Exponential Sums and Circuits with a Single Threshold Gate and Mod-Gates

Frederic Green
Department of Mathematics and Computer Science
Clark University
Worcester, Massachusetts 01610
fgreen@black.clarku.edu

Abstract

Consider circuits consisting of a threshold gate at the top, Mod_m gates at the next level (for a fixed m), and polylog fan-in AND gates at the lowest level. It is a difficult and important open problem to obtain exponential lower bounds for such circuits. Here we prove exponential lower bounds for restricted versions of this model, in which each Mod_m-of-AND subcircuit is a symmetric function of the inputs to that subcircuit. We show that if q is a prime not dividing m, the Mod_q function requires exponential size circuits of this type. This generalizes recent results and techniques of Cai, Green and Thierauf [CGT] (which held only for $q = 2$) and Goldmann (which held only for depth two threshold over Mod_m circuits). As a further generalization of the [CGT] result, the symmetry condition on the Mod_m sub-circuits can be relaxed somewhat, still resulting in an exponential lower bound. The basis of the proof is to reduce the problem to estimating an exponential sum, which generalizes the notion of "correlation" studied by [CGT]. This identifies the type of exponential sum that will be instrumental in proving the general case. Along the way we substantially simplify previous proofs.

1 Introduction

Proving exponential lower bounds against families of certain bounded depth circuits presents complexity theory with some formidable problems. There was much success in the early and middle 80's with the breakthrough results of Furst, Saxe, and Sipser [FSS], Yao [Yao 85], Cai [Cai] and Håstad [Has] for $AC^{(0)}$ circuits, and those of Razborov [Raz] and Smolensky [Sm] for $AC^{(0)}$ with Mod_{p^k} gates for a fixed prime power p^k. Since then, there has been little progress for bounded-depth circuits containing Mod_m gates for general composite m (called ACC circuits [Bar]). On the other hand, there have been some quite dramatic upper bounds for ACC ([Yao 90], [BT], [GKRST]). Building on work of Toda [Tod], Yao [Yao 90] showed that ACC can be simulated by depth 2 probabilistic circuits consisting of a symmetric gate over polylog-size AND's. Beigel and Tarui [BT] improved this by showing that the simulation can be deterministic. Green, Köbler, Regan, Schwentick and Torán [GKRST] further showed that the top gate can be one that computes a particular symmetric function (the so-called "MidBit" function that outputs the middle bit of the sum of the inputs).

In Allender's original work [Al] proving the circuit analogue of Toda's theorem, it was shown that $AC^{(0)}$ can be simulated by quasi-polynomial size circuits with a threshold at the root, a layer of parity gates below that, and a layer of polylog-size AND's attached to the inputs. (In terms of polynomial-time classes, this is equivalent to $PH \subseteq PP^{\oplus P}$ relative to all oracles). Such circuits appear to be weaker than the symmetric over AND's of Beigel and Tarui [BT] or the MidBit over AND's of [GKRST]. Nevertheless, in view of the surprising power of the middle bit representation of ACC, it is natural to ask, can the upper bounds on ACC be improved? In particular, can ACC be simulated by depth three circuits of the type just mentioned, of quasi-polynomial size, with a threshold on top, parity gates at the next level, and small AND's at the bottom?

The intuitive answer to this question is "no." The purpose of this note is to report on progress towards proving this conjecture. In fact, it is natural to suspect that even the Mod_3 function cannot be so computed. (This would imply an oracle A such that $Mod_3P^A \not\subseteq PP^{\oplus P^A}$, which is significant since no oracle is known separating $PP^{\oplus P}$ from PSPACE.) There is some evidence to support this. Cai, Green and Thierauf [CGT] proved an exponential bound for the parity function against circuits with a threshold over Mod_m's (for odd m) over small AND's, but with a severe restriction on the Mod-of-AND subcircuits: Each Mod-of-AND subcircuit computes a *symmetric* function of the inputs. As severe as this restriction is, it is reasonable to determine if our circuit can take direct advantage of the fact that parity is a symmetric function, and the lower bound says that it cannot. The key contribution of [CGT] was to show that the general problem (i.e., with unrestricted polynomials) reduces to an important type of problem that often arises in number theory: estimating an exponential sum.

Unfortunately, it was not at all clear from [CGT] how to obtain lower bounds for any mod function other than parity. The quantity computed in [CGT], the "correlation," is really only appropriate for lower bounds on the parity function. The main contribution of this paper is to develop a technique for treating other mod functions, again via exponential sums. Using this formulation, before generalizing [CGT], we present a new proof of a result of Goldmann [Go]: The Mod_q function requires exponential size on circuits consisting of a threshold over Mod_m gates, where q and m are relatively prime. (This theorem was extended to apply to thresholds of unbounded weight by Krause and Pudlák [KP].) The advantage of this approach is that the proof reduces to a routine calculation. A very simple proof of the main lemma of [CGT] is presented here, which makes this paper almost entirely self-contained. Finally, the main theorem shows that the symmetry constraint on the Mod-of-AND subcircuits can be relaxed a little bit. One may allow the Mod-of-AND gates to be symmetric in pairwise disjoint subsets of inputs, while still obtaining an exponential lower bound. In a final section the outlook for the general case is discussed in light of these new results.

The introduction concludes with some remarks regarding notation and conventions. As is standard, any circuit should be understood to be an element of an infinite family of circuits, one for each number of inputs n. A Mod_m gate over n Boolean inputs $x_1, ..., x_n$ computes the Mod_m function, which returns 1 iff $\sum_{i=1}^{n} x_i \not\equiv 0\ (m)$ (we use the slightly more compact notation $a \equiv b\ (m)$ for $a \equiv b \pmod{m}$). A threshold gate over n Boolean inputs $x_1, ..., x_n$ with weights $w_i \in \mathbf{Z}$ $(1 \leq i \leq n)$ and threshold $t \in \mathbf{Z}$, with $0 \leq t \leq n$, returns 1 iff $\sum_{i=1}^{n} w_i x_i \geq t$. Our results apply to threshold gates with weights polynomially bounded in n. As in [GKRST], if G is a Boolean gate, G^+ denotes a family of circuits with G at the root and AND gates of polylog fan-in at the input level. If G is a Boolean gate and $\mathcal{C}$ is a circuit class, G-of-$\mathcal{C}$ denotes the class of circuits with $\mathcal{C}$-type circuits serving as inputs to G-type gates. Thus the main concern of this paper is threshold-of-Mod_m^+ circuits. Note that the size of the $\mathcal{C}$ subcircuits is to be regarded as a function of the number of inputs to the global G-of-$\mathcal{C}$ circuit.

2 Results

The proof of the main theorem is based on the "ϵ-discriminator" lemma of Hajnal, Maass, Pudlák, Szegedy and Turán [HMPST]. It is the basis for most of the lower bound proofs on depth three (or more) circuits with a threshold on top (e.g., [HMPST], [Gr], [HG]).

Lemma 2.1. Let T be a threshold circuit consisting of a threshold gate over subcircuits $c_1, ..., c_s$, each taking up to n inputs. Thus, T outputs 1 on input $\mathbf{x} \in \{0,1\}^n$ if and only if $\sum_{i=1}^{s} c_i(\mathbf{x}) \geq t$, for some integer t which is fixed for the circuit T. Let T compute the Boolean function $f : \{0,1\}^n \rightarrow \{0,1\}$. Let $A \subseteq f^{-1}(1)$ and $B \subseteq f^{-1}(0)$, and define the quantity,

$$\Delta_i = Pr(c_i(\mathbf{x}) = 1 | \mathbf{x} \in A) - Pr(c_i(\mathbf{x}) = 1 | \mathbf{x} \in B),$$

where Pr denotes the probability over the uniform distribution over all input settings. If $|\Delta_i| \leq \epsilon$ for all $1 \leq i \leq s$, then $s \geq \epsilon^{-1}$.

The lemma is useful because if we can show that $|\Delta_i|$ is exponentially small, then the circuit T computing f is exponentially large. Most of the work is in computing a good upper bound for $|\Delta_i|$. But the first crucial step is finding suitable sets A and B. In the case considered here, in which f is the Mod_q function, we choose sets which correspond to the distribution used by Goldmann [Go], giving the most unbiased possible situation. Let

$$\begin{aligned} A &= \{\mathbf{x} | \sum_{i=1}^{n} x_i \equiv 1\ (q)\} \\ B &= \{\mathbf{x} | \sum_{i=1}^{n} x_i \equiv 0\ (q)\} \end{aligned}$$

Thus, for the type of circuits considered here, namely threshold-of-Mod_m^+, the quantity Δ has the form,

$$\Delta = Pr(c = 1 | \sum_i x_i = 1\ (q)) - Pr(c = 1 | \sum_i x_i = 0\ (q)), \tag{1}$$

where c denotes a Mod_m^+ circuit.

For Mod_m^+ circuits, the evaluation of Δ reduces to the evaluation of a certain exponential sum. The next lemma gives this reduction. Before stating it, we need some notation. For any natural number m, $\zeta_m = e^{\frac{2\pi i}{m}}$ denotes the complex primitive m^{th} root of unity. Recall that,

$$\sum_{a=0}^{m-1} \zeta_m^{ak} = \begin{cases} m & \text{if } k \equiv 0\ (m) \\ 0 & \text{otherwise.} \end{cases}$$

Let $t : \{0,1\}^n \to \mathbf{Z}$ be an integer polynomial. The exponential sum $S(n, m, q; t)$ is defined as,

$$S(n, m, q; t) = \frac{1}{m2^n} \sum_{\mathbf{x} \in \{0,1\}^n} \sum_{a=0}^{m-1} \sum_{b=0}^{q-1} (\zeta_q^{-b} - 1)\zeta_q^{b\sum_i x_i} \zeta_m^{at(\mathbf{x})} \tag{2}$$

Since $S(n, m, q; t)$ has so many arguments, when it is clear from the context we will simply write S.

Lemma 2.2. Let n be any natural number, q a prime which does not divide m, and c a Mod_m-of-AND circuit over n inputs. Let $t : \{0,1\}^n \to \mathbf{Z}$ denote the integer polynomial such that, for $\mathbf{x} \in \{0,1\}^n$, $t(\mathbf{x}) \not\equiv 0\ (m)$ if and only if $c(\mathbf{x})$ outputs 1. Define Δ as in equation (1) and $S(n, m, q; t)$ as in equation (2). Then,

$$|\Delta| \leq |S| + 2^{-\Omega(n)}.$$

Proof. To put Δ in the correct form, first consider the individual terms. From the definition of conditional probability, for any integer k,

$$Pr(c = 1 | \sum_i x_i \equiv k\ (q)) = \frac{Pr(c = 1\ \wedge\ \sum_i x_i \equiv k\ (q))}{Pr(\sum_i x_i \equiv k\ (q))}.$$

Now $Pr(\sum_i x_i \equiv 1\ (q))$ and $Pr(\sum_i x_i \equiv 0\ (q))$ are very close; in fact, it is well known [Go] that they are exponentially close to $\frac{1}{q}$:

$$\frac{1}{q} - 2^{-\alpha n} \leq Pr(\sum_i x_i \equiv k\ (q)) \leq \frac{1}{q} + 2^{-\alpha n},$$

for some constant α depending on k. (This fact can be derived from the techniques of this proof, in particular, those which give inequality (4) below.) It follows that,

$$|\Delta| \leq q \left| Pr(c = 1\ \wedge\ \sum_i x_i \equiv 1\ (q)) - Pr(c = 1\ \wedge\ \sum_i x_i \equiv 0\ (q)) \right| + 2^{-\Omega(n)}.$$

For any function $f : \{0,1\}^n \to \{true, false\}$, write $[f(\mathbf{x})] = 1$ if $f(\mathbf{x})$ is *true* and $[f(\mathbf{x})] = 0$ if $f(\mathbf{x})$ is *false*. (Henceforth in this proof, square brackets will be used only this context.) Thus, for example,

$$Pr(c = 1 \ \wedge \ \sum_i x_i \equiv 1 \ (q)) = 2^{-n} \sum_{\mathbf{x} \in \{0,1\}^n} [c = 1 \ \wedge \ \sum_i x_i \equiv 1 \ (q)].$$

Then we may re-write the bound on $|\Delta|$ as,

$$\begin{aligned} |\Delta| &\leq q2^{-n} \left| \sum_{\mathbf{x} \in \{0,1\}^n} \left([c = 1 \ \wedge \ \sum_i x_i \equiv 1 \ (q)] - [c = 1 \ \wedge \ \sum_i x_i \equiv 0 \ (q)] \right) \right| + 2^{-\Omega(n)} \\ &= q2^{-n} \left| \sum_{\mathbf{x} \in \{0,1\}^n} \left([c = 1] \cdot [\sum_i x_i \equiv 1 \ (q)] - [c = 1] \cdot [\sum_i x_i \equiv 0 \ (q)] \right) \right| + 2^{-\Omega(n)} \end{aligned}$$

By hypothesis, there is a polynomial t in $x_1, ..., x_n$ with integer coefficients such that $c(\mathbf{x}) = 1$ iff $t(\mathbf{x}) \not\equiv 0 \ (m)$. Observe, then, that

$$[c(\mathbf{x}) = 1] = 1 - \frac{1}{m} \sum_{a=0}^{m-1} \zeta_m^{at(\mathbf{x})}.$$

Similarly,

$$[\sum_i x_i \equiv k \ (q)] = \frac{1}{q} \sum_{b=0}^{q-1} \zeta_q^{b(\sum_i x_i - k)},$$

for any k. Using these identities,

$$|\Delta| \leq 2^{-\Omega(n)} +$$

$$q2^{-n} \left| \sum_{\mathbf{x} \in \{0,1\}^n} \left((1 - \frac{1}{m} \sum_a \zeta_m^{at(\mathbf{x})}) (\frac{1}{q} \sum_b \zeta_q^{b(\sum_i x_i - 1)}) - (1 - \frac{1}{m} \sum_a \zeta_m^{at(\mathbf{x})}) (\frac{1}{q} \sum_b \zeta_q^{b \sum_i x_i}) \right) \right|. \quad (3)$$

Expand the expression in the summand, and note the terms that appear. We now claim that,

$$2^{-n} \left| \sum_{\mathbf{x} \in \{0,1\}^n} \frac{1}{q} \sum_{b=0}^{q-1} (\zeta_q^{b \sum_i x_i} - \zeta_q^{b(\sum_i x_i - 1)}) \right| \leq 2^{-\Omega(n)}. \quad (4)$$

To see this, perform the sum over the x_i's to obtain,

$$\begin{aligned} \sum_{\mathbf{x} \in \{0,1\}^n} \zeta_q^{b \sum_i x_i} &= (1 + \zeta_q^b)^n, \\ \sum_{\mathbf{x} \in \{0,1\}^n} \zeta_q^{b(\sum_i x_i - 1)} &= \zeta_q^{-b} (1 + \zeta_q^b)^n \end{aligned}$$

and hence,

$$2^{-n} \left(\sum_{\mathbf{x} \in \{0,1\}^n} \frac{1}{q} \sum_{b=0}^{q-1} (\zeta_q^{b \sum_i x_i} - \zeta_q^{b(\sum_i x_i - 1)}) \right) = \frac{1}{q2^n} \sum_{b=0}^{q-1} (1 - \zeta_q^{-b})(1 + \zeta_q^b)^n.$$

Thus, by the triangle inequality,

$$2^{-n}\left|\sum_{\mathbf{x}\in\{0,1\}^n}\frac{1}{q}\sum_{b=0}^{q-1}(\zeta_q^{b\sum_i x_i}-\zeta_q^{b(\sum_i x_i-1)})\right|\le\frac{1}{q2^n}\sum_{b=0}^{q-1}|1-\zeta_q^{-b}|\cdot|1+\zeta_q^b|^n.$$

Note that the $b=0$ term does not contribute. Thus the maximum value that $|1+\zeta_q^b|$ can take on is $|1+\zeta_q|=2\cos(\pi/q)$. Including the factor of 2^{-n} and noting that $\cos(\pi/q)$ is a constant less than 1, inequality (4) follows.

After simplifying inequality (3), applying the triangle inequality, and taking into account equation (4), the lemma is proved.

□

Lemmas 2.1 and 2.2 show that obtaining an exponential lower bound on threshold-of-Mod_m^+ circuits reduces to finding an exponentially small upper bound on $|S(n,m,q;t)|$ where t is a low degree polynomial. Now S is of the form,

$$S(n,m,q;t)=\frac{1}{m}\sum_{a=0}^{m-1}\sum_{b=0}^{q-1}(\zeta_q^{-b}-1)\sigma(n,m,q,a,b;t),$$

where we define,

$$\sigma(n,m,q,a,b;t)=\frac{1}{2^n}\sum_{\mathbf{x}\in\{0,1\}^n}\zeta_q^{b\sum_i x_i}\zeta_m^{at(\mathbf{x})}. \tag{5}$$

(Again, we often omit the arguments of $\sigma(n,m,q,a,b;t)$, and simply write σ.) Thus the circuit lower bound reduces to finding exponentially small upper bounds on $|\sigma(n,m,q,a,b;t)|$. Note that we may assume that $b\not\equiv 0\pmod q$, since for $b\equiv 0\pmod q$ the terms in the sum for S vanish.

Before treating the more general case, using the methods of this paper we prove the result of Goldmann [Go], that the Mod_q function requires exponential-size threshold-of-Mod_m circuits. Although it is much simpler, the proof contains some of the essential ingredients of the proof of the main theorem. The key is to show that σ factorizes into n factors of norm less than 1.

Theorem 2.3. Let q be prime and m be a number such that q does not divide m. Let C be a threshold-of-Mod_m circuit which computes the Mod_q function. Then the number of gates in C is $2^{\Omega(n)}$.

Proof. Define $S(n,m,q;t)$ and $\sigma(n,m,q,a,b;t)$ as above. In this case, the polynomial t corresponding to a Mod_m subcircuit is *linear*, so we may write $t(x_1,...,x_n)=\sum_{i=1}^n k_ix_i$ for integers $k_i, 1\le i\le n$. By Lemmas 2.1 and 2.2, it suffices to prove an exponentially small upper bound on $|\sigma|$ for $b\not\equiv 0\ (q)$ (since $b=0$ does not contribute to S). Then, we can explicitly perform the sum over the x_i's,

$$\begin{aligned}\sigma &= \frac{1}{2^n}\sum_{\mathbf{x}\in\{0,1\}^n}\zeta_q^{b\sum_i x_i}\zeta_m^{a\sum_{i=1}^n k_ix_i}\\ &= \frac{1}{2^n}\prod_{i=1}^n(1+\zeta_q^b\zeta_m^{ak_i})\\ &= \frac{1}{2^n}\prod_{i=1}^n(1+\zeta_{qm}^{bm+aqk_i}).\end{aligned}$$

If $bm + aqk_i \equiv 0\ (qm)$, it follows that $b \equiv 0\ (q)$, contrary to what we assumed. Hence $bm + aqk_i \not\equiv 0\ (qm)$, and hence,

$$\zeta_{qm}^{bm+aqk_i} \neq 1.$$

Thus the greatest possible value for $|1 + \zeta_{qm}^{bm+aqk_i}|$ is $|1 + \zeta_{qm}| = 2\cos\frac{\pi}{qm}$, which implies,

$$|\sigma| \leq \left(\cos\frac{\pi}{qm}\right)^n.$$

Since $\cos\frac{\pi}{qm} < 1$ is a constant, the result follows. $\square$

It appears to be a difficult problem to prove the required upper bound on $|\sigma|$ for polynomials t of degree more than 1. The following gives a general estimate which doesn't depend on any assumptions about t. Although this estimate is weak, we will make use of it. Subsequently, stronger bounds are obtained when we place constraints on t.

Lemma 2.4. Let q be a prime dividing neither m nor b. Then,

$$|\sigma(n, m, q, a, b; t)| \leq \cos\left(\frac{\pi}{qm}\right).$$

Proof. The proof is similar to that of Theorem 2.3, the difference being that we only perform the sum over x_1. Details are given in the Appendix. $\square$

Suppose $t : \{0,1\}^n \to \mathbf{Z}$ is a symmetric polynomial. Since it is then a sum of elementary symmetric polynomials, it is a function only of the sum of the inputs $\sum_i x_i$. Thus t can be represented by a function $r : \{0, ..., n\} \to \mathbf{Z}$, where $r(k)$ is the value of t when $\sum_i x_i = k$. In this case we say t is *symmetric via* r. It is known (see, e.g., [BBR]) that $r(k)$ is periodic in $\mathbf{Z}/m\mathbf{Z}$. The period D of r mod m is defined to be the minimum $D > 0$ such that for all k, $r(k + D) \equiv r(k)\ (m)$. D is bounded by a polynomial in the degree of t, and can be easily expressed in terms of the prime factors of m [CGT].

We can obtain much stronger upper bounds than Lemma 2.4 for $|\sigma(n, m, q, a, b; t)|$ when t is symmetric. In order to obtain these bounds, we will need the following lemma, originally obtained in [CGT]. Here we also present a much-simplified proof.

Lemma 2.5. For any natural numbers D, k, n,

$$\sum_{j \equiv k\ \ (D)}^{n} \binom{n}{j} = \frac{1}{D}\sum_{\ell=0}^{D-1} \zeta_D^{k\ell}(1 + \zeta_D^{-\ell})^n.$$

Proof. Using the fact that,

$$\frac{1}{D}\sum_{\ell=0}^{D-1} \zeta_D^{\ell(k-j)} = \begin{cases} 1 & \text{if } j \equiv k\ \ (D) \\ 0 & \text{otherwise.} \end{cases}$$

it follows that,

$$\sum_{j \equiv k\ \ (D)}^{n} \binom{n}{j} = \sum_{j=0}^{n} \frac{1}{D}\sum_{\ell=0}^{D-1} \zeta_D^{\ell(k-j)}\binom{n}{j}$$

$$
\begin{aligned}
&= \frac{1}{D}\sum_{\ell=0}^{D-1}\zeta_D^{k\ell}\sum_{j=0}^{n}\zeta_D^{-j\ell}\binom{n}{j} \\
&= \frac{1}{D}\sum_{\ell=0}^{D-1}\zeta_D^{k\ell}(1+\zeta_D^{-\ell})^n,
\end{aligned}
$$

where the last equality follows from the binomial theorem. □

The following is the main technical lemma, which gives an upper bound for $|\sigma|$ in the event that t is symmetric. Note that this is an exponentially small upper bound as long as the period D grows slowly enough. Fortunately, this is the case in the main theorem. It would be nice to obtain a stronger upper bound (e.g., $2^{-\Omega(n)}$ even when D is not a constant), but it is not clear how to eliminate the factors of $1/(qD)$ in the exponent.

Lemma 2.6. Let n be any natural number, q a prime dividing neither m nor b, and let $t : \{0,1\}^n \to \mathbf{Z}$ be symmetric via r. Let the function $r(k)$ have period D mod m. Then,

$$
|\sigma(n,m,q,a,b;t)| \le \left(\cos\left(\frac{\pi}{qm}\right)\right)^{\frac{n}{(qD)^3}}.
$$

Proof. When $n < (qD)^3$, the conclusion is clear from Lemma 2.4, since in that case, noting that $\cos\left(\frac{\pi}{qm}\right) < 1$,

$$
|\sigma| \le \cos\left(\frac{\pi}{qm}\right) \le \left(\cos\left(\frac{\pi}{qm}\right)\right)^{\frac{n}{(qD)^3}}.
$$

Therefore, assume for the remainder of the proof that $n \ge (qD)^3$.

Using the definition of σ, the fact that the function $\zeta_q^{bk}\zeta_m^{ar(k)}$ has period qD, and Lemma 2.5,

$$
\begin{aligned}
\sigma &= \frac{1}{2^n}\sum_{j=0}^{n}\zeta_q^{bj}\zeta_m^{ar(j)}\binom{n}{j} \\
&= \frac{1}{2^n}\sum_{k=0}^{qD-1}\zeta_q^{bk}\zeta_m^{ar(k)}\sum_{j\equiv k\ (qD)}^{n}\binom{n}{j} \\
&= \frac{1}{2^n}\sum_{k=0}^{qD-1}\zeta_q^{bk}\zeta_m^{ar(k)}\frac{1}{qD}\sum_{\ell=0}^{qD-1}\zeta_{qD}^{k\ell}(1+\zeta_{qD}^{-\ell})^n \\
&= \frac{1}{2^n}\frac{1}{qD}\sum_{\ell=0}^{qD-1}(1+\zeta_{qD}^{-\ell})^n\sum_{k=0}^{qD-1}\zeta_q^{bk}\zeta_m^{ar(k)}\zeta_{qD}^{k\ell} \\
&= \frac{1}{2^n}\frac{1}{qD}\sum_{\ell=0}^{qD-1}(1+\zeta_{qD}^{-\ell})^n\sigma^{(\ell)},
\end{aligned}
$$

where $\sigma^{(\ell)}$ is defined as,

$$
\sigma^{(\ell)} \equiv_{def} \sum_{k=0}^{qD-1}\zeta_q^{bk}\zeta_m^{ar(k)}\zeta_{qD}^{k\ell}.
$$

We determine those values of ℓ such that $\sigma^{(\ell)} = 0$ as follows. Using the period of r once again, and re-defining variables of summation,

$$\begin{aligned}
\sigma^{(\ell)} &= \sum_{k=0}^{qD-1} \zeta_q^{-bD}\zeta_q^{b(k+D)}\zeta_m^{ar(k+D)}\zeta_{qD}^{-\ell D}\zeta_{qD}^{\ell(k+D)} \\
&= \zeta_q^{-bD}\zeta_{qD}^{-\ell D}\sum_{k=0}^{qD-1} \zeta_q^{b(k+D)}\zeta_m^{ar(k+D)}\zeta_{qD}^{\ell(k+D)} \\
&= \zeta_q^{-bD}\zeta_{qD}^{-\ell D}\sigma^{(\ell)}.
\end{aligned}$$

Thus,

$$(1-\zeta_q^{-bD}\zeta_{qD}^{-\ell D})\sigma^{(\ell)} = 0,$$

or,

$$(1-\zeta_q^{-(\ell+bD)})\sigma^{(\ell)} = 0.$$

Now $\zeta_q^{-(\ell+bD)} = 1$ iff $\ell + bD \equiv 0$ (q), which is true for exactly $1/q$ of the qD values of ℓ. For any such value of ℓ, trivially, $|\sigma^{(\ell)}| \le qD$. For any ℓ such that $\zeta_q^{-(\ell+bD)} \ne 1$, the above equation implies $\sigma^{(\ell)} = 0$. In particular, note that $\sigma^{(0)} = 0$, so that in the sum for σ, the $\ell = 0$ term does not contribute. This is significant, as in [CGT], because the $\ell \ne 0$ terms are exponentially small (in n) compared with the $\ell = 0$ term.

Using the triangle inequality, and making use of the observations above and our expression for σ up to this point, we obtain,

$$\begin{aligned}
|\sigma| &\le \frac{1}{2^n qD}\sum_{\ell=0}^{qD-1} |1+\zeta_{qD}^{-\ell}|^n \cdot |\sigma^{(\ell)}| \\
&\le \frac{1}{2^n qD}\cdot |1+\zeta_{qD}^{-1}|^n \cdot \frac{1}{q}\cdot qD\cdot qD \\
&= \frac{|1+\zeta_{qD}^{-1}|^n}{2^n}\cdot D \\
&= \left(\cos\left(\frac{\pi}{qD}\right)\right)^n \cdot D \\
&= \left(D^{1/n}\cos\frac{\pi}{qD}\right)^n.
\end{aligned}$$

From the well-known inequalities $e^x \ge 1+x$ and $1-\frac{1}{2}x^2 \ge \cos(x)$, we conclude that $\cos(\frac{\pi}{qD}) \le e^{-\frac{1}{2}(\frac{\pi}{qD})^2}$. Since $n \ge (qD)^3$,

$$D^{1/n}\cos\left(\frac{\pi}{qD}\right) \le D^{1/(qD)^3}e^{-\frac{1}{2}(\frac{\pi}{qD})^2} = e^{\frac{\ln(D)}{(qD)^3}-\frac{1}{2}(\frac{\pi}{qD})^2}.$$

But it is easy to show that,

$$\frac{\ln(D)}{(qD)^3} - \frac{1}{2}(\frac{\pi}{qD})^2 \le -\frac{1}{(qD)^2},$$

and thus,

$$D^{1/n}\cos\left(\frac{\pi}{qD}\right) \le e^{-\frac{1}{(qD)^2}}.$$

Therefore,

$$|\sigma| \leq \left(D^{1/n}\cos\left(\frac{\pi}{qD}\right)\right)^n \leq e^{-\frac{n}{(qD)^2}} \leq e^{-\frac{n}{(qD)^3}}.$$

Finally, for $q, m \geq 2$, $\cos\left(\frac{\pi}{qm}\right) \geq \cos\left(\frac{\pi}{4}\right) = \frac{1}{\sqrt{2}} > \frac{1}{e}$, so that,

$$|\sigma| \leq \left(\cos\left(\frac{\pi}{qm}\right)\right)^{\frac{n}{(qD)^3}}.$$

□

Say a polynomial $t : \{0,1\}^n \to \mathbf{Z}$ is *block symmetric* if it is a sum of polynomials which are symmetric in pairwise disjoint subsets of the inputs. More precisely, there are subsets $I_s \subseteq \{1, ..., n\}$, where $s \neq s' \implies I_s \cap I_{s'} = \emptyset$, and corresponding polynomials t_s, such that $t(x_1, ..., x_n) = \sum_s t_s(x_{i_1}, x_{i_2}, ..., x_{i_{n_s}})$, where $n_s = |I_s|$ and $I_s = \{i_1, ..., i_{n_s}\}$. The polynomials t_s are called the *symmetric polynomials of* t. We are now in a position to prove our main theorem. As in the proof of Theorem 2.3, the goal is to factorize σ into many factors of norm less than 1. However, in this case, due to the presence of the block-symmetric polynomials, we must be content with n^ϵ factors for some $0 < \epsilon < 1$.

Theorem 2.7. Let q be prime and m be a number such that q does not divide m. Let C be a threshold-of-Mod_m^+ circuit, in which the polynomials defining the Mod_m^+ circuits are block-symmetric (of degree $\log^{O(1)} n$), and where C computes the Mod_q function. Then for some real number $0 \leq \epsilon < 1$, the number of gates in C is $2^{\Omega(n^\epsilon)}$.

Proof. Consider one of the Mod_m^+ subcircuits, and let t be the polynomial in the inputs associated with that subcircuit. From Lemmas 2.1 and 2.2, it is sufficient to show that $|S(n, m, q; t)| \leq 2^{-n^\epsilon}$. By hypothesis, t is block symmetric, so we may write, $t(x_1, ..., x_n) = \sum_s t_s(x_{i_1}, x_{i_2}, ..., x_{i_{n_s}})$, where the t_s are the symmetric polynomials of t. Since t_s is symmetric, it is periodic mod m. Denote the period of t_s mod m as D_s. Since D_s is a polynomial in the degree of t_s, and since the degree of t_s is bounded by $O(\log^{O(1)} n)$, we know that $D_s = O(\log^{O(1)} n)$. From the definitions of S and σ, noting that $n = \sum_s n_s$, the sum S factorizes as follows,

$$S(n, m, q; t) = \frac{1}{m}\sum_{a=0}^{m-1}\sum_{b=0}^{q-1}(\zeta_q^{-b} - 1)\prod_s \sigma(n_s, m, q, a, b; t_s).$$

Hence, using Lemma 2.6 and bounding $|\zeta_q^b - 1|$ by 2,

$$\begin{aligned}
|S(n, m, q; t)| &\leq 2\sum_{b=1}^{q-1}\max_{0\leq a\leq m-1}\prod_s |\sigma(n_s, m, q, a, b; t_s)| \\
&\leq 2\sum_{b=1}^{q-1}\prod_s \left(\cos\frac{\pi}{qm}\right)^{\frac{n_s}{(qD_s)^3}} \\
&\leq 2(q-1)\prod_s \left(\cos\frac{\pi}{qm}\right)^{\frac{n_s}{\log^{O(1)} n}}
\end{aligned}$$

$$
\begin{aligned}
&= 2(q-1)\left(\cos\frac{\pi}{qm}\right)^{\frac{1}{\log^{O(1)} n}\sum_s n_s} \\
&= 2(q-1)\left(\cos\frac{\pi}{qm}\right)^{\frac{n}{\log^{O(1)} n}} \\
&\leq 2^{-\Omega(n^{\epsilon})},
\end{aligned}
$$

for some $\epsilon < 1$. $\square$

3 Discussion

It would of course be most desirable to obtain exponentially small upper bounds for the quantity $|S(n,q,m;t)|$ for *any* low-degree polynomial t. While obtaining good bounds for exponential sums can be very difficult, the more general results entailing elaborate techniques and deep results from algebraic geometry [Ka] and/or analysis [AKC], the results of this paper offer encouragement that the required techniques will not be quite that hard. This is suggested by the following observations.

The techniques for estimating exponential sums in number theory are quite sophisticated even for univariate polynomials. In a sense the polynomials here are like univariate polynomials, since, being symmetric, they only depend on the sum of the inputs. Nevertheless, the techniques employed here are of an elementary nature, relying (for example) on the Chinese remainder theorem and properties of the binomial coefficients in $\mathbf{Z}/m\mathbf{Z}$. One of the apparent reasons for the differing techniques is that here we are interested in a different type of asymptotic estimate. We fix the degree of t (or let it grow very slowly), allowing the number of variables to approach infinity. By contrast in the theory of equations over finite fields, the interest is in obtaining estimates in successively larger finite extensions of a fixed base field, the polynomial (along with the number of inputs) remaining fixed [LN]. More importantly, for the known techniques to apply, the polynomials have different kinds of constraints imposed on them. For example, a wide class of results (see [LN], [Sch]) holds if the polynomial is "absolutely irreducible," i.e., irreducible in every finite extension of the base field. No such assumptions were required here or in [CGT]. The techniques employed here hold for any low-degree block-symmetric polynomial, including those that are reducible. While the known algebraic techniques will very likely play a role, we are clearly facing a problem of a different nature.

Acknowledgements: I wish to thank Jin-Yi Cai and Thomas Thierauf for helpful conversations on this subject. The hospitality of Steve Homer and the Computer Science Department of Boston University, where part of this work was done while the author was on sabbatical from Clark University, is gratefully acknowledged.

References

[AKC] G. I. Arhipov, A. A. Karacuba, V. N. Cubarikov, Multiple trigonometric sums, *Proceedings of the Steklov Institute of Mathematics,* volume 151, American Mathematical Society, Providence, 1982.

[Al] E. ALLENDER, A note on the power of threshold circuits. In *Proceedings of the 30th Symposium on Foundations of Computer Science*, (1989), 580-584.

[Bar] D. BARRINGTON, Bounded-width polynomial-size branching programs recognize exactly those languages in NC^1, in *Journal of Computer and System Sciences* **38**, (1989), 150-164.

[BBR] D. A. M. BARRINGTON, R. BEIGEL, AND S. RUDICH, Representing Boolean functions as polynomials modulo composite numbers, in *Computational Complexity* **4** (1994) 367-382.

[BT] R. BEIGEL AND J. TARUI, On ACC, in *Computational Complexity* **4** (1994) 350–366.

[Cai] J.-Y. CAI, With probability one, a random oracle separates PSPACE from the polynomial-time hierarchy *Journal of Computer and System Science* **38** (1989) 68-85.

[CGT] J.-Y. CAI, F. GREEN, AND T. THIERAUF, On the correlation of symmetric functions, in *Mathematical Systems Theory* **29** (1996) 245-258.

[FSS] M. FURST, J. B. SAXE, AND M. SIPSER, Parity, circuits, and the polynomial-time hierarchy, in *Mathematical Systems Theory,* **17** (1984) 13-27.

[Go] M. GOLDMANN, A note on the power of majority gates and modular gates, in *Information Processing Letter,* **53** (1995), 321-327.

[GKRST] F. GREEN, J. KÖBLER, K. REGAN, T. SCHWENTICK, AND J. TORÁN, The power of the middle bit of a #P function, in *Journal of Computer and System Sciences* **50** (1995) 456-467.

[Gr] F. GREEN, An oracle separating $\oplus$P from PP^{PH}, *Information Processing Letters* **37** (1991) 149-153.

[Has] J. HÅSTAD, Computational limitations of small-depth circuits, the MIT press, Cambridge, 1987.

[HG] J. HÅSTAD AND M. GOLDMANN, On the power of small-depth threshold circuits, in *Computational Complexity,* **1** (1991) 113-129.

[HMPST] A. HAJNAL, W. MAASS, P. PUDLÁK, M. SZEGEDY, AND G. TURÁN, Threshold circuits of bounded depth, in *Proceedings 28th Annual IEEE Symposium on Foundations of Computer Science*, IEEE Computer Society Press (1987) 99-110.

[Ka] N. KATZ, Sommes exponentielles, *Astérisque* **79** (1980).

[KP] M. KRAUSE AND P. PUDLÁK, On the computational power of depth 2 circuits with threshold and modulo gates, in *Proceedings of the Twenty-Sixth Annual ACM Symposium on the Theory of Computing,* ACM Press (1994) 48-57.

[LN] R. LIDL AND H. NIEDERREITER, Finite Fields, *Encyclopedia of Mathematics and its Applications,* Vol. 20, Cambridge University Press, Cambridge, 1983.

[Raz] A. A. RAZBOROV, Lower bounds on the size of bounded depth networks over a complete basis with logical addition, *Matematicheskie Zametki* **41** (1987) 598-607. English translation in *Mathematical Notes of the Academy of Sciences of the USSR* **41** (1987) 333-338.

[Sch] W. M. Schmidt, Equations over finite fields: An elementary approach, *Lecture Notes in Mathematics*, vol. 536, Springer, New York, 1976.

[Sm] R. Smolensky, Algebraic methods in the theory of lower bounds for Boolean circuit complexity, in *Proceedings of the 19th Annual ACM Symposium on Theory of Computing* (1987) 77-82.

[Tod] S. Toda. PP is as hard as the polynomial-time hierarchy. In *SIAM Journal on Computing* **20**, (1991) 865-877.

[Yao 85] A.C. Yao, Separating the polynomial-time hierarchy by oracles, in *Proceedings of the 26th Annual IEEE Symposium on Foundations of Computer Science* (1985) 1-10.

[Yao 90] A. Yao, On ACC and threshold circuits. In *Proceedings of the 31st Symposium on Foundations of Computer Science*, (1990), 619-627.

Appendix: Proof of Lemma 2.4

Performing the sum over x_1,

$$\begin{aligned}
\sigma &= \frac{1}{2^n} \sum_{(x_2,\ldots,x_n)\in\{0,1\}^{n-1}} \left(\zeta_q^{(b\sum_{i=2}^n x_i)} \zeta_m^{at(0,x_2,\ldots,x_n)} + \zeta_q^b \zeta_q^{(b\sum_{i=2}^n x_i)} \zeta_m^{at(1,x_2,\ldots,x_n)} \right) \\
&= \frac{1}{2^n} \sum_{(x_2,\ldots,x_n)\in\{0,1\}^{n-1}} \zeta_q^{(b\sum_{i=2}^n x_i)} \zeta_m^{at(0,x_2,\ldots,x_n)} \cdot \left(1 + \zeta_q^b \zeta_m^{a(t(1,x_2,\ldots,x_n)-t(0,x_2,\ldots,x_n))}\right).
\end{aligned}$$

Then by the triangle inequality,

$$|\sigma| \leq \frac{1}{2^n} \sum_{(x_2,\ldots,x_n)\in\{0,1\}^{n-1}} \left|1 + \zeta_q^b \zeta_m^{a(t(1,x_2,\ldots,x_n)-t(0,x_2,\ldots,x_n))}\right|.$$

For notational convenience, let Δt denote $t(1,x_2,\ldots,x_n) - t(0,x_2,\ldots,x_n)$. Now $\zeta_q^b \zeta_m^{a\Delta t} = \zeta_{qm}^{bm+qa\Delta t}$. Furthermore, $\zeta_{qm}^{bm+qa\Delta t} = 1$ if and only if $bm + qa\Delta t \equiv 0$ (qm). But if the latter holds, then q must divide either b or m, neither of which is possible. Hence,

$$\zeta_{qm}^{bm+qa\Delta t} \neq 1,$$

so the maximum value that $|1 + \zeta_{qm}^{bm+qa\Delta t}|$ can be is $|1 + \zeta_{qm}|$. Hence, from the foregoing inequality for $|\sigma|$,

$$\begin{aligned}
|\sigma| &\leq \frac{1}{2^n}|1+\zeta_{qm}| \sum_{(x_2,\ldots,x_n)\in\{0,1\}^{n-1}} 1 \\
&= \frac{|1+\zeta_{qm}|}{2} \\
&= \cos\left(\frac{\pi}{qm}\right).
\end{aligned}$$