# Exponential Lower Bounds for Depth 3 Boolean Circuits

Ramamohan Paturi*, Michael E. Saks†, and Francis Zane‡

## Abstract

We consider the class $\Sigma_3^k$ of unbounded fan-in depth-3 boolean circuits, for which the bottom fan-in is limited by $k$ and the top gate is an OR. It is known that the smallest such circuit computing the parity function has $\Omega(2^{\varepsilon n/k})$ gates (for $k = O(n^{1/2})$) for some $\varepsilon > 0$, and this was the best lower bound known for explicit (P-time computable) functions. In this paper, for $k = 2$, we exhibit functions in uniform $NC^1$ that requires $2^{n-o(n)}$ size depth 3 circuits. The main tool is a theorem that shows that any $\Sigma_3^2$ circuit on $n$ variables that accepts $a$ inputs and has size $s$ must be constant on a projection (subset defined by equations of the form $x_i = 0$, $x_i = 1$, $x_i = x_j$ or $x_i = \bar{x}_j$) of dimension at least $\frac{\log(a/s)}{\log n}$.

*Department of Computer Science and Engineering, University of California, San Diego, La Jolla, Ca 92093

†Dept. of Mathematics, Rutgers University, New Brunswick, NJ 08903. This work was supported in part by NSF grant CCR-9215293 and by DIMACS (Center for Discrete Mathematics & Theoretical Computer Science), through NSF grant NSF-STC91-19999 and by the New Jersey Commission on Science and Technology.

‡Department of Computer Science and Engineering, University of California, San Diego, La Jolla, Ca 92093

## 1 Introduction

Considerable progress has been made in understanding the limitations of unbounded fan-in Boolean circuits of bounded depth. The results of Ajtái, Furst, Saxe, Sipser, Yao, Håstad, Razborov, and Smolensky [1, 6, 17, 8, 10, 12], among others, show that if the size of the circuit is not too large, then any function computed by such circuit must be constant on a large subcube or can be approximated by a small degree polynomial. Such limitations of small size bounded depth circuits can be used to show that certain explicit functions such as parity and majority require a large number of gates. More precisely, a result of Håstad [8] says that computing the parity function in depth $d$ requires $\Omega(2^{\varepsilon n^{1/(d-1)}})$ gates for some $\varepsilon < 1$. Except for the constant $\varepsilon$ this result is essentially tight.

Recently, Håstad, Jukna, and Pudlák [9] described a top down approach for proving lower bounds on depth 3 circuits. However, these and other techniques seem incapable of proving a lower bound on depth 3 circuits of the form $\Omega(2^{h(n)\sqrt{n}})$ with $h(n)$ unbounded, for any explicit Boolean function. Here, as usual, the term "explicit function" is a somewhat informal term, which is taken to mean "uniformly and efficiently computable", in, say $P$ or $NC$.

To clarify the situation, it is useful to parameterize the lower bound in terms of the maximum fan-in of the bottom gates. Define $\Sigma_d^k$ to be the set of depth $d$ circuits with top gate $OR$ such that each bottom gate has fan-in at most $k$. Then it follows from known results that there is a constant $\varepsilon \leq 1$ such that for any $k \geq 1$, any $\Sigma_3^k$ circuit for the parity function or the majority function requires $\Omega(2^{\varepsilon n/k})$ gates at level 2, and such bounds are tight for $k = O(\sqrt{n})$.

As in Håstad, Jukna, and Pudlák [9], our motivation is to prove stronger lower bounds on depth–3 circuits that go beyond the above trade-off between bottom fan-in and size. We note that even for constant bottom fan-in $k \geq 2$, currently known lower bound techniques seem incapable of providing a lower bound better than $2^{n/k}$ on the number of gates at level 2. There is another independent compelling motivation for studying the depth–3 model with limited fan-in. Valiant [14] showed that linear–size logarithmic–depth Boolean circuits with bounded fan–in can be computed by depth–3 unbounded fan-in circuits of size $O(2^{n/\log\log n})$ and bottom fan-in limited by $n^{\varepsilon}$ for arbitrarily small $\varepsilon$. If we consider linear–size logarithmic–depth circuits with the additional restriction that the graph of the connections is series–parallel, then such circuits can be computed by depth–3 unbounded fan-in circuits of size $2^{n/2}$ with bounded bottom fan–in [14]. Thus, strong exponential lower bounds on depth 3 circuits would imply nonlinear lower bounds on size of fan-in 2 Boolean circuits with logarithmic–depth, an open problem proposed some twenty years ago [14].

In this paper, we take a modest step towards proving such strong bounds on depth–3 circuits. We show that for some explicit func-

tion, contained in log-space uniform $NC^1$, any $\Sigma_3^2$ circuit that computes it must have at least $2^{n-o(n)}$ gates. We obtain this result by showing that the function computed by a small $\Sigma_3^2$ circuit must be constant on a large "nicely structured" subset of the cube. These subsets, called projections, are defined by equating literals to each other or to constants.

The starting point for our argument is the top-down approach used in [9], which says that if the number of gates at level 2 is small, there must be a depth 2 subcircuit that accepts a large number of inputs. We prove that such a depth–2 subcircuit (which is a 2-CNF formula) must accept a projection of large size. We then give two constructions of functions such that any $\Sigma_3^2$ or $\Pi_3^2$ circuit computing them requires $2^{n-o(n)}$ size. For the first construction, we first show that, with high probability, a randomly chosen homogeneous multilinear $n$-variable polynomial of degree 2 over $GF(2)$ is non-constant on every large projection. We then use derandomization techniques to construct a specific Boolean function with the property that it has a subfunction on a large enough set of variables which is not constant on any large projection. This property is stronger than what we needed to prove lower bounds on depth 3 fan–in 2 circuits. In the second construction, we obtain a simpler function $f$ on $n$ variables using error–correcting codes which is not identically 1 on any large projection and hence requires a large size $\Sigma_3^2$ circuits. It then follows that the $n+1$ variable function $g = x_{n+1}f + \overline{x}_{n+1}\overline{f}$ requires large size $\Sigma_3^2$ and $\Pi_3^2$ circuits.

The rest of the paper is organized as follows: In section 2, we review some basic definitions and results, including a proof that any symmetric function can be computed by a $\Sigma_3^2$ circuit of size at most $poly(n)2^{0.59n}$. In section 3, we show that any 2-CNF which accepts a large number of inputs must necessarily accept a projection with large dimension. Using this result, in sections 4 and 5 we construct functions which do not have depth–3 bottom fan-in 2 circuits of size less than $2^{n-o(n)}$.

## 2 Preliminaries

### 2.1 Boolean variables, literals and assignments

Let $X$ denote the set $\{x_1, x_2, \ldots, x_n\}$ of variables and $L$ denote the set $\{x_1, \overline{x}_1, x_2, \overline{x}_2, \ldots, x_n, \overline{x}_n\}$ of literals. If $V$ is a subset of $L$, then $\overline{V}$ denotes the set $\{\overline{v} | v \in V\}$. An assignment of $X$ is a function $\alpha : X \longrightarrow \{0,1\}$, and a partial assignment is a function $\alpha$ from a subset of $X$ to $\{0,1\}$. Associated to any partial assignment $\alpha$ is the subset $X(\alpha) \subseteq X$ of variables set to 1 by $\alpha$ and the set $L(\alpha) \subseteq L$ of literals set to 1 by that assignment.

### 2.2 2-CNF Formulae

We briefly review some basic facts about 2-CNF formulae. A 2-CNF formula $\Phi$ on variable set $X$ can be associated naturally with its implication digraph $D(\Phi)$, whose vertex set is $L$. Each clause $v \vee w$ (where $v$ and $w$ are literals) gives rise to two edges $\overline{v} \longrightarrow w$ and $\overline{w} \longrightarrow v$. Each singleton clause $v$ gives rise to the edge $\overline{v} \longrightarrow v$. Note that the map that exchanges each pair of complementary literals and reverses the direction of all edges is an isomorphism of $D(\Phi)$.

We say that literal $v$ *implies* literal $w$ if there is a directed path in $D(\Phi)$ from $v$ to $w$. The *implies* relation is clearly transitive. The digraph $D(\Phi)$ defines a partition of $L$ into strong components, i.e., maximal subsets $V$ with the property that for any two vertices $v$ and $w$ in $V$ $v$ implies $w$ and $w$ implies $v$. Note that $V \subseteq L$ is a strong component if and only if $\overline{V}$ is a strong component. A subset $V$ of literals is said to be *initial* in $D(\Phi)$ if there is no edge entering $V$ from outside $V$, and is said to be *final* if there is no edge from a vertex in $V$ to a vertex outside of $V$. Trivially each initial set and each final set is a union of strong components. If $\Phi$ is satisfiable, we say that the literal $v$ is fixed by $\Phi$ if the value of $v$ is the same for every satisfying assignment of $\Phi$.

We state without proof the following facts, which are easy to prove and belong to the folklore about 2-CNF formulae.

**Proposition 1** *Let $\Phi$ be a 2-CNF formula on $\{x_1, x_2, \ldots, x_n\}$. Then:*

1. *An assignment $\alpha$ satisfies $\Phi$ if and only if $L(\alpha)$ is a final set in $D(\Phi)$.*
2. *If the relation "$v$ implies $w$" holds in $D(\Phi)$ then in any satisfying assignment of $\Phi$, $v = 0$ or $w = 1$.*
3. *$\Phi$ is satisfiable if and only if for each variable $x_i$, the literals $x_i$ and $\overline{x}_i$ lie in different strong components.*
4. *If $V$ is a strong component of $D$ then in any satisfying assignment of $\Phi$, either all literals in $V$ are true or all literals in $V$ are false.*
5. *If $V$ is a strong component of $D$, and $\Phi$ is satisfiable, then one of the following two situations holds: either $V$ consists entirely of fixed literals or there exist two satisfying assignments of $\Phi$ that differ precisely on the variables of $V$.*

A strong component consisting entirely of fixed literals is a *fixed component* otherwise it is an *unfixed component.*

### 2.3 Circuits

As usual, for an integer $d$, $\Sigma_d$ (resp. $\Pi_d$) denotes the class of layered unbounded fan-in boolean circuits with $d$ alternating levels of ANDs and ORs, and a single OR gate (resp. AND gate) at the top. The inputs are viewed as feeding into the first level, and the top gate is at the $d^{th}$ level. Similar to [9], we define $\Sigma_d^k$ (resp. $\Pi_d^k$) to be the class of circuits in $\Sigma_d$ (resp. $\Pi_d$) such that all gates at the first level have fan-in at most $k$. For a boolean function $f$, we define $s_d^k(f)$ to be the size (number of gates) of the smallest $\Sigma_d^k$ circuit computing $f$ (here we assume $d \geq 3$, so that $s_d^k(f)$ is well defined). We are interested in computing lower bounds on $s_3^k(f)$ for explicit functions $f$, and will obtain such bounds for the case $k = 2$.

If $C$ is a $\Sigma_3$ circuit having $M$ AND gates at level 2, we write $C^1, C^2, \ldots, C^M$ for the $\Pi_2$ subcircuits at level 2. Each of the circuits $C^i$ is equivalent to a CNF formula of the inputs. If $C$ is a $\Sigma_3^k$ circuit, then each of the $C^i$ computes a $k$-CNF formula. If $f$ is the function computed by $C$, then $f$ is the OR of the functions $f_1, f_2, \ldots, f_M$ computed by the circuits $C^1, C^2, \ldots, C^M$. Let $\kappa(f)$ be the minimum number $M$ such that $f$ can be written as an OR of $M$ 2-CNF functions. Trivially, $s_3^2(f) \geq \kappa(f)$ and since any 2-CNF on $n$ variables can be expressed as a $\Pi_2^2$ circuit with at most $4n^2$ gates we have:

**Proposition 2** *Let $f$ be a boolean function on $n$ variables. Then:*

$$\kappa(f) \leq s_3^2(f) \leq \kappa(f)4n^2.$$

So to approximate $s_3^2(f)$ it suffices to analyze $\kappa(f)$. It is useful to think of the determination of $\kappa(f)$ as a cover problem: we want to cover the subset $A = f^{-1}(1)$ of $\{0,1\}^n$ by subsets of $A$ each of which can be expressed as the accepting set of a 2-CNF.

As an example, consider $s_3^2(f)$ for symmetric boolean functions. Consider first the slice functions: $S_k^n$ is the $n$ variable function that is one on inputs of weight $k$ (where the weight is the number of 1's in the input). It is easy to see that $\kappa(S_k^n) = \kappa(S_{n-k}^n)$ (given a circuit for $S_k^n$, replace all literals by their complements to get one for $S_{n-k}^n$), so assume $k \leq n/2$.

We want to cover the set of assignments of weight $k$ by 2-CNFs. We can only use 2-CNFs whose accepting set consists of inputs of weight $k$.

To get an upper bound, consider the set $\mathcal{G}$ of boolean formula that can be constructed in the following way. Partition the variables arbitrarily into $k+1$ sets $V, P_1, P_2, \ldots, P_k$ where each of the $P_i$ is of size 2 and $V$ is of size $n-2k$. Define the formula $\Phi$ having clauses $\bar{x_i}$ for $x_i \in V$ and clauses $x_i \vee x_j$ and $\bar{x_i} \vee \bar{x_j}$ for each $P_r = \{x_i, x_j\}$. Then the assignment $\alpha$ satisfies $\Phi$ if and only if $\alpha$ is 0 on all variables in $V$ and for each $P_i$, one variable is set to 1 and the other is set to 0. Hence each formula in $\mathcal{G}$ accepts only inputs of weight $k$.

We claim that there exists a set of such formulae having size $M \leq n2^{0.59n}$ that cover all assignments of weight $k$. Let $\alpha$ be an assignment of weight $k$. If $\Phi$ is a formula chosen uniformly at random from $\mathcal{G}$ then the probability that $\Phi$ covers $\alpha$, i.e., that $\alpha$ satisfies $\Phi$, is the probability that the $k$ variables set to 1 by $\alpha$ belong to $k$ distinct pairs $P_i$, which is easily shown to be $2^k/\binom{n}{k}$, Therefore, if we choose $\Phi_1, \Phi_2, \ldots, \Phi_M$ independently and uniformly from $\mathcal{G}$, the probability that none of them cover $\alpha$ is $(1-2^k/\binom{n}{k})^M \leq e^{-M2^k/\binom{n}{k}}$. Since there are $\binom{n}{k}$ such assignments, the probability that there is an assignment that is uncovered is at most

$$\binom{n}{k} e^{-M2^k/\binom{n}{k}}$$

Thus, if $M$ is at least $m(k) = 2^{-k}\binom{n}{k}\ln\binom{n}{k}$, then this probability is less than one, and some choice of $\Phi_1, \ldots, \Phi_M$ is a cover. $m(k)$ is maximized when $k = \frac{n}{3} - c$ for some constant $c$, and is at most $2^{0.59n}$.

Now any symmetric boolean function is the OR of at most $n$ slice functions and so if $f$ is a symmetric boolean function then $\kappa(f) \leq n^2 2^{0.59n}$ and $s_3^2(f) \leq 2^{0.59n + O(\log n)}$.

Our goal in this paper is to exhibit concrete functions which require circuits of much larger size S, that is, circuits of size $S$ such that $\frac{\log_2 S}{n}$ approaches 1.

## 3 Projections

In this section, we prove that if a 2–CNF formula accepts many inputs, then it must accept a projection of large dimension.

A projection for variable set $X$ is a subset of the set of all assignments (or, equivalently, a subset of $\{0,1\}^n$), defined by equations of the form $v_i = 0$, $v_i = 1$, or $v_i = v_j$ where $v_i$ and $v_j$ are literals. Trivially, the condition $v_i = 0$ is equivalent to $\bar{v_i} = 1$ and the condition $v_i = v_j$ is equivalent to $\bar{v_i} = \bar{v_j}$. A projection is an affine subspace of $GF(2)^n$, and the dimension of a projection is its dimension as an affine subspace. A projection of dimension $d$ can be specified by $2(d+1)$ sets $(A_0, B_0, A_1, B_1, A_2, B_2, \cdots, A_d, B_d)$ where $A_i \cup B_i$ are disjoint for $i \geq 0$, $\cup_{i \geq 0}(A_i \cup B_i) = X$, and $A_i \cup B_i$ are nonempty for $i \geq 1$. The projection $P$ specified by such a sequence of sets consists of all assignments $\alpha$ which are 0 on the variables of $A_0$, 1 on the variables of $B_0$, and such that for each $j \geq 1$, all the variables in $A_j$ are equal and all the variables in $B_j$ are equal to the negation of the variables in $A_j$. For a subset $S$ of assignments, we define $\pi(S)$ to be the dimension of the largest projection $P$ such that $P \subseteq S$. If $f$ is a boolean function, we write $\pi(f)$ for $\pi(f^{-1}(1))$.

The following result gives a lower bound on the number of gates at level 2, $\kappa(f)$, (and hence on the circuit size $s_3^2(f)$) in terms of $\pi(f)$:

**Theorem 1** *Let $f$ be a boolean function on $n$ variables and suppose that $\pi(f) \leq d$. Then*

$$s_3^2(f) \geq \kappa(f) \geq \frac{|f^{-1}(1)|}{\sum_{i=0}^{d}\binom{n}{i}}.$$

Theorem 1 is an immediate consequence of the following:

**Lemma 1** *If $\Phi$ is a 2-CNF formula on $n$ variables then $\Phi$ accepts at most $\sum_{i=0}^{\pi(\Phi)}\binom{n}{i}$ assignments.*

Theorem 1 follows since if $f$ is covered by 2-CNFs $\Phi_1, \Phi_2, \ldots, \Phi_M$, then $\pi(\Phi_i) \leq \pi(f)$, and so the lemma implies that each $\Phi_i$ accepts at most $\sum_{i=0}^{d}\binom{n}{i}$ assignments and hence $M$ is at least $|f^{-1}(1)|/(\sum_{i=0}^{d}\binom{n}{d})$.

So it suffices to prove the lemma. We begin with a definition. A set $Y = \{x_{j_1}, x_{j_2}, \ldots, x_{j_k}\}$ of variables is said to be *free* with respect to the set of assignments $S$, if any assignment to the variables in $Y$ can be extended to an assignment in $S$, i.e., for any assignment $\beta$ to the variables in $Y$, there exists $\alpha \in S$ such that $\alpha(x_{j_i}) = \beta(x_{j_i})$ for $i \in [k]$. Define $\phi(S)$ to be the size of the largest set of free variables with respect to $S$.

If $P$ is a projection of dimension $d$, and $V = \{x_{j_1}, x_{j_2}, \ldots, x_{j_d}\}$ is a set of representatives from the non-constant classes of $P$, then it is easy to see that $V$ is free with respect to $P$, and hence also free with respect to any superset of $P$. Hence we have:

**Proposition 3** *For any set $S \subseteq \{0,1\}^n$, $\phi(S) \geq \pi(S)$.*

In general $\phi(S)$ can be much larger than $\pi(S)$, but the following lemma shows that if $S$ is the set of inputs accepted by a 2-CNF formula then equality holds:

**Lemma 2** *Let $S \subseteq \{0,1\}^n$ be the set of inputs accepted by a 2-CNF formula $\Phi$. Then if $V$ is a set of variables that is free with respect to $S$ then there exists a projection $P \subseteq S$ for which the variables in $V$ are in distinct non-constant classes. Hence $\pi(S) = \phi(S)$.*

**Proof:**

We will call a literal *free* if the associated variable is free and non-free otherwise. Consider the implication digraph $D(\Phi)$. By definition, no free literal can imply another. Since the implies relation is transitive, we have that for each non-free literal $y$ exactly one of the following holds:

1. $y$ is in the same strong component as some free literal
2. $y$ is implied by one or more free literals.
3. $y$ implies one or more free literals.
4. $y$ neither implies nor is implied by a free literal.

We now construct a projection that satisfies all the clauses. Let $\alpha$ be any satisfying assignment. For each variable $x_i$ of type (4), assign it according to $\alpha$. For each variable of type (2), set it equal to 1. For each variable of type (3), set it equal to 0. Each remaining literal is set equal to the free literal to whose strong component it belongs. It is easily verified that every assignment consistent with this projection satisfies the formula $\Phi$. □

To complete the proof of the theorem 1, observe that $\phi(S)$ is the VC-dimension [16] of $S$ when considered as a family of subsets of an $n$ element set. Lemma 1 now follows from $\phi(S) = \pi(S)$ and the following standard result from the theory of VC-dimension (see, e.g., [11]):

**Lemma 3** *If $A$ is a family of subsets of an $n$ element set, and $A$ has VC-dimension at most $d$ then:*

$$|A| \leq \sum_{i=0}^{d} \binom{n}{i}.$$

## 4 Constructing Hard Functions

In this section, we will exhibit an explicit function in log-space uniform $NC^1$ for which $s_3^2(f) = 2^{n-o(n)}$. The main idea is to consider the set $H_2(X)$ of multilinear $GF(2)$ polynomials in the variable set $X$ that are homogeneous of degree 2. Each such polynomial is specified by a function $a$ defined on the set $E(X)$ of edges of the complete graph on $\{1, 2, \ldots, |X|\}$, where, for $e = \{i, j\}$, $a_e \in \{0, 1\}$ is the coefficient of $x_i x_j$ in the polynomial. First we will prove:

**Lemma 4** *Let $\epsilon > 0$ and $X$ be sufficiently large (depending on $\epsilon$). If $f$ is a polynomial chosen uniformly at random from $H_2(X)$ then the probability that $\pi(f) \geq |X|^{1/2+\epsilon}$ is strictly less than 1.*

Now, this fact, Theorem 1, and the easily proved and well known fact that a nonzero degree 2 polynomial over $GF(2)$ is 1 on at least $2^{|X|-2}$ inputs implies that for $|X|$ sufficiently large, there is a degree 2 $GF(2)$ polynomial $f$ for which $s_3^2(f) \geq \kappa(f) \geq 2^{|X|-|X|^{1/2+\epsilon}\log_2|X|}$. In fact, the proof of Lemma 4 shows that for sufficiently large $X$, almost all functions in $H_2(X)$ satisfy this inequality. The problem, as usual, is to give a uniform construction of such polynomials, which we don't know how to do. Instead we proceed as follows. Lemma 4 can be strengthened to show that one can get good upper bounds on $\pi(f)$ if $f$ is chosen from a $k$-wise independent distribution.

**Lemma 5** *Let $\epsilon > 0$ and $k$ be sufficiently large (depending on $\epsilon$). Let $X$ be a set of size at least $k$ and let $D$ be a probability distribution over $H_2(X)$ such that for any set $\{e_1, e_2, \ldots, e_k\}$ of $k$ edges in $E(X)$, we have that $a_{e_1}, a_{e_2}, \ldots, a_{e_k}$ are independent and unbiased. If $f$ is a polynomial chosen from $H_2(X)$ according to $D$ then the probability that $\pi(f) \geq |X|/k^{1/2-\epsilon}$ is strictly less than 1.*

It is well known (see e.g., [2] ) that for any integers $k \leq m$, there is an explicitly constructible set $S(m, k)$ of vectors in $\{0, 1\}^m$ having size at most $(2m)^{\lceil (k+1)/2 \rceil}$ such that for a vector $v$ chosen uniformly at random from $S(m, k)$, the coordinates of $v$ are $k$-wise independent random variables. Furthermore, using the construction in [2], the basis vectors which generate this set can be computed in logarithmic space. Noting that each function in $H_2(X)$ is specified by a vector in $\{0,1\}^m$ with $m = \binom{|X|}{2}$, we define $H_2(X, k)$ to be the subset of $H_2(X)$ consisting of those polynomials whose coefficient vector is chosen from $S(m, k)$. Each function in $H_2(X, k)$ can be explicitly indexed by a sequence of at most $b(X, k) = (k + 2)(\log |X|)$ bits.

Again, by theorem 1 and lemma 5 we have:

**Corollary 1** *Given $\epsilon > 0$ and $k$ sufficiently large, then for $|X| \geq k$ there exists a function $g$ in $H_2(X, k)$ for which*

$$s_3^2(g) \geq 2^{|X|(1-k^{-1/2+\epsilon}\log_2|X|)}$$

Now define the function $f_{X,k}$ on variable set $X \cup Y$ where $|Y| = b(X, k)$ as follows: for an assignment $\alpha$ of $X$ and $\beta$ of $Y$, the assignment $\beta$ of the variables in $Y$ indexes a function $g_\beta$ in $H_2(X, k)$, and $f_{X,k}(\alpha, \beta) = g_\beta(\alpha)$. Trivially, $s_3^2(f_{X,k}) \geq s_3^2(g)$ for any $g \in H_2(X, k)$. By the above corollary, for $k$ sufficiently large, $s_3^2(f_{X,k}) \geq 2^{|X|(1-k^{-1/2+\epsilon}\log_2|X|)}$. For fixed $\delta > 0$ and all sufficiently large $n$, we define the Boolean function $f_n$ on $n$ variables as follows. View the first $n - n^{2/3+\delta/2}$ variables as $X$ and the last $n^{2/3+\delta/2}$ variables as $Y$. $Y$ is large enough to specify a function in $H_2(X, k)$ for $k = n^{2/3}$. Then we have:

**Corollary 2** *For any $\delta > 0$, $f_n$ is logspace-uniformly computable in $NC^1$ and*

$$s_3^2(f_n) \geq 2^{n-n^{2/3+\delta}}.$$

The fact that $f_n$ is logspace-uniformly computable in $NC^1$ follows from the observation that the basis for the space of vectors with limited independence can be generated by a logspace machine. So it remains to prove lemma 4 and its generalization lemma 5.

**Proof of Lemma 4**. We need to upper bound the probability that a random function in $H_2(X)$ has a large projection on which it is 1. Fix an integer $d$ and let $P$ be a projection of dimension $d$. As described in section 3, we can represent $P$ by a sequence $(A_0, B_0, A_1, B_1, \ldots, A_d, B_d)$ of subsets of the variables. If $f$ is a polynomial in $H_2(X)$ and $G_f$ is the corresponding graph defined on $X$, let $f_P$ be the function on variables $y_1, y_2, \ldots, y_d$ obtained from $f$ by substituting 1 for each variable in $A_0$, 0 for each variable in $B_0$, and for $i \in [d]$ substituting $y_i$ for each variable in $A_i$ and $1 + y_i$ for each variable in $B_i$. Then $f$ is constant on $P$ if and only if $f_P$ is a constant polynomial. We upper bound the probability that $f_P$ is constant by upper bounding the probability that its degree is at most 1. Let $b_{i,j}$ be the coefficient of $y_i y_j$ in $f_P$. Then the event that $f_P$ has degree at most 1 is the event that all of $b_{i,j}$ are 0. Now $b_{i,j}$ is just the number (mod 2) of edges in $G_f$ between the sets $A_i \cup B_i$ and $A_j \cup B_j$. For a randomly chosen function in $H_2(X)$, $b_{i,j}$ is uniformly random and the $b_{i,j}$ are mutually independent. Hence the probability that $f_P$ has degree at most 1 is $2^{-\binom{d}{2}}$. Note that the event that $f_P$ has degree at most 1 only depends on the sequence of sets $(A_0 \cup B_0, A_1 \cup B_1, A_2 \cup B_2, \ldots, A_d \cup B_d)$ representing the projection. Since the number of ways to choose such a sequence is at most $(d+1)^n$ we can upper bound the probability that there exists a projection such that $f_P$ has degree at most 1 by $2^{-\binom{d}{2}}(d+1)^{|X|}$. For $d = |X|^{1/2+\epsilon}$, this is less than 1.

**Proof of Lemma 5**. To show that the probability that $\pi(f) \geq |X|/k^{1/2-\epsilon}$ is strictly less than 1, we need the following:

**Claim.** Let $f$ be a boolean function on variable set $X$ and $h, d \leq |X|$ be positive integers. If there is a projection of dimension $d$ on which $f$ is constant than there is a projection of dimension at least $dh/|X| - 1$ on which $f$ is constant and such that the number of unfixed variables is at most $h$.

To see the claim, consider a projection $\phi$ of dimension $d$ on which $f$ is constant, and let $P = (A_0, B_0, A_1, B_1, \ldots, A_d, B_d)$ be a sequence of sets representing the projection, with the parts ordered so that $|A_1 \cup B_1| \leq |A_2 \cup B_2| \leq \ldots |A_d \cup B_d|$. Let $j$ be the largest integer such that the number of the variables in the smallest $j$ parts is at most $h$. Consider the projection $\phi'$ obtained from $\phi$ by fixing, for each $i > j$, all the variables in $A_i$ to 1 and all variables in $B_i$ to 0. Then $\phi'$ has at most $h$ unfixed variables. Also it is a subset of $\phi$, and so $f$ is fixed on $\phi'$. It can be easily be seen that $j$, the dimension of $\phi'$, is at least $h/(|X|/d) - 1$ since $|X|/d$ is the average part size.

Returning to the proof of the lemma, let $D$ be a $k$-wise independent distribution on $H_2(X)$ and suppose $f$ is selected according to $D$. By the claim, to upper bound the probability that $f$ has a

projection of dimension $d$ it suffices to upper bound the probability that it has a projection with $h = \lceil k^{1/2} \rceil$ unfixed variables of dimension at least $d' = dh/|X| - 1$. Consider a projection $P$ with $h$ unfixed variables, Note that for such a projection, the number of pairs of unfixed variables is $\binom{h}{2} \leq k$. Hence, the set of random variables $a_{i,j}$ where $x_i, x_j$ are unfixed are mutually independent. Thus we can now proceed exactly as in the previous lemma and say that the probability that $f_P$ has degree at most 1 is at most $2^{-\binom{d'}{2}}$. As before we note that the event that $f_P$ has degree at most 1 only depends on the $d'$ parts $\{A_1 \cup B_1, \ldots, A_{d'} \cup B_{d'}\}$. Now we only need to count the $d'$-part partitions with at most $h$ unfixed variables, and there are at most $(|X|d')^h$ of these and so the probability that for $f$ chosen according to $D$, there exists a dimension $d'$ projection $P$ with $h$ unfixed variables on which $f$ is constant is at most $(|X|d')^h 2^{-\binom{d'}{2}}$. For $d' \geq |X|/k^{1/2-\epsilon}$ and $k$ sufficiently large, this probability is less than 1.

□

## 5 Constructing Hard Functions Using Codes

In this section, we will give a simpler construction of a hard function using BCH codes. Observe that in the previous section we constructed a function with the property that if $o(n)$ bits are instantiated in a certain fashion, then we will get a function which is not constant under large projections. In particular, any small circuit computing the hard function cannot be constant under large projections. If we show that small circuits always have large constant projections, then we get the desired contradiction. However, for our case of $\Sigma_3^2$ circuits, we proved that if the size is small, then the circuit outputs 1 on a large dimensional projection. We will use this stronger property to construct a hard function in the following way. We will construct a function $f$ such that $f^{-1}(1)$ does not contain any large dimensional projections, though $f^{-1}(0)$ may contain large dimensional projections. Hence, any $\Sigma_3^2$ circuit must have large size to compute $f$. We will then construct a function $g$ on $n+1$ inputs to index $f$ and $\overline{f}$. By instantiating the index bit appropriately, we can show that both $\Sigma_3^2$ and $\Pi_3^2$ circuits need a large number of gates to compute $g$.

We start with a simple observation: If a set $A$ contains a $d$–dimensional projection, then the set $A$ has two points at a Hamming distance of at most $n/d$: If $P$ is a $d$–dimensional projection, then it must contain a part with at most $n/d$ variables and by fixing all the variables outside the part consistent with the projection we get two points which are at a distance of at most $n/d$. If $A$ is a set of codewords for a code with rate $r$ and distance $\delta$, then $A$ has size $2^{rn}$ and cannot contain a projection of dimension larger than $n/\delta$. We can use constructions of linear codes to come up with 'dense' sets with no large projections [15]. For example, one can construct binary BCH codes with codeword length $n$, dimension $n - 1 - t\log n$ and distance $2t+1$. Let $f_t$ be the Boolean function which is 1 on the codewords of a BCH code with dimension $n - 1 - t\log n$. $f_t$ is not identically 1 on any projection of dimension larger than $n/(2t+1)$. On the other hand, by theorem 1, any $\Sigma_3^2$ circuit computing $f_t$ in size $S$ must accept a projection of dimension at least $\log(|f_t^{-1}(1)|/S)/\log n$. Hence, by taking $t = \sqrt{n/2}$, it follows that $S$ must be at least $\Omega(2^{n-\sqrt{2n}\log n})$.

## 6 Conclusions and Open Problems

The obvious question that is suggested by this work is whether a large set accepted by a $k$–CNF ($k > 2$) must necessarily contain a projection of large dimension. However, it can be shown that there are large sets defined by even linear size 3–CNF which can only contain projections of dimension bounded by a constant. This follows from the existence of sparse parity check matrices which define codes with linear distance and constant rate [7, 13]. It can even be shown that there are sparse parity check matrices which have at most 3 1's in each row. The set of codewords defined by such a parity check matrix can be accepted by a 3–CNF and such 3-CNF cannot contain a projection whose size large is larger than some fixed constant. This implies that using the idea of projections to prove nonlinear lower bounds on circuit size using Valiant's reduction to depth 3 unbounded fan-in circuits cannot work.

However, it may still be possible to apply the technique directly to linear size and logarithmic depth circuits. In particular, we do not know the answer to the following question: Let $S \subseteq \{0,1\}^n$ be recognizable by a linear size and logarithmic depth (or just even linear size) circuit. Does $S$ or $S$ contain a projection of dimension $\Omega(n^\epsilon)$ for some $\epsilon > 3/4$? If we have an affirmative answer to the question, then it follows that the hard function constructed in section 4 would require nonlinear circuit size. The codes discussed in section 5 would not suffice since their complements contain large dimensional projections.

One can also consider more general types of nice subsets of $\{0,1\}^n$. For instance: consider the set of subsets of $\{0,1\}^n$ that are affine subspaces. Is it true that for constant $k$, every $\Sigma_3^k$ circuit is constant on an affine subspace of dimension $\Omega(n^\epsilon)$ for some $\epsilon$ (or even $\Omega(n)$)? Can one construct an explicit function which has no such subspace? A counting argument shows that almost all homogeneous multilinear polynomials of degree 3 over $GF(2)$ have the property that they are not constant on any affine subspace of dimension more than $\Omega(n^{2/3})$, but we don't yet know how to make this explicit.

**Acknowledgments:** The authors would like to thank Johan Håstad for pointing out an error in the example in section 2. The authors would also like to thank Russell Impagliazzo, Pavel Pudlák and Jiří Sgall for useful discussions.

## References

[1] Ajtái M., $\Sigma_1^1$-Formulae on Finite Structures, *Annals of Pure and Applied Logic*, **24**, pp. 1–48, 1983.

[2] Alon, N., Spencer, J., and Erdös, P., (1992), "The Probabilistic Method", John Wiley & Sons, Inc.

[3] Boppana, R. and Sipser, M. (1990), The Complexity of Finite Functions, *in* "The Handbook of Theoretical Computer science", Vol. A, Elsevier Science Publishers.

[4] Chandra, A. and Stockmeyer, L. (1984), Constant Depth Reducibility, *SIAM Journal on Computing*, **13**, pp. 423–439.

[5] Dunne, P.E. (1988), "The Complexity of Boolean Functions", Academic Press.

[6] Furst, M., Saxe, J.B. and Sipser, M. (1984), Parity, Circuits, and the Polynomial Time Hierarchy, *Mathematical Systems Theory*, **17**, pp. 13–28.

[7] Gallager, R.G., *Low Density Parity–Check Codes*, MIT Press, 1963.

[8] Håstad, J., (1986), Almost Optimal Lower Bounds for Small Depth Circuits, *in* " Proceedings of the 18th ACM Symposium on Theory of Computing", pp. 6–20.

[9] Håstad, J., Jukna, S., and Pudlák, P., (1993), Top–Down Lower Bounds for Depth 3 Circuits, "Proceedings of the 34th Annual IEEE Symposium on Foundations of Computer Science", pp. 124–129.

[10] Razborov, A.A. (1986), Lower Bounds on the Size of Bounded Depth Networks over a Complete Basis with Logical Addition, *Mathematische Zametki* **41** pp. 598–607 (in Russian). English Translation in*Mathematical Notes of the Academy of Sciences of the USSR* **41**, pp. 333–338.

[11] Sauer, N. (1972), On the Density of Families of Sets, *Journal of Combinatorial Theory, series A*, vol. **13**, pp. 145–147.

[12] Smolensky, R. (1987), Algebraic Methods in the Theory of Lower Bounds for Boolean Circuit Complexity *in* "Proceedings of the 19th ACM symposium on Theory of Computing", pp. 77–82.

[13] Sipser, M., and Spielman, D., Expander Codes, *in* 35th IEEE Symposium on Foundations of Computer Science, pp 566–576, 1994.

[14] Valiant, L.G., (1977), Graph–theoretic arguments in low–level complexity, in *Proceedings of the 6th Symposium on Mathematical Foundations of Computer Science*, Springer–Verlag, Lecture Notes in Computer Science, vol. ??, pp. 162–176.

[15] Van Lint, J.H., Introduction to Coding Theory, 2nd Edition, Springer–Verlag, 1992.

[16] Vapnik, V.N., and Chervonenkis, A. Ya, (1971), On the Uniform Convergence of Relative Frequencies of Events to their Probabilities, *Theory of Probability Applications*, vol. **16**, pp. 264–280.

[17] Yao, A. C–C. (1985), Separating the Polynomial Hierarchy by Oracles, *in* "Proceedings of the 31st Annual IEEE Symposium on Foundations of Computer Science", pp. 1–10.