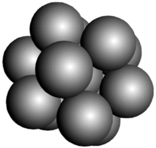
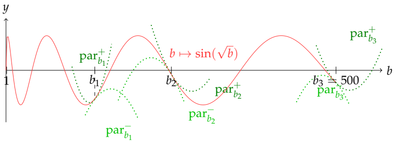
Formal Proofs, Program Analysis and Moment-SOS Relaxations

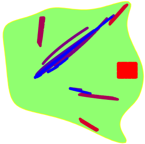
Victor Magron, Postdoc LAAS-CNRS

15 July 2014

Imperial College
Department of Electrical and Electronic Eng.







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

Computer Science and Mathematics

- PhD on Formal Proofs for Global Optimization: Templates and Sums of Squares
- Collaboration with:

  LIX Benjamin Werner (LIX Polytechnique)

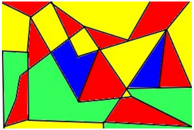
  CMAP Stéphane Gaubert (Maxplus Team CMAP/INRIA Polytechnique)

  CMAP Xavier Allamigeon (Maxplus Team)

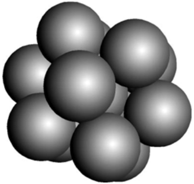
Complex Proofs

- Complex mathematical proofs / mandatory computation

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



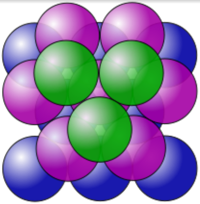
T. Hales, A Proof of the Kepler Conjecture, 1994.



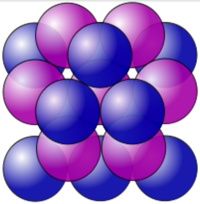
From Oranges Stack...

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]: **F**ormal **P**roof of **K**epler Conjecture

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 $|\cdot|, \sqrt{}, +, -, \times, /, \sup, \inf, \dots$

$$p(\mathbf{x}) := \partial_4 \Delta \mathbf{x} \qquad q(\mathbf{x}) := 4 x_1 \Delta \mathbf{x}$$
$$r(\mathbf{x}) := p(\mathbf{x}) / \sqrt{q(\mathbf{x})}$$

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

A “Simple” Example

In the computational part:

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

A "Simple" Example

In the computational part:

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

Lemma$_{9922699028}$ from Flyspeck:

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

New Framework (in my PhD thesis)

- Certificates for lower bounds of Global Optimization Problems using SOS and new ingredients in Global Optimization:
  - Maxplus approximation (Optimal Control)
  - Nonlinear templates (Static Analysis)
- Verification of these certificates inside Coq
- Implementation: `NLCertify`
  `http://nl-certify.forge.ocamlcore.org/`

Moment-SOS relaxations

- Semialgebraic set $\mathbf{K} := \{\mathbf{x} \in \mathbb{R}^n : g_1(\mathbf{x}) \geqslant 0, \dots, g_m(\mathbf{x}) \geqslant 0\}$
- $p^* := \min\limits_{\mathbf{x} \in \mathbf{K}} p(\mathbf{x})$: NP hard
- Sums of squares $\Sigma[\mathbf{x}]$
- $\mathcal{Q}(\mathbf{K}) := \Big\{ \sigma_0(\mathbf{x}) + \sum_{j=1}^m \sigma_j(\mathbf{x}) g_j(\mathbf{x}), \text{ with } \sigma_j \in \Sigma[\mathbf{x}] \Big\}$

Moment-SOS relaxations

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

Polynomial Optimization Problems (POP)

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

Moment-SOS relaxations

- Truncated quadratic module $\mathcal{Q}_k(\mathbf{K}) := \mathcal{Q}(\mathbf{K}) \cap \mathbb{R}_{2k}[\mathbf{x}]$

Polynomial Optimization Problems (POP)

(Moment) $\qquad\qquad$ (SOS)

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

Practical Computation

- Hierarchy of SOS relaxations:
  $\lambda_k := \sup_{\lambda} \Big\{ \lambda : p - \lambda \in \mathcal{Q}_k(\mathbf{K}) \Big\}$

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

- Can be computed with SOS solvers (CSDP, SDPA)

- Extension to semialgebraic functions $r(\mathbf{x}) = p(\mathbf{x}) / \sqrt{q(\mathbf{x})}$ [Lasserre-Putinar 10]

The 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 $f^* \geqslant 0$

- 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)

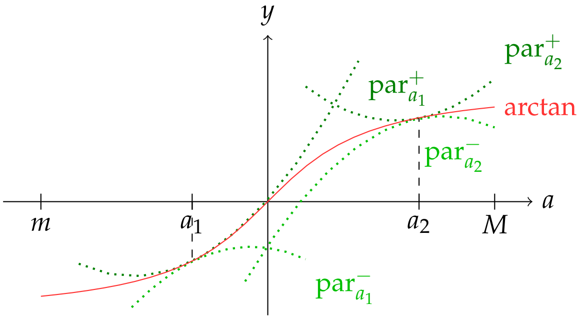
Maxplus Approximation

- Initially introduced to solve Optimal Control Problems [Fleming-McEneaney 00]
- **Curse of dimensionality** reduction [McEaneney 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

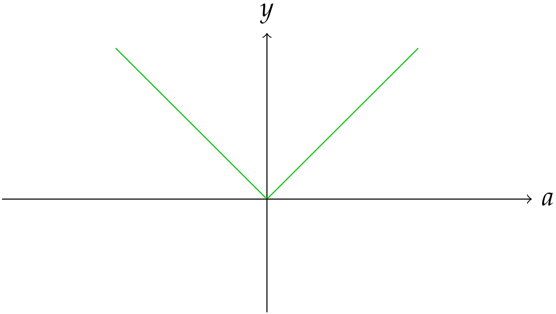
Definition

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



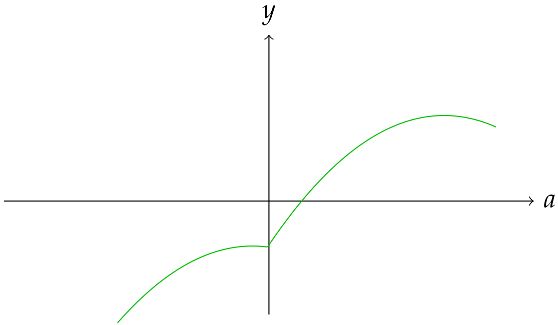
Nonlinear Function Representation

Exact parsimonious maxplus representations



Nonlinear Function Representation

Exact parsimonious maxplus representations



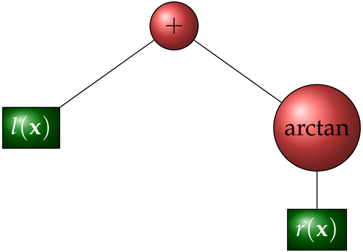
Nonlinear Function Representation

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

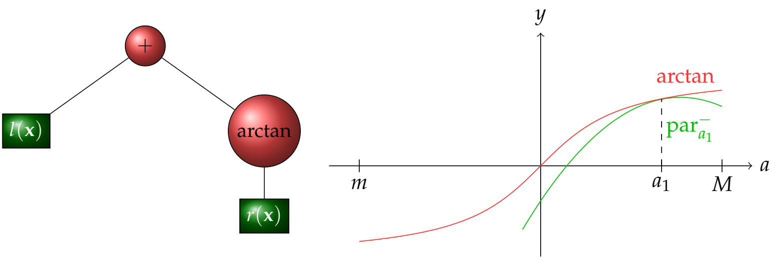
Nonlinear Function Representation

- For the “Simple” Example from Flyspeck:



Maxplus Optimization Algorithm

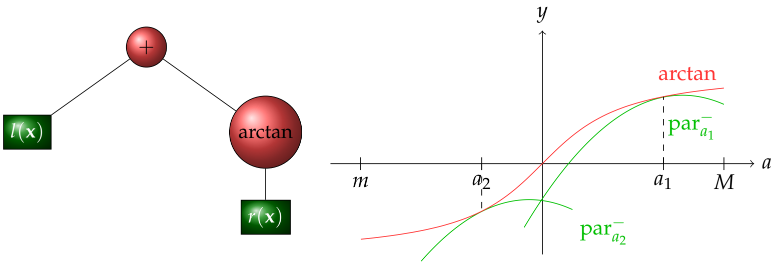
First iteration:



1. 1 control point $\{a_1\}$ SOS Computation: $m_1 = -0.746$

Maxplus Optimization Algorithm

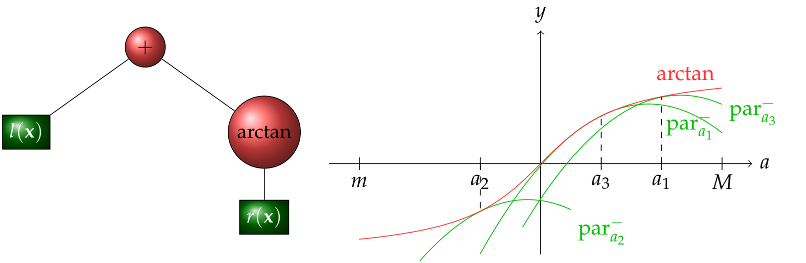
Second iteration:



2. 2 control points $\{a_1, a_2\}$: $m_2 = -0.112$

Maxplus Optimization Algorithm

Third iteration:



3. 3 control points $\{a_1, a_2, a_3\}$: $m_3 = -0.04$

Contributions

For more details:

- X. Allamigeon, S. Gaubert, V. Magron, and B. Werner. Certification of inequalities involving transcendental functions: combining sdp and max-plus approximation. In *Proceedings of the European Control Conference (ECC) Zurich*, pages 2244-2250, 2013.

- X. Allamigeon, S. Gaubert, V. Magron, and B. Werner. Certification of bounds of non-linear functions: the templates method. In *Proceedings of Conferences on Intelligent Computer Mathematics, CICM Calculemus, Bath*, pages 51-65. LNAI 7961 Springer, 2013.

In revision:

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

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** $\boxed{q = q'}$ 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

Benchmarks for Flyspeck Inequalities

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! $\Longrightarrow$ Current bottleneck is to check polynomial equalities

Contribution: Publications and Software

For more details:

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

X. Allamigeon, S. Gaubert, V. Magron, and B. Werner. Certification of bounds of non-linear functions: the templates method. In *Proceedings of Conferences on Intelligent Computer Mathematics, CICM Calculemus, Bath*, pages 51-65. LNAI 7961 Springer, 2013.

Contribution: Publications and Software

Software Implementation `NLCertify`**:**

- `https://forge.ocamlcore.org/projects/nl-certify/`

15 000 lines of OCAML code

4000 lines of COQ code

V. Magron NLCertify: A Tool for Formal Nonlinear Optimization. To appear in the *Proceedings of the 4th International Congress on Mathematical Software*, ICMS 2014, Séoul, arxiv:1405.5668

Postdoc Research

1. Approximating Pareto curves, image of semialgebraic sets. With people from LAAS-CNRS:
   - Didier Henrion
   - Jean-Bernard Lasserre
2. Static analysis. With people from Onera:
   - Assalé Adjé
   - Pierre-Loic Garoche

Bicriteria Optimization Problems

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

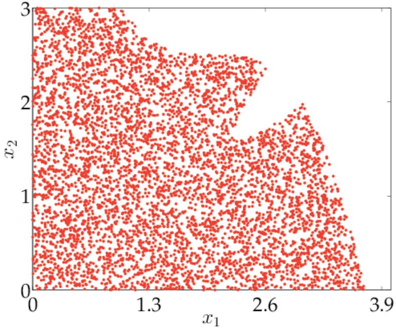
$g_1 := -(x_1 - 2)^3/2 - x_2 + 2.5$,

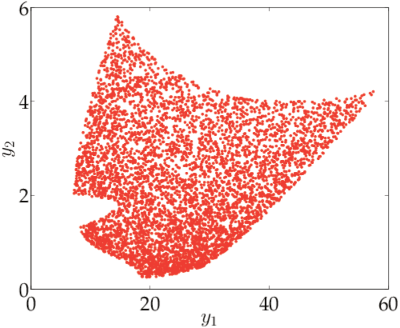
$g_2 := -x_1 - x_2 + 8(-x_1 + x_2 + 0.65)^2 + 3.85$,

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

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

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





Parametric sublevel set approximation

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

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

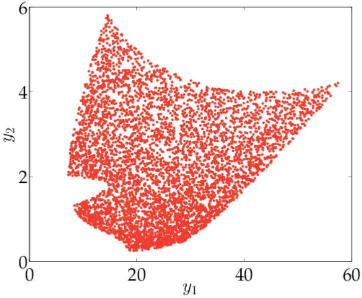
A Hierarchy of Polynomial underestimators

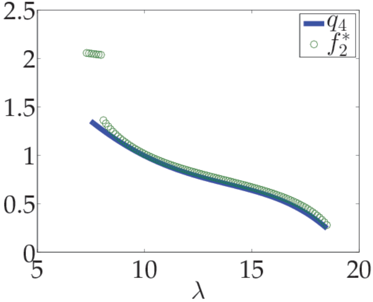
Moment-SOS approach [Lasserre 10]:

$$(D_d) \begin{cases} \max\limits_{q \in \mathbb{R}_{2d}[\lambda]} & \sum\limits_{k=0}^{2d} q_k / (1+k) \\ \text{s.t.} & f_2(\mathbf{x}) - q(\lambda) \in \mathcal{Q}_{2d}(\mathbf{K}) \end{cases} .$$

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

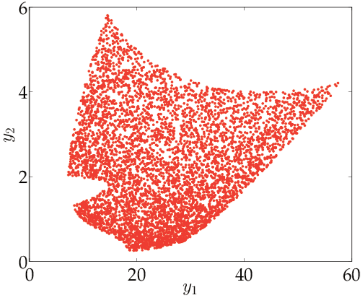
A Hierarchy of Polynomial underestimators

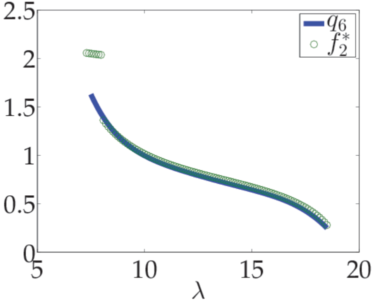




Degree 4

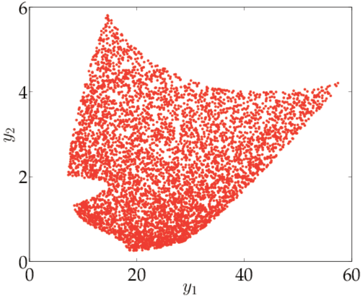
A Hierarchy of Polynomial underestimators

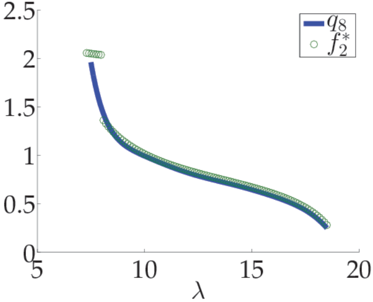




Degree 6

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. Accepted pending minor revisions in *Operations Research Letters*. arxiv:1404.4772, April 2014.

Approximation of sets defined with "∃"

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

- Another point of view:

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

  with

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

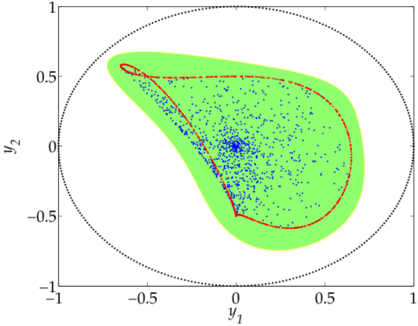
- Approximate $f(\mathbf{S})$ as closely as desired by a sequence of sets of the form :

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

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

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

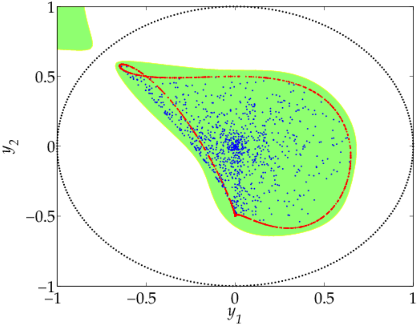
$$f(\mathbf{x}) := (x_1 + x_1 x_2, x_2 - x_1^3)/2$$



Degree 4

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

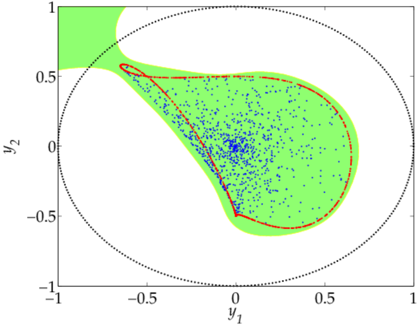
$$f(\mathbf{x}) := (x_1 + x_1 x_2, x_2 - x_1^3)/2$$



Degree 6

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

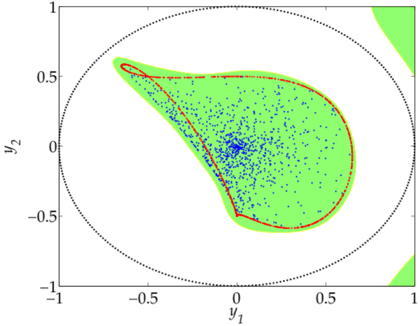
$$f(\mathbf{x}) := (x_1 + x_1 x_2, x_2 - x_1^3)/2$$



Degree 8

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

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



Degree 10

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

```
x = x0;
while (r(x) ⩽ 0){
   if (s(x) ⩽ 0){
         x = T^i(x);
         }
   else{
         x = T^e(x);
         }
}
```

Bounding Template using SOS

Sufficient condition to get bounding inductive invariant:

$$\alpha := \min_{q \in \mathbb{R}[\mathbf{x}]} \quad \sup_{\mathbf{x} \in \mathbf{X}_0} q(\mathbf{x})$$

$$\text{s.t.} \quad q - q \circ T^i \geqslant 0 \ ,$$

$$q - q \circ T^e \geqslant 0 \ ,$$

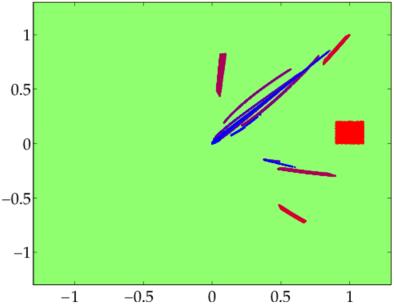
$$q - \| \cdot \|_2^2 \geqslant 0 \ .$$

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

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

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

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

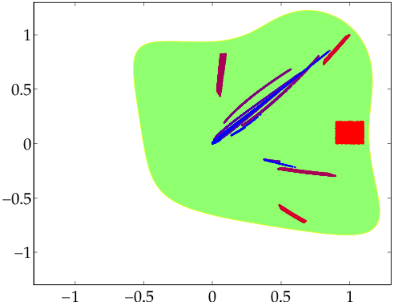


Degree 6

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

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

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

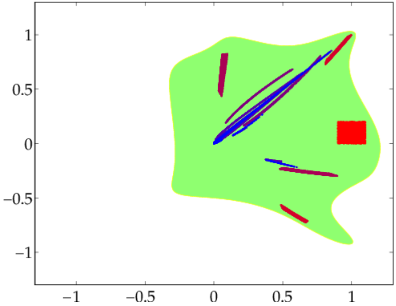


Degree 8

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

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

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



Degree 10

Conclusion

- New framework for nonlinear optimization
- Formal nonlinear optimization: `NLCertify`
- Approximation of Pareto Curves, images and projections of semialgebraic sets
- Program Analysis with polynomial templates

Conclusion

Further research:

- Improve formal polynomial checker
- Alternative Polynomials bounds using geometric programming (T. de Wolff, S. Iliman)
- Programs analysis with transcendental assignments/conditions

End

Thank you for your attention!

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