

Sum of Squares: Part 1

Apoorv Vikram Singh

February 3, 2021

Introduction

Reason two basic problems about polynomial inequalities

1. Feasibility of polynomial system.

$$\begin{aligned} p_1(\mathbf{x}) &\geq 0 \\ &\vdots \\ p_m(\mathbf{x}) &\geq 0 \end{aligned} \tag{1}$$

Is there an \mathbf{x} such that (1) is satisfied?

2. Checking non-negativity: Is $q(\mathbf{x}) \geq 0$, $\forall \mathbf{x}$ satisfying (1) ?

Justification: Feasibility

Feasibility checking is highly expressive.

- Example: MaxCut.
- . Input: $G = (V, E)$, $|V| = n$.
- . Goal: Find $S \subseteq V$ such that $|E(S, \bar{S})|$ is maximized.
- . Polynomial Feasibility: For some $\beta \in \mathbb{Z}^+$,

$$\frac{1}{4} \sum_{(i,j) \in E} (\mathbf{x}_i - \mathbf{x}_j)^2 = \beta$$

$$\mathbf{x}_i^2 = 1, \quad \forall 1 \leq i \leq n.$$

- . $n + 1$ degree-2 polynomials. Enumerate over polynomial number of values of β and solve MaxCut.

Other examples: MaxClique, Max 3-SAT, Knapsack. Therefore, polynomial feasibility checking problem is *NP*-Hard.

Goal

- ▶ Analyze a relaxation for the feasibility problem, and try to find interesting situations where one can get a poly-time algorithms.

Justification: Non-Negativity

- ▶ Checking non-negativity: Given $f : \{-1, 1\}^n \rightarrow \mathbb{R}$ with rational coefficients, decide if:
 - $f \geq 0$, $\forall \mathbf{x} \in \{-1, 1\}^n$, or,
 - find an $\mathbf{x} \in \{-1, 1\}^n$ such that $f(\mathbf{x}) < 0$.
- ▶ Example MaxCut: Decide if $\text{MaxCut} \leq c$.
 - . Let $f_G(\mathbf{x}) \stackrel{\text{def}}{=} \frac{1}{4} \sum_{(i,j) \in E} (\mathbf{x}_i - \mathbf{x}_j)^2$.
 - . Decide if $c - f_G(\mathbf{x}) \geq 0$, $\forall \mathbf{x} \in \{-1, 1\}^n$.

Certifying Non-Negativity

Given $f : \{-1, 1\}^n \rightarrow \mathbb{R}$, find an “efficiently verifiable” certificate of non-negativity.

SoS Certificates

Definition (SoS cert of non-neg (or) SoS proof of non-neg)

A degree- d SoS certificate of non-negativity of $f : \{-1, 1\}^n \rightarrow \mathbb{R}$ is a list of polynomials $g_1, \dots, g_r : \{-1, 1\}^n \rightarrow \mathbb{R}$, such such that

- $\deg(g_i) \leq d/2$, and
- $f(\mathbf{x}) = \sum_{i \leq r} g_i^2(\mathbf{x}), \forall \mathbf{x} \in \{-1, 1\}^n$.

Efficiently Verifiable (?)

Polynomials f, g_1, \dots, g_r are represented as a vector of coefficients.

1. How large if r ? ($\leq n^d$, see later)
2. How large are coefficients of g_i ?

Efficiently Verifiable

Proposition (Efficiently Verifiable)

Suppose $r \leq n^d$, all coefficients of g_i are bounded in magnitude by $2^{\text{poly}(n^d)}$. Then the identity $f = \sum_{i \leq r} g_i^2$ over all $\mathbf{x} \in \{-1, 1\}^n$ can be checked in $\text{poly}(n^d)$ time.

Proof.

- Given g_i , can compute g_i^2 , and $\sum_{i \leq r} g_i^2$ in polynomial time.
- Check if $(f - \sum_{i \leq r} g_i^2)(\mathbf{x}) = 0, \forall \mathbf{x} \in \{-1, 1\}^n$.
- Using the fact that coefficient vector representation is unique, just check if $f - \sum_{i \leq r} g_i^2 = \mathbf{0}$



Fact (Unique Representation)

$\forall f : \{-1, 1\}^n \rightarrow \mathbb{R}$, there exists a unique representation of f : The multi-linear representation of f

$$f = \sum_{S \subseteq [n]} \hat{f}(S) \prod_{i \in S} x_i.$$

(this representation is its Fourier transform)

\implies coefficient vector representation is unique.

(multilinear representation exists because $x_i^2 = 1$)

Are non-negative functions always certifiable?

Proposition (Certifiability of non-negative functions)

Let $f : \{-1, 1\}^n \rightarrow \mathbb{R}$ be non-negative over $\{-1, 1\}^n$. Then, there exists a $\deg(2n)$ -SoS certificate of non-negativity.

Proof.

- Consider $g : \{-1, 1\}^n \rightarrow \mathbb{R}$, and $g(\mathbf{x}) = \sqrt{f(\mathbf{x})}$.
- Every function on $\{-1, 1\}^n$ is a polynomial of $\deg \leq n$.
- $f = g^2 \implies \deg(2n)$ -SoS Certificate.



Tensor Notation

- Suppose vector $\mathbf{v} \in \mathbb{R}^n$.
- $\mathbf{v}^{\otimes 2} \in \mathbb{R}^{n^2}$, where $\mathbf{v}(i, j) = \mathbf{v}_i \mathbf{v}_j$.
- $\mathbf{v}^{\otimes k} \in \mathbb{R}^{n^k}$.

Proving Efficient Verifiability

Theorem (PSD Matrices and SoS Certificates)

$f : \{-1, 1\}^n \rightarrow \mathbb{R}$ has a $\deg(d)$ -SoS certificate of non-negativity iff there exists a matrix A such that $A \succcurlyeq 0$, and

$$f(\mathbf{x}) = \left\langle (1, \mathbf{x})^{\otimes \frac{d}{2}}, A \cdot (1, \mathbf{x})^{\otimes \frac{d}{2}} \right\rangle .$$

► Parsing Notation:

- $(1, \mathbf{x}) \in \mathbb{R}^{(n+1)}$.
- $(1, \mathbf{x})^{\otimes \frac{d}{2}}$: populate in a vector all possible monomials in the variable \mathbf{x} of degree at most $d/2$.
- $A \in \mathbb{R}^{(n+1)^{d/2} \times (n+1)^{d/2}}$.

Proof

- If Part:

$$\begin{aligned} f(\mathbf{x}) &= \left\langle (1, \mathbf{x})^{\otimes \frac{d}{2}}, A \cdot (1, \mathbf{x})^{\otimes \frac{d}{2}} \right\rangle \\ &= \left\langle (1, \mathbf{x})^{\otimes \frac{d}{2}}, (B^\top B) \cdot (1, \mathbf{x})^{\otimes \frac{d}{2}} \right\rangle \\ &= \left\langle B \cdot (1, \mathbf{x})^{\otimes \frac{d}{2}}, B \cdot (1, \mathbf{x})^{\otimes \frac{d}{2}} \right\rangle \\ &= \left\| B \cdot (1, \mathbf{x})^{\otimes \frac{d}{2}} \right\|^2. \end{aligned} \tag{2}$$

- Let $g_i(\mathbf{x}) = \left\langle e_i, B \cdot (1, \mathbf{x})^{\otimes \frac{d}{2}} \right\rangle$, i.e., i -th entry of the vector.
- B is a matrix of constants, applied to monomials of degree at most $d/2$, therefore, $\deg(g_i) \leq d/2$.
- Therefore,

$$f(\mathbf{x}) = \sum_{i=1}^{(n+1)^{\otimes d/2}} g_i^2(\mathbf{x}).$$

Proof Cont...

- Only if Part: Suppose f has a degree- d SoS certificate.

$$f = \sum_{i \leq r} g_i^2, \text{ and } \underbrace{g_i(\mathbf{x})}_{\deg \leq d/2} = \left\langle \mathbf{v}_i, (1, \mathbf{x})^{\otimes \frac{d}{2}} \right\rangle.$$

$$\begin{aligned} f(\mathbf{x}) &= \sum_{i \leq r} \left\langle \mathbf{v}_i, (1, \mathbf{x})^{\otimes \frac{d}{2}} \right\rangle^2 \\ &= \left\langle (1, \mathbf{x})^{\otimes \frac{d}{2}}, \underbrace{\left(\sum_{i \leq r} \mathbf{v}_i \mathbf{v}_i^\top \right)}_{\succeq 0} \cdot (1, \mathbf{x})^{\otimes \frac{d}{2}} \right\rangle. \end{aligned}$$



Efficient Verifiability

Corollary (Bound on r)

If $f : \{-1, 1\}^n \rightarrow \mathbb{R}$ has a degree- d SoS certificate, then it has a certificate with $r \leq (n+1)^{d/2}$.

Proof.

Follows from (2):

$$f(\mathbf{x}) = \left\| B \cdot (1, \mathbf{x})^{\otimes \frac{d}{2}} \right\|^2.$$



Efficient Verifiability

Lemma (Bit-complexity of SoS proofs)

Suppose f has a degree- d SoS certificate over $\{-1, 1\}^n$. Then, we can find a degree- d SoS certificate for $f + \varepsilon$ in time $\text{poly}(n^d, \log 1/\varepsilon, \text{size}(f))$.

- ▶ $\{-1, 1\}^n$ is important, and doesn't necessarily hold for other domains.

Proof

Since we are given f , we know that it can be efficiently represented. Therefore, we try to bound the entries of A in terms of f .

- $f(\mathbf{x}) = \left\langle (1, \mathbf{x})^{\otimes \frac{d}{2}}, A \cdot (1, \mathbf{x})^{\otimes \frac{d}{2}} \right\rangle$, for some A .
- $f = \sum_u \hat{f}_u \mathbf{x}_u$, where $\mathbf{x}_u = \prod_{i \in u} \mathbf{x}_i$.
- Expanding the inner product, we see $\hat{f}_u = \sum_{S, T} A_{S, T}$, such that $\text{odd}(S + T) = u$, and $|S|, |T| \leq d/2$.

$$\hat{f}_{\emptyset} = \sum_S A_{S, S} = \text{tr}(A) = \sum_i \underbrace{\lambda_i(A)}_{\geq 0}.$$

$$\|A\|_F^2 = \sum_{S, T} A_{S, T}^2 = \sum_i \lambda_i^2(A) \leq \hat{f}_{\phi}^2.$$

We do not know if entries of A are rational, therefore, above proof doesn't suffice. We now try to find an A .

Connection between SDP and SoS: Find A

Proof Cont...

Recall

$$f(\mathbf{x}) = \left\langle \underbrace{(1, \mathbf{x})^{\otimes \frac{d}{2}}, A \cdot (1, \mathbf{x})^{\otimes \frac{d}{2}}}_{\stackrel{\text{def}}{=} g_A(\mathbf{x})} \right\rangle,$$

Then, we form the following constraints:

1. $A \succcurlyeq 0$.
2. $\text{mult}(g_A(\mathbf{x})) = \text{mult}(f(\mathbf{x}))$.

Therefore, we get the following SDP feasibility problem:

$$\forall u \subseteq [n] : \hat{f}_u = \sum_{\text{odd}(S+T)=u} A_{S,T}; \quad \left((n+1)^d \text{constraints} \right)$$
$$A \succcurlyeq 0.$$

It is unknown if we can decide the feasibility of this system.

Therefore, we try to solve it approximately.

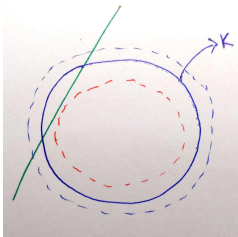
Tools for Approximately Solving SDP

Definition (Weak Separation Oracle)

Let $K \subseteq \mathbb{R}^N$ be a convex set. Weak Separation Oracle:

- Input: Rational vector $\mathbf{x} \in \mathbb{R}^N$, and $\varepsilon > 0$.
- Output: Either
 - Correctly asserts that $\mathbf{x} \in K + \mathcal{B}(0, \varepsilon)$, or,
 - Returns an “almost separating hyperplane”, i.e., returns $\mathbf{y} \neq \mathbf{0} \in \mathbb{R}^N$, such that

$$\langle \mathbf{y}, \mathbf{x} \rangle > \langle \mathbf{y}, \mathbf{z} \rangle - \varepsilon \|\mathbf{y}\|_2, \forall \mathbf{z} \in K.$$



Tools for Approximately Solving SDP

Theorem (Grötschel, Lovász, Schrijver '81)

Let K be a closed, convex, and bounded set. Suppose there exists $R > r > 0$, such that $\mathcal{B}(\mathbf{p}, r) \subseteq K \subseteq \mathcal{B}(\mathbf{0}, R)$. Assume that we have a poly-time weak-separation oracle for K . Then given any rational vector $\mathbf{v} \in \mathbb{R}^N$, we can compute a rational vector $\mathbf{x} \in \mathbb{R}^N$ such that

1. $\mathbf{x} \in K$.
2. $\langle \mathbf{v}, \mathbf{x} \rangle \geq \langle \mathbf{v}, \mathbf{z} \rangle - \varepsilon, \forall \mathbf{z} \in K$.

Running time: $\text{poly}(\log R/r + \log 1/\varepsilon + N)$.

- Interpreting theorem: If I have a convex set K with non-empty interior, with a weak separation oracle, then I can approximately maximize $\mathbf{v}^\top \mathbf{x}$ over K .

Proof Cont...

Applying this to our problem. We define the following:

$$S = \left\{ A \mid A \succcurlyeq 0, \langle C_i, A \rangle = b_i, \forall i, \|A\|_F^2 \leq \hat{f}_\emptyset^2 \right\}.$$

We note that S is convex, bounded, closed. Now,

$$\mathcal{B}(\mathbf{p}, r) \not\subseteq S.$$

Therefore, relax the equality constraints. And find a point in S' ,

$$S' = \left\{ A \mid A \succcurlyeq 0, \langle C_i, A \rangle = [b_i - \varepsilon, b_i + \varepsilon], \forall i, \|A\|_F^2 \leq \hat{f}_\emptyset^2 \right\}.$$

Now, $\mathcal{B}(\mathbf{p}, r) \subseteq S'$ because for any point $A \in S$, then $A + \delta I \in S'$ for δ small enough.

Applying the Theorem

Proof Cont...

- We can find some f' such that f' has a degree- d SoS certificate and $\left| \hat{f}_u - \hat{f}'_u \right| \leq \varepsilon$.
- Note: $f(\mathbf{x}) = f'(\mathbf{x}) + (f - f')(\mathbf{x})$.
- Small coefficient: $\sum_{|u| \leq d} \left| \hat{f}_u - \hat{f}'_u \right| \leq \varepsilon(n+1)^d$.
- Let $L = \sum_{|u| \leq d} \left| \hat{f}_u - \hat{f}'_u \right|$.
- Then, $L + f - f'$ has a degree d -SoS certificate.
- Then, it implies, $L + f$ has a degree- d SoS certificate, i.e., $\varepsilon(n+1)^d + f$ has a degree- d SoS certificate.
- $\varepsilon = \mathcal{O}(n^{-d})$ finishes the proof.



$L + f - f'$ has a degree d -SoS certificate

Proof.

Claim

Let $f = \sum_{|S| \leq d} \hat{f}_S \mathbf{x}_S$. Then $(1 - \mathbf{x}_S)$ and $(1 + \mathbf{x}_S)$ has a degree- d SoS certificate.

- $f = \sum_S |\hat{f}_S| \left(\text{sign}(\hat{f}_S) \mathbf{x}_S \right).$
- By **claim**: $\sum_S |\hat{f}_S| \left(1 + \text{sign}(\hat{f}_S) \mathbf{x}_S \right)$ has a degree- d SoS certificate.

$$\begin{aligned} \sum_S |\hat{f}_S| \left(1 + \text{sign}(\hat{f}_S) \mathbf{x}_S \right) &= \sum_S |\hat{f}_S| + \sum_S |\hat{f}_S| \text{sign}(\hat{f}_S) \mathbf{x}_S \\ &= \sum_S |\hat{f}_S| + f. \end{aligned}$$



Proof of Claim

Proof.

- Let $|S| \leq d$.
- $S = T_1 \cup T_2$, $|T_1| \leq |T_2| \leq d/2$.
- Then $\mathbf{x}_S = \mathbf{x}_{T_1} \cdot \mathbf{x}_{T_2}$.

$$\begin{aligned}(\mathbf{x}_{T_1} - \mathbf{x}_{T_2})^2 &= \mathbf{x}_{T_1}^2 + \mathbf{x}_{T_2}^2 - 2\mathbf{x}_{T_1}\mathbf{x}_{T_2} \\ &= 2 - 2\mathbf{x}_{T_1}\mathbf{x}_{T_2} \\ \therefore (1 - \mathbf{x}_S) &= \frac{1}{2}(\mathbf{x}_{T_1} - \mathbf{x}_{T_2})^2.\end{aligned}$$

- Similarly for $(1 + \mathbf{x}_S)$.



Halfway Through

What if Degree- d SoS Certificate Doesn't Exist for f ?

In that case

1. $\exists \mathbf{x}$, such that $f(\mathbf{x}) < 0$, or,
 2. If $d \leq 2n$, then f may be non-negative and yet a degree- d SoS certificate doesn't exist.
- ▶ **Ideally**, if f does not have a degree- d SoS certificate, we would like the “algorithm” to output an \mathbf{x} such that $f(\mathbf{x}) < 0$.
 - ▶ However, that may not always be possible.
 - ▶ To achieve that **ideal** aim, we construct an object called *Pseudo-distribution*.

Towards Constructing Pseudo-distribution

Fact

The set $SoS_d \subseteq \mathbb{R}^{2^n}$, where

$$SoS_d \stackrel{\text{def}}{=} \{f \mid f \text{ has a degree-}d \text{ SoS certificate}\} ,$$

is a closed, convex cone.

Theorem (Hyperplane Separation Theorem)

Suppose $K \subseteq \mathbb{R}^N$ is a convex set. Let $\mathbf{v} \notin K$. Then there exists a hyperplane $\mathcal{H} = \{\mathbf{x} \mid \langle \mathbf{u}, \mathbf{x} \rangle \geq 0\}$, such that $K \subseteq \mathcal{H}$, and $\mathbf{v} \notin \mathcal{H}$.

Towards Constructing Pseudo-distribution

Suppose $p \notin \text{SoS}_d$. Then, there exists μ such that p is on one side of the hyperplane (defined by μ) and SoS_d on the other side, i.e.,

$$\sum_{\mathbf{x} \in \{-1,1\}^n} \mu(\mathbf{x}) \cdot p(\mathbf{x}) < 0,$$

$$\sum_{\mathbf{x} \in \{-1,1\}^n} \mu(\mathbf{x}) \cdot f(\mathbf{x}) \geq 0, \quad \forall f \in \text{SoS}_d$$

$$\sum_{\mathbf{x} \in \{-1,1\}^n} \mu(\mathbf{x}) = 1, \quad (\text{by scaling}).$$

- Hypothetical: Suppose $\mu \geq 0$, then it describes a probability distribution over $\{-1, 1\}^n$.
- . Therefore, there exists a distribution such that for $p \notin \text{SoS}_d$

$$\mathbb{E}_{\mu} p < 0,$$

$$\mathbb{E}_{\mu} f \geq 0, \quad \forall f \in \text{SoS}_d.$$

Pseudo-distribution

- Notation: Pseudo-expectation (when μ is not non-negative)

$$\tilde{\mathbb{E}}_{\mu} f = \sum_{\mathbf{x} \in \{-1, 1\}^n} \mu(\mathbf{x}) \cdot f(\mathbf{x}).$$

Definition (Pseudo-distribution)

A degree- d pseudo-distribution over $\{-1, 1\}^n$ is a function $\mu : \{-1, 1\}^n \rightarrow \mathbb{R}$ such that $\tilde{\mathbb{E}}_{\mu}$ satisfies:

1. $\tilde{\mathbb{E}}_{\mu} \mathbf{1} = 1$.
2. $\forall f : \deg(f) \leq \frac{d}{2}, \tilde{\mathbb{E}}_{\mu} f^2 \geq 0$.

Fact

Every degree $\geq 2n$ pseudo-distribution μ is an actual probability distribution.

Proof Sketch.

- Define indicator polynomial $f_{\mathbf{y}} : \{-1, 1\}^n \rightarrow \mathbb{R}$, such that $f_{\mathbf{y}}(\mathbf{y}) = 1$, and $f_{\mathbf{y}}(\mathbf{x}) = 0$, $\forall \mathbf{x} \neq \mathbf{y}$. Moreover, $\deg(f) \leq n$.
- By definition $\tilde{\mathbb{E}}_{\mu} f_{\mathbf{y}}^2 \geq 0$.
- Construct the distribution $\text{prob}(\mathbf{y}) \stackrel{\text{def}}{=} \tilde{\mathbb{E}}_{\mu} f_{\mathbf{y}}^2$. And this distribution is the pseudo-distribution μ .



Specifying Pseudo-distributions

Pseudo-distributions can be specified as a vector $\mu' \in \mathbb{R}^{n^d}$.

Claim

For all degree- d pseudo-distributions, there exists a degree- d multi-linear polynomial $\mu' : \{-1, 1\}^n \rightarrow \mathbb{R}$, such that $\tilde{\mathbb{E}}_{\mu} p = \tilde{\mathbb{E}}_{\mu'} p$ for all p such that $\deg(p) \leq d$.

Proof Sketch.

Write μ and p in multilinear form.

$$\begin{aligned}\mu(\mathbf{x}) &= \sum_{S \subseteq [n]} \hat{\mu}_S \mathbf{x}_S \\ p(\mathbf{x}) &= \sum_{S \subseteq [n], |S| \leq d} \hat{p}_S \mathbf{x}_S.\end{aligned}$$

$\tilde{\mathbb{E}}_{\mu} p = \langle \mu, p \rangle = \langle \mu', p \rangle$, where μ' is a degree $\leq d$ part of μ . □

- ▶ Notation- Pseudo-moments: $\tilde{\mathbb{E}}_{\mu}(1, \mathbf{x})^{\otimes d}$.
(Expectation of a vector) expectations of degree $\leq d$ monomials.

Pseudo-distribution and PSD Matrices

Proposition

μ is a degree- d pseudo-distribution iff

1. $\tilde{\mathbb{E}}_{\mu} \mathbf{1} = 1$,
2. $\underbrace{\tilde{\mathbb{E}}_{\mu} (1, \mathbf{x})^{\otimes \frac{d}{2}} \left((1, \mathbf{x})^{\otimes \frac{d}{2}} \right)^{\top}}_{\text{degree-}d \text{ pseudo-moment matrix}} \succcurlyeq 0$.

- Parsing notation: The S, T -th entry of the pseudo-moment matrix is $\tilde{\mathbb{E}}_{\mu} \mathbf{x}_S \mathbf{x}_T = \tilde{\mathbb{E}}_{\mu} \mathbf{x}_{\text{odd}(S+T)}$.

Proof.

- Let f be a degree- $d/2$ polynomial.
- $\tilde{\mathbb{E}}_\mu f^2 = c \left(\hat{f} \right)^\top M_{d/2} \hat{f}$, because

$$\begin{aligned}\tilde{\mathbb{E}}_\mu f^2 &= \tilde{\mathbb{E}}_\mu \left(\sum_S \hat{f}_S \mathbf{x}_S \right)^2 = \tilde{\mathbb{E}}_\mu \sum_{S,T} \hat{f}_S \hat{f}_T \mathbf{x}_S \mathbf{x}_T \\ &= \sum_{S,T} \hat{f}_S \hat{f}_T \tilde{\mathbb{E}}_\mu \mathbf{x}_S \mathbf{x}_T.\end{aligned}$$

- But since f^2 is a degree- d SoS, the quadratic form is positive, and therefore, $M_{d/2} \succcurlyeq 0$.



Formal Pseudo-expectation Proof

Theorem

For every f , every even $d \in \mathbb{Z}_{\geq 0}$, there exists a degree- d SoS certificate of f iff

$$\forall \text{ degree-}d \text{ pseudo-distribution over } \{-1, 1\}^n, \tilde{\mathbb{E}}_{\mu} f \geq 0.$$

Proof

- If Part: If f has a degree- d SoS certificate,
 $\tilde{\mathbb{E}}_{\mu} f = \sum_{i \leq r} \tilde{\mathbb{E}}_{\mu} g_i^2 \geq 0.$
- Only If Part: Suppose $f \notin \text{SoS}_d$, then there exists a hyperplane with μ as the normal vector such that

$$\begin{aligned} \tilde{\mathbb{E}}_{\mu} f &< 0, \\ \tilde{\mathbb{E}}_{\mu} g^2 &\geq 0, \quad \forall g \text{ such that } \deg(g) \leq d/2. \end{aligned}$$

Need: $\tilde{\mathbb{E}}_{\mu} \mathbf{1} > 0.$

$$\tilde{\mathbb{E}}_{\mu} \mathbf{1} > 0$$

Proof Cont...

- We know, $\exists L > 0$ such that $L + f$ has a degree- d SoS certificate. We have $\tilde{\mathbb{E}}_{\mu} f < 0$, and

$$\tilde{\mathbb{E}}_{\mu}(L + f) \geq 0$$

$$\tilde{\mathbb{E}}_{\mu} L \geq -\tilde{\mathbb{E}}_{\mu} f$$

$$L \tilde{\mathbb{E}}_{\mu} \mathbf{1} \geq -\tilde{\mathbb{E}}_{\mu} f$$

$$\tilde{\mathbb{E}}_{\mu} \mathbf{1} \geq \frac{-\tilde{\mathbb{E}}_{\mu} f}{L} > 0.$$



Pseudo-distribution and SDP Connection

Theorem

Suppose f does not have a degree- d SoS certificate. Then, there exists a poly $\left(n^d, \log 1/\varepsilon, \text{size}(f)\right)$ -time algorithm to compute a pseudo-distribution μ such that $\tilde{\mathbb{E}}_\mu f < \varepsilon$.

Proof Sketch.

Form a SDP system- Make variables for all possible monomials:
 $\tilde{\mathbb{E}}_\mu \mathbf{x}_S, |S| \leq d$.

- Constraint 1: $\tilde{\mathbb{E}}_\mu \mathbf{1} = 1$.
- Constraint 2: $\tilde{\mathbb{E}}_\mu (1, \mathbf{x})^{\otimes \frac{d}{2}} \left((1, \mathbf{x})^{\otimes \frac{d}{2}} \right)^\top \succeq 0$.
- Constraint 3: $\tilde{\mathbb{E}}_\mu f < 0$.

Use SDP solving up to slack ε to obtain the theorem statement.



SoS Algorithms

Algorithms for computing
SoS certificates, and Pseudo-distributions
are called *SoS algorithms*.