DSOS and SDSOS: More Tractable Alternatives to SOS

Amir Ali Ahmadi

Goldstine Fellow, IBM Watson Research Center

Joint work with:

Anirudha Majumdar

MIT

SOS Optimization—The Good

- Arguably, the strongest tool we have to tackle a fundamental problem in computational mathematics:

optimization over nonnegative polynomials

- Essentially *any* semialgebraic problem can be reduced to it
- Diverse applications:
  - Polynomial optimization, control and robotics, combinatorial optimization, approximation algorithms, robust optimization, quantum computation, continuous games, filter design, automated theorem proving,…, …, ….
- Beautiful theory:
  - Classical problems in algebra revisited through the lens of computation
  - Strongest known bounds for several intractable problems

SOS Optimization—The Bad

- **Scalability** is a pain in the (_I_)

Thm: ***p(x)*** of degree ***2d*** is sos if and only if

$$p(x) = z^T Q z \quad Q \succeq 0$$

$$z = [1, x_1, x_2, \ldots, x_n, x_1 x_2, \ldots, x_n^d]^T$$

- The size of the Gram matrix is:

$$\binom{n+d}{d} \times \binom{n+d}{d}$$

- Polynomial in n for fixed d, but grows quickly
  - **The semidefinite constraint is expensive**
- E.g., a quartic in 50 variables has 316251 coefficients and the Gram matrix (that needs to be psd) has 879801 entries…

SOS Optimization—~~The Ugly~~ The Hope

Make SOS scalable through a number of interesting approaches:

- Techniques for exploiting structure (e.g., symmetry and sparsity)
  - [Gatermann, Parrilo], [Vallentin], [de Klerk, Sotirov], [Riener et al.], …
- Customized algorithms (e.g., first order methods)
  - [Bertsimas, Freund, Sun], [Nie, Wang], …

Our approach:

- Let's not work with SOS to begin with…
- Give other sufficient conditions for nonnegativity that are **perhaps stronger than SOS, but hopefully cheaper**

Not totally clear a priori how to do this…

Consider, e.g., the following two sets:

1) All polynomials that are **sums of 3 squares of polynomials**
2) All polynomials that are **sums of 4th powers of polynomials**

Both sets are clearly inside the SOS cone

- But linear optimization over either set is intractable!
- So set inclusion doesn't mean anything in terms of complexity
- We have to work a bit harder…

dsos and sdsos

Defn. A polynomial ***p*** is *diagonally-dominant-sum-of-squares* ***(dsos)*** if it can be written as:

$$p = \sum_i \alpha_i m_i^2 + \sum_{i,j} \beta_{ij}(m_i \pm m_j)^2,$$

for some monomials m_i, m_j
and some nonnegative constants $\alpha_i, \beta_{i,j}$.

Defn. A polynomial ***p*** is *scaled-diagonally-dominant-sum-of-squares* ***(sdsos)*** if it can be written as:

$$p = \sum_i \alpha_i m_i^2 + \sum_{i,j} (\beta_i m_i \pm \gamma_j m_j)^2,$$

for some monomials m_i, m_j
and some constants $\alpha_i \geq 0, \beta_i, \gamma_i$.

Obvious: $DSOS_{n,d} \subseteq SDSOS_{n,d} \subseteq SOS_{n,d} \subseteq POS_{n,d}$

r-dsos and r-sdsos

Defn. A polynomial p is *r-diagonally-dominant-sum-of-squares* ***(r-dsos)*** if

$$p \cdot \left(\sum_i x_i^2\right)^r$$

is dsos.

Defn. A polynomial p is *r-scaled-diagonally-dominant-sum-of-squares* ***(r-sdsos)*** if

$$p \cdot \left(\sum_i x_i^2\right)^r$$

is sdsos.

Easy: $rDSOS_{n,d} \subseteq rSDSOS_{n,d} \subseteq POS_{n,d}, \forall r.$

dd and sdd matrices

Defn. A symmetric matrix A is *diagonally dominant* ***(dd)*** if

$$a_{ii} \geq \sum_{j \neq i} |a_{ij}| \ for \ all \ i.$$

Defn*. A symmetric matrix A is *scaled diagonally dominant* ***(sdd)*** if there exists a diagonal matrix $D>0$ s.t.

DAD is dd.

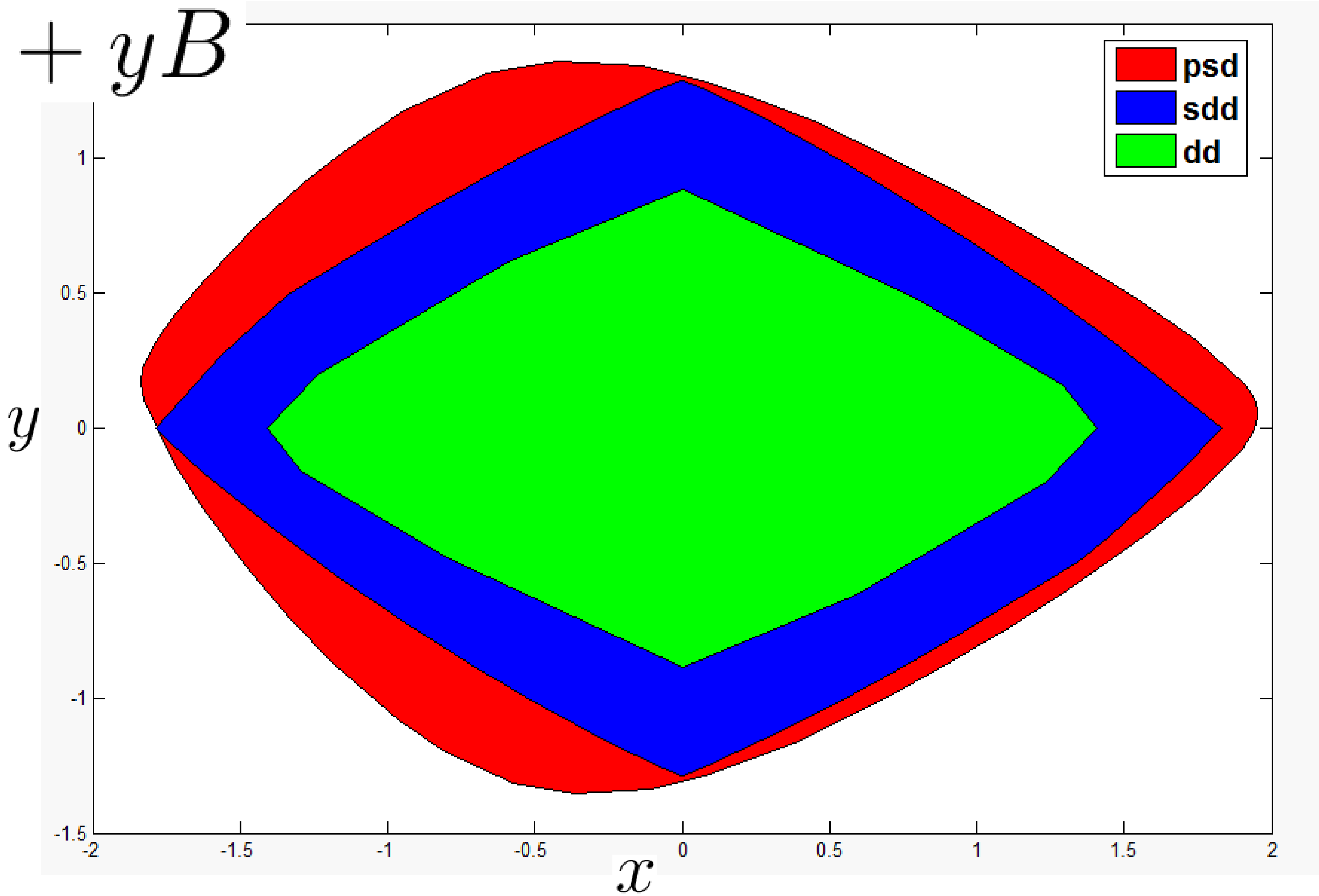
$$dd \Rightarrow sdd \Rightarrow psd$$

Greshgorin's circle theorem

*Thanks to Pablo Parrilo for telling us about sdd matrices.

$$I + xA + yB$$

A, B
10×10
random



Optimization over these sets is an SDP, SOCP, LP !!

(code courtesy of Pablo Parrilo)

Two natural matrix programs: DDP and SDPP

LP:

$$\min\langle C, X\rangle$$
$$A(X) = b$$
$$X \quad \text{diagonal\&nonnegative}$$

DDP:

$$\min\langle C, X\rangle$$
$$A(X) = b$$
$$X \quad \text{dd}$$

SDDP:

$$\min\langle C, X\rangle$$
$$A(X) = b$$
$$X \quad \text{sdd}$$

SDP:

$$\min\langle C, X\rangle$$
$$A(X) = b$$
$$X \succeq 0$$

From matrices to polynomials

Thm. A polynomial ***p*** is ***dsos***

$$p = \sum_i \alpha_i m_i^2 + \sum_{i,j} \beta_{ij} (m_i \pm m_j)^2,$$

if and only if

$$p(x) = z^T(x) Q z(x)$$

$$Q \quad dd$$

Thm. A polynomial ***p*** is ***sdsos***

$$p = \sum_i \alpha_i m_i^2 + \sum_{i,j} (\beta_i m_i \pm \gamma_j m_j)^2,$$

if and only if

$$p(x) = z^T(x) Q z(x)$$

$$Q \quad sdd$$

Optimization over r-dsos and r-dsos polynomials

Thm. For any integer r, the set $\boldsymbol{rDSOS_{n,d}}$ **is polyhedral** and the set $\boldsymbol{rSDSOS_{n,d}}$ **has a second order cone representation.**

- For any fixed r and d, optimization over $rDSOS_{n,d}$ (resp. $rSDSOS_{n,d}$) can be done with an **LP (resp. SOCP) of size polynomial in** $\boldsymbol{n}$.

- Commercial solvers such as CPLEX and GUROBI are very mature (very fast, deal with numerical issues)
- We have automated the generation of these LPs and SOCPs in SPOTless (package by Megretski, Tobenkin –MIT)

How well does it do?!

- We’ll show preliminary experiments
  - From combinatorial optimization, control, copositive programming, polynomial optimization, etc.
- And we’ll give Positivstellensatz results (converse results)

First observation: r-dsos can outperform sos

The Motzkin polynomial:

$$M(x_1, x_2, x_3) = x_1^4 x_2^2 + x_1^2 x_2^4 - 3x_1^2 x_2^2 x_3^2 + x_3^6$$

psd but *not* sos!

…but it's 2-dsos.

(certificate of nonnegativity using LP)

Another ternary sextic:

$$p(x_1, x_2, x_3) = x_1^4 x_2^2 + x_2^4 x_3^2 + x_3^4 x_1^2 - 3x_1^2 x_2^2 x_3^2$$

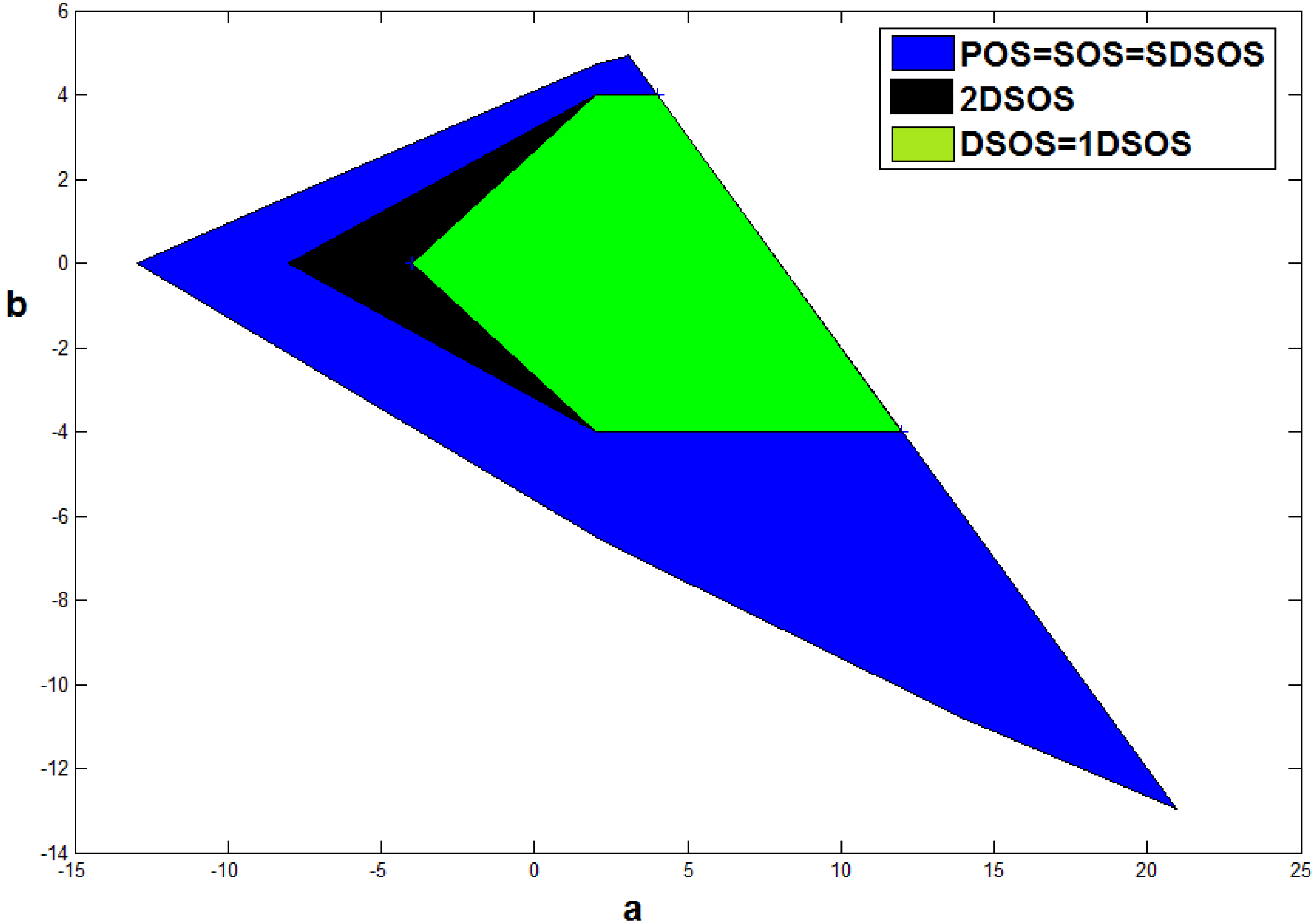
***not* sos but 1-dsos (hence psd)**

A parametric family of polynomials

$$p(x) = 2x_1^4 + cx_2^4 + ax_1^2x_2^2 + bx_1^3x_2$$

Compactify:

$$p(x) = 2x_1^4 + (8-a-b)x_2^4 + ax_1^2x_2^2 + bx_1^3x_2$$



Minimizing a form on the sphere

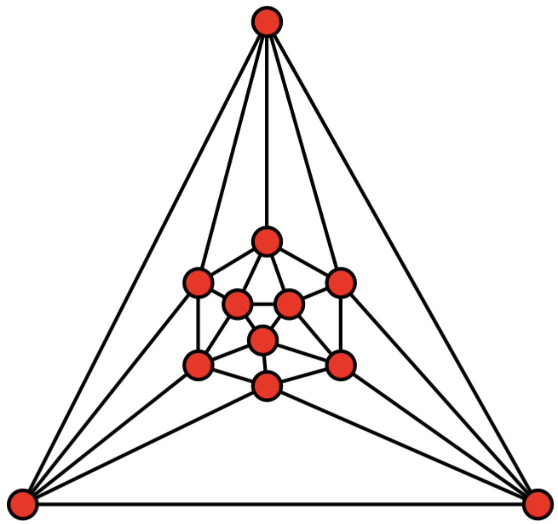
$$\min_{x \in \mathcal{S}^{n-1}} p(x)$$

- degree=4; all coefficients present – generated randomly
- PC: 3.4 GHz, 16 Gb RAM

n=10	Lower bound	Run time (secs)	n=15	Lower bound	Run time (secs)	n=20	Lower bound	Run time (secs)
sos (sedumi)	-1.920	1.01	**sos (sedumi)**	-3.263	165.3	**sos (sedumi)**	-3.579	5749
sos (mosek)	-1.920	0.184	**sos (mosek)**	-3.263	5.537	**sos (mosek)**	-3.579	79.06
sdsos	-5.046	0.152	**sdsos**	-10.433	0.444	**sdsos**	-17.333	1.935
dsos	-5.312	0.067	**dsos**	-10.957	0.370	**dsos**	-18.015	1.301
BARON	-175.4	0.35	**BARON**	-1079.9	0.62	**BARON**	-5287.9	3.69

n=30	Lower bound	Run time (secs)	n=40	Lower bound	Run time (secs)	n=50	Lower bound	Run time (secs)
sos (sedumi)	---------	∞	**sos (sedumi)**	---------	∞	**sos (sedumi)**	---------	∞
sos (mosek)	---------	∞	**sos (mosek)**	---------	∞	**sos (mosek)**	---------	∞
sdsos	-36.038	9.431	**sdsos**	-61.248	53.95	**sdsos**	-93.22	100.5
dsos	-36.850	8.256	**dsos**	-62.2954	26.02	**dsos**	-94.25	72.79
BARON	-28546.1							

Maximum Clique



Can be reformulated (via Motzkin-Straus) as a copositive program
→ positivity of a quartic

Polynomial optimization problem in 12 variables

Upper bound on max clique:

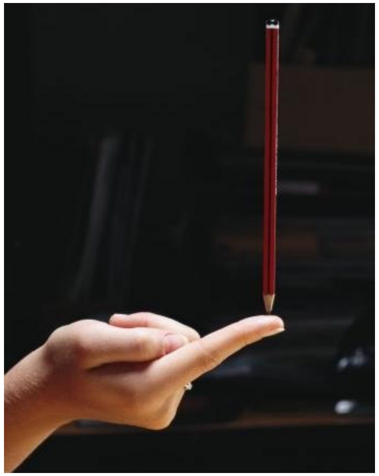
- **dsos:** 6.0000 ⟷
- **sdsos:** 6.0000
- **1-dsos:** 4.3333 ⟷
- **1-sdsos:** 4.3333
- **2-dsos:** 3.8049 ⟷
- **2-sdsos:** 3.6964
- **sos:** 3.2362

LP of de Klerk and Pasechnik

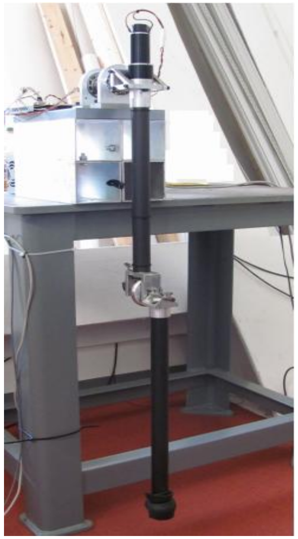
- **Level 0:** ∞
- **Level 1:** ∞
- **Level 2:** 6.0000

- **r-dsos LP *guaranteed* to give exact value for r=(max clique)^2**

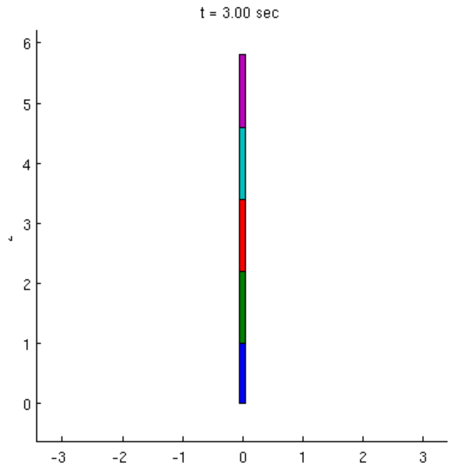
Stabilizing the inverted N-link pendulum (2N states)



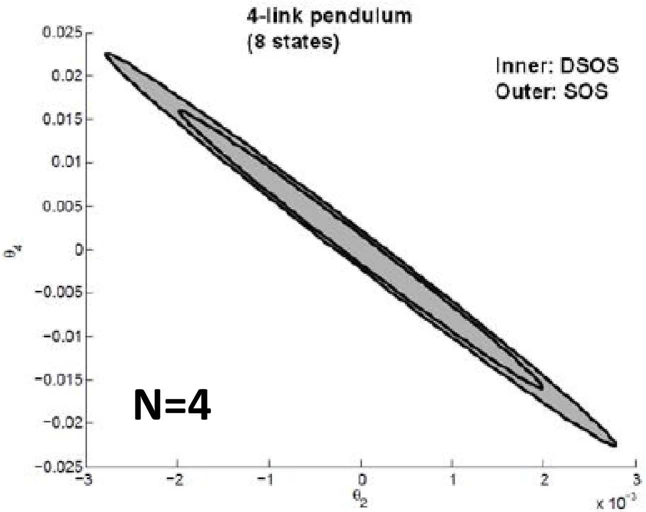
N=1

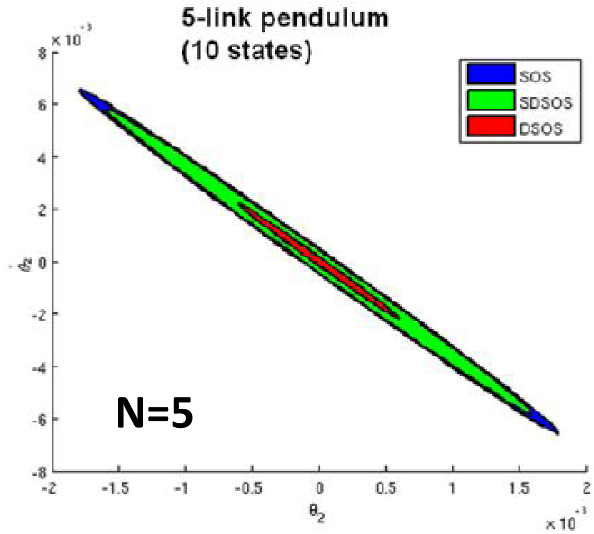


N=2



N=5





$$\left(\frac{Vol^{1/dim}_{\text{ROA-sdsos}}}{Vol^{1/dim}_{\text{ROA-sos}}}\right) = 78.54\%$$

Running time (N=5):

dsos: 37 secs   sdsos: 40 secs   sos: 1.6 hours   (speed-up: x150)

Converse results 1&2

Thm. Any **even** positive definite form $\boldsymbol{p}$ **is r-dsos** for some $\boldsymbol{r}$.

- Hence proof of positivity can always be found with LP
- Even forms (of degree 4) include **copositive programming**!
- Proof follows from a result of **Polya (1928)** on Hilbert's 17th problem:
  - $p.||x||^r$ will always become a sum of squares of monomials for sufficiently large r.

Thm. Any positive definite **bivariate** form $\boldsymbol{p}$ **is r-sdsos** for some $\boldsymbol{r}$.

- Proof follows from a result of **Reznick (1995)**
  - $p.||x||^r$ will always become a sum of powers of linear forms for sufficiently large r.

Converse result 3

Thm. For any positive definite form $\boldsymbol{p}$, there exists an integer r and a **polynomial $\boldsymbol{q}$ of degree $\boldsymbol{r}$** such that

$\boldsymbol{q}$ is dsos

and

$\boldsymbol{pq}$ is dsos.

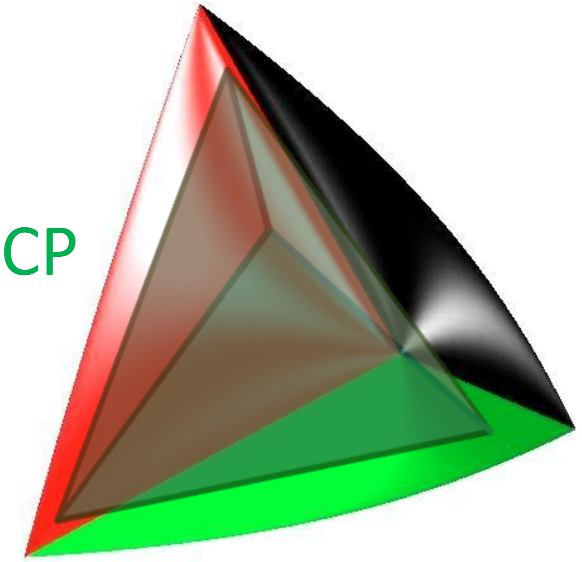
- Search for q is an LP
- Such a q is a certificate of nonnegativity of p
- Proof follows from a result of **Habicht (1940)** on Hilbert's 17th problem:
    - Every positive definite form is a quotient of two sums of squares of monomials

The general polynomial optimization problem (POP)

$$\min_{x} p(x)$$
$$f_i(x) \leq 0$$
$$h_i(x) = 0$$

- Similar to the **Lasserre/Parrilo hierarchies of SDP** (based on Putinar, Schmüdgen or Stengle's Positivstellensatz) that solve POP to global optimaility,
  our converse results imply that **POP can be solved to global optimality using hierarchies of LP and SOCP coming from dsos and sdsos.**

Main messages…



- Inner approximations to SOS cone and move away from SDP towards LP and SOCP
- Many of the guarantees still go through!
- This can be used *anywhere* SOS is used!
- Potential to integrate with CVXGEN for real-time algebraic optimization
- There is a range of problems with 20-100 variables that are within reach

- **60 variables is becoming a piece of (birthday) cake!**
  - Happy birthday JB Lasserre!!

Thank you for your attention!

Questions?

Want to know more?

http://aaa.lids.mit.edu/