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

Massimo Lauria
Universitat Politécnica de Catalunya

Feb 1st, 2016 — IMS, Singapore

This talk is mostly about the paper

- ***Tight Size-Degree Lower Bounds for Sums-of-Squares Proofs***
  presented at Computational Complexity Conference, 2015

joint work with Jakob Nordström
(KTH, Stockholm)

i.

intro to proof complexity

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 → proof of unsatisfiability.

Solver running time → proof length.

Applications of SAT solving

- Model checking for hardware and software verification
- Combinatorial designs (Ramsey, Latin squares …)
- Planning
- Cryptography
- Scheduling
- Testing security protocols
- …

Proof System

A language to express proofs of UNSAT

Expressiveness: stronger algorithms ⇒ shorter proofs

⇒ hard to use

⇒ hard to analyze

Simplicity: weaker algorithms ⇒ simpler search space

⇒ better heuristics

Resolution

Proof Lines $C = l_1 \vee l_2 \vee \ldots l_w$

Resolution rule $$\frac{A \vee x \qquad B \vee \neg x}{A \vee B}$$

Refutation derive the empty clause $\bot$

Resolution is the most popular proof format

- **Simple** enough to prove size lower bounds
- Powerful enough to be **useful** in practice
- Davis-Putnam-Logemann-Loveland (DPLL) algorithm
- CDCL SAT solvers (Conflict-Driven Clause Learning)

Polynomial Calculus

Proof Lines

$$p = 0 \qquad\qquad p \in \mathbb{F}[x_1, \ldots, x_n]$$

Inference Rules

$$\frac{p = 0 \qquad q = 0}{\alpha p + \beta q = 0} \quad \alpha, \beta \in \mathbb{F} \qquad\qquad \frac{p = 0}{xp = 0}$$

Refutation

$$\frac{\vdots}{1{=}0}$$

Sherali-Adams

The proofs is a single equation, not a sequence of deductive steps

Refutation

$$p_i \geq 0 \qquad\qquad p_i \in \mathbb{R}[x_1, \ldots, x_n]$$

$$\sum_{i=1}^{t} g_i p_i = -1 \qquad \text{where} \qquad g_i = \alpha_i \prod_{\ell} x_{i_\ell} \prod_{\ell} (1 - x_{j_\ell})$$

$$\alpha_i \geq 0$$

Polynomial inequalities over the reals

Polynomial inequalities over the reals

E.g. Propositional theorem proving

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

Polynomial inequalities over the reals

E.g. Propositional theorem proving

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

E.g. Optimization and approximation

$\min_{x\in D} F(x) \geq c$

Sums-of-squares (SOS)

A method to prove polynomial inequalities

$$F \geq 0 \ ?$$

Sums-of-squares (SOS)

A method to prove polynomial inequalities

$$F \geq 0 \ ?$$

$$F = A^2 + B^2 + C^2 + \ldots$$

[Shor, ’87, Nesterov ’00, Parrilo ’00, Lasserre ’01]

Approximation Algorithms

Max Cut [Goemans, Williamson, 1995]

Sparsest Cut [Arora, Rao, Vazirani, 2009]

Machine Learning

Sparse dictionary learning [Barak, Kelner, Steurer , 2014]

Others

Aircraft research [Chakraborty et al., 2011]

Sphere packing [de Laat et al., 2014]

Power Flow optimization [Ghaddar et al., 2014]

Refutation of $\bigwedge_{j\in[m]} C_j$

a proof of

$$-1 \geq 0$$

assuming

$$x_i^2 - x_i = 0 \qquad \text{i.e. } x_i \in \{0,1\}$$

$$1 - C_j(x) = 0 \qquad \text{Clause } C_j \text{ is satisfied}$$

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

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

$$+$$

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

$$+$$

$$(-1) = (H_1)^2 + (H_2)^2 + (H_3)^2 + \ldots$$

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

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

$$+$$

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

$$+$$

$$(-1) \quad = (H_1)^2 + (H_2)^2 + (H_3)^2 + \ldots$$

Degree: max degree among summands $\leq n+1$

Size: #monomials (before cancellation) $\leq n^{\text{Degree}}$

The problem of proof search

Finding short proofs seems hard, e.g. [Alekhnovich, Razborov, 08]

Search space gets smaller if we restrict some proof **parameter**

The maximum **width** (i.e. # literals) of the clauses in a resolution proof

The maximum **degree** among the polynomials in a PC/SA/SOS proof

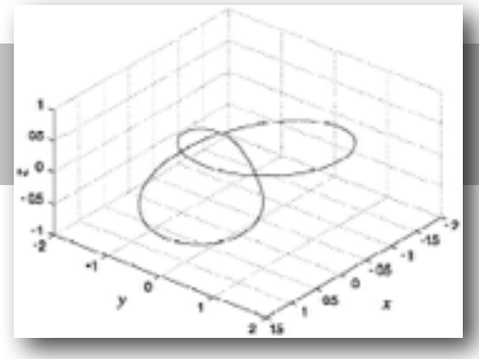
Note: This morning Albert Atserias discussed connections between these measures and length/size lower bounds. I won't discuss them further.

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

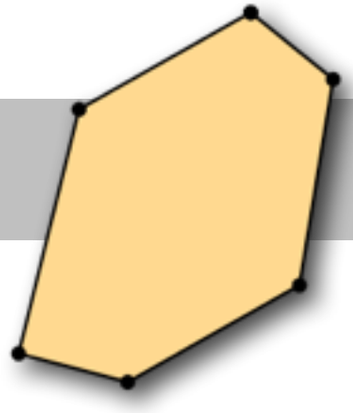
Resolution proof of width $O(k)$

Exhaustive search of length $n^{O(k)}$

Polynomial Calculus of degree $O(k)$



Gröbner basis computation of length $n^{O(k)}$



Sherali-Adams proofs of degree $O(k)$

Linear Program of size $n^{O(k)}$

Sums-of-Squares proofs of degree $O(k)$

Semidefinite Program of size $n^{O(k)}$

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

Resolution proof of width $O(k)$

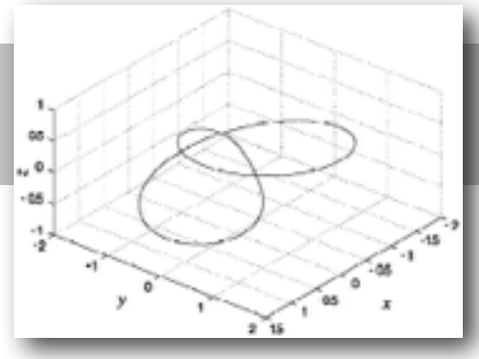
Tight by [ALN,14]

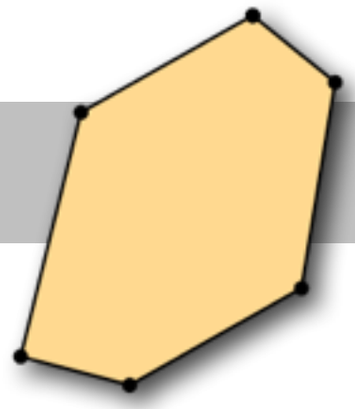
of length $n^{O(k)}$

Polynomial Calcul… $O(k)$

Tight by [ALN,14]

Gröbner basis compu… $O(k)$





Tight by [ALN,14]

…i-Adams proofs of degree $O(k)$

…rogram of size $n^{O(k)}$

Sums-of-Squares proofs of degree $O(k)$

Semidefinite Program of size $n^{O(k)}$

ii.

our result

For sums-of-squares we know that

$$\textbf{Size} \leq n^{\textbf{Degree}}$$

$$\textbf{Running-Time} \leq n^{\textbf{Degree}}$$

Can this bound be improved?

We build 3-CNF formulas $F_{k;n}$ for $k \ll n^{\delta}$

- polynomial size in k and n,

with sums-of-squares proof of

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

and which require sums-of-squares proofs of

- size $n^{\Omega(k)}$, no matter the degree

Step 1, we start with:

symmetric formula, hard for degree

Step 2, by "relativization" we build:

a formula hard for size

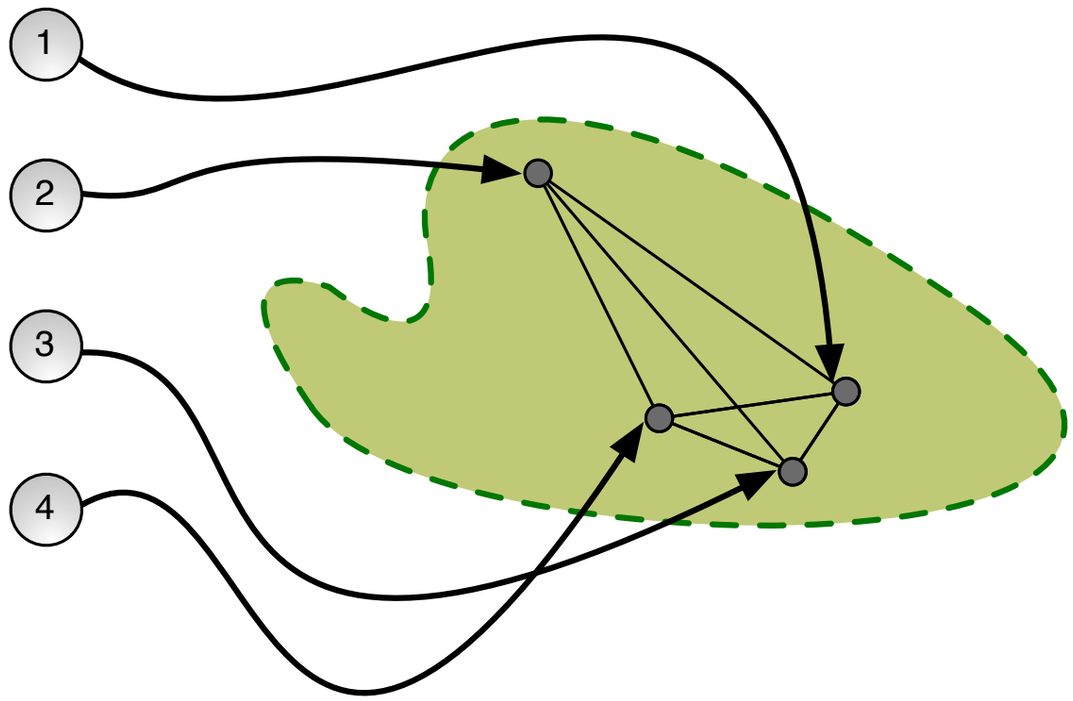
iii.

symmetric formula, hard for degree

“Graph G has a k-clique”

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

$\neg x_{i,v} \vee \neg x_{j,w}$ for $i \neq j$ in $[k]$ and $\{v,w\} \notin E$



The formula is symmetric w.r.t. indices in $[k]$

We need a k-clique formulas for a graph G such that

- G has $O(k)$ vertices
- G has no k-cliques
- any refutation has a monomial which mention $\Omega(k)$ indices

Theorem [Gri01, Sch08]: For every $k > 0$ we can sample a random $3\text{-}XOR$ formula ϕ in $O(k)$ variables and $O(k)$ constraints so that the following properties hold with positive probability.

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

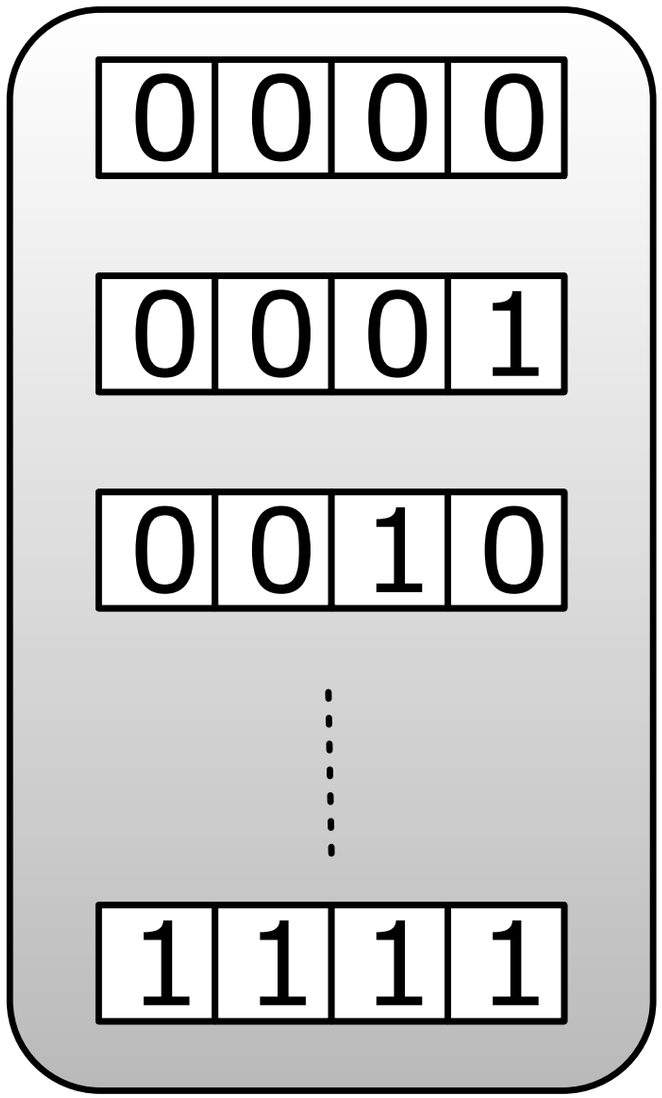
Theorem [Gri01, Sch08]: For every $k > 0$ we can sample a random $3\text{-}XOR$ formula ϕ in $O(k)$ variables and $O(k)$ constraints so that the following properties hold with positive probability.

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

Pick such ϕ and chop it into k parts of size $O(1)$ each

$$\phi = \phi_1 \ \wedge \ \phi_2 \ \wedge \cdots \wedge \ \phi_k$$

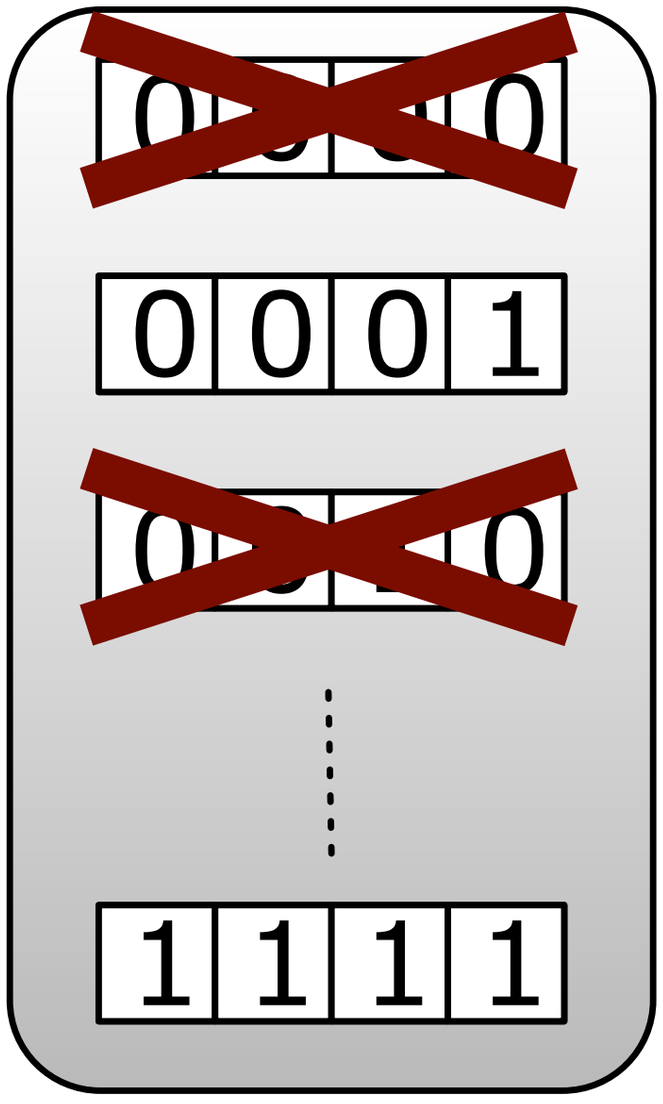
Vertices for all assignments to the variables in ϕ_i



$\leq O(1)$

$z_3\ z_7\ z_{12}\ z_{29}$

Vertices for all assignments to the variables in ϕ_i

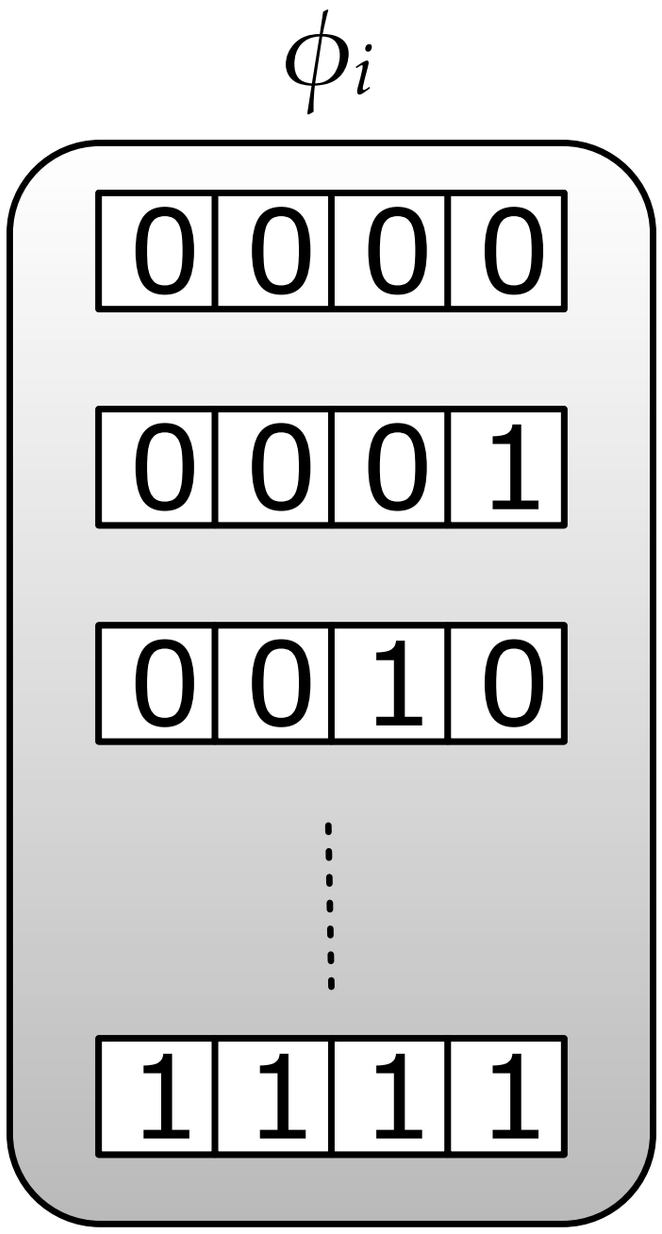


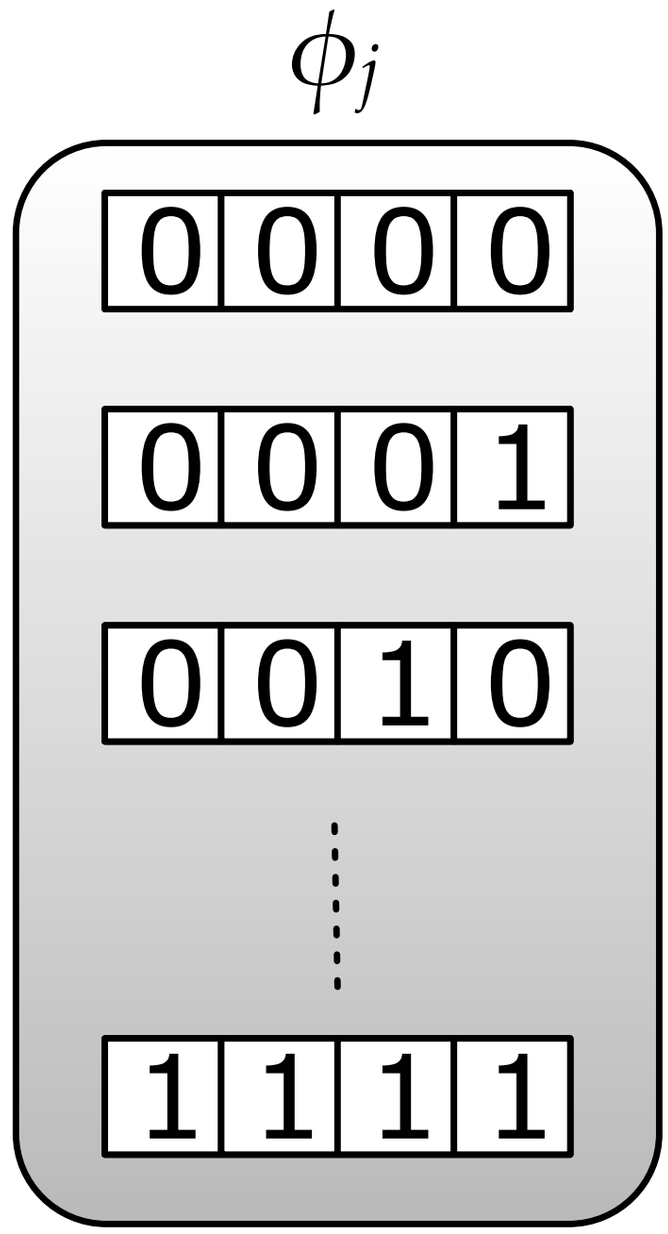
$\leq O(1)$

$z_3\ z_7\ z_{12}\ z_{29}$

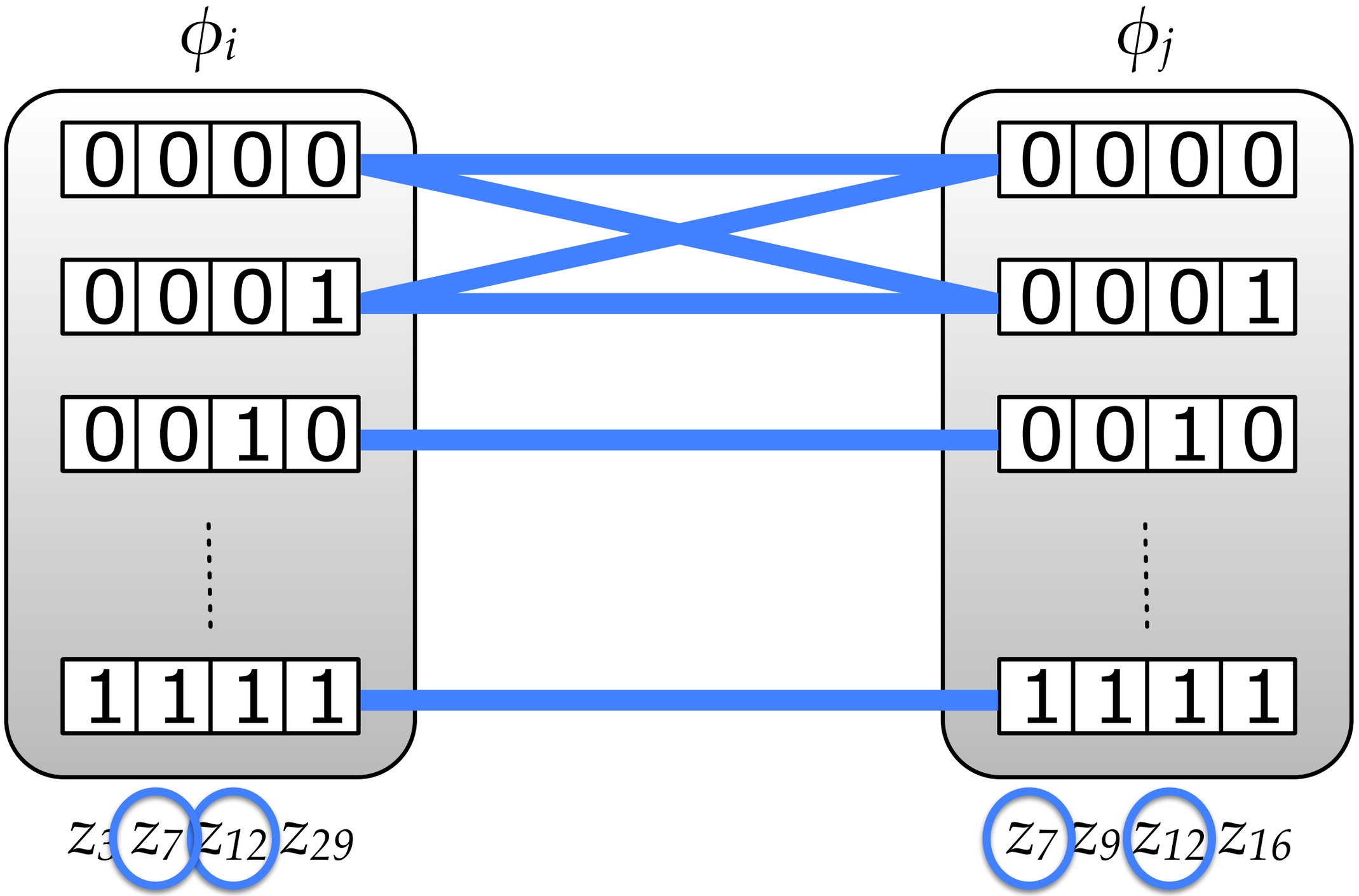
We remove the vertices not compatible with constraints in ϕ_i

E.g. $z_3 \oplus z_7 \oplus z_{29} = 1$



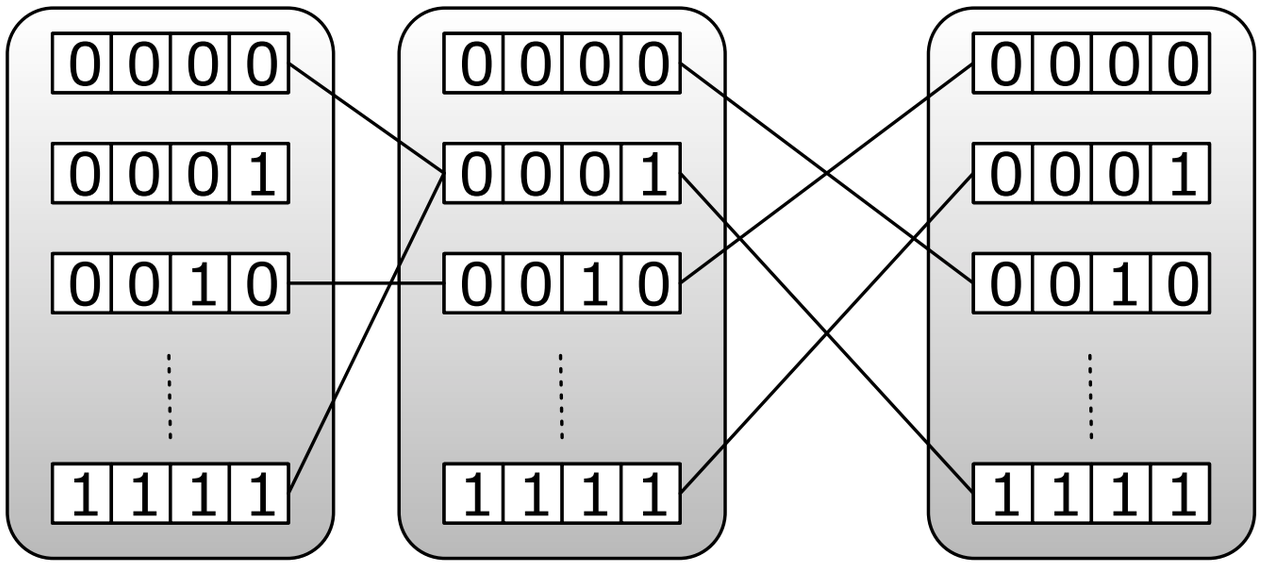


Edge iff assignments are compatible



Edge iff assignments are compatible

$G_\phi =$



- G_ϕ has $O(k)$ vertices, since each block has $O(1)$ vertices

- G_ϕ has no k-cliques, since ϕ is unsatisfiable

[FGLSS reduction]

Lemma 1: any sums-of-square refutation for the k-Clique formula over graph G_ϕ has a monomial which mentions $\Omega(k)$ indices.

proof idea: from a refutation of k-Clique over graph G_ϕ with at most t indices we extract a refutation for ϕ of degree $\Omega(t)$.

Using **Theorem [Gri01, Sch08]** we get that $t = \Omega(k)$.

iv.

relativization

Relativization [Krajíček, 2004; Dantchev, Riis, 2003]

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

[Atserias, L. , Nordström, 2014]: $n^{\Omega(k)}$ lower bound for

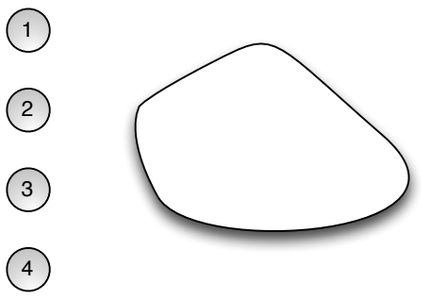
- resolution of width k
- polynomial calculus of degree k
- Sherali-Adams proof of rank k

F_k alias of the k-clique formula on G_ϕ on variables "$x_{i,v}$"

F_k alias of the k-clique formula on G_ϕ on variables "$x_{i,v}$"

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

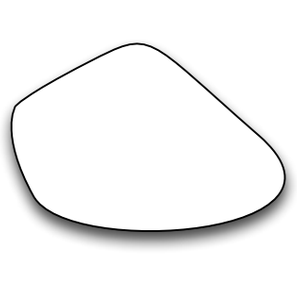
$$F_k = \bigwedge_{S \subseteq [k]} F_S$$



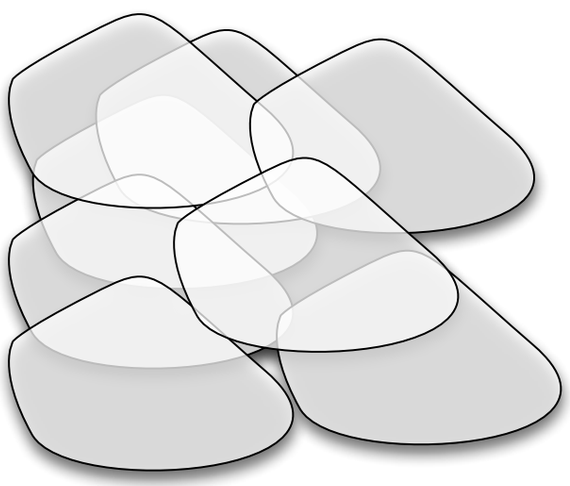
F_k alias of the k-clique formula on G_ϕ on variables "$x_{i,v}$"

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

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



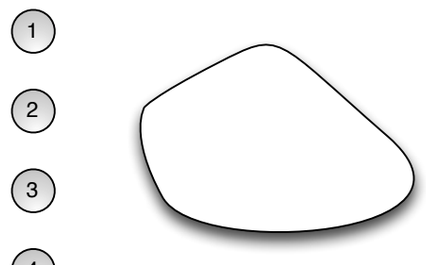
$$F_n = \bigwedge_{S \subseteq [n], |S| \leq k} F_S$$



F_k alias of the k-clique formula on G_ϕ on variables "$x_{i,v}$"

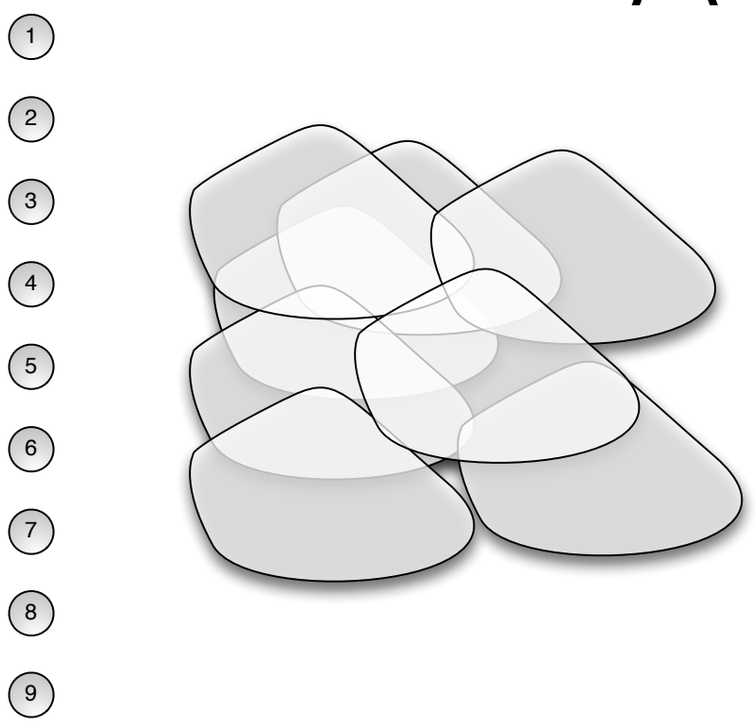
Let F_S be the $\bigwedge$ of clauses that mention the indices in S

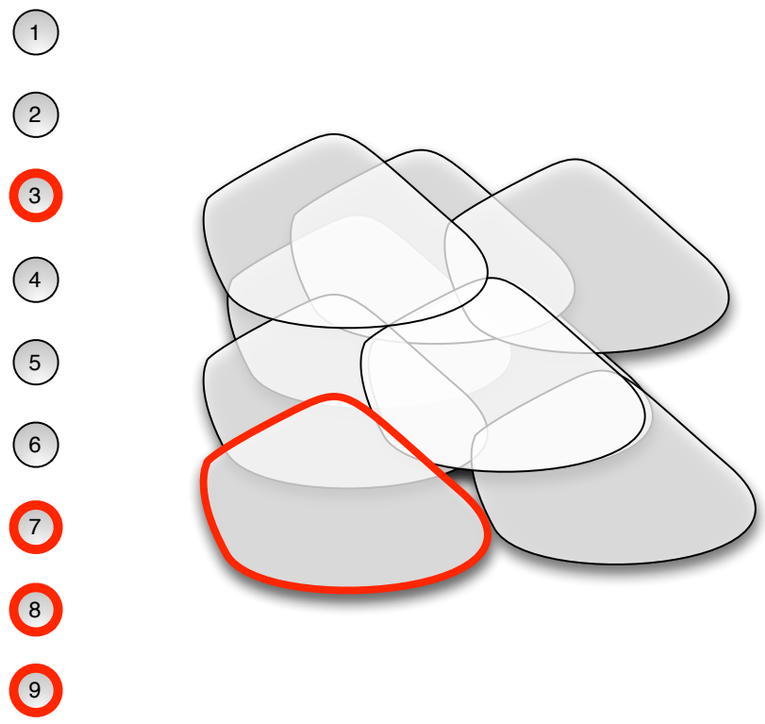
$$F_k = \bigwedge_{S \subseteq [k]} F_S$$



$$F_{k;n} = \begin{cases} \sum_{i \in [n]} s_i \;\geq\; k \\ \bigwedge_{S \subseteq [n],\, |S| \leq k} \left(\bigvee_{i \in S} \neg s_i \right) \vee F_S \end{cases}$$

$$F_n = \bigwedge_{S \subseteq [n], |S| \leq k} F_S$$

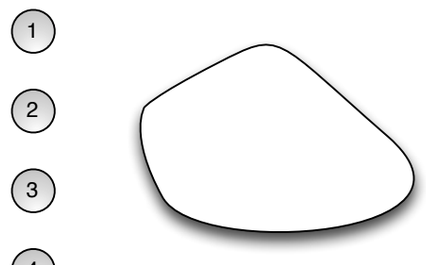




F_k alias of the k-clique formula on G_ϕ on variables "$x_{i,v}$"

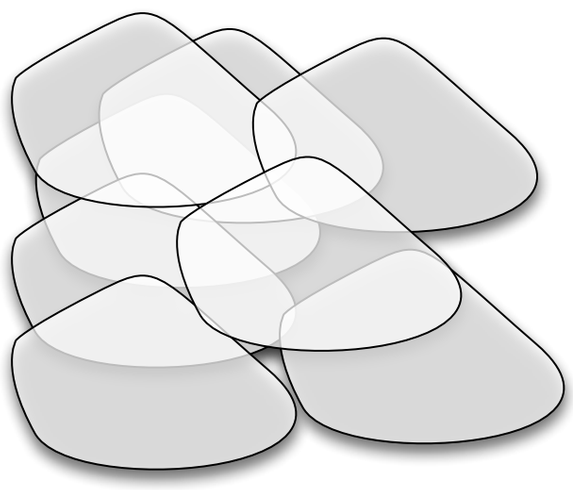
Let F_S be the $\bigwedge$ of clauses that mention the indices in S

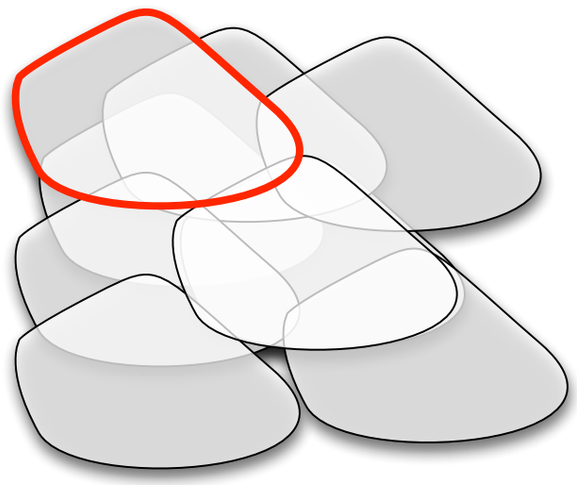
$$F_k = \bigwedge_{S \subseteq [k]} F_S$$



$$F_{k;n} = \begin{cases} \sum_{i \in [n]} s_i \;\geq\; k \\ \bigwedge_{S \subseteq [n],\, |S| \leq k} \left(\bigvee_{i \in S} \neg s_i \right) \vee F_S \end{cases}$$

$$F_n = \bigwedge_{S \subseteq [n], |S| \leq k} F_S$$



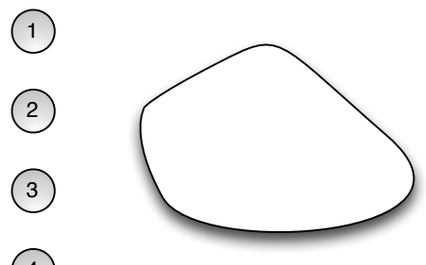


F_k alias of the k-clique formula on G_ϕ on variables "$x_{i,v}$"

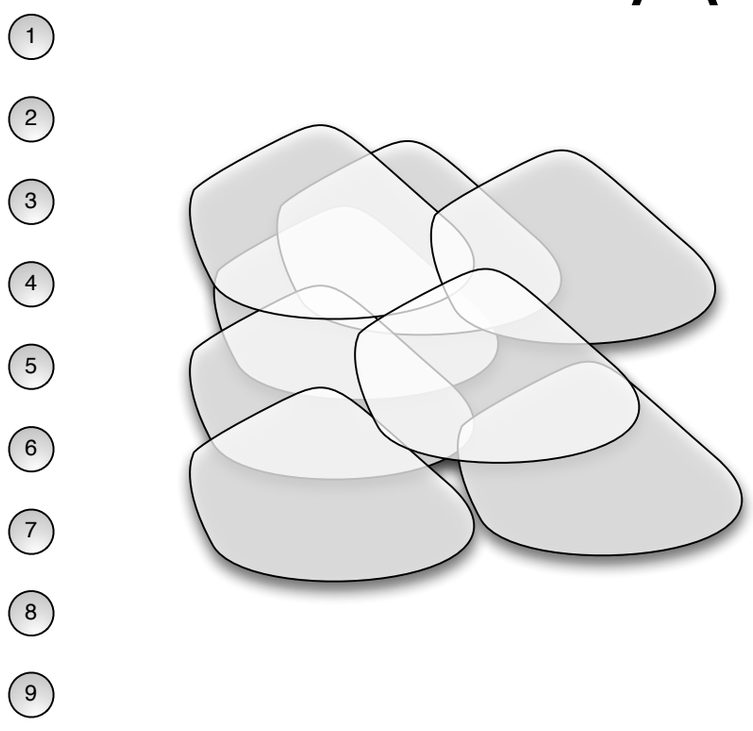
Let F_S be the $\bigwedge$ of clauses that mention the indices in S

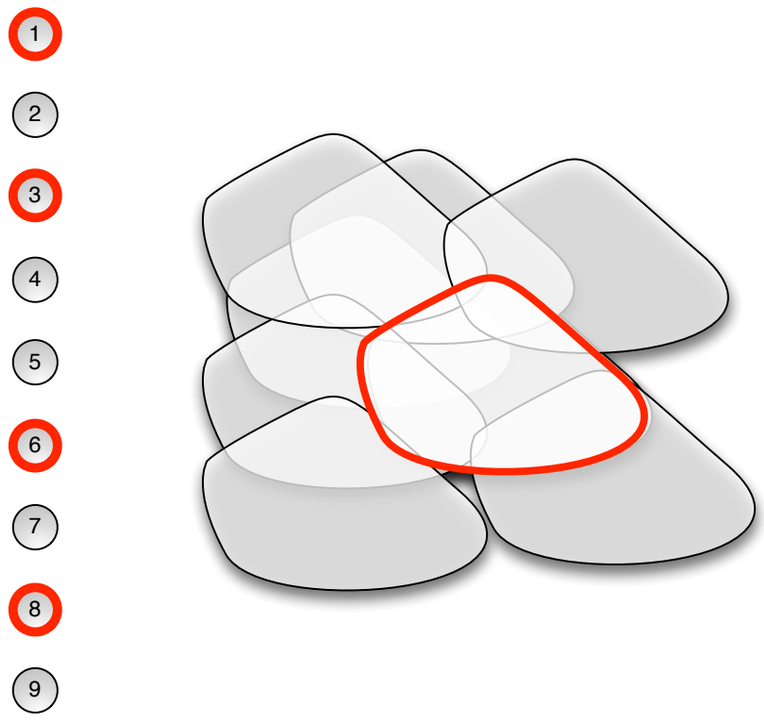
$$F_k = \bigwedge_{S \subseteq [k]} F_S$$

$$F_{k;n} = \begin{cases} \sum_{i \in [n]} s_i \;\; \geq \;\; k \\ \bigwedge_{S \subseteq [n],\, |S| \leq k} \left(\bigvee_{i \in S} \neg s_i \right) \vee F_S \end{cases}$$

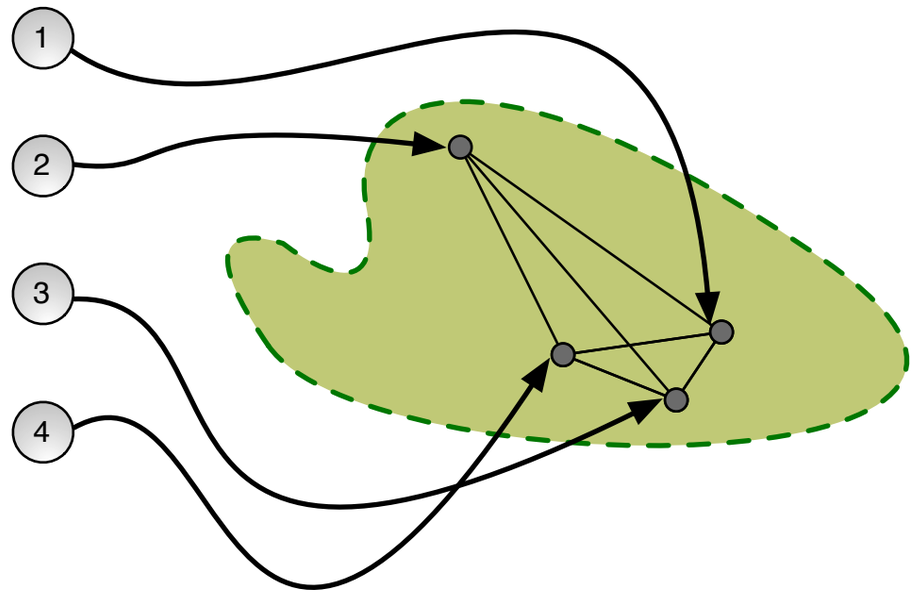


$$F_n = \bigwedge_{S \subseteq [n], |S| \leq k} F_S$$





Relativized k-clique formula $F_{k;n}$



$\bigvee_{v \in V} x_{i,v}$

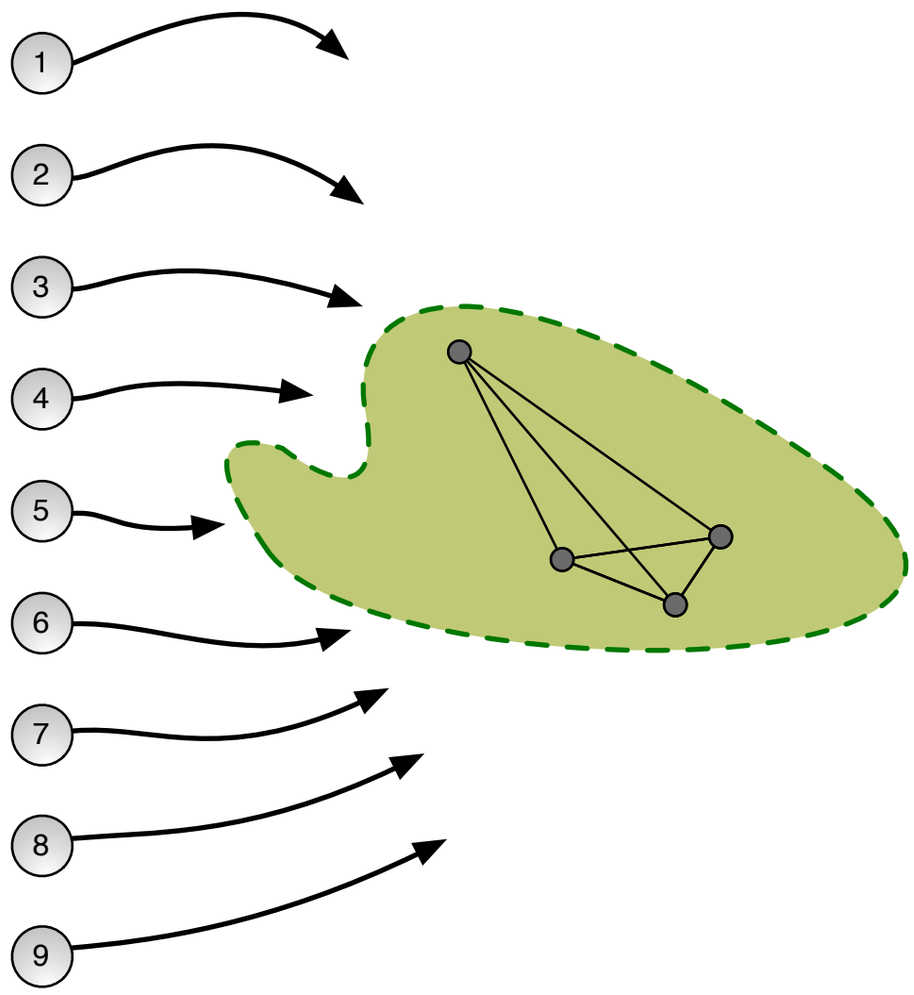
for $i \in [k]$

$\neg x_{i,v} \vee \neg x_{j,w}$

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

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

Relativized k-clique formula $F_{k;n}$



$$\bigvee_{v \in V} x_{i,v}$$

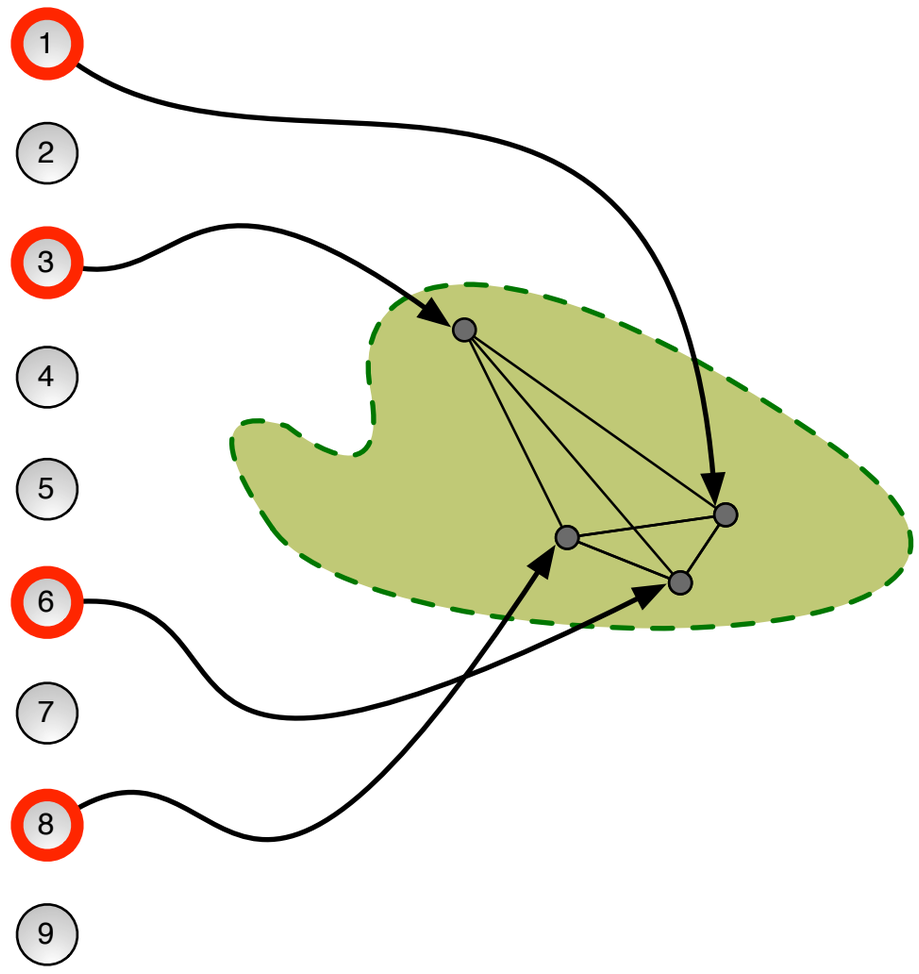
for $i \in [n]$

$$\neg x_{i,v} \vee \neg x_{j,w}$$

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

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

Relativized k-clique formula $F_{k;n}$



$\neg s_i \vee \bigvee_{v \in V} x_{i,v}$

for $i \in [n]$

$\neg s_i \vee \neg s_j \vee \neg x_{i,v} \vee \neg x_{j,w}$

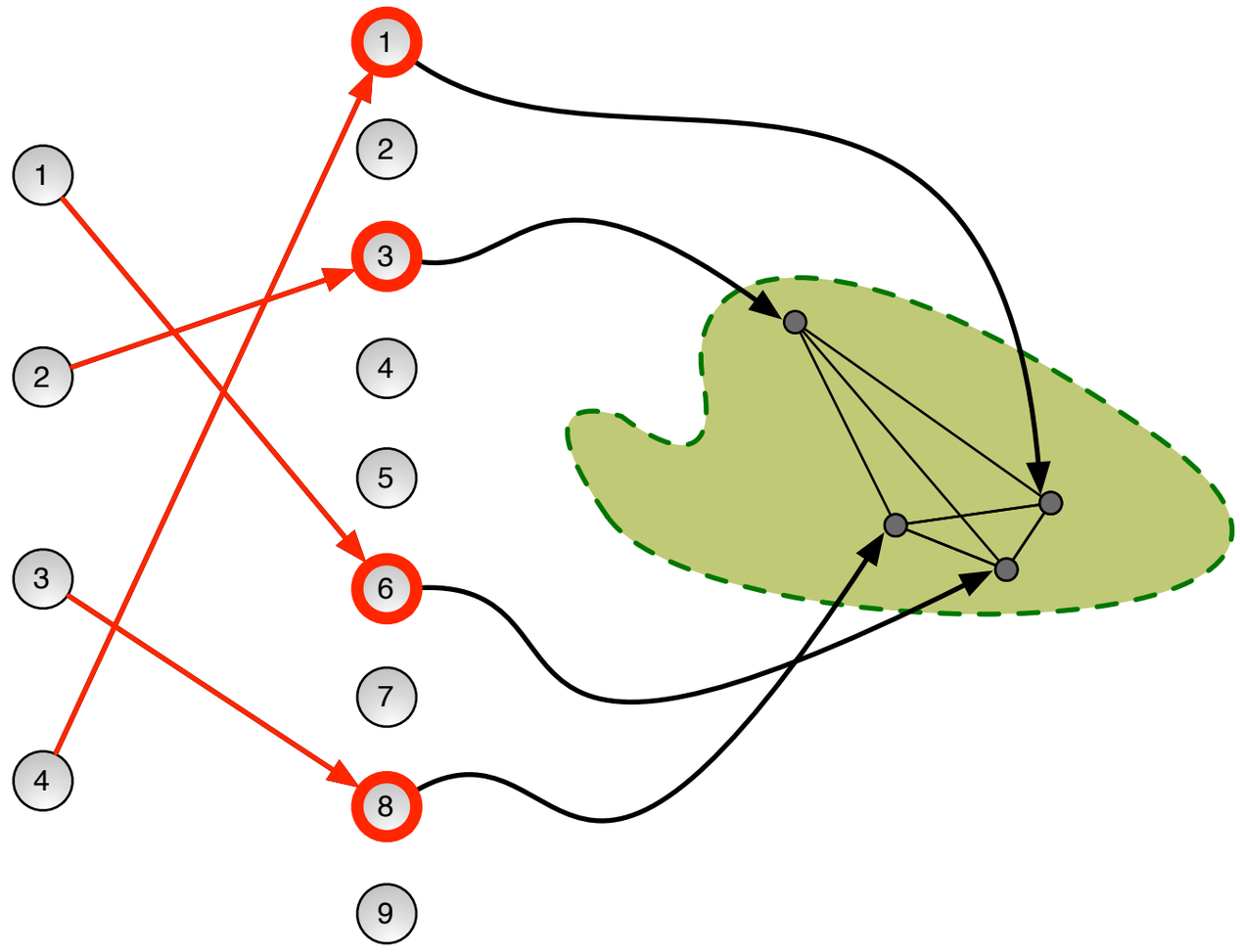
for $i \neq j$ in $[n]$

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 $[n]$



$\neg s_i \;\vee\; \bigvee_{v\in V} x_{i,v}$

for $i\in[n]$

$\neg s_i \;\vee\; \neg s_j \vee\; \neg x_{i,v} \vee \neg x_{j,w}$

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

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

The formula $F_{k;n}$ has a refutation of size $n^{O(k)}$

proof idea: brute force over all possible ways to

- choose k indices from $[n]$

- point them to k vertices in the graph G_ϕ

V.

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

Key tool: random restriction

A partial assignment ρ

$$F_{k;n} \xrightarrow{\rho} F_{k;n}\,|_{\rho}$$ simplifies formula

$$\Pi \xrightarrow{\rho} \Pi\,|_{\rho}$$ proof of simplied formula

Idea:
Simplified formula requires proof with large monomial
If proof is small, restriction removes all large monomials

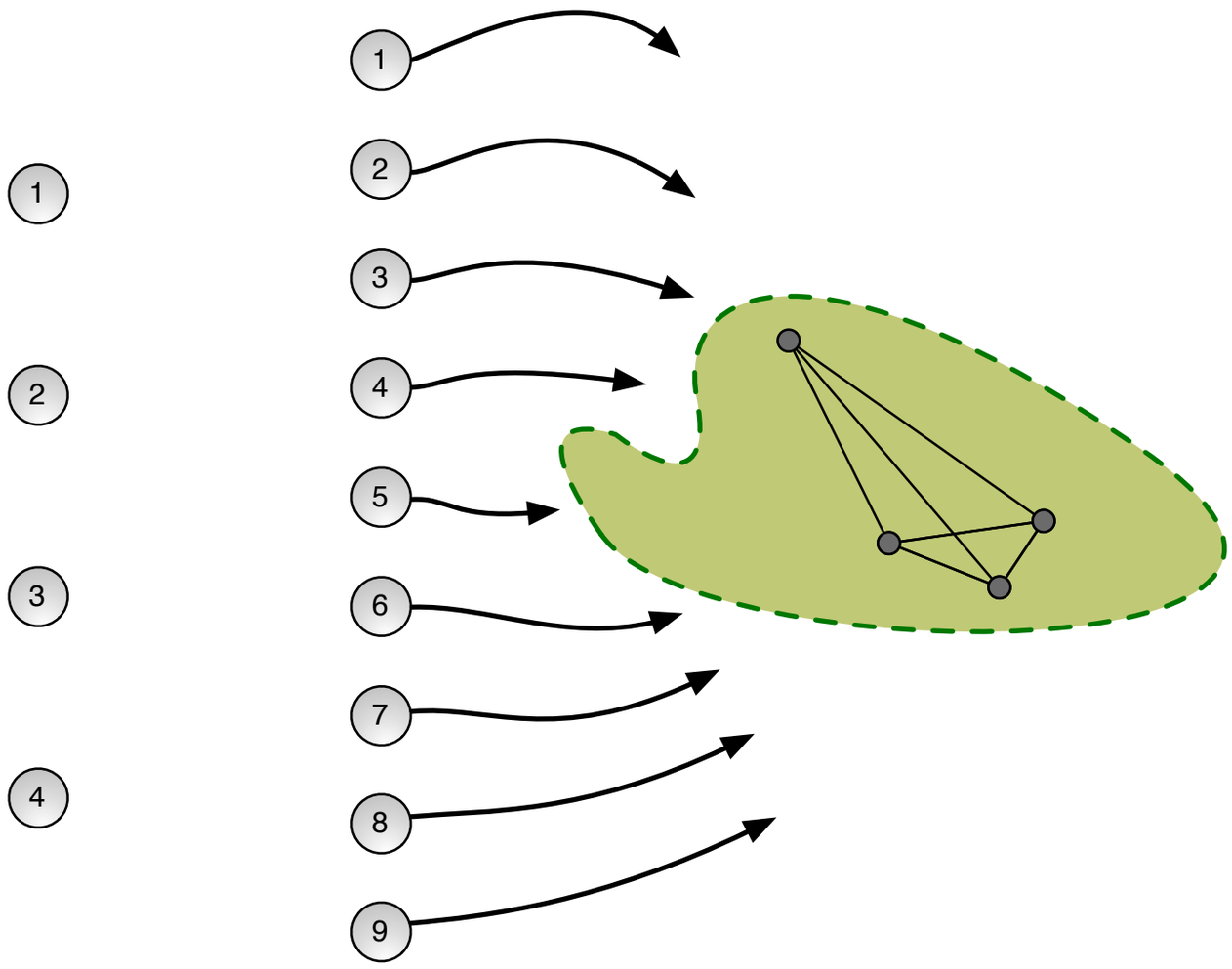
Usually, restriction arguments give exponential lower bounds, which **cannot work here**...

... we need to fine tune the restriction to make it work in the right range of parameters.

For the experts:

Furst-Saxe-Sipser style instead of Håstad style
(see also [Atserias, Müller, Oliva, 2013])

Random restriction ρ as follows



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

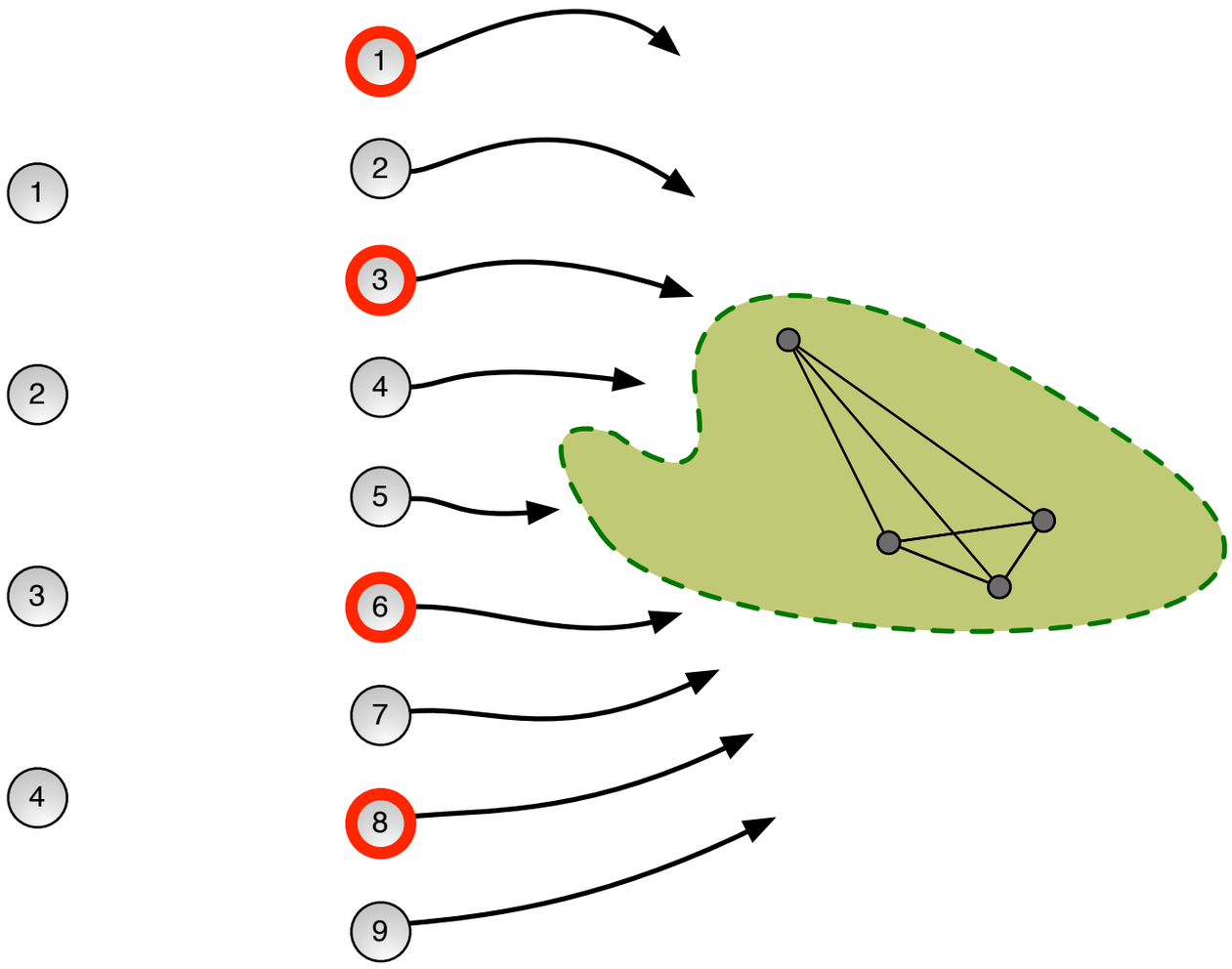
   Fix s_i to 1 iff $i \in S$, to 0 ow

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 ρ as follows



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

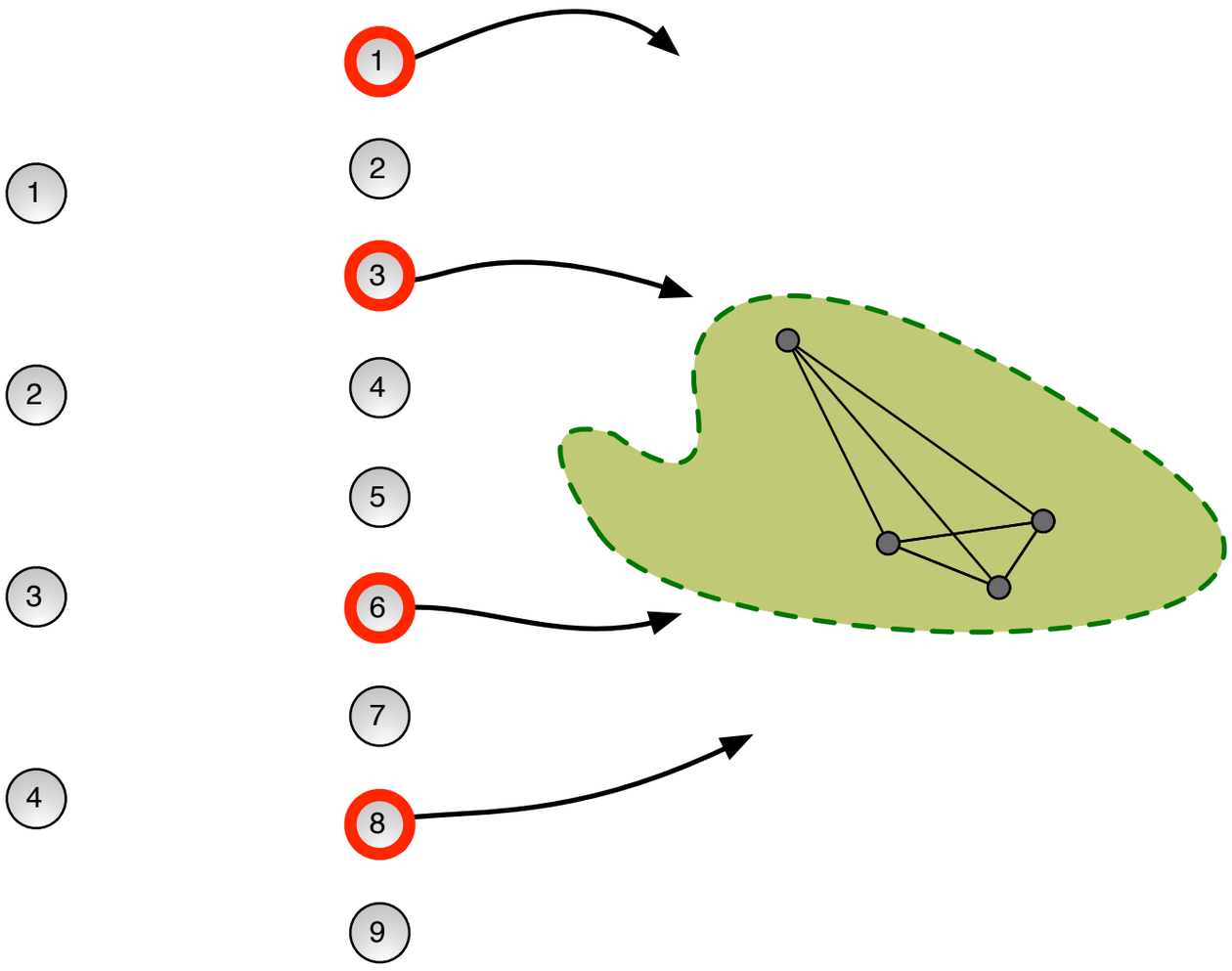
   Fix s_i to 1 iff $i \in S$, to 0 ow

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 ρ as follows



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

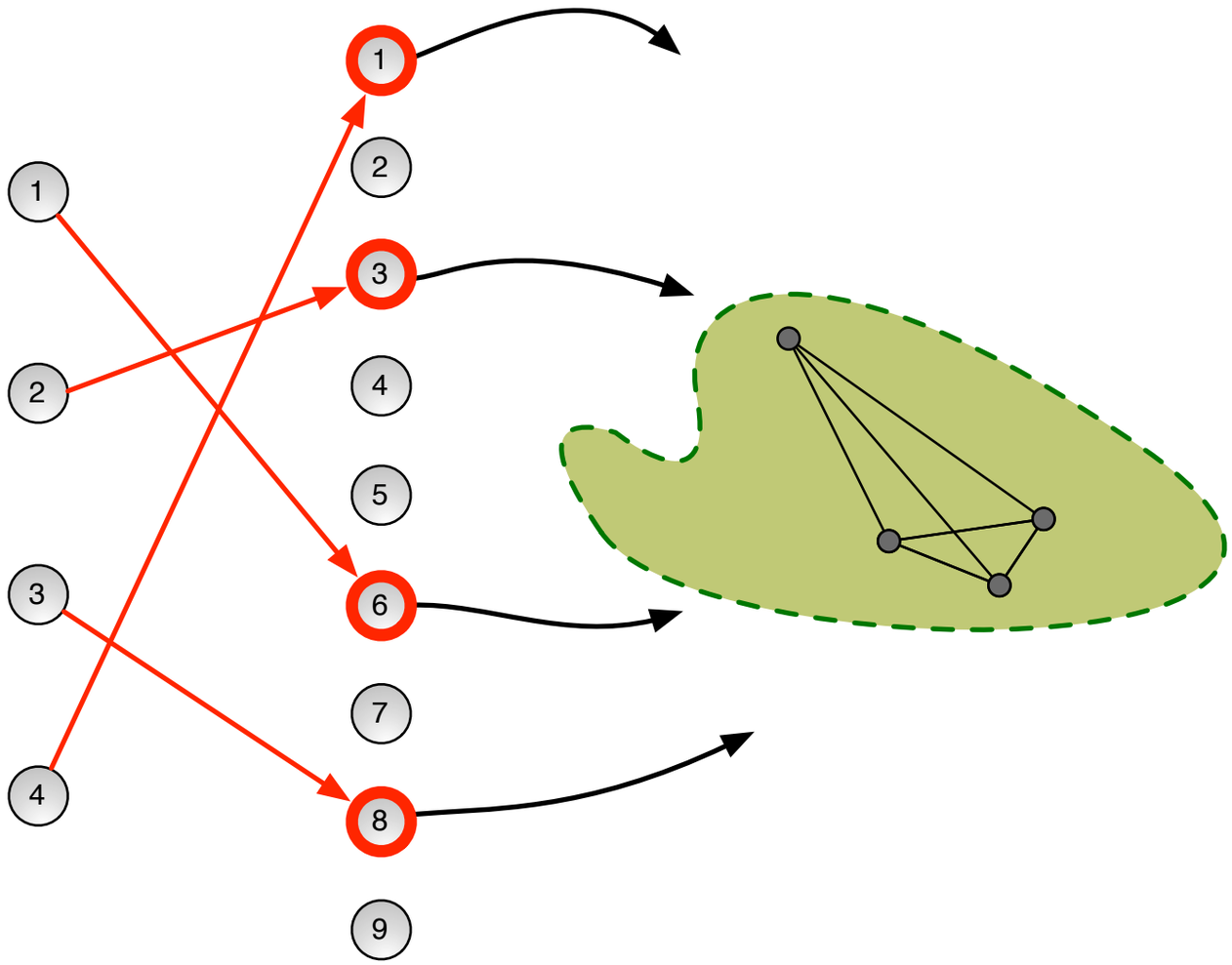
   Fix s_i to 1 iff $i \in S$, to 0 ow

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 ρ as follows



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

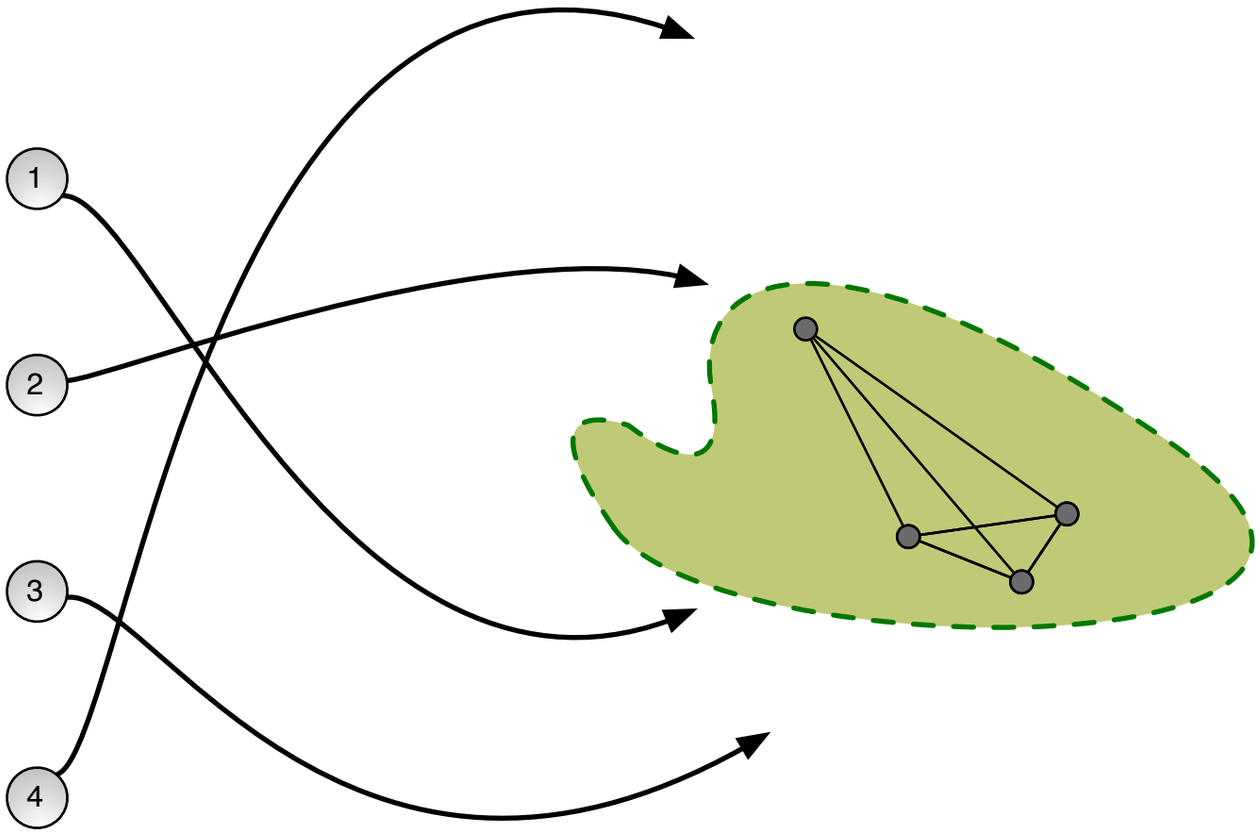
   Fix s_i to 1 iff $i \in S$, to 0 ow

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 ρ 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

Recall

Since $F_{k;n} \mid_{\rho}$ is a copy of k-Clique formula for G_ϕ

Lemma I: Any sums-of-squares refutation of it has a monomial that mentions $\Omega(k)$ indices.

Lemma 2. After restriction, a monomial mentions $\Omega(k)$ indices with probability $< n^{-\Omega(k)}$

Hence, if proof size $< n^{\Omega(k)}$ there is restriction yielding a proof with no monomial with $\Omega(k)$ indices.

$$F_{k;n} \xrightarrow{\rho} F_{k;n}\restriction_\rho$$

$$\Pi \xrightarrow{\rho} \Pi\restriction_\rho$$

m a monomial in the unrestricted refutation Π

r # of indices mentioned in m

$$r \geq k \log n + k$$

$$\Pr[m \text{ is not set to zero}] \leq \left(\frac{1}{2}\right)^{k \log n} \leq n^{-k}$$

$$r < k \log n + k$$

$$\Pr[m \text{ mentions } \alpha k \text{ surviving indices}] \lessapprox \binom{k}{\alpha k} \binom{k \log n + k}{\alpha k} n^{-\Omega(k)}$$

Proof recap

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

We restrict and we get a refutation $\Pi \mid_\rho$ *for* $F_{k;n} \mid_\rho$

by **Lemma 1,** $F_{k;n} \mid_\rho$ is k-Clique formula on G_ϕ, for all ρ

$\Pi \mid_\rho$ must mention $\Omega(k)$ indices in some monomial

by **Lemma 2**, for some choice of the restriction ρ

$\Pi \mid_\rho$ does mention $o(k)$ indices in every monomial

q.e.d.

"Didn't you promise a 3-CNF?"

$F_{k;n}$ can be turned into a 3-CNF using extension variables

- the lower bound applies with a fix to the argument
- the upper bound now has size $n^{O(k)}$ and degree $O(k)$

∞.

conclusion

Our result

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

- polynomial size in k and n,

with sums-of-squares proof of

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

and which require sums-of-squares proofs of

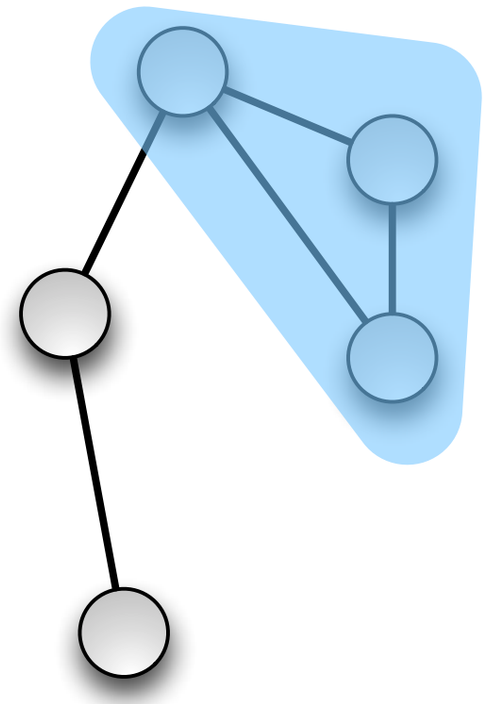
- size $n^{\Omega(k)}$, no matter the degree.

Open problem: *k*-Clique formula

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?

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

Random $3\text{-}XOR$ is hard for Sums-of-Squares w.r.t. size

[Kojevnikov, Itsykson, 2006]

Our reduction to $3\text{-}XOR$ is also size efficient…

Random $3\text{-}XOR$ is hard for Sums-of-Squares w.r.t. size

[Kojevnikov, Itsykson, 2006]

Our reduction to $3\text{-}XOR$ is also size efficient…

…assuming no negative variables

Corollary: lower bound if no negative variables.

K-Clique is still open even for resolution:

[Beyersdorff, Galesi, **L**, Razborov, 2012] conjecture length $|V|^{\Omega(k)}$

[Beyersdorff, Galesi, **L**, 2013] prove it for treelike resolution

[**L**, Pudlák, Rödl, Thapen, 2013] prove it for the "wrong" encoding

[Beame, Impagliazzo, Sabharwal, 2007] size $>2^{\Omega(|V|)}$ even for $k \approx |V|$

Thank you