Lecture 8: Constrained SoS & Reductions

Last time: Grigoriev's thm
$\Omega(n)$-degree lower bound on
val_I for k-XOR

No lecture on Nov 3

informally: SoS requires $\deg \geq cn$ (c const.) to "realize" that m-constraint, n variable random instance I of k-XOR is unsatisfiable.

today: general framework of reductions for lower bounds (ex. MIS)

building a theory of constrained SoS proofs / pseudo-dist^{n}s

HW: more examples of using reductions.

Constrained SoS proofs & pseudo-dist^{n}s

$f_1, f_2, \ldots, f_m$ be polynomials in $\underbrace{\mathbb{R}[x]}_{x=(x_1,\ldots,x_n)}$

"semialgebraic sets"

$\mathcal{A} = \{ f_1 \geq 0, f_2 \geq 0, \ldots, f_m \geq 0, h_1 = 0 \}$

$h_2 = 0$
$\vdots$
$h_m = 0$

"System of poly inequalities"!

<u>Goal</u>: to solve two natural problems about $\mathcal{A}$

① given $g \in \mathbb{R}[X]$, decide if g is non-neg over $\mathcal{A}$.

think of $\mathcal{A}$ as a subset of $\mathbb{R}^n$: set of $x \in \mathbb{R}^n$ s.t. $f_i(x) \geq 0 \quad \forall\, 1 \leq i \leq m$.

① $\Leftrightarrow$ is $g(x) \geq 0 \quad \forall x : \{f_i(x) \geq 0\}_{i=1}^m$

<u>Ex</u>.: $f_i(x) = x_i^2 - 1$

$$\mathcal{A} = \{ f_i(x) = 0 \}_{i=1}^n \equiv \{\pm 1\}^n$$

② decide if $\mathcal{A}$ is empty.

i.e. does $\exists\, x \in \mathbb{R}^n$ s.t. $\{f_i(x) \geq 0\}_{i=1}^m$

Why is ② more general than ① ?

$$\mathcal{A} : \{ f_i \geq 0 ,\ g < 0 \}$$

Sum-of-Squares Proofs

Def: "$\exists$ a degree ℓ SoS proof that $\mathcal{A}$ implies $\{g \succcurlyeq 0\}$" if $\exists$ SoS polynomials $(p_S)_{S \subseteq [m]}$ s.t.

multiset

arbitrary polynomials

$$g = \sum_{S \subseteq [m]} p_S \cdot \prod_{i \in S} f_i + \sum_{S \subseteq [m]} p'_S \prod_{i \in S} h_i$$

& every summand $p_S \cdot \prod_{i \in S} f_i$ has deg $\le \ell$.

Notation: $\mathcal{A} \; \underbrace{\vdash_{x}^{\ell}}_{\text{"derives"}} \{g \succcurlyeq 0\}$ "$\mathcal{A}$ derives $g \succcurlyeq 0$ in deg ℓ"

degree of the proof

"variable" of the proof

Suppress when clear from context

Ex: SoS proofs over hypercube

$$\underbrace{g = \Sigma_i f_i^2}_{\text{on the hypercube}} \quad \text{for deg} \le \tfrac{\ell}{2} \text{ poly } f_i$$

$\Updownarrow$

Same coefficients after multilinear

same coeff...
reduction.

$$\mathcal{A} = \{(x_i^2 - 1)\}_{i=1}^{n}$$

arbitrary polynomials

$$\boxed{g = \sum_{S \subseteq [n]} p_S \cdot \prod_{i \in S} (x_i^2 - 1) + \sum h_i^2}$$

Fact: If $g = h$ ($\deg \ell$) over hypercube

then $g = h + \underbrace{\sum_{S \subseteq [n]} P_S \prod_{i \in S} (1 - x_i^2)}$ $\deg \leq \ell$

over the reals.

Pf sketch: $x_1^3 = x_1 - x_1 + x_1^3$

$$= \underline{x_1} + \underbrace{x_1 (x_1^2 - 1)}$$

$\vdots$

Can repeat for larger deg monomials

Theorem: [Krivine'64, Stengle'74]

$\mathcal{A} : \{f_1 \ldots f_m \geq 0\}$

then either there's a solution to $\mathcal{A}$

Then ... ,

or $\exists$ a SoS proof that $\mathcal{A} \underset{\text{refutation}}{\underbrace{\vdash^{\ell} \{-1 \geq 0\}}}$

for some $\ell \in \mathbb{N}$.

"Positivstellensatz"

highly non-constructive proof, no bounds on ℓ

Calculus of SoS proofs

$\mathcal{A}, \mathcal{B}, \mathcal{C}$ systems..

Addition rule:

$$\frac{\mathcal{A} \vdash^{\ell} \{f \geq 0,\ g \geq 0\}}{\mathcal{A} \vdash^{\ell} \{f+g \geq 0\}}$$

Mult rule:

$$\frac{\mathcal{A} \vdash^{\ell} \{f \geq 0\},\ \mathcal{A} \vdash^{\ell'} \{g \geq 0\}}{\mathcal{A} \vdash^{\ell+\ell'} \{f \cdot g \geq 0\}}$$

transitivity rule:

$$\frac{\mathcal{A} \vdash^{\ell} \mathcal{B},\ \mathcal{B} \vdash^{\ell'} \mathcal{C}}{\mathcal{A} \vdash^{\ell \cdot \ell'} \mathcal{C}}$$

Substitution:

$$\frac{\{F \geq 0\} \vdash^{\ell} \{G \geq 0\}}{\{F(H) \geq 0\} \vdash_{\ell \cdot \deg(H)} \{G(H) \geq 0\}}$$

F: m-polynomials

G: p-~~po~~polynomials

$H: \mathbb{R}^n \to \mathbb{R}^p$

Pseudo-distributions

$\mu: \mathbb{R}^n \to \mathbb{R}$, finitely supported

$\equiv$ μ is 0 on all but a finite set of points

$$\tilde{\mathbb{E}}_\mu f = \sum_{x:\, \mu(x) \neq 0} \mu(x) \cdot f(x)$$

"formal expectation"

Def (Pseudo-dist^n):

A finitely supp $\mu: \mathbb{R}^n \to \mathbb{R}$ is a deg d p.d. if ① $\tilde{\mathbb{E}}_\mu 1 = 1$ ② $\tilde{\mathbb{E}}_\mu f^2 \geq 0$

$\forall f$ of $\deg \leq \frac{d}{2}$.

$\forall f$ of $\deg \le \frac{d}{2}$

If $\mathrm{Supp}(\mu) \subseteq \Omega \subseteq \mathbb{R}^n$, then μ is a "a deg d p.d. over Ω"

<u>Lemma:</u> $\mu : \mathbb{R}^n \to \mathbb{R}$, finitely supported.

$$\tilde{\mathbb{E}}_{\mu} 1 = 1. = \sum_{x:\, \mu(x) \neq 0} \mu(x) = 1.$$

<u>Then</u>, μ is a deg d p.d. $\Leftrightarrow$

$$\underbrace{\tilde{\mathbb{E}} \left((1,x)^{\otimes d/2} \right) \cdot \left((1,x)^{\otimes d/2} \right)^T}_{\text{"Pseudo-moment matrix of deg } d.\text{"}} \succcurlyeq 0$$

<u>Satisfying Constraints</u>

$\mu : \mathbb{R}^n \to \mathbb{R}$, finitely supp, p.d. of <u>deg</u> d.

$\mathcal{A}$: System of poly inequalities

extra parameter

μ satisfies $\mathcal{A}$ at $\boxed{\deg \ell}$

if $\forall\, S \subseteq [m]$, every sos poly h on $\mathbb{R}^n$

s.t. $\deg(h) + \sum_{i\in S} \max\{\deg(f_i), \ell\} \leq d$.

it holds that $\tilde{\mathbb{E}}_\mu\, h \cdot \prod_{i \in S} f_i \geq 0$.

__Ex.__ μ of deg d satisfies $\{f = 0\}$ at deg $d - \deg(f)$ if $\forall$ polynomials h, $\tilde{\mathbb{E}}_\mu\, h \cdot f = 0$ whenever

$$\deg(h) + \deg(f) \leq d$$

Notice if μ was a distⁿ supported on $\{f=0\}$

$$\mathbb{E}_\mu\, h \cdot f = 0$$

μ a distⁿ supp on $\underline{\{f \geq 0\}}$: h: non-neg

$$\mathbb{E}_\mu\, h \cdot f \geq 0$$

Notation

$\mu \overset{\ell}{\models} \mathcal{A}$

"is a model of"

__Corollary:__ If $\underline{\mathcal{A} \vdash^{\ell} \{g \geq 0\}}$

& $\mu \overset{d}{\models} \mathcal{A}$ then $\tilde{\mathbb{E}}_\mu\, g \geq 0$.

__Pf sketch__

$\mathcal{A} \vdash^{\ell} \{g \geq 0\}$

$$g = \sum_S p_S \cdot \underbrace{\prod_{i \in S} f_i}_{} \to \leq \ell$$

<u>$\mu \models^{\ell} \mathcal{A}$</u> : $\tilde{\mathbb{E}}_{\mu} \cdot p_S \cdot \underbrace{\prod_{i \in S} f_i}_{\ell} \geq 0$: $\deg(p_S) + \max\{\deg f_i, \ell\}$

$$\Rightarrow \tilde{\mathbb{E}}_{\mu} \cdot g \geq 0.$$

<u>Corollary</u>: $\mathcal{A} \vdash^{\ell} \{g \geq 0\}$

$\mu \models^{0} \mathcal{A}$, deg d

then if $\ell \leq d$, $\tilde{\mathbb{E}}_{\mu} \cdot g \geq 0$.

<u>Pf</u>: $\mathcal{A} \vdash^{\ell} \{g \geq 0\}$

$$g = \sum_S p_S \cdot \underbrace{\prod_{i \in S} f_i}_{\leq \text{ total deg} \leq \ell.}$$

μ is a deg d p.d. & $\mu \models^{0} \mathcal{A}$

$\Rightarrow \forall h, S$ s.t

$$\deg\left(h \cdot \prod_{i \in S} f_i\right) \leq d$$

$$\tilde{\mathbb{E}}_\mu \, h \cdot \prod_{i \in S} f_i \geq 0$$

$\Rightarrow \quad \tilde{\mathbb{E}}_\mu \, g \geq 0.$

i.e. "whatever poly g can be proven non-neg via low deg get non-negative non-neg. p. expectations if μ sat $\models \mathcal{A}$"

Duality

"explicitly bounded semialgebraic sets"

Theorem: $\mathcal{A}$, $\underline{\sum_{i=1}^{n} x_i^2 \leq M}$ is included in $\mathcal{A}$

Then $\forall d \in \mathbb{N}$, $\forall f$ poly of deg $\leq d$,

exactly one of following holds:

① $\underline{\forall \varepsilon > 0}$, $\mathcal{A} \vdash_{d} \{ f \geq -\varepsilon \}$

② $\exists$ a μ of deg d, $\mu \models^{0} \mathcal{A}$,

$$\tilde{\mathbb{E}}_\mu f \leq 0.$$

$\boxed{\begin{array}{l}\text{\underline{Analog for dist}}^n\text{: either } \forall x \text{ that satisfy } \mathcal{A}, \ f(x) \geq 0 \\ \text{or } \exists \mu \text{ on } x \text{ satisfying } \mathcal{A} \text{ s.t. } \mathbb{E}_\mu f < 0.\end{array}}$

<u>Algorithms</u>

<u>Thm</u>: Given $d \in \mathbb{N}$, $\mathcal{A}$: poly system, rational coeffs over $\mathbb{R}^n$ explicitly bounded,

then $\exists$ $\underbrace{N}_{\text{size of input}}^{O(d)}$ time algo to output a

deg d p.d. μ : $\underline{\mu \models \mathcal{A}}$ up to an additive

error of 2^{-N}:

error of ...

$\forall$ SoS poly h,

$$\tilde{\mathbb{E}}_{\mu} \cdot h \cdot \prod_{i \in S} f_i \geq -2^{-N} \cdot \left(\begin{array}{l}\text{Sum of}\\ \ell_1 \text{norms}\\ \text{of coeffs}\\ \text{of } h, f_i\end{array}\right)$$

3-XOR

$x_{C_i} = b_i \equiv b_i \cdot x_{C_i} = 1$

$\mathcal{I}$: $C_1, \ldots, C_m \rightarrow$ random

$b_1, \ldots, b_m \rightarrow$ random indep

Sat version: $\exists x$: $b_i \cdot x_{C_i} = 1 \quad \forall i \leq m$

Max-3XOR: $\mathrm{val}_{\mathcal{I}} = \frac{1}{2} + \frac{1}{2m} \underbrace{\sum_i b_i \cdot x_{C_i}}_{\mathcal{I}}$,

Last time: Cert for $\underbrace{1 - \varepsilon - \mathrm{val}_{\mathcal{I}}^{\approx \frac{1}{2}+\varepsilon}}$

today: $\mathcal{A} = \{ x_i^2 = 1 \ \forall i \} \cup \{ b_i \cdot x_{C_i} = 1 \}_{i=1}^{m}$

Notice for random $\mathcal{I}$, with $m \gg n/\varepsilon^2$,

w.p. ≥ 0.99 $\mathcal{A}$ is empty.

Krivine Stengle: for some $l \in \mathbb{N}$ large enough,

$$\mathcal{A} \vdash^{l} \{-1 \geq 0\}$$

Is there a low-deg refutation of the system $\mathcal{A}$?

Observation: Suppose there exists a p.d. μ of deg $\geq c \cdot n$. (c = const) that satisfies $\mathcal{A}$. Then $\mathcal{A} \not\vdash^{d} \{-1 \geq 0\}$.

Proof: Suppose $\mathcal{A} \vdash^{d} \{-1 \geq 0\}$

$$\underline{\deg d}: \mu \models \mathcal{A}, \quad \tilde{\mathbb{E}}_\mu(-1) \geq 0$$
$$= -\tilde{\mathbb{E}}_\mu 1 = -1 \not\geq 0.$$

"if you want to prove the absence of a low-deg refutation, exhibit a p.d. satisfying $\mathcal{A}$"

Last time: μ on $\{\pm 1\}^n$ that sat

deg $d = \Omega(n)$

$$\underline{\tilde{\mathbb{E}}_\mu \cdot \mathrm{val}_I = 1.}$$

<u>Observation:</u> The Grigoriev construction of μ in fact sat: $\mu \models \mathcal{A}$.

<u>Pf:</u> recall $\mu \models \mathcal{A} := \{ \underset{=0}{(x_i^2 - 1)} \}_{i=1}^{n} \cup \{ \underset{=0}{\overline{(b_i \cdot x_{c_i} - 1)}} \}$

<u>for all polynomials</u> h, of <u>deg $\le d-3$</u>,

$$\tilde{\mathbb{E}}_\mu \cdot \underline{h \cdot (b_i x_{c_i} - 1)} = 0.$$

<u>Let's verify.</u> enough to verify for monomials

$$h = X_S. \quad |S| \le d-3$$

S is a set.

$$\tilde{\mathbb{E}}_\mu \cdot X_S \cdot (b_i \cdot X_{c_i} - 1) = 0$$

$$\Leftrightarrow b_i \tilde{\mathbb{E}}_\mu \cdot \underbrace{X_S \cdot X_{c_i}} = \tilde{\mathbb{E}}_\mu \cdot X_S.$$

Note: $|S \oplus C_i| \le d.$

$\underline{X_{C_i}} \in \mathrm{Der}_d$

$\underline{X_S \cdot X_{C_i}}$

$\in \mathrm{Der}_d$

$\tilde{\mathbb{E}}_\mu X_{\dots} = 0.$

Suppose $\tilde{\mathbb{E}}_\mu X_S = 0$ $\in \mathrm{Der}_d$

then $\tilde{\mathbb{E}}_\mu X_S X_{C_i} = 0$. $X_S \notin \mathrm{Der}_d$

Suppose $\tilde{\mathbb{E}}_\mu X_S \neq 0$.

$$\tilde{\mathbb{E}}_\mu X_S X_{C_i} = (\tilde{\mathbb{E}} X_S)\underbrace{(\tilde{\mathbb{E}} X_{C_i})}_{= b_i}.$$

$$= b_i \tilde{\mathbb{E}} X_S$$

Thus $b_i \tilde{\mathbb{E}}_\mu X_S X_{C_i} = \tilde{\mathbb{E}} X_S$

□.

Der_d system $\equiv$ all constraints that can be derived via SoS proofs of deg $\leq d$ from $\mathcal{A}$.

"refuting random 3-XOR"

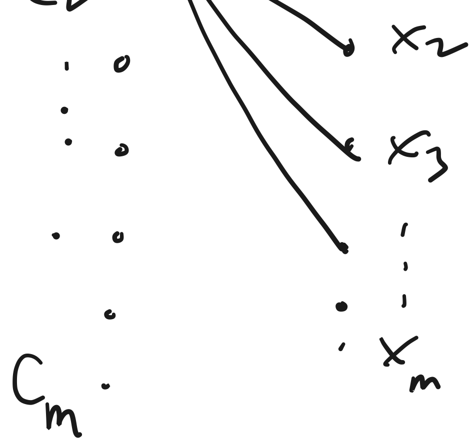
CSP(P) : $P : \{\pm 1\}^k \to \{0,1\}$ (0: "unsat", 1: "sat")

I: instance of CSP(P)

C_1 $(i,j) : l_i \in \{\pm 1\}^k$ x_1

C_2

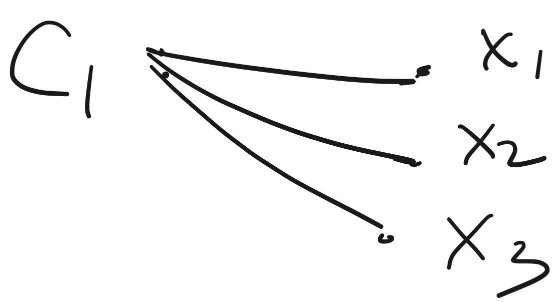
(i,j) : j^{th} variable in i^{th} constraint



→ instance specified by bip graph with labeled edges

& $l_1, l_2, \ldots, l_m \in \{\pm 1\}^k$

3-SAT: $x_1 \vee \bar{x}_2 \vee x_3$



$l_1(1) = 1$

$l_1(2) = -1$

$l_1(3) = 1$

l_1

Max-CSP

Goal: $\max_{y \in \{\pm 1\}^n} \frac{1}{m} \sum_{i=1}^{m} P(\underline{C_i(l_i \cdot y)})$

Sat: is there an x: $\{P(C_i(l_i \cdot x)) = 1\}_{i=1}^{m}$

Refuting Random CSPs (P)

"the goal is to "prove" or "certify" that a random instance of CSP(P) is unsatisfiable"

Random CSP(P) : Choose C_i at random & indep $[n]^k$

l_i at random, indep

Prop: $m \geq O_k\left(\frac{n}{\varepsilon^2}\right)$ then $opt_{\mathcal{I}} \leq \frac{|P^{-1}(1)|}{2^k} + \varepsilon$

$P^{-1}(1)$: set of all assignment sat P

Today: we asked for SoS certificates of feasibility for random 3-XOR CSP.

Def (Ref Algorithm) : Feige'01

A refutation algorithm for R-CSP(P) takes input $\mathcal{I}$. instance & outputs a value $U \in [0,1]$ (U= "upper bound")

(1) "Correctness" : $U \geq opt_{\mathcal{I}}$ for every $\mathcal{I}$.

(2) W.p. 0.99, over the choice $\mathcal{I}$ from R-CSP(P)

dist^n, $\quad u \leq 1-\delta$ for some $\delta > 0$.

we analyzed

Ref algorithm for k-XOR: output $\max_{\substack{\mu: \text{p.d.} \\ \text{of deg } d}} \tilde{\mathbb{E}}_\mu \text{val}_{\mathcal{I}}$

Algorithm never incorrectly outputs unsat

$\longrightarrow$ certifying algorithm.

Fact: there's no known poly time refutation algorithm for R-3-SAT with $m \ll n^{1.5}$ constraints.

Feige ('01) (R-3SAT Hypothesis): For some $C \gg 2^k/\varepsilon^2$, & $m = Cn$, there's no poly time refutation algo for R-3SAT with $m \sim C \cdot n$ constraints.

Hardness of approximation:

Hardness of approxim...

Conj for other CSPs :

- Learning theory hardness ~ 2013
- Cryptography (pseudorandom generators)

In general, we would like to study these conjectures..

best known algo for this problem: SoS.

Thm: If $m \gg n^{1.5}$, there's a poly time refutation algo for R-3SAT (also R-3XOR) $\rightarrow$ uses deg 4 SoS.

At $m = Cn$: R-3XOR needs deg $\Omega(n)$ SoS for refutation.

R-3SAT ... needs deg $\Omega(n)$

R-3SAT also needs ...
SoS for refutation.

You could ask.

what happens when $Cn << m << n^{1.5}$?

" " Predicate P arbitrary

we know complete answer to this question.

Def (Pairwise Uniformity)

A $dist^n$ μ on $\{\pm 1\}^k$ is pairwise uniform

if $\mathbb{E}_\mu x = 0$, $\mathbb{E}_\mu x x^T = \mathbb{I}$.

Let $\delta_2 :\overset{def}{=} \max_{\substack{\mu : \text{ pairwise } \mu \\ \text{uniform} \\ \text{over } \{\pm 1\}^k}} \mathbb{E} P(x)$

Thm: If $m = Cn$ (const), then I: R-CSP(P) with m constraints

Fc': $\max_{\mu : \text{p.d.f.}} \tilde{\mathbb{E}}_\mu \, val_{I,P} = \delta_2 \pm o(1)$

$\deg \le c'n$ on $\{\pm 1\}^n$

<u>Corollary</u>: if $\exists\, \mu$: p.u. & $\mathbb{E}_{\mu} P(x) = 1$

$\equiv$ μ is supp on sat assignments for P

<u>then</u>. for $m \ll n^{1.5}$, there's no

<u>$O(1)$-deg</u> SoS refutation for $\mathcal{A}_I$.

<u>Obs^n</u>: $P(x)=1$ iff $x_1x_2x_3=-1$

Uniform dist^n on all $x: \{x_1x_2x_3=-1\}$

$\Downarrow$

pairwise uniform.

$\Rightarrow$ $\delta_2(P) = 1$.

Same holds for 3-SAT too.

<u>In general</u>: $\frac{|P^{-1}(1)|}{2^k} \sim \mathrm{opt}_I$

$$\geq \mathop{\mathbb{E}}_{\text{Uniform}} P(x)$$

$$\delta_2 \geq \mathop{\mathbb{E}}_{\text{Uniform}} P(x)$$

Reference for algorithm: Allen, O'donnell, Witmer

$\exists$ a $O(1)$ deg cert. for

$\delta_2 + o(1)$ - $\mathrm{val}_{I,P}$.

Note: $1 > \delta_2 > \frac{|\bar{P}^{-1}(1)|}{2^k}$.

Claim: Suppose $\delta_2 = 1$

Then $\mathcal{A}_I = \{x_i^2 = 1\} \cup \{P(C_i(l_i \cdot x)) = 1\}_{i=1}^{n}$

$\exists$ a deg d refutation of $\mathcal{A}_I$

for ... $m \leq C \cdot n$

for $d = c'n$ when $m = \ldots$

<u>Claim</u>: Suppose $\delta_2 < 1$. $m = Cn$

then $\max_{\substack{\mu: \deg \leq c'n \\ \text{p.d.}}} \tilde{\mathbb{E}}_\mu \cdot \mathrm{val}_{I,P} \leq \delta_2 + o(1) < 1$

<u>implies that</u>: $\exists$ a $O(1)$ deg refutation of $\mathcal{A}_I$.

<u>Claim 3</u>: Suppose $\max_{\mu: \deg d \text{ on } \{\pm 1\}^n} \tilde{\mathbb{E}}_\mu \cdot \mathrm{val}_{I,P} < 1$.

<u>Then</u>, $\underline{\mathcal{A}_I \vdash_d \{-1 \geq 0\}}$.

<u>Pf</u>: Suppose $\exists$ μ of deg d that ~~is~~ satisfies $\mathcal{A}_I$

$$\tilde{\mathbb{E}}_\mu \cdot \mathrm{val}_{I,P} = 1$$

that's a contradiction.