem. More precisely, we consider the minimum size $m = m(n, \varepsilon)$ to be the

[4] Recall that a Boolean function $f : \{0,1\}^n \to \{0,1\}$ is said to be symmetric if its output depends only on the Hamming weight of its input.

minimum size of a family $\mathcal{F}$ of $(n/2)$-sized subsets of $[n]$ such that for all but an ε-fraction of sets S of size $n/2$, there is a set $T \in \mathcal{F}$ such that $|S \cap T| = n/4$.

Following the proof of Galvin's theorem from Hegedűs's lemma, we can prove a lower bound of $\Omega(\sqrt{n \log(1/\varepsilon)})$ for the above version of Galvin's problem for any $\varepsilon \in [2^{-n}, 1/2]$. Note that this interpolates smoothly between a bound of $\Omega(\sqrt{n})$ for constant ε and $\Omega(n)$ for $\varepsilon = 2^{-\Omega(n)}$, both of which are tight. For general ε in between these two extremes, we do not know if our bounds are tight (we suspect they are). However, our bounds *are* tight for every ε for a natural generalization of the above problem, where we allow intersections of any size (and not just $n/4$). We refer the reader to Section 4.3 for details.

1.1 Proof Outline

The Main Lemma. Assume for simplicity that the error parameter ε is a constant and that $k+q = n/2$. In this case, the main lemma says the following. If $P \in \mathbb{F}[x_1, \ldots, x_n]$ is a polynomial that vanishes on most inputs of weight $(n/2) - q$ and is non-zero on most inputs of weight $n/2$, where $q = \Theta(\sqrt{n})$ is a power of p, then the degree of P is $\Omega(q)$.

We view the lemma as a strengthening of the classical lower bounds of Smolensky [Smo87a, Smo93]. Smolensky's results imply a tight lower bound on the degree of any polynomial that approximates certain Boolean functions such as the Majority function.[5] In our setting too, we are proving lower bounds on the degree of polynomials that approximate certain symmetric Boolean functions.[6] However, these are 'promise' symmetric Boolean functions: we require that they take the value 0 at Hamming weight $(n/2)-q$ and 1 at Hamming weight $n/2$. Clearly, the Majority function is one such function, but there are many others. (This increased flexibility of our lower bound will serve us well later, when we try to prove lower bounds on the probabilistic degrees of symmetric Boolean functions.)

Given the parallel to Smolensky's theorem, it is natural that we follow the strategy of Smolensky's [Smo87a] proof (we follow the 'dual' form of this proof strategy as in [ABFR94, KS18]). In the case of the Majority function, the crux of this proof is to show that if $E \subseteq \{0,1\}^n$ is a set of size at most $2^n/100$ consisting of points of weight less than $n/2$, then there is a polynomial Q of degree at most $(n/2) - \Omega(q)$ that vanishes on E, while at the same time not vanishing at *some* point a of Hamming weight more than $n/2$. This is easy to ensure: as $q = \Theta(\sqrt{n})$, the number of monomials of degree at most $D = (n/2) - \Omega(q)$ is greater than $|E|$, and standard linear algebra then yields that there is a non-zero polynomial Q of degree at most D that vanishes at all the points of E; further, Q cannot vanish at all points of weight greater than $n/2$, since it is known that any polynomial of degree at most D cannot vanish at all points of a Hamming ball of radius D.

In our setting, however, we end up having the stronger constraint that Q must not vanish at some point a of weight *exactly* $n/2$. This is not clear, since it is not true that every non-zero polynomial of degree at most D is non-zero at some point $a \in \{0,1\}^n_{n/2}$; in fact, this is not even true for polynomials of degree 1. However, since we are free to choose our polynomial of degree at most D as we wish, it is conceivable that we can avoid such bad polynomials. To do this, we try to understand the *degree-D closure* $\mathrm{cl}_D(E)$ of the set E, which is the set of points where any degree-D polynomial Q vanishing throughout E is forced to vanish. This is a well-studied object in coding theory [Wei91] and combinatorics [CL69, KS05, NW15], and we know how large $\mathrm{cl}_D(E)$ can be, given our upper bounds on $|E|$. We use a theorem of Nie and

[5] This is the symmetric function which accepts exactly those inputs that have weight at least $n/2$. Strictly speaking, Smolensky [Smo87a] proved lower bounds for the MOD^q functions, which accept inputs whose Hamming weight is divisible by q, where q is relatively prime to p. However, his proof can be easily adapted to the Majority function. As far as we know, this fact first appeared in Szegedy's PhD thesis [Sze89].

[6] Here, the polynomial approximates the Boolean function in the sense that it is zero at most zeroes of the Boolean function and non-zero at most ones of the Boolean function. This is similar to the notion of approximation considered in [Smo93].

Wang [NW15] to bound $|\mathrm{cl}_D(E)|$. Note that if $|\mathrm{cl}_D(E)| < \binom{n}{n/2}$, then we would be done as then there would be a point $a \in \{0,1\}^n_{n/2}$ that is not in the degree-D closure of E, meaning that there is a polynomial Q that vanishes at all points of E but not at a.

It turns out that $|\mathrm{cl}_D(E)|$ is not much larger than E in this setting, but this is still much larger than $|\{0,1\}^n_{n/2}| = \binom{n}{n/2}$, so this does not yet yield the kind of polynomial Q we need. However, this is where the characteristic p and the choice of q come in. We note that (by a polynomial construction from [Lu01, STV19]) there is a simple polynomial Q_1 of degree at most q that vanishes exactly at points of weights $w \not\equiv n/2 \pmod q$. Hence, it suffices to choose a polynomial Q_2 of degree at most $D' = D - q = (n/2) - \Theta(\sqrt{n})$ that vanishes at all points in

$$E' := E \cap \left(\bigcup_{\substack{j<n/2: \\ j \equiv n/2 \pmod q}} \{0,1\}^n_j \right).$$

Now, for the parameters we have chosen $|E'|$ is much smaller[7] than $\binom{n}{n/2}$ and hence an application of the result of Nie and Wang, yields that $|\mathrm{cl}_{D'}(E')| < \binom{n}{n/2}$. In particular, we thus get that there is a polynomial Q_2 of degree at most D' that vanishes at all points of E' but not at *some* point of weight $n/2$. Setting $Q = Q_1 \cdot Q_2$ yields the polynomial Q we need for the proof.

Degree lower bounds for the Coin Problem. The lower bounds for the Coin problem follow almost immediately from the statement of the robust Hegedűs lemma. Consider again the case of constant ε discussed above (note that this is for illustration only: tight lower bounds for solving the coin problem with constant error are already known from [LSS+18, CHLT19, Agr19]). Assume that there is a polynomial $Q(x_1, \ldots, x_N)$ of degree d for solving the δ-coin problem with error ε. We show how to use Q to distinguish between $\{0,1\}^n_{n/2}$ and $\{0,1\}^n_{(n/2)-q}$ where $q = \Theta(\sqrt{n})$ is a power of p and n is chosen suitably.

This is done by sampling. More precisely, imagine that we sample N uniformly random bits (chosen with replacement) from an n-bit input a and feed them into Q. If $|a| = n/2$, then the input to Q are independent uniformly distributed bits; if $|a| = (n/2) - q$, however, the input to Q is $((1/2) - (q/n))$-biased, which is $((1/2) - \delta)$-biased if we choose $n = \Theta(1/\delta^2)$. Thus, using the fact that Q solves the δ-coin problem, we obtain a *probabilistic polynomial* $\boldsymbol{P}$ on n variables of degree at most d that accepts inputs of weight $n/2$ with probability $1 - \varepsilon$ and rejects inputs of weight $((n/2) - q)$ with probability $1 - \varepsilon$. By averaging, there is a fixed polynomial P of degree at most d that accepts most inputs of weight $n/2$ and rejects most inputs of weight $((n/2) - q)$. However, the main lemma now implies that the degree of P must be at least $\Omega(\sqrt{n}) = \Omega(1/\delta)$. This implies the desired lower bound on d.

For smaller ε, this basic template remains the same, except that we choose $n = \Theta(\frac{1}{\delta^2}\log(1/\varepsilon))$.

Comparison with previous bounds. It may be worth noting that, given the robust Hegedűs lemma, even in the constant-error regime, this proof is quite a bit simpler than the corresponding proofs in [LSS+18, CHLT19]. The former required a considerable amount of computation, while to deduce the degree lower bound from the latter result, some non-trivial ideas are necessary [Agr19].

Furthermore, it is not clear how to use the proof methods of [LSS+18, CHLT19] to deduce a stronger lower bound depending on the error parameter ε. The informal reason for this is as follows. For a degree-d polynomial $Q(x_1, \ldots, x_N)$ solving the δ-coin problem with error ε, let us define $\Gamma_Q : [0,1] \to [0,1]$ so that $\Gamma_Q(\alpha)$ is the probability that Q accepts an N-bit input where each bit is set to 1 independently with

[7] Actually, this is not true for every set E of size $2^n/100$. However, it is true for, say, random sets E of this size. In the actual application, we will have that $|E \cap \{0,1\}^n_{n/2-q}|$ is quite small and this will be enough to guarantee that $|E'|$ is small.

probability α. The results of [CHLT19, LSS$^+$18] essentially imply that the derivative Γ'_Q is bounded pointwise by $O(d)$ in absolute value. As $|\Gamma_Q(1/2) - \Gamma_Q(1/2-\delta)| \geq (1-2\varepsilon) = \Omega(1)$, the mean-value theorem from Calculus implies that $d = \Omega(1/\delta)$. However, note that this proof cannot take advantage of the small error: it is not clear how to modify it to use the fact that $|\Gamma_Q(1/2) - \Gamma_Q(1/2-\delta)| \geq 1-2\varepsilon$.

Probabilistic Degree lower bounds for symmetric Boolean functions. In a recent paper with Tripathi and Venkitesh [STV19], we gave upper and lower bounds on the probabilstic degree of any Boolean function over any field $\mathbb{F}$. Unfortunately, however, these bounds differed from each other by logarithmic factors over *every* field, even for simple families of functions such as, say, the Exact Threshold function $\mathrm{EThr}_n^{n/2}$ which accepts only inputs of weight $n/2$. As far as we know, before this result, there was no proof (over any field) that the probabilistic degree of this function is $\Omega(\sqrt{n})$ for constant error.

Note that the robust Hegedűs lemma immediately implies such a lower bound for $\mathrm{EThr}_n^{n/2}$ over fixed positive characteristic. As noted above, the robust Hegedűs lemma implies that any polynomial P that accepts most inputs of weight $n/2$ and rejects most inputs of weight $(n/2)-q$ (where $q = \Theta(\sqrt{n})$ is a power of p) must have degree $\Omega(\sqrt{n})$. But any probabilistic polynomial $\boldsymbol{P}$ of degree d for $\mathrm{EThr}_n^{n/2}$ yields such a polynomial P (also of degree d) by averaging. This immediately implies that the probabilistic degree of $\mathrm{EThr}_n^{n/2}$ is $\Omega(\sqrt{n})$. Similarly, we can also obtain optimal lower bounds for Boolean functions that count modulo r for r that is relatively prime to p (previous lower bounds [Smo87a] were optimal when r is a constant, but not for growing r as far as we know).

We can extend these ideas to prove tight lower bounds for all symmetric Boolean functions f and all errors ε. To do this, we follow the proof ideas of [STV19] but with one crucial change. The bottleneck to proving tight lower bounds in [STV19] was that we used classical probabilistic degree lower bounds [Smo87a, Smo93] for the Majority and MOD^r functions as a starting point, with an overall strategy of 'reducing' one of these hard functions to the symmetric Boolean function f at hand. In this paper, our starting point is an equally strong lower bound for a much easier 'promise' symmetric function. More precisely, the robust Hegedűs lemma implies that any symmetric function h on n Boolean variables that takes different values at weights $n/2$ and $(n/2)-q$ where q is a power of p, must have probabilistic degree $\Omega(q)$. Note that this assumes very little about the function h (only its values at two different weights). Consequently, it becomes much easier to use this to prove lower bounds for other symmetric functions. Indeed, in all cases, by suitably restricting the symmetric function f, we are able to obtain a function h of the type above, leading to the desired lower bound. In particular, unlike in [STV19], we do not need to consider many distinct restrictions of f to construct the hard function, which makes it possible to avoid the log-factor losses in the lower bounds from that paper.

2 Preliminaries

We use the notation $[a,b]$ to denote an interval in $\mathbb{R}$ as well as an interval in $\mathbb{Z}$. The distinction will be clear from context.

Multilinear polynomials and Multilinearization. Fix any field $\mathbb{F}$. Throughout, we work with functions $f : \{0,1\}^n \to \mathbb{F}$ which are represented by multilinear polynomials. Recall that each such function has a *unique* multilinear polynomial representation. Further, given a (possibly non-multlinear) polynomial $P(x_1,\ldots,x_n)$ representing f (i.e. $P(a) = f(a)$ for all $a \in \{0,1\}^n$), we can obtain a multilinear representation Q by simply replacing each x_i^r for $r > 1$ by x_i in the polynomial P. This preserves the underlying function as $b^r = b$ for $b \in \{0,1\}$. We call this process *multilinearization.*

Bernstein's inequality. The following standard deviation bound can be found in, e.g., the book of Dubhashi and Panconesi [DP09, Theorem 1.2].

Lemma 4 (Bernstein's inequality)**.** *Let $X_1, \ldots, X_m$ be independent and identically distributed Bernoulli random variables with mean q. Let $X = \sum_{i=1}^{m} X_i$. Then for any $\theta > 0$,*

$$\Pr\left[|X - mq| > \theta\right] \leq 2\exp\left(-\frac{\theta^2}{2mq(1-q) + 2\theta/3}\right).$$

2.1 Symmetric Boolean functions

Let n be a growing integer parameter which will always be the number of input variables. We use $s\mathcal{B}_n$ to denote the set of all symmetric Boolean functions on n variables. Note that each symmetric Boolean function $f : \{0,1\}^n \to \{0,1\}$ is uniquely specified by a string $\mathrm{Spec}\, f : [0, n] \to \{0, 1\}$, which we call the *Spectrum* of f, in the sense that for any $a \in \{0,1\}^n$, we have

$$f(a) = \mathrm{Spec}\, f(|a|).$$

Given a $f \in s\mathcal{B}_n$, we define the *period of* f, denoted $\mathrm{per}(f)$, to be the smallest positive integer b such that $\mathrm{Spec}\, f(i) = \mathrm{Spec}\, f(i + b)$ for all $i \in [0, n - b]$. We say f is *k-bounded* if $\mathrm{Spec}\, f$ is constant on the interval $[k, n - k]$; let $B(f)$ denote the smallest k such that f is k-bounded.

Standard decomposition of a symmetric Boolean function [Lu01]. Fix any $f \in s\mathcal{B}_n$. Among all symmetric Boolean functions $f' \in s\mathcal{B}_n$ such that $\mathrm{Spec}\, f'(i) = \mathrm{Spec}\, f(i)$ for all $i \in [\lceil n/3 \rceil + 1, \lfloor 2n/3 \rfloor]$, we choose a function g such that $\mathrm{per}(g)$ is as small as possible. We call g the *periodic part* of f. Define $h \in s\mathcal{B}_n$ by $h = f \oplus g$. We call h the *bounded part* of f.

We will refer to the pair (g, h) as a *standard decomposition* of the function f. Note that we have $f = g \oplus h$.

Observation 5. *Let $f \in s\mathcal{B}_n$ and let (g, h) be a standard decomposition of f. Then,* $\mathrm{per}(g) \leq \lfloor n/3 \rfloor$ *and* $B(h) \leq \lceil n/3 \rceil$.

Some symmetric Boolean functions. Fix some positive $n \in \mathbb{N}$. The *Majority* function Maj_n on n Boolean variables accepts exactly the inputs of Hamming weight greater than $n/2$. For $t \in [0, n]$, the *Threshold* function Thr_n^t accepts exactly the inputs of Hamming weight at least t; and similarly, the *Exact Threshold* function EThr_n^t accepts exactly the inputs of Hamming weight exactly t. Finally, for $b \in [2, n]$ and $i \in [0, b - 1]$, the function $\mathrm{MOD}_n^{b,i}$ accepts exactly those inputs a such that $|a| \equiv i \pmod b$. In the special case that $i = 0$, we also use MOD_n^b.

2.2 Probabilistic polynomials

Definition 6 (Probabilistic polynomial and Probabilistic degree)**.** *A* probabilistic polynomial *is a random multilinear polynomial $\boldsymbol{P}$ (with some distribution having finite support) over $\mathbb{F}[x_1, \ldots, x_n]$. We say that the degree of $\boldsymbol{P}$, denoted $\deg(\boldsymbol{P})$, is at most d if the probability distribution defining $\boldsymbol{P}$ is supported on polynomials of degree at most d*

Given a Boolean function $f : \{0,1\}^n \to \{0,1\}$ and an $\varepsilon > 0$, an ε-error probabilistic polynomial *for f is a probabilistic polynomial $\boldsymbol{P}$ such that for each $a \in \{0,1\}^n$,*

$$\Pr_{\boldsymbol{P}}\left[\boldsymbol{P}(a) \neq f(a)\right] \leq \varepsilon.$$

We define the ε-error probabilistic degree *of* f*, denoted* $\mathrm{pdeg}^{\mathbb{F}}_{\varepsilon}(f)$*, to be the least* d *such that* f *has an* ε*-error probabilistic polynomial of degree at most* d*.*

When the field $\mathbb{F}$ *is clear from context, we use* $\mathrm{pdeg}_{\varepsilon}(f)$ *instead of* $\mathrm{pdeg}^{\mathbb{F}}_{\varepsilon}(f)$*.*

Fact 7. *We have the following simple facts about probabilistic degrees of Boolean functions. Let* $\mathbb{F}$ *be any field.*

1. *(Error reduction [HS16]) For any* $\delta < \varepsilon \leq 1/3$ *and any Boolean function* f*, if* $\boldsymbol{P}$ *is an* ε*-error probabilistic polynomial for* f*, then* $\boldsymbol{Q} = M(\boldsymbol{P}_1, \ldots, \boldsymbol{P}_\ell)$ *is a* δ*-error probabilistic polynomial for* f *where* $\ell = O(\log(1/\delta)/\log(1/\varepsilon))$*,* M *is the exact multilinear polynomial for* Maj_ℓ*, and* $\boldsymbol{P}_1, \ldots, \boldsymbol{P}_\ell$ *are independent copies of* $\boldsymbol{P}$*. In particular, we have* $\mathrm{pdeg}^{\mathbb{F}}_{\delta}(f) \leq \mathrm{pdeg}^{\mathbb{F}}_{\varepsilon}(f) \cdot O(\log(1/\delta)/\log(1/\varepsilon))$*. (Note that the reason this is not obvious is that the polynomial* $\boldsymbol{P}$ *is not necessarily Boolean-valued at points when* $\boldsymbol{P}(a) \neq f(a)$*. Hence, it is not clear that composing with a polynomial that computes the Boolean Majority function achieves error-reduction.)*

2. *(Composition) For any Boolean function* f *on* k *variables and any Boolean functions* $g_1, \ldots, g_k$ *on a common set of* m *variables, let* h *denote the natural composed function* $f(g_1, \ldots, g_k)$ *on* m *variables. Then, for any* $\varepsilon, \delta > 0$*, we have* $\mathrm{pdeg}^{\mathbb{F}}_{\varepsilon+k\delta}(h) \leq \mathrm{pdeg}^{\mathbb{F}}_{\varepsilon}(f) \cdot \max_{i\in[k]} \mathrm{pdeg}^{\mathbb{F}}_{\delta}(g_i)$*.*

3. *(Sum) Assume that* $f, g_1, \ldots, g_k$ *are all Boolean functions on a common set of* m *variables such that* $f = \sum_{i\in[k]} g_i$*. Then, for any* $\delta > 0$*, we have* $\mathrm{pdeg}^{\mathbb{F}}_{k\delta}(f) \leq \max_{i\in[k]} \mathrm{pdeg}^{\mathbb{F}}_{\delta}(g_i)$*.*

Building on work of Alman and Williams [AW15] and Lu [Lu01], Tripathi, Venkitesh and the author [STV19] gave upper bounds on the probabilistic degree of any symmetric function. We recall below the statement in the case of fixed positive characteristic.

Theorem 8 (Known upper bounds on probabilistic degree of symmetric functions [STV19])**.** *Let* $\mathbb{F}$ *be a field of constant characteristic* $p > 0$ *and* $n \in \mathbb{N}$ *be a growing parameter. Let* $f \in s\mathcal{B}_n$ *be arbitrary and let* (g, h) *be a standard decomposition of* f*. Then we have the following for any* $\varepsilon > 0$*.*

1. *If* $\mathrm{per}(g) = 1$*, then* $\mathrm{pdeg}_{\varepsilon}(g) = 0$*.*

   If $\mathrm{per}(g)$ *is a power of* p*, then* $\mathrm{pdeg}^{\mathbb{F}}_{\varepsilon}(g) \leq \mathrm{per}(g)$*,*

2. $\mathrm{pdeg}_{\varepsilon}(h) = O(\sqrt{B(h)\log(1/\varepsilon)} + \log(1/\varepsilon))$ *if* $B(h) \geq 1$ *and* 0 *otherwise, and*

3. $$\mathrm{pdeg}_{\varepsilon}(f) = \begin{cases} O(\sqrt{n\log(1/\varepsilon)}) & \text{if } \mathrm{per}(g) > 1 \text{ and not a power of } p, \\ O(\min\{\sqrt{n\log(1/\varepsilon)}, \mathrm{per}(g)\}) & \text{if } \mathrm{per}(g) \text{ a power of } p \text{ and } B(h) = 0, \\ O(\min\{\sqrt{n\log(1/\varepsilon)}, \mathrm{per}(g) + \\ \quad \sqrt{B(h)\log(1/\varepsilon)} + \log(1/\varepsilon)\}) & \text{otherwise.} \end{cases}$$

2.3 A string lemma

Given a function $w : I \to \{0,1\}$ where $I \subseteq \mathbb{N}$ is an interval, we think of w as a string from the set $\{0,1\}^{|I|}$ in the natural way. For an interval $J \subseteq I$, we denote by $w|_J$ the substring of w obtained by restriction to J.

The following simple lemma can be found, e.g. as a special case of [BK03, Theorem 3.1]. For completeness, we give a short proof in Appendix B.

Lemma 9. *Let* $w \in \{0,1\}^+$ *be any non-empty string and* $u, v \in \{0,1\}^+$ *such that* $w = uv = vu$*. Then there exists a string* $z \in \{0,1\}^+$ *such that* w *is a power of* z *(i.e.* $w = z^k$ *for some* $k \geq 2$*).*

Corollary 10. *Let $g \in s\mathcal{B}_n$ be arbitrary with $\mathrm{per}(g) = b > 1$. Then for all $i, j \in [0, n-b+1]$ such that $i \not\equiv j \pmod b$, we have $\mathrm{Spec}\, g|_{[i,i+b-1]} \neq \mathrm{Spec}\, g|_{[j,j+b-1]}$.*

Proof. Suppose $\mathrm{Spec}\, g|_{[i,i+b-1]} = \mathrm{Spec}\, g|_{[j,j+b-1]}$ for some $i \not\equiv j \pmod b$. Assume without loss of generality that $i < j < i + b$. Let $u = \mathrm{Spec}\, g|_{[i,j-1]}, v = \mathrm{Spec}\, g|_{[j,i+b-1]}, w = \mathrm{Spec}\, g|_{[i+b,j+b-1]}$. Then $u = w$ and the assumption $uv = vw$ implies $uv = vu$. By Lemma 9, there exists a string z such that $uv = z^k$ for $k \geq 2$ and therefore $\mathrm{per}(g) < b$. This contradicts our assumption on b. □

2.4 Lucas's theorem

Theorem 11 (Lucas's theorem)**.** *Let A, B be any non-negative integers and p any prime. Then*

$$\binom{A}{B} = \prod_{i \geq 0} \binom{A_i}{B_i} \pmod p$$

where A_i (resp. B_i) is the $(i+1)$th least significant digit of A (resp. B) in base p.

The following is a standard application of Lucas's theorem, essentially observed by Lu [Lu01] and Hegedűs [Heg09], showing that Hegedűs's lemma is tight.

Corollary 12. *Fix any prime p and positive integer n. Assume i is a non-negative integer and q a power of p such that $i + q \leq n$. Then, there is a symmetric multilinear polynomial $Q \in \mathbb{F}_p[x_1, \ldots, x_n]$ of degree q such that Q vanishes at all points of $\{0,1\}^n_i$ but at no point of $\{0,1\}^n_{i+q}$.*

Proof. Assume $q = p^\ell$. Let $a_\ell, b_\ell \in \{0, \ldots, p-1\}$ be the $(\ell+1)$th least significant digit of i and $i+q$ respectively in base p. Note that $b_\ell = a_\ell + 1 \pmod p$.

Define the polynomial

$$Q(x_1, \ldots, x_n) = \left(\sum_{S \subseteq [n] : |S| = q} \prod_{i \in S} x_i \right) - a_\ell,$$

which we consider an element of $\mathbb{F}_p[x_1, \ldots, x_n]$. Note that at any input $c \in \{0,1\}^n$ of Hamming weight w, we have

$$Q(c) = \binom{|c|}{q} - a_\ell$$

where the right hand side is interpreted modulo p. Lucas's theorem then easily implies that $Q(c) = 0$ if $w = i$ and 1 if $w = i + q$. □

3 The Main Lemma

In this section, we prove the main lemma, which is a robust version of Lemma 1.

Lemma 13 (A Robust Version of Hegedűs's Lemma)**.** *Assume that $\mathbb{F}$ is a field of characteristic p. Let n be a growing parameter and assume we have positive integer parameters k, q such that $100q < k < n - 100q$ and q is a power of p. Define $\alpha = \min\{k/n, 1 - (k/n)\}$ and $\delta = q/n$. Assume $P \in \mathbb{F}[x_1, \ldots, x_n]$ is a multilinear polynomial such that for some $K \in \{k+q, k-q\}$,*

$$\Pr_{\boldsymbol{a} \sim \{0,1\}^n_k} [P(\boldsymbol{a}) \neq 0] \leq \min\{e^{-100\delta^2 n/\alpha}, 1/1000\} \tag{1a}$$

$$\Pr_{\boldsymbol{a} \sim \{0,1\}^n_K} [P(\boldsymbol{a}) \neq 0] \geq e^{-\delta^2 n/100\alpha}. \tag{1b}$$

Then, $\deg(P) = \Omega(q)$, where the $\Omega(\cdot)$ hides an absolute constant.

One can ask if the above lemma can be proved under weaker assumptions: specifically, if the upper bound in (1a) can be relaxed. It turns out that it cannot (up to changing the constant in the exponent) because for larger error parameters, there is a sampling-based construction of a polynomial with smaller degree that is zero on most of $\{0,1\}^n_k$ and non-zero on most of $\{0,1\}^n_K$. We discuss this construction in Section 3.3.

We first prove a special case of the lemma which corresponds to the case when $K = k + q = \lfloor n/2 \rfloor$ and q sufficiently larger than $\sqrt{n}$. This case suffices for most of our applications. The general case is a straightforward reduction to this special case.

3.1 A special case

Lemma 14 (A special case of Lemma 13)**.** *Let n be a growing parameter and assume $\varepsilon \in [2^{-n/100}, e^{-200}]$. Assume t is an integer such that t is a power of p and furthermore, $t = \sqrt{n\ell}$ for some $\ell \in \mathbb{R}$ such that $100 \leq \ell \leq \frac{1}{2} \cdot \ln(1/\varepsilon)$. Let $P \in \mathbb{F}[x_1, \ldots, x_n]$ be any polynomial such that*

$$\Pr_{\boldsymbol{a} \sim \{0,1\}^n_{\lfloor n/2 \rfloor - t}} [P(\boldsymbol{a}) \neq 0] \leq \varepsilon \tag{2a}$$

$$\Pr_{\boldsymbol{a} \sim \{0,1\}^n_{\lfloor n/2 \rfloor}} [P(\boldsymbol{a}) \neq 0] \geq e^{-\ell/2}. \tag{2b}$$

Then, $\deg(P) \geq t/25$.

Remark 15. *By negating inputs (i.e. replacing x_i with $1 - x_i$ for each i), the above lemma also implies the analogous statements where $\lfloor n/2 \rfloor - t$ and $\lfloor n/2 \rfloor$ are replaced by $\lceil n/2 \rceil + t$ and $\lceil n/2 \rceil$ respectively.*

Before we prove this lemma, we need to collect some technical facts and lemmas.

The following is standard. See, e.g., [KS18, Lemma 3.3] for a proof.

Fact 16. *Let $R \in \mathbb{F}[x_1, \ldots, x_n]$ be a non-zero multilinear polynomial of degree at most $d \leq n$. Then R cannot vanish at all points in any Hamming ball of radius d in $\{0,1\}^n$.*

Lemma 17. *Let n, r, s be any non-negative integers with $r \leq s \leq n/4$. Then we have*

$$e^{-8s(r-s)/n} \leq \frac{\binom{n}{\lfloor n/2 \rfloor - s}}{\binom{n}{\lfloor n/2 \rfloor - r}} \leq e^{-2r(r-s)/n}.$$

Proof. Note that

$$\frac{\binom{n}{\lfloor n/2 \rfloor - s}}{\binom{n}{\lfloor n/2 \rfloor - r}} = \frac{(\lfloor n/2 \rfloor - s + 1) \cdots (\lfloor n/2 \rfloor - r)}{(\lceil n/2 \rceil + s) \cdots (\lceil n/2 \rceil + r + 1)} \leq \left(\frac{\lfloor n/2 \rfloor - r}{\lceil n/2 \rceil + r} \right)^{r-s} \leq \left(1 - \frac{2r}{n} \right)^{r-s} \leq e^{-2r(r-s)/n},$$

which implies the right inequality in the statement of the claim. We have used the inequality $1 - x \leq e^x$ to deduce the final inequality above.

For the left inequality, we similarly have

$$\frac{\binom{n}{\lfloor n/2 \rfloor - s}}{\binom{n}{\lfloor n/2 \rfloor - r}} \geq \left(\frac{\lceil n/2 \rceil - s}{\lceil n/2 \rceil + s} \right)^{r-s} \geq \left(\left(1 - \frac{2s}{n} \right)^2 \right)^{r-s} \geq e^{-8s(r-s)/n}.$$

where the final inequality follows from the fact that $(1 - x) \geq e^{-2x}$ for $x \in [0, 1/2]$. □

Given a set $E \subseteq \{0,1\}^n$, and a parameter $D \leq n$, we define $\mathcal{I}_D(E)$ to be the set of all multilinear polynomials Q of degree at most D that vanish at all points of E. Further, we define the *degree-D closure of E*, denoted $\mathrm{cl}_D(E)$ as follows.

$$\mathrm{cl}_D(E) := \{a \in \{0,1\}^n \mid Q(a) = 0 \ \ \forall Q \in \mathcal{I}_D(E)\}.$$

Note that $\mathrm{cl}_D(E) \supseteq E$ but could be much bigger than E. The following result of Nie and Wang [NW15] gives a bound on $|\mathrm{cl}_D(E)|$ in terms of $|E|$. (This particular form is noted and essentially proved in [NW15], and is explicitly stated and proved in [KS18, Theorem A.1].)

Theorem 18. *For any $E \subseteq \{0,1\}^n$ and any $D \leq n$, we have*

$$\frac{\mathrm{cl}_D(E)}{2^n} \leq \frac{|E|}{N_D}$$

where $N_D = \sum_{j=0}^{D} \binom{n}{j}$, the number of multilinear monomials of degree at most D.

We now begin the proof of the Lemma 14.

Proof of Lemma 14. Assume that P is as given. Let $m = \lfloor n/2 \rfloor$.

Let E_0, E_1 be defined as follows.

$$E_0 = \{a \in \{0,1\}^n_{m-t} \mid P(a) \neq 0\}$$
$$E_1 = \{a \in \{0,1\}^n_m \mid P(a) = 0\}$$

We show that there are polynomials $Q_1, Q_2 \in \mathbb{F}[x_1, \ldots, x_n]$ such that the following conditions hold.

(Q1.1) $Q_1(a) \neq 0$ if and only if $|a| \equiv m \pmod t$.

(Q2.1) $Q_2(a) = 0$ for all $a \in E_0$.

(Q2.2) $Q_2(a) = 0$ for all a such that $|a| < m - t$ and $|a| \equiv m \pmod t$.

(Q2.3) $Q_2(a) \neq 0$ for some $a \in \{0,1\}^n_m \setminus E_1$.

Given polynomials Q_1, Q_2 as above, we construct the polynomial R to be the multilinear polynomial obtained by computing the formal product $P \cdot Q_1 \cdot Q_2$ and replacing x_i^r by x_i for each $r > 1$. Note that $R(a) = P(a)Q_1(a)Q_2(a)$ for any $a \in \{0,1\}^n$.

We observe that $R(a) = 0$ for all $|a| < m$. This is based on a case analysis of whether $|a| \equiv m \pmod t$ or not. In the latter case, we see that $Q_1(a) = 0$ and hence $R(a) = 0$. In the former case, we have either $a \in \{0,1\}^n_{m-t} \setminus E_0$, in which case $P(a) = 0$, or not, in which case $Q_2(a) = 0$. Hence, $R(a) = 0$ for all $|a| < m$.

On the other hand, we note that R is a non-zero polynomial. This is because by (Q2.3), we know that there is some $a' \in \{0,1\}^n_m \setminus E_1$ where $Q_2(a') \neq 0$. Further, $Q_1(a') \neq 0$ and $P(a') \neq 0$ by (Q1.1) and the definition of E_1 respectively. Hence, $R(a') \neq 0$, implying that R is a non-zero multilinear polynomial.

By Fact 16, we thus know that R has degree at least m. In particular, we obtain

$$\deg(P) \geq \deg(R) - \deg(Q_1) - \deg(Q_2) \geq m - \deg(Q_1) - \deg(Q_2).$$

Hence, to finish the proof of the lemma, it suffices to prove the following claims.

Claim 19. *There is a Q_1 of degree at most t satisfying property (Q1.1).*

Claim 20. *There is a Q_2 of degree at most $m - t - t_1$ satisfying properties (Q2.1)-(Q2.3), where $t_1 = \lceil t/25 \rceil$.*

We now prove the above claims.

Proof of Claim 19. This follows immediately from the upper bound for periodic functions in Theorem 8. Consider the t-periodic function that takes the value 1 at point $a \in \{0,1\}^n$ if and only if $|a| \equiv m \pmod t$. Since this function is t-periodic, it can be represented exactly as a polynomial of degree at most t. This yields the claim. □

Proof of Claim 20. Let D denote $m - t - t_1$. Let $E = E_0 \cup \bigcup_{j<m-t:j\equiv m \pmod t} \{0,1\}^n_j$. We want to show the existence of a polynomial Q_2 of degree at most D such that Q_2 vanishes at all points of E but Q_2 does not vanish at some point in $\bar{E}_1 := \{0,1\}^n_m \setminus E_1$. Note that this is equivalent to saying that $\mathrm{cl}_D(E) \not\supseteq \bar{E}_1$. To show this, it suffices to show that

$$|\mathrm{cl}_D(E)| < e^{-\ell/2} \cdot \binom{n}{m} \tag{3}$$

since by hypothesis we have $|\bar{E}_1| \geq e^{-\ell/2} \cdot \binom{n}{m}$.

To do this, we use Theorem 18. Note that we have

$$\begin{aligned}
|E| &\leq |E_0| + \sum_{j<m-t:j\equiv m \pmod t} \binom{n}{j} \\
&\leq \varepsilon \cdot \binom{n}{m-t} + \sum_{k\geq 1} \binom{n}{m-t-k\cdot t} \\
&\leq \varepsilon \cdot \binom{n}{m-t} + \binom{n}{m-t} \cdot \left(e^{-2\ell} + e^{-4\ell} + \ldots\right) \\
&\leq \binom{n}{m-t} \cdot (\varepsilon + 2 \cdot e^{-2\ell}) \leq \binom{n}{m-t} \cdot (3e^{-2\ell})
\end{aligned} \tag{4}$$

where the third inequality is a consequence of Lemma 17 (with $r = t$ and $s = (k+1)t$ for various k) and the final inequality uses $\varepsilon \leq e^{-2\ell}$.

On the other hand, the parameter N_D from the statement of Theorem 18 can be lower bounded as follows.

$$\begin{aligned}
N_D = \sum_{j=0}^{D} \binom{n}{D-j} &\geq t_1 \binom{n}{m-t-2t_1} \\
&\geq t_1 e^{-\ell} \cdot \binom{n}{m-t} > e^{-\ell} \cdot \frac{\sqrt{n}}{3} \cdot \binom{n}{m-t}
\end{aligned}$$

where the second inequality follows from Lemma 17 (with $r = t$ and $s = t + 2t_1$) and the final inequality uses the fact that $t_1 > t/30 = \sqrt{n\ell}/30 \geq \sqrt{n}/3$.

Putting the above together with (4) immediately yields

$$\frac{|E|}{N_D} < 9e^{-\ell} \cdot \frac{\binom{n}{m-t}}{\sqrt{n} \cdot \binom{n}{m-t}} = 9e^{-\ell} \cdot n^{-1/2}.$$

Using Theorem 18, we thus obtain

$$\mathrm{cl}_D(E) < 9e^{-\ell} \cdot \frac{2^n}{\sqrt{n}} \leq e^{-\ell/2} \cdot \frac{2^n}{2\sqrt{n}} \leq e^{-\ell/2} \cdot \binom{n}{m}$$

where the last inequality follows from Stirling's approximation. Having shown (3), the claim now follows. □

□

3.2 The General Case

We start with some preliminaries.

We first show a simple 'error-reduction' procedure for polynomials. For any multilinear polynomial $P \in \mathbb{F}[x_1, \ldots, x_n]$ and any $m \in [0, n]$, let $\mathrm{NZ}_m(P)$ denote the set of points of $\{0,1\}^n_m$ where P does not vanish. Let $\psi_m(P)$ denote $|\mathrm{NZ}_m(P)|/\binom{n}{m}$.

Lemma 21. *For any multilinear $Q \in \mathbb{F}[x_1, \ldots, x_n]$ and any $r \geq 1$, there is a probabilistic polynomial $\boldsymbol{Q}^{(r)}$ of degree at most $r \cdot \deg(Q)$ such that for all $m \in [0, n]$, $\boldsymbol{E}_{\boldsymbol{Q}^{(r)}}[\psi_m(\boldsymbol{Q}^{(r)})] = \psi_m(Q)^r$.*

Proof. For a permutation $\pi \in S_n$, and $a \in \{0,1\}^n$, define $a^\pi = (a_{\pi(1)}, \ldots, a_{\pi(n)})$. Also, define $Q^\pi(x_1, \ldots, x_n) = Q(x^\pi) = Q(x_{\pi(1)}, \ldots, x_{\pi(n)})$.

For a uniformly random $\boldsymbol{\pi} \in S_n$, and any $a \in \{0,1\}^n_m$, the probabilistic polynomial $Q^{\boldsymbol{\pi}}$ satisfies

$$\Pr_{\boldsymbol{\pi}}[Q^{\boldsymbol{\pi}}(a) \neq 0] = \Pr_{\boldsymbol{\pi}}[Q(a^{\boldsymbol{\pi}}) \neq 0] = \Pr_{\boldsymbol{\pi}}[a^{\boldsymbol{\pi}} \in \mathrm{NZ}_m(Q)] = \psi_m(Q)$$

as $a^{\boldsymbol{\pi}}$ is uniformly distributed over $\{0,1\}^n_m$.

Choose $\boldsymbol{\pi_1}, \ldots, \boldsymbol{\pi_r}$ i.u.a.r. from S_n, and define $\boldsymbol{Q}^{(r)} = \prod_{i=1}^r Q^{\boldsymbol{\pi_i}}$. For any $a \in \{0,1\}^n_m$

$$\Pr_{\boldsymbol{Q}^{(r)}}\left[\boldsymbol{Q}^{(r)}(a) \neq 0\right] = (\psi_m(Q))^r.$$

In particular, the above holds for a uniformly random $\boldsymbol{a}$ chosen from $\{0,1\}^n_m$. Hence, we have

$$\mathop{\mathbf{E}}_{\boldsymbol{Q}^{(r)}}[\psi_m(\boldsymbol{Q}^{(r)})] = \Pr_{\boldsymbol{Q}^{(r)}, \boldsymbol{a} \sim \{0,1\}^n_m}\left[\boldsymbol{Q}^{(r)}(\boldsymbol{a}) \neq 0\right] = \psi_m(Q)^r.$$

□

We are now ready to prove the main lemma in its full generality.

Proof of Lemma 13. W.l.o.g. we assume that $k \leq n/2$. (To prove the lemma for $k > n/2$, consider the polynomial $Q(x) = P(1 - x_1, \ldots, 1 - x_n)$ instead.)

We first reduce to the case where $K = n/2$.

More precisely, note that there exist non-negative integers $r \leq 2q$ and s so that $2(K - r) = n - r - s$. This can be seen by a simple case analysis. If $K = k - q$, we can choose $r = 0$, $s = n - 2k + 2q$; if $K = k + q$ and $n - 2k \geq 2q$, we can choose $r = 0$ and $s = n - 2k - 2q$; and if $K = k + q$ and $n - 2k < 2q$, we can choose $r = 2q - (n - 2k)$ and $s = 0$.

Having chosen r, s as above, we set $K' = K - r$, $k' = k - r$ and $n' = n - r - s$. Let $\boldsymbol{S}$ be a uniformly random subset of $[n]$ of size $r + s$ and $\boldsymbol{y}$ a uniformly random point in $\{0,1\}^{r+s}_r$. We set $P_{\boldsymbol{S},\boldsymbol{y}}(x_i : i \notin S)$ to be the probabilistic polynomial obtained by setting all the variables indexed by $\boldsymbol{S}$ according to $\boldsymbol{y}$. Note that we have

$$\mathop{\mathbf{E}}_{\boldsymbol{S},\boldsymbol{y}}[\psi_{k'}(P_{\boldsymbol{S},\boldsymbol{y}})] = \psi_k(P) =: \varepsilon_0 \qquad \text{and} \qquad \mathop{\mathbf{E}}_{\boldsymbol{S},\boldsymbol{y}}[\psi_{K'}(P_{\boldsymbol{S},\boldsymbol{y}})] = \psi_K(P) =: \varepsilon_1.$$

By Markov's inequality, we have

$$\Pr_{\boldsymbol{S},\boldsymbol{y}}\left[\psi_{k'}(P_{\boldsymbol{S},\boldsymbol{y}}) > \frac{2\varepsilon_0}{\varepsilon_1}\right] < \frac{\varepsilon_1}{2} \qquad \text{and} \qquad \Pr_{\boldsymbol{S},\boldsymbol{y}}\left[\psi_{K'}(P_{\boldsymbol{S},\boldsymbol{y}}) > \frac{\varepsilon_1}{2}\right] \geq \frac{\varepsilon_1}{2}.$$

Hence, with positive probability over the choice of $\boldsymbol{S}$ and $\boldsymbol{y}$, we have both $\psi_{k'}(P_{\boldsymbol{S},\boldsymbol{y}}) \leq 2\varepsilon_0/\varepsilon_1$ and $\psi_{K'}(P_{\boldsymbol{S},\boldsymbol{y}}) > \varepsilon_1/2$. We fix such a choice S, y for $\boldsymbol{S}, \boldsymbol{y}$ and let P' denote $P_{S,y}$. Clearly, $\deg(P) \geq \deg(P')$ and hence it suffices to lower bound $\deg(P')$.

We will now use Lemma 14 to obtain the desired lower bound on $\deg(P')$. First of all, note that $\ell' := q^2/n'$ satisfies

$$\ell' = \frac{q^2}{n'} \leq \frac{k^2}{10000n'} \leq \frac{n'}{10000},$$

by the bounds on q in the statement of the lemma and the fact that $k \leq 2K = n'$.

We consider now two cases.

Case 1: Assume first that $\ell' \geq 100$. Using the bounds on ε_0 and ε_1 in the lemma statement and the bounds above, P' is a polynomial in n' variables satisfying

$$\psi_{k'}(P') \leq \frac{2\varepsilon_0}{\varepsilon_1} \leq 2\varepsilon_0^{0.99} \leq 2\exp(-99\delta^2 n/\alpha) = 2\exp(-99\delta^2 n^2/(\alpha n)) \leq 2\exp(-99q^2/n'), \text{ and}$$
$$\psi_{K'}(P') \geq \frac{\varepsilon_1}{2} \geq \frac{1}{2}\exp(-(1/100)\cdot\delta^2 n/\alpha) = \frac{1}{2}\exp(-(1/100)\cdot\delta^2 n^2/(\alpha n)) \geq \frac{1}{2}\exp(-(1/25)\cdot q^2/n').$$

where we have used the inequalities $n' \geq 2(k-q) \geq \alpha n$ and $n' = 2K' \leq 2(k+q) \leq 4\alpha n$.

Define $\varepsilon = \exp(-2\ell')$. Note that we have $\varepsilon \geq \exp(-n'/5000)$ by the bound on ℓ' above. Further,

$$\psi_{(n'/2)-q}(P') = \psi_{k'}(P') \leq 2\exp(-99q^2/n') = 2\exp(-99\ell') \leq \exp(-2\ell') = \varepsilon, \text{ and}$$
$$\psi_{n'/2}(P') = \psi_{K'}(P') \geq \frac{1}{2}\exp(-(1/25)\cdot q^2/n') \geq \exp(-\ell'/2).$$

Applying Lemma 14 to P' (see also Remark 15), we immediately obtain $\deg(P') \geq q/25$ and hence we are done in this case.

Case 2: Now consider the case when $\ell' < 100$. In this case, the hypothesis of the lemma assures us that $\varepsilon_0 \leq 1/1000$ and $\varepsilon_1 \geq \exp(-q^2/100\alpha n) \geq \exp(-\ell'/25) \geq e^{-4}$ where the second inequality uses $n' \leq 4\alpha n$ as argued above. Then, we have

$$\psi_{k'}(P') \leq \frac{2\varepsilon_0}{\varepsilon_1} \leq 2\varepsilon_0^{0.99} \leq \frac{1}{400}, \tag{5a}$$
$$\psi_{K'}(P') \geq \frac{\varepsilon_1}{2} \geq \frac{\varepsilon_0^{0.01}}{2} \geq \frac{1}{2^{100/99}}\cdot\psi_{k'}(P')^{1/99} \geq \psi_{k'}(P')^{1/7}, \tag{5b}$$
$$\psi_{K'}(P') \geq \frac{\varepsilon_1}{2} \geq e^{-5}. \tag{5c}$$

where (5b) uses $\varepsilon_0^{0.01} \geq (\psi_{k'}(P')/2)^{1/99}$ and $\psi_{k'}(P') \leq 1/400$, both of which follow from (5a).

Let r be a large constant that will be fixed below. By Lemma 21, we know that there is a probabilistic polynomial $\boldsymbol{P'^{(r)}}$ of degree at most $r\cdot\deg(P')$ such that for each $m \in \{k', K'\}$, we have $\mathbf{E}_{\boldsymbol{P'^{(r)}}}[\psi_m(\boldsymbol{P'^{(r)}})] = \psi_m(P')^r$.

The proof will proceed by another restriction to n'' variables, where n'' is defined to be the largest even integer such that $100n'' \leq q^2$. We assume that n'' is greater than a large enough absolute constant, since otherwise q is upper bounded by a fixed constant, in which case the degree bound to be proved is trivial. Note that $\ell'' := q^2/n'' \geq 100$ by definition. We also have $n'' = (q^2/100) - 2$, which implies that $\ell'' \leq 100 + O(1)/q^2 \leq 101$, as long as q is greater than a large enough absolute constant.

Relabel the variables so that P' is a polynomial in $x_1, \ldots, x_{n'}$. Let $\boldsymbol{T}$ be a uniformly random subset of $[n']$ of size $n' - n''$ and let $\boldsymbol{z}$ be a uniformly random point in $\{0,1\}^{n'-n''}_{(n'-n'')/2}$. Define the probabilistic

polynomial $\boldsymbol{P}^{\boldsymbol{\prime}(\boldsymbol{r})}_{\boldsymbol{T},\boldsymbol{z}}$ obtained by setting the variables indexed by $\boldsymbol{T}$ according to $\boldsymbol{z}$ in the probabilistic polynomial $\boldsymbol{P}^{\boldsymbol{\prime}(\boldsymbol{r})}$. Let $K'' := n''/2$ and $k'' := k' - (n' - n'')/2$. As above, we have

$$\mathop{\mathbf{E}}_{\boldsymbol{P}^{\boldsymbol{\prime}(\boldsymbol{r})},\boldsymbol{T},\boldsymbol{z}}[\psi_{k''}(\boldsymbol{P}^{\boldsymbol{\prime}(\boldsymbol{r})}_{\boldsymbol{T},\boldsymbol{z}})] = \psi_{k'}(P')^r =: \varepsilon'_0 \qquad \text{and} \qquad \mathop{\mathbf{E}}_{\boldsymbol{P}^{\boldsymbol{\prime}(\boldsymbol{r})},\boldsymbol{T},\boldsymbol{z}}[\psi_{K''}(\boldsymbol{P}^{\boldsymbol{\prime}(\boldsymbol{r})}_{\boldsymbol{T},\boldsymbol{z}})] = \psi_{K'}(P')^r =: \varepsilon'_1.$$

Let r be the smallest positive integer so that $\varepsilon'_0 = \psi_{k'}(P')^r \le e^{-300}$. Note that r is upper bounded by an absolute constant, as $\psi_{k'}(P') \le 1/400$ by (5a). Further, we have $\psi_{k'}(P')^{r-1} > e^{-300}$ and hence

$$\varepsilon'_1 = \psi_{K'}(P')^r = \psi_{K'}(P')^{r-1} \cdot \psi_{K'}(P') \ge \left((\psi_{k'}(P'))^{r-1}\right)^{1/7} \cdot e^{-5} > e^{-48}$$

where the first inequality uses (5).

By Markov's inequality as above, there is a fixed choice of $\boldsymbol{P}^{\boldsymbol{\prime}(\boldsymbol{r})}, \boldsymbol{T}$, and $\boldsymbol{z}$ such that the corresponding polynomial P'' is a polynomial on n'' variables satisfying

$$\psi_{k''}(P'') \le \frac{2\varepsilon'_0}{\varepsilon'_1} < e^{-210} < e^{-2\ell''} \qquad \text{and} \qquad \psi_{K''}(P'') \ge \frac{\varepsilon'_1}{2} > e^{-50} \ge e^{-\ell''/2}.$$

Applying Lemma 14 to P'' with error parameter $\varepsilon = \frac{2\varepsilon'_0}{\varepsilon'_1}$ yields $\deg(P'') \ge q/25$. As $\deg(P'') \le r \cdot \deg(P')$, we also get $\deg(P') = \Omega(q)$, finishing the proof in this case as well. (Note that the $\Omega(\cdot)$ hides an absolute constant.) □

3.3 Tightness of the Main Lemma (Lemma 13)

In this section, we discuss the near-optimality of Lemma 13 w.r.t. to the various parameters. Fix n, k, q, α, δ and $\mathbb{F}$ as in the statement of Lemma 13. Assume that $K = k + q$ (the case when $K = k - q$ is similar) and that $k \le n/2$. Let $\varepsilon \in (0, 1)$ be arbitrary.

First of all, we note that the degree lower bound obtained cannot be larger than q, because by Corollary 12, it follows that there is a degree-q polynomial that vanishes at all points of weight k but no points of weight K.

In Lemma 13, we try to understand if this continues to be true (up to constant factors in the degree) even if the polynomial is forced to be zero only on most (say a $1 - \varepsilon$ fraction) of $\{0,1\}^n_k$ and non-zero on most (say a $1 - \varepsilon$ fraction) of $\{0,1\}^n_K$. (Lemma 13 is a stronger statement, but the parameters are even tight for this weaker version.)

To understand why ε is set as in (1), we analyze a different polynomial construction to achieve this based on sampling. We will need the following interpolation lemma that can be found in a paper of Alman and Williams [AW15].

Lemma 22. *Let n be arbitrary and $I \subseteq [0, n]$ be any interval of integers. Given any $f : I \to \{0, 1\}$, there is a multilinear polynomial $Q \in \mathbb{Z}[x_1, \ldots, x_n]$ of degree at most $|I| - 1$ such that $Q(a) = f(|a|)$ for each $a \in \bigcup_{i \in I} \{0,1\}^n_i$.*

Fix any positive integer m. By Lemma 22, it follows that there is a multilinear polynomial $Q \in \mathbb{Z}[y_1, \ldots, y_m]$ of degree $O(\delta m)$ such that $Q(b) = 0$ for each $b \in \{0,1\}^m$ such that $|b| \in ((\alpha - \delta/2)m, (\alpha + \delta/2)m)$ and $Q(b) = 1$ for each $b \in \{0,1\}^m$ such that $|b| \in ((\alpha + \delta/2)m, (\alpha + 3\delta/2)m)$. Reducing the coefficients modulo p, we obtain a polynomial $\tilde{Q} \in \mathbb{F}[y_1, \ldots, y_m]$ with the same property. Fix this $\tilde{Q}$.

Consider the probabilistic polynomial $\boldsymbol{P}(x_1, \ldots, x_n)$ defined as follows. Choose $\boldsymbol{i_1}, \ldots, \boldsymbol{i_m}$ i.u.a.r. from $[n]$ where $m = C \cdot (\alpha/\delta^2) \log(1/\varepsilon)$ for a large enough constant C we will fix below. We define $\boldsymbol{P}(x_1, \ldots, x_n)$ to be the multilinear polynomial obtained from $\tilde{Q}(x_{\boldsymbol{i_1}}, \ldots, x_{\boldsymbol{i_m}})$ by multilinearization (i.e. replacing each x_i^r by x_i for $r > 1$). Note that $\deg(\boldsymbol{P}) \le \deg(Q) = O(\delta m) = O((\alpha/\delta) \log(1/\varepsilon))$.

Let $a \in \{0,1\}_k^n$ be arbitrary. We analyze the random variable $\boldsymbol{P}(a)$. Note that as long as the Hamming weight of $\boldsymbol{b} = (a_{\boldsymbol{i_1}}, \ldots, a_{\boldsymbol{i_m}})$ is in the interval $((\alpha - \delta/2)m, (\alpha + \delta/2)m)$, we have $\boldsymbol{P}(a) = 0$. As each co-ordinate of $\boldsymbol{b}$ is 1 with probability $k/n = \alpha \in [0, 1/2]$, Bernstein's inequality (Lemma 4) yields

$$\Pr_{\boldsymbol{P}}[\boldsymbol{P}(a) \neq 0] \leq \Pr_{\boldsymbol{i_1},\ldots,\boldsymbol{i_m}}[||\boldsymbol{b}| - \alpha m| > \delta m/3] \leq \exp(-\Omega(\delta^2 m/\alpha)) < \varepsilon/2$$

as long as C is a large enough constant. In a similar way, we also see that for any $a \in \{0,1\}_K^n$, we have $\Pr_{\boldsymbol{P}}[\boldsymbol{P}(a) \neq 1] < \varepsilon/2$ and hence, in particular, $\Pr_{\boldsymbol{P}}[\boldsymbol{P}(a) \neq 0] > 1 - (\varepsilon/2)$, as long as C is a large enough constant.

In particular, by Markov's inequality and the union bound, we see that there is a P of degree at most $\deg(\boldsymbol{P})$ such that

$$\psi_k(P) \leq \varepsilon \qquad \text{and} \qquad \psi_K(P) \geq 1 - \varepsilon.$$

Thus, we have a polynomial P that satisfies conditions similar to the hypothesis of Lemma 13. Note, however, that the degree of P is $O((\alpha/\delta)\log(1/\varepsilon))$ which can be much smaller than the trivial upper bound of q unless $\varepsilon < \exp(-\Omega(\delta^2 n/\alpha))$. This motivates why the error parameters are set as they are in (1) (specifically (1a)).

4 Applications

4.1 Tight Degree Lower Bounds for the Coin Problem

We start with a definition.

Definition 23 (The δ-Coin Problem)**.** *For any $\alpha \in [0,1]$ and integer $n \geq 1$, let μ_α^n be the product distribution over $\{0,1\}^n$ obtained by setting each bit to 1 independently with probability α. Let $\delta \in (0,1)$ be a parameter.*

Given a function $g : \{0,1\}^n \to \{0,1\}$, we say that g solves the δ-coin problem with error ε *if*

$$\Pr_{\boldsymbol{x} \sim \mu_{(1/2)-\delta}^n}[g(\boldsymbol{x}) = 1] \leq \varepsilon \text{ and } \Pr_{\boldsymbol{x} \sim \mu_{1/2}^n}[g(\boldsymbol{x}) = 1] \geq 1 - \varepsilon. \tag{6}$$

(This definition is sometimes [LSS+18] stated in terms of the distributions $\mu_{(1/2)-\delta}$ and $\mu_{(1/2)+\delta}$. This is essentially equivalent to the definition above.)

Let $\mathbb{F}$ be a prime field of characteristic p, where p is a fixed constant. We consider here the minimum degree of a polynomial $P \in \mathbb{F}[x_1, \ldots, x_n]$ that solves the δ-coin problem with error ε.

By Lemma 22, for any $n \geq 1$, there is a polynomial $P \in \mathbb{F}[x_1, \ldots, x_n]$ of degree $O(\delta n)$ that outputs 0 on all inputs of weight $w \in (n((1/2) - 3\delta/2), n(1/2 - \delta/2))$ and 1 on all inputs of weight $(n(1/2 - \delta/2), n(1/2 + \delta/2))$. Using Lemma 4 (Bernstein's inequality), it can be easily checked that P solves the δ-coin problem with error ε as long as $n \geq C\frac{1}{\delta^2}\log(1/\varepsilon)$ for some large enough constant $C > 0$. This yields a polynomial P of degree $O(\frac{1}{\delta}\log(1/\varepsilon))$.

In earlier work [LSS+18], we showed that this was tight for constant ε. That is, we showed that any polynomial P that solves the δ-coin problem with error at most $1/10$ (say) must have degree $\Omega(1/\delta)$. This was also implied by an independent result of Chattopadhyay, Hosseini, Lovett and Tal [CHLT19] (see [Agr19]). Both proofs relied on slight strengthenings of Smolensky's [Smo87b] lower bound on polynomials approximating the Majority function. It is not clear from these proofs, however, if this continues to be true for subconstant ε. The main lemma (Lemma 13), or even its simpler version Lemma 14, shows that this is indeed true.

Theorem 24 (Tight Degree Lower Bound for the δ-coin problem for all errors)**.** *Assume $\mathbb{F}$ is as above and δ, ε are parameters going to 0. Let $N \geq 1$ be any positive integer. Any polynomial $P \in \mathbb{F}[x_1, \ldots, x_N]$ that solves the δ-coin problem with error ε must have degree $\Omega(\frac{1}{\delta}\log(1/\varepsilon))$.*

Proof. We assume that ε is smaller than some small enough constant ε_0 (for larger ε, we can just appeal to the lower bound of [LSS+18]).

Assume for now that $\delta = 1/k$ for some integer $k \geq 1$. Fix n to be the least even integer such that $n \geq \frac{C}{\delta^2}\log(1/\varepsilon)$ for a large constant C and $q := \delta n$ is a power of the characteristic p. Note that $n \leq O(p) \cdot \frac{C}{\delta^2}\log(1/\varepsilon) = O(\frac{1}{\delta^2}\log(1/\varepsilon))$ as p is a constant. Define the probabilistic polynomial $\boldsymbol{Q} \in \mathbb{F}[y_1, \ldots, y_n]$ obtained from P by randomly replacing each variable of P by a uniformly random variable among $y_1, \ldots, y_n$ (followed by multilinearization). For any $a \in \{0,1\}^n_{n/2}$, we have

$$\Pr_{\boldsymbol{Q}}[\boldsymbol{Q}(a) = 0] = \Pr_{\boldsymbol{b} \sim \mu_{1/2}}[P(\boldsymbol{b}) = 0] \leq \varepsilon,$$

and similarly for $a \in \{0,1\}^n_{(n/2)-q}$, we have $\Pr_{\boldsymbol{Q}}[\boldsymbol{Q}(a) \neq 0] \leq \varepsilon$. In particular, by Markov's inequality, there is a fixed polynomial Q of degree at most $\deg(P)$ that satisfies

$$\Pr_{\boldsymbol{a} \sim \{0,1\}^n_{n/2}}[Q(a) = 0] \leq 2\varepsilon \text{ and } \Pr_{\boldsymbol{a} \sim \{0,1\}^n_{(n/2)-q}}[Q(a) \neq 0] \leq 2\varepsilon.$$

Hence, by Lemma 14, we have $\deg(P) = \Omega(\delta n) = \Omega(\frac{1}{\delta}\log(1/\varepsilon))$.

Now, if δ is not of the assumed form, we consider k be the largest integer such that $\delta \leq 1/k$ and set $\delta' := 1/k$. Define $\alpha \in (0,1)$ by $\alpha = \delta/\delta'$. Note that if $\boldsymbol{a}, \boldsymbol{b} \in \{0,1\}$ are sampled independently from the distributions $\mu^1_{1/2-\delta'}$ and $\mu^1_{1/2-(\alpha/2)}$ respectively, then their parity $\boldsymbol{a} \oplus \boldsymbol{b}$ has the distribution $\mu^1_{1/2-\delta}$. Now, if we define the probabilistic polynomial $\boldsymbol{R}(x_1, \ldots, x_n)$ by

$$\boldsymbol{R}(x_1, \ldots, x_n) = P(x_1 \oplus \boldsymbol{y}_1, \ldots, x_n \oplus \boldsymbol{y}_n)$$

where $\boldsymbol{y} = (\boldsymbol{y}_1, \ldots, \boldsymbol{y}_n)$ is sampled from $\mu^n_{1/2-(\alpha/2)}$, then $\boldsymbol{R}$ solves the δ'-coin problem with error at most ε. Note also that $\deg(\boldsymbol{R}) \leq \deg(P)$ as for each fixed $\boldsymbol{y}$, each $x_i \oplus \boldsymbol{y}_i$ is a linear function of x_i.

Repeating the above argument with $\boldsymbol{R}$ instead of P yields that $\deg(\boldsymbol{R}) = \Omega(\frac{1}{\delta'}\log(1/\varepsilon)) = \Omega(\frac{1}{\delta}\log(1/\varepsilon))$. We thus get the same lower bound for $\deg(P)$. □

4.2 Tight Probabilistic Degree Lower bounds for Positive Characteristic

We start with some basic notation and definitions and then state our result.

Throughout this section, let $\mathbb{F}$ be a field of fixed (i.e. independent of n) characteristic $p > 0$. The main theorem of this section characterizes (up to constant factors) the ε-error probabilistic degree of every symmetric function and for almost all interesting values of ε.

Theorem 25 (Probabilistic Degree lower bounds over positive characteristic)**.** *Let $n \in \mathbb{N}$ be a growing parameter. Let $f \in s\mathcal{B}_n$ be arbitrary and let (g, h) be a standard decomposition of f (see Section 2 for the definition). Then for any $\varepsilon \in [1/2^n, 1/3]$, we have*

$$\mathrm{pdeg}^{\mathbb{F}}_{\varepsilon}(f) = \begin{cases} \Omega(\sqrt{n\log(1/\varepsilon)}) & \text{if } \mathrm{per}(g) > 1 \text{ and not a power of } p, \\ \Omega(\min\{\sqrt{n\log(1/\varepsilon)}, \mathrm{per}(g)\}) & \text{if } \mathrm{per}(g) \text{ a power of } p \text{ and } B(h) = 0, \\ \Omega(\min\{\sqrt{n\log(1/\varepsilon)}, \mathrm{per}(g) & \\ \quad + \sqrt{B(h)\log(1/\varepsilon)} + \log(1/\varepsilon)\}) & \text{otherwise.} \end{cases}$$

Here the $\Omega(\cdot)$ notation hides constants depending on the characteristic p of the field $\mathbb{F}$.

Note that this matches the upper bound construction from Theorem 8.

4.2.1 Some Preliminaries

Definition 26 (Restrictions)**.** *Given functions $f \in s\mathcal{B}_n$ and $g \in s\mathcal{B}_m$ where $m \leq n$, we say that g is a restriction of f if there is some $a \in [0, n-m]$ such that the identity*

$$g(x) = f(x1^a0^{n-m-a})$$

holds for every $x \in \{0,1\}^n$. Or equivalently, that g can be obtained from f by setting some inputs to 0 and 1 respectively.[8]

We will use the following obvious fact freely.

Observation 27. *If g is a restriction of f, then for any $\delta > 0$, $\mathrm{pdeg}_\delta(g) \leq \mathrm{pdeg}_\delta(f)$.*

In earlier work with Tripathi and Venkitesh [STV19], we showed the following near-optimal lower bound on the probabilistic degrees of Threshold functions.

Lemma 28 (Lemma 27 in [STV19])**.** *Assume $t \geq 1$. For any $\varepsilon \in [2^{-n}, 1/3]$,*

$$\mathrm{pdeg}_\varepsilon(\mathrm{Thr}_n^t) = \Omega(\sqrt{\min\{t, n+1-t\}\log(1/\varepsilon)} + \log(1/\varepsilon)).$$

(The corresponding lemma in [STV19] is only stated for $t \leq n/2$. However, as $\mathrm{Thr}_n^{n+1-t}(x) = 1 - \mathrm{Thr}_n^t(1-x_1, \ldots, 1-x_n)$, the above lower bound holds for $t > n/2$ also.)

The following classical results of Smolensky prove optimal lower bounds on the probabilistic degrees of some interesting classes of symmetric functions.

Lemma 29 (Smolensky's lower bound for close-to-Majority functions [Sze89, Smo93])**.** *For any field $\mathbb{F}$, any $\varepsilon \in (1/2^n, 1/5)$, and any Boolean function g on n variables that agrees with Maj_n on a $1-\varepsilon$ fraction of its inputs, we have*

$$\mathrm{pdeg}_\varepsilon^{\mathbb{F}}(g) = \Omega(\sqrt{n\log(1/\varepsilon)}).$$

Lemma 30 (Smolensky's lower bound for MOD functions [Smo87a])**.** *For $2 \leq b \leq n/2$, any $\mathbb{F}$ such that $\mathrm{char}(\mathbb{F})$ is either zero or coprime to b, any $\varepsilon \in (1/2^n, 1/(3b))$, there exists an $i \in [0, b-1]$ such that*

$$\mathrm{pdeg}_\varepsilon^{\mathbb{F}}(\mathrm{MOD}_n^{b,i}) = \Omega(\sqrt{n\log(1/b\varepsilon)}).$$

We now show how to use our robust version of Hegedűs's lemma to prove Theorem 25. In fact, Lemma 14 will suffice for this application.

4.2.2 Strategy and two simple examples

The probabilistic degree lower bounds below will use the following corollary of Lemma 14.

Corollary 31. *Let n be a growing parameter and assume $\varepsilon \in [2^{-n/100}, e^{-200}]$. Assume t is an integer such that t is a power of p and furthermore, $t = \sqrt{n\ell}$ for some $\ell \in \mathbb{R}$ such that $100 \leq \ell \leq \frac{1}{2}\cdot\ln(1/\varepsilon)$. Let $h \in s\mathcal{B}_n$ be any function such that $\mathrm{Spec}\, h(\lfloor n/2 \rfloor) \neq \mathrm{Spec}\, h(\lfloor n/2 \rfloor - t)$. Then, $\mathrm{pdeg}_\varepsilon(h) = \Omega(t)$.*

Proof. By error reduction for probabilistic polynomials (Fact 7 item 1), it suffices to prove an $\Omega(t)$ lower bound on $\mathrm{pdeg}_{\varepsilon/2}(h)$.

[8] Note that exactly which inputs are set to 0 or 1 is not important, since we are dealing with *symmetric* Boolean functions.

Assume without loss of generality that $\operatorname{Spec} h(\lfloor n/2 \rfloor) = 1$ and $\operatorname{Spec} h(\lfloor n/2 \rfloor - t) = 0$. Let $\boldsymbol{P}$ be an $(\varepsilon/2)$-error probabilistic polynomial for h. Then, we have

$$\Pr_{\boldsymbol{P}, \boldsymbol{a} \sim \{0,1\}^n_{\lfloor n/2 \rfloor}} [\boldsymbol{P}(\boldsymbol{a}) \neq 1] \leq \varepsilon/2$$

$$\Pr_{\boldsymbol{P}, \boldsymbol{b} \sim \{0,1\}^n_{\lfloor n/2 \rfloor - t}} [\boldsymbol{P}(\boldsymbol{b}) \neq 0] \leq \varepsilon/2$$

Thus, we have

$$\mathop{\mathbf{E}}_{\boldsymbol{P}} [\Pr_{\boldsymbol{a} \sim \{0,1\}^n_{\lfloor n/2 \rfloor}} [\boldsymbol{P}(\boldsymbol{a}) \neq 1] + \Pr_{\boldsymbol{b} \sim \{0,1\}^n_{\lfloor n/2 \rfloor - t}} [\boldsymbol{P}(\boldsymbol{b}) \neq 0]] \leq \varepsilon,$$

and hence, by averaging, there is a polynomial P in the support of the distribution of $\boldsymbol{P}$ such that

$$\Pr_{\boldsymbol{a} \sim \{0,1\}^n_{\lfloor n/2 \rfloor}} [P(\boldsymbol{a}) \neq 1] + \Pr_{\boldsymbol{b} \sim \{0,1\}^n_{\lfloor n/2 \rfloor - t}} [P(\boldsymbol{b}) \neq 0] \leq \varepsilon.$$

Applying Lemma 14 to P yields

$$\deg(\boldsymbol{P}) \geq \deg(P) = \Omega(t).$$

□

To illustrate the usefulness of Corollary 31, we prove optimal lower bounds on the probabilistic degrees for two interesting classes of functions (both of which will be subsumed by Theorem 25).

Corollary 32. *Let $\varepsilon \in (0, 1/3]$ be a constant. Let q be any integer relatively prime to p such that $q \leq 0.99n$. Then the ε-error probabilistic degrees of $\mathrm{EThr}_n^{\lfloor n/2 \rfloor}$ and MOD_n^q are $\Omega(\sqrt{n})$.*

Known lower bounds (Lemmas 29 and 30) can be used to prove similar lower bounds to the one given above, but with additional log-factor losses (see Lemma 30, which requires the error to be subconstant, and [STV19]). However, we do not know how to prove the above tight (up to constants) lower bound without appealing to Lemma 14. In particular, we do not know how to prove the above in characteristic 0.

Proof. We use Corollary 31. We will use $\mathrm{EThr}_n^{\lfloor n/2 \rfloor}$ and MOD_n^q to construct functions that distinguish between weights $\lfloor n/2 \rfloor$ and $\lfloor n/2 \rfloor - t$ for suitable $t = \Omega(\sqrt{n})$. Corollary 31 then implies the required lower bound.

For $h = \mathrm{EThr}_n^{\lfloor n/2 \rfloor}$, note that $\operatorname{Spec} h(\lfloor n/2 \rfloor) \neq \operatorname{Spec} h(\lfloor n/2 \rfloor - t)$ for any $t < \lfloor n/2 \rfloor$. In particular, setting t to be the smallest power of p such that $t \geq \sqrt{100n}$ and $\varepsilon_0 = e^{-2t^2/n}$, we get by Corollary 31 that $\mathrm{pdeg}_{\varepsilon_0}(h) = \Omega(t) = \Omega(\sqrt{n})$. By error-reduction for probabilistic polynomials (Fact 7 item 1), we also have the same lower bound (up to constant factors) for any $\varepsilon \leq 1/3$. This proves the claim in the case that $h = \mathrm{EThr}_n^{\lfloor n/2 \rfloor}$.

For $h = \mathrm{MOD}_n^q$, we make some minor modifications to the above idea. Let $r \in [0, q-1]$ be such that $r + \lfloor (n-q)/2 \rfloor \equiv 0 \pmod{q}$. Define $h' \in s\mathcal{B}_{n-q}$ by

$$h'(x) = h(x 1^r 0^{q-r}).$$

Set t to be the smallest power of p such that $t \geq \sqrt{100(n-q)}$ and $\varepsilon_0 = e^{-2t^2/(n-q)}$. Note that $\operatorname{Spec} h'(\lfloor (n-q)/2 \rfloor) = \operatorname{Spec} h(r + \lfloor (n-q)/2 \rfloor) = 1$ as $r + \lfloor (n-q)/2 \rfloor \equiv 0 \pmod{q}$. On the other hand, $r + \lfloor (n-q)/2 \rfloor - t \not\equiv 0 \pmod{q}$ as t is a power of p and hence not divisible by q, which implies that $\operatorname{Spec} h'(\lfloor (n-q)/2 \rfloor - t) = 0$. Thus, by Corollary 31, we get $\mathrm{pdeg}_{\varepsilon_0}(h') = \Omega(t) = \Omega(\sqrt{n})$. □

4.2.3 Proof of Theorem 25

The proof of this theorem closely follows our probabilistic degree lower bounds in [STV19] with careful modifications to avoid the log-factor losses therein.

Let $f \in s\mathcal{B}_n$ be arbitrary and let (g, h) be a standard decomposition of f.

We start with a lemma that proves lower bounds on $\mathrm{pdeg}_\varepsilon(f)$ as long as $\mathrm{per}(g)$ is large.

Lemma 33. *Fix any $\varepsilon \in [2^{-n}, 1/3]$. Assume that f is such that $\mathrm{per}(g) > \sqrt{n \log(1/\varepsilon)}$. Then*

$$\mathrm{pdeg}_\varepsilon(f) = \Omega(\sqrt{n \log(1/\varepsilon)}).$$

Proof. We first prove the lemma under the assumption that $\varepsilon \in [2^{-n/1000}, e^{-10000p^2}]$.

Fix m to be the largest power of p upper bounded by $\frac{1}{4}\sqrt{n \log(1/\varepsilon)}$.

Since $\mathrm{per}(g) > \sqrt{n \log(1/\varepsilon)} \geq m$, there is no function $g' \in s\mathcal{B}_n$ that has period m and agrees with f on the interval $I := [\lceil n/3 \rceil + 1, \lfloor 2n/3 \rfloor]$. Thus, there exists some $r \in I$ such that $r + m \in I$ and $\mathrm{Spec}\, f(r) \neq \mathrm{Spec}\, f(r + m)$.

Let $k = \lceil n/2 \rceil$. Note that $r \geq \lceil n/3 \rceil \geq k/2$ and $r + m \leq \lfloor 2n/3 \rfloor$. Define $F \in s\mathcal{B}_k$ by setting

$$F(x) = f(x1^a0^b)$$

where $a = r + m - \lfloor k/2 \rfloor$ and $b = n - k - a$ (it can be checked that a, b are non-negative for parameters r, m, k as above). Note that $\mathrm{Spec}\, F(\lfloor k/2 \rfloor) = \mathrm{Spec}\, f(\lfloor k/2 \rfloor + a) = \mathrm{Spec}\, f(r + m)$ and similarly that $\mathrm{Spec}\, F(\lfloor k/2 \rfloor - m) = \mathrm{Spec}\, f(r)$. We thus obtain $\mathrm{Spec}\, F(\lfloor k/2 \rfloor) \neq \mathrm{Spec}\, F(\lfloor k/2 \rfloor - m)$.

Note that by the bounds on ε assumed above

$$m \geq \frac{1}{4p}\sqrt{n \log(1/\varepsilon)} \geq 20\sqrt{n}. \tag{7}$$

Using Corollary 31, we hence get

$$\mathrm{pdeg}_\varepsilon(f) \geq \mathrm{pdeg}_{\varepsilon/2}(F) = \Omega(m) = \Omega(\sqrt{n \log(1/\varepsilon)})$$

which proves the lemma under the assumption on ε above. (We use the bounds on ε to ensure that $2^{-k/200} \leq \varepsilon \leq e^{-2m^2/k}$, which is part of the hypothesis of Corollary 31.)

If $\varepsilon \in [2^{-n}, 2^{-n/10000p^2}]$, then for $\varepsilon_0 = 2^{-n/10000p^2}$, we have

$$\mathrm{pdeg}_\varepsilon(f) \geq \mathrm{pdeg}_{\varepsilon_0}(f) = \Omega(\sqrt{n \log(1/\varepsilon_0)}) = \Omega(\sqrt{n \log(1/\varepsilon)})$$

which implies the desired lower bound.[9]

On the other hand, if $\varepsilon > e^{-10000p^2}$, we proceed as follows. We construct F as above, but we may no longer have $m \geq 20\sqrt{n}$ as implied by (7). However, for $F' \in s\mathcal{B}_{k'}$ defined by

$$F'(x) = F(x0^t1^t)$$

for suitably chosen $t \leq k/2$, we can ensure that $m \in [10\sqrt{k'}, 20\sqrt{k'}]$. Note that $\mathrm{Spec}\, F'(\lfloor k'/2 \rfloor) = \mathrm{Spec}\, F(\lfloor k/2 \rfloor)$ and $\mathrm{Spec}\, F'(\lfloor k'/2 \rfloor - m) = \mathrm{Spec}\, F(\lfloor k/2 \rfloor - m)$. Hence, for $\varepsilon_1 = e^{-10000}$, Corollary 31 implies

$$\mathrm{pdeg}_{\varepsilon_1}(f) \geq \mathrm{pdeg}_{\varepsilon_1}(F') = \Omega(m) = \Omega(\sqrt{n \log(1/\varepsilon_1)}).$$

By error reduction (Fact 7 item 1), the same lower bound holds for $\mathrm{pdeg}_\varepsilon(f)$ as well. □

[9] Note that we assume that the characteristic is a fixed positive constant and hence the $\Omega(\cdot)$ can hide constants depending on p.

The next lemma allows us to prove a weak lower bound on $\mathrm{pdeg}_\varepsilon(f)$ depending only on its periodic part g.

Lemma 34. *For any $\varepsilon \in [2^{-n}, 1/3]$,*

$$\mathrm{pdeg}_\varepsilon(f) \geq \begin{cases} \Omega(\sqrt{n\log(1/\varepsilon)}), & \text{if } \mathrm{per}(g) \text{ is not a power of } p \\ \Omega(\min\{\mathrm{per}(g), \sqrt{n\log(1/\varepsilon)}\}), & \text{if } \mathrm{per}(g) \text{ is a power of } p. \end{cases}$$

Proof. By Fact 7 item 1 (error reduction), we know that $\mathrm{pdeg}_\varepsilon(g) = \Theta(\mathrm{pdeg}_\delta(g))$ as long as $\delta = \varepsilon^{\Theta(1)}$. In particular, we may assume without loss of generality that $\varepsilon \in [2^{-n/10000}, e^{-10000p^2}]$.

Let $b := \mathrm{per}(g)$. If $\mathrm{per}(g) > \sqrt{n\log(1/\varepsilon)}$, we are done by Lemma 33. So we assume that $b \leq \sqrt{n\log(1/\varepsilon)}$. In particular, this implies that $b \leq n/100$.

We have two cases.

b is not a power of p. Let n_1 be the largest power of p upper bounded by $\frac{1}{4}\sqrt{n\log(1/\varepsilon)}$. By the constraints on ε, we have $10\sqrt{n} \leq n_1 \leq n/100$.

Let $b_1 \in [0, b-1]$ such that $b_1 \equiv n_1 \pmod b$; note that $b_1 \neq 0$ as b is not a power of p. As b_1 is smaller than $b = \mathrm{per}(g)$, there must exist $r \in [0, n - b_1]$ such that

$$\mathrm{Spec}\, g(r) \neq \mathrm{Spec}\, g(r + b_1).$$

Assume that we choose the smallest $r \geq n/2$ so that this condition holds. Then we have $r \leq n/2 + b \leq 51 \cdot n/100$. Fix this r. As $\mathrm{Spec}\, g(r) \neq \mathrm{Spec}\, g(r + b_1)$, we also have $\mathrm{Spec}\, g(r) \neq \mathrm{Spec}\, g(r + b_1 + k \cdot b)$ for any integer k such that $0 \leq r + b_1 + kb \leq n$. In particular, as $b_1 \equiv n_1 \pmod b$, we note that $\mathrm{Spec}\, g(r) \neq \mathrm{Spec}\, g(r + n_1)$. As $n_1 \leq n/100$, we have

$$n/2 \leq r \leq r + n_1 \leq n/2 + n/50.$$

As $\mathrm{Spec}\, g(i) = \mathrm{Spec}\, f(i)$ for all $i \in [\lceil n/3 \rceil + 1, \lfloor 2n/3 \rfloor]$, we have $\mathrm{Spec}\, f(r) \neq \mathrm{Spec}\, f(r + n_1)$. Without loss of generality, we assume that $\mathrm{Spec}\, f(r) = 0$ and $\mathrm{Spec}\, f(r + n_1) = 1$.

Let $m = \lceil n/2 \rceil$. We define $F \in s\mathcal{B}_m$ as follows.

$$F(x) = f(x1^a0^{n-m-a})$$

where a is chosen so that $\mathrm{Spec}\, F(\lfloor m/2 \rfloor) = \mathrm{Spec}\, f(r + n_1) = 1$. This also implies that $\mathrm{Spec}\, F(\lfloor m/2 \rfloor - n_1) = \mathrm{Spec}\, f(r) = 0$. By Corollary 31, we get $\mathrm{pdeg}_\varepsilon(F) = \Omega(n_1) = \Omega(\sqrt{n\log(1/\varepsilon)})$, proving the lemma in this case.

b is a power of p. In this case, we first choose parameters m, δ with the following properties.

(P1) $m \in [n]$ with $m \geq 20b$ and $m \equiv n \pmod 2$.

(P2) $1/3 \geq \delta \geq \max\{\varepsilon, 1/2^m\}$.

(P3) $\sqrt{m\log(1/\delta)} < b$.

(P4) $\sqrt{m\log(1/\delta)} = \Omega(\min\{b, \sqrt{n\log(1/\varepsilon)}\}) = \Omega(b)$. (Recall that $b \leq \sqrt{n\log(1/\varepsilon)}$.)

We will show below how to find m, δ satisfying these properties. Assuming this for now, we first prove the lower bound on $\mathrm{pdeg}_\varepsilon(f)$.

Define $F \in s\mathcal{B}_m$ as follows.

$$F(x) = f(x0^t1^t)$$

for $t = (n-m)/2$. We observe that if (G, H) is a standard decomposition of F, then $\mathrm{per}(G) \geq b$. To see this, note that by Corollary 10, we have

$$\operatorname{Spec} g|_{[\lfloor n/2 \rfloor, \lfloor n/2 \rfloor + b - 1]} \neq \operatorname{Spec} g|_{[\lfloor n/2 \rfloor + i, \lfloor n/2 \rfloor + i + b - 1]}$$

for any $i \in [b-1]$. As f and g agree on inputs of weight from $[\lfloor n/3 \rfloor + 1, \lfloor 2n/3 \rfloor]$, the same non-equality holds for $\operatorname{Spec} f$ also. Further, as $\operatorname{Spec} F(\lfloor m/2 \rfloor + j) = \operatorname{Spec} f(\lfloor n/2 \rfloor + j)$ for $j \leq m/2$, we also get

$$\operatorname{Spec} F|_{[\lfloor m/2 \rfloor, \lfloor m/2 \rfloor + b - 1]} \neq \operatorname{Spec} F|_{[\lfloor m/2 \rfloor + i, \lfloor m/2 \rfloor + i + b - 1]}.$$

for any $i \in [b-1]$ (we have used here the fact that $m \geq 20b$ which holds by (P1)). Finally, as F and G agree on inputs of weight from $[\lfloor m/3 \rfloor + 1, \lfloor 2m/3 \rfloor] \supseteq [\lfloor m/2 \rfloor, \lfloor m/2 \rfloor + 2b]$, the above non-equality holds for G as well. This implies that G cannot have period smaller than b.

By (P3), we have $\mathrm{per}(G) > \sqrt{m \log(1/\delta)}$. Lemma 33 above and (P4) now imply that $\mathrm{pdeg}_\delta(F) = \Omega(\min\{b, \sqrt{n \log(1/\varepsilon)}\})$. However, as $\delta \geq \varepsilon$ (by (P2)) and F is a restriction of f, the same lower bound holds for $\mathrm{pdeg}_\varepsilon(f)$ as well. This proves the lemma modulo the existence of m, δ as above. We justify this now.

1. If $b \leq 10\sqrt{n}$, we take m to be the largest integer such that $m \equiv n \pmod 2$ and $m \leq b^2/100$. The parameter δ is set to $1/3$.
2. If $10\sqrt{n} < b \leq n/100$, then we take m to be the largest integer such that $m \equiv n \pmod 2$ and $m \leq n/2$. The parameter $\delta = \max\{\varepsilon, 2^{-b^2/2m}\}$.

Note that as observed above, we have $b \leq n/100$, and hence, the above analysis subsumes all cases.

In each case, the verification of properties (P1)-(P4) is a routine computation. (We assume here that b is greater than a suitably large constant, since otherwise the statement of the lemma is trivial.) This concludes the proof. □

We now prove a lower bound on $\mathrm{pdeg}_\varepsilon(h)$.

Lemma 35. *Assume $B(h) \geq 1$. Then, $\varepsilon \in [2^{-n}, 1/3]$,*

$$\mathrm{pdeg}_\varepsilon(h) = \Omega(\sqrt{B(h) \log(1/\varepsilon)} + \log(1/\varepsilon)).$$

Proof. Similar to the proof of Lemma 34, we may assume without loss of generality that $\varepsilon \in [2^{-n/10000}, e^{-10000p^2}]$.

Let $B(h) = b$. Recall (Observation 5) that $B(h) \leq \lceil n/3 \rceil$. Further, by definition of $B(h)$, we have either $\operatorname{Spec} h(b-1) = 1$ or $\operatorname{Spec} h(n-b+1) = 1$. We assume that $\operatorname{Spec} h(n-b+1) = 1$ (the other case is similar).

The lemma is equivalent to showing that $\mathrm{pdeg}_\varepsilon(h) = \Omega(\max\{\sqrt{B(h) \log(1/\varepsilon)}, \log(1/\varepsilon)\})$. We do this based on a case analysis based on the relative magnitudes of $\log(1/\varepsilon)$ and b.

Assume for now that $\varepsilon \leq 2^{-b/1000}$. In this case, we show a lower bound of $\Omega(\log(1/\varepsilon))$. To see this, set $m = \lceil n/4 \rceil$ and consider the restriction $H \in s\mathcal{B}_m$ obtained as follows.

$$H(x) = h(x 1^{n-b+1-m} 0^{b-1}).$$

Note that as $\operatorname{Spec} h$ is the constant 0 function on the interval $[b, n-b]$, the function H is computing the AND function on m inputs. By Lemma 28, we immediately have $\mathrm{pdeg}_\varepsilon(h) \geq \mathrm{pdeg}_\varepsilon(H) = \Omega(\log(1/\varepsilon))$ proving the lemma in this case.

Now assume that $\varepsilon > 2^{-b/1000}$. In this case, we need to show a lower bound of $\Omega(\sqrt{b \log(1/\varepsilon)})$. To prove this, consider the restriction $H \in s\mathcal{B}_{2b-2}$ defined by $H(x) = h(x 1^{n-2b+2})$. Since $\operatorname{Spec} h$ is the constant 0 function on the interval $[b, n-b]$ and $\operatorname{Spec} h(n-b+1) = 1$, it follows that the periodic part of H has period $\Omega(b)$. It then follows from Lemma 33 that $\mathrm{pdeg}_\varepsilon(b) = \Omega(\sqrt{b \log(1/\varepsilon)})$. This concludes the proof of the lemma. □

Now, we are ready to prove Theorem 25.

Proof of Theorem 25. By Lemma 34, we already have the desired lower bound on $\mathrm{pdeg}_\varepsilon(f)$ in any of the following scenarios.

- $\mathrm{per}(g)$ is not a power of p, or
- $\mathrm{per}(g)$ is a power of p and $\mathrm{per}(g) \geq \sqrt{n \log(1/\varepsilon)}$, or
- $B(h) = 0$.

So from now, we assume that $\mathrm{per}(g)$ is a power of p upper-bounded by $\sqrt{n \log(1/\varepsilon)}$ and that $B(h) \geq 1$. In this case, Lemma 34 shows that $\mathrm{pdeg}(f) = \Omega(\mathrm{per}(g))$. On the other hand, since $B(h) \leq n$ and $\varepsilon \geq 2^{-n}$, the lower bound we need to show is $\Omega(\mathrm{per}(g) + \sqrt{B(h) \log(1/\varepsilon)} + \log(1/\varepsilon))$. By Lemma 35, it suffices to show a lower bound of $\Omega(\mathrm{per}(g) + \mathrm{pdeg}_\varepsilon(h))$.

The analysis splits into two simple cases.

Assume first that $\mathrm{pdeg}_\varepsilon(h) \leq 4 \cdot \mathrm{per}(g)$. In this case, we are trivially done, because we already have $\mathrm{pdeg}(f) = \Omega(\mathrm{per}(g))$, which is $\Omega(\mathrm{pdeg}(g) + \mathrm{pdeg}_\varepsilon(h))$ as a result of our assumption.

Now assume that $\mathrm{pdeg}_\varepsilon(h) > 4 \cdot \mathrm{per}(g)$. We know that $f = g \oplus h$ and hence $h = f \oplus g$. Hence, we have

$$\mathrm{pdeg}_\varepsilon(h) \leq 2(\mathrm{pdeg}_{\varepsilon/2}(f) + \mathrm{pdeg}_{\varepsilon/2}(g)) \leq O(\mathrm{pdeg}_\varepsilon(f)) + 2\mathrm{per}(g),$$

where the first inequality is a consequence of Fact 7 item 2 and the second follows from error-reduction and Theorem 8. The above yields

$$\mathrm{pdeg}_\varepsilon(f) = \Omega((\mathrm{pdeg}_\varepsilon(h) - 2 \cdot \mathrm{per}(g))) = \Omega(\mathrm{pdeg}_\varepsilon(h)) = \Omega(\mathrm{per}(g) + \mathrm{pdeg}_\varepsilon(h)).$$

This finishes the proof. □

4.3 A Robust Version of Galvin's Problem

We recall here a combinatorial theorem of Hegedűs [Heg09] regarding set systems. The theorem (and also our robust generalization given below) is easier to prove in the language of indicator vectors, so we state it in this language.

Given any vectors $u, v \in \mathbb{F}^k$ for any field $\mathbb{F}$, we define $\langle u, v \rangle := \sum_{j \in [n]} u_j v_j$.

Theorem 36. *Assume $n = 4p$, for a large enouh prime p. Let $u^{(1)}, \ldots, u^{(m)} \in \{0,1\}^n_{n/2} \subseteq \mathbb{Z}^n$ and $b_1, \ldots, b_m \leq (3/2 - \Omega(1)) \cdot p$ be such that for each $v \in \{0,1\}^n_{n/2}$, there is an $i \in [m]$ such that $\langle u^{(i)}, v \rangle = b_i$. Then $m \geq p$.*

The above theorem is nearly tight as can be seen by taking the indicator vectors of the sets $S_i = \{i \pmod n, (i+1) \pmod n, \ldots, i + (n/2) - 1 \pmod n\}$ for $i \in [n/2]$ and setting $b_1 = \cdots = b_m = n/4$. Improvements on the above theorem (some of them asymptotically tight) were proved recently by Alon et al. [AKV18] and Hrubeš et al. [HRRY19].

Using the robust version of Hegedűs's lemma, we can prove tight robust versions of the above statement.

Remark 37. *The statement given above is a slight generalization of the theorem in the paper of Hegedűs [Heg09], which proves the above only for the case when $b_1 = b_2 = \cdots = b_m = n/4$. However, the stronger result stated above follows immediately from the proof techniques. Our robust generalization (stated below) is tight for this stronger statement. However, if one we consider the robust versions of the original statements in [Heg09], then while our lower bound continues to hold, it is not clear whether it is tight (except in the settings where ε is either a constant or $2^{-\Omega(n)}$). We conjecture that it is.*

We now prove a robust version of Theorem 36.

Theorem 38. *Assume n is a growing even integer parameter and $\varepsilon \in [2^{-n}, 1/2]$.. Let $u^{(1)}, \ldots, u^{(m)} \in \{0,1\}^n_{n/2} \subseteq \mathbb{Z}^n$ and $b_1, \ldots, b_m \leq n$ be such that*

$$\Pr_{\boldsymbol{v} \sim \{0,1\}^n_{n/2}} \left[\exists i \in [m] \text{ s.t. } \langle u^{(i)}, \boldsymbol{v} \rangle = b_i \right] \geq 1 - \varepsilon.$$

Then $m = \Omega(\sqrt{n \log(1/\varepsilon)})$.

The theorem can easily seen to be tight up to constant factors. For $t = C \cdot \sqrt{n \log(1/\varepsilon)}$, set $m = 2t+1$ and take $u^{(1)} = u^{(2)} = \cdots = u^{(m)} = 1^{n/2}0^{n/2}$ and $b_1 = (n/4) - t, b_2 = (n/4) - t + 1, \ldots, b_m = (n/4) + t$. By standard Chernoff bounds for the Hypergeometric distribution, we immediately get that this set of hyperplanes satisfy the above condition for a large enough choice of the constant C.

We need the following standard bound on binomial coefficients. For completeness, we include the proof in Appendix C.

Claim 39. *Let n be an even integer and m a non-negative integer with $m \leq n/2$. Then, for any $k, \ell \in \{0, \ldots, \lfloor m/2 \rfloor\}$ with $\ell \leq k$, we have*

$$\frac{\binom{n/2}{\lfloor m/2 \rfloor - k}\binom{n/2}{\lceil m/2 \rceil + k}}{\binom{n/2}{\lfloor m/2 \rfloor - \ell}\binom{n/2}{\lceil m/2 \rceil + \ell}} \leq \exp(-\Omega((k^2 - \ell^2)/m)).$$

Given the above, we can prove Theorem 38 as follows.

Proof of Theorem 38. Recall that for any fixed $u \in \{0,1\}^n_{n/2}$ and any $b \in \mathbb{Z}$, the probability that a uniformly random $v \in \{0,1\}^n$ satisfies $\langle u, v \rangle = b$ is at most $O(1/\sqrt{n})$. In particular, we must have $m = \Omega(\sqrt{n})$ for any $\varepsilon \leq 1/2$. This proves the result for $\varepsilon = \Omega(1)$.

Hence, we may assume that ε is smaller than any fixed constant. We can also assume that $\varepsilon \geq 2^{-\delta n}$ for a small enough constant δ. Assume that $m \leq \sqrt{n \log(1/\varepsilon)}$.

We call $i \in [m]$ *balanced* if $|b_i - \frac{n}{4}| \leq t$ where $t := C\sqrt{n \log(1/\varepsilon)}$ for a large enough constant C. If i is not balanced, then we have for a uniformly random $\boldsymbol{v} \sim \{0,1\}^n_{n/2}$,

$$\Pr_{\boldsymbol{v}} \left[\langle u^{(i)}, \boldsymbol{v} \rangle = b_i \right] \leq \frac{\binom{n/2}{n/4+t}\binom{n/2}{n/4-t}}{\binom{n}{n/2}} \leq \exp(-\Omega(t^2/n)) \frac{\binom{n/2}{n/4}^2}{\binom{n}{n/2}} < \frac{\varepsilon^2}{\sqrt{n}}.$$

The second inequality above follows from Claim 39, and the third follows from the Stirling approximation and using the fact that C is a large enough constant. In particular, if B is the set of balanced i, we have

$$\Pr_{\boldsymbol{v}} \left[\exists i \notin B, \ \langle u^{(i)}, \boldsymbol{v} \rangle = b_i \right] \leq m \cdot \frac{\varepsilon^2}{\sqrt{n}} < \varepsilon$$

where we used the fact that $m \leq \sqrt{n \log(1/\varepsilon)}$. We can thus consider only $\{u^{(i)} \mid i \in B\}$, which satisfy the hypothesis with error probability $\varepsilon_1 := 2\varepsilon$.

Now consider the polynomial

$$P(x_1, \ldots, x_n) = \prod_{i \in B} (\langle u^{(i)}, x \rangle - b_i).$$

We know that P vanishes at a random point of $\{0,1\}^n_{n/2}$ with probability at least $1-\varepsilon_1$. Now, fix any prime $p \in [10t, 20t]$ (such a prime exists by standard number-theoretic results). We claim that for any $i \in B$ and a uniformly random point $\boldsymbol{v} \in \{0,1\}^n_{n/2-p}$, we have

$$\Pr_{\boldsymbol{v}}\left[\langle u^{(i)}, \boldsymbol{v}\rangle \equiv b_i \pmod p\right] \leq \frac{\varepsilon^2}{\sqrt{n}} \tag{8}$$

for a large enough constant C. Informally speaking, the reason for this inequality is as follows: the expected value of $\langle u^{(i)}, \boldsymbol{v}\rangle$ is $(n/4) - p/2$ and any number $b \equiv b_i \pmod p$ is far from this expectation. To prove this, let $s = n/2 - p$. Using the fact that i is balanced, we note that

$$\begin{aligned}\Delta_i := b_i - \left\lceil \frac{s}{2} \right\rceil &\leq \frac{n}{4} + t - \left(\frac{n}{4} - \left\lceil \frac{p}{2} \right\rceil\right) \leq \frac{2p}{3}\\ \Delta_i &\geq \frac{n}{4} - t - \left(\frac{n}{4} - \left\lceil \frac{p}{2} \right\rceil\right) \geq \frac{p}{3}.\end{aligned}$$

We thus have

$$\begin{aligned}\Pr_{\boldsymbol{v}}\left[\langle u^{(i)}, \boldsymbol{v}\rangle \equiv b_i \pmod p \wedge \langle u^{(i)}, \boldsymbol{v}\rangle \geq \left\lceil \frac{s}{2} \right\rceil\right] &= \sum_{j\geq 0} \Pr_{\boldsymbol{v}}\left[\langle u^{(i)}, \boldsymbol{v}\rangle = b_i + jp\right]\\ &= \sum_{j\geq 0} \Pr_{\boldsymbol{v}}\left[\langle u^{(i)}, \boldsymbol{v}\rangle - \left\lceil \frac{s}{2} \right\rceil = \Delta_i + jp\right]\\ &= \sum_{j\geq 0} \frac{\binom{n/2}{\lceil \frac{s}{2}\rceil + \Delta_i + jp}\binom{n/2}{\lfloor \frac{s}{2}\rfloor - \Delta_i - jp}}{\binom{n}{s}}\\ \text{(by Claim 39)} &= \frac{\binom{n/2}{\lceil \frac{s}{2}\rceil}\binom{n/2}{\lfloor \frac{s}{2}\rfloor}}{\binom{n}{s}} \cdot \sum_{j\geq 0} \exp(-\Omega((\Delta_i + jp)^2)/s))\\ \text{(Stirling approximation and } s = \Omega(n) \text{ as long as } \sqrt{\delta}C \leq 1/100) &\leq O\left(\frac{1}{\sqrt{n}}\right) \cdot \sum_{j\geq 0} \exp(-\Omega(\Delta_i^2 + jp^2)/s)\\ (p^2/s \geq C^2\ ) &\leq O\left(\frac{1}{\sqrt{n}}\right) \cdot \exp(-\Omega(\Delta_i^2/s)) \cdot \sum_{j\geq 0} \exp(-\Omega(C^2 j))\\ \text{(for large enough } C\ ) &\leq O\left(\frac{1}{\sqrt{n}}\right) \cdot \exp(-\Omega(\Delta_i^2/s)) \cdot 2\\ &= O\left(\frac{1}{\sqrt{n}}\right) \cdot \exp(-\Omega(p^2/s)).\end{aligned}$$

In a similar way, we also get

$$\Pr_{\boldsymbol{v}}\left[\langle u^{(i)}, \boldsymbol{v}\rangle \equiv b_i \pmod p \wedge \langle u^{(i)}, \boldsymbol{v}\rangle \leq \left\lfloor \frac{s}{2} \right\rfloor\right] \leq O\left(\frac{1}{\sqrt{n}}\right) \cdot \exp(-\Omega(p^2/s)).$$

Overall, we thus obtain for any $i \in B$,

$$\Pr_{\boldsymbol{v}}\left[\langle u^{(i)}, \boldsymbol{v}\rangle \equiv b_i \pmod p\right] \leq O\left(\frac{1}{\sqrt{n}}\right) \cdot \exp(-\Omega(p^2/s)) \leq \frac{\varepsilon^2}{\sqrt{n}}$$

as long as C is a large enough constant. Union bounding over the at most $m \leq \sqrt{n\log(1/\varepsilon)}$ elements of B, we see that

$$\Pr_{\boldsymbol{v}\in\{0,1\}^n_{n/2-p}}[P(\boldsymbol{v}) \equiv 0 \pmod p] \leq \varepsilon.$$

From now on, we consider the polynomial P as an element of $\mathbb{F}_p[x_1, \ldots, x_n]$. At this point, we would like to apply Lemma 13 to the polynomial P and finish the proof. Unfortunately, the error parameter ε_1 above is not small enough to apply Lemma 13 directly (we need $\varepsilon_1 \leq \exp(-200p^2/n)$). However, we can do a simple error reduction as in Lemma 21 to ensure that Lemma 13 is applicable. More precisely, choose r to be a large enough absolute constant so that $\varepsilon_1^r \leq \frac{1}{2}\exp(-200p^2/n)$. Now, by Lemma 21 there is a probabilistic polynomial $\boldsymbol{P^{(r)}}$ of degree at most $r \cdot \deg(P)$ such that

$$\Pr_{\boldsymbol{v}\sim\{0,1\}^n_{n/2},\boldsymbol{P^{(r)}}}\left[\boldsymbol{P^{(r)}}(\boldsymbol{v}) = 0\right] \leq \varepsilon_1^{2r} \leq \frac{1}{2}\exp(-200p^2/n), \text{ and}$$

$$\Pr_{\boldsymbol{v}\sim\{0,1\}^n_{n/2-p},\boldsymbol{P^{(r)}}}\left[\boldsymbol{P^{(r)}}(\boldsymbol{v}) \neq 0\right] \geq (1-\varepsilon)^r \geq 1 - r\varepsilon \geq \frac{9}{10}$$

where for the last inequality we used the fact that ε is smaller than some absolute constant.

By a simple union bound, there is a fixed polynomial $P' \in \mathbb{F}_p[x_1, \ldots, x_n]$ of degree $r \cdot \deg(P) = O(m)$ such that

$$\Pr_{\boldsymbol{v}\sim\{0,1\}^n_{n/2}}\left[P'(\boldsymbol{v}) = 0\right] \leq \exp(-200p^2/n), \text{ and}$$

$$\Pr_{\boldsymbol{v}\sim\{0,1\}^n_{n/2-p}}\left[P'(\boldsymbol{v}) \neq 0\right] \geq \frac{1}{2}$$

Hence, applying Lemma 13 to the polynomial P', we get $\deg(P') = \Omega(p) = \Omega(\sqrt{n\log(1/\varepsilon)})$. This yields the desired lower bound on m. □

Acknowledgements. The author is grateful to Mrinal Kumar, Nutan Limaye, Utkarsh Tripathi and S. Venkitesh for useful discussions, feedback, and encouragement. The author thanks Nutan Limaye for suggesting the robust version of Galvin's problem as an application. The author is also grateful to the anonymous referees of STOC 2020 for their corrections and suggestions.

References

[ABFR94] James Aspnes, Richard Beigel, Merrick L. Furst, and Steven Rudich. The expressive power of voting polynomials. *Combinatorica*, 14(2):135–148, 1994.

[ABO84] Miklós Ajtai and Michael Ben-Or. A theorem on probabilistic constant depth computations. In *STOC*, pages 471–474, 1984.

[Agr19] Rohit Agrawal. Coin theorems and the fourier expansion. *Electronic Colloquium on Computational Complexity (ECCC)*, 26:87, 2019.

[AKV18] Noga Alon, Mrinal Kumar, and Ben Lee Volk. Unbalancing sets and an almost quadratic lower bound for syntactically multilinear arithmetic circuits. In *33rd Computational Complexity Conference, CCC 2018, June 22-24, 2018, San Diego, CA, USA*, pages 11:1–11:16, 2018.

[AW15] Josh Alman and Ryan Williams. Probabilistic polynomials and hamming nearest neighbors. In *2015 IEEE 56th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 136–150. IEEE, 2015.

[AWY15] Amir Abboud, Richard Ryan Williams, and Huacheng Yu. More applications of the polynomial method to algorithm design. In *Proceedings of the Twenty-Sixth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2015, San Diego, CA, USA, January 4-6, 2015*, pages 218–230, 2015.

[Bei93] Richard Beigel. The polynomial method in circuit complexity. *[1993] Proceedings of the Eigth Annual Structure in Complexity Theory Conference*, pages 82–95, 1993.

[BHMS18] Siddharth Bhandari, Prahladh Harsha, Tulasimohan Molli, and Srikanth Srinivasan. On the Probabilistic Degree of OR over the Reals. In Sumit Ganguly and Paritosh Pandya, editors, *38th IARCS Annual Conference on Foundations of Software Technology and Theoretical Computer Science (FSTTCS 2018)*, volume 122 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 5:1–5:12, Dagstuhl, Germany, 2018. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik.

[BK03] Jean Berstel and Juhani Karhumäki. Combinatorics on words: a tutorial. *Bulletin of the EATCS*, 79:178, 2003.

[Bra10] Mark Braverman. Polylogarithmic independence fools AC^0 circuits. *J. ACM*, 57(5), 2010.

[BV10] Joshua Brody and Elad Verbin. The coin problem and pseudorandomness for branching programs. In *FOCS*, pages 30–39. IEEE Computer Society, 2010.

[CGR14] Gil Cohen, Anat Ganor, and Ran Raz. Two sides of the coin problem. In *APPROX-RANDOM*, volume 28 of *LIPIcs*, pages 618–629. Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik, 2014.

[CHLT19] Eshan Chattopadhyay, Pooya Hatami, Shachar Lovett, and Avishay Tal. Pseudorandom generators from the second fourier level and applications to AC0 with parity gates. In *10th Innovations in Theoretical Computer Science Conference, ITCS 2019, January 10-12, 2019, San Diego, California, USA*, pages 22:1–22:15, 2019.

[CL69] G.F. Clements and B. Lindström. A generalization of a combinatorial theorem of macaulay. *Journal of Combinatorial Theory*, 7(3):230 – 238, 1969.

[CLP17] Ernie Croot, Vsevolod F Lev, and Péter Pál Pach. Progression-free sets in are exponentially small. *Annals of Mathematics*, pages 331–337, 2017.

[DP09] Devdatt P Dubhashi and Alessandro Panconesi. *Concentration of measure for the analysis of randomized algorithms*. Cambridge University Press, 2009.

[EFIN87] Hikoe Enomoto, Peter Frankl, Noboru Ito, and Kazumasa Nomura. Codes with given distances. *Graphs and Combinatorics*, 3(1):25–38, 1987.

[EG17] Jordan S Ellenberg and Dion Gijswijt. On large subsets of with no three-term arithmetic progression. *Annals of Mathematics*, pages 339–343, 2017.

[GII+19] Alexander Golovnev, Rahul Ilango, Russell Impagliazzo, Valentine Kabanets, Antonina Kolokolova, and Avishay Tal. Ac^0[p] lower bounds against MCSP via the coin problem. In *46th International Colloquium on Automata, Languages, and Programming, ICALP 2019, July 9-12, 2019, Patras, Greece.*, pages 66:1–66:15, 2019.

[Gut16] Larry Guth. *Polynomial methods in combinatorics*, volume 64. American Mathematical Soc., 2016.

[Heg09] Gábor Hegedűs. Balancing sets of vectors. *Studia Scientiarum Mathematicarum Hungarica*, 47(3):333–349, 2009.

[HRRY19] Pavel Hrubes, Sivaramakrishnan Natarajan Ramamoorthy, Anup Rao, and Amir Yehudayoff. Lower bounds on balancing sets and depth-2 threshold circuits. *Electronic Colloquium on Computational Complexity (ECCC)*, 26:26, 2019.

[HS16] Prahladh Harsha and Srikanth Srinivasan. On Polynomial Approximations to AC^0. In Klaus Jansen, Claire Mathieu, José D. P. Rolim, and Chris Umans, editors, *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques (APPROX/RANDOM 2016)*, volume 60 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 32:1–32:14, Dagstuhl, Germany, 2016. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik.

[KOS04] Adam R. Klivans, Ryan O'Donnell, and Rocco A. Servedio. Learning intersections and thresholds of halfspaces. *J. Comput. Syst. Sci.*, 68(4):808–840, 2004.

[KP17] Alexander S. Kulikov and Vladimir V. Podolskii. Computing majority by constant depth majority circuits with low fan-in gates. In *34th Symposium on Theoretical Aspects of Computer Science, STACS 2017, March 8-11, 2017, Hannover, Germany*, pages 49:1–49:14, 2017.

[KS04] Adam R. Klivans and Rocco A. Servedio. Learning DNF in time $2^{\tilde{O}(n^{1/3})}$. *J. Comput. Syst. Sci.*, 68(2):303–318, 2004.

[KS05] Peter. Keevash and Benny. Sudakov. Set systems with restricted cross-intersections and the minimum rank ofinclusion matrices. *SIAM Journal on Discrete Mathematics*, 18(4):713–727, 2005.

[KS18] Swastik Kopparty and Srikanth Srinivasan. Certifying polynomials for $\mathrm{AC}^0[\oplus]$ circuits, with applications to lower bounds and circuit compression. *Theory of Computing*, 14(1):1–24, 2018.

[LMN93] Nathan Linial, Yishay Mansour, and Noam Nisan. Constant depth circuits, fourier transform, and learnability. *J. ACM*, 40(3):607–620, 1993.

[LSS+18] Nutan Limaye, Karteek Sreenivasaiah, Srikanth Srinivasan, Utkarsh Tripathi, and S. Venkitesh. The coin problem in constant depth: Sample complexity and parity gates. *Electronic Colloquium on Computational Complexity (ECCC)*, 25:157, 2018.

[Lu01] Chi-Jen Lu. An exact characterization of symmetric functions in $\mathrm{qAC}^0[2]$. *Theoretical Computer Science*, 261(2):297–303, 2001.

[MNV16] Raghu Meka, Oanh Nguyen, and Van Vu. Anti-concentration for polynomials of independent random variables. *Theory of Computing*, 12(1):1–17, 2016.

[NW15] Zipei Nie and Anthony Y. Wang. Hilbert functions and the finite degree zariski closure in finite field combinatorial geometry. *Journal of Combinatorial Theory, Series A*, 134:196 – 220, 2015.

[Pot19] Aditya Potukuchi. On the $\text{AC}^0[\oplus]$ complexity of andreev's problem. *Electronic Colloquium on Computational Complexity (ECCC)*, 26:96, 2019.

[Raz87] Alexander A. Razborov. Lower bounds on the size of bounded depth circuits over a complete basis with logical addition. *Mathematicheskie Zametki*, 41(4):598–607, 1987. (English translation in *Mathematical Notes of the Academy of Sciences of the USSR*, 41(4):333–338, 1987).

[RSY08] Ran Raz, Amir Shpilka, and Amir Yehudayoff. A lower bound for the size of syntactically multilinear arithmetic circuits. *SIAM J. Comput.*, 38(4):1624–1647, 2008.

[Smo87a] Roman Smolensky. Algebraic methods in the theory of lower bounds for boolean circuit complexity. In *Proceedings of the nineteenth annual ACM symposium on Theory of computing*, pages 77–82. ACM, 1987.

[Smo87b] Roman Smolensky. Algebraic methods in the theory of lower bounds for boolean circuit complexity. In *Proceedings of the 19th Annual ACM Symposium on Theory of Computing*, pages 77–82, 1987.

[Smo93] Roman Smolensky. On representations by low-degree polynomials. In *Proceedings of 1993 IEEE 34th Annual Foundations of Computer Science*, pages 130–138. IEEE, 1993.

[STV19] Srikanth Srinivasan, Utkarsh Tripathi, and S. Venkitesh. On the probabilistic degrees of symmetric boolean functions. *Electronic Colloquium on Computational Complexity (ECCC)*, 26:138, 2019.

[SV10] Ronen Shaltiel and Emanuele Viola. Hardness amplification proofs require majority. *SIAM J. Comput.*, 39(7):3122–3154, 2010.

[SY10] Amir Shpilka and Amir Yehudayoff. Arithmetic circuits: A survey of recent results and open questions. *Foundations and Trends in Theoretical Computer Science*, 5(3-4):207–388, 2010.

[Sze89] M. Szegedy. Algebraic methods in lower bounds for computational models with limited communication. *PhD thesis*, The University of Chicago, 1989.

[Val84] Leslie G. Valiant. Short monotone formulae for the majority function. *J. Algorithms*, 5(3):363–366, 1984.

[Vio09] Emanuele Viola. On approximate majority and probabilistic time. *Computational Complexity*, 18(3):337–375, 2009.

[Wei91] V. K. Wei. Generalized hamming weights for linear codes. *IEEE Transactions on Information Theory*, 37(5):1412–1418, Sep. 1991.

[Wil14a] Richard Ryan Williams. The polynomial method in circuit complexity applied to algorithm design (invited talk). In *34th International Conference on Foundation of Software Technology and Theoretical Computer Science (FSTTCS 2014)*. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, 2014.

[Wil14b] Ryan Williams. New algorithms and lower bounds for circuits with linear threshold gates. *CoRR*, abs/1401.2444, 2014.

[Wil14c] Ryan Williams. Nonuniform ACC circuit lower bounds. *J. ACM*, 61(1):2:1–2:32, 2014.

[Wil18] R. Ryan Williams. Faster all-pairs shortest paths via circuit complexity. *SIAM J. Comput.*, 47(5):1965–1985, 2018.

A Lemma 1 is implied by Lemma 13 (up to constant factors)

The following claim shows that if there is a P satisfying the hypotheses of Lemma 1, then there is also a polynomial Q of degree at most $\deg(P)$ satisfying a stronger property, namely, that of not vanishing at too many points of $\{0,1\}^n_{k+q}$.

Claim 40. *Let $\mathbb{F}$ be a field of characteristic $p > 0$. Fix any positive integers n, k, q such that $k \in [q, n-q]$, and q a power of p. If there is a polynomial $P \in \mathbb{F}[x_1, \ldots, x_n]$ is any multilinear polynomial that vanishes at* all *$a \in \{0,1\}^n_k$ but does not vanish at* some *$b \in \{0,1\}^n_{k+q}$, then there is a multilinear $Q \in \mathbb{F}[x_1, \ldots, x_n]$ of degree at most* $\deg(P)$ *such that Q vanishes at all $a \in \{0,1\}^n_k$ but is non-zero at at least a $(1 - 1/p)$ fraction of the points in $\{0,1\}^n_{k+q}$.*

Proof. Let $d = \deg(P)$. Assume without loss of generality that $P(b) = 1$. Note that P is the solution to the system of linear equations defined by the following constraints on polynomials of degree at most d.

$$|a| = k \Rightarrow P(a) = 0$$
$$P(b) = 1.$$

As the above linear system is over $\mathbb{F}_p \subseteq \mathbb{F}$, we note that we may assume that $P \in \mathbb{F}_p[x_1, \ldots, x_n]$. From now on, we assume that $\mathbb{F} = \mathbb{F}_p$.

Consider the degree-d closure $C = \mathrm{cl}_d(\{0,1\}^n_k)$. By the existence of P, we see that $b \notin C$. However, by symmetry, this implies that no point $b' \in \{0,1\}^n_{k+q}$ lies in C.

Let $V_{d,k}$ denote the vector space of all multilinear polynomials of degree at most d that vanish at all points in $\{0,1\}^n_k$. Let $\boldsymbol{Q}$ be a uniformly random element of $V_{d,k}$. For any $c \in \{0,1\}^n \setminus C$, standard linear algebra implies that $\boldsymbol{Q}(c)$ is a uniformly random element of $\mathbb{F} = \mathbb{F}_p$. In particular, for any $b' \in \{0,1\}^n_{k+q}$, we see that

$$\Pr_{\boldsymbol{Q}}\left[\boldsymbol{Q}(b') \neq 0\right] = 1 - 1/p.$$

In particular, there is a $Q \in V_{d,k}$ that is non-zero at at least a $(1 - 1/p)$ fraction of points in $\{0,1\}^n_{k+q}$. This yields the statement of the claim. □

B Proof of Lemma 9 (the string lemma)

We begin by recalling the statement of the lemma.

Lemma 9 (Restated)**.** *Let $w \in \{0,1\}^+$ be any non-empty string and $u, v \in \{0,1\}^+$ such that $w = uv = vu$. Then there exists a string $z \in \{0,1\}^+$ such that w is a power of z (i.e. $w = z^k$ for some $k \geq 2$).*

Proof. Assume that $|u| = \ell, |v| = m$ and $|w| = \ell + m = n$. We will show in fact that both u and v are powers of the same non-empty string z. This will clearly imply the lemma.

The proof is by induction on the length of w. The base case of the induction corresponds to $n = 2$, which is obvious.

We now proceed with the inductive case. Assume w.l.o.g. that $\ell \leq m$. As $uv = vu$, we see that the first ℓ symbols in v match those of u, and hence we have $v = uv'$ for some $v' \in \{0,1\}^{m-\ell}$. If $\ell = m$, this implies that $u = v$ and we are immediately done. Otherwise, we see that $w = uv'u = v'uu$ for a non-empty string v'. Hence, we have $uv' = v'u$. By the induction hypothesis, we know that both u and v' are powers of some non-empty z. Hence, so is v. This concludes the proof. □

C Proof of Claim 39

Claim 39 (Restated)**.** *Let n be an even integer and m a non-negative integer with $m \le n/2$. Then, for any $k, \ell \in \{0, \ldots, \lfloor m/2 \rfloor\}$ with $\ell \le k$, we have*

$$\frac{\binom{n/2}{\lfloor m/2 \rfloor - k}\binom{n/2}{\lceil m/2 \rceil + k}}{\binom{n/2}{\lfloor m/2 \rfloor - \ell}\binom{n/2}{\lceil m/2 \rceil + \ell}} \le \exp(-\Omega((k^2 - \ell^2)/m)).$$

Proof. It suffices to show that for each $k \in \{0, \ldots, \lfloor m/2 \rfloor - 1\}$,

$$\frac{\binom{n/2}{\lfloor m/2 \rfloor - k - 1}\binom{n/2}{\lceil m/2 \rceil + k + 1}}{\binom{n/2}{\lfloor m/2 \rfloor - k}\binom{n/2}{\lceil m/2 \rceil + k}} \le \exp(-\Omega(k/m)). \tag{9}$$

The claim then follows by a simple induction on $k - \ell$.

To prove (9), we proceed as follows. By an expansion of binomial coefficients in terms of factorials, we see that

$$\begin{aligned}
\frac{\binom{n/2}{\lfloor m/2 \rfloor - k - 1}\binom{n/2}{\lceil m/2 \rceil + k + 1}}{\binom{n/2}{\lfloor m/2 \rfloor - k}\binom{n/2}{\lceil m/2 \rceil + k}} &= \frac{(\lfloor m/2 \rfloor - k)(n/2 - (\lceil m/2 \rceil + k))}{(n/2 - (\lfloor m/2 \rfloor - k - 1))(\lceil m/2 \rceil + k + 1)} \\
&\le \frac{\lfloor m/2 \rfloor - k}{\lceil m/2 \rceil + k + 1} \\
&\le \frac{(m/2) - k}{(m/2)} \le 1 - 2k/m \le \exp(-2k/m).
\end{aligned}$$

□

ECCC ISSN 1433-8092
https://eccc.weizmann.ac.il