Lower Bounds for Circuits with Few Modular Gates using Exponential Sums

Kristoffer Arnsfelt Hansen
Department of Computer Science
University of Aarhus
arnsfelt@daimi.au.dk

Abstract

We prove that any $\mathbf{AC^0}$ circuit augmented with $\epsilon \log^2 n$ MOD_m gates and with a MAJORITY gate at the output, require size $n^{\Omega(\log n)}$ to compute MOD_l, when l has a prime factor not dividing m and ϵ is sufficiently small. We also obtain that the MOD_2 function is hard on the average for $\mathbf{AC^0}$ circuits of size $n^{\epsilon \log n}$ augmented with $\epsilon \log^2 n$ MOD_m gates, for every odd integer m and any sufficiently small ϵ. As a consequence, for every odd integer m, we obtain a pseudorandom generator, based on the MOD_2 function, for circuits of size S containing $\epsilon \log S$ MOD_m gates.

Our results are based on recent bounds of exponential sums that were previously introduced for proving lower bounds for $\mathbf{MAJ} \circ \mathbf{MOD}_m \circ \mathbf{AND}_d$ circuits.

1 Introduction

One of the most important challenges in circuit complexity is to prove lower bounds for circuits containing MOD_m gates for a fixed integer m. Indeed the circuit class $\mathbf{ACC^0}$, consisting of constant depth circuits ($\mathbf{AC^0}$ circuits) augmented with MOD_m gates is the smallest natural circuit class, for which we have no nontrivial lower bounds. In the special case when the modulus m is a prime power, exponential lower bounds holds for computing MAJORITY, and if l has a prime divisor not dividing m also for computing MOD_l, as proved by Razborov [16] and Smolensky [17]. It is however unclear if the techniques used for proving these lower bounds can be extended to the case when m is not a prime power.

One possible direction for approaching lower bounds for $\mathbf{ACC^0}$ is to prove lower bounds for $\mathbf{AC^0}$ circuits augmented with *few* MOD_m gates. For circuits with MOD_2 gates, this approach was in fact already taken by Håstad in his thesis [14], although his results were soon overshadowed by the results of Razborov and Smolensky. To be more precise, Håstad proved that constant depth circuits augmented with $\epsilon \log^{\frac{3}{2}} n$ MOD_2 gates require size $2^{\Omega(\log^{\frac{3}{2}} n)}$ to compute MAJORITY, for any sufficiently small ϵ. His techniques can in fact be generalized to prove that when m is a prime power, constant depth circuits augmented with n^ϵ MOD_m gates require exponential size to compute MAJORITY, for any sufficiently small ϵ, but like the lower bounds of Razborov and Smolensky it does not directly extend to the case when m is not a prime power.

This same approach has also been successfully applied to the circuit class $\mathbf{TC^0}$, consisting of constant depth circuits augmented with MAJORITY gates, in a series of papers [1, 6, 5, 3]. In particular must constant depth circuits augmented with n^ϵ MAJORITY gates be of exponential size to compute the MOD_m function for any constant m and any sufficiently small ϵ.

Recently the approach has also been successfully applied to circuits augmented with few SYM gates (i.e. gates computing arbitrary symmetric functions). This class of circuits generalize each of the above classes of circuits since both MOD_m and MAJORITY are symmetric functions. In [8] it was proved that constant depth circuits augmented with $\epsilon \log^2 n$ SYM gates require size $n^{\Omega(\log n)}$ to compute a certain (complicated) function in $\mathbf{ACC^0}$. This function is based on the so-called generalized inner product function and the circuit lower bound rely on a lower bound on the distributional multi party communication complexity for this function [2].

The lower bound for circuit with few SYM gates is currently also the best known lower bound for circuits with few MOD_m gates. However it is also interesting to obtain such bounds for computing simpler functions such as MAJORITY and MOD_l for the following reasons. First of all because we believe that such lower bounds holds even without restrictions on the number of MOD_m gates. Secondly, because it is conceivably that better lower bounds can be obtained by techniques that do not rely on multi party communication complexity, but instead take advantage of the fact that the circuits only contain MOD_m gates rather than more powerful gates such as SYM gates. Such lower bounds was also obtained in [8], although they are weaker than the lower bounds given for circuits with SYM gates. Specifically, let m be a positive integer with $r \geq 2$ distinct prime factors. Then constant depth circuits augmented with s MOD_m gates must have size $n^{\Omega(\frac{1}{s} log^{\frac{1}{r-1}} n)}$ to compute MAJORITY or MOD_l, if l has a prime factor not dividing m.

Here we obtain improved lower bounds for computing the MOD_l functions, matching the strongest lower bounds known for circuits with few SYM gates. Additionally our lower bound holds even if we allow a MAJORITY gate at the output.

Theorem 1 *Any* $\mathbf{MAJ} \circ \mathbf{AC^0}$ *circuit containing* $\epsilon \log^2 n$ MOD_m *gates computing* MOD_l*, where* l *has a prime factor not dividing* m*, must have size at least* $n^{\Omega(\log n)}$ *for any sufficiently small* ϵ*.*

Recently Viola [18] showed that the function used for proving lower bounds for constant depth circuit with few SYM gates is in fact also hard on the average for the same class of circuits. He could then apply the Nisan-Wigderson pseudorandom generator construction [15] to obtain a pseudorandom generator stretching l bits to $n = l^{\epsilon \log l}$ bits that fools constant depth circuits of size n containing $\log n$ SYM gates.

Here we show that for constant depth circuits with few MOD_m gates for an odd integer m, we can in a similar way base the Nisan-Wigderson construction on the MOD_2 function. Our motivation for presenting this generator is the same as for our circuit lower bounds. First of all the generator is simpler to compute. Secondly it is conceivable that better pseudorandom generators can be constructed using techniques that take advantage of the fact that the circuits only contain MOD_m gates rather than more powerful gates such as SYM gates. To be precise we obtain the following average case hardness result.

Theorem 2 *For every odd integer* m *and every* h *there exists* $\epsilon > 0$ *such that for every sufficiently large* n *and for every depth* h *circuit* C *on* n *inputs of size* $n^{\epsilon \log n}$ *containing* $\epsilon \log^2 n$ MOD_m *gates we have*

$$\Pr\left[C(x) \neq \text{MOD}_2(x)\right] \geq \frac{1}{2} - \frac{1}{n^{\epsilon \log n}}$$

Applying the Nisan-Wigderson construction we obtain the following pseudorandom generator.

Theorem 3 *For every odd integer* m *and every* h *there exists* $\epsilon > 0$ *such that for all sufficiently large* l *there is a generator* $G : \{0,1\}^l \to \{0,1\}^n$*, where* $n = l^{\epsilon \log l}$*, such that for every depth* h *circuit* C *on* n *inputs of size* n *containing* $\log n$ MOD_m *gates we have*

$$\left| \Pr_{x \in \{0,1\}^n}\left[C(x) = 1\right] - \Pr_{x \in \{0,1\}^l}\left[C(G(x)) = 1\right] \right| \leq \frac{1}{n}$$

Furthermore every output bit of $G(x)$ *is the* MOD_2 *function taken on a subset of the input bits.*

2 Preliminaries

2.1 Exponential sums

Let $m, l > 1$ and let P by a polynomial of degree d over $\mathbf{Z}_m$. The following exponential sum S was defined by Green [10].

$$S = \frac{1}{2^n} \sum_{x \in \{0,1\}^n} e_m(P(x)) e_l(a \sum_{i=1}^n x_i)$$

where $e_k(x)$ denotes $e^{\frac{2\pi i}{k} x}$. Recently Bourgain [7] and Green et al [11] obtained the following estimate on the absolute value of the exponential sum S.

Theorem 4 *Assume m and l are relatively prime and $0 < a < l$. Then there exists $0 < \mu_d < 1$ such that*

$$|S| < (\mu_d)^n$$

for all $n > 0$. Furthermore, for all $0 < \epsilon < 1$, there exists $c > 0$ such that

$$(\mu_{c \log n})^n < 2^{-n^\epsilon}$$

for all sufficiently large n.

2.2 Circuit classes

We consider circuits built from families of unbounded fanin gates. Inputs are allowed to be boolean variables and their negations as well as the constants 0 and 1. Let $x_1, \ldots, x_n$ be boolean inputs. For a positive integer m, the MOD_m function outputs 1 if and only if $\sum_{i=1}^n x_i \not\equiv 0 \pmod m$. The MAJORITY function is 1 if and only if $\sum_{i=1}^n x_i \geq \frac{n}{2}$.

Let **AND** and **OR** denote the families of unbounded fanin AND and OR gates. Similarly, let $\mathbf{MOD}_m$ and **MAJ** denote the families of MOD_m, MAJORITY gates. If G is a family of boolean gates and $\mathcal{C}$ is a family of circuits we let $G \circ \mathcal{C}$ denote the class of circuits consisting of a G gate taking circuits from $\mathcal{C}$ as inputs. If we need to specify a specific bound on the fanin of some of the gates, this will be specified by a subscript.

$\mathbf{AC^0}$ is the class of functions computed by constant depth circuits built from AND and OR gates. $\mathbf{ACC^0}$ is the the analogous class of functions computed when we also allow unbounded fanin MOD_m gates for constants m, and similarly is $\mathbf{TC^0}$ the class of functions computed when we instead allow unbounded fanin MAJORITY gates.

2.3 Discriminator lemma

Let C be a circuit taking n inputs and f a boolean function on n variables. Let $A \subseteq f^{-1}(1)$ and $B \subseteq f^{-1}(0)$. We say that C is an ϵ-discriminator for f with respect to A and B if

$$\Pr[C(x) = 1 | x \in A] - \Pr[C(x) = 1 | x \in B] \geq \epsilon$$

The so-called discriminator lemma by Hajnal et al [12], states that if a circuit with a MAJORITY gate at the output computes a boolean function f, then one of the inputs to the output gate is an ϵ-discriminator for f.

Lemma 5 *Let f be a boolean function computed by a circuit C with a* MAJORITY *gate as the output gate, and let $C_1, \ldots, C_s$ be the subcircuits of C whose output gates are the inputs to the output of C. Let $A \subseteq f^{-1}(1)$ and $B \subseteq f^{-1}(0)$ be arbitrary. Then for some i, C_i is an $\frac{1}{s}$-discriminator for f with respect to A and B.*

2.4 The switching lemma

A *restriction* on a set V of boolean variables is a map $\rho : V \to \{0, 1, \star\}$. It acts on a boolean function f in the variables V, creating a new boolean function f_ρ on the set of variables for which $\rho(x) = \star$, obtained by substituting $\rho(x)$ for $x \in V$ whenever $\rho(x) \neq \star$. The variables x for which $\rho(x) = \star$ are called *free*. Let R_n^l denote the set of all restriction ρ leaving l of n variables free.

A decision tree is a binary tree, where the internal nodes are labeled by variables and leafs are labeled by either 0 or 1. On a given input x, its value is the value of the leaf reached by starting at the root, and at any internal node labeled by x_i proceeding to the left child if $x_i = 0$ and to the right child otherwise. We will use the following version of Håstads Switching Lemma due to Beame [4].

Lemma 6 *Let f be a DNF formula in n variables with terms of length at most r. Let $l = pn$ and pick ρ uniformly at random from R_n^l. Then the probability that f_ρ does not have a decision tree of depth at most d is less than $(7pr)^d$.*

The advantage of using Beame's switching lemma is that it directly gives us a decision tree. If we convert a decision tree into a DNF, we in fact obtain a *disjoint* DNF, i.e. a DNF where all terms are *mutually contradictory*. We can then view it as a sum of terms, instead as an OR of AND's, and this will allow us to absorb the sum into MOD_m and MAJORITY gates.

Proposition 7 *Let h be any integer and let $c > 0$. Then there exists $\epsilon > 0$ with the following property. Let C be a $\mathbf{AC^0}$ circuit of depth h and size $S = n^{\epsilon \log n}$. Choose a restriction $\rho \in R_n^{\sqrt{n}}$ at random. Then with probability at least $1 - n^{-\Omega(\log n)}$, after applying the restriction ρ, every function computed at any gate of C has a decision tree of depth at most $d = c \log n$.*

Proof We view ρ as a composition of several restrictions $\rho_1, \ldots, \rho_h$, where $\rho_i \in R_{n_{i-1}}^{n_i}$ and $n_i = n\left(n^{\frac{1}{2h}}\right)^{-i}$. Assume that after having applied the first $i-1$ restrictions, that all functions computed by gates at level $i-1$ of C are computed by decision trees of depth at most d. They are then also computed by DNF's with terms of size at most d. Now, assuming without loss of generality that the gates at level i are OR gates, the functions computed by these gates are also computed by DNF's with terms of size of most d. By Lemma 6, the probability that the function computed by such an OR gate can not be computed by a decision tree of depth at most d after applying ρ_i is then at most

$$\left(7\frac{m_i}{m_{i-1}}d\right)^d = \left(7n^{-\frac{1}{2h}}c \log n\right)^{c \log n} = n^{-\Omega(\log n)}$$

Since we have at most this probability of error at each of the $n^{\epsilon \log n}$ gates of C, the result follows for ϵ sufficiently small. $\square$

3 Circuit lower bounds

Theorem 8 *Let C be a $\mathbf{AND}_t \circ \mathbf{MOD}_m \circ \mathbf{AND}_d$ circuit and let l be relatively prime to m. Then*

$$|\Delta| \leq 2^t (\mu_d)^n$$

where

$$\Delta = \Pr\left[C(x) = 1 \wedge \sum_{i=1}^n x_i \equiv 1 \pmod{l}\right] - \Pr\left[C(x) = 1 \wedge \sum_{i=1}^n x_i \equiv 0 \pmod{l}\right]$$

and μ_d is given by Theorem 4.

Proof Let $C_1, \ldots, C_t$ be the subcircuits of C feeding the output. Let $P_1, \ldots, P_t$ be polynomials over $\mathbf{Z}_m$ of degree d such that $C_i(x) = 1$ if and only if $P_i(x) \not\equiv 0 \pmod{m}$. We can then rewrite the terms of Δ as exponential sums.

$$\Pr\left[C(x) = 1 \wedge \sum_{i=1}^n x_i \equiv 1 \pmod{l}\right] =$$

$$\Pr\left[\bigwedge_{i=1}^t P_i(x) \not\equiv 0 \pmod{m} \wedge \sum_{i=1}^n x_i \equiv 1 \pmod{l}\right] =$$

$$\frac{1}{2^n} \sum_{x \in \{0,1\}^n} \left[\prod_{i=1}^t \left(1 - \frac{1}{m}\sum_{b_i=0}^{m-1} e_m(b_i P_i(x))\right) \frac{1}{l}\sum_{a=0}^{l-1} e_l\left(a\left(\sum_{i=1}^n x_i - 1\right)\right)\right] =$$

$$\sum_{A \subseteq [t]} \frac{1}{l}\left(\frac{-1}{m}\right)^{|A|} \sum_{b_{i_1}=0}^{m-1} \cdots \sum_{b_{i_h}=0}^{m-1} \sum_{a=0}^{l-1} \left[\frac{1}{2^n}\sum_{x \in \{0,1\}^n} e_m\left(\sum_{j=1}^h b_{i_j} P_{i_j}(x)\right) e_l\left(a\left(\sum_{i=1}^n x_i - 1\right)\right)\right]$$

Where, in the sum $A = \{i_1, \ldots, i_h\}$. Similarly we obtain

$$\Pr\left[C(x) = 1 \wedge \sum_{i=1}^{n} x_i \equiv 0 \pmod{l}\right] =$$

$$\sum_{A \subseteq [t]} \frac{1}{l}\left(\frac{-1}{m}\right)^{|A|} \sum_{b_{i_1}=0}^{m-1} \cdots \sum_{b_{i_h}=0}^{m-1} \sum_{a=0}^{l-1} \left[\frac{1}{2^n} \sum_{x \in \{0,1\}^n} e_m\left(\sum_{j=1}^{h} b_{i_j} P_{i_j}(x)\right) e_l\left(a \sum_{i=1}^{n} x_i\right)\right]$$

And hence

$$\Delta \; = \sum_{A \subseteq [t]} \frac{1}{l}\left(\frac{-1}{m}\right)^{|A|} \sum_{b_{i_1}=0}^{m-1} \cdots \sum_{b_{i_h}=0}^{m-1} \sum_{a=0}^{l-1} (e_l(-a) - 1) \left[\frac{1}{2^n} \sum_{x \in \{0,1\}^n} e_m\left(\sum_{j=1}^{h} b_{i_j} P_{i_j}(x)\right) e_l\left(a \sum_{i=1}^{n} x_i\right)\right]$$

Since $\sum_{j=1}^{h} b_{i_j} P_{i_j}(x)$ is a polynomial of degree d, and noting that when $a = 0$ the term vanishes, we may bound $|\Delta|$ using Theorem 4 and the triangle inequality, to obtain

$$|\Delta| \; \leq 2^t (\mu_d)^n$$

□

Theorem 9 *Let C be a $\mathbf{MAJ} \circ \mathbf{AND}_t \circ \mathbf{MOD}_m \circ \mathbf{AND}_d$ circuit computing* MOD_l*, where l has a prime factor not dividing m. If d is constant and $t = \epsilon n$ for sufficiently small ϵ the size of C must be $2^{\Omega(n)}$. If $t = n^{\epsilon}$ and $d = c \log n$ for sufficiently small ϵ and c the size of C must be at least $2^{n^{\Omega(1)}}$.*

Proof We may without loss of generality assume that m and l are relatively prime. Otherwise, let p be a prime dividing l but not dividing m. Group the input variables into $\frac{np}{l}$ groups of size $\frac{l}{p}$, and consider only inputs giving the same value to all variables in each of these groups. In this way we obtain a $\mathbf{MAJ} \circ \mathbf{AND}_t \circ \mathbf{MOD}_m \circ \mathbf{AND}_d$ circuit computing MOD_p on $\frac{np}{l}$ inputs.

Let S be the size of C. Then from lemma 5 we have an $\mathbf{AND}_t \circ \mathbf{MOD}_m \circ \mathbf{AND}_d$ circuit C' of size S such that

$$\Delta \; = \Pr\left[C'(x) = 1 \mid \sum_{i=1}^{n} x_i \equiv 1 \pmod{l}\right] - \Pr\left[C'(x) = 1 \mid \sum_{i=1}^{n} x_i \equiv 0 \pmod{l}\right] \geq \frac{1}{S}$$

We will rewrite Δ using the fact [9] that for all integers l and a

$$\left|\Pr\left[\sum_{i=1}^{n} x_i \equiv a \pmod{l}\right]\right| \; = \frac{1}{l} + 2^{-\Omega(n)}$$

Thus

$$\Delta \; \leq l\,|\Delta'| + 2^{-\Omega(n)}$$

where

$$\Delta' \; = \Pr\left[C'(x) = 1 \wedge \sum_{i=1}^{n} x_i \equiv 1 \pmod{l}\right] - \Pr\left[C'(x) = 1 \wedge \sum_{i=1}^{n} x_i \equiv 0 \pmod{l}\right]$$

From Theorem 8 we have

$$|\Delta'| \; \leq 2^t (\mu_d)^n$$

And thus

$$\frac{1}{S} \; \leq l 2^t (\mu_d)^n + 2^{-\Omega(n)}$$

from which the result follows using Theorem 4. □

Using the switching lemma we can then obtain the following 'meet-in-the-middle' [13] lower bound.

Theorem 10 *Any depth $h+3$ $\mathbf{MAJ} \circ \mathbf{AND}_t \circ \mathbf{MOD}_m \circ \mathbf{AC^0}$ circuit computing MOD_l, where l has a prime factor not dividing m, must have size at least $n^{\Omega(\log n)}$, when $t = n^\epsilon$ for any sufficiently small ϵ.*

Proof Assume C is a $\mathbf{MAJ} \circ \mathbf{AND}_t \circ \mathbf{MOD}_m \circ \mathbf{AC^0}$ circuit of depth $h+3$ and size $n^{\epsilon \log n}$ computing MOD_l. Given $c > 0$, we apply Proposition 7 to the circuit consisting of the lowest h levels of C. If ϵ is sufficiently small, we thus have a restriction ρ such that after applying ρ, every input of a MOD_m gate in C is computed by a decision tree of depth at most $d = c \log n$. Each of these can be rewritten as a disjoint DNF with terms of size at most d and we can absorb the OR gates of these in the MOD_m gates, thus obtaining a $\mathbf{MAJ} \circ \mathbf{AND}_t \circ \mathbf{MOD}_m \circ \mathbf{AND}_d$ circuit of size $n^{\epsilon \log n}$ computing MOD_l on at least $\sqrt{n} - l$ inputs, contradicting Theorem 8. □

Using this theorem we can obtain the same lower bound for circuits with few MOD_m gates, even when allowing a MAJORITY gate at the output, i.e. the lower bound stated as Theorem 1.

Proposition 11 *Let C be a depth h $\mathbf{AC^0}$ circuit of size S containing s MOD_m gates. Then the function computed by C is also computed by a depth $h+3$ $\mathbf{OR}_{2^s} \circ \mathbf{AND}_{O(s)} \circ \mathbf{MOD}_m \circ \mathbf{AC^0}$ circuit C' of size $O(2^s S)$, and furthermore, the two top layers of C' define a disjoint DNF.*

Proof Let $g_1, \ldots, g_s$ be the MOD_m gates of C' such that there is no path from the output of g_i to g_j if $j < i$. For $\alpha \in \{0,1\}^s$ let C_i^α be the $\mathbf{MOD}_m \circ \mathbf{AC^0}$ subcircuit of C with g_i as output, where every g_j for $j < i$ is replaced by the constant α_j. Similarly, let C^α be the $\mathbf{AC^0}$ circuit obtained from C where every g_i is replaced by α_i. Note now that $\neg C_i^\alpha$ can be computed by a $\mathbf{AND}_{m-1} \circ \mathbf{MOD}_m \circ \mathbf{AC^0}$ circuit of size $O(S)$. We can now construct a $\mathbf{OR}_{2^s} \circ \mathbf{AND}_{(m-1)s+1} \circ \mathbf{MOD}_m \circ \mathbf{AC^0}$ circuit C'' of size $O(2^s S)$ computing the same function as C' as follows: The output OR gate is fed by AND's corresponding to all $\alpha \in \{0,1\}^s$. The AND gate corresponding to α takes C^α as input, as well as C_i^α if $\alpha_i = 1$ and the inputs of the output AND gate of the $\mathbf{AND}_{m-1} \circ \mathbf{MOD}_m \circ \mathbf{AC^0}$ circuit for $\neg C_i^\alpha$ if $\alpha_i = 0$. □

Proof (Theorem 1) Let C be $\mathbf{MAJ} \circ \mathbf{AC^0}$ circuit of size S containing $s = \epsilon \log^2 n$ MOD_m gates computing MOD_l. Using Proposition 11 we may replace each subcircuit feeding the output of C by a $\mathbf{OR}_{2^s} \circ \mathbf{AND}_{O(s)} \circ \mathbf{MOD}_m \circ \mathbf{AC^0}$ circuit of size $O(2^s S)$. Since furthermore the two top layers of these circuits define disjoint DNF's we can absorb the OR gates in the top MAJORITY gate, obtaining a $\mathbf{MAJ} \circ \mathbf{AND}_{O(s)} \circ \mathbf{MOD}_m \circ \mathbf{AC^0}$ circuit of size $O(2^s S^2)$ computing MOD_l. The result then follows from Theorem 10. □

4 A pseudo-random generator

We have the following corollary to Theorem 8, from which we can obtain Theorem 2 using Proposition 11.

Corollary 12 *Let C be an $\mathbf{OR}_s \circ \mathbf{AND}_t \circ \mathbf{MOD}_m \circ \mathbf{AND}_d$ circuit, such that the two top layers of the circuit defines a disjoint DNF, and let l be relatively prime to m. Then*

$$|\Delta| \ \leq s 2^t (\mu_d)^n$$

where

$$\Delta \ = \Pr\left[C(x) = 1 \wedge \sum_{i=1}^n x_i \equiv 1 \pmod{l}\right] - \Pr\left[C(x) = 1 \wedge \sum_{i=1}^n x_i \equiv 0 \pmod{l}\right]$$

and μ_d is given by Theorem 4.

Proof Let $C_1, \ldots, C_s$ be the subcircuits of C feeding the output. Since at most one of these subcircuit can evaluate to 1 at the same time, we have

$$\Delta \ = \sum_{i=1}^s \left(\Pr\left[C_i(x) = 1 \wedge \sum_{i=1}^n x_i \equiv 1 \pmod{l}\right] - \Pr\left[C_i(x) = 1 \wedge \sum_{i=1}^n x_i \equiv 0 \pmod{l}\right]\right)$$

The result then follows from the triangle inequality and Theorem 8. □

Proof (Theorem 2) Let C be depth h $\mathbf{AC^0}$ circuit of size $S = n^{\epsilon \log n}$ containing $s = \epsilon \log^2 n$ MOD_m gates. Using Proposition 11 we then have a depth $h+3$ $\mathbf{OR}_{2^s} \circ \mathbf{AND}_{O(s)} \circ \mathbf{MOD}_m \circ \mathbf{AC^0}$ circuit C'

of size $O(2^s S) = O(n^{2\epsilon \log n})$ computing the same function as C, and such that the two top layers of C' defines a disjoint DNF. Given $c > 0$, we apply Proposition 7 to the circuit consisting of the lowest h levels of C'. If ϵ is sufficiently small, then with probability $1 - n^{-\Omega(\log n)}$ a random restriction ρ satisfies that after applying ρ, every input of a MOD_m gate in C' is computed by a decision tree of depth at most $d = c \log n$. Each of these can be rewritten as a disjoint DNF with terms of size at most d and we can absorb the OR gates of these in the MOD_m gates, thus obtaining a $\mathbf{OR}_{2^s} \circ \mathbf{AND}_{O(s)} \circ \mathbf{MOD}_m \circ \mathbf{AND}_d$ circuit of size $n^{O(\epsilon \log n)}$ on $k = \sqrt{n}$ inputs. Assuming that ρ allows us to make this transformation we obtain

$$\begin{aligned}
&\Pr\left[C_\rho(x) \neq \mathrm{MOD}_2(x)\right] = \Pr\left[C'_\rho(x) \neq \mathrm{MOD}_2(x)\right] = \\
&\Pr\left[C'_\rho(x) = 0 \wedge \sum_{i=1}^{k} x_i \equiv 1 \pmod 2\right] + \Pr\left[C'_\rho(x) = 1 \wedge \sum_{i=1}^{k} x_i \equiv 0 \pmod 2\right] = \\
&\frac{1}{2} - \left(\Pr\left[C'_\rho(x) = 1 \wedge \sum_{i=1}^{k} x_i \equiv 1 \pmod 2\right] - \Pr\left[C'_\rho(x) = 1 \wedge \sum_{i=1}^{k} x_i \equiv 0 \pmod 2\right]\right) \geq \\
&\frac{1}{2} - 2^{O(s)}(\mu_d)^k \geq \frac{1}{2} - n^{-\Omega(\log n)}
\end{aligned}$$

for c and ϵ sufficiently small, using Corollary 12. Likewise we obtain

$$\Pr\left[C_\rho(x) \neq \neg\mathrm{MOD}_2(x)\right] \ \geq \frac{1}{2} - n^{-\Omega(\log n)}$$

And thus

$$\Pr\left[C_\rho(x) \neq \mathrm{MOD}_{2,\rho}(x)\right] \ \geq \frac{1}{2} - n^{-\Omega(\log n)}$$

Since we can generate a random input to C, by first choosing a restriction ρ at random, and then a random input to the remaining free variables we finally obtain

$$\Pr\left[C(x) \neq \mathrm{MOD}_2(x)\right] \geq \left(\frac{1}{2} - n^{-\Omega(\log n)}\right)\left(1 - n^{-\Omega(\log n)}\right) \geq \frac{1}{2} - n^{-\Omega(\log n)}$$

□

Proof (Theorem 3) Nisan and Wigderson [15] constructed, from any function $f : \{0,1\}^{\sqrt{\frac{l}{2}}} \rightarrow \{0,1\}$ and parameter n, a generator $G : \{0,1\}^l \rightarrow \{0,1\}^n$, by using f on n different subsets of the input variables. The generator thus obtained satisfies the following property. Let C be a circuit such that

$$\left| \Pr_{x \in \{0,1\}^n}\left[C(x) = 1\right] - \Pr_{x \in \{0,1\}^l}\left[C(G(x)) = 1\right] \right| \ > \frac{1}{n}$$

Then C can be transformed into another circuit C' such that

$$\Pr\left[C'(x) = f(x)\right] \ > \frac{1}{2} + \frac{1}{n^2}$$

The transformation is done by adding one more more layer of AND or OR gates at the bottom of C and possibly negating the output, and thereby increasing the size by at most a polynomial in n.

In our case f is the MOD_2 function and C would be a depth h circuit of size $n = l^{\epsilon \log l}$ containing $\epsilon \log^2 l$ MOD_m gates. The circuit C' constructed above can then be further transformed into a depth $O(h)$ circuit C'' of size $l^{O(\epsilon \log l)}$ with at most $O(\epsilon \log^2 l)$ MOD_m gates such that

$$\Pr\left[C''(x) = f(x)\right] \ > \frac{1}{2} + \frac{1}{l^{2\epsilon \log l}}$$

thus contradicting Theorem 2 for any sufficiently small ϵ. □

Acknowledgment

The author is supported by BRICS, Basic Research in Computer Science, a Centre of the Danish National Research Foundation.

References

[1] J. Aspnes, R. Beigel, M. L. Furst, and S. Rudich. The expressive power of voting polynomials. *Combinatorica*, 14(2):135–148, 1994.

[2] L. Babai, N. Nisan, and S. Szegedy. Multiparty protocols, pseudorandom generators for logspace, and time-space trade-offs. *Journal of Computer and System Sciences*, 45(2):204–232, Oct. 1992.

[3] D. A. M. Barrington and H. Straubing. Complex polynomials and circuit lower bounds for modular counting. *Computational Complexity*, 4(4):325–338, 1994.

[4] P. Beame. A switching lemma primer. Technical Report UW-CSE-95-07-01, Department of Computer Science and Engineering, University of Washington, 1994. Availible online at www.cs.washington.edu/homes/beame.

[5] R. Beigel. When do extra majority gates help? Polylog(n) majority gates are equivalent to one. *Computational Complexity*, 4(4):314–324, 1994.

[6] R. Beigel, N. Reingold, and D. A. Spielman. PP is closed under intersection. *Journal of Computer and System Sciences*, 50(2):191–202, 1995.

[7] J. Bourgain. Estimation of certain exponential sums arising in complexity theory. *C. R. Acad. Sci. Paris, Ser. I*, 341:627–631, 2005.

[8] A. Chattopadhyay and K. A. Hansen. Lower bounds for circuits with few modular and symmetric gates. In *32nd Annual International Colloquium on Automata, Languages and Programming*, volume 3580 of *Lecture Notes in Computer Science*, pages 994–1005. Springer, 2005.

[9] M. Goldmann. A note on the power of majority gates and modular gates. *Information Processing Letters*, 53(6):321–327, 1995.

[10] F. Green. Exponential sums and circuits with a single threshold gate and mod-gates. *Theory of Computing Systems*, 32(4):453–466, 1999.

[11] F. Green, A. Roy, and H. Straubing. Bounds on an exponential sum arising in boolean circuit complexity. *C. R. Acad. Sci. Paris, Ser. I*, 341:279–282, 2005.

[12] A. Hajnal, W. Maass, P. Pudlák, M. Szegedy, and G. Turán. Threshold circuits of bounded depth. *Journal of Computer and System Sciences*, 46(2):129–154, 1993.

[13] K. A. Hansen and P. B. Miltersen. Some meet-in-the-middle circuit lower bounds. In *29th International Symposium on Mathematical Foundations of Computer Science*, volume 3153 of *Lecture Notes in Computer Science*, pages 334–345. Springer, 2004.

[14] J. Håstad. *Computational limitations of small-depth circuits.* MIT Press, 1987.

[15] N. Nisan and A. Wigderson. Hardness vs randomness. *Journal of Computer and System Sciences*, 49(2):149–167, 1994.

[16] A. A. Razborov. Lower bounds for the size of circuits of bounded depth with basis ($\wedge$, $\oplus$). *Mathematical Notes of the Academy of Science of the USSR*, 41(4):333–338, 1987.

[17] R. Smolensky. Algebraic methods in the theory of lower bounds for Boolean circuit complexity. In *19th Annual ACM Symposium on Theory of Computing*, pages 77–82, 1987.

[18] E. Viola. Pseudorandom bits for constant depth circuits with few arbitrary symmetric gates. In *Proceedings of the 20th Annual IEEE Conference on Computational Complexity*, pages 198–209. IEEE Computer Society, 2005.