Formal Proofs, Program Analysis and Moment-SOS Relaxations

Victor Magron, Postdoc LAAS-CNRS

15 July 2014

Imperial College
Department of Electrical and Electronic Eng.
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, ...
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:
  - Benjamin Werner (LIX Polytechnique)
  - Stéphane Gaubert (Maxplus Team CMAP/INRIA Polytechnique)
  - Xavier Allamigeon (Maxplus Team)
Complex Proofs

- Complex mathematical proofs / mandatory computation

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

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
A “Simple” Example

In the computational part:

- **Multivariate Polynomials:**

  \[ \Delta x := x_1 x_4 (-x_1 + x_2 + x_3 - x_4 + x_5 + x_6) + x_2 x_5 (x_1 - x_2 + x_3 + x_4 - x_5 + x_6) + x_3 x_6 (x_1 + x_2 - x_3 + x_4 + x_5 - x_6) - x_2 (x_3 x_4 + x_1 x_6) - x_5 (x_1 x_3 + x_4 x_6) \]
A “Simple” Example

In the computational part:

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

$$p(x) := \partial_4 \Delta x \quad q(x) := 4x_1 \Delta x$$

$$r(x) := p(x) / \sqrt{q(x)}$$

$$l(x) := -\frac{\pi}{2} + 1.6294 - 0.2213 \left( \sqrt{x_2} + \sqrt{x_3} + \sqrt{x_5} + \sqrt{x_6} - 8.0 \right) + 0.913 \left( \sqrt{x_4} - 2.52 \right) + 0.728 \left( \sqrt{x_1} - 2.0 \right)$$
A “Simple” Example

In the computational part:

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

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

Lemma\textsubscript{9922699028} from Flyspeck:

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

**Moment-SOS relaxations and Maxplus approximation**

Formal Nonlinear Optimization

Pareto Curves and Images of Semialgebraic Sets

Program Analysis with Polynomial Templates

Conclusion
Semialgebraic set \( K := \{ x \in \mathbb{R}^n : g_1(x) \geq 0, \ldots, g_m(x) \geq 0 \} \)

- \( p^* := \min_{x \in K} p(x) \): NP hard

- Sums of squares \( \Sigma[x] \)

- \( Q(K) := \left\{ \sigma_0(x) + \sum_{j=1}^{m} \sigma_j(x)g_j(x), \text{ with } \sigma_j \in \Sigma[x] \right\} \)
Moment-SOS relaxations

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

Polynomial Optimization Problems (POP)

$\begin{align*}
\text{(Primal)} & \quad \inf \int_K p \, d\mu = \sup \lambda \\
\text{s.t.} & \quad \mu \in \mathcal{M}_+(K) \quad \text{s.t.} \quad \lambda \in \mathbb{R}, \\
& \quad p - \lambda \in \mathcal{Q}(K)
\end{align*}$
Moment-SOS relaxations

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

Polynomial Optimization Problems (POP)

(Moment) \hspace{1cm} (SOS)

$$\inf \int_K p \, d\mu \quad \geq \quad \sup \lambda$$

s.t. $\mu \in M_+(K)$ \hspace{1cm} s.t. $\lambda \in \mathbb{R}$, $p - \lambda \in Q_k(K)$
Hierarchy of SOS relaxations:
\[ \lambda_k := \sup_{\lambda} \left\{ \lambda : p - \lambda \in Q_k(K) \right\} \]

Convergence guarantees \( \lambda_k \uparrow p^* \) [Lasserre 01]

Can be computed with SOS solvers (CSDP, SDPA)

Extension to semialgebraic functions \( r(x) = p(x) / \sqrt{q(x)} \)
[Lasserre-Putinar 10]
The General “Informal Framework”

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

- \( f \) is underestimated by a \textit{semialgebraic} function \( f_{sa} \)
- Reduce the problem \( f_{sa}^* := \inf_{x \in K} f_{sa}(x) \) to a \textit{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 $\gamma \geq 0$. A function $\phi : \mathbb{R}^n \rightarrow \mathbb{R}$ is said to be $\gamma$-semiconvex if the function $x \mapsto \phi(x) + \frac{\gamma}{2} \|x\|^2$ is convex.
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 $A$

- nodes are univariate functions of $D$ or binary operations
For the “Simple” Example from Flyspeck:

\[ l(x) + \arctan(r(x)) \]
Maxplus Optimization Algorithm

First iteration:

1 control point \( \{a_1\} \) SOS Computation: \( m_1 = -0.746 \)
Maxplus Optimization Algorithm

Second iteration:

2 control points \( \{a_1, a_2\} \): \( m_2 = -0.112 \)
Third iteration:

\[ l(x) + \arctan(r(x)) \]

3 control points \( \{a_1, a_2, a_3\} \): 
\[ m_3 = -0.04 \]
Contributions

For more details:


In revision:

Introduction

Moment-SOS relaxations and Maxplus approximation

**Formal Nonlinear Optimization**

Pareto Curves and 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 $q \in \mathbb{Q}(K)$, $\sigma_0, \ldots, \sigma_m$ 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 $\text{bigQ}$ [Grégoire-Théry 06]

- Much simpler to verify certificates using **sceptical approach**

- Extends also to **semialgebraic functions**
## Benchmarks for Flyspeck Inequalities

<table>
<thead>
<tr>
<th>Inequality</th>
<th>#boxes</th>
<th>Time</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>9922699028</td>
<td>39</td>
<td>190s</td>
<td>2218s</td>
</tr>
<tr>
<td>3318775219</td>
<td>338</td>
<td>1560s</td>
<td>19136s</td>
</tr>
</tbody>
</table>

- 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
Contribution: Publications and Software

For more details:


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

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
Introduction

Moment-SOS relaxations and Maxplus approximation

Formal Nonlinear Optimization

Pareto Curves and Images of Semialgebraic Sets

Program Analysis with Polynomial Templates

Conclusion
Bicriteria Optimization Problems

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

\[
\begin{align*}
(P) \left\{ \min_{x \in S} (f_1(x) f_2(x))^\top \right\}
\end{align*}
\]

**Assumption**

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

\[ g_1 := -(x_1 - 2)^3/2 - x_2 + 2.5 , \]
\[ g_2 := -x_1 - x_2 + 8(-x_1 + x_2 + 0.65)^2 + 3.85 , \]
\[ S := \{ x \in \mathbb{R}^2 : g_1(x) \geq 0, g_2(x) \geq 0 \} . \]

\[ f_1 := (x_1 + x_2 - 7.5)^2/4 + (-x_1 + x_2 + 3)^2 , \]
\[ f_2 := (x_1 - 1)^2/4 + (x_2 - 4)^2/4 . \]
Parametric sublevel set approximation

- Inspired by previous research on multiobjective linear optimization [Gorissen-den Hertog 12]

- Workaround: reduce $P$ to a parametric POP

$$(P_{\lambda}) : \quad f^*(\lambda) := \min_{x \in S} \{ f_2(x) : f_1(x) \leq \lambda \} ,$$
A Hierarchy of Polynomial underestimators

Moment-SOS approach [Lasserre 10]:

\[
(D_d) \begin{cases} 
\max_{q \in \mathbb{R}_{2d}[\lambda]} & \sum_{k=0}^{2d} q_k / (1 + k) \\
\text{s.t.} & f_2(x) - q(\lambda) \in Q_{2d}(K).
\end{cases}
\]

- 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

![Diagram showing a scatter plot and a line graph representing degree 4 polynomial underestimators. The scatter plot on the left has a red distribution with axes labeled $y_2$ and $y_1$. The line graph on the right has a blue line labeled $q_4$ and green circles labeled $f_2^*$. The x-axis is labeled $\lambda$ with values 5 to 20 and the y-axis is labeled from 0 to 2.5.]
A Hierarchy of Polynomial underestimators

Degree 6
A Hierarchy of Polynomial underestimators

Degree 8

Degree 8
Contributions

- Numerical schemes that avoid computing finitely many points.

- Pareto curve approximation with polynomials, convergence guarantees in $L_1$-norm

Approximation of sets defined with “∃”

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

- Another point of view:

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

with

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

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

$$\mathcal{G}_d := \{ y \in \mathcal{B} : q_d(y) \leq 0 \} ,$$

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


A Hierarchy of Outer approximations for $f(S)$

$$f(x) := \left( x_1 + x_1 x_2, x_2 - x_1^3 \right) / 2$$

Degree 4
A Hierarchy of Outer approximations for $f(S)$

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

Degree 6
A Hierarchy of Outer approximations for $f(S)$

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

Degree 8
A Hierarchy of Outer approximations for $f(S)$

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

Moment-SOS relaxations and Maxplus approximation

Formal Nonlinear Optimization

Pareto Curves and Images of Semialgebraic Sets

Program Analysis with Polynomial Templates

Conclusion
One-loop with Conditional Branching

- \( r, s, T^i, T^e \in \mathbb{R}[x] \)

- \( x_0 \in X_0 \), with \( X_0 \) semialgebraic set

\[
\begin{aligned}
x &= x_0; \\
\text{while } (r(x) \leq 0) \{ \\
&\quad \text{if } (s(x) \leq 0) \{ \\
&\quad\quad x = T^i(x); \\
&\quad \} \\
&\quad \text{else} \{ \\
&\quad\quad x = T^e(x); \\
&\quad \} \\
&\}\end{aligned}
\]
Bounding Template using SOS

Sufficient condition to get bounding inductive invariant:

\[ \alpha := \min_{q \in \mathbb{R}[x]} \sup_{x \in X_0} q(x) \]

s.t. \[ q - q \circ T^i \geq 0 , \]
    \[ q - q \circ T^e \geq 0 , \]
    \[ q - \| \cdot \|_2^2 \geq 0 . \]

- Nontrivial correlations via polynomial templates \( q(x) \)
  \[ \{ x : q(x) \leq \alpha \} \supset \bigcup_{k \in \mathbb{N}} X_k \]
Bounds for $\bigcup_{k \in \mathbb{N}} X_k$

\[ X_0 := [0.9, 1.1] \times [0, 0.2] \quad r(x) := 1 \quad s(x) := 1 - x_1^2 - x_2^2 \]

\[ T^i(x) := (x_1^2 + x_2^3, x_1^3 + x_2^2) \quad T^e(x) := \left( \frac{1}{2} x_1^2 + \frac{2}{5} x_2^3, -\frac{3}{5} x_1^3 + \frac{3}{10} x_2^2 \right) \]
Bounds for $\bigcup_{k \in \mathbb{N}} X_k$

$$X_0 := [0.9, 1.1] \times [0, 0.2] \quad r(x) := 1 \quad s(x) := 1 - x_1^2 - x_2^2$$

$$T^i(x) := (x_1^2 + x_2^3, x_1^3 + x_2^2) \quad T^e(x) := \left( \frac{1}{2} x_1^2 + \frac{2}{5} x_2^3, -\frac{3}{5} x_1^3 + \frac{3}{10} x_2^2 \right)$$
Bounds for $\bigcup_{k \in \mathbb{N}} X_k$

$X_0 := [0.9, 1.1] \times [0, 0.2]$ \hspace{1cm} $r(x) := 1$ \hspace{1cm} $s(x) := 1 - x_1^2 - x_2^2$

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

Moment-SOS relaxations and Maxplus approximation

Formal Nonlinear Optimization

Pareto Curves and Images of Semialgebraic Sets

Program Analysis with Polynomial Templates

Conclusion
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
Thank you for your attention!

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