Depth lower bounds in Stabbing Planes

Nicola Galesi

Department of Computer, Control and Management Engineering "A.Ruberti"
Sapienza, Rome

Mathematical approaches to lower bounds: complexity of proofs and computations - ICMS, Edinburgh

July 6, 2022

Joint work with: Stefan Dantchev, Abdul Ghani, Barnaby Martin

Branch-&-Cut

Let $\mathbf{Ax} \geq \mathbf{b}$ be a system of linear inequalities defining a polytope $\mathcal{P}$.

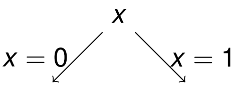
Stabbing Planes (also Branch-&-Cut) is a method implementing the search for an integer point inside $\mathcal{P}$ by :

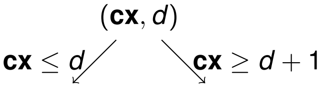
- *branching* $\mathcal{P}$ into smaller polytopes $P_1 \ldots P_k$ such that every integer solution of $\mathcal{P}$ lies in at least one of $P_1, \ldots, P_k$;
- *cut*. Add further cutting planes to refine $P_1, \ldots, P_k$ and recurse the search on the smaller refined polytopes.

Stabbing Planes

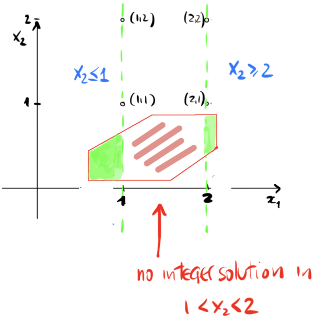
Assume $\mathbf{Ax} \geq \mathbf{b}$ does not admit integral points.

As a proof system *Stabbing Planes* can be seen as *DPLL* where instead of querying a variable x and splitting the two cases $x = 1$ and $x = 0$, we query a pair $Q = (\mathbf{cx}, d)$ with $c_i, b \in \mathbb{Z}$ and split according to the two cases $cx \leq d$ and $cx \geq d + 1$. The search terminates when we reach the empty polytope $\emptyset$.

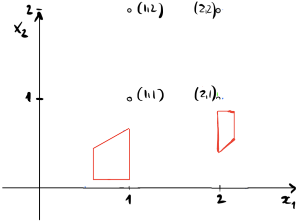




Example



x_2
2
1
(1,2)
(2,2)
(1,1)
(2,1)
$x_2 \leq 1$
$x_2 \geq 2$
1
2
x_1
no integer solution in
$1 < x_2 < 2$



x_2
2
1
(1,2)
(2,2)
(1,1)
(2,1)
1
2
x_1

Definition of Stabbing Planes

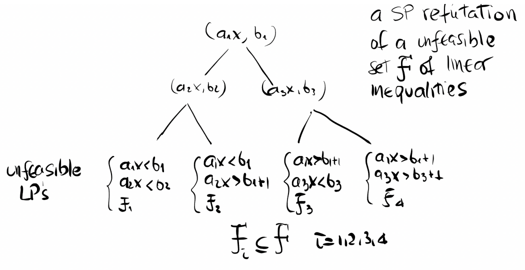
[Beame Fleming Impagliazzo Kolokolova Pankratov Pitassi Robere 19], [Fleming Göös, Impagliazzo Pitassi Robere Tan Wigderson 21]

Let $\mathcal{F} := \mathbf{Ax} \geq \mathbf{b}$ be an unsatisfiable system of linear inequalities. A *Stabbing Planes (SP)* refutation of $\mathcal{F}$ is a directed binary tree $\mathcal{T}$ such that

- *Internal Nodes*, are labelled with a pair $(\mathbf{c}, d)$ with $\mathbf{c} \in \mathbb{Z}^n, d \in \mathbb{Z}$. The right outgoing edge is labelled with $\mathbf{cx} \leq d$, and the left outgoing edge is labelled with its integer negation $\mathbf{cx} \geq d + 1$.
- *Leaves*. Each leaf node ℓ of $\mathcal{T}$ is labelled with a conic combination of inequalities in $\mathcal{F}$ with the inequalities along the path leading to an unfeasible *LP*, equivalent to $0 \geq 1$.

Example and measures of complexity

An example of a tree proof with the LPs on the leaves.



$\mathcal{F}$ the initial inequalities,

SP Complexity measures

Size of $\mathcal{T}$ = # of nodes in the tree $\mathcal{T}$

Depth of $\mathcal{T}$ = depth of the tree $\mathcal{T}$

Relation with Cutting Planes *CP*

Cutting Planes (CP) refutation system for unfeasible families of integer linear inequalities, is a Hilbert-style system equipped with boolean axioms $0 \leq x \leq 1$ and two inference rules:

Linear Combination $$\frac{\mathbf{ax} \geq c \qquad \mathbf{bx} \geq d}{\alpha\mathbf{ax} + \beta\mathbf{bx} \geq \alpha c + \beta d}$$

Rounding $$\frac{\alpha\mathbf{ax} \geq b}{\mathbf{ax} \geq \lceil b/\alpha \rceil}$$

with $\mathbf{a}, \mathbf{b} \in \mathbb{Z}^n$, and $c, d, \alpha, \beta \in \mathbb{Z}$.

CP complexity measures

Proof: $L_1, \ldots, L_{k-1}, 0 \geq 1$, or the usual associated DAG with $0 \geq 1$ as the only sink node.

Size: # of inequalities in the proof.

Rank = maximal number of applications of the rounding rule along a path from an axiom to $0 \geq 1$ in the DAG.

Known results and *SP* proof strenght

Important SP results obtained in [Beame et al.], [Fleming et al.]

- *SP* poly simulates *CP*
- *CP* quasipoly simulates *SP**
- *SP* poly equivalent to treelike-Res(*CP*)

To our work, more important, are the following results

- There are quasipolynomial size and $O(\log^2 n)$-depth *SP* proofs for Tseitin contradictions Ts(G) over a graph G with n nodes.
- There exists a family of formulas which requires *SP* proofs of depth $\Omega(\frac{n}{\log^2 n})$.

  This result uses similar techniques used for treelike *CP*: reduce shallow *SP* proofs to efficient real communication protocols for certain functions, which instead does not admit efficient protocols.

Motivations of our work

- PHP_n^m can be refuted in *SP* with depth $O(\log n)$ (and poly size, of course)
- $\mathrm{Ts}(G)$ can be refuted in depth $O(\log^2 n)$.This bound is conjectured optimal in [Beame et al.]
- The $\Omega(\frac{n}{\log^2 n})$ lower bound is for a lifted family of formulas $\mathrm{Ts}(G) \circ g$, with G an expander graph.

No technique not using communication complexity was known to prove depth lower bounds, for example for PHP or $\mathrm{Ts}(G)$.

CNF Formulas and inequalities

From CNFs to set of linear inequalities

$$\bigvee_i x_i \vee \bigvee_j \bar{y}_j \mapsto \sum_i x_i + \sum_j (1 - y_j) \geq 1$$

Tseitin contradictions

For a graph $G = (V, E)$ with a charging function $\omega : V \to \{0, 1\}$ satisfying $\sum_{v \in V} \omega(v) = 1 \mod 2$, the *Tseitin contradiction* $\mathrm{Ts}(G, \omega)$ is the CNF equivalent of

$$\sum_{e \in E, e \ni v} x_e = \omega(v) \mod 2 \qquad v \in V,$$

where x_e ranges over $e \in E$.

Our Results

- $\Omega(\log n)$ depth lower bounds for PHP^m_n, $\mathrm{Ts}(K_n)$, $\mathrm{Ts}(H_n)$ and for LOP_n.
- an incomparability result for rank in *CP* and depth in *SP*.

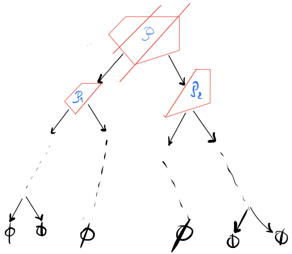
Despite the fact that *SP proofs cannot be balanced* [Beame et al.], that is

size S *SP* proofs cannot be transformed into $poly(S)$ size and $polylog(S)$ depth *SP* proofs

Yet, since *SP* is a treelike system,

One can prove depth lower bounds by proving size lower bounds

High level idea



$\mathcal{F}$ the initial inequalities,
$Q = (\mathbf{cx}, d)$ and $slab(Q) = \{x \in \mathbb{R}^n | d < \mathbf{cx} < d + 1\}$

1. Figure out a *large* family of non integral points satisfying $\mathcal{F}$ (*admissible points*), i.e. in the initial polytope;
2. Argue that each slab excludes only a *limited* number of admissible points;
3. Observe that at the leaves the polytopes are $\emptyset$.

First approach: the antichain method

A *toy example*: the simple *PHP*.
SPHP_n is the following set of unsatisfiable inequalities:

$$\sum_{i=1}^{n} x_i \geq 2$$
$$x_i + x_j \leq 1 \text{ (for all } i \neq j \in [n])$$

Lemma

For $n \geq 3$, SPHP_n

1. *can be refuted in depth $\Theta(\log n)$ in SP, and*
2. *has a rank 1 CP refutation.*

A toy example: SPHP_n

CP rank is 1

Let $S := \sum_{i=1}^{n} x_i$ (so we have $S \geq 2$).

$$\frac{\begin{array}{c} x_2 + x_1 \\ \vdots \\ x_n + x_1 \end{array}}{S + (n-2)x_1 \leq n-1} \quad \cdots \quad \frac{\begin{array}{c} x_1 + x_i \\ \vdots \\ x_{i-1} + x_i \\ x_{i+1} + x_i \\ \vdots \\ x_n + x_1 \end{array}}{S + (n-2)x_i \leq n-1} \quad \cdots \quad \frac{\begin{array}{c} x_1 + x_n \\ \vdots \\ x_{n-1} + x_n \end{array}}{S + (n-2)x_n \leq n-1}$$

$$\frac{-S \leq -2}{x_1 \leq (n-3)/(n-2)} \quad \cdots \quad \frac{-S \leq -2}{x_i \leq (n-3)/(n-2)} \quad \cdots \quad \frac{-S \leq -2}{x_n \leq (n-3)/(n-2)}$$

$$\downarrow \qquad\qquad\qquad \downarrow \qquad\qquad\qquad \downarrow$$

$$x_1 \leq 0 \quad \cdots \quad x_i \leq 0 \quad \cdots \quad x_n \leq 0$$

$$\searrow \qquad \downarrow \qquad \swarrow$$

$$\frac{\begin{array}{c} \sum_{i=1}^{n} x_i \leq 0 \\ -\sum_{i=1}^{n} x_i \leq -2 \end{array}}{0 \leq -2}$$

SP depth is $\Omega(\log n)$

We prove that *SP* size is $\Omega(\sqrt[4]{n})$. The proof consists of four main ingredients

Fact (1)

Define the set of admissible non-integral points and prove its largeness .

for SPHP_n.

$D = \{0, 1/2\}$
$\mathcal{A}_n = \{\mathbf{s} \in D^n | \text{at least 4 coordinates in } \mathbf{s} \text{ are } \frac{1}{2}\}$
$|\mathcal{A}_n| \geq 2^n - 4n^3$

□

SPHP$_n$ - Fact 2

Fact (2)

Argue that at each slab not many admissible points are lost.

for SPHP$_n$.

$\mathbf{a} \in \mathbb{R}^n$, $Q = (\mathbf{a}, b)$, $D = \{0, \frac{1}{2}\}$.

$w(\mathbf{a})$= number of of non-zero coordinates in $\mathbf{a}$

We want to bound the number of $\mathbf{s} \in A_n$ such that $b < \mathbf{as} < b + 1$ and we count the number of $\mathbf{s} \in A_n$ such that $\mathbf{as} = q$, $q \in \mathbb{Q}$.

Sperner's theorem Let $[t]^n$ be equipped with the pointwise ordering $\preceq$ ($\mathbf{a} \preceq \mathbf{b}$ iff $a_i \leq b_i$ for all i). Any antichain A on $[t]^n$ has size $|A| \leq \frac{t^n}{\sqrt{n}}$.

Observation Let $I_{\mathbf{a}} = \{i \in [n] | a_i \neq 0\}$, so that $|I_{\mathbf{a}}| = w(\mathbf{a})$. The solutions $\mathbf{s}$ to $\mathbf{as} = q$ form an antichain on $D^{I_{\mathbf{a}}}$ on the following pointwise order $\preceq_D$:

$$\mathbf{s} \preceq_D \mathbf{t} \text{ iff } \begin{cases} s_i \leq t_i & \text{if } a_i > 0 \\ t_i \leq s_i & \text{if } a_i < 0 \end{cases}$$

Conclusion Each slab Q removes at most $\frac{|D|^{w(\mathbf{a})}}{\sqrt{w(\mathbf{a})}} = \frac{2^{w(\mathbf{a})}}{\sqrt{w(\mathbf{a})}}$ points from A_n. □

SPHP_n - Fact 3

Let $\mathcal{T}$ be a *SP* proof of SPHP_n.
Let $w_{\mathcal{T}} = \min_{Q \in \mathcal{T}} w(Q)$

Fact (3)

Lower bound the size of a SP proof $\mathcal{T}$ in terms of the $w_{\mathcal{T}}$.

$$|\mathcal{T}| \geq \Omega(\sqrt{w_{\mathcal{T}}})$$

for SPHP_n.

$$|\mathcal{T}| \geq \frac{|A_n|}{\frac{2^n}{\sqrt{w_{\mathcal{T}}}}}, \qquad \text{By Fact 2}$$

□

SPHP_n - Fact 4

Fact (4)

Conclude the size lower bound proof showing that $w_{\mathcal{T}} \geq \Omega(t(n))$, for a suitable function t.

for SPHP_n.

Let $t = t(n)$ ($\sqrt[4]{n}$ for SPHP_n) be a parameter.

$$\Sigma_{\mathcal{T}} = \{Q \in \mathcal{T} | w(Q) \leq t^2\}$$

1. $|\Sigma_{\mathcal{T}}| \geq t$: ✓;
2. $|\Sigma_{\mathcal{T}}| = 0$: ✓, by Fact 3;
3. $0 < |\Sigma_{\mathcal{T}}| < t$.

□

SPHP_n - Fact 4

for SPHP_n.

Case 3

1. Each query $Q \in \Sigma_{\mathcal{T}}$ involves at most t^2 variables, hence in total at most t^3 variables X. Define ρ by setting $x = 0$ for $x \in X$.
2. Consider $\mathcal{T}\restriction\rho$ and reiterate the argument.
3. At each iteration at least one query disappears
4. Set $t(n)$ in such a way, that the number of iterations is at least $\Omega(t(n))$.

□

Other results obtained by the antichain method

$PHP_n^m, m > n$

- n. of vars $O(mn)$
- Size bound : $\Omega(n^{1/4})$
- Depth lower bounds: $\Omega(\log n)$
- $D = \{0, \frac{1}{2}\}$
- A_n: set of points with at least two coordinates set to $1/2$

$\mathrm{Ts}(K_n)$

- n. of vars $O(n^2)$
- Size lower bound: $\Omega(n^{1/4})$
- Depth lower bounds: $\Omega(\log n)$
- $D = \{0, \frac{1}{2}, 1\}$
- A_n: set of points such that each nodes in K_n has at least two incident edges set to $1/2$.

LOP_n

- n. of vars $O(n^2)$
- Size lower bound: $\Omega(n^{\frac{1-\epsilon}{4}})$
- Depth lower bounds: $\Omega(\log n)$
- $D = \{0, \frac{1}{2}, 1\}$
- A_n: Given $X \subseteq [n]$ of size $\leq n-3$, coordinates $x_{i,j} = \frac{1}{2}$ if $i, j \notin X$ and $x_{i,j} = 0, 1$ according to the order or if one of $i, j \notin X$.

Second approach: the covering method

[Linial and Radhakrishnan05] studied the minimal number of hyperplanes covering all points of $\{0,1\}^n$. To make the problem meaningful they define *essential coverings* of $\{0,1\}^n$.

Definition

A set L of linear polynomials with real coefficients is said to be an *essential cover* of the cube $\{0,1\}^n$ if

(E1) for each $v \in \{0,1\}^n$, there is a $p \in L$ such that $p(v) = 0$,

(E2) no proper subset of L satisfies (E1), that is, for every $p \in L$, there is a $v \in \{0,1\}^n$ such that p alone takes the value 0 on v, and

(E3) every variable appears (in some monomial with non-zero coefficient) in some polynomial of L.

Theorem (Linial and Radhakrishnan 05)

Any essential cover L of the cube with n coordinates satisfies $|L| \in \Omega(\sqrt{n})$.

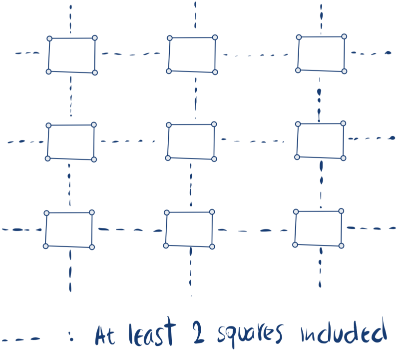
Polynomials

Let $\mathcal{T}$ a *SP* refutation of $\mathrm{Ts}(H_n)$.

We consider the set of polynomials

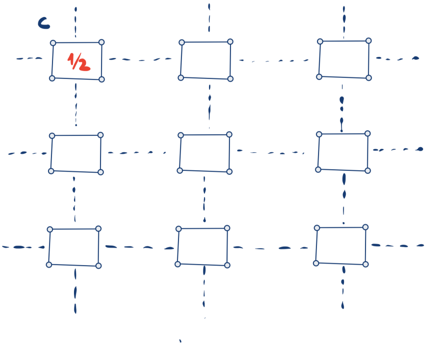
$$\tau = \{\mathbf{ax} = b + 1/2 \mid Q = (\mathbf{ax}, b) \in \mathcal{T}\}$$

Grids and 4-cycles



Let C be the set of such cycles. Notice that $|C| = (n/3)^2$

Admissible points



Lemma

Given $c \in C$, there is an admissible point α^c such that in α^c

- *the variables of the edges in c are set to $1/2$;*
- *the rest of variables are in $\{0, 1\}$.*

Essentialization of a covering

Fact

Let L be a covering of $\{0,1\}^{|A|}$ (hence verifying only (E1)). There is a $L' \subset L$ and a $A' \subseteq A$ such that L' is an essential covering $\{0,1\}^{|A'|}$.

Proof.

Force $(E2)$ and $(E3)$ by choosing the minimal $L' \subseteq L$ covering $\{0,1\}^{|A|}$ and limits the cube to the only variables in A with non-zero coefficients in polynomials of L'. □

We say that (L', A') is an essentialization of (L, A).

Coverings

Definition

Let $c \in C$ and $p \in \tau$ with $p = \sum_{e \in E' \subseteq E(H_n)} a_e x_e$. We say that *p is odd on c* if $\sum_{e \in E' \cap c} a_e = 1 \mod 2$.

Definition

Let $c \in C$. We define $\tau^c = \{p \in \tau \mid p \text{ odd on } c\}$.

Lemma

τ^c covers $\{0,1\}^{|C - \{c\}|}$.

Proof.

Since α^c is admissible, it must be necessarily covered by some $p \in \tau$. Notice that p must be odd on c since c has 4 edges , hence p on α^c can be $1/2 + b$ only if p is odd on c. Hence $p \in \tau^c$. □

Main argument

Notice that $|\tau| \geq |\tau^c|$. We prove that $|\tau^c| = \Omega(n)$.

1. if τ^c is an essential cover of $\{0,1\}^{|C-\{c\}|}$. Then by [LN] and since $|C| = O(n^2)$, $|\tau^c| = \Omega(n)$.
2. If τ^c is only a cover of $\{0,1\}^{|C-\{c\}|}$. We extract an essentialization (τ_1, C_1) of $(\tau^c, C - \{c\})$ and reiterate the argument choosing another $c_1 \in C - (C_1 \cup \{c\})$ until (1) holds or no cycle remain.

Let $(\tau_1, C_1), \ldots, (\tau_q, C_q)$ be the list of refined essentializations. Observe that $\tau^c \geq q$ by def of essentilization. Then

- if $q \geq (n/3)^2/2$, we have done
- if $q < (n/3)^2/2$, then
  - $\sum_{i=1}^q |C_i| \geq (n/3)^2/2$. This is because $|C| = q + \sum_i |C_i|$
  - $\tau = \sum_{i=1}^q |\tau_i|$. This is because τ_i's partitions τ.
  - Hence

$$|\tau| \geq \sum_i |\tau_i| \geq \sum_i \sqrt{|C_i|} \geq \sqrt{\sum_i |C_i|} = \Omega(n)$$

Further devolpements

Size lower bounds are poor:

1. [Yehuda, Yehudayoff 22]. Improve [Linial Radakrishnam 05] lower bound to

Theorem (Yehuda, Yehudayoff 22)

Any essential cover L of the cube with n coordinates satisfies $|L| \in \Omega(n^{0.52})$.

This allow to push our lower bound to $\Omega(n^{1.04})$

2. We have new different results for $\mathrm{Ts}(H_n)$ getting a truly linear size lower bound $\Omega(n^2)$. However still significantly far from proving that *SP* proofs of $\mathrm{Ts}(H_n)$ in [Beame et al.] are optimal wrt size and depth.