On ACC and Threshold Circuits*

Andrew Chi-Chih Yao
Department of Computer Science
Princeton University
Princeton, New Jersey 08544

Abstract

We prove that any language in ACC can be approximately computed by two-level circuits of size $2^{(\log n)^k}$, with a symmetric-function gate at the top and only AND gates on the first level. This implies that any language in ACC can be recognized by depth-3 threshold circuits of size $2^{(\log n)^k}$. This result gives the first nontrivial upper bound on the computing power of ACC circuits.

1 Introduction

The class AC^0, which consists of languages accepted by bounded-depth polynomial-size circuits with {AND, OR, NOT } gates, has been studied extensively. By now, several characterizations and limitations have been found for languages in this class: it is known that AC^0 does not include the PARITY or MAJORITY functions ([Aj][FSS][Ya][Ha]), that every language in AC^0 can be approximated by a polynomial of low degree over any finite field ([Ra][Sm]) and over the reals ([Br][LMN][BS]), and that any language in AC^0 can be recognized by depth-3 threshold circuits of size $2^{(\log n)^k}$ ([All]).

In contrast, ACC, which is the class of languages accepted by bounded-depth polynomial-size circuits with AND, OR, NOT, plus any finite set of MOD_p gates (first considered in [Ba]), is much less understood. The only limitation result known is that, if the modular gates are restricted to a single prime p, then one cannot compute functions such as MAJORITY or modular functions of other primes. To underline the state of ignorance on this subject, note that it is consistent with our knowledge at the moment that the Hamiltonian Circuit problem might be computable with depth-3 polynomial-size circuits with alternating MOD_2 and MOD_3 gates.

In this paper, we will prove several results on the characterization of ACC, analogous to those results for AC^0 as mentioned above. We will show that any language in ACC can be approximated by a polynomial of low degree over some suitable domain, and that any language in ACC can be recognized by depth-3 threshold circuits of size $2^{(\log n)^k}$.

At the present time, an important direction in boolean circuit theory is to prove lower bounds for non-monotone circuits. However, as the eventual goal of proving superpolynomial bounds for general circuits remains elusive, a systematic study of the more restricted models, such as ACC, polynomial threshold circuits, NC^1, etc., is an attractive way to proceed. The present work is encouraging in that it shows progress can be made on the ACC problem, which has heretofore been a serious barrier in this line of investigation.

Remarks All logarithms in this paper are taken with base 2. The MOD_m gates are defined by $MOD_m(y_1, y_2, \cdots, y_\ell) = 1$ if $\sum_{1 \le i \le \ell} y_i \bmod m = 0$ and 0 otherwise.

*This research was supported in part by the National Science Foundation under Grant CCR-8813283.

2 Results

A language $L \subseteq \{0,1\}^*$ is said to be computed by a sequence of Boolean circuits $\{C_n \mid n\}$ if, given n Boolean variables $x_1, x_2, \cdots, x_n$ as inputs, C_n will output 1 if and only if $(x_1, x_2, \cdots, x_n) \in L^{\{n\}}$, where $L^{\{n\}} = L \cap \{0,1\}^n$.

Let S be a finite set of positive integers. A *modular circuit* (of *type* S) is a boolean circuit with gates AND, OR, NOT, and MOD_m ($m \in S$) of unbounded fan-in. As usual, we will assume that the NOT gates appear only in the form as negated input variables. We will say the modular circuit is *pure*, if it does not contain AND or OR gates. A language $L \subseteq \{0,1\}^*$ is said to be in *ACC* if there exist an integer $k \geq 1$, a finite set S, and a sequence of depth-k modular circuits $\{C_n \mid n\}$ of type S computing L with $\mathrm{size}(C_n) \leq q(n)$ for some polynomial q; L is in *pure ACC*, if the C_n's are pure modular circuits.

A modular circuit is said to be *canonical*, if all the gates at any given level are of the same kind and all MOD_m gates have prime m. The *signature* of a depth-k canonical modular circuit C is a k-tuple $\sigma(C) = (\sigma_1, \sigma_2, \cdots, \sigma_k)$, where $\sigma_i \in \{\mathrm{AND},\ \mathrm{OR}\} \cup \{\mathrm{MOD}_p \mid p : \text{primes}\}$ is the gate used on level i.

For any integers y and $a_1, a_2, \cdots, a_k$ where $a_i > 1$, let $\mathrm{mod}(y; a_1, a_2, \cdots, a_k)$ denote the expression:

$$\mathrm{mod}_{a_k}(\cdots(\mathrm{mod}_{a_2}(\mathrm{mod}_{a_1}(y)))\cdots).$$

For any language L, let χ_L denote its characteristic function, i.e., $\chi_L(\tilde{x}) = 1$ if $\tilde{x} \in L$ and 0 otherwise.

Let $\mathbf{Z}^{[n]}$ denote the algebra $\mathbf{Z}[x_1, x_2, \cdots, x_n]/\{x_i^2 - x_i \mid i\}$, i.e., the set of polynomials with integer coeffients over n indeterminates, in which the identities $x_i^2 = x_i$ are valid. For any $F \in \mathbf{Z}^{[n]}$, *degree*(F) denotes the degree of F, and *size*(F) is defined as $\log w$, where w is the sum of the absolute values of the coefficients in F. Let $\|F\|$ denote the maximum of degree(F) and size(F).

Theorem 1 Let $k, \beta \geq 1$, $\lambda \geq 7$ be fixed integers, and $p_1, p_2, \cdots, p_k$ be any fixed sequence of primes. If L can be computed by a sequence of depth-k canonical modular circuits of size less than n^β and with signature $(\mathrm{MOD}_{p_1}, \mathrm{MOD}_{p_2}, \cdots, \mathrm{MOD}_{p_k})$, then there exist (a) polynomials $F_n \in \mathbf{Z}^{[n]}$ with $\|F_n\| = O((\log n)^{\lambda^k})$, and (b) positive integers $m_{n,i}$, $1 \leq i \leq k$ with $\log m_{n,i} = \Theta((\log n)^{\lambda^{k-i}})$ for large n such that $\chi_{L^{\{n\}}}(\tilde{x}) = \mathrm{mod}(F_n(\tilde{x}); p_1^{m_{n,1}}, p_2^{m_{n,2}}, \cdots, p_k^{m_{n,k}})$.

Corollary If L is in pure ACC, then there exist a constant $c > 0$, polynomials $F_n \in \mathbf{Z}^{[n]}$ with $\|F_n\| \leq (\log n)^c$, and functions $\Psi_n : \{0, 1, \cdots, \lceil 4^{(\log n)^{c+1}} \rceil\} \to \{0,1\}$ such that $\chi_{L^{\{n\}}} = \Psi_n(F_n(x_1, x_2, \cdots, x_n))$ for all n and $x_i \in \{0,1\}$.

Theorem 2 Let $\gamma, k, r > 0$, $\lambda \geq 7$ be fixed integers and $p_1, p_2, \cdots, p_r$ be any fixed sequence of r primes. If L can be computed by a sequence of depth-k canonical modular circuits with $(\mathrm{MOD}_{p_1}, \mathrm{MOD}_{p_2}, \cdots, \mathrm{MOD}_{p_r})$ being the portion of the signature not involving AND-OR gates, then there exist (a) a probability distribution ρ_n over $\mathbf{Z}^{[n]} \cap \{F \mid \text{polynomials } F \text{ with } \|F\| = O((\log n)^{\lambda^k})\}$, and (b) positive integers $m_{n,i}$, $1 \leq i \leq k$ such that the following is true: for any $\tilde{x} \in \{0,1\}^n$, if we take a random F distributed according to ρ_n, then $\chi_{L^{\{n\}}}(\tilde{x}) = \mathrm{mod}(F(\tilde{x}); p_1^{m_{n,1}}, p_2^{m_{n,2}}, \cdots, p_k^{m_{n,k}})$ with probability at least $1 - n^{-\gamma}$.

Corollary If L is in ACC, then there exist a constant $c > 0$, a probability distribution ρ_n over $\mathbf{Z}^{[n]} \cap \{F \mid \text{polynomials } F \text{ with } \|F\| \leq (\log n)^c\}$, and functions $\Psi_n : \{0, 1, \cdots, \lceil 4^{(\log n)^{c+1}} \rceil\} \to \{0,1\}$ such that the following is true: for every $\tilde{x} \in \{0,1\}^n$, a random F distributed according to ρ_n satisfies the equality $\chi_{L^{\{n\}}}(\tilde{x}) = \Psi_n(F(x_1, x_2, \cdots, x_n))$ with probability $1 - o(1)$ for large n.

Theorem 3 If L is in pure ACC, then there exists a sequence of depth-3 threshold circuits $\{C_n \mid n\}$ computing L with $\mathrm{size}(C_n) \leq 2^{(\log n)^c}$ for some fixed c; furthermore, the bottom gates are AND gates with fan-in not exceeding $(\log n)^c$, and the top gate is an OR gate.

Theorem 4 If L is in ACC, then there exists a sequence of depth-3 threshold circuits $\{C_n \mid n\}$ computing L with $\mathrm{size}(C_n) \leq 2^{(\log n)^c}$ for some fixed c; furthermore, the bottom gates are AND gates with

fan-in not exceeding $(\log n)^c$.

Remarks Theorem 2 can be regarded as an extension of the Razborov-Smolensky result ([Ra][Sm]) that every language in AC^0 can be approximated by low-degree polynomials over any finite field; it also can be viewed as an extension of the approximation results of AC^0 by low-degree polynomials over the reals (see Bruck [Br], Linial, Mansour and Nisan [LMN], and Bruck and Smolensky [BS]). Theorem 4 extends a result by Allender [All] for languages in AC^0.

The rest of this extended abstract is devoted to a detailed proof of Theorem 1 and its Corollary. We first discuss some useful facts in Section 3. In Section 4, we introduce the concept of *supermodular representation* of numbers and functions which is essential to our proof. Making use of the machinery developed, we prove Theorem 1 in Section 5. !! The proof of Theorem 2 and its Corollary is similar to that of Theorem 1, except that one needs to replace all the AND and OR gates probabilistically by gates computing low-degree polynomials using the Razborov-Smolensky [Ra][Sm] or the Valiant-Vazirani [VV] methods (see Allender and Hertrampf [AH] for futher discussions). To prove Theorems 3 and 4, we first rewrite F_n in Theorem 1 and Theorem 2 as a polynomial in x_i and $\bar{x}_i$ with all coefficients nonnegative; the proof then proceeds in essentially the same way as the corresponding result for AC^0 languages (see Allender [All]). Details will be given in the complete paper.

3 Preliminaries

Let $h(x) = 3x^2 - 2x^3$. Define a sequence of polynomials $h^{(0)}(x) = x$, and $h^{(t)}(x) = h(h^{(t-1)}(x))$ for all integers $t > 0$.

Lemma 1 Let $t \geq 0$, $m > 1$, and N be integers. If $N \bmod m \in \{0, 1\}$, then $h^{(t)}(N) \bmod (m^{2^t}) = N \bmod m$.

Proof We proceed by induction on t. The case $t = 0$ is trivially true. Let $t > 0$ and assume that the lemma is true for all smaller values of t. Write $N = N_0 m + j$, where N_0 is an integer and $j = N \bmod m$. By the induction hypothesis, we have $h^{(t-1)}(x) = q + j$ where $q = N_1 m^{2^{t-1}}$ for some integer N_1. Thus, $h^{(t)}(N) = 3(j+q)^2 - 2(j+q)^3$, which is equal to $(3j^2 - 2j^3) - q^2(3j + 2q)$. As $3j^2 - 2j^3 = j$ for $j \in \{0, 1\}$ and m^{2^t} divides q^2, we have $h^{(t)}(N) \bmod (m^{2^t}) = N \bmod m$. This completes the proof. □

Lemma 2 For $t \geq 0$, $h^{(t)}$ has degree 3^t and size not exceeding $\frac{1}{2}(3^t - 1)\log 5$.

Proof The degree of $h^{(t)}$ is clearly 3^t, since the degree of h is 3. To bound the size of $h^{(t)}$, let $u^{(1)}(x) = 3x^2 + 2x^3$, and $u^{(t)}(x) = u^{(1)}(u^{(t-1)}(x))$ for all integers $t > 1$. Clearly, the size of $h^{(t)}$ cannot be greater than the size of $u^{(t)}$. Since all the coefficients of $u^{(t)}$ are nonnegative, the size of $u^{(t)}$ is equal to $u^{(t)}(1)$. Now $u^{(t)}(1) \leq 5(u^{(t-1)}(1))^3$. A simple inductive argument shows that $\log u^{(t)}(1) \leq ((3^t - 1)\log 5)/2$. □

Remarks A similar function was used in Toda [To]; let $f(x) = 3x^4 + 4x^3$ and $f^{(t)}(x)$ be defined as $f(f^{(t-1)}(x))$ with $f^{(0)}(x) = x$. It was observed that $N \bmod m \in \{0, -1\}$ implies $f^{(t)}(N) \bmod (m^{2^t}) = N \bmod m$, a fact used by Toda to collapse mod 2 operations in the context of Turing machine simulations. Our application of Lemma 1 and 2 in the next section extends that usage, and h is defined in a way convenient for that purpose.

The next lemma is a simple fact concerning the composition of polynomials.

Lemma 3 Let ℓ, n, d, s, d', s' be positive integers, $A \in \mathbf{Z}^{[\ell]}$, and $g_i \in \mathbf{Z}^{[n]}$ for $1 \leq i \leq \ell$. If degree$(A) \leq d$, size$(A) \leq s$, degree$(g_i) \leq d'$, and size$(g_i) \leq s'$ for $1 \leq i \leq k$, then the polynomial $A(g_1(\tilde{x}), g_2(\tilde{x}), \cdots, g_\ell(\tilde{x}))$ has degree at most dd' and size at most $s + ds'$.

Proof Let $A_g(\tilde{x})$ denote the polynomial $A(g_1(\tilde{x}), g_2(\tilde{x}), \cdots, g_\ell(\tilde{x}))$. Clearly, A_g has degree at most dd'. We now prove that the size of A_g is at most $s + ds'$. Without loss of generality, we can assume that A and g_i have only nonnegative coefficients. The size of A_g is then equal to $\log A_g(1, 1, \cdots, 1)$, which is $\log A(g_1(1, 1, \cdots, 1), g_2(1, 1, \cdots, 1), \cdots, g_\ell(1, 1, \cdots, 1))$,

and thus at most $\log A(2^{s'}, 2^{s'}, \cdots, 2^{s'})$. Write $A(y_1, y_2, \cdots, y_\ell) = \sum_{T \subseteq \{1,2,\cdots,\ell\}} \mu_T(\prod_{i \in T} y_i)$. Then

$$\begin{aligned} A(2^{s'}, 2^{s'}, \cdots, 2^{s'}) &= \sum_{T \subseteq \{1,2,\cdots,\ell\}} \mu_T(\prod_{i \in T} 2^{s'}) \\ &\leq (2^{s'})^d \sum_{T \subseteq \{1,2,\cdots,\ell\}} \mu_T \\ &\leq 2^{s'd} 2^s. \end{aligned}$$

It follows that the size of A_g is at most $s + ds'$. □

4 Supermodular Representations

Let $\mathcal{R}_p^n$ denote the algebra $\mathbf{Z}_p[x_1, x_2, \cdots, x_n]/\mathcal{J}$, where p is a prime and $\mathcal{J} = \{x_i^2 - x_i \mid i\}$. In [Ra][Sm], a way of representing Boolean functions by elements of $\mathcal{R}_p^n$ was defined as a tool for exploring AC^0. A major obstacle in trying to use this algebraic approach to study ACC is the following. Suppose $f_i \in \mathcal{R}_p^n$ is represented by its coefficients, then in general, the sum $\sum_i f_i$ does not have an easily constructible representation over $\mathcal{R}_{p'}^n$ for a prime p' distinct from p. A crucial step in our proof is to develop a representation of Boolean functions under which interwoven operations can be carried out efficiently within the context of ACC computations.

Let us view the Razborov-Smolensky method of representing Boolean functions in $\mathcal{R}_p^n$ as follows: a pair (p, f), where p is prime and $f \in \mathbf{Z}^{[n]}$ is a representative for a Boolean function q if $q = f \bmod p$. We will consider objects a more general than the primes p, and consider pairs of the form (a, f) as possible representatives of q. Thus, in some sense, we are constructing a representation which includes and extends $\cup_p \mathcal{R}_p^n$. The goal is to have a representation flexible enough to carry out arithmetic operations, especially the MOD_p operations (for all p), in an efficient manner. The most important result to be derived in this section is Lemma 8 (Section 4.2).

Before introducing the desired representation for Boolean functions, we first discuss the corresponding representation for numbers.

4.1 Numbers

Let $k > 0$ and $a = (a_1, a_2, \cdots, a_k; b_1, b_2, \cdots, b_k)$. We call a a *base* if a_i and b_j are positive integers and $a_k = 1$. If an integer N can be written as $\sum_{1 \leq i \leq t} a_i z_i$, where z_i are integers satisfying $|z_i| \leq b_i/a_i$ for $1 \leq i \leq k$, we write $z_k \in M(a, N)$.

A base a is called an *admissible base*, if the diophantine equation $\sum_{1 \leq i \leq t} a_i z_i = 0$ has no integer solution $(z_1, z_2, \cdots, z_t)$ with $|z_i| \leq 2b_i/a_i$ other than the trivial all-zero solution. Clearly, for any integer N, either there exists *no* $(t+1)$-tuple of integers $z = (z_1, z_2, \cdots, z_t)$ satisfying $|z_i| \leq b_i/a_i$ and $N = \sum_{1 \leq i \leq t} a_i z_i$, or there exists a *unique* such z. In the latter case, we write $mod_a(N) = z_k$; otherwise $\text{mod}_a(N)$ is defined to be ∞. We regard (a, N) as a *representative* of integer z_k. Call this the *supermodular representation* of integers.

Lemma 4 Let $a = (a_1, a_2, \cdots, a_k, b_1, b_2, \cdots, b_k)$ be a base. If $a_i > 4b_{i+1}$ for $1 \leq i < k$, then a is an admissible base.

Proof Immediate. □

In the next three lemmas, we will show that the supermodular representation allows us to perform arithmetic operations easily on elements from $\{0, 1\}$ by manipulating the representatives of these elements. (One can extend the results to domains larger than $\{0, 1\}$, but we will not discuss the extensions here.)

Let k, d, s be positive integers, and $a = (a_1, \cdots, a_k; b_1, \cdots, b_k)$, $a' = (a_1, \cdots, a_k; b'_1, \cdots, b'_k)$ be bases. We say that a, a' satisfy *Property* $P_1(k, d, s)$ if the following conditions are true: $b_i > 2b_{i+1}$ for $1 \leq i < k$, and $b'_i \geq 2^s(2b_i)^d$ for $1 \leq i \leq k$.

Let $\ell > 0$, $N_1, N_2, \cdots, N_\ell$ be integers and $j_1, j_2, \cdots, j_\ell \in \{0, 1\}$. Let $u \in \mathbf{Z}^{[\ell]}$ be a polynomial with degree and size not greater than d and s, respectively.

Lemma 5 Let a, a' be bases satisfying $P_1(k, d, s)$. If $j_r \in M(a, N_r)$ for $1 \leq r \leq \ell$, then $u(j_1, j_2, \cdots, j_\ell) \in M(a', u(N_1, N_2, \cdots, N_\ell))$.

Corollary Let a, a' be admissible bases satisfying

$P_1(k,d,s)$. If $j_r = \mathrm{mod}_a(N_r)$ for $1 \le r \le \ell$, then $u(j_1, j_2, \cdots, j_\ell) = \mathrm{mod}_{a'}(u(N_1, N_2, \cdots, N_\ell))$.

Proof For each $1 \le r \le \ell$, write

$$N_r = q_{r,1} + q_{r,2} + \cdots + q_{r,k}, \tag{1}$$

where, for $1 \le i \le k$,

$$a_i \mid q_{r,i}, \tag{2}$$

$$|q_{r,i}| \le b_i, \tag{3}$$

and

$$q_{r,k} = j_r \tag{4}$$

Then, for each $T \subseteq \{1, 2, \cdots, \ell\}$ with $0 < |T| \le d$, we have

$$\prod_{r \in T} N_r = \sum_{1 \le i \le k} q'_{T,i}, \tag{5}$$

where

$$q'_{T,i} = \prod_{r \in T} \Big(\sum_{i \le \alpha \le k} q_{r,\alpha} \Big) - \prod_{r \in T} \Big(\sum_{i < \alpha \le k} q_{r,\alpha} \Big). \tag{6}$$

Using (2), (3), (6) and the fact $b_\alpha > 2b_{\alpha+1}$, we have

$$a_i \mid q'_{T,i}, \tag{7}$$

and

$$\begin{aligned} |q'_{T,i}| &\le \Big(\sum_{i \le \alpha \le k} b_\alpha \Big)^{|T|} \\ &\le (2b_i)^d. \end{aligned} \tag{8}$$

Now, write

$$u(y_1, y_2, \cdots, y_\ell) = \sum_{T \subseteq \{1,2,\cdots,\ell\}} \mu_T \Big(\prod_{r \in T} y_r \Big).$$

Then we have from (5)

$$\begin{aligned} u(N_1, N_2, \cdots, N_\ell) &= \mu_\emptyset + \sum_{\substack{T \subseteq \{1,2,\cdots,\ell\} \\ |T| > 0}} \mu_T \Big(\prod_{r \in T} N_r \Big) \\ &= \sum_{1 \le i \le k} Q'_i, \end{aligned} \tag{9}$$

where

$$Q'_i = \sum_{\substack{T \subseteq \{1,2,\cdots,\ell\} \\ 0 < |T| \le d}} \mu_T q'_{T,i} + \delta_{i,k} \mu_\emptyset. \tag{10}$$

Let $1 \le i \le k$. It follows from (7) and (10) that

$$a_i \mid Q'_i, \tag{11}$$

and from (8) and (10) that

$$\begin{aligned} |Q'_i| &\le (2b_i)^d \sum_{\substack{T \subseteq \{1,2,\cdots,\ell\} \\ 0 < |T| \le d}} |\mu_T| + |\mu_\emptyset| \\ &\le (2b_i)^d 2^s \\ &\le b'_i. \end{aligned} \tag{12}$$

Also, from (4), (6) and (10), we have

$$\begin{aligned} Q'_k &= \mu_\emptyset + \sum_{\substack{T \subseteq \{1,2,\cdots,\ell\} \\ |T| > 0}} \mu_T \Big(\prod_{r \in T} j_r \Big) \\ &= u(j_1, j_2, \cdots, j_\ell). \end{aligned} \tag{13}$$

Lemma 5 follows from (9), (11), (12) and (13). □

Let k, t be positive integers, and

$$\begin{aligned} a &= (a_1, a_2, \cdots, a_k; b_1, b_2, \cdots, b_k), \\ a' &= (a_1, a_2, \cdots, a_{k-1}, p^{2^t}, 1; b'_1, b'_2, \cdots, b'_k, 1) \end{aligned}$$

be bases. We say that a, a' satisfy *Property* $P_2(k,t)$ if the following conditions are true: $b_i > 2b_{i+1}$ for $1 \le i < k$, and $b'_i \ge (8b_i)^{4^t}$ for $1 \le i \le k$.

Let N, j be integers where $j \in \{0, 1\}$. Let p be a prime.

Lemma 6 Let a, a' be bases satisfying $P_2(k,t)$. If $j \in M(a, N)$, then $j \bmod p \in M(a', h^{(t)}(N))$.

Corollary Let a, a' be admissible bases satisfying $P_2(k,t)$. If $j = \mathrm{mod}_a(N)$, then $j \bmod p = \mathrm{mod}_{a'}(h^{(t)}(N))$.

Proof By Lemma 2, we can write

$$h^{(t)}(y) = \sum_{0 \le \ell \le 3^t} \mu_\ell y^\ell, \tag{14}$$

where

$$\sum_\ell |\mu_\ell| \le 5^{(3^t - 1)/2}. \tag{15}$$

Write $N = q_1 + q_2 + \cdots + q_k$, where $q_k = j$ and

$$a_i \mid q_i \text{ and } |q_i| \le b_i \tag{16}$$

for $1 \le i \le k$. Then

$$(q_1 + q_2 + \cdots + q_k)^\ell = \sum_{1 \le i \le k} Q_{\ell,i}, \tag{17}$$

where

$$Q_{\ell,i} = \sum_{1\le r\le \ell} \binom{\ell}{r} q_i^r (q_{i+1} + q_{i+2} + \cdots + q_k)^{\ell - r}. \quad (18)$$

Note that

$$q_i \mid Q_{\ell,i}. \quad (19)$$

It follows from (14) and (17) that

$$\begin{aligned} h^{(t)}(N) &= \sum_{0\le \ell\le 3^t} \mu_\ell (q_1 + q_2 + \cdots + q_k)^\ell \\ &= \sum_{0\le \ell\le 3^t} \mu_\ell \sum_{1\le i\le k} Q_{\ell,i} \\ &= \sum_{1\le i\le k} q_i', \end{aligned} \quad (20)$$

where

$$q_i' = \sum_{0\le \ell\le 3^t} \mu_\ell Q_{\ell,i}. \quad (21)$$

As $b_\alpha > 2b_{\alpha+1}$, we have from (16) and (18) that

$$\begin{aligned} |Q_{\ell,i}| &\le \sum_{1\le r\le \ell} \binom{\ell}{r} (b_i)^r (b_{i+1} + \cdots + b_k)^{\ell - r} \\ &\le (b_i)^\ell \sum_{1\le r\le \ell} \binom{\ell}{r} \\ &\le (2b_i)^\ell. \end{aligned} \quad (22)$$

It follows from (15), (21) and (22) that

$$\begin{aligned} |q_i'| &\le \sum_{0\le \ell\le 3^t} |\mu_\ell| (2b_i)^\ell \\ &\le (2b_i)^{3^t} \sum_{0\le \ell\le 3^t} |\mu_\ell| \\ &\le (2b_i)^{3^t} 5^{(3^t-1)/2} \\ &\le b_i' - 1. \end{aligned} \quad (23)$$

Note that by (16), (19) and (21)

$$a_i \mid q_i'. \quad (24)$$

Now by (14), (18) and (21) $q_k' = h^{(t)}(q_k) = h^{(t)}(j)$. By Lemma 1, we have

$$\mathrm{mod}_{p^{2^t}}(h^{(t)}(j)) = j \bmod p. \quad (25)$$

Let $q_i'' = q_i'$ for $1 \le i < k$, $q_{k+1}'' = j \bmod p$, and $q_k'' = h^{(t)}(j) - j \bmod p$. Then by (20)

$$h^{(t)}(N) = \sum_{1\le i\le k+1} q_i''. \quad (26)$$

Also, from (23)-(25), we have

$$a_i \mid q_i'' \text{ and } |q_i''| \le b_i', \quad (27)$$

for $1 \le i \le k+1$. This means $j \bmod p \in M(a', h^{(t)}(N))$, and proves the lemma. The corollary follows immediately. □

Let k, ℓ, t be positive integers and p be a prime. Let $a = (a_1, a_2, \cdots, a_k; b_1, b_2, \cdots, b_k)$, $a' = (a_1, a_2, \cdots, a_{k-1}, p^{2^t}, 1; b_1', b_2', \cdots, b_k', 1)$ be bases. We say that a, a' satisfy *Property* $P_3(k, \ell, t, p)$ if the following conditions are true: $b_i > 2b_{i+1}$ for $1 \le i < k$, $b_k = 1$, and $\log b_i' \ge p4^{t+3} \max\{1, \log \ell, \log b_i\}$ for $1 \le i \le k$.

Let $N_1, N_2, \cdots, N_\ell$ be integers and $j_1, j_2, \cdots, j_\ell \in \{0, 1\}$.

Lemma 7 Let a, a' be admissible bases satisfying $P_3(k, \ell, t, p)$. If $j_r = \mathrm{mod}_a(N_r)$ for $1 \le r \le \ell$, then $\mathrm{MOD}_p(j_1, j_2, \cdots, j_\ell) = \mathrm{mod}_{a'}(h^{(t)}(1 - (\sum_{1\le r\le \ell} N_r)^{p-1}))$.

Proof Let $u(y_1, y_2, \cdots, y_\ell) = 1 - (\sum_{1\le r\le \ell} y_r)^{p-1}$. Then $\mathrm{MOD}_p(j_1, j_2, \cdots, j_\ell) = 1$ if and only if $u(j_1, j_2, \cdots, j_\ell) \bmod p = 1$.

Let $a'' = (a_1, a_2, \cdots, a_k; b_1'', b_2'', \cdots, b_k'')$, where $b_i'' = \ell^p (2b_i)^{p-1}$ for $1 \le i \le k$. Then a, a'' clearly satisfy $P_1(p-1, p\log \ell, k)$.

Now, u is a polynomial of degree $p-1$ and of size no greater than $p \log \ell$. By Lemma 5, we have

$$u(j_1, j_2, \cdots, j_\ell) \in M(a'', u(N_1, N_2, \cdots, N_\ell)).$$

Now, let $j = u(j_1, j_2, \cdots, j_\ell)$ and $N = u(N_1, N_2, \cdots, N_\ell)$. Then the above equation can be written as $j \in M(a'', N)$. Note that $b_i'' > 2b_{i+1}''$ for $1 \le i < k$. For $1 \le i \le k$, we have

$$\log b_i' \ge p4^{t+2}(\log \ell + \log b_i + 1),$$

and

$$\log b_i'' \le p(\log \ell + \log b_i + 1).$$

This implies that, for $1 \le i \le k$, $b_i' \ge (8b_i'')^{4^t}$. Thus, a'', a' satisfy $P_2(k, t)$. By Lemma 6, we have $j \bmod p \in M(a', h^{(t)}(N))$. As a' is admissible, we have $j \bmod p = \mathrm{mod}_{a'}(h^{(t)}(N))$. This proves Lemma 7. □

4.2 Functions

Let k, n be positive integers. Let g, G be functions, where $g : \{0,1\}^n \to \mathbf{Z}$ and $G : \{0,1\}^n \to \{0,1\}$. Let $a = (a_1, \cdots, a_k; b_1, \cdots, b_k)$ be a base. We write $G \in M(a, g)$ if $G(\tilde{x}) \in M(a, g(\tilde{x}))$ for all $\tilde{x} \in \{0,1\}^n$. If, furthermore, a is admissible, then we write $G = \mathrm{mod}_a(g)$. We regard (a, g) as a *representative* of G. Call this the *supermodular representation* of Boolean functions.

All the three lemmas in Section 4.1 have analogs in the above representation of Boolean functions. For the present purpose, we will discuss the analog for Lemma 7.

Let k, ℓ, t be positive integers and p be a prime. Let $a = (a_1, a_2, \cdots, a_k; b_1, b_2, \cdots, b_k)$, $a' = (a_1, a_2, \cdots, a_{k-1}, p^{2^t}, 1; b'_1, b'_2, \cdots, b'_k, 1)$ be bases satisfying *Property* $P_3(k, \ell, t, p)$.

Let $g_1, g_2, \cdots, g_\ell$ be integer-valued functions

$$g_i : \{0,1\}^n \to \mathbf{Z},$$

and $G_1, G_2, \cdots, G_\ell$ be Boolean functions

$$G_i : \{0,1\}^n \to \{0,1\}.$$

Lemma 8 Let a, a' be admissible bases satisfying $P_3(k, \ell, t, p)$. If $G_r = \mathrm{mod}_a(g_r)$ for $1 \le r \le \ell$, then $\mathrm{MOD}_p(G_1, G_2, \cdots, G_\ell) = \mathrm{mod}_{a'}(h^{(t)}(1 - (\sum_{1 \le r \le \ell} g_r)^{p-1}))$.

Proof The lemma follows from Lemma 7 immediately. □

5 Proof of Theorem 1

We first outline the approach. Let C be a canonical ACC-circuit of size at most n^β and with signature $(\mathrm{MOD}_{p_1}, \mathrm{MOD}_{p_2}, \cdots, \mathrm{MOD}_{p_k})$, where the inputs to the bottom gates are of the form x_i and $1 - x_i$, $1 \le i \le n$. We will first construct a sequence of bases $a^{(0)}, a^{(1)}, \cdots, a^{(k)}$, then show that they are admissible bases, and that $a^{(i)}$ and $a^{(i-1)}$ satisfy property P_3 with suitable parameters. By Lemma 8, the output of any i-th level gate can be written as $\mathrm{MOD}_{a^{(i)}}(g(x_1, x_2, \cdots, x_n))$, where $g = u(g_1, g_2, \cdots, g_{n^\beta})$ for some polynomial u of small degree and size; for each j, $\mathrm{MOD}_{a^{(i-1)}}(g_j)$ is the output of some $(i-1)$-st level gate. This allows us to prove inductively that the output of any i-th level gate in C can be written as $\mathrm{MOD}_{a^{(i)}}(g(x_1, x_2, \cdots, x_n))$, where g is a polynomial of small degree and size. We now give the details.

Without loss of generality, we can assume that n is a power of two and is large enough so that

$$\log n \ge \prod_{1 \le j \le k} ((64)^4 \beta^2 p_j (\log p_j)^2). \qquad (28)$$

Let $m_i = 8\beta(\log n)^{\lambda^{k-i}}$ and $t_i = \lceil \log m_i \rceil$ for $1 \le i \le k$. For $1 \le r \le i \le k$, let $b_{i,r} = 2^{e_{i,r}}$, where $e_{i,r} = \beta(\log n) \prod_{r \le j \le i} (8^4 m_j^2 p_j (\log p_j)^2)$; let $b_{i,i+1} = 1$. Define $b_{0,1} = 1$.

Let $a^{(i)} = (a_{i1}, a_{i2}, \cdots, a_{i,i+1}; b_{i1}, b_{i2}, \cdots, b_{i,i+1})$ for $0 \le i \le k$, where $a_{i,i+1} = 1$ if $0 \le i \le k$, and $a_{i,r} = 2^{m_r \log p_r}$ if $1 \le r \le i$.

Fact 1 $a^{(i)}$ is an admissible base for $0 \le i \le k$.

Proof This is obviously true for $i = 0$. Let $i \ge 1$. With the use of (28), it is not hard to verify that $a_{i,r} > 4b_{i,r+1}$ for $1 \le r \le i$. By Lemma 4, $a^{(i)}$ is admissible. □

Fact 2 $\log b_{i,r} \ge p_i 4^{t_i+3} \max\{1, \beta \log n, \log b_{i-1,r}\}$ for $1 \le r \le i \le k$.

Proof If $r = i$, then $\max\{1, \beta \log n, \log b_{i-1,r}\} = \beta \log n$. Since $4m_i^2 \ge 4^{t_i}$, we have $e_{i,r} = \beta(\log n)(64m_i)^2 p_i (\log p_i)^2 \ge p_i 4^{t_i+3} \beta \log n$, as was claimed.

If $1 \le r < i$, then $e_{i,r} = (64m_i)^2 p_i (\log p_i)^2 e_{i-1,r}$. Since $e_{i-1,r} > \beta \log n \ge 1$ and $4m_i^2 \ge 4^{t_i}$, we have $e_{i,r} \ge 4^{t_i+3} p_i \max\{1, \beta \log n, e_{i-1,r}\}$, as was to be proved. □

We now prove the following proposition. Let $D_i = \prod_{1 \le j \le i} (p_j 3^{t_j+2})$ and $S_i = (2\beta \log n)^i D_i$ for $1 \le i \le k$, and $D_0 = 1$, $S_0 = 1$.

Proposition Let $0 \le i \le k$. The output of any i-th level gate in C can be written in the form mod

$_{a^{(i)}}(g(x_1, x_2, \cdots, x_n))$, where $g \in \mathbf{Z}^{[n]}$ is of degree at most D_i and size at most S_i.

Proof We prove the Proposition by induction. Any 0-th level output is either x_i or $1 - x_i$, hence the Proposition is satisfied for $i = 0$. For the inductive step, let $i > 0$ and assume that the Proposition is true for all smaller values of i. Let $K(x_1, x_2, \cdots, x_n)$ be the output of an i-th level $\mathrm{MOD}_{p_{(i)}}$ gate.

Clearly, $K(\tilde{x}) = \mathrm{MOD}_{p_i}(G_1(\tilde{x}), G_2(\tilde{x}), \cdots, G_\ell(\tilde{x}))$ where $\ell \leq n^\beta$ and each $G_j(\tilde{x})$ is the output of some gate on the $(i-1)$-st level. By the induction hypothesis, $G_j(\tilde{x}) = \mathrm{mod}_{a^{(i-1)}}(g_j(\tilde{x}))$, where $g_j \in \mathbf{Z}^{[n]}$ has degree and size at most D_{i-1} and S_{i-1}.

Let $v(y_1, y_2, \cdots, y_\ell) = h^{(t_i)}(1 - (\sum_{1 \leq r \leq \ell} y_i)^{p_i - 1})$. Using Lemma 2 and Lemma 3, we see that v is a polynomial of degree $d_v \leq 3^{t_i}(p-1)$ and size $s_v \leq \frac{1}{2}(3^{t_i} - 1)\log 5 + 3^{t_i} p_i \log \ell \leq 3^{t_i+2} p_i \log \ell$.

Now, by Fact 1, 2 and Lemma 8, we can write $K(\tilde{x}) = \mathrm{mod}_{a^{(i)}}(v(g_1(\tilde{x}), g_2(\tilde{x}), \cdots, g_\ell(\tilde{x})))$. Using Lemma 3, we see that $v(g_1(\tilde{x}), g_2(\tilde{x}), \cdots, g_\ell(\tilde{x}))$ has degree at most $d_v D_{i-1} \leq D_i$ and size $s_v + d_v S_{i-1} \leq S_i$. This completes the inductive proof of Proposition. □

Let $\nu = \log 3/(\lambda - 1)$; clearly, $\nu\lambda^k + k < \lambda^k$. A simple calculation shows that $D_k = O((\log n)^{\nu\lambda^k})$ and $S_k = O((\log n)^{\nu\lambda^k + k})$. Setting $i = k$ in the Proposition, we conclude that there exists a polynomial $f_n \in \mathbf{Z}^{[n]}$ with $\|f_n\| = O((\log n)^{\lambda^k})$ such that circuit C computes $\mathrm{mod}_{a^{(k)}}(f_n(x_1, x_2, \cdots, x_n))$. This is almost the statement of Theorem 1, except that $\mathrm{mod}_{a^{(k)}}(f_n(\tilde{x}))$ is in general not equal to $\mathrm{mod}(f_n(\tilde{x}); a_1, a_2, \cdots, a_k)$.

Let w_i be integers such that $b_{k,i} < w_i a_{k,i} \leq 2b_{k,i}$, $1 \leq i \leq k$. Define $F_n(\tilde{x}) = f_n(\tilde{x}) + \sum_{1 \leq i \leq k} w_i a_i$. We will prove that

$$\mathrm{mod}_{a^{(k)}}(F_n(\tilde{x})) = \mathrm{mod}(F_n(\tilde{x}); a_1, a_2, \cdots, a_k), \tag{29}$$

which clearly will establish Theorem 1.

Write

$$f_n(\tilde{x}) = \sum_{1 \leq i \leq k+1} z_i(\tilde{x}) a_{k,i},$$

where

$$|z_i(\tilde{x})| \leq b_i/a_i$$

and

$$z_{k+1}(\tilde{x}) = \mathrm{mod}_{a^{(k)}}(f_n(\tilde{x})). \tag{30}$$

Then

$$F_n(\tilde{x}) = \sum_{1 \leq i \leq k} (w_i + z_i(\tilde{x}) a_{k,i}) + z_{k+1}(\tilde{x}). \tag{31}$$

As $0 < w_r + z_r(\tilde{x}) a_{k,r} \leq 3b_{k,r}$ for each $1 \leq r \leq k$, we have, for each $1 < i \leq k$,

$$\begin{aligned} 0 &\leq \sum_{i \leq r \leq k} (w_r + z_r(\tilde{x}) a_{k,i}) + z_{k+1}(\tilde{x}) \\ &\leq \sum_{i \leq r \leq k} 3b_{k,r} + 1 \\ &\leq 6b_{k,i} + 1 \\ &< a_{k,i-1}. \end{aligned} \tag{32}$$

Equation (29) follows immediately from (30), (31) and (32). We have completed the proof of Theorem 1.

6 Concluding Remarks

In this paper we have proved several characterization theorems for ACC. It would be of great interest if these results could lead to a proof that certain languages are outside of ACC.

One approach is to study the power of depth-3 threshold circuits of size $O(2^{(\log n)^c})$ with the bottom gates being AND gates of fan-in $O((\log n)^c)$. Recently and independently of the present work, Hastad and Goldman [HG] showed that certain languages cannot be computed by depth-3 threshold circuits of subexponential size with bottom gates (not necessarily being AND gates) having fan-in $(\log n)/4$. The proof makes use of the results by Babai, Nisan, and Szegedy [BSS] on multiparty communication complexity. Further progress along this line might be possible.

Another approach is to focus on the language class $\mathcal{I} = \cup_{c \geq 1} I_c$, where I_c is defined as the set of languages representable in the form $\{\tilde{x} \mid \Psi_n(F_n(\tilde{x})) = 1\}$, with $F_n \in \mathbf{Z}^{[n]}$, $\|F_n\| \leq (\log n)^c$, and Ψ_n :

$\{0,1,\cdots,\lceil 4^{(\log n)^{c+1}}\rceil\} \to \{0,1\}$. Theorem 1 states that pure ACC is contained in $\mathcal{I}$, and Theorem 2 states that any ACC language can be approximated by some language in $\mathcal{I}$. We conjecture that the canonical NC^1 language is not in $\mathcal{I}$, and hence not in ACC. A number of interesting combinatorial questions arise in pursuing this approach, which we will report in a future paper.

References

[Aj] M. Ajtai, "Σ_1^1 formulae on finite structures," *Annals of Pure and Applied Logic* **24** (1983), 1-48.

[A11] E. Allender, "A note on the power of threshold circuits," *Proceedings of the 30th Annual IEEE Symposium on Foundations of Computer Science* (1989), 580-584.

[AH] E. Allender and U. Hertrampf, "On the power of uniform families of constant depth threshold circuits," draft, January 1990.

[BSS] L. Babai, N. Nisan, and M. Szegedy, "Multiparty protocols and logspace-hard pseudorandom sequences," *Proceedings of the 21st ACM Symposium on Theory of Computing* (1989), 1-11.

[Ba] D. A. Barrington, "Bounded-width polynomial size branching programs recognize exactly those languages in NC^1," *Proceedings of the 18th ACM Symposium on Theory of Computing* (1986), 1-5.

[Br] J. Bruck, "Harmonic analysis of polynomial threshold functions," preprint, IBM Almaden Research Center, July 1988.

[BS] J. Bruck and R. Smolensky, "Polynomial, threshold functions, AC^0 functions and spectral norms," IBM Research Report RJ 7140, November 1989.

[FSS] M. Furst, J. Saxe, and M. Sipser, "Parity, circuits, and the polynomial-time hierarchy," *Proceedings of the 22th Annual IEEE Symposium on Foundations of Computer Science* (1981), 260-270.

[Ha] J. Hastad, "Almost optimal lower bounds for small depth circuits," *Proceedings of the 18th ACM Symposium on Theory of Computing* (1986), 6-20.

[HG] J. Hastad and M. Goldman, "On the power of small-depth threshold circuits," the present proceedings.

[LMN] N. Linial, Y. Mansour, and N. Nisan, "Constant depth circuits, Fourier transform, and learnability," *Proceedings of the 30th Annual IEEE Symposium on Foundations of Computer Science* (1989), 574-579.

[Ra] A. A. Razborov, "Lower bounds for the size of circuits of bounded depth with basis $\{\wedge,\oplus\}$," *Mat. Zametki.* **41 (4)** (1987), 598-607. (English translation: *Mathematical Notes of the Soviet Academy of Science of USSR* **41 (4)**, 333-338.)

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

[To] S. Toda, "On the computational power of PP and $\oplus$PP," *Proceedings of the 30th Annual IEEE Symposium on Foundations of Computer Science* (1989), 514-519.

[VV] L. Valiant and V. Vazirani, "NP is as easy as detecting unique solutions," *Theoretical Computer Science* **47** (1986), 85-93.

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