A note on lower bounds on constant-depth modular circuits

Tatsuie Tsukiji

School of Informatics and Sciences, Nagoya University, Nagoya 464-01, JAPAN
(tskijit@info.human.naoya-u.ac.jp)

Keywords: constant-depth circuits, lower bounds, modular gates

1 Introduction

A Boolean function AND outputs 1 if and only if all the input variables are 1, and a Boolean function MOD$_m$ for a constant $m \geq 2$ outputs 1 if and only if the number of 1 in the input variables is equal to a multiple of $m$. In the early of 80's Ajtai [Ajt83] and Furst, Saxe and Sipser [FSS84] have shown that AND-type gates cannot compute MOD$_m$ gate efficiently in a model of constant-depth circuits. This note proves a converse: MOD$_m$ gates cannot compute AND gate efficiently.

A fundamental task in computational complexity theory is to reveal intrinsic computational difficulty of finite problems appeared in nowadays computer and cryptographic systems. Currently, the class of constant-depth circuits with unbounded fan-in is widely acknowledged as a model for the first step of proving complexity lower bounds of concrete problems. As usual, a Boolean circuit of depth $d$ is a parallel computing network (an unbounded fan-in undirected graph) consisting from $d$ layers. Each layer contains nodes called Boolean gates whose input wires are coming from gates in the previous layer and the output wires are going into gates in the next layer, except that the input wires of the gates in the initial (bottom) layer are issued from input Boolean variables or their negations (i.e. literals.) Each gate computes a designated Boolean function and the full circuit computes a Boolean function at the unique gate in the end (top) layer. AC$^0$-circuits are constant-depth circuits using logical gates \{AND, OR\} and AC$^0$ is the class of languages recognized by a sequence of polynomial-size AC$^0$ circuits. This class contains some basic functions in
computing, e.g. the addition of two $n$ bit numbers. Limitation of the computing power of $AC^0$ is widely known. As we have mentioned, Ajtai [Ajt83] and Furst Saxe and Sipser [FSS84] have proved that $MOD_m \not\subseteq AC^0$. Later on, Yao [Yao85] and Håstad [Has86] have improved it to exponential lower bounds. Thus, logical gates cannot compute a modular gate efficiently. Adding $MOD_m$-gates to $AC^0$ circuits defines $AC^0(m)$-circuits, hence $AC^0(m)$ is a super class of $AC^0$ in a strict sense. This extension of computing power seems a mere matter at first glance, yet previous lower bounds on the class $AC^0(m)$ are limited in case that $m$ is a power of a prime number. Razborov [Raz87] has proved an exponential lower bound of the majority function on $AC^0(p^k)$ and Smolensky [Smo87] has proved an exponential lower bound of $MOD_r$ on $AC^0(p^k)$ if $r$ is not a power of $p$. If the circuit depth is restricted as 2, then we have lower bounds for $m$ that is not a prime power. A $MOD_m \circ MOD_{m'}$-gate at the top followed by $MOD_{m'}$-gates in the bottom. Krawse and Waack [KW91] have proved an exponential lower bound of the equality function on $MOD_m \circ MOD_{m'}$-circuits for any $m$ and $m'$ (more generally bottom gates can be any kind of symmetric gates.) Krawse and Pudlák have proved an exponential lower bound of $MOD_q$ on $MOD_{p^k} \circ MOD_{r}$-circuits, where $p$ and $q$ are distinct primes and $r$ is any integer. However, for depth-3 circuits consisting from modular gates, even $MOD_6 \circ MOD_6 \circ MOD_6 \neq NP$ has been a long-standing open conjecture.

This note attacks lower bounds on circuits consisting from purely modular gates. The class $CC^0(m)$ is the class of languages recognized by a sequence of polynomial-size constant-depth circuits consisting from $MOD_m$-gates and $CC^0 = \cup_{m \geq 2} CC^0(m)$ [MPT91] (Yao called the class pure-ACC [Yao90].) We shall prove the next theorem.

**Theorem 1** $\text{AND} \not\subseteq CC^0$.

## 2 Proof

We suppose that $\text{AND} \in CC^0$ and derive a contradiction. We fix a large integer $n$ for the dimension (bit length) of input Boolean assignment, hence we take assignments from the $n$-dimensional Boolean cube $N = \{0, 1\}^n$. We sometimes use $N$ for the number $2^n$. In fact we show that the equality (or identity) function $EQ(x, y)$ is hard to compute on $CC^0$, where

$$EQ(x, y) = \begin{cases} 1 & \text{if } x = y \\ 0 & \text{otherwise} \end{cases}$$
We shall evaluate a bilinear form producing the characteristic vector $\mathrm{EQ}(y) = \langle \mathrm{EQ}(x, y) : x \in N \rangle$ of $\mathrm{EQ}$ (the $y$-th coordinate vector) and a given $N$-dimensional vector $z$ in two different ways and obtain conflicting values. The direct evaluation yields

$$q(\mathrm{EQ}(y), z) = \sum_{x \in N} z(x) \mathrm{EQ}(x, y) = z(y),$$

the projection of $z$ to the $y$th coordinate. On the other hand, we shall show that the assumption $\mathrm{EQ} \in \mathrm{CC}^0$ would derive a small set $D \subseteq N$ and a prime number $p$ such that for any $z$ and almost all $y$ an appropriate variation of $z$ on $D$ makes $q(\mathrm{EQ}(y), z) \equiv 0 \pmod{p}$. Strictly speaking, we say that $D$ spans $N$ modulo $p$ almost everywhere if for any $\epsilon > 0$ there exists $n_0$ such that for any $n \geq n_0$, any $z$ and at least $1 - \epsilon$ fraction of $y \in N$, there exists an integer vector $\tau(y, z)$ that satisfies both $\tau(y, z)(N - D) = \{0\}$ and $q(\mathrm{EQ}(y), z + \tau(y, z)) \equiv 0 \pmod{p}$. We call $y$ in the fraction good for $D$. Then we shall show that if $\mathrm{EQ} \in \mathrm{CC}^0$ then there is a small set $D$ that spans $N$ modulo a certain prime number $p$ almost everywhere. This claim derives a contradiction in the following way: Take any good $y \in N - D$ and any $z$ such that $z(y) = 1$. Then we must conclude that $q(\mathrm{EQ}(y), z + \tau(y, z)) = (z + \tau(y, z))(y) = z(y) = 1 \equiv 0 \pmod{p}$. A contradiction.

Razborov and Smolensky have used low degree polynomials over finite fields for obtaining their lower bounds. Here we use low degree polynomials over the integer ring of characteristic $0$. Therefore a Boolean polynomial $P$ is a linear combination of AND's of some (positively occurred) Boolean variables with integer coefficients. Each AND is called a term whose cardinality (the number of variables in it) is called the degree. As usual, we call the maximum of the degree of a term with non-zero coefficient the degree of $P$ and write $d(P)$, and the sum of the absolutes of the coefficients the norm of $P$ and write $n(P)$. We denote by $\text{mod}_m(x)$ the unique modulo of $x$ in the range $-\left\lfloor \frac{m}{2} \right\rfloor + 1 \leq \text{mod}_m(x) \leq \left\lfloor \frac{m}{2} \right\rfloor$ divided by $m$. As usual $f \circ g(x)$ of functions $f$ and $g$ (from the set of integers to the set of integers) denotes the composite function $f(g((x)))$. Our proof is founded on Yao's simulation of $\mathrm{CC}^0$-circuits by low degree polynomials over the integer ring [Yao90]. He has used modulus amplification investigated by Toda [Tod89] and collapsed modular hierarchies of different primes (see also [BT94]).

**Theorem 2 (Yao)** Given a language $L$ that is recognized by a sequence of depth-$d$ $\mathrm{CC}^0$ circuits and a polynomial $Q$. There is a constant $c > 1$ and for any $n > 0$ there are prime powers $q_1, \ldots, q_d > n$ and a Boolean polynomial $P$ of degree $O((\log n)^c)$ and norm $O(n^{(\log n)^c})$ such that

$$L(x) = \text{mod}_{q_d} \circ \cdots \circ \text{mod}_{q_1}(P(x))$$
holds for any $x$.

We apply this theorem for rewriting $\mathrm{EQ}(x, y)$ as

$$\mathrm{EQ}(x, y) = \mod_{q_d} \circ \cdots \circ \mod_{q_1}(P(x, y))$$

(1)

where $q_1, q_2, \ldots, q_d$ are prime powers greater than $n$ and $P(x, y)$ is a Boolean polynomial of degree $O((\log n)^c)$ and norm $O(n^{(\log n)^c})$ for a constant $c > 0$. A merit of this expression of $\mathrm{EQ}$ is that we can decompose $P(x, y)$ into a small number of productions of $x$-functions (functions that depends only on $x$) and $y$-functions in the following way:

$$P(x, y) = \sum_{t} t(x) P_t(y)$$

where the degrees of $t$ are at most $d(P)$ and the sum of the norms of $P_t$ is at most $n(P)$. If there were no barrier of moduli functions and $\mathrm{EQ}(x, y) = P(x, y)$ held then a rank argument on communication matrices would immediately derive evaluations of $q(\mathrm{EQ}(y), z)$ that conflict to the direct one. Thus we undertake to transport the above decomposition of $P(x, y)$ through moduli functions until reaching to a decomposition of $\mathrm{EQ}(x, y)$ that can derive unexpected evaluations of $q(\mathrm{EQ}(y), z)$.

We prepare a terminology. For a term $t$ let $t^{\ast}$ be the minimal satisfiable assignment of $t$ (\((x_1 x_2 x_3)^{\ast} = 101010^{\ast-5})\) and call $t^{\ast}$ the dual assignment of $t$. We denote by $D$ the set of $x$-terms of degree at most $d(P)$ and $D^{\ast}$ the set of the dual assignments of terms in $D$. We may assume that all $t \in D$ appears in $P(x, y)$ by allowing $P_t(y) = 0$.

Now we prove the claim for $D^{\ast}$ and finish the proof.

Claim 1 $D^{\ast}$ spans $N$ modulo a certain prime number $p \geq 2$ almost everywhere.

Proof of Claim. $D$ is a basis of the field of the real functions defined on $D^{\ast}$, hence we can normalize it as follows. We denote terms in $D$ as $t, u$ and $v$. Let $Q_t(x) = \sum_{t \subseteq u} (-1)^{d(t) - d(u)} u(x)$. Then we obtain

$$Q_t(u^{\ast}) = \delta_{t,u}$$

(2)

because if $t(u^{\ast}) = 0$ then $Q_t(u^{\ast}) = 0$ and otherwise we have

$$Q_t(u^{\ast}) = \sum_{t \subseteq u \subseteq v} (-1)^{d(t) - d(u)} = \sum_{i=0}^{\text{card}(u) - \text{card}(t)} (-1)^i (\text{card}(u) - \text{card}(t)) = \delta_{t,u}$$
In order to implement this normalization in the evaluation of $P(x, y)$ we need to linearly transform $y$-polynomials as $R_t(y) = \sum_{u \subseteq t} P_u(y)$. Then we obtain

$$P(x, y) = \sum_i Q_i(x) R_i(y)$$

(3)

because

$$\sum_t Q_t R_t = \sum_t P_t \sum_{u \subseteq v} Q_u = \sum_t P_t \sum_{u \subseteq v} \sum_{i=0}^{\text{card}(v) - \text{card}(t)} (-1)^{i} (\text{card}(v) - \text{card}(t))$$

$$= \sum_t P_t \sum_{i=0}^{\text{card}(v) - \text{card}(t)} \sum_{v \delta_i, v} = \sum_t P_t t$$

We transport this decomposition of $P(x, y)$ through moduli functions by evaluating bilinear forms invoked from the ith remainders appeared in (2). For an integer $a$ let $h_i(a) = \text{mod}_{q_i} \cdots \text{mod}_{q_1} a$ ($h_0(a) = a$) and call its value the ith remainder of $a$. Let $Q(x) = (Q_i(x) : t \in D)$ and $h_i(R(y)) = (h_i(R_t(y)) : t \in D)$. We wish to evaluate $P(x, y)$ by using $r(Q(x), h_i(R(y)))$ so we wish to switch order of the summation $\sum_t$ and the modular function $h_d$. For it we obtain switchings between $s_i$ and $\text{mod}_{q_i}(s_{i-1})$ beginning from $i = 1$ up to $i = d$ by adding up phase vectors to the first factor of $r$. Precisely, we claim that there are $\text{card}(D)$-dimensional integer vectors $\theta_i(x, y)$ (we call them phase vectors) such that for all $1 \leq i \leq d$ we have

$$\text{mod}_{q_i}(r(Q(x) + \theta_{\leq i-1}, h_{i-1}(R(y)))) = r(Q(x) + \theta_{\leq i}, h_i(R(y)))$$

(4)

where $\theta_{\leq i} = \sum_{j \leq i} \theta_j$. These consecutive switchings derive required remote switchings

$$h_i(P(x, y)) = r(Q(x), R(y))$$

(5)

for all $i$. Moreover, these hold for any $x \in D$ and $y$ with phase free (all $\theta_i(x, y) = 0$) because $R_t(y) = P(t^*, y)$ holds due to (2).

Now we fix arbitrary $x$ and $y$ and prove (4) by induction on $i$. At the $i$th stage we have already defined $\theta_j$ for all $j \leq i - 1$ and will define $\theta_i$ for getting (5).
We are enough to show the followings:

\[
\begin{align*}
    r(\theta_i, h_i(R)) &\equiv 0 \quad \text{(mod } q_i) \quad (6) \\
    -\left\lfloor \frac{q_i}{2} \right\rfloor + 1 < r(Q + \theta_{\leq i}, h_i(R)) < \left\lfloor \frac{q_i}{2} \right\rfloor \\
    &\equiv \left\lceil \frac{q_i}{2} \right\rceil - 1 \quad (7)
\end{align*}
\]

We call the greatest common divider of the components of an integer vector the period of the vector. We prove (6) and (7) by dividing into cases distinguished on the period of \( h_i(R_i(y)) \). First of all, if the period is divisible by \( q_i \) then (4) trivially holds. Hence we may assume that the period is not a multiple of \( q_i \). Secondly, if the period is 1 then we can define the \( i \)th phase so that we have

\[
r(\theta_i, h_i(R)) = \mod_{q_i} (r(Q + \theta_{\leq i-1}, h_i(R))) - r(Q(x) + \theta_{\leq i-1}, h_i(R))
\]

hence (6) holds. Moreover the linearity of the quadratic form derives

\[
r(Q + \theta_{\leq i}, h_i(R)) = r(Q + \theta_{\leq i-1}, h_i(R)) + (\theta_i, h_i(R)) = \mod_{q_i} (r(Q + \theta_{\leq i-1}, h_i(R)))
\]

hence (7) holds, too.

Thus (4) holds when the period of the vector \( h_i(R(y)) \) is equal to 1. We can forth this period in the following probabilistic argument. Choose one prime number \( p \) uniformly at random from the first \( \left\lfloor \frac{n}{(2 + o(1)) \log n} \right\rfloor \)  prime numbers.

The prime number theorem guarantees that the \( \left\lfloor \frac{n}{(2 + o(1)) \log n} \right\rfloor \)  th smallest prime number is smaller than \( \frac{q_i}{2 + o(1)} < \frac{q_i}{2 + o(1)} \), so \( h_i(p) = p \) hold for all \( i \). Moreover for every \( y \) the number of different prime factors in the period of \( h_i(R(y)) \) is \( O(\log(n(P))) \) because the grade of components of \( h_i(R(y)) \) is at most \( n(P) \) in absolute, hence the probability that \( p \) does not touch to any of these factors for \( 1 \leq i \leq d \) is smaller than \( O \left( \frac{d \log n \log(n(P))}{n} \right) = o(1) \). Therefore we have \( p \) that is relatively prime with all the periods of \( h_i(R(y)) \) with \( 1 \leq i \leq d \) for more than \( 1 - O \left( \frac{d \log n \log(n(P))}{n} \right) = 1 - o(1) \) fraction of \( y \in N \). We call \( y \) in the fraction good. In order to put the constant function \( p \) in the list of \( y \)-polynomials of a decomposition of \( P(x, y) \), we consider an abstract term \( s \) and define \( Q_s(x) = 0 \) and \( R_s(y) = p \). Thus we have \( P(x, y) = \sum Q_s(x) R_s(y) + Q_s(x) R_s. \) For this decomposition, if \( y \) is good then the period of the vector of \( y \)-polynomials (the concatenation of \( h_i(R(y)) \) and \( R_s = p \)) is 1. Therefore we have reduced the case of the general period to the case of period 1 and obtained (4) for almost all \( y \).

Finally, we show that \( D^* \) spans \( N \) modulo \( p \) for any good \( y \). We apply (5) for \( i = d \) and obtain

\[
q(EQ(y), z) = \sum z(x) r_d(Q(x) + \theta_{\leq d}(x, y), R(y))
\]
where \( \theta_{\leq d}(x, y) = 0 \) for all \( x \in D \) so we have

\[
q(\text{EQ}(y), z) = \sum_{t \in D} h_d(R_t(y)) \left( z(t^*) + \sum_{t \in N - D^*} z(t) (Q_t(z) + \theta_{\leq d}(x,y)(t)) h_d(R_t(y)) \right)
\]

\[
+ p \sum_{t \in N - D^*} z(t) \theta_{d}(x,y)(t)
\]

thus letting

\[
\tau(y, z)(x) = \begin{cases} 
- z(t^*) - \sum_{t \in N - D^*} (Q_t(z) + \theta_{\leq d}(x,y)(t)) & \text{if } x = t^* \in D^* \\
0 & \text{otherwise}
\end{cases}
\]

we obtain

\[
q(\text{EQ}(y), z + \tau(y, z)) = p \sum_{t \in N - D^*} z(t) \theta_{\leq d}(x)(t) \equiv 0 \pmod{q_i}
\]

Claim 1

References


