Announcements

- Sign-up for slot to discuss your presentation (if you havent already) (see email)

  April 17 presentations : April 10 5-7:30

  April 24 presentations : April 17 5-7:30

- HW2 will be posted next week

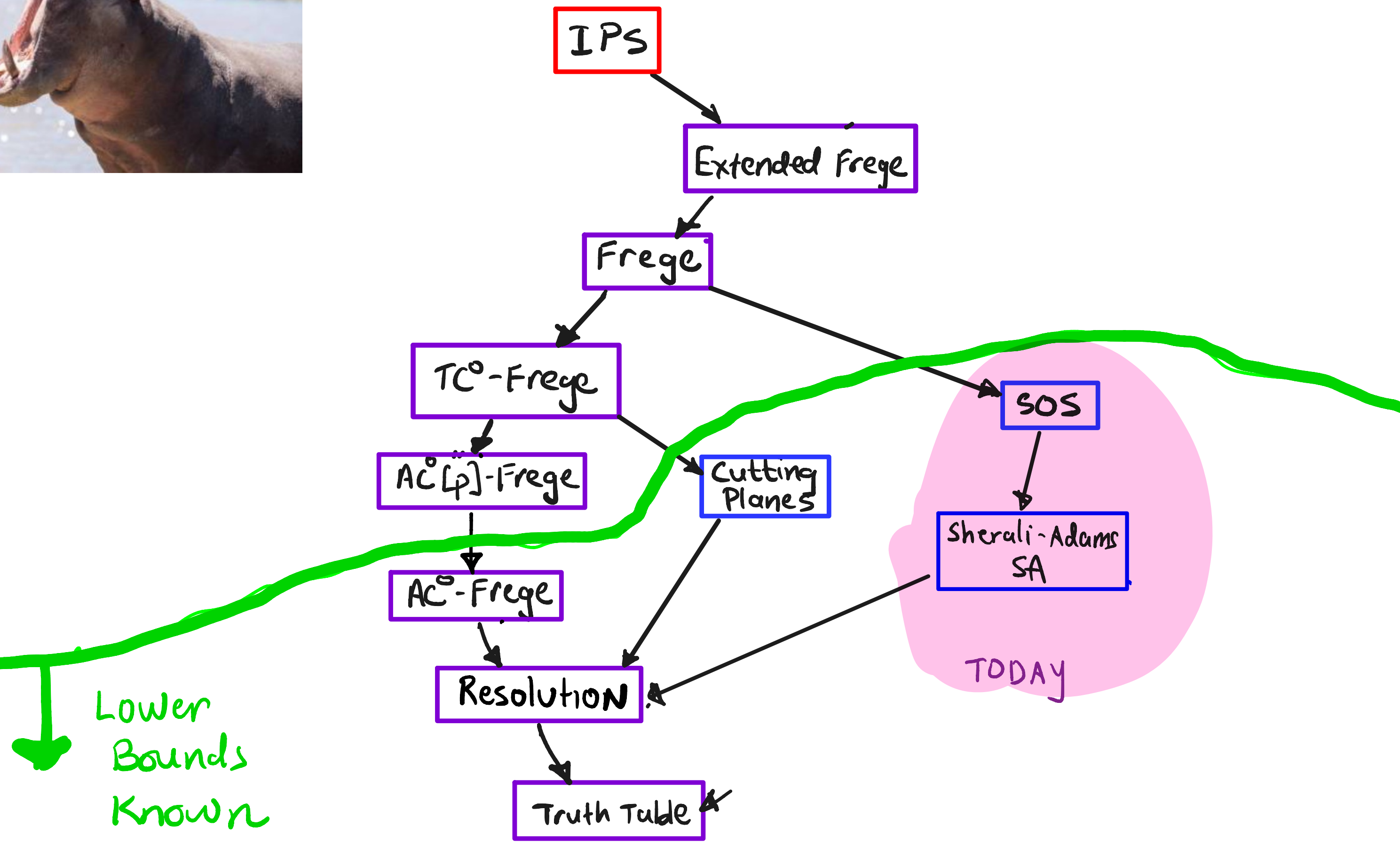
- Turn in your lecture notes — 1st draft due one week after lecture, then final draft due 1 week after you receive feedback. Thanks!

TODAY

Semi-Algebraic Proof Systems

- Sherali Adams (SA)
- Sum-of-squares (SOS)

THE PROOF COMPLEXITY ZOO



Sherali-Adams

SA is a sound + complete proof system for refuting a family of polynomial inequalities over the reals.

For refuting UNSAT CNF formulas:

Let $f = C_1 \wedge C_2 \wedge \cdots \wedge C_m$

Let $C_i = (x_1 \vee \bar{x}_2 \vee x_3)$. Then $\tilde{C}_i = x_1 + (1-x_2) + x_3$

System of inequalities corresponding to f:

$$\tilde{C}_i - 1 \geq 0 \quad \forall i \in [m]$$

plus $x_i^2 - x_i = 0 \quad \forall i \in [n]$

Sherali-Adams (SA)

Defn. A conjunction is an AND of Boolean literals $D = \ell_1 \wedge \ell_2 \wedge \dots \wedge \ell_k$

We can encode a conjunction as a polynomial over the reals as follows:

If $D = \bigwedge_{i \in S} x_i \wedge \bigwedge_{j \in T} \bar{x}_j$, then can write D equivalently as:

$$D = \prod_{i \in S} x_i \prod_{j \in T} (1 - x_j)$$

Defn A conical junta is a non-negative linear combination of juntas

$$J = \sum_i \lambda_i D_i \,, \quad \text{where } \lambda_i \geq 0 \,,\ D_i \text{ a conjunction}$$

* Conical juntas are non-negative over $\{0,1\}$ assignments

Sherali-Adams (SA)

Defn. Let $f = C_1 \wedge C_2 \wedge \dots \wedge C_m$ be an unsat CNF over $x_1, \dots, x_n$.

A Sherali-Adams refutation of f is given by:

$$\underbrace{\{J_0, J_1, \dots, J_m\}}_{\text{conical juntas}} + \underbrace{\{q_1, q_2, \dots, q_n\}}_{\text{polynomials}} \quad \text{such that}$$

$$J_0 + \sum_{i=1}^{m} J_i \cdot (\tilde{C}_i - 1) + \underbrace{\sum_{i=1}^{n} q_i \cdot (x_i^2 - x_i)}_{(*)} = -1$$

Multilinearization: Equivalently we can drop the last part (*) and write a SA refutation as simply $J_0 + \sum_{i=1}^{m} J_i \cdot (\tilde{C}_i - 1)$, where arithmetic is done as multilinear polynomials (replace x^i, $i \geq 1$ by x_i)

Sherali-Adams (SA)

Soundness Let CNF F have a SA refutation. Then F is UNSAT.

Pf: Let $\Pi = J_0 + \sum_{i=1}^{m} J_i (\tilde{C}_i - 1) + \sum_{i=1}^{n} q_i (x_i^2 - x_i) = -1$

assume for contradiction $\alpha \in \{0,1\}^n$ satisfies F. Then

$$\left.\begin{array}{l} \bullet\ q_i(\alpha_i^2 - \alpha_i) = 0 \quad \forall i \\ \bullet\ \tilde{C}_i - 1 \geq 0 \\ \bullet\ \text{all } J_i(\alpha) \geq 0 \end{array}\right\} \text{imply } \Pi(\alpha) \geq 0. \quad \#$$

Sherali-Adams (SA)

Completeness

Idea: Partition the set of all assignments $\alpha \in \{0,1\}^n$ into m groups, $g_1, \dots, g_m$ groups such that $\alpha \in g_i$ iff $C_i(\alpha)$ is false and $C_{i'}(\alpha) = 1 \ \forall i' < i$

Let P_i be the conical junta corresponding to g_i

example: Let $g_1 = \{000, 011, 100\}$

then $P_1 = (1-x_1)(1-x_2)(1-x_3) + (1-x_1)x_2x_3 + x_1(1-x_2)(1-x_3)$

SA refutation: $\sum_{i=1}^{m} P_i (\tilde{C}_i - 1)$

Claim: $\sum_{i=1}^{m} P_i (\tilde{C}_i - 1) = -1$

Lower Bounds for SA via Pseudo Distributions

Defn A degree-d pseudodistribution on $\{0,1\}^n$ is a family of probability distributions $\mathcal{D} = \{ \mathcal{D}_S \mid S \subseteq [n], |S| \le d \}$ satisfying:

(1) $\forall S \subseteq [n]$, $\mathcal{D}_S$ is supported on $\underbrace{\{0,1\}^S}_{\text{assignments for the vars } X_S = \{x_i \mid i \in S\}}$

(2) Marginals property: $\forall S, T \subseteq [n]$, $|S|, |T| \le d$

$$\mathcal{D}_S\big|_{S\cap T} = \mathcal{D}_T\big|_{S \cap T} = \mathcal{D}_{S\cap T}$$

Example $n=3$ $\quad S = \{1,2\}$ $\quad T = \{2,3\}$

$\mathcal{D}_S$:

$x_1\, x_2$	
0 0	.2
0 1	.1
1 0	.1
1 1	.6

$\mathcal{D}_T$:

$x_2\, x_3$	
00	.3
01	0
10	.5
11	.2

$\mathcal{D}_{S\cap T}$:

x_2	
0	.3
1	.7

$\mathcal{D}_S\big|_{S\cap T}$:

x_2	
0	.3
1	.7

$\mathcal{D}_T\big|_{S\cap T}$:

x_2	
0	.3
1	.7

Lower Bounds for SA via Pseudo Distributions

Defn A degree-d pseudoexpectation operator $\widetilde{\mathbb{E}}$ is a functional that maps all degree $\leq d$ multilinear polys over $x_1 .. x_n$ to $\mathbb{R}$, + satisfies:

(1) $\widetilde{\mathbb{E}}[1] = 1$

(2) $\widetilde{\mathbb{E}}$ is Linear: if $\deg(p), \deg(q) \leq d$, $\alpha, \beta \in \mathbb{R}$ then

$$\widetilde{\mathbb{E}}[\alpha p + \beta q] = \alpha \widetilde{\mathbb{E}}[p] + \beta \widetilde{\mathbb{E}}[q]$$

(3) for every conjunction D of width $\leq d$, $\widetilde{\mathbb{E}}[D] \geq 0$

Lemma (Degree-d pseudoexpectations take expectations over a degree-d pseudodistrib)

(1) Let $\mathcal{D}$ be a degree-d pseudo-distribution then the functional $\widetilde{\mathbb{E}}$ defined by $\widetilde{\mathbb{E}}[\sum \alpha_S \prod_{i \in S} x_i] := \sum_S \alpha_S \mathop{\mathbb{E}}_{x \sim \mathcal{D}} [\prod_{i \in S} x_i]$ is a degree-d pseudoexpectation

(2) If $\widetilde{\mathbb{E}}$ is a degree-d pseudoexpectation. Then the distributions $\mathcal{D}_S$ defined as

$$\mathcal{D}_S(y) := \Pr_{x \sim \mathcal{D}_S}[x = y] = \widetilde{\mathbb{E}}\Big[\prod_{i : y_i = 1} x_i \prod_{j : y_j = 0} (1 - x_j)\Big]$$

is a degree-d pseudodistribution

$y \in \{0,1\}^S$

Lower Bounds for SA via Pseudo Distributions

Defn Let $F = C_1 \wedge \dots \wedge C_m$ be an unsat CNF. A degree-d pseudoexpectation $\tilde{\mathbb{E}}$ is a pseudoexpectation for F if it is a degree-d pseudoexpectation and $\forall$ conjunction D of degree $\leq d-1$, and for every $C_i \in F$, $\tilde{\mathbb{E}}[D(\tilde{C}_i - 1)] \geq 0$

Theorem (SA Duality) Let $F = C_1 \wedge \dots \wedge C_m$ be UNSAT CNF. Then F has no degree d SA refutation iff there exists a degree-d pseudoexpectation for F.

Lower Bounds for SA via Pseudo Distributions

Theorem (SA Duality) Let $F = C_1 \wedge \dots \wedge C_m$ be UNSAT CNF.
Then F has no degree d SA refutation iff there exists a degree-d pseudoexpectation for F.

Proof (one direction: degree-d SA refutation $\Rightarrow \nexists$ a degree-d pseudoexp.)

Assume F has a degree-d SA refutation: $J_0 + \sum_{j=1}^{m} J_i (\tilde{C}_i - 1) = -1 \quad (*)$

and assume $\tilde{\mathbb{E}}$ is a degree-d pseudo-expectation for F.

Apply $\tilde{\mathbb{E}}$ to both sides of $(*)$:

$$\tilde{\mathbb{E}}[-1] = -1 = \sum_{i=1}^{m} \tilde{\mathbb{E}}[J_i(\tilde{C}_i - 1)] + \tilde{\mathbb{E}}[J_0] \geq 0$$

which is a contradiction.

For other direction, see [Fleming, Kothari, Pitassi] Book.

Sherali Adams (SA) Some (equivalent) views

- As a proof system
- Pseudo distributions
- LP tightening

SA as LP tightening (degree d)

- Add new variables to represent all degree $\leq d$ terms
- This "lifts" polytope from n dimensions to $n^{O(d)}$ dimensions.
- Projection back to $x_1 \ldots x_n$ preserves all 0/1 solutions (+ removes some fractional ones)

Background: Let's consider Linear Programming as a Proof System

Sound, complete proof system for linear inequalities over $\mathbb{R}$

LP: $\max \; c^T x$

s.t. $\left. Ax \leq b \right\}$ (*) Linear constraints

Decision version: Is there a value of x satisfying (*) ?

Farkas' Lemma (Soundness + completeness of Decision version)

A set $\{Ax - b \geq 0\}$ of linear inequalities is unsat over $\mathbb{R}$ iff

$\exists y \geq 0$ s.t. $y^T A = 0$, $y^T b = -1$

LP Duality

(P) Primal

$$\max \; c^T x \quad \text{s.t.} \quad Ax \le b, \quad x \ge 0$$

(D) Dual:

$$\min \; b^T y \quad \text{s.t.} \quad A^T y \ge c, \quad y \ge 0$$

Duality Theorem (Implied by Farkas' Lemma)

Exactly one of the following holds:

(i) Neither (P) nor (D) have a feasible solution

(ii) (P) has solns with arbitrarily large values & (D) is unsat

(iii) (P) unsat & (D) has arb. large solutions

(iv) Both (P) & (D) have optimal solns, x^* & y^*. Then $c^T x^* = b^T y^*$

So there is a solution to dual that witnesses tight bound

So an LP "refutation" of $\{Ax \leq b,\ x \geq 0\}$ is a nonneg linear combination of these inequalities that equals -1

An LP "derivation" of $\{Ax \leq b,\ x \geq 0\} \Longrightarrow c^T x \leq c_0$ is a nonneg y^* s.t. $(y^*)^T b = c_0$

[since $c^T x \leq (y^*)^T A x \leq (y^*)^T b = c_0$]

Soundness / completeness: Farkas' Lemma / Duality Thm

Polynomial Automatizability of LP

Satisfiability of linear inequalities over $\mathbb{R}$:

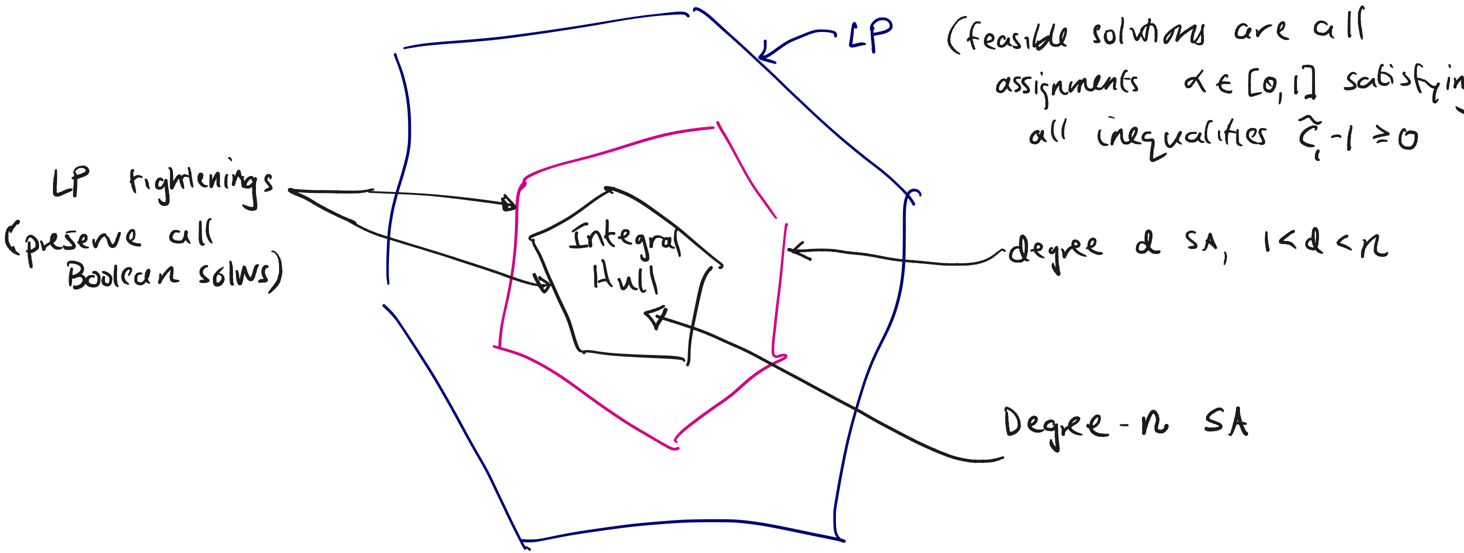
- in NP
- By Farkas' Lemma, in coNP

$\therefore$ in $NP \cap coNP$

<u>Ellipsoid Algorithm</u> [Khachiyan '79] $\quad LP \in P$

SA: A proof system for integer programming obtained by successive LP tightenings

Consider unsat F -- so we want to determine if there is an integer solution where all variables $x_i \in \{0,1\}$



SA degree d tightening

Original LP: [ignore max $c^T x$]

$$\underset{1\geq 0}{Ax \geq b} \quad , 0 \leq x \leq 1$$

add new variables Y_S $\forall S \subseteq [n], |S| \leq d$

Impose constraints $\underbrace{\prod_{i\in S} x_i \cdot \prod_{i \in T} (1-x_i)}_{\substack{S\cap T = \phi \quad |S\cup T| \leq d \\ \text{"junta"}}} \cdot (a^T x - b) \geq 0$ $\forall$ rows $a \in A$

New constraints:

lifted SA_d constraints (*)

$$Y_\phi = 1$$

$$Y_{\{i\}} = X_i$$

$$0 \leq Y_S \leq 1$$

$$\underbrace{\sum_{T' \subseteq T} (-1)^{|T'|} \left(\sum_{i=1}^{n} a_i Y_{S\cup T' \cup \{i\}} - b Y_{S \cup T'} \right) \geq 0}_{} \quad \forall \text{ rows } a \in A$$

above constraints translated to linear inequalities using new Y vars plus multilinearization $(x_i^2 - x_i = 0)$

Ex $\underbrace{(x_1 \vee \bar{x}_2 \vee x_3)}_{C_1 = x_1 + \cancel{1} - x_2 + x_3 \geq 1} \underbrace{(x_2 \vee x_4 \vee x_5)}_{x_2 + x_4 + x_5 \geq 1} \quad , \quad - \quad - \quad - (\qquad)$

assignment $(\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$ satisfies F

So does $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$

But if we multiply $(C_1) \cdot D$

$$\underbrace{x_4 x_5 x_6}_{D} \left[x_1 + (1 - x_2) + x_3 - 1 \right]$$

since we have $1 \geq 0$ as an initial constraint this gives $\forall S,T$ $|S \cup T| \leq d$, $S \cap T = \phi$:

$$\prod_{i \in S} x_i \prod_{j \in T} (1 - x_j) \geq 0$$

translates to: $\displaystyle\sum_{T' \subseteq T} (-1)^{|T'|} \left(\quad Y_{S \cup T'} \quad \right) \geq 0$

Example: $S = 1,2$ $T = 3,4,5$ $x_1 x_2 (1-x_3)(1-x_4)(1-x_5) \geq 0$

Multiplying out: $x_1 x_2 - x_1 x_2 x_3 - x_1 x_2 x_4 - x_1 x_2 x_5 + x_1 x_2 x_3 x_4 + x_1 x_2 x_3 x_5 + x_1 x_2 x_4 x_5 - x_1 x_2 x_3 x_4 x_5$

$$Y_{12} - Y_{123} - Y_{124} - Y_{125} + Y_{1234} + Y_{1235} + Y_{1245} - Y_{12345} \geq 0$$

Lemma Let $\{Ax - b \geq 0, 1 \geq 0, x \geq 0\} = \mathcal{H}$ Then the degree-d SA LP has NO feasible solution iff there is a degree-d SA refutation of $\mathcal{H}$

Pf

(1) degree-d SA LP has a feas. soln $\Rightarrow$ there is NO degree-d SA refutation

plug in feas. soln into alleged SA ref.

The LHS will evaluate to something ≥ 0

(2) degree-d SA LP has no feas soln $\Rightarrow$ $\exists$ a degree-d SA refutation

By Farkas Lemma, $\exists$ non-neg linear combination of inequalities (*) summing to -1

They convert to a degree-d SA refutation (use $x_i^2 - x_i$ to multilinearize)

Degree Automatizability of SA

Theorem degree d SA LP has $n^{O(d)}$ linear constraints, so solvable in time poly(d)

Equivalently degree-d SA refutations can be found in time $n^{O(d)}$

Q: What about size automatizability?

UB: size $s \Rightarrow 2^{\sqrt{n \log s}}$ for $s = \mathrm{poly}(n)$ this is expl time!

similar to automatizability of Resolution

Sum-of-Squares (SOS)

SOS generalizes SA - instead of $J_0 + \sum_{j=1}^{m} J_i(\tilde{\hat{C}}_i - 1) = -1$ (*)

where the J_i's are conical juntas, we want to allow J_i's to be arbitrary (but verifiable) polynomials that are always non-negative.

Defn A polynomial q over $x_1 .. x_n$ is a sum-of-squares if $q = \sum_i p_i^2$ for some polynomials p_i

Sum-of-squares (SOS)

Defn Let $F = C_1 \wedge \dots \wedge C_m$ be UNSAT CNF.

A sum-of-squares (SOS) refutation of F is a set of sum-of-squares polynomials $\{q_0, q_1, \dots, q_m\}$ such that

$$\sum_{i=1}^{m} q_i (\tilde{C}_i - 1) + q_0 = -1$$

where we assume multilinear arithmetic

Soundness: Similar to proof of soundness for SA

Completeness: Follows from completeness of SA since any conjunction (+ thus any conical junta) can be written as a sum-of-squares:

Let $D = \prod_{i \in S} x_i \prod_{j \in T} (1 - x_j)$

Then $D^2 = \left(\prod_{i \in S} x_i \prod_{j \in T} (1 - x_j) \right)^2 = \prod_{i \in S} x_i^2 \prod_{j \in T} (1 - x_j)^2 = \prod_{i \in S} x_i \prod_{j \in T} (1 - x_j) = D$

SUM-OF-SQUARES (SOS) Some (equivalent) views:

[See Fleming-Kothari-P Book for full treatment + examples]

- As a proof system
- Pseudo distributions
- SDP tightening

- Like SA, we can define suitable notion of pseudodistribution and pseudo-expectation so that $\exists$ degree-d SOS refutation of F iff $\exists$ a degree-d SOS-pseudoexpectation for F.
  This gives us a complete method for proving SOS degree lower bounds
- SOS can be viewed as a tightening of SDP (semidef. program)
- Efficient alg for SDP $\Rightarrow$ degree-d SOS refutations are automatizable* in time $n^{O(d)}$

* ignoring coefficient size

SOS is stronger than SA

Theorem There are polysize, degree-3 SOS refutations of $BPHP_n$, but SA refutations require $\Omega(n)$ degree

$BPHP_n$ (binary PHP). Let $n = 2^d$

Variables : $x_{i,j}$, $i \in [0,..,n-1]$, $j \in [0,..,d-1]$ $x_{i,j}$ = j^{th} bit of binary representation of where pigeon i is mapped

We write $\langle x_i \rangle = j$ to denote the width-d conjunction which is true iff pigeon i maps to hole j

Clauses:

(1) No pigeon maps to hole 0 : $\bigvee_{j=1}^{d} x_{i,j}$

(2) No 2 pigeons map to same hole : $\neg(\langle x_i \rangle = j) \vee \neg(\langle x_{i'} = j \rangle)$ $\forall i \neq i' \in [0..n-1]$

SA Lower bound for $BPHP_n$

<u>Thm</u> Degree of any SA refutation of $BPHP_n$ is $\Omega(n)$.

<u>Proof sketch</u>: Let $d = \frac{n}{4}$. Let $S \subseteq [n+1]$ be a subset of $\leq d$ pigeons

$\mathcal{D}_S$: choose random set H of $|S|$ holes from $[n]$, & a random bijection from $S \to H$

given subset $\mathcal{U}$ of variables from $BPHP_n$, let S_u = pigeons mentioned by u

$\mathcal{D}_u$: distrib over $\{0,1\}^u$ obtained by picking random matching from $\mathcal{D}_{S_u}$ and set variables indexed by u accordingly

- marginals property : given 2 subsets S, T of pigeons, $\mathcal{D}_S|_{S\cap T} = \mathcal{D}_T|_{S\cap T} = \mathcal{D}_{S\cap T}$
- Let $\tilde{\mathbb{E}}$ be associated pseudo-expection.

Need to show $\tilde{\mathbb{E}}[D(\tilde{C}_i - 1)] \geq 0$ for all clauses C_i, and all conjunctions D, $|D| \leq d - |C_i|$:

recall $\tilde{C}_i(\alpha) \geq 1$ if α satisfies C_i, and is 0 otherwise

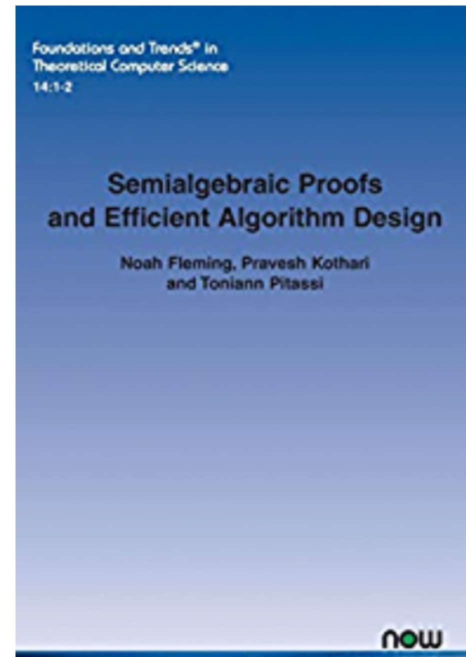
since all α in support satisfy C_i, $(\tilde{C}_i - 1)(\alpha) \geq 0$ and $D(\alpha) \geq 0$.

$\therefore \tilde{\mathbb{E}}[D \cdot (\tilde{C}_i - 1)] \geq 0$

- A flurry of degree Lower Bounds for Nullstellensatz, Poly Calculus, SA, SOS

- SOS mania:

  SOS upper bounds $\Rightarrow$ learning algs (via automatizability)



THE AMAZING USEFULNESS OF SOS: UPPER BOUNDS

UPPER BOUNDS CAN AUTOMATICALLY GENERATE EFFICIENT ALGS!

- PC/SA/SOS are automatizable: degree d proofs can be found in time $n^{O(d)}$

$\therefore$ Low degree proofs certifying the mere existence of a solution automatically give ptime algorithms

- Dictionary Learning [BKS'15]
- Tensor completion [BM16, PS17]
- Tensor decomposition [MSS16]
- Robust moment estimation [KS17]
- Clustering [HL18] [KS17]
- Robust linear regression [KKM18]

THE AMAZING USEFULNESS OF SOS: LOWER BOUNDS

LOWER BOUNDS IMPLY LOWER BOUNDS FOR A BROAD CLASS OF ALGORITHMS

[LRS'15, CLRS'16] LP/SDP EXTENSION COMPLEXITY OF Δ $\approx$ SA/SOS DEGREE OF P_Δ

[RPRC'16, PR'18] MONOTONE FORMULA SIZE/SPAN PROGRAM SIZE $\approx$ NSATZ DEGREE

[GGKS'18] MONOTONE CIRCUIT SIZE $\approx$ PC DEGREE

Lower Bounds [Buss - Grigoriev - Impagliazzo - Pitassi], [Grigoriev], [Schoenebeck]

See [Fleming - Kothari - Pitassi] for simplified proof

There exist unsat 3CNFs (Tseitin over degree-3 expander graphs) that require $\Omega(n)$ degree refutations in SOS (& therefore also $\Omega(n)$ degree Nsatz, PC, and SA refutations)

Reference Fleming - Kothari - Pitassi
Semialgebraic Proofs and Efficient Algorithm Design