

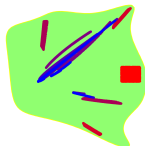
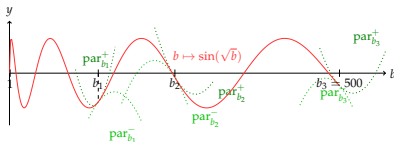
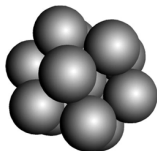
Semialgebraic Relaxations using Moment-SOS Hierarchies

Victor Magron, Postdoc LAAS-CNRS

17 September 2014

SIERRA Seminar

Laboratoire d'Informatique de l'Ecole Normale Supérieure



Personal Background

- 2008 – 2010: Master at Tokyo University
HIERARCHICAL DOMAIN DECOMPOSITION METHODS
(S. Yoshimura)
- 2010 – 2013: PhD at Inria Saclay LIX/CMAP
FORMAL PROOFS FOR NONLINEAR OPTIMIZATION
(S. Gaubert and B. Werner)
- 2014–now: Postdoc at LAAS-CNRS
MOMENT-SOS APPLICATIONS
(J.B. Lasserre and D. Henrion)

Errors and Proofs

- Mathematicians want to eliminate all the uncertainties on their results. Why?



M. Lecat, Erreurs des Mathématiciens des origines à nos jours, 1935.

130 pages of errors! (Euler, Fermat, Sylvester, ...)

Errors and Proofs

- Possible workaround: proof assistants

COQ (Coquand, Huet 1984) 🐦

HOL-LIGHT (Harrison, Gordon 1980)



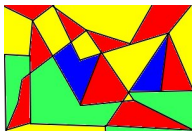
Built in top of OCAML 🐪

Complex Proofs

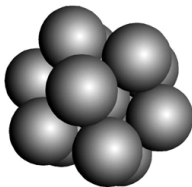
- Complex mathematical proofs / mandatory computation



K. Appel and W. Haken , Every Planar Map is Four-Colorable, 1989.



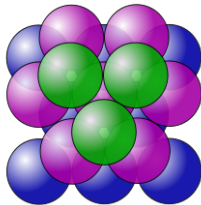
T. Hales, A Proof of the Kepler Conjecture, 1994.



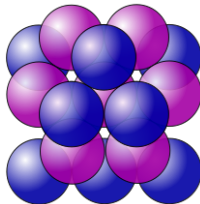
From Oranges Stack...

Kepler Conjecture (1611):

The maximal density of sphere packings in 3D-space is $\frac{\pi}{\sqrt{18}}$



Face-centered cubic Packing



Hexagonal Compact Packing

...to Flyspeck Nonlinear Inequalities

- The proof of T. Hales (1998) contains mathematical and computational parts
- Computation: check thousands of nonlinear inequalities
- Robert MacPherson, editor of The Annals of Mathematics: “[...] the mathematical community will have to get used to this state of affairs.”
- **Flyspeck** [Hales 06]: **Formal Proof of Kepler Conjecture**

...to Flyspeck Nonlinear Inequalities

- The proof of T. Hales (1998) contains mathematical and computational parts
- Computation: check thousands of nonlinear inequalities
- Robert MacPherson, editor of The Annals of Mathematics: “[...] the mathematical community will have to get used to this state of affairs.”
- **Flyspeck [Hales 06]: Formal Proof of Kepler Conjecture**
- **Project Completion on 10 August by the Flyspeck team!!**

A “Simple” Example

In the computational part:

■ Multivariate Polynomials:

$$\Delta \mathbf{x} := x_1 x_4 (-x_1 + x_2 + x_3 - x_4 + x_5 + x_6) + x_2 x_5 (x_1 - x_2 + x_3 + x_4 - x_5 + x_6) + x_3 x_6 (x_1 + x_2 - x_3 + x_4 + x_5 - x_6) - x_2 (x_3 x_4 + x_1 x_6) - x_5 (x_1 x_3 + x_4 x_6)$$

A “Simple” Example

In the computational part:

- **Semialgebraic** functions: composition of polynomials with $|\cdot|, \sqrt{\cdot}, +, -, \times, /, \sup, \inf, \dots$

$$\begin{aligned} p(\mathbf{x}) &:= \partial_4 \Delta \mathbf{x} & q(\mathbf{x}) &:= 4x_1 \Delta \mathbf{x} \\ r(\mathbf{x}) &:= p(\mathbf{x}) / \sqrt{q(\mathbf{x})} \end{aligned}$$

$$l(\mathbf{x}) := -\frac{\pi}{2} + 1.6294 - 0.2213 (\sqrt{x_2} + \sqrt{x_3} + \sqrt{x_5} + \sqrt{x_6} - 8.0) + 0.913 (\sqrt{x_4} - 2.52) + 0.728 (\sqrt{x_1} - 2.0)$$

A “Simple” Example

In the computational part:

- **Transcendental** functions \mathcal{T} : composition of semialgebraic functions with \arctan , \exp , \sin , $+$, $-$, \times , \dots

A “Simple” Example

In the computational part:

- Feasible set $\mathbf{K} := [4, 6.3504]^3 \times [6.3504, 8] \times [4, 6.3504]^2$

Lemma₉₉₂₂₆₉₉₀₂₈ from Flyspeck:

$$\forall \mathbf{x} \in \mathbf{K}, \arctan\left(\frac{p(\mathbf{x})}{\sqrt{q(\mathbf{x})}}\right) + l(\mathbf{x}) \geq 0$$

Existing Formal Frameworks

Formal proofs for Global Optimization:

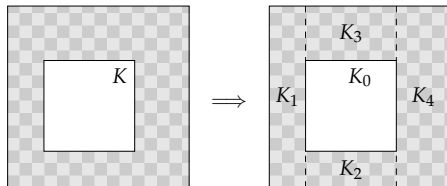
- Bernstein polynomial methods [Zumkeller 08]
restricted to polynomials
- Taylor + Interval arithmetic [Melquiond 12, Solovyev 13]
robust but subject to the **CURSE OF DIMENSIONALITY**

Existing Formal Frameworks

Lemma₉₉₂₂₆₉₉₀₂₈ from Flyspeck:

$$\forall \mathbf{x} \in \mathbf{K}, \arctan\left(\frac{\partial_4 \Delta \mathbf{x}}{\sqrt{4x_1 \Delta \mathbf{x}}}\right) + l(\mathbf{x}) \geq 0$$

- Dependency issue using Interval Calculus:
 - One can bound $\partial_4 \Delta \mathbf{x} / \sqrt{4x_1 \Delta \mathbf{x}}$ and $l(\mathbf{x})$ separately
 - Too coarse lower bound: -0.87
 - Subdivide \mathbf{K} to prove the inequality



Introduction

Moment-SOS relaxations

Another look at Nonnegativity

New Applications of Moment-SOS Hierarchies

Conclusion

Polynomial Optimization Problems

- Semialgebraic set $\mathbf{K} := \{\mathbf{x} \in \mathbb{R}^n : g_1(\mathbf{x}) \geq 0, \dots, g_m(\mathbf{x}) \geq 0\}$
- $p^* := \min_{\mathbf{x} \in \mathbf{K}} p(\mathbf{x})$: NP hard
- Sums of squares $\Sigma[\mathbf{x}]$
e.g. $x_1^2 - 2x_1x_2 + x_2^2 = (x_1 - x_2)^2$
- $\mathcal{Q}(\mathbf{K}) := \left\{ \sigma_0(\mathbf{x}) + \sum_{j=1}^m \sigma_j(\mathbf{x})g_j(\mathbf{x}), \text{ with } \sigma_j \in \Sigma[\mathbf{x}] \right\}$

Polynomial Optimization Problems

Archimedean module

The set \mathbf{K} is compact and the polynomial $N - \|\mathbf{x}\|_2^2$ belongs to $\mathcal{Q}(\mathbf{K})$ for some $N > 0$.

- Assume that \mathbf{K} is a box: product of closed intervals
- Normalize the feasibility set to get $\mathbf{K}' := [-1, 1]^n$
 $\mathbf{K}' := \{\mathbf{x} \in \mathbb{R}^n : g_1 := 1 - x_1^2 \geq 0, \dots, g_n := 1 - x_n^2 \geq 0\}$
- $n - \|\mathbf{x}\|_2^2$ belongs to $\mathcal{Q}(\mathbf{K}')$

Convexification and the K Moment Problem

- Borel σ -algebra \mathcal{B} (generated by the open sets of \mathbb{R}^n)
- $\mathcal{M}_+(\mathbf{K})$: set of probability measures supported on \mathbf{K} .
If $\mu \in \mathcal{M}_+(\mathbf{K})$ then
 - 1 $\mu : \mathcal{B} \rightarrow [0, 1], \mu(\emptyset) = 0, \mu(\mathbb{R}^n) < \infty$
 - 2 $\mu(\bigcup_i B_i) = \sum_i \mu(B_i)$, for any countable $(B_i) \subset \mathcal{B}$
 - 3 $\int_{\mathbf{K}} \mu(d\mathbf{x}) = 1$
- $\text{supp}(\mu)$ is the smallest set \mathbf{K} such that $\mu(\mathbb{R}^n \setminus \mathbf{K}) = 0$

Convexification and the K Moment Problem

$$p^* = \inf_{\mathbf{x} \in \mathbf{K}} p(\mathbf{x}) = \inf_{\mu \in \mathcal{M}_+(\mathbf{K})} \int_{\mathbf{K}} p d\mu$$

Convexification and the K Moment Problem

- Let $(\mathbf{x}^\alpha)_{\alpha \in \mathbb{N}^n}$ be the monomial basis

Definition

A sequence \mathbf{y} has a representing measure on \mathbf{K} if there exists a finite measure μ supported on \mathbf{K} such that

$$\mathbf{y}_\alpha = \int_{\mathbf{K}} \mathbf{x}^\alpha \mu(d\mathbf{x}), \quad \forall \alpha \in \mathbb{N}^n.$$

Convexification and the K Moment Problem

$$L_{\mathbf{y}}(q) : q \in \mathbb{R}[\mathbf{x}] \mapsto \sum_{\alpha} q_{\alpha} \mathbf{y}_{\alpha}$$

Theorem [Putinar 93]

Let \mathbf{K} be compact and $\mathcal{Q}(\mathbf{K})$ be Archimedean.

Then \mathbf{y} has a representing measure on \mathbf{K}

iff

$$L_{\mathbf{y}}(\sigma) \geq 0, \quad L_{\mathbf{y}}(g_j \sigma) \geq 0, \quad \forall \sigma \in \Sigma[\mathbf{x}].$$

Lasserre's Hierarchy of SDP relaxations

- Moment matrix

$$\mathbf{M}(\mathbf{y})_{u,v} := L_{\mathbf{y}}(u \cdot v), \quad u, v \text{ monomials}$$

- Localizing matrix $M(g_j \mathbf{y})$ associated with g_j

$$\mathbf{M}(g_j \mathbf{y})_{u,v} := L_{\mathbf{y}}(u \cdot v \cdot g_j), \quad u, v \text{ monomials}$$

Lasserre's Hierarchy of SDP relaxations

- $\mathbf{M}_k(\mathbf{y})$ contains $\binom{n+2k}{n}$ variables, has size $\binom{n+k}{n}$
- Truncated matrix of order $k = 2$ with variables x_1, x_2 :

$$\mathbf{M}_2(\mathbf{y}) = \begin{matrix} & \begin{matrix} 1 & | & x_1 & x_2 & | & x_1^2 & x_1x_2 & x_2^2 \end{matrix} \\ \begin{matrix} 1 \\ - \\ x_1 \\ x_2 \\ - \\ x_1^2 \\ x_1x_2 \\ x_2^2 \end{matrix} & \left(\begin{array}{c|c|c} 1 & y_{1,0} & y_{0,1} & y_{2,0} & y_{1,1} & y_{0,2} \\ \hline y_{1,0} & y_{2,0} & y_{1,1} & y_{3,0} & y_{2,1} & y_{1,2} \\ \hline y_{0,1} & y_{1,1} & y_{0,2} & y_{2,1} & y_{1,2} & y_{0,3} \\ \hline y_{2,0} & y_{3,0} & y_{2,1} & y_{4,0} & y_{3,1} & y_{2,2} \\ \hline y_{1,1} & y_{2,1} & y_{1,2} & y_{3,1} & y_{2,2} & y_{1,3} \\ \hline y_{0,2} & y_{1,2} & y_{0,3} & y_{2,2} & y_{1,3} & y_{0,4} \end{array} \right) \end{matrix}$$

Lasserre's Hierarchy of SDP relaxations

- Consider $g_1(\mathbf{x}) := 2 - x_1^2 - x_2^2$. Then $v_1 = \lceil \deg g_1 / 2 \rceil = 1$.

$$\mathbf{M}_1(g_1 \mathbf{y}) = \begin{matrix} & \begin{matrix} 1 & x_1 & x_2 \end{matrix} \\ \begin{matrix} 1 \\ x_1 \\ x_2 \end{matrix} & \begin{pmatrix} 2 - y_{2,0} - y_{0,2} & 2y_{1,0} - y_{3,0} - y_{1,2} & 2y_{0,1} - y_{2,1} - y_{0,3} \\ 2y_{1,0} - y_{3,0} - y_{1,2} & 2y_{2,0} - y_{4,0} - y_{2,2} & 2y_{1,1} - y_{3,1} - y_{1,3} \\ 2y_{0,1} - y_{2,1} - y_{0,3} & 2y_{1,1} - y_{3,1} - y_{1,3} & 2y_{0,2} - y_{2,2} - y_{0,4} \end{pmatrix} \end{matrix}$$

$$\begin{aligned} \mathbf{M}_1(g_1 \mathbf{y})(3,3) &= L(g_1(\mathbf{x}) \cdot x_2 \cdot x_2) = L(2x_2^2 - x_1^2x_2^2 - x_2^4) \\ &= 2y_{0,2} - y_{2,2} - y_{0,4} \end{aligned}$$

Lasserre's Hierarchy of SDP relaxations

- Truncation with moments of order at most $2k$
- $v_j := \lceil \deg g_j / 2 \rceil$
- Hierarchy of semidefinite relaxations:

$$\left\{ \begin{array}{lcl} \inf_{\mathbf{y}} L_{\mathbf{y}}(p) & = & \sum_{\alpha} \int_{\mathbf{K}} p_{\alpha} \mathbf{x}^{\alpha} \mu(d\mathbf{x}) = \sum_{\alpha} p_{\alpha} \mathbf{y}_{\alpha} \\ \mathbf{M}_k(\mathbf{y}) & \succcurlyeq & 0, \\ \mathbf{M}_{k-v_j}(g_j \mathbf{y}) & \succcurlyeq & 0, \quad 1 \leq j \leq m, \\ \mathbf{y}_1 & = & 1. \end{array} \right.$$

Semidefinite Optimization

- F_0, F_α symmetric real matrices, cost vector c

Primal-dual pair of semidefinite programs:

$$(SDP) \begin{cases} \mathcal{P} : & \inf_{\mathbf{y}} \quad \sum_{\alpha} c_{\alpha} \mathbf{y}_{\alpha} \\ & \text{s.t.} \quad \sum_{\alpha} F_{\alpha} \mathbf{y}_{\alpha} - F_0 \succcurlyeq 0 \\ \\ \mathcal{D} : & \sup_{\mathbf{Y}} \quad \text{Trace} (F_0 \mathbf{Y}) \\ & \text{s.t.} \quad \text{Trace} (F_{\alpha} \mathbf{Y}) = c_{\alpha} , \quad \mathbf{Y} \succcurlyeq 0 . \end{cases}$$

- Freely available SDP solvers (CSDP, SDPA, SEDUMI)

Primal-dual Moment-SOS

- $\mathcal{M}_+(\mathbf{K})$: space of probability measures supported on \mathbf{K}

Polynomial Optimization Problems (POP)

$$\begin{array}{ll} \text{(Primal)} & \text{(Dual)} \\ \inf \int_{\mathbf{K}} p d\mu & = \sup \lambda \\ \text{s.t. } \mu \in \mathcal{M}_+(\mathbf{K}) & \text{s.t. } \lambda \in \mathbb{R}, \\ & p - \lambda \in \mathcal{Q}(\mathbf{K}) \end{array}$$

Primal-dual Moment-SOS

- Truncated quadratic module $\mathcal{Q}_k(\mathbf{K}) := \mathcal{Q}(\mathbf{K}) \cap \mathbb{R}_{2k}[\mathbf{x}]$
- For large enough k , **zero duality gap** [Lasserre 01]:

Polynomial Optimization Problems (POP)

(Moment)		(SOS)
$\inf \sum_{\alpha} p_{\alpha} \mathbf{y}_{\alpha}$	$=$	$\sup \lambda$
s.t. $\mathbf{M}_{k-v_j}(g_j \mathbf{y}) \succcurlyeq 0, \quad 0 \leq j \leq m,$		s.t. $\lambda \in \mathbb{R},$
$y_1 = 1$		$p - \lambda \in \mathcal{Q}_k(\mathbf{K})$

Practical Computation

- Hierarchy of SOS relaxations:

$$\lambda_k := \sup_{\lambda} \left\{ \lambda : p - \lambda \in \mathcal{Q}_k(\mathbf{K}) \right\}$$

- Convergence guarantees $\lambda_k \uparrow p^*$ [Lasserre 01]

- If $p - p^* \in \mathcal{Q}_k(\mathbf{K})$ for some k then:

$$\mathbf{y}^* := (1, x_1^*, x_2^*, (x_1^*)^2, x_1^* x_2^*, \dots, (x_1^*)^{2k}, \dots, (x_n^*)^{2k})$$

is a global minimizer of the primal SDP [Lasserre 01].

Practical Computation

- *Caprasse Problem*

$$\forall \mathbf{x} \in [-0.5, 0.5]^4, -x_1x_3^3 + 4x_2x_3^2x_4 + 4x_1x_3x_4^2 + 2x_2x_4^3 + 4x_1x_3 + 4x_3^2 - 10x_2x_4 - 10x_4^2 + 5.1801 \geq 0.$$

- Scale on $[0, 1]^4$

- SOS of degree at most 4

- Redundant constraints $x_1^2 \leq 1, \dots, x_4^2 \leq 1$

The “No Free Lunch” Rule

- Exponential dependency in
 - 1 Relaxation order k (SOS degree)
 - 2 number of variables n
- Computing λ_k involves $\binom{n+2k}{n}$ variables
- At fixed k , $O(n^{2k})$ variables

Introduction

Moment-SOS relaxations

Another look at Nonnegativity

New Applications of Moment-SOS Hierarchies

Conclusion

Another look at Nonnegativity

- Knowledge of \mathbf{K} through $\mu \in \mathcal{M}_+(\mathbf{K})$
- **Independent** of the representation of \mathbf{K}
- Typical from **INVERSE PROBLEMS**
“reconstruct” \mathbf{K} from measuring moments of $\mu \in \mathcal{M}_+(\mathbf{K})$

Another look at Nonnegativity

Borel σ -algebra \mathcal{B} (generated by the open sets of \mathbb{R}^n)

Lemma

A continuous function $p : \mathbb{R}^n \rightarrow \mathbb{R}$ is nonnegative on \mathcal{K}

iff

the set function $\nu : B \in \mathcal{B} \mapsto \int_{\mathcal{K} \cap B} p(\mathbf{x}) \, d\mu(\mathbf{x})$ belongs to \mathcal{M}_+ .

Another look at Nonnegativity

$$\nu : B \in \mathcal{B} \mapsto \int_{\mathbf{K} \cap B} p(\mathbf{x}) d\mu(\mathbf{x})$$

Proof

1 “Only if” part is straightforward

2 “If part”

- If $\nu \in \mathcal{M}_+$ then $p(\mathbf{x}) \geq 0$ for μ -almost all $\mathbf{x} \in \mathbf{K}$, i.e. there exists $G \in \mathcal{B}$ such that $\mu(G) = 0$ and $p(\mathbf{x}) \geq 0$ on $\mathbf{K} \setminus G$.
- $\mathbf{K} = \overline{\mathbf{K} \setminus G}$ (from the support definition).
- Let $\mathbf{x} \in \mathbf{K}$. There is a sequence $(\mathbf{x}_l) \subset \mathbf{K} \setminus G$ such that $\mathbf{x}_l \rightarrow \mathbf{x}$, as $l \rightarrow \infty$. By continuity of p and $p(\mathbf{x}_l) \geq 0$, one has $p(\mathbf{x}) \geq 0$.

The K Moment Problem

Definition

A sequence \mathbf{y} has a representing measure on \mathbf{K} if there exists a finite measure μ supported on \mathbf{K} such that

$$\mathbf{y}_\alpha = \int_{\mathbf{K}} \mathbf{x}^\alpha \mu(d\mathbf{x}), \quad \forall \alpha \in \mathbb{N}^n.$$

The K Moment Problem

$$L_{\mathbf{y}}(q) : q \in \mathbb{R}[\mathbf{x}] \mapsto \sum_{\alpha} q_{\alpha} \mathbf{y}_{\alpha}$$

Theorem [Putinar 93]

Let \mathbf{K} be compact and $\mathcal{Q}(\mathbf{K})$ be Archimedean.
Then \mathbf{y} has a representing measure on \mathbf{K}

iff

$$L_{\mathbf{y}}(\sigma) \geq 0, \quad L_{\mathbf{y}}(g_j \sigma) \geq 0, \quad \forall \sigma \in \Sigma[\mathbf{x}].$$

The K Moment Problem

Theorem [Lasserre 11]

Let \mathbf{K} be compact and $\mu \in \mathcal{M}_+$ be arbitrary fixed with moments

$$\mathbf{y}_\alpha = \int_{\mathbf{K}} \mathbf{x}^\alpha \mu(d\mathbf{x}), \quad \forall \alpha \in \mathbb{N}^n.$$

Then a polynomial p is nonnegative on \mathbf{K}

iff

$$\mathbf{M}_k(p \mathbf{y}) \succcurlyeq 0, \quad \forall k \in \mathbb{N}.$$

Hierarchy of Outer Approximations

Cone of nonnegative polynomials $\mathcal{C}(\mathbf{K})_d \subset \mathbb{R}_d[\mathbf{x}]$

- Each entry of the matrix $\mathbf{M}_k(p \mathbf{y})$ is linear in the coefficients of p
- The set $\Delta_k := \{p \in \mathbb{R}_d[\mathbf{x}] : \mathbf{M}_k(p \mathbf{y}) \succcurlyeq 0\}$ is closed and convex, called **spectrahedron**
- Nested outer approximations:
 $\Delta_0 \supset \Delta_1 \cdots \supset \Delta_k \cdots \supset \mathcal{C}(\mathbf{K})_d$

Hierarchy of Outer Approximations

Hierarchy of upper bounds for $p^* := \inf_{\mathbf{x} \in \mathbf{K}} p(\mathbf{x})$

Theorem [Lasserre 11]

Let \mathbf{K} be closed and $\mu \in \mathcal{M}_+(\mathbf{K})$ be arbitrary fixed with moments (y_α) , $\alpha \in \mathbb{N}^n$. Consider the hierarchy of SDP:

$$\begin{aligned} u_k &:= \max\{\lambda : \mathbf{M}_k((p - \lambda) \mathbf{y}) \succcurlyeq 0\} \\ &= \max\{\lambda : \lambda \mathbf{M}_k(\mathbf{y}) \preccurlyeq \mathbf{M}_k(p \mathbf{y})\}. \end{aligned}$$

Then $u_k \downarrow p^*$, as $k \rightarrow \infty$.

Hierarchy of Outer Approximations

Hierarchy of upper bounds for $p^* := \inf_{\mathbf{x} \in \mathbf{K}} p(\mathbf{x})$

- Computing u_k : **GENERALIZED EIGENVALUE PROBLEM** for the pair $[\mathbf{M}_k(p \mathbf{y}), \mathbf{M}_k(\mathbf{y})]$.
- Index the matrices in the basis of orthonormal polynomials w.r.t. μ :

$$\begin{aligned} u_k &= \max \{ \lambda : \lambda I \preceq \mathbf{M}_k(p \mathbf{y}) \} \\ &= \lambda_{\min}(\mathbf{M}_k(p \mathbf{y})), \end{aligned}$$

a standard **MINIMAL EIGENVALUE PROBLEM**.

Hierarchy of Outer Approximations

Primal-Dual Moment-SOS

(Primal)	(Dual)
$u'_k := \min \int_{\mathbf{K}} p(\mathbf{x}) \sigma(\mathbf{x}) d\mu(\mathbf{x})$	$u_k := \max \lambda$
s.t. $\int_{\mathbf{K}} \sigma(\mathbf{x}) d\mu(\mathbf{x}) = 1,$	s.t. $\lambda \in \mathbb{R},$
$\sigma \in \Sigma_k[\mathbf{x}]$	$\lambda \mathbf{M}_k(\mathbf{y}) \preceq \mathbf{M}_k(p \mathbf{y}).$

Then $u'_k, u_k \rightarrow p^*$, as $k \rightarrow \infty$.

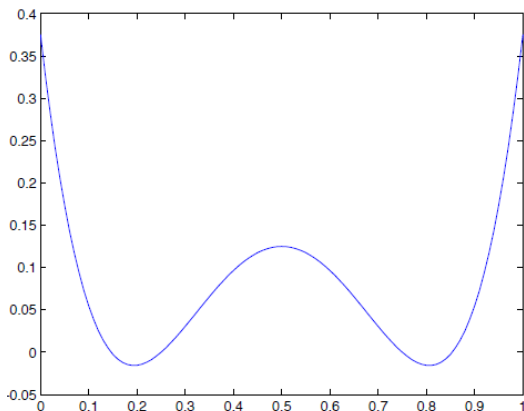
Interpretation of the Primal Problem

$$u_k^* := \min_{\nu} \left\{ \int_{\mathbf{K}} p \underbrace{\sigma d\mu}_{d\nu} : \nu(\mathbf{K}) = 1, \nu(\mathbb{R}^n \setminus \mathbf{K}) = 0, \sigma \in \Sigma_k[\mathbf{x}] \right\}$$

The measure ν approximates the **DIRAC** measure $\delta_{\mathbf{x}=\mathbf{x}^*}$ at a global minimizer $\mathbf{x}^* \in \mathbf{K}$ and

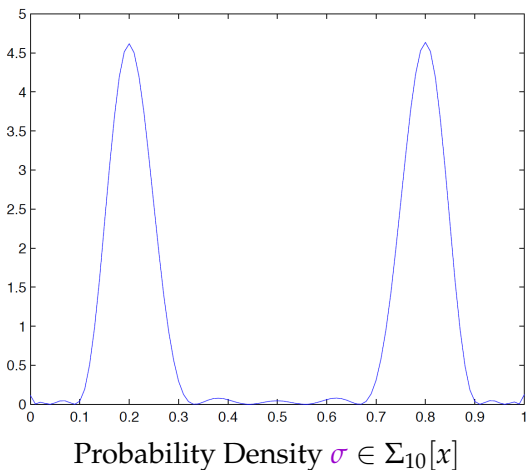
- 1 ν is absolutely continuous w.r.t. μ
- 2 the density of ν is $\sigma \in \Sigma_k[\mathbf{x}]$

Example with uniform probability measure



Polynomial $p := 0.375 - 5x + 21x^2 - 32x^3 + 16x^4$ on $\mathbf{K} := [0, 1]$

Example with uniform probability measure



Extension to the non-compact case

In this case, the measure μ

- needs to satisfy **CARLEMAN**-type sufficient condition to limit the growth of its moments (y_α):

$$\sum_{t=1}^{\infty} L_{\mathbf{y}}(x_i^{2t})^{-1/(2t)} = +\infty, \quad i = 1, \dots, n.$$

- e.g. $d\mu(\mathbf{x}) := \exp(-\|\mathbf{x}\|_2^2/2) d\mu_0$, with $\mu_0 \in \mathcal{M}_+(\mathbf{K})$ finite

Extension to the non-compact case

- $d\mu(\mathbf{x}) := \exp(-\|\mathbf{x}\|_2^2/2) d\mathbf{x}$, for $\mathbf{K} = \mathbb{R}^n$
- $d\mu(\mathbf{x}) := \exp(-\sum_{i=1}^n x_i) d\mathbf{x}$, for $\mathbf{K} = \mathbb{R}_+^n$
- $d\mu(\mathbf{x}) := d\mathbf{x}$, when \mathbf{K} is a box, simplex

Introduction

Moment-SOS relaxations

Another look at Nonnegativity

New Applications of Moment-SOS Hierarchies

Conclusion

Introduction

Moment-SOS relaxations

Another look at Nonnegativity

New Applications of Moment-SOS Hierarchies

Semialgebraic Maxplus Optimization

Formal Nonlinear Optimization

Pareto Curves

Polynomial Images of Semialgebraic Sets

Program Analysis with Polynomial Templates

Conclusion

General informal Framework

Given \mathbf{K} a compact set and f a **transcendental** function, bound $f^* = \inf_{\mathbf{x} \in \mathbf{K}} f(\mathbf{x})$ and prove $f^* \geq 0$

- f is underestimated by a **semialgebraic** function f_{sa}
- Reduce the problem $f_{\text{sa}}^* := \inf_{\mathbf{x} \in \mathbf{K}} f_{\text{sa}}(\mathbf{x})$ to a **polynomial optimization problem (POP)**

General informal Framework

Approximation theory: Chebyshev/Taylor models

- mandatory for non-polynomial problems
- hard to combine with Sum-of-Squares techniques (degree of approximation)

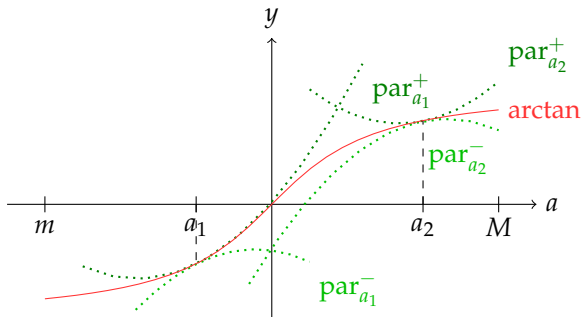
Maxplus Approximation

- Initially introduced to solve Optimal Control Problems [Fleming-McEneaney 00]
- **Curse of dimensionality** reduction [McEneaney Kluberg, Gaubert-McEneaney-Qu 11, Qu 13].
Allowed to solve instances of dim up to 15 (inaccessible by grid methods)
- In our context: approximate **transcendental** functions

Maxplus Approximation

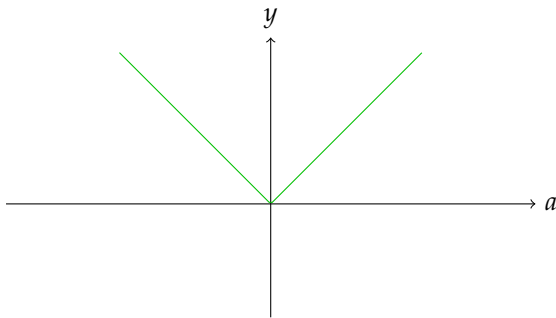
Definition

Let $\gamma \geq 0$. A function $\phi : \mathbb{R}^n \rightarrow \mathbb{R}$ is said to be γ -semiconvex if the function $\mathbf{x} \mapsto \phi(\mathbf{x}) + \frac{\gamma}{2} \|\mathbf{x}\|_2^2$ is convex.



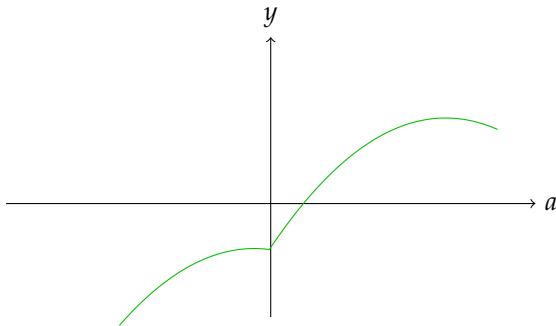
Nonlinear Function Representation

Exact parsimonious maxplus representations



Nonlinear Function Representation

Exact parsimonious maxplus representations



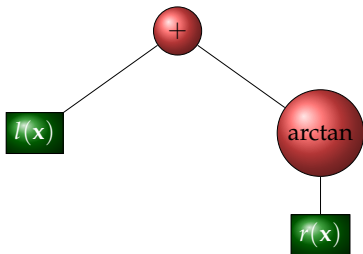
Nonlinear Function Representation

Abstract syntax tree representations of multivariate transcendental functions:

- leaves are **semialgebraic** functions of \mathcal{A}
- nodes are univariate functions of \mathcal{D} or binary operations

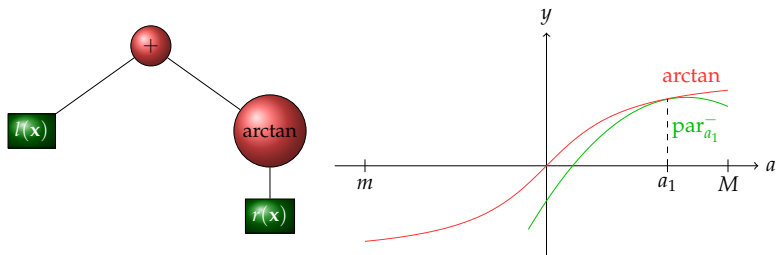
Nonlinear Function Representation

- For the “Simple” Example from Flyspeck:



Maxplus Optimization Algorithm

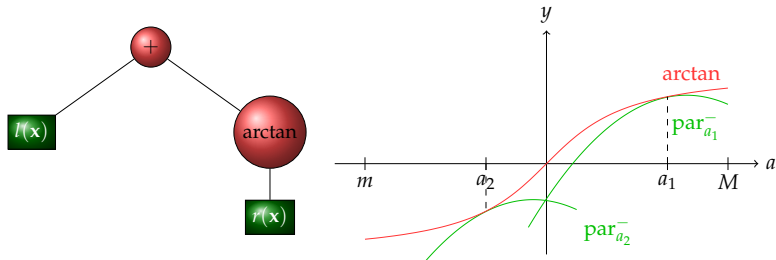
First iteration:



1 control point $\{a_1\}$: $m_1 = -4.7 \times 10^{-3} < 0$

Maxplus Optimization Algorithm

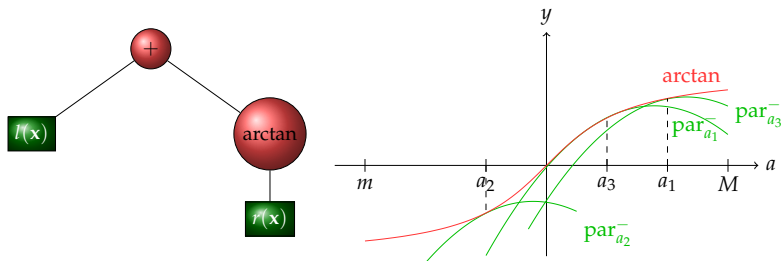
Second iteration:



2 control points $\{a_1, a_2\}$: $m_2 = -6.1 \times 10^{-5} < 0$

Maxplus Optimization Algorithm

Third iteration:



3 control points $\{a_1, a_2, a_3\}$: $m_3 = 4.1 \times 10^{-6} > 0$

OK!

Contributions

Published:



X. Allamigeon, S. Gaubert, V. Magron, and B. Werner.
Certification of inequalities involving transcendental functions:
combining sdp and max-plus approximation, *ECC Conference*
2013.



X. Allamigeon, S. Gaubert, V. Magron, and B. Werner.
Certification of bounds of non-linear functions: the templates
method, *CICM Conference*, 2013.

In revision:



X. Allamigeon, S. Gaubert, V. Magron, and B. Werner.
Certification of Real Inequalities – Templates and Sums of
Squares, arxiv:1403.5899, 2014.

Introduction

Moment-SOS relaxations

Another look at Nonnegativity

New Applications of Moment-SOS Hierarchies

Semialgebraic Maxplus Optimization

Formal Nonlinear Optimization

Pareto Curves

Polynomial Images of Semialgebraic Sets

Program Analysis with Polynomial Templates

Conclusion

The General “Formal Framework”



We check the correctness of SOS certificates for **POP**



We build certificates to prove interval bounds for **semialgebraic** functions



We bound formally **transcendental** functions with semialgebraic approximations

Formal SOS bounds

When $q \in \mathcal{Q}(\mathbf{K})$, $\sigma_0, \dots, \sigma_m$ is a positivity certificate for q

Check **symbolic polynomial equalities** $q = q'$ in COQ



Existing tactic `ring` [Grégoire-Mahboubi 05]



Polynomials coefficients: arbitrary-size rationals `bigQ`
[Grégoire-Théry 06]





Much simpler to verify certificates using *sceptical approach*



Extends also to **semialgebraic** functions

Formal Nonlinear Optimization

Inequality	#boxes		
		Time	Time
9922699028	39	190 s	2218 s
3318775219	338	1560 s	19136 s

- Comparable with Taylor interval methods in HOL-LIGHT [Hales-Solovyev 13]



Bottleneck of informal optimizer is SOS solver



22 times slower! \implies Current bottleneck is to check polynomial equalities

For more details on the formal side:



X. Allamigeon, S. Gaubert, V. Magron and B. Werner. Formal Proofs for Nonlinear Optimization. Submitted for publication, arxiv:1404.7282

Introduction

Moment-SOS relaxations

Another look at Nonnegativity

New Applications of Moment-SOS Hierarchies

Semialgebraic Maxplus Optimization

Formal Nonlinear Optimization

Pareto Curves

Polynomial Images of Semialgebraic Sets

Program Analysis with Polynomial Templates

Conclusion

Bicriteria Optimization Problems

- Let $f_1, f_2 \in \mathbb{R}_d[\mathbf{x}]$ two conflicting criteria
- Let $\mathbf{S} := \{\mathbf{x} \in \mathbb{R}^n : g_1(\mathbf{x}) \geq 0, \dots, g_m(\mathbf{x}) \geq 0\}$ a semialgebraic set

$$(\mathbf{P}) \left\{ \min_{\mathbf{x} \in \mathbf{S}} (f_1(\mathbf{x}) \ f_2(\mathbf{x}))^\top \right\}$$

Assumption

The image space \mathbb{R}^2 is partially ordered in a natural way (\mathbb{R}_+^2 is the ordering cone).

Bicriteria Optimization Problems

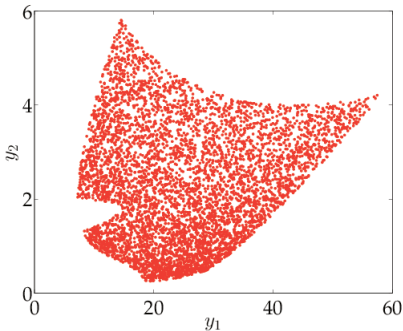
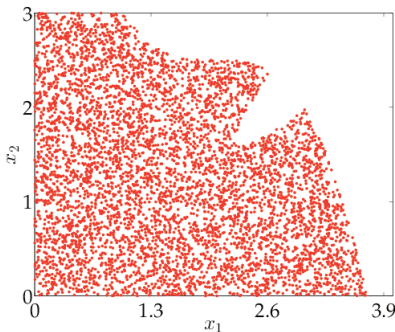
$$g_1 := -(x_1 - 2)^3/2 - x_2 + 2.5 ,$$

$$g_2 := -x_1 - x_2 + 8(-x_1 + x_2 + 0.65)^2 + 3.85 ,$$

$$\mathbf{S} := \{\mathbf{x} \in \mathbb{R}^2 : g_1(\mathbf{x}) \geq 0, g_2(\mathbf{x}) \geq 0\} .$$

$$f_1 := (x_1 + x_2 - 7.5)^2/4 + (-x_1 + x_2 + 3)^2 ,$$

$$f_2 := (x_1 - 1)^2/4 + (x_2 - 4)^2/4 .$$



Parametric sublevel set approximation

- Inspired by previous research on multiobjective linear optimization [Gorissen-den Hertog 12]
- Workaround: reduce \mathbf{P} to a **parametric POP**

$$(\mathbf{P}_\lambda) : \quad f^*(\lambda) := \min_{\mathbf{x} \in \mathbf{S}} \{ f_2(\mathbf{x}) : f_1(\mathbf{x}) \leq \lambda \} ,$$

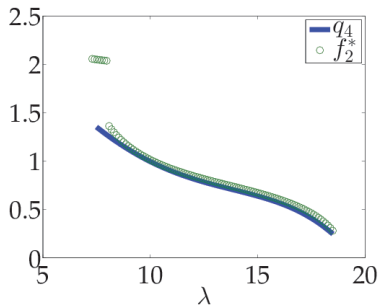
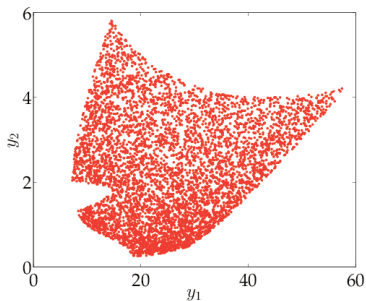
A Hierarchy of Polynomial underestimators

Moment-SOS approach [Lasserre 10]:

$$(D_d) \left\{ \begin{array}{ll} \max_{q \in \mathbb{R}_{2d}[\lambda]} & \sum_{k=0}^{2d} q_k / (1+k) \\ \text{s.t.} & f_2(\mathbf{x}) - q(\lambda) \in \mathcal{Q}_{2d}(\mathbf{K}) \end{array} \right. .$$

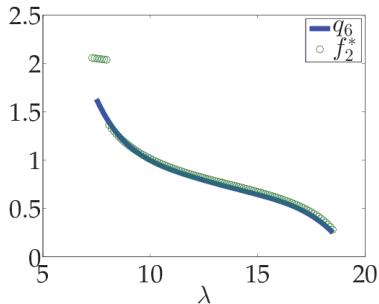
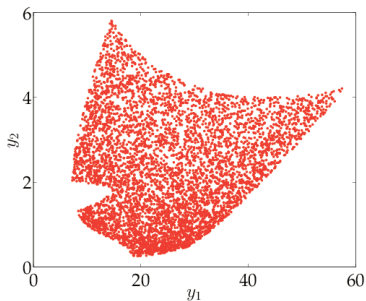
- The hierarchy (D_d) provides a sequence (q_d) of **polynomial underestimators** of $f^*(\lambda)$.
- $\lim_{d \rightarrow \infty} \int_0^1 (f^*(\lambda) - q_d(\lambda)) d\lambda = 0$

A Hierarchy of Polynomial underestimators



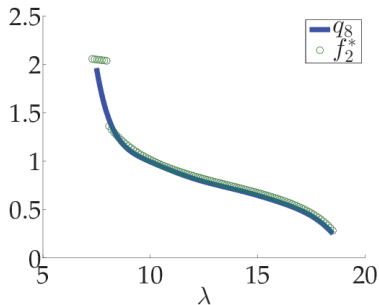
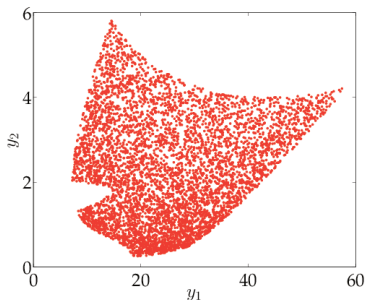
Degree 4

A Hierarchy of Polynomial underestimators



Degree 6

A Hierarchy of Polynomial underestimators



Degree 8

Contributions

- Numerical schemes that **avoid computing finitely many points**.
- Pareto curve approximation with polynomials, **convergence guarantees** in L_1 -norm



V. Magron, D. Henrion, J.B. Lasserre. Approximating Pareto Curves using Semidefinite Relaxations. *Operations Research Letters*. arxiv:1404.4772, April 2014.

Introduction

Moment-SOS relaxations

Another look at Nonnegativity

New Applications of Moment-SOS Hierarchies

Semialgebraic Maxplus Optimization

Formal Nonlinear Optimization

Pareto Curves

Polynomial Images of Semialgebraic Sets

Program Analysis with Polynomial Templates

Conclusion

Approximation of sets defined with “ \exists ”

Let $\mathbf{B} \subset \mathbb{R}^2$ be the unit ball and assume that $f(\mathbf{S}) \subset \mathbf{B}$.

- Another point of view:

$$f(\mathbf{S}) = \{\mathbf{y} \in \mathbf{B} : \exists \mathbf{x} \in \mathbf{S} \text{ s.t. } h(\mathbf{x}, \mathbf{y}) \leq 0\} ,$$

with

$$h(\mathbf{x}, \mathbf{y}) := \|\mathbf{y} - f(\mathbf{x})\|_2^2 = (y_1 - f_1(\mathbf{x}))^2 + (y_2 - f_2(\mathbf{x}))^2 .$$

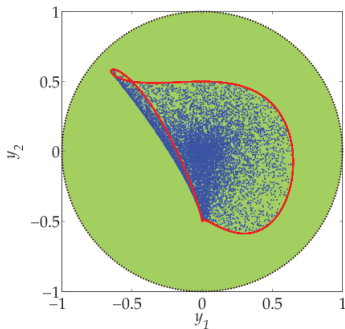
- Approximate $f(\mathbf{S})$ as closely as desired by a sequence of sets of the form :

$$\Theta_d := \{\mathbf{y} \in \mathbf{B} : q_d(\mathbf{y}) \leq 0\} ,$$

for some polynomials $q_d \in \mathbb{R}_{2d}[\mathbf{y}]$.

A Hierarchy of Outer approximations for $f(\mathbf{S})$

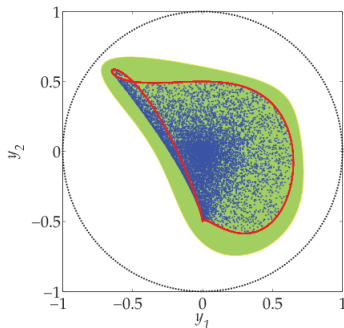
$$f(\mathbf{x}) := (x_1 + x_1x_2, x_2 - x_1^3)/2$$



Degree 2

A Hierarchy of Outer approximations for $f(\mathbf{S})$

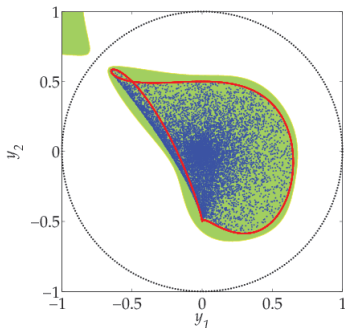
$$f(\mathbf{x}) := (x_1 + x_1x_2, x_2 - x_1^3)/2$$



Degree 4

A Hierarchy of Outer approximations for $f(\mathbf{S})$

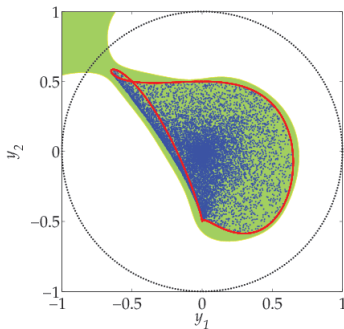
$$f(\mathbf{x}) := (x_1 + x_1x_2, x_2 - x_1^3)/2$$



Degree 6

A Hierarchy of Outer approximations for $f(\mathbf{S})$

$$f(\mathbf{x}) := (x_1 + x_1x_2, x_2 - x_1^3)/2$$



Degree 8

Introduction

Moment-SOS relaxations

Another look at Nonnegativity

New Applications of Moment-SOS Hierarchies

Semialgebraic Maxplus Optimization

Formal Nonlinear Optimization

Pareto Curves

Polynomial Images of Semialgebraic Sets

Program Analysis with Polynomial Templates

Conclusion

One-loop with Conditional Branching

■ $r, s, T^i, T^e \in \mathbb{R}[\mathbf{x}]$

■ $\mathbf{x}_0 \in \mathbf{X}_0$, with \mathbf{X}_0 semialgebraic set

```
 $\mathbf{x} = \mathbf{x}_0;$   
while ( $r(\mathbf{x}) \leq 0$ ) {  
  if ( $s(\mathbf{x}) \leq 0$ ) {  
     $\mathbf{x} = T^i(\mathbf{x});$   
  }  
  else {  
     $\mathbf{x} = T^e(\mathbf{x});$   
  }  
}
```

Bounding Template using SOS

Sufficient condition to get bounding inductive invariant:

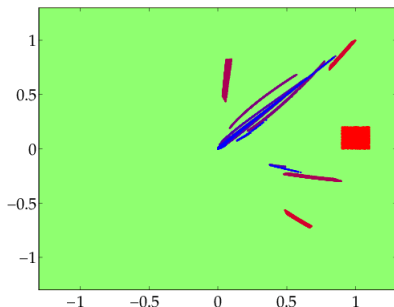
$$\begin{aligned} \alpha := \min_{q \in \mathbb{R}[\mathbf{x}]} \quad & \sup_{\mathbf{x} \in \mathbf{X}_0} q(\mathbf{x}) \\ \text{s.t.} \quad & q - q \circ T^i \geq 0, \\ & q - q \circ T^e \geq 0, \\ & q - \|\cdot\|_2^2 \geq 0. \end{aligned}$$

- Nontrivial correlations via polynomial templates $q(\mathbf{x})$
- $\{\mathbf{x} : q(\mathbf{x}) \leq \alpha\} \supset \bigcup_{k \in \mathbb{N}} \mathbf{X}_k$

Bounds for $\bigcup_{k \in \mathbb{N}} \mathbf{X}_k$

$$\mathbf{X}_0 := [0.9, 1.1] \times [0, 0.2] \quad r(\mathbf{x}) := 1 \quad s(\mathbf{x}) := 1 - x_1^2 - x_2^2$$

$$T^i(\mathbf{x}) := (x_1^2 + x_2^3, x_1^3 + x_2^2) \quad T^e(\mathbf{x}) := \left(\frac{1}{2}x_1^2 + \frac{2}{5}x_2^3, -\frac{3}{5}x_1^3 + \frac{3}{10}x_2^2\right)$$

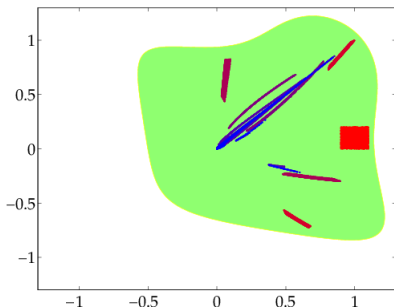


Degree 6

Bounds for $\bigcup_{k \in \mathbb{N}} \mathbf{X}_k$

$$\mathbf{X}_0 := [0.9, 1.1] \times [0, 0.2] \quad r(\mathbf{x}) := 1 \quad s(\mathbf{x}) := 1 - x_1^2 - x_2^2$$

$$T^i(\mathbf{x}) := (x_1^2 + x_2^3, x_1^3 + x_2^2) \quad T^e(\mathbf{x}) := \left(\frac{1}{2}x_1^2 + \frac{2}{5}x_2^3, -\frac{3}{5}x_1^3 + \frac{3}{10}x_2^2\right)$$

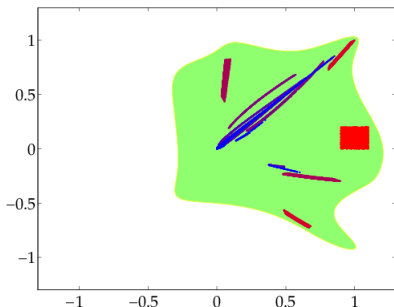


Degree 8

Bounds for $\bigcup_{k \in \mathbb{N}} \mathbf{X}_k$

$$\mathbf{X}_0 := [0.9, 1.1] \times [0, 0.2] \quad r(\mathbf{x}) := 1 \quad s(\mathbf{x}) := 1 - x_1^2 - x_2^2$$

$$T^i(\mathbf{x}) := (x_1^2 + x_2^3, x_1^3 + x_2^2) \quad T^e(\mathbf{x}) := \left(\frac{1}{2}x_1^2 + \frac{2}{5}x_2^3, -\frac{3}{5}x_1^3 + \frac{3}{10}x_2^2\right)$$



Degree 10

Contribution

For more details:



A. Adjé and V. Magron. Polynomial Template Generation using Sum-of-Squares Programming. Submitted for publication, arxiv:1409.3941

Introduction

Moment-SOS relaxations

Another look at Nonnegativity

New Applications of Moment-SOS Hierarchies

Conclusion

Conclusion

With **MOMENT-SOS HIERARCHIES**, you can

- Optimize nonlinear (transcendental) functions
- Approximate Pareto Curves, images and projections of semialgebraic sets
- Analyze programs

Conclusion

Further research:

- Alternative polynomial bounds using geometric programming (T. de Wolff, S. Ilmanen)
- Mixed LP/SOS certificates (trade-off CPU/precision)

End

Thank you for your attention!

<http://homepages.laas.fr/vmagron/>