LP and SOCP-based Approximations to SOS

(initially funded by AFOSR YIP)

Amir Ali Ahmadi
Princeton, ORFE
Affiliated member of PACM, COS, MAE, CSML

MURI Kick-off Meeting
AFOSR, February 2019

Optimization over nonnegative polynomials

Is $p(x) \geq 0$ on $\{g_1(x) \geq 0, \ldots, g_m(x) \geq 0\}$?

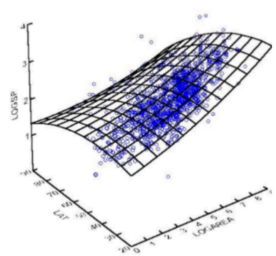
Optimization

- Lower bounds on polynomial optimization problems

$$\min_{x} \quad p(x)$$
$$\text{s.t. } g_i(x) \geq 0$$

Statistics/ML

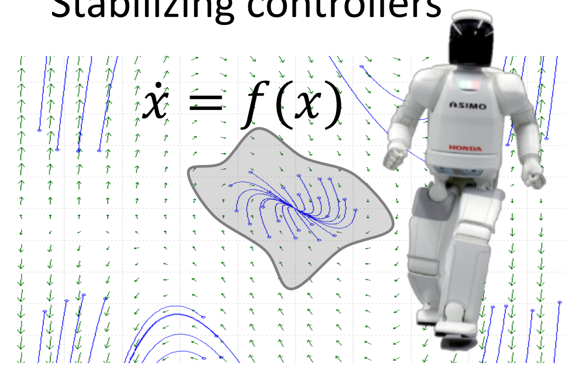
- Fitting a polynomial to data subject to shape constraints (e.g., convexity, or monotonicity)



$$\frac{\partial p(x)}{\partial x_j} \geq 0, \forall x \in B$$

Control

- Stabilizing controllers



$V(x) > 0,$
$V(x) \leq \beta \Rightarrow \nabla V(x)^T f(x) < 0$

Implies that
$\{x \mid V(x) \leq \beta\}$
is in the region of attraction

How to prove nonnegativity?

$$\begin{aligned} p(x) \;&=\; x_1^4 - 6x_1^3x_2 + 2x_1^3x_3 + 6x_1^2x_3^2 + 9x_1^2x_2^2 - 6x_1^2x_2x_3 \\ &\quad -14x_1x_2x_3^2 + 4x_1x_3^3 + 5x_3^4 - 7x_2^2x_3^2 + 16x_2^4 \end{aligned}$$

Nonnegative

$\Uparrow$

$$\begin{aligned} p(x) \;&=\; (x_1^2 - 3x_1x_2 + x_1x_3 + 2x_3^2)^2 + (x_1x_3 - x_2x_3)^2 \\ &\quad +(4x_2^2 - x_3^2)^2. \end{aligned}$$

SOS

- Optimization over sum of squares (SOS) polynomials is a semidefinite program (SDP)!

- SDPs can be solved in poly-time to arbitrary accuracy.

Practical limitations of SOS

- **Scalability** is a nontrivial challenge!

A polynomial p of degree $2d$ is SOS if and only if $\exists Q \succcurlyeq 0$ such that

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

where $z = [1, x_1, \dots, x_n, x_1 x_2, \dots, x_n^d]^T$ is the vector of monomials of degree up to d.

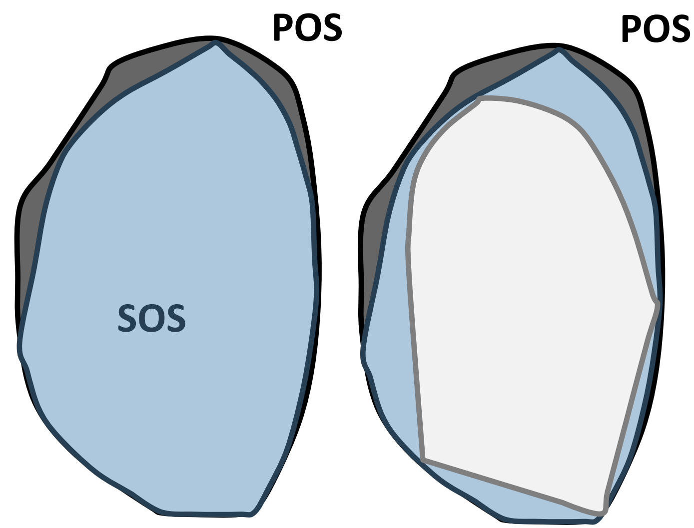
- The size of the Gram matrix is:

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

- Polynomial in n for fixed d, but grows quickly
  - **The semidefinite constraint is expensive**
- E.g., local stability analysis of a 20-state cubic vector field is typically an SDP with ~1.2M decision variables and ~200k constraints

Simple idea…

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



Not any set inside SOS would work!

1) **sums of 4^{th} powers of polynomials**
2) **sums of 3 squares of polynomials**

Both sets are clearly inside the SOS cone,
but linear optimization over them is **intractable**.

[AAA, Majumdar, *SIAM J. on Applied Algebra and Geometry*]

PRINCETON UNIVERSITY ORFE

dsos and sdsos polynomials (1/3)

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

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

for some monomials m_i, m_j
and some nonnegative constants $\alpha_i, \beta_{ij}^+, \beta_{ij}^-$.

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

$$p(x) = \sum_i \alpha_i m_i^2(x) + \sum_{i,j} (\hat{\beta}_{ij}^+ m_i(x) + \tilde{\beta}_{ij}^+ m_j(x))^2 + \sum_{i,j} (\hat{\beta}_{ij}^- m_i(x) - \tilde{\beta}_{ij}^- m_j(x))^2,$$

for some monomials m_i, m_j
and some constants $\alpha_i, \hat{\beta}_{ij}^+, \tilde{\beta}_{ij}^+, \hat{\beta}_{ij}^-, \tilde{\beta}_{ij}^-$ *with* $\alpha_i \geq 0.$

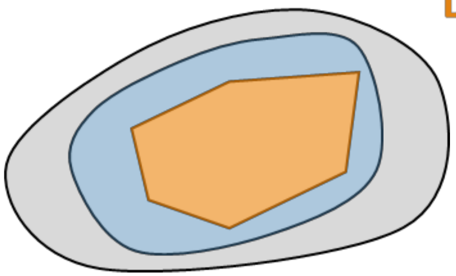
Note: $DSOS_{n,d} \subseteq SDSOS_{n,d} \subseteq SOS_{n,d} \subseteq POS_{n,d}$

dsos and sdsos polynomials (2/3)

Sum of squares **(sos)**	$p(x) = z(x)^T Q z(x), Q \succcurlyeq 0$	**SDP**

DD cone $:= \{Q \mid Q_{ii} \geq \sum_{j \neq i} \lvert Q_{ij} \rvert, \forall i\}$

PSD cone$:= \{Q \mid Q \succcurlyeq 0\}$



SDD cone $:= \{Q \mid \exists$ **diagonal** D **with** $D_{ii} > 0$ **s.t.** DQD $dd\}$

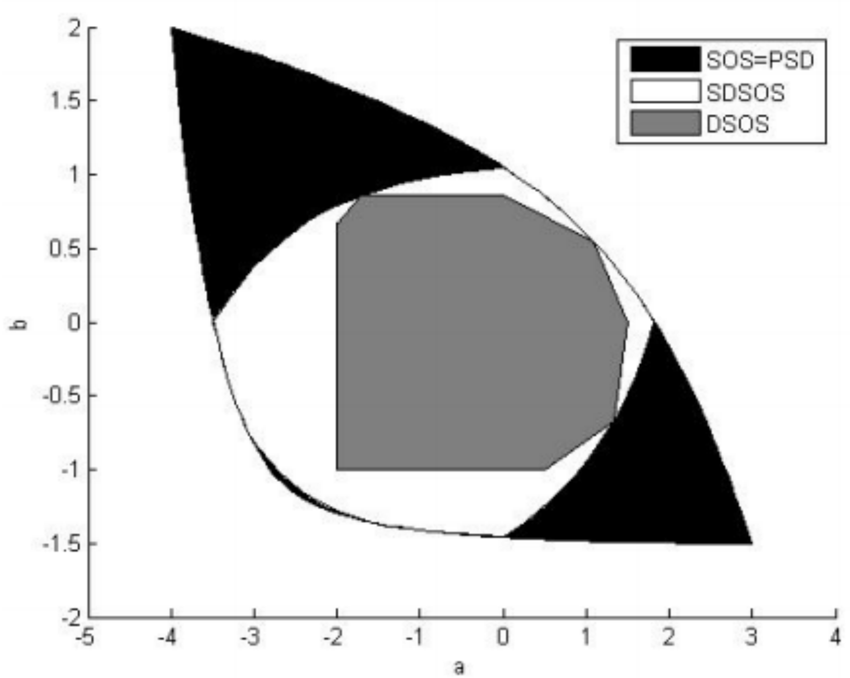
Diagonally dominant sum of squares **(dsos)**	$p(x) = z(x)^T Q z(x), Q$ *diagonally dominant* (dd)	**LP**

Scaled diagonally dominant sum of squares **(sdsos)**	$p(x) = z(x)^T Q z(x), Q$ *scaled diagonally dominant* (sdd)	**SOCP**

PRINCETON UNIVERSITY ORFE

[AAA, Majumdar, *SIAM J. on Applied Algebra and Geometry*]

dsos and sdsos polynomials (3/3)



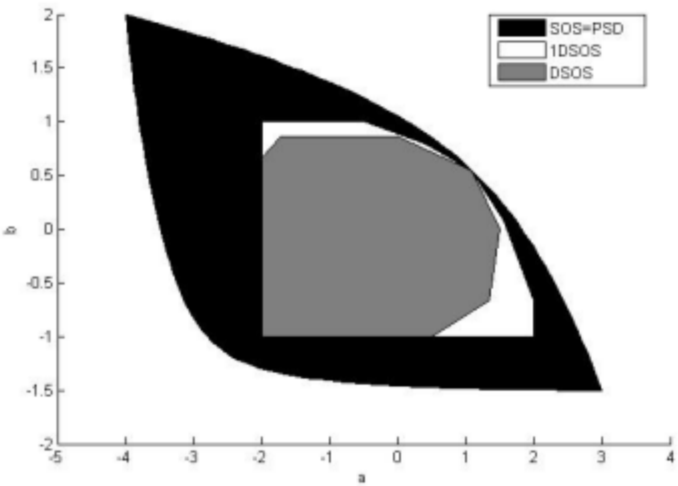
$$p(x_1, x_2) = x_1^4 + x_2^4 + a x_1^3 x_2 + (1 - \frac{1}{2}a - \frac{1}{2}b) x_1^2 x_2^2 + 2b x_1 x_2^3$$

How to do better?

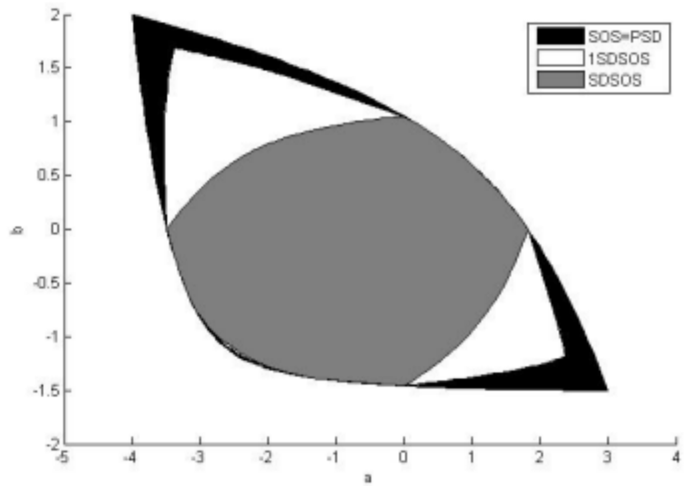
Method #1: r-dsos and r-sdsos polynomials (1/2)

Defn.

- A polynomial p is **r-dsos** if $p(x) \cdot \left(\sum_i x_i^2\right)^r$ is dsos.
- A polynomial p is **r-sdsos** if $p(x) \cdot \left(\sum_i x_i^2\right)^r$ is sdsos.



(a) The LP-based r-dsos hierarchy.



(b) The SOCP-based r-sdsos hierarchy.

$$p(x_1, x_2) = x_1^4 + x_2^4 + a x_1^3 x_2 + (1 - \frac{1}{2}a - \frac{1}{2}b) x_1^2 x_2^2 + 2b x_1 x_2^3$$

Method #1: r-dsos and r-sdsos polynomials (2/2)

- r-dsos can outperform sos!

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

is 1-dsos but not sos.

Theorem: Any even positive definite form is r-dsos for some r.

- Even forms include *copositive programming* *(and all problems in NP)*.

Theorem: Any form can be made even, while preserving positivity, by doubling the number of variables and degree.

- Leads to arbitrarily tight lower bounds on any polynomial optimization problem (with a compact feasible set).

PRINCETON UNIVERSITY ORFE

[AAA, Hall, *Math of OR*] (2018 INFORMS Young Researchers Prize)

Method #2: dsos/sdsos + change of basis (1/2)

$$\begin{aligned} p(x) \;=\;& x_1^4 - 6x_1^3x_2 + 2x_1^3x_3 + 6x_1^2x_3^2 + 9x_1^2x_2^2 - 6x_1^2x_2x_3 - 14x_1x_2x_3^2 + 4x_1x_3^3 \\ & +5x_3^4 - 7x_2^2x_3^2 + 16x_2^4 \end{aligned}$$

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

$$Q = \begin{pmatrix} 1 & -3 & 0 & 1 & 0 & 2 \\ -3 & 9 & 0 & -3 & 0 & -6 \\ 0 & 0 & 16 & 0 & 0 & -4 \\ 1 & -3 & 0 & 2 & -1 & 2 \\ 0 & 0 & 0 & -1 & 1 & 0 \\ 2 & -6 & 4 & 2 & 0 & 5 \end{pmatrix}$$

$$z(x) = (x_1^2, x_1x_2, x_2^2, x_1x_3, x_2x_3, x_3^2)^T$$

$$p(x) = \tilde{z}^T(x) \begin{pmatrix} \frac{1}{2} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 4. \end{pmatrix} \tilde{z}(x)$$

$$\tilde{z}(x) = \begin{pmatrix} 2x_1^2 - 6x_1x_2 + 2x_1x_3 + 2x_3^2 \\ x_1x_3 - x_2x_3 \\ x_2^2 - \frac{1}{4}x_3^2 \end{pmatrix}$$

Goal: iteratively improve $z(x)$

[AAA, Hall, *Contemporary Mathematics*]

Method #2: dsos/sdsos + change of basis (2/2)

LP
$$\max_{\vec{p},Q} \ \ell(\vec{p}) \quad \text{s.t. } p(x) = z^T(x) Q z(x) \ \forall x, \quad Q \text{ dd}$$

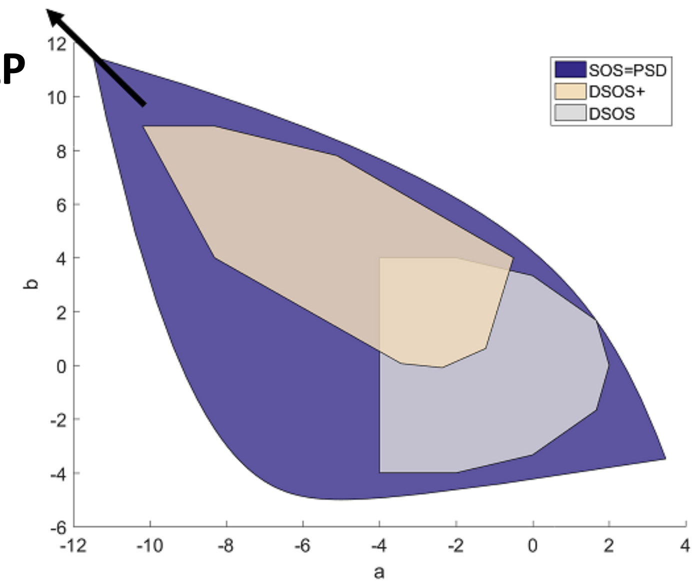
→ Optimal soln. Q^*

→ Cholesky: $Q^* = U^T U$

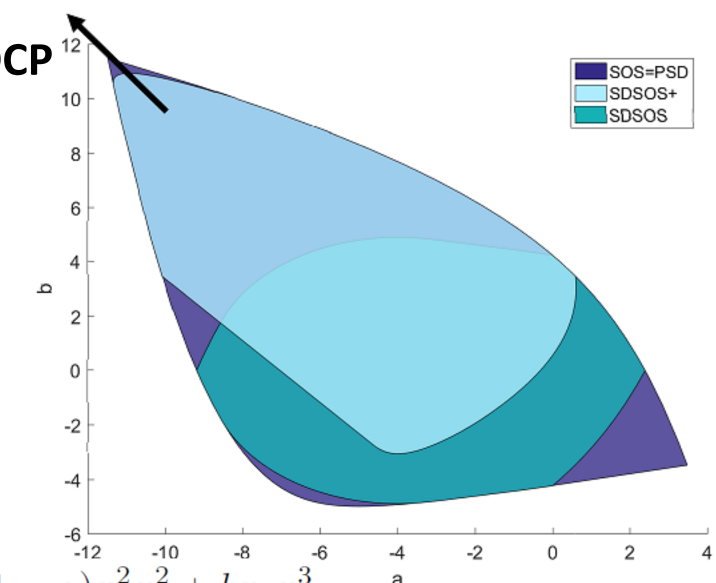
LP_+
$$\max_{\vec{p},Q} \ \ell(\vec{p}) \quad \text{s.t. } p(x) = z^T(x) U^T Q U z(x) \ \forall x, \quad Q \text{ dd}$$

Works beautifully!

LP



SOCP



$$p_{a,b}(x_1, x_2) = 2x_1^4 + 2x_2^4 + a x_1^3 x_2 + (1-a) x_2^2 x_2^2 + b x_1 x_2^3$$

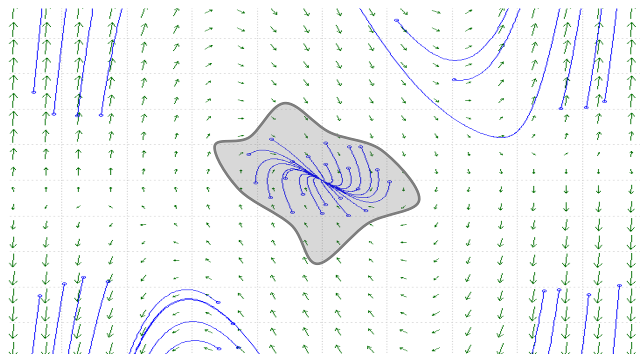
Applications in control

(see paper for applications in statistics, finance, combinatorial and polynomial optimization)

Reminder

$$\dot{x} = f(x, u)$$

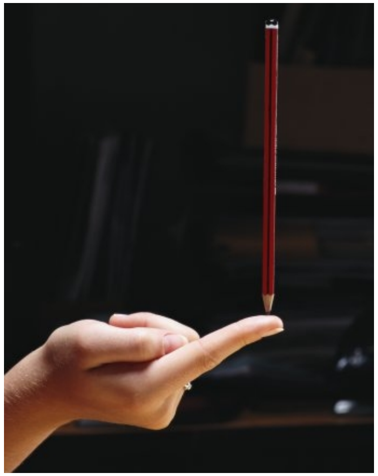
Stability of equilibrium points



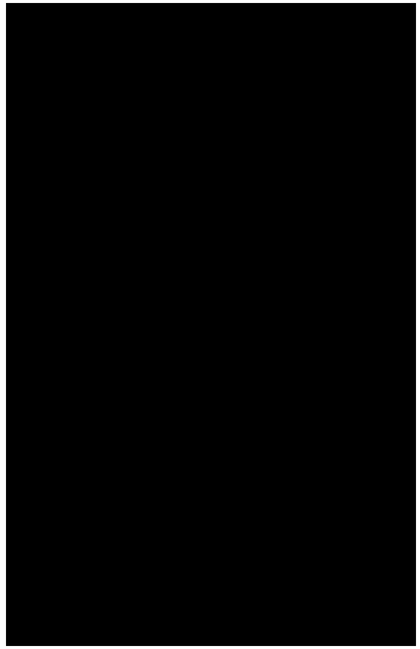
$$V(x) > 0,$$
$$V(x) \leq \beta \Rightarrow \dot{V}(x) < 0$$

implies $\{x \mid V(x) \leq \beta\}$ is in the region of attraction (ROA)

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

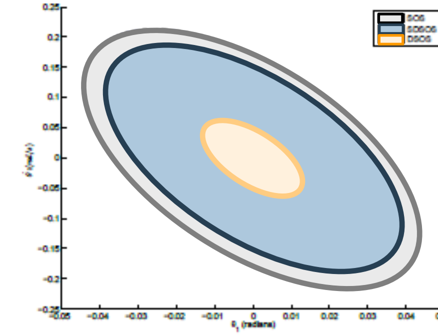


N=1

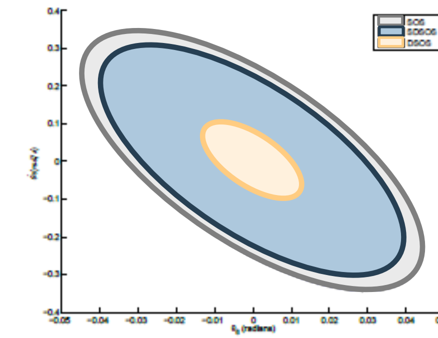


N=2

N=6



(a) θ_1-$\dot{\theta}_1$ subspace.



(b) θ_6-$\dot{\theta}_6$ subspace.

Runtime:

2N (# states)	4	6	8	10	12	14	16	18	20	22
DSOS	< 1	0.44	2.04	3.08	9.67	25.1	74.2	200.5	492.0	823.2
SDSOS	< 1	0.72	6.72	7.78	25.9	92.4	189.0	424.74	846.9	1275.6
SOS (SeDuMi)	< 1	3.97	156.9	1697.5	23676.5	∞	∞	∞	∞	∞
SOS (MOSEK)	< 1	0.84	16.2	149.1	1526.5	∞	∞	∞	∞	∞

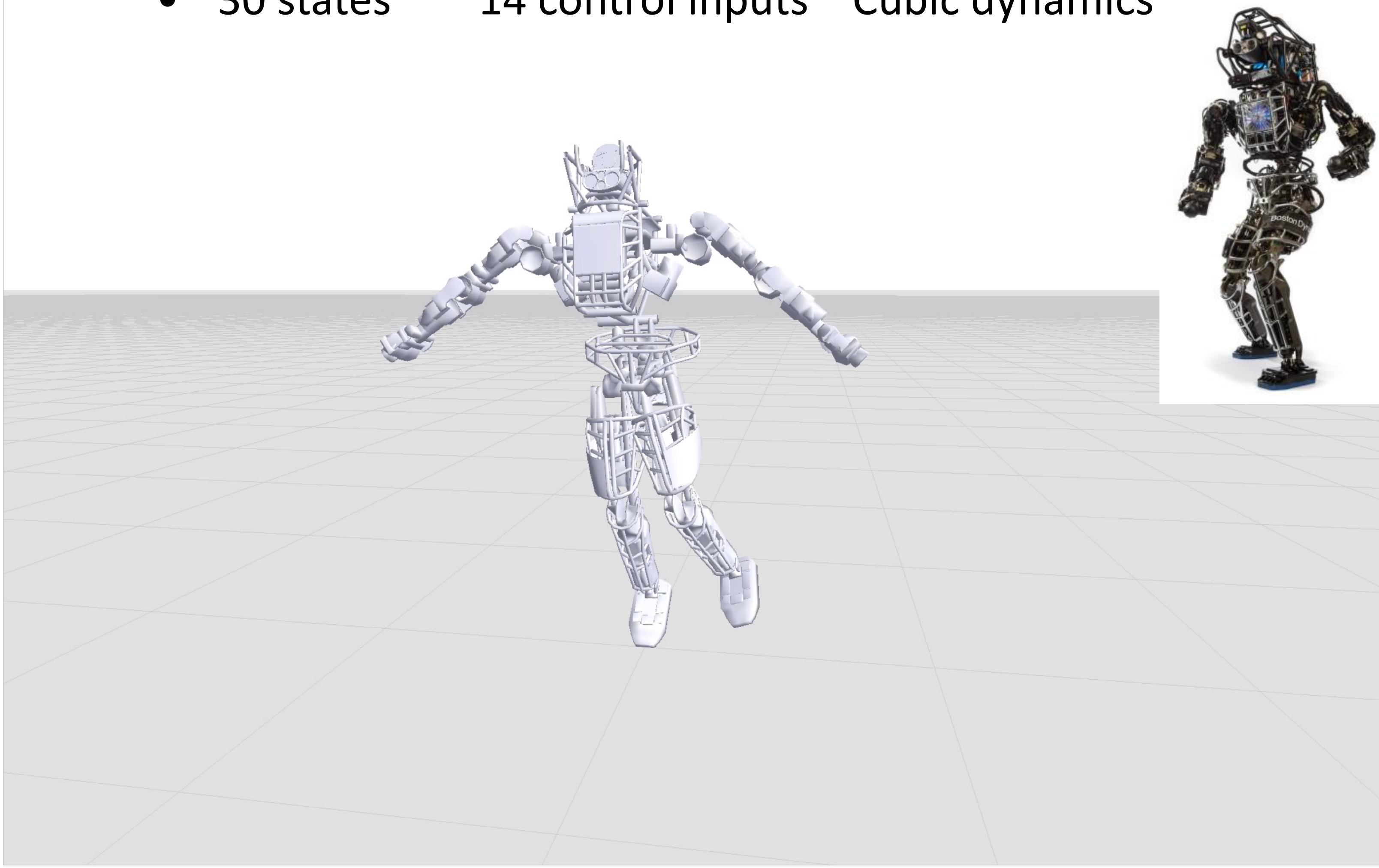
ROA volume ratio:

2N (states)	4	6	8	10	12
ρ_{dsos}/ρ_{sos}	0.38	0.45	0.13	0.12	0.09
ρ_{sdsos}/ρ_{sos}	0.88	0.84	0.81	0.79	0.79

[Majumdar, AAA, Tedrake, CDC]

Stabilizing ATLAS

- 30 states   14 control inputs   Cubic dynamics



Done by **SDSOS Optimization**

[Majumdar, AAA, Tedrake, CDC]

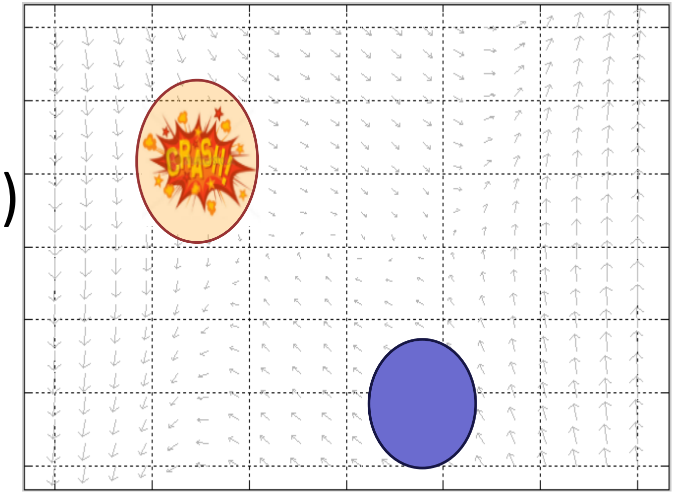
https://github.com/spot-toolbox/spotless

Real-time barrier certificates

Collision avoidance

$$\dot{x} = f(x)$$

$$(f: \mathbb{R}^n \to \mathbb{R}^n)$$



***S*:** needs safety verification

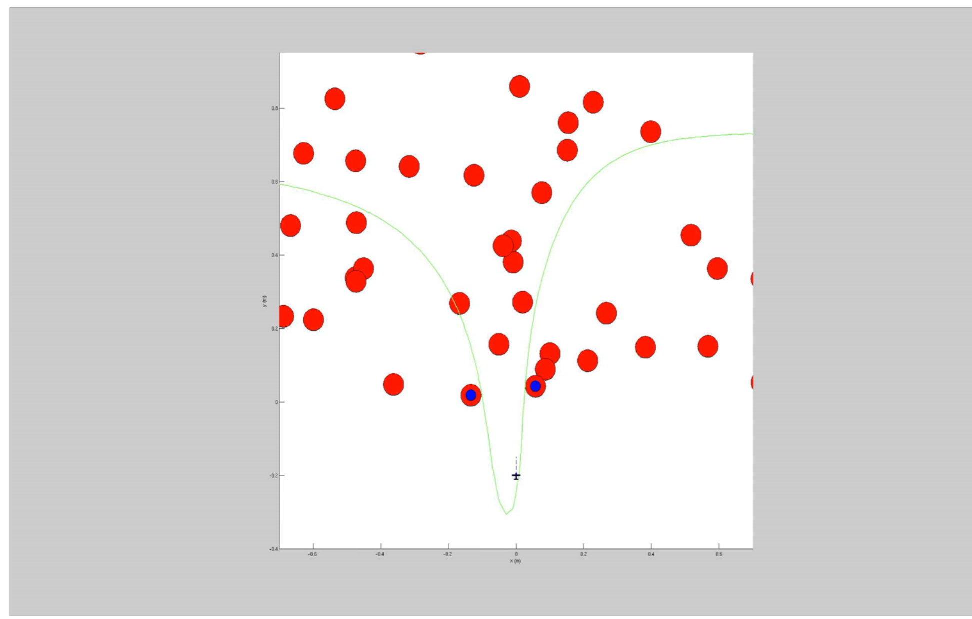
***U*:** unsafe (or forbidden) set

$(V: \mathbb{R}^n \to \mathbb{R})$

$$V(x) < 0, \forall x \in S$$

$$V(x) > 0, \forall x \in U$$

$$\dot{V}(x) = \langle \nabla V(x), f(x) \rangle \leq 0$$

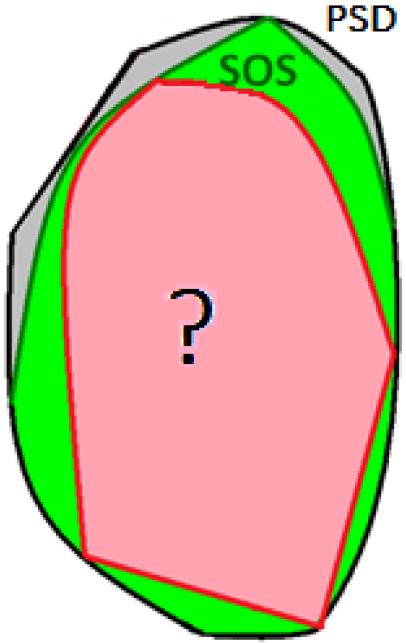


SDOS run time: ~ms

[AAA, Majumdar, *Optimization Letters*]

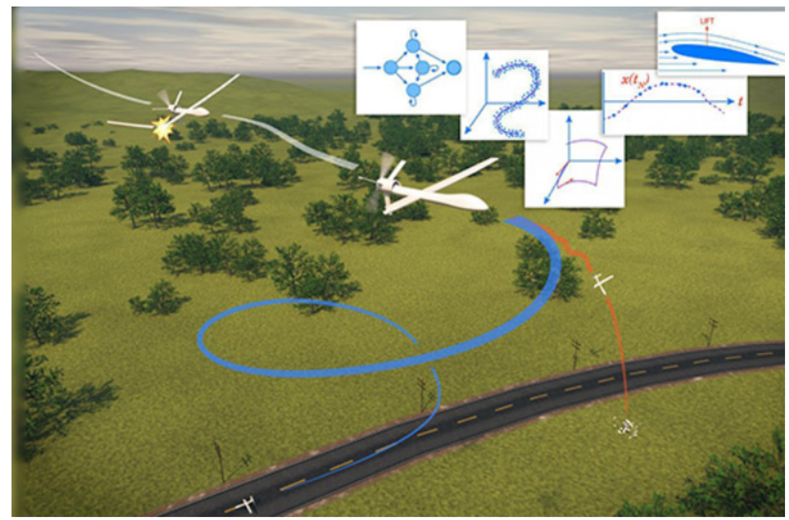
PRINCETON UNIVERSITY ORFE

Take-away message



There is plenty of room inside the SOS cone!

But we have a long way to go!



- Trade-offs between the different approximations?
- What algorithm to use when? Can ML tell us?
- How to exploit sparsity, symmetry, etc. on top of this?
- Real-time implementations?

Backup slides…

What can DSOS/SDSOS do in theory?

$$p(x) > 0, \forall x \in \{x \in \mathbb{R}^n \mid g_i(x) \geq 0, i = 1, \dots, m\}$$

- Is there always an SOS proof?

Yes, e.g. based on Putinar's Psatz.
(under a compactness assumption)

Putinar

If $p(x) > 0, \forall x \in S$,
then $p(x) = \sigma_0(x) + \sum_i \sigma_i(x) g_i(x)$,
where σ_0, σ_i are sos

- Is there always an SDSOS proof?

- Is there always an DSOS proof?

Yes! In fact, a much stronger statement is true.

An optimization-free Positivstellensatz (1/2)

$$p(x) > 0, \forall x \in \{x \in \mathbb{R}^n \mid g_i(x) \geq 0, i = 1, \dots, m\}$$

$2d$ =maximum degree of p, g_i

$\Updownarrow$ Under compactness assumptions, i.e., $\{x \mid g_i(x) \geq 0\} \subseteq B(0, R)$

$\exists\, r \in \mathbb{N}$ such that

$$\left(f(v^2 - w^2) - \frac{1}{r}\left(\textstyle\sum_i \left(v_i^2 - w_i^2\right)^2\right)^d + \frac{1}{2r}\left(\textstyle\sum_i \left(v_i^4 + w_i^4\right)\right)^d\right) \cdot \left(\textstyle\sum_i v_i^2 + \sum_i w_i^2\right)^{r^2}$$

has **nonnegative coefficients**,

where f is a form in $n + m + 3$ variables and of degree $4d$, which can be explicitly written from p, g_i and R.

[AAA, Hall, *Math of OR*] (2018 INFORMS Young Researchers Prize)

An optimization-free Positivstellensatz (2/2)

$$p(x) > 0 \text{ on } \{x \mid g_i(x) \geq 0\} \Leftrightarrow$$
$$\exists r \in \mathbb{N} \text{ s.t. } \left(f(v^2 - w^2) - \frac{1}{r}\left(\sum_i (v_i^2 - w_i^2)^2\right)^d + \frac{1}{2r}\left(\sum_i (v_i^4 + w_i^4)\right)^d \right) \cdot \left(\sum_i v_i^2 + \sum_i w_i^2\right)^{r^2}$$
has ≥ 0 coefficients

- $p(x) > 0$ on $\{x \mid g_i(x) \geq 0\} \Leftrightarrow f$ is pd
- **Result by Polya (1928):**

f **even** and pd $\Rightarrow \exists r \in \mathbb{N}$ such that $f(z) \cdot \left(\sum_i z_i^2\right)^r$ has nonnegative coefficients.

- Make $f(z)$ even by considering $f(v^2 - w^2)$. We lose positive definiteness of f with this transformation.
- Add the positive definite term $\frac{1}{2r}\left(\sum_i (v_i^4 + w_i^4)\right)^d$ to $f(v^2 - w^2)$ to make it positive definite. **Apply Polya's result.**
- The term $-\frac{1}{r}\left(\sum_i (v_i^2 - w_i^2)^2\right)^d$ ensures that the converse holds as well.

As a corollary, gives LP/SOCP-based converging hierarchies…

(Even forms with nonnegative coefficients are trivially dsos.)