Tight Size-Degree Lower Bounds for Sums-of-Squares Proofs

Massimo Lauria
KTH Royal Institute of Technology (Stockholm)

1st Computational Complexity Conference, 2015 — Portland —

Joint work with

Jakob Nordström (KTH, Stockholm)
i.

the length of refutations
Proof complexity: 
Study of succinct, polynomial-time verifiable *proofs of unsatisfiability* (i.e. refutations) for CNF formulas

Original motivation: super-polynomial size lower bounds would imply $coNP \neq NP$ and hence $P \neq NP$ 

(quite a remote goal...
Recent motivation:

Study of potential and limitations of current methods for SAT solving and combinatorial optimization.

Solver outputs UNSAT $\rightarrow$ proof of unsatisfiability.
Polynomial inequalities over the reals are expressive

\[ F \geq 0 \]
Polynomial inequalities over the reals are expressive

\[ F \geq 0 \]

E.g. Propositional theorem proving

\[ \bigwedge_{j \in [m]} C_j \text{ is UNSAT} \iff (m-1) - \sum_{j \in [m]} C_j(x) \geq 0 \]
Polynomial inequalities over the reals are expressive

\[ F \geq 0 \]

E.g. Propositional theorem proving

\[ \bigwedge_{j \in \{m\}} C_j \text{ is UNSAT} \quad \text{iff} \quad (m-1) - \sum_{j \in \{m\}} C_j(x) \geq 0 \]

E.g. Optimization and approximation

\[ \min_{x \in D} F(x) \geq c \]
How do you prove an inequality

\[ F \geq 0 \]?
How do you prove a inequality

\[ F \geq 0 ? \]

(sum of squares) \[ F = A^2 + B^2 + C^2 + \ldots \]

\[ F \geq 0 ! \]

[Shor, ’87, Nesterov ’00, Parrilo ’00, GV ’01, Lasserre ’01]
Not suited for general purpose SAT solving but…

<table>
<thead>
<tr>
<th>Approximation Algorithms</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Max Cut</td>
<td>[Goemans, Williamson, 1995]</td>
</tr>
<tr>
<td>Sparsest Cut</td>
<td>[Arora, Rao, Vazirani, 2009]</td>
</tr>
<tr>
<td>Unique games</td>
<td>[Barak et al., 2012]</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Machine Learning</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sparse dictionary learning</td>
<td>[Barak, Kelner, Steurer, 2014]</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Others</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aeronautic research</td>
<td>[Chakraborty et al., 2011]</td>
</tr>
<tr>
<td>Sphere packing</td>
<td>[de Laat et al., 2014]</td>
</tr>
<tr>
<td>Power Flow optimization</td>
<td>[Ghaddar et al., 2014]</td>
</tr>
</tbody>
</table>
Refutation of $\bigwedge_{j\in[m]} C_j$

a proof of

$$-1 \geq 0$$

assuming

$$x_i^2 - x_i = 0$$

i.e. $x_i \in \{0,1\}$

$$1-x_i - x_i' = 0$$

$x_i'$ is the negation of $x_i$

$$1 - C_j(x) = 0$$

Clause $C_j$ is satisfied
A sums-of-squares refutation of $\bigwedge_{j \in [m]} C_j$ over $n$ variables

$$\sum_{j \in [m]} P_j \cdot (1 - C_j(x))$$
$$+$$
$$\sum_{i \in [n]} R_i \cdot (x_i^2 - x_i)$$
$$+$$
$$\sum_{i \in [n]} S_i \cdot (1 - x_i - x_i') - 1 = (H_1)^2 + (H_2)^2 + (H_3)^2 + \ldots$$
A sums-of-squares refutation of $\bigwedge_{j \in [m]} C_j$ over $n$ variables

$$\sum_{j \in [m]} P_j \cdot (1-C_j(x))$$

$$+$$

$$\sum_{i \in [n]} R_i \cdot (x_i^2-x_i)$$

$$+$$

$$\sum_{i \in [n]} S_i \cdot (1-x_i-x_i') - 1 = (H_1)^2 + (H_2)^2 + (H_3)^2 + \ldots$$

Size: #monomials (before cancellation)

Degree: max degree among summands
Size $\leq n^{\text{Degree}}$

Can this bound be improved?
Our work

There are 3-CNF formulas $F_{k,n}$ for $k \ll n^{\delta}$

- polynomial size in $k$ and $n$,

- degree $O(k)$ proofs, thus of size $n^{O(k)}$,

- proof size is at least $n^{\Omega(k)}$, no matter the degree.
Step 1.
symmetric formula
— degree $\Theta(k)$

Step 2.
relativization
— size $n^{\Omega(k)}$
Step 1.

**symmetric formula**
— degree $\Theta(k)$

Step 2.

**relativization**
— size $n^{\Omega(k)}$

---

[ALN ’14]

Pigeonhole Principle
from $k$ pigeons
to $k-1$ holes

- resolution
- polynomial calculus
- Sherali-Adams
Step 1.
symmetric formula
— degree $\Theta(k)$

Step 2.
relativization
— size $n^{\Omega(k)}$

[ALN ’14]
Pigeonhole Principle
from $k$ pigeons
to $k-1$ holes

This work
$k$-clique

- resolution
- polynomial calculus
- Sherali-Adams

- Sums-of-squares
Step 1.

**symmetric formula** — degree $\Theta(k)$

Step 2.

**relativization** — size $n^{\Omega(k)}$

---

**[ALN '14]**

$k$-clique

$K_{(k-1)}$

---

**This work**

- resolution
- polynomial calculus
- Sherali-Adams

- Sums-of-squares
ii.

**symmetric formula**, hard for degree
“Graph $G$ has a $k$-clique”

\[ \bigvee_{v \in V} x_{i,v} \quad \text{for } i \in [k] \]

\[ \neg x_{i,v} \lor \neg x_{j,w} \quad \text{for } i \neq j \text{ in } [k] \]

and $\{v,w\} \notin E$

The formula is symmetric w.r.t. indices in $[k]$
Objective: we want a graph $G$ so that any refutation of the $k$-clique formulas has a monomial that mention $\Omega(k)$ indices

#mention < degree
There are 3-XOR formulas with $O(k)$ clauses and variables.

1. unsatisfiable,
2. any sums-of-squares refutation requires degree $\Omega(k)$.

[Gri01, Sch08]
There are 3-XOR formulas with $O(k)$ clauses and variables.

1. unsatisfiable,
2. any sums-of-squares refutation requires degree $\Omega(k)$.

$\phi = \phi_1 \land \phi_2 \land \ldots \land \phi_k$

[Gri01, Sch08]
$\phi_i$

\[ \begin{array}{cccc}
0 & 0 & 0 & 0 \\
0 & 0 & 1 & 1 \\
0 & 1 & 0 & 1 \\
1 & 1 & 1 & 1 \\
\end{array} \]

$\mathbb{Z}_3 \mathbb{Z}_7 \mathbb{Z}_{12} \mathbb{Z}_{29}$

One vertex per assignment
One vertex per assignment

remove assignments that violate $\phi_i$
Edge iff assignments are compatible
Edge iff assignments are compatible
The graph $G$ has no $k$-Clique, since formula $\phi$ is UNSAT
The graph $G$ has no $k$-Clique, since formula $\phi$ is UNSAT.

**Lemma I:** any SOS refutation for the $k$-Clique formula over the graph $G$ has a monomial which mention $\Omega(k)$ indices.

**proof:** few #mentions implies degree proof of the 3-XOR formula. This contradicts [Gri01, Sch08].
iii.

size $n^{\Omega(k)}$ lower bound
Relativization

[Krajíček, 2004; Dantchev, Riis, 2003]: build hard formulas

[Atserias, Müller, Oliva, 2013]: lower bound for Depth-2 Frege

[Atserias, Lauria, Nordström, 2014]: $n^{\Omega(k)}$ lower bound for
• resolution of width $k$
• polynomial calculus of degree $k$
• Sherali-Adams proof of rank $k$
$F_k$ — the “semi-hard” $k$-clique formula, on variables $x_{i,v}$
$F_k$ — the “semi-hard” $k$-clique formula, on variables $x_{i,v}$

Let $F_S$ be the $\bigwedge$ of clauses that mention the indices in $S$

$$F_k = \bigwedge_{S \subseteq [k]} F_S$$
$F_k$ — the “semi-hard” $k$-clique formula, on variables $x_{i,v}$

Let $F_S$ be the $\bigwedge$ of clauses that mention the indices in $S$

$$F_k = \bigwedge_{S \subseteq [k]} F_S$$

$$F_n = \bigwedge_{S \subseteq [n], |S| \leq k} F_S$$
Let $F_S$ be the $\bigwedge$ of clauses that mention the indices in $S$

$$F_k = \bigwedge_{S \subseteq [k]} F_S$$

$$F_{k;n} = \left\{ \begin{array}{l} \sum_{i \in [n]} s_i \geq k \\
\bigwedge_{S \subseteq [n], |S| \leq k} \left( \bigvee_{i \in S} \neg s_i \right) \lor F_S \end{array} \right.$$
$F_k$ — the “semi-hard” $k$-clique formula, on variables $x_{i,v}$

Let $F_S$ be the $\bigwedge$ of clauses that mention the indices in $S$

$$F_k = \bigwedge_{S \subseteq [k]} F_S$$

$$F_{k;n} = \left\{ \begin{array}{l}
\sum_{i \in [n]} s_i \geq k \\
\bigwedge_{S \subseteq [n], |S| \leq k} \left( \bigvee_{i \in S} \neg s_i \right) \lor F_S
\end{array} \right.$$
Let $F_s$ be the $\bigwedge$ of clauses that mention the indices in $S$.
Relativized $k$-clique formula $F_{k;n}$

\[ \forall v \in V \; x_{i,v} \quad \text{for} \; i \in [k] \]

\[ \neg x_{i,v} \lor \neg x_{j,w} \quad \text{for} \; i \neq j \; \text{in} \; [k] \]

and \( \{v,w\} \notin E \)
Relativized $k$-clique formula $F_{k;n}$

$$\forall v \in V \; x_{i,v}$$

for $i \in [n]$  

for $i \neq j$ in $[n]$  

and $\{v,w\} \notin E$
Relativized $k$-clique formula $F_{k;n}$

\[ \neg s_i \lor \bigvee_{v \in V} x_{i,v} \quad \text{for } i \in [n] \]

\[ \neg s_i \lor \neg s_j \lor \neg x_{i,v} \lor \neg x_{j,w} \quad \text{for } i \neq j \text{ in } [n] \]

and $\{v,w\} \notin E$
Relativized $k$-clique formula $F_{k;n}$

\[ \sum_{i \in [n]} s_i \geq k \]

\{ i : s_i = 1 \} is the range of an injective (multi)function from $[k]$ to $[n]$

\[ \neg s_i \lor \bigvee_{v \in V} x_{i,v} \]
for $i \in [n]$

\[ \neg s_i \lor \neg s_j \lor \neg x_{i,v} \lor \neg x_{j,w} \]
for $i \neq j$ in $[n]$

and \{v,w\} $\not\in E$
The lower bound
Random restriction $\rho$ as follows

1. Select $S \subseteq [n], |S| = k$
   
   Fix $s_i$ to 1 iff $i \in S$, to 0 ow

2. If $i \notin S$ set all $x_{i,v}$ at random

3. Match $[k]$ with $S$ arbitrarily

4. we get a copy of the original $k$-Clique formula
Random restriction $\rho$ as follows

1. Select $S \subseteq [n], |S| = k$
   
   Fix $s_i$ to 1 iff $i \in S$, to 0 ow

2. If $i \notin S$ set all $x_{i,v}$ at random

3. Match $[k]$ with $S$ arbitrarily

4. we get a copy of the original $k$-Clique formula
Random restriction $\rho$ as follows

1. Select $S \subseteq [n], |S| = k$
   
   Fix $s_i$ to 1 iff $i \in S$, to 0 ow

2. If $i \not\in S$ set all $x_{i,v}$ at random

3. Match $[k]$ with $S$ arbitrarily

4. we get a copy of the original $k$-Clique formula
Random restriction $\rho$ as follows

1. Select $S \subseteq [n], |S| = k$

   Fix $s_i$ to 1 iff $i \in S$, to 0 ow

2. If $i \notin S$ set all $x_{i,v}$ at random

3. Match $[k]$ with $S$ arbitrarily

4. we get a copy of the original $k$-Clique formula
Random restriction $\rho$ as follows

1. Select $S \subseteq [n], |S| = k$

   Fix $s_i$ to 1 iff $i \in S$, to 0 ow

2. If $i \notin S$ set all $x_{i,v}$ at random

3. Match $[k]$ with $S$ arbitrarily

4. we get a copy of the original $k$-Clique formula
Lemma 2. After restriction, a monomial mentions $\Omega(k)$ indices with probability $< n^{-\Omega(k)}$

many indices before restriction:
• satisfied with high probability

few indices before restriction:
• unlikely to contain $\Omega(k)$ surviving indices
Proof conclusion

Consider a refutation $\Pi$ of size $n^{o(k)}$ for the formula $F_{k;n}$

Random restrict with $\rho$ and get a refutation $\Pi'$ for $F_k$

(by Lemma 1) $\Pi'$ must mention $\Omega(k)$ indices in some monomial
∞.

conclusion
Our work

There are 3-CNF formulas $F_{k,n}$ for $k \ll n^\delta$

- polynomial size in $k$ and $n$,
- degree $O(k)$ proofs, thus of size $n^{O(k)}$,
- proof size is at least $n^{\Omega(k)}$, no matter the degree.
Open problem — $k$-Clique

Fix $G=(V,E)$ with no $k$-clique

$$\sum_{v \in V} x_v \geq k$$

$$x_v x_w = 0 \quad \text{for } \{u,v\} \notin E$$

Does sums-of-squares require $|V|^{\Omega(k)}$ size proofs?

Worst case*: $G$

Average case: $G=G(n,p)$ for $p \approx n^{-2/(k-1)}$

* solved under some condition (unpublished)
Thank you