A Switching Lemma for Small Restrictions and Lower Bounds for k-DNF Resolution (extended abstract)

Nathan Segerlind* Sam Buss† Russell Impagliazzo‡

Abstract

We prove a new switching lemma that works for restrictions that set only a small fraction of the variables and is applicable to DNFs with small conjunctions. We use this to prove lower bounds for the Res(k) propositional proof system, an extension of resolution which works with k-DNFs instead of clauses. We also obtain an exponential separation between depth d circuits of bottom fan-in k and depth d circuits of bottom fan-in $k+1$.

Our results for Res(k) are:

1. *The $2n$ to n weak pigeonhole principle requires exponential size to refute in Res(k), for $k \leq \sqrt{\log n / \log\log n}$.*
2. *For each constant k, there exists a constant $w > k$ so that random w-CNFs require exponential size to refute in Res(k).*
3. *For each constant k, there are sets of clauses which have polynomial size Res$(k+1)$ refutations, but which require exponential size Res(k) refutations.*

1 Introduction

This is an extended abstract. For a full version of the paper, please visit the web page http://www.cs.ucsd.edu/~nsegerli

This paper studies the complexity of $\mathrm{Res}(k)$, a propositional refutation system that extends resolution by allowing k-DNFs instead of clauses [24]. The complexity of propositional proof systems has close connections to open problems in computational and circuit complexity (see [15, 23, 27, 7]), as well as implications for the run times of satisfiability algorithms and automated theorem provers. Resolution is one of the most studied proof systems, and is used as the basis for many satisfiability algorithms. Back-tracking algorithms such as DPLL that branch on a single variable provide tree-like resolution refutations on unsatisfiable formulas. General resolution proofs correspond to adding a limited form of memoization (previously refuted subproblems are saved for reuse rather than refuted again) to DPLL. $\mathrm{Res}(k)$ corresponds to algorithms that branch on more general conditions: the value of any function of up to k variables.

The $\mathrm{Res}(k)$ systems are also interesting as intermediates between previously studied proof systems. Resolution can be thought of as $\mathrm{Res}(1)$ and depth-two Frege can be thought of as $\mathrm{Res}(n)$ (where n is the number of variables). In this way, the $\mathrm{Res}(k)$ systems provide a transition between resolution and depth-two Frege. Moreover, statements provable in the theory $T^2_2(\alpha)$ (a fragment of Peano's arithmetic that allows induction only on Σ^b_2 predicates) correspond to propositional statements with quasi-polynomial size $\mathrm{Res}(\mathrm{polylog}(n))$ refutations [24]. T^2_2 is the weakest fragment of Peano's arithmetic known to be able to use counting arguments such as the weak pigeonhole principle [25]. On the other hand, these counting tautologies are known to be hard for resolution. Thus, there must be a critical range for k between 1 and $\mathrm{polylog}(n)$ where these arguments become possible in sub-exponential size. More generally, we can ask: when does increasing k give the $\mathrm{Res}(k)$ system more power? Is there a reason to want to branch on more complex functions in satisfiability algorithms? Does such branching give algorithms better performance in the average case?

We give partial answers to all of these questions. In particular we prove:

1. The $2n$ to n weak pigeonhole principle requires size $2^{\Omega(n^\epsilon)}$ to refute in $\mathrm{Res}(\sqrt{\log n/\log\log n})$. So

*Supported in part by NSF grant DMS-0100589 and CCR-0098197.

†Supported in part by NSF grant DMS-0100589.

‡Supported in part by NSF grant CCR-0098197 and USA-Israel BSF Grant 97-00188.

branching on a super-constant number of variables is necessary for counting arguments such as the weak pigeonhole principle.

2. For each k, there exists a constant $w > k$ so that random w-CNFs require size $2^{\Omega(n^\epsilon)}$ to refute in $\mathrm{Res}(k)$. Thus, extending DPLL algorithms to branch on multiple (but a constant number of) variables, will not make run times sub-exponential on average.

3. $\mathrm{Res}(k+1)$ has exponential speedup over $\mathrm{Res}(k)$: there are sets of clauses which have polynomial size $\mathrm{Res}(k+1)$ refutations, but which require size $2^{\Omega(n^\epsilon)}$ to refute in $\mathrm{Res}(k)$. Therefore, increasing the complexity of branching conditions can give an exponential speed up.

Our lower bounds are proved using a new kind of switching lemma. A switching lemma provides conditions under which a OR of small ANDs can be rewritten as an AND of small ORs after the application of a random restriction [1, 18, 21, 4]. Our switching lemma differs from previous switching lemmas in that the random restriction is allowed to set a small number of the variables, even as few as $n^{1-\epsilon}$ out of n. The trade-off is that ORs of *extremely* small ANDs are transformed into ANDs of modestly small ORs. Therefore, our switching lemma cannot be iterated to prove lower bounds for proof systems of depth more than two. However, one application of our small fan-in switching lemma suffices to prove lower bounds for the $\mathrm{Res}(k)$ proof systems, because each line in such a proof has depth two and fan-in k.

Our switching lemma also gives an exponential separation between depth d circuits with bottom fan-in k from depth d circuits with bottom fan-in $k+1$ (for constant k). This refines a result of Håstad [22], who showed that for all d there exist $\epsilon, \delta > 0$ so that there are functions on n variables, computable with polynomial size, depth d circuits of bottom fan-in n^ϵ but which require exponential size to compute with depth d circuits of bottom fan-in n^δ, and later results of Cai, Chen and Håstad [13], who showed that for each constant d, there exist functions computable with polynomial size, depth d, bottom fan-in 2 circuits that require exponential size to compute with depth d circuits with bottom fan-in 1, and that for each constant k, there exists a function of n variables computable by depth d circuits of polynomial size and bottom fan-in $O(\log n)$ that requires exponential size to compute with depth d circuits of bottom fan-in k.

Because resolution may be viewed as $\mathrm{Res}(1)$, our results for $\mathrm{Res}(k)$ generalize known results for resolution. The weak pigeonhole principle (for any number of pigeons) is known to require an exponential number of steps to refute in resolution [31, 20, 32, 12, 6, 16, 26, 28, 29], and we generalize these lower bounds for the case of the cn to n pigeonhole principle. Resolution refutations of randomly chosen sets of clauses are also known to require exponential size [14, 6, 5, 9]. We extend these results to general $\mathrm{Res}(k)$ systems, although as k increases, so does the width of the random CNFs for which our lower bounds apply.

Our work also extends previous research on the $\mathrm{Res}(k)$ system. The complexity of $\mathrm{Res}(k)$ refutations was first studied by Krajíček [24], who was motivated by the connection between $\mathrm{Res}(\mathrm{polylog}(n))$ and the provability of combinatorial statements in the arithmetic theory $T_2^2(\alpha)$. Atserias, Bonet and Esteban [3] gave exponential lower bounds for $\mathrm{Res}(2)$ refutations of the $2n$ to n weak pigeonhole principle and of random 3-CNFs. They also proved a quasi-polynomial separation between $\mathrm{Res}(2)$ and resolution; this separation was later strengthened to almost-exponential by Atserias and Bonet [2]. Esteban, Galesi and Messner [17] showed that that there is an exponential separation between treelike $\mathrm{Res}(k)$ and treelike $\mathrm{Res}(k+1)$.

The lower bounds for $\mathrm{Res}(k)$ refutations of the weak pigeonhole principle given by Atserias, Bonet and Esteban [3] apply only for $k = 2$; our lower bound works for non-constant k, up to $\sqrt{\log n / \log \log n}$. On the other hand, Maciel, Pitassi and Woods [25] give quasi-polynomial size refutations in $\mathrm{Res}(\mathrm{polylog}(n))$. Therefore, among depth two, quasi-polynomial size refutations of the weak pigeonhole principle, the refutations the refutations of Maciel, Pitassi and Woods have bottom fan-in that is almost optimal.

Our lower bounds for $\mathrm{Res}(k)$ refutations of random w-CNFs are the first such lower bounds for $\mathrm{Res}(k)$ with $k \geq 3$. Atserias, Bonet and Esteban [3] gave exponential lower bounds for random 3-CNFs in $\mathrm{Res}(2)$. We extend these results to $\mathrm{Res}(k)$, although the width w increases with k (it is $4k^2+2$). At present, the $\mathrm{Res}(k)$ systems are the strongest fragments of bounded-depth Frege systems for which we know there are super-polynomial size lower bounds for refutations of random sets of clauses.

The separation between $\mathrm{Res}(k+1)$ from $\mathrm{Res}(k)$ is the first for $k \geq 3$. Earlier work of Atserias and Bonet [2] gave a $2^{O\left(2^{\log^\epsilon n}\right)}$ separation of $\mathrm{Res}(2)$ from $\mathrm{Res}(1)$, and our result improves this to $2^{O(n^\epsilon)}$.

2 Definitions and Conventions

We will use the notation $[k] = \{i \mid 1 \leq i \leq k\}$.

A *literal* is a variable or its negation. A *term* is a constant 0 or 1 or a conjunction of literals. Our convention is that a term is specified as a set of literals, with 1 corresponding to the empty set and 0 to any literal and its negation. We say that a term T contains a literal l if $l \in T$, and that a term T contains a variable x if either $x \in T$ or $\neg x \in T$. We will often identify literals with terms of size one, and will write l instead of $\{l\}$. A *DNF* is a disjunction of terms, specified as a set of terms. A k*-DNF* is a DNF whose terms are each of size at most k. A *clause* is a 1-DNF, i.e. a disjunction of literals. The width of a clause C, written $w(C)$, is the number of literals appearing in C. The width of a set of clauses is the maximum width of any clause in the set. A *CNF* is a conjunction of clauses, specified as a set of clauses. A k*-CNF* is a CNF whose clauses are each of width at most k. Two terms t and t' are *consistent* if there is no literal l with $l \in t$ and $\neg l \in t'$.

A *restriction* ρ is a map from a set of variables to $\{0, 1, *\}$. For a formula F, *the restriction of* F *by* ρ, $F \restriction_\rho$ is defined as usual, simplifying only when a subexpression has become explicitly constant. For any restriction ρ, let $D(\rho)$ denote the set of variables to which ρ assigns the value 0 or 1.

Resolution is a refutation system for propositional logic. The input to a resolution refutation is a set $\mathcal{C}$ of initial clauses; a resolution refutation consists of a derivation of the empty clause from clauses in $\mathcal{C}$ using only the resolution inference: $\frac{A \vee x \quad \neg x \vee B}{A \vee B}$. Every line in a resolution refutation is a clause, i.e., a 1-DNF.

Definition 2.1 *Res*(k) *is the refutation system whose lines are* k*-DNFs and whose inference rules are given below (*A*,* B *are* k*-DNFs,* $1 \leq j \leq k$*, and* $l, l_1, \ldots, l_j$ *are literals): Subsumption,* $\frac{A}{A \vee l}$*, AND-introduction,* $\frac{A \vee l_1 \cdots A \vee l_j}{A \vee \bigwedge_{i=1}^{j} l_i}$*, Cut,* $\frac{A \vee \bigwedge_{i=1}^{j} l_i \quad B \vee \bigvee_{i=1}^{j} \neg l_i}{A \vee B}$*, AND-elimination,* $\frac{A \vee \bigwedge_{i=1}^{j} l_i}{A \vee l_i}$*.*

Let $\mathcal{C}$ *be a set of* k*-DNFs. A* Res(k) derivation from $\mathcal{C}$ *is a sequence of* k*-DNFs* $F_1, \ldots, F_m$ *so that each* F_i *either belongs to* $\mathcal{C}$ *or follows from the preceding lines by an application of one of the inference rules. For a set of* k*-DNFs* $\mathcal{C}$*, a* Res(k) refutation of $\mathcal{C}$ *is a derivation from* $\mathcal{C}$ *whose final line is the empty clause.* $w_R(\mathcal{C})$ *denotes the minimum width of a resolution refutation of* $\mathcal{C}$*, and* $s_k(\mathcal{C})$ *denotes the minimum size of a Res*(k) *refutation of* $\mathcal{C}$*. If* $\mathcal{C}$ *is satisfiable, then* $\mathcal{C}$ *has no refutation and we use the convention that* $w_R(\mathcal{C})$ *and* $s_k(\mathcal{C})$ *are* ∞*.*

We do not use the exact definition of the $Res(k)$ system in our arguments; the main property we use is *strong soundness*: if F is inferred from $F_1, \ldots, F_j$, and $t_1, \ldots, t_j$ are mutually consistent terms of $F_1, \ldots, F_j$ respectively, then there is a term t of F implied by $\wedge_{i=1}^{j} t_i$. In other words, any reason why $F_1, \ldots, F_k$ are true implies a reason why F is true.

Lemma 1 *$Res(k)$ is strongly sound.*

3 The Switching Lemma

A switching lemma is a guarantee that after the application of a randomly chosen restriction, a disjunction of small ANDs can be represented by a conjunction of small ORs, thus "switching" an OR into an AND. We use a slightly stronger variation: after the application of a random restriction, a k-DNF can be "strongly represented" by a short decision tree.

Definition 3.1 *A* decision tree *is a rooted binary tree in which every internal node is labeled with a variable, the edges leaving a node correspond to whether the variable is set to* 0 *or* 1*, and the leaves are labeled with either* 0 *or* 1*. Every path from the root to a leaf may be viewed as a partial assignment. For a decision tree* T *and* $v \in \{0, 1\}$*, we write the set of paths (partial assignments) that lead from the root to a leaf labeled* v *as* $Br_v(T)$*. We say that a decision tree* T strongly represents *a DNF* F *if for every* $\pi \in Br_0(T)$*, for all* $t \in F$*,* $t \restriction_\pi = 0$ *and for every* $\pi \in Br_1(T)$*, there exists* $t \in F$*,* $t \restriction_\pi = 1$*. The representation height of* F*,* $h(F)$*, is the minimum height of a decision tree strongly representing* F*.*

Notice that the function computed by a decision tree of height h can also be computed by both an h-CNF and an h-DNF.

Our switching lemma will exploit a trade-off based on the minimum size of a set of variables that meets each term of a k-DNF. When this quantity is small, we can build a decision tree by querying these variables and recursing on the $(k-1)$-DNFs created. When this quantity is large, the DNF has many disjoint terms and is likely to be satisfied by a random restriction.

Definition 3.2 *Let* F *be a DNF, and let* S *be a set of variables. If every term of* F *contains a variable from* S*, then we say that* S is a cover of F*. The* covering number of F, $c(F)$*, is the minimum cardinality of a cover of* F*.*

For example, the 3-DNF $xyz \vee \neg x \vee yw$ has covering number two.

First, we give a general condition on the distributions of partial assignments for which our switching lemma holds, namely that the distribution almost always satisfies k-DNFs with large cover number. Later, we will show that this condition holds for the distributions used in our applications.

Theorem 2 *Let $k \geq 1$, let $s_0, \ldots, s_{k-1}$ and $p_1, \ldots, p_k$ be sequences of positive numbers, and let $\mathcal{D}$ be a distribution on partial assignments so that for every $i \leq k$ and every i-DNF G, if $c(G) > s_{i-1}$, then $Pr_{\rho \in \mathcal{D}}\left[G \upharpoonright_\rho \neq 1\right] \leq p_i$. Then for every k-DNF F:*

$$Pr_{\rho \in \mathcal{D}}\left[h(F \upharpoonright_\rho) > \sum_{i=0}^{k-1} s_i\right] \leq \sum_{i=1}^{k} 2^{\left(\sum_{j=i}^{k-1} s_j\right)} p_i$$

Proof: We proceed by induction on k. First consider $k = 1$. If $c(F) \leq s_0$, then at most s_0 variables appear in F. We can construct a height $\leq s_0$ decision tree that strongly represents $F \upharpoonright_\rho$ by querying all of the variables of $F \upharpoonright_\rho$. On the other hand, if $c(F) > s_0$ then $\Pr_{\rho \in \mathcal{D}}\left[F \upharpoonright_\rho \neq 1\right] \leq p_1$. Therefore, $h(F \upharpoonright_\rho)$ is non-zero with probability at most $p_1 = p_1 2^{\sum_{j=1}^{k-1} s_j}$ (because $k = 1$).

For the induction step, assume that the theorem holds for all k-DNFs, and let F be a $(k+1)$-DNF. If $c(F) > s_k$, then by the conditions of the lemma, $\Pr_{\rho \in \mathcal{D}}\left[F \upharpoonright_\rho \neq 1\right] \leq p_{k+1}$. Because $p_{k+1} \leq \sum_{i=1}^{k+1} 2^{\sum_{j=i}^{k} s_j} p_i$, we have that $h(F \upharpoonright_\rho)$ is non-zero with probability at most $\sum_{i=1}^{k+1} 2^{\sum_{j=i}^{k} s_j} p_i$.

Consider the case when $c(F) \leq s_k$. Let S be a cover of F of size at most s_k. Let π be any assignment to the variables in S. Because each term of F contains at least one variable from S, $F \upharpoonright_\pi$ is a k-DNF, By combining the induction hypothesis with the union bound, we have that

$$\Pr_{\rho \in \mathcal{D}}\left[\exists \pi \in \{0,1\}^S \;\; h((F \upharpoonright_\pi) \upharpoonright_\rho) > \sum_{i=0}^{k-1} s_i\right]$$

$$\leq 2^{s_k}\left(\sum_{i=1}^{k} 2^{\left(\sum_{j=i}^{k-1} s_j\right)} p_i\right) < \sum_{i=1}^{k+1} 2^{\left(\sum_{j=i}^{k} s_j\right)} p_i$$

In the event that $\forall \pi \in \{0,1\}^S$, $h((F \upharpoonright_\pi) \upharpoonright_\rho) \leq \sum_{i=0}^{k-1} s_i$, we construct a decision tree for $F \upharpoonright_\rho$ as follows. First, query all variables in S unset by ρ, and then underneath each branch, β, simulate a decision tree of minimum height strongly representing $(F \upharpoonright_\beta) \upharpoonright_\rho$. Notice that for each such branch β, there is a unique assignment π to the variables of S so that π agrees with β on $S \setminus D(\rho)$ and π agrees with ρ on $D(\rho)$. In particular, $(F \upharpoonright_\beta) \upharpoonright_\rho = (F \upharpoonright_\pi) \upharpoonright_\rho$. Therefore, $h((F \upharpoonright_\beta) \upharpoonright_\rho) = h((F \upharpoonright_\pi) \upharpoonright_\rho)$, and the height of the resulting decision tree is at most $s_k + \max_{\pi \in \{0,1\}^S} h((F \upharpoonright_\pi) \upharpoonright_\rho) \leq \sum_{i=0}^{k} s_i$.

In the full paper, we show that the decision tree constructed above strongly represents $F \upharpoonright_\rho$. ■

In this abstract, we will always use the switching lemma in the following form:

Corollary 3 *Let $k \geq 1$, $d > 0, 1 \geq \delta > 0, 1 \geq \gamma > 0$, s, and let $\mathcal{D}$ be a distribution on partial assignments so that for every k-DNF G, $Pr_{\rho \in \mathcal{D}}\left[G \upharpoonright_\rho \neq 1\right] \leq d 2^{-\delta (c(G))^\gamma}$. Then for every k-DNF F:*

$$Pr_{\rho \in \mathcal{D}}\left[h(F \upharpoonright_\rho) > 2s\right] \leq dk 2^{-\delta' s^{\gamma'}}$$

where $\delta' = 2(\delta/4)^k$ and $\gamma' = \gamma^k$.

The proof of corollary 3 appears in the full version. The idea is to apply theorem 2 with $s_i = (\delta/4)^i (s^{\gamma^i})$, and $p_i = d 2^{-4 s_i}$.

3.1 Switching with Small Restrictions

In this subsection, we show that small, uniform restrictions meet the conditions of the switching lemma. Using corollary 3, k-DNFs can then be converted into decision trees – *using restrictions that set only a polynomially small fraction of the bits*. We include it here for comparison with previous switching lemmas, and later it will be used to prove the lower bound on Res(k) refutations of random CNFs. More complicated distributions will be used for our other results.

Definition 3.3 *Let $n > 0$ and $p \in [0,1]$. Define $\mathcal{D}_p$ to be the family of random restrictions which arises by assigning variables $*$ with probability $1-p$, and 0, 1 each with probability $\frac{p}{2}$.*

Lemma 4 *Let $k \geq 1$, G be an k-DNF, and ρ be chosen from $\mathcal{D}_p$. Then $Pr[G \upharpoonright_\rho \neq 1] \leq e^{-\frac{c(G) p^k}{k 2^k}}$.*

Proof: Because every covering set of G has size at least $c(G)$, there is a set of variable-disjoint terms of size at least $c(G)/k$ (such a set can be found by greedily choosing a maximal set of disjoint terms). Each of these variable-disjoint terms is satisfied with independent probability at least $(p/2)^k$. Therefore,

$\Pr_{\rho \in \mathcal{D}_p}[G \upharpoonright_\rho \neq 1] \leq \left(1 - \left(\frac{p}{2}\right)^k\right)^{\frac{c(G)}{k}} \leq e^{-\left(\frac{p}{2}\right)^k \frac{c(G)}{k}} = e^{-\frac{c(G) p^k}{k 2^k}}$ ■

Corollary 5 *Let $k \geq 1$ be given. For all $\beta > 0$, there exists $\gamma > 0$ so that for any k-DNF F, $c > 0$, $w > 0$, $p > \beta n^{-1/(ck^2)}$, $Pr_{\rho \in \mathcal{D}_p}[h(F \upharpoonright_\rho) > w] \leq 2^{-\gamma w n^{-1/c}}$.*

Proof:

Combining lemma 4 with corollary 3, with $p = \beta n^{-1/ck^2}$, $d = 1$, $\gamma = 1$, $s = w/2$ and $\delta = \frac{p^k}{k 2^k} = \frac{\beta^k n^{-1/ck}}{k 2^k}$, shows that for every k-DNF F: $\Pr_{\rho \in \mathcal{D}_p}\left[h(F \upharpoonright_\rho) > w\right] \leq k 2^{-2(w/2)(\delta^k/4^k)} =$

$k2^{-w(\beta^{k^2} n^{-1/c})/4^k k^k 2^{k^2}}$. Because k and β are fixed, we may choose the constant γ as necessary. ∎

The interesting case for corollary 5 is for c to be a small constant and for w to be at least $n^{\delta+c}$ for some $\delta > 0$.

4 An Application to Circuit Bottom Fan-in

Our first application of the switching lemma is an exponential size separation between depth d circuits of bottom fan-in k and depth d circuits of bottom fan-in $k+1$.

All circuits are organized into alternating layers of AND and OR gates, with connections appearing only between adjacent levels. NOT gates may have only variables as their inputs. The output gate is said to be at level one, the gates feeding into the output gate are said to be at level two and so forth. The depth of a circuit is the maximum depth of an AND or OR gate in the circuit. The size of a circuit is the number of AND and OR gates appearing in it. The *bottom fan-in* of a depth d circuit is the maximum number of inputs of a gate at level d. For more detail on the basics of constant depth circuits, consult the survey by Boppana and Sipser [11].

4.1 The Functions

Definition 4.1 *[30, 11] Let integers d and $m_1, \ldots m_d$ be given, and let there be variables $x_{i_1,\ldots,i_d}$ for $1 \leq i_j \leq m_j$.*

$$f_d^{m_1,\ldots m_d} = \bigwedge_{i_1 \leq m_1} \bigvee_{i_2 \leq m_2} \cdots \bigodot_{i_d \leq m_d} x_{i_1,\ldots,i_d}$$

Where $\bigodot = \bigvee$ if d is even, and $\bigodot = \bigwedge$ if d is odd.

The Sipser function *f_d^m is $f_d^{m_1,\ldots m_d}$ with $m_1 = \sqrt{m/logm}$, $m_2 = \cdots = m_{d-1} = m$ and $m_d = \sqrt{dm \log m/2}$.*

The modified Sipser function *$g_d^{m,k}$ is $f_{d+1}^{m_1,\ldots m_d,k}$, with $m_1 = \sqrt{m/logm}$, $m_2 = \cdots = m_{d-1} = m$, and $m_d = 4\sqrt{dm \log m/2}$.*

Notice that the function f_d^m depends on $m^{d-1}\sqrt{d/2}$ many variables and it can be computed by a circuit of depth d and size linear in the number of variables. However, it is impossible to reduce the bottom fan-in dramatically without increasing the size or the depth. Moreover, an OR of depth d small bottom fan-in circuits also requires exponential size to compute f_d^m.

Theorem 6 *[22] For all $d \geq 1$, there exists $\epsilon_d > 0$ so that if a depth d, bottom fan-in k circuit with an AND gate at the output and at most S gates in levels 1 through $d-1$ computes f_d^m, then either $k \geq m^{\epsilon_d}$ or $S \geq 2^{m^{\epsilon_d}}$.*

For all $d \geq 1$, there exists $\beta_d > 0$ so that if a depth $d+1$, bottom fan-in k circuit with an OR gate at the output and at most S gates in levels 1 through d computes f_d^m, then either $S \geq 2^{m^{\beta_d}}$ or $k \geq m^{\beta_d}$.

The modified Sipser function $g_d^{m,k+1}$ is used to obtain the exponential separation between depth $d+1$, bottom fan-in $k+1$ and depth $d+1$, bottom fan-in k circuits. Notice that the function $g_d^{m,k}$ has $4m^{d-1}\sqrt{d/2}$ many blocks and $4km^{d-1}\sqrt{d/2}$ many variables. Moreover, it can be computed by a circuit of depth $d+1$, bottom fan-in k and size linear in the number of variables. For each $i_1, \ldots i_d$, we say that the variables $x_{i_1,\ldots i_d,1}, \ldots x_{i_1,\ldots i_d,k}$ come from *block* $(i_1, \ldots i_d)$. The idea is that these variables occur in the same bottom level conjunction of $g_d^{m,k}$.

4.2 The Lower Bounds

We will show that depth $d+1$ circuits with bottom fan-in k require exponential size to compute $g_d^{m,k+1}$. In light of theorem 6, it suffices to consider only circuits with an AND gate at the output level. Furthermore, in this extended abstract, we will consider only the case when d is even. This ensures that all gates at depth d are OR gates. The case for odd d is dual and we simply invert the random restriction used; for more details see the full paper. Each gate at depth d computes a k-DNF, and we will apply a random restrictions which almost certainly collapse all of the k-DNFs to narrow CNFs and thus collapse the circuits to depth d circuits with small bottom fan-in. On the other hand, the random restrictions will probably leave $g_d^{m,k+1}$ containing f_d^m as a sub-function, and thus we obtain a contradiction to theorem 6.

Definition 4.2 *Let m, d and k be given. Set $m_1 = \sqrt{m/logm}$, $m_2 = \cdots = m_{d-1} = m$ and $m_d = 4\sqrt{dm \log m/2}$.*

Let $\mathcal{B}_{d,0}^{m,k+1}$ be the random distribution on partial assignments given by the following experiment: for each $i_1 \leq m_1, \ldots, i_d \leq m_d$, with independent probability probability $\frac{1}{2}$ either set $x_{i_1,\ldots,i_d,j} = *$, for all $j \in [k+1]$, or uniformly choose a 0/1 assignment to $\{x_{i_1,\ldots,i_d,j} \mid j \in [k+1]\}$ which sets at least one of the variables to 0.

The following lemma is proved in a manner similar to lemma 4. The only major difference is that we use "block disjointness", rather than variable disjointness, to obtain independence between the events of satisfying terms. Terms T_1 and T_2 are block disjoint no variable of T_1 comes from the same block as some variable of T_2. See the full version for the details of the proof.

Lemma 7 *Let $k \geq 1$ be given. There exists a constant $\gamma_k > 0$ so that for every k-DNF F:*

$$Pr_{\rho \in \mathcal{B}^{m,k+1}_{d,0}} [F \upharpoonright_\rho \neq 1] \leq 2^{-\gamma_k c(F)}$$

Lemma 8 *Let $k \geq 1$ be given. There exists a constant ϵ_k so that for all d, for all w sufficiently large with respect to k, and for every depth $d+1$, bottom fan-in k circuit C of size $S \leq 2^{\epsilon_k w}$, when by ρ is chosen from $\mathcal{B}^{m,k+1}_{d,0}$, with probability at least $3/4$, $C \upharpoonright_\rho$ is is equivalent to a depth d, bottom fan-in w circuit with at most S gates in levels 1 through $d-1$.*

Proof:

We will solve for the particular values of ϵ_k and w after going through the calculations.

For each OR gate g at depth d, we let F_g denote the k-DNF computed by the sub-circuit at g.

Suppose that there is a partial assignment $\rho \in \mathcal{B}^{m,k}_{d,0}$ so that for each depth d gate g of C, $h(F_g \upharpoonright_\rho) < w$. For each g at depth d, let T_g be the shortest decision tree representing $F_g \upharpoonright_\rho$. We can compute $C \upharpoonright_\rho$ with a depth d, bottom fan-in w circuit with at most S gates in levels 1 through $d-1$ by starting with C, replacing each level d gate g the conjunction of the negated branches of $\mathrm{Br}_0(T_g)$ and then merging these conjuncts with the AND gate at depth $d-1$ to which g sends its output.

We now show that for ρ chosen according to the distribution $\mathcal{B}^{m,k}_{d,0}$, with probability at least $3/4$, every depth d gate g of C has $h(F_g \upharpoonright_\rho) < w$.

Let g be a depth d gate of the circuit. By combining lemma 7 with corollary 3, setting $d=1, \gamma = 1$ $s = w/2$ and $\delta = \gamma_k$ shows that

$$\Pr_{\rho \in \mathcal{B}^{m,k}_{d,0}} [h(F_g) > w] \leq k 2^{-w\gamma_k^k/4^k}$$

Because there are at most $S = 2^{\epsilon_k w}$ many gates at depth d, by the union bound, there exists a gate with $h(F_g) > w$ with probability at most $2^{w(\epsilon_k - \gamma_k^k/4^k) + \log k}$. We simply take ϵ_k sufficiently small so that this probability is less than $1/4$. ∎

Theorem 9 *For all $k \geq 1$, $d \geq 1$, there exists $\epsilon_k, \epsilon_d > 0$ so that for every m sufficiently large, every size S, depth $d+1$ bottom fan-in k circuit for $g^{m,k+1}_d$ has $S \geq 2^{\epsilon_k m^{\epsilon_d}}$.*

Proof: We will have to take m sufficiently large to apply theorem 6 and lemma 8, and large enough for an application of the Chernoff bounds. Set $w = m^{\epsilon_d}$ (with ϵ_d from theorem 6) and $S = 2^{\epsilon_k w}$ (with ϵ_k from lemma 8).

Suppose, for the sake of contradiction, that C is a size S, depth d, bottom fan-in k circuit computing $g^{m,k+1}_d$.

Fix an OR gate at depth d in $g^{m,k+1}_d$. When ρ is chosen from the distribution $\mathcal{B}^{m,k+1}_{d,0}$, the expected number of blocks underneath this gate that are left unset is $2\sqrt{dm \log m/2}$. By the Chernoff bounds, with probability at most $e^{-\sqrt{dm \log m/2}/4}$ are there fewer than $\sqrt{dm \log m/2}$ blocks left unset by ρ underneath this gate.

Because there are $m^{d-3/2}/\sqrt{\log m}$ many depth d gates in $g^{m,k+1}_d$, by the union bound, the probability that there exists a depth d gate underneath which there are fewer than $\sqrt{dm \log m/2}$ many blocks unset is at most $(m^{d-3/2}/\sqrt{\log m})e^{-\sqrt{dm \log m/2}/4}$. This tends to 0 as m tends to infinity.

On the other hand, by lemma 8, with probability at least $3/4$, $C \upharpoonright_\rho$ is is equivalent to a depth d, bottom fan-in w circuit with at most S gates in levels 1 through $d-1$.

Therefore we may choose $\rho \in \mathcal{B}^{m,k+1}_{0,d}$ so that underneath each depth d gate of $g^{m,k+1}_d$ there are at least $\sqrt{dm \log m/2}$ many blocks unset by ρ, and $C \upharpoonright_\rho$ is equivalent to a depth d, bottom fan-in w circuit with $\leq S$ gates in levels $1, \ldots d-1$.

Because $C \upharpoonright_\rho$ computes $g^{m,k+1}_d \upharpoonright_\rho$, a restriction of it computes f^m_d: set some blocks to 0 and collapse the other blocks to a single variable. This gives a depth d circuit with $\leq S$ gates in levels $1, \ldots d-1$, and bottom fan-in w that computes f^m_d, a contradiction to theorem 6. ∎

5 Lower Bounds for Res(k) Refutations

We give three size lower bounds for Res(k) refutations of sets of clauses: lower bounds for random, constant width CNFs, lower bounds for the cn to n weak pigeonhole principle and lower bounds that give an exponential speedup of Res($k+1$) over Res(k). We cannot give much of the proofs for these results because of space considerations. We give more details for the lower bounds for random CNFs, and outline the other results.

All of our lower bounds use the fact that when the lines of a Res(k) refutation can be strongly represented by short decision trees, the refutation can be converted into a narrow resolution refutation.

Theorem 10 *Let $\mathcal{C}$ be a set of clauses of width $\leq h$. If $\mathcal{C}$ has a Res(k) refutation so that for each line F of the refutation, $h(F) \leq h$, then $w_R(\mathcal{C}) \leq kh$.*

Proof: We will use the short decision trees to construct a narrow refutation of $\mathcal{C}$ in resolution augmented with subsumption inferences: whenever $A \subseteq B$, $\frac{A}{B}$. These new inferences simplify our proof, but they may be removed from the resolution refutation without increasing the size or the width.

For each initial clause $C \in \mathcal{C}$, we let T_C be the decision tree that queries the (at most h) variables in C, stopping with a 1 if the clause becomes satisfied and stopping with a 0 if the clause becomes falsified. For the other lines, F, let T_F be a shortest decision tree that strongly represents F.

For any partial assignment π let C_π be the clause of width $\leq h$ that contains the negation of every literal in π, i.e., the clause that says that branch π was not taken.

We construct a resolution proof of width $\leq kh$ by deriving C_π for each line F of the refutation and each $\pi \in \mathrm{Br}_0(T_F)$.

Notice that for $\pi \in \mathrm{Br}_0(T_\emptyset)$, $C_\pi = \emptyset$, and for each $C \in \mathcal{C}$, for the unique $\pi \in \mathrm{Br}_0(T_C)$, $C_\pi = C$.

Let F be a line of the refutation that is inferred from the previously derived formulas $F_1, \ldots, F_j$, $j \leq k$. Assume we have derived all $C_\pi \in \mathrm{Br}_0(T_{F_i})$ for $1 \leq i \leq j$.

To guide the derivation of $\{C_\pi \mid \pi \in \mathrm{Br}_0(T_F)\}$, we construct a decision tree that represents the the conjunction of $F_1, \ldots F_j$. The tree (call it T) begins by simulating, T_{F_1} and outputting 0 on any 0-branch of T_{F_1}. On any 1-branch, it then simulates T_{F_2}, etc. If all j branches are 1, T outputs 1; otherwise T outputs 0. The height of T is at most $jh \leq kh$, so the width of any such C_π, with $\pi \in \mathrm{Br}(T)$ is at most kh.

Every $\sigma \in \mathrm{Br}_0(T)$ contains some $\pi \in \bigcup_{i=1}^{j} \mathrm{Br}_0(T_{F_i})$. Therefore, $\{C_\sigma \mid \sigma \in \mathrm{Br}_0(T)\}$ can be derived from the previously derived clauses by subsumption inferences.

On the other hand, if $\sigma \in \mathrm{Br}_1(T)$, there exists $\pi_1 \in \mathrm{Br}_1(T_{F_1}), \ldots, \pi_j \in \mathrm{Br}_1(T_{F_j})$ so that $\pi_1 \cup \cdots \cup \pi_j = \sigma$. Because the decision trees $T_{F_1}, \ldots T_{F_j}$ strongly represent the k-DNFs $F_1, \ldots, F_j$, there exist terms $t_1 \in F_1, \ldots, t_j \in F_j$ so that $\bigwedge_{i=1}^{j} t_i$ is satisfied by σ. By strong soundness of $\mathrm{Res}(k)$, there exists $t \in F$ so that σ satisfies t.

Let $\sigma \in \mathrm{Br}_0(T_F)$ be given. Because T_F strongly represents F, σ sets all terms of F to 0. So by the preceding paragraph, for all $\pi \in \mathrm{Br}(T)$, if π is consistent with σ, then $\pi \in \mathrm{Br}_0(T)$.

We now begin the derivation of $\mathrm{Br}_0(T_F)$. Let $\sigma \in \mathrm{Br}_0(T_F)$ be given. For each node v in T, let π_v be the path (viewed as a partial assignment) from the root to v. Bottom-up from leaves to root, we inductively derive $C_{\pi_v} \vee C_\sigma$, for each v so that π_v is consistent with σ. When we reach the root, we will have derived C_σ.

If v is a leaf, then $\pi_v \in \mathrm{Br}_0(T)$ (because it is consistent with σ), and it has already been derived.

If v is labeled with a variable that appears in σ, call it x, then there is a child u of v with $\pi_u = \pi_v \cup \{x\}$. Therefore, $C_{\pi_v} \vee C_\sigma = C_{\pi_u} \vee C_\sigma$. By induction, the clause $C_{\pi_u} \vee C_\sigma$ has already been derived.

If v is labeled with a variable x that does not appear in σ, then for both of the children of v, call them v_1, v_2, the paths π_{v_1} and π_{v_2} are consistent with σ. Moreover, $C_{\pi_{v_1}} \vee C_\sigma = x \vee C_{\pi_v} \vee C_\sigma$ and $C_{\pi_{v_2}} \vee C_\sigma = \neg x \vee C_{\pi_v} \vee C_\sigma$. Resolving these two previously derived clauses gives us $C_{\pi_v} \vee C_\sigma$. ∎

We will use this theorem after we apply a random restriction which simultaneously collapses every line of a $\mathrm{Res}(k)$ refutation to a short decision tree. Hence, we can use a width lower bound for resolution proofs of a restricted tautology to give a size lower bound for $\mathrm{Res}(k)$ proofs of the original tautology.

Corollary 11 *Let $\mathcal{C}$ be a set of clauses of width $\leq h$, let Γ be a Res(k) refutation of $\mathcal{C}$, and let ρ be a partial assignment so that for every line F of Γ, $h(F \upharpoonright_\rho) \leq h$. Then $w_R(\mathcal{C} \upharpoonright_\rho) \leq kh$.*

5.1 Lower Bounds for Random CNFs

Definition 5.1 *Let n, Δ and w be given. The distribution $\mathcal{F}_w^{n,\Delta}$ is defined by choosing $\Delta \cdot n$ many clauses independently, with repetitions, from the set of all $\binom{n}{w}2^w$ clauses of width w.*

Theorem 12 *For any $\epsilon \in (1/3, 1/2]$, there exists $\delta > 0$, so that for n sufficiently large and for $\Delta = n^{\frac{1}{2}-\epsilon}$,*

$$Pr_{F \in \mathcal{F}^{n,\Delta}_{4k^2+2}}\left[s_k(F) \leq 2^{n^\delta}\right] = o(1).$$

The reason that our proof does not give lower bounds for refutations of random 3-CNFs in $\mathrm{Res}(k)$ is that on one hand, we want our random restrictions to have a good chance of satisfying a fixed k-term (so we can apply the switching lemma), but on the other hand, the restrictions should have little probability of falsifying any of the initial clauses (this would make the restricted set of clauses trivial to refute). Because satisfying a k-term is equivalent to falsifying a k-clause, we can only work with initial clauses width far larger than k.

A set of clauses that, with constant probability, requires high width to refute after random restriction is called *robust*. Recall the distribution $\mathcal{D}_p$ from definition 3.3.

Definition 5.2 *Let F be a CNF in variables $x_1, \ldots x_n$.*

We say that F is (p, r) robust *if $Pr_{\rho \in \mathcal{D}_p}\left[w_R(F \upharpoonright_\rho) \geq r\right] \geq 1/2$.*

It turns out that a random w-CNF is almost surely robust when w is sufficiently large compared to p (the probability of fixing a bit to either 0 or 1). The proof appears in the full version. The idea is to show that when we consider the joint distribution on w-CNFs F and random restrictions ρ, with high probability, $F\restriction_\rho$ is implied by (contains) a random 3-CNF. We then apply the width bounds for the resolution refutations of random 3-CNFs [9].

Lemma 13 *There exists a constant c so that for any constants w and t, with $w \geq 2t+2$, for every n sufficiently large, and every $\epsilon \in [0,1/2]$, if we set $\Delta = n^{\frac{1}{2}-\epsilon}$ then the following inequality holds:*

$$Pr_{F\in\mathcal{F}_w^{n,\Delta}}\left[F \textit{is not } \left(n^{-1/t}, cn\cdot\Delta^{-\frac{2}{1-\epsilon}}\right)\textit{-robust}\right] = o(1)$$

We now prove the size lower bound. We set bits with probability $n^{-1/2k^2}$ so we can collapse k-DNFs but still have that most $4k^2+2$ CNFs are robust. For each $k \geq 1$, let γ_k be the constant of corollary 5 (with $\beta = 1$).

Lemma 14 *Let n, r, w, and k be given. For sufficiently large n, if F is a $(n^{-1/2k^2}, r)$-robust w-CNF, then $s_k(F) \geq 2^{(\gamma_k(r-1)/k\sqrt{n})-2}$.*

Proof: Suppose that Γ is a Res(k) refutation of F of size at most $2^{(\gamma_k(r-1)/k\sqrt{n})-2}$.

By corollary 5, with $p = n^{-1/2k^2}$, $c = 2$, $w = (r-1)/k$, we have that for every line F of Γ,

$$\Pr_{\rho\in\mathcal{D}_p}\left[h(F\restriction_\rho) > (r-1)/k\right] \leq 2^{-\gamma_k(r-1)/k\sqrt{n}}$$

By the union bound we have

$$\Pr_{\rho\in\mathcal{D}_p}\left[\exists F\in\Gamma\ h(F\restriction_\rho) > (r-1)/k\right]$$

$$\leq |\Gamma| 2^{-\gamma_k(r-1)/k\sqrt{n}}$$

$$\leq 2^{(\gamma_k(r-1)/k\sqrt{n})-2}2^{-\gamma_k(r-1)/k\sqrt{n}} = \frac{1}{4}$$

Because F is (p,r)-robust, with probability at least $1/2$ over choices of ρ, $w_R(F\restriction_\rho) \geq r$. Therefore, we may choose $\rho \in \mathcal{D}_p$ so that $w_R(F\restriction_\rho) \geq r$ and for all $F \in \Gamma$, $h(F\restriction_\rho) \leq (r-1)/k$. This is a contradiction because by corollary 11 there should be a width $r-1$ resolution refutation of $F\restriction_\rho$. ■

Combining lemmas 13 and 14 with $t = 2k^2$, $w = 4k^2+2$ and $r = cn\cdot\Delta^{-\frac{2}{1-\epsilon}}$ shows that a random $(4k^2+2)$-CNF almost surely requires exponential size to refute in Res(k).

Corollary 15 *There exists a constant c so that for every k, for every n sufficiently large and $\epsilon \in [0,1/2]$, if we set $\Delta = n^{\frac{1}{2}-\epsilon}$, then the following inequality holds.*

$$Pr_{F\in\mathcal{F}_{4k^2+2}^{n,\Delta}}\left[s_k(F) \leq 2^{\gamma_k(cn\cdot\Delta^{-\frac{2}{1-\epsilon}}-1)/k\sqrt{n}}\right] = o(1)$$

This gives an exponential lower bound only when $\Delta^{2/(1-\epsilon)}$ is substantially less than $\sqrt{n}$. Because $\Delta = n^{\frac{1}{2}-\epsilon}$, $\Delta^{2/(1-\epsilon)} = \left(n^{(1-2\epsilon)/2}\right)^{2/(1-\epsilon)} = n^{1-(\epsilon/(1-\epsilon))}$. Solving for ϵ shows that for any $\epsilon > 1/3$, there exists $\gamma > 0$ so that $\Delta^{2/(1-\epsilon)} < n^{\frac{1}{2}-\gamma}$

Theorem 16 *For any $\epsilon \in (1/3, 1/2]$, there exists $\delta > 0$, so that for n sufficiently large and for $\Delta = n^{\frac{1}{2}-\epsilon}$,*

$$Pr_{F\in\mathcal{F}_{4k^2+2}^{n,\Delta}}\left[s_k(F) \leq 2^{n^\delta}\right] = o(1)$$

5.2 Lower Bounds for the Weak Pigeonhole Principle

Definition 5.3 *The* m to n pigeonhole principle, *PHP_n^m, is the following set of clauses: (1) For each $i \in [m]$, $\bigvee_{j\in[n]} x_{i,j}$. (2) For each $i, i' \in [m]$ with $i \neq i'$, $\neg x_{i,j} \vee \neg x_{i',j}$.*

Theorem 17 *For $c > 1$, there exists $\epsilon > 0$ so that for all n sufficiently large, if $k \leq \sqrt{\log n/\log\log n}$, then every Res$(k)$ refutation of PHP_n^{cn} has size at least 2^{n^ϵ}.*

The general idea of the proof of Theorem 17 is as follows: Suppose there are small Res(k) refutations of the pigeonhole principles. Then, by applying random restrictions we obtain low width resolution refutations of restricted pigeonhole principles. This can be shown to be impossible, using the well-known lower bounds on the width of resolution refutations of the pigeonhole principle. See the full version for more details.

5.3 Separation Between Res(k) and Res$(k+1)$

In the full version of this paper, we show that for each constant k, there is an $\epsilon_k > 0$, such that there is a family of unsatisfiable CNFs which have polynomial size Res$(k+1)$ refutations but which require size $2^{n^{\epsilon_k}}$ to refute in Res(k). The unsatisfiable clauses are a variation of the graph ordering tautologies of [19, 10].

Definition 5.4 *Let G be an undirected graph. For each vertex u of G, let $N(u)$ denote the set of neighbors of u in G. For each ordered pair of vertices $(u,v) \in V(G)^2$, with $u \neq v$, let there be a propositional variable $X_{u,v}$.*

The graph ordering principle for G, *$GOP(G)$, is the following set of clauses: (1) The relation X is transitive: for all $u, v, w \in V(G)$, $X_{u,v} \wedge X_{v,w} \rightarrow X_{u,w}$ (2)*

The relation X is anti-symmetric: for all $u, v \in V(G)$ with $u \neq v$, $\neg X_{u,v} \vee \neg X_{v,u}$ (3) There is no locally X-minimal element: for every $u \in V(G)$, $\bigvee_{v \in N(u)} X_{v,u}$.

The k-fold graph ordering principle of G, $GOP^k(G)$, is obtained by replacing each variable $X_{u,v}$ by a conjunction of k variables, $X^1_{u,v}, \ldots, X^k_{u,v}$, and then using the distributive rule and DeMorgan's law to express this as a set of clauses.

Notice that for a graph G on n vertices with maximum degree $d \geq 3$, the principle $GOP(G)$ has width d and the principle $GOP^k(G)$ has size $O(n^3 k^d)$.

It is easily shown that, for any graph G, the principle $GOP(G)$ has polynomial size resolution refutations. Furthermore, these refutations can be transformed into Res$(k+1)$ refutations of $GOP^{k+1}(G)$. On the other hand:

Theorem 18 *Let k be given. There exist constants $c > 0$ and $\epsilon_k > 0$, and a family of graphs G on n vertices (for n sufficiently large) with maximum degree $c \log n$ so that Res(k) refutations of $GOP^{k+1}(G)$ require size at least $2^{O(n^{\epsilon_k})}$.*

The proof of theorem 18 uses a random restriction that collapses a small Res(k) refutation into a narrow resolution refutation. Then, using a width bound for resolution refutations of the graph ordering principle, we show that no small Res(k) refutation can exist. The width bound for resolution refutations of the graph ordering principle is a generalization of lower bounds of Bonet and Galesi [10], but applies to randomly chosen, sparse graphs that satisfy a certain expansion-like property. See the full version for more details.

6 Conclusions and Open Problems

Switching with small restrictions seems to be a promising technique for analyzing the power of bottom fan-in in proof and circuit complexity. Our results could not have been obtained by switching with larger restrictions. For example, the lower bounds for random w-CNFs could not be proved using restrictions that set a constant fraction of the variables because some initial clause is falsified with high probability. Also, this method is relatively easy to apply because you do not have to reprove the switching lemma for every lower bound, but only check that the restrictions in question are likely to satisfy k-DNFs with high cover number.

However, switching with small restrictions still suffers from the limitations of random restriction method. In particular, it seems ineffective against random 3-CNFs and very weak pigeonhole principles. The only techniques for understanding the refutation complexity of such CNFs seem specific to resolution [9, 8, 28, 29]. Understanding the refutation complexity of these principles in Res(k) is a necessary step before understanding them in more powerful systems, and the Res(k) systems might be simple enough for the development of new techniques.

With this in mind, we suggest the following open problems as particularly relevant: (1) Do random 3-CNFs almost surely require exponential size refutations in Res(k) for all k? (2) Does there exist a family of 3-CNFs that require exponential size to refute in Res(k) but have (quasi-) polynomial size proofs in Res$(k+1)$? (3) Do Res(2) refutations of PHP^m_n require size exponential in n for all m? (3) Do there exist polynomial size depth-two Frege refutations PHP^{2n}_n? (4) Let $0 < \epsilon \leq 1/2$. Do there exist sub-exponential size refutations for $PHP^{n+n^{1-\epsilon}}_n$ in Res(polylog(n))? or even in depth-two Frege? (5) Does there exist a family of CNFs that require exponential size refutations in Res(polylog(n)) but have (quasi-) polynomial size depth-two Frege refutations? (6) For given $\epsilon < \delta \leq 1$, does there exist a family of CNFs that require exponential size refutations in Res(n^ϵ) but have (quasi-) polynomial size Res(n^δ) refutations?

References

[1] M. Ajtai. Σ^1_1-formulae on finite structures. *Annals of Pure and Applied Logic*, 24:1–48, 1983.

[2] A. Atserias and M. L. Bonet. On the automatizability of resolution and related propositional proof systems. In *Electronic Colloquium on Computational Complexity, technical reports*, 2002.

[3] A. Atserias, M. L. Bonet, and J. L. Esteban. Lower bounds for the weak pigeonhole principle beyond resolution. *Lecture Notes in Computer Science*, 2076, 2001.

[4] P. Beame. A switching lemma primer. Technical report, Department of Computer Science and Engineering, University of Washington, 1994.

[5] P. Beame, R. Karp, T. Pitassi, and M. Saks. On the complexity of unsatisfiability proofs for random k-CNF formulas. In *ACM Symposium on Theory of Computing (STOC)*, 1998.

[6] P. Beame and T. Pitassi. Simplified and improved resolution lower bounds. In *37th Annual Symposium on Foundations of Computer Science*, pages 274–282. IEEE, 1996.

[7] P. Beame and T. Pitassi. Propositional proof complexity: Past, present, and future. *Bulletin of the EATCS*, 65:66–89, 1998.

[8] E. Ben-Sasson. Hard examples for bounded depth frege. In *ACM Symposium on Theory of Computing (STOC)*, 2002.

[9] E. Ben-Sasson and A. Wigderson. Short proofs are narrow — resolution made simple. *Journal of the ACM*, 48(2):149–169, 2001.

[10] M. L. Bonet and N. Galesi. A study of proof search algorithms for resolution and polynomial calculus. In *40th Annual Symposium on Foundations of Computer Science*, pages 422–431. IEEE Computer Society Press, 1999.

[11] R. Boppana and M. Sipser. The complexity of finite functions. In J. van Leeuwen, editor, *Handbook of Theoretical Computer Science*, volume A. Elsevier and MIT Press, 1990.

[12] S. R. Buss and G. Turán. Resolution proofs of generalized pigeonhole principles. *Theoretical Computer Science*, 62(3):311–317, December 1988.

[13] L. Cai, J. Chen, and J. Håstad. Circuit bottom fan-in and computational power. *SIAM Journal on Computing*, 27(2):341–355, March 1998.

[14] V. Chvátal and E. Szemerédi. Many hard examples for resolution. *Journal of the ACM*, 35(4):759–768, October 1988.

[15] S. Cook and R. Reckhow. The relative efficiency of propositional proof systems. *The Journal of Symbolic Logic*, 44(1):36 – 50, March 1979.

[16] S. Dantchev and S. Riis. Tree resolution proofs of the weak pigeon-hole principle. In *Annual Conference on Structure in Complexity Theory*, 2001.

[17] J. L. Esteban, N. Galesi, and J. Messner. On the complexity of resolution with bounded conjunctions. In *29th International Colloquium on Automata, Languages and Programming*, 2002.

[18] M. Furst, J. Saxe, and M. Sipser. Parity, circuits, and the polynomial-time hierarchy. *Mathematical Systems Theory*, 17(1):13–27, April 1984.

[19] A. Goerdt. Unrestricted resolution versus N-resolution. In *Mathematical Foundations of Computer Science 1990*, volume 452 of *Lecture Notes in Computer Science*, pages 300–305. Springer, 1990.

[20] A. Haken. The intractability of resolution. *Theoretical Computer Science*, 39(2-3):297–308, August 1985.

[21] J. Håstad. Almost optimal lower bounds for small depth circuits. In *Proceedings of the Eighteenth Annual ACM Symposium on Theory of Computing (STOC)*, pages 6–20, 1986.

[22] J. Håstad. Almost optimal lower bounds for small depth circuits. *ADVCR: Advances in Computing Research*, 5, 1989.

[23] J. Krajíček. *Bounded Arithmetic, Propositional Logic and Complexity Theory*. Cambridge University Press, 1995.

[24] J. Krajíček. On the weak pigeonhole principle. *Fudamenta Mathematicae*, 170:123–140, 2001.

[25] A. Maciel, T. Pitassi, and A. Woods. A new proof of the weak pigeonhole principle. In *ACM Symposium on Theory of Computing (STOC)*, 2000.

[26] T. Pitassi and R. Raz. Regular resolution lower bounds for the weak pigeonhole principle. In *ACM Symposium on Theory of Computing (STOC)*, 2001.

[27] P. Pudlák. The lengths of proofs. In S. R. Buss, editor, *Handbook of Proof Theory*, pages 547–637. Elsevier North-Holland, 1998.

[28] R. Raz. Resolution lower bounds for the weak pigeonhole principle. In *Proceedings of the Thirty-Fourth Annual ACM Symposium on Theory of Computing (STOC)*, 2002.

[29] A. Razborov. Improved resolution lower bounds for the weak pigeonhole principle. In *Electronic Colloquium on Computational Complexity, technical reports*, 2001.

[30] M. Sipser. Borel sets and circuit complexity. In *Proceedings of the Fifteenth Annual ACM Symposium on Theory of Computing (STOC)*, pages 61–69, 1983.

[31] G. Tseitin. On the complexity of proofs in propositional logics. *Seminars in Mathematics*, 8, 1970.

[32] A. Urquhart. Hard examples for resolution. *Journal of the ACM*, 34(1):209–219, 1987.