On the Degree of Boolean Functions as Polynomials over $\mathbb{Z}_m$

Xiaoming Sun
Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China
sunxiaoming@ict.ac.cn

Yuan Sun
Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China
sunyuan2016@ict.ac.cn

Jiaheng Wang
School of Electronics Engineering and Computer Science, Peking University, Beijing, China
pw384@hotmail.com

Kewen Wu
School of Electronics Engineering and Computer Science, Peking University, Beijing, China
shlw_kevin@hotmail.com

Zhiyu Xia
Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China
xiazhiyu@ict.ac.cn

Yufan Zheng
Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China
yufan.zheng@pku.edu.cn

Abstract

Polynomial representations of Boolean functions over various rings such as $\mathbb{Z}$ and $\mathbb{Z}_m$ have been studied since Minsky and Papert (1969). From then on, they have been employed in a large variety of areas including communication complexity, circuit complexity, learning theory, coding theory and so on. For any integer $m \geq 2$, each Boolean function has a unique multilinear polynomial representation over ring $\mathbb{Z}_m$. The degree of such polynomial is called *modulo-m degree*, denoted as $\deg_m(\cdot)$.

In this paper, we investigate the lower bound of modulo-m degree of Boolean functions. When $m = p^k$ $(k \geq 1)$ for some prime p, we give a tight lower bound $\deg_m(f) \geq k(p-1)$ for any non-degenerate function $f : \{0,1\}^n \to \{0,1\}$, provided that n is sufficient large. When m contains two different prime factors p and q, we give a nearly optimal lower bound for any symmetric function $f : \{0,1\}^n \to \{0,1\}$ that $\deg_m(f) \geq \frac{n}{2+\frac{1}{p-1}+\frac{1}{q-1}}$.

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

Keywords and phrases Boolean function, polynomial, modular degree, Ramsey theory

Digital Object Identifier 10.4230/LIPIcs.ICALP.2020.100

Category Track A: Algorithms, Complexity and Games

Funding This work was supported in part by the National Natural Science Foundation of China Grants No. 61832003, 61761136014, the 973 Program of China Grant No. 2016YFB1000201, the Strategic Priority Research Program of Chinese Academy of Sciences Grant No. XDB28000000, and K. C. Wong Education Foundation.

Acknowledgements We thank Noga Alon and Xiaoxu Guo for pointing us to relevant references [1, 9]. We also appreciate anonymous reviewers' careful feedback and constructive advice.

1 Introduction

Given a Boolean function $f : \{0,1\}^n \to \{0,1\}$, the *degree* (resp., *modulo-m degree*), denoted as $\deg(f)$ (resp., $\deg_m(f)$), is the degree of the unique[1] multilinear polynomial representation of f over $\mathbb{R}$ (resp., $\mathbb{Z}_m$). These complexity measures and related notions have been studied extensively since the work of Minsky and Papert [23]. The polynomial representation of a Boolean function has found numerous applications in the study of query complexity (see e.g. [5]), communication complexity [4, 28, 31, 30, 29, 24, 10], learning theory [17, 20, 16, 25], explicit combinatorial constructions [13, 14, 11, 7], circuit lower bounds [33, 27, 1, 12] and coding theory [36, 37, 15, 21], etc.

In this paper, we focus on modulo-m degree of Boolean functions. Throughout, all Boolean functions are assumed to be *non-degenerate*[2], if not specifically mentioned. One of the complexity theoretic motivations of studying $\deg_m(f)$ is to understand the power of modular counting. For example, the famous Razborov–Smolensky polynomial method [27, 33] reduces the task of proving size lower bounds for $\mathsf{AC}^0[p]$ circuits to proving a lower bound of approximate modulo-p degree of the target Boolean function. However, this approach mainly works when p is a prime.[3] Another example, in which m can be composite, is that a $(1/2 + o(1))$-inapproximability of a Boolean function f by degree-$O(1)$ polynomials over $\mathbb{Z}_m$ implies that f cannot be computed by $\mathsf{MAJ}_{O(1)} \circ \mathsf{MOD}_m \circ \mathsf{AND}_{O(1)}$ circuits [1]. In general, it has been proved important to understand the computational power of polynomials over $\mathbb{Z}_m$ for general m.

Towards the complexity measure $\deg_m(f)$ itself, the case when m is a prime has been studied a lot in previous works. For example, one natural question is whether $\deg_m(f)$ is polynomially related to $\deg(f)$ for general m, as other complexity measures like decision tree complexity $\mathrm{D}(f)$ do? The answer is NO according to the parity function $\mathsf{PARITY}(x) := \bigoplus_{i=1}^n x_i$. That is, $\deg_2(\mathsf{PARITY}) = 1$ but $\deg(\mathsf{PARITY}) = n$. Though this function works as a counterexample for the relationship between $\deg_2(f)$ and $\deg(f)$, it is still inspiring because its modulo-3 degree is large. By writing PARITY as $\frac{1}{2} - \frac{1}{2}\prod_{i=1}^n (1 - 2x_i)$ and taking modulo 3, one can get $\deg_3(\mathsf{PARITY}) = n$. Actually, Gopalan et al. [12] give the following relationship between the polynomial degrees modulo two different primes p and q:

$$\deg_q(f) \geq \frac{n}{\lceil \log_2 p \rceil \deg_p(f) p^{2\deg_p(f)}}.$$

Daunting at the first glance, the inequality implies an essential fact that, as long as $\deg_p(f) = o(\log n)$, a lower bound of $\Omega(n^{1-o(1)})$ for $\deg_q(f)$ follows. Moreover, if m has at least two different prime factors p and q, then $\deg_m(f) \geq \max\{\deg_p(f), \deg_q(f)\} = \Omega(\log n)$.

Having negated the possibility for the case of prime m, it is natural to study the case of composite number. The systematic study of this case was initiated by Barrington et al. [3]. Alas, whether $\deg_m(f)$ is polynomially related to $\deg(f)$ is still a widely open problem. Though the case m being a prime power is proved to be not true in Gopalan's thesis [10], we are unable to find better separation between $\deg_m(f)$ and $\deg(f)$, for $m = pq$ with p and q being two distinct primes, than the quadratic one given by Li and Sun [19]. This leads to the following conjecture:

---

[1] The existence and uniqueness are guaranteed by the Möbius inversion, see e.g. [12].

[2] A Boolean function is called non-degenerate if it depends on all its n variables.

[3] It is a folklore that $\mathsf{AC}^0[m] = \mathsf{AC}^0[\mathrm{rad}(m)]$, where $\mathrm{rad}(m)$ is the square-free part of m. Therefore in fact we are able to handle $\mathsf{AC}^0[q]$ circuits for any prime power q.

▶ **Conjecture 1.1.** *Let f be a Boolean function. If m has at least two distinct prime factors, then*

$$\deg(f) = O(\mathrm{poly}(\deg_m(f))).$$

Towards this conjecture, the first step is to deal with *symmetric* Boolean functions. Lee et al. [18] proves that $2\deg_p(f)\deg_q(f) > n$ for any distinct primes p, q and non-trivial symmetric Boolean function $f : \{0,1\}^n \to \{0,1\}$, implying the correctness of Conjecture 1.1 in symmetric cases. Li and Sun [19] improved their bound to $p\deg_p(f) + q\deg_q(f) > n$, which implies $\deg_{pq}(f) > \frac{n}{p+q}$. This is far from being tight; actually, as we will present later, the denominator $p+q$ can be reduced to 3.5.

On the tight lower bound of $\deg(f)$, Nisan and Szegedy [26] give the bound $\deg(f) \geq \log_2 n - O(\log\log n)$ as long as f is non-degenerate. Very recently, this bound is improved to $\deg(f) \geq \log_2 n - O(1)$ by [6, 35], which is tight up to the additive $O(1)$-term by the *address function*. Gathen and Roche [34] show that $\deg(f) \geq \deg_{p(n)}(f) \geq p(n) - 1$ for any non-trivial *symmetric* Boolean function, where $p(n)$ is the largest prime below $n+2$. (Notice that the module degree gives a lower bound on the degree.) Using the currently best result on prime gaps [2], this gives an $n - O(n^{0.525})$ lower bound. On the other side, Gathen and Roche give a polynomial family with $\deg(f) = n - 3$, and they propose Conjecture 1.2 below with a probabilistic heuristic argument:

▶ **Conjecture 1.2.** *For any non-trivial symmetric Boolean function $f : \{0,1\}^n \to \{0,1\}$,*

$$\deg(f) \geq n - O(1).$$

Our Results. In this work, we prove the following four theorems, giving better lower bounds for $\deg_m(f)$. As we have already mentioned, the gap between $\deg(f)$ and $\deg_{p^k}(f)$ can be arbitrarily large. Nevertheless, we claim that $\deg_{p^k}(f)$ cannot be too small either. This begins with symmetric functions:

▶ **Theorem 1.3.** *For any prime p, positive integer k, and non-trivial symmetric function $f : \{0,1\}^n \to \{0,1\}$,*

$$\deg_{p^k}(f) \geq (p-1)\cdot k$$

when $n \geq (k-1)\varphi(p^\mu) + p^\mu - 1 \in O(p^2k^2)$ where $\mu = \lceil \log_p((p-1)k - 1)\rceil$. The bound $(p-1)\cdot k$ is tight.

The proof of Theorem 1.3 is centered around Mahler expansion [22], which has been deemed useful in several fields of study, from analytic functions to combinatorics. Wilson [36] studied Mahler coefficients and related degree to period of symmetric functions. However, by introducing some more insights, we are able to give a stronger analysis to settle this case once for all. To be a bit more concrete, our argument (i) introduces the base-p period to replace normal period, and then (ii) spans every symmetric functions into two fashions, by MODs or binomials, and then (iii) introduces Mahler coefficient matrix and determines its kernel.

In addition, Theorem 1.3 can be extended to non-degenerate Boolean functions. We achieve this by showing that one can embed an $\omega(1)$-size non-trivial symmetric Boolean function into any non-degenerate functions by applying Erdős–Rado Theorem from Ramsey theory.[4] This leads to the same tight bound, provided that the input size is sufficiently large.

[4] We note that a similar embedding argument has appeared before in [1].

▶ **Theorem 1.4.** *For any prime p, positive integer k, and non-degenerate function $f : \{0,1\}^n \to \{0,1\}$ with sufficiently large n,*

$$\deg_{p^k}(f) \geq (p-1) \cdot k.$$

The bound $(p-1) \cdot k$ is tight.

Now turn to the case of non-prime-power composite m. The following theorem provides a lower bound on $\deg_m(f)$.

▶ **Theorem 1.5.** *For any composite number m with at least two different prime factors p, q and any non-trivial symmetric Boolean function $f : \{0,1\}^n \to \{0,1\}$,*

$$\deg_m(f) \geq \frac{1}{2 + \frac{1}{p-1} + \frac{1}{q-1}} \cdot n.$$

Note that this bound approaches $n/2$ when p and q become larger and larger. It improves the $n/(p+q)$ bound in [19]. To prove this theorem, we show a stronger version of Theorem 1.3 for $k = 1$, which requires a more elaborate analysis. Then we utilizes Periodicity Lemma [9] to obtain the desired lower bound.

On the other hand, the next theorem shows that the bound in Theorem 1.5 cannot be larger than $(1 + o(1))n/2$:

▶ **Theorem 1.6.** *Let m be a square-free composite number. There exists a symmetric Boolean function $f : \{0,1\}^n \to \{0,1\}$ with arbitrarily large n, such that $\deg_m(f) \leq n/2 + o_m(n)$.*[5]

Organization. In Section 2, we give necessary definitions and concepts. Then we give the proofs of Theorem 1.3 and Theorem 1.4 respectively in Section 3.1 and Section 3.2. In Section 4.2 we prove Theorem 1.5, and in Section 4.3 we prove Theorem 1.6. Finally, we conclude the paper in Section 5.

2 Preliminaries

We denote $\{1, 2, \ldots, n\}$ as $[n]$ throughout this paper. $\varphi(\cdot)$ denotes Euler's totient function. Notation $\log^{\circ k}(n)$ is defined as $\underbrace{\log \log \cdots \log}_{k} n$, and $\log^*(n)$ is for the iterated logarithm, that is, $\min\{k : \log^{\circ k}(n) \leq 1\}$.

2.1 Basics of Boolean Functions

An n-bit *Boolean function* $f(x)$ is a mapping from $\{0,1\}^n$ to $\{0,1\}$. Sometimes we write $\boldsymbol{x}$ to indicate the n-dimensional 0-1 vector corresponding to string $x \in \{0,1\}^n$. The following operation will be frequently used: Suppose $x \in \{0,1\}^n$ is a (input) string, and $S \subseteq [n]$ is a set of indices. Denote the string obtained by flipping all bits in x whose indices are in S as $x^{\oplus S}$. As a common practice, $x^{\oplus \{i\}}$ is abbreviated as $x^{\oplus i}$.

[5] The subscript "m" in the $o(\cdot)$ notation means that the hidden factor depends on m.

Here we list some subclasses of Boolean functions, which we will frequently deal with later:

- A Boolean function is called *non-trivial* if it is not a constant.
- A Boolean function is called *non-degenerate* if its value depends on all input bits. In other words, there does not exist such t that, for every $x \in \{0,1\}^n$ the equality $f(x) = f(x^{\oplus t})$ holds. Such bit, if exists, is also known as *dumb* bit.
- A Boolean function is called *symmetric*, if $f(x) = f(y)$ for any x, y satisfying $|x| = |y|$. Here $|x|$ denotes the Hamming weight of x, i.e., number of 1's.

There exists a unique polynomial representing f over $\mathbb{Z}_m$ or $\mathbb{Z}$. More formally:

▶ **Fact 2.1.** *For any Boolean function $f : \{0,1\}^n \to \{0,1\}$, the unique polynomial*

$$\sum_{a\in\{0,1\}^n} f(a) \prod_{i=1}^{n} ((2a_i - 1)x_i + 1 - a_i) =: \sum_{S\subseteq[n]} c_S \prod_{i\in S} x_i$$

represents f over $\mathbb{Z}$. On top of this, the polynomial $\sum_{S\subseteq[n]}(c_S \bmod m)\prod_{i\in S} x_i$ represents f over $\mathbb{Z}_m$.

▶ **Definition 2.2.** *The degree (resp., modulo-m degree) of a Boolean function f, denoted by $\deg(f)$ (resp., $\deg_m(f)$), is the degree of the polynomial representing f over $\mathbb{Z}$ (resp., $\mathbb{Z}_m$).*

This measure has some simple but useful properties. The following fact is a consequence of the Chinese Remainder Theorem; see [19, Fact 5].

▶ **Fact 2.3.** *Suppose $f : \{0,1\}^n \to \{0,1\}$ is a Boolean function, and m, m' are coprime. Then $\deg_{m'm}(f) = \max\{\deg_m(f), \deg_{m'}(f)\}$.*

With some input bits fixed, the degree of a Boolean function may decrease. This can be easily derived by substituting those variables with their values. More formally, we define the *restriction* of Boolean functions and restate this fact below.

▶ **Definition 2.4** (Restriction)**.** *Suppose $f : \{0,1\}^n \to \{0,1\}$ is a Boolean function, $S \subseteq [n]$ is a set of indices, and there is a mapping $\sigma : [n]\backslash S \to \{0,1\}$. For every $i \notin S$, fix the i-th bit in the input of f to be $\sigma(i)$ to obtain a new Boolean function with input size $|S|$. We call it the restriction of f over σ, denoted as $f|_\sigma$.*

▶ **Fact 2.5.** *Suppose $f : \{0,1\}^n \to \{0,1\}$ is a Boolean function. For any integer $m \geq 2$ and restriction $f|_\sigma$, we have $\deg_m(f) \geq \deg_m(f|_\sigma)$.*

A common complexity measure, the *sensitivity*, will be used in Section 3.2. Simon gave a lower bound on this measure [32].

▶ **Definition 2.6** (Sensitivity)**.** *Given a Boolean function $f : \{0,1\}^n \to \{0,1\}$ and an input x, we say bit i is sensitive if $f(x) \neq f(x^{\oplus i})$. The sensitivity of f on input x is $s(f,x) := |\{i : i \in [n], f(x) \neq f(x^{\oplus i})\}|$. The sensitivity of f is then defined as $s(f) := \max_x s(f,x)$.*

▶ **Theorem 2.7** ([32])**.** *For any non-degenerate Boolean function $f : \{0,1\}^n \to \{0,1\}$, we have*

$$s(f) \geq \frac{1}{2}\log n - \frac{1}{2}\log\log n + \frac{1}{2}.$$

2.2 Periodicity and Mahler Expansion

We consider symmetric Boolean functions in this section. For a symmetric function $f : \{0,1\}^n \to \{0,1\}$, clearly there exists a unique $F : \{0,\ldots,n\} \to \{0,1\}$, called the *univariate version of* f, such that $f(x) = F(|x|)$ for every x. We call f (and F) *ℓ-periodic*, if $F(a) = F(b)$ for any $0 \le a, b \le n$ satisfying $\ell \mid a - b$. For example, f is trivially ℓ-periodic for any $\ell > n$. We are also interested in integer power period length. Hence we introduce the following definition.

▶ **Definition 2.8** (Base-m period)**.** *Assume $f : \{0,1\}^n \to \{0,1\}$ is a symmetric Boolean function. The base-m period is the minimum ℓ such that ℓ is a power of m, and f is ℓ-periodic. Denote it as $\pi_m(f)$.*

Here are some concrete examples for a clearer illustration.

- The not-all-equal NAE function is defined as $\mathsf{NAE}_n(x_1,\ldots,x_n) := \mathbb{I}[\exists i,j \text{ s.t. } x_i \neq x_j]$. Then $\pi_3(\mathsf{NAE}_3) = 3$ while $\pi_3(\mathsf{NAE}_4) = 9$. That is, $\pi_m(f)$ may be larger than n.
- If f is a trivial function, then $\pi_3(f) = 1$.

One may write F as a univariate polynomial over $\mathbb{Q}$, but it will not always induce a polynomial when we move to work on $\mathbb{Z}_m$, like what f does. Fortunately, the following representation, also known as *Mahler expansion* [36], serves the purpose similar to polynomial representation.

▶ **Theorem 2.9** (Mahler expansion)**.** *Assume that $f : \{0,1\}^n \to \{0,1\}$ is a symmetric Boolean function, and F is the corresponding univariate version. Let $d := \max\{n, m-1\}$. Then there exists a unique sequence $\alpha_0, \alpha_1, \cdots, \alpha_d \in \mathbb{Z}_m$ such that*

$$\sum_{j=0}^{d} \alpha_j \binom{t}{j} = \begin{cases} F(t), & 0 \le t \le n; \\ 0, & n < m-1 \text{ and } n < t < m. \end{cases}$$

We call $\sum_{j=0}^{d} \alpha_j \binom{t}{j}$ the Mahler expansion of F over $\mathbb{Z}_m$, and α_j the j-th Mahler coefficient.

There are some connections between polynomial degree and Mahler expansion. Over ring $\mathbb{Z}_m$, let $d^* := \max\{\ell : \alpha_\ell \not\equiv 0 \pmod m, \ell \le n\}$. If we only take 0-th to d^*-th terms in the Mahler expansion to get $\hat{F}(t) = \sum_{j=0}^{d^*} \alpha_j \binom{t}{j}$, then $\hat{F}(|x|) = F(|x|)$ for all $x \in \{0,1\}^n$, which implies

▶ **Fact 2.10.** $\deg_m(f) = \max\{\ell : \alpha_\ell \not\equiv 0 \pmod m, \ell \le n\}$.

▶ Remark. The fact above does not hold if we take away the condition $\ell \le n$. The next example shows that on $\mathbb{Z}_m$, the existence of high-order non-zero Mahler coefficient does not imply high degree, if the input length is too short.

▶ Example 2.11. Let $n = 2$ and $f(x) = -x_0x_1 + x_0 + x_1 = x_0 \vee x_1$. On $\mathbb{Z}_5$, one can verify its Mahler expansion is

$$f(x) = \binom{|x|}{1} + 4\binom{|x|}{2} + 2\binom{|x|}{4}.$$

But $\deg_5(f) = 2$.

This phenomenon does not come from nowhere; intuitively speaking, in the Mahler expansion over $\mathbb{Z}_m$, one may need to imitate Lagrange-style interpolation for $|x| > n$, and hence introduce some high-order terms. (Although $|x|$ can never be above n, we need to utilize MOD functions later in this paper, which requires $F(t)$ to be zero for $n < t < m$.)

Wilson [36] showed the following result about the degree and Mahler expansion of symmetric Boolean functions, given the base-p period.

▶ **Theorem 2.12** ([36, Lemma 1]). *Let p be a prime, and t, k be positive integers. Assume $f : \{0,1\}^n \to \{0,1\}$ is a symmetric Boolean function, and $\{\alpha_\ell\}$ are its Mahler coefficients over $\mathbb{Z}_{p^k}$. If f is p^t-periodic, then*

$$\deg_{p^k}(f) \le (k-1)\cdot\varphi(p^t) + p^t - 1.$$

In addition, for any positive integer j and $\ell \ge j\cdot\varphi(p^t)+p^t$, we have $\alpha_\ell \equiv 0 \pmod{p^j}$.

2.3 MOD and Its Mahler Expansion over $\mathbb{Z}_{p^k}$

We first look into the Mahler expansion of weight modular functions. This special case is illuminating in our later proofs. The MOD function is defined as

$$\mathsf{MOD}_n^{c,m}(x) := \mathbb{I}\big[|x| \equiv c \pmod m\big] \in \{0,1\},$$

where $n \ge m-1$ denotes the length of input x, and $\mathbb{I}[\cdot]$ is the indicator function. The following theorem gives the degree of MOD_n^{0,p^t}.

▶ **Theorem 2.13** ([36, Theorem 10]). *Let p be a prime, and t, k be positive integers. Denote $d := (k-1)\cdot\varphi(p^t)+p^t-1$. Then for any $n \ge d$, we have*

$$\deg_{p^k}(\mathsf{MOD}_n^{0,p^t}) = d.$$

In fact, we can achieve a more general result by further analysis. Fix n, p, t and k. Notate the Mahler coefficient of MOD_n^{a,p^t} over $\mathbb{Z}_{p^k}$ as $\alpha_\ell^{(a,p^t)}$ i.e., $\mathsf{MOD}_n^{a,p^t}(x) = \sum_{j=0}^{d}\alpha_j^{(a,p^t)}\binom{|x|}{j}$. Moreover, MOD_n^{a,p^t} can also be represented with $\alpha_\ell^{(0,p^t)}$ as

$$\mathsf{MOD}_n^{a,p^t}(x) = \sum_{j=0}^{d}\alpha_j^{(0,p^t)}\binom{|x|-a}{j}.$$

Then expand each $\binom{|x|-a}{j}$ by Vandermonde convolution to get

$$\alpha_\ell^{(a,p^t)} = \sum_{i=0}^{d-\ell}\binom{-a}{i}\alpha_{i+\ell}^{(0,p^t)}. \tag{1}$$

Specially, by setting $\ell = d$, we get $\alpha_d^{(a,p^t)} = \alpha_d^{(0,p^t)}$. This equation generalizes the theorem to all remainders.

▶ **Corollary 2.14.** *Let p be a prime, and t and k be positive integers. Denote $d := (k-1)\cdot\varphi(p^t)+p^t-1$. For any $n \ge d$ and $0 \le a < p^t$, we have*

$$\deg_{p^k}(\mathsf{MOD}_n^{a,p^t}) = d.$$

3 Lower Bound of $\deg_{p^k}(f)$

By identifying the degree of MOD_n^{i,p^t} over $\mathbb{Z}_{p^k}$, we show that the degree of all p^t-periodic functions is constantly small since they can be spanned by $\{\mathsf{MOD}_n^{j,p^t}\}_{j=0}^{p^t-1}$. In Section 3.1, we prove that the degree of any p^t-periodic (but not p^{t-1}-periodic) function will not decrease too much from $(k-1)\cdot\varphi(p^t)+p^t-1$ during the spanning, despite the cancellation of the high-order coefficients. By a Ramsey-type argument in Section 3.2, we further extend our lower bound to all non-degenerate Boolean function with sufficiently many input bits.

3.1 Symmetric Functions – Proof of Theorem 1.3

We begin with the periodicity of symmetric Boolean functions with low degree. In our proof, the following Lucas's Theorem is important.

▶ **Theorem 3.1** (Lucas). *Let $n, m \in \mathbb{N}$, and p be a prime. Assume in base p, n and m can be represented as $n = (n_v n_{v-1} \cdots n_0)_p$ and $m = (m_v m_{v-1} \cdots m_0)_p$ (the number with fewer digits are padded with 0). Then*

$$\binom{n}{m} \equiv \binom{n_v}{m_v}\binom{n_{v-1}}{m_{v-1}} \cdots \binom{n_0}{m_0} \pmod{p}.$$

The next lemma indicates the periodicity of symmetric Boolean functions with low degree.

▶ **Lemma 3.2.** *Let $f : \{0,1\}^n \to \{0,1\}$ be a symmetric Boolean function. For prime p and positive integers t and k, if $\deg_{p^k}(f) \le p^t - 1$, then f is p^t-periodic.*

Proof. Denote $d := \deg_{p^k}(f)$, and suppose α_ℓ are the Mahler coefficients of f over ring $\mathbb{Z}_{p^k}$, i.e., $f(x) = \sum_{j=0}^{d} \alpha_j \binom{|x|}{j} \bmod p^k$. According to Lucas's Theorem, if $a \equiv b \pmod{p^t}$, then for any $0 \le j \le p^t - 1$, we have $\binom{a}{j} \equiv \binom{a_v}{j_v} \cdots \binom{a_t}{j_t}\binom{a_{t-1}}{j_{t-1}} \cdots \binom{a_0}{j_0} \pmod{p}$. Here, a_i (resp. b_i and j_i) is the representation of a (resp. b and j) in the base p. Note that $j_i = 0$ for any $i \ge t$. Hence $\binom{a}{j} \equiv \binom{a_{t-1}}{j_{t-1}} \cdots \binom{a_0}{j_0} \pmod{p}$. For the same reason, $\binom{b}{j} \equiv \binom{b_{t-1}}{j_{t-1}} \cdots \binom{b_0}{j_0} \pmod{p}$. In addition, a and b's last t digits are the same as $a \equiv b \pmod{p^t}$. Thus $\sum_{j=0}^{d} \alpha_j \binom{a}{j} \bmod p = \sum_{j=0}^{d} \alpha_j \binom{b}{j} \bmod p$. By the definition that $f(x) \in \{0,1\}$, we get $\sum_{j=0}^{d} \alpha_j \binom{a}{j} \bmod p^k = \sum_{j=0}^{d} \alpha_j \binom{b}{j} \bmod p^k$. ◀

Next, provided $\pi_p(f)$, we give some lower bounds on the degree by the following two lemmas, conditioned on that n is large enough. Together with the lemma above we lead to a contradiction, and our theorem follows eventually. Before continuing, notice again that the value $f(x)$ is related only to the Hamming weight of x, and thus $f(x) = \sum_{0 \le i \le n : F(i)=1} \mathsf{MOD}_n^{i,n+1}(x)$. Specially, if $f(x)$ is t-periodic, then we can write $f(x)$ as $f(x) = \sum_{0 \le i \le t-1 : F(i)=1} \mathsf{MOD}_n^{i,t}(x)$ and apply Corollary 2.14.

▶ **Lemma 3.3.** *Assume p is a prime, and $f : \{0,1\}^n \to \{0,1\}$ is a symmetric Boolean function of period p. If for some positive integer k, it holds that $n \ge (p-1) \cdot k$, then*

$$\deg_{p^k}(f) = (p-1) \cdot k.$$

Proof. Expand $f(x)$ to get $\sum_{0 \le i \le p-1 : F(i)=1} \mathsf{MOD}_n^{i,p}(x)$. According to Corollary 2.14, the degree of each term in the summation is $\deg_{p^k}\left(\mathsf{MOD}_n^{i,p}\right) = (p-1) \cdot k =: d$. On the other hand, Section 2.3 says the coefficient of degree d term is identical in the polynomial of every MOD. As $f(x)$ is non-trivial, the number of terms in the summation, denoted as N, is neither 0 nor p. Therefore, the degree d term in $f(x)$ has coefficient $N \cdot \alpha_d^{(0,p)} \not\equiv 0 \pmod{p}$, implying $\deg_{p^k}(f) = d$ as desired. ◀

▶ **Lemma 3.4.** *Assume p is a prime, and $f : \{0,1\}^n \to \{0,1\}$ is a symmetric Boolean function with $\pi_p(f) = p^t$. Here $t \ge 1$ and $n \ge (k-1) \cdot \varphi(p^t) + p^t - 1$. Then*

$$\deg_{p^k}(f) \ge (k-2) \cdot \varphi(p^t) + p^t.$$

Proof. Consider over the ring $\mathbb{Z}_{p^k}$. Let $d := (k-1) \cdot \varphi(p^t) + p^t - 1$. Provided p^t, we abbreviate $\alpha_\ell^{(j,p^t)}$, the ℓ-th Mahler coefficients of MOD_n^{j,p^t}, as $\alpha_\ell^{(j)}$ for convenience. According to Corollary 2.14, we have $\deg_{p^k}\left(\mathsf{MOD}_n^{j,p^t}\right) = d$.

For all $0 \le i \le p^t - 1$, Theorem 2.12 implies the fact that $\alpha_\ell^{(i)}$ can be divided by p^{k-2} when $\ell \ge (k-2)\cdot\varphi(p^t) + p^t = d - \varphi(p^t) + 1$. Therefore, we divide every such coefficient by p^{k-2}, and then take the remainder modulo p to get $\tilde{\alpha}_\ell^{(i)} := \left(\alpha_\ell^{(i)}/p^{k-2}\right) \bmod p$, where $d - \varphi(p^t) + 1 \le \ell \le d$. All these $\tilde{\alpha}_\ell^{(i)}$ forms the matrix

$$\boldsymbol{S} := \begin{bmatrix} \tilde{\alpha}_d^{(0)} & \cdots & \tilde{\alpha}_d^{(p^t-1)} \\ \vdots & \ddots & \vdots \\ \tilde{\alpha}_{d-\varphi(p^t)+1}^{(0)} & \cdots & \tilde{\alpha}_{d-\varphi(p^t)+1}^{(p^t-1)} \end{bmatrix} \in \mathbb{F}_p^{\varphi(p^t)\cdot p^t}.$$

Take its first $\varphi(p^t)$ columns to get a square matrix

$$\boldsymbol{S}' := \begin{bmatrix} \tilde{\alpha}_d^{(0)} & \cdots & \tilde{\alpha}_d^{(\varphi(p^t)-1)} \\ \vdots & \ddots & \vdots \\ \tilde{\alpha}_{d-\varphi(p^t)+1}^{(0)} & \cdots & \tilde{\alpha}_{d-\varphi(p^t)+1}^{(\varphi(p^t)-1)} \end{bmatrix}.$$

When $d-\varphi(p^t)+1 \le \ell \le d$, divide both sides of Equation (1) by p^{k-2} and take the remainder modulo p to get

$$\tilde{\alpha}_\ell^{(a)} = \sum_{j=0}^{d-\ell} \binom{-a}{j} \tilde{\alpha}_{j+\ell}^{(0)},$$

which leads to the following decomposition of matrix $\boldsymbol{S}'$:

$$\boldsymbol{S}' = \begin{bmatrix} \tilde{\alpha}_d^{(0)} & & \\ \vdots & \ddots & \\ \tilde{\alpha}_{d-\varphi(p^t)+1}^{(0)} & \cdots & \tilde{\alpha}_d^{(0)} \end{bmatrix} \cdot \begin{bmatrix} \binom{0}{0} & \cdots & \binom{1-\varphi(p^t)}{0} \\ \vdots & \ddots & \vdots \\ \binom{0}{\varphi(p^t)-1} & \cdots & \binom{1-\varphi(p^t)}{\varphi(p^t)-1} \end{bmatrix} =: \boldsymbol{T}\cdot\boldsymbol{C}.$$

As $\tilde{\alpha}_d^{(0)} \ne 0$, the first matrix has determinant $\det(\boldsymbol{T}) \ne 0$. The latter one, consisting of binomial coefficients, is also invertible. We will prove this fact later. Eventually, $\operatorname{rank}(\boldsymbol{S}') = \varphi(p^t)$, so the kernel of $\boldsymbol{S}$ has dimension $\dim\ker \boldsymbol{S} = (p^t) - \varphi(p^t) = p^{t-1}$.

On one hand, because $f(x)$ is p^t-periodic, we can expand it by $\{\mathsf{MOD}_n^{j,p^t}\}_{j=0}^{p^t-1}$ with coefficients w_j.

$$f(x) = \sum_{j=0}^{p^t-1} w_j \mathsf{MOD}_n^{j,p^t}(x) = \sum_{j=0}^{p^t-1} \left(w_j \sum_{\ell=0}^{d} \alpha_\ell^{(j)} \binom{|x|}{\ell} \right) = \sum_{\ell=0}^{d} \left(\left(\sum_{j=0}^{p^t-1} w_j \alpha_\ell^{(j)} \right) \binom{|x|}{\ell} \right). \quad (2)$$

On the other hand, $\mathsf{MOD}_n^{i,p^{t-1}}$ can also be spanned by $\{\mathsf{MOD}_n^{j,p^t}\}$. In other words, assume $w_j^{(i)} := \mathbb{I}[i \equiv j \pmod{p^{t-1}}]$, then

$$\mathsf{MOD}_n^{i,p^{t-1}} = \sum_{j=0}^{p^t-1} w_j^{(i)} \mathsf{MOD}_n^{j,p^t} = \sum_{\ell=0}^{d} \left(\left(\sum_{j=0}^{p^t-1} w_j^{(i)} \alpha_\ell^{(j)} \right) \binom{|x|}{\ell} \right). \quad (3)$$

We claim $\deg_{p^k}(\mathsf{MOD}_n^{i,p^{t-1}}) \le d - \varphi(p^t)$, and because $n \ge d$, the coefficients of highest $\varphi(p^t)$ terms in its Mahler expansion (i.e., from degree $d - \varphi(p^t) + 1$ to degree d) are all zero. This is due to Corollary 2.14, where we have

$$\begin{aligned}\deg_{p^k}(\mathsf{MOD}_n^{i,p^{t-1}}) &= (k-1)\varphi(p^{t-1}) + p^{t-1} - 1\\ &= p^{t-2}\left((k-1)(p-1)+p\right) - 1\\ &\le p^{t-2}\left((k-1)(p^2-p)+p\right) - 1\\ &= d - \varphi(p^t).\end{aligned}$$

Further, if we set column vector $\boldsymbol{w}^{(i)} = \left(w_0^{(i)}, ..., w_{p^t-1}^{(i)}\right)^\top$ where $0 \le i \le p^{t-1}-1$, then the above fact, together with Equation (3), indicates the following equation:

$$\boldsymbol{S}\boldsymbol{w}^{(i)} = \left[\sum_{j=0}^{p^t-1} w_j^{(i)}\tilde{\alpha}_d^{(j)}, \sum_{j=0}^{p^t-1} w_j^{(i)}\tilde{\alpha}_{d-1}^{(j)}, \cdots, \sum_{j=0}^{p^t-1} w_j^{(i)}\tilde{\alpha}_{d-\varphi(p^t)+1}^{(j)}\right]^\top = \mathbf{0}.$$

One can verify that $\left\{\boldsymbol{w}^{(0)}, \cdots, \boldsymbol{w}^{(p^{t-1}-1)}\right\}$ is linear independent, and therefore they form a base of the kernel of $\boldsymbol{S}$ i.e., $\ker \boldsymbol{S} = \operatorname{span}\left\{\boldsymbol{w}^{(0)}, \cdots, \boldsymbol{w}^{(p^{t-1}-1)}\right\}$.

However, $f(x)$ is not p^{t-1}-periodic because $\pi_p(f) = p^t$. This means $f(x)$ cannot be written as the sum of some $\mathsf{MOD}_n^{i,p^{t-1}}$. Thus $\boldsymbol{w} := (w_0, ..., w_{p^t-1})^\top \notin \operatorname{span}\left\{\boldsymbol{w}^{(0)}, \cdots, \boldsymbol{w}^{(p^{t-1}-1)}\right\}$ i.e., $\boldsymbol{S}\boldsymbol{w} \neq \mathbf{0}$. This leads to the existence of $D \in [d - \varphi(p^t) + 1, d]$ such that $\sum_{j=0}^{p^t-1} w_j^{(i)}\tilde{\alpha}_D^{(j)} \neq 0$. Namely, the degree D term in Equation (2) exists as desired. ◀

To finish the proof of this lemma, we show that the matrix $\boldsymbol{C}$ is invertible, due to the following proposition.

▶ **Proposition 3.5.** $\det(\boldsymbol{C}) = \pm 1$.

Proof. In fact,

$$\boldsymbol{C} = \operatorname{diag}\{1, -1, 1, -1, ...(-1)^{m-1}\} \cdot \begin{bmatrix} 1 & 1 & \binom{1}{1} & \cdots & \binom{m-2}{m-2} \\ 0 & \binom{1}{0} & \binom{2}{1} & \cdots & \binom{m-1}{m-2} \\ 0 & \binom{2}{0} & \binom{3}{1} & \cdots & \binom{m}{m-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \binom{m-1}{0} & \binom{m}{1} & \cdots & \binom{2m-3}{m-2} \end{bmatrix}$$

where $m := \varphi(p^t)$. Denote the second matrix as $\boldsymbol{C}'$. Take Row m of $\boldsymbol{C}'$ and subtract Row $m-1$ from it. Then take Row $m-1$ and subtract Row $m-2$ from it. $\cdots$ Take Row 3 and subtract Row 2 from it. Then C' has been transformed to

$$\begin{bmatrix} 1 & 1 & \binom{1}{1} & \cdots & \binom{m-2}{m-2} \\ 0 & \binom{1}{0} & \binom{2}{1} & \cdots & \binom{m-1}{m-2} \\ 0 & 0 & \binom{2}{0} & \cdots & \binom{m-1}{m-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \binom{m-1}{0} & \cdots & \binom{2m-4}{m-3} \end{bmatrix}.$$

Keep repeating this step, and $\boldsymbol{C}'$ will be transformed into an upper triangular matrix, whose diagonal elements are all 1. ◀

Now we are ready to prove Theorem 1.3.

Proof of Theorem 1.3. Note that Lemma 3.3 provides tight instances, so below we are going to prove the inequality.

Assume towards contradiction that there exists $f : \{0,1\}^n \to \{0,1\}$ satisfying $\deg_{p^k}(f) < (p-1)\cdot k$, and $n \geq (k-1)\cdot\varphi(p^\mu) + p^\mu - 1$ where $\mu = \lceil \log_p((p-1)k-1)\rceil$. Lemma 3.2 tells us that f is p^μ-periodic. The non-triviality of f indicates that there exists $1 \leq t \leq \mu$ such that $\pi_p(f) = p^t$.

If $t = 1$, then according to Lemma 3.3 we have $\deg_{p^k}(f) = (p-1)\cdot k$. If $t \geq 2$, then Lemma 3.4 indicates

$$\begin{aligned}\deg_{p^k}(f) &\geq (k-2)\cdot\varphi(p^t) + p^t \\ &\geq (k-2)\cdot\varphi(p^2) + p^2 \\ &= k\cdot(p-1)^2 + 2p - p^2 + (p-1)\cdot k \\ &\geq (p-1)^2 + 2p - p^2 + (p-1)\cdot k \\ &> (p-1)\cdot k.\end{aligned}$$

Both cases lead to contradiction. ◀

3.2 Non-degenerate Functions – Proof of Theorem 1.4

For the general non-degenerate case, our key idea is to embed a symmetric Boolean function into it, and then apply Theorem 1.3. The following lemma is crucial.

▶ **Lemma 3.6.** *There exists a monotone increasing function $r(n) = \omega(1)$ such that the following holds. Let $f : \{0,1\}^n \to \{0,1\}$ be a non-degenerate Boolean function. Then there exists a set of indices $S \subseteq [n]$ with $|S| \geq r(n)$, and a mapping $\sigma : [n]\backslash S \to \{0,1\}$ such that $f|_\sigma$ is a non-trivial symmetric Boolean function.*

Generally speaking, with this lemma in hand, every $h(n)$ lower bound on complexity measure of symmetric functions that is monotone decreasing w.r.t. restriction (e.g., Fact 2.5) can yield an $h(r(n))$ lower bound on that of all non-degenerate functions. In the setting of modulo-p^k degree, the bound $h(n)$ is a constant function (when n is large than some threshold), so we can get the same bound, except that the threshold for n blows up. However, as indicated by our proof, the function $r(n)$ grows extraordinary slow (approximately the square root of iterated logarithm of n).

First, let us see how to utilize this lemma in proving Theorem 1.4.

Proof of Theorem 1.4. As long as Lemma 3.6 holds, by Fact 2.5 and Theorem 1.3, if $n \geq r^{-1}((k-1)\cdot\varphi(p^\mu) + p^\mu - 1)$, then $\deg_{p^k}(f) \geq (p-1)\cdot k$, deriving the desired lower bound. In addition, any symmetric function with period p (and input long enough) still serves as an instance with $\deg_{p^k}(f) = (p-1)\cdot k$. ◀

In the rest part of this section we prove Lemma 3.6.

For convenience, we introduce the following notation. If $x_i = 1$ for every $i \in S \subseteq [n]$, then define $\mathrm{DOWN}(S, x, k) := \left\{x^{\oplus T} \mid T \subseteq S, |T| = k\right\}$. Intuitively speaking, it is the set of strings obtained by flipping k bits, whose indices are in S, of x from 1 to 0.

According to Theorem 2.7, there exists $\tilde{x}$ such that $s(f,\tilde{x}) = \Omega(\log n)$. Without loss of generality we assume the set $\{i \in [n] : \tilde{x}_i = 1, f(\tilde{x}) \neq f(\tilde{x}^{\oplus i})\}$, defined as S_0, is of cardinality $\Omega(\log n)$. Recursively define S_t to be the largest set satisfying the following two conditions:

- $S_t \subseteq S_{t-1}$.
- The value $f(y)$ are identical for any $y \in \mathrm{DOWN}(S_t, \tilde{x}, t)$.

We then make the following claim, and prove it.

▷ Claim 3.7.

$$|S_t| = \Omega\left(\log^{\circ(t-1)}(|S_{t-1}|)\right).$$

Our proof relies on Erdős–Rado Theorem [8] from Ramsey theory on hypergraphs.

▶ **Definition 3.8** (k-Uniform Hypergraph Ramsey Number)**.** *Suppose V is a set of vertices, and all size-k subsets of V forms $\mathcal{F}_k(V)$. If $E \subseteq \mathcal{F}_k(V)$, then we call (V, E) as a k-uniform hypergraph of order $|V|$. Naturally, we call $(V, \mathcal{F}_k(V))$ a complete k-uniform hypergraph.*

If the following property holds for complete k-uniform hypergraph of order M but not $M-1$, then $r_k(s,t) := M$ is called the k-uniform hypergraph Ramsey number: color every k-hyperedge red or blue arbitrarily, then there must exist a complete red hyper-subgraph of order s, or a complete blue hyper-subgraph of order t.

▶ **Theorem 3.9** ([8])**.** $r_2(s,t) \le \binom{s+t-2}{t-1}$ *and*

$$r_k(s,t) \le 2^{\binom{r_{k-1}(s-1,t-1)}{k-1}}, \quad (k > 2).$$

This theorem implies the following fact: if we color k-uniform hypergraph of order n with two colors, then there exists a monochromatic clique of size $\Omega(\log^{\circ k}(n))$.

Proof of Claim 3.7. Construct the following complete t-uniform hypergraph. The vertex set is S_{t-1}. For any $x \in \mathrm{DOWN}(S_{t-1}, \tilde{x}, t)$, we color the hyperedge $\{i : x_i \neq \tilde{x}_i\}$ with red if $f(x) = 1$, or blue otherwise. Take the largest monochromatic clique and suppose its vertex set is S. According to Theorem 3.9, $|S| = \Omega(\log^{\circ k}(n))$. Furthermore, the monochromaticity implies the value $f(z)$ is identical for any $z \in \mathrm{DOWN}(S, \tilde{x}, t)$. Hence S satisfies the desired conditions where our claim follows immediately. ◁

With the help of the claim and notations above, we can complete our proof of Lemma 3.6.

Proof of Lemma 3.6. By invoking Claim 3.7 iteratively,

$$|S_t| \ge Z \cdot \log^{\circ((t-1)t/2+1)}(n),$$

where Z is a positive constant irrelevant to n. Take $t' = t'(n) := \lfloor\sqrt{\log^*(n) - 2}\rfloor$. Then $(t'-1)t'/2 + 1 < \log^*(n)$ and $|S_{t'}| = \omega(1)$. In addition, define

$$r = r(n) := \min\left\{Z \cdot \log^{\circ(t'(t'-1)/2+1)}(n), t'(n)\right\} = \omega(1).$$

Then we have $r(n) \le |S_{r(n)}|$. This is because

- if $t'(n) \le Z \cdot \log^{\circ(t'(t'-1)/2+1)}(n)$, then we have $r(n) = t'(n) \le Z \cdot \log^{\circ(r(r-1)/2+1)}(n) \le |S_{r(n)}|$;
- if $t'(n) > Z \cdot \log^{\circ(t'(t'-1)/2+1)}(n)$, then $t'(n) > r(n)$, so $r(n) = Z \cdot \log^{\circ(t'(t'-1)/2+1)}(n) \le Z \cdot \log^{\circ(r(r-1)/2+1)}(n) \le |S_{r(n)}|$.

Now take a size $r(n)$ subset T of $S_{r(n)}$ arbitrarily. Define the mapping $\sigma : [n] \backslash T \to \{0,1\}$ such that $\sigma(i) = \tilde{x}_i$. Restrict f over σ to obtain a new function $g := f|_\sigma$. We will prove g is symmetric and non-trivial, and then the lemma follows immediately.

Symmetric. Assume $x, y \in \{0,1\}^{r(n)}$ such that $|x| = |y|$. Define x' (resp. y') to be the string of size n obtained from x (resp. y) and σ. Recall the definition of $S_{r(n)}$. That is, for any $i \in S_{r(n)}$, it holds that $\tilde{x}_i = 1$. Therefore,

$$x', y' \in \mathrm{DOWN}(S_{r(n)}, \tilde{x}, |x|) \subseteq \mathrm{DOWN}(S_{|x|}, \tilde{x}, |x|).$$

By definition of DOWN, we have $f(x') = f(y')$ i.e., $g(x) = g(y)$.

Non-trivial. Assume $z, w \in \{0,1\}^{r(n)}$ such that $|z| = 0$ and $|w| = 1$. We define z' and w' similarly to x' and y'. Suppose $z' = w' \oplus e_i$. As $i \in T \subseteq S_0$, the i-th bit is sensitive. Therefore, $f(z') \neq f(w')$ i.e., $g(z) \neq g(w)$. ◀

4 Lower Bounds of $\deg_{pq}(f)$ for Symmetric Functions

We first go further with the analysis of Mahler coefficients of MOD functions, then prove Theorem 1.5 in Section 4.2, and give an instance in Section 4.3, showing one can never improve the constant factor $1/2$.

4.1 More Analyses of MOD and Its Mahler Coefficients

Let p be a prime and t be a positive integer. Consider over the ring $\mathbb{Z}_p$. Recall the notation $\alpha_\ell^{(a,p^t)}$: it is the ℓ-th Mahler coefficient of MOD_n^{a,p^t}. Since $\deg_p(\mathsf{MOD}_n^{a,p^t}) = p^t - 1$, the following $p^t \times p^t$ matrix collects all the coefficients of MOD_n^{a,p^t}:

$$\boldsymbol{A}_{p^t} := \begin{bmatrix} \alpha_0^{(0,p^t)} & \cdots & \alpha_0^{(p^t-1,p^t)} \\ \vdots & \ddots & \vdots \\ \alpha_{p^t-1}^{(0,p^t)} & \cdots & \alpha_{p^t-1}^{(p^t-1,p^t)} \end{bmatrix}.$$

In fact,

$$(\boldsymbol{A}_{p^t})_{i,j} = \alpha_i^{(j,p^t)} = \binom{p^t - 1 - j}{p^t - 1 - i}.$$

This is because

$$\sum_{i=0}^{p^t-1} \binom{p^t-1-j}{p^t-1-i}\binom{|x|}{i} = \binom{p^t-1-j+|x|}{p^t-1} = \begin{cases} 1 & \text{if } |x| \equiv j \pmod{p^t}, \\ 0 & \text{otherwise,} \end{cases} \tag{4}$$

by Vandermonde's convolution.

The matrix $\boldsymbol{A}_{p^t}$ has many elegant properties. For example, the following one shows the relationship between $\boldsymbol{A}_{p^t}$ and $\boldsymbol{A}_p$. We use $\otimes$ to denote matrix tensor product.

▶ **Proposition 4.1.** *On the ring $\mathbb{Z}_p$,*

$$\boldsymbol{A}_{p^t} = \boldsymbol{A}_p^{\otimes t} := \underbrace{\boldsymbol{A}_p \otimes \boldsymbol{A}_p \otimes \cdots \otimes \boldsymbol{A}_p}_{t}.$$

Proof. Let i_ℓ and j_ℓ be the representation of i and j in base p. Then by Lucas's Theorem,

$$\binom{p^t-1-j}{p^t-1-i} \equiv \binom{\sum_{\ell=0}^{t-1}(p-1-j_\ell)\cdot p^\ell}{\sum_{\ell=0}^{t-1}(p-1-i_\ell)\cdot p^\ell} \equiv \prod_{\ell=0}^{t-1}\binom{p-1-j_\ell}{p-1-i_\ell} \equiv \prod_{\ell=0}^{t-1}(\boldsymbol{A}_p)_{i_\ell,j_\ell} \pmod{p}.$$ ◀

Below we give another observation, which assists with our proof of Theorem 1.5.

▶ **Lemma 4.2.** *Suppose p is a prime, and $n < p-1$ is a positive integer. Then for any $v \in \{0,1\}^p$ satisfying $v_i \neq v_j$ for some $0 \leq i < j \leq n$, there exists $\lfloor n/2 \rfloor + 1 \leq \ell \leq n$ such that $(\boldsymbol{A}_p \boldsymbol{v})_\ell \neq 0$.*

Our proof of Lemma 4.2 utilizes the following proposition on another binomial coefficient matrix.

▶ **Proposition 4.3.** *For any prime p, integers j, k with $j+k < p$ and distinct $a_0, \ldots, a_k \in \mathbb{F}_p$ satisfying $a_0, \ldots, a_k \geq j$, the matrix*

$$\begin{bmatrix} \binom{a_0}{j} & \binom{a_1}{j} & \cdots & \binom{a_k}{j} \\ \binom{a_0}{j+1} & \binom{a_1}{j+1} & \cdots & \binom{a_k}{j+1} \\ \vdots & \vdots & \ddots & \vdots \\ \binom{a_0}{j+k} & \binom{a_1}{j+k} & \cdots & \binom{a_k}{j+k} \end{bmatrix}$$

is invertible over $\mathbb{F}_p$.

Proof. One can verify that

$$\operatorname{diag}\left\{ \frac{(j+0)^{\underline{0}}}{\binom{a_0}{j}}, \ldots, \frac{(j+k)^{\underline{k}}}{\binom{a_k}{j}} \right\} \cdot S \cdot \begin{bmatrix} \binom{a_0}{j} & \cdots & \binom{a_k}{j} \\ \vdots & \ddots & \vdots \\ \binom{a_0}{j+k} & \cdots & \binom{a_k}{j+k} \end{bmatrix} = \begin{bmatrix} (a_0-j)^0 & \cdots & (a_k-j)^0 \\ \vdots & \ddots & \vdots \\ (a_0-j)^k & \cdots & (a_k-j)^k \end{bmatrix},$$

where S is the second Stirling number matrix, i.e., $S_{ij} = \left\{ {i \atop j} \right\}$, and the notation $x^{\underline{y}}$ stands for the falling factorial power $x(x-1)\cdots(x-y+1)$. The Vandermonde matrix on the R.H.S. is also invertible because $a_0, \ldots, a_k$ are distinct. ◀

Proof of Lemma 4.2. Assume towards contradiction that there exists some v satisfying the condition, but $(\boldsymbol{A}_p v)_\ell = 0$ for all $\lfloor n/2 \rfloor + 1 \leq \ell \leq n$. In other words, if we take row $\lfloor n/2 \rfloor + 1$ to n and column 0 to n from $\boldsymbol{A}_p$ to get another $\lceil n/2 \rceil \times (n+1)$ matrix $\boldsymbol{B}$ i.e.,

$$\boldsymbol{B} := \begin{bmatrix} \binom{p-1-0}{p-1-(\lfloor n/2 \rfloor+1)} & \binom{p-1-1}{p-1-(\lfloor n/2 \rfloor+1)} & \cdots & \binom{p-1-n}{p-1-(\lfloor n/2 \rfloor+1)} \\ \vdots & \vdots & \ddots & \vdots \\ \binom{p-1-0}{p-1-n} & \binom{p-1-1}{p-1-n} & \cdots & \binom{p-1-n}{p-1-n} \end{bmatrix},$$

then $\boldsymbol{B}\boldsymbol{v}' = \boldsymbol{0}$ where $\boldsymbol{v}' = \{v_0, ..., v_n\}^T$. (This is because $(\boldsymbol{A}_p)_{i,j} = \binom{p-1-j}{p-1-i} = 0$ for all $\lfloor n/2 \rfloor + 1 \leq i \leq n$ and $n+1 \leq j \leq p-1$.)

Next, for any $t \in [(\lfloor n/2 \rfloor + 1), n]$, the sum of row t is $\binom{p-1-0}{p-1-t} + \binom{p-1-1}{p-1-t} + \cdots + \binom{p-1-n}{p-1-t} = \binom{p}{p-t} \equiv 0 \pmod p$. Therefore $\boldsymbol{B}\boldsymbol{1} = \boldsymbol{0}$, so we can assume the number of 1's in $\boldsymbol{v}'$ is no more that $\lceil n/2 \rceil$, without loss of generality. (Otherwise, subtract $\boldsymbol{v}'$ from $\boldsymbol{1}$.) This means that we can take $s \leq \lceil n/2 \rceil$ column vectors of $\boldsymbol{B}$, the summation of which is $\boldsymbol{0}$, and furthermore, the last s dimensions of these vectors form a singular matrix with form $\boldsymbol{B}'_{i,j} = \binom{a_j}{p-1-n+s-1+i}$. However, by flipping it upside down and applying Proposition 4.3, this matrix is invertible. ◀

4.2 Proof of Theorem 1.5

Our proof requires the following two lemmas. The first one is often referred to as Periodicity Lemma. It says any function with coprime periods is constant, if the domain is large enough.

▶ **Lemma 4.4** (Periodicity Lemma, [9]). *Let g be an a-periodic and b-periodic function on domain $\{0,1,\ldots,n\}$ with $gcd(a,b)=1$ and $n \geq a+b-2$. Then g is a constant function.*

The next one can be regarded as a stronger version of Theorem 1.3 with $k=1$.

▶ **Lemma 4.5.** *Assume p is a prime. For any non-trivial symmetric $f:\{0,1\}^n \to \{0,1\}$,*

$$\deg_p(f) \geq \min\left\{\frac{n}{2}, \left(1-\frac{1}{p}\right)\pi_p(f)\right\}.$$

Note that the base-p period appear explicitly in the lower bound. This allows us to apply Lemma 4.4. We prove Theorem 1.5 first.

Proof of Theorem 1.5. If $\max\{\deg_p(f), \deg_q(f)\} \geq \frac{n}{2}$, the theorem then follows naturally. Otherwise, according to Fact 2.3 and Lemma 4.5, we have

$$\deg_m(f) \geq \deg_{pq}(f) = \max\{\deg_p(f), \deg_q(f)\} \geq \max\left\{\left(1-\frac{1}{p}\right)\pi_p(f), \left(1-\frac{1}{q}\right)\pi_q(f)\right\}. \tag{5}$$

On the other hand, the non-triviality of $f(x)$ implies $\pi_p(f)+\pi_q(f) > n+2$ owing to Lemma 4.4. The last term in Inequality (5) is $> \frac{n+2}{2+1/(p-1)+1/(q-1)} > \frac{n}{2+1/(p-1)+1/(q-1)}$, and hence the theorem is also true. ◀

It remains to show why Lemma 4.5 is true.

Proof of Lemma 4.5. Consider over the ring $\mathbb{Z}_p$. Suppose $\pi_p(f)=p^t$. We write $f(x)$ as we did in Equation (2), and let α_ℓ be the ℓ-th Mahler coefficient of $f(x)$. Then $\sum_{j=0}^{p^t-1} w_j \alpha_\ell^{(j,p^t)} = \alpha_\ell$, or

$$\boldsymbol{\alpha} = \boldsymbol{A}_{p^t}\boldsymbol{w} \tag{6}$$

if we set $\boldsymbol{w} := (w_0,\ldots,w_{p^t-1})^\top$ and $\boldsymbol{\alpha} := (\alpha_0,\ldots,\alpha_{p^t-1})^\top$.

Divide $\boldsymbol{w}$ and $\boldsymbol{\alpha}$ into blocks of length p^{t-1} as $\boldsymbol{w} = (\boldsymbol{w}^{\langle 0\rangle},\ldots,\boldsymbol{w}^{\langle p-1\rangle})^\top$ and $\boldsymbol{\alpha} = (\boldsymbol{\alpha}^{\langle 0\rangle},\ldots,\boldsymbol{\alpha}^{\langle 0\rangle})^\top$ where $\boldsymbol{w}^{\langle i\rangle} \in \{0,1\}^{p^{t-1}}, \boldsymbol{\alpha}^{\langle i\rangle} \in \mathbb{F}_p^{p^{t-1}}$. By Proposition 4.1, we have

$$\boldsymbol{\alpha}^{\langle i\rangle} = \boldsymbol{A}_{p^{t-1}} \sum_{j=0}^{p-1}\left((\boldsymbol{A}_p)_{ij}\boldsymbol{w}^{\langle j\rangle}\right). \tag{7}$$

Consider two cases. One deals with the case $\pi_p(f)=p^t<n$, where we show $\deg_p(f) > \frac{p-1}{p}\cdot\pi_p(f)$; another deals with $\pi_p(f)\geq n$, where we can obtain $\deg_p(f)\geq n/2$.

Case I ($p^t < n$). First, assume $\boldsymbol{\alpha}^{\langle p-1\rangle} = \boldsymbol{0}$. Note that $\boldsymbol{A}_{p^{t-1}}$ is full-rank according to Proposition 4.3, which allows Equation (7) to be transformed into

$$\boldsymbol{A}_{p^{t-1}}^{-1}\boldsymbol{\alpha}^{\langle p-1\rangle} = \sum_{j=0}^{p-1}\left((\boldsymbol{A}_p)_{p-1,j}\boldsymbol{w}^{\langle j\rangle}\right).$$

Since $(\boldsymbol{A}_p)_{p-1,j}=1$, we have $\sum_{j=0}^{p-1}\boldsymbol{w}^{\langle j\rangle}=\boldsymbol{0}$. This implies $\boldsymbol{w}^{\langle 0\rangle}=\cdots=\boldsymbol{w}^{\langle p-1\rangle}$ as $\boldsymbol{w}^{\langle i\rangle}\in\{0,1\}^{p^{t-1}}$, and further, $f(x)$ becomes p^{t-1}-periodic, conflicting with $\pi_p(f)=p^t$. Eventually, we have $\boldsymbol{\alpha}^{\langle p-1\rangle}\neq\boldsymbol{0}$. Because $n>p^t$, the highest non-zero Mahler coefficient indicates the degree of f (see the remark below Fact 2.10), and then $\deg_p(f) > (p-1)p^{t-1} = \frac{p-1}{p}\cdot\pi_p(f)$.

Case II ($p^t \geq n$). By Equation (6),

$$\left(\boldsymbol{I}_p \otimes (\boldsymbol{A}_{p^{t-1}})^{-1}\right) \boldsymbol{A}_{p^t} \boldsymbol{w} = \left(\boldsymbol{I}_p \otimes (\boldsymbol{A}_{p^{t-1}})^{-1}\right) \boldsymbol{\alpha}. \tag{8}$$

The R.H.S. of (8) is just

$$\left(\boldsymbol{I}_p \otimes (\boldsymbol{A}_{p^{t-1}})^{-1}\right) \boldsymbol{\alpha} = \begin{bmatrix} (\boldsymbol{A}_{p^{t-1}})^{-1}\boldsymbol{\alpha}^{\langle 0\rangle} \\ \vdots \\ (\boldsymbol{A}_{p^{t-1}})^{-1}\boldsymbol{\alpha}^{\langle p-1\rangle} \end{bmatrix} =: \begin{bmatrix} \boldsymbol{\beta}^{\langle 0\rangle} \\ \vdots \\ \boldsymbol{\beta}^{\langle p-1\rangle} \end{bmatrix}.$$

The L.H.S. of (8) can be written as

$$\begin{aligned} \left(\boldsymbol{I}_p \otimes (\boldsymbol{A}_{p^{t-1}})^{-1}\right) \boldsymbol{A}_{p^t} \boldsymbol{w} &= \left(\boldsymbol{I}_p \otimes (\boldsymbol{A}_{p^{t-1}})^{-1}\right) (\boldsymbol{A}_p \otimes \boldsymbol{A}_{p^{t-1}}) \boldsymbol{w} \\ &= (\boldsymbol{I}_p \boldsymbol{A}_p) \otimes ((\boldsymbol{A}_{p^{t-1}})^{-1} \boldsymbol{A}_{p^{t-1}}) \boldsymbol{w} \\ &= (\boldsymbol{A}_p \otimes \boldsymbol{I}_{p^{t-1}}) \boldsymbol{w}. \end{aligned}$$

Therefore,

$$(\boldsymbol{A}_p \otimes \boldsymbol{I}_{p^{t-1}}) \boldsymbol{w} = \left(\boldsymbol{\beta}^{\langle 0\rangle}, \cdots, \boldsymbol{\beta}^{\langle p-1\rangle}\right)^\top. \tag{9}$$

For $0 \leq j < p^{t-1}$ we define

$$\tilde{\boldsymbol{\beta}}^{\langle j\rangle} := \left(\boldsymbol{\beta}_j^{\langle 0\rangle}, \cdots, \boldsymbol{\beta}_j^{\langle p-1\rangle}\right)^\top \text{ and } \tilde{\boldsymbol{w}}^{\langle j\rangle} := \left(\boldsymbol{w}_j^{\langle 0\rangle}, \cdots, \boldsymbol{w}_j^{\langle p-1\rangle}\right)^\top,$$

Intuitively, vectors with tildes here contain entries taken from the original vector with stride p^{t-1}. Then Equation (9) implies

$$\boldsymbol{A}_p \tilde{\boldsymbol{w}}^{\langle j\rangle} = \tilde{\boldsymbol{\beta}}^{\langle j\rangle}.$$

Let $n' = \lfloor (n+1)/p^{t-1} \rfloor - 1$, $n'' = \lceil (n+1)/p^{t-1} \rceil - 1$, and $m' = n \bmod p^{t-1}$. Consider the following two subcases:

Subcase II-1. Suppose there exists $\ell \leq m'$ and $i, j \in [0, n'']$ such that $\tilde{\boldsymbol{w}}_i^{\langle \ell\rangle} \neq \tilde{\boldsymbol{w}}_j^{\langle \ell\rangle}$. According to Lemma 4.2, there exists $i' \in [\lfloor n''/2 \rfloor + 1, n'']$ satisfying $0 \neq \left(\boldsymbol{A}_p \tilde{\boldsymbol{w}}^{\langle j\rangle}\right)_{i'} = \tilde{\boldsymbol{\beta}}_{i'}^{\langle \ell\rangle} = \boldsymbol{\beta}_\ell^{\langle i'\rangle}$. Because $\boldsymbol{A}_{p^{t-1}}$ is invertible, we have $\boldsymbol{\alpha}^{\langle i'\rangle} = \boldsymbol{A}_{p^{t-1}} \boldsymbol{\beta}^{\langle i'\rangle} \neq \boldsymbol{0}$. What's more,

- if $i' < n''$ and recall Fact 2.10, we have $\deg_p(f) \geq (\lfloor n''/2 \rfloor + 1) \cdot p^{t-1} \geq n/2$;
- if $i' = n''$, we select the minimum ℓ such that $\boldsymbol{\beta}_\ell^{\langle i'\rangle} \neq 0$. Due to the fact $(\boldsymbol{A}_{p^{t-1}})_{\ell,j} = \binom{p^{t-1}-1-j}{p^{t-1}-1-\ell} = 0$ when $j > \ell$, it follows that

$$\boldsymbol{\alpha}_\ell^{\langle i'\rangle} = \sum_{j=0}^{\ell} (\boldsymbol{A}_{p^{t-1}})_{\ell,j} \boldsymbol{\beta}_j^{\langle i'\rangle} = (\boldsymbol{A}_{p^{t-1}})_{\ell,\ell} \boldsymbol{\beta}_\ell^{\langle i'\rangle} \neq 0. \tag{10}$$

Eventually $\deg_p(f) \geq n'' \cdot p^{t-1} \geq n/2$.

Subcase II-2. Otherwise, there exists a minimum $\ell \in [m'+1, p^{t-1}-1]$ and $i, j \in [0, n']$ such that $\tilde{\boldsymbol{w}}_i^{\langle \ell\rangle} \neq \tilde{\boldsymbol{w}}_j^{\langle \ell\rangle}$ as $f(x)$ is non-trivial. The same argument shows that there exists $i' \in [\lfloor n'/2 \rfloor + 1, n']$ satisfying $\boldsymbol{\beta}_\ell^{\langle i'\rangle} \neq 0$. In addition, the condition $\tilde{\boldsymbol{w}}_0^{\langle \ell'\rangle} = \cdots = \tilde{\boldsymbol{w}}_{n'}^{\langle \ell'\rangle}$ for all $\ell' < \ell$ implies

$$\boldsymbol{\beta}_{\ell'}^{\langle i'\rangle} = \tilde{\boldsymbol{\beta}}_{i'}^{\langle \ell'\rangle} = \sum_{j=0}^{p-1} (\boldsymbol{A}_p)_{i',j} \tilde{\boldsymbol{w}}_j^{\langle \ell'\rangle} = \boldsymbol{w}_0^{\langle \ell'\rangle} \cdot \sum_{j=0}^{i'} \binom{p-1-j}{p-1-i'} = \boldsymbol{w}_0^{\langle \ell'\rangle} \cdot \binom{p}{p-i'} = 0.$$

Hence, by imitating (10) we can obtain $\boldsymbol{\alpha}_\ell^{\langle i'\rangle} \neq 0$, which leads to $\deg_p(f) \geq \lfloor n'/2 \rfloor \cdot p^{t-1} + m' + 1 \geq n/2$. ◀

4.3 Proof of Theorem 1.6

We will later apply the following lemma about Diophantine approximation, which is an immediate corollary of Kronecker's Theorem.

▶ **Lemma 4.6.** *Suppose real numbers $a_1, \ldots, a_k$ satisfy that $1, a_1, \ldots, a_k$ are linearly independent over $\mathbb{Q}$. Then, for any $\varepsilon > 0$, there exist infinitely many positive integers ℓ such that $\ell a_i \bmod 1 \in (1-\varepsilon, 1)$ for each $i = 1, \ldots, k$.*

Now we prove Theorem 1.6.

Proof of Theorem 1.6. Write $m = p_1 p_2 \cdots p_k$ for p_i being primes. Choose a prime q different from all p_i. Fix an arbitrary $\varepsilon > 0$. Let $a_i = \log q / \log p_i$ for $i = 1, \ldots, k$. Then $1, a_1, \ldots, a_k$ are linearly independent over $\mathbb{Q}$, otherwise a nontrivial linear relation can be exponentiated to contradict the unique factorization theorem over $\mathbb{Z}^+$. Applying Lemma 4.6 we get infinitely many ℓ that satisfy the condition $\ell \cdot \log q / \log p_i \bmod 1 \in (1-\varepsilon, 1)$, which implies $p_i^{r_i}/q^\ell \in (1, p_i^\varepsilon)$ where $r_i = \lceil \ell \log q / \log p_i \rceil$.

Now, choose a sufficiently large ℓ, let $n = 2q^\ell$ and define $f : \{0,1\}^n \to \{0,1\}$ by $f(x) = \mathbb{I}[|x| = q^\ell]$. Then f is $p_i^{r_i}$-periodic since $p_i^{r_i} > q^\ell$. Therefore $\deg_{p_i}(f) \le p_i^{r_i} - 1$ by Theorem 2.12. Thus,

$$\deg_m(f) \le \max_{1 \le i \le k} \{p_i^{r_i}\} \le \frac{n}{2} \max_{1 \le i \le k} \{p_i^\varepsilon\}.$$

The theorem follows by letting $\varepsilon \to 0$. ◀

5 Conclusion

In a nutshell, we explore and exploit the matrices consisting of Mahler coefficients of the MOD function, serving as a significant extension of Wilson's arguments. This approach fully characterizes the modulo degree of Boolean functions when the base is prime or prime power, and provides good lower bounds for the composite case with the help of periodicity lemma. In addition, we also show a practical way to generalize properties of symmetric functions to non-degenerate ones by a Ramsey-type argument.

Nevertheless, there is still ample room for further discussion. First and foremost, we conjecture that the constant factor in Theorem 1.5 can be improved to $1/2$ in correspondence with Theorem 1.6. Moreover, an anonymous reviewer also raises a good question with regard to Theorem 1.4: Could the extraordinary large prerequisite $n \ge \text{tower}(\text{poly}(p,k))$ (which is implicit in the proof) be improved to something like $n \ge \exp(\text{poly}(p,k))$? We also wonder if it is possible to embed other kinds of functions to derive similar results. Above all, both Conjecture 1.1 and Conjecture 1.2 remain open.

References

1 Noga Alon and Richard Beigel. Lower bounds for approximations by low degree polynomials over $\mathbb{Z}_m$. In *Proceedings of the 16th Annual IEEE Conference on Computational Complexity, Chicago, Illinois, USA, June 18-21, 2001*, pages 184–187. IEEE Computer Society, 2001. doi:10.1109/CCC.2001.933885.

2 Roger C Baker, Glyn Harman, and János Pintz. The difference between consecutive primes, ii. *Proceedings of the London Mathematical Society*, 83(3):532–562, 2001.

3 David A. Mix Barrington, Richard Beigel, and Steven Rudich. Representing boolean functions as polynomials modulo composite numbers. *Computational Complexity*, 4:367–382, 1994. doi:10.1007/BF01263424.

4 Harry Buhrman and Ronald de Wolf. Communication complexity lower bounds by polynomials. In *Proceedings of the 16th Annual IEEE Conference on Computational Complexity, Chicago, Illinois, USA, June 18-21, 2001*, pages 120–130. IEEE Computer Society, 2001. doi:10.1109/CCC.2001.933879.

5 Harry Buhrman and Ronald de Wolf. Complexity measures and decision tree complexity: a survey. *Theor. Comput. Sci.*, 288(1):21–43, 2002. doi:10.1016/S0304-3975(01)00144-X.

6 John Chiarelli, Pooya Hatami, and Michael E. Saks. An asymptotically tight bound on the number of relevant variables in a bounded degree boolean function. *Combinatorica*, 40(1):237–244, 2020. doi:10.1007/s00493-019-4136-7.

7 Klim Efremenko. 3-query locally decodable codes of subexponential length. *SIAM J. Comput.*, 41(6):1694–1703, 2012. doi:10.1137/090772721.

8 P. Erdős and R. Rado. Combinatorial theorems on classifications of subsets of a given set. *Proceedings of the London Mathematical Society*, s3-2(1):417–439, 1952.

9 Nathan J Fine and Herbert S Wilf. Uniqueness theorems for periodic functions. *Proceedings of the American Mathematical Society*, 16(1):109–114, 1965.

10 Parikshit Gopalan. *Computing with polynomials over composites*. PhD thesis, Georgia Institute of Technology, 2006.

11 Parikshit Gopalan. Constructing ramsey graphs from boolean function representations. *Combinatorica*, 34(2):173–206, 2014. doi:10.1007/s00493-014-2367-1.

12 Parikshit Gopalan, Shachar Lovett, and Amir Shpilka. On the complexity of boolean functions in different characteristics. In *Proceedings of the 24th Annual IEEE Conference on Computational Complexity, CCC 2009, Paris, France, 15-18 July 2009*, pages 173–183. IEEE Computer Society, 2009. doi:10.1109/CCC.2009.14.

13 Vince Grolmusz. Superpolynomial size set-systems with restricted intersections mod 6 and explicit ramsey graphs. *Combinatorica*, 20(1):71–86, 2000. doi:10.1007/s004930070032.

14 Vince Grolmusz. Constructing set systems with prescribed intersection sizes. *J. Algorithms*, 44(2):321–337, 2002. doi:10.1016/S0196-6774(02)00204-3.

15 D. J. Katz. p-adic valuation of weights in abelian codes over $\mathbb{Z}_{p^d}$. *IEEE Trans. Inf. Theory*, 51(1):281–305, 2005. doi:10.1109/TIT.2004.839495.

16 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:10.1016/j.jcss.2003.07.007.

17 Eyal Kushilevitz and Yishay Mansour. Learning decision trees using the fourier spectrum. *SIAM J. Comput.*, 22(6):1331–1348, 1993. doi:10.1137/0222080.

18 Chia-Jung Lee, Satyanarayana V. Lokam, Shi-Chun Tsai, and Ming-Chuan Yang. Restrictions of nondegenerate boolean functions and degree lower bounds over different rings. In *IEEE International Symposium on Information Theory, ISIT 2015, Hong Kong, China, June 14-19, 2015*, pages 501–505. IEEE, 2015. doi:10.1109/ISIT.2015.7282505.

19 Qian Li and Xiaoming Sun. On the modulo degree complexity of boolean functions. *Theor. Comput. Sci.*, 818:32–40, 2020. doi:10.1016/j.tcs.2018.04.049.

20 Nathan Linial, Yishay Mansour, and Noam Nisan. Constant depth circuits, fourier transform, and learnability. *J. ACM*, 40(3):607–620, 1993. doi:10.1145/174130.174138.

21 Xiaoyu Liu. An equivalence of ward's bound and its application. *Des. Codes Cryptogr.*, 58(1):1–9, 2011. doi:10.1007/s10623-010-9380-1.

22 Kurt Mahler. An interpolation series for continuous functions of a p-adic variable. *J. reine angew. Math*, 199:23–34, 1958.

23 Marvin Minsky and Seymour Papert. An introduction to computational geometry. *Cambridge tiass., HIT*, 1969.

24 Ashley Montanaro and Tobias Osborne. On the communication complexity of XOR functions. *CoRR*, abs/0909.3392, 2009. arXiv:0909.3392.

25 Elchanan Mossel, Ryan O'Donnell, and Rocco A. Servedio. Learning juntas. In Lawrence L. Larmore and Michel X. Goemans, editors, *Proceedings of the 35th Annual ACM Symposium on Theory of Computing, June 9-11, 2003, San Diego, CA, USA*, pages 206–212. ACM, 2003. doi:10.1145/780542.780574.

26 Noam Nisan and Mario Szegedy. On the degree of boolean functions as real polynomials. *Computational Complexity*, 4:301–313, 1994. doi:10.1007/BF01263419.

27 Alexander A Razborov. Lower bounds for the size of circuits of bounded depth with basis $\{\wedge, \oplus\}$. *Math. notes of the Academy of Sciences of the USSR*, 41(4):333–338, 1987.

28 Alexander A Razborov. Quantum communication complexity of symmetric predicates. *Izvestiya: Mathematics*, 67(1):145, 2003.

29 Alexander A. Sherstov. Communication lower bounds using dual polynomials. *Bulletin of the EATCS*, 95:59–93, 2008.

30 Alexander A. Sherstov. Strong direct product theorems for quantum communication and query complexity. *SIAM J. Comput.*, 41(5):1122–1165, 2012. doi:10.1137/110842661.

31 Yaoyun Shi and Yufan Zhu. Quantum communication complexity of block-composed functions. *Quantum Information & Computation*, 9(5):444–460, 2009. URL: http://www.rintonpress.com/xxqic9/qic-9-56/0444-0460.pdf.

32 Hans Ulrich Simon. A tight $\Omega(\log \log n)$-bound on the time for parallel RAM's to compute nondegenerated boolean functions. *Information and Control*, 55(1-3):102–106, 1982. doi:10.1016/S0019-9958(82)90477-6.

33 Roman Smolensky. Algebraic methods in the theory of lower bounds for boolean circuit complexity. In Alfred V. Aho, editor, *Proceedings of the 19th Annual ACM Symposium on Theory of Computing, 1987, New York, New York, USA*, pages 77–82. ACM, 1987. doi:10.1145/28395.28404.

34 Joachim von zur Gathen and James R. Roche. Polynomials with two values. *Combinatorica*, 17(3):345–362, 1997. doi:10.1007/BF01215917.

35 Jake Wellens. A tighter bound on the number of relevant variables in a bounded degree boolean function. *CoRR*, abs/1903.08214, 2019. arXiv:1903.08214.

36 Richard M. Wilson. A lemma on polynomials modulo p^m and applications to coding theory. *Discrete Mathematics*, 306(23):3154–3165, 2006. doi:10.1016/j.disc.2004.10.030.

37 Bahattin Yildiz. Weights modulo p^e of linear codes over rings. *Des. Codes Cryptogr.*, 43(2-3):147–165, 2007. doi:10.1007/s10623-007-9076-3.