A Lower Bound for k-DNF Resolution on Random CNF Formulas via Expansion

Anastasia Sofronova, Dmitry Sokolov

EPFL

Proofs and their complexity

φ is a CNF boolean formula; we want to prove its unsatisfiability efficiently.

Proofs and their complexity

φ is a CNF boolean formula; we want to prove its unsatisfiability efficiently.

Proof: can check in polytime.

Proofs and their complexity

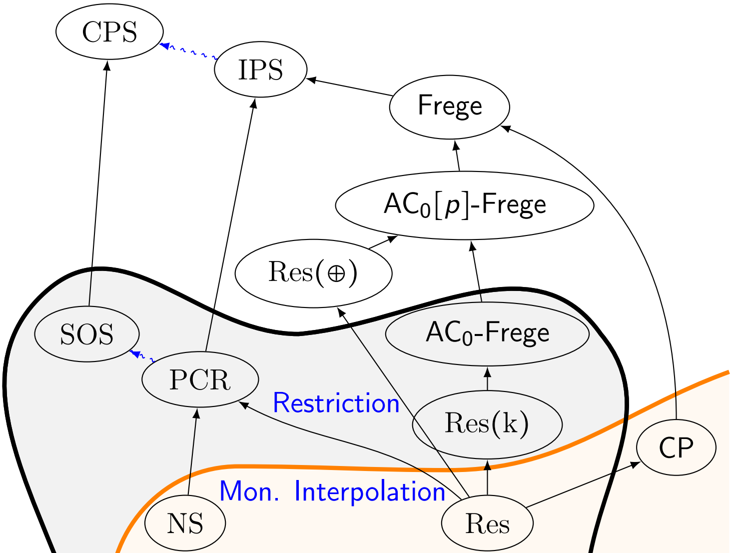
φ is a CNF boolean formula; we want to prove its unsatisfiability efficiently.

Proof: can check in polytime.

Example: Resolution

$$\textit{Weakening: } \frac{F}{F \vee \ell};$$
$$\textit{Resolution rule: } \frac{F \vee \ell_i, G \vee \neg \ell_i}{F \vee G}.$$

Where we're at



Res(k)

Weakening: $\frac{F}{F \vee \ell}$;

AND-*introduction:* $\frac{F \vee \ell_1, \ldots, F \vee \ell_w}{F \vee (\bigwedge_{i=0}^{w} \ell_i)}$;

AND-*elimination:* $\frac{F \vee (\bigwedge_{i=0}^{w} \ell_i)}{F \vee \ell_i}$;

Cut: $\frac{F \vee (\bigwedge_{i=0}^{w} \ell_i), G \vee (\bigvee_{i=0}^{w} \neg \ell_i)}{F \vee G}$.

A subsystem of AC_0-Frege.

Res(k)

Weakening: $\frac{F}{F \vee \ell}$;

AND-*introduction:* $\frac{F \vee \ell_1, \ldots, F \vee \ell_w}{F \vee (\bigwedge_{i=0}^{w} \ell_i)}$;

AND-*elimination:* $\frac{F \vee (\bigwedge_{i=0}^{w} \ell_i)}{F \vee \ell_i}$;

Cut: $\frac{F \vee (\bigwedge_{i=0}^{w} \ell_i), G \vee (\bigvee_{i=0}^{w} \neg \ell_i)}{F \vee G}$.

A subsystem of AC_0-Frege.

Known lower bounds for random CNFs are based on Switching Lemma.

Res(k)

Weakening: $\frac{F}{F \vee \ell}$;

AND-*introduction:* $\frac{F \vee \ell_1, \ldots, F \vee \ell_w}{F \vee (\bigwedge_{i=0}^{w} \ell_i)}$;

AND-*elimination:* $\frac{F \vee (\bigwedge_{i=0}^{w} \ell_i)}{F \vee \ell_i}$;

Cut: $\frac{F \vee (\bigwedge_{i=0}^{w} \ell_i), G \vee (\bigvee_{i=0}^{w} \neg \ell_i)}{F \vee G}$.

A subsystem of AC_0-Frege.

Known lower bounds for random CNFs are based on Switching Lemma.

From $k = \log^{1+\varepsilon} n$ would follow for all k.

Random Δ-CNFs

n variables, m clauses.
Density $\frac{m}{n}$ threshold for SAT/UNSAT.
Believed to be hard for any proof system.

Underlying graph is a bipartite expander.

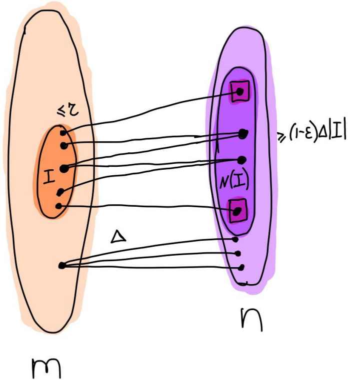
Random Δ-CNFs

n variables, m clauses.
Density $\frac{m}{n}$ threshold for SAT/UNSAT.
Believed to be hard for any proof system.

Underlying graph is a bipartite expander.

Any small subset of vertices has a lot of (unique) neighbours.

$(r, \Delta, (1-\varepsilon)\Delta)$-*(boundary)* expander: $(1-\varepsilon)\Delta|I|$ *(unique)* neighbours for $I \subseteq L, |I| \leq r$.



Random Δ-CNFs: what is known?

$\mathfrak{D} \coloneqq \frac{m}{n}$, clause density.

$\mathfrak{D} = \mathcal{O}(1)$	$\Delta \geq 3$	$k = \mathcal{O}\left(\sqrt{\frac{\log n}{\log \log n}}\right)$	[Ale11]
$\mathfrak{D} = n^{1/6}$	$\Delta = \mathcal{O}\left(k^2\right)$	$k = \mathcal{O}(1)$	[SBI04]

Random Δ-CNFs: what is known?

$\mathfrak{D} \coloneqq \frac{m}{n}$, clause density.

$\mathfrak{D} = \mathcal{O}(1)$	$\Delta \geq 3$	$k = \mathcal{O}\left(\sqrt{\frac{\log n}{\log\log n}}\right)$	[Ale11]
$\mathfrak{D} = n^{1/6}$	$\Delta = \mathcal{O}(k^2)$	$k = \mathcal{O}(1)$	[SBI04]

$\mathfrak{D} = \mathcal{O}(1)$	$\Delta = \mathcal{O}(1)$	$k = \mathcal{O}(\sqrt{\log n})$
$\mathfrak{D} = \mathrm{poly}(n)$	$\Delta = \mathcal{O}(1)$, ind of k	k = const

Main result

Theorem

φ is a Δ-CNF and its dependency graph G is an $(r, \Delta, 0.95\Delta)$-boundary expander. Then for $\delta > 0$ if:

$$n^{\delta}\left(\frac{n}{0.4r}\right)^{20k^2} = o(r/k)$$

then $\mathrm{Res(k)}$ proof of φ has size $\geq 2^{n^{\delta}}$.

Expanders from proof complexity point of view

Applying restriction:

- preserve the structure of the formula;
- decrease some parameters of the proof predictably.

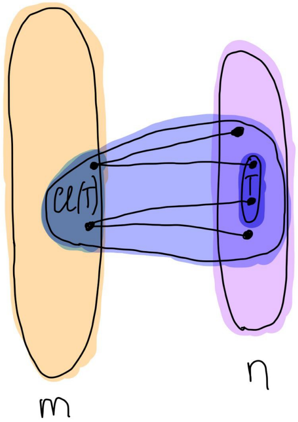
How to make restrictions to expander-based formulas?

Expanders from proof complexity point of view

Applying restriction:

- preserve the structure of the formula;
- decrease some parameters of the proof predictably.

How to make restrictions to expander-based formulas?



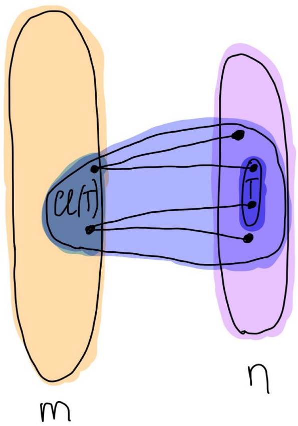
Closure: delete small part of the graph T, then delete something else to make it expander again.

Widely used to prove lower bounds in Res, PCR, SOS, etc.

Closure

Can do it differently:

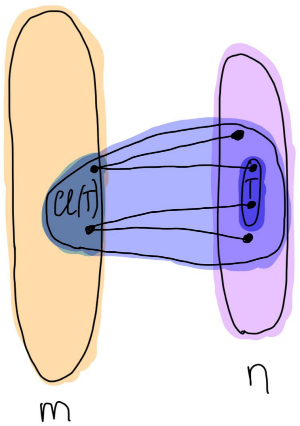
1. [AR03, Ale+04] Delete the set that violates expansion.



Closure

Can do it differently:

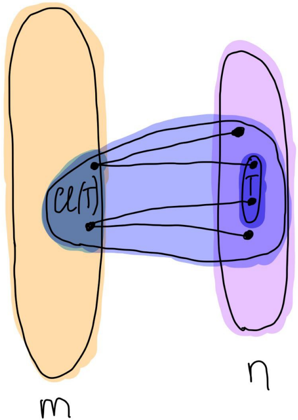
1. [AR03, Ale+04] Delete the set that violates expansion.
   Not too big, loss of parameters: $(r - \mathcal{O}(|T|), \Delta, (1 - 2\varepsilon)\Delta)$.



Closure

Can do it differently:

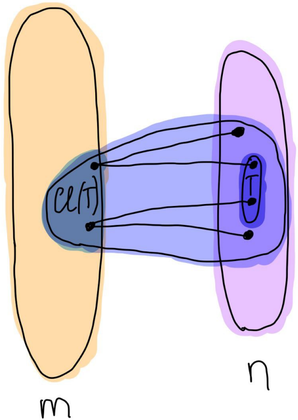
1. [AR03, Ale+04] Delete the set that violates expansion.
   Not too big, loss of parameters: $(r - \mathcal{O}(|T|), \Delta, (1 - 2\varepsilon)\Delta)$.
2. [Rez+19, Sok20] Delete the maximal set that violates expansion.



Closure

Can do it differently:

1. [AR03, Ale+04] Delete the set that violates expansion.
   Not too big, loss of parameters: $(r - \mathcal{O}(|T|), \Delta, (1-2\varepsilon)\Delta)$.
2. [Rez+19, Sok20] Delete the maximal set that violates expansion.
   Still not too big, $(r - \mathcal{O}(|T|), \Delta, (1-2\varepsilon)\Delta)$, we can repeat with the same guarantee.

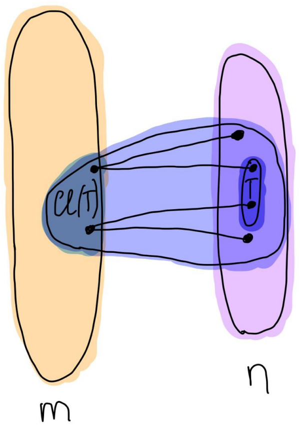


Closure

Can do it differently:

1. [AR03, Ale+04] Delete the set that violates expansion.
   Not too big, loss of parameters: $(r - \mathcal{O}(|T|), \Delta, (1-2\varepsilon)\Delta)$.
2. [Rez+19, Sok20] Delete the maximal set that violates expansion.
   Still not too big, $(r - \mathcal{O}(|T|), \Delta, (1-2\varepsilon)\Delta)$, we can repeat with the same guarantee.

[!] 3. Delete the maximal sequence of vertices s.t. each next violates expansion.

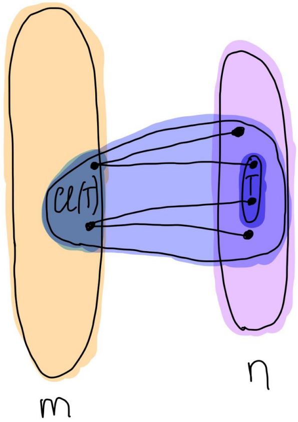


Closure

Can do it differently:

1. [AR03, Ale+04] Delete the set that violates expansion.
   Not too big, loss of parameters: $(r - \mathcal{O}(|T|), \Delta, (1-2\varepsilon)\Delta)$.
2. [Rez+19, Sok20] Delete the maximal set that violates expansion.
   Still not too big, $(r - \mathcal{O}(|T|), \Delta, (1-2\varepsilon)\Delta)$, we can repeat with the same guarantee.

[!] 3. Delete the maximal sequence of vertices s.t. each next violates expansion.
   $(r - \mathcal{O}(|T|), \Delta, (1-2\varepsilon)\Delta)$, is uniquely defined.



k-DNFs and coverings

$$(x_1 \wedge x_2 \wedge x_3) \vee (\neg x_2 \wedge x_4) \vee (x_3 \wedge \neg x_5 \wedge x_6)$$

k-DNFs and coverings

$$(x_1 \wedge x_2 \wedge x_3) \vee (\neg x_2 \wedge x_4) \vee (x_3 \wedge \neg x_5 \wedge x_6)$$

A covering: $\{x_2, x_6\}$.

k-DNFs and coverings

$$(x_1 \wedge x_2 \wedge x_3) \vee (\neg x_2 \wedge x_4) \vee (x_3 \wedge \neg x_5 \wedge x_6)$$

A covering: $\{x_2, x_6\}$.

Covering $\geq q$: $\frac{q}{k}$ non-intersecting terms.

k-DNFs and coverings

$$(x_1 \wedge x_2 \wedge x_3) \vee (\neg x_2 \wedge x_4) \vee (x_3 \wedge \neg x_5 \wedge x_6)$$

A covering: $\{x_2, x_6\}$.

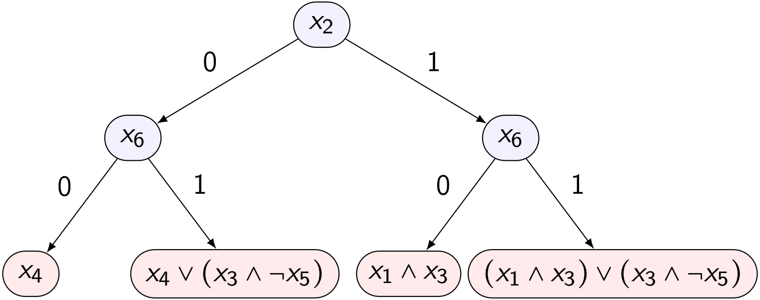
Covering $\geq q$: $\frac{q}{k}$ non-intersecting terms.
Covering $\leq q$:

k-DNFs and coverings

$$(x_1 \wedge x_2 \wedge x_3) \vee (\neg x_2 \wedge x_4) \vee (x_3 \wedge \neg x_5 \wedge x_6)$$

A covering: $\{x_2, x_6\}$.

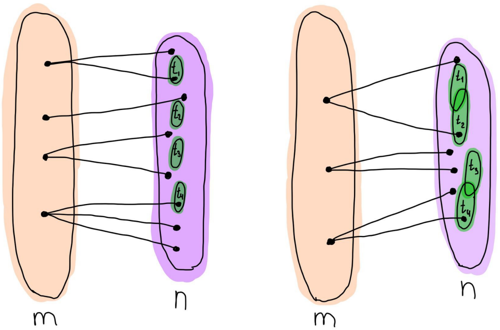
Covering $\geq q$: $\frac{q}{k}$ non-intersecting terms.
Covering $\leq q$:

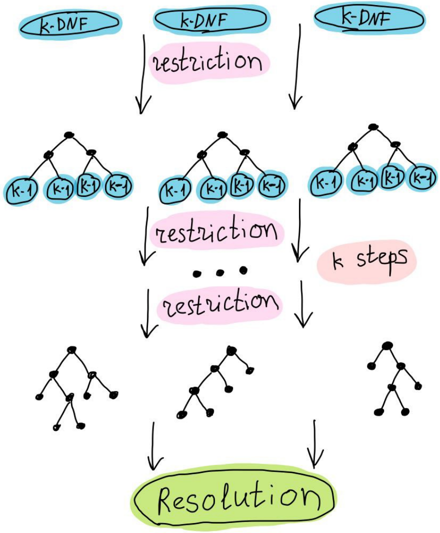


k-DNFs under random restrictions

Ideas from [Segerlind, Buss, Impagliazzo '04; Alekhnovich '11]:

- Big covering number → a lot of "independent terms";
- otherwise equivalent to a decision tree + small collection of $(k-1)$-DNFs;
- iterate that k times, what's left is a Resolution proof.




k-DNF
k-DNF
k-DNF
restriction
k-1
k-1
k-1
k-1
k-1
k-1
k-1
k-1
k-1
k-1
k-1
k-1
restriction
k steps
restriction
Resolution

Independence issue

Restrictions in expanders need to be closed.

How to argue “independence”?

Independence issue

Restrictions in expanders need to be closed.

How to argue "independence"?

In [Alekhnovich '11]: low right degree, delete small neighbourhood.

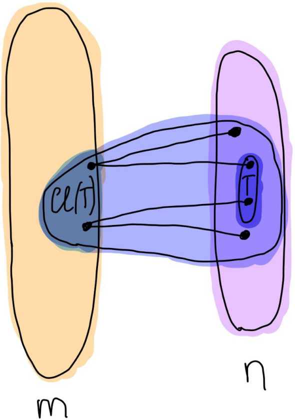
Independence issue

Restrictions in expanders need to be closed.

How to argue "independence"?

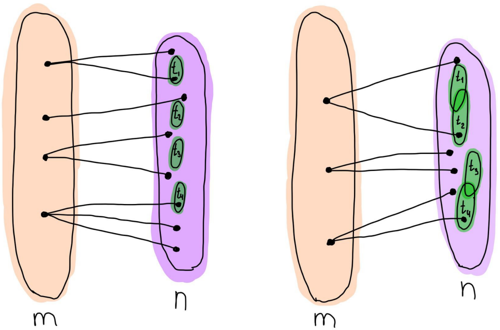
In [Alekhnovich '11]: low right degree, delete small neighbourhood.

What part of the graph actually depends on a term? Closure!



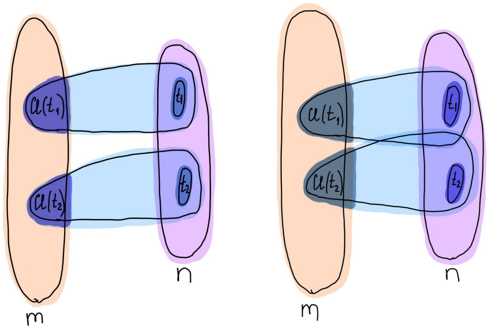
k-DNFs under random restrictions

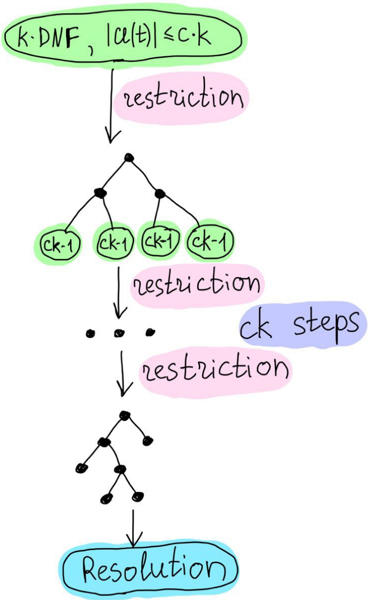
- Big covering number → a lot of "independent terms";
- otherwise equivalent to a decision tree + small collection of $(k-1)$-DNFs;
- iterate k times, what's left is a Resolution proof.



k-DNFs under random restrictions

- Big closure covering number → a lot of "closure independent terms";
- otherwise equivalent to a decision tree + small collection of DNFs where terms have smaller closure;
- iterate $\mathcal{O}(k)$ times, what's left is a Resolution proof.





Important property of individual closure: subgraph-preserving.

Open problems

- Lower bounds for larger k.
- WPHP for $k = 2$.