Correlation lower bounds from correlation upper bounds

Shiteng Chen Periklis A. Papakonstantinou
Tsinghua University

Abstract

We show that for any coprime $m, r$ there is a circuit of the form $\text{MOD}_m \circ \text{AND}_{d(n)}$ whose correlation with $\text{MOD}_r$ is at least $2^{-O\left(\frac{n}{\log n}\right)}$. This is the first correlation lower bound for arbitrary $m, r$, whereas previously lower bounds were known for prime $m$. Our motivation is the question posed by Green et al. [11] to which the $2^{-O\left(\frac{n}{\log n}\right)}$ bound is a partial negative answer. We first show a $2^{-\Omega(n)}$ correlation upper bound that implies a $2^{\Omega(n)}$ circuit size lower bound. Then, through a reduction we obtain a $2^{-O\left(\frac{n}{\log n}\right)}$ correlation lower bound. In fact, the $2^{\Omega(n)}$ size lower bound is for $\text{MAJ} \circ \text{ANY}_{o(n)} \circ \text{AND} \circ \text{MOD}_r \circ \text{AND}_{O(1)}$ circuits, which can be of independent interest.

1 Introduction

Understanding the power of small-depth circuits that have $\text{MOD}_m$ gates, in addition to the usual boolean gates, is one of the most fascinating areas of computational complexity. $\text{MOD}_m$ is the boolean function that outputs 1 if and only if the number of 1s in its input is a multiple of $m$. The computational limitations of $\text{MOD}_m$ gates for prime $m = p$ is well-understood since 1980s through the seminal works of Razborov [14] and Smolensky [15]. They proved that no constant depth polynomial size circuit with $\{\text{MOD}_p, \text{AND}, \text{OR}, \text{NOT}\}$ gates can compute the $\text{MOD}_q$ function, for primes $p \neq q$. Smolensky further conjectured that the same holds true for composite moduli, which remains an important open question.

A main tool in the study of small-depth circuit lower bounds is via correlation upper bounds [2, 3, 7, 8, 9, 11, 13]. The notion of correlation quantifies the distance of two functions and was introduced by Hajnal et al. [13]; see p. 2 for definitions. The smaller the correlation between the circuit and a function the larger the circuit size to compute this function.

In this note we show a limitation of the correlation method, aiming to answer the question of Green et al. [11]. They asked whether it is possible to prove correlation upper bounds that yield size lower bounds for circuits of the form $\text{MOD}_m \circ \text{AND}_{\omega(\log n)}$, which correspond to functions $\text{MOD}_m(P(x))$, for a polynomial $P$ of degree $\omega(\log n)$. We show a correlation lower bound between $\text{MOD}_r$ and $\text{MOD}_m(P(x))$ where $m \in \mathbb{Z}$ is anything and $P$ is of any degree. Previously, Green [10] and Viola [17] discussed correlation lower bounds that differ from ours. Viola’s argument is for the correlation between symmetric functions and polynomials of degree $\sqrt{n}$ (i.e. high degree) over $\text{GF}(2)$ (in fact, $\text{GF}(p)$ for prime $p$ and his result is incomparable to ours), whereas Green’s argument is only about $\text{MOD}_2$ and $\text{MOD}_3$.

Our goal is to lower bound the correlation between $\text{MOD}_r$ and any circuit $C_{\text{simple}}$ with a single layer of $\text{MOD}_m$. This is shown in two steps. In the first step we obtain a correlation upper bound
but for more complicated circuits $C_{\text{multi-layer}}$, which in particular includes circuits with two MOD layers. This correlation upper bound implies a circuit size lower bound for $C_{\text{multi-layer}}$. In the second step we do a reduction to obtain the lower bound on the correlation of a specific $C_{\text{simple}}$ and $\text{MOD}_r$.

There is considerable success in using correlation upper bounds in obtaining circuit lower bounds. In our argument we need to lower bound the size of circuits of the form $\text{MAJ} \circ \text{ANY}_{o(n)} \circ \text{AND} \circ \text{MOD}_r \circ \text{AND}_d(n)$, for which no previous lower bounds were known.

Hajnal et al. [13] showed the discriminator lemma, according to which upper bounded correlation of $f, g$ implies a lower bound for circuits of the form $\text{MAJ} \circ f$ that compute $g$. $\text{MAJ}$ outputs 1 if and only if the majority of input bits is 1. Cai et al. [3] studied depth 3 circuits of the form $\text{MAJ} \circ f \circ g$, where $f$ and $g$ are of AND, $\text{MOD}_r$ and introduced the analytic study of exponential sums, which is important for our work as well. Their results were for symmetric MOD functions, later generalized by Green [9], whereas Bourgain [2] (for odd moduli) and Green et al [11], and Chattopadhyay [5] (best known constants), finally showed an exponential size lower bound for $\text{MAJ} \circ \text{MOD}_m \circ \text{AND}_{O(1)}$ computing $\text{MOD}_q$, when $m, q$ are coprime, i.e. $(m, q) = 1$. For OR $\circ$ MOD circuits, i.e. linear systems over mod $m$, Chattopadhyay and Wigderson [8] showed exponential small correlation with $\text{MOD}_q$ for restricted $m$ and the general abelian case was handled by Chattopadhyay and Lovett [7].

For two layers of MOD gates, Grolmusz et al. [12] and Caussinus [4] studied $\text{MOD}_m \circ \text{MOD}_r$ circuits computing the AND function and proved, for any $m, r$, exponential circuit size lower bounds. Barrington and Straubing [1] considered $\text{MOD}_p \circ \text{MOD}_m$ circuits and proved an exponential size lower bound for such circuits computing $\text{MOD}_q$, where $p$ is a prime and $(p, q) = (m, q) = 1$. Straubing [10] introduce a finite field representation of MOD gates and simplified the previous proofs [11, 12]. Chattopadhyay et al. [6] studied $\text{MOD}_r \circ \text{MOD}_m$ to compute $\text{MOD}_q$, where $(r, q) = (m, q) = 1$, for composite $r$. The authors proved that the fan-in of the output $\text{MOD}_r$ gate, or any ANY gate, must be $\Omega(n)$.

## 2 Notations and prerequisites

All operations in this note are over $\mathbb{C}$, e.g. in evaluating a polynomial function $P : \{0, 1\}^n \to \mathbb{Z}$ with integer coefficients the operations treat the inputs 0, 1 as integers. We write $||x||_1 := \sum_{i=1}^{n} x_i$ for $x \in \{0, 1\}^n$ and denote by $\text{MOD}_m$ the boolean function (gate), where $\text{MOD}_m(||x||_1) = 1$ if $m||x||_1$ and 0 otherwise; not to be confused with the modulus over $\mathbb{X}$. Thus, polynomial functions take inputs $\{0, 1\}^n$ and MOD functions take inputs from $\mathbb{Z}$. For $X \in \mathbb{Z}$ we write $e_m(X) := e^{X \frac{2\pi i}{m}}$, where $e^{2\pi i/m}$ is the $m$-th primitive root of 1. Then, $\text{MOD}_m(X) = \frac{1}{m} \sum_{0 \leq k < m} e_m(kX)$. The correlation of the boolean functions $f, g : \{0, 1\}^n \to \{0, 1\}$ is defined as $\text{Corr}(f, g) = \left| \Pr_x[f(x) = 1 \mid g(x) = 1] - \Pr_x[f(x) = 1 \mid g(x) = 0] \right| = \frac{E_x[f(x) \cdot g(x)] - E_x[f(x) \cdot (1-g(x))]}{Pr_x(g(x)=1) - Pr_x(g(x)=0)}$. We extend the definition for $f : \{0, 1\}^n \to \mathbb{C}$ and $g : \{0, 1\}^n \to \{0, 1\}$ so that $\text{Corr}(f, g) = \left| \frac{E_x[f(x) \cdot g(x)] - E_x[f(x) \cdot (1-g(x))]}{Pr_x(g(x)=1) - Pr_x(g(x)=0)} \right|$. Now, let us state an observation we made, which is repeatedly used later on.

**Observation 1** (sub-additivity). Let functions $f_1, f_2 : \{0, 1\}^n \to \mathbb{C}$ and boolean function $g$. Then, $\text{Corr}(f_1 + f_2, g) \leq \text{Corr}(f_1, g) + \text{Corr}(f_2, g)$ and $\text{Corr}(c \cdot f, g) = |c| \cdot \text{Corr}(f, g)$, for constant $c \in \mathbb{C}$.

The main tool for proving $\text{MAJ} \circ \text{ANY}$ circuit lower bounds is the following lemma [13]. In fact, this lemma applies not only to MAJ but to any threshold gate.

**Lemma 2** (discriminator lemma [13]). Let $T$ be a circuit consisting of a majority gate over sub-circuits $C_1, C_2, \ldots, C_s$ each taking $n$-bit inputs. Let $f$ be the function computed by this circuit. If $\text{Corr}(C_i(x), f(x)) \leq \epsilon$ for each $i = 1, \ldots, s$, then $s \geq 1/\epsilon$. 

2
We use the above lemma together with elementary analytic techniques. The analytic machinery is explicit in the statement of the following Lemma 3.

**Lemma 3.** [see [5]] For any \( m, q, k \in \mathbb{Z}^+ \), \( (m, q) = 1 \), \( P \) a polynomial function with integer coefficients, \( \deg(P) = O(1) \), and \( x \in \{0, 1\}^n \), then \( \text{Corr}(e_m(P(x)), \text{MOD}_q(||x||_1)) \leq 2^{-\Omega(n)} \).

We represent functions \( f : \{0, 1\}^n \to \{0, 1\} \) as \( f(x) = \sum_{S \subseteq \{1, 2, \ldots, n\}} \alpha_S \prod_{i \in S} x_i \). This representation is unique since the functions \( \{\prod_{i \in S} x_i | S \subseteq \{1, 2, \ldots, n\}\} \) form a function basis\(^{1}\) for \( \{0, 1\}^n \to \mathbb{C} \). These basis functions are not to be confused with the fourier basis, which consists of the characters written multiplicatively (\( \{\{\{-1, 1\}^n \to \{-1, 1\}\} \)). We also introduce the definition of \( \text{norm}(f) := \sum_S |\alpha_S| \), which is particularly useful for our purposes.

### 3 Our results: statements and proofs

Our main results are Theorem 4 which states the circuit lower bound, and Theorem 5 which states the correlation lower bound. Note that Theorem 4 is used to show Theorem 5.

To simplify expression we represent a family of functions \( \{g_m\}_m \) by one \( g \in \{g_m\}_m \).

**Theorem 4.** Let \( n \) be the input length to circuits and \( \deg_g = o(n) \). Fix arbitrary \( g : \{0, 1\}^{\deg_g} \to \{0, 1\} \) and \( m, q \in \mathbb{Z}^+ \), where \( (m, q) = 1 \). If a \( \text{MAJ} \circ g \circ \text{AND} \circ \text{MOD}_m \circ \text{AND}_{O(1)} \) circuit computes \( \text{MOD}_q \), then the fanin of the MAJ gate on the top is \( 2^{\Omega(n)} \).

**Theorem 5.** For every \( d \in \mathbb{Z}^+ \) and every \( m, q \in \mathbb{Z}^+, (m, q) = 1 \) there exists a degree \( d \) polynomial \( P \) such that \( \text{Corr}(\text{MOD}_m(P(x)), \text{MOD}_q(||x||_1)) \geq 2^{-O(\frac{d}{q})} \).

### 3.1 Proof of Theorem 4 via a correlation upper bound

First, the sub-additive properties of correlation (Observation 1) yield the following lemma.

**Lemma 6** (bounded correlation amplifier). For every \( d, m, q \in \mathbb{Z}^+, (m, q) = 1 \) and every \( g : \{0, 1\}^{\deg_g} \to \{0, 1\} \) and polynomial functions \( P_i(x), x \in \{0, 1\}^n \), whose degrees are \( \deg(P_i(x)) \leq d \) we have

\[
\text{Corr}(g(\text{MOD}_m(P_1(x))), \text{MOD}_m(P_2(x)), \ldots, \text{MOD}_m(P_{\deg_g}(x))), \text{MOD}_q(||x||_1)) \\
\leq \text{norm}(g) \cdot \max_{P(x) \in \mathbb{Z}^n, \deg(P) \leq d} (\text{Corr}(e_m(P(x)), \text{MOD}_q(||x||_1))) \]

In particular, for \( P_i(x) = O(1) \) we have

\[
\text{Corr}(g(\text{MOD}_m(P_1(x))), \text{MOD}_m(P_2(x)), \ldots, \text{MOD}_m(P_{\deg_g}(x))), \text{MOD}_q(||x||_1)) \leq \text{norm}(g) \cdot 2^{-\Omega(n)}
\]

**Proof.** Let \( y_i = \text{MOD}_m(P_i(x)) \) and \( y = (y_1, y_2, \ldots, y_{\deg_g}) \) the input to \( g \). Now, let \( g(y) = \sum_{S \subseteq \{1, \ldots, \deg_g\}} \alpha_S \prod_{i \in S} y_i \). Therefore we have the following.

\[
\text{Corr}(g(\text{MOD}_m(P_1(x))), \text{MOD}_m(P_2(x)), \ldots, \text{MOD}_m(P_{\deg_g}(x))), \text{MOD}_q(||x||_1)) \\
= \text{Corr}(g(y), \text{MOD}_q(||x||_1)) \\
= \text{Corr}(\sum_{S \subseteq \{1, \ldots, \deg_g\}} \alpha_S \prod_{i \in S} y_i, \text{MOD}_q(||x||_1))
\]

\(^{1}\)Since \( \prod_{i \in S} x_i \prod_{i \notin S}(1 - x_i) \) is the standard basis and the dimension of the function space is \( 2^n \).
Proof. Fix arbitrary \( g \) and suppose \( \deg g \leq 3 \). Then, \( g \) is simply a monomial on \( x \). Finally, by Lemma 6 we have that \( \operatorname{Corr}(g, \operatorname{MOD}_m \circ \operatorname{AND}_O(1), \operatorname{MOD}_q(||x||_1)) \leq \norm(g) \cdot \max_{P(x) \in \mathbb{Z}[x], \deg(P) \leq 2} \left( \operatorname{Corr}(e_m(P(x)), \operatorname{MOD}_q(||x||_1)) \right) = \norm(g) \cdot \max_{P(x) \in \mathbb{Z}[x], \deg(P) \leq 2} \left( \operatorname{Corr}(e_m(P(x)), \operatorname{MOD}_q(||x||_1)) \right) \\

The second part of the statement follows by Lemma 3.

The above lemma shows the relation between correlation bounds and norm bounds. Now, we show a norm bound, which together with Lemma 6 concludes Theorem 8 below.

Lemma 7. For every \( g : \{0,1\}^{\deg g} \to \{0,1\} \) we have \( \norm(g) \leq 3^{\deg g} \).

Proof. We proceed by induction on \( \deg g \). If \( \deg g = 0 \) then \( g = 0 \) or \( g = 1 \), that is \( \norm(g) = 0 \) or 1. Suppose the predicate holds for \( \deg g \leq k \). For \( \deg g = k + 1 \) let the polynomial representation of \( g \) be \( g(x_1, x_2, \ldots, x_{k+1}) = P_1(x_1, x_2, \ldots, x_k) + x_{k+1} \cdot P_2(x_1, x_2, \ldots, x_k) \), i.e. \( g|_{x_{k+1}=0} = P_1 \), \( g|_{x_{k+1}=1} = P_1 + P_2 \). Then, \( P_1 = g|_{x_{k+1}=0} \) and \( P_2 = g|_{x_{k+1}=1} - g|_{x_{k+1}=0} \). Since \( g|_{x_{k+1}=1} \) and \( g|_{x_{k+1}=0} \) are boolean function on \( k \) variables, by the induction hypothesis we have \( \norm(g|_{x_{k+1}=1}) \leq 3^k \) and \( \norm(g|_{x_{k+1}=0}) \leq 3^k \). Then, \( \norm(g) = \norm(g|_{x_{k+1}=0} + x_{k+1} \cdot (g|_{x_{k+1}=1} - g|_{x_{k+1}=0})) \leq 3 \cdot 3^k = 3^{k+1} \).

Theorem 8. Fix arbitrary \( g : \{0,1\}^{\deg g} \to \{0,1\} \) where \( \deg g = o(n) \) and \( (m, q) = 1 \). Then, the correlation between \( g \circ \operatorname{MOD}_m \circ \operatorname{AND}_O(1) \) circuit and \( \operatorname{MOD}_q(||x||_1) \) is \( 2^{-\Omega(n)} \).

Proof. By Lemma 7 we have \( \norm(g) \leq 2^{\Omega(\deg g)} = 2^{o(n)} \), and thus the above observation yields \( \norm(g \circ \operatorname{AND}) \leq \norm(g) \leq 2^{o(n)} \). Finally, by Lemma 6 we have that \( \operatorname{Corr}(g \circ \operatorname{MOD}_m \circ \operatorname{AND}_O(1), \operatorname{MOD}_q(||x||_1)) \leq \norm(g) \cdot 2^{-\Omega(n)} \leq 2^{-\Omega(n)} \).
3.2 Proof of Theorem 5: the correlation lower bound

We stated the lower bound of Theorem 4 in the most general form we could obtain (since it is also of independent interest). Now, we give the proof of Theorem 5 where we only need to show how to write MOD_q as a ANY \circ MOD_m \circ AND_d circuit, for a function ANY = g that we determine later.

Here is the main tool used to obtain Theorem 5.

**Theorem 9.** For every \( d \in \mathbb{Z}^+ \) and \( m, q \in \mathbb{Z}^+, (m, q) = 1 \) there exists a degree \( d \) polynomial \( P \), such that \( \text{Corr}(e_m(P(x)), \text{MOD}_q(||x||_1)) \geq 2^{-O\left(\frac{d}{4}\right)} \).

**Proof.** Let \( M_d \) be such that for every \( d \in \mathbb{Z}^+ \) and \( m, q \in \mathbb{Z}^+, (m, q) = 1 \) we have

\[
\max_{P(x) \in \mathbb{Z}[x], \deg(P) \leq d} (\text{Corr}(e_m(P(x)), \text{MOD}_q(||x||_1))) = M_d
\]

Split \( \{x_1, x_2, \ldots, x_n\} \) into \( n/d \) subsets \( S_i = \{x_{id+1}, x_{id+2}, \ldots, x_{(i+1)d}\} \) for \( i = 1, 2, \ldots, n/d \), where for simplicity we assume \( d|n \). Now, use \( \log q \) bits (all logarithms are of base 2) to encode the value of each \((\sum_{j \in S_i} x_j) \mod q\). Thus, using \( \frac{n\log q}{d} \) bits denoted by \( b_{1,1}, b_{1,2}, \ldots, b_{1,\log q}, b_{2,1}, \ldots, b_{n,\log q} \) we can compute \( \text{MOD}_q(||x||_1) \). We define \( g \) such that \( \text{MOD}_q(||x||_1) = g(b_{1,1}, b_{1,2}, \ldots, b_{1,\log q}, b_{2,1}, \ldots, b_{n,\log q}) \).

Since \( \text{MOD}_m(1-y) = y \) for any \( y \in \{0, 1\} \), we have \( \text{MOD}_q(||x||_1) = g(\text{MOD}_m(1-b_{1,1}), \text{MOD}_m(1-b_{1,2}), \ldots, \text{MOD}_m(1-b_{\log q,1})) \). Since \( b_{i,j} \) is a function on variables \( \{x_{id+1}, x_{id+2}, \ldots, x_{(i+1)d}\} \), we can represent \( 1-b_{i,j} \) as a polynomial \( P_{i,j} \) on \( d \) variables and hence \( \deg(P_{i,j}) \leq d \). Thus, \( \text{MOD}_q(||x||_1) = g(\text{MOD}_m(P_{1,1}), \text{MOD}_m(P_{1,2}), \ldots, \text{MOD}_m(P_{\log q,1})) \), which we use to obtain the following.

\[
\begin{align*}
\text{Corr}(\text{MOD}(||x||_1), \text{MOD}(||x||_1)) &= \text{Corr}(g(\text{MOD}_m(P_{1,1}), \text{MOD}_m(P_{1,2}), \ldots, \text{MOD}_m(P_{\log q,1})), \text{MOD}(||x||_1)) \\
&\leq \text{norm}(g)M_d \leq 2^{\Omega\left(\frac{d}{4}\right)}M_d \quad \text{(by Lemma 7 – used with different parameters than in Theorem 8)}
\end{align*}
\]

On the other hand, by the definition of correlation we have that \( \text{Corr}(\text{MOD}(||x||_1), \text{MOD}(||x||_1)) = 1 \), and thus \( 1 \leq 2^{\Omega\left(\frac{d}{4}\right)}M_d \) that implies \( M_d \geq 2^{-O\left(\frac{d}{4}\right)} \). \( \square \)

Since \( e_m(X) \) is a linear combination of \( \text{MOD}(X), \text{MOD}(X-1), \ldots, \text{MOD}(X-m+1) \) we conclude Theorem 5.

**Proof of Theorem 5.** Let \( P' \) be a polynomial of degree at most \( d \) such that \( \text{Corr}(e_m(P'(x)), \text{MOD}_q(||x||_1)) \geq 2^{-O\left(\frac{d}{4}\right)} \). Since \( e_m(P'(x)) = \sum_{0 \leq i < m} e_m(i) \text{MOD}(P'(x) - i) \), by Observation 1 we have \( \frac{1}{2^{O\left(\frac{d}{4}\right)}} \leq \text{Corr}(e_m(P'(x)), \text{MOD}_q(||x||_1)) \leq \sum_{0 \leq i < m} \text{Corr}(\text{MOD}(P'(x) - i), \text{MOD}(||x||_1)) \).

Then, there exists \( 0 \leq i < m \) such that \( \text{Corr}(\text{MOD}_m(P'(X) - i), \text{MOD}_q(||x||_1)) \geq 2^{-O\left(\frac{d}{4}\right)} \).

\( \square \)

**Acknowledgments**

We wish to thank Arkadev Chattopadhyay, Shachar Lovett, and Emanuele Viola for useful suggestions.
References


