SUM-OF-SQUARES LOWER BOUNDS

So far: degree 2 certificates & Max-Cut, Q.P.

deg 4 certificates & Expansion

higher degree certificates & application

degree 2 certificate analysis is tight for odd-cycle

but odd cycle is not a hard example for deg 4 SoS.

How do we prove strong, convincing lower bounds on the sum-of-squares method? → Today's Goal.

(Average-Case Problems)

Running Example: k-LIN over $\mathbb{F}_2$

$\equiv$ k-XOR

K-XOR Problem

k-sparse linear equations over $\mathbb{F}_2$.

$$z_1 \oplus z_2 \oplus \cdots \oplus z_k = +1 \pmod 2$$
$$= 0$$

an instance: Collection of k-sparse lin equations

Goal: satisfy max possible # of constraints

Goal: notational switch to work over the reals

$z_1 \longrightarrow (1-2z_1) \rightarrow x_1$ (over the reals)

$0 \longrightarrow +1$

$1 \longrightarrow -1$

$$\{b = \underbrace{z_1 \oplus \cdots \oplus z_k}_{=+1}\} \equiv \underbrace{\left(\prod_{i=1}^{k} x_i\right)}_{=-1}(1-2b) = 1$$

$b \in \{0, 1\}$

when odd number of z_is are 1

iff odd # of x_is are -1.

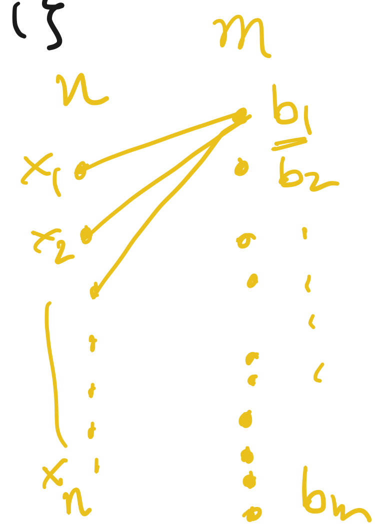
$$\boxed{\text{For the rest of the class, } b \in \{\pm 1\}.}$$

k-XOR / k-LIN problem

input: $C_1, \ldots, C_m$: k-tuples of variables $\subseteq [n]$ of size k

$b_1, \ldots, b_m \in \{\pm 1\}$

$$\mathcal{I} : \{C_i, b_i\}_{i=1}^{m} : \; b_i \prod_{j \in C_i} x_j = 1$$



m linear equations each being k-sparse

Notation:

$$\mathcal{I}(x) = \frac{1}{m} \sum_{i=1}^{m} b_i \cdot \prod_{j \in C_i} x_j$$

$\leftarrow$ deg k polynomial over $\mathbb{R}$

$\mathcal{I}(x)$
$x \in \{\pm 1\}^n$: $\mathcal{I}(x) = 1$ if x SAT all constraints

$\mathcal{I}(x) = -1$ if x violates all constraints.

frac of constraints SAT by x $= \underline{\mathrm{Val}_{\mathcal{I}}(x)} = \underline{\frac{1}{2} + \frac{\mathcal{I}(x)}{2}}$.

$\mathrm{opt}(\mathcal{I})$: $\max_{x \in \{\pm 1\}^n} \mathrm{val}_{\mathcal{I}}(x)$

<u>K-XOR</u>

<u>Goal</u>: Given $\mathcal{I}$, approximate $\mathrm{opt}(\mathcal{I})$.

<u>Algorithms for approx. $\mathrm{opt}(\mathcal{I})$</u>

Suppose $\mathrm{opt}(\mathcal{I}) = 1$.
Then there's a polynomial time algo for finding a SAT assignment.

<u>Algorithm</u>: Simply return a random assignment.

$$\frac{1}{2} + \frac{1}{2} \mathbb{E}_{y \in \{\pm 1\}^n} \mathcal{I}(y) = \text{expected frac of constraints satisfied.}$$

$$\mathbb{E}\, I(y) = \frac{1}{m} \sum_{i=1}^{m} \mathbb{E}\Big[b_i \cdot \prod_{j \in C_i} y_j\Big]$$

$$= \frac{1}{m} \sum_{i=1}^{m} b_i \cdot \mathbb{E}\Big[\prod_{j \in C_i} y_j\Big]$$

$$= 0.$$

k=2: $C_i = \{i_1, i_2, \ldots, i_k\}$

$$\underbrace{x_{i_1} \cdot x_{i_2}} = -1 \qquad b_i = -1$$

For max-cut, we were able to beat $\frac{1}{2}$ by a whole lot! 0.878 approx.

$\forall k \geq 3$

Håstad '99: $\forall \varepsilon > 0$, it is NP-hard to decide if k-XOR instance is

1. $\geq 1-\varepsilon$ satisfiable ($I(x) \geq 1 - 2\varepsilon$ for some x)
2. $\leq \frac{1}{2} + \varepsilon$ sat ($\forall x$: $I(x) < 2\varepsilon$)

Corollary: If $P \neq NP$ then $\nexists$ poly time ... finding 0.51 sat assignment to algo for finding

a 0.99999 satisfiable instance - (for $k \geq 3$)

"Approx. Resistance": random assignment is essentially optimal for k-XOR.

"maximally hard":

Today: Unconditionally verify for SoS.

Theorem: (Grigoriev '01, Schoenebeck '08)

Let $c < 2$. Then there's a constant $c'(k, \varepsilon)$ & a family of k-XOR instances: $\{I_n\}_{n \geq 1}$

s.t. $c \cdot opt(I_n) - Val_I(x)$ doesn't have a degree $c'n$ certificate $\forall$ large enough n.

Approx. $c \cdot opt(I) - val_I(x) \geq 0$

cannot beat random assignment using $c'n$ degree SoS.

What family?

$\mathcal{I}_n$: random instance $\approx \Delta n$ Constraints.

Sample constraints in $\mathcal{I}_n$ as follows:

1) pick every k-tuple from $[n]$ independently to be in $\mathcal{I}_n$ w.p. $\frac{\Delta n}{\binom{n}{k}}$. $\to$ $m = \#$ of k-tuples.

(Note that in expectation $\sim \Delta n$ constraints)

2) For every k-tuple C_i, choose $b_i \in \{\pm 1\}$ uniformly & independently at random.

Output $\{(C_i, b_i)\}_{i=1}^{m}$

Next goal: to prove that there's no $c'n$ deg certificate for $c \cdot opt(\mathcal{I}_n) - Val_{\mathcal{I}_n}(x)$.

Duality: equivalent to proving that there's a pseudo-distn of degree $c'n$ s.t.

$$\tilde{\mathbb{E}}_\mu \left(c \cdot \mathrm{opt}(I_n) - \mathrm{val}_{I_n}(x) \right) < 0.$$

$$\equiv \quad \tilde{\mathbb{E}}_\mu \cdot \mathrm{val}_{I_n}(x) \geq \underset{\underset{(2-\delta)\cdot \underbrace{\mathrm{opt}(I_n)}}{\parallel}}{c \cdot \mathrm{opt}(I_n)}$$

Proposition (value of opt(I_n))

For some constant D, whenever $\Delta \geq D/\varepsilon^2$.

$\mathrm{opt}(I_n) \leq \frac{1}{2} + \varepsilon$ w.p. ≥ 0.99.

Proof: fix any $y \in \{\pm 1\}^n$. $\{C_i, b_i\}_{i=1}^m$

Consider the r.v. $\underbrace{b_i}\cdot \prod_{j \in C_i} y_j$.

$$I(y) = \frac{1}{m} \sum_{i=1}^m \underbrace{b_i \cdot \prod_{j \in C_i} y_j}$$

$$\mathbb{E}[I(y)] = 0.$$

$$\mathop{\mathbb{E}}_{b} I(y)^2 = \frac{1}{m}.$$

Chernoff-Hoeffding. $\Pr\left[I(y) \geq \frac{t}{\sqrt{m}} \right]$ ($\frac{t}{\sqrt{m}} = \varepsilon$)

$\leq 2 \cdot e^{-\Omega(t^2)}$

$t = \varepsilon\sqrt{m}$. $\sim 2^{-\Omega(\varepsilon^2 \cdot m)}$.

Union bound over 2^n assignments:

$$\Pr[\exists y : I(y) \geq \varepsilon] \leq \underbrace{2^n}\cdot 2 \cdot \underbrace{2^{-\Omega(\varepsilon^2 m)}}$$

If $m \gg n/\varepsilon^2$ then this prob $\ll 0.01$.

Thus. w.p. 0.99 over the draw of I_n,

$\mathrm{Opt}(I_n) \leq 1/2 + \varepsilon$. $\square$

Theorem (Grigoriev '01)

of k-XOR

Let $\{I_n\}_{n\geq 1}$ be random instance with Δn constraints in expectation. Then for some $c' = c'(k, \Delta)$, w.p. 0.99 over I_n, there's a degree $c'n$ p.d. μ s.t. $\underset{\mu}{\tilde{\mathbb{E}}}\, \mathrm{val}_I = 1.$

__Note__: this can be interpreted as saying that "μ thinks all constraints in I_n" are satisfiable

"SoS can't even solve linear equations".

Our goal: Construct such a μ.

__Recall__: enough to describe $\boxed{\underset{\mu}{\tilde{\mathbb{E}}}\, X_S,\ |S| \leq c'n}$.

__Why?__ μ is p.d. of deg $c'n$ iff $\tilde{\mathbb{E}}_\mu = 1$ ✓

& $\forall$ poly p, $\tilde{\mathbb{E}}\, p^2 \geq 0$ whenever $\deg(p) \leq \frac{c'n}{2}$.

$$\tilde{\mathbb{E}}_{\mu} \cdot val_{I} = 1$$

$$= \tilde{\mathbb{E}}_{\mu} \cdot \left(\frac{1}{2} + \frac{I(x)}{2}\right) = 1 \Leftrightarrow \underbrace{\tilde{\mathbb{E}}_{\mu}\, I(x) = 1.}$$

$$\{C_i, b_i\}_{i=1}^{m} : b_i \prod_{j \in C_i} x_j = 1$$

$$\cdot \underline{\tilde{\mathbb{E}}_{\mu} \frac{1}{m} \sum_{i=1}^{m} b_i \cdot \prod_{j \in C_i} x_j} = 1 \quad \text{---} \quad *$$

deg k monomial

I must give $\tilde{\mathbb{E}}_{\mu} \prod_{j \in C_i} x_j = b_i$

Why can't I set $\tilde{\mathbb{E}}_{\mu} \cdot \underline{x_{C_i}} = \underline{1 \cdot 1\, b_i}$?

<u>Recall (Local Distributions)</u>.

Let μ be a deg d p.d. $\{\pm 1\}^n$.

Then $\forall\, S \subseteq [n]$, $|S| \leq \frac{d}{2}$, the $\mu|_S$

is an actual distribution over $\{\pm 1\}^S$.

Conclusion: $\mathbb{E}_{\mu} X_{C_i} = b_i$

$\left\{ \begin{matrix} C_i \\ b_i \end{matrix} \right\}, \left\{ \begin{matrix} C_j \\ b_j \end{matrix} \right\}$ $\tilde{\mathbb{E}}_{\mu} X_{C_i} X_{C_j} = b_i \cdot b_j$

Def: (degree d derivation)

Define Der_d be the output of the following process:

1. In steps $1 \to m$ add $\{X_{C_i} = b_i\}_{i=1}^{m}$ to Der_d.

2. traverse all monomials is fixed order graded by degree ($\leq d$) for every X_U in this traversal check if $\{X_S = b_S\}, \{X_T = b_T\}$ are included in Der_d for some S, T such that $S \oplus T = U$.

Symmetric diff of S & T.

Then add $\{X_u = b_S \cdot b_T\}$ to Der_d.

Note: every $\{X_S = b_S\}$ is associated to a union of constraints from I.

Def: For every $\{X_S = b_S\} \in \mathrm{Der}_d$, set $\tilde{\mathbb{E}}_\mu X_S = b_S$.

Issue: What if there are 2 different conflicting derivations for X_u?

We'll show that we will never run into a conflict (whp over the choice of $\underset{\sim}{I_n}$)

Def (Hypergraph Expansion)

A k-uniform hypergraph H on $[n]$ (subset of k-size sets in $[n]$) is (t, β)-expanding if for every subset

(1) of at most t-hyperedges in H

$$\left|\bigcup_{i\in \mathbb{C}} C_i\right| \geq \beta \cdot |\mathbb{C}|$$

$$|\{1,2,3\} \cup \{1,2,4\}| = 4$$

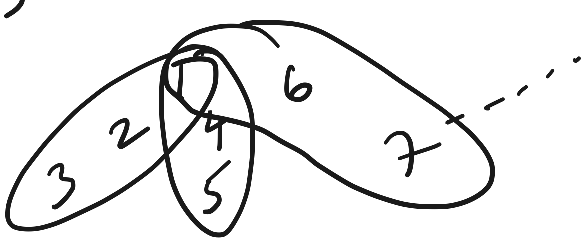
Let's focus on $k=3$

Proposition: Let I_n be random instance with m constraints. Then $\forall \delta>0$, $\exists$ $\eta = \eta(\Delta, \delta)$ s.t. w.p. 0.99, the hypergraph $\{C_1 - C_m\}$ is $(\eta \cdot n, 2-\delta)$ expanding.

$k-1-\delta$

"each constraint contributes $\sim 2-\delta$ unique variables on an average"

$$\left.\begin{array}{l}\{1,2,3\}\\ \{1,4,5\}\end{array}\right\}$$ 5 unique variables : ~ 2.5 expansion.



only need > 1.5 < 2

Lemma (no conflict in expanding I)

Suppose $\{C_1, \ldots, C_m\}$ is $(\eta \cdot n, 1.7)$-expanding. Then for $\boxed{d < \eta \cdot n / 100}$,

$$\underbrace{\nexists\ S_1, T_1, S_2, T_2 \in \mathrm{Der}_d}\ \text{s.t.}$$

$$S_1 \oplus T_1 = S_2 \oplus T_2.$$

Proof: Suppose

$$\left.\begin{array}{l} \{X_{S_1} = b_{S_1}\} : U_1 \\ \{X_{T_1} = b_{T_1}\} : U_2 \\ \{X_{S_2} = b_{S_2}\} : U_1' \\ \{X_{T_2} = b_{T_2}\} : U_2' \end{array}\right\}$$

Claim: Suppose $U_1 \ldots U_r$ are all the subset of constraints from I that appear in Der_d.

Then, $\forall i$: $|U_i| \leq 10d$.

Pf of claim: Let's pick the smallest i s.t.

$|U_i| > 10d$. $\quad U_i : \underbrace{\{X_{S_i} = b_{S_i}\}}$

then $j, k <$ ...

s.t. $\quad U_j: \{x_{S_j} = b_{S_j}\}$

$\quad\quad U_k: \{x_{S_k} = b_{S_k}\}$.

$$x_{S_j} \cdot x_{S_k} = x_{S_i} : \underbrace{U_j \oplus U_k}$$

I know: $|U_j| \leq 10d$, $|U_k| \leq 10d$.

$$|U_i| \leq 20d \leq 20 \cdot \frac{\eta \cdot n}{100} < \eta \cdot n.$$

U_i is a collection of at most $\eta \cdot n$ constraints of I.

$$\Rightarrow \underbrace{\left|\bigcup_{j \in U_i} C_j\right|} \geq \underbrace{1.7 \cdot |U_i|}.$$

$$> 1.7 \cdot 10d = 17d \gg d.$$

$$\underbrace{\left|\oplus C_j\right|} \geq 0.2 \cdot |U_i|.$$

$$\geq 0.2 \cdot 10d \geq 2d$$

$\prod_{j \in U_i} \prod_{k \in C_j} x_j$ is monomial of deg $> d$

$\longrightarrow$ Contradiction...

Pf of Lemma: $\quad \{x_{S_1} = b_{S_1}\} : U_1$

$\{x_{T_1} = b_{T_1}\} : U_2$

$\{x_{S_2} = b_{S_2}\} : U_1'$

$\{x_{T_2} = b_{T_2}\} : U_2'$

$$|\underline{U_1 \oplus U_2 \oplus U_1' \oplus U_2'}| \leq 40d < \eta \cdot n.$$

$$S_1 \oplus T_1 = S_2 \oplus T_2.$$

$\Rightarrow$ every variable must appear an even # of times in the union of clauses in U_1, U_2, U_1', U_2'

$\Rightarrow$ # variables in set of constraints $\underline{U_1 \oplus U_2 \oplus U_1' \oplus U_2'}$ is at most $\frac{3}{2} |U_1 \oplus U_2 \oplus U_1' \oplus U_2'|$

$= 1.5 |\_\_\_\_\_\_|$

$\Rightarrow$ <u>no conflict</u>

Der_d is well defined.

$\{ X_S = b_S \} \in \mathrm{Der}_d$

then we must set $\tilde{\mathbb{E}}_\mu X_S = b_S$.

$\{X_{S_1} = b_{S_1}\}$, $\{X_{S_2} = b_{S_2}\}$

$$\Downarrow$$

$\{ X_{S_1} \cdot X_{S_2} = b_{S_1} \cdot b_{S_2} \}$

$$\Rightarrow \quad \tilde{\mathbb{E}}_\mu X_{S_1} \cdot X_{S_2} = (\tilde{\mathbb{E}}_\mu X_{S_1})(\tilde{\mathbb{E}}_\mu X_{S_2})$$

What about the monomials that don't appear in Der_d?

"Guess": $\tilde{\mathbb{E}}_\mu X_S = 0$ if X_S wasn't derived in Der_d.

"max-entropy principle"

With this guess the μ is completely defined.

Lemma (μ is a p.d.)

Let $d < \eta \cdot n/100$. Then μ is a degree d p.d.

1. $\tilde{\mathbb{E}}_\mu 1 = 1$
2. $\tilde{\mathbb{E}}_\mu \operatorname{val}_I(X) = 1$.

Proof: 1 & 2 satisfied by construction.

What remains is $\tilde{\mathbb{E}}_\mu p^2 \geq 0$ for p of $\deg \leq \frac{d}{2}$.

Define an equivalence relation on $\{ S \subseteq [n] \mid |S| \leq \frac{d}{2} \}$.

$S \sim T$ if $\tilde{\mathbb{E}}_\mu X_S \cdot X_T \neq 0$

iff $X_S \cdot X_T$ appears in Der_d.

1. $S \sim S$? $\hat{\mathbb{E}} X_S \cdot X_S = \tilde{\mathbb{E}}\, 1 = 1$

② $S \cup S'$ $\Rightarrow S \cup S''$ $\rightarrow |S \oplus S'| \le d.$

$S' \cup S''$

$S' \cup S'' : |S \oplus S''| \le d.$

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

$$\underbrace{\tilde{\mathbb{E}}\, X_S \cdot X_{S''}} = \tilde{\mathbb{E}}\big(\underbrace{X_S \cdot X_{S'}} \cdot \underbrace{X_{S'} \cdot X_{S''}}\big)$$

$$= \tilde{\mathbb{E}}[X_S \cdot X_{S'}] \cdot \tilde{\mathbb{E}}[X_{S'} \cdot X_{S''}]$$

$$\neq 0.$$

For every class $R_1, \ldots, R_N$

choose rep: $S_1, \ldots, S_N.$

Then observe: ① $S, T \in R_i$

then $\tilde{\mathbb{E}}(X_S \cdot X_T)$ $(X_S \cdot X_{S_i})$ $(X_{S_i} \cdot X_T)$

$$= (\tilde{\mathbb{E}}\, X_S \cdot X_{S_i})(\tilde{\mathbb{E}}\, X_{S_i} \cdot X_T)$$

② $S \in R_i, T \in R_j$ for $i \neq j$

$$\tilde{\mathbb{E}}\, X_S \cdot X_T = 0.$$

$$p = \sum_{\substack{S \subseteq [n] \\ |S| \leq \frac{d}{2}}} p_S \cdot X_S$$

$$= \sum_i \underbrace{\sum_{S \in R_i} p_S \cdot X_S}_{p_i} = \sum_i p_i$$

$$p_i p_j = \left(\sum_{S \in R_i} p_S \cdot X_S\right)\left(\sum_{S \in R_j} p_S \cdot X_S\right)$$

$$\tilde{\mathbb{E}} \cdot p^2 = \tilde{\mathbb{E}}\left(\sum_i p_i\right)^2$$

$$= \sum_{i,j} \tilde{\mathbb{E}} p_i \cdot p_j$$

$$= \underbrace{\sum_i \tilde{\mathbb{E}} p_i^2}_{} + \sum_{i \neq j} 0$$

$$\tilde{\mathbb{E}} \cdot p_i^2 = \sum_{S,T \in R_i} p_S \cdot p_T \tilde{\mathbb{E}} X_S \cdot X_T$$

$$= \sum_{S,T \in R_i} p_S \cdot p_T \left(\hat{\mathbb{E}} \cdot X_S \cdot X_{S_i}\right)\left(\hat{\mathbb{E}} \cdot X_{S_i} \cdot X_T\right)$$

$$= \left(\sum_S \cdot p_S \cdot \mathbb{E}\, X_S \cdot X_{S_i} \right)$$

$$\geq 0.$$

---

where did we use randomness of $C_1 \dots C_m$

→ <u>expansion</u> of the hypergraph.

Where did we use the randomness in RHS $b_1 \dots b_m$:

→ upper bound on $\mathrm{opt}(I)$.

Construction of μ works for arbitrary b_is

→ doesn't need randomness.

We showed that random K-XOR is hard. → "maximally hard".

<u>CSPs</u>: $P: \{\pm 1\}^K \to \{0,1\}$

<u>each instance</u>: $C_1, \dots, C_m \subseteq [n]$

negation pattern for each constraint

<u>Goal</u>: $x \in \{\pm 1\}^n$ $\quad P(C_i(x)) = 1$

$\forall\ 1 \leq i \leq m.$

Can completely characterize the set of predicates P for which

<u>random CSP (P)</u> is maximally hard

worst-case CSP(P) is maximally hard.

<u>We know</u>: Complete answers for both. worst-case(P), conditioned on U.G.C.

<u>Next time</u>: we'll study the answers to both.