Joint Advanced Student School

Explanation for 'Lower bounds for k-DNF Resolution on random 3-CNFs' talk

by Sergey Nurk (sergeynurk@gmail.com)

Saint-Petersburg, Russia
2009

Abstract

In this talk we proved exponential lower bounds for the refutation of a random 3-CNF with linear number of clauses by k-DNF Resolution for $k \leq \sqrt{\log n / \log \log n}$. This result was achieved by M. Alekhnovich. For this he introduced a specially tailored random restrictions that preserve the structure of the input random 3-CNF while mapping every k-DNF with large covering number to 1 with high probability. Next he made use of the switching lemma for small restrictions by Segerlind, Buss and Impagliazzo [1] to prove the lower bound.

The main part of this talk is based on [2].

1 Introduction

Random 3-CNFs are an interesting candidate for proving lower bounds for propositional proof systems. No short refutations of these formulas are known even for strong systems as Frege or Extended Frege. While there is little hope to prove lower bounds for such strong systems at the moment, one of more feasible goals is to prove combinatorially the hardness of refuting a random 3-CNF for those proof systems for which lower bounds are known.

Here we investigate the complexity of random 3-CNFs for Res(k) systems. k-DNF Resolution is a generalized variant of Resolution which operates with k-DNFs instead of clauses. This is an interesting intermediate system between Resolution and depth two Frege. It was shown by Segerlind, Buss and Impagliazzo [1] that random $O(k^2)$-CNFs are hard for Res(k) for $k \leq \sqrt{\log n / \log \log n}$. As one of the open questions they asked whether this bound may be improved for random 3-CNF. A positive answer to this question was given by M. Alekhnovich [2]. He proved that random 3-CNFs are hard for Res(k) for the same range of k.

2 Preliminaries

2.1 Resolution and k-DNF Resolution

Resolution is a simple propositional proof system that operates with clauses and has one rule of inference called resolution rule:

$$\frac{A \vee x \quad \neg x \vee B}{A \vee B}$$

A *resolution refutation* of a CNF formula τ is a resolution proof of the empty clause from the clauses appearing in τ. The *size* of a resolution proof is the number of different clauses in it. The *width* $\omega(C)$ of a clause C is the number of literals in C. The *width* $\omega(\tau)$ of a set of clauses τ (in particular, the width of a resolution proof) is the maximal width of the clauses appearing in this set. For an unsatisfiable set of clauses τ denote by $S_R(\tau)$ the size of minimal refutation in Resolution. Denote by $\omega_R(\tau)$ the minimal refutation width over all possible proofs of τ.

We will extend *Resolution* with *weakening* inferences: A, B-clauses. If $A \subseteq B$, then $\frac{A}{B}$.

k-DNF Resolution (or Res(k)) is a generalization of Resolution that operates with k-DNFs and has the following inference rules:

A, B are k-DNFs, $1 \leq j \leq k$ and $l, l_1, \ldots, l_j$ are literals

- Weakening: $\dfrac{A}{A \vee l}$
- Cut: $\dfrac{A \vee \bigwedge_{i=1}^{j} l_i \quad B \vee \bigvee_{i=1}^{j} \neg l_i}{A \vee B}$
- AND-introduction: $\dfrac{A \vee l_1 \ldots A \vee l_j}{A \vee \bigwedge_{i=1}^{j} l_i}$
- AND-elimination: $\dfrac{A \vee \bigwedge_{i=1}^{j} l_i}{A \vee l_i}$

Res(k) refutation of unsatisfiable CNF τ is the inference of the empty clause from the clauses in τ using the rules 1-4. Similarly to Resolution, define the size of Res(k) refutation as the number of lines it contains. $S_{R(k)(\tau)}$ is the size of the smallest Res(k) refutation of CNF τ.

Important property: Res(k) is *strongly sound.*
In our case it means that if k-DNF F is inferred from k-DNFs $F_1, \ldots, F_j$, 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 $F_1, \ldots, F_j$ are true implies a reason why F is true.

2.2 Decision Trees

Definition 1. *A* decision tree *is a rooted binary tree such that every internal node is labeled with a variable, the edges leaving this 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 $\upsilon \in \{0, 1\}$, we write the set of paths that lead from the root to a leaf labeled υ as $Br_\upsilon(T)$. We say that a decision tree T* strongly represents *a DNF F if for every $\pi \in Br_0(T)$ and for all $t \in F, t|_\pi = 0$ and for every $\pi \in Br_1(T)$ there exists $t \in F$, such that $t|_\pi = 1$. Let the* representation height *of F, $h(F)$ be 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.

Definition 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.*

2.3 Switching Lemma

The following results in were proved in [SBI02].

Theorem 1. (Switching Lemma)
Let $k \geq 1$, let $s_0, \ldots, s_{k-1}$ and $p_1, \ldots, p_k$ be sequences of positive numbers, and let 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 $\Pr_{\rho \in D}[G|_\rho \neq 1] \leq p_i$. Then for every k-DNF F:

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

Proof. 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 $F|_\rho$ by querying all of the variables in $F|_\rho$. If $c(F) > s_0$ then $\Pr(h(F|_\rho) \neq 0) \leq \Pr_{\rho \in D}[F|_\rho \neq 1] \leq p_1$.

Induction step. $k \to k+1$ Consider $(k+1)$-DNF F.

If $c(F) > s_k$ then $\Pr(h(F|_\rho) \neq 0) \leq \Pr_{\rho \in D}[F|_\rho \neq 1] \leq p_{k+1} \leq \sum_{i=1}^{k+1} 2^{\sum_{j=i}^{k} s_j} p_i$

Consider $c(F) \leq s_k$. Let S be a cover of size at most s_k. π–assignment to the variables in S. Because each term of F contains at least one variable from S, $F|_\pi$ is a k-DNF. By combining the induction hypothesis with the union bound we achieve $\Pr_{\rho \in D}\left[\exists \pi \in \{0,1\}^S : h((F|_\rho)|_\pi) > \sum_{i=0}^{k-1} s_i\right] \leq 2^{s_k}(\sum_{i=1}^{k} 2^{(\sum_{j=i}^{k-1} s_j)} p_i) < \sum_{i=1}^{k+1} 2^{(\sum_{j=i}^{k} s_j)} p_i$

In the event that $\forall \pi \in \{0,1\}^S \; h((F|_\rho)|_\pi) \leq \sum_{i=0}^{k-1} s_i$ we may easily construct a decision tree of height at most $\sum_{j=i}^{k} s_j$ strongly representing $F|_\rho$. (more formal proof can be found in [SBI02]) □

Corollary 1. (One more Switching Lemma)
k, s, d are positive integers, $\gamma, \delta \in (0,1]$. D is a distribution on partial assignments s.t. $\forall$ k-DNF G $\Pr_{\rho \in D}[G|_\rho \neq 1] \leq d2^{-\delta(c(G))^\gamma}$. For every k-DNF F:

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

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 $s_{i-1}/4 \geq (\delta/4)s_{i-1} = (\delta/4)^i s^{\gamma^{i-1}} \geq s_i$. 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 $c(G) \geq s_{i-1}$

$$\Pr_{\rho \in D}[G|_\rho \neq 1] \leq d2^{-\delta(c(G))^\gamma} \leq d2^{-\delta s_{i-1}^\gamma} = 2^{-\delta(\delta/4)^{i-1}(s^{\gamma^{i-1}})^\gamma} = d2^{-4s_i}$$

After applying previous theorem we have that for every k-DNF F
$\Pr_{\rho \in D}[h(F|_\rho) > 2s] \leq \Pr_{\rho \in D}\left[h(F|_\rho) > \sum_{i=0}^{k-1} s_i\right] \leq \sum_{i=1}^{k} 2^{(\sum_{j=i}^{k-1} s_j)} p_i \leq \sum_{i=1}^{k} 2^{2s_i}(d2^{-4s_i}) \leq$
$dk2^{-2s_k} = dk2^{-\delta' s^{\gamma'}}$ □

2.4 Decision Trees and Res(k) Refutations

The following result also can be found in [SBI02].

Theorem 2. *Let τ be a set of clauses s.t. $\omega(\tau) \leq h$. If τ has a Res(k) refutation s.t. for each line F of the refutation $h(F) \leq h$, then $\omega_R(\tau) \leq kh$.*

Proof. For each clause $C \in \tau$, T_C is a decision tree for C. For any line F of Res(k) refutation that is not a hypothesis let T_F be a decision tree of minimum height that strongly represents F. For any partial assignment π, C_π is a clause that contains negations of every literal in π. Notice that for $\pi \in Br_0(T_\emptyset), C_\pi = \emptyset$ and for each $C \in \tau$ for the unique $\pi \in Br_0(T_C), C_\pi = C$. We construct narrow resolution refutation by deriving C_π for each line F and each $\pi \in Br_0(T_F)$.

Consider F inferred from previously derived $F_1, \ldots, F_j$, $j \leq k$. We construct a decision tree T of height $\leq kh$ that represents $\bigwedge_{i=1}^{j} F_i$.

The set $\{C_\sigma | \sigma \in Br_0(T)\}$ can be derived using the weakening rule.

Because of the strong soundness of Res(k) for every $\sigma \in Br_1(T)$ one can find term $t \in F$ satisfied by σ.

Let $\pi \in Br_0(T_F)$ be given. Because T_F strongly represents F, π falsifies all terms of F. For all $\sigma \in Br(T)$ if σ consistent with π, then $\sigma \in Br_0(T)$ (otherwise, π would not falsify the term of F satisfied by σ).

For each node ν in T, let σ_ν be the path from the root to ν (viewed as partial assignment). Bottom-up, from the leaves to the root, we recursively derive (in resolution) $C_{\sigma_\nu} \vee C_\pi$, for each ν such that σ_ν is consistent with π. When we reach the root we will have derived C_π. (more formal proof can be found in [SBI02]) □

2.5 Random 3-CNFs and Linear Systems

Definition 3. *Denote by $\phi_{n,\Delta n}$ the random $3-CNF$ with Δn clauses and n variables, in which every clause is chosen independently from the set of all $2^3 C_n^3$ clauses.*

Definition 4. *For each $\phi_{n,\Delta}$ we consider a $\Delta n \times n$ matrix $A_{n,\Delta n}$ and a vector $b \in \{0,1\}^{\Delta n}$ s.t.:*

- $A_{n,\Delta n}[i,j] = 1$ *iff the i-th clause of $\phi_{n,\Delta}$ contains the variable x_j.*
- $b[i] =$ *(number of positive variables in the i-th clause) mod 2.*

Remark: Each clause of $\phi_{n,\Delta n}$ is a semantical corollary of some linear equation of the system $A_{n,\Delta n} x = b$

Instead of proving a lower bound on the size of the refutation of $\phi_{n,\Delta}$ we will prove a stronger bound on the refutation of the system $A_{n,\Delta n} = b$. Note that $A_{n,\Delta n}$ is a random matrix in which each row contain 3 ones.

3 Expanders

3.1 Basic Definitions

A boundary expander is a bounded-degree graph that has many neighbors for every subset of its nodes. We use a more general notion of expander as an $m \times n$ matrix.
$A \in \{0,1\}^{m \times n}$
We denote the i-th row of A by A_i and identify it with the set $\{j | A_{ij} = 1\}$

Definition 5. *For a set of rows $I \subseteq [m]$ of $m \times n$ matrix A, we define its* boundary, *$\partial_A I$ (or just ∂I), as a set of all $j \in [n]$ (called* boundary elements*) s.t. there exists exactly one row $i \in I$ that contains j.*

Definition 6. *We say that A is an (r, c)-*boundary expander *if*

$$\forall I \subseteq [m] \ (|I| \leq r \Rightarrow |\partial I| \geq c|I|)$$

Fact: $\forall \Delta > 0$, $c < 1$ $\exists \delta$ s.t. with probability $1 - o(1)$ $A_{n,\Delta n}$ is $(\delta n, c)$-boundary expander.

3.2 Various closures of a set of columns

We make use of two different operations of taking closures.

Definition 7. $A \in \{0,1\}^{m \times n}$ *For a set of columns $J \subseteq [n]$ define the following inference relation $\vdash_J$ on the set $[m]$ of rows of A:*

$$I \vdash_J I_1 \iff |I_1| \leq r/2 \ \wedge \ \partial(I_1) \subseteq \left[\bigcup_{i \in I} A_i \cup J \right]$$

Let the closure *$Cl(J)$ of J be a set of all rows which can be inferred via $\vdash_J$ from the empty set.*

Lemma 1. *If $|J| \leq cr/2$ then $|Cl(J)| \leq c^{-1}|J|$*

Proof. Assume the controrary and choose a chain of subsets $I_1, I_2, \ldots$ such that $I_1 \cup \cdots \cup I_{\nu-1} \vdash_J I_\nu$ and $|\cup_\nu I_\nu| > c^{-1}|J|$. Consider the smallest k s.t. $|\bigcup_{\nu=1}^k I_\nu| > c^{-1}|J|$ Since $|J| \leq cr/2$ $|\bigcup_{\nu=1}^k I_\nu| \leq r$ and since we are dealing with expander $|\partial(\bigcup_{\nu=1}^k I_\nu)| > c(c^{-1}|J|) = |J|$. On the other hand every new boundary element that results from appending via $\vdash_J$ some set of rows must belong to J, therefore $\partial(\bigcup_{\nu=1}^k I_\nu) \subseteq J$ □

We also need another closure operation the intuitive sense of which is to extract a good expander out of a given matrix by removing rows and columns.

Definition 8. $A \in \{0,1\}^{m\times n}$ *For a set of columns* $J \subseteq [n]$ *we define the following inference relation* $\vdash^e_J$ *on the set* $[m]$ *of rows of* A*:*

$$I \vdash^e_J I_1 \iff |I_1| \le r/2 \ \wedge \ \left|\partial(I_1) \setminus \left[\bigcup_{i\in I} A_i \cup J\right]\right| < (c/2)|I_1|$$

By $Cl^e(J)$ *we denote the result achieved in the following algorithm:*
Algorithm $Cl^e(J)$
$I := \emptyset$
$R := [m]$
while (there exists $I_1 \in R$ s.t. $I \vdash^e_J I_1$)
/*cleaning step*/
$I := I \cup I_1$
$R := R \setminus I_1$
end
output I;

Lemma 2. *If* $|J| < cr/4$ *then* $|Cl^e(J)| < 2c^{-1}|J|$ *no matter in what order have cleaning steps taken place.*

Proof. Assume the controrary and consider the sequence $I_1 \dots I_t$ appearing in the cleaning procedure. These sets are pairwise disjoint. $C_t := \bigcup_{k=1}^{t} I_k$. By T denote the first t : $|C_t| > 2c^{-1}|J|$ Since $|J| < cr/4 \quad |C_T| \le r$ Due to expansion $|\partial C_T| > c|C_T|$, which implies

$$|\partial C_T \setminus J| > c|C_T| - |J| \ge c|C_T|/2$$

On the other hand, every time we add some I_{t+1} to C_t during the cleaning procedure, only $(c/2)|I_{t+1}|$ new elements can be added to $\partial C_t \setminus J$ (of those elements that have never been there before). This implies

$$|\partial C_T \setminus J| \le c/2 \sum_{k=1}^{T} |I_k| = c/2|C_T|$$

□

Lemma 3. $A \in \{0,1\}^{m\times n}$, $J \subseteq [n]$
$I' = Cl^e(J)$, $J' = \bigcup_{i\in I'} A_i$.
Obtain $\hat{A}$ *by removing the rows corresponding to* I' *and columns to* J'*.* $\hat{A}$ *is either empty or* $(r/2, c/2)$*-boundary expander.*

Proof. Consider a set of rows I in $\hat{A}$ such that $|I| \le r/2$.
It is easy to understand that $\partial_A I \subseteq \partial_{\hat{A}} I \cup J \cup J'$. Therefore if $|\partial_{\hat{A}} I| < (c/2)|I|$ then $|\partial_A I \setminus (J' \cup J)| < (c/2)|I|$ and $I' \vdash^e_J I$. □

4 Random Restriction Lemma

In this section we define the special random restrictions and prove the main technical result (*Random Restriction Lemma*) that will be used in the next section to lower bound the size of Res(k) refutation.

For a term t we will wite $Cl(t)$ and $Cl^e(t)$ for $Cl(\mathrm{Vars}(t))$ and $Cl^e(\mathrm{Vars}(t))$ respectively.

Definition 9. *$A \in \{0,1\}^{m\times n}$, $b \in \{0,1\}^m$, a term t is locally consistent w.r.t. $Ax = b$ iff the formula*

$$t \wedge [A_I x = b_I]$$

is satisfiable, where $I = Cl(t)$.

Lemma 4. *If term t is locally consistent then for any set of rows I with $|I| < r/2$ the formula*

$$t \wedge [A_I x = b_I]$$

is satisfiable.

Proof. Let us regard the conjunction of literals in t as a conjunction of linear equations (having one variable each). Assume for the contradiction that linear equations corresponding to t and $A_I x | b_I$ are inconsistent. From the basic linear algebra this implies the existence of two sets $t' \subseteq t$, and $I' \subseteq I$ s.t.

$$\sum_{i\in I'}(A_i x - b_i) + \sum_{x_j^\epsilon \in t'} (x_j - \epsilon) \equiv 1$$

Then $\partial(I') \in Vars(t')$, hence $I' \in Cl(t)$. This however contradicts to the fact that t is locally consistent. □

Definition 10. *For $A \in \{0,1\}^{m\times n}$ let $G(A)$ be the corresponding bipartite graph between m row vertices and n column vertices with incidence matrix A. For two vertices v_1, v_2 each of which is either a colrmn or a row, the distance $d_A(v_1, v_2)$ is the length of the shortest path between them.*
For two subsets $V_1, V_2 \subseteq [m] \sqcup [n]$ we let

$$d_A(V_1, V_2) = \min_{v_1\in V_1, v_2 \in V_2} d_A(v_1, v_2).$$

Lemma 5. *Let A be (r, c)-boundary expander. Assume that I is a set of rows of A, $|I| < r/2$ and term t is a term such that the formula*

$$t \wedge [A_I x = b_I]$$

is satisfiable. Then for any locally consistent term t_1 with $|t_1| \le k$ s.t.

$$d_A(Cl^e(t), t_1) > 4c^{-1}k$$

the formula $t_1 \wedge t \wedge [A_I x = b_I]$ is also satisfiable.

Proof. Let us regard the conjunction of literals in t, t_1 as a conjunction of linear equations (having one variable each). Assume for the contradiction that linear equations corresponding to t, t_1 and $A_I x | b_I$ are inconsistent. As in the proof of the previous lemma this implies the existence of three sets $t' \subseteq t$, $t'_1 \subseteq t_1$ and $I' \subseteq I$ s.t.

$$\sum_{i \in I'} (A_i x - b_i) + \sum_{x_j^\epsilon \in t'} (x_j - \epsilon) + \sum_{x_l^\epsilon \in t'_1} (x_l - \epsilon) \equiv 1$$

Definition 11. *Assume that L contains linear equations of $Ax = b$ and literals. Denote by G_L the induced subgraph of $G(A)$ that contains the corresponding row vertex for each linear equation in L and the corresponding column vertex for each variable contained in L (either in literals or in equations).*

Let us consider a subcombination $L = (I, t', t'_1)$ of the sum above with the minimal number of equations that sum up to one. In this case the graph G_L is connected (otherwise one may split the equations in our sum into two smaller sets corresponding to the connected components in G_L which both sum up to a constant, this would contradict to the minimality condition).

It follows from the equation above that $\partial_A(I') \subseteq Vars(t') \cup Vars(t'_1)$. By the assumption in the statement t is consistent with $A_I x = b_I$, by the previous lemma t_1 is also consistent with $A_I x = b_I$. Thus, the equation above can hold only when t, t_1 are both non-empty. Note that $\partial(I') \backslash \mathrm{Vars}(t) \subseteq \mathrm{Vars}(t_1)$.

Case 1. $|I' \setminus Cl^e(t)| > 2c^{-1}k$ In this case

$$\left| \partial(I') \setminus \left[\bigcup_{i \in Cl^e(t)} A_i \cup Vars(t) \right] \right| \leq k \leq (c/2)|I' \setminus Cl^e(t)|$$

thus $Cl^e(t)$ is not closed since we may add $I' \setminus Cl^e(t)$.

Case 2. $|I' \setminus Cl^e(t)| \leq 2c^{-1}k$. Consider the minimal path in G_L that connects equations corresponding to t with those corresponding to t_1, this path goes along the equations in I'. Given this path one may construct a path of length $2|I' \setminus Cl^e(t)|$ that connects $Cl^e(t)$ with t_1 in $G(A)$. This however contradicts to the assumption of the lemma. □

We will need one more lemma which proof can be found in [A05].

Lemma 6. *Let $Y \subset X$ be a set of variables. Assume that b is a partial assignment on Y uniformly distributed on some affine subspace $A \subseteq \{0,1\}^Y$. Then for any term t in Y variables either* $\Pr[t|_b \equiv 1] = 0$ *or* $\Pr[t|_b \equiv 1] \geq 2^{-|t|}$.

Now it is time to introduce our special random restrictions and prove the main result of this section.

Definition 12. *Let $A \in \{0,1\}^{m\times n}$ be (r,c)-boundary expander and $b \in \{0,1\}^m$, let $X = \{x_1,\ldots,x_n\}$ be the set of all variables. Denote by $\rho_{A,b}$ a random restriction on X that results from the following experiment. Choose a random $X_1 \subset X$ of size $cr/4$. Denote by $\hat{I} = Cl^e(X_1), \quad \hat{X} = X_1 \cup \{x_j | \exists i \in \hat{I} : A_{ij} = 1\}$. The restriction $\rho_{A,b}$ assignes a random partial assignment chosen uniformly from all $\hat{x} \in \{0,1\}^{\hat{X}}$ satisfying*

$$A_{\hat{I}}\hat{x} = b_{\hat{I}}.$$

Definition 13. *A DNF ϕ is in* normal form *with respect to A, b if each of its terms is locally consistent.*

Theorem 3. (Restriction Lemma)
Assume that every column of A contains at most $\hat{\Delta}$ ones, b is arbitrary vector and $r = \Omega(n/\hat{\Delta})$. For any k-DNF ϕ in normal form holds:

$$\Pr[\phi|_\rho \neq 1] < \left(1 - 2^{-k}\right)^{c(\phi)/\hat{\Delta}^{O(k)}}$$

Proof. Consider some k-DNF ϕ in normal form and the restriction $\rho_{A,b}$. Let $X_1, \hat{X}, \hat{I}$ be the corresponding random variables from Definition 12. We may extract at least $C_0 = c(\phi)/k$ variable disjoint terms (where $c(\phi)$ is a covering number of ϕ), each of which will be covered by X_1 with probability at least $(cr/4n)^k$, thus the expected number of covered disjoint terms is $C_0(cr/4n)^k$. By Chernoff bound we may assume that there exist $C_1 = C_0/\hat{\Delta}^{O(k)}$ variable disjoint terms covered by X_1.

Let us observe the partial assignment $\hat{x} \in \{0,1\}^{\hat{X}}$ given by $\rho_{A,b}$. This is a random variable distributed uniformly on some affine subspace $A \subset \{0,1\}^{\hat{X}}$. We define the following experiment. Assume that all bits of x are hidden. In the base consider a term t_1 of ϕ. Since it is locally consistent there exists a satisfying assignment for the formula

$$t_1 \wedge [A_{\hat{I}}x = b_{\hat{I}}]$$

Let us reveal the bits of $\hat{x}$ corresponding to t_1. Denote by E_1 the event that t_1 is satisfied, by Lemma 6 $\Pr[E_1] \geq 2^{-k}$. If E_1 occurs then we terminate the process successfully, otherwise we proceed according to the following inductive step.

Assume that at step l we revealed values of terms $t_1,\ldots,t_l$ of ϕ. Let $t^{(l)}$ be the term corresponding to the partial assignment of revealed bits of $\hat{x}$ on $\mathrm{Vars}(t_1) \cup \cdots \cup \mathrm{Vars}(t_l)$, thus $\hat{x}_{Vars(t^{(l)})} = t^{(l)}$, $|t^{(l)}| \leq lk$. Consider the set of variables $Y^{(l)} \subseteq \hat{X}$ located within the distance $4c^{-1}k$ from $Cl^e(t^{(l)})$. If there is no term t_{l+1} in ϕ free of $Y^{(l)}$-variables then we terminate the process unsuccessfully. Otherwise, consider the term t_{l+1} and reveal the corresponding bits of $\hat{x}$. Let E_{l+1} be the event that t_{l+1} is satisfied. We apply Lemma 5 to show that

$$\Pr[E_{l+1} | \hat{x}_{Vars(t^{(l)})} = t^{(l)}] > 0$$

and then use Lemma 6 to conclude that $\Pr[E_{l+1} | \hat{x}_{Vars(t^{(l)})} = t^{(l)}] \geq 2^{-k}$.

Let T be the stopping time of the inductive process, i.e. the last index l for which we did the inductive step. If at least one event in the list $E_1, E_2, \ldots, E_T$ has occurred

then ϕ is killed to 1. Thus, all we need is to show that T cannot be small in case of unsuccessful termination. First consider the case when $kT \leq cr/4$. Then by Lemma 2 $Cl^e(t^{(T)}) \leq 2c^{-1}Tk$, and since variables $Y^{(T)}$ are located within the distance $4c^{-1}k$ from $Cl^e(t^{(T)})$

$$|Y^{(T)}| \leq 2c^{-1}Tk\hat{\Delta}^{4c^{-1}k}.$$

Since $Y^{(T)}$ is a hitting set for ϕ we have

$$C_1 \leq |Y^{(T)}| \leq 2c^{-1}Tk\hat{\Delta}^{4c^{-1}k}$$

and

$$T \geq C_1/\hat{\Delta}^{O(k)}$$

In the other case $T > cr/(4k)$, thus $T \geq C_2$, where $C_2 = \min(C_1/\hat{\Delta}^{O(k)}, cr/(4k))$. Because $r = \Omega(n/\hat{\Delta})$ and $c(\phi) \leq n$ we infer that $C_2 = c(\phi)/\hat{\Delta}^{O(k)}$. The theorem follows. □

In future we will use the following corollary from this theorem:

Corollary 2. *There exists an absolute constant D s.t. under the assumption of the theorem for any normal form k-DNF ϕ*

$$\Pr[\phi|_\rho \neq 1] < \ 2^{-c(\phi)/\hat{\Delta}^{Dk}}$$

5 Lower bound for Res(k)

Here we finally prove a lower bound for Res(k) based on the random restriction constructed in the previous section.

Recall the Definition 4 of the random matrix $A_{n,\Delta}$. With high probability $A_{n,\Delta}$ is $(r, 0.8)$-boundary expander for some $r = \Omega(n)$. From this an now on, let us fix r, $A_{n,\Delta}$ and assume that it is $(r, 0.8)$-boundary expander. To apply the Corollary 2 we need our matrix contain finitely many ones in each column.

Definition 14. *For matrix $A_{n,\Delta}$ let J be a set of $0.2r$ columns of the largest hamming weight. $I' = Cl^e(J)$, $J' = \bigcup_{i \in I'} A_i$. By $\hat{A}_{n,\Delta}$ we denote matrix $A_{n,\Delta}$ with all columns corresponding to J' and all rows corresponding to I' removed. Similarly define $\hat{b} = b_{[\Delta n] \setminus I'}$.*

Lemma 7. *The matrix $\hat{A}_{n,\Delta}$ is $(r/2, 0.4)$-boundary expander in which every column has weight bounded by some $\hat{\Delta}$ that depends on Δ only.*

Proof. $\hat{A}_{n,\Delta}$ is $(r/2, 0.4)$-boundary expander by Lemma 3. Since the matrix $A_{n,\Delta}$ contains exactly $3\Delta n$ ones and we removed the columns of the largest hamming weight, the matrix $\hat{A}_{n,\Delta}$ may contain at most $3\Delta n/(0.2r)$ ones in each column. Since $r = \Omega(n)$ the lemma follows. □

Lemma 8. *Every Res(k) refutation of $\phi_{n,\Delta}$ can be transformed into Res(k) refutation of the system*

$$\hat{A}_{n,\Delta}x = \hat{b}$$

in which every line is in normal form with only polynomial increase of the size.

Proof. Clearly every clause in $\phi_{n,\Delta}$ may be easily inferred from the encoding of the system $A_{n,\Delta}x = b$. Let I', J' be defined as in Definition 14. It is not hard to understand that the empty term is locally consistent. By this fact and Lemma 4 we may assign values to $x_{J'}$ so that all the equations in $A_{I'}x = b_{I'}$ are satisfied. Then we get a Res(k) refutation of the system $\hat{A}_{n,\Delta}x = \hat{b}$. It is left to show that this refutation can be transformed into normal form.

For every term t that is not locally consistent we may infer $\bar{t}$ from $2k$ axioms in polynomial size in Resolution. Thus we may substitute any occurrence of locally inconsistent terms with $\perp$ with the polynomial increase of the proof. The only subtle point are the singletons x_j^{ϵ}, however our expansion condition on $\hat{A}_{n,\Delta}$ implies that $\forall j\ Cl(\{j\}) = \emptyset$, thus all singletons are locally consistent. □

The following lemma was proved in [BW01] [3].

Lemma 9. (Ben-Sasson–Wigderson)
Assume that the matrix $\hat{A}_{n,\Delta}$ is (r, c)-boundary expander, c is a positive constant. Then every resolution refutation of the system $\hat{A}_{n,\Delta}x = \hat{b}$ requires width ϵr, where ϵ depends only on c.

Now we are ready to prove the final theorem of this talk.

Theorem 4. (Lower Bound)
For any constant Δ with probability $1 - o(1)$ every Res(k) refutation of $\phi_{n,\Delta}$ for $k < \sqrt{\log n / \log\log n}$ has size $2^{n^{1-o(1)}}$.

Proof. Consider some outcome of the random variable $\phi_{n,\Delta}$, let $\hat{A}_{n,\Delta}$ and $\hat{b}$ be defined according to Definitions 4, 14. If there exists $Res(k)$ refutation of $\phi_{n,\Delta}$ of size S then due to Lemma 8 there exists $Res(k)$ refutation P of the system $\hat{A}_{n,\Delta}x = \hat{b}$ of size $Sn^{O(1)}$ which is in normal form. Next we apply restriction $\rho_{\hat{A}_{n,\Delta},b}$ constructed in the previous section to the whole refutation P. Due to the Corollary of Restiction Lemma for each line F of P $Pr[F|_\rho \neq 1] < 2^{-c(F)/\hat{\Delta}^{Dk}}$. Applying Switching Lemma (Corollary 1) plugging in parameters $d = 1, \gamma = 1, \delta = (1/\hat{\Delta})^{Dk}, s = \epsilon r/(2k)$ we have that for every line F of $Pr\left[h(F|_\rho) > \epsilon r/k\right] \leq k2^{-\epsilon r(1/\hat{\Delta})^{2Dk^2}}$. Now applying Theorem 2 we get that with probability at least

$$1 - Sk2^{-\epsilon r/k(1/\hat{\Delta})^{2Dk^2}} > 1 - S2^{-n/2^{O(k^2)}}$$

we will be able to convert our restricted Res(k) refutation into the Resolution refutation which width is at most ϵr.

On the other hand by Lemma 9 the probability of this event must be 0. Indeed, the formula resulting after the restriction still encodes a linear system over $(r/4, 0.2)$-boundary

expander matrix, thus the restricted system has always high resolution refutation width. Altogether this implies $S > 2^{n/2^{O(k^2)}}$ and the theorem follows. □

References

[1] [SBI02] N. Segerlind, S. R. Buss, R. Impagliazzo.
A switching lemma for small restrictions and lower bounds for k-DNF Resolution.
Proc. of the 43nd IEEE Symposium on Foundations of Computer Science, 2002.

[2] [A05] M. Alechnovich, *Lower bounds for k-DNF Resolution on random 3-CNFs* ,2005

[3] [BW01] E. Ben-Sasson, A. Wigderson.
Short proofs are narrow Journal of ACM, 48(2):149-169, 2001.