MINI-TUTORIAL ON SEMI-ALGEBRAIC PROOF SYSTEMS

Albert Atserias
Universitat Politècnica de Catalunya
Barcelona

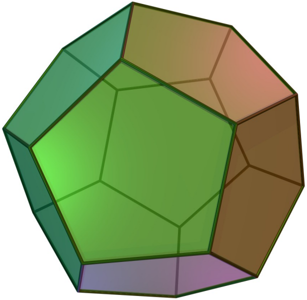
Part I

CONVEX POLYTOPES

Convex polytopes as linear inequalities

Polytope:

$$P = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{A}\mathbf{x} \geq \mathbf{b}\}$$

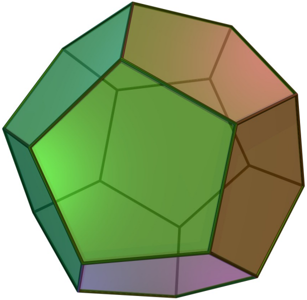


Created by Wikipedia User:Cyp

Convex polytopes as convex hulls

Polytope:

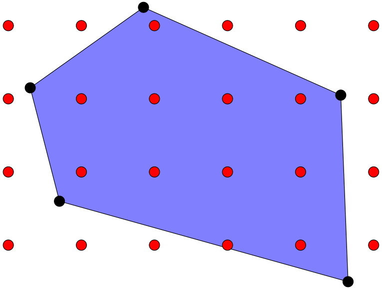
$$P = \text{conv}(\{\mathbf{x}_1, \ldots, \mathbf{x}_m\})$$



Created by Wikipedia User:Cyp

Integer hull

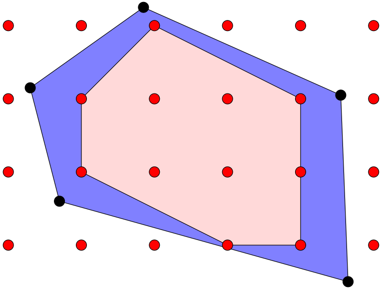
Integer hull of $P \subseteq \mathbb{R}^n$:



Integer hull

Integer hull of $P \subseteq \mathbb{R}^n$:

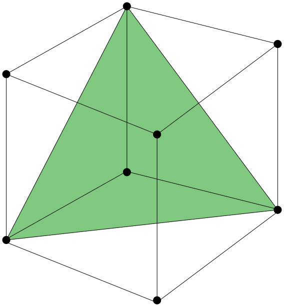
$$P_I = \mathrm{conv}(P \cap \mathbb{Z}^n)$$



Case of special interest: relaxations of 0-1 problems

Polytopes inscribed in the unit cube:

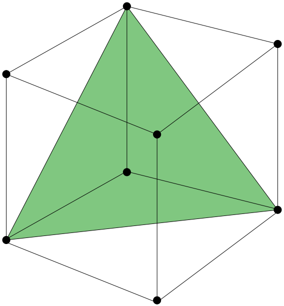
$$\mathrm{conv}\{\mathbf{x} \in \{0,1\}^n : \mathbf{A}\mathbf{x} \geq \mathbf{b}\} = \mathrm{conv}(\{\mathbf{x}_1, \ldots, \mathbf{x}_t\})$$



Case of special interest: relaxations of 0-1 problems

Polytopes inscribed in the unit cube:

$$\mathrm{conv}\{\mathbf{x} \in \{0,1\}^n : \mathbf{A}\mathbf{x} \geq \mathbf{b}\} = \mathrm{conv}(\{\mathbf{x}_1, \ldots, \mathbf{x}_t\})$$



Obvious relaxation:

- What's available: $P = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{A}\mathbf{x} \geq \mathbf{b},\ \mathbf{0} \leq \mathbf{x} \leq \mathbf{e}\}$
- What we want: P_I

Part II

EXPLICIT REPRESENTATIONS OF P_I

Gomory-Chvátal cuts: $C(P)$

Inference rules:

$$\frac{\mathbf{a}_1^{\mathrm{T}}\mathbf{x} \geq b_1 \qquad \cdots \qquad \mathbf{a}_m^{\mathrm{T}}\mathbf{x} \geq b_m}{\sum_{i=1}^{m} c_i \mathbf{a}_i^{\mathrm{T}}\mathbf{x} \geq \sum_{i=1}^{m} c_i b_i} \quad (c_1, \ldots, c_m \in \mathbb{R}^+) \tag{1}$$

$$\frac{\mathbf{a}^{\mathrm{T}}\mathbf{x} \geq b}{\mathbf{a}^{\mathrm{T}}\mathbf{x} \geq \lceil b \rceil} \quad (\mathbf{a} \in \mathbb{Z}^n) \tag{2}$$

New polytope:

1. start at inequalities defining P
2. first close them under (1)
3. then close them under (2)

$C(P)$ is defined by resulting inequalities

Completeness

Completeness [Chvátal 1973]:

$$P \supseteq C(P) \supseteq C(C(P)) \supseteq \cdots \supseteq C^{(t)}(P) = P_I$$

Completeness

Completeness [Chvátal 1973]:

$$P \supseteq C(P) \supseteq C(C(P)) \supseteq \cdots \supseteq C^{(t)}(P) = P_I$$

(with $t \leq n^2 \log n$ if $P \subseteq [0,1]^n$ [ES03]).

Chvátal's "slogan"

Slogan:

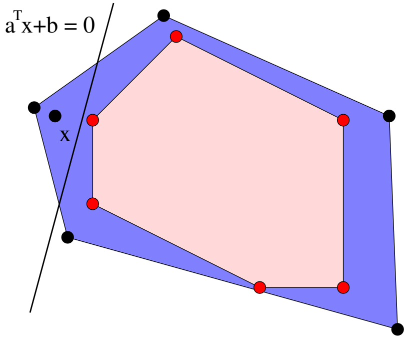
$$\boxed{\text{combinatorics} = \text{linear programming} + \text{number theory}}$$

(the box is Chvátal's)

A little problem

Theorem [Eisenbrand 1999]:

Given P as input, the separation problem for $C(P)$ is NP-hard



Lift-and-project methods and semialgebraic proofs

In the 1990's:

- Sherali and Adams. "A hierarchy of relaxations between the continuous and [...] 0-1 programming problems", 1990.
- Lovász and Schrijver. "Cones of Matrices and Set-Functions and 0-1 Optimization", 1991.
- Balas, Ceria, and Cornuéjols. "A lift-and-project cutting plane algorithm for mixed 0-1 programs", 1993.

Lift-and-project methods and semialgebraic proofs

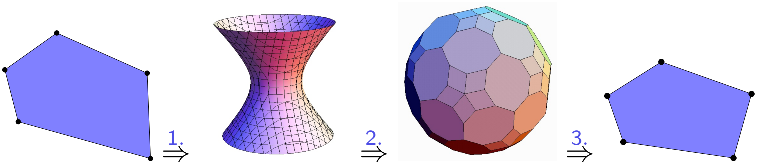
In the 1990's:

- Sherali and Adams. "A hierarchy of relaxations between the continuous and [...] 0-1 programming problems", 1990.
- Lovász and Schrijver. "Cones of Matrices and Set-Functions and 0-1 Optimization", 1991.
- Balas, Ceria, and Cornuéjols. "A lift-and-project cutting plane algorithm for mixed 0-1 programs", 1993.

Semi-algebraic proof systems:

- Grigoriev and Vorobyov. "Complexity of Null- and Positivstellensatz Proofs", 2001.
- Grigoriev, Hirsch, and Pasechnik. "Complexity of semi-algebraic proofs", 2002.

Lift-and-project cuts, graphically



3D-graphics by Mathematica

Steps:

1. lift by products and new variables y_{ij} $(= x_i x_j)$
2. linearize by using $x_i = x_i^2 = y_i$ and forgetting products
3. project by a linear map that eliminates y-variables

Lift-and-project cuts: $N(P)$

Inference rules:

$$\frac{L(\mathbf{x}) \geq 0}{L(\mathbf{x})x_i \geq 0} \qquad \frac{L(\mathbf{x}) \geq 0}{L(\mathbf{x})(1 - x_i) \geq 0} \tag{3}$$

$$\frac{\emptyset}{x_i^2 - x_i \geq 0} \qquad \frac{\emptyset}{x_i - x_i^2 \geq 0} \tag{4}$$

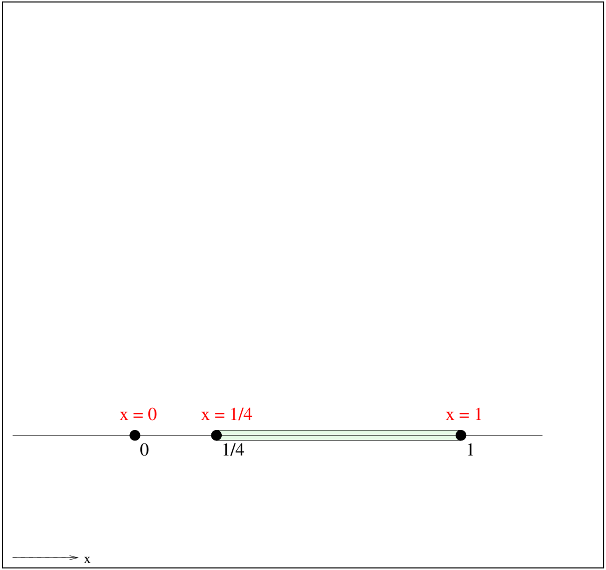
$$\frac{Q_1(\mathbf{x}) \geq 0 \quad \cdots \quad Q_m(\mathbf{x}) \geq 0}{\sum_{i=1}^m c_i Q_i(\mathbf{x}) \geq 0} \quad (c_1, \ldots, c_m \in \mathbb{R}^+) \tag{5}$$

New polytope:

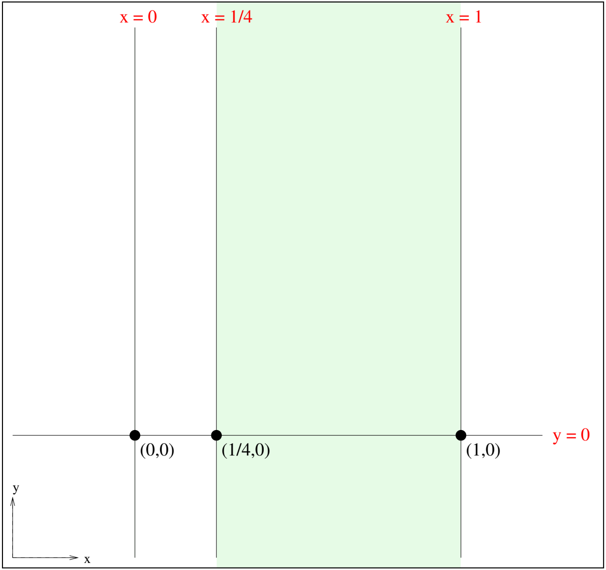
1. Start at inequalities defining P,
2. first lift them through (3) and (4) to degree 2,
3. then project them through (5):

$N(P)$ is defined by resulting linear inequalities.

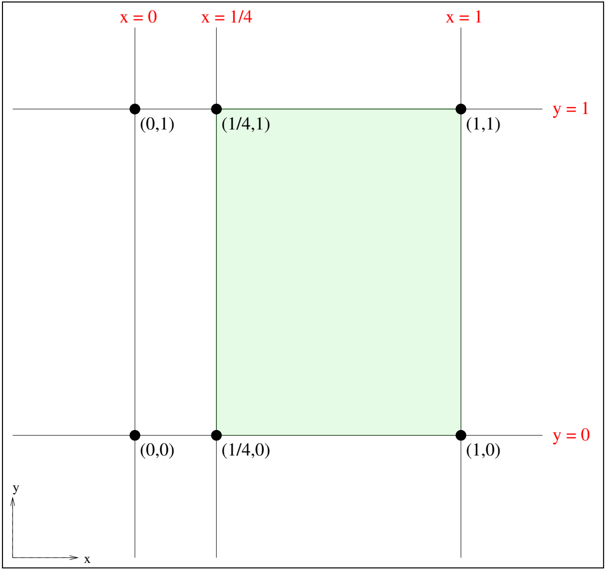
Example: $x - 1/4 \geq 0$ with $x \geq 0$ and $1 - x \geq 0$



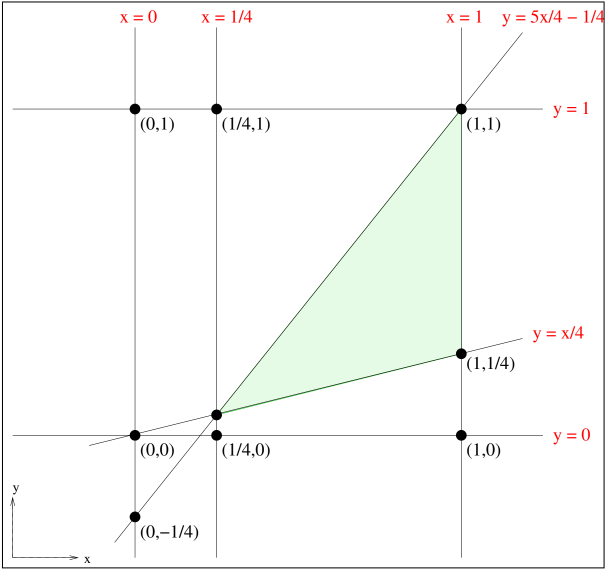
Add a new dimension $y(= x^2)$



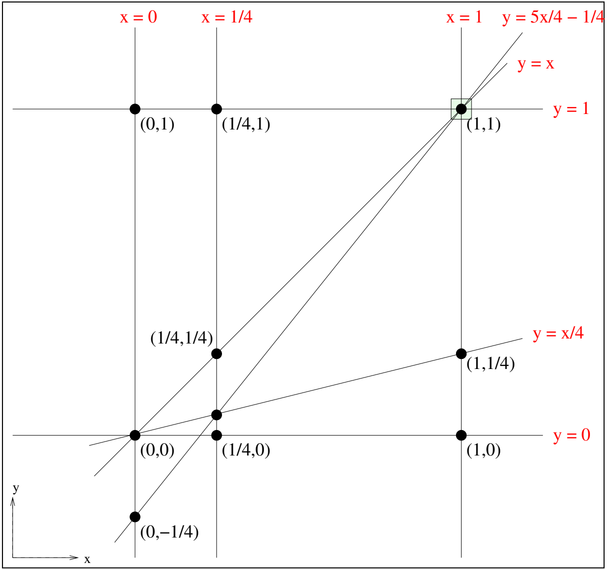
Add $y \geq 0$ and $1 - y \geq 0$



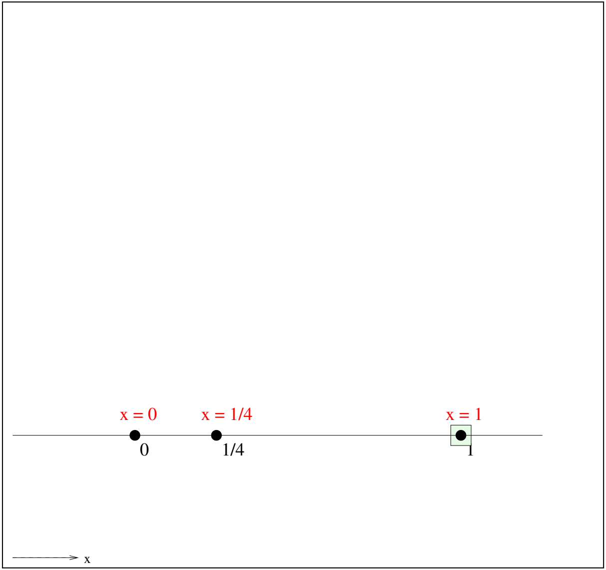
Add $(x - 1/4)x \geq 0$ and $(x - 1/4)(1 - x) \geq 0$



Add $y = x$ to enforce $x^2 = x$



Project back to dimension x



Completeness and algorithmic goodness

Completeness [Lovász-Schrijver]:

$$P \supseteq N(P) \supseteq N(N(P)) \supseteq \cdots \supseteq N^{(n)}(P) = P_I$$

Tractable separation problem [Lovász-Schrijver]:

For $N(P)$, solvable in time $\mathrm{poly}(s+n)$.
For $N^{(d)}(P)$, solvable in time $\mathrm{poly}(s+n^d)$.

(s is the bit-size of the given representation of P)

Lift-and-project degree-d cuts: $N_d(P)$

Lift-and-project degree-d cuts: $N_d(P)$

Inference rules:

$$\frac{Q(\mathbf{x}) \geq 0}{Q(\mathbf{x})x_i \geq 0} \qquad \frac{Q(\mathbf{x}) \geq 0}{Q(\mathbf{x})(1 - x_i) \geq 0} \tag{6}$$

$$\frac{\emptyset}{x_i^2 - x_i \geq 0} \qquad \frac{\emptyset}{x_i - x_i^2 \geq 0} \tag{7}$$

$$\frac{Q_1(\mathbf{x}) \geq 0 \quad \cdots \quad Q_m(\mathbf{x}) \geq 0}{\sum_{i=1}^m c_i Q_i(\mathbf{x}) \geq 0} \quad (c_1, \ldots, c_m \in \mathbb{R}^+) \tag{8}$$

New polytope:

1. Start at inequalities defining P,
2. first lift them through (6) and (7) up to degree d,
3. then project them through (8):

$N_d(P)$ is defined by resulting linear inequalities.

Lift-and-project degree-d semidefinite cuts: $N_{d,+}(P)$

Inference rules:

$$\frac{Q(\mathbf{x}) \geq 0}{Q(\mathbf{x})x_i \geq 0} \qquad \frac{Q(\mathbf{x}) \geq 0}{Q(\mathbf{x})(1 - x_i) \geq 0} \tag{9}$$

$$\frac{\emptyset}{x_i^2 - x_i \geq 0} \qquad \frac{\emptyset}{x_i - x_i^2 \geq 0} \qquad \frac{\emptyset}{Q(\mathbf{x})^2 \geq 0} \tag{10}$$

$$\frac{Q_1(\mathbf{x}) \geq 0 \quad \cdots \quad Q_m(\mathbf{x}) \geq 0}{\sum_{i=1}^m c_i Q_i(\mathbf{x}) \geq 0} \quad (c_1, \ldots, c_m \in \mathbb{R}^+) \tag{11}$$

New polytope:

1. Start at inequalities defining P,
2. first lift them through (9) and (10) up to degree d,
3. then project them through (11):

$N_{d,+}(P)$ is defined by resulting linear inequalities.

Comparison

Sandwich:

$$P \supseteq N^{(d)}(P) \supseteq N_d(P) \supseteq N_{d,+}(P) \supseteq P_I$$

for every $d \geq 2$.

Tractable separation problem:

For $N_{d,+}(P)$, solvable in time $\mathrm{poly}(s + n^d)$.

(again s is the bit-size of the given representation of P)

Measures

Lovász-Schrijver rank / **LS semidefinite rank**:

- min k such that $N^{(k)}(P) = \emptyset$
- min k such that $N_+^{(k)}(P) = \emptyset$

Sherali-Adams degree / **Lasserre degree**:

- min d such that $N_d(P) = \emptyset$
- min d such that $N_{d,+}(P) = \emptyset$

Measures

Lovász-Schrijver rank / **LS semidefinite rank**:

- min k such that $N^{(k)}(P) = \emptyset$
- min k such that $N_+^{(k)}(P) = \emptyset$

Sherali-Adams degree / **Lasserre degree**:

- min d such that $N_d(P) = \emptyset$
- min d such that $N_{d,+}(P) = \emptyset$

YOU NAME IT (LS size, LS$_+$ tree-size, SOS, etc...)

Part III

UPPER BOUNDS

Stable set polytope

$\mathrm{STAB}(G)$ **and** $\mathrm{FRAC}(G)$ **for a graph** $G = (V, E)$:

$$0 \leq x_u \leq 1 \qquad \text{for every vertex } u \in V$$

$$1 - x_u - x_v \geq 0 \qquad \text{for every edge } \{u, v\} \in E$$

Clique constraints are valid for $\mathrm{STAB}(G)$:

$$1 - \sum_{u \in S} x_u \geq 0 \qquad \text{for every clique } S \text{ in } G$$

Question:

What is smallest $d \geq 1$ so that
all clique constraints are valid in $N_{d,+}(\mathrm{FRAC}(G))$?

Stable set polytope (cntd)

Answer is $d = 2$**!** [Lovász-Schrijver]:

Stable set polytope (cntd)

Answer is $d = 2$**!** [Lovász-Schrijver]:

$$(1 - x_u - x_v)x_u \qquad (x_u^2 - x_u) \qquad (1 - \textstyle\sum_u x_u)^2$$

Stable set polytope (cntd)

Answer is $d = 2$**!** [Lovász-Schrijver]:

$$\sum_u \sum_{v:v\neq u} (1 - x_u - x_v)x_u + \sum_u (x_u^2 - x_u)(n-2) + (1 - \sum_u x_u)^2$$

Stable set polytope (cntd)

Answer is $d = 2$**!** [Lovász-Schrijver]:

$$\sum_u \sum_{v:v\neq u} (1 - x_u - x_v)x_u + \sum_u (x_u^2 - x_u)(n-2) + (1 - \sum_u x_u)^2$$
$$=$$

Stable set polytope (cntd)

Answer is $d = 2$**!** [Lovász-Schrijver]:

$$\sum_u \sum_{v:v\neq u} (1 - x_u - x_v)x_u + \sum_u (x_u^2 - x_u)(n-2) + (1 - \sum_u x_u)^2$$
$$=$$
$$1 - \sum_u x_u$$

Stable set polytope (cntd)

Answer is $d = 2$**!** [Lovász-Schrijver]:

$$\sum_u \sum_{v:v\neq u} (1 - x_u - x_v)x_u + \sum_u (x_u^2 - x_u)(n-2) + (1 - \sum_u x_u)^2$$
$$=$$
$$1 - \sum_u x_u$$

Corollary [Grötschel-Lovász-Schrijver 1981]:

The weighted maximum independent set problem
is solvable in polynomial time on perfect graphs.

Pigeonhole principle $n+1$ to n

Representing the usual clauses:

$$\begin{array}{lll} \text{a.} & x_{i,1} \vee \cdots \vee x_{i,n} & \Longrightarrow \quad \sum_k x_{i,k} - 1 \geq 0 \\ \text{b.} & \neg x_{i,k} \vee \neg x_{j,k} & \Longrightarrow \quad 1 - x_{i,k} - x_{j,k} \geq 0 \end{array}$$

Pigeonhole principle $n+1$ to n

Representing the usual clauses:

a. $x_{i,1} \vee \cdots \vee x_{i,n} \implies \sum_k x_{i,k} - 1 \geq 0$

b. $\neg x_{i,k} \vee \neg x_{j,k} \implies 1 - x_{i,k} - x_{j,k} \geq 0$

But wait!:

Pigeonhole principle $n+1$ to n

Representing the usual clauses:

a. $x_{i,1} \vee \cdots \vee x_{i,n} \implies \sum_k x_{i,k} - 1 \geq 0$
b. $\neg x_{i,k} \vee \neg x_{j,k} \implies 1 - x_{i,k} - x_{j,k} \geq 0$

But wait!:

$1 - \sum_i x_{i,k} \geq 0$ from b. in one N_+ round as in clique

Pigeonhole principle $n+1$ to n

Representing the usual clauses:

$$\begin{array}{lll} \text{a.} & x_{i,1} \vee \cdots \vee x_{i,n} & \Longrightarrow & \sum_k x_{i,k} - 1 \geq 0 \\ \text{b.} & \neg x_{i,k} \vee \neg x_{j,k} & \Longrightarrow & 1 - x_{i,k} - x_{j,k} \geq 0 \end{array}$$

But wait!:

$1 - \sum_i x_{i,k} \geq 0$	from b. in one N_+ round as in clique
$n - \sum_k \sum_i x_{i,k} \geq 0$	from previous by addition

Pigeonhole principle $n+1$ to n

Representing the usual clauses:

$$\begin{array}{lll} \text{a.} & x_{i,1} \vee \cdots \vee x_{i,n} & \Longrightarrow \quad \sum_k x_{i,k} - 1 \geq 0 \\ \text{b.} & \neg x_{i,k} \vee \neg x_{j,k} & \Longrightarrow \quad 1 - x_{i,k} - x_{j,k} \geq 0 \end{array}$$

But wait!:

$1 - \sum_i x_{i,k} \geq 0$	from b. in one N_+ round as in clique
$n - \sum_k \sum_i x_{i,k} \geq 0$	from previous by addition
$\sum_i \sum_k x_{i,k} - (n+1) \geq 0$	from a. by addition

Pigeonhole principle $n + 1$ to n

Representing the usual clauses:

$$\begin{array}{lll} \text{a.} & x_{i,1} \vee \cdots \vee x_{i,n} & \Longrightarrow & \sum_k x_{i,k} - 1 \geq 0 \\ \text{b.} & \neg x_{i,k} \vee \neg x_{j,k} & \Longrightarrow & 1 - x_{i,k} - x_{j,k} \geq 0 \end{array}$$

But wait!:

$1 - \sum_i x_{i,k} \geq 0$	from b. in one N_+ round as in clique
$n - \sum_k \sum_i x_{i,k} \geq 0$	from previous by addition
$\sum_i \sum_k x_{i,k} - (n+1) \geq 0$	from a. by addition
$-1 \geq 0$	from previous two by addition

Some additional facts

Proof complexity:

- width-w resolution ref. $\Rightarrow N_w = \emptyset$
- size-s resolution ref. $\Rightarrow$ size-$O(s)$ LS ref. [Pudlák 1999]
- tree-size-s LS ref. $\Rightarrow N^{(\sqrt{n \log s})} = \emptyset$ [Pitassi-Segerlind 2012]

Some additional facts

Proof complexity:

- width-w resolution ref. $\Rightarrow N_w = \emptyset$
- size-s resolution ref. $\Rightarrow$ size-$O(s)$ LS ref. [Pudlák 1999]
- tree-size-s LS ref. $\Rightarrow N^{(\sqrt{n \log s})} = \emptyset$ [Pitassi-Segerlind 2012]

Combinatorial problems:

- $N_{2,+}$ on MAX-CUT gives 0.878-approximation [GW96]
- $N_{9,+}$ solves all its known gap examples [Mossel 2013]
- N_{15} solves graph isomorphism on planar graphs [AM12]

Some additional facts

Proof complexity:

- width-w resolution ref. $\Rightarrow N_w = \emptyset$
- size-s resolution ref. $\Rightarrow$ size-$O(s)$ LS ref. [Pudlák 1999]
- tree-size-s LS ref. $\Rightarrow N^{(\sqrt{n \log s})} = \emptyset$ [Pitassi-Segerlind 2012]

Combinatorial problems:

- $N_{2,+}$ on MAX-CUT gives 0.878-approximation [GW96]
- $N_{9,+}$ solves all its known gap examples [Mossel 2013]
- N_{15} solves graph isomorphism on planar graphs [AM12]

Interpolation:

LS has feasible interpolation [Pudlák 1999]
LS_+ has feasible interpolation [Dash 2001]

Part IV

LOWER BOUNDS

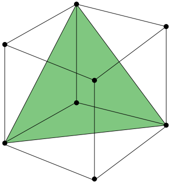
How to prove lower bounds?

Goal:

Build a feasible solution for $N_d(P)$
by patching together local (i.e. partial) fractional solutions

Useful observation:

local fractional solution $\equiv$ prob. dist. on local 0-1 solutions



How to prove lower bounds? (cntd)

System of d-local distributions for P:

$$H = \{\mu_X : X \subseteq [n],\ |X| \leq d\}$$

such that

1. μ_X: a prob. dist. on $\{0,1\}^X$ with support in $P|_X \cap \{0,1\}^X$
2. $\mu_X(\mathbf{x}) = \sum_{\mathbf{y}:\mathbf{y}\supseteq\mathbf{x}} \mu_Y(\mathbf{y})$ for each $X \subseteq Y$ and $\mathbf{x} \in \{0,1\}^X$

How to prove lower bounds? (cntd)

System of d-local distributions for P:

$$H = \{\mu_X : X \subseteq [n],\ |X| \leq d\}$$

such that

1. μ_X: a prob. dist. on $\{0,1\}^X$ with support in $P|_X \cap \{0,1\}^X$
2. $\mu_X(\mathbf{x}) = \sum_{\mathbf{y}:\mathbf{y}\supseteq\mathbf{x}} \mu_Y(\mathbf{y})$ for each $X \subseteq Y$ and $\mathbf{x} \in \{0,1\}^X$

Theorem: The following are equivalent:

1. there is a system of d-local distributions for P,
2. $N_d(P) \neq \emptyset$.

(analogue for $N_{d,+}$ too)

3-XOR-SAT

Systems of linear equations mod 2:

$$\left[\begin{array}{ccc} x_{i_1} \oplus x_{j_1} \oplus x_{k_1} & = & a_1 \\ & \vdots & \\ x_{i_m} \oplus x_{j_m} \oplus x_{k_m} & = & a_m \end{array}\right]$$

Encoding:

Each equation in CNF, then as a polytope in $\mathbb{R}^3$.

From Gaussian-width to N_+-degree

Gaussian calculus:

$$\frac{\bigoplus_{i \in I} x_i = a \qquad \bigoplus_{j \in J} x_j = b}{\bigoplus_{k \in I \triangle J} x_k = a \oplus b}$$

Lemma [Schoenebeck 2008]

If refuting S requires Gaussian-width $> d$,
then $N_{d/2,+}(S) \neq \emptyset$.

Corollary [Schoenebeck 2008, Grigoriev 2001]:

Tseitin formulas, random systems mod 2, etc require
Lasserre degree $\Omega(n)$ and tree-like LS_+ size $2^{\Omega(n)}$.

Schoenebeck's construction

Define:

- Let $\mathcal{C}$ be all (A, a) such that $S \vdash_d \bigoplus_{i \in A} x_i = a$,
- let $\pi(A) := (-1)^a$ if $(A, a) \in \mathcal{C}$ (note: $(A, 1-a) \notin \mathcal{C}$),
- let $A \sim B$ if $(A \triangle B, c) \in \mathcal{C}$ for some c for $|A|, |B| \leq d/2$,
- and

$$\mu_X(\mathbf{x}) := \sum_{[A]} \left(\sum_{B \sim A} \pi(B) \widehat{I}_{X=\mathbf{x}}(B) \right)^2$$

Part V

SOME OPEN PROBLEMS

Open problems

Use it for SAT:

Can we integrate semialgebraic methods into symbolic solvers?

Open problems

Use it for SAT:

Can we integrate semialgebraic methods into symbolic solvers?

Lower bounds on LS size:

Prove a superpolynomial lower bound for dag-like LS_+ (or LS)

Open problems

Use it for SAT:

Can we integrate semialgebraic methods into symbolic solvers?

Lower bounds on LS size:

Prove a superpolynomial lower bound for dag-like LS_+ (or LS)

"Learning" the linear transformation?:

Under $y_i := 1 - 2x_i$, parities are $\prod_{i \in I} y_i = \pm 1$. Useful?

Open problems

Use it for SAT:

Can we integrate semialgebraic methods into symbolic solvers?

Lower bounds on LS size:

Prove a superpolynomial lower bound for dag-like LS_+ (or LS)

"Learning" the linear transformation?:

Under $y_i := 1 - 2x_i$, parities are $\prod_{i \in I} y_i = \pm 1$. Useful?

Find MAX-CUT gaps or improve over GW:

Does degree-$n^{o(1)}$ Lasserre leave a 0.878 gap for MAX-CUT?