**YALE UNIVERSITY**
**Department of Computer Science**

# EXPONENTIAL LOWER BOUNDS FOR CONSTANT DEPTH CIRCUITS IN THE PRESENCE OF HELP BITS

Jin-yi Cai, Yale University*

YALEU/DCS/TR-670
January 1989

* Research supported by NSF grant CCR-8709818.

# Exponential Lower Bounds for Constant Depth Circuits in the Presence of Help Bits

*Jin-yi Cai**
Department of Computer Science
Yale University
New Haven, CT 06520

**(Extended Abstract)**

**Abstract**

Constant depth unbounded fan-in Boolean circuits have been the success story in our never ending quest for lower bounds in complexity theory. Exponential lower bounds were obtained by Yao, that established the first oracle separating the Polynomial time hierarchy and PSPACE. Recently the algebraic approach by Razborov, and subsequently by Smolensky, has stripped away much of the convoluted combinatorial argument in favor of its inherent structural clarity. Meanwhile there is a renewed interest in various complexity classes of restricted number of queries. We consider the problem of how many extra bits of "help" does a constant depth circuit need in order to compute $m$ sets of parity functions. We prove an exponential lower bound in the presence of $m - 1$ help bits. The proof is carried out using the algebraic machinery of Razborov and Smolensky. A by-product of the proof is that the same bound holds for circuits with $Mod_p$ gates for a fixed prime $p > 2$. We will also discuss some tight approximation results in the associated algebra and its relationship to the Hadamard matrix.

## 1 Introduction

Computational complexity theory is the study of minimal resources required for computational problems. The theory originated in the work of Hartmanis, Lewis and Stearns [HLS]

*Research supported by NSF grant CCR-8709818.

concerning time and space bounded Turing machine computations, which culminated in the famous conjecture $NP \neq P$ [Co][Ka].

Perhaps the most "promising" approach to this conjecture today is the Boolean circuit model; a super-polynomial lower bound on the circuit size for any NP language will settle $NP \neq P$. However, for the complexity of general Boolean circuits, little is known concerning any natural function beyond linear lower bound.

On the other hand, the class of constant depth unbounded fan-in Boolean circuits has been a success story in our quest for lower bounds in complexity theory. The famous Russian theoretician Lupanov [Lu] observed a $\Omega(2^{n-1})$ lower bound for depth 2 Boolean circuits computing the parity function $\oplus$, where

$$\oplus(x_1, x_2, \ldots, x_n) = 1 \textit{ iff} \text{ the number of 1's in } x_i \text{ is odd.}$$

More substantial lower bounds followed for circuits of depth $k > 2$. Furst, Saxe and Sipser [FSS], and independently Ajtai [Aj], obtained super-polynomial lower bound on the circuit size for any constant depth unbounded fan-in Boolean circuit computing the parity function $\oplus$. The original motivation of Furst, Saxe and Sipser, from a complexity theory point of view, was to construct oracles that separate the Meyer-Stockmeyer polynomial time hierarchy PH and PSPACE, which was a long standing open problem [St]. Their dream was fulfilled with the breakthrough of Yao, who established exponential lower bounds on the size, which in turn produced the first oracle separating PH and PSPACE [Ya]. A sequence of improvements followed quickly. Cai showed that with the same exponential bounds as the one Yao established on the size, the constant depth circuits fail to compute the parity function $\oplus$ on asymptotically 50% of the inputs [Ca]. His result implies that a random oracle separates PH from PSPACE with probability one. Babai gave a remarkably short proof to the probability one separation result [Ba]. Hastad improved Yao's lower bound to almost optimal [Ha].

All these proofs mentioned are based on certain probabilistic counting argument. Recently an algebraic approach developed by Razborov, and subsequently by Smolensky, has stripped away much of the combinatorics in favor of its inherent structural clarity [Ra] [Sm]. In their approach, a certain commutative algebra was used as the representation of the ring of Boolean functions. The isomorphism is highly suggestive of a coordinate ring from the viewpoint of algebraic geometry.

In the field of structural complexity, meanwhile, there is a renewed interest in various complexity classes defined by restricted queries. For instance, Kadin showed that the Boolean Hierarchy BH is proper if PH does not collapse [Cai et al.][Kad]. Much work (and heat) was generated over concepts such as terseness.

In this paper, we consider a similar problem in the low-level complexity setting. We ask

how many extra bits of "help"—arbitrary Boolean functions—does a constant depth circuit need in order to compute $m$ sets of parity functions. We prove an exponential lower bound on the size of the circuits in the presence of $m-1$ help bits; $m-1$ is certainly tight. Unlike similar results involving terseness at the polynomial time level, the result here is *absolute*, i.e. it does not depend on any unproven hypothesis. The proof is easily carried out using the algebraic machinery of Razborov and Smolensky. A by-product of the proof is that the same bound holds for circuits with $Mod_p$ gates for a fixed prime $p > 2$. We will also discuss some tight approximation results in the associated algebra and its relationship to the Hadamard matrix.

## 2 Definitions and Preliminaries

Consider a multi-output Boolean circuit $C$ of unbounded fan-in and constant depth $k$. Assume $C$ has input variables $X \cup B$, and $m$ outputs $\{f_1, \ldots, f_m\}$, where $X = \{x_1, \ldots, x_N\}$ is a set of Boolean variables, and $B = \{b_1, \ldots, b_{m-1}\}$ will be used as "help bits". The size of the circuit $|C|$ is the number of gates in $C$.

We recast briefly the Razborov-Smolensky algebraic set-up [Ra][Sm]. For indeterminates $X \cup B$, define

$$\mathcal{R} = \mathbf{Z}_3[X] \;/\; ( \; x_i^2 - x_i \mid 1 \leq i \leq N \; )$$

and

$$\mathcal{R}^B = \mathbf{Z}_3[X \cup B] \;/\; ( \; x_i^2 - x_i, b_j^2 - b_j \mid 1 \leq i \leq N, 1 \leq j \leq m-1 \; ).$$

**Lemma 2.1** *$\mathcal{R}$ is a commutative algebra of dimension $2^N$.*

$\mathcal{R}$ has the following basis.

$$\mathbf{\Pi} = \{ \; \Pi_{i \in \omega}(1 + x_i) \mid \omega \subseteq \{1, \ldots, N\} \; \}, \text{ (parity basis)}$$

$$\mathbf{A} = \{ \; \Pi_{i \in \omega} \hat{x_i} \mid \omega \subseteq \{1, \ldots, N\}, \text{ and } \hat{x_i} = \begin{cases} x_i & \text{if } i \in \omega \\ 1 - x_i & \text{otherwise} \end{cases} \; \}, \text{ (assignment basis).}$$

We have the following basis transformation

$$H\mathbf{\Pi} = (-1)^N \mathbf{A},$$

where the matrix $H$ is the $N$th Hadamard matrix defined inductively by $H_0 = 1$, and

$$H_n = \begin{pmatrix} H_{n-1} & H_{n-1} \\ H_{n-1} & -H_{n-1} \end{pmatrix}.$$

We note that for any variable $v \in X \cup B$, $v^2 = v$ in $\mathcal{R}^B$. Next we define $\deg p$, for

$p \in \mathcal{R}$, to be $\min\{ \deg \hat{p} \mid \hat{p} \in \mathbf{Z}_3[X]$ and $\hat{p} \bmod (x_i^2 - x_i) = p \}$. The degree for $p^B \in \mathcal{R}^B$ is similarly defined. We have $\deg(f \cdot g) \leq \deg f + \deg g$.

Let $\mathcal{F}$ be the class of Boolean functions on $X$, we have an injection $\phi$ from $\mathcal{F}$ to $\mathcal{R}$ satisfying:

$$\begin{aligned} 0^\phi = 0, 1^\phi = 1, x_i^\phi &= x_i, \\ (\neg f)^\phi &= 1 - f^\phi, \\ (f \vee g)^\phi &= f^\phi + f^\phi - f^\phi \cdot g^\phi, \\ (f \wedge g)^\phi &= f^\phi \cdot g^\phi. \end{aligned}$$

**Lemma 2.2** *An ideal of $\mathcal{R}$ is generated by a subset of $\mathbf{A}$.*

**Proof** Let $\mathcal{R}'$ be an ideal of $\mathcal{R}$, let $\mathbf{A}' = \mathbf{A} \cap \mathcal{R}'$. We claim $\mathcal{R}' = \langle \mathbf{A} \rangle$. Let $p \in \mathcal{R}'$, $p = \sum_{a_\omega \in \mathbf{A}} \lambda_\omega a_\omega$. Since $\mathcal{R}'$ is an ideal, if $\lambda_\omega \neq 0$, then $a_\omega = \lambda_\omega^{-1} a_\omega p \in \mathcal{R}'$. **QED**

**Lemma 2.3** *For $f, g \in \mathcal{F}$, $f^\phi = g^\phi \bmod K$ for some ideal $K$ of dimension $d$ iff $f = g$ for $2^N - d$ assignments as Boolean functions.*

## 3 $m - 1$ bits are not enough for $m$ parities

The following approximation lemma is an easy consequence of a result of Barrington [Bar] (stated as Lemma 1 in [Sm]).

**Lemma 3.1** *Given any $\varepsilon(N)$ and $d(N)$, any Boolean circuit $C$ of unbounded fan-in and constant depth $k$, and size $|C| \leq E_k(N)$, where $E_k(N) = \varepsilon(N)\sqrt{3}^{d(N)^{1/k}}$. Let $h_i(X), 1 \leq i \leq m-1$, be any Boolean functions, and $f_i(X), 1 \leq i \leq m$, be the outputs of $C$, when substituting $h_i(X)$ for $b_i$, $1 \leq i \leq m-1$. Then there exist an ideal $K$ in $\mathcal{R}$ of $\dim K \leq \varepsilon(N)2^N$, and polynomials $p_i(X, B) \in \mathcal{R}^B$ of degree $\leq d(N)$, such that $f_i(X) = p_i(X, h_1(X), \ldots, h_{m-1}(X))$ in $\mathcal{R}/K$, $1 \leq i \leq m$.*

We note that the approximation is on the variables $X$ and not on $X \cup B$, although our degree measure is taken in $\mathcal{R}^B$.

Now suppose $N = mn$, and write $X = \{ x_{ij} \mid 1 \leq i \leq m, 1 \leq j \leq n \}$. We consider $m$ parity functions $\oplus(x_{i1}, \ldots, x_{in})$. In the algebra $\mathcal{R}$ they assume the form $-1 - \Pi_i$, where $\Pi_i = \prod_j (1 + x_{ij})$. We prove an exponential lower bound on $|C|$ with the help of arbitary $m - 1$ Boolean functions, if it computes all $\Pi_i, 1 \leq i \leq m$.

**Theorem 3.2** *Let $\varepsilon > 0$ and $m \leq N^{1/5-\varepsilon}$. Suppose a Boolean circuit $C$ of unbounded fan-in and constant depth $k$ computes all parity functions $\oplus\{x_{i1}, \ldots, x_{in}\}$ with the help of some $m - 1$ Boolean functions $h_i(X), 1 \leq i \leq m - 1$. Then $\exists \lambda = \lambda_{\varepsilon,k}$, such that $|C| = \Omega(2^{N^\lambda})$.*

**Proof** Let $d(N) = N^\varepsilon$. Suppose $|C| \leq \varepsilon(N)\sqrt{3}^{N^{\frac{\varepsilon}{k}}}$, where $\varepsilon(N)$ is to be determined. We have $p_i(X, B) \in \mathcal{R}^B$ of degree $\leq N^\varepsilon$, and an ideal $\mathcal{K}$ of dimension $\leq \varepsilon(N)2^N$, such that

$$\begin{aligned} \Pi_1 &= p_1(X, h_1(X), \ldots, h_{m-1}(X)) \\ &\vdots \\ \Pi_m &= p_m(X, h_1(X), \ldots, h_{m-1}(X)) \end{aligned}$$

in $\mathcal{R}/\mathcal{K}$.

We observe that each $h_i$ is $\{0,1\}$-valued, thus $h_i^2 = h_i$, for all $i$. Hence there is a canonical representation for each $p_i$,

$$p_i = \sum_{\Lambda \subseteq \{1,\ldots,m-1\}} f_i^\Lambda(X) \cdot \prod_{j \in \Lambda} h_j(X),$$

where $\deg f_i^\Lambda \leq N^\varepsilon$.

Consider the parity basis $\mathbf{\Pi}$ of $\mathcal{R}$. Let $\omega \subseteq X$ and

$$\prod_{x \in \omega}(1+x) = \prod_{x_{1j} \in \omega_1}(1+x_{1j}) \cdots \prod_{x_{mj} \in \omega_m}(1+x_{mj}),$$

where $\omega = \omega_1 \cup \cdots \cup \omega_m$, and $\omega_i \subseteq \{x_{i1}, \ldots, x_{in}\}$.

If $|\omega_i| > n/2$, following Smolensky, we replace $\prod_{\omega_i}(1+x_{ij})$ by

$$\prod_{\omega_i^c}(1+x_{ij}) \cdot \Pi_i = \sum_{\Lambda \subseteq \{1,\ldots,m-1\}} g_i^\Lambda(X) \cdot \prod_{j \in \Lambda} h_j(X)$$

in $\mathcal{R}/\mathcal{K}$.

We note that each $g_i^\Lambda$ has degree $\leq n/2 + N^\varepsilon$. *But more can be said:* $g_i^\Lambda$ has degree $\leq n/2 + N^\varepsilon$ in $\{x_{i1}, \ldots, x_{in}\}$, and $\leq N^\varepsilon$ for all other $x$'s.

Thus we get

$$\prod_{x \in \omega}(1+x) = \sum_{\Lambda \subseteq \{1,\ldots,m-1\}} p_\omega^\Lambda(X) \cdot \prod_{j \in \Lambda} h_j(X),$$

where each $p_\omega^\Lambda(X)$ has degree $\leq n/2 + mN^\varepsilon$ in each group of variables $\{x_{i1}, \ldots, x_{in}\}$.

Note that the number of monomials in $X$ satisfying the restriction is bounded by

$$\left(\sum_{i=0}^{n/2+mN^\varepsilon} \binom{n}{i}\right)^m,$$

which can be estimated by $2^{N-m}(1-o(1))$. Since there are at most $2^{m-1}$ such terms $p_\omega^\Lambda(X)$ in $\prod_{x \in \omega}(1+x)$, we have shown that

$$\dim \mathcal{R}/\mathcal{K} \leq (1/2 - o(1))2^N.$$

Now let $\varepsilon(N) = o(1)$ we get a contradiction. **QED**

By setting $\varepsilon(N)$ more carefully, we have

**Corollary 3.3** *The same bound holds for circuits approximating $m$ sets of parity functions to within any fraction $\alpha > 1/2$ of all assignments, with the help of arbitrary $m-1$ Boolean functions.*

A by-product is that $Mod_3$ gates can be assumed for free, and the same bound holds. A similar bound holds for any $Mod_p$ gate for fixed prime $p > 2$.

**Corollary 3.4** *For any fixed prime $p > 2$, circuits with $Mod_p$ gates approximating $m$ sets of parity functions to within any fraction $\alpha > 1/2$ of all assignments, with the help of arbitrary $m-1$ Boolean functions, must have size $|C| = 2^{N^{\Omega(1)}}$.*

# 4 A probability one separation

As in [Ca], the lower bound translates to a probability one separation. Here we consider how much extra information does the Polynomial time hierarchy need to subsume PSPACE. Surely sufficiently dense oracles would suffice.

**Theorem 4.1** *For any $\varepsilon > 0$,*

$$\forall A, \exists B, |B^{\leq n}| \leq 2^{n^{\varepsilon}}, PSPACE^A \subseteq P^{A \oplus B},$$

*(here $\oplus$ denotes disjoint sum, as it is customary in structural complexity.)*

The proof is, of course, to encode the PSPACE computation relative to $A$ in the oracle set $B$, sufficiently far away, and yet accessible in polynomial time.

More interestingly, we claim that the above density bound is tight:

**Theorem 4.2**

$$Pr.(\ A \mid \forall B, |B^{\leq n}| = 2^{n^{o(1)}}, PSPACE^A \not\subseteq PH^{A \oplus B}\ ) = 1.$$

For a proof sketch, we note that it suffices to show for the parity language $L^A = \{\ x \mid$ the number of $xy \in A$, $|y| = 6|x|$, is odd. $\} \in PSPACE^A$,

$$\exists \varepsilon > 0, \forall i, Pr.(\ A \mid \forall B, |B^{\leq n}| = 2^{n^{o(1)}}, L^A \notin PH^{A \oplus B}\ ) \geq \varepsilon,$$

where $M_i$ is an enumeration of all alternating Turing machines [BG].

Using the reduction in [FSS], we get a family of constant depth exponential poly-logarithmic size $\exp(\log(N)^{O(1)})$ circuits with unbounded fan-in, computing various parity functions. The input to the circuits are oracle memberships of $A$ and $B$ of strings upto the appropriate length. If we combine circuits for all computations $x \in^? L^A$ at the same length $|x| = n$, we get a circuit still of exponential poly-logarithmic size $\exp(\log(N)^{O(1)})$, where $N = 2^{6n}$, and computing $2^n = N^{1/6}$ sets of parity. However, if $N$ is large, the number of help bits from $B$ ($\ll N^{1/6}$) is nowhere nearly sufficient.

## 5 A tight approximation in $\mathcal{R}$

The result by Razborov and Smolensky that the parity function $\Pi_{i=1}^{n}(1+x_i)$ can not be approximated by small degree polynomials can be re-phrased in terms of Hadamard matrices as follows. Recall that the $n$th Hadamard matrix $H_n$ is the Kronecker product $\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \otimes H_{n-1}$. Let $i = i_1 \ldots i_n$ be the binary representation of $i$, and $w(i)$ be the Hamming weight of $i$, the number of 1's in $i_1 \ldots i_n$. Let $a_i = a_i^{(n)}$ be the $i$th row vector of $H_n$, $0 \leq i \leq 2^n - 1$.

**Theorem 5.1 (Smolensky)** *The number of non-zeros in*

$$a_{2^n-1} + \sum_{i \in S} \lambda_i a_i, \qquad \lambda_i \in \mathbf{Z}_3,$$

*is $\Omega(2^n)$, for any $S$ consisting of $i$ with $w(i) \leq O(\sqrt{n})$. Here the arithmetic is over $\mathbf{Z}_3$.*

A natural approach to Theorem 3.2 is to try to "solve" and "eliminate" the $m-1$ unknowns $h_i$ one-by-one, via the $m$ equations. In such an approach, as well as in other circumstances, one is faced with a similar sum as above $\sum_{i \in S} \lambda_i a_i$, where not all $\lambda_i = 0$ and $i \in S \Longrightarrow w(i) \leq w$ or $w(i) \geq n - w$, for some $w$.

Let $S_w(n) = \{\, i \mid 0 \leq i < 2^n, w(i) \leq w \,\}$. We have the following tight bound:

**Theorem 5.2** *Over any field $F$ of* char.$F \neq 2$, *the number of non-zeros in $\sum_{i \in S_w(n)} \lambda_i a_i$, where $\lambda_i \in F$ and not all $\lambda_i = 0$, is bounded below by $2^{n-w}$.*

*The situation with $w(i) \geq n - w$ is dual.*

**Proof** Let $N(n,w)$ be the least number of nonzeros in a nontrivial sum $\sum_{i \in S_w(n)} \lambda_i a_i$. If $w = 0$ then clearly $N(n,0) = 2^n$. Also, if $w = n$, then $N(n,w) = 1$, since $H_n$ is nonsingular, and $\sum_{i=0}^{2^n-1} a_i$ has but one nonzero. Suppose $0 < w < n$, inductively let $S_0$ be an optimal set, such that $N(n-1, w-1) = 2^{n-w}$ is achieved by $\sum_{i \in S_0} a_i^{(n-1)}$ in $H_{n-1}$. Then the sum $\sum_{i \in S} a_i^{(n)}$, where $S = \{\, i \mid i = i_1 \ldots i_n \;\&\; j = i_2 \ldots i_n \in S_0 \,\}$, shows that $N(n,w) \leq 2^{n-w}$.

Now, suppose a sum $\sum_{i \in S} \lambda_i a_i$ achieves $N(n,w)$, where $\emptyset \neq S \subseteq S_w(n)$, and all $\lambda_i \neq 0$. Let $S_0 = \{\, j \mid j = j_2 \ldots j_n \;\&\; 0 j_2 \ldots j_n \in S \,\}$. Similarly we define $S_1$. Then $S_0 \subseteq S_w(n-1)$ and $S_1 \subseteq S_{w-1}(n-1)$. Clearly if either $S_0$ or $S_1 = \emptyset$, the Theorem is proved. Let $T_\ell$ be the nonzero columns in $\sum_{i \in S_\ell} \lambda_i a_i^{(n-1)}$, $\ell = 0, 1$. Assume neither $S_0$ nor $S_1 = \emptyset$, by induction, $|T_0| \geq N(n-1,w) \geq 2^{n-1-w}$, and $|T_1| \geq N(n-1, w-1) \geq 2^{n-w}$. Now there are exactly two nonzeros resulting from every $k \in T_0 \triangle T_1$, and exactly one for every $k \in T_0 \cap T_1$, in set of nonzeros of $\sum_{i \in S} \lambda_i a_i$. Here we take into account the defintion of $H_n$ and char.$F \neq 2$.

Hence $N(n,w) = 2|T_0 \triangle T_1| + |T_0 \cap T_1| \geq \max\{\, |T_0|, |T_1| \,\} \geq 2^{n-w}$. **QED**

## Acknowledgement

The author thanks Lane Hemachandra for interesting discussions on some closely related research. Theorem 4.2 directly stems from a discussion on an analogous problem in the P v.s NP level. He also thanks Professors Sandeep Bhatt, Mike Fischer, Merrick Furst, Roger Howe and Herb Scarf for encouraging conversations. Prof. Fischer has obtained a purely combinatorial proof in case of $m = 2$ in Theorem 3.2.

## References

[Aj] Ajtai, M., "$\Sigma_1^1$ -formulae on finite structures," *Ann. Pure and Applied Logic*, 24:1–48, 1983.

[Ba] Babai, L., "A random oracle separates PSPACE from Polynomial Hierarchy," *Notes*, 1986.

[Bar] Barrington, D., "A note on the theorem of Razborov," *Notes*, 1987.

[BG] Bennet, C., Gill, J., "Relative to a random oracle $A$, $P^A \neq NP^A$ with probability 1," *SIAM J. on Computing*, 10:96–113, 1981.

[Ca] Cai, J., "With probability one, a random oracle separates PSPACE from the polynomial time hierarchy," Proc. 18th ACM Sym. on Theory of Computation, 1986, pp 21-29.

[Cai et al.] Cai, J., T. Gundermann, J. Hartmanis, L. Hemachandra, V. Sewelson, K. Wagner, and G. Wechsung. The Boolean Hierarchy I: Structural Properties, and The Boolean Hierarchy II: Applications, To appear in *SIAM J. Computing*, 1988 and 1989.

[FSS] Furst, M., Saxe, J., Sipser, M., "Parity, circuits and the polynomial-time hierarchy," Mathematical Systems Theory, 17, 1984, pp 13-27.

[HLS] Hartmanis, J., Lewis, P., Stearns, R., "Hierarchies of memory limited computations", Proc. 6th Symp. on Switching Circuit Theory and Logical Design, 1965, pp 179-190.

[Ha] Hastad, J., "Almost optimal lower bounds for small depth circuits", Proc. 18th ACM Sym. on Theory of Computation, 1986, pp 6–20.

[Ka] Kadin, J., "The Polynomial time hierarchy collapses if the Boolean Hierarchy collapses," Proc. 3rd Structure in Complexity Theory Conference, pp 278–292.

[Ra] Razborov, A., "Lower bounds for the size of circuits of bounded depth with basis AND, XOR," *Notes*, 1987.

[Sm] Smolensky, R., "Algebraic methods in the theory of lower bounds for Boolean circuit complexity," Proc. 19th ACM Sym. on Theory of Computation, 1987, pp 77–82.

[St] Stockmeyer, L., "The Polynomial-time hierarchy," *Theoretical Computer Science*, 3:1–22, 1977.

[Y] Yao, A., "Separating the polynomial-time hierarchy by oracles," Proc. 26th IEEE Foundations of Computer Science, 1985, pp 1-10.