Electronic Colloquium on Computational Complexity, Report No. 22 (2004)

The Degree of Threshold mod 6 and Diophantine Equations

Nayantara Bhatnagar
College of Computing
Georgia Tech
nand@cc.gatech.edu

Parikshit Gopalan
College of Computing
Georgia Tech
parik@cc.gatech.edu

Richard J. Lipton
College of Computing
Georgia Tech
rjl@cc.gatech.edu

March 31, 2004

Abstract

We continue the study of the degree of polynomials representing threshold functions modulo 6 initiated by Barrington *et al.*[1]. We use the framework established in [4] relating representations by symmetric polynomials to simultaneous protocols. We show that proving bounds on the degree of Threshold functions is equivalent to counting the number of solutions to certain Diophantine equations. This allows us to use tools from number theory to study such polynomial representations.

When k is a fixed constant, we show that the degree of the Threshold-k function (T_k) is $O(n^{\frac{1}{2}+\varepsilon})$ for any $\varepsilon > 0$. The proof uses a result of Filaseta [6] on factors of numbers of the form $N(N+d)$. We show an upper bound of $O(n^{\frac{1}{t}+\varepsilon})$ when m has t distinct prime factors using a theorem due to Granville [8], which generalizes Filaseta's result but which assumes the *abc* conjecture [9]. We show that for $t = 2$, the *abc* conjecture implies that the degree of T_k is $O(nk)^{\frac{1}{2}+\varepsilon}$.

We show a lower bound of $\Omega(n^{\frac{1}{t}}k^{\frac{t-1}{t}})$ for strong representations of T_k over $\mathbb{Z}_m$. This improves the previous bound of $\Omega(\max(k, n^{\frac{1}{t}}))$ [17]. When $t = 2$, it nearly matches the upper bound of $O(nk)^{\frac{1}{2}+\varepsilon}$. Further, when $t = 2$, we also show a similar weak lower bound. These lower bounds are proved by constructing solutions to the equations in question using a pigeonhole argument.

The $O(\sqrt{n})$ upper bound for the OR function can be interpreted as follows: For suitably chosen parameters (k_2, k_3) if $0 \leq w \leq n$ and $w \bmod 2^{k_2}$ and $w \bmod 3^{k_3}$ are both zero, then in fact $w = 0$. Our bounds for T_k give a similar result about the size of w: For n sufficiently large and suitably chosen parameters (k_2, k_3), if the residues $w \bmod 2^{k_2}$ and $w \bmod 3^{k_3}$ are both less than k, then in fact they are both equal to w itself and $w < k$. Conversely, if $w \geq k$, then one of the residues must be large.

ISSN 1433-8092

1 Introduction

Representations of Boolean functions as polynomials over various rings such as $\mathbb{R}$ and $\mathbb{Z}_m$ have been well studied in computer science starting with the work of Minsky and Papert [13]. In addition to having several applications to complexity theory and learning theory, this study has produced many surprising results and challenging open questions (see the survey by Beigel [2]). One of the complexity theoretic motivations for studying polynomials over $\mathbb{Z}_m$ is to understand the power of modular counting. Razborov [14] and Smolensky [15] prove strong lower bounds for AC_0 with Mod-p gates when p is a prime. In contrast, proving lower bounds for circuits with mod_6 gates is an important open problem. A first step towards this problem might be to better understand the computational power of polynomials over $\mathbb{Z}_6$.

The study of the degree of polynomials that represent threshold functions modulo 6 was initiated by the seminal work of Barrington, Beigel and Rudich [1] who proved the surprising result that the OR function can be strongly represented over $\mathbb{Z}_6$ by a polynomial of degree $\Theta(\sqrt{n})$. It can be seen that AND requires degree $\Omega(n)$ for strong representation but can be weakly represented by polynomials of degree $\Theta(\sqrt{n})$. Define the threshold k function T_k to be 1 if the input contains k or more 1s and 0 otherwise. OR and AND correspond to T_1 and T_n respectively. This raises the following natural question:

What is the (strong/weak) degree of Threshold k is for $1 < k < n$?

1.1 Our Results

We show that proving bounds on the degree of T_k for $1 < k < n$ is equivalent to showing that certain Diophantine equations have only finitely many solutions. More precisely, we show that there exists a strong protocol for T_k on n variables with parameters k_2, k_3 iff there are no non-trivial solutions to the equation

$$|a2^{k_2} - b3^{k_3}| = \ell \qquad\qquad a2^{k_2} \leq n,\ b3^{k_3} \leq n,\ \ \ell < k$$

Bounds on the degree of the OR and AND functions follow directly from the Chinese Remainder Theorem (CRT). Resolving the strong degree for other values of k is equivalent to much harder questions about Diophantine equations.

When k is a fixed constant, we show an upper bound of $O(n^{\frac{1}{2}+\varepsilon})$ for any $\varepsilon > 0$. The proof uses a result of Filaseta [6] on factors of numbers of the form $N(N+d)$ which implies that the above equation has only finitely many solutions. We show an upper bound of $O(n^{\frac{1}{t}+\varepsilon})$ when m has t distinct prime factors using a theorem due to Granville [8], which is a generalization of Filaseta's result but which assumes the *abc* conjecture from number theory [9]. We also show that for all values of k, when $t = 2$, the *abc* conjecture implies that the degree of T_k is bounded by $O(nk)^{\frac{1}{2}+\varepsilon}$.

The $O(\sqrt{n})$ upper bound for the OR function can be interpreted as follows: For suitably chosen parameters (k_2, k_3) if w mod 2^{k_2} and w mod 3^{k_3} are both zero, then in fact the number w must equal 0. Our bounds for T_k give a similar result about the size of w: For suitably chosen parameters (k_2, k_3) if the residues w mod 2^{k_2} and w mod 3^{k_3} are both less than k, then in fact they are both equal to w itself and $w < k$. Conversely, if $w \geq k$, then one of the residues must be large.

We show a lower bound of $\Omega(n^{\frac{1}{t}}k^{\frac{t-1}{t}})$ for strong representations of T_k over $\mathbb{Z}_m$. This improves the previous bound of $\Omega(\max(k, n^{\frac{1}{t}}))$ [17]. When $t = 2$, the lower bound nearly matches the upper bound of $O(nk)^{\frac{1}{2}+\varepsilon}$ for all values of k. Further, when $t = 2$, we show a weak lower bound of $O(nk)^{\frac{1}{2}+\varepsilon}$. The lower bounds are proved through a pigeonhole argument which shows that for appropriate settings of parameters, the equations in question do have solutions. In [4], it is shown that T_k can be represented by probabilistic polynomials of deg $O(\max(k, \sqrt{n}))$ over $\mathbb{Z}_6$. Hence our deterministic lower bound of $\Omega(\sqrt{nk})$ shows that even for symmetric polynomials, probabilistic representations can have lower degree than deterministic representations.

1.2 Previously Known Results

In this section we give the basic definitions and survey some known results about of representations of Boolean functions as polynomials over $\mathbb{Z}_m$.

There are many possible definitions of what it means for a polynomial to represent a Boolean function over $\mathbb{Z}_m$. In what follows, the inputs to the polynomial are 0-1 inputs denoted by $X = X_1, X_2, \cdots, X_n$.

Definition 1.1 *Polynomial P 0-1 represents function f if $P(X) = f(X)$.*

Definition 1.2 *Polynomial P strongly represents function f if $f(X) = 0 \Rightarrow P(X) = 0$ and $f(X) = 1 \Rightarrow P(X) \neq 0$.*[1]

Definition 1.3 *Polynomial P weakly represents function f if $f(X) \neq f(Y) \Rightarrow P(X) \neq P(Y)$.*

Such representations are easy to understand when p is a prime and $\mathbb{Z}_p$ is a field. Using the fact that every function from $\mathbb{Z}_p \rightarrow \mathbb{Z}_p$ is a polynomial, we can obtain a 0-1 representation from either a strong or a weak representation while increasing the degree only by a constant. While there could be many polynomials that 0-1 represent f, their degrees can vary by at most a constant. This can be proved using Fermat's little theorem. Hence proving lower bounds is not a challenge, since any polynomial representing f is only a constant factor away from optimal.

Over $\mathbb{Z}_m$ when m is not a prime power however, things are very different. For convenience, we take $m = 6$. Let $P(X)$ be a 0-1 representation of OR. Let $P_2(X) \equiv P(X) \bmod 2$ and $P_3(X) \equiv P(X) \bmod 3$. Then both $P_2(X)$ and $P_3(X)$ are 0-1 representations of OR over Z_2 and Z_3 and have degree $\Omega(n)$. For a strong or a weak representation however, this no longer holds. Indeed Barrington *et al.* [1] proved the surprising result that the OR function can be strongly represented by a symmetric polynomial of degree $O(\sqrt{n})$ over $\mathbb{Z}_6$ showing that such representations are greatly different from 0-1 representation mod 6. They prove that $O(\sqrt{n})$ is the best possible for symmetric polynomials. It is not known if this bound is optimal for general polynomials. While there is no better upper bound, the best lower bound is $O(\log n)$ due to Tardos and Barrington [16]. Grolmusz uses this upper bound to construct a super-polynomial size set system where the size of each set is 0 mod 6 but all pairwise intersections are nonzero mod6. He uses this to construct explicit Ramsey graphs whose parameters approach those of the best known construction [12].

There has been considerable success in proving lower bounds in the strong representation. Lower bounds of $\Omega(n)$ are known in the strong representation for some functions using general (not just symmetric) polynomials [1, 17, 10]. Tsai shows a lower bound of k on the degree of T_k [17]. However as pointed out by [16] the task of proving lower bounds for strong representations is simplified by the fact that P must output 0 whenever f is 0. The weak representation seems a more natural definition and here far less is known with regard to lower bounds. The best lower bound known in this case for general polynomials is $\Omega(\log n)$ [11, 16].

Representations using symmetric polynomials were studied by the authors in [4]. There it was shown that representations using symmetric polynomials are equivalent to certain two player simultaneous protocols. The protocol framework allows the use of ideas from communication complexity to show $\Omega(n)$ lower bounds on weak representations of some Threshold and Mod functions by symmetric polynomials. In addition, it shows that questions about the degree of threshold functions are equivalent to questions in number theory about exponential Diophantine equations. Using this connection, an upper bound of $o(n)$ was shown on the strong degree of T_c for any constant c. A lower bound of $\max(k, \sqrt{n})$ was shown for T_k. Since Tsai shows a lower bound of k for T_k with general polynomials, this bound can be inferred by

[1] Tardos and Barrington [16] use the terminology *one-sided representation* for what we call *strong representation*

combining his result with the lower bound for the OR function [1]. Using this framework, Beigel [3] shows that an upper bound of $O(\sqrt{nk})$ holds for infinitely many n for $k \leq c\sqrt{n}$. His result is unconditional.

The remainder of this paper is organized as follows. In section 2, we state without proof some results from [4] relating symmetric polynomial representations to simultaneous protocols. We show upper bounds in Section 3. We prove lower bounds in Section 4.

2 Symmetric Polynomials and Simultaneous Protocols

Simultaneous communication protocols were first defined by Yao in [18]. In this model Alice receives an input x, Bob receives an input y and they wish to compute $f(x, y) \in \{0, 1\}$. They cannot directly communicate with each other. They simultaneously write messages on a blackboard. A referee reads the messages and decides the output. There is an equivalence between symmetric polynomials representing f over $\mathbb{Z}_m$ and certain simultaneous communication protocols for computing the function f.

As a first step towards showing this equivalence, consider symmetric polynomials over $\mathbb{Z}_p$. Every symmetric polynomial $P(X)$ over $\mathbb{Z}_p$ computes a function $f : \{0, 1, \cdots n\} \to \mathbb{Z}_p$ where $f(w)$ is the value of P on a 0-1 input of weight w. Functions that can be computed by a low degree symmetric polynomial in $\mathbb{Z}_p[X]$ are exactly those than can be computed from the first few digits of the base p representation of the weight. This is a consequence of a classical result in number theory called Lucas' Theorem [7] which tells us how to evaluate binomial coefficients modulo p. This is made precise by the following theorem.

Theorem 2.1 *The symmetric functions $f : \{0, 1\}^n \to \mathbb{Z}_p$ that can be computed by a symmetric polynomial $P[X] \in \mathbb{Z}_p[X]$ of degree $d < p^l$ are exactly those functions that can be computed from $w(X) \bmod p^l$.*

Equivalently these functions can be computed from the l least significant digits of $w(X)$ in base p. A similar equivalence holds over $\mathbb{Z}_{p^a}$. We now introduce the notion of a protocol for computing a Boolean function $f : w \in \{0, \cdots n\} \to \{0, 1\}$. For convenience, we take $m = 6$.

Definition 2.2 *A strong protocol for computing f is a simultaneous protocol involving two players P_2 and P_3. P_2 is given $i \equiv w \bmod 2^{k_2}$ as input and outputs $P_2(i)$ in $\mathbb{Z}_2$. P_3 is given $j \equiv w \bmod 3^{k_3}$ as input and outputs $P_3(j)$ in $\mathbb{Z}_3$. If $f(w) = 0$, then both players must output 0. If $f(w) = 1$, at least one player must output a non-zero value. The cost of the protocol is $\max(2^{k_2}, 3^{k_3})$.*

Definition 2.3 *A weak protocol is defined similarly except that if $f(w) \neq f(w')$ then at least one player should give different outputs on w and w'.*

In both cases, the two players cannot communicate but they can agree on a procedure beforehand. They compute their outputs independently of one another and write them on a blackboard. A referee then reads their outputs and decides if the value of the function is 0 or 1. In a strong protocol, the referee's strategy is fixed, he outputs 0 iff both players say 0. In a weak protocol, the referee can choose any strategy. For m with t distinct prime factors $p_1, \cdots, p_t$, we define protocols with t players where player P_i reads the input in base p_i.

We can now make the connection between symmetric polynomials and simultaneous protocols. By the CRT, a degree d symmetric polynomial $P(X)$ over $\mathbb{Z}_6$ corresponds to symmetric polynomials $P_2(X)$ and $P_3(X)$ over $\mathbb{Z}_2$ and $\mathbb{Z}_3$ respectively whose degrees are at most d. By Theorem 2.1 this means that the function computed by P can be computed from the residues of w mod 2^{k_2} and 3^{k_3} where these are the smallest powers of 2 and 3 which exceed d. Conversely assume there exists a low cost protocol for f. By Theorem 2.1 the function computed by each player can be represented by a low degree symmetric polynomial. We now use the CRT to combine these polynomials and get a low degree polynomial over $\mathbb{Z}_6$. This gives us the following theorem.

Theorem 2.4 *There exists a symmetric polynomial over $\mathbb{Z}_6$ of degree d that strongly (weakly) represents f iff there exists a strong (weak) protocol of cost $\Theta(d)$ for computing f.*

This theorem allows us to prove both upper and lower bounds on the degrees of polynomials for both representations by viewing them as simultaneous communication protocols. We first need some notation. Player P_2 receives $i \equiv w \bmod 2^{k_2}$ and P_3 receives $j \equiv w \bmod 3^{k_3}$. They wish to compute $f(w)$. If $2^{k_2}3^{k_3} \leq n$ there might be multiple values of w between 0 and n satisfying the congruences for i and j. If $f(w)$ is not the same for all these values, then clearly no protocol with parameters k_2, k_3 exists.

Assume that the value of f is well defined for every pair (i, j). We define a $2^{k_2} \times 3^{k_3}$ input matrix $A = a_{ij},\ 0 \leq i < 2^{k_2},\ 0 \leq j < 3^{k_3}$.

$$\begin{aligned} a_{ij} &\equiv i \bmod 2^{k_2} \\ a_{ij} &\equiv j \bmod 3^{k_3} \end{aligned}$$

P_2 receives the same input i for all inputs in the same row of A and hence outputs the same value. Similarly inputs in a column are indistinguishable to P_3. For a function f, we then define the $2^{k_2} \times 3^{k_3}$ matrix A^f as below.

$$A^f_{ij} = \begin{cases} f(a_{ij}) & 0 \leq a_{ij} \leq n \\ \text{x} & a_{ij} > n \end{cases}$$

The symbol x indicates that the function is not defined for this value of weight. We wish to know whether for given parameters (k_2, k_3), there exists a protocol for f. We can give a combinatorial characterization in terms of the matrix A^f. The next two lemmas are proved by adapting standard results about deterministic simultaneous protocols.

Lemma 2.5 *There is a strong protocol for f with parameters k_2, k_3 iff $\forall\ i,\ j$ such that $f(a_{ij}) = 1$, either there are no 0s in row i or there are no 0s in column j of A^f.*

Definition 2.6 *Two rows i, i' in the matrix A^f are distinct, if there exists a column index j such that $a_{ij}, a_{i'j} \leq n$ and $f(a_{ij}) \neq f(a_{i'j})$. Rows $i_1, \cdots, i_k$ are said to be distinct if they are pairwise distinct.*

Lemma 2.7 *For a weak protocol for f over $\mathbb{Z}_{pq}$ with parameters (k_p, k_q) to exist, the matrix A^f must have at most p distinct rows and q distinct columns.*

3 Upper Bounds for Threshold

We wish to know if there exists a protocol for T_k on n variables with parameters k_2, k_3. The following lemma gives a necessary condition on k_2, k_3.

Lemma 3.1 *[4] Any strong protocol for T_k has cost $\Omega(\max(k, \sqrt{n}))$.*

Proof: Suppose $2^{k_2}3^{k_3} \leq n$. Choose $a < k \leq a + 2^{k_2}3^{k_3}$. Both players receive the same inputs for these weights but $T_k(a) = 0$ while $T_k(a + 2^{k_2}3^{k_3}) = 1$. This proves a lower bound of $\sqrt{n}$.

Now suppose $\max(2^{k_2}, 3^{k_3}) < k$. Consider any $w \geq k$. Since $j \equiv w \bmod 2^{k_2}$ and $2^{k_2} < k, j < k$. Similarly $i < k$. The entry i lies in the same row as a while j lies in the same column.

$$\begin{aligned} T_k(w) &= 1 \\ T_k(i) &= 0 \\ T_k(j) &= 0 \end{aligned}$$

Now apply lemma 2.5. Hence $\max(2^{k_2}, 3^{k_3}) \geq k$. ∎

We now prove a theorem that equates showing degree bounds on threshold to the number of solutions to certain families of equations. Note that we already have a lower bound of $\max(k, \sqrt{n})$ by Corollary 3.1. Since we wish to minimize the cost of the protocol which is defined as $\max(2^{k_2}, 3^{k_3})$, we will assume that 2^{k_2} and 3^{k_3} are both greater than $\max(k, \sqrt{n})$.

Theorem 3.2 *There exists a strong protocol for T_k on n variables with parameters k_2, k_3 iff the equation*

$$|a2^{k_2} - b3^{k_3}| = \ell \qquad a2^{k_2} \leq n,\ b3^{k_3} \leq n,\ \ \ell < k \tag{1}$$

has no non-trivial solutions.

Proof: As a first step, we show that it suffices to analyze the following strong protocol.

Protocol 3.3 Strong Protocol for Threshold-k

- Let $i \equiv w \bmod 2^{k_2}$. If $i \geq k$, P_2 outputs 1, else P_2 outputs 0.
- Let $j \equiv w \bmod 3^{k_3}$. If $j \geq k$, P_3 outputs 1, else P_3 outputs 0.

In a strong protocol, if $f(w) = 0$ both players must output 0. Hence when $i < k$, P_2 must output 0 since the input could be i. If $i \geq k$, then clearly $w \geq k$, hence P_2 can output 1. Similarly, this is also the best strategy for P_3.

We now analyze inputs on which the protocol fails. Let $w \geq k,\ i < k,\ j < k$. On such inputs, both players output 0 whereas the value of the function is 1, and so the protocol is incorrect. Note that $i \neq j$ since if $i = j$, by the CRT $w = i$, which contradicts the fact that $w \geq k$. But now

$$w = a2^{k_2} + i = b3^{k_3} + j$$

Assume that $i > j$ and let $i - j = \ell$ where $0 < \ell < k$. Then, we have

$$\begin{aligned} b3^{k_3} - a2^{k_2} &= \ell \\ a2^{k_2}, b3^{k_3} \leq w &\leq n \end{aligned}$$

Hence any such input gives a solution to Equation (1).

Conversely, we will show that any solution to Equation (1) for fixed n gives an input w so that the protocol is incorrect. Assume that we have

$$|a2^{k_2} - b3^{k_3}| = \ell \quad s.t.\ \ a2^{k_2} \leq n,\ b3^{k_3} \leq n,\ \ \ell < k$$

Assume $b3^{k_3} > a2^{k_2}$. Set $w = b3^{k_3} = a2^{k_2} + \ell$. From this setting, we obtain

$$\begin{aligned} i &\equiv w \bmod 2^{k_2} = \ell \\ j &\equiv w \bmod 3^{k_3} = 0 \end{aligned}$$

Hence we have $w > 2^{k_2} \geq k$ whereas $i, j < k$ and hence the protocol is incorrect. ∎

In using this theorem, $a = b = 0$ is considered a trivial solution. We will only be interested in non-trivial solutions. As an example, suppose we were trying to show a bound of $n^{\frac{3}{4}}$ on T_2. We set $2^{k_2}, 3^{k_3} > n^{\frac{3}{4}}$. This implies that $a, b < n^{\frac{1}{4}} = (2^{k_2})^{\frac{1}{3}}$. We are looking for solutions to

$$|a2^{k_2} - b3^{k_3}| = 1 \qquad\qquad a < (3^{k_3})^{\frac{1}{3}} \quad b < (2^{k_2})^{\frac{1}{3}}$$

If we relax the constraints on a, b to $a < 3^{k_3}$ and $b < 2^{k_2}$, by the GCD equation, we will have a solution for every value of k_2, k_3 Since $(2^{k_2}, 3^{k_3}) = 1$. We are asking how many solutions exist with the constraint that $a, b < (2^{k_2})^{\frac{1}{3}}$. We will show that the answer is only finitely many.

3.1 Constant Threshold with Two Players

We now prove an upper bound of $O(n^{\frac{1}{2}+\varepsilon})$ for constant threshold when m has two prime factors. This improves the bound of $o(n)$ in [4]. We set $m = 6$ for convenience. We will use the following Theorem by Filaseta which builds on work by Mahler [6].

Theorem 3.4 *Let ℓ be a fixed non-zero integer. Let M be a fixed positive integer. Let $\varepsilon > 0$. Let D be the largest divisor of $N(N - \ell)$ which is relatively prime to M. If N is sufficiently large (depending on ℓ, M and ε), then $D > N^{1-\varepsilon}$.*

Theorem 3.5 *Any symmetric polynomial that strongly represents T_c over $\mathbb{Z}_{pq}$ has degree $O(n^{\frac{1}{2}+\varepsilon})$ $\forall \varepsilon > 0$ for any fixed constant c.*

Proof: We prove the theorem over $\mathbb{Z}_6$.

Set $2^{k_2} \cdot 3^{k_3} > n^{1+\varepsilon}$. We will show with this setting of parameters, Protocol 3.3 works for sufficiently large n. By Theorem 3.2, the protocol for n fails iff there is a solution to

$$|a2^{k_2} - b3^{k_3}| \;=\; \ell \qquad\qquad a2^{k_2} \le n,\ b3^{k_3} \le n,\ \ell < c \tag{2}$$

We first show that for each $\ell < c$, this equation has only finitely many solutions. Set $M = 6$. Take

$$\begin{aligned} N \;&=\; a2^{k_2} \;=\; b3^{k_3} + \ell \\ \Rightarrow\quad N(N-\ell) \;&=\; ab2^{k_2}3^{k_3} \end{aligned}$$

Let D be largest divisor of $N(N - \ell)$ relatively prime to 6. It follows that $D \le ab$. By our setting of parameters,

$$\begin{aligned} 2^{k_2}3^{k_3} > n^{1+\varepsilon} \;&\ge\; N^{1+\varepsilon} \\ ab2^{k_2}3^{k_3} = N(N-\ell) \;&<\; N^2 \\ \Rightarrow \qquad D \le ab &< N^{1-\varepsilon} \end{aligned}$$

By Theorem 3.4, this is possible for only finitely many N. Hence, with fixed ℓ, there are only finitely many solutions. There are only finitely many possibilities for ℓ since $1 \le \ell < c$. Hence Equation 2 has only finitely many solutions in $a2^{k_2}, b3^{k_3}$. This implies an upper bound on n since

$$2^{k_2} \cdot 3^{k_3} \ge n^{1+\varepsilon} \quad \Rightarrow \quad n \le ab2^{k_2}3^{k_3}$$

Hence there are only finitely many solutions in n. Hence Protocol 3.3 works for all sufficiently large n. We can take 2^{k_2} and 3^{k_3} approximately equal to give the desired degree bound. ∎

Theorem 3.6 *For any constant c, the degree of threshold T_c over Z_{pq} is $O(n^{\frac{1}{2}+\varepsilon})$ $\forall \varepsilon > 0$.*

By the CRT, we know that if $2^{k_2}3^{k_3} > n$, and if $w \equiv 0$ modulo 2^{k_2} and 3^{k_3} then in fact $w = 0$. The above theorem states that if $2^{k_2}3^{k_3} > n^{1+\varepsilon}$ for any positive ε, and if the residues of w modulo 2^{k_2} and 3^{k_3} are both less than c then in fact $w < c$ for sufficiently large n.

3.2 Constant Threshold with Multiple Players

In this section we consider the case when m has t distinct prime divisors $p_1, p_2, \cdots p_t$. For T_c with c constant, it is easy to show a lower bound of $\Omega(n^{\frac{1}{t}})$. We will show an upper bound of $O(n^{\frac{1}{t}+\varepsilon})$ for all $\varepsilon > 0$. No upper bounds better than $o(n)$ were previously known for this class of functions.

We will use a result due to Granville which generalizes Filaseta's result. But this result holds only under the assumption of the *abc* conjecture. This is a very powerful conjecture which has many important implications, including an asymptotic version of Fermat's Last Theorem. A survey about the conjecture and its consequences can be found in [9].

Definition 3.7 *The Radical of M denoted by $R(M)$ is the product of distinct primes dividing M.*

Conjecture 3.8 The *abc*-conjecture *[9]: Let $\varepsilon > 0$. If a, b, c are coprime positive integers satisfying $a + b = c$, then*

$$c < D \cdot R(abc)^{1+\varepsilon}$$

where D is constant that depends only on ε.

Theorem 3.9 *[8] Assume the abc-conjecture is true. Suppose that $g(X) \in \mathbb{Z}[X]$ has no repeated roots. For $\varepsilon > 0$, and w sufficiently large,*

$$R(g(w)) > |w|^{deg(g)-1-\varepsilon}$$

Using this result, we analyze the following protocol which is the natural generalization of Protocol 3.3.

Protocol 3.10 Threshold-c with multiple players

- Take $p_1^{k_1} p_2^{k_2} \cdots p_t^{k_t} > n^{1+\varepsilon}$.
- Set $w_i \equiv w \bmod p_i^{k_i}$. If $w_i < c$, Player i outputs 0 else player i outputs 1.

Theorem 3.11 *Let $c > 0$ be a constant and let t be the number of distinct prime factors of m. If the* abc *conjecture is true, there exist symmetric degree $O(n^{\frac{1}{t}+\varepsilon})$ that strongly represent T_c over $\mathbb{Z}_m$.*

Proof: Fix a value of n. If Protocol 3.10 is incorrect for some $w \in \{0, \cdots n\}$. Then it must be that $w \geq c$ but $w_i < c$ for all i. Let $w = a_i p_i^{k_i} + w_i$. Note that we must have $a_i > 0$ for all i since otherwise $w < c$. This gives us a non-trivial solution to the following system of equations in the a_is and n.

$$\begin{aligned} a_i p_i^{k_i} &\leq n \qquad \forall\, i \in \{1, \cdots t\} \\ |a_i p_i^{k_i} - a_j p_j^{k_j}| &< c \qquad \forall\, i < j \end{aligned} \tag{3}$$

We now set

$$g(X) \;=\; X(X-1)\cdots(X-c+1)$$

Clearly $g(X)$ has no repeated roots and we can apply Theorem 3.9. Hence, $\forall \varepsilon > 0$, for all but finitely many n,

$$R(g(w)) > w^{c-1-\varepsilon} \tag{4}$$

We will show that if Protocol 3.10 is incorrect on w, then $g(w)$ is divisible by high prime powers, and so $R(g(w))$ is small, which contradicts Equation (4).

$$g(w) = w(w-1)\cdots(w-c+1)$$

We know that $w - a_i p_i^{k_i} = w_i$ where $0 \leq w_i < c$. Hence for all i,

$$\begin{aligned} & w - w_i \mid g(w) \\ & w - w_i \;=\; a_i p_i^{k_i} \\ \Rightarrow \quad & p_i^{k_i} | g(w) \end{aligned}$$

By the CRT, for a suitable constant C,

$$g(w) = C \prod_i p_i^{k_i}$$

We now bound the size of C.

$$\begin{aligned} \prod_i p_i^{k_i} &> n^{1+\varepsilon} \geq w^{1+\varepsilon} \\ g(w) &= w(w-1)\cdots(w-c+1) < w^c \\ \Rightarrow\; C &= \frac{g(w)}{\prod_i p_i^{k_i}} < w^{c-1-\varepsilon} \end{aligned}$$

This gives an upper bound on $R(g(w))$.

$$\begin{aligned} R(g(w)) &< C p_1 p_2 \cdots p_t \\ &< w^{c-1-\varepsilon} p_1 p_2 \cdots p_t \\ &= w^{c-1-\varepsilon'} \end{aligned}$$

This gives a contradiction to Equation 4. Hence w must be one of only finitely many exceptions. This bounds the value of n since

$$\begin{aligned} w \geq a_i p_i^{k_i} \geq p_i^{k_i} \\ \prod_i p_i^{k_i} > n^{1+\varepsilon} \\ \Rightarrow w^t > n^{1+\varepsilon} \\ \Rightarrow n < w^{\frac{t}{1+\varepsilon}} \end{aligned}$$

Hence there are only finitely many solutions in n and the protocol works correctly for n sufficiently large. The degree bound follows by taking nearly equal powers of p_i. ∎

3.3 Upper Bounds for General Threshold Functions

We now return to the case when m has two prime divisors and show that the $abc-$conjecture implies an upper bound of $O(nk)^{\frac{1+\varepsilon}{2}}$ on T_k for all values of k in the strong representation. We begin with the following technical lemma.

Lemma 3.12 *Assume the* abc *conjecture holds for some $\varepsilon > 0$. There exists a constant $n_0(\varepsilon)$ such that for $n > n_0(\varepsilon)$, the equation*

$$|a2^{k_2} - b3^{k_3}| = \ell \qquad\qquad a2^{k_2} \leq n, \quad b3^{k_3} \leq n, \quad 2^{k_2}3^{k_3} \geq (n\ell)^{1+\varepsilon}$$

has no solutions with $a2^{k_2}, b3^{k_3}, \ell$ relatively prime.

Proof: Assume that we have a solution where $a2^{k_2} > b3^{k_3}$. Applying the *abc* conjecture to the equation $a2^{k_2} = b3^{k_3} + \ell$, we must have

$$D \cdot R(a2^{k_2}, b3^{k_3}, \ell)^{1+\varepsilon} > a2^{k_2} \geq (a2^{k_2}b3^{k_3})^{\frac{1}{2}} \tag{5}$$

where the last inequality holds since $a2^{k_2} > b3^{k_3}$. We can bound $R(a2^{k_2}, b3^{k_3}, \ell)$ by $6ab\ell$. Plugging this bound into (5), for a suitable constant D' depending only on ε, we get

$$D' \cdot (ab\ell)^{1+\varepsilon} > (ab2^{k_2}3^{k_3})^{\frac{1}{2}} \geq (ab)^{\frac{1}{2}}(n\ell)^{\frac{1+\varepsilon}{2}}$$

The last inequality uses the fact that $2^{k_2}3^{k_3} \geq (n\ell)^{1+\varepsilon}$. Rearranging terms,

$$D' \cdot (ab)^{\frac{1}{2}+\varepsilon}\ell^{1+\varepsilon} > (n\ell)^{\frac{1+\varepsilon}{2}} \tag{6}$$

We now upper bound the size of ab.

$$a2^{k_2}b3^{k_3} \leq n^2, \quad 2^{k_2}3^{k_3} \geq (n\ell)^{1+\varepsilon} \quad \Rightarrow ab \leq \frac{n^{1-\varepsilon}}{\ell^{1+\varepsilon}}$$

A calculation now gives the following bound on the LHS of (6).

$$D' \cdot (ab)^{\frac{1}{2}+\varepsilon}\ell^{1+\varepsilon} \leq D'n^{\frac{1+\varepsilon}{2}-\varepsilon^2}\ell^{\frac{1-\varepsilon}{2}-\varepsilon^2} \tag{7}$$

Plugging this bound into (6), we have

$$D'n^{\frac{1+\varepsilon}{2}-\varepsilon^2}\ell^{\frac{1-\varepsilon}{2}-\varepsilon^2} > (n\ell)^{\frac{1+\varepsilon}{2}}$$

For all $n > n_0(\varepsilon)$, this gives a contradiction. Hence for sufficiently large n, the equation has no solutions. ∎

Theorem 3.13 *If the abc– conjecture is true for some $\varepsilon > 0$, there exist symmetric polynomials of degree $O((nk)^{\frac{1+\varepsilon}{2}})$ that strongly represent T_k over Z_6.*

Proof: Note that for a non-trivial bound, we need $\varepsilon < 1$, else $(nk)^{\frac{1+\varepsilon}{2}} = \Omega(n)$ for all k. Take $n > n_0(\varepsilon)$ as in Lemma 3.12. Set $2^{k_2}3^{k_3} \geq (nk)^{1+\varepsilon}$. We claim that there are no solutions to

$$|a2^{k_2} - b3^{k_3}| = \ell \qquad\qquad a2^{k_2} \leq n, \; b3^{k_3} \leq n, \quad \ell < k \tag{8}$$

Assume that a solution exists. If $a2^{k_2}, b3^{k_3}, \ell$ are coprime, then we get a contradiction to Lemma 3.12. Assume that $a2^{k_2}, b3^{k_3}, \ell$ are not coprime. Their GCD can be written as $2^{t_2}3^{t_3}g$ where g is relatively prime to 2 and 3. Dividing throughout we get

$$|a'2^{k_2-t_2} - b'3^{k_3-t_3}| = \ell' \qquad\qquad a'2^{k_2-t_2} \leq n, \; b'3^{k_3} \leq n, \quad \ell' < \frac{k}{2^{t_2}3^{t_3}}$$

Further, we now have that $a'2^{k_2-t_2}, b'3^{k_3-t_3}, \ell'$ are relatively prime. To apply Lemma 3.12, we need to check that $2^{k_2-t_2}3^{k_3-t_3} \geq (n\ell')^{1+\varepsilon}$. It is easy to see that this condition does hold.

$$2^{k_2-t_2}3^{k_3-t_3} \geq \frac{(nk)^{1+\varepsilon}}{2^{t_2}3^{t_3}} \geq \left(\frac{nk}{2^{t_2}3^{t_3}}\right)^{1+\varepsilon} \geq (n\ell')^{1+\varepsilon}$$

However, by Lemma 3.12, our choice of n guarantees that such a solution cannot exist. Hence in fact Equation (8) has no solutions. The degree bound then follows by taking 2^{k_2} and 3^{k_3} nearly equal and applying Theorem 3.2. ∎

While the techniques in this section will allows us to show the same bound for $\mathbb{Z}_{pq}$, we do know how to show an upper bound for the t-player case for $t \geq 3$.

4 Lower Bounds for Threshold-k Functions

4.1 Strong Representations

In this section, we will show a $\Omega(\sqrt{kn})$ lower bound on the strong degree of the T_k function over $\mathbb{Z}_{pq}$. For small ε, this matches the upper bound of the previous section. Over $\mathbb{Z}_m$, when m has t distinct prime factors, we show a lower bound of $\Omega(n^{\frac{1}{t}}k^{1-\frac{1}{t}})$ on the strong degree of T_k.

Theorem 4.1 *The strong degree of T_k over $\mathbb{Z}_6$ is $\Omega(\sqrt{nk})$.*

Proof: Set $2^{k_2}, 3^{k_3} \leq \frac{\sqrt{kn}}{2}$. We will construct solutions to the following equation for all n.

$$|a2^{k_2} - b3^{k_3}| = \ell \qquad\qquad a2^{k_2}, b3^{k_3} \leq n, \ell < k \tag{9}$$

By Theorem 3.2 this implies that the lower bound on the degree of T_k.

We construct the solutions by a pigeonhole argument. By Lemma 3.1 we may assume $2^{k_2}, 3^{k_3} \geq \max(k, \sqrt{n})$. Consider all pairs (u, v) such that $u2^{k_2} \leq n,\ v3^{k_3} \leq n$. We map the pair (u, v) to the point $P_{uv} = u2^{k_2} - v3^{k_3}$, so that $P_{uv} \in [-n, n]$. Each pair u, v is mapped to a distinct point, since if

$$\begin{aligned} & P_{uv} = P_{st},\ (u,v) \neq (s,t) \\ \Rightarrow\ & (u-s)2^{k_2} - (v-t)3^{k_3} = 0 \\ \Rightarrow\ & 2^{k_2}3^{k_3} | (u-s)2^{k_2} \\ \Rightarrow\ & |(u-s)2^{k_2}| > n \end{aligned}$$

However, $|(u-s)2^{k_2}| \leq n$ by our choice of u and s.

We can now count the total number of points $P_{u,v}$. We can take $0 \leq u, v < 2\sqrt{\frac{n}{k}}$. Hence there are $4\frac{n}{k}$ points lying in the interval $[-n, n]$, and hence by the pigeonhole principle, there are two points within a distance of $\frac{(2n+1)k}{4n} < k$. Call them P_{uv} and P_{st}. Hence

$$|(u-s)2^{k_2} - (v-t)3^{k_3}| = \ell \qquad\qquad \ell < k$$

Set $a = u - s$, and $b = v - t$. Assume that $a \geq 0$. This implies that $b \geq 0$, since $2^{k_2} > k, 3^{k_3} > k$ so we cannot add multiples of 2^{k_2} and 3^{k_3} to get $\ell < k$. Also, $a2^{k_2} \leq u2^{k_2} \leq n$ and similarly $b3^{k_3} \leq v3^{k_3} \leq n$. Hence a, b, ℓ give the desired solution to Equation 9. ∎

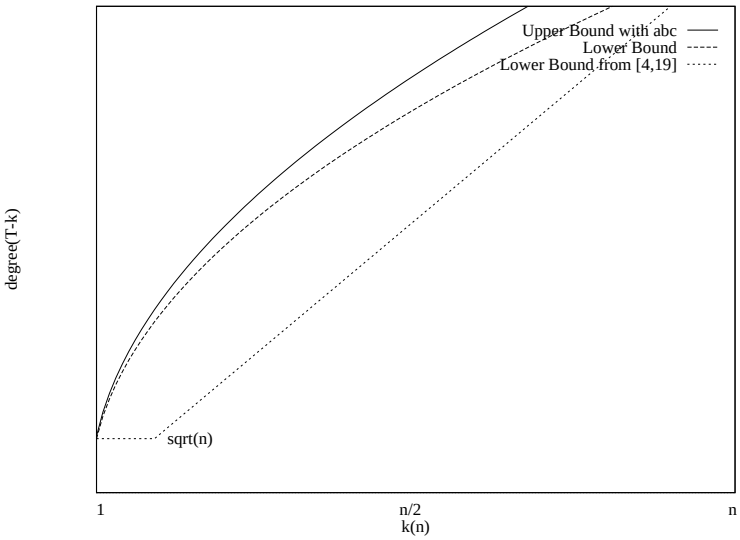

Figure 1: Strong Degree of T_k over $\mathbb{Z}_6$

Theorem 4.2 *The strong degree of T_k over $\mathbb{Z}_{pq}$ is $\Omega(\sqrt{nk})$.*

Note that the lower bound of $\sqrt{nk}$ almost matches the upper bound of $(nk)^{\frac{1}{2}+\varepsilon}$ implied by the $abc-$conjecture. In personal communication, Beigel shows unconditionally that the bound of $O(\sqrt{nk})$ holds for *infinitely many* n for $k \leq c\sqrt{n}$ for some constant c [3].

We now generalize this proof to get a lower bound for the t- player case. We will construct a *lattice* in $t-1$ dimensional space and argue that it contains short vectors using the pigeonhole principle. The set of points is not really a lattice because we will not allow all integer linear combinations, but place restrictions on the size of the scalars.

Theorem 4.3 *Let $p_1, \cdots p_t$ be the prime divisors of m. The strong degree of T_k over $\mathbb{Z}_m$ is $\Omega(n^{\frac{1}{t}}k^{1-\frac{1}{t}})$.*

Proof: Let $p_i^{k_i} < \frac{1}{3}n^{\frac{1}{t}}k^{1-\frac{1}{t}}$ $\forall i$. We will construct solutions to the equation

$$\begin{aligned} \forall\ i, \quad a_i p_i^{k_i} &\leq n \\ \forall\ i \neq j, \quad |a_i p_i^{k_i} - a_j p_j^{k_j}| &< k \end{aligned} \tag{10}$$

By Theorem 3.2, this will imply the desired lower bound.

By Lemma 3.1 we may assume that $p_i^{k_i} > k, n^{\frac{1}{t}}$ $\forall i$. We define t vectors $v_1, \cdots v_t$ in $t-1$ dimensions.

$$\begin{aligned} v_1 &= (p_1^{k_1}, p_1^{k_1}, \cdots, p_1^{k_1}) \\ v_2 &= (p_2^{k_2}, 0, \cdots, 0) \\ v_i &= (0,\ 0,\ p_i^{k_i}, 0) \\ v_t &= (0, 0, \cdots, p_t^{k_t}) \end{aligned}$$

For $i = 1, \cdots t$, consider b_i such that $b_i p_i^{k_i} \leq n$. We map every such t-tuple $b = (b_1, b_2, \cdots, b_t)$ to a point P_b in $t-1$ dimensional space.

$$\begin{aligned} P_b &= b_1 v_1 - b_2 v_2 \cdots - b_t v_t \\ &= (b_1 p_1^{k_1} - b_2 p_2^{k_2}, b_1 p_1^{k_1} - b_3 p_3^{k_3}, \cdots, b_1 p_1^{k_1} - b_t p_t^{k_t}) \end{aligned}$$

We can use the fact that $p_1^{k_1}p_2^{k_2}\cdots p_t^{k_t} > n$ to show that if $b \neq c$, then $P_b \neq P_c$. For each i we can take $0 \leq b_i < 3(\frac{n}{k})^{1-\frac{1}{t}}$. This gives a total of $3^t(\frac{n}{k})^{t-1}$ points. Since each co-ordinate of P_b lies between $[-n, n]$, every point lies in $[-n, n]^{t-1}$ which is a cube of volume $(2n+1)^{t-1}$. We can partition this cube into $\lceil\frac{2n+1}{k-1}\rceil^{t-1} < (\frac{3n}{k})^{t-1}$ smaller cubes with each side of length $k-1$. However there are $3^t(\frac{n}{k})^{t-1}$ distinct points. By the pigeonhole principle, two points lie in the same cube of side $k-1$. Call these points P_b and P_c. This implies for $2 \leq i \leq t$ we have

$$|(b_1 - c_1)p_1^{k_1} - (b_i - c_i)p_i^{k_i}| \leq k-1$$

Assume that $b_1 \geq c_1$. Since $p_i^{k_i} > k\ \forall i$, this implies $b_i \geq c_i\ \forall i$. We set $a_i = b_i - c_i$.

$$\begin{aligned} \forall i, \quad a_i p_i^{k_i} &\leq b_i p_i^{k_i} &&\leq n \\ \forall\ i \neq j, \quad |a_i p_i^{k_i} - a_j p_j^{k_j}| &&&< k \end{aligned}$$

Hence we get a solution to Equation 10. ∎

4.2 Weak Representations

In [4] we show a lower bound of $\Omega(\max(k, \sqrt{n}))$ on the degree of symmetric polynomials weakly representing T_k over $\mathbb{Z}_6$ when $k \leq \frac{n}{6}$. In what follows, we improve this bound by applying the results obtained above on the strong degree of T_k.

Theorem 4.4 *Any symmetric polynomial weakly representing the threshold function T_k over $\mathbb{Z}_6$ has degree $\Omega(\sqrt{nk})$ for $k \leq \frac{n}{2}$.*

Proof: We apply the construction in the proof of Theorem 4.1 with $\frac{n}{2}$ and $\frac{k}{2}$. Set $2^{k_2}, 3^{k_3} \leq \frac{\sqrt{kn}}{4}$. There exist a, b and ℓ satisfying the following equation.

$$|a2^{k_2} - b3^{k_3}| = \ell \qquad\qquad a2^{k_2}, b3^{k_3} \leq \frac{n}{2}, \ell < \frac{k}{2} \tag{11}$$

We show that there does not exist a weak protocol for T_k of cost $\max(2^{k_2}, 3^{k_3})$. By Lemma 2.7 it suffices to show that A^{T_k} has a submatrix with 3 distinct rows. We use solutions to Equation (11) to construct this submatrix. Assume $a2^{k_2} \geq b3^{k_3}$. By Lemma 3.1 we may assume $2^{k_2}, 3^{k_3} \geq \max(k, \sqrt{n})$ and hence $a2^{k_2} \geq k$. We choose the submatrix V of A

$$\begin{aligned} V &= \begin{pmatrix} 0 & a2^{k_2} & 2\cdot a2^{k_2} \\ \times & a2^{k_2} - b3^{k_3} & 2\cdot a2^{k_2} - b3^{k_3} \\ \times & \times & 2(a2^{k_2} - b3^{k_3}) \end{pmatrix} \\ &= \begin{pmatrix} 0 & a2^{k_2} & 2\cdot a2^{k_2} \\ \times & \ell & \ell + a2^{k_2} \\ \times & \times & 2\ell \end{pmatrix} \\ \Rightarrow\ V^{T_k} &= \begin{pmatrix} 0 & 1 & 1 \\ \times & 0 & 1 \\ \times & \times & 0 \end{pmatrix} \end{aligned}$$

We need to ensure that all entries in the fooling set are valid. The largest entry in the fooling set is $2\cdot a2^{k_2}$. From Equation (11), we have $2\cdot a2^{k_2} \leq n$. By Lemma 2.7 a weak protocol cannot exist since V^{T_k} has at

least 3 distinct columns. Hence $\max(2^{k_2}, 3^{k_3}) > \frac{\sqrt{nk}}{4}$. Note that $2a2^{k_2} \leq n$, on the other hand, $a2^{k_2} \geq k$. Combining the inequalities, we obtain $k \leq \frac{n}{2}$. ∎

In general, over $\mathbb{Z}_{pq}$ we have the following theorem:

Theorem 4.5 *Any symmetric polynomial weakly representing the threshold function T_k over $\mathbb{Z}_{pq}$ where $p < q$ has degree $\Omega(\sqrt{nk})$ for $k \leq \frac{n}{p}$.*

We believe that this bound too holds for $k \leq \frac{n}{2}$. It is natural to ask if one can show linear bounds for all $k > \frac{n}{2}$. The next theorem shows that the answer is no (see Figure 2). It explains the remark in the introduction that the weak degree of the AND function is $\Theta(\sqrt{n})$.

Theorem 4.6 *The weak degree of the threshold function T_k is equal to the weak degree of T_{n-k+1}.*

Proof: Assume that there is a weak protocol for T_k where the players read k_2 and k_3 digits respectively. On an input w, let

$$i \equiv w \bmod 2^{k_2}, \quad j \equiv w \bmod 3^{k_3}$$

Since both players know the value of n, they can compute

$$i' \equiv (n-i) \bmod 2^{k_2} \equiv (n-w) \bmod 2^{k_2}$$
$$j' \equiv (n-j) \bmod 3^{k_3} \equiv (n-w) \bmod 3^{k_3}$$

Now if the players use the protocol for T_k with the values i' and j' instead, they can differentiate the values w such that $n - w < k$ and $n - w \geq k$. This is then a weak protocol differentiating values of $w \geq n-k+1$ and $w < n-k+1$ of cost $\max(2^{k_2}, 3^{k_3})$. A symmetric argument shows that a weak protocol for T_{n-k+1} gives a weak protocol for T_k. By Theorem 2.4 shows the weak degrees of T_k and T_{n-k+1} are equal. ∎

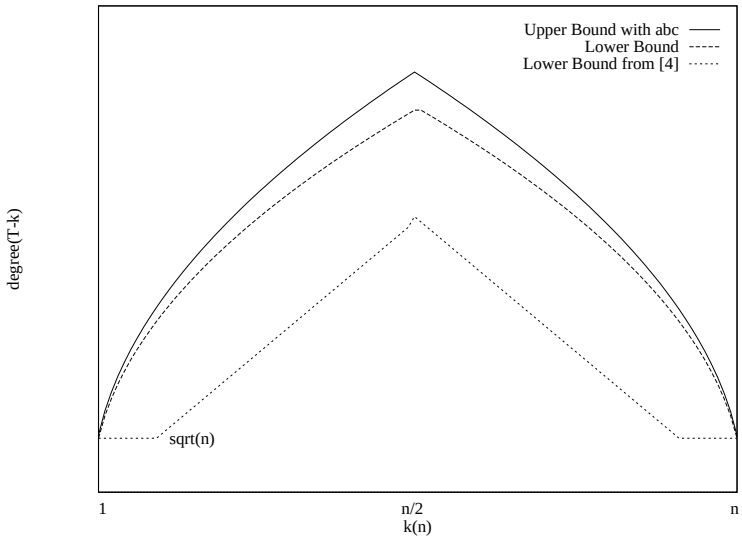

Figure 2: Weak Degree of T_k over $\mathbb{Z}_6$

Corollary 4.7 *Assume the abc-conjecture is true for some $0 < \varepsilon < 1$. The weak degree of the threshold function T_k over $\mathbb{Z}_6$ for $k > \frac{n}{2}$ is bounded by $(n(n-k+1))^{\frac{1+\varepsilon}{2}}$.*

This shows that for $k > \frac{n}{2}$, there is a gap between the strong and weak degree. This raises the question: Is there a $k \leq \frac{n}{2}$ where there is a gap between the strong and weak degree? Also, is there a k such that the degree of T_k over $\mathbb{Z}_6$ and $\mathbb{Z}_{15}$ differ significantly? To answer these questions, we will use the following upper bound for weak protocols (also observed by Beigel in [3]).

Theorem 4.8 *The weak degree of T_c over $\mathbb{Z}_{pq}$ where $c < \min(p, q)$ is $O(\sqrt{n})$.*

Proof: We prove the bound for T_2 over $\mathbb{Z}_{15}$, the general bound is similar. We give a weak protocol for T_2, where the players read k_3 and k_5 digits of the weight in base 3 and 5 respectively. On input w, let $i \equiv w \bmod 3^{k_3}$ and $j \equiv \bmod 5^{k_5}$.

Weak Protocol for T_2

- Let $\sqrt{n} < 3^{k_3} \leq 3\sqrt{n}$ and $\sqrt{n} < 5^{k_5} \leq 5\sqrt{n}$.
- If $i = 0$, $P_3(i) = 0$. If $i = 1$, $P_3(i) = 1$. If $i \geq 2$, $P_3(i) = 2$.
- If $j = 0$, $P_5(j) = 0$. If $j = 1$, $P_5(j) = 1$. If $j \geq 2$, $P_3(j) = 2$.
- The referee outputs 0 if $P_3(i) = P_5(j) < 2$, and 1 otherwise.

We argue that this is a weak protocol protocol for T_2. If the referee answers 0, either $i = j = 0$ or $i = j = 1$. In both cases, since $3^{k_3}, 5^{k_5} > \sqrt{n}$, by the CRT, $w = 0$ or $w = 1$ respectively. On the other hand, if the referee answers 1, there are two cases. In the first case $i \neq j$, in which case w was not 0 or 1. In the second case, $i = j \geq 2$. Thus we have a weak protocol of cost at most $5\sqrt{n}$. ∎

In comparison, the best strong upper bound we have for T_2 over $\mathbb{Z}_{15}$ is $n^{\frac{1}{2}+\varepsilon}$ for all sufficiently large n. The best strong lower bound is $\sqrt{n}$ so this does not prove that there is a gap. But clearly, the weak upper bound is much easier to prove. A similar comparison can be made between the weak degree of T_2 over $\mathbb{Z}_6$ and $\mathbb{Z}_{15}$ since the only upper bound we have over $\mathbb{Z}_6$ is the strong upper bound of $O(n^{\frac{1}{2}+\varepsilon})$.

5 Conclusions

We have shown that resolving the degree of Threshold functions for symmetric polynomials is equivalent to questions regarding Diophantine equations. These are rather hard questions and it does not seem that tight upper bounds can be shown unconditionally. Is showing tight bounds on threshold for general polynomials as hard? Perhaps we run into hard number theoretic questions because we are restricted to symmetric polynomials and proving upper bounds with general polynomials is easy. We do not believe that this is the case, but we cannot rule out this possibility. Proving lower bounds on the other hand can only be harder for general polynomials. The fact that the best known lower bound for OR is $\Omega(\log n)$ suggests that indeed lower bounds are much harder for general polynomials.

We conclude with some open problems.

- Is it possible to improve on the lower bound of k due to Tsai [17] for T_k with general polynomials? A bound of $\Omega(\sqrt{nk})$ for all k would resolve the degree of OR over $\mathbb{Z}_6$, which has been open for a while.
- Show upper bounds on T_k over Z_{pq} either unconditionally or with weaker assumptions (For instance Beigel [3] shows an upper bound of $O(\sqrt{nk})$ unconditionally for $k = O(\sqrt{n})$. Show an upper bound (even conditional) on T_k for 3 or more players that beats $O(\sqrt{nk})$ (The bound of $O(\sqrt{nk})$ follows from the 2 player case) .

- Grolmusz [12] uses the $O(\sqrt{n})$ upper bound on OR to construct a super polynomial size set system with restricted intersection mod 6 and explicit Ramsey Graphs. Can one obtain similar constructions by using upper bounds on T_k for $k > 1$?
- The following question is related to a question raised by Grolmusz in [12]. In all our strong protocols, each player outputs either 0 or 1. What about protocols where both players cannot simultaneously say 1? It is not hard to show an $\Omega(n)$ lower bound for symmetric polynomials representing OR with this restriction. Can one show a better lower bound for general polynomials representing OR with this restriction?
- Consider the permutation on $\{1, \cdots, 3^{k_3} - 1\}$ defined by $\sigma(a) = a2^{k_2} \bmod 3^{k_3}$. The bounds in this paper can be interpreted as saying that this permutation behaves like a random permutation in some respects. Permutations of the kind $\sigma(a) = k \cdot a \bmod m$, where $k \in \mathbb{Z}_m^*$, are well studied and are known to be quasirandom [5]. Can this be used to say anything interesting about the degree of T_k?

Acknowledgments

We thank Ernie Croot for many enlightening discussions and pointers to literature. He also suggested the proof of Theorem 3.11. We thank Michael Filaseta for the pointer to [6] and for help with the proof of Theorem 3.5. We thank Richard Beigel for telling us about his results [3].

References

[1] David A. Barrington, Richard Beigel, and Steven Rudich. Representing boolean functions as polynomials modulo composite numbers. *Computational Complexity*, 4:367–382, 1994.

[2] Richard Beigel. The polynomial method in circuit complexity. *Structures in Complexity Theory: 8th Annual Conference*, pages 82–95, 1993.

[3] Richard Beigel. Personal communication. 2003.

[4] Nayantara Bhatnagar, Parikshit Gopalan, and Richard J. Lipton. Symmetric polynomials over $\mathbb{Z}_m$ and simultaneous communication protocols. *Proceedings of the 44^{th} Annual Symposium on the Foundations of Computer Science, 450-459. Available as ECCC report TR03-047.*, 2003.

[5] Joshua Cooper. Survey of quasirandomness in number theory. *Available online at http://arxiv.org/abs/math.NT/0310384*, 2003.

[6] Michael Filaseta. A generalization of an irreducibility theorem of I. Schur. *Acta Arithmetica*, 58(3):251–272, 1991.

[7] Andrew Granville. Arithmetic properties of binomial coefficients. *Canadian Mathematical Society Conference Proceedings*, 20:253–275, 1997.

[8] Andrew Granville. *abc* means we can count squarefrees. *International Mathematical Research Notices*, 19:1224–1231, November 1998.

[9] Andrew Granville and Thomas J. Tucker. It's as easy as *abc*. *Notices of the AMS*, 49(10):991–1009, 2002.

[10] Frederic Green. Complex fourier technique for lower bounds on the mod-m degree. *Computational Complexity*, 9:16–38, 2000.

[11] Vince Grolmusz. On the weak mod m representation of boolean functions. *Chicago Journal of Theoretical Computer Science*, 2, 1995.

[12] Vince Grolmusz. Superpolynomial size set-systems with restricted intersections mod 6 and explicit ramsey graphs. *Combinatorica*, 20(1):71–86, 2000.

[13] Marvin Minsky and Seymour Papert. *Perceptrons: an Introduction to Computational Geometry.* MIT Press, 1968.

[14] Alexander Razborov. Lower bounds for the size of circuits of bounded depth with basis $\{\wedge, \oplus\}$. *Methematical Notes of the Academy of Science of the USSR*, (41):333–338, 1987.

[15] Roman Smolensky. Algebraic methods in the theory of lower bounds for boolean circuit complexity. *Proceedings of the Nineteenth Annual ACM Symposium on Theoretical Computer Science.*, pages 77–82, 1987.

[16] Gabor Tardos and David Barrington. A lower bound on the mod 6 degree of the or function. *Computational Complexity*, 7:99–108, 1998.

[17] Shi-Chun Tsai. Lower bounds on representing boolean functions as polynomials in $\mathbb{Z}_m$. *SIAM Journal of Discrete Mathematics*, 9:55–62, 1996.

[18] Andrew C. Yao. Some complexity questions related to distributive computing. *Proceedings of the 11th Annual ACM Symposium on Theory of Computation*, pages 209–213, 1979.

ECCC ISSN 1433-8092
http://www.eccc.uni-trier.de/eccc
ftp://ftp.eccc.uni-trier.de/pub/eccc
ftpmail@ftp.eccc.uni-trier.de, subject 'help eccc'