

A Robust Version of Hegedűs's Lemma, with Applications

Received Feb 28, 2022
 Revised Oct 21, 2022
 Accepted Dec 20, 2022
 Published Feb 24, 2023

Key words and phrases
 Polynomial approximation,
 Boolean functions, Probabilistic
 degree, Coin Problem

Srikanth Srinivasan^a  

^a Department of Computer
 Science, Aarhus University,
 Denmark

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 $P \in \mathbb{F}[x_1, \dots, x_n]$ of degree less than q that vanishes at all points in $\{0, 1\}^n$ of some fixed 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

Work done while at the Department of Mathematics, Indian Institute of Technology Bombay, Mumbai, India. Supported by MATRICS grant MTR/2017/000958 awarded by SERB, Government of India. A preliminary version of this paper appeared at STOC 2020.

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 [38, 41, 8, 52], Pseudorandom generators [11]), Algorithm design (Learning Algorithms [32, 28, 27], Satisfiability algorithms [52, 51], Combinatorial algorithms [49, 1, 4]), and Extremal Combinatorics [21, 16, 18].

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.

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 $m \in \{0, \dots, n\}$, let $\{0, 1\}_m^n$ be the elements of $\{0, 1\}^n$ of Hamming weight exactly m . As a first approximation, let us say that a polynomial $P \in \mathbb{F}[x_1, \dots, x_n]$ (here \mathbb{F} is some field) distinguishes between level sets $\{0, 1\}_k^n$ and $\{0, 1\}_K^n$ 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 $\{0, 1\}_k^n$ and $\{0, 1\}_K^n$. A very interesting lemma of Hegedűs [23] shows that this is tight even if we only require P to be non-zero at *some* point of $\{0, 1\}_K^n$. More precisely, Hegedűs’s lemma shows the following.¹

LEMMA 1.1 (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 $P \in \mathbb{F}[x_1, \dots, x_n]$ is any polynomial that vanishes at all $a \in \{0, 1\}_k^n$ but does not vanish at some $b \in \{0, 1\}_{k+q}^n$, then $\deg(P) \geq q$.*

¹ The lemma is usually stated [23, 5, 25] for a more restricted choice of parameters. However, the known proofs extend to yield the stronger statement given here. A proof of a more general statement can be found in [44, Theorem 1.5].

This lemma was first proved in [23] using Gröbner basis techniques. An elementary proof of this was recently given by the author and independently by Alon (see [25]) 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 Nomura [19] 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. [5] and more recently to a near-tight lower bound of $(n/2) - o(n)$ for all t by Hrubeš et al. [25]. Both these results used the lemma above.
- Alon et al. [5] 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 *syntactically multilinear algebraic circuits*. These are algebraic circuits that compute multilinear polynomials in a “transparently multilinear” way (see e.g. [40] 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 [37].
- Hrubeš et al. [25] also used Hegedűs's lemma to answer the following question of Kulikov and Podolskii [30] on depth-2 threshold circuits. What is the smallest $k = k(n)$ such that there is a depth-2 circuit made up of Majority² 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 $k(n)$.

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 $\{0, 1\}_k^n$ and $\{0, 1\}_{k+q}^n$: this means that the polynomial P vanishes at ‘most’ points of weight k but is non-zero at ‘many’

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

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 1.1 above.

LEMMA 1.2 (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 $P \in \mathbb{F}[x_1, \dots, x_n]$ that vanishes at a $(1 - \varepsilon)$ -fraction of points of $\{0, 1\}_k^n$ but does not vanish at an $\varepsilon^{0.0001}$ fraction of points of $\{0, 1\}_{k+q}^n$, then $\deg(P) = \Omega(q)$.*

REMARK 1.3. 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 Hegedús’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 1.1. 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 1.1 (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 1.1 (up to the constant-factor losses in the degree lower bound).

Applications. Our investigations into robust versions of Hegedús’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 . (See Definition 4.1 for the formal definition.)

The coin problem has been studied in a variety of settings in complexity theory (see, e.g. [3, 46, 47, 39, 12, 15]) 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. [31] proved optimal lower bounds on the size of $AC^0[\oplus]$ ³ circuits

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

solving the δ -coin problem with constant error, strengthening an earlier lower bound of Shaltiel and Viola [39]. This led to the first class of explicit functions for which we have tight (up to polynomial factors) $AC^0[\oplus]$ lower bounds. These bounds were in turn used by Golovnev, Ilango, Impagliazzo, Kabanets, Kolokolova and Tal [20] to resolve a long-standing open problem regarding the complexity of MCSP in the $AC^0[\oplus]$ model, and by Potukuchi [36] to prove lower bounds for Andreev's problem.

A key result in the lower bound of Limaye et al. [31] was a tight lower bound on the degree of any polynomial $P \in \mathbb{F}[x_1, \dots, x_N]$ 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 [2], this is essentially equivalent to a recent result of Chattopadhyay, Hatami, Lovett and Tal [13] on the level-1 Fourier coefficients of low-degree polynomials over finite fields, which in turn is connected to an intriguing new approach [13] toward constructing pseudorandom generators secure against $AC^0[\oplus]$.

Using the robust Hegedűs lemma, we are able to strengthen the degree lower bound of [31] 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 [38], 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 $f : \{0, 1\}^n \rightarrow \{0, 1\}$ is a random polynomial \mathbf{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 [8] and algorithm design [50] 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 [43], 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., [4, 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.

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

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 [34, 22, 10]. However, in positive (fixed) characteristic, there are no obvious barriers. 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 [43]) 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 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.

Proof Outline. We observe that the main lemma (Lemma 1.2) is quite similar to classical polynomial approximation results of Razborov [38] and Smolensky [41, 42] (see also [45]). The main difference is that while these results hold for polynomials approximating some function on the whole cube $\{0, 1\}^n$, the lemma deals with polynomial approximations that are more 'local' in that they are restricted on just two layers of the cube. Nevertheless, we can show that the basic proof strategy of Smolensky (or more specifically a variant as in [6, 29]) can be used to prove our lemma as well.

The main point of difference from these standard proofs is the employment of a result from discrete geometry due to Nie and Wang [35], that allows us to bound the size of the *closure*⁵ of a small set of points in the cube. This is a well-studied object in coding theory [48] and combinatorics [14, 26, 35], and turns out to be a crucial ingredient in our proof.

For the application to the coin problem, we show that if a polynomial P solves the coin problem (see Definition 4.1 for the formal definition of this), then it can be used to distinguish

5 The *degree- D closure* $\text{cl}_D(E)$ of a set E is the set of points where any degree- D polynomial Q vanishing throughout E is forced to vanish.

between Hamming weights k and $k + q$ for k and q as in Lemma 1.2. This reduction is done by a simple sampling argument. The degree lower bound in Lemma 1.2 then implies the desired degree lower bound on the degree of P .

In the other applications to probabilistic degree and the robust version of Galvin's problem, the idea is to follow the proofs of the previous best results in this direction and apply the main lemma at suitable points. We defer more details to the actual proofs.

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 $f : \{0, 1\}^n \rightarrow \mathbb{F}$ which are represented by multilinear polynomials. Recall that each such function has a *unique* multilinear polynomial representation. Further, given a (possibly non-multilinear) polynomial $P(x_1, \dots, x_n)$ 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 $b \in \{0, 1\}$. Any polynomial P can be *multilinearized* this way without increasing the degree.

Bernstein's inequality. The following standard deviation bound can be found in, e.g., the book of Dubhashi and Panconesi [17, Theorem 1.2].

LEMMA 2.1 (Bernstein's inequality). *Let X_1, \dots, X_m be independent and identically distributed Bernoulli random variables with mean q . Let $X = \sum_{i=1}^m X_i$. Then for any $\theta > 0$,*

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

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 $f : \{0, 1\}^n \rightarrow \{0, 1\}$ is uniquely specified by a string $\text{Spec } f : [0, n] \rightarrow \{0, 1\}$, which we call the *Spectrum* of f , in the sense that for any $a \in \{0, 1\}^n$, we have

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

Given a $f \in s\mathcal{B}_n$, we define the *period* of f , denoted $\text{per}(f)$, to be the smallest positive integer b such that $\text{Spec } f(i) = \text{Spec } f(i + b)$ for all $i \in [0, n - b]$. We say f is *k-bounded* if $\text{Spec } f$ 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 [33]. Fix any $f \in s\mathcal{B}_n$. Among all symmetric Boolean functions $f' \in s\mathcal{B}_n$ such that $\text{Spec } f'(i) = \text{Spec } f(i)$ for all $i \in [\lceil n/3 \rceil + 1, \lfloor 2n/3 \rfloor]$, we choose a function g such that $\text{per}(g)$ 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 2.2. Let $f \in s\mathcal{B}_n$ and let (g, h) be a standard decomposition of f . Then, $\text{per}(g) \leq \lfloor n/3 \rfloor$ 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 $\text{MOD}_n^{b,i}$ 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 2.3 (Probabilistic polynomial and Probabilistic degree). A *probabilistic polynomial* is a random polynomial \mathbf{P} (with some distribution having finite support) over $\mathbb{F}[x_1, \dots, x_n]$. We say that the degree of \mathbf{P} , denoted $\text{deg}(\mathbf{P})$, is at most d if the probability distribution defining \mathbf{P} is supported on polynomials of degree at most d .

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

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

We define the ε -*error probabilistic degree* of f , denoted $\text{pdeg}_{\varepsilon}^{\mathbb{F}}(f)$, 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 $\text{pdeg}_{\varepsilon}(f)$ instead of $\text{pdeg}_{\varepsilon}^{\mathbb{F}}(f)$.

FACT 2.4. We have the following simple facts about probabilistic degrees of Boolean functions. Let \mathbb{F} be any field.

1. (Error reduction [22]) For any $\delta < \varepsilon \leq 1/3$ and any Boolean function f , if \mathbf{P} is an ε -error probabilistic polynomial for f , then $\mathbf{Q} = M(\mathbf{P}_1, \dots, \mathbf{P}_{\ell})$ 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 $\mathbf{P}_1, \dots, \mathbf{P}_{\ell}$ are independent copies of \mathbf{P} . In particular, we have $\text{pdeg}_{\delta}^{\mathbb{F}}(f) \leq \text{pdeg}_{\varepsilon}^{\mathbb{F}}(f) \cdot O(\log(1/\delta)/\log(1/\varepsilon))$.

2. (Composition) For any Boolean function f on k variables and any Boolean functions g_1, \dots, g_k on a common set of m variables, let h denote the natural composed function $f(g_1, \dots, g_k)$ on m variables. Then, for any $\varepsilon, \delta > 0$, we have $\text{pdeg}_{\mathbb{F}}^{\varepsilon+k\delta}(h) \leq \text{pdeg}_{\mathbb{F}}^{\varepsilon}(f) \cdot \max_{i \in [k]} \text{pdeg}_{\mathbb{F}}^{\delta}(g_i)$.
3. (Sum) Assume that f, g_1, \dots, g_k are all Boolean functions on a common set of m variables such that the functions g_1, \dots, g_k are mutually exclusive and $f = \sum_{i \in [k]} g_i$. Then, for any $\delta > 0$, we have $\text{pdeg}_{\mathbb{F}}^{\delta}(f) \leq \max_{i \in [k]} \text{pdeg}_{\mathbb{F}}^{\delta}(g_i)$.

The first item above is not entirely obvious, as the polynomial \mathbf{P} is not necessarily Boolean-valued at points when $\mathbf{P}(a) \neq f(a)$. Hence, it is not clear that composing with a polynomial that computes the Boolean Majority function achieves error-reduction. The second and third items above are trivial.

Building on work of Alman and Williams [4] and Lu [33], Tripathi, Venkitesh and the author [43] gave upper bounds on the probabilistic degree of any symmetric function. We recall below the statement in the case of fixed positive characteristic.

THEOREM 2.5 (Known upper bounds on probabilistic degree of symmetric functions [43]).

Let \mathbb{F} be a field of constant characteristic $p > 0$ and $n \in \mathbb{N}$ be a growing parameter. Let $f \in \mathcal{SB}_n$ be arbitrary and let (g, h) be a standard decomposition of f . Then we have the following for any $\varepsilon > 0$.

1. If $\text{per}(g) = 1$, then g is a constant and hence $\text{pdeg}_{\mathbb{F}}^{\varepsilon}(g) = 0$.
If $\text{per}(g)$ is a power of p , then g can be exactly represented⁶ as a polynomial of degree at most $\text{per}(g)$, and hence $\text{pdeg}_{\mathbb{F}}^{\varepsilon}(g) \leq \text{per}(g)$,
2. $\text{pdeg}_{\mathbb{F}}^{\varepsilon}(h) = O(\sqrt{B(h) \log(1/\varepsilon)} + \log(1/\varepsilon))$ if $B(h) \geq 1$ and 0 otherwise, and
3. $\text{pdeg}_{\mathbb{F}}^{\varepsilon}(f) = \begin{cases} O(\sqrt{n \log(1/\varepsilon)}) & \text{if } \text{per}(g) > 1 \text{ and not a power of } p, \\ O(\min\{\sqrt{n \log(1/\varepsilon)}, \text{per}(g)\}) & \text{if } \text{per}(g) \text{ a power of } p \text{ and } B(h) = 0, \\ O(\min\{\sqrt{n \log(1/\varepsilon)}, \text{per}(g) + \sqrt{B(h) \log(1/\varepsilon)} + \log(1/\varepsilon)\}) & \text{otherwise.} \end{cases}$

2.3 A string lemma

Given a function $w : I \rightarrow \{0, 1\}$ 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 [9, Theorem 3.1]. For completeness, we give a short proof in Appendix B.

⁶ While this is not part of the formal theorem statement from [43], it follows readily from the proof.

LEMMA 2.6. *Let $w \in \{0, 1\}^+$ be any non-empty string⁷ 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 2.7. *Let $g \in s\mathcal{B}_n$ be arbitrary with $\text{per}(g) = b > 1$. Then for all $i, j \in [0, n - b + 1]$ such that $i \not\equiv j \pmod{b}$, we have $\text{Spec } g|_{[i, i+b-1]} \neq \text{Spec } g|_{[j, j+b-1]}$.*

PROOF. Suppose $\text{Spec } g|_{[i, i+b-1]} = \text{Spec } g|_{[j, j+b-1]}$ for some $i \not\equiv j \pmod{b}$. Assume without loss of generality that $i < j < i + b$. Let $u = \text{Spec } g|_{[i, j-1]}$, $v = \text{Spec } g|_{[j, i+b-1]}$, $w = \text{Spec } g|_{[i+b, j+b-1]}$. Then $u = w$ and the assumption $uv = vw$ implies $uv = vu$. By Lemma 2.6, there exists a string z such that $uv = z^k$ for $k \geq 2$ and therefore $\text{per}(g) < b$. This contradicts our assumption on b . ■

2.4 Lucas's theorem

THEOREM 2.8 (Lucas's theorem). *Let A, B be any non-negative integers and p any prime. Then*

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

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 [33] and Hegedűs [23], showing that Hegedűs's lemma is tight.

COROLLARY 2.9. *Fix any prime p and positive integer n . Assume i is a non-negative integer and q a positive integer such that $i + q \leq n$. Let p^ℓ be the largest power of p dividing q . Then, there is a symmetric multilinear polynomial $Q \in \mathbb{F}_p[x_1, \dots, x_n]$ of degree p^ℓ such that Q vanishes at all points of $\{0, 1\}_i^n$ but at no point of $\{0, 1\}_{i+q}^n$.*

PROOF. Assume $q = p^\ell s$ where s is not divisible by p . Let $a_\ell, b_\ell \in \{0, \dots, p - 1\}$ be the $(\ell + 1)$ th least significant digit of i and $i + q$ respectively in base p . Note that $b_\ell = a_\ell + s_0 \pmod{p}$ where s_0 is the least significant digit of s in base p (s_0 is non-zero as s is not divisible by p).

Define the polynomial

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

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

$$Q(c) = \binom{w}{p^\ell} - a_\ell$$

where the right hand side is interpreted modulo p . Lucas's theorem then easily implies that $Q(c) = 0$ if $w = i$ and s_0 if $w = i + q$. ■

⁷ Recall that, for any alphabet Σ , the notation Σ^+ denotes the set of non-empty strings over this alphabet.

3. The Main Lemma

In this section, we prove the main lemma, which is a robust version of Lemma 1.1.

LEMMA 3.1 (A Robust Version of Hegedús'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 $P \in \mathbb{F}[x_1, \dots, x_n]$ is a polynomial such that for some $K \in \{k + q, k - q\}$,*

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

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

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 $\{0, 1\}_k^n$ and non-zero on most of $\{0, 1\}_K^n$. 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 3.2 (A special case of Lemma 3.1). *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 $P \in \mathbb{F}[x_1, \dots, x_n]$ be any polynomial such that*

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

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

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

REMARK 3.3. 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., [29, Lemma 3.3] for a proof.

FACT 3.4. Let $R \in \mathbb{F}[x_1, \dots, x_n]$ 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 3.5. 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 $\text{cl}_D(E)$ as follows.

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

Note that $\text{cl}_D(E) \supseteq E$ but could be much bigger than E . The following result of Nie and Wang [35] gives a bound on $|\text{cl}_D(E)|$ in terms of $|E|$. (This particular form is noted and essentially proved in [35], and is explicitly stated and proved in [29, Theorem A.1] for all fields.)

THEOREM 3.6. For any $E \subseteq \{0, 1\}^n$ and any $D \leq n$, we have

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

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

REMARK 3.7. It should be noted that the above lemma generalizes the standard linear-algebraic fact that for any E such that $|E| < N_D$, there is a non-zero multilinear polynomial of degree D that vanishes on E . Or equivalently,

$$|E| < N_D \implies \text{cl}_D(E) < 2^n.$$

The inequality stated in the lemma is tight for certain sets E of size N_D (a good example of such a set is any Hamming ball of radius D). However, when $|E|$ is much smaller than N_D , the parameters can be tightened. A tight form of this lemma, that gives the best possible parameters

depending on $|E|$, was proved in earlier work of Keevash and Sudakov [26] (see also the works of Clements and Lindström [14], Wei [48], Heijnen and Pellikaan [24], and Beelen and Dutta [7] that prove similar results). However, we don't need this general form of the lemma here.

We now begin the proof of the Lemma 3.2.

PROOF OF LEMMA 3.2. Assume that P is as given. Let $m = \lfloor n/2 \rfloor$.

Let E_0, E_1 be defined as follows. (Here, the notation “ E ” stands for “error sets”.)

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

$$E_1 = \{a \in \{0, 1\}_m^n \mid P(a) = 0\}$$

We show that there are polynomials $Q_1, Q_2 \in \mathbb{F}[x_1, \dots, x_n]$ 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 $a \in \{0, 1\}_m^n \setminus E_1$.

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 $a \in \{0, 1\}_{m-t}^n \setminus E_0$, 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 $a' \in \{0, 1\}_m^n \setminus E_1$ 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 3.4, 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 3.8. *There is a Q_1 of degree at most t satisfying property (Q1.1).*

CLAIM 3.9. *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 3.8. This follows immediately from the upper bound for periodic functions in Theorem 2.5. 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. \blacklozenge

Proof of Claim 3.9. Let D denote $m - t - t_1$. Let $E = E_0 \cup \bigcup_{j < m-t: j \equiv m \pmod{t}} \{0, 1\}_j^n$. 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 $E'_1 := \{0, 1\}_m^n \setminus E_1$. Note that this is equivalent to saying that $\text{cl}_D(E) \not\supseteq E'_1$. To show this, it suffices to show that

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

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

To do this, we use Theorem 3.6. Note that we have

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

where the third inequality is a consequence of Lemma 3.5 (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 3.6 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 3.5 (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 3.6, we thus obtain

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

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 polynomial $P \in \mathbb{F}[x_1, \dots, x_n]$ and any $m \in [0, n]$, let $\text{NZ}_m(P)$ denote the set of points of $\{0, 1\}_m^n$ where P does not vanish. Let $\psi_m(P)$ denote $|\text{NZ}_m(P)| / \binom{n}{m}$.

LEMMA 3.10. *For any $Q \in \mathbb{F}[x_1, \dots, x_n]$ and any $r \geq 1$, there is a probabilistic polynomial $Q^{(r)}$ of degree at most $r \cdot \deg(Q)$ such that for all $m \in [0, n]$, $E_{Q^{(r)}}[\psi_m(Q^{(r)})] = \psi_m(Q)^r$.*

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

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

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

as a^π is uniformly distributed over $\{0, 1\}_m^n$.

Choose π_1, \dots, π_r i.u.a.r. from S_n , and define $Q^{(r)} = \prod_{i=1}^r Q^{\pi_i}$. For any $a \in \{0, 1\}_m^n$

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

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

$$E_{Q^{(r)}} [\psi_m(Q^{(r)})] = \Pr_{Q^{(r)}, \mathbf{a} \sim \{0, 1\}_m^n} [Q^{(r)}(\mathbf{a}) \neq 0] = \psi_m(Q)^r.$$

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

PROOF OF LEMMA 3.1. W.l.o.g. we assume that $k \leq n/2$. (To prove the lemma for $k > n/2$, consider the polynomial $Q(x) = P(1 - x_1, \dots, 1 - x_n)$ 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 \mathbf{S} be a uniformly random subset of $[n]$ of size $r + s$ and \mathbf{y} a uniformly random point in $\{0, 1\}_{r+s}^{r+s}$. We set

$P_{\mathbf{S},\mathbf{y}}(x_i : i \notin S)$ to be the probabilistic polynomial obtained by setting all the variables indexed by \mathbf{S} according to \mathbf{y} . Note that we have

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

By Markov's inequality, we have

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

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

We will now use Lemma 3.2 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 that follow from the lemma statement and the bounds above, P' is a polynomial in n' variables satisfying

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

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,

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

Applying Lemma 3.2 to P' (see also Remark 3.3), 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

$$\psi_{k'}(P') \leq \frac{2\varepsilon_0}{\varepsilon_1} \leq 2\varepsilon_0^{0.99} \leq \frac{1}{400}, \quad (5a)$$

$$\psi_{K'}(P') \geq \frac{\varepsilon_1}{2} \geq \frac{\varepsilon_0^{0.01}}{2} \geq \frac{1}{2^{100/99}} \cdot \psi_{k'}(P')^{1/99} \geq \psi_{k'}(P')^{1/7}, \quad (5b)$$

$$\psi_{K'}(P') \geq \frac{\varepsilon_1}{2} \geq e^{-5}. \quad (5c)$$

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 3.10, we know that there is a probabilistic polynomial $\mathbf{P}'^{(r)}$ of degree at most $r \cdot \deg(P')$ such that for each $m \in \{k', K'\}$, we have $\mathbb{E}_{\mathbf{P}'^{(r)}}[\psi_m(\mathbf{P}'^{(r)})] = \psi_m(P')^r$.

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 $x_1, \dots, x_{n'}$. Let \mathbf{T} be a uniformly random subset of $[n']$ of size $n' - n''$ and let \mathbf{z} be a uniformly random point in $\{0, 1\}_{(n'-n'')/2}$. Define the probabilistic polynomial $\mathbf{P}'_{\mathbf{T}, \mathbf{z}}^{(r)}$ obtained by setting the variables indexed by \mathbf{T} according to \mathbf{z} in the probabilistic polynomial $\mathbf{P}'^{(r)}$. Let $K'' := n''/2$ and $k'' := k' - (n' - n'')/2$. As above, we have

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

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

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

where the first inequality uses (5).

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

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

Applying Lemma 3.2 to P'' with error parameter $\varepsilon = \frac{2\varepsilon'_0}{\varepsilon'_1}$ yields $\deg(P'') \geq q/25$. As $\deg(P'') \leq r \cdot \deg(P')$, 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 3.1)

In this section, we discuss the near-optimality of Lemma 3.1 w.r.t. to the various parameters. Fix n, k, q, α, δ and \mathbb{F} as in the statement of Lemma 3.1. Assume that $K = k + q$ (the case when $K = k - q$ is similar) and that $k \leq n/2$. 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 2.9, it follows that there is a degree- q polynomial that vanishes at all points of weight k but no points of weight K .

So, the statement of Lemma 3.1 proves a lower bound on the degree that nearly (up to constant factors) matches this trivial upper bound, under the weaker assumption that the polynomial is forced to be zero only on most (say a $1 - \varepsilon$ fraction) of $\{0, 1\}_k^n$ and non-zero on most (say a $1 - \varepsilon$ fraction) of $\{0, 1\}_K^n$. (Lemma 3.1 is a stronger statement, but we will show that even this weaker statement is tight.)

In this section, we show that the value of ε cannot be increased beyond $\varepsilon = \exp(-O(\delta^2 n/\alpha))$, if we want to prove a lower bound of $\Omega(q)$ on the degree. More precisely, we show the following.

THEOREM 3.11. *Assume that $\varepsilon = \exp(-o(\delta^2 n/\alpha))$. Then, there is a polynomial P of degree $o(q)$ such that*

$$\Pr_{\mathbf{a} \sim \{0,1\}_k^n} [P(\mathbf{a}) \neq 0] \leq \varepsilon$$

$$\Pr_{\mathbf{a} \sim \{0,1\}_K^n} [P(\mathbf{a}) \neq 0] \geq 1 - \varepsilon.$$

PROOF. To prove this theorem, 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 [4].⁸

LEMMA 3.12. *Let n be arbitrary and $I \subseteq [0, n]$ be any interval of integers. Given any $f : I \rightarrow \{0, 1\}$, there is a multilinear polynomial $Q \in \mathbb{Z}[x_1, \dots, x_n]$ of degree at most $|I| - 1$ such that $Q(\mathbf{a}) = f(|\mathbf{a}|)$ for each $\mathbf{a} \in \bigcup_{i \in I} \{0, 1\}_i^n$.*

Fix any positive integer m . By Lemma 3.12, it follows that there is a multilinear polynomial $Q \in \mathbb{Z}[y_1, \dots, y_m]$ of degree $O(\delta m)$ such that $Q(\mathbf{b}) = 0$ for each $\mathbf{b} \in \{0, 1\}^m$ such that $|\mathbf{b}| \in ((\alpha - \delta/2)m, (\alpha + \delta/2)m)$ and $Q(\mathbf{b}) = 1$ for each $\mathbf{b} \in \{0, 1\}^m$ such that $|\mathbf{b}| \in ((\alpha + \delta/2)m, (\alpha + 3\delta/2)m)$. Reducing the coefficients modulo p , we obtain a polynomial $\tilde{Q} \in \mathbb{F}[y_1, \dots, y_m]$ with the same property. Fix this \tilde{Q} .

Consider the probabilistic polynomial $\mathbf{P}(x_1, \dots, x_n)$ defined as follows. Choose $\mathbf{i}_1, \dots, \mathbf{i}_m$ 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 $\mathbf{P}(x_1, \dots, x_n)$ to be the polynomial $\tilde{Q}(x_{\mathbf{i}_1}, \dots, x_{\mathbf{i}_m})$. Note that

⁸ This lemma has a trivial proof via univariate polynomial interpolation if we only want the polynomial Q to have rational coefficients. However, here it is important that Q has integer coefficients.

$$\deg(\mathbf{P}) \leq \deg(Q) = O(\delta m) = O((\alpha/\delta) \log(1/\varepsilon)) = o(\delta n) = o(q)$$

where the second-last equality uses our assumption that $\varepsilon = \exp(-o(\delta^2 n/\alpha))$.

Let $a \in \{0, 1\}_k^n$ be arbitrary. We analyze the random variable $\mathbf{P}(a)$. Note that as long as the Hamming weight of $\mathbf{b} = (a_{i_1}, \dots, a_{i_m})$ is in the interval $((\alpha - \delta/2)m, (\alpha + \delta/2)m)$, we have $\mathbf{P}(a) = 0$. As each co-ordinate of \mathbf{b} is 1 with probability $k/n = \alpha \in [0, 1/2]$, Bernstein's inequality (Lemma 2.1) yields

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

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 $\Pr_{\mathbf{P}} [\mathbf{P}(a) \neq 1] < \varepsilon/2$ and hence, in particular, $\Pr_{\mathbf{P}} [\mathbf{P}(a) \neq 0] > 1 - (\varepsilon/2)$, 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 $\deg(\mathbf{P})$ such that

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

Thus, we have a polynomial P as claimed in Theorem 3.11. ■

3.4 An extension to the case when q is not a power of p

An anonymous reviewer suggested the following extension of the main lemma (Lemma 3.1). We prove this by a simple reduction to the main lemma. (This leads to a worsening in the constants involved.)

LEMMA 3.13 (An extension to the case when q is not a power of p). *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 $200q < k < n - 200q$. Let q' be the largest power of p that divides q and assume $q = q's$. Define $\alpha = \min\{k/n, 1 - (k/n)\}$ and $\delta = q/n$. Assume that $Q \in \mathbb{F}[x_1, \dots, x_n]$ is a polynomial such that for some $K \in \{k + q, k - q\}$,*

$$\Pr_{\mathbf{a} \sim \{0,1\}_k^n} [Q(\mathbf{a}) \neq 0] \leq \min\{e^{-1000\delta^2 n/s\alpha}, 1/2000\} \tag{6a}$$

$$\Pr_{\mathbf{a} \sim \{0,1\}_K^n} [Q(\mathbf{a}) \neq 0] \geq e^{-\delta^2 n/1000s\alpha}. \tag{6b}$$

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

REMARK 3.14. The 'non-robust' version of this lemma (when Q vanishes everywhere on $\{0, 1\}_k^n$ but not on some point in $\{0, 1\}_K^n$) yields a degree lower bound of q' , and can be proved using similar techniques to those used in proving Hegedűs's lemma. A proof can be found in [44].

REMARK 3.15. As in the case of the main lemma, the degree lower bound obtained above is tight, using the same reasoning as in Section 3.3.

PROOF. W.l.o.g. assume $K = k + q$.

Let $k' = \lfloor k/s \rfloor$ and $n' = \lfloor n/s \rfloor - 1$. Our aim will be to show using the polynomial Q that there is a polynomial P on n' variables that distinguishes between Hamming weights k' and $K' := k' + q'$. We will then appeal to Lemma 3.1 to get the degree lower bound.

It is easy to check that $100q' < k' < n' - 100q'$ as

$$\begin{aligned} 100q's &= 100q < k - q < (k' + 1)s - q \leq k's \\ k's &\leq k < n - 102q < (n' + 2)s - 102q \leq n's - 100q = (n' - 100q')s \end{aligned}$$

where we used the hypotheses that $200q < k < n - 200q$.

We construct the polynomial P as follows. Assume that $k = k's + r_1$ and $n = n's + s + r_2$ for $r_1, r_2 \in \{0, \dots, s-1\}$. On an input $x \in \{0, 1\}^{n'}$, we consider the *random* input $\mathbf{y} \in \{0, 1\}^n$ defined as follows.

- Each co-ordinate of x is repeated s times to get an $X \in \{0, 1\}^{sn'}$.
- We concatenate X with the string $1^{r_1}0^{s+r_2-r_1}$ to get a string $Y \in \{0, 1\}^n$.
- A uniformly random permutation π is applied to the n coordinates of Y to get \mathbf{y} .

Finally, we define the probabilistic polynomial $\mathbf{P}(x) := Q(\mathbf{y})$. For a fixed permutation π , each coordinate of \mathbf{y} is a polynomial of degree at most 1 in the variables $x_1, \dots, x_{n'}$, and hence, $\deg(\mathbf{P}) \leq \deg(Q)$. We will show that there is some polynomial P in the support of \mathbf{P} that has the desired properties.

Let ε_0 and ε_1 denote the right hand sides of inequalities (6a) and (6b) respectively. Observe that when $x \in \{0, 1\}_w^{n'}$, then the random $\mathbf{y} \in \{0, 1\}^n$ is uniformly distributed over $\{0, 1\}_{ws+r_1}^n$. In particular, setting $w = k'$ and $k' + q'$, we get the following.

$$\begin{aligned} \Pr_{\mathbf{a} \sim \{0,1\}_{k'}^{n'}, \mathbf{P}} [\mathbf{P}(\mathbf{a}) \neq 0] &= \Pr_{\mathbf{b} \sim \{0,1\}_k^n} [Q(\mathbf{b}) \neq 0] \leq \varepsilon_0 \\ \Pr_{\mathbf{a} \sim \{0,1\}_{k'+q'}^{n'}, \mathbf{P}} [\mathbf{P}(\mathbf{a}) \neq 0] &= \Pr_{\mathbf{b} \sim \{0,1\}_{k+q}^n} [Q(\mathbf{b}) \neq 0] \geq \varepsilon_1. \end{aligned}$$

To find a suitable fixing of \mathbf{P} , we consider two cases.

- **Case 1:** $e^{-\delta^2 n/250s\alpha} \geq 1/2$: In this case, define two events \mathcal{E}_0 and \mathcal{E}_1 (depending only on the probabilistic polynomial \mathbf{P}) as follows.

$$\mathcal{E}_0 := \Pr_{\mathbf{a} \sim \{0,1\}_{k'}^{n'}} [\mathbf{P}(\mathbf{a}) \neq 0] \geq 2\varepsilon_0, \text{ and } \mathcal{E}_1 := \Pr_{\mathbf{a} \sim \{0,1\}_{k'+q'}^{n'}} [\mathbf{P}(\mathbf{a}) = 0] \geq 2.5\zeta_1$$

where $\zeta_1 := 1 - \varepsilon_1$. Note that $\varepsilon_1 = e^{-\delta^2 n/1000s\alpha} \geq 2^{-0.25} > 0.8$ and hence $\zeta_1 < 0.2$.

By Markov's inequality, with positive probability over the choice of \mathbf{P} , neither of the above events occurs. Fix such a polynomial P . Then, we have

$$\begin{aligned} \Pr_{\mathbf{a} \sim \{0,1\}_{k'}^{n'}} [P(\mathbf{a}) \neq 0] &< 2\varepsilon_0 = \min\{2e^{-1000\delta^2 n/s\alpha}, 2/2000\} \\ &\leq \min\{e^{-200\delta^2 n/s\alpha}, 1/1000\}, \end{aligned} \quad (7)$$

where we used the simple fact that for any non-negative real number γ , we have the inequality $\min\{2\gamma, 1/1000\} \leq \min\{\gamma^{0.2}, 1/1000\}$. We also have

$$\Pr_{\mathbf{a} \sim \{0,1\}_{k'+q'}^{n'}} [P(\mathbf{a}) \neq 0] > 1 - 2.5\zeta_1 \geq (1 - \zeta_1)^5 = \varepsilon_1^5 \geq e^{-\delta^2 n/200s\alpha}, \quad (8)$$

where the second inequality uses the fact that $\zeta_1 \leq 0.2$ for any $\gamma \in [0, 1]$, we have⁹ $(1 - \gamma)^5 \leq 1 - 5\gamma + 10\gamma^2$.

— **Case 2:** $e^{-\delta^2 n/250s\alpha} < 1/2$: In this case, we proceed analogously, but define the 'bad' events as follows.

$$\mathcal{E}'_0 := \Pr_{\mathbf{a} \sim \{0,1\}_{k'}^{n'}} [P(\mathbf{a}) \neq 0] \geq \frac{2\varepsilon_0}{\varepsilon_1}, \text{ and } \mathcal{E}'_1 := \Pr_{\mathbf{a} \sim \{0,1\}_{k'+q'}^{n'}} [P(\mathbf{a}) \neq 0] < \frac{\varepsilon_1}{2}.$$

By Markov's inequality, there is again a fixing P of \mathbf{P} such that neither of the above two events occurs. For such a polynomial P , we have

$$\Pr_{\mathbf{a} \sim \{0,1\}_{k'+q'}^{n'}} [P(\mathbf{a}) \neq 0] \geq \frac{\varepsilon_1}{2} \geq e^{-\delta^2 n/1000s\alpha} \cdot e^{-\delta^2 n/250s\alpha} = e^{-\delta^2 n/200s\alpha}, \quad (9)$$

where the second inequality used our assumption that $e^{-\delta^2 n/250s\alpha} < 1/2$. We also have

$$\begin{aligned} \Pr_{\mathbf{a} \sim \{0,1\}_{k'}^{n'}} [P(\mathbf{a}) \neq 0] &< \frac{\varepsilon_0}{\varepsilon_1/2} \leq e^{\delta^2 n/200s\alpha} \cdot \varepsilon_0 \leq e^{-999\delta^2 n/s\alpha} \\ &\leq e^{200\delta^2 n/s\alpha} \leq \min\{e^{-200\delta^2 n/s\alpha}, 1/1000\}, \end{aligned} \quad (10)$$

where the second inequality used (9) above and the third and last inequalities use the fact that $e^{-\delta^2 n/250s\alpha} < 1/2$ to deduce that $e^{-1000\delta^2 n/s\alpha} \leq 1/2000$ and $e^{-200\delta^2 n/s\alpha} < 1/1000$ respectively.

Putting (7), (8), (10) and (9) together gives us that in both cases we have

$$\Pr_{\mathbf{a} \sim \{0,1\}_{k'}^{n'}} [P(\mathbf{a}) \neq 0] \leq \min\{e^{-200\delta^2 n/s\alpha}, 1/1000\} \quad (11a)$$

$$\Pr_{\mathbf{a} \sim \{0,1\}_{k'+q'}^{n'}} [P(\mathbf{a}) \neq 0] \geq e^{-\delta^2 n/200s\alpha}. \quad (11b)$$

⁹ This is a special case of the Boole-Bonferroni inequalities, which are closely related to the Principle of Inclusion-Exclusion.

To apply Lemma 3.1 to P , we need to relate the above bounds to quantities defined in terms of $\delta' := q'/n'$ and $\alpha' := k'/n'$. We claim that

$$\frac{\delta^2 n}{s\alpha} \leq \frac{(\delta')^2 n'}{\alpha'} \leq \frac{2\delta^2 n}{s\alpha}. \quad (12)$$

Assuming these inequalities, we observe that P satisfies the hypotheses of Lemma 3.1. Applying this lemma gives us

$$\deg(Q) \geq \deg(\mathbf{P}) \geq \deg(P) = \Omega(q'),$$

finishing the proof of Lemma 3.13.

It remains to prove (12), which is a simple calculation.

$$\begin{aligned} \frac{(\delta')^2 n'}{\alpha'} &= \frac{(\delta' n')^2}{\alpha' n'} = \frac{(q')^2}{k'} = \frac{(q's)^2}{k's^2} \geq \frac{q^2}{sk} = \frac{\delta^2 n}{s\alpha}, \\ \frac{(\delta')^2 n'}{\alpha'} &= \frac{(q')^2}{k'} = \frac{(q's)^2}{k's^2} \leq \frac{q^2}{s(k-s)} = \frac{\delta^2 n}{s\alpha} \cdot \left(1 - \frac{s}{k}\right)^{-1} \leq \frac{2\delta^2 n}{s\alpha} \end{aligned}$$

where the final inequality uses the fact that $\frac{s}{k} \leq \frac{q}{k} \leq 0.01$. ■

4. Applications

4.1 Tight Degree Lower Bounds for the Coin Problem

We start with a definition.

DEFINITION 4.1 (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 $g : \{0, 1\}^n \rightarrow \{0, 1\}$, we say that g solves the δ -coin problem with error ε if

$$\Pr_{\mathbf{x} \sim \mu_{(1/2)-\delta}^n} [g(\mathbf{x}) = 1] \leq \varepsilon \text{ and } \Pr_{\mathbf{x} \sim \mu_{1/2}^n} [g(\mathbf{x}) = 1] \geq 1 - \varepsilon. \quad (13)$$

(This definition is sometimes [31] 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 $P \in \mathbb{F}[x_1, \dots, x_n]$ that solves the δ -coin problem with error ε .

By Lemma 3.12, for any $n \geq 1$, there is a polynomial $P \in \mathbb{F}[x_1, \dots, x_n]$ 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 2.1 (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 [31], 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 [13] (see [2]). Both proofs relied on slight strengthenings of Smolensky's [41] 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 3.1), or even its simpler version Lemma 3.2, shows that this is indeed true.

THEOREM 4.2 (Tight Degree Lower Bound for the δ -coin problem for all errors). *Assume \mathbb{F} has characteristic p and δ, ε are parameters going to 0. Let $N \geq 1$ be any positive integer. Any polynomial $P \in \mathbb{F}[x_1, \dots, x_N]$ 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 [31]).

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 $Q \in \mathbb{F}[y_1, \dots, y_n]$ obtained from P by randomly replacing each variable of P by a uniformly random variable among y_1, \dots, y_n . For any $a \in \{0, 1\}_{n/2}^n$, we have

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

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

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

Hence, by Lemma 3.2, 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 $\mathbf{a}, \mathbf{b} \in \{0, 1\}^n$ are sampled independently from the distributions $\mu_{1/2-\delta'}^1$ and $\mu_{1/2-(\alpha/2)}^1$ respectively, then their parity $\mathbf{a} \oplus \mathbf{b}$ has the distribution $\mu_{1/2-\delta}^1$. Now, if we define the probabilistic polynomial $\mathbf{R}(x_1, \dots, x_n)$ by

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

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

Repeating the above argument with \mathbf{R} instead of P yields that $\deg(\mathbf{R}) = \Omega(\frac{1}{\delta'} \log(1/\varepsilon)) = \Omega(\frac{1}{\delta} \log(1/\varepsilon))$. 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 4.3 (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*

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

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 2.5.

4.2.1 Some Preliminaries

DEFINITION 4.4 (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.¹⁰

We will use the following obvious fact freely.

OBSERVATION 4.5. *If g is a restriction of f , then for any $\delta > 0$, $\text{pdeg}_{\delta}(g) \leq \text{pdeg}_{\delta}(f)$.*

In earlier work with Tripathi and Venkitesh [43], we showed the following near-optimal lower bound on the probabilistic degrees of Threshold functions.

LEMMA 4.6 (Lemma 27 in [43]). *Assume $t \geq 1$. For any $\varepsilon \in [2^{-n}, 1/3]$,*

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

(The corresponding lemma in [43] is only stated for $t \leq n/2$. However, as $\text{Thr}_n^{n+1-t}(x) = 1 - \text{Thr}_n^t(1 - x_1, \dots, 1 - x_n)$, the above lower bound holds for $t > n/2$ also.)

¹⁰ Note that exactly which inputs are set to 0 or 1 is not important, since we are dealing with *symmetric* Boolean functions.

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

LEMMA 4.7 (Smolensky's lower bound for Majority function [45, 42]). *For any field \mathbb{F} , any $\varepsilon \in (1/2^n, 1/5)$, we have*

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

LEMMA 4.8 (Smolensky's lower bound for MOD functions [41]). *For $2 \leq b \leq n/2$, any \mathbb{F} such that $\text{char}(\mathbb{F})$ 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*

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

We now show how to use our robust version of Hegedús's lemma to prove Theorem 4.3. In fact, Lemma 3.2 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 3.2.

COROLLARY 4.9. *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 $h \in \mathcal{SB}_n$ be any function such that $\text{Spec } h(\lfloor n/2 \rfloor) \neq \text{Spec } h(\lfloor n/2 \rfloor - t)$. Then, $\text{pdeg}_{\varepsilon}(h) = \Omega(t)$.*

PROOF. By error reduction for probabilistic polynomials (Fact 2.4 item 1), it suffices to prove an $\Omega(t)$ lower bound on $\text{pdeg}_{\varepsilon/2}(h)$.

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

$$\begin{aligned} \Pr_{\mathbf{P}, \mathbf{a} \sim \{0,1\}_{\lfloor n/2 \rfloor}} [\mathbf{P}(\mathbf{a}) \neq 1] &\leq \varepsilon/2 \\ \Pr_{\mathbf{P}, \mathbf{b} \sim \{0,1\}_{\lfloor n/2 \rfloor - t}} [\mathbf{P}(\mathbf{b}) \neq 0] &\leq \varepsilon/2 \end{aligned}$$

Thus, we have

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

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

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

Applying Lemma 3.2 to P yields

$$\deg(\mathbf{P}) \geq \deg(P) = \Omega(t). \quad \blacksquare$$

To illustrate the usefulness of Corollary 4.9, we prove optimal lower bounds on the probabilistic degrees for two interesting classes of functions (both of which will be subsumed by Theorem 4.3).

COROLLARY 4.10. *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 $\text{EThr}_n^{\lfloor n/2 \rfloor}$ and MOD_n^q are $\Omega(\sqrt{n})$.*

Known lower bounds (Lemmas 4.7 and 4.8) can be used to prove similar lower bounds to the one given above, but with additional log-factor losses (see Lemma 4.8, which requires the error to be subconstant, and [43]). However, we do not know how to prove the above tight (up to constants) lower bound without appealing to Lemma 3.2. In particular, we do not know how to prove the above in characteristic 0.

PROOF. We use Corollary 4.9. We will use $\text{EThr}_n^{\lfloor n/2 \rfloor}$ 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 4.9 then implies the required lower bound.

For $h = \text{EThr}_n^{\lfloor n/2 \rfloor}$, note that $\text{Spec } h(\lfloor n/2 \rfloor) \neq \text{Spec } h(\lfloor n/2 \rfloor - t)$ 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 4.9 that $\text{pdeg}_{\varepsilon_0}(h) = \Omega(t) = \Omega(\sqrt{n})$. By error-reduction for probabilistic polynomials (Fact 2.4 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 $h = \text{EThr}_n^{\lfloor n/2 \rfloor}$.

For $h = \text{MOD}_n^q$, 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

$$h'(x) = h(x1^r 0^{q-r}).$$

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 $\text{Spec } h'(\lfloor (n-q)/2 \rfloor) = \text{Spec } h(r + \lfloor (n-q)/2 \rfloor) = 1$ 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 $\text{Spec } h'(\lfloor (n-q)/2 \rfloor - t) = 0$. Thus, by Corollary 4.9, we get $\text{pdeg}_{\varepsilon_0}(h') = \Omega(t) = \Omega(\sqrt{n})$. ■

4.2.3 Proof of Theorem 4.3

The proof of this theorem closely follows our probabilistic degree lower bounds in [43] 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 $\text{pdeg}_{\varepsilon}(f)$ as long as $\text{per}(g)$ is large.

LEMMA 4.11. *Fix any $\varepsilon \in [2^{-n}, 1/3]$. Assume that f is such that $\text{per}(g) > \sqrt{n \log(1/\varepsilon)}$. Then*

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

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 $\text{per}(g) > \sqrt{n \log(1/\varepsilon)} \geq m$, 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 $\text{Spec } f(r) \neq \text{Spec } f(r + m)$.

Let $k = \lfloor n/2 \rfloor$. 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

$$F(x) = f(x1^a 0^b)$$

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 $\text{Spec } F(\lfloor k/2 \rfloor) = \text{Spec } f(\lfloor k/2 \rfloor + a) = \text{Spec } f(r + m)$ and similarly that $\text{Spec } F(\lfloor k/2 \rfloor - m) = \text{Spec } f(r)$. We thus obtain $\text{Spec } F(\lfloor k/2 \rfloor) \neq \text{Spec } F(\lfloor k/2 \rfloor - m)$.

Note that by the bounds on ε assumed above

$$m \geq \frac{1}{4p} \sqrt{n \log(1/\varepsilon)} \geq 20\sqrt{n}. \quad (14)$$

Using Corollary 4.9, we hence get

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

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 4.9.)

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

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

which implies the desired lower bound.¹¹

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 (14). However, for $F' \in s\mathcal{B}_{k'}$ defined by

$$F'(x) = F(x0^t 1^t)$$

for suitably chosen $t \leq k/2$, we can ensure that $m \in [10\sqrt{k'}, 20\sqrt{k'}]$. Note that $\text{Spec } F'(\lfloor k'/2 \rfloor) = \text{Spec } F(\lfloor k/2 \rfloor)$ and $\text{Spec } F'(\lfloor k'/2 \rfloor - m) = \text{Spec } F(\lfloor k/2 \rfloor - m)$. Hence, for $\varepsilon_1 = e^{-10000}$, Corollary 4.9 implies

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

By error reduction (Fact 2.4 item 1), the same lower bound holds for $\text{pdeg}_\varepsilon(f)$ as well. ■

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

¹¹ Note that we assume that the characteristic is a fixed positive constant and hence the $\Omega(\cdot)$ can hide constants depending on p .

LEMMA 4.12. For any $\varepsilon \in [2^{-n}, 1/3]$,

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

PROOF. By Fact 2.4 item 1 (error reduction), we know that $\text{pdeg}_\varepsilon(g) = \Theta(\text{pdeg}_\delta(g))$ 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 $b := \text{per}(g)$. If $\text{per}(g) > \sqrt{n \log(1/\varepsilon)}$, we are done by Lemma 4.11. 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 $b = \text{per}(g)$, there must exist $r \in [0, n-b_1]$ such that

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

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 $\text{Spec } g(r) \neq \text{Spec } g(r+b_1)$, we also have $\text{Spec } g(r) \neq \text{Spec } g(r+b_1+k \cdot b)$ 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 $\text{Spec } g(r) \neq \text{Spec } g(r+n_1)$. As $n_1 \leq n/100$, we have

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

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

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

$$F(x) = f(x1^a 0^{n-m-a})$$

where a is chosen so that $\text{Spec } F(\lfloor m/2 \rfloor) = \text{Spec } f(r+n_1) = 1$. This also has the consequence that $\text{Spec } F(\lfloor m/2 \rfloor - n_1) = \text{Spec } f(r) = 0$. By Corollary 4.9, we get $\text{pdeg}_\varepsilon(F) = \Omega(n_1) = \Omega(\sqrt{n \log(1/\varepsilon)})$, 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 $\text{pdeg}_\varepsilon(f)$.

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 $\text{per}(G) \geq b$. To see this, note that by Corollary 2.7, we have

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

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 $\text{Spec } f$ also. Further, as $\text{Spec } F(\lfloor m/2 \rfloor + j) = \text{Spec } f(\lfloor n/2 \rfloor + j)$ for $j \leq m/2$, we also get

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

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 $\text{per}(G) > \sqrt{m \log(1/\delta)}$. Lemma 4.11 above and (P4) now imply that $\text{pdeg}_\delta(F) = \Omega(\min\{b, \sqrt{n \log(1/\varepsilon)}\})$. However, as $\delta \geq \varepsilon$ (by (P2)) and F is a restriction of f , the same lower bound holds for $\text{pdeg}_\varepsilon(f)$ 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 $\text{pdeg}_\varepsilon(h)$.

LEMMA 4.13. *Assume $B(h) \geq 1$. Then, $\varepsilon \in [2^{-n}, 1/3]$,*

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

PROOF. Similar to the proof of Lemma 4.12, we may assume without loss of generality that $\varepsilon \in [2^{-n/10000}, e^{-10000p^2}]$.

Let $B(h) = b$. Recall (Observation 2.2) that $B(h) \leq \lceil n/3 \rceil$. Further, by definition of $B(h)$, we have either $\text{Spec } h(b - 1) = 1$ or $\text{Spec } h(n - b + 1) = 1$. We assume that $\text{Spec } h(n - b + 1) = 1$ (the other case is similar).

The lemma is equivalent to showing that $\text{pdeg}_\varepsilon(h) = \Omega(\max\{\sqrt{B(h)\log(1/\varepsilon)}, \log(1/\varepsilon)\})$. 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.

$$H(x) = h(x1^{n-b+1-m}0^{b-1}).$$

Note that as $\text{Spec } h$ is the constant 0 function on the interval $[b, n-b]$, the function H is computing the AND function on m inputs. By Lemma 4.6, we immediately have $\text{pdeg}_\varepsilon(h) \geq \text{pdeg}_\varepsilon(H) = \Omega(\log(1/\varepsilon))$ proving the lemma in this case.

Now assume that $\varepsilon > 2^{-b/1000}$. In this case, we need to show that $\text{pdeg}_\varepsilon(h)$ is lower bounded by $\Omega(\sqrt{b\log(1/\varepsilon)})$. To prove this, consider the restriction $H \in s\mathcal{B}_{2b-2}$ defined by $H(x) = h(x1^{n-2b+2})$. Since $\text{Spec } h$ is the constant 0 function on the interval $[b, n-b]$ and $\text{Spec } h(n-b+1) = 1$, it follows that the periodic part of H has period $\Omega(b)$. It then follows from Lemma 4.11 that $\text{pdeg}_\varepsilon(h) = \Omega(\sqrt{b\log(1/\varepsilon)})$. This concludes the proof of the lemma. ■

Now, we are ready to prove Theorem 4.3.

PROOF OF THEOREM 4.3. By Lemma 4.12, we already have the desired lower bound on $\text{pdeg}_\varepsilon(f)$ in any of the following scenarios.

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

So from now, we assume that $\text{per}(g)$ is a power of p upper-bounded by $\sqrt{n\log(1/\varepsilon)}$ and that $B(h) \geq 1$. In this case, Lemma 4.12 shows that $\text{pdeg}(f) = \Omega(\text{per}(g))$. On the other hand, since $B(h) \leq n$ and $\varepsilon \geq 2^{-n}$, the lower bound we need to show is $\Omega(\text{per}(g) + \sqrt{B(h)\log(1/\varepsilon)} + \log(1/\varepsilon))$. By Lemma 4.13, it suffices to show a lower bound of $\Omega(\text{per}(g) + \text{pdeg}_\varepsilon(h))$.

The analysis splits into two simple cases.

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

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

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

where the first inequality is a consequence of Fact 2.4 item 2 and the second follows from error-reduction and Theorem 2.5. The above yields

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

This finishes the proof. ■

4.3 A Robust Version of Galvin's Problem

We recall here a combinatorial theorem of Hegedűs [23] 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 $u, v \in \mathbb{F}^n$ for any field \mathbb{F} , we define $\langle u, v \rangle := \sum_{j \in [n]} u_j v_j$.

THEOREM 4.14. *Assume $n = 4p$, for a large enough prime p . Let $u^{(1)}, \dots, u^{(m)} \in \{0, 1\}_{n/2}^n \subseteq \mathbb{Z}^n$ be such that for each $v \in \{0, 1\}_{n/2}^n$, there is an $i \in [m]$ such that $\langle u^{(i)}, v \rangle = p$. Then $m \geq p$.*

The above theorem is nearly tight as can be seen by taking the indicator vectors of the sets $S_i = \{i, (i+1), \dots, i + (n/2) - 1\}$ for $i \in [n/2]$. Improvements on the above theorem (some of them asymptotically tight) were proved recently by Alon et al. [5] and Hrubeš et al. [25].

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

REMARK 4.15. We can prove a robust generalization (stated below) in a slightly more general setting where the i th inner product $\langle u^{(i)}, v \rangle$ is supposed to take a value b_i (which is not necessarily p). Similar to Theorem 4.14 above, it is easy to note that our robust version is tight up to constant factors.

However, if we consider the robust version of the original statement of Theorem 4.14 (where all the inner products take value p), 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 4.14.

THEOREM 4.16. *Assume n is a growing even integer parameter and $\varepsilon \in [2^{-n}, 1/2]$. Let $u^{(1)}, \dots, u^{(m)} \in \{0, 1\}_{n/2}^n \subseteq \mathbb{Z}^n$ and $b_1, \dots, b_m \leq n$ be such that*

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

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

The theorem can easily be seen to be tight up to constant factors. For $t = C \cdot \sqrt{n \log(1/\varepsilon)}$, set $m = 2t+1$ and take $u^{(1)} = u^{(2)} = \dots = u^{(m)} = 1^{n/2} 0^{n/2}$ and $b_1 = (n/4) - t, b_2 = (n/4) - t + 1, \dots, b_m = (n/4) + t$. 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 4.17. *Let n be an even integer and m a non-negative integer with $m \leq n/2$. Then, for any $k, \ell \in \{0, \dots, \lfloor m/2 \rfloor\}$ 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 4.16 as follows.

PROOF OF THEOREM 4.16. Recall that for any fixed $u \in \{0, 1\}_{n/2}^n$ 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 $v \sim \{0, 1\}_{n/2}^n$,

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

The second inequality above follows from Claim 4.17, 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

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

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

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

We know that P vanishes at a random point of $\{0, 1\}_{n/2}^n$ 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 $v \in \{0, 1\}_{n/2-p}^n$, we have

$$\Pr_v \left[\langle u^{(i)}, v \rangle \equiv b_i \pmod{p} \right] \leq \frac{\varepsilon^2}{\sqrt{n}} \tag{15}$$

for a large enough constant C . Informally speaking, the reason for this inequality is as follows: the expected value of $\langle u^{(i)}, v \rangle$ is $(n/4) - p/2$ and any number $b \equiv b_i \pmod{p}$ is far from this expectation. To prove this, let $s = n/2 - p$. Note that $s = \Omega(n)$ as long as t is small enough in relation to n , which happens if δ is assumed to be a small enough constant. Using the fact that i

is balanced, we note that

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

We thus have

$$\begin{aligned}\Pr_{\mathbf{v}} \left[\langle u^{(i)}, \mathbf{v} \rangle \equiv b_i \pmod{p} \wedge \langle u^{(i)}, \mathbf{v} \rangle \geq \left\lfloor \frac{s}{2} \right\rfloor \right] &= \sum_{j \geq 0} \Pr_{\mathbf{v}} \left[\langle u^{(i)}, \mathbf{v} \rangle = b_i + jp \right] \\ &= \sum_{j \geq 0} \Pr_{\mathbf{v}} \left[\langle u^{(i)}, \mathbf{v} \rangle - \left\lfloor \frac{s}{2} \right\rfloor = \Delta_i + jp \right] \\ &= \sum_{j \geq 0} \frac{\binom{n/2}{\left\lfloor \frac{s}{2} \right\rfloor + \Delta_i + jp} \binom{n/2}{\left\lfloor \frac{s}{2} \right\rfloor - \Delta_i - jp}}{\binom{n}{s}} \\ &\text{(by Claim 4.17)} = \frac{\binom{n/2}{\left\lfloor \frac{s}{2} \right\rfloor} \binom{n/2}{\left\lfloor \frac{s}{2} \right\rfloor}}{\binom{n}{s}} \cdot \sum_{j \geq 0} \exp(-\Omega((\Delta_i + jp)^2/s)) \\ &\text{(Stirling approximation and } s = \Omega(n) \leq O\left(\frac{1}{\sqrt{n}}\right) \cdot \sum_{j \geq 0} \exp(-\Omega(\Delta_i^2 + jp^2)/s)) \\ &\quad (p^2/s \geq C^2) \leq O\left(\frac{1}{\sqrt{n}}\right) \cdot \exp(-\Omega(\Delta_i^2/s)) \cdot \sum_{j \geq 0} \exp(-\Omega(C^2 j)) \\ &\text{(for large enough } C) \leq O\left(\frac{1}{\sqrt{n}}\right) \cdot \exp(-\Omega(\Delta_i^2/s)) \cdot 2 \\ &= O\left(\frac{1}{\sqrt{n}}\right) \cdot \exp(-\Omega(p^2/s)).\end{aligned}$$

In a similar way, we also get

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

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

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

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

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

From now on, we consider the polynomial P as an element of $\mathbb{F}_p[x_1, \dots, x_n]$. At this point, we would like to apply Lemma 3.1 to the polynomial P and finish the proof. Unfortunately, the error parameter ε_1 above is not small enough to apply Lemma 3.1 directly (we need $\varepsilon_1 \leq \exp(-200p^2/n)$). However, we can do a simple error reduction as in Lemma 3.10 to ensure that

Lemma 3.1 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 3.10 there is a probabilistic polynomial $\mathbf{P}^{(r)}$ of degree at most $r \cdot \deg(P)$ such that

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

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

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 $P' \in \mathbb{F}_p[x_1, \dots, x_n]$ of degree $r \cdot \deg(P) = O(m)$ such that

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

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

Hence, applying Lemma 3.1 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 and TheoretCS for their corrections and suggestions. In particular, a referee for the TheoretCS submission pointed out an extension to the main lemma (Lemma 3.13).

References

- [1] Amir Abboud, Richard Ryan Williams, and Huacheng Yu. More applications of the polynomial method to algorithm design. *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 [DOI](#) (2).
- [2] Rohit Agrawal. Coin theorems and the fourier expansion. *Chic. J. Theor. Comput. Sci.* 2020, 2020 (5, 23).
- [3] Miklós Ajtai and Michael Ben-Or. A theorem on probabilistic constant depth computations. *Proceedings of the 16th Annual ACM Symposium on Theory of Computing, STOC 1984, April 30 - May 2, 1984, Washington, DC, USA*, pages 471–474. ACM, 1984 [DOI](#) (4).
- [4] Josh Alman and Ryan Williams. Probabilistic polynomials and hamming nearest neighbors. *IEEE 56th Annual Symposium on Foundations of Computer Science, FOCS 2015, Berkeley, CA, USA, 17-20 October, 2015*, pages 136–150. IEEE Computer Society, 2015 [DOI](#) (2, 5, 9, 18).
- [5] Noga Alon, Mrinal Kumar, and Ben Lee Volk. Unbalancing sets and an almost quadratic lower bound for syntactically multilinear arithmetic circuits. *33rd Computational Complexity Conference, CCC 2018, June 22-24, 2018, San Diego, CA, USA*, 11:1–11:16, 2018 [DOI](#) (2, 3, 31).
- [6] James Aspnes, Richard Beigel, Merrick L. Furst, and Steven Rudich. The expressive power of voting polynomials. *Combinatorica*, 14(2):135–148, 1994 (6).

- [7] Peter Beelen and Mrinmoy Datta. Generalized Hamming weights of affine Cartesian codes. *Finite Fields Appl.* 51:130–145, 2018 [DOI](#) (13).
- [8] Richard Beigel. The polynomial method in circuit complexity. *Proceedings of the Eighth Annual Structure in Complexity Theory Conference, CCC 1993, San Diego, CA, USA, May 18-21, 1993*, pages 82–95. IEEE Computer Society, 1993 [DOI](#) (2, 5).
- [9] Jean Berstel and Juhani Karhumäki. Combinatorics on words: a tutorial. *Bulletin of the EATCS*, 79:178, 2003 (9).
- [10] Siddharth Bhandari, Prahladh Harsha, Tulasimohan Molli, and Srikanth Srinivasan. On the Probabilistic Degree of OR over the Reals. *38th IARCS Annual Conference on Foundations of Software Technology and Theoretical Computer Science (FSTTCS 2018)*, volume 122 of *Leibniz International Proceedings in Informatics (LIPIcs)*, 5:1–5:12, Dagstuhl, Germany. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, 2018 [DOI](#) (6).
- [11] Mark Braverman. Polylogarithmic independence fools AC^0 circuits. *J. ACM*, 57(5), 2010 [DOI](#) (2).
- [12] Joshua Brody and Elad Verbin. The coin problem and pseudorandomness for branching programs. *51th Annual IEEE Symposium on Foundations of Computer Science, FOCS 2010, October 23-26, 2010, Las Vegas, Nevada, USA*, pages 30–39. IEEE Computer Society, 2010 [DOI](#) (4).
- [13] Eshan Chattopadhyay, Pooya Hatami, Shachar Lovett, and Avishay Tal. Pseudorandom generators from the second fourier level and applications to AC^0 with parity gates. *10th Innovations in Theoretical Computer Science Conference, ITCS 2019, January 10-12, 2019, San Diego, California, USA*, 22:1–22:15, 2019 [DOI](#) (5, 23).
- [14] G.F. Clements and B. Lindström. A generalization of a combinatorial theorem of Macaulay. *Journal of Combinatorial Theory*, 7(3):230–238, 1969 [DOI](#) (6, 13).
- [15] Gil Cohen, Anat Ganor, and Ran Raz. Two sides of the coin problem. *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques, APPROX/RANDOM 2014, September 4-6, 2014, Barcelona, Spain*, volume 28 of *LIPIcs*, pages 618–629. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2014 [DOI](#) (4).
- [16] Ernie Croot, Vsevolod F Lev, and Péter Pál Pach. Progression-free sets in are exponentially small. *Annals of Mathematics*:331–337, 2017 (2).
- [17] Devdatt P Dubhashi and Alessandro Panconesi. Concentration of measure for the analysis of randomized algorithms. Cambridge University Press, 2009 (7).
- [18] Jordan S Ellenberg and Dion Gijswijt. On large subsets of with no three-term arithmetic progression. *Annals of Mathematics*:339–343, 2017 (2).
- [19] Hikoe Enomoto, Peter Frankl, Noboru Ito, and Kazumasa Nomura. Codes with given distances. *Graphs and Combinatorics*, 3(1):25–38, 1987 (3).
- [20] Alexander Golovnev, Rahul Ilango, Russell Impagliazzo, Valentine Kabanets, Antonina Kolokolova, and Avishay Tal. $AC^0[p]$ lower bounds against MCSP via the coin problem. *46th International Colloquium on Automata, Languages, and Programming, ICALP 2019, July 9-12, 2019, Patras, Greece*, volume 132 of *LIPIcs*, 66:1–66:15. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2019 [DOI](#) (5).
- [21] Larry Guth. Polynomial methods in combinatorics, volume 64 of *University Lecture Series*. American Mathematical Society, Providence, RI, 2016, pages ix+273 [DOI](#) (2).
- [22] Prahladh Harsha and Srikanth Srinivasan. On polynomial approximations to AC . *Random Struct. Algorithms*, 54(2):289–303, 2019 [DOI](#) (6, 8).
- [23] Gábor Hegedűs. Balancing sets of vectors. *Studia Scientiarum Mathematicarum Hungarica*, 47(3):333–349, 2009 (2, 3, 10, 31).
- [24] Petra Heijnen and Ruud Pellikaan. Generalized Hamming weights of q -ary Reed-Muller codes. *IEEE Trans. Inform. Theory*, 44(1):181–196, 1998 [DOI](#) (13).
- [25] Pavel Hrubes, Sivaramakrishnan Natarajan Ramamoorthy, Anup Rao, and Amir Yehudayoff. Lower bounds on balancing sets and depth-2 threshold circuits. *46th International Colloquium on Automata, Languages, and Programming, ICALP 2019, July 9-12, 2019, Patras, Greece*, volume 132 of *LIPIcs*, 72:1–72:14. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2019 [DOI](#) (2, 3, 31).
- [26] Peter Keevash and Benny Sudakov. Set systems with restricted cross-intersections and the minimum rank of inclusion matrices. *SIAM Journal on Discrete Mathematics*, 18(4):713–727, 2005 [DOI](#) (6, 13).
- [27] 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 [DOI](#) (2).
- [28] 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 [DOI](#) (2).
- [29] Swastik Kopparty and Srikanth Srinivasan. Certifying polynomials for $AC^0[\oplus]$ circuits, with applications to lower bounds and circuit compression. *Theory of Computing*, 14(1):1–24, 2018 [DOI](#) (6, 11, 12).
- [30] Alexander S. Kulikov and Vladimir V. Podolskii. Computing majority by constant depth majority circuits with low fan-in gates. *34th Symposium on Theoretical Aspects of Computer Science, STACS 2017, March 8-11, 2017, Hannover, Germany*, 49:1–49:14, 2017 [DOI](#) (3).
- [31] Nutan Limaye, Karteek Sreenivasaiyah, Srikanth Srinivasan, Utkarsh Tripathi, and S. Venkitesh. A fixed-depth size-hierarchy theorem for $AC^0[\oplus]$ via the coin problem. *SIAM J. Comput.* 50(4):1461–1499, 2021 [DOI](#) (4, 5, 22, 23).

- [32] Nathan Linial, Yishay Mansour, and Noam Nisan. Constant depth circuits, fourier transform, and learnability. *J. ACM*, 40(3):607–620, 1993 [DOI](#) (2).
- [33] Chi-Jen Lu. An exact characterization of symmetric functions in $\text{qAC}^0[2]$. *Theoretical Computer Science*, 261(2):297–303, 2001 (8–10).
- [34] Raghu Meka, Oanh Nguyen, and Van Vu. Anti-concentration for polynomials of independent random variables. *Theory of Computing*, 12(1):1–17, 2016 [DOI](#) (6).
- [35] 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 [DOI](#) (6, 12).
- [36] Aditya Potukuchi. On the $\text{AC}^0[\oplus]$ complexity of andreev's problem. *39th IARCS Annual Conference on Foundations of Software Technology and Theoretical Computer Science, FSTTCS 2019, December 11–13, 2019, Bombay, India*, volume 150 of *LIPICs*, 25:1–25:14. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2019 [DOI](#) (5).
- [37] 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 [DOI](#) (3).
- [38] Alexander A. Razborov. Lower bounds on the size of bounded depth circuits over a complete basis with logical addition. *Russian. Matematicheskije Zametki*, 41(4):598–607, 1987 [DOI](#). (English translation in *Mathematical Notes of the Academy of Sciences of the USSR*, 41(4):333–338, 1987) (2, 5, 6).
- [39] Ronen Shaltiel and Emanuele Viola. Hardness amplification proofs require majority. *SIAM J. Comput.* 39(7):3122–3154, 2010 [DOI](#) (4, 5).
- [40] 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 [DOI](#) (3).
- [41] Roman Smolensky. Algebraic methods in the theory of lower bounds for boolean circuit complexity. *Proceedings of the nineteenth annual ACM symposium on Theory of computing*, pages 77–82. ACM, 1987 (2, 6, 23, 25).
- [42] Roman Smolensky. On representations by low-degree polynomials. *Proceedings of 1993 IEEE 34th Annual Foundations of Computer Science, FOCS 1993*, pages 130–138. IEEE, 1993 (6, 25).
- [43] Srikanth Srinivasan, Utkarsh Tripathi, and S. Venkitesh. On the probabilistic degrees of symmetric boolean functions. *SIAM J. Discret. Math.* 35(3):2070–2092, 2021 [DOI](#) (5, 6, 9, 24, 26).
- [44] Srikanth Srinivasan and S. Venkitesh. On vanishing properties of polynomials on symmetric sets of the boolean cube, in positive characteristic, 2021 [DOI](#). Available at <https://arxiv.org/abs/2111.05445> (2, 19).
- [45] M. Szegedy. Algebraic methods in lower bounds for computational models with limited communication. *PhD thesis*, The University of Chicago, 1989 (6, 25).
- [46] Leslie G. Valiant. Short monotone formulae for the majority function. *J. Algorithms*, 5(3):363–366, 1984 [DOI](#) (4).
- [47] Emanuele Viola. On approximate majority and probabilistic time. *Computational Complexity*, 18(3):337–375, 2009 (4).
- [48] V. K. Wei. Generalized hamming weights for linear codes. *IEEE Transactions on Information Theory*, 37(5):1412–1418, September 1991 [DOI](#) (6, 13).
- [49] R. Ryan Williams. Faster all-pairs shortest paths via circuit complexity. *SIAM J. Comput.* 47(5):1965–1985, 2018 [DOI](#) (2).
- [50] Richard Ryan Williams. The polynomial method in circuit complexity applied to algorithm design (invited talk). *34th International Conference on Foundation of Software Technology and Theoretical Computer Science, FSTTCS 2014, December 15–17, 2014, New Delhi, India*, volume 29 of *LIPICs*, pages 47–60. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2014 [DOI](#) (5).
- [51] Ryan Williams. New algorithms and lower bounds for circuits with linear threshold gates. *Symposium on Theory of Computing, STOC 2014, New York, NY, USA, May 31 - June 03, 2014*, pages 194–202. ACM, 2014 [DOI](#) (2).
- [52] Ryan Williams. Nonuniform ACC circuit lower bounds. *J. ACM*, 61(1):2:1–2:32, 2014 [DOI](#) (2).

A. Lemma 1.1 is implied by Lemma 3.1 (up to constant factors)

The following claim shows that if there is a P satisfying the hypotheses of Lemma 1.1, 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 $\{0, 1\}_{k+q}^n$.

CLAIM A.1. *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 $P \in \mathbb{F}[x_1, \dots, x_n]$ is any polynomial that vanishes at all $a \in \{0, 1\}_k^n$ but does not vanish at some $b \in \{0, 1\}_{k+q}^n$, then there is a $Q \in$*

$\mathbb{F}[x_1, \dots, x_n]$ of degree at most $\deg(P)$ such that Q vanishes at all $a \in \{0, 1\}_k^n$ but is non-zero at at least a $(1 - 1/p)$ fraction of the points in $\{0, 1\}_{k+q}^n$.

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 .

$$\begin{aligned} |a| = k &\Rightarrow P(a) = 0 \\ P(b) &= 1. \end{aligned}$$

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

Consider the degree- d closure $C = \text{cl}_d(\{0, 1\}_k^n)$. By the existence of P , we see that $b \notin C$. However, by symmetry, this implies that no point $b' \in \{0, 1\}_{k+q}^n$ 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 $\{0, 1\}_k^n$. Let Q be a uniformly random element of $V_{d,k}$. For any $c \in \{0, 1\}^n \setminus C$, standard linear algebra implies that $Q(c)$ is a uniformly random element of $\mathbb{F} = \mathbb{F}_p$. In particular, for any $b' \in \{0, 1\}_{k+q}^n$, we see that

$$\Pr_Q [Q(b') \neq 0] = 1 - 1/p.$$

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 $\{0, 1\}_{k+q}^n$. This yields the statement of the claim. ■

B. Proof of Lemma 2.6 (the string lemma)

We begin by recalling the statement of the lemma.

LEMMA 2.6. *Let $w \in \{0, 1\}^+$ be any non-empty string¹² 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,

¹² Recall that, for any alphabet Σ , the notation Σ^+ denotes the set of non-empty strings over this alphabet.

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 4.17

We first restate the claim.

CLAIM 4.17. *Let n be an even integer and m a non-negative integer with $m \leq n/2$. Then, for any $k, \ell \in \{0, \dots, \lfloor m/2 \rfloor\}$ with $\ell \leq k$, we have*

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

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

$$\frac{\binom{n/2}{\lfloor m/2 \rfloor - k - 1} \binom{n/2}{\lceil m/2 \rceil + k + 1}}{\binom{n/2}{\lfloor m/2 \rfloor - k} \binom{n/2}{\lceil m/2 \rceil + k}} \leq \exp(-\Omega(k/m)). \quad (16)$$

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

To prove (16), we proceed as follows. By an expansion of binomial coefficients in terms of factorials, we see that

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