INFORMATION AND COMPUTATION 112, 117–130 (1994)

Exponential Size Lower Bounds for Some Depth Three Circuits

P. Y. YAN*

Department of Mathematics, The Pennsylvania State University

AND

IAN PARBERRY†

Department of Computer Science, The Pennsylvania State University

Exponential size lower bounds are obtained for some depth three circuits computing conjunction using one layer each of gates which compute Boolean functions of low total degree when expressed as polynomials, parity-modulo-2 gates, and parity-modulo-q gates, where q is prime. One of these results implies a special case of the *constant degree hypothesis* of Barrington *et al.* The lower bounds are obtained from an algebraic characterization of the functions computed by the circuits: it is shown that certain integer multiples of these functions can be expressed as the sum of a lattice element and a function of small value. It is conjectured that this characterization can be used to resolve the constant degree hypothesis.

1. INTRODUCTION

The complexity of unbounded fan-in Boolean circuits has been studied extensively for over a decade because of its mathematical elegance and strong connections with other sequential and parallel models of computation. Though the Boolean circuit model is considered by many as one of the most promising approaches to the $\mathcal{P} \neq \mathcal{NP}$ problem, there has not been great success in proving lower bounds for general Boolean circuits. The best known size lower bound for Boolean circuits which compute an $\mathcal{NP}$ complete problem is only linear, which is far away from the superpolynomial lower bound needed to separate $\mathcal{P}$ and $\mathcal{NP}$.

* Author's current adress: Department of Mathematics, Lycoming College, Williamsport, Pennsylvania 17701.

† Author's current address: Department of Computer Sciences, P. O. Box 13886, University of North Texas, Denton, Texas 76203-3886. Research supported by the Air Force Office of Scientific Research, Air Force System Command, USAF, under Grant AFOSR 87-0400.

0890-5401/94 $6.00

As with any difficult problem, it is prudent to begin with restricted cases. One of the most successful of these is the case of *constant depth* circuits. The study of constant depth unbounded fan-in Boolean circuits began with Furst, *et al* [4], and independently Ajtai [1], in an attempt to separate the polynomial time hierarchy from $\mathcal{PSPACE}$ by oracles. Since then, exponential lower bounds have been shown for constant depth circuits of conjunction, disjunction, and Boolean complementation gates computing the parity modulo 2 of n inputs (Hastad [6], Yao [16]), and circuits of conjunction, disjunction, Boolean complementation, and parity-modulo-2 gates computing the parity modulo 3 of n inputs (Razborov [10], Smolensky [12]). These circuit lower bounds often have important applications, including the separation of the polynomial time hierarchy by oracles (Sipser [11], Yao [16]), and proofs of lower bounds for finite automata over groups (Barrington *et al.* [3]).

Another motivation for the study of bounded depth unbounded fan-in circuits is their connection with neural networks. It was shown in Parberry and Schnitger [8, 9] that the complexity class TC^0, consisting of functions computable by polynomial size circuits of bounded depth with threshold gates, is equivalent to the class of functions computable by a certain connectionist model of neutral networks. Recent lower bound results concerning TC^0 include the separation of threshold circuits of depth 2 from depth 3 (Hajnal *et al.* [5]), the separation of monotone threshold circuits of depth k from depth $k+1$ for all k (Yao [15]), and lower bounds for some special depth 3 threshold circuits (Hastad and Goldmann [7]).

Currently, there are no nontrivial size lower bounds for constant depth circuits of conjunction, disjunction, and parity-modulo-2 and parity-modulo-3 gates. Barrington *et al.* recently showed an exponential lower bound for some special cases; they showed that depth 2 circuits with only parity-modulo-p and parity-modulo-q gates which compute the Boolean conjunction of n inputs must have exponential size [3]. Even this very restricted case requires a fairly deep proof. They conjecture that the result remains true even if the circuit is extended to depth 3 by inserting an extra level of gates which compute constant-degree polynomials of the inputs. They call this conjecture the *constant degree hypothesis.*

We provide a new algebraic technique for proving size lower bounds for special depth 3 circuits which consist of one layer each of parity-mod-2 gates, parity-mod-q gates, and a layer of gates which compute functions whose degrees (when expressed as polynomials) are a small linear function of the number of inputs to the circuit. We derive exponential size lower bounds for circuits computing conjunction in which the latter layer is either the first or second layer. We reduce the more general case in which parity-mod-2 gates are replaced by parity-mod-p gates, and the constant degree hyothesis, to plausible algebraic questions which nonetheless remain open.

The remainder of this paper is divided into four sections. The first contains background and definitions. The second contains a new algebraic characterization of circuits of parity-modulo-2 and parity-modulo-q gates. The third uses this characterization to prove an exponential size lower bound for a special case of the constant degree hypothesis: the case in which the sum of the degrees of the first layer of gates is bounded above by a small linear function in n (the second and third layers are parity-modulo-2 and parity-modulo-q gates, respectively). It also contains an exponential size lower bound for a similar circuit in which the order of the layers is changed. The fourth and final section uses the characterization to reduce the case in which parity-mod-2 gates are replaced by parity-mod-p gates, and the constant degree hypothesis to an algebraic question which is left as an open problem. A preliminary version of the result in this paper forms part of the first author's Ph.D. thesis [14].

2. Background and Definitions

A *circuit* is a directed acyclic graph. The sources of the graph are called *inputs*, the sinks are called *outputs*, and the other nodes are called *gates*. Each gate is labelled with a function f. We will call any gate which is labelled with f an *f-gate*. A circuit calculates a function of the input variables in the natural way. The value of an f-gate g is defined to be the function f applied to the values of those gates h such that there is an edge from h to g in the graph. Thus given a sequence of values for the inputs, the corresponding output of the circuit is defined to be the sequence of values of the outputs in some fixed order. We will consider circuits with a single output gate, that is, circuits that compute Boolean functions. The maximum in-degree of gates and outputs in the circuit is called the *fan-in* of the circuit. The number of nodes is called the *size* of the circuit, and the length of the longest path from an input to an output is called the *depth* of the circuit. Equivalently, the depth of a circuit is the number of layers of gates, where each gate in a given layer receives inputs only from the gates in the preceding one.

If p is a prime number, the *parity-modulo-p* function, written $\mathrm{Mod}[p]$, is defined as

$$\mathrm{Mod}[p](x_1, x_2, \ldots, x_n) = \begin{cases} 0 & \text{if } x_1 + \cdots + x_n = 0 \pmod p \\ 1 & \text{otherwise,} \end{cases}$$

where $n \geqslant 1$. In the future, we will use p, q to represent arbitrary primes unless otherwise stated.

We represent each n-input Boolean function as a polynomial in $x_1, x_2, \ldots, x_n$. Any n-input Boolean function can be uniquely written as a

polynomial in $x_1, x_2, ..., x_n$ over the ring of integers Z, where x_i satisfies the relation $x_i = x_i^2$ for $1 \leqslant i \leqslant n$. For example, we can write the conjunction of x and y as xy, and the disjunction of x and y as $x + y - xy$ (in both cases, concatenation denotes integer multiplication and "+" and "−" integer addition and subtraction, respectively). More formally, any Boolean function can be identified with an element $\mathbf{Z}[x_1, x_2, ..., x_n]/\langle x_1 - x_1^2, ..., x_n - x_n^2 \rangle$ by identifying the Boolean value "true" as the integer 1 and the Boolean value "false" as the integer 0, where $\mathbf{Z}[x_1, x_2, ..., x_n]$ is the polynomial ring of n indeterminates $x_1, x_2, ..., x_n$ over Z and $\langle x_1 - x_1^2, ..., x_n - x_n^2 \rangle$ is the ideal in $\mathbf{Z}[x_1, x_2, ..., x_n]$ generated by $x_i - x_i^2$ for $1 \leqslant i \leqslant n$. Note that such polynomials do not contain variables of exponent greater than unity because $x_i = x_i^2$, and therefore the term of highest possible degree is $x_1 x_2 \cdots x_n$. This representation of Boolean functions as polynomials over Z can be generalized in the sense that Z can be replaced by any other ring (with identity) or field. Let A be a ring or a field. We use U_A^n to denote $A[x_1, x_2, ..., x_n]/\langle x_1 - x_1^2, ..., x_n - x_n^2 \rangle$. If A is a field, then U_A^n is an algebra of dimension 2^n over A under the usual addition and multiplication. A complete description of the representation of Boolean functions by polynomials over fields can be found in Smolensky [12].

3. Circuits and Lattices

Let V be an n-dimensional vector space over the rationals Q. A *lattice* (of dimension $m \leqslant n$) is a Z-module generated by m vectors $v_1, v_2, ..., v_m$, that is, the set $\{\sum_{i=1}^{m} z_i v_i \mid z_i \in \mathbf{Z}\}$.

We will derive our lower bounds by expressing a certain integer multiple of the function computed by a circuit as a sum of a function from a lattice M and a function e, where the range of the function e consists of values bounded above by a polynomial in the size of the circuit. The expressing power of e is limited by the bound on its range. If the expressing power of M is also limited in a certain way then the expressing power of the sum is also limited.

Let $\mathrm{Poly}(f(i))$ be the class of polynomials in $U_{\mathbf{Z}}^n$ where each monomial of degree i has a coefficient which is an integer multiple of $f(i)$. In this paper we consider mainly lattices of the type $\mathrm{Poly}(p^{\lfloor i/c_1 - c_2 \rfloor})$, where $c_1, c_2 \in \mathbf{Z}$. The reason for considering lattices of this type is that they correspond to Z-modules generated by certain $\mathrm{Mod}[p]$ functions. In this and the next section, we consider the special case where $p = 2$ and $c_1 = 1$. The general case is considered in Section 5.

We define M_p to be the module over Z generated by all functions of the form

$$\mathrm{Mod}[p](y_1, y_2, ..., y_t),$$

where $y_i \in \{x_1, x_2, ..., x_n\}$ for $1 \leqslant i \leqslant t$. We first consider M_2, the simplest such module.

Mod[2]$(x_1, x_2, ..., x_n)$ can be written as

$$\text{Mod}[2](x_1, x_2, ..., x_n) = \frac{1 - \prod_{i=1}^{n} (1 - 2x_i)}{2}.$$

This expression is a polynomial of $x_1, x_2, ..., x_n$ over $\mathbf{Z}$ which has the property that the coefficient of a monomial of degree i is an integer multiple of 2^{i-1}. Since all other Mod[2] functions of input variables can be obtained by setting some of the variables in the above expression to 0 or 1, they all have the same property. Thus

$$M_2 \subseteq \text{Poly}(2^{i-1}).$$

Poly(2^{i-1}) is a lattice of dimension 2^n in U_0^n which has limited expressing power because the coefficients of high degree terms are restricted to be in some proper subset of $\mathbf{Z}$. We will use this limitation later to put restrictions on the power of depth 2 circuits of Mod[2] and Mod[q] gates.

In the next lemma, we observe a property of the function Mod[q] which will be used to construct a decomposition of circuits with Mod[q] gates. In the statement of the lemma, $|x|$ denotes the absolute value of x.

LEMMA 3.1. *Let $g_1, g_2, ..., g_s$ be Boolean functions of $x_1, x_2, ..., x_n$. For all primes q, there exists $e \in \mathbf{Z}[x_1, x_2, ..., x_n]$ such that $|e(x_1, x_2, ..., x_n)| < s^{q-1} + 1$ for all $(x_1, x_2, ..., x_n) \in \mathbf{B}^n$, and*

$$\text{Mod}[q](g_1, g_2, ..., g_s) = (g_1 + g_2 + \cdots + g_s)^{q-1} + qe.$$

Proof. Let $\mathbf{Z}_q$ denote the finite yield of q elements. Since $x^{q-1} = 1$ for any $x \neq 0 \in \mathbf{Z}_q$, we see that over $\mathbf{Z}_q$,

$$\text{Mod}[q](g_1, g_2, ..., g_s) = (g_1 + g_2 + \cdots + g_s)^{q-1}.$$

Therefore over $\mathbf{Z}$, there must exist a polynomial $e \in \mathbf{Z}[x_1, x_2, ..., x_n]$ such that

$$\text{Mod}[q](g_1, g_2, ..., g_s) = (g_1 + g_2 + \cdots + g_s)^{q-1} + qe. \tag{1}$$

It remains to show that for all $(x_1, x_2, ..., x_n) \in \mathbf{B}^n$, $|e(x_1, x_2, ..., x_n)| < s^{q-1} + 1$. Rewrite (1) as

$$qe = \text{Mod}[q](g_1, g_2, ..., g_s) - (g_1 + g_2 + \cdots + g_s)^{q-1}.$$

Taking the absolute value of both sides, we have

$$|qe| \leqslant |\text{Mod}[q](g_1, g_2, ..., g_s)| + |(g_1 + g_2 + \cdots + g_s)^{q-1}|.$$

For all $(x_1, x_2, ..., x_n) \in \mathsf{B}^n$, the first term on the right is no larger than 1 and the second term is no larger than s^{q-1} (since Mod$[q]$ and g_i are Boolean functions for $1 \leqslant i \leqslant s$). Since $q > 1$, the required result follows immediately. ∎

We are now in a position to introduce the basic idea behind our lower bound technique. We decompose a certain integer multiple of the function computed by a circuit into a sum of a lattice element and a function with small range.

Let M be a lattice, we denote by M^k, where $k \in \mathsf{N}$, the lattice generated by all elements of the form $m_1 m_2 \cdots m_k$, where $m_i \in M$, $1 \leqslant i \leqslant k$.

THEOREM 3.2. *Let C be a circuit with a* Mod$[q]$ *gate on the last level, and above it s subcircuits which compute polynomials from* Poly(2^{i-c}) *for some $c \in \mathsf{Z}$ (c may be a function of n). Let $T = \lfloor 2^{n-c(q-1)}/q \rfloor$. If C calculates a function f, then there exists $l \in$* Poly$(2^{i-c(q-1)})$ *and a polynomial $r \in \mathsf{Z}[x_1, x_2, ..., x_n]$ with $|r(x_1, x_2, ..., x_n)| = O(s^{q-1})$ for all $(x_1, x_2, ..., x_n) \in \mathsf{B}^n$, such that*

$$Tf = l + r.$$

Proof. f can be expressed as

$$f = \text{Mod}[q](g_1, g_2, ..., g_s),$$

where g_i is a subcircuit which computes a polynomial in Poly(2^{i-c}). By Lemma 3.1, there exists a polynomial $e \in \mathsf{Z}[x_1, x_2, ..., x_n]$ such that

$$f = (g_1 + g_2 + \cdots + g_s)^{q-1} + qe \tag{2}$$

and $|e(x_1, x_2, ..., x_n)| < s^{q-1} + 1$ for all $(x_1, x_2, ..., x_n) \in \mathsf{B}^n$.

Let $T = \lfloor 2^{n-c(q-1)}/q \rfloor$ and $d = 2^{n-c(q-1)} \bmod q$. Multiplying both sides of (2) by T, we have

$$Tf = T(g_1 + g_2 + \cdots + g_s)^{q-1} + 2^{n-c(q-1)}e - de.$$

Since $g_1 + g_2 + ... + g_s \in$ Poly(2^{i-c}) and $[\text{Poly}(2^{i-c})]^{q-1} = \text{Poly}(2^{i-c(q-1)})$, both $T(g_1 + g_2 + ... + g_s)^{q-1}$ and $2^{n-c(q-1)}e$ are in Poly$(2^{i-c(q-1)})$. Taking $l = T(g_1 + g_2 + ... + g_s)^{q-1} + 2^{n-c(q-1)}e$ and $r = -de$, the required result follows. ∎

Theorem 3.2 will be the crucial tool used to show lower bounds for some special depth 2 or 3 circuits of Mod$[2]$ and Mod$[q]$ gates in the next section.

4. Circuit Lower Bounds

Before we meet the circuit size lower bounds, we will make some observations about the lattice $\mathrm{Poly}(2^{i-c})$, where $c \in \mathbf{Z}$. Let $\mathscr{M}$ be the lattice generated over $\mathbf{Z}$ by functions

$$\prod_{i \in S} (1 - 2x_i),$$

where $S \subseteq \{1, 2, ..., n\}$. It is easy to see that $\mathscr{M} = \mathrm{Poly}(2^i)$ and $\mathrm{Poly}(2^{i-c}) \subseteq \mathscr{M}/2^c$. Note that the functions $\prod_{i \in S}(1 - 2x_i)$ form an orthogonal basis for $U_{\mathbf{Q}}^n$ with respect to the inner product

$$\langle f, g \rangle = \sum_{x \in \mathbf{B}^n} f(x)g(x).$$

We will use this fact later to analyze the minimum distance between the lattice $\mathrm{Poly}(2^{i-c})$ and a function outside the lattice.

$\mathrm{Poly}(2^{i-c})$ can also be characterized in terms of the discrete Fourier transform. A function can be uniquely expressed as a linear combination of $\{\prod_{i \in S}(1 - 2x_i)\}_S$, the coefficients of which are called the *spectral coefficients*. $\mathrm{Poly}(2^{i-c})$ is the lattice whose elements have spectral coefficients of the form $t/2^c$, where t is an integer.

Let f_1, f_2 be two functions in $U_{\mathbf{Z}}^n$. We define the distance $d(f_1, f_2)$ between f_1 and f_2 to be the L_∞ norm of $f_1 - f_2$, that is,

$$d(f_1, f_2) = \max_{x \in \mathbf{B}^n} |f_1(x) - f_2(x)|.$$

In Theorem 3.2, if we can show a large distance from Tf to all functions in the lattice $\mathrm{Poly}(2^{i-c(q-1)})$, then r must be large and therefore s, the number of subcircuits above the last level, must be large since r is bounded by $O(s^{q-1})$. The next lemma shows that this is indeed the case with the conjunction. Throughout the paper, S denotes a subset of the natural numbers and $\|S\|$ denotes the cardinality of S.

Lemma 4.1. *Let $f = x_1 \cdots x_n$ be the n-input conjunction, and let c be an integer which satisfies $c(q-1) \leqslant n/2$. Then for large enough n and all $l \in \mathrm{Poly}(2^{i-c(q-1)})$,*

$$d(Tf, l) \geqslant \frac{2^{n/2 - c(q-1)}}{2q},$$

where $T = \lfloor 2^{n-c(q-1)}/q \rfloor$.

Proof. Let L_2 norm of a function f is defined to be

$$\|f\|_2 = \sqrt{\sum_{x \in \mathbf{B}^n} f(x)^2}.$$

Let l be an arbitrary element in $\text{Poly}(2^{i-c(q-1)})$. Although the L_∞ norm $d(Tf, l)$ seems to be hard to analyze, we will use the L_2 norm of $Tf - l$ to bound it.

To estimate $\|Tf - l\|_2$, we write both $Tf = Tx_1x_2 \cdots x_n$, and l as linear combinations of $\{\prod_{i \in S}(1-2x_i)\}_S$. Denote the coefficient of $\prod_{i \in S}(1-2x_i)$ in the expansion of Tf and l by a_S and b_S respectively. Then a_S can be evaluated by taking the inner product of Tf and $\prod_{i \in S}(1-2x_i)$:

$$a_S = \frac{\langle Tf, \prod_{i \in S}(1-2x_i)\rangle}{\|\prod_{i \in S}(1-2x_i)\|_2^2}.$$

A simple calculation shows that

$$a_S = (-1)^{\|S\|} T/2^n = (-1)^{\|S\|} \frac{1}{q2^{c(q-1)}} \pm O(1/2^n).$$

For b_S, we note that $\text{Poly}(2^{i-c(q-1)}) \subseteq \text{Poly}(2^i)/2^{c(q-1)}$ and $\text{Poly}(2^i)$ is generated over $\mathbf{Z}$ by $\{\prod_{i \in S}(1-2x_i)\}_S$. Therefore $b_S = t/2^{c(q-1)}$ for $t \in \mathbf{Z}$. Since $\{\prod_{i \in S}(1-2x_i)\}_S$ is an orthogonal basis and $\|\prod_{i \in S}(1-2x_i)\|_2^2 = 2^n$, we now have

$$\|Tf - l\|_2^2 = \sum_{S \subseteq \{1, 2, \ldots, n\}} (a_S - b_S)^2 2^n.$$

The difference between a_S and b_S is at least $(1/q2^{c(q-1)}) \pm O(1/2^n)$, which is greater than $1/2q2^{c(q-1)}$ for large enough n. Therefore

$$\|Tf - l\|_2^2 = \sum_{S \subseteq \{1, 2, \ldots, n\}} (a_S - b_S)^2 2^n \geqslant \frac{1}{4q^2 2^{2c(q-1)}} 2^n 2^n = \frac{1}{4q^2} 2^{n - 2c(q-1)} 2^n.$$

Since the domain of a function f in U_Q^n has only 2^n members, we can bound $d(Tf, l)$ as follows:

$$d(Tf, l) \geqslant \|Tf - l\|_2 / 2^{n/2} \geqslant \frac{2^{n/2 - c(q-1)}}{2q}. \quad \blacksquare$$

We now prove our lower bound results.

THEOREM 4.2. *Let C be a circuit with a* Mod[q] *gate on the last level, and above it s subcircuits which compute polynomials from* $\text{Poly}(2^{i-c})$ *for some $c \in \mathbf{Z}$ (c may be a function of n). If C calculates the conjunction of n inputs, then the size of the circuit is $\Omega(2^{n/2(q-1)-c})$.*

Proof. Let $f = x_1 x_2 \cdots x_n$ and $T = \lfloor 2^{n - c(q-1)}/q \rfloor$. By Theorem 3.2,

$$Tf = l + r,$$

where $l \in \text{Poly}(2^{i - c(q-1)})$ and $r \in \mathbf{Z}[x_1, x_2, ..., x_n]$ with $|r(x_1, x_2, ..., x_n)| = O(s^{q-1})$ for all $(x_1, x_2, ..., x_n) \in \mathbf{B}^n$. By Lemma 4.1, $d(Tf, l) \geq 2^{n/2 - c(q-1)}/2q$, for large enough n. Thus

$$|r(x_1, x_2, ..., x_n)| = O(s^{q-1}) \geq \frac{2^{n/2 - c(q-1)}}{2q}.$$

This implies that $s = \Omega(2^{n/2(q-1) - c})$. Since the output gate is a Mod[q] gate, we may assume that no more than q subcircuits input to it are identical. Therefore, the size must be $\Omega(2^{n/2(q-1)-c})$. ∎

Applying Theorem 4.2 to (Mod[2], Mod[q])-circuits give us the result below.

COROLLARY 4.3. *Any* (Mod[2], Mod[q])*-circuit calculating the conjunction of n inputs must have size* $\Omega(2^{n/2(q-1)})$.

Proof. By Theorem 4.2, it is sufficient to show that a Mod[2] function of the input variables is in $\text{Poly}(2^{i-1})$. This is true since $M_2 \subseteq \text{Poly}(2^{i-1})$. ∎

Barrington [2] has an alternative proof of Theorem 4.2 based on a *unique representation* argument. Our proof provides a new method for the analysis of circuits of Mod[2] and Mod[q] gates. It has the advantage that it can be extended to show lower bounds on some other slightly more complicated circuits as we shall see in Corollary 4.4 and 4.5 below.

COROLLARY 4.4. *Let C be a depth* 3 *circuit with first level gates calculating* $g_1, g_2, ..., g_m$, *input to a* (Mod[2], Mod[q])*-circuit. Suppose that*

$$\sum_{1 \leq i \leq m,\ \deg(g_i) \geq 1} (\deg(g_i) - 1) \leq \frac{n}{2(q-1)} - 1 - d,$$

where $\deg(g_i)$ *denotes the degree of* g_i *as a polynomial. If C calculates the conjunction of n inputs, then C must have size* $\Omega(2^d)$.

Proof. Let $h_1, h_2, ..., h_s$ be the Mod[2] functions on the second level. By Theorem 4.2, it is sufficient to show that

$$h_i = \text{Mod}[2](g_{i_1}, g_{i_2}, ..., g_{i_{t_i}}) \in \text{Poly}(2^{i - n/2(q-1) + d}).$$

Each term in h_i has the form $(-1)^{\|S\|-1} 2^{\|S\|-1} \prod_{i_j \in S} g_{i_j}$. The degree of $\prod_{i_j \in S} g_{i_j}$ is at most $\|S\| + n/2(q-1) - 1 - d$ because

$$\begin{aligned}\deg\left(\prod_{i_j \in S} g_{i_j}\right) &\leqslant \sum_{i_j \in S,\, \deg(g_{i_j}) \geqslant 1} \deg(g_{i_j}) \\ &= \sum_{i_j \in S,\, \deg(g_{i_j}) \geqslant 1} (\deg(g_{i_j}) - 1) + \sum_{i_j \in S,\, \deg(g_{i_j}) \geqslant 1} 1 \\ &\leqslant \frac{n}{2(q-1)} - 1 - d + \|S\|.\end{aligned}$$

This implies that $h_i \in \text{Poly}(2^{i - n/2(q-1)+d})$. ∎

Note that this is in a sense slightly more general than the constant degree hypothesis, since the polynomial degrees of the functions may be non-constant, provided the sum of the degrees is a small linear function of n. We can also prove the following variant.

COROLLARY 4.5. *Let C be a circuit with the first layer* Mod[2] *gates, the second layer s gates which compute $g_1, g_2, ..., g_s$, and the third layer a* Mod[q] *gate, where g_i is a polynomial of degree no larger than $n/2(q-1) - d$. Suppose C calculates the conjunction of n inputs. Then the size of the circuit is $\Omega(2^d)$.*

Proof. It is sufficient to show that $g_i(h_{i_1}, h_{i_2}, ..., h_{i_{i_i}}) \in \text{Poly}(2^{i - n/2(q-1)+d})$, where h's are the Mod[2] gates on the first level. Since $\deg(g_i) \leqslant n/2(q-1) - d$ and each Mod[2] function is in $\text{Poly}(2^{i-1})$, $g_i(h_{i_1}, h_{i_2}, ..., h_{i_{i_i}})$ must be in $[\text{Poly}(2^{i-1})]^{n/2(q-1)-d} = \text{Poly}(2^{i - n/2(q-1)+d})$. ∎

5. THE GENERAL CASE

In this section we generalize the algebraic characterization of the last section to (Mod[p], Mod[q])-circuits, and examine a possible method of using a similar characterization to prove the constant degree hypothesis. Barrington *et al.* conjectured in [3] that circuits of the form (constant degree polynomial, Mod[p], Mod[q]) calculating conjunction must have exponential size, and they call this conjecture the *constant degree hypothesis*. In both the (Mod[p], Mod[q]) and (constant degee polynomial, Mod[p], Mod[q]) cases, we are able to reduce the circuit complexity question to an algebraic one. These questions remain open for conjunction.

We first study the analogue of M_2 for the Mod[p] function. In the case of the Mod[2] function, M_2 can be characterized as $\text{Poly}(2^{i-1})$. The next

lemma gives us a similar characterization of the lattice generated by all Mod[p] functions of input variables.

LEMMA 5.1. $\text{Mod}[p](x_1, x_2, ..., x_n) \in \text{Poly}(p^{\lfloor (i-1)/(p-1) \rfloor})$.

Proof. Let ξ_p be the pth root of unity and let $Z[\xi_p]$ be the ring generated by ξ_p over Z. Let ψ_k denote the function

$$1 - \text{Mod}[p](x_1, x_2, ..., x_n, \underbrace{0, ..., 0}_{k}, \underbrace{1, ..., 1}_{p-k}).$$

Define function ψ as

$$\psi = \sum_{k=0}^{p-1} \xi_p^k \psi_k = \prod_{i=1}^{n} (1 - (1 - \xi_p) x_i).$$

The equality between the sum and the product in the above is true since they both have value ξ^k for an input satisfying $x_1 + \cdots + x_n = k \pmod p$. The roots of unity $1, \xi_p, \xi_p^2, ..., \xi_p^{p-1}$ satisfy the equation $1 + \xi_p + \xi_p^2 + \cdots + \xi_p^{p-1} = 0$. Therefore, ψ can be rewritten as

$$\psi = \sum_{k=0}^{p-2} \xi_p^k (\psi_k - \psi_{p-1}) = \prod_{i=1}^{n} (1 - (1 - \xi_p) x_i).$$

The coefficient of a monomial of degree i in ψ is divisible by $(1 - \xi_p)^i$, and therefore divisible by $p^{\lfloor i/(p-1) \rfloor}$ since p factors as $\eta(1 - \xi_p)^{p-1}$ in $Z[\xi_p]$, where η is an invertible element in $Z[\xi_p]$(see, for example, Weiss [13, p. 262]). Since $1, \xi_p, \xi_p^2, ..., \xi_p^{p-2}$ are linearly independent, this implies that a monomial of degree i in $\psi_k - \psi_{p-1}$ is divisible by $p^{\lfloor i/(p-1) \rfloor}$ for all $0 \leqslant k \leqslant p-2$. Write $\text{Mod}[p](x_1, x_2, ..., x_n, x_{n+1})$ as $\sum_{S \subseteq \{1, 2, ..., n, n+1\}} \alpha_S X_S$, where $X_S = \prod_{i \in S} x_i$. We show that α_S is divisible by $p^{\lfloor (\|S\| - 1)/(p-1) \rfloor}$. By setting $x_{n+1} = 0$ and $x_{n+1} = 1$ we can write ψ_0 and ψ_{p-1} in terms of $1 - \text{Mod}[p](x_1, x_2, ..., x_n, x_{n+1})$:

$$\psi_0 - \psi_{p-1} = \sum_{n+1 \in S} \alpha_S X_{S - \{n+1\}}.$$

For all S containing x_{n+1}, α_S is the coefficient of a monomial of degree $\|S\| - 1$ in $\psi_0 - \psi_{p-1}$, so it is divisible by $p^{\lfloor (\|S\| - 1)/(p-1) \rfloor}$. Since the Mod[$p$] function is symmetric, the restriction of containing x_{n+1} merely says that $\|S\| \geqslant 1$. ∎

Since all Mod[p] functions of the input variables can be obtained by setting some of the variables equal to other variables or to 0 or 1 in the gernaral expression $\text{Mod}[p](x_1, x_2, ..., x_n)$, they all have the property

stated in Lemma 5.1 and therefore so does every element in M_p. This characterization of M_p is similar to that of M_2, the difference being that in M_p, the coefficient of a monomial of degree i is an integer multiple of $p^{\lfloor (i-1)/(p-1) \rfloor}$ instead of 2^{i-1}. Next, we generalize this characterization to the case of (constant degree polynomial, Mod[p], Mod[q])-circuits.

LEMMA 5.2. *Let M' be the module generated by all functions of the form*

$$\text{Mod}[p](g_1, g_2, ..., g_m),$$

where g_i are polynomials of degree $\leqslant c$. Let $M = (M')^{q-1}$. Then

$$M \subseteq \text{Poly}(p^{\lfloor i/c(p-1) \rfloor - 2(q-1)}).$$

Proof. Since $\text{Mod}[p](x_1, x_2, ..., x_n)$ is in $\text{Poly}(p^{\lfloor (i-1)/(p-1) \rfloor})$, it is obvious that $M' \subseteq \text{Poly}(p^{\lfloor (i/(c-1))/(p-1) \rfloor})$. Raising $\text{Poly}(p^{\lfloor (i/(c-1))/(p-1) \rfloor})$ to the $(q-1)$th power, we get

$$\begin{aligned}[\text{Poly}(p^{\lfloor (i/(c-1))/(p-1) \rfloor})]^{q-1} &\subseteq \text{Poly}(p^{\lfloor i/c(p-1) - (q-1)/(p-1) \rfloor - (q-1)}) \\ &\subseteq \text{Poly}(p^{\lfloor i/c(p-1) \rfloor - 2(q-1)}),\end{aligned}$$

which is the desired result. ∎

Lemma 5.2 gives us the following result about the constant degree hypothesis:

THEOREM 5.3. *Let f be a Boolean function. If*

$$d(Tf, l) = 2^{\Omega(n)}$$

for all $l \in \text{Poly}(p^{\lfloor i/c(p-1) \rfloor - 2(q-1)})$, where $T = \lfloor p^{\lfloor n/c(p-1) \rfloor - 2(q-1)}/q \rfloor$, then the constant degree hypothesis is true for f.

We conjecture that for $f = x_1 x_2 \cdots x_n$, $d(Tf, l) = 2^{\Omega(n)}$. Although we cannot prove this conjecture, we can prove that $d(Tf, l) = 2^{\Omega(n)}$ is true for almost every Boolean function f. For simplicity, the following theorem is stated only for $p = 2$, $c = 2$.

THEOREM 5.4. *Let q be an arbitrary prime different from 2, $M = \text{Poly}(2^{\lfloor i/2 \rfloor - 2(q-1)})$ and $T = \lfloor 2^{\lfloor n/2 \rfloor - 2(q-1)}/q \rfloor$. If f is a random Boolean function then*

$$\lim_{n \to \infty} P\{d(Tf, l) \geqslant \frac{1}{3q} 2^{\lfloor n/8 \rfloor}, \forall l \in M\} = 1.$$

Proof. The natural map from $\mathbf{Z}$ to $\mathbf{Z}' = \mathbf{Z}_{2^{\lfloor n/8 \rfloor}}$ extends to a map from $U_{\mathbf{Z}}^n$ to $U_{\mathbf{Z}'}^n$. The image of M under this map is denoted by M'. It is easy to see that M' is generated over $\mathbf{Z}'$, as a $\mathbf{Z}'$-module, by all monomials of degree i with $\lfloor i/2 \rfloor - 2(q-1) < \lfloor n/8 \rfloor$. The number of such monomials is bounded by a^n for large enough n and some positive constant $a < 2$, hence the number of functions in M' is bounded by $(2^{\lfloor n/8 \rfloor})^{a^n}$.

Let f_1, f_2 be the two different Boolean functions. If there are $l_1, l_2 \in M$ such that $d(Tf_1, l_1) < (1/3q)\, 2^{\lfloor n/8 \rfloor}$ and $d(Tf_2, l_2) < (1/3q)\, 2^{\lfloor n/8 \rfloor}$, then $l_1' \neq l_2'$, where l_1', l_2' are the images of l_1, l_2 under the map from $U_{\mathbf{Z}}^n$ to $U_{\mathbf{Z}'}^n$. This is true because if not, then

$$(Tf_1 - l_1) - (Tf_2 - l_2) = 2^{\lfloor n/8 \rfloor} g + (Tf_1 - Tf_2)$$

for some nonzero polynomial g over $\mathbf{Z}$. Taking the absolute value and using the assumption that $d(Tf_1, l_1) < (1/3q)\, 2^{\lfloor n/8 \rfloor}$, $d(Tf_2, l_2) < (1/3q)\, 2^{\lfloor n/8 \rfloor}$, we would have

$$\left| \frac{2}{3q} 2^{\lfloor n/8 \rfloor} \right| \geqslant |2^{\lfloor n/8 \rfloor} g + T(f_1 - f_2)|.$$

This is a contradiction since $T = \lfloor 2^{\lfloor n/2 \rfloor - 2(q-1)}/q \rfloor = 2^{\lfloor n/2 \rfloor - 2(q-1)}/q \pm O(1)$ and it can only be approximated by a number of the form $t2^{\lfloor n/8 \rfloor}$ for some integer t with error no less than $(1/q)\, 2^{\lfloor 2/8 \rfloor} \pm O(1)$.

The above argument implies that the number of Boolean functions f for which there is an $l \in M$ such that $d(Tf, l) < (1/3q)\, 2^{\lfloor n/8 \rfloor}$ is no more than the number of functions in M'.
Therefore

$$\lim_{n \to \infty} P\{d(Tf, l) \geqslant \frac{1}{3q} 2^{\lfloor n/8 \rfloor}, \forall l \in M\} \geqslant 1 - \lim_{n \to \infty} \frac{(2^{\lfloor n/8 \rfloor})^{a^n}}{2^{2^n}} = 1.$$

RECEIVED November 16, 1990; FINAL MANUSCRIPT RECEIVED April 8, 1992

REFERENCES

1. AJTAI, M. (1983) Σ_1^1-formulae on finite structures, *Ann. Pure Appl. Logic* **24**, 1–48.
2. BARRINGTON, D. (1983) "With 3 Permutation Branching Programs," Technical report TM-291, MIT Laboratory for Computer Science.
3. BARRINGTON, D. A. M., STRAUBING, H., AND THERIEN, D. (1988) "Non-uniform automata over groups," Technical Report 88-77, Department of Computer and Information Science, University of Massachusetts at Amherst.
4. FURST, M., SAXE, J. B., AND SIPSER, M. Parity, circuits and the polynomial time hierarchy, *Math. Systems Theory* **17**(1), 13–27.
5. HAJNAL, A., MAASS, W., PUDLAK, P., SZEGEDY, M. AND TURAN, G. (1987), Threshold circuits of bounded depth, *in* "Proceedings, 28th Annual IEEE Symposium on Foundations of Computer Science," pp. 99–110.

6. HASTAD, J. (1986) Improved lower bounds for small depth circuits, *in* "Proceedings 18th Annual ACM Symposium on Theory of Computing," pp. 6–20.
7. HASTAD, J., AND GOLDMANN, M. On the power of small-depth threshold circuits, *in* "Proceedings 31st Annual IEEE Symposium on Foundations of Computer Science," pp. 610–618.
8. PARBERRY, I., AND SCHNITGER, G. (1988), Parallel computation with threshold functions, *J. Comput. System Sci.* **36**(3), 278–302.
9. PARBERRY, I., AND SCHNITGER, G. (1989), Relating Boltzmann machines to conventional models of computation, *Neural Networks* **2**(1), 59–67.
10. RAZBOROV, A. A. (1987), Lower bounds for the size of circuits of bounded depth with basis $\{\wedge, \oplus\}$, *Mat. Zametki*, **41**(4), 598–607; English translation, *Math. Notes* **41**(4), 333–338.
11. SIPSER, M. (1983), Borel sets and circuit complexity, *in* "Proceedings, 15th Annual ACM Symposium on Theory of Computing", pp. 61–69.
12. SMOLENSKY, R. (1987), Algebraic methods in the theory of lower bounds for Boolean circuit complexity, "Proceedings, 19th Annual ACM Symposium on Theory of Computing," pp. 77–82.
13. WEISS, E. "Algebraic Number Theory," 2nd ed., Chelsea, New York.
14. YAN, P. Y. "Lower Bound Techniques in Some Parallel Models of Computation," Ph.D. Thesis, Department of Mathematics, Penn State University.
15. YAO, A. C. (1989), Circuits and local computation, *in* "Proceedings, 21st Annual ACM Symposium on Theory of Computing," pp. 186–196.
16. YAO, A. C. (1985), Separating the polynomial-time hierarchy by oracles. *in* "Proceedings, 26th Annual IEEE Symposium on Foundations of Computer Science," pp. 1–10.