A SWITCHING LEMMA FOR SMALL RESTRICTIONS AND LOWER BOUNDS FOR K-DNF RESOLUTION

NATHAN SEGERLIND*, SAM BUSS†, AND 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 terms. 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.

Key words. propositional proof complexity, Boolean circuit complexity, switching lemmas, lower bounds, k-DNFs, resolution, res(k), circuit bottom fan-in, random restriction, Sipser functions, weak pigeonhole principles, random CNFs

AMS subject classifications.
03F20, 68Q15

Preferred hort title: Lower Bounds for K-DNF Resolution

1. Introduction.

This paper studies the complexity of 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 [14, 23, 29, 6], 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. Res(k) corresponds to algorithms that branch on more general conditions: the value of any function of up to k variables.

The Res(k) systems are also interesting as intermediates between previously studied proof systems. Resolution can be thought of as Res(1) and depth two Frege can be thought of as Res(n) (where n is the number of variables). The Res(k) systems provide a transition between these systems. Moreover, statements provable in the theory $T_2^2(\alpha)$ (a fragment of Peano's arithmetic that allows induction only on Σ_2^b predicates) correspond to propositional statements with quasi-polynomial size Res(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 require exponential size resolution refutations. Thus, there must be a critical range for k between 1 and $polylog(n)$ where these arguments become possible in sub-exponential size. More generally, we can ask:

*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.

when does increasing k give the 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 Res(k) for all $k \leq \sqrt{\log n / \log \log n}$. So having large bottom fan-in is necessary for counting arguments.
2. For each k, there exists a constant $w > k$ so that random w-CNFs require size $2^{\Omega(n^\epsilon)}$ to refute in Res(k). Therefore, extending DPLL algorithms to branch on multiple (but a constant number of) variables, will not make run times sub-exponential on average.
3. Res($k+1$) is exponentially more powerful than Res(k). We demonstrate sets of clauses that have polynomial size Res($k+1$) refutations, but require size $2^{\Omega(n^\epsilon)}$ to refute in Res(k). Therefore, increasing the complexity of branching conditions can make proofs exponentially smaller.

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 switching lemma suffices to prove lower bounds for the Res(k) proof systems, because each line in such a proof is of depth two with small bottom fan-in.

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 $k+1$. This refines a previous result of Håstad [21], 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. Our result also refines results of Cai, Chen and Håstad [12]. They showed that for each constant d, there exist functions computable with polynomial size, depth $d+1$, bottom fan-in 2 circuits that require exponential size to compute with depth d circuits, 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 Res(1), our results for 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 [35, 20, 36, 11, 5, 15, 28, 30, 31], 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 [13, 5, 8]. We extend these results to general 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 Res(k) system. The complexity of Res(k) refutations was first studied by Krajíček [24], who was motivated by the connection between Res(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 Res(2) refutations of the $2n$ to n weak pigeonhole principle and of random 3-CNFs. They also proved a quasi-polynomial separation between 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 Res(k) and treelike Res($k+1$), and that general (DAG-like) Res(k) requires high *space* to refute random CNFs and the weak pigeonhole principle.

The lower bounds for 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 quasipolynomial size refutations in Res(polylog(n)). Our results show that super-constant bottom fan-in is necessary for sub-exponential size proofs of the weak pigeonhole principle. Indeed, after the preliminary version of this paper appeared [33], our techniques were extended by Alexander Razborov to prove that the weak pigeonhole principle requires exponential size to refute in Res($\epsilon \log n / \log\log n$) [32].

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

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

In the time since the preliminary version of this paper appeared [33], we have extended one of our results. The original separation of Res($k+1$) from Res(k) was based on clauses of width $O(\log n)$, whereas the new separation uses clauses of constant width. We include proofs for both results.

1.1. Outline of the Paper. Background material, including the basics of the Res(k) proof system, is given in section 2. In section 3 we prove the switching lemma. Section 4 applies the switching lemma to prove a separation between constant-depth circuits of bottom fan-in $k+1$ and constant-depth circuits of bottom fan-in k. In section 5 we prove that Res(k) refutations of sets of clauses whose lines are represented by short decision trees can be transformed into narrow resolution refutations. This conversion is used in combination with the switching lemma to show prove lower bounds for Res(k) refutations. Lower bounds for Res(k) refutations of the weak pigeonhole principle are proved in section 6 and lower bounds for Res(k) refutations of random CNFs are proved in section 7. The separations between Res($k+1$) and Res(k) are proved in sections 8 and 9.

2. Background. 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 sub-expression has become explicitly constant. For any restriction ρ, let $\mathrm{dom}(\rho)$ 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 of clauses $\mathcal{C}$; a resolution refutation consists of a derivation of the empty clause from the clauses in $\mathcal{C}$ using only the resolution inference: $\frac{A \vee x \quad \neg x \vee B}{A \vee B}$. Notice that every line in a resolution refutation is a clause. $w_R(\mathcal{C})$ denotes the minimum width of a resolution refutation of $\mathcal{C}$; if $\mathcal{C}$ is satisfiable then there is no refutation and we use the convention that $w_R(\mathcal{C})$ is ∞.

The $\mathrm{Res}(k)$ refutation system is a generalization of resolution that can reason using k-DNFs.

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-DNF's,* $1 \leq j \leq k$*, and* $l, l_1, \ldots, l_j$ *are literals):*

$$\textit{Subsumption:} \quad \frac{A}{A \vee l} \qquad \textit{AND-introduction:} \quad \frac{A \vee l_1 \ \cdots \ A \vee l_j}{A \vee \bigwedge_{i=1}^{j} l_i}$$

$$\textit{Cut:} \quad \frac{A \vee \bigwedge_{i=1}^{j} l_i \quad B \vee \bigvee_{i=1}^{j} \neg l_i}{A \vee B} \qquad \textit{AND-elimination:} \quad \frac{A \vee \bigwedge_{i=1}^{j} l_i}{A \vee l_i}$$

Let $\mathcal{C}$ *be a set of k-DNFs. A* $\mathrm{Res}(k)$ 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* $\mathrm{Res}(k)$ refutation of $\mathcal{C}$ *is a derivation from* $\mathcal{C}$ *whose final line is the empty clause. The size of a Res(k) refutation is the number of lines it contains.* $S_k(\mathcal{C})$ *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* $S_k(\mathcal{C})$ *is* ∞*.*

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 $\bigwedge_{i=1}^{j} t_i$. In other words, any reason why $F_1, \ldots, F_k$ are true implies a reason why F is true. This is stronger than mere soundness.

LEMMA 2.2. *Res(k) is strongly sound.*

We also use a well-known interpolation property for resolution:

LEMMA 2.3. *Let* $\mathcal{C}_1$ *and* $\mathcal{C}_2$ *be unsatisfiable sets of clauses on disjoint sets of variables. If there is a resolution refutation* Γ *of* $\mathcal{C}_1 \cup \mathcal{C}_2$*, then there is a refutation* Γ' *of either* $\mathcal{C}_1$ *or* $\mathcal{C}_2$*. Moreover,* $w(\Gamma') \leq w(\Gamma)$*.*

2.1. The Chernoff Bounds. Throughout this paper we will repeatedly make use of a simplified form of the Chernoff bounds. These formulations come from standard references on applying such bounds in algorithmics, c.f. [26, 37].

LEMMA 2.4. *Let* $X_1, \ldots, X_n$ *be independent random indicator variables. Let* $\mu = E[\sum_{i=1}^{n} X_i]$.

$Pr\left[\sum_{i=1}^{n} X_i < \frac{\mu}{2}\right] \leq e^{-\mu/8}$ *and* $Pr[\sum_{i=1}^{n} X_i > 2\mu] \leq e^{-\mu/4}$.

2.2. Miscellaneous Notation. We will use the notation $[k] = \{i \mid 1 \leq i \leq k\}$. For graphs $G = (V, E)$ and $S \subseteq V$ we will write $G - S$ to denote the induced subgraph

on $V \setminus S$.

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 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 $Br_v(T)$. For a partial assignment ρ, $T \restriction_\rho$ is the decision tree obtained by deleting from T every edge whose label conflicts with ρ and contracting along each edge whose label belongs to ρ. 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 be computed both by an h-CNF and by 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.

The switching lemma is shown to hold for all distributions which satisfy certain properties. When we apply the switching lemma, we will show that the random restrictions used satisfy these properties.

THEOREM 3.3. *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 $Pr_{\rho \in \mathcal{D}}[G \restriction_\rho \neq 1] \leq p_i$. Then for every k-DNF F:*

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

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 $F \restriction_\rho$ by querying all of the variables of $F \restriction_\rho$. On the other hand, if $c(F) > s_0$ then $\Pr_{\rho \in \mathcal{D}}[F \restriction_\rho \neq 1] \leq p_1$. Therefore, $h(F \restriction_\rho)$ is non-zero with probability at most $p_1 2^{\sum_{j=1}^{k-1} s_j} = p_1$ (because $k = 1$).

For the induction step, assume that the theorem holds for all k-DNFs, let F be a $(k+1)$-DNF, and let $s_0, \ldots, s_k$ and $p_1, \ldots, p_{k+1}$ be sequences of positive numbers satisfying the hypotheses of the theorem. If $c(F) > s_k$, then by the conditions of the lemma, $\Pr_{\rho \in \mathcal{D}}[F \restriction_\rho \neq 1] \leq p_{k+1}$. Because $p_{k+1} \leq \sum_{i=1}^{k+1} 2^{\sum_{j=i}^{k} s_j} p_i$, we have that $h(F \restriction_\rho)$ 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, $F \restriction_\pi$ is a k-DNF. By combining the induction hypothesis with the union bound, we have that

$$\begin{aligned} \Pr_{\rho\in\mathcal{D}}\left[\exists\pi\in\{0,1\}^S \;\; h((F\restriction_\rho)\restriction_\pi) > \textstyle\sum_{i=0}^{k-1} s_i\right] &\leq 2^{s_k}(\sum_{i=1}^{k} 2^{\left(\sum_{j=i}^{k-1} s_j\right)} p_i) \\ &< \sum_{i=1}^{k+1} 2^{\left(\sum_{j=i}^{k} s_j\right)} p_i \end{aligned}$$

In the event that $\forall\pi\in\{0,1\}^S$, $h((F\restriction_\rho)\restriction_\pi) \leq \sum_{i=0}^{k-1} s_i$, we construct a decision tree for $F\restriction_\rho$ 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 $(F\restriction_\rho)\restriction_\beta$. For each such β, let $\pi = (\rho\cup\beta)\restriction_S$, and note that $h((F\restriction_\rho)\restriction_\beta) = h((F\restriction_\rho)\restriction_\pi)$. Therefore the height of the resulting decision tree is at most $s_k + \max_{\pi\in\{0,1\}^S} h((F\restriction_\rho)\restriction_\pi) \leq \sum_{i=0}^{k} s_i$.

Now we show that the decision tree constructed above strongly represents $F\restriction_\rho$. Let π be a branch of the tree. Notice that $\pi = \beta\cup\sigma$, where β is an assignment to the variables in $S\setminus\mathrm{dom}(\rho)$ and σ is a branch of a tree that strongly represents $(F\restriction_\rho)\restriction_\beta$. Consider the case that π leads to a leaf labeled 1. In this case, σ satisfies a term t' of $(F\restriction_\rho)\restriction_\beta$. We may choose a term t of F so that $t' = (t\restriction_{\rho\cup\beta})$, and $\pi = \beta\cup\sigma$ satisfies the term $t\restriction_\rho$ of $F\restriction_\rho$. Now consider the case that π leads to a leaf labeled 0. There are two cases, $(F\restriction_\rho)\restriction_\beta = 0$ and $(F\restriction_\rho)\restriction_\beta \neq 0$. If $(F\restriction_\rho)\restriction_\beta = 0$, then for every term t of $F\restriction_\rho$, t is inconsistent with β and hence with π. If $(F\restriction_\rho)\restriction_\beta \neq 0$ then because the sub-tree underneath β strongly represents $(F\restriction_\rho)\restriction_\beta$, for every term t of $(F\restriction_\rho)\restriction_\beta$, t is inconsistent with σ. Therefore, every term of $F\restriction_\rho$ is inconsistent with either β or σ, and thus with $\pi = \beta\cup\sigma$. □

We usually use this theorem in the following normal form for the parameters:

COROLLARY 3.4. *Let k, s and d be positive integers, let γ and δ be real numbers from the range $(0,1]$ and let $\mathcal{D}$ be a distribution on partial assignments so that for every k-DNF G, $Pr_{\rho\in\mathcal{D}}[G\restriction_\rho \neq 1] \leq d2^{-\delta(c(G))^\gamma}$. For every k-DNF F:*

$$Pr_{\rho\in\mathcal{D}}[h(F\restriction_\rho) > 2s] \leq dk2^{-\delta' s^{\gamma'}}$$

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

Proof. Let $s_i = (\delta/4)^i(s^{\gamma^i})$, and $p_i = d2^{-4s_i}$. Note that $s_{i-1}/4 \geq (\delta/4)s_{i-1} = (\delta/4)(\delta/4)^{i-1}s^{\gamma^{i-1}} \geq (\delta/4)^i s^{\gamma^i} = s_i$. It follows that $\sum_{j=i}^{k} s_j \leq \sum_{j\geq i} s_i/4^{j-i} \leq 2s_i$. Also, for any i-DNF G, with $c(G) \geq s_{i-1}$, $\Pr_{\rho\in\mathcal{D}}[G\restriction_\rho \neq 1] \leq d2^{-\delta(c(G))^\gamma} \leq d2^{-\delta s_{i-1}^\gamma} = 2^{-\delta(\delta/4)^{i-1}(s^{\gamma^{i-1}})^\gamma} = d2^{-4s_i}$. Thus, we can apply theorem 3.3 with parameters $p_1,\ldots,p_k,s_0,\ldots,s_{k-1}$. For every k-DNF F:

$$\begin{aligned} \Pr_{\rho\in\mathcal{D}}[h(F\restriction_\rho) > 2s] &\leq \Pr_{\rho\in\mathcal{D}}\left[h(F\restriction_\rho) > \textstyle\sum_{i=0}^{k-1} s_i\right] \leq \sum_{i=1}^{k} 2^{\left(\sum_{j=i}^{k-1} s_j\right)} p_i \\ &\leq \sum_{i=1}^{k} 2^{2s_i}(d2^{-4s_i}) \leq dk2^{-2s_k} = dk2^{-\delta' s^{\gamma'}} \end{aligned}$$

□

3.1. Switching with Small Restrictions. In this subsection, we show that small, uniform restrictions meet the conditions for the switching lemma. Using corollary 3.4, 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. Later, it will be used to prove the lower bound on

Res(k) refutations of random CNFs. More complicated distributions are used for our other results.

DEFINITION 3.5. *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 3.6. *Let $i \geq 1$, G be an i-DNF, and ρ be chosen from $\mathcal{D}_p$. Then $Pr[G \restriction_\rho \neq 1] \leq e^{-\frac{c(G)p^i}{i2^i}}$.*

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 $c(G)/i$ (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 $(p/2)^i$. Therefore,

$$\Pr_{\rho\in\mathcal{D}_p}[G \restriction_\rho \neq 1] \leq \left(1-\left(\tfrac{p}{2}\right)^i\right)^{\frac{c(G)}{i}} \leq e^{-\left(\frac{p}{2}\right)^i \frac{c(G)}{i}} = e^{-\frac{c(G)p^i}{i2^i}} \qquad \square$$

Combining this with the switching lemma shows that a k-DNF is strongly represented by a short decision tree when restricted.

COROLLARY 3.7. *Let $k \geq 1$ be given. There exists $\gamma > 0$ so that for any k-DNF F, $w > 0$, $p \geq n^{-1/(2k^2)}$, $Pr_{\rho\in\mathcal{D}_p}[h(F \restriction_\rho) > w] \leq k2^{-\gamma w n^{-1/2}}$.*

Proof. In the notation of corollary 3.4, set $p = n^{-1/2k^2}$, $d = 1$, $\gamma = 1$, $s = w/2$ and $\delta = (\log e)\frac{p^k}{k2^k} = (\log e)\frac{n^{-1/2k}}{k2^k}$. Combining lemma 3.6 with corollary 3.4 shows that for every k-DNF F:

$$\Pr_{\rho\in\mathcal{D}_p}\left[h(F \restriction_\rho) > w\right] \leq k2^{-2(w/2)(\delta^k/4^k)} = k2^{-w(\log e)^k n^{-1/2}/4^k k^k 2^{k^2}}$$

Because k is fixed, we may take $\gamma = (\log e)^k/4^k k^k 2^{k^2}$.
$\square$

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 $k+1$.

Our 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 [10].

4.1. The Functions. DEFINITION 4.1. *[34, 10] 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. Furthermore, we will often identify these functions with the circuits defining them.

Our result builds upon the earlier result that it is impossible to decrease the bottom fan-in of a circuit computing a Sipser function without increasing the size or the depth. Moreover, an OR of depth d, small bottom fan-in circuits requires exponential size to compute f_d^m.

THEOREM 4.2. *[21] 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 $k \geq m^{\beta_d}$.

We use the modified Sipser function $g_d^{m,k+1}$ to obtain the exponential separation between depth $d+1$, bottom fan-in $k+1$ and depth $d+1$, bottom fan-in k circuits. 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* $(i_1, \ldots, i_d)$. Variables in the same block occur in the same bottom-level conjunction of $g_d^{m,k}$. 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.

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 4.2, it suffices to consider only circuits with an AND gate at the output level. Furthermore, we 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. Each gate at depth d computes a k-DNF, and we will apply 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 4.2.

DEFINITION 4.3. *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 $\frac{1}{2}$ 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 dual distribution, $\mathcal{B}_{d,1}^{m,k+1}$, selects a restriction according to $\mathcal{B}_{d,0}^{m,k+1}$ and then inverts the 0s and 1s.

LEMMA 4.4. *Let $k \geq 1$ be given. There exists a constant $\gamma_k > 0$ so that for every k-DNF F:*

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

Proof. We say that two terms T and T' are *block-disjoint* if no variable of T shares a block with a variable of T'. More formally, whenever a variable $x_{i_1,\ldots,i_{d+1}}$ appears in T and a variable $x_{j_1,\ldots,j_{d+1}}$ appears in T', we have that $(i_1, \ldots, i_d) \neq (j_1, \ldots, j_d)$. Because each term involves at most k variables, there must be a set of $c(F)/k$ many variable-disjoint terms, and hence a set of $c(F)/(k(k+1))$ many block-disjoint terms.

We now show that each term is satisfied with probability at least $\frac{1}{6^k}$. Because the literals of a term come from at most k distinct blocks, the chance that every variable in the term is set to 0 or 1 is at least $1/2^k$. Conditioned on this event, the probability of satisfying the term is at least $1/3^k$. To see this, consider the chance of satisfying each literal of the term in turn, conditioned on the event of satisfying the previous literals. If a variable from that block has already been set to 0, then clearly the probability of satisfying the current literal is 1/2. If not, then suppose there l variables in the block of the current variable that have not yet been set to a value. The probability of satisfying the current literal is at least $(2^{l-1}-1)/(2^l-1)$. Because there are $k+1$ variables and the term has size at most k, $l \geq 2$, and thus the probability is at least 1/3.

The events of satisfying block-disjoint terms are independent, therefore we have:

$$\Pr_{\rho \in \mathcal{B}_{d,0}^{m,k+1}} [F \restriction_\rho \neq 1] \leq \left(1 - \frac{1}{6^k}\right)^{c(F)/(k(k+1))}$$

Set $\gamma_k = -\log_2(1-\frac{1}{6^k})/(k(k+1))$. □

Symmetrically, the dual result holds for k-CNFs when we apply a random restriction from $\mathcal{B}_{d,1}^{m,k+1}$.

LEMMA 4.5. *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 ρ is chosen from $\mathcal{B}_{d,0}^{m,k+1}$ ($\mathcal{B}_{d,1}^{m,k+1}$), with probability at least 3/4 , $C \restriction_\rho$ 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.

We consider the case when d is even; the case when d is odd is handled by using the restrictions $\mathcal{B}_{d,1}^{m,k+1}$ instead of $\mathcal{B}_{d,0}^{m,k+1}$. Each gate at depth d is an OR gate and its inputs are AND gates of fan-in at most k. For each 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 $\rho \in \mathcal{B}_{d,0}^{m,k}$ so that for each depth d gate g of C, $h(F_g \restriction_\rho) < w$. For each g at depth d, let T_g be the shortest decision tree representing $F_g \restriction_\rho$. We can compute $C \restriction_\rho$ 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 with 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 $\mathcal{B}_{d,0}^{m,k}$, with probability at least 3/4, every depth d gate g of C has $h(F_g \restriction_\rho) < w$.

Let g be a depth d gate of the circuit. By combining lemma 4.4 with corollary 3.4, setting $d=1$, $\gamma = 1$ $s = w/2$ and $\delta = \gamma_k$ shows that

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

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 4.6. *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 $g_d^{m,k+1}$ has $S \geq 2^{\epsilon_k m^{\epsilon_d}}$.*

Proof. We will have to take m sufficiently large to apply theorem 4.2 and lemma 4.5, and large enough for an application of the Chernoff bounds. Set $w = m^{\epsilon_d}$ (with ϵ_d from theorem 4.2) and $S = 2^{\epsilon_k w}$ (with ϵ_k from lemma 4.5). Furthermore, we consider the case when d is even; the case when d is odd is handled by using the restrictions $\mathcal{B}_{d,1}^{m,k+1}$ instead of $\mathcal{B}_{d,0}^{m,k+1}$.

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

Fix an OR gate at depth d in $g_d^{m,k+1}$. When ρ is chosen from the distribution $\mathcal{B}_{d,0}^{m,k+1}$, 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 $g_d^{m,k+1}$, 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 4.5, with probability at least 3/4 , $C \restriction_\rho$ 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 $\rho \in \mathcal{B}_{0,d}^{m,k+1}$ so that underneath each depth d gate of $g_d^{m,k+1}$ there are at least $\sqrt{dm \log m/2}$ many blocks unset by ρ, and $C \restriction_\rho$ is equivalent to a depth d, bottom fan-in w circuit with $\leq S$ gates in levels $1, \ldots, d-1$.

Because $C \restriction_\rho$ computes $g_d^{m,k+1} \restriction_\rho$, a restriction of it computes f_d^m: 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_d^m, a contradiction to theorem 4.2. □

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

THEOREM 5.1. *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 $A \subseteq B$, infer B from A. These new inferences simplify our proof, but they may be removed from the resolution refutation without increasing the size or the width.

For a line F of the Res(k) refutation, let T_F be a decision tree of minimum height that strongly represents F. Notice that for each initial clause $C \in \mathcal{C}$, T_C is the 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 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 refutation 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 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

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

We now show that for every $\sigma \in \mathrm{Br}_1(T)$, there exists a $t \in F$ so that σ satisfies t. Choose $\pi_1 \in \mathrm{Br}_1(T_{F_1}), \ldots, \pi_j \in \mathrm{Br}_1(T_{F_j})$ 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 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 falsifies all terms of F. By the preceding paragraph, for all $\pi \in \mathrm{Br}(T)$, if π is consistent with σ, then $\pi \in \mathrm{Br}_0(T)$ (otherwise, σ would not falsify the term of F satisfied by π). 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 the leaves to the root, we recursively 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 $\pi_v \in \mathrm{Br}_0(T)$ so 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 Res(k) refutation to a short decision tree. Hence, we can use a width lower bound for resolution refutations of a restricted tautology to give a size lower bound for Res(k) refutations of the original tautology.

COROLLARY 5.2. *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, $h(F \restriction_\rho) \leq h$. Then $w_R(\mathcal{C} \restriction_\rho) \leq kh$.*

6. Lower Bounds for the Weak Pigeonhole Principle. DEFINITION 6.1. *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 6.2. *For every $c > 1$, 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 idea of the proof for Theorem 6.2 is as follows: Suppose there is a small Res(k) refutation of the weak pigeonhole principle. Then, by applying a random restriction we obtain a low width resolution refutation of the restricted pigeonhole principle. By the well-known lower bounds on the width of resolution refutations of the pigeonhole principle, this is impossible.

In order to make the random restriction method work, we prove lower bounds for for the pigeonhole principle restricted to a low degree graph. Because these principles reduce to the pigeonhole principle by setting some variables to 0, this suffices to prove lower bounds for the pigeonhole principle. The difficulty with applying random restrictions directly to the clauses of the pigeonhole principle is that there are clauses of high width which are not satisfied with very high probability. If we were to choose

a random subset of the holes and place into each hole a randomly chosen pigeon, then a clause of the form $\bigvee_{i=1}^{m} x_{i,j}$ would be satisfied with probability no better than the chance that hole j is in the random subset (this will be no better than a constant in our proof). At the heart of this problem is that each hole j appears in cn distinct variables, $x_{1,j}, \ldots, x_{cn,j}$, and restricting the principle to low degree graph solves this.

DEFINITION 6.3. *Let $G = (U \cup V, E)$ be a bipartite graph. The* pigeonhole principle of G, PHP(G), *is the set of clauses*
1. *For each $u \in U$*

$$\bigvee_{\substack{v \in V \\ \{u,v\} \in E}} x_{u,v}$$

2. *For each $u, u' \in [m]$, with $u \neq u'$, and each $v \in V$ with $\{u, v\} \in E$ and $\{u', v\} \in E$*

$$\neg x_{u,v} \vee \neg x_{u',v}$$

DEFINITION 6.4. *Let $G = (U \cup V, E)$ be a bipartite graph. The* maximum degree of G, $\Delta(G)$, *is defined to be* $\max_{v \in V} \deg v$.

Furthermore, we assume that all Res(k) refutations have been put into a normal form in which no term of any DNF asks that two pigeons be mapped to the same hole. See, for example, [3].

DEFINITION 6.5. *Let $G = (U \cup V, E)$ be a bipartite graph. A term is said to be in* pigeon-normal-form *if it does not contain two literals $x_{u,v}$ and $x_{u',v}$ with $u \neq u'$. A DNF is said to be in pigeon-normal-form if all of its terms are in pigeon-normal-form and a Res(k) refutation is said to be in pigeon normal form if every line is in pigeon-normal-form.*

Every Res(k) refutation of $PHP(G)$ can be transformed into a refutation in pigeon normal form which at must doubles the number of lines in the proof. When there is an AND-introduction inference that creates a line not in pigeon normal form, say

$$\frac{(A \vee x_{u,v}) \quad (A \vee x_{u',v}) \quad \cdots \quad (A \vee l_j)}{A \vee \left(x_{u,v} \wedge x_{u',v} \wedge \bigwedge_{i=3}^{j} l_i \right)}$$

Replace the inference by a derivation that cuts $A \vee x_{u',v}$ with $\neg x_{u,v} \vee \neg x_{u',v}$ to obtain $A \vee \neg x_{u,v}$. Cut this with $A \vee x_{u,v}$ to obtain A. We may proceed through the rest of the proof with A because it subsumes $A \vee x_{u,v} \wedge x_{u',v} \wedge \bigwedge_{i=3}^{j} l_i$.

6.1. Random Restrictions. DEFINITION 6.6. *For a bipartite graph $G = (U \cup V, E)$ and a real number $p \in [0, 1]$, let $\mathcal{M}_p(G)$ denote the distribution on partial assignments which arises from the following experiment:*

Independently, for each $v \in V$, with probability $1 - p$ choose to match v and with probability p leave v unmatched. If v is matched, uniformly select a neighbor u of v, set $x_{u,v}$ to 1, and for every $w \neq u$ that is a neighbor of v, set $x_{w,v}$ to 0. Moreover, for each $v' \neq v$, set $x_{u,v'} = 0$.

Let V_ρ be the set of vertices of V matched by ρ, let U_ρ be the set of vertices of U matched by ρ, and let $S_\rho = U_\rho \cup V_\rho$.

These restrictions randomly associate pigeons with holes in an injective way. While some pigeons can be associated with multiple holes, no two pigeons can be associated with the same hole. It is easy to check that for any $\rho \in \mathcal{M}_p(G)$, we have that $PHP(G) \restriction_\rho = PHP(G - S_\rho)$.

LEMMA 6.7. *Let $p \in [0,1]$, $i \in [k]$ be given. Let $G = (U \cup V, E)$ be a bipartite graph with $\Delta = \Delta(G)$. Let F be an i-DNF in pigeon-normal-form.*

$$Pr_{\rho \in \mathcal{M}_p(G)} [F \restriction_\rho \neq 1] \leq 2^{-\frac{(\log e)(1-p)^i c(F)}{i\Delta^{i+1}}}$$

Proof. For a term T, define the *holes of* T as $\text{Holes}(T) = \{v \mid x_{u,v} \in T \text{ or } \neg x_{u,v} \in T\}$. We say that two terms T and T' are *hole-disjoint* if $\text{Holes}(T) \cap \text{Holes}(T') = \emptyset$.

Because F contains at least $c(F)/i$ many variable-disjoint terms, and each hole $v \in V$ appears in at most Δ many variables, F must contain at least $c(F)/i\Delta$ many hole-disjoint terms.

The events of satisfying hole-disjoint terms are independent, and for a given term, T, the probability that $T \restriction_\rho = 1$ is at least $(1-p)^i/\Delta^i$. This is because with probability $(1-p)^i$, every hole of T is matched, and with probability at least $1/\Delta^i$ the holes are matched in a way that satisfies T (here we use that F is in pigeon-normal-form). Therefore, we have the following inequalities:

$$\Pr_\rho [F \restriction_\rho \neq 1] \leq \left(1 - (1-p)^i/\Delta^i\right)^{\frac{c(F)}{i\Delta}} \leq \left(e^{-(1-p)^i/\Delta^i}\right)^{\frac{c(F)}{i\Delta}} = 2^{-\frac{(\log e)(1-p)^i c(F)}{i\Delta^{i+1}}}$$

□

6.2. Width Lower Bounds for Resolution. For the lower bound proof to work, we need a graph G so that after the application of a random restriction ρ , with high probability, $PHP(G) \restriction_\rho$ requires high width to refute in resolution. We call such graphs *robust*, and in this subsection we probabilistically demonstrate robust, low degree graphs.

DEFINITION 6.8. *A bipartite graph G is said to be* (p, w)-robust*, if when ρ is selected from $\mathcal{M}_p(G)$, with probability at least $\frac{1}{2}$, $w_R(PHP(G) \restriction_\rho) \geq w$.*

All we need for the size lower bound is the following lemma, which is proven probabilistically in the next sub-subsection. Readers who believe that random graphs should be robust can skip to the proof of the lower bound.

LEMMA 6.9. *For all $c > 1$, there exists $d > 0$, so that for n sufficiently large, there exists a $(3/4, n/24)$-robust graph with $\Delta(G) \leq d \log n$ on the vertex sets $[cn]$ and $[n]$.*

6.2.1. Existence of Robust Graphs. As a starting point, we use a now standard lower bound of $w_R(PHP(G))$ in terms of the expansion of G.

DEFINITION 6.10. *For a vertex $u \in U$, let $N(u)$ be its set of neighbors. For a subset $V' \subseteq V$, let its boundary be*

$$\partial V' = \{u \in U \mid |N(u) \cap V'| = 1\}$$

A bipartite graph G is an (m, n, r, f)-expander if $|V| = m$, $|U| = n$, and for all $V' \subseteq V$, $|V'| \leq r$, $|\partial V'| \geq f|V'|$.

THEOREM 6.11. *[8] If G is a bipartite graph that is an (m, n, r, f)-expander, then $w_R(PHP(G)) \geq rf/2$.*

DEFINITION 6.12. *Let $G_{m,n,p}$ be the distribution on bipartite graphs with vertex sets $[m]$ and $[n]$ in which every edge is included with independent probability p.*

The following lemma was proven by Atserias, Bonet and Esteban [3].

LEMMA 6.13. *[3] Let $m = cn$, $q = \frac{48c \ln m}{m}$, $\alpha = \frac{1}{mq}$ and $f = \frac{nq}{6}$. Let G be selected according to the distribution $G_{m,n,q}$.*

$$Pr_G\left[G \text{ is an } (m, n, \alpha m, f) \text{ expander }\right] \geq \frac{2}{3}$$

LEMMA 6.14. *Let $m = cn$, let $q \geq \frac{48c \ln m}{m}$ and let G be selected according to the distribution $G_{m,n,q}$.*

$$Pr_G\left[w_R(PHP(G)) \geq n/12\right] \geq \frac{2}{3}$$

Proof. Let $\alpha = \frac{1}{mq}$ and $f = \frac{nq}{6}$. Because $\alpha m f/2 = (1/mq)m(nq/6)/2 = n/12$, an application of theorem 6.11 shows that when G is selected according to $G_{m,n,\frac{48c \ln m}{m}}$, with probability at least 2/3, $w_R(PHP(G)) \geq n/12$.

Now consider G selected according to $G_{m,n,q}$, with $q \geq \frac{48c \ln m}{m}$. Whenever G_0 is an edge-induced subgraph of G_1, $w_R(PHP(G_1)) \geq w_R(PHP(G_0))$ because a refutation of $PHP(G_1)$ can always be transformed into a refutation of $PHP(G_0)$ by setting some variables to 0. Therefore, by increasing the probability of including an edge, the probability of having no small resolution refutation for $PHP(G)$ only increases. □

We now prove lemma 6.9.

Proof. Set $m = cn$, $p = 3/4$ and $q = \frac{192c \ln m}{m}$. Consider the joint distribution that arises by selecting G according to $G_{m,n,q}$ and ρ according to $\mathcal{M}_{3/4}(G)$. We will bound the probability that the degree is too large, that too many holes are matched, and that the restricted graph is expanding.

By the Chernoff bounds, for each $v \in [n]$ the probability that v has degree in excess of $2mq$ is at most $e^{-mq/4}$. By the union bound, the probability that there exists some $v \in [n]$ of degree in excess of $2mq$ is at most $ne^{-mq/4}$. Similarly, the probability that there exists some $v \in [m]$ of degree in excess of $2mq$ is at most $me^{-nq/4}$. Therefore, the probability that the maximum degree of G exceeds $2mq$ is bounded as follows:

$$ne^{-mq/4} + me^{-nq/4} = ne^{-m192c \ln m/m} + me^{-n192c \ln m/m} = O(n^{-191})$$

Remember that V_ρ is the set of holes matched by the restriction ρ. By the Chernoff bounds, the probability that $|V_\rho| \geq 2n(1-p) = n/2$ is at most $e^{-\frac{n(1-p)}{4}} = e^{-n/16}$.

We now bound the probability that $G - S_\rho$ is an expander. First, up to renaming vertices, $G - S_\rho$ is distributed as $G_{m_\rho,n_\rho,q}$, with $n_\rho = n - |V_\rho|$ and $m_\rho = m - |U_\rho|$. This is because for fixed sets of vertices $V_0 \subseteq V$ and $U_0 \subseteq U$, when we condition on the event that $V_\rho = V_0$ and $U_\rho = U_0$, the edges $\{u, v\}$ with $u \in U \setminus U_0$ and $v \in V \setminus V_0$ are included in $G - S_\rho$ with independent probability q. Now, condition on the event that $|V_\rho| \leq n/2$. We have that $m_\rho = m - |U_\rho| \geq m - n/2 \geq m/2$ and thus $q = 192c \ln m/m \geq 192c \ln m_\rho/m \geq 192c \ln m_\rho/2m_\rho = 48 \cdot 2c \ln m_\rho/m_\rho$. Because $\frac{m_\rho}{n_\rho} \leq \frac{cn}{n/2} = 2c$, we can apply lemma 6.14 and deduce that $w_R(PHP(G - S_\rho)) \geq n_\rho/12$ with probability at least $\frac{2}{3}$. Because $n_\rho \geq n/2$, with the same probability, $w_R(G - S_\rho) \geq n/24$.

Combining the three inequalities from the preceding paragraphs shows that the probability that G contains a vertex of degree in excess of $192c \ln m$, that V_ρ contains more than $n/2$ vertices, or that $w_R(PHP(G - S_\rho)) < n/24$, is at most

$$\frac{1}{3} + O(n^{-191}) + e^{-n/16}$$

For sufficiently large n, this probability is bounded above by $\frac{1}{2}$. By averaging over the choices of the edges, there exists a bipartite graph G on vertex sets $[cn]$ and $[n]$ with $\Delta(G) \leq 2mq = 384c\ln(cn)$, so that upon selection of $\rho \in \mathcal{R}_{3/4}(G)$, $w_R(G - S_\rho) \geq n/24$ with probability at least $\frac{1}{2}$. □

6.3. Size Lower Bounds for Res(k). To prove the size lower bounds for Res(k) refutations of PHP_n^{cn} we first prove size lower bounds for the weak pigeonhole principle restricted to a robust graph, and then we reduce these principles to PHP_n^{cn}.

LEMMA 6.15. *For any $c > 1$ and $d > 0$, there exists $\epsilon > 0$ so that for all n sufficiently large, if $k \leq \sqrt{\log n / \log\log n}$ and G is a $(3/4, n/24)$-robust bipartite graph with vertex sets of sizes cn and n and $\Delta(G) \leq d\log n$, then $S_k(PHP(G)) \geq 2^{n^\epsilon}$.*

Proof. By lemma 6.7, for each $i \in [k]$ and every i-DNF F,

$$\Pr_{\rho\in\mathcal{M}_{3/4}(G)}\left[F \restriction_\rho \neq 1\right] \leq 2^{-\frac{(\log e)(1-3/4)^i c(F)}{i(d\log n)^{i+1}}} = 2^{-\frac{(\log e)c(F)}{i\cdot 4^i(d\log n)^{i+1}}}.$$

In the interest of obtaining a better bound, we will not appeal to corollary 3.4, but directly apply the theorem 3.3. We define sequences $s_0, \ldots, s_k$ and $p_1, \ldots, p_k$ for use in the switching lemma. Set $s_0 = \frac{3}{4k}(n/24 - 1)$. For each $i \in [k]$, set

$$s_i = \left(\frac{\log e}{2i4^i(d\log n)^{i+1}}\right) s_{i-1}$$

For each $i \in [k]$ set $p_i = 2^{-2s_i}$. For any i-DNF F so that $c(F) > s_{i-1}$, we have the following inequality:

$$\Pr_{\rho\in\mathcal{M}_{3/4}(G)}\left[F \restriction_\rho \neq 1\right] < 2^{-\frac{(\log e)s_{i-1}}{i\cdot 4^i(d\log n)^{i+1}}} = 2^{-2\frac{(\log e)s_{i-1}}{2i4^i(d\log n)^{i+1}}} = 2^{-2s_i} = p_i$$

It can be shown that there exists $\epsilon > 0$ so that for sufficiently large n, $s_k \geq n^\epsilon$. To avoid distraction, we show this in in lemma 6.17, at the end of this subsection. Suppose that Γ is a Res(k) refutation of $PHP(G)$ of size less than 2^{n^ϵ}.

By an application of theorem 3.3 and the union bound, we have:

$$\Pr_{\rho\in\mathcal{M}_{3/4}(G)}\left[\exists F \in \Gamma,\ h(F \restriction_\rho) > \sum_{i=0}^{k-1} s_i\right] \leq 2^{n^\epsilon}\sum_{i=1}^{k} p_i 2^{\sum_{j=i}^{k-1} s_j}$$

$$\leq 2^{s_k}\sum_{i=1}^{k} p_i 2^{\sum_{j=i}^{k-1} s_j} = \sum_{i=1}^{k} p_i 2^{\sum_{j=i}^{k} s_j}$$

We now bound $p_i 2^{\sum_{j=i}^{k} s_j}$ for each $i > 0$. For each i, $s_{i+1} < \frac{1}{4}s_i$ so $\sum_{j=i}^{k-1} s_j \leq \frac{4}{3}s_i$. This gives us the following inequality:

$$p_i 2^{\sum_{j=i}^{k-1} s_j} = 2^{\sum_{j=i}^{k-1} s_j - 2s_i} \leq 2^{(4/3-2)s_i} = 2^{-(2/3)s_i} \leq 2^{-(2/3)s_k} \leq 2^{-(2/3)n^\epsilon}$$

Therefore:

$$\Pr_{\rho\in\mathcal{M}_{3/4}(G)}\left[\exists F \in \Gamma,\ h(F \restriction_\rho) > (n/24-1)/k\right] \leq \Pr_{\rho\in\mathcal{M}_{3/4}(G)}\left[\exists F \in \Gamma,\ h(F \restriction_\rho) > \sum_{i=0}^{k-1} s_i\right]$$

$$\leq \sum_{i=1}^{k} p_i 2^{\sum_{j=i}^{k-1} s_j} \leq \sum_{i=1}^{k} 2^{-(2/3)n^\epsilon} \leq k 2^{-(2/3)n^\epsilon} = 2^{\log k - (2/3)n^\epsilon}$$

For n sufficiently large, this probability is strictly less than $1/2$. Because G is $(3/4, n/24)$-robust, for $\rho \in \mathcal{M}_{3/4}(G)$, with probability at least $1/2$, $w_R(PHP(G) \restriction_\rho) \geq n/24$. Thus, there is a ρ so that $w_R(PHP(G) \restriction_\rho) \geq n/24$ and $\forall F \in \Gamma$, $h(F \restriction_\rho) \leq \frac{1}{k}(n/24 - 1)$. This is a contradiction because by corollary 5.2, there is a resolution refutation of $PHP(G) \restriction_\rho$ of width $\leq n/24 - 1$. □

THEOREM 6.16. *For each $c > 1$, 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}.*

Proof. Apply lemma 6.9 and choose d so that for sufficiently large n, there exists a $(3/4, n/24)$-robust graph G on vertex sets cn and n, with $\Delta(G) \leq d \log n$. By lemma 6.15, there exists $\epsilon > 0$ so that for $k \leq \sqrt{\log n / \log\log n}$, $S_k(PHP(G)) \geq 2^{n^\epsilon}$. Because $PHP(G)$ can be obtained by setting some of the variables of PHP_n^{cn} to 0, every Res(k) refutation of PHP_n^{cn} can be converted into a Res(k) refutation of $PHP(G)$ of the same or lesser size. Therefore, all Res(k) refutations of PHP_n^{cn} must have size at least 2^{n^ϵ}. □

Now we prove the lower bound on the number s_k that we used in lemma 6.15. The constants are *not* optimized.

LEMMA 6.17. *There exists $\epsilon > 0$, so that all n sufficiently large, with $k \leq \sqrt{\log n / \log\log n}$ and $s_0, \ldots, s_k$ defined as in the proof of lemma 6.15, $s_k \geq n^\epsilon$.*

Proof. Unwinding the recursive definition of the s_i's, gives the following equality:

$$s_k = \frac{1}{2^k}(\log e)^k \frac{1}{k!}\left(\frac{1}{4}\right)^{\sum_{j=1}^{k} j} \left(\frac{1}{d \log n}\right)^{\sum_{j=2}^{k+1} j} \frac{3}{4k}(n/24 - 1)$$

Because $k \leq \sqrt{\log n / \log\log n}$, we have that $\frac{1}{2^k}(\log e)^k \frac{1}{k!}\left(\frac{1}{4}\right)^{\sum_{j=1}^{k} j} \frac{3}{4k} = n^{-o(1)}$.

$$s_k = n^{-o(1)}(1/d \log n)^{(k+2)(k+1)/2}(n/24 - 1) = n^{-o(1)} 2^{-(\log(d \log n))(k^2+3k+2)/2}(n/24 - 1)$$

Because $k \leq \sqrt{\log n / \log\log n}$ and d is a constant, for n sufficiently large, $(\log(d \log n))(k^2 + 3k + 2)/2 = (\log n)(1 + o(1))/2$. Therefore,

$$s_k = n^{-o(1)} 2^{-(\log n)(1+o(1))/2}(n/24 - 1)$$

and there exists $\epsilon > 0$ so that for all n sufficiently large, $s_k \geq n^\epsilon$. □

7. Lower Bounds for Random CNFs. It is well-known that, in some cases, randomly generated sets of clauses require exponentially large resolution refutations, see [13, 5, 8]. We extend these results by giving exponential lower bounds for the size of Res(k) refutations of randomly chosen sets of width $4k^2 + 2$ clauses.

DEFINITION 7.1. *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.*

Our main result for this section is:

THEOREM 7.2. *For any $\epsilon \in [0, \frac{1}{6})$, there exists $\delta > 0$, so that for n sufficiently large and for $\Delta = n^\epsilon$,*

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

The reason that our proof does not give lower bounds for refutations of random 3-CNFs in 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 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 3.5.

DEFINITION 7.3. *Let F be a CNF in variables $x_1, \ldots, x_n$. We say that F is* (p, r) robust *if*
$Pr_{\rho \in \mathcal{D}_p}[w_R(F \restriction_\rho) \geq r] \geq 1/2$.

It turns out that for sufficiently large w, a random w-CNF is almost surely robust. We state the result below and prove it in the following subsection.

LEMMA 7.4. *There exists a constant c so that for any constants w and t, $w \geq 2t+2$, for every n sufficiently large, and every $\epsilon \in [0, 1/2]$, if we set $\Delta = n^\epsilon$ then the following inequality holds:*

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

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 3.7.

LEMMA 7.5. *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 $S_k(F) \geq \frac{1}{4k} 2^{(\gamma_k (r-1)/k\sqrt{n})}$.*

Proof. Suppose that Γ is a Res(k) refutation of F of size at most $\frac{1}{4k} 2^{(\gamma_k (r-1)/k\sqrt{n})}$. By corollary 3.7, with $p = n^{-1/2k^2}$ and $w = (r-1)/k$, we have that for every line F of Γ, $\Pr_{\rho \in \mathcal{D}_p}[h(F \restriction_\rho) > (r-1)/k] \leq k 2^{-\gamma_k (r-1)/k\sqrt{n}}$. By the union bound we have that

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

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

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 5.2 there should be a width $r-1$ resolution refutation of $F \restriction_\rho$. □

Combining lemmas 7.4 and 7.5 with $t = 2k^2$, $w = 4k^2+2$ and $r = cn^{\frac{1-2\epsilon}{1+2\epsilon}}$ shows that a random $(4k^2+2)$-CNF almost surely requires exponential size to refute in Res(k).

COROLLARY 7.6. *There exists a constant c so that for every k, for every n sufficiently large and $\epsilon \in [0, 1/2]$, if we set $\Delta = n^\epsilon$, then the following inequality holds.*

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

This gives an exponential lower bound only when $\frac{1-2\epsilon}{1+2\epsilon} > \frac{1}{2}$. This holds exactly for $\epsilon \in [0, \frac{1}{6})$.

THEOREM 7.7. *For any $\epsilon \in [0, \frac{1}{6})$, there exists $\delta > 0$, so that for n sufficiently large and for $\Delta = n^{\epsilon}$,*

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

7.1. Robustness of Random CNFs. In this section we show that for appropriate clause densities, a random w-CNF is almost surely robust.

We begin with a width bound for resolution refutations of random 3-CNFs given by Ben-Sasson and Wigderson.

THEOREM 7.8. *[8] There exists a constant c, so that for all n, and all $\epsilon \in [0, 1/2]$ with $\Delta = n^{\epsilon}$, the following inequality holds.*

$$Pr_{F \in \mathcal{F}_3^{n,\Delta}} \left[w_R(F) \leq cn^{\frac{1-2\epsilon}{1+2\epsilon}} \right] = o(1)$$

LEMMA 7.9. *There exists a constant c so that for any constants w and t with $w \geq 2t+2$, for every n sufficiently large and $\epsilon \in [0, 1/2]$, if we set $\Delta = n^{\epsilon}$ and $p = n^{-1/t}$, then the following inequality holds:*

$$Pr_{\substack{F \in \mathcal{F}_w^{n,\Delta} \\ \rho \in \mathcal{D}_p}} \left[w_R(F \restriction_\rho) \leq cn^{\frac{1-2\epsilon}{1+2\epsilon}} \right] = o(1)$$

Proof. Let w, t, n, ϵ be given as above and set $\Delta = n^{\epsilon}$ and $p = n^{-1/t}$.

Because the expected size of $\mathrm{dom}(\rho)$ is pn, the Chernoff bounds show that the size of $\mathrm{dom}(\rho)$ exceeds $2pn = 2n^{1-\frac{1}{t}}$ with probability at most $e^{-n^{1-1/t}/4} = o(1)$.

Let C be a fixed clause of width w that contains no opposite literals. When we choose $\rho \in \mathcal{D}_p$, the probability that the domain of ρ contains at least $w-2$ variables of C is at most $\binom{w}{2} p^{w-2}$. Because w is a constant, this probability is $O\left(n^{-(w-2)/t}\right)$. Because $w \geq 2t+2$, this probability is $O\left(n^{-2}\right)$. For any fixed w-CNF F on Δn many clauses, an application of the union bound shows that there is some clause with $\geq w-2$ of its variables in the restriction with probability $O(\Delta n \cdot n^{-2}) = o(1)$. Because this calculation holds for every w-CNF of Δn many clauses, we have that

$$\Pr_{\substack{F \in \mathcal{F}_w^{n,\Delta} \\ \rho \in \mathcal{D}_p}} \left[\exists C \in F,\ |vars(C) \setminus \mathrm{dom}(\rho)| \leq 2 \right] = o(1)$$

Fix a restriction ρ so that $\mathrm{dom}(\rho) \leq 2n^{1-\frac{1}{t}}$ and let $n' = n - |\mathrm{dom}(\rho)|$. Conditioned on the event that $\forall i \in [\Delta]$, $|vars(C_i) \setminus \mathrm{dom}(\rho)| \geq 3$, $F \restriction_\rho$ is subsumed by a random 3-CNF distributed as $\mathcal{F}_3^{n',\Delta}$. (To see this, consider the distribution on 3-CNFs that chooses three literals unset by ρ from each C_i.) Choose ϵ' so that $n^{\epsilon} = (n')^{\epsilon'}$. Adding up the conditional probabilities and applying theorem 7.8 shows that

$$\Pr_{\substack{F \in \mathcal{F}_w^{n,\Delta} \\ \rho \in \mathcal{D}_p}} \left[w_R(F \restriction_\rho) \leq cn'^{\frac{1-2\epsilon'}{1+2\epsilon'}} \right] = o(1)$$

Because $n' \geq n - 2n^{1-1/t}$, we may choose c' so that

$$\Pr_{\substack{F\in\mathcal{F}_w^{n,\Delta}\\ \rho\in\mathcal{D}_p}}\left[w_R(F\restriction_\rho)\leq c'n^{\frac{1-2\epsilon}{1+2\epsilon}}\right]=o(1)$$

□

From lemma 7.9, an averaging argument yields the following phrasing of lemma 7.4.

LEMMA 7.10. *There exists a constant c so that for any constants w and t, $w\geq 2t+2$, for every n sufficiently large and $\epsilon\in[0,1/2]$, if we set $\Delta=n^\epsilon$ and let $p=n^{-1/t}$, then the following inequality holds:*

$$Pr_{F\in\mathcal{F}_w^{n,\Delta}}\left[Pr_{\rho\in\mathcal{D}_p}\left[w_R(F\restriction_\rho)\leq cn^{\frac{1-2\epsilon}{1+2\epsilon}}\right]\geq 1/2\right]=o(1)$$

8. Separation Between Res(k) and Res$(k+1)$. In this section we show that for each constant k, there is an $\epsilon_k>0$ and 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 [19, 9].

DEFINITION 8.1. *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, the principle $GOP(G)$ consists of $O(n^3)$ many clauses each of width at most $\max\{3,d\}$. Therefore, for any graph G on n vertices with maximum degree d, the principle $GOP^k(G)$ has size $O(n^3k^d)$.

It is readily 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 $GOP^{k+1}(G)$, as shown in lemma 8.4. On the other hand, we will also prove that Res(k) refutations of $GOP^{k+1}(G)$ require exponential size for certain graphs:

THEOREM 8.2. *Let k be a positive integer. 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 $2^{\Omega(n^{\epsilon_k})}$.*

8.1. The Upper Bounds. We build Res(k) refutations for $GOP^k(G)$ from resolution refutations of $GOP(G)$.

The resolution refutation of $GOP(G)$ is a slight variation of the resolution refutation of GT_n [19, 9].

LEMMA 8.3. *Let G be an n vertex graph. There is a resolution refutation of $GOP(G)$ of size $O(n^3)$.*

Proof. To construct the resolution refutation of $GOP(G)$, we iteratively derive the formulas $\bigvee_{\substack{i\in[l,n]\\ i\neq j}}X_{i,j}$ for all i,j with $1\leq l\leq j\leq n$. The clauses $\bigvee_{\substack{i\in[n]\\ i\neq j}}X_{i,j}$ are

derived by weakening the hypotheses. We proceed in stages as l ranges from 1 up to n. At stage l, for $j = l$, we have $\bigvee_{i\in[l+l,n]} X_{i,l}$. For $j \neq l$, we resolve $\bigvee_{i\in[l+1,n]} X_{i,l}$ with the transitivity axioms $\neg X_{i,l} \vee \neg X_{l,j} \vee X_{i,j}$ to obtain $\neg X_{l,j} \vee X_{j,l} \bigvee_{\substack{i\in[l+1,n]\\ i\neq j}} X_{i,j}$. This clause is resolved with $\neg X_{l,j} \vee \neg X_{j,l}$ and $\bigvee_{\substack{i\in[l,n]\\ i\neq j}} X_{i,j}$ to obtain $\bigvee_{i=l+1}^{n} X_{i,j}$. At stage n, with $j = n$, we have derived the empty clause. This refutation clearly has size $O(n^3)$. □

LEMMA 8.4. *For each k, and every G with n vertices and degree at most $d \geq 3$, $GOP^k(G)$ has a Res(k) refutation of size $O(n^3k^d)$.*

Proof. Let τ be the operation that replaces $X_{u,v}$ by $\bigwedge_{i=1}^{k} X^i_{u,v}$ and $\neg X_{u,v}$ by $\bigvee_{i=1}^{k} \neg X^i_{u,v}$.

Let Γ be the size $O(n^3)$ resolution refutation of $GOP(G)$ given above, and remove all of its weakening inferences. If we apply the transformation τ to the refutation, we obtain a Res(k) refutation of $\tau(GOP(G))$.

From the clauses of $GOP^k(G)$ we can derive the k-DNFs of $\tau(GOP(G))$ by a sequence of $O(k^d)$ many AND-introduction inferences per formula. Thus, we have a Res(k) refutation of $GOP^k(G)$ of the claimed size. □

8.2. Random Restrictions. In this subsection we define a distribution on partial assignments so that i-DNFs with high cover number are satisfied with high probability. The idea is to randomly color the graph with $4k$ many colors, and then between vertices u and v of distinct color classes, uniformly choose an assignment to $X^1_{u,v}, \ldots, X^{k+1}_{u,v}, X^1_{v,u}, \ldots, X^{k+1}_{v,u}$ which makes both $\bigwedge_{i=1}^{k+1} X^i_{u,v}$ and $\bigwedge_{i=1}^{k+1} X^i_{v,u}$false.

DEFINITION 8.5. *Let $k \geq 1$ be given. Let G be a graph. The distribution $\mathcal{P}_{k+1}(G)$ on partial assignments ρ to the variables of $GOP^{k+1}(G)$ is given by the following experiment.*

For each $(u, v) \in V(G)^2$, let $\sigma^{u,v}_\rho$ be chosen uniformly among $0, 1$ assignments to $X^1_{u,v}, \ldots, X^{k+1}_{u,v}$ so that for at least one $i \in [k + 1]$, $\sigma^{u,v}_\rho(X^i_{u,v}) = 0$.

Select a random coloring of $V(G)$ by $4k$ many colors, $c_\rho : V(G) \to [4k]$.

The partial assignment ρ is defined as:

$$\rho = \bigcup_{\substack{(u,v)\in V(G)^2\\ c_\rho(u)\neq c_\rho(v)}} \sigma^{u,v}_\rho$$

The auxiliary total assignment, σ_ρ, is defined as:

$$\sigma_\rho = \bigcup_{(u,v)\in V(G)^2} \sigma^{u,v}_\rho$$

If we let B_ρ be the set of edges which are bichromatic under the coloring c_ρ, then $GOP^{k+1}(G) \restriction_\rho$ is $GOP^{k+1}(G \setminus B_\rho)$. Moreover, we have the following lemma, which the reader can easily check.

LEMMA 8.6. *Let G be a graph. Let $\rho \in \mathcal{P}_{k+1}(G)$ be given. Let B_ρ be the set of edges of G that are bichromatic under c_ρ. Let $G_1, \ldots, G_m$ be the connected components of $G \setminus B_\rho$.*

$$GOP^{k+1}(G) \restriction_\rho = \bigcup_{j=1}^{m} GOP^{k+1}(G_j)$$

Formulas with high cover number contain many variable-disjoint terms, but the events of satisfying these terms with $\rho \in \mathcal{P}_{k+1}(G)$ are not necessarily independent. To obtain independence, we look at the pairs of vertices involved with the literals of the terms. Remember that in the definition of $GOP(G)$, there are no variables $X_{u,u}$.

DEFINITION 8.7. *Let $X^i_{u,v}$ be a variable of $GOP^{k+1}(G)$. The* underlying pair *of $X^i_{u,v}$ is the set $\{u,v\}$. The* underlying ordered pair *of $X^i_{u,v}$ is (u,v). Let T be a term. The set of* vertex pairs of T, P_T, *is defined as*

$$P_T = \{\{u,v\} \mid \ \{u,v\} \text{ is the underlying pair of a variable in } T\}$$

The set of vertices of T, S_T, *is defined as $S_T = \bigcup P_T$.*

We use combinatorial sunflowers to obtain independence between the events of satisfying terms of an i-DNF with high cover number. To guarantee that such a system exists, we apply the Erdös-Rado lemma.

DEFINITION 8.8. *A (p,l)* sunflower *is a collection of sets $P_1, \ldots, P_p$, each of size $\leq l$, so that there exists a set C so that $P_i \cap P_j = C$ for all $i,j \in [p]$, $i \neq j$. The set C is called the* core *of the sunflower.*

THEOREM 8.9. *([16], c.f. [22]) Let l be given. Let $\mathcal{Z}$ be a family of M distinct sets, each with cardinality $\leq l$. $\mathcal{Z}$ contains a (p,l) sunflower where $p \geq \left(\frac{M}{l!}\right)^{\frac{1}{l}}$.*

DEFINITION 8.10. *Let $T_1, \ldots, T_t$ be terms in the variables of $GOP^{k+1}(G)$. We say that the terms are* sufficiently independent *if the following conditions hold:*
1. *For $i,j \in [t]$, if $i \neq j$ then $S_{T_i} \neq S_{T_j}$.*
2. *The family $\{S_{T_i} \mid 1 \leq i \leq t\}$ forms a sunflower with core C.*
3. *For each $i \in [t]$, each $\{u,v\} \in P_{T_i}$, $\{u,v\} \not\subseteq C$.*

LEMMA 8.11. *Let $T_1, \ldots, T_t$ be a sufficiently independent set of terms. The sets P_{T_i}, $1 \leq i \leq t$, are disjoint.*

Proof. Let i,j, $1 \leq i < j \leq t$ be given and let C denote the core of the sunflower. Suppose that $\{u,v\} \in P_{T_i} \cap P_{T_j}$. We then have that $\{u,v\} \subseteq S_{T_i} \cap S_{T_j}$, so $\{u,v\} \subseteq C$. Therefore, by the third property of sufficient independence, $\{u,v\} \notin P_{T_i}$ – contradiction. □

We begin the task of showing that a DNF with high cover number is likely to be satisfied by a random restriction. The quality of our bounds is most affected by the use of the sunflower lemma, and the particular constants we obtain at other points have limited impact. Therefore, to conserve space and readability, we will not optimize many of the probabilities involved.

LEMMA 8.12. *Let k be given. There exist a constants $\beta_k > 0$ and $c_k > 0$ so that for every k-DNF F in the variables of $GOP^{k+1}(G)$, F contains a sufficiently independent set of size at least $\beta_k (c(F))^{\frac{1}{2k}} - c_k$.*

Proof. F contains a set of $s = c(F)/k$ many variable-disjoint terms, $T_1, \ldots, T_s$. It is possible that $S_{T_m} = S_{T_l}$ for some $m \neq l$. However, because all terms have size at most k, for each i, $|S_{T_i}| \leq 2k$, and a set of $\leq 2k$ many vertices can be the underlying set of fewer than $\left((k+1)4k^2\right)^k$ many different variable-disjoint terms (since a variable $X^i_{u,v}$ is determined by an ordered pair $(u,v) \in [S_T]^2$ and $i \in [k+1]$). Therefore, there is a sub-collection of of $\frac{s}{((k+1)4k^2)^k}$ many variable-disjoint terms whose underlying sets of vertices are distinct.

Because the underlying sets of vertices have size at most $2k$, we can apply the sunflower lemma to find $s' = \left(\frac{s}{((k+1)4k^2)^k(2k)!}\right)^{\frac{1}{2k}} = \left(\frac{c(F)}{k((k+1)4k^2)^k(2k)!}\right)^{\frac{1}{2k}}$ many terms whose sets of underlying vertices form an $(s', 2k)$ sunflower. We rename these terms $T_1, \ldots, T_{s'}$.

Let C be the core of the sunflower $S_{T_1}, \ldots, S_{T_{s'}}$. Notice that $|C| \leq 2k$. Call a variable *bad* if both of its underlying vertices belong to C. There are fewer than $2k^2$ many unordered pairs of vertices contained in C, and each is the underlying pair of exactly $2(k+1)$ many variables. Therefore, there are fewer than $2(k+1) \cdot 2k^2 = 4k^2(k+1)$ many variables whose underlying vertices are both in C. The terms $T_1, \ldots, T_{s'}$ are variable-disjoint, so each bad variable appears in at most one term, and when we remove all terms containing a bad variable, we obtain a sufficiently independent set of terms of size $s' - 4k^2(k+1) = \left(\frac{c(F)}{k(2k)!((k+1)4k^2)^k}\right)^{\frac{1}{2k}} - 4k^2(k+1)$. □

Before we bound the probability of satisfying a DNF with high covering number, we make a few observations.

FACT 1. *Let T be a term, and let $\rho \in \mathcal{P}_{k+1}(G)$ be given. $T \restriction_\rho = 1$ if and only if the following two events occur: (i) $T \restriction_{\sigma_\rho} = 1$ and (ii) For each $\{u, v\} \in P_T$, $c_\rho(u) \neq c_\rho(v)$.*

For each term T in a Res(k) refutation of $GOP^{k+1}(G)$, because T contains at most k literals, there is a nonzero chance that it will be satisfied by σ_ρ when ρ is a random restriction chosen according to $\mathcal{P}_{k+1}(G)$. This is made precise in the following lemma. Recall that by construction, $\bigwedge_{i=1}^{k+1} X^i_{u,v} \restriction_{\sigma_\rho} = 0$, so the lemma fails for terms of size $k+1$ (as it should, since we are separating Res($k+1$) from Res(k)). The argument is similar to that in lemma 4.4.

LEMMA 8.13. *Let T be a term of size at most k.*

$$Pr_{\rho \in \mathcal{P}_{k+1}(G)} \left[T \restriction_{\sigma_\rho} = 1 \right] \geq \frac{1}{3^k}$$

Proof. Order the literals of T as $l_1, \ldots, l_k$. For each j, $1 \leq j \leq k$, if we condition on the event that each of $l_1, \ldots, l_{j-1}$ is satisfied, then the probability of l_j being satisfied is at least 1/3. This is because in the worst case, l_j is a literal $X^{i_j}_{u,v}$ and the other literals are $X^{i_1}_{u,v}, \ldots, X^{i_{j-1}}_{u,v}$, and in this case, the probability that $X^{i_j}_{u,v}$ is satisfied by σ_ρ is at least 1/3. □

LEMMA 8.14. *Let G be graph and let k be a positive integer. Let F be a k-DNF which contains contains t sufficiently independent terms.*

$$Pr_{\rho \in \mathcal{P}_{k+1}(G)} \left[F \restriction_\rho \neq 1 \right] \leq \left(1 - \frac{1}{3^k 2^{2k}}\right)^t$$

Proof. Let $T_1, \ldots, T_t$ be the sufficiently independent terms of F. Let C be the core of the sunflower $S_{T_1}, \ldots, S_{T_t}$. Fix a coloring of the vertices in the core, $\chi : C \to [4k]$. Condition on the event that $c_\rho \restriction_C = \chi$.

We now lower-bound the probability that a given term T of the sufficiently independent set is satisfied. First, we bound the probability that every underlying edge of T is bichromatic. Note that by property (3) of sufficient independence, for all $\{u, v\} \in P_T$, $\{u, v\} \not\subseteq C$, so it suffices to bound the probability that the vertices in $S_T \setminus C$ receive distinct colors not in the range of χ. Therefore, the probability that every pair in P_T is bichromatic, conditioned on $c_\rho \restriction_C = \chi$, is at least $1/2^{2k}$. Because T contains at most k literals, the probability that $T \restriction_{\sigma_\rho} = 1$ is at least $\frac{1}{3^k}$. These two events are independent, so we have $\Pr_{\rho \in \mathcal{P}_{k+1}(G)} \left[T \restriction_\rho = 1 \mid c_\rho \restriction_C = \chi \right] \geq \frac{1}{2^{2k}} \frac{1}{3^k}$.

Now we show that (when we condition on the event that $c_\rho \restriction_C = \chi$) the events $T_i \restriction_\rho = 1$ are totally independent. Because the terms share no underlying pairs, the

events $T_i \restriction_{\sigma_\rho}$ are independent of the satisfaction of other terms. The events "for each $\{u, v\} \in P_{T_i}$, $c_\rho(u) \neq c_\rho(v)$" are independent of the satisfaction of other terms. This is because once we condition on the event $c_\rho \restriction_C = \chi$, the probability that every pair of P_{T_i} is bichromatic under c_ρ depends only on the values that c_ρ takes on $S_{T_i} \setminus C$ and, for all $i \neq j$, $S_{T_i} \cap S_{T_j} = C$.

Combining the results of the previous two paragraphs shows that $\Pr_{\rho \in \mathcal{P}_{k+1}(G)}[F \restriction_\rho \neq 1 \mid c_\rho \restriction_C = \chi] \leq \left(1 - \frac{1}{3^k 2^{2k}}\right)^t$. Because this holds for all colorings $\chi : C \to [4k]$, we have that $\Pr_{\rho \in \mathcal{P}_{k+1}(G)}[F \restriction_\rho \neq 1] \leq \left(1 - \frac{1}{3^k 2^{2k}}\right)^t$ □

We now have the lemma relating cover number to the probability that a restriction satisfies a k-DNF.

LEMMA 8.15. *For each k there exist positive constants δ, γ and d so that for any k-DNF F:*

$$Pr_{\rho \in \mathcal{P}_{k+1}(G)}[F \restriction_\rho \neq 1] \leq d 2^{-\delta(c(F))^\gamma}$$

Proof. By lemma 8.12 F contains a sufficiently independent set of size at least $\beta_k(c(F))^{\frac{1}{2k}} - c_k$

By lemma 8.14:

$$\Pr_{\rho \in \mathcal{P}_{k+1}(G)}[F \restriction_\rho \neq 1] \leq \left(1 - \frac{1}{3^k 2^{2k}}\right)^{\beta_k(c(F))^{\frac{1}{2k}} - c_k}$$

Because k is fixed, this concludes the proof with $\delta = -\beta_k \log\left(1 - \frac{1}{3^k 2^{2k}}\right)$, $\gamma = 1/2k$ and $d = \left(1 - \frac{1}{3^k 2^{2k}}\right)^{-c_k}$.
□

8.3. Width Lower Bound for Resolution. In this subsection we show that for each n, if G is a graph on n vertices satisfying a certain expansion-like property, then $w_R(GOP^{k+1}(G)) = \Omega(n)$. Combining this with a probabilistic calculation will that there exist graphs G so that for $\rho \in \mathcal{P}_{k+1}(G)$, with probability at least $1/2$, $w_R(GOP^{k+1}(G) \restriction_\rho) = \Omega(n)$.

The proof of the resolution width bound is similar to the one used by Bonet and Galesi for the GT_n principles [9]. They worked with complete graphs, but we do not because the principles $GOP^2(K_n)$ have size in excess of 2^n. Fortunately, for the proof technique to work, G need not be complete but instead have the following property:

DEFINITION 8.16. *Let G be an undirected graph on n-vertices. We say that G is ϵ-neighborly if between every pair of disjoint sets of vertices, $A, B \subseteq V(G)$ with $|A|, |B| \geq \epsilon n$, there exists an edge joining A and B.*

We now show that resolution refutations of $GOP(G)$ require large width when G is a connected, neighborly graph.

LEMMA 8.17. *If G is a connected graph of n vertices that is ϵ-neighborly, then every resolution refutation of $GOP(G)$ contains a clause of width $\left(\frac{1-3\epsilon}{6}\right) n$.*

Proof. We begin by defining the "measure" of a clause. A *critical truth assignment*, or *cta*, is an assignment to the variables of $GOP(G)$ which forms a total order on $V(G)$. For each $v \in V(G)$, let $C_v := \bigvee_{u \in N(v)} X_{u,v}$, and for each $I \subseteq V(G)$, $C_I := \bigwedge_{v \in I} C_v$. Let C be a clause. The *measure of C, $\mu(C)$*, is the minimum cardinality of a set $I \subseteq V(G)$ so that for every cta α, if α satisfies C_I then α satisfies C.

Notice that if a clause $A \vee B$ is the resolvent of $A \vee x$ and $B \vee \neg x$ then $\mu(A \vee B) \leq \mu(A \vee x) + \mu(A \vee \neg x)$. Because of this, we say that μ is *subadditive with respect to resolution*. If $A \subseteq B$, then we have that $\mu(B) \leq \mu(A)$, so μ is *decreasing with respect to subsumption*.

We now show that $\mu(\emptyset) = n$. Suppose otherwise, and let I be a subset of $V(G)$ with $|I| \leq n-1$. Choose one vertex $v_0 \in V(G) \setminus I$ and let α be a total order which arises by taking a depth-first search of G starting with v_0. Clearly α satisfies C_I but α does not satisfy $\emptyset$.

Because every clause of $GOP(G)$ has measure either 0 or 1, the empty clause has measure n and the measure is both subadditive with respect to resolution and decreasing with respect to subsumption, there must exist a clause C so that $\frac{n}{3} \leq \mu(C) \leq \frac{2n}{3}$. Suppose for the sake of contradiction that $w(C) < \frac{n-3\epsilon n}{6}$.

Let I be a minimal subset of $V(G)$ so that for every critical truth assignment α, if α satisfies C_I then α satisfies C. Let $J = V(G) \setminus I$. Notice that $|I|, |J| \geq \frac{n}{3}$.

Let S be the set of vertices mentioned by variables of C. Clearly, $|S| \leq 2w(C) < 2\left(\frac{n-3\epsilon n}{6}\right) = \frac{n-3\epsilon n}{3}$. Therefore, $|I \setminus S| \geq \frac{n}{3} - \frac{n-3\epsilon n}{3} = \epsilon n$. Similarly, $|J \setminus S| \geq \epsilon n$. Because G is ϵ-neighborly, we may choose $u \in I \setminus S$ and $v \in J \setminus S$ so that $\{u, v\}$ is an edge of G.

Let α be a critical truth assignment so that α satisfies $C_{I \setminus \{u\}}$ but α does not satisfy C_u and α does not satisfy C. Let β be the critical truth assignment which arises by moving v to the front of the order given by α. For $w \in I,\ w \neq u$, β satisfies C_w because every predecessor of w in α is a predecessor of w in β. For u, β satisfies C_u because β satisfies $X_{v,u}$. However, β does not satisfy C because α does not satisfy C and no variable mentioning u or v appears in C. Therefore, β satisfies C_I but β does not satisfy C, a contradiction to the choice of I. □

A resolution refutation of $GOP^k(G)$, $k \geq 1$, can be transformed into a resolution refutation of $GOP(G)$ by setting the appropriate variables to 1. Applying a restriction does not increase the width of a resolution refutation, so we have the following corollary:

COROLLARY 8.18. *If G is a connected graph of n vertices that is ϵ-neighborly, then for all $k \geq 1$, $w_R(GOP^k(G)) \geq \left(\frac{1-3\epsilon}{6}\right) n$.*

8.4. Robust Graphs. DEFINITION 8.19. *We say that a graph G is r-robust if for ρ selected at random by $\mathcal{P}_{k+1}(G)$, with probability at least $\frac{3}{4}$, $w_R(GOP(G) \restriction_\rho) \geq r$.*

To guarantee that the restricted principle will require high width to refute, it suffices that the graph obtained by deleting the bichromatic edges should consist of large, neighborly connected components. Random graphs of degree $\Theta(\log n)$ have this property with high probability. This is shown in the following subsection.

LEMMA 8.20. *There exists a constant c, so that for sufficiently large n, there exists an $\frac{n}{96k}$-robust graph G on n vertices with degree at most $c \log n$.*

The proof of this lemma is a standard probabilistic argument. The reader may skip its proof in the following sub-subsection, and move directly to the proof of the lower bound in the next subsection.

8.4.1. Demonstration of Robust Graphs. An easy probabilistic argument shows that with very high probability, a random graph of expected degree $\Theta((1/\epsilon)\log n)$ is almost surely ϵ-neighborly.

LEMMA 8.21. *Let n and d be positive integer so that $d \leq n$ and let $p = d/n$. Let $G_{n,p}$ be the distribution on graphs on n vertices in which every edge is included with independent probability p. With probability $\leq e^{2\epsilon n(1+\ln(1/\epsilon))-d\epsilon^2 n}$, a graph selected*

according to $G_{n,p}$ is not *ϵ-neighborly.*

Proof. There are fewer than $\binom{n}{\epsilon n}^2 \leq \left(\frac{en}{\epsilon n}\right)^{2\epsilon n} = e^{2\epsilon n(1+\ln(1/\epsilon))}$ many pairs of disjoint sets of ϵn many vertices. Each such pair has a chance of at most $(1-p)^{\epsilon^2 n^2}$ of being unconnected. However, $(1-p)^{\epsilon^2 n^2} = (1-\frac{d}{n})^{\epsilon^2 n^2} \leq e^{-d\epsilon^2 n}$, so by the union bound the probability is at most $e^{2\epsilon n(1+\ln(1/\epsilon))} e^{-d\epsilon^2 n} = e^{2\epsilon n(1+\ln(1/\epsilon))-d\epsilon^2 n}$. □

We now show that a random graph will probably have each component large and neighborly if we randomly partition it into vertex-induced subgraphs.

LEMMA 8.22. *For all $\epsilon > 0$, and all integers $k \geq 1$, there exists a constant c, so that for sufficiently large n, there exists graph G with $\Delta(G) \leq 2c \log n$ so that upon the random partition of G into $4k$ many vertex-induced subgraphs, with probability at least $1/2$, each component has size at least $n/8k$ and is ϵ-neighborly.*

Proof. Let $p = \frac{c \log n}{n}$. We will solve for the value of c at the end. Consider the following experiment: select a graph G according to the distribution $G_{n,p}$, and then independently color each vertex with one of $4k$ colors, then remove all bichromatic edges to form $4k$ vertex induced subgraphs, $G_1, \ldots, G_{4k}$.

Let P be the probability that G has a vertex of degree $> 2c \log n$, or that one of the induced subgraphs has size $< \frac{n}{8k}$, is disconnected or is not ϵ-neighborly. We now bound this probability.

Consider the probability that G has a vertex of degree $\geq 2c \log n$. By the Chernoff bounds, the probability of any one vertex having degree in excess of $2p(n-1)$ is no more than $e^{-p(n-1)/4} = e^{-c(\log n)(n-1)/4n}$. Therefore, the probability of there existing a vertex with degree in excess of $2p(n-1)$ is no more than $ne^{-c(\log n)(n-1)/4n}$.

The Chernoff bounds also allow us to bound the probability that any of the G_i's contain too few vertices. The probability that a given color class of the partition contains fewer than $\frac{1}{2} \cdot \frac{n}{4k} = \frac{n}{8k}$ vertices is bounded by $e^{-\frac{n}{64k}}$.

Once we condition upon all pieces of the partition containing at least $\frac{n}{8k}$ vertices, we can bound the probability that any induced subgraph is disconnected. Consider a fixed set of $s \geq \frac{n}{8k}$ many vertices, and condition upon the event those vertices receive the same color in the partition. Each edge internal to the set is included with probability $\frac{c \log n}{n} = \frac{(cs/n) \log n}{s} \geq \frac{(c/8k) \log s}{s}$. By a standard result on the connectivity of random graphs (c.f. [27]), each color class is disconnected with probability bounded by $O\left(1/n^{(c/8k)-1}\right)$.

Finally, we consider the probability that each of the components G_i is ϵ-neighborly. For a fixed set of $s \geq \frac{n}{8k}$ vertices, if we condition on the event that set forms a component after partition, each internal edge is included with probability $\frac{c \log n}{n} \geq \frac{(c/8k) \log s}{s}$. By lemma 8.21, that means that the component is *not* ϵ-neighborly with probability at most $e^{2\epsilon s(1+\ln(1/\epsilon))-(c/8k)(\log s)\epsilon^2 s} = e^{-\Omega(n \log n)}$.

Therefore,

$$P \leq ne^{-c(\log n)(n-1)/4n} + 4ke^{-\frac{n}{64k}} + O\left(4k/n^{(c/8k)-1}\right) + e^{-\Omega(n \log n)}$$

For a sufficiently large constant c, dependent only on k and ϵ, this is below $\frac{1}{4}$.

Therefore, by an averaging argument on the edge choices, there exists a graph G of maximum degree $\leq 2c \log n$ so that upon random partition of its vertices into $4k$ color classes, its induced subgraphs are each connected, of size $\geq \frac{n}{8k}$, and ϵ-neighborly with probability $\geq \frac{3}{4}$. □

We now prove lemma 8.20.

Proof. Using lemma 8.22, choose c so that for sufficiently large n, there exists graph G so that upon the random partition of G into $4k$ many vertex-induced subgraphs, with probability at least $3/4$, each component has size at least $n/8k$ and is $(1/6)$-neighborly. Therefore we may choose $\rho \in \mathcal{P}_{k+1}(G)$ so that for each $i \in [4k]$, $w_R(GOP^{k+1}(G_i)) \geq \left(\frac{n}{8k}\right)\left(\frac{1-3\frac{1}{6}}{6}\right) = \frac{n}{96k}$ (by lemma 8.18).

Let Γ be a resolution refutation of $GOP^{k+1}(G) \restriction_\rho$. By lemma 8.6, $GOP^{k+1}(G) \restriction_\rho = \bigcup_{i=1}^{4k} GOP^{k+1}(G_i)$. However, for each $i, j \in [4k]$, $i \neq j$, we have that $GOP^{k+1}(G_i)$ and $GOP^{k+1}(G_j)$ are variable disjoint. Therefore, by lemma 2.3, for some $i \in [4k]$, there exists a resolution refutation of $GOP^{k+1}(G_i)$ of width is at most $w(\Gamma)$. However, by the preceding paragraph, each $GOP^{k+1}(G_i)$ requires width $\frac{n}{96k}$ to refute in resolution. Therefore $w(\Gamma) \geq \frac{n}{96k}$. □

8.5. The Lower Bound. THEOREM 8.23. *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 $2^{\Omega(n^{\epsilon_k})}$.*

Proof. Let k be given. Apply lemma 8.20 and choose c so that for sufficiently large n, there exists a $\frac{n}{96k}$-robust graph G on n vertices with degree at most $c \log n$. By lemma 8.15, there are positive constants d, δ and γ so that for every k-DNF F $\Pr_{\rho \in \mathcal{P}_{k+1}(G)}[F \restriction_\rho \neq 1] \leq d2^{-\delta(c(F))^\gamma}$. By corollary 3.4, with $s = (\frac{n}{96k} - 1)/k$, for every k-DNF F:

$$\Pr_{\rho \in \mathcal{P}_{k+1}(G)}[h(F \restriction_\rho) > (n/96k - 1)/k] \leq dk2^{-2\delta^k((n/96k-1)/k)^{\gamma^k}/4^k}$$

Because d, γ and δ depend only on k, there exists ϵ_k so that

$$\Pr_{\rho \in \mathcal{D}}[h(F \restriction_\rho) > (n/96k - 1)/k] \leq \frac{1}{2}2^{-n^{\epsilon_k}}$$

Suppose for the sake of contradiction that Γ is a Res(k) refutation of $GOP^{K+1}(G)$ of size less than $2^{n^{\epsilon_k}}$, By the union bound, with probability at least $1/2$, every line F of Γ has $h(F \restriction_\rho) \leq (n/96k - 1)/k$. On the other hand, because G is $(n/96k)$-robust, $w_R(GOP^{k+1}(G) \restriction_\rho) \geq n/96k$ with probability at least $3/4$. So we may choose $\rho \in \mathcal{P}_{k+1}(G)$ so that $w_R(GOP^{k+1}(G) \restriction_\rho) \geq n/96k$, and for all lines F of Γ, $h(F \restriction_\rho) \leq (n/96k - 1)/k$. By corollary 5.2, $GOP^{k+1}(G) \restriction_\rho$ has a resolution refutation of width at most $n/96k - 1$. Contradiction. □

9. Separating Res(k) and Res($k+1$) with Constant Width Clauses. The separation between Res($k+1$) and Res(k) given by theorem 8.23 uses sets of clauses whose maximum width is $\Theta(\log n)$. In this section we present a similar result which separates Res(k) and Res($k+1$) using constant width clauses.

DEFINITION 9.1. *Let $X_1, \ldots, X_k$ be propositional variables. The formula Odd($X_1, \ldots, X_k$) is the k-DNF expressing that the number of satisfied variables of $X_1, \ldots, X_k$ is odd. The formula Even($X_1, \ldots, X_k$) is the k-DNF expressing that the number of satisfied variables of $X_1, \ldots, X_k$ is even.*

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

Because every clause of $GOP(G)$ contains at most $\max(d, 3)$ literals, every k-DNF in $GOP(G)[X_{u,v} \leftarrow \text{Odd}(X^1_{u,v}, \ldots X^k_{u,v}), \neg X_{u,v} \leftarrow \text{Even}(X^1_{u,v}, \ldots X^k_{u,v})]$ contains at

most dk variables. When such a DNF is expressed as a set of clauses using the distributive rule, the set of clauses has size at most $2^{O(dk)}$ and each clause has width at most dk. Therefore, $GOP^{\oplus k}(G)$ contains at most $2^{O(dk)}n^3$ clauses, each of width at most dk.

For any graph G, the polynomial-size refutations of $GOP(G)$ can be transformed into Res$(k+1)$ refutations of $GOP^{\oplus(k+1)}(G)$. On the other hand, Res(k) refutations of $GOP^{\oplus(k+1)}(G)$ require exponential size for certain graphs:

THEOREM 9.2. *Let k be given. There exist constants $d > 0$ and $\epsilon_k > 0$, and a family of graphs G on n vertices (for n sufficiently large) with maximum degree d so that Res(k) refutations of $GOP^{\oplus(k+1)}(G)$ require size at least $2^{\Omega(\epsilon_k n)}$.*

9.1. The Upper Bounds. We build Res(k) refutations for $GOP^{\oplus k}(G)$ from resolution refutations of $GOP(G)$.

DEFINITION 9.3. *Let k be a positive integer and let $X_1, \ldots, X_n$ be propositional variables. Let $X_1^1, \ldots, X_1^k, X_2^1, \ldots, X_n^k$ be new variables. Let σ be the mapping given by $\sigma(X_i) = Even(X_i^1, \ldots, X_i^k)$ and $\sigma(\neg X_i) = Odd(X_i^1, \ldots, X_i^k)$. For a clause $C = \bigvee_i l_i$, let $\sigma(C) = \bigvee_i \sigma(l_i)$.*

LEMMA 9.4. *Let k be a constant. There exists a constant c (dependent only on k) so that for all clauses $A \vee X_i$ and $B \vee \neg X_i$ be clauses in the variables $X_1, \ldots, X_n$, there is a derivation of $\sigma(A) \vee \sigma(B)$ from $\{\sigma(A) \vee \sigma(X_i), \sigma(B) \vee \sigma(\neg X_i)\}$ of size $\leq c$.*

Proof. By the completeness of Res(k), there is a Res(k) refutation of the pair of k-DNFs $\{\text{Even}(X_1, \ldots, X_k), \text{Odd}(X_1, \ldots, X_k)\}$. Let c be the minimum size of such a refutation. Because $\sigma(X_i) = \text{Odd}(X_i^1, \ldots, X_i^k)$ and $\sigma(\neg X_i) = \text{Even}(X_i^1, \ldots, X_i^k)$, there is a derivation of $\sigma(A) \vee \sigma(B)$ from $\{\sigma(A) \vee \sigma(X_i), \sigma(B) \vee \sigma(\neg X_i)\}$ of size $\leq c$. □

LEMMA 9.5. *For each k, there exists a constant c so that for every G with n vertices and degree at most $d \geq 3$, $GOP^{\oplus k}(G)$ has a Res(k) refutation of size $2^{O(dk)}n^3$.*

Proof. With the repeated application of AND-introduction inferences, $\sigma(GOP(G))$ can be derived from $GOP^{\oplus k}(G)$ in $2^{O(dk)}n^3$ many inferences. By lemma 8.3, $GOP(G)$ has a refutation of size $O(n^3)$ so by lemma 9.4, $\sigma(GOP(G))$ has a Res(k) refutation of size $O(cn^3)$. Therefore, $GOP^{\oplus k}(G)$ has a refutation of size $2^{O(dk)}n^3$. □

9.2. Random Restrictions. DEFINITION 9.6. *Let $k \geq 1$ be given. Let G be a graph. The distribution $\mathcal{P}_{k+1}(G)$ on partial assignments ρ to the variables of $GOP^{\oplus(k+1)}(G)$ is given by the following experiment:*

For each $(u, v) \in V(G)^2$, choose $i \in \{1, \ldots, k+1\}$ uniformly and independently. For each $j \in \{1, \ldots, k\}$, $j \neq i$, set $X_{u,v}^j$ to 0 or 1, uniformly and independently.

LEMMA 9.7. *Let k be given, and let F be a k-DNF in the variables of $GOP^{\oplus(k+1)}(G)$. There exist constants $\delta > 0$, dependent only on k, so that $Pr_{\rho \in \mathcal{P}_{k+1}(G)}[F \restriction_\rho \neq 1] \leq 2^{-\delta \cdot c(F)}$*

Proof. We will say that two terms T and T' are *underlying-variable-disjoint* if whenever $X_{u,v}^i \in T$ and $X_{u',v'}^{i'} \in T'$ we have that $(u, v) \neq (u', v')$. Because F is a k-DNF, it contains at least $c(F)/k(k+1)$ many underlying-variable-disjoint terms. Each of these terms is satisfied with independent probability at least $1/4^k$ (consider setting each variable of a term in turn, the probability that a variable is set to 0 or 1 is always $\geq 1/(k+1-(k-1)) = 1/2$). Therefore, $\Pr_{\rho \in \mathcal{P}_{k+1}(G)}[F \restriction_\rho \neq 1] \leq (1 - 1/4^k)^{c(F)/k(k+1)}$. □

When we apply a random restriction from $\mathcal{P}_{k+1}(G)$ to $GOP^{\oplus(k+1)}(G)$, we do not necessarily obtain an instance of $GOP(G)$. It possible that some of the edge variables

will become inverted. However, inverting some variables does not affect the width required for a resolution refutation and we may apply lemma 8.17.

COROLLARY 9.8. *If G is a connected graph that is ϵ-neighborly, then for all $k \geq 1$, for all $\rho \in \mathcal{P}_{k+1}(G)$, $GOP^{\oplus(k+1)}(G) \restriction_\rho$ requires width $\left(\frac{1-3\epsilon}{6}\right) n$.*

9.3. The Lower Bound. THEOREM 9.9. *Let k be given. There exist constants $d > 0$ and $\epsilon_k > 0$, and a family of graphs G on n vertices (for n sufficiently large) with maximum degree c so that Res(k) refutations of $GOP^{\oplus(k+1)}(G)$ require size at least $2^{\Omega(\epsilon_k n)}$.*

Proof. Let k be given. Set $p = 15 \ln 6/n$. Consider a random graph selected according to $G_{n,p}$; by lemma 8.21, G is almost certainly $\frac{1}{6}$-neighborly and by the Chernoff bounds, it has maximum degree $\leq 2pn = 26 \ln 6$. Let G be a graph that is both $\frac{1}{6}$-neighborly and has maximum degree $\leq 26 \ln 6$.

By lemma 9.7, we have that for every k-DNF F $\Pr_{\rho \in \mathcal{P}_{k+1}(G)}[F \restriction_\rho \neq 1] \leq 2^{-\delta \cdot c(F)}$. Now apply corollary 3.4 with $s = (n/12 - 1)/k$, $d = 1$. For every k-DNF F:

$$\Pr_{\rho \in \mathcal{P}_{k+1}(G)}[h(F \restriction_\rho) > (n/12 - 1)/k] \leq k 2^{-2\delta^k ((n/12-1)/k)/4^k}$$

Because k is fixed and δ depends only on k, there exists ϵ_k so that

$$\Pr_{\rho \in \mathcal{D}}[h(F \restriction_\rho) > (n/12 - 1)/k] < 2^{-\epsilon_k n}$$

Suppose for the sake of contradiction that Γ is a Res(k) refutation of $GOP^{\oplus(k+1)}(G)$ of size less than $2^{\epsilon_k n}$, By the union bound, with probability > 0, every line F of Γ has $h(F \restriction_\rho) \leq (n/12 - 1)/k$. By corollary 5.2, $GOP^{k+1}(G) \restriction_\rho$ has a resolution refutation of width at most $n/12 - 1$. On the other hand, because G is $\frac{1}{6}$-neighborly, $w_R(GOP(G)) \geq \left(\frac{1-3(1/6)}{6}\right) n = n/12$, and therefore $w_R(GOP^{\oplus(k+1)}(G) \restriction_\rho) \geq w_R(GOP(G)) \geq n/12 - 1$. Contradiction. □

10. 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 clause of the hypothesis would be 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 [8, 7, 30, 31]. 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_n^m require size exponential in n for all m? (3) Do there exist

polynomial size depth-two Frege refutations PHP_n^{2n}? (4) Let $0 < \epsilon \leq 1/2$. Do there exist sub-exponential size refutations for $PHP_n^{n+n^{1-\epsilon}}$ 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?

Nathan Segerlind
School of Mathematics
Institute for Advanced Study
Einstein Drive
Princeton, NJ 08540
nsegerli@math.ias.edu
office phone: 609-734-8246
fax: 609-951-4459
Samuel R. Buss
Department of Mathematics
University of California, San Diego
La Jolla, CA 92093
sbuss@herbrand.ucsd.edu
office phone: 858-534-6455
fax: 858-534-5273
Russell Impagliazzo
Department of Computer Science
University of California, San Diego
La Jolla, CA 92093
russell@cs.ucsd.edu
office phone: 858-534-1332
fax: 858-534-7029

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. Bonet. On the automatizability of resolution and related propositional proof systems. In *Sixteenth International Workshop on Computer Science Logic*, volume 2471 of *Lecture Notes in Computer Science*, pages 569–583. Springer, 2002.

[3] A. Atserias, M. Bonet, and J. Esteban. Lower bounds for the weak pigeonhole principle and random formulas beyond resolution. *Information and Computation*, 176:136–152, August 2002.

[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. The efficiency of resolution and Davis–Putnam procedures. *SIAM Journal on Computing*, 31(4):1048–1075, August 2002. Preliminary versions in FOCS 1996 and STOC 1998.

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

[7] E. Ben-Sasson. Hard examples for bounded depth frege. In *Proceedings of the Thirty-fourth Annual ACM Symposium on Theory of Computing*, pages 563–572, 2002.

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

[9] M. Bonet and N. Galesi. A study of proof search algorithms for resolution and polynomial calculus. In *Proceedings of the Fortieth Annual IEEE Symposium on Foundations of Computer Science*, pages 422–431, 1999.

[10] 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.

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

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

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

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

[15] S. Dantchev and S. Riis. Tree resolution proofs of the weak pigeon-hole principle. In *Proceedings of the Sixteenth Annual IEEE Conference on Computational Complexity*, pages 69–75, 2001.

[16] P. Erdös and R. Rado. Intersection theorems for finite sets. *Journal of the London Math Society*, 35:85–90, 1960.

[17] J. L. Esteban, N. Galesi, and J. Messner. On the complexity of resolution with bounded conjunctions. In *Proceedings of the Twenty-ninth International Colloquium on Automata, Languages and Programming*, pages 220–231, 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. *Theoretical Computer Science*, 93(1):159–167, February 1992.

[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 *Advances in Computing Research*, volume 5, pages 143–170. JAI Press, 1989.

[22] S. Jukna. *Extremal Combinatorics: with applications to computer science*. Springer-Verlag, 2001.

[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. *JCSS: Journal of Computer and System Sciences*, 64:843–872, 2002.

[26] R. Motwani and P. Raghavan. *Randomized Algorithms*. Cambridge University Press, 1995.

[27] E. Palmer. *Graphical Evolution: An Introduction to the Theory of Random Graphs*. Wiley Interscience, 1985.

[28] T. Pitassi and R. Raz. Regular resolution lower bounds for the weak pigeonhole principle. In *Proceedings of the Thirty-third Annual ACM Symposium on Theory of Computing*, pages 347–355, 2001.

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

[30] R. Raz. Resolution lower bounds for the weak pigeonhole principle. In *Proceedings of the Thirty-Fourth Annual ACM Symposium on Theory of Computing*, pages 553–562, 2002.

[31] A. Razborov. Improved resolution lower bounds for the weak functional pigeonhole principle. *Theoretical Computer Science*, 303(1):233–243, 2001.

[32] A. Razborov. Pseudorandom generators hard for k-DNF resolution and polynomial calculus resolution. Available at `http://genesis.mi.ras.ru/~razborov/`, 2003.

[33] N. Segerlind, S. Buss, and R. Impagliazzo. A switching lemma for small restrictions and lower bounds for k-dnf resolution. In *Forty-third Annual Symposium on Foundations of Computer Science*, pages 604–613. IEEE, 2002.

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

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

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

[37] V. Vazirani. *Approximation Algorithms*. Springer-Verlag, 2001.