Complexity of Boolean Functions

Om Swostik
Guide: Srikanth Srinivasan

Autumn 2024

Contents

1 Introduction

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 many areas, including communication complexity, circuit complexity, learning theory, coding theory, etc. In this report, we summarize various results on the degree and other complexity measures of Boolean functions and also outline some open problems and conjectures in this area.

2 Preliminaries

We denote $\{1, 2, \ldots, n\}$ as $[n]$ throughout this report. 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 $x^{\oplus S}$. In what follows, we will abbreviate $x^{\oplus\{i\}}$ as $x^{\oplus i}$.

2.1 Basics of Boolean Functions

An n-bit Boolean function f is a mapping from $\{0,1\}^n$ to $\{0,1\}$. Here's a list of common subclasses of Boolean functions:

- 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.
- 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 $\mathbb{Z}$ or $\mathbb{Z}_m$. More formally,

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

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

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

Let $P_m(x)$ denote the mod m polynomial representation of a Boolean function f.

Definition 2.1.1. *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., $\mathbb{Z}_m$).*

Fact. *For any Boolean function f , we have $\deg(f) \geq \deg_m(f)$ for all m. Similarly $\deg_m(f) \geq \deg_{m'}(f)$ if $m' \mid m$.*

The above fact implies $\deg_m(f) \leq \deg_{m^k}(f)$. The following fact shows that they are always within a factor $2k-1$ of each other, for a proof refer to section 2 of [6].

Fact. *For any Boolean function f , and any integers $m \geq 2, k \geq 1$, we have*

$$\deg_m(f) \leq \deg_{m^k}(f) \leq (2k-1)\deg_m(f).$$

Let m be prime, consider the function $f(x) = (x_1 + \cdots + x_n)^{m-1} \mod m$ with $\deg_m(f) \le m-1$ and $\deg(f) = \Omega(n)$. Such functions also exist for powers of primes.

Fact. *For any prime power m, there exists a sequence of functions f with n variables such that $\deg_m(f) = O(1)$ and $\deg(f) = \Omega(n)$.*

The following fact is a consequence of Chinese Remainder Theorem,

Fact. *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 2.1.2 (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 $\sigma : [n] \setminus S \to \{0,1\}$. 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. *Suppose $f : \{0,1\}^n \to \{0,1\}$ is a Boolean function. For any integer $m \ge 2$ and restriction $f|_\sigma$, we have $\deg_m(f) \ge \deg_m(f|_\sigma)$.*

Definition 2.1.3 (Projection)**.** *A projection from n variables to m variables is a map $\nu : [n] \to [m]$. Given a function $f : \{0,1\}^n \to \{0,1\}$ and a projection ν from n variables to m variables, we get a function $f|_\nu : \{0,1\}^m \to \{0,1\}$ by identifying variables of f that map to the same image under ν.*

2.2 The Method of Symmetrization

We will state the method of symmetrization, first used by Minsky and Papert [13]. Let R denote any commutative ring,

Definition 2.2.1. *If $p : R^n \to R$ is a multivariate polynomial, then the symmetrization of p is defined as follows:*

$$p^{sym}(x_1, \ldots, x_n) = \frac{\sum_{\pi \in S_n} p(x_{\pi(1)}, \ldots, x_{\pi(n)})}{n!}.$$

The important point is that if we are only interested in inputs $x \in \{0,1\}^n$, then $p^{\text{sym}}(x)$ only depends upon $x_1 + \cdots + x_n$, i.e. only on the Hamming weight of x. Based on this, we can represent it as an univariate polynomial of $x_1 + \cdots + x_n$,

Lemma 2.2.2. *If $p : R^n \to R$ is a multivariate polynomial, then there exists a unique univariate polynomial $\tilde{p} : R \to R$ of degree at most n such that for all $x = (x_1, \ldots, x_n) \in \{0,1\}^n$, we have*

$$p^{sym}(x_1, \ldots, x_n) = \tilde{p}(x_1 + \cdots + x_n).$$

Moreover, $\deg(\tilde{p}) \le \deg(p)$.

2.3 The Parity and MOD functions

Consider the function $\text{PARITY}(x) = \oplus_{i=1}^{n} x_i$, where $x = (x_1, \ldots, x_n)$. Over $\mathbb{Z}_2$, $\text{PARITY}(x) = \sum_{i=1}^{n} x_i$ which implies $\deg_2(f) = 1$. We have the following claim,

Claim. *For an input $x \in \{0,1\}^n$, let $|x|$ denote its Hamming weight (i.e. its number of 1s). The function f has degree n over $\mathbb{Z}_m$ if and only if*

$$\sum_{|x| \text{ even}} f(x) - \sum_{|x| \text{ odd}} f(x)$$

is non-zero in $\mathbb{Z}_m$.

Proof. $\sum_{|x| \text{ even}} f(x) - \sum_{|x| \text{ odd}} f(x)$ is the coefficient of the term $x_1 x_2 \ldots x_n$ in $P_m(x)$, where P_m denotes the polynomial representing f over $\mathbb{Z}_m$. $\square$

Using the above claim, it follows that PARITY has degree n over $\mathbb{Z}$. Further, by writing PARITY as $\frac{1}{2} - \frac{1}{2}\prod_{i=1}^{n}(1 - 2x_i)$ and taking modulo 3, one can get $\deg_3(\text{PARITY}) = n$. Another interesting example is of the MOD function which is defined as,

$$\text{MOD}_n^{c,m}(x) \coloneqq \mathbb{I}[|x| \equiv c \mod m] \in \{0,1\},$$

where $n \geq m-1$ denotes the length of the input x, and $\mathbb{I}[.]$ is the indicator function. Whenever the context is clear, we will abbreviate $\text{MOD}_n^{0,p}$ as MOD_p. The following theorem gives the degree of MOD_n^{0,p^t},

Theorem 2.3.1. *Let p be a prime and t, k be positive integers. Denote $d \coloneqq (k-1).\varphi(p^t) + p^t - 1$. Then for any $n \geq d$, we have*

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

We also have the following generalization to all remainders,

Theorem 2.3.2. *Let p be a prime and t, k be positive integers. Denote $d \coloneqq (k-1).\varphi(p^t) + p^t - 1$. For any $n \geq d$ and $0 \leq a < p^t$, we have*

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

The following result might also be interesting to note,

Lemma 2.3.3. *Let p, q be distinct primes. We have,*

$$\deg_q(MOD_n^{0,p}) = \Omega(n).$$

Proof. Let $g(x) = w^{|x|} = w^{\sum_i x_i} = \prod_i (1 - (1-w)x_i)$. Consider the following sum $\sum_x \omega^{|x|}(-1)^{|x|}$, where ω is chosen to be a pth root of unity. The sum evaluates to $(1-\omega)^n$, which is non-zero in char q. By the Claim above, $\deg_q(g) = n$. Further, consider the equality,

$$g(x_1, x_2, \ldots, x_{n-p}) = \sum_{r=0}^{p-1} \text{MOD}_n^{0,p}(x_1, x_2, \ldots, x_{n-p}, 1, \ldots, 1, 0, \ldots, 0)\omega^r,$$

where the rth term in the RHS has exactly r zeroes in the last r bits of $\mathrm{MOD}_n^{0,p}$. This gives,

$$\deg_q(\mathrm{MOD}_n^{0,p}) \geq \deg_q(g(x_1, \ldots, x_{n-p})) = n - p = \Omega(n).$$

$\square$

2.4 Complexity measures

Definition 2.4.1 (Sensitivity)**.** *The sensitivity complexity of f on input x is defined as $s(f,x) := |\{i : f(x) \neq f(x^{\oplus i})\}|$, and the sensitivity complexity of the function f is defined as $s(f) := \max_x s(f,x)$.*

Simon gave a lower bound on this measure [8],

Theorem 2.4.2. *For any non-degenerate Boolean function $f : \{0,1\}^n \to \{0,1\}$, we have*

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

Definition 2.4.3 (Block Sensitivity)**.** *The block sensitivity $bs(f,x)$ of f on input x is the maximum number of disjoint subsets $B_1, B_2, \ldots, B_r$ of $[n]$ such that for all j, $f(x) \neq f(x^{\oplus B_j})$. The block sensitivity of f is defined as $bs(f) = \max_x bs(f,x)$, and the minimum block sensitivity of f is defined as $bs_{\min}(f) = \min_x bs(f,x)$.*

Note that the above definitions imply the obvious bound $bs(f) \geq s(f)$. Nisan pointed out [14] that for monotone Boolean functions sensitivity and block sensitivity are equal.
A *deterministic decision tree* on n variables $x_1, \ldots, x_n$ is a rooted binary tree, whose internal nodes are labeled with variables, and the leaves are labeled 0 or 1. Edges are also labeled 0 or 1. To evaluate such a tree on input x, start at the root and query the corresponding variable, then move to the next node along the edge labeled with the outcome of the query. Repeat until a leaf is reached, at which point the label of the leaf is declared to be the output of the evaluation. A decision tree computes a Boolean function f if it agrees with f on all inputs.

Definition 2.4.4 (Decision tree complexity)**.** *The deterministic decision tree complexity of a Boolean function f, denoted by $D(f)$, is the depth of a minimum-depth decision tree that computes f.*

Let C be an assignment $C : S \to \{0,1\}$ of values to some subsets $S \subseteq [n]$. We say C is consistent with $x \in \{0,1\}^n$ if $x_i = C(i)$ for all $i \in S$. For $b \in \{0,1\}$, a b-certificate for f is an assignment C such that $f(x) = b$ whenever x is consistent with C . The size of C is $|S|$.

Definition 2.4.5 (Certificate Complexity)**.** *The certificate complexity $C(f,x)$ of f on input x is the size of a smallest $f(x)$-certificate that is consistent with x. The certificate complexity of f is $C(f) = \max_x C(f,x)$. The minimum certificate complexity of f is $C_{\min}(f) = \min_x C(f,x)$.*

For simplicity, we will refer to the largest degree monomials in the polynomial representation of a function as 'maxonomials'.

Definition 2.4.6 (Rank)**.** *Let $m \geq 2$ be an integer, the mod-m rank of a Boolean function f, denoted by rank$_m(f)$, is the minimum integer r s.t. f can be expressed as,*

$$f = x_{i_1} f_1 + \cdots + x_{i_r} f_r + f_0 \pmod m$$

where $\deg_m(f_i) < \deg_m(f)$ for all $0 \leq i \leq r$. Equivalently, rank$_m(f)$ is the minimum number of variables to hit all maxonomials in the $(\bmod\ m)$ polynomial representation for f.

For simplicity, we will abbreviate $\text{rank}_m(f)$ as $r_m(f)$. We will now introduce a new variant of rank, which has several useful properties.

Definition 2.4.7. *Let $m \geq 2$ be an integer, define $r'_m(f)$ to be the maximal number of disjoint maxonomials occuring in the $\bmod\ m$ polynomial representation of f.*

Lemma 2.4.8. *For any Boolean function f,*

$$r'_m(f) \leq r_m(f) \leq r'_m(f) \deg_m(f).$$

Proof. Consider any set of $r'_m(f)$ disjoint maxonomials in $P_m(x)$, and note that the $r'_m(f) \deg_m(f)$ variables in these maxonomials must hit all maxonomials in $P_m(x)$. $\square$

2.5 Relations between complexity measures

The block sensitivity is known to be polynomially related to the decision tree complexity, the certificate complexity, and the degree of the boolean function [8]. The following was proved by Huang [9],

Theorem 2.5.1. *For every Boolean function f,*

$$bs(f) \leq s(f)^4.$$

The above bound confirms the sensitivity conjecture, and places sensitivity among all the polynomially related classes mentioned above!

3 A General Lower bound on Degree

In this section, we will establish a tight lower bound on the degree of general non-degenerate Boolean functions over $\mathbb{Z}$.

3.1 Nisan-Szegedy Lower Bound [15]

Theorem 3.1.1. *Every Boolean function f that depends on n variables has degree $\deg(f) \geq \log_2(n) - O(\log\log(n))$.*

For the proof of this theorem, it will be convenient to use the Fourier transform representation, i.e., -1 for true and 1 for false (used in this subsection only). Thus, a Boolean function will be viewed as a real function $f : \{-1,1\}^n \to \{-1,1\}$. For a subset $S \subseteq [n]$, we will denote $X_S = \prod_{i \in S} x_i$. We will require the following set of results,

Lemma 3.1.2 (Parseval's Equality)**.** *If we represent a Boolean function f as $f = \sum_S \alpha_S X_S$, then*

$$\sum_S \alpha_S^2 = 1.$$

Definition 3.1.3 (Influence)**.** *For a Boolean function f and a variable x_i, the influence of x_i on f (denoted by $Inf_i(f)$) is defined to be the following probability*

$$Pr[f(x_1, \ldots, x_{i-1}, true, x_{i+1}, \ldots, x_n) = f(x_1, \ldots, x_{i-1}, false, x_{i+1}, \ldots, x_n)]$$

where $x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n$ are chosen at random in $\{false, true\}$.

Lemma 3.1.4. *For any Boolean function f on n variables, if we represent $f = \sum_S \alpha_S X_S$, then*

$$\sum_{i=1}^{n} Inf_i(f) = \sum_S |S| \alpha_S^2.$$

We have the following useful corollary from the above,

Corollary 3.1.5. *For any Boolean function f,*

$$\sum_{i=1}^{n} Inf_i(f) \leq \deg(f).$$

We will also use the Schwartz-Zippel lemma to generate an upper bound on the number of zeroes of multilinear polynomial over $\{-1, 1\}^n$.

Lemma 3.1.6 (Schwartz)**.** *Let $p(x_1, \ldots, x_n)$ be a multilinear polynomial of degree d. If we choose $x_1, \ldots, x_n$ independently at random in $\{-1, 1\}$, then the following inequality holds,*

$$Pr[p(x_1, \ldots, x_n) \neq 0] \geq 2^{-d}.$$

Proof. The proof is by induction on n. For $n = 1$, we just have a linear function of one variable which may have only one zero. For the induction step, write

$$p(x_1, \ldots, x_n) = x_n g(x_1, \ldots, x_{n-1}) + h(x_1, \ldots, x_{n-1}).$$

Note that if $p(x_1, \ldots, x_{n-1}, 1) \neq 0$ then $h(x_1, \ldots, x_{n-1}) + g(x_1, \ldots, x_{n-1}) \neq 0$, and if $p(x_1, \ldots, x_{n-1}, -1) \neq 0$ then $h(x_1, \ldots, x_{n-1}) - g(x_1, \ldots, x_{n-1}) \neq 0$. We now distinguish between three cases,

1. $h + g$ is identically equal to zero. In this case, $p = (x_n - 1)g$, where $\deg(g) = d - 1$ and we use the induction hypothesis on g for the x's satisfying $x_n = -1$.

2. $h - g$ is identically equal to zero. In this case, $p = (1 + x_n)g$, where $\deg(g) = d - 1$, and again we use the induction hypothesis on g for the x's satisfying $x_n = 1$.

3. Both $h + g$ and $h - g$ are not identically equal to zero. The degrees of $h + g$ and of $h - g$ are both bounded by d and thus we use the induction hypothesis on $h + g$ for the x's satisfying $x_n = 1$ and on $h - g$ for the x's satisfying $x_n = -1$.

The lemma follows. ☐

Now we will prove the main theorem of this section,

Proof of Theorem 3.1.1. For each $1 \leq i \leq n$, define a function f^i on $n-1$ variables as follows,

$$f^i(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n) = f(x_1, \ldots, x_{i-1}, -1, x_{i+1}, \ldots, x_n) - f(x_1, \ldots, x_{i-1}, 1, x_{i+1}, \ldots, x_n).$$

Observe that,

$$\mathrm{Inf}_i(f) = \Pr[f^i(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n) \neq 0],$$

where $x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n$ are chosen at random in $\{-1, 1\}$. Since f depends on all of its variables, we have that for every i, f^i is not identically zero, and thus, we can apply Lemma 3.1.6 and conclude that for all i, $\mathrm{Inf}_i(f) \geq 2^{-d}$. On the other hand, it follows from Corollary 3.1.5 that $\sum_i \mathrm{Inf}_i(f) \leq d$. Combining these two bounds, we get,

$$n/2^d \leq \sum_i \mathrm{Inf}_i(f) \leq d.$$

Thus $d2^d \geq n$, and the theorem follows. ☐

3.2 The Addressing function

The Addressing function Addr_r has $n = r + 2^r$ variables. We think of the input variables as being divided into two parts: there are r 'addressing' variables $y_1, \ldots, y_r$ and 2^r 'addressed' variables $\{z_a \mid a \in \{0,1\}^r\}$ (the latter part of the input is thus indexed by elements of $\{0,1\}^r$). On an input $(a, A) \in \{0,1\}^r \times \{0,1\}^{2^r}$, the output of the function is defined to be A_a (i.e. the ath co-ordinate of the vector A). The Addressing function satisfies $\deg(\mathrm{Addr}_r) = r + 1 = O(\log n)$. This construction implies the bound in Theorem 3.1.1 is tight.
Considering the role of 'Influence' in the proof of Theorem 3.1.1, it makes sense to try and see what the influence of the addressing function could be! To compute the total influence of the Addr_r function, split the variables into two: addressing variables X and the addressed variables Y. We have, $|X| = r$ and $|Y| = 2^r$. Each X variable influence is $1/2$, so that gives a total of $r/2$ as the influence of the X variables in the sum. The Y variables contribute quite less, something like $1/2^r$ for each, and the sum over influence of Y variables comes out to be 1. This gives $r/2 + 1 < r$ as the total influence. Compare this with the bound on total influence in the proof of Theorem 3.1.1.

3.3 Another Proof of Schwartz-Zippel Lemma

As a slight digression, we will look at a proof of the Schwartz lemma by Combinatorial Nullstellensatz! We restate the lemma,

Lemma 3.3.1 (Schwartz-Zippel)**.** *Let $f : \{0,1\}^n \to \{0,1\}$ be a Boolean function with $\deg(P(x)) = d$, where P represents f over $\mathbb{Z}$. We have,*

$$Pr[f(x_1, \ldots, x_n) \neq 0] \geq 2^{-d},$$

where $x_1, \ldots, x_n$ are chosen independently at random in $\{0,1\}$.

Proof. The polynomial $P(x_1, x_2, ..., x_n)$ representing f is multilinear on $\{0,1\}^n$ with degree d. Let $x_{i_1}x_{i_2}...x_{i_k}$ be a maxonomial in P with non-zero coefficient. Consider the 'box' $S_1 \times S_2 \times ... \times S_n$, where S_i is singleton $\{1\}$ or $\{0\}$ if the index i isn't part of the chosen maxonomial, or choose S_i to be $\{0,1\}$ if i is some i_k, $1 \leq k \leq d$. By Nullstellensatz, an assignment must exist for the variables involved in the maxonomial for which the whole polynomial evaluates to some non-zero value. We are fixing some assignment to the $(n-d)$ variables (which aren't involved in the maxonomial) at the start and then applying Nullstellensatz. This 'fixing' happens by the choice of the singleton $\{0\}$ or $\{1\}$ for the variables which aren't in the maxonomial. For each of the 2^{n-d} choices, we are guaranteed an assignment for the other d variables for which the polynomial evaluates to a non-zero value. Therefore, the non-vanishing set of the polynomial P has cardinality $\geq 2^{n-d}$, the bound follows. $\square$

4 Huang's proof of the Sensitivity Conjecture

This section is adapted from [9].

4.1 The Buildup

Let Q^n be the n-dimensional hypercube graph, whose vertex set consists of vectors in $\{0,1\}^n$. Two vectors are adjacent if they differ in exactly one coordinate. For an undirected graph G, we use the standard graph-theoretic notation $\Delta(G)$ for its maximum degree, and we use $\lambda_1(G)$ for the largest eigenvalue of its adjacency matrix. We will show the following,

Theorem 4.1.1. *For every integer $n \geq 1$, let H be an arbitrary $(2^{n-1}+1)$-vertex induced subgraph of Q^n. Then*

$$\Delta(H) \geq \sqrt{n}.$$

Moreover this inequality is tight when n is a perfect square.

The following was a major open problem posed by Nisan and Szegedy [15],

Conjecture 4.1.2 (Sensitivity Conjecture)**.** *There exists an absolute constant $C > 0$, such that for every boolean function f,*

$$bs(f) \leq s(f)^C.$$

For an induced graph H of Q^n, let $Q^n \setminus H$ denote the subgraph of Q^n induced on the vertex set $V(Q^n) \setminus V(H)$. Let $\Gamma(H) = \max(\Delta(H), \Delta(Q^n \setminus H))$. Gotsman and Linial [7] proved the following remarkable equivalence using Fourier analysis.

Theorem 4.1.3. *The following are equivalent for any monotone function $h : \mathbb{N} \to \mathbb{R}$:*

- *For any induced subgraph H of Q^n with $|V(H)| \neq 2^{n-1}$, we have $\Gamma(H) \geq h(n)$.*
- *For any boolean function f, we have $s(f) \geq h(\deg(f))$.*

Note that Theorem 4.1.1 implies that $h(n)$ can be taken as $\sqrt{n}$, since one of H and $Q^n \setminus H$ must contain at least $2^{n-1}+1$ vertices, and the maximum degree Δ is monotone. As a corollary, we have

Corollary 4.1.4. *For every Boolean function f,*

$$s(f) \geq \sqrt{\deg(f)}.$$

This inequality is tight for the AND-of-ORs boolean function [8, Example 5.2]. Tal [18] showed that $bs(f) \leq \deg(f)^2$, combining this with the above gives,

Theorem 4.1.5. *For every boolean function f,*

$$bs(f) \leq s(f)^4.$$

Therefore Conjecture 4.1.2 holds!

4.2 Proof of Main Theorem

To establish Theorem 4.1.1, we will first prove a series of lemmas. Recall that given an $n \times n$ matrix A, a *principal submatrix* of A is obtained by deleting the same set of rows and columns from A.

Lemma 4.2.1 (Cauchy's Interlace theorem)**.** *Let A be a symmetric $n \times n$ matrix, and let B be a $m \times m$ principal submatrix of A for some $m < n$. If the eigenvalues of A are $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$, and the eigenvalues of B are $\mu_1 \geq \mu_2 \geq \cdots \geq \mu_m$, then for all $1 \leq i \leq m$,*

$$\lambda_i \geq \mu_i \geq \lambda_{i+n-m}.$$

For a direct proof, refer to [4].

Lemma 4.2.2. *We define a sequence of symmetric square matrices iteratively as follows:*

$$A_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \; A_n = \begin{bmatrix} A_{n-1} & I \\ I & -A_{n-1} \end{bmatrix}.$$

Then A_n is a $2^n \times 2^n$ matrix whose eigenvalues are $\sqrt{n}$ of multiplicity 2^{n-1}, and $-\sqrt{n}$ of multiplicity 2^{n-1}.

Proof. We will show by induction that $A_n^2 = nI$. For $n = 1$ (base case), $A_1^2 = I$. Suppose the statement holds for $n-1$, i.e. $A_{n-1}^2 = (n-1)I$, then

$$A_n^2 = \begin{bmatrix} A_{n-1}^2 + I & 0 \\ 0 & A_{n-1}^2 + I \end{bmatrix} = nI.$$

Therefore, the eigenvalues of A_n are either $\sqrt{n}$ or $-\sqrt{n}$. Since $\mathrm{Tr}[A_n] = 0$, the lemma follows. $\square$

Lemma 4.2.3. *Suppose H is an m-vertex undirected graph and A is a symmetric matrix whose entries are in $\{-1, 0, 1\}$ and whose rows and columns are indexed by $V(H)$, and whenever u and v are non-adjacent in H, $A_{u,v} = 0$. Then*

$$\Delta(H) \geq \lambda_1 \coloneqq \lambda_1(A).$$

Proof. Suppose v is the eigenvector corresponding to the eigenvalue λ_1. Then $\lambda_1 v = Av$. WLOG assume v_1 is the coordinate of v with the largest absolute value, then

$$|\lambda_1 v_1| = |(Av)_1| = \Big|\sum_{j=1}^{m} A_{1,j} v_j\Big| \leq \sum_{j=1}^{m} |A_{1,j}||v_1| \leq \Delta(H)|v_1|.$$

Therefore, $|\lambda_1| \leq \Delta(H)$. $\square$

Now we will proceed to show the main theorem.

Proof of Theorem 4.1.1. Let A_n be the sequence of matrices defined in Lemma 4.2.2. Note that the entries of A_n are in $\{-1, 0, 1\}$. By the iterative construction of A_n, it is not hard to see that when changing every (-1)-entry of A_n to 1, we get exactly the adjacency matrix of Q^n, and thus A_n and Q^n satisfy the conditions in Lemma 4.2.3. For example, we may let the upper-left and lower-right blocks of A_n correspond to the two $(n-1)$-dimensional subcubes of Q^n, and the two identity blocks correspond to the perfect matching connecting these two subcubes. Therefore, a $(2^{n-1}+1)$-vertex induced subgraph H of Q^n and the principal submatrix A_H of A_n naturally induced by H also satisfy the conditions of Lemma 4.2.3. As a result,

$$\Delta(H) \geq \lambda_1(A_H).$$

On the other hand, from Lemma 4.2.2, the eigenvalues of A_n are,

$$\sqrt{n}, \ldots, \sqrt{n}, -\sqrt{n}, \ldots, -\sqrt{n}.$$

Note that A_H is a $(2^{n-1}+1) \times (2^{n-1}+1)$ submatrix of the $2^n \times 2^n$ matrix A_n. By Lemma 4.2.1,

$$\lambda_1(A_H) \geq \lambda_{2^{n-1}}(A_n) = \sqrt{n}.$$

Combining the two inequalities obtained,

$$\Delta(H) \geq \sqrt{n}.$$

$\square$

5 Some Conjectures and Separation Theorems

In this section, we focus on the relation between $\deg(f)$ and $\deg_m(f)$ when m has atleast two distinct prime factors. For the case when m is a prime power, the separation between $\deg_m(f)$ and $\deg(f)$ can be arbitrarily large.

5.1 A Conjecture

Conjecture 5.1.1. *Let f be a Boolean function and m be an integer which has at least two distinct prime factors, then*

$$\deg(f) \leq \mathit{poly}(\deg_m(f)).$$

Towards resolving the above, we establish some equivalent conjectures which might be easier to prove.

Theorem 5.1.2. *The following four conjectures are all equivalent to Conjecture 5.1.1:*

1. *For any Boolean function f, and two distinct primes p and q,*

$$rank_p(f) \leq poly(\deg_p(f), \deg_q(f)).$$

2. *For any Boolean function f, and two distinct primes p and q,*

$$C_{\min}(f) \leq poly(\deg_p(f), \deg_q(f)).$$

3. *For any Boolean function f, and two distinct primes p and q,*

$$bs_{\min}(f) \leq poly(\deg_p(f), \deg_q(f)).$$

4. *For any Boolean function f, and two distinct primes p and q,*

$$s(f) \leq poly(\deg_p(f), \deg_q(f)).$$

Towards establishing Theorem 5.1.2, we will prove a series of lemmas.

Lemma 5.1.3. *Conjecture 1 of Theorem 5.1.2 $\Leftrightarrow$ Conjecture 5.1.1.*

Proof. ($\Leftarrow$) Follows as $\mathrm{rank}_p(f) = O(\deg(f)^3)$.
($\Rightarrow$) We design an algorithm to query f, which contains at most $\deg_p(f)$ rounds and each round reduces $\deg_p$ by at least one. Denote the function at round t by $f^{(t)}$. Note that $f^{(t)}$ is a subfunction of f, hence $\deg_p(f^{(t)}) \leq \deg_p(f)$ and $\deg_q(f^{(t)}) \leq \deg_q(f)$. For each round, we can query $\mathrm{rank}_p(f^{(t)})$ variables to make the largest monomials in $P_p(x)$ vanish, which means $\deg_p(f^{(t)})$ is reduced by at least one. Therefore assuming Conjecture 1, we have $\mathrm{rank}_p(f^{(t)}) \leq poly(\deg_p(f^{(t)}), \deg_q(f^{(t)})) \leq poly(\deg_p(f), \deg_q(f))$, which implies $\deg(f) \leq D(f) \leq poly(\deg_p(f), \deg_q(f))$. $\square$

Since $\mathrm{rank}_p(f) \leq C_{\min}(f) = O(\deg(f)^3)$, we get Conjecture 2 $\Leftrightarrow$ Conjecture 5.1.1.

Lemma 5.1.4. *Conjecture 3 of Theorem 5.1.2 $\Leftrightarrow$ Conjecture 5.1.1.*

Proof. ($\Leftarrow$) Follows as $bs_{\min}(f) \leq bs(f) = O(\deg(f)^2)$.
($\Rightarrow$) We call monomial M maximal in $P_p(x)$ if no other monomials contains it. Observe that for any input x and any maximal monomial M, there exist a block $B \subseteq \mathrm{supp}(M)$ such that $f(x) \neq f(x^{\oplus B})$, because for any restriction S: $[n] \setminus M \to \{0,1\}$ monomial M can't be cancelled, which implies $f|_S$ is a nonconstant function. In addition, according to the definition of $\mathrm{rank}_p(f)$, there exists at least $\mathrm{rank}_p(f)/\deg_p(f)$ disjoint largest monomials in $P_p(x)$. Therefore we get $bs_{min}(f) \geq \mathrm{rank}_p(f)/\deg_p(f)$, which implies Conjecture 5.1.1 assuming Conjecture 3. $\square$

Lemma 5.1.5. *Conjecture 5.1.1 $\Leftrightarrow$ Conjecture 4 of Theorem 5.1.2.*

Proof. ($\Rightarrow$) Since $s(f) = O(\deg(f)^2)$, we have Conjecture 5.1.1 $\Rightarrow$ Conjecture 4.
($\Leftarrow$) WLOG assume $bs(f, \overrightarrow{0}) = bs(f) = r$, where $\overrightarrow{0} = (0, \ldots, 0)$. There exist r disjoint blocks $B_1, B_2, \ldots, B_r \subseteq [n]$ such that for all i, $f(0) \neq f(0^{\oplus B_i})$. Further assume WLOG that $i \in B_i$. Replace (substitute) all variables in B_i to get a new function f'. Observe that $f'(\overrightarrow{0}) = f(\overrightarrow{0}) \neq f(\overrightarrow{0}^{B_i}) = f'(\overrightarrow{0}^i)$. This gives,

$$bs(f) = s(f') \leq poly(\deg_p(f'), \deg_q(f')) \leq poly(\deg_p(f), \deg_q(f)).$$

The conclusion follows since $bs(f)$ and $\deg(f)$ are polynomially related. $\square$

Note that Theorem 2.5.1 implies sensitivity and degree are polynomially related, and immediately implies the conclusion of the above lemma. However, the above proof introduces some new ideas (eg method of 'replacing' variables), and might be instructive. We confirm Conjecture 5.1.1 for the case of symmetric functions.

Lemma 5.1.6. *Let $f : \{0,1\}^n \to \{0,1\}$ be symmetric and nonconstant, and p_1, p_2 are two distinct primes, then*

$$\deg(f) \leq n < p_1 \deg_{p_1}(f) + p_2 \deg_{p_2}(f).$$

Proof. Let $d_i = \deg_{p_i}(f)$, $L_i = p_i^{1+\lfloor \log_{p_i} d_i \rfloor}$, and $P_{p_i}(x)$ refer to the mod p_i polynomial representation of f $(i = 1, 2)$. Since f is symmetric, each $P_{p_i}(x)$ can be written as $\sum_{k=0}^{d_i} c_{i,k} \binom{|x|}{k}$. Then according to Lucas formula, for any nonnegative integers s, j and $k \leq d_i$, we have

$$\binom{sL_i + j}{k} \equiv_{p_i} \binom{j}{k}.$$

Define $g(|x|) = f(x)$, the above equality says $g(k + L_i) = g(k)$. It suffices to show $n < L_1 + L_2$, which implies $n < p_1 d_1 + p_2 d_2$ proving the lemma. Assume for sake of contradiction that $L_1 + L_2 \leq n$. Since $L_1 \neq L_2$, assume WLOG $L_1 < L_2$. We claim that $\forall k \leq L_2$, $g(k) = g(k + L_1 \bmod L_2)$. If $k + L_1 \leq L_2$, we are done since $g(k + L_1) = g(k)$. If $k + L_1 > L_2$, we have $g(k) = g(k + L_1) = g(k + L_1 - L_2) = g(k + L_1 \bmod L_2)$, since $L_1 + L_2 \leq n$ and $L_1 < L_2$. Moreover, $\gcd(L_1, L_2) = 1$, hence $\forall l \leq L_2$, there exists an integer t such that $l - k \equiv_{L_2} tL_1$, i.e. $g(k) = g(k + tL_1 \bmod L_2) = g(l)$, which means f is constant, a contradiction. $\square$

5.2 A separation theorem

In the direction of disproving Conjecture 5.1.1, the following gives a quadratic separation between $\deg_{p_1 p_2}(f)$ and $\deg(f)$.

Theorem 5.2.1. *For any two distinct prime p_1 and p_2, there exists a sequence of Boolean functions f, s.t.:*

$$\deg_{p_1 p_2}(f) = O(\deg(f)^{1/2}).$$

Proof. Let $f = \mathrm{Mod}_{p_1}(\mathrm{Mod}_{p_2}(x_1, \ldots, x_{\sqrt{n}}), \ldots, \mathrm{Mod}_{p_2}(x_{n-\sqrt{n}+1}, \ldots, x_n))$. Here, $\mathrm{Mod}_{p_i}(.) = 0$, if the sum of inputs can be divided by p_i, otherwise $\mathrm{Mod}_{p_i}(.) = 1$. Over $\mathbb{Z}$, Mod_{p_1} with $\sqrt{n}$ arguments has deg $= \Omega(\sqrt{n})$, further each Mod_{p_2} function in the maximum degree monomial (with respect to Mod_{p_1}) will have deg $= \Omega(\sqrt{n})$. This gives a $\Omega(n)$ bound on $\deg(f)$. Now note that $\deg_{p_1p_2}(f) = \max(\deg_{p_1}(f), \deg_{p_2}(f))$. We claim that $\deg_{p_i}(f) = O(\sqrt{n})$, for $i = 1, 2$. This follows as in char p_1, Mod_{p_2} with $\sqrt{n}$ arguments will have $\deg = \Omega(\sqrt{n})$ and similarly in char p_2, Mod_{p_1} with $\sqrt{n}$ arguments will have $\deg = \Omega(\sqrt{n})$. $\square$

5.3 $bs_{\min}$ vs rank

Is there a function with low hitting set size and high block sensitivity? The answer to this question is affirmative; we will provide a construction illustrating this.

6 Complexity in different characteristics

In this section, we analyze how the degree of a function in one characteristic affects its complexity in other characteristics. We establish the following general principle: *functions with low degree modulo p must have high complexity in every other characteristic q.*

6.1 $\deg_p(f)$ vs $\deg_q(f)$

Theorem 6.1.1. *Let $f : \{0,1\}^n \to \{0,1\}$ be a Boolean function which depends on all n variables. Let $p \neq q$ be distinct primes. Then*

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

6.2 Tightness of Degree bounds

Theorem 6.2.1. *For any Boolean function f,*

$$\deg(f) \geq \frac{n}{2^{\deg_p(f)}}.$$

Sketch. A slight modification of the proof of Theorem 3.1.1 gives the result, note that Schwartz-Zippel lemma holds over any characteristic. $\square$

In the context of the above, we will attempt to see if there are constructions which achieve degree d over char 3 and degree $n/2^d$ over char 0. Consider the Boolean function $g = \mathrm{MOD}_3(f_1, f_2, ..., f_k)$, where $k = n/2^{d/p}$ and each f_i is addressing function with $\deg(f_i) = d/p$. This gives,

$$\deg(g) = \sum_i \deg(f_i) = \frac{dn}{2^{d/p}},$$

since $\deg(\mathrm{MOD}_n^{0,3}) = n$. Further,

$$\deg_3(g) = \frac{2d}{p},$$

since $\deg_3(\mathrm{MOD}_3) = 2$. Setting $p = 2$ gives $\deg_3(g) = d$ and $\deg(g) = O(n/2^{d/2})$, clearly this doesn't work!

We will try induction on d to build a function which achieves degree $O(d)$ over char 3 and degree $O(n/2^d)$ over char 0. Set $f_1 = \mathrm{MOD}_n^{0,1}$ for the base case, this has degree $\Omega(n)$ over $\mathbb{Z}$ and degree 2 over char 3. Since we are building the function inductively, the idea is to check if there's a way to relate f_{d+1} and f_d, where f_d is the function with degree $O(d)$ in char 3 and degree $O(n/2^d)$ in char 0. Therefore, let $f_{d+1} = x_1 f_d + (1 - x_1) f_d$, where each f_d has $n/2$ variables, is of degree d in char 3 and degree $(n/2)/2^d = n/2^{d+1}$ in char 0. The function f_{d+1} has n variables, is of degree $d+1$ in char 3 and degree $n/2^{d+1}$ in char 0.

Remark. *The function just constructed is basically the addressing function but with* $MOD_3(x_1, ..., x_{n/2^d})$ *at the leaves. Everytime we decrease the parameter* d *by* 1*, we halve the number of variables. So there are* $n/2^d$ *variables in total at the leaves, and* MOD_3 *functions at the leaves because that's the base case for the inductive construction.*

7 Functions with Transitive Symmetries

This section is adapted from [2].

7.1 Some Definitions

We call the elements of $\{0,1\}^n$ 'words'. For any word x and $1 \le i \le n$ we denote by x^i the word obtained by switching the ith bit of x. For a word x and $A \subseteq [n]$ we use x^A to denote the word obtained from x by switching all the bits in A. For a word $x = x_1, x_2, \ldots, x_n$ we define $\text{supp}(x)$ as $\{i | x^i = 1\}$. Weight of x, denoted $wt(x)$, is $|\text{supp}(x)|$, i.e., number of 1's in x.

Definition 7.1.1 (0-sensitivity)**.** *We define* 0*-sensitivity of* f *as* $s^0(f) = \max\{s(f,x) \ : \ x \in \{0,1\}^n, \ f(x) = 0\}$.

Similarly, we define 1-sensitivity of a Boolean function,

Definition 7.1.2 (1-sensitivity)**.** *We define* 1*-sensitivity of* f *as* $s^1(f) = \max\{s(f,x) \ : \ x \in \{0,1\}^n, \ f(x) = 1\}$.

Definition 7.1.3 (Partial Assignment)**.** *A partial assignment is a function* $p : S \to \{0,1\}$ *where* $S \subseteq [n]$*. We call* S *the support of this partial assignment. The weight of a partial assignment is the number of elements in* S *that is mapped to* 1*. We call* x *a (full) assignment if* $x : [n] \to \{0,1\}$*. (Note that any word* $x \in \{0,1\}^n$ *can be thought of as a full assignment.) We say* $p \subseteq x$ *if* x *is an extension of* p*, i.e., the restriction of* x *to* S *denoted* $x|_S = p$*.*

Definition 7.1.4 (1-certificate)**.** *A 1-certificate is a partial assignment,* $p : S \to \{0,1\}$*, which forces the value of the function to* 1*. Thus if* $x|_S = p$ *then* $f(x) = 1$*.*

Definition 7.1.5. *If* $\mathcal{F}$ *is a set of partial assignments then we define* $m_{\mathcal{F}} : \{0,1\}^n \to \{0,1\}$ *as* $m_{\mathcal{F}}(x) = 1 \Leftrightarrow (\exists p \in \mathcal{F})$ *such that* $(p \subseteq x)$*.*

Note that each member of $\mathcal{F}$ is a 1-certificate for $m_{\mathcal{F}}$ and $m_{\mathcal{F}}$ is the unique smallest such function.

Definition 7.1.6 (Minterms)**.** *A minterm is a minimal* 1*-certificate, that is, no sub-assignment is a* 1*- certificate.*

Definition 7.1.7. *Let* $S \subseteq [n]$ *and let* $\pi \in S_n$*. Then we define* S^π *to be* $\{\pi(i) \,|\, i \in S\}$*.*

Let G be a permutation group acting on $[n]$. Then the sets S^π, where $\pi \in G$, are called the G-shifts of S. If $p : S \to \{0,1\}$ is a partial assignment then we define $p^\pi : S^\pi \to \{0,1\}$ as $p^\pi(i) = p(\pi^{-1}i)$.

Definition 7.1.8 (G-invariant functions)**.** *Let G be a subgroup of S_n, i.e., a permutation group acting on $[n]$. A function $f : \{0,1\}^n \to \{0,1\}$ is said to be invariant under the group G if for all permutations $\pi \in G$ we have $f(x^\pi) = f(x)$ for all $x \in \{0,1\}^n$.*

Definition 7.1.9 (Transitive group)**.** *Let G be a permutation group on $[n]$. G is called transitive if for all $1 \leq i, j \leq n$ there exists a $\pi \in G$ such that $\pi(i) = j$.*

Definition 7.1.10 (Transitive-invariant functions)**.** *A Boolean function which is G-invariant for some transitive permutation group G is a transitive-invariant function.*

We can analogously define cyclic-invariant functions.

Proposition 7.1.11. *Let G be a permutation group. Let $p : S \to \{0,1\}$ be a partial assignment and let $\mathcal{F} = p^\pi \mid \pi \in G$. Then p is a minterm for the function $m_{\mathcal{F}}$.*

The function $m_{\mathcal{F}}$ will be denoted p^G. Note that the function p^G is invariant under the group G. When G is the group of cyclic shifts we denote the function p^{cyc}. The function p^{cyc} is cyclically invariant.

Proof of Proposition 7.1.11. If p has k zeros then for any word x with fewer than k zeros $m_{\mathcal{F}}(x) = 0$, since all the element of $\mathcal{F}$ has same number of 1's and 0's. But if q is a 1-certificate with fewer than k zeros we can have a word x by extending q to a full assignment by filling the rest with 1's, satisfying $f(x) = 1$ (since $q \subseteq x$). But x contains fewer than k zeros, a contradiction. So no minterm of $m_{\mathcal{F}}$ has fewer than k zeros. Similarly no minterm of $\mathcal{F}$ has weight less than p. So no proper sub-assignment of p can be a 1-certificate. Hence p is a minterm of $m_{\mathcal{F}}$. ☐

Definition 7.1.12 (Minterm-cyclic functions)**.** *Let $C(n,k)$ be the set of Boolean functions f on n variables such that there exists a partial assignment $p : S \to \{0,1\}$ with support k ($\neq 0$) for which $f = p^{cyc}$. Let $C(n) = \cup_{k=1}^{n} C(n,k)$. We will call the functions in $C(n)$ minterm-cyclic.*

Definition 7.1.13 (Minterm-transitive functions)**.** *Let G be a permutation group on $[n]$. We define $D_G(n,k)$ (for $k \neq 0$) to be the set of Boolean functions f on n variables such that there exists a partial assignment $p : S \to \{0,1\}$ with support k for which $f = p^G$. We define $D_G(n)$ to be $\cup_{k=1}^{n} D_G(n,k)$. We define $D(n)$ to be $\cup_G D_G(n)$ where G ranges over all transitive groups. We call these functions minterm-transitive.*

Note that the class of minterm-cyclic functions is a subset of the class of minterm-transitive functions.

7.2 Minterm-transitive functions have sensitivity $\Omega(n^{1/3})$

Theorem 7.2.1. *If f is a minterm-transitive function on n variables then $s(f) = \Omega(n^{1/3})$ and $s^0(f)s^1(f) = \Omega(\sqrt{n})$.*

To prove the theorem, the following three lemmas are used. Since f is a minterm-transitive function, i.e., $f \in D(n)$, we can say $f \in D_G(n,k)$ for some transitive group G and some $k \neq 0$.

Lemma 7.2.2. *If $f \in D_G(n,k)$, then $s^1(f) \geq k/2$.*

Proof. Let y be the minterm defining f. Without loss of generality $wt(y) \geq k$. Let us extend y to a full assignment x by assigning zeros everywhere outside the support of y. Then switching any 1 to 0 changes the value of the function from 1 to 0. So we obtain $s(f,x) \geq k$. Hence $s^1(f) \geq k$. $\square$

Lemma 7.2.3. *If S is a subset of $[n]$, $|S| = k$ then there exist at least n/k^2 disjoint G-shifts of S.*

Proof. Let T be a maximal union of G-shifts of S. Since T is maximal T intersects with all G-shifts of S. So we must have $|T| \geq n/k$. So T must be a union of at least n/k^2 disjoint G-shifts of S. And this proves the lemma. $\square$

Lemma 7.2.4. *If $f \in D_G(n,k)$ then $s^0(f) = \Omega(n/k^2)$.*

Proof. Let y be the minterm defining f. By Lemma 7.2.3 we can have $\Omega(n)$ disjoint G-shifts of y. The union of these disjoint G-shifts of y defines a partial assignment. Let $S = \{s_1, s_2, \ldots, s_r\}$ be the support of the partial assignment. And let Y_{s_i} be the value of the partial assignment in the s_i-th entry. Since $k \neq 0$ the function f is not a constant function. Thus there exists a word z such that $f(z) = 0$. The i-th bit of z is denoted by z_i. We define,

$$T = \{j \mid z_j \neq Y_{s_m}, s_m = j\}.$$

Now let $P \subseteq T$ be a maximal subset of T such that $f(z^P) = 0$. Since P is maximal,if we switch any other bit in $T \setminus P$ the value of the function f will change to 1. So $s(f, z^P) \geq |(T \setminus P)|$. Now since $f(z^P) = 0$ we note that z^P does not contain any G-shift of y. But from Lemma 7.2.3 we know that z^T contains $\Omega(n)$ disjoint G-shifts of y. So $|(T \setminus P)|$ is $\Omega(n/k^2)$ and thus $s^0(f) \geq s(f, z^P) = \Omega(n/k^2)$. $\square$

Proof of Theorem 7.2.1. From the Lemma 7.2.2 and Lemma 7.2.4 we obtain,

$$s(f) = \max(s^0(f), s^1(f)) = \max(\Omega(\frac{n}{k}), k/2).$$

This implies $s(f) = \Omega(n^{1/3})$. Now since $s^0(f)$ and $s^1(f)$ cannot be smaller than 1, it follows from Lemma 7.2.2 and 7.2.4 that,

$$s^0(f)s^1(f) = \max(\Omega(n/k), k/2).$$

So $s^0(f)s^1(f) = \Omega(\sqrt{n})$. $\square$

The following is a corollary to Theorem 7.2.1,

Corollary 7.2.5. *If f is minterm-transitive, then $bs(f) = O(s(f)^3)$.*

Hence for minterm-transitive functions, sensitivity and block sensitivity are polynomially related.

Remark. *Note that Huang's result already implies the above!*

7.3 An Open Question

The following is stated as an open question in [2],

Problem: If f is a Boolean function invariant under a transitive group of permutations then is it true that $s(f) \geq n^c$ for some constant $c > 0$? We will show the answer is affirmative! Consider the following,

Proposition 7.3.1. *Any Boolean function that has a transitive group of symmetries has sensitivity (and hence degree) at least n^ϵ for some constant $\epsilon > 0$.*

Proof. Let the Boolean function f be invariant under the transitive group of permutations $G \subseteq S_n$. By Theorem 2.5.1, it suffices to show that $C(f) \geq n^\epsilon$ for some constant $\epsilon > 0$. Let p be a 1-certificate of size k. WLOG let $p = \{1, 2, ...k\}$. Consider the set $S = \{\sigma(p) = (\sigma(1), ..., \sigma(k)) : \sigma \in G\}$, i.e. the set of all G-shifts of the certificate p. We pick a permutation σ uniformly at random from group G. Consider the following claim,

Claim. *For all $1 \leq i, j \leq n$, $Pr(\sigma(i) = j) = 1/n$.*

Proof of Claim. The action of group G on the set $[n]$ is transitive which implies a single orbit, and therefore all stabilizers must have the same size $|G|/n$. $\square$

Now, note that p and $\sigma(p)$ have non-zero intersection if there exists $1 \leq i, j \leq k$ such that $\sigma(i) = j$. Taking union over these k^2 events, and applying union bound gives that the probability of p and $\sigma(p)$ having non-zero intersection is $\leq k^2/n$. Say we have generated r disjoint certificates from the set S until a certain point. If $k^2 r/n < 1$, we can pick σ uniformly at random from G, and $\sigma(p)$ will gives us another certficate disjoint from all the others with non-zero probability. Therefore, for any set of r maximal number of disjoint certificates, we must have $k^2 r \geq n$. We have the following claim,

Claim. *If p is any 0-certificate for f and q is any 1-certificate for f, then p and q must 'conflict': that is there exists $i \in domain(p) \cap domain(q)$ such that $p(i)$ and $q(i)$ are different.*

Proof of Claim. Assume for contradiction that p and q are disjoint certificates forcing the function f to be 0 and 1 respectively. The function f must evaluate to 0 for all inputs agreeing with the certificate p, and evaluate to 1 for all inputs agreeing with the certificate q. Since p and q are disjoint, there exists $x \in \{0, 1\}^n$ which agrees with both p, q. Therefore we must have that $f(x)$ is both 0 and 1, which is a contradiction. $\square$

It follows from the claim that the minimal 0-certificate must be of size atleast $r \geq n/k^2$, since we generated r disjoint 1-certificates for f. We can conclude,

$$C(f) \geq \max(k, n/k^2) \geq n^{1/3}.$$

The proposition follows. $\square$

Remark. *If the Conjecture 5.1.1 is true, then for any function f with a transitive group of symmetries, it must be the case that either $\deg_p(f)$ or $\deg_q(f) > n^\delta$ for some constant $\delta > 0$.*

8 The Case of Symmetric Functions

In this section, we establish lower bounds for mod m degree of symmetric functions and lower bounds for non-degenerate functions provided the number of inputs (n) is sufficiently large.

8.1 Lower bounds on degree

Gathen and Roche show that $\deg(f) \geq \deg_{p(n)}(f) \geq p(n) - 1$ for any non-trivial symmetric Boolean function f, where $p(n)$ is the largest prime below $n+2$. Using the current best result on prime gaps, this implies $\deg(f) \geq n - O(n^{0.525})$. Following this we have the conjecture,

Conjecture 8.1.1. *For any non-trivial symmetric Boolean function $f : \{0,1\}^n \to \{0,1\}$,*

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

8.2 Weakened Symmetry

Will the results in the ICALP paper still go through if the symmetry conditions are weakened, say the function is C_n symmetric?

9 Rational Degree

9.1 Introduction

A natural measure of Boolean function complexity is the minimal degree of a rational polynomial which represents the function exactly, called the *rational degree* (denoted rdeg). However, rdeg is not known to be either polynomially related to or separated from the complexity measures mentioned above. In fact, this was the other open question posed over 30 years ago in the paper of Nisan and Szegedy (via personal communication with Fortnow) [15]. This question was reiterated by Aaronson et al. [1] yet very little progress has been made toward its resolution. The following is an interesting open problem,

Question: Does there exist $c > 1$ such that for all total Boolean functions f, $\deg(f) \leq O(\mathrm{rdeg}(f)^c)$?

de Wolf defined the non-deterministic degree $\mathrm{ndeg}(f)$ of a Boolean function f as the minimal degree of a polynomial whose zero set is precisely the set of inputs on which f evaluates to false [3], and related it to the rational degree through the identity $\mathrm{rdeg}(f) = \max(\mathrm{ndeg}(f), \mathrm{ndeg}(\bar{f}))$, where $\bar{f}$ is the (Boolean) function $(1 - f)$. Observe that this implies,

$$\mathrm{rdeg}(f) = \mathrm{rdeg}(\bar{f}),$$

for all Boolean functions f.

Remark. *Let P_1 be a non-deterministic polynomial for f, and let P_2 be a non-deterministic polynomial for $\bar{f}$. We have,*

$$f = \frac{P_1}{P_1 + P_2},$$

and this acts as a rational representation for the (Boolean) function f.

In the same paper, de Wolf stated the following equivalent conjecture,

Conjecture 9.1.1 ([3])**.** *For all Boolean functions f, $D(f) \leq O(ndeg(f), ndeg(\bar{f}))$.*

9.2 A Tight Lower Bound for Symmetric Functions

Recall that we have the relation,

$$\mathrm{rdeg}(f) = \max(\mathrm{ndeg}(f), \mathrm{ndeg}(\bar{f}))$$

for any Boolean function f. First we will show the following,

Proposition 9.2.1. *If $f : \{0,1\}^n \to \{0,1\}$ is symmetric and non-constant, then $rdeg(f) \geq n/4$.*

Proof. Define $S_0 = \{k \in [n] : |x| = k \Rightarrow f(x) = 0\}$ and $S_1 = \{k \in [n] : |x| = k \Rightarrow f(x) = 1\}$. WLOG let $|S_0| \geq |S_1|$. If $f = p/q$ is a rational representation for f, symmetrize the polynomial p^2. The symmetrization of p^2 is univariate, non-zero, and must have atleast $n/2$ roots. By Lemma 2.2.2, it follows that $\deg(p^2) \geq n/2$ which implies $\mathrm{rdeg}(f) \geq \deg(p) \geq n/4$. □

How tight is the above bound? Can we construct a symmetric function f such that $\mathrm{rdeg}(f) = n/3$? The characterization of Zariski closures of symmetric subsets of the hypercube [19] turns out to be quite useful! The following construction is motivated from [19, Section 4, Lemma 28(b)].

Proposition 9.2.2. *Let $n + 1 = 3k$, for some $k \in \mathbb{N}$. There exists a symmetric, non-constant Boolean function f for which $rdeg(f) = (n + 1)/3$.*

Proof. Since $\mathrm{rdeg}(f) = \max(\mathrm{ndeg}(f), \mathrm{ndeg}(\bar{f}))$, it suffices to find f for which $\mathrm{ndeg}(f) \leq (n+1)/3$ and $\mathrm{ndeg}(\bar{f}) \leq (n+1)/3$ with equality holding in atleast one of the two inequalities. Let f be the Boolean function which vanishes on all Hamming weights in the interval $[(n + 1)/3, (2n - 1)/3]$, and is 1 otherwise. Consider the polynomial

$$P(x) = \prod_{k \in [(n+1)/3, (2n-1)/3]} \Big(|x| - k\Big).$$

This is of degree $(n + 1)/3$ and vanishes on all inputs with Hamming weight in the interval $[(n + 1)/3, (2n - 1)/3]$. To construct a non-deterministic polynomial for $\bar{f}$, we want a polynomial which is non-zero on all inputs with Hamming weight in the interval $[(n + 1)/3, (2n - 1)/3]$, and vanishes on all inputs with Hamming weight in $[0, (n - 2)/3] \cup [(2n + 2)/3, n]$. To satisfy the constraint $\mathrm{ndeg}(\bar{f}) \leq (n+1)/3$, we further require the degree of this polynomial to be $\leq (n+1)/3$. Consider the polynomial

$$Q(X) = (X_1 - X_2) \ldots (X_{(2n-1)/3} - X_{(2n+2)/3})$$

of degree $\leq (n+1)/3$, where $X = (X_1, \ldots, X_n)$. This vanishes on all inputs with Hamming weights in $[0, (n - 2)/3] \cup [(2n + 2)/3, n]$, but might also vanish if the Hamming weight of some input lies in the 'middle' interval $[(n + 1)/3, (2n - 1)/3]$. To fix this issue, we consider the sum

$$R(X) = \sum_{\sigma \in S_n} \lambda_\sigma Q(X_\sigma),$$

where $X_\sigma = (X_{\sigma(1)}, \ldots, X_{\sigma(n)})$ and $\{\lambda_\sigma\}_{\sigma \in S_n}$ is a set of linearly independent reals over $\mathbb{Q}$.

Claim. *If $|x| < (n+1)/3$ or $|x| > (2n-1)/3$, then $R(x) = 0$. Conversely, if $(n+1)/3 \le |x| \le (2n-1)/3$, then $R(x) \ne 0$.*

Proof of Claim. Note that $\forall \sigma \in S_n$, $|x_\sigma| = |x|$. Therefore, $R(x)$ vanishes if $|x| \in [0, (n-2)/3] \cup [(2n+2)/3, n]$. Assume now that for some input x, $|x| \in [(n+1)/3, (2n-1)/3]$. If $R(x) = 0$, by linear independence of $\{\lambda_\sigma\}_{\sigma \in S_n}$, it must follow that $Q(x_\sigma) = 0$ $\forall \sigma \in S_n$. However, note that we can choose $\sigma \in S_n$ such that $Q(x_\sigma) \ne 0$. This is a contradiction and the claim follows. ☐

The proposition follows. ☐

Remark. *We can (explicitly) choose λ_σ as 2^i (powers of 2) starting from $i = 0$ to $n! - 1$. This works because $Q(X) \in \{-1, 0, 1\}$ for all $X \in \mathbb{F}_2^n$, and base-2 representation is unique.*

Remark. *Proceeding similarly as above, we can show that for $n = 3k-2$ there exists a symmetric function f for which rdeg$(f) = (n+2)/3$. For $n = 3k$, there exists a construction for f which achieves rdeg$(f) = (n+3)/3$.*

Note that so far we have established $\mathrm{rdeg}(f) \ge n/4$ for any symmetric Boolean function f, and also specified a construction for which $\mathrm{rdeg}(f) \le (n+1)/3$. This motivates us to improve the lower bound even further,

Proposition 9.2.3. *If $f : \{0,1\}^n \to \{0,1\}$ is symmetric and non-constant, then rdeg$(f) \ge (n+1)/3$.*

Proof of Proposition 9.2.3. Given function $f : \{0,1\}^n \to \{0,1\}$, symmetric and non-constant. Assume for contradiction that $\mathrm{rdeg}(f) \le d$, for $d \in [n/4, (n+1)/3)$. Since $\mathrm{rdeg}(f) = \max(\mathrm{ndeg}(f), \mathrm{ndeg}(\bar{f}))$, we have $\mathrm{ndeg}(f) \le d$ and $\mathrm{ndeg}(\bar{f}) \le d$. Since f is symmetric, there exists $E \subseteq [0,n]$ such that f evaluates to 1 on all inputs with Hamming weight in E, and is 0 otherwise. There exist non-deterministic polynomials P_1 and P_2 for f and $\bar{f}$ (resp.) with $\deg(P_1) \le d$ and $\deg(P_2) \le d$. This gives,

$$\mathrm{Z}^* - \mathrm{cl}_{G,d}(E) = E \text{ and } \mathrm{Z}^* - \mathrm{cl}_{G,d}\Big([0,n] \setminus E\Big) = \Big([0,n] \setminus E\Big)$$

where $G = \{0,1\}^n$. Assume WLOG $|E| \le \big|[0,n] \setminus E\big|$, i.e. $|E| \le (n+1)/2$. From [19, Theorem 7, Section 3.2], we get

$$\mathrm{Z}^* - \mathrm{cl}_{G,d}(E) = \bar{L}_{n,d}(E) \text{ and } \mathrm{Z}^* - \mathrm{cl}_{G,d}\Big([0,n] \setminus E\Big) = \bar{L}_{n,d}\Big([0,n] \setminus E\Big).$$

Therefore,

$$\bar{L}_{n,d}(E) = E \text{ and } \bar{L}_{n,d}\Big([0,n] \setminus E\Big) = \Big([0,n] \setminus E\Big). \tag{1}$$

We will use Proposition 8(b) of [19] which states,

Proposition 9.2.4. *For every $d \in [0,N]$ and $E \subseteq [0,N]$, $\bar{L}_{N,d}(E) = E$ if and only if $T_{N,|E|-d} \subseteq E$, where $T_{n,i} \coloneqq [0, i-1] \cup [n-i+1, n]$ for $i \ge 1$ and $T_{n,i} = \emptyset$ for $i \le 0$.*

Note that by the assumption before, $|E| - d \leq n + 1 - |E| - d = |[0, n] \setminus E| - d$. We take cases on the signs of $n + 1 - |E| - d$, $|E| - d$ and show that each case leads to a contradiction.
Case-1 ($n + 1 - |E| - d \leq 0$): Note that $|E| - d \leq n + 1 - |E| - d \leq 0$. This gives,

$$d \geq |E| \text{ and } d \geq n + 1 - |E|.$$

Therefore,

$$d \geq (n + 1)/2.$$

This is not possible as we assumed $d \in [n/4, (n + 1)/3)$.
Case-2 ($n + 1 - |E| - d > 0$ and $|E| - d > 0$): Since $E \cap ([0, n] \setminus E) = \emptyset$, we have

$$T_{n,|E|-d} \cap T_{n,|[0,n]\setminus E|-d} = \emptyset,$$

as $T_{n,|E|-d} \subseteq E$ and $T_{n,|[0,n]\setminus E|-d} \subseteq [0, n] \setminus E$ from Proposition 9.2.4. This gives,

$$\min\Big(|E| - d - 1, n - |E| - d\Big) < 0 \ \Rightarrow \ |E| - d - 1 < 0,$$

which isn't possible as $|E| - d > 0$.
Case-3 ($n + 1 - |E| - d > 0$ and $|E| - d \leq 0$): Since $|E| \leq d$, all inputs with Hamming weight in E can be captured as the zero set of the polynomial,

$$P(x) = \prod_{k \in E} \Big(|x| - k\Big)$$

such that $\deg(P) = |E| \leq d$. We also have $\Big|[0, n] \setminus E\Big| \geq (n + 1 - d)$, by hypothesis a degree d polynomial has to capture this set. Note that

$$T_{n,|[0,n]\setminus E|-d} \subseteq \Big([0, n] \setminus E\Big),$$

from Proposition 9.2.4 and this gives,

$$\Big[0, n - |E| - d\Big] \cup \Big[|E| + d, n\Big] \subseteq [0, n] \setminus E.$$

We can conclude,

$$\begin{aligned} & 2(n - |E| - d + 1) \leq n + 1 - |E| \\ \Rightarrow\ & \frac{(n - |E| + 1)}{2} \leq d \\ \Rightarrow\ & d \geq \max\Big(|E|, \frac{n - |E| + 1}{2}\Big). \end{aligned}$$

The above implies $d \geq (n + 1)/3$, which contradicts $d \in [n/4, (n + 1)/3)$. Since all of the cases above give a contradiction, we conclude $\mathrm{rdeg}(f) \geq (n+1)/3$ for all non-constant, symmetric f. $\square$

Remark. *When $n = 3k - 1$, the bound rdeg$(f) \geq (n + 1)/3$ is tight. For $n = 3k$ and $n = 3k - 2$, Proposition 9.2.3 implies the lower bounds $(n + 3)/3$ and $(n + 2)/3$ (resp.), and constructions* *similar to Proposition 9.2.2 show that these bounds are tight!*

9.3 General Lower Bound for Non-degenerate functions

We will show the following theorem,

Theorem 9.3.1. *For any non-degenerate Boolean function f, rdeg$(f) \geq (\log(n))^{(1/2)-o(1)}$.*

The proof of the above will require the following two lemmas.

Lemma 9.3.2. *For any non-degenerate Boolean function f with sensitivity s, we have*

$$\mathit{rdeg}(f) \geq \Omega\Big(\frac{\log(s)}{\log\log(s)}\Big).$$

Lemma 9.3.3. *For any non-degenerate Boolean function f with sensitivity s, we have*

$$\mathit{rdeg}(f) \geq \Omega\Big(\frac{\log(n/s^{O(1)})}{\log(s)}\Big).$$

It is clear that the Lemmas 9.3.2 and 9.3.3 imply Theorem 9.3.1. We will use the following (slightly modified) definition of rank from [12],

Definition 9.3.4 (Rank)**.** *For a degree d multilinear polynomial of the form,*

$$P(x_1, \dots, x_n) \coloneqq \sum_{S\subset[n];|S|\leq d} a_S \prod_{j\in S} x_j \tag{2}$$

the rank of P, denoted by rank(P), is the largest integer r such that there exist disjoint sets $S_1, ..., S_r \subseteq [n]$ of size d with $|a_{S_j}| \geq 1$, for $j \in [r]$.

Note that the above definition of rank is similar to the definition of r' in Section 2.4. We will need the following result from [12, Theorem 1.6],

Theorem 9.3.5. *There is an absolute constant B such that the following holds for all d, n. Let P be a polynomial of the form Equation (2) whose rank $r \geq 2$. Then for any interval I of length 1,*

$$Pr\Big(P(\xi_1, \dots, \xi_n) \in I\Big) \leq \min\Big(\frac{Bd^{4/3}\sqrt{\log(r)}}{r^{1/4d+1}}, \frac{\exp(Bd^2(\log\log r)^2)}{\sqrt{r}}\Big).$$

For the application of the above in the proof of Lemma 9.3.2, we will restrict ourselves only to the case when $I = \{0\}$, i.e. I is an interval of length 0. WLOG, we can further scale the nonzero coefficients of the polynomial so that all coefficients have modulus at least 1. We will also need the following result, which directly follows from Theorem 2.5.1:

Theorem 9.3.6 ([8])**.** *For every Boolean function f,*

$$D(f) \leq s(f)^{12} \leq D(f)^{12}.$$

The above simply states that the decision tree height and sensitivity of a Boolean function are polynomially related.

Proof of Lemma 9.3.2. Assume for contradiction,

$$\text{rdeg}(f) \leq d < \frac{\log(s)}{11 \log\log(s)}.$$

This gives, $2d^{10d+2} < s$. Since $\text{rdeg}(f) = \max(\text{ndeg}(f), \text{ndeg}(\bar{f}))$, there exist non-deterministic polynomials P_1 and P_2 for f and $\bar{f}$ (resp.) such that $\deg(P_1) \leq d$ and $\deg(P_2) \leq d$. By definition of rank in Definition 9.3.4, it follows that each maxonomial of P_i ($i = 1, 2$) intersects atleast one of the disjoint $\text{rank}(P_i)$ maxonomials. If the rank of either polynomial P_1 or P_2 is less than d^{10d}, randomly assign all variables in the (disjoint) rank maxonomials of that polynomial to 0 or 1. This random restriction forces the degree of one of P_1 or P_2 to decrease by atleast 1, and number of variables assigned to 0 or 1 is at most $\text{rank}(P_i)\deg(P_i) \leq d^{10d}.d \leq d^{10d+1}$. Repeat this process: whenever one of the polynomials P_1 or P_2 has small rank (i.e. $\leq d^{10d}$), randomly assign all variables in the disjoint rank maxonomials to 0 or 1, and consider the function obtained from these restrictions. This process terminates because $\deg(P_1) + \deg(P_2)$ decreases at each step. Ultimately, we either reach a saturation point where $\text{rank}(P_i) \geq d^{10d}$ for $i = 1, 2$, or one of P_1, P_2 becomes constant. The latter case is not possible as $D(f) \geq s > 2d^{10d+2}$, and height of the decision tree of f decreases by at most the number of variables (randomly) assigned at each step. Specifically, the height reduces by at most the number of variables in the disjoint rank maxonomials $\leq d^{10d}.d = d^{10d+1}$ in each step of the process. As the process terminates in at most $2d$ steps, the decision tree has height atleast $s - 2d.d^{10d+1} > 0$ at the end, implying neither P_1 nor P_2 can become constant. Meanwhile, the former case leads to a contradiction via the anti-concentration bound:

Claim. *At any step in the sequence of random restrictions, $\text{rank}(P_i) \geq d^{10d}$ for $i = 1, 2$ is not possible.*

Proof of Claim. Assume for contradiction that $\text{rank}(P_i) \geq d^{10d}$ for $i = 1, 2$. By Theorem 9.3.5,

$$\Pr(P_i = 0) \leq \frac{Bd^{4/3}\sqrt{\log(r)}}{r^{1/4d+1}} \leq \frac{Bd^{4/3}\sqrt{10d\log(d)}}{d^{5/2}} < \frac{1}{2}$$

for d large enough, and $i = 1, 2$. However,

$$\Pr(P_1 = 0) + \Pr(P_2 = 0) = 1,$$

since P_1 and P_2 are non-deterministic polynomials for f and $\bar{f}$, respectively. This is a contradiction, and the claim follows. □

As neither case is possible at the saturation step following the sequence of random restrictions, this contradicts our initial hypothesis. Therefore, $\text{rdeg}(f) \geq \Omega(\log(s)/\log\log(s))$ holds for all Boolean and non-degenerate f. □

The proof of Lemma 9.3.3 essentially mimics the proof of [17, Lemma 16, Section 3.2].

Proof of Lemma 9.3.3. The proof of the lemma is in two steps. In the first step, we use a counting argument to prove a lower bound on the rational degrees of random functions $\boldsymbol{F} : \{0,1\}^m \to \{0,1\}$ which are chosen from a distribution such that for a large subset $X \subseteq \{0,1\}^m$, the random variables

$\{\boldsymbol{F}(x) \mid x \in X\}$ are independently and uniformly chosen random bits. In the second step, we show how any f as in the statement of Lemma 9.3.3 can be randomly restricted to a random function $\boldsymbol{F}$ where the lower bound for random functions applies.

We now show the lower bound for random functions and use it to prove Lemma 9.3.3.

Lemma 9.3.7. *The following holds for positive integer parameters m, M and d such that $M > m^{10d}$. Let $\boldsymbol{F} : \{0,1\}^m \to \{0,1\}$ be a random function such that for some $X \subseteq \{0,1\}^m$ with $|X| = M$, the random variables $(\boldsymbol{F}(x))_{x \in X}$ are independent and uniformly distributed random bits. Then we have,*

$$\Pr_{\boldsymbol{F}}\Big(\mathrm{rdeg}(\boldsymbol{F}) \leq d\Big) < \frac{1}{10}.$$

Proof of Lemma 9.3.7. The proof is via a counting argument. We will show the stronger inequality,

$$\Pr_{\boldsymbol{F}}\Big(\mathrm{ndeg}(\boldsymbol{F}) \leq d\Big) < \frac{1}{10}.$$

Consider the following proposition (similar to [17, Lemma 21]),

Proposition 9.3.8. *Consider the system of linear equations $\{Ax = 0\}$ over the rational numbers, where A is an $(p_1 \times q)$-Boolean matrix. Let A' be another $(p_2 \times q)$-Boolean matrix, and consider the system of inequations $(A'x)_i \neq 0$ for $1 \leq i \leq p_2$. If the combined system of equations and inequations admits a non-trivial real solution, it has a non-trivial rational solution that can be specified (as a list of numerator-denominator pairs in binary) by at most $10q^7$ bits.*

Proof of Proposition 9.3.8. We will use the following [16, Corollary 3.2d],

Theorem 9.3.9 (Schrijver)**.** *Let A be a rational $(m \times n)$-matrix, and let b be a rational column vector such that each row of the matrix $\begin{bmatrix} A & b \end{bmatrix}$ has size atmost φ. If $\{Ax = b\}$ has a solution, then*

$$\{x \mid Ax = b\} = \{x_0 + \lambda_1 x_1 + \cdots + \lambda_t x_t \mid \lambda_1, \ldots, \lambda_t \in \mathbb{R}\}$$

for certain rational vectors $x_0, x_1, \ldots, x_t$ of size at most $4n^2\varphi$.

For the given system of equations $\{Ax = 0\}$, note that $x_0 = 0$ and therefore,

$$\{x \mid Ax = 0\} = \{\lambda_1 x_1 + \cdots + \lambda_t x_t \mid \lambda_1, \ldots, \lambda_t \in \mathbb{R}\},$$

where $t = \dim(\mathrm{Null}(A)) \leq q$ and the vectors x_i's are all of bit complexity $\leq 10q^3$. By the given hypothesis, there exists non-zero $\lambda = (\lambda_1, \ldots, \lambda_t) \in \mathbb{R}^t$ such that

$$(A'(\lambda_1 x_1 + \cdots + \lambda_t x_t))_i \neq 0 \tag{3}$$

for all $1 \leq i \leq p_2$. Set $y_i = A'x_i$, and note that each vector y_i has bit complexity $\leq 20q^3$ since the matrix A' is Boolean. Equivalently, we can rewrite Equation (3) as the condition that each coordinate of the vector

$$v_\lambda := \sum_{i=1}^{t} \lambda_i y_i$$

is non-zero. Depending on the sign of each (non-zero) coordinate of the vector v_λ, we can build an $(p_2 \times t)$-matrix B with rational entries such that

$$B\lambda < 0. \tag{4}$$

Note that the matrix B has bit complexity $\leq \sum_{i=1}^{t} \text{size}(y_i) \leq 20q^4$. Since λ is a non-trivial real solution to Equation (4), there exists $n \in \mathbb{N}$ large enough such that

$$B(n\lambda) = n(B\lambda) \leq \begin{bmatrix} -1, & \dots & ,-1 \end{bmatrix}^\top.$$

We will need the following proposition,

Proposition 9.3.10. *Let A be a rational $(m \times n)$-matrix, and let b be a rational column vector such that each row of the matrix $\begin{bmatrix} A & b \end{bmatrix}$ has size atmost φ. If the system of inequalities $\{Ax \leq b\}$ has a non-trivial solution in $\mathbb{R}^n$, then it has a non-trivial solution in $\mathbb{Q}^n$ of bit complexity $\leq 4n^2\varphi$.*

Proof of Proposition 9.3.10. If $Ax \leq b$ has a real solution, then there is a solution inside a minimal face of the polyhedron $P = \{x \in \mathbb{R}^n \mid Ax \leq b\}$ defined by the system of linear inequalities $Ax \leq b$ [16, Theorem 8.5]. A minimal face is a non-empty subset $F \subseteq P$ of the form

$$F = \{x \in \mathbb{R}^n \mid A'x = b'\}$$

where $A'x \leq b'$ is a subset of the inequalities in $Ax \leq b$ such that the matrix A' has the same rank as A [16, Theorem 8.4]. The proposition now follows from Theorem 9.3.9, by considering a rational solution of small bit complexity on one of the minimal faces of the polyhedron P. □

By Proposition 9.3.10, we can find $\lambda' = (\lambda'_1, \dots, \lambda'_t) \in \mathbb{Q}^t$ of small bit complexity in terms of the matrix B (i.e. $\leq 200q^6$) which satisfies

$$B\lambda' \leq \begin{bmatrix} -1, & \dots & ,-1 \end{bmatrix}^\top < \begin{bmatrix} 0, & \dots & ,0 \end{bmatrix}^\top.$$

In particular, each coordinate of the vector $v_{\lambda'} = \sum_{i=1}^{t} \lambda'_i y_i$ is non-zero. Therefore,

$$\sum_{i=1}^{t} \lambda'_i x_i$$

is a rational solution to the combined system of equations and inequations specified by the matrices A and A' (resp.), and this solution has bit complexity $\leq 400q^6 \leq 10q^7$. Proposition 9.3.8 follows. □

Let $F : \{0,1\}^m \to \{0,1\}$ be such that $\text{ndeg}(F|_X) \leq d$, i.e. there exists a polynomial P of degree $\leq d$ which acts as a non-deterministic polynomial for $F|_X$ on the set X. To apply Proposition 9.3.8, consider the task of finding a non-deterministic polynomial Q of degree d that vanishes precisely when $F|_X = 0$ and remains non-zero elsewhere on the set X. Let $X' \subseteq X$ denote the set where $F|_X$ vanishes, while $F|_X$ is non-zero at every point in $X \setminus X'$. The coefficients of such a polynomial Q solve a linear system of $p_1 := |X'|$ many linear equations in $q := \binom{m}{\leq d}$ variables, and also satisfy the system of inequations $(A'x)_i \neq 0$ for $1 \leq i \leq p_2$ where $p_2 := |X \setminus X'|$. By the existence of the polynomial P, this combined system of equations and inequations has a non-trivial real solution. By Proposition 9.3.8, we know that there is a rational solution to the whole system of bit-complexity at most $10q^7 \leq m^{8d} < 9M/10$. Let $\mathcal{B}$ be the set of all polynomials of deg $\leq d$, and of bit complexity $\leq 9M/10$. Given a polynomial $Q \in \mathcal{B}$, there exists a unique $f : X \to \{0,1\}$ for which Q acts as a non-deterministic polynomial. We conclude,

$$\Pr_{\boldsymbol{F}}\Big(\mathrm{ndeg}(\boldsymbol{F}) \leq d\Big) \leq \Pr_{\boldsymbol{F}}\Big(\mathrm{ndeg}(\boldsymbol{F}|_X) \leq d\Big) \leq \frac{|\mathcal{B}|}{2^M}, \tag{5}$$

since $\mathrm{ndeg}(F|_X) \leq \mathrm{ndeg}(F)$ and the random variables $(F(x) : x \in X)$ are independently and uniformly distributed. By definition, $|\mathcal{B}| \leq 2^{9M/10}$. Plugging this into Equation (5), we get

$$\Pr_{\boldsymbol{F}}\Big(\mathrm{ndeg}(\boldsymbol{F}) \leq d\Big) \leq \frac{2^{9M/10}}{2^M} < \frac{1}{10}.$$

This finishes the proof of Lemma 9.3.7. □

We will use Lemma 9.3.7 to prove Lemma 9.3.3. This proof again breaks into two smaller steps.

1. Show that, after a projection, f turns into something similar to an addressing function, that we will call a *Pseudoaddressing function*.

2. Show that any pseudoaddressing function has large rational degree.

As any projection g of f satisfies $\mathrm{rdeg}(g) \leq \mathrm{rdeg}(f)$, the above implies a lower bound on $\mathrm{rdeg}(f)$, hence proving Lemma 9.3.3. To make the above precise, we need the following definition.

Definition 9.3.11 (Pseudoaddressing function)**.** *A function $g : \{0,1\}^{r+t} \to \{0,1\}$ is an (r,t)-Pseudoaddressing function if the input variables to g can be partitioned into two sets $Y = \{y_1, \ldots, y_r\}$ and $Z = \{z_1, \ldots, z_t\}$ and g can be computed by a decision tree T with the following properties.*

$(P1)$ *For each $z_j \in Z$, there are two root-to-leaf paths π_j^0 and π_j^1 in T that diverge at a node labeled z_j and lead to outputs 0 and 1 respectively.*

$(P2)$ *All the other nodes on these paths are labeled by variables in Y, and further these variables take the same values on both paths. In particular, π_j^0 and π_j^1 differ only on the value of z_j.*

In analogy with the Addressing function, given an (r,t)-pseudoaddressing function as above, we refer to the variables in Y as the addressing variables and the variables in Z as the addressed variables. The two steps of the proof as outlined above can now be formalized as follows.

Proposition 9.3.12. *Let f be as in the statement of Lemma 9.3.3. Then, there exist $r \leq s^{O(1)}$ and $t \geq n/s^{O(1)}$ and a projection $\nu : [n] \to [r+t]$ such that $g = f|_\nu$ is an (r,t)-pseudoaddressing function.*

Proposition 9.3.13. *Let g be any (r,t)-pseudoaddressing function. Then, rdeg$(g) = \Omega(\log t/\log r)$.*

The above propositions immediately imply Lemma 9.3.3. For the proofs of Propositions 9.3.12 and 9.3.13, refer to Claims 19 and 20 (resp.) in [17, Lemma 16]. □

9.4 Product of Large Rank Polynomials is Non-Zero

We will show the following result assuming Ramsey theory.

Theorem 9.4.1. *For any $d \in \mathbb{N}$, there exists (large enough) $r = r(d) \in \mathbb{N}$ such that for any two polynomials P, Q satisfying $\deg(P) \leq d$ and $\deg(Q) \leq d$:*

$$\Big[\mathit{rank}(P) \geq r(d),\ \mathit{rank}(Q) \geq r(d)\Big] \;\Rightarrow\; \Big[PQ \neq 0\Big].$$

Remark. *Note that $r(d)$ in the above theorem depends only on the degree parameter d, and is independent of the number of variables in the polynomials P, Q.*

Before proving the theorem, we will look at the simpler case of the above when $d = 1$. In this case, $\operatorname{rank}(P) \geq 5$ and $\operatorname{rank}(Q) \geq 5$ implies $PQ \neq 0$, where P and Q are linear polynomials. Assume not, i.e. there exist linear polynomials P, Q such that $\operatorname{rank}(P) \geq 5$, $\operatorname{rank}(Q) \geq 5$ and $PQ = 0$. First, note that if a variable x_i only occurs in the polynomial P and not in Q, then the product $PQ \neq 0$. Therefore, both P and Q share the same set of variables with non-zero coefficients. Applying pigeonhole principle twice with respect to P and Q gives that there exist distinct variables x_i and x_j such that the corresponding linear terms have non-zero coefficients in both P and Q, and further these coefficients have the same signs in each of the polynomials P and Q. Observe that this implies the coefficient of the term $x_i x_j$ in PQ is non-zero, which is a contradiction to our initial assumption. To prove Theorem 9.4.1 for higher degree polynomials, the goal would be to extend this 'sign argument'.

Proof of Theorem 9.4.1. □

References

[1] Scott Aaronson, Shalev Ben-David, Robin Kothari, Shravas Rao, and Avishay Tal. Degree vs. approximate degree and quantum implications of huang's sensitivity theorem. In *Proceedings of the 53rd Annual ACM SIGACT Symposium on Theory of Computing*, pages 1330–1342, 2021.

[2] Sourav Chakraborty. On the sensitivity of cyclically-invariant boolean functions. In *20th Annual IEEE Conference on Computational Complexity (CCC'05)*, pages 163–167. IEEE, 2005.

[3] Ronald De Wolf. Nondeterministic quantum query and communication complexities. *SIAM Journal on Computing*, 32(3):681–699, 2003.

[4] Steve Fisk. A very short proof of cauchy's interlace theorem for eigenvalues of hermitian matrices. *arXiv preprint math/0502408*, 2005.

[5] Parikshit Gopalan. *Computing with polynomials over composites.* Georgia Institute of Technology, 2006.

[6] Parikshit Gopalan, Venkatesan Guruswami, and Richard J Lipton. Algorithms for modular counting of roots of multivariate polynomials. In *LATIN 2006: Theoretical Informatics:*

7th Latin American Symposium, Valdivia, Chile, March 20-24, 2006. Proceedings 7, pages 544–555. Springer, 2006.

[7] Craig Gotsman and Nathan Linial. The equivalence of two problems on the cube. *Journal of Combinatorial Theory, Series A*, 61(1):142–146, 1992.

[8] Pooya Hatami, Raghav Kulkarni, and Denis Pankratov. Variations on the sensitivity conjecture. *arXiv preprint arXiv:1011.0354*, 2010.

[9] Hao Huang. Induced subgraphs of hypercubes and a proof of the sensitivity conjecture. *Annals of Mathematics*, 190(3):949–955, 2019.

[10] Vishnu Iyer, Siddhartha Jain, Matt Kovacs-Deak, Vinayak M Kumar, Luke Schaeffer, Daochen Wang, and Michael Whitmeyer. On the rational degree of boolean functions and applications. *arXiv preprint arXiv:2310.08004*, 2023.

[11] Qian Li and Xiaoming Sun. On the modulo degree complexity of boolean functions. *Theoretical Computer Science*, 818:32–40, 2020.

[12] Raghu Meka, Oanh Nguyen, and Van Vu. Anti-concentration for polynomials of independent random variables. *arXiv preprint arXiv:1507.00829*, 2015.

[13] Marvin L Minsky and Seymour A Papert. Perceptrons: expanded edition, 1988.

[14] Noam Nisan. Crew prams and decision trees. In *Proceedings of the twenty-first annual ACM symposium on Theory of computing*, pages 327–335, 1989.

[15] Noam Nisan and Mario Szegedy. On the degree of boolean functions as real polynomials. *Computational complexity*, 4:301–313, 1994.

[16] A Schrijver. *Theory of linear and integer programming.* John Wiley & Sons, 1998.

[17] Srikanth Srinivasan and S Venkitesh. On the probabilistic degree of an n-variate boolean function. *arXiv preprint arXiv:2107.03171*, 2021.

[18] Avishay Tal. Properties and applications of boolean function composition. In *Proceedings of the 4th conference on Innovations in Theoretical Computer Science*, pages 441–454, 2013.

[19] S Venkitesh. Covering symmetric sets of the boolean cube by affine hyperplanes. *arXiv preprint arXiv:2107.10385*, 2021.