Sum-of-Squares Algorithm

Computer Science, Yonsei University
Changyeol Lee

Problem

- Given $f: \{\pm 1\}^n \to \mathbb{R}$,

  verify if $f \geq 0$, i.e., $f(x) \geq 0$ for all $x \in \{\pm 1\}^n$.

- $\{\pm 1\}^n$ is called a hypercube
  - I found out that $\{0,1\}$ is more widely used definition for a hypercube…
    But note that there is no difference .

Problem

- Given $f: \{\pm 1\}^n \to \mathbb{R}$,
  efficiently verify if $f \geq 0$, i.e., $f(x) \geq 0$ for all $x \in \{\pm 1\}^n$.

- $\{\pm 1\}^n$ is called a hypercube
- "Given f" = Given a vector of coefficients of f
  - but... there are infinite number of representation of f...

Multilinear Reduction

- Reduction: $x_i^{\mathrm{odd}} \to x_i$ and $x_i^{\mathrm{even}} \to 1$.
- After the reduction, we can view f as a vector $\hat{f} \in \mathbb{R}^{2^n}$ where $\hat{f}(S)$ is the coefficient of $\prod_{i \in S} x_i$.

- Fact. Regardless of representation of f, the reduction outputs a unique vector $\hat{f}$, which implies $f \equiv f'$ iff $\hat{f} \equiv \widehat{f'}$.

Multilinear Reduction

- If the input f satisfies $\deg(f) \leq d$ (after the reduction), the dimension of $\hat{f}$ can be $\sum_{i=0}^{\ell} \binom{n}{i} \leq (d+1)n^d = n^{O(d)}$.

Problem

- Given $f: \{\pm 1\}^n \to \mathbb{R}$ s.t. $\deg(f) \leq d$,
  efficiently verify if $f \geq 0$, i.e., $f(x) \geq 0$ for all $x \in \{\pm 1\}^n$.

- $\{\pm 1\}^n$ is called a hypercube
- "Given f" = "Given $\hat{f}$ of f (WLOG)"

Problem

- Given $f: \{\pm 1\}^n \to \mathbb{R}$ s.t. $\deg(f) \leq d$ and has rational coeffs., efficiently verify if $f \geq 0$, i.e., $f(x) \geq 0$ for all $x \in \{\pm 1\}^n$.

- $\{\pm 1\}^n$ is called a hypercube
- "Given f" = "Given $\hat{f}$ of f (WLOG)"

Problem

- Given $f: \{\pm 1\}^n \rightarrow \mathbb{R}$ s.t. $\deg(f) \leq d$ and has rational coeffs., efficiently verify if $f \geq 0$

(Our) Application: Max-Cut Problem

- Let $f_G(x) := \frac{1}{4}\sum_{(i,j)\in E}\left(x_i - x_j\right)^2$ be the function over $\{\pm 1\}^{|V|}$.
  - find a bipartition of vertices; count the #edges on the cut.
- We want to find x^* that maximizes $f_G(x)$.
- Or... we want to find the smallest c s.t. $c - f_G(x) \geq 0$.
- Naive randomized alg.: for each $i \in V$, assign $+1$ or -1 u.a.r.
  - $|E|/2$ number of edges are on the cut (in expectation)

(Athor) Applications

- Graph densities; flag algebras (Cauchy-Schwarz proof);
- Quantum information
- and so on…

SoS (Sum-of-Squares) Certificate

d is even

- A degree d SoS certificate for f is

  a list of polynomial functions $g_1, \dots, g_r : \{\pm 1\}^n \to \mathbb{R}$ s.t.

  - $\deg(g_i) \le d/2$ for all $i = 1, \dots, r$ and

    e.g., $\deg(x_1 x_3 x_7 + x_2 x_7 + 2) = 3$

  - $f(x) = g_1^2(x) + \cdots + g_r^2(x)$ for all $x \in \{\pm 1\}^n$.

To Answer the Problem with SoS Cert…

- Suppose someone gives a degree d SoS certificate $g_1, \dots, g_r$.
- Does this efficiently verify $f \geq 0$?
- Requirement:
  - d and r should not be too large.
  - #bits needed to represent coeffs. of g_i should not be too large.
  - Testing $f = g_1^2 + \cdots + g_r^2$ is done efficiently.
    - Note that testing for all $x \in \{\pm 1\}^n$ is not efficient.

To Answer the Problem with SoS Cert…

- Suppose someone gives a degree d SoS certificate $g_1, \dots, g_r$.
- Does this efficiently verify $f \geq 0$?
- Requirement:
  - d and r should not be too large.
  - #bits needed to represent coeffs. of g_i should not be too large.
  - Testing $f = g_1^2 + \cdots + g_r^2$ is done efficiently.
    - Compare the coeffs. of $\sum g_i^2$ after the multilinear reduction ($n^{O(d)}$ comparisons)

Does $f \geq 0$ always have a SoS cert.?

- Yes, but with high degree.
- There exists a deg $2n$ SoS cert. of $f \geq 0$ where $f: \{\pm 1\}^n \to \mathbb{R}$.
- Consider $g \coloneqq \sqrt{f}$.

  Since f has degree at most n WLOG, g is a deg $2n$ SoS cert.

Does $f \geq 0$ always have a SoS cert.?

- Yes, but with high degree.
- There exists a deg $2n$ SoS cert. of $f \geq 0$ where $f: \{\pm 1\}^n \to \mathbb{R}$.
- Consider $g \coloneqq \sqrt{f}$.

  Since f has degree at most n WLOG, g is a deg $2n$ SoS cert.
- WRONG… since g can be non-polynomial.

Does $f \geq 0$ always have a SoS cert.?

- Correct proof) Consider $\{g_z\}_{z\in\{\pm 1\}^n}$ s.t. $g_z(x) = f(z)$ iff $x = z$.
- Then $\sum_{z\in\{\pm 1\}^n} g_z^2 = f$.
- Let $g_z(x) \coloneqq \sqrt{f(z)\left(\frac{1+z_1x_1}{2}\right)^2\left(\frac{1+z_2x_2}{2}\right)^2\cdots\left(\frac{1+z_nx_n}{2}\right)^2}$.
- $\deg(g_z) \leq n$ and g_z is a polynomial.
- $\{g_z\}_{z\in\{\pm 1\}^n}$ is a degree $2n$ SoS certificate. Q.E.D.

“Smaller” SoS certificate?

- If there exists a deg d SoS cert., coeffs $\leq 2^{\mathrm{poly}(n^d)}$.
  - Proof uses the fact that f is on the hypercube.

- How large is r if there exists a deg d SoS cert.?

Notations and Definitions

- Every vector is a column vector by default.
- $v^{\otimes 2} = v \otimes v \in \mathbb{R}^{n^2}$, $v^{\otimes 2}(i,j) = v_i \cdot v_j$
- $(1, v) = (1, v_1, \ldots, v_n) \in \mathbb{R}^{n+1}$
- $(1, x)^{\otimes d}$: all possible monomials with deg $\leq d$ with redundancy.
  - Suppose $x = (x_1, x_2)$ and $d = 4$.

    $(1,x)^{\otimes d/2} = (1,x)^{\otimes 2} = \left(1, x_1, x_2, x_1, x_1^2 (= 1), x_1 x_2, x_2, x_2 x_1, (x_2^2 = 1)\right)$

Notations and Definitions

- A matrix $M \in R^{n\times n}$ is positive semidefinite (PSD) (or $M \succcurlyeq 0$) iff M is symmetric and $v^T M v \geq 0$ for all $v \in \mathbb{R}^n$.

- Fact. For all positive semidefinite matrix $A \in R^{n\times n}$, $\exists B \in R^{m\times n}$ $(m \leq n)$ s.t. $A = B^T B$.

Theorem. f has a deg d SoS cert. iff $\exists A \succcurlyeq 0$ s.t. $\left((1,x)^{\otimes d/2}\right)^T A(1,x)^{\otimes d/2} = f$.

- Proof, $\Rightarrow$ part) Let $\{g_i\}$ be a deg d SoS cert.
- Let v_i be a vector s.t. $g_i = v_i^T(1,x)^{\otimes d/2}$. Intuitively, v_i is a coeff. vector of g_i.
- $f = \sum_i \left(v_i^T(1,x)^{\otimes d/2}\right)^2 = \sum_i \left(v_i^T(1,x)^{\otimes d/2}\right)^T v_i^T(1,x)^{\otimes d/2}$

$$= \sum_i \left((1,x)^{\otimes d/2}\right)^T v_i v_i^T (1,x)^{\otimes d/2}$$

$$= \left((1,x)^{\otimes d/2}\right)^T \left(\sum_i v_i v_i^T\right)(1,x)^{\otimes d/2}$$

This matrix is PSD.

Theorem. f has a deg d SoS cert. iff $\exists A \succcurlyeq 0$ s.t. $\left((1,x)^{\otimes d/2}\right)^T A(1,x)^{\otimes d/2} = f$.

- Proof, $\Leftarrow$ part) $f = \left((1,x)^{\otimes d/2}\right)^T B^T B(1,x)^{\otimes d/2}$ for some B.

- $f = \left(B(1,x)^{\otimes d/2}\right)^T B(1,x)^{\otimes d/2} = \left\|B(1,x)^{\otimes d/2}\right\|_2^2$ sum of squares of each entry of $B(1,x)^{\otimes d/2}$

- Let g_i be the i-th entry of $B(1,x)^{\otimes d/2}$.

- Note $\deg(g_i) \le d/2$ and $f = \sum_i g_i$ and g_i is polynomial. Q.E.D.

Corollary

- If f has a deg d SoS cert.,

  then $\exists$ a deg d SoS cert. $g_1, \ldots, g_r$ where $r \leq (n+1)^{d/2}$.

  - Recall "Let g_i be the i-th entry of $B(1,x)^{\otimes d/2}$" part.

So far…

- If $\exists A \succcurlyeq 0$ s.t. $\left((1,x)^{\otimes d/2}\right)^T A(1,x)^{\otimes d/2} = f$,

  we can construct a deg d SoS cert. $g_1, \ldots, g_r$ w/ $r \leq (n+1)^{d/2}$

- How to find such A?

Understanding $\left((1,x)^{\otimes d/2}\right)^T A(1,x)^{\otimes d/2}$

- Each element of $(1,x)^{\otimes d/2}$ corresponds to $x^S \coloneqq \prod_{i \in S} x_i$ for some set S, i.e., $(1,x)^{\otimes d/2}$ can be indexed by monomials.
  w/ redundancy
- Similarly, each row and column of A can be indexed by a set.
- Consider $A_{S,S'}$ which is S-th row and S'-th column of A.
  It contributes to $\hat{f}(U)$, coeff. of x^U in f where $x^U \equiv x^S \cdot x^{S'}$.

equivalence after the
multilinear reduction

Equivalent Statements

- Find $A \succcurlyeq 0$ s.t. $\left((1,x)^{\otimes d/2}\right)^T A(1,x)^{\otimes d/2} = f$.
- Find $A \succcurlyeq 0$ s.t. for all $U \subseteq \{1,\ldots,n\}$ s.t. $|U| \leq d$,

$$\sum_{\substack{\forall S,S': \\ x^U \equiv x^S \cdot x^{S'}}} A_{S,S'} = \hat{f}(U).$$

Semidefinite Programming (SDP)

- We can efficiently find $A \succcurlyeq 0$ s.t. $\forall U \subseteq \{1, \dots, n\}$ s.t. $|U| \leq d$,

  $$\sum_{\substack{\forall S, S': \\ x^U \equiv x^S \cdot x^{S'}}} A_{S,S'} \in \left[\hat{f}(U) - \epsilon, \hat{f}(U) + \epsilon\right] \text{ for some } \epsilon > 0.$$

  - Using SDP, we can efficiently solve the "relaxed" one.

$$
\begin{aligned}
&\text{maximize or minimize} && \sum_{i,j} c_{ij} x_{ij} && \\
&\text{subject to} && \sum_{i,j} a_{ijk} x_{ij} = b_k, && \forall k, \\
& && x_{ij} = x_{ji}, && \forall i, j, \\
& && X = (x_{ij}) \succeq 0. &&
\end{aligned}
$$

linear objective

linear constraints

w/ additional constraint that
a square symmetric matrix of variables is PSD

Semidefinite Programming (SDP)

- Let A' be the returned matrix.
  We know f' s.t. $\widehat{f'}(U) = \sum_{\substack{\forall S,S': \\ x^U \equiv x^S \cdot x^{S'}}} A'_{S,S'}$ has a deg d SoS cert.
- Claim. $f + \epsilon(n+1)^d$ also has a deg d SoS cert.
  - Proof) Since for any U s.t. $|U| \leq d$, $\left|\hat{f}(U) - \widehat{f'}(U)\right| \leq \epsilon$
    we have $\sum_{U:|U|\leq d} \left|\hat{f}(U) - \widehat{f'}(U)\right| \leq \epsilon(n+1)^d$.
  - Claim. For any $h: \{\pm 1\}^n \to \mathbb{R}$ w/ $\deg(h) \leq d$,
    we can show $h + \sum_{U:|U|\leq d} \left|\hat{h}(U)\right|$ has a deg d SoS cert. (proof omitted)
  - $f - f' + \epsilon(n+1)^d$ and f' both have a deg d SoS cert. Q.E.D.

Theorem.

- There is an efficient algorithm that
  if the given $f: \{\pm 1\}^n \to \mathbb{R}$ (with $\deg(f) \leq d$) has a deg d SoS cert., then the algorithm outputs a deg d SoS cert. for $f + \epsilon'$.

Max-Cut Problem

- Recall $f_G(x) \coloneqq \frac{1}{4}\sum_{(i,j)\in E}\left(x_i - x_j\right)^2$ is the function over $\{\pm 1\}^{|V|}$.
- Let c' be the (almost) smallest value s.t.

  the alg. outputs a deg 2 SoS cert. for $c' - f_G(x) + \epsilon'$.
  - The alg. didn't return a deg 2 SoS cert. for $c'' - f_G(x) + \epsilon'$ for $c'' < c'$.
- Our hope is that $\alpha \cdot f_G(x^*) \le c'\big(\le f_G(x^*)\big)$ for some $\alpha \lesssim 1$.

Theorem.

- There is an efficient algorithm that if the given $f:\{\pm 1\}^n \to \mathbb{R}$ (with $\deg(f) \leq d$) has a deg d SoS cert., then the algorithm outputs a deg d SoS cert. for $f + \epsilon'$.

- Even if f does not admit a deg d SoS cert., we want some (efficient) "dual object" that we can utilize.

Duality and
Sum-of-Squares Algorithm

Geometric Intuition

- What does it mean that $\nexists$ deg d SoS cert. for some $\bar{f}$?
- Let $SoS_d \coloneqq \{f : \{\pm 1\}^n \to \mathbb{R} \mid f \text{ has a deg } d \text{ SoS cert.}\}$
- Observe that SoS_d is a convex cone, i.e., closed under convex combination & nonnegative scaling.
- If $\bar{f} \notin SoS_d$, then there should be a separating hyperplane!

through the origin

Geometric Intuition

- We can represent a hyperplane by $\mu: \{\pm 1\}^n \to \mathbb{R}$.
  - WLOG, $\sum_{x \in \{\pm 1\}^n} \mu(x) = 1$.
- Consider the halfspace H that contains SoS_d but not $\bar{f}$, i.e.,

  $H = \{h: \{\pm 1\}^n \to \mathbb{R} \mid \sum_{x \in \{\pm 1\}^n} \mu(x) h(x) \geq 0\}$.
- If $\mu(x) \geq 0$ for all x, μ can be seen as a prob. distribution.

  For all $f \in SoS_d$, $E_{x \sim \mu}[f(x)] \geq 0$ but $E_{x \sim \mu}[\bar{f}(x)] < 0$.

μ serves as a dual object!

Geometric Intuition

- We can represent a hyperplane by $\mu: \{\pm 1\}^n \to \mathbb{R}$.
  - WLOG, $\sum_{x \in \{\pm 1\}^n} \mu(x) = 1$.
- Consider the halfspace H that contains SoS_d but not $\bar{f}$, i.e.,

  $H = \{h: \{\pm 1\}^n \to \mathbb{R} \mid \sum_{x \in \{\pm 1\}^n} \mu(x) h(x) \geq 0\}$.
- What if $\mu(x) < 0$ for some x, i.e., not a prob. dist.?

  We will see that μ still behaves prob.-like distribution.

  (pseudo)

(More) Notations and Definitions

- Formal expectation of $f: \{\pm 1\}^n \to \mathbb{R}$ w.r.t. a distribution μ

$$\tilde{E}_\mu[f] = \sum_x \mu(x) f(x)$$

  - μ is not necessarily a probability distribution!
    (Therefore, $x \sim \mu$ is not well defined.)

(More) Notations and Definitions

- A deg d pseudo-distribution (over the hypercube) is a function μ (over the hypercube) s.t.
  - $\tilde{E}_\mu[1] = 1$, (i.e., the sum of entries of μ is 1) and
  - for every polynomial g of $\deg(g) \leq d/2$, $\tilde{E}_\mu[g^2] \geq 0$.

Captures the "separating hyperplane condition"

- A deg d pseudo-expectation is a formal expectation w.r.t. a deg d pseudo-distribution.

Pseudo-distribution as a dual object

- We will see a pseudo-distribution will serve as a dual object.
  - Technical issue: How can we represent a pseudo-distribution?
  - Answer: For all deg d pseudo-dist. μ,

    there is a deg d pseudo-dist. μ' with $\mathrm{deg}(\mu') \leq d$.

    -> use multilinear reduction

Duality of SoS Cert. and Pseudo-dist.

- Theorem. For every $f:\{\pm 1\}^n \to \mathbb{R}$ and every even $d \in \mathbb{N}$,

  $f \in SoS_d$ iff every deg d pseudo-dist. μ (over $\{\pm 1\}^n$) satisfies $\tilde{E}_\mu[f] \geq 0$.
  - Proof, ⇒ part) $f = \sum_i g_i$ where $\deg(g_i) \leq d/2$. $\tilde{E}_\mu[f] = \sum_i \tilde{E}_\mu[g^2] \geq 0$.
  - Proof, ⇐ part) Sps $\bar{f} \notin SoS_d$. Consider a separating hyperplane μ.

    We show μ is a deg d pseudo-dist. Suffices to show $\tilde{E}_\mu[1] > 0$.

    We have $\tilde{E}_\mu[\bar{f}] < 0$. Choose large enough L s.t. $\tilde{E}_\mu[\bar{f} + L] \geq 0$.

    Since $L \cdot \tilde{E}_\mu[1] = \tilde{E}_\mu[L]$ we have $\tilde{E}_\mu[1] \geq -\tilde{E}_\mu[\bar{f}]/L$ and RHS>0. Q.E.D.

SoS Algorithm

- Theorem. There is an efficient algorithm that, given f,

  either outputs a deg d SoS cert. for $f + \epsilon'$ or

  outputs a pseudo-dist. μ s.t. $\tilde{E}_\mu[f] < \epsilon'$.

  - We know there is a pseudo-dist. μ s.t. $\tilde{E}_\mu[f] < 0$ if $f \notin SoS_d$ but we don't know how to find it efficiently. Instead, as in finding a SoS cert., we can efficiently compute the "approximate" pseudo-dist. using SDP.

Side note) Equiv. Defn. of Pseudo-dist.

- Let $M_{d/2} \coloneqq (1,x)^{\otimes d/2}\left((1,x)^{\otimes d/2}\right)^T$.

element-wise pseudo-expectation

- μ is a deg d pseudo-dist. iff $\tilde{E}_\mu[1] = 1$ and $\tilde{E}_\mu\left[M_{d/2}\right] \succcurlyeq 0$.
  - proof) Consider any polynomial g w/ $\deg(g) \le d/2$.
  - $\tilde{E}_\mu[g^2] = \tilde{E}_\mu\left[\left(\sum_{S:|S|\le d/2} \hat{g}(S)x^S\right)^2\right] = \tilde{E}_\mu\left[\sum_{S,S':|S|,|S'|\le d/2} \hat{g}(S)\hat{g}(S')x^S x^{S'}\right]$

$= \sum_{S,S':|S|,|S'|\le d/2} \hat{g}(S)\hat{g}(S')\tilde{E}_\mu\left[x^S x^{S'}\right] = (\hat{g})^T M_{d/2}(\hat{g})$. Q.E.D.

changing the order
of summation

Max-Cut Algorithm

Max-Cut Problem

- Suppose $c' - f_G \in SoS_2$ but $c' - \epsilon' - f_G \notin SoS_2$.
- The sum-of-squares alg. outputs a deg 2 pseudo-dist. μ over the hypercube s.t. $\tilde{E}_\mu[c' - \epsilon' - f_G] < \epsilon'$, i.e., $\tilde{E}_\mu[f_G] > c' - 2\epsilon'$.
- If μ is an actual prob. dist., we are happy.
  - $E_\mu[f_G] = \tilde{E}_\mu[f_G] > f_G(x^*) - 2\epsilon'$
- If not, "round" a pseudo-dist. into an actual prob. dist. μ'.

Max-Cut Problem

- Theorem. For every G and deg 2 pseudo-dist. μ over the hypercube, there exists a prob. dist. μ' over the hypercube s.t.

$$E_{\mu'}[f_G] \geq \alpha \tilde{E}_{\mu}[f_G]$$

where $\alpha = 0.878\ldots$ (to be defined more precisely later).

  - Corollary 1. Finding μ' efficiently leads to an $(\alpha - 2\epsilon')$-approx. alg.

(or 0.878-approx. alg.)

Max-Cut Problem

- Theorem. For every G and deg 2 pseudo-dist. μ over the hypercube, there exists a prob. dist. μ' over the hypercube s.t.

$$E_{\mu'}[f_G] \geq \alpha \tilde{E}_\mu[f_G]$$

where $\alpha = 0.878 \ldots$ (to be defined more precisely later).

  - Corollary 2. $f_G(x^*)/\alpha - f_G \in SoS_2$.
    - Sps not. Then there is a deg 2 pseudo-dist. μ s.t. $\tilde{E}_\mu[f_G(x^*)/\alpha - f_G] < 0$.
    - By the above thm, $\exists \mu'$ s.t. $\alpha \tilde{E}_\mu[f_G] \leq E_{\mu'}[f_G]$. Contradiction. Q.E.D.

Proof of the Theorem.

- (To be filled out)

Final Remarks
CSP, UGC and SoS

2-CSP (Constraint Satisfaction Problems)

- $x_1, \ldots, x_n$ are variables; take values in some finite alphabet Σ.
- $\left(x_{i_1}, x_{i_2}\right), A_i$ are i-th constraints; $1 \leq i \leq m$.
  - I will call A_i "allowable assignment" of x_{i_1} and x_{i_2}.

The number 2 of 2-CSP comes from here.
Constraints consist of at most 2 variables.

- We give an assignment $x' \in \Sigma^n$.

  We say i-th constraint is satisfied if $\left(x'_{i_1}, x'_{i_2}\right) \in A_i$.
- Goal: find an assignment that maximizes #satisfied constraints.

2-CSP (Constraint Satisfaction Problems)

- Max-Cut
  - Variable for each vertex; $\Sigma = \{\pm 1\}$
    for each $e = (i, j)$, constraint = $\left(x_i, x_j\right), \{(1, -1), (-1, 1)\}$.
- Max-2-Sat
  - Variable for each vertex; $\Sigma = \{T, F\}$;
    for each constraint, $\left(x_{i_1}, x_{i_2}\right), \{(T, F), (F, T), (T, T)\}$
- Max-3-Coloring
  - Variable for each vertex; $\Sigma = \{R, G, B\}$; for each $e = (i, j)$,
    constraint = $\left(x_i, x_j\right), \{(R, G), (R, B), (G, R), (G, B), (B, R), (B, G)\}$.

Unique 2-CSP

- A constraint $(x_{i_1}, x_{i_2}), A_i$ is unique if for any assignment to one variable (x_{i_1}), there is exactly one assignment to the other variable (x_{i_2}) that makes the constraint satisfied.
- 2-CSP is called unique, if all constraints are unique.
  - Max-Cut is unique but Max-2-Sat or Max-3-coloring are not unique.
- Unique game (UG) is another name for unique 2-CSP.

(c, s)-CSP

- Given a CSP instance, decide if
  - there is an assignment that satisfies at least $c \cdot m$ constraints
  - for any assignment, #satisfied constraints is at most $s \cdot m$.

- Obs. $(1, 1 - 1/m)$-CSP is checking the satisfiability.
- Obs. $(1, 1 - 1/m)$-UG is polytime solvable.

Unique Game Conjecture (Khot'02)

- For every sufficiently small $\epsilon > 0$,
  there is a large enough constant k such that
  $(1-\epsilon, \epsilon)$-UG is NP-hard where $|\Sigma| = k$.

- Obs. Uniform random assignment satisfies $1/k$ fraction.
  - This implies $k \gg 1/\epsilon$.

Under UGC,

- for every $\epsilon > 0$, $(\alpha + \epsilon)$-approx for Max-Cut is NP-hard
  - There are many similar results such as
    $\forall\epsilon > 0$, $(2 - \epsilon)$-approx for Vertex Cover is NP-hard, …
- It turns out that (by Raghavendra'08)
  for all CSP, a "natural" SDP (deg 2 SoS) + "natural" rounding
  is optimal.
  + And it matches the integrality gap.

Conclusion

- For deg 2 SoS, we know the exact power for CSP.
- deg d SoS for $d = 4,6,\ldots$? Nothing known.

- SoS proof technique in other fields?