Sum of Squares Optimization and Applications

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

CDC'17, Tutorial Lecture
Melbourne

Optimization over nonnegative polynomials

Defn. A polynomial $p(x) := p(x_1, \dots, x_n)$ is nonnegative if $p(x) \geq 0, \forall x \in \mathbb{R}^n$.

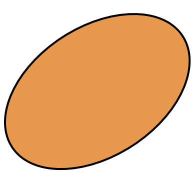
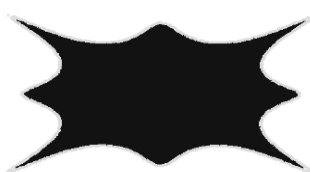
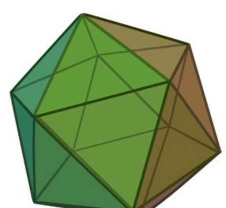
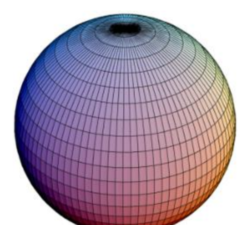
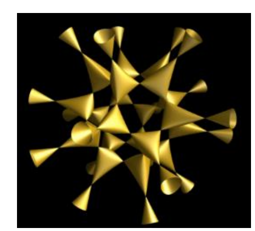
Example: When is

$$p(x_1, x_2) = c_1 x_1^4 - 6x_1^3 x_2 - 4x_1^3 + c_2 x_1^2 x_2^2 + 10x_1^2 + 12x_1 x_2^2 + c_3 x_2^4$$

nonnegative?
nonnegative over a given basic semialgebraic set?

Basic semialgebraic set: $\{x \in \mathbb{R}^n | \, g_i(x) \geq 0\}$

Ex:
$$x_1^3 - 2x_1 x_2^4 \geq 0$$
$$x_1^4 + 3x_1 x_2 - x_2^6 \geq 0$$

Why do this?!

Optimization over nonnegative polynomials

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

Optimization	Statistics	Control

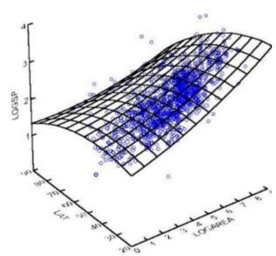
Optimization

- Lower bounds on polynomial optimization problems

$$\max_{\gamma} \gamma$$
$$\text{s.t.} \quad p(x) - \gamma \geq 0,$$
$$\forall x \in \{g_i(x) \geq 0\}$$

Statistics

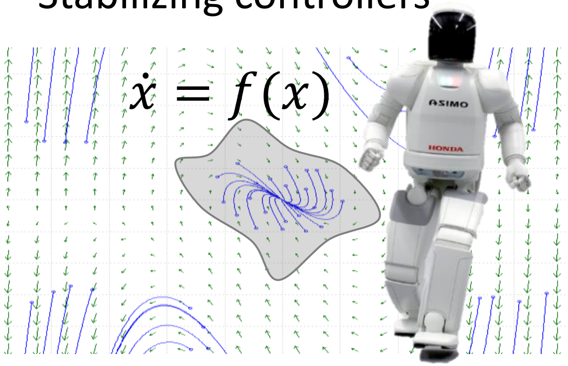
- Fitting a polynomial to data subject to shape constraints (e.g., convexity, or monotonicity)



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

Control

- Stabilizing controllers



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

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

How would you prove nonnegativity?

Ex. Decide if the following polynomial is nonnegative:

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

- Not so easy! (In fact, **NP-hard for degree ≥ 4**)
- But what if I told you:

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

- Is it any easier to test for a sum of squares (SOS) decomposition?

SOS→SDP

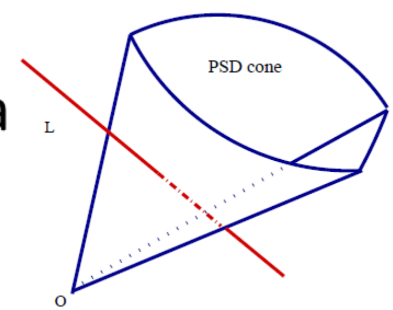
Thm: A polynomial ***p(x)*** of degree ***2d*** is sos if and only if there exists a matrix Q such that

$$Q \succcurlyeq 0,$$
$$p(x) = z(x)^T Q z(x),$$

where

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

The set of such matrices Q forms the feasible set of a semidefinite program.



Example coming up in Antonis' talk

Fully automated in YALMIP, SOSTOOLS, SPOTLESS, GloptiPoly, …

How to prove nonnegativity over a basic semialgebraic set?

Positivstellensatz: Certifies that

$$p(x) > 0 \text{ on } \{g_1(x) \geq 0, \dots, g_m(x) \geq 0\}$$

Putinar's Psatz: (1993)

$$p(x) > 0 \text{ on } \{g_1(x) \geq 0, \dots, g_m(x) \geq 0\}$$

under Archimedean condition

$$\Downarrow$$

$$p(x) = \sigma_0(x) + \textstyle\sum_i \sigma_i(x) g_i(x),$$

where $\sigma_i, i = 0, \dots, m$ are sos

Search for σ_i is an SDP when we bound the degree.

[Lasserre, Parrilo]

Stengle's Psatz (1974)

Schmudgen's Psatz (1991)

...

All use sos polynomials...

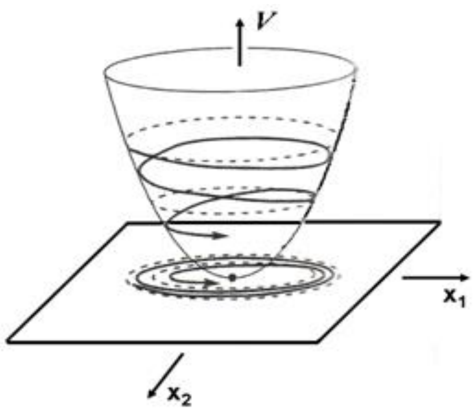
Dynamics and Control

Lyapunov theory with sum of squares (sos) techniques

Questions about dynamical systems (e.g., stability, safety) → **Lyapunov Theory** → Search for sign-definite functions → Polynomial inequalities → **SOS** → Semidefinite programming

Ex. Lyapunov's stability theorem.

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



Lyapunov function

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

$$\dot{V}(x) = \langle \frac{\partial V}{\partial x}, f(x) \rangle$$

$$\begin{matrix} V(x) & \text{sos} \\ -\dot{V}(x) & \text{sos} \end{matrix} \Rightarrow \begin{matrix} V(x) > 0 \\ -\dot{V}(x) > 0 \end{matrix} \Rightarrow \text{GAS}$$

(similar local version)

Global stability

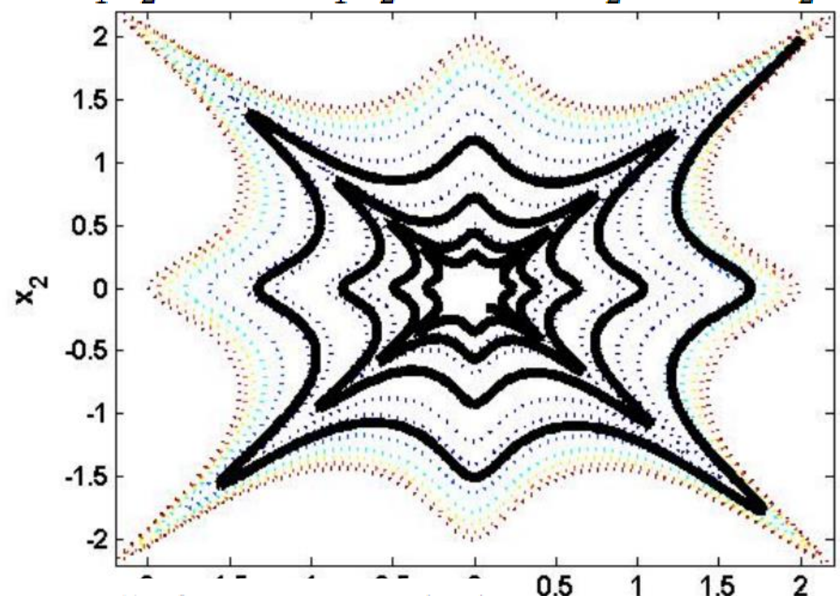
$$\begin{array}{ll} V(x) & \text{sos} \\ -\dot{V}(x) & \text{sos} \end{array} \Rightarrow \begin{array}{l} V(x) > 0 \\ -\dot{V}(x) > 0 \end{array} \Rightarrow \text{GAS}$$

Example.

$$\begin{aligned} \dot{x}_1 &= -0.15x_1^7 + 200x_1^6x_2 - 10.5x_1^5x_2^2 - 807x_1^4x_2^3 + 14x_1^3x_2^4 + 600x_1^2x_2^5 - 3.5x_1x_2^6 + 9x_2^7 \\ \dot{x}_2 &= -9x_1^7 - 3.5x_1^6x_2 - 600x_1^5x_2^2 + 14x_1^4x_2^3 + 807x_1^3x_2^4 - 10.5x_1^2x_2^5 - 200x_1x_2^6 - 0.15x_2^7 \end{aligned}$$

Couple lines of code in SOSTOOLS, YALMIP, SPOTLESS, etc.

Output of SDP solver:

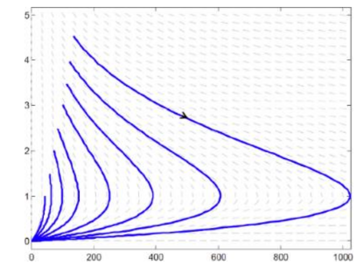


$$\begin{aligned} V &= 0.02x_1^8 + 0.015x_1^7x_2 + 1.743x_1^6x_2^2 - 0.106x_1^5x_2^3 - 3.517x_1^4x_2^4 \\ &\quad + 0.106x_1^3x_2^5 + 1.743x_1^2x_2^6 - 0.015x_1x_2^7 + 0.02x_2^8. \end{aligned}$$

Theoretical limitations: converse implications may fail

- Testing asymptotic stability of cubic vector fields is strongly NP-hard. [AAA]

$$\begin{aligned}\dot{x} &= -x + xy \\ \dot{y} &= -y\end{aligned}$$

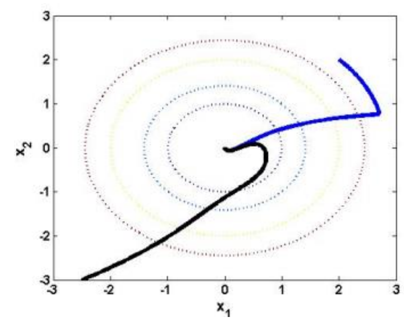


- Globally asymptotically stable.
- But no polynomial Lyapunov function of any degree!

[AAA, Krstic, Parrilo]

$$\begin{aligned}\dot{x}_1 &= -x_1^3x_2^2 + 2x_1^3x_2 - x_1^3 + 4x_1^2x_2^2 - 8x_1^2x_2 + 4x_1^2 - x_1x_2^4 + 4x_1x_2^3 - 4x_1 + 10x_2^2 \\ \dot{x}_2 &= -9x_1^2x_2 + 10x_1^2 + 2x_1x_2^3 - 8x_1x_2^2 - 4x_1 - x_2^3 + 4x_2^2 - 4x_2\end{aligned}$$

- $V(x) = x_1^2 + x_2^2$ proves GAS.
- SOS fails to find *any* quadratic Lyapunov function.



[AAA,Parrilo]

Converse statements possible in special cases

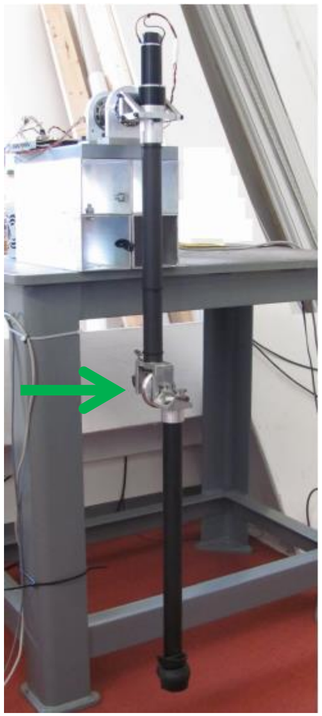
1. Asymptotically stable homogeneous polynomial vector field→ Rational Lyapunov function with an SOS certificate.
   [AAA, El Khadir]

2. Exponentially stable polynomial vector field on a compact set→ Polynomial Lyapunov function.
   [Peet, Papachristodoulou]

3. Asymptotically stable switched linear system→ Polynomial Lyapunov function with an SOS certificate.
   [Parrilo, Jadbabaie]

4. Asymptotically stable switched linear system→ Convex polynomial Lyapunov function with an SOS certificate.
   [AAA, Jungers]

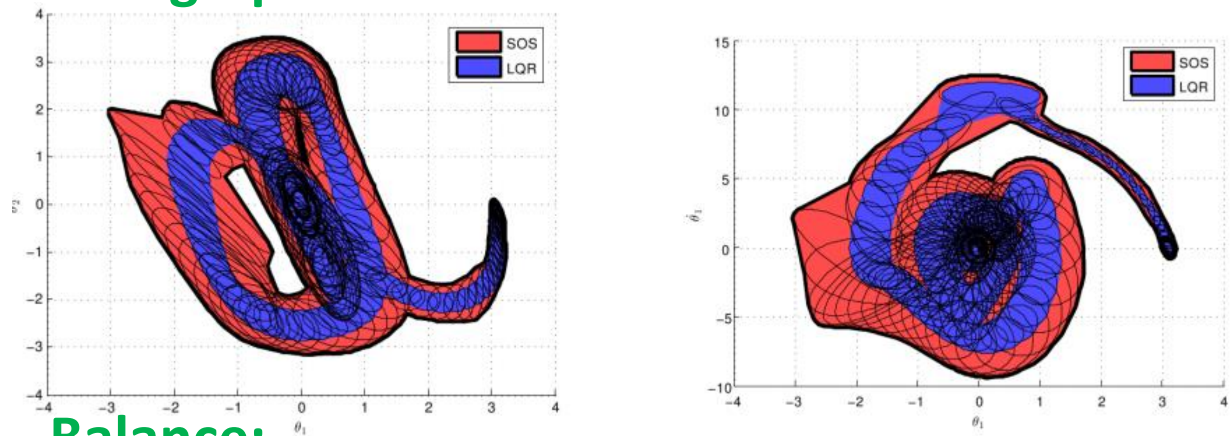
Local stability – SOS on the Acrobot



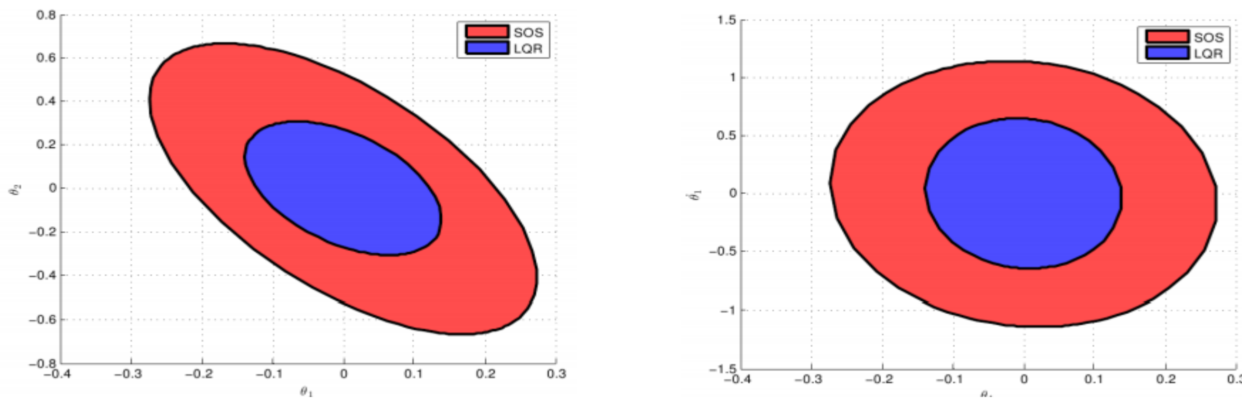
(4-state system)

Controller designed by SOS

Swing-up:



Balance:

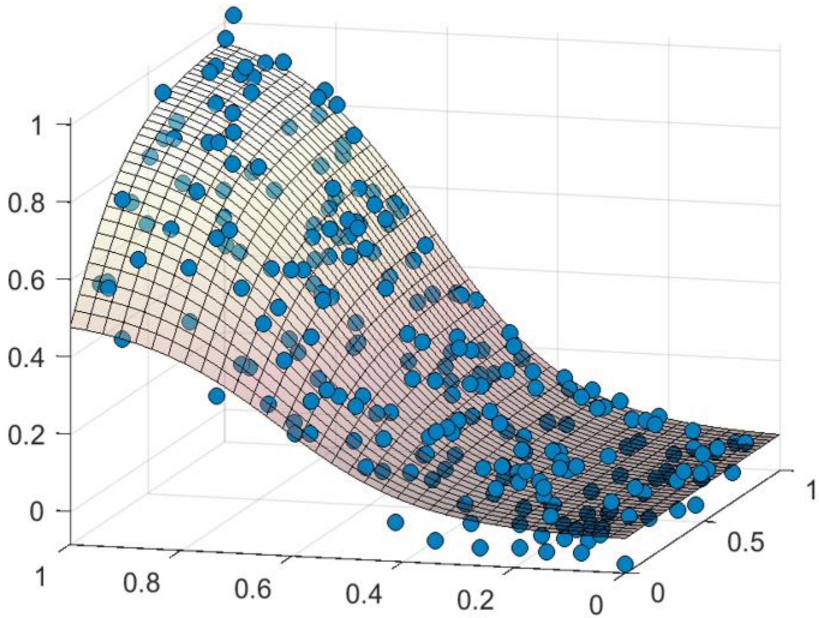


[Majumdar, AAA, Tedrake]

(Best paper award - *IEEE Conf. on Robotics and Automation*)

Statistics and Machine Learning

Monotone regression: problem definition



- ***N* data points:** (x_i, y_i) with $x_i \in \mathbb{R}^n, y_i \in \mathbb{R}$, noisy measurements of a monotone function

$$y_i = f(x_i) + \epsilon_i$$

- **Feature domain**: box $B \subseteq \mathbb{R}^n$

Monotonicity profile:

$$\rho_j = \begin{cases} 1 & \text{if } f \text{ is monotonically increasing w.r.t. } x_j \\ -1 & \text{if } f \text{ is monotonically decreasing w.r.t. } x_j \\ 0 & \text{if no monotonicity requirements on } f \text{ w.r.t. } x_j \end{cases}$$

for $j = 1, \ldots, n$.

Goal: Fit a polynomial to the data that has monotonicity profile ρ over B.

NP-hardness and SOS relaxation

Theorem: Given a cubic polynomial p, a box B, and a monotonicity profile ρ, it is NP-hard to test whether p has profile ρ over B.

[AAA, Curmei, Hall]

SOS relaxation:

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

where

$$\mathrm{B} = [b_1^-, b_1^+] \times \cdots \times [b_n^-, b_n^+]$$

⇐

$$\frac{\partial p(x)}{\partial x_j} = \sigma_0(x) + \sum_i \sigma_i(x)(b_i^+ - x_i)(x_i - b_i^-)$$

where $\sigma_i, i = 0, \ldots, n$ are sos polynomials

Approximation theorem

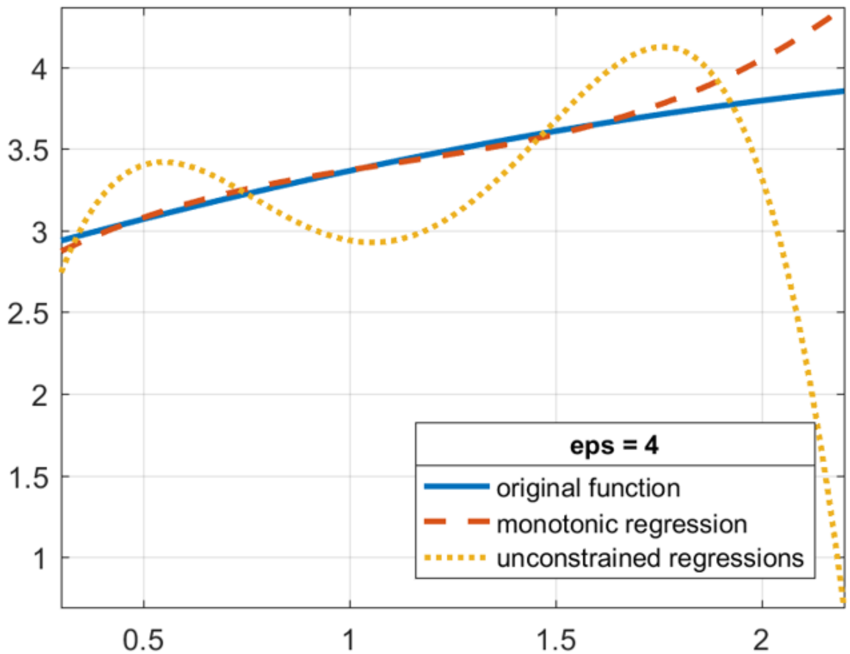
Theorem: For any $\epsilon > 0$, and any C^1 function f with monotonicity profile ρ, there exists a polynomial p with the same profile ρ, such that

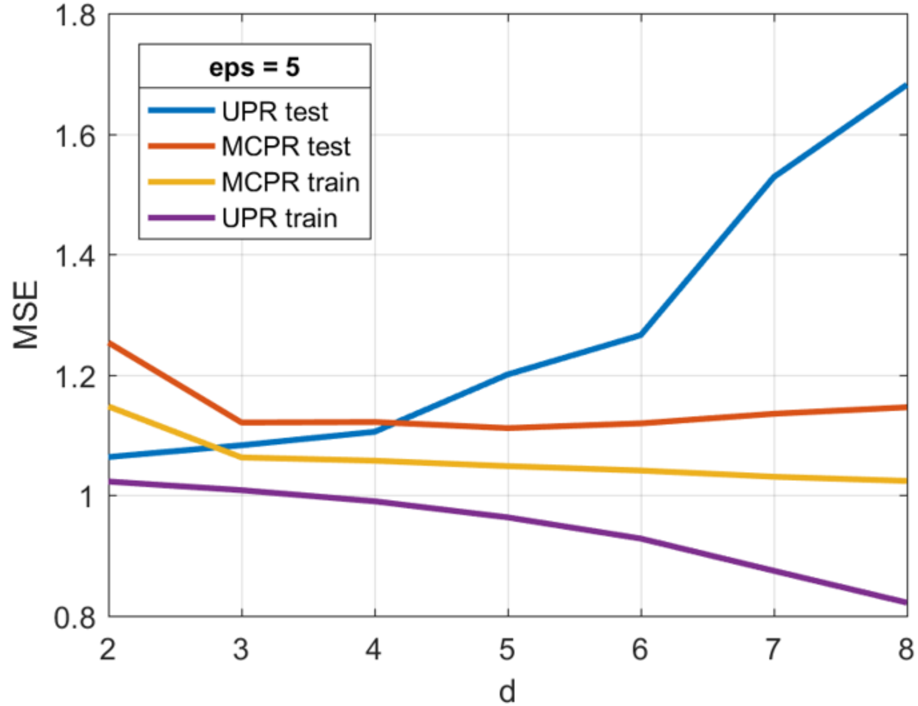
$$\max_{x \in B} |f(x) - p(x)| < \epsilon.$$

Moreover, one can certify its monotonicity profile using SOS.

[AAA, Curmei, Hall]

Numerical experiments





Polynomial Optimization

A meta-theorem for producing hierarchies

[AAA, Hall]

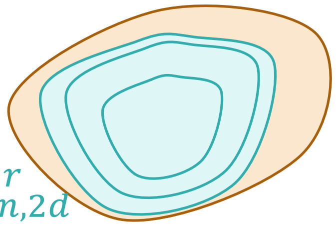
Theorem: Let $K_{n,2d}^r$ be a sequence of sets of homogeneous polynomials in n variables and of degree $2d$. If:

(1) $K_{n,2d}^r \subseteq P_{n,2d}\ \forall r$ and $\exists\ s_{n,2d}$ pd in $K_{n,2d}^0$

(2) $p > 0 \Rightarrow \exists r \in \mathbb{N}$ s.t. $p \in K_{n,2d}^r$

(3) $K_{n,2d}^r \subseteq K_{n,2d}^{r+1}\ \forall r$

(4) $p \in K_{n,2d}^r \Rightarrow p + \epsilon s_{n,2d} \in K_{n,2d}^r, \forall \epsilon \in [0,1]$



Then,

POP: $\min_{x \in \mathbb{R}^n} p(x)$ $s.t.\ g_i(x) \geq 0, i = 1, \dots, m$

$2d =$ maximum degree of p, g_i

$r \uparrow \infty$
Compactness assumptions
$=$
opt. val.

$$\max_{\gamma} \gamma$$
$$s.t.\ f_\gamma(z) - \frac{1}{r} s_{n+m+3,4d}(z) \in K_{n+m+3,4d}^r$$

where f_γ is a form which can be written down explicitly from p, g_i.

Example: Artin cones $A_{n,2d}^r = \{p \mid p \cdot q \text{ is sos for some sos } q \text{ of degree } 2r\}$

An optimization-free converging hierarchy

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

$2d$ =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

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

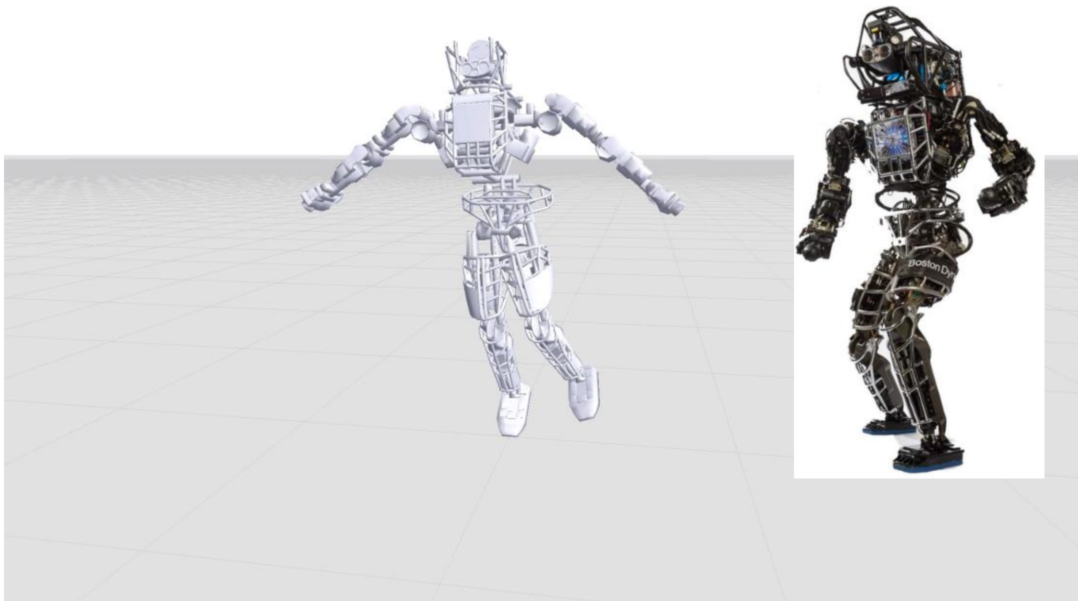
has **nonnegative coefficients**,

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]

PRINCETON UNIVERSITY ORFE (also leads to DSOS/SDSOS-based converging hierarchies for POP)

Ongoing directions: large-scale/real-time verification



- 30 states, 14 control inputs, cubic dynamics
- Done with SDSOS optimization (see Georgina's talk)

Two promising approaches:

1. LP and SOCP-based alternatives to SOS, Georgina's talk
   Less powerful than SOS (James' talk), but good enough for some applications
2. Exploiting problem structure and designing customized algorithms

   Antonis' talk (next), and Pablo Parrilo's plenary (Thu. 8:30am)