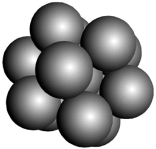
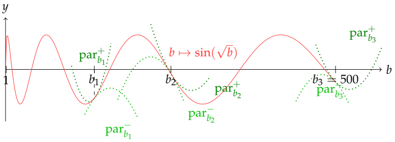
NLCertify: A Tool for Formal Nonlinear Optimization

Victor Magron, Postdoc LAAS-CNRS

18 September 2014

Aric Seminar
Lyon





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

- Tool: Formal Bounds for Global Optimization

- Collaboration with:

  Benjamin Werner (LIX Polytechnique)

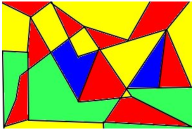
  Stéphane Gaubert (Maxplus Team CMAP/INRIA Polytechnique)

  Xavier Allamigeon (Maxplus Team)

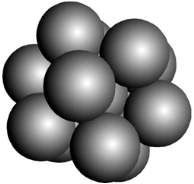
Complex Proofs

- Complex mathematical proofs / mandatory computation

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



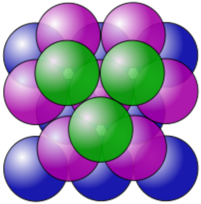
T. Hales, A Proof of the Kepler Conjecture, 1994.



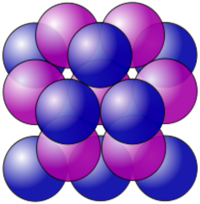
From Oranges Stack...

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."
- **F**lys**p**ec**k** [Hales 06]: **F**ormal **P**roof of **K**epler 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]: **F**ormal **P**roof of **K**epler Conjecture
- **Project Completion on 10 August by the Flyspeck team!!**

...to Flyspeck Nonlinear Inequalities

- Nonlinear inequalities: quantified reasoning with "$\forall$"

$$\forall \mathbf{x} \in \mathbf{K}, f(\mathbf{x}) \geqslant 0$$

- NP-hard optimization problem

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

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

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

A "Simple" Example

In the computational part:

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

A "Simple" Example

In the computational part:

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

Lemma$_{9922699028}$ from Flyspeck:

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

Existing Formal Frameworks

Formal proofs for Global Optimization:

- Bernstein polynomial methods [Zumkeller's PhD 08]
- SMT methods [Gao et al. 12]
- Interval analysis and Sums of squares

Existing Formal Frameworks

Interval analysis

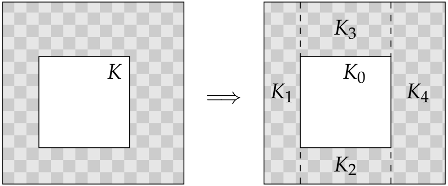
- Certified interval arithmetic in Coq [Melquiond 12]
- Taylor methods in HOL Light [Solovyev thesis 13]
  - Formal verification of floating-point operations
- robust but subject to the **Curse of Dimensionality**

Existing Formal Frameworks

Lemma$_{9922699028}$ from Flyspeck:

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

- 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



Existing Formal Frameworks

Sums of squares techniques

- Formalized in HOL-LIGHT [Harrison 07] COQ [Besson 07]
- Precise methods but scalibility and robustness issues (numerical)
- powerful: global optimality certificates without branching

but

- not so robust: handles moderate size problems
- Restricted to polynomials

Existing Formal Frameworks

Approximation theory: Chebyshev/Taylor models

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

Existing Formal Frameworks

Can we develop a new approach with both keeping the respective strength of interval and precision of SOS?

Proving Flyspeck Inequalities is challenging: medium-size and tight

New Framework (in my PhD thesis)

- Certificates for lower bounds of Nonlinear optimization using:
  - Moment-SOS hierarchies
  - Maxplus approximation (Optimal Control)
- Verification of these certificates inside COQ

New Framework (in my PhD thesis)

Software Implementation `NLCertify`:

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

15 000 lines of OCAML code

4000 lines of COQ code

Polynomial Optimization Problems

- Semialgebraic set $\mathbf{K} := \{\mathbf{x} \in \mathbb{R}^n : g_1(\mathbf{x}) \geqslant 0, \dots, g_m(\mathbf{x}) \geqslant 0\}$
- $p^* := \min_{\mathbf{x} \in \mathbf{K}} p(\mathbf{x})$: 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 $\mathbf{K}' := [-1, 1]^n$
  $\mathbf{K}' := \{\mathbf{x} \in \mathbb{R}^n : g_1 := 1 - x_1^2 \geqslant 0, \cdots, g_n := 1 - x_n^2 \geqslant 0\}$
- $n - \|\mathbf{x}\|_2^2$ 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. $\mu : \mathcal{B} \to [0, 1]$, $\mu(\emptyset) = 0$, $\mu(\mathbb{R}^n) < \infty$
  2. $\mu(\bigcup_i B_i) = \sum_i \mu(B_i)$, for any countable $(B_i) \subset \mathcal{B}$
  3. $\int_{\mathbf{K}} \mu(d\mathbf{x}) = 1$

- $\mathrm{supp}(\mu)$ is the smallest set $\mathbf{K}$ such that $\mu(\mathbb{R}^n \backslash \mathbf{K}) = 0$

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

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

Lasserre's Hierarchy of SDP relaxations

- Moment matrix
  $\mathbf{M}(\mathbf{y})_{u,v} := L_{\mathbf{y}}(u \cdot v), \; u, v$ monomials

- Localizing matrix $M(g_j \, \mathbf{y})$ associated with g_j
  $\mathbf{M}(g_j \, \mathbf{y})_{u,v} := L_{\mathbf{y}}(u \cdot v \cdot g_j), \; u, v$ 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:

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

Lasserre's Hierarchy of SDP relaxations

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

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

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

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:

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

Semidefinite Optimization

- F_0, F_α symmetric real matrices, cost vector c

Primal-dual pair of semidefinite programs:

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

- 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)

$$\begin{array}{llcll}
\multicolumn{2}{c}{\text{(Primal)}} & & \multicolumn{2}{c}{\text{(Dual)}} \\
\inf & \displaystyle\int_{\mathbf{K}} p \, d\mu & = & \sup & \lambda \\
\text{s.t.} & \mu \in \mathcal{M}_+(\mathbf{K}) & & \text{s.t.} & \lambda \in \mathbb{R} \,, \\
 & & & & p - \lambda \in \mathcal{Q}(\mathbf{K})
\end{array}$$

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)

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

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]

- 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

  $\forall \mathbf{x} \in [-0.5, 0.5]^4, -x_1 x_3^3 + 4x_2 x_3^2 x_4 + 4x_1 x_3 x_4^2 + 2x_2 x_4^3 + 4x_1 x_3 + 4x_3^2 - 10x_2 x_4 - 10x_4^2 + 5.1801 \geqslant 0.$

- `scale_pol = true`: scaled on $[0, 1]^4$

- `relax_order = 2`: SOS of degree at most 4

- `bound_squares_variables = true`:
  redundant constraints $x_1^2 \leqslant 1, \ldots, x_4^2 \leqslant 1$

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

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]

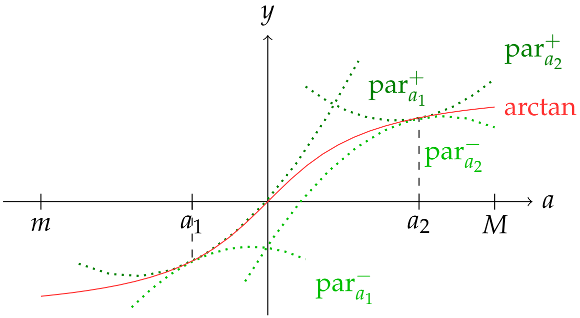
- **Curse of dimensionality** reduction [McEaneney Kluberg, Gaubert-McEneaney-Qu 11, Qu 13].
  Allowed to solve instances of dim up to 15 (inaccessible by grid methods)

- In our context: approximate transcendental functions

Maxplus Approximation

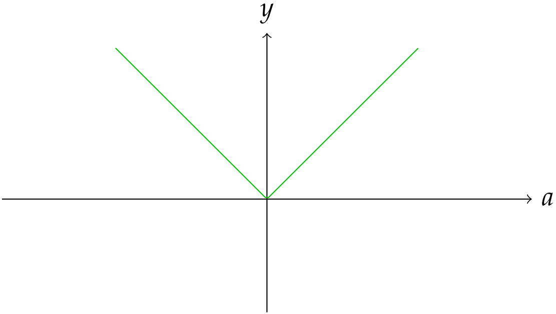
Definition

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.



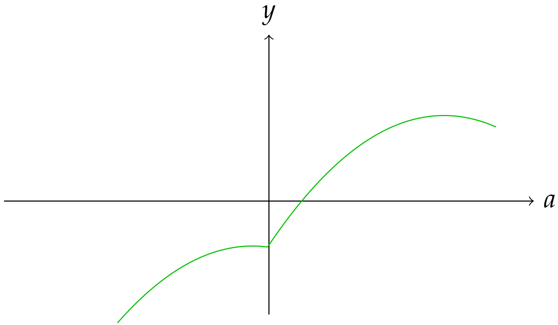
Nonlinear Function Representation

Exact parsimonious maxplus representations



Nonlinear Function Representation

Exact parsimonious maxplus representations



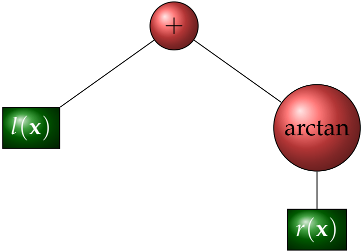
Nonlinear Function Representation

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

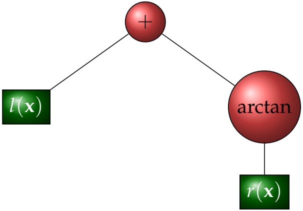
Nonlinear Function Representation

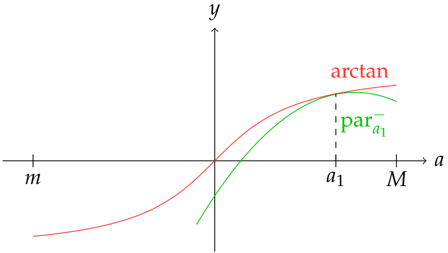
- For the "Simple" Example from Flyspeck:



Maxplus Optimization Algorithm

First iteration:

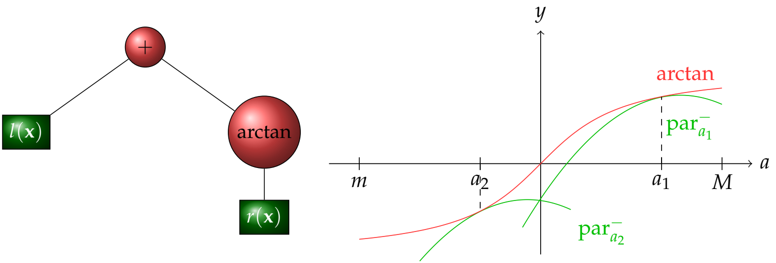




1 control point $\{a_1\}$: $m_1 = -4.7 \times 10^{-3} < 0$

Maxplus Optimization Algorithm

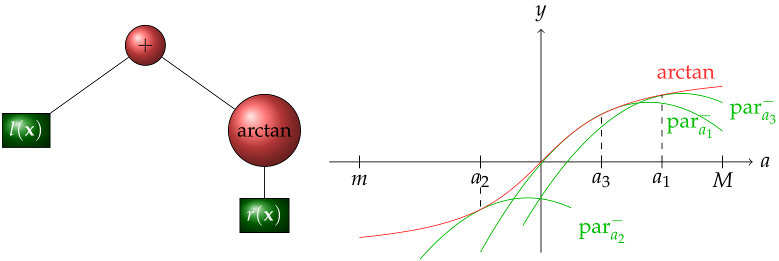
Second iteration:



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!

Maxplus Optimization Algorithm

Input: tree t, box $\mathbf{K}$, SOS relaxation order k, precision p
Output: bounds m and M, approximations t_2^- and t_2^+

if $t \in \mathcal{A}$ **then** $t^- := t, t^+ := t$
else if $u := \texttt{root}(t) \in \mathcal{D}$ with child c **then**
  $m_c, M_c, c^-, c^+ := \texttt{samp\_approx}(c, \mathbf{K}, k, p)$
  $I := [m_c, M_c]$
  $u^-, u^+ := \texttt{unary\_approx}(u, I, c, p)$
  $t^-, t^+ := \texttt{compose\_approx}(u, u^-, u^+, I, c^-, c^+)$
else if $\texttt{bop} := \texttt{root}(t)$ with children c_1 and c_2 **then**
  $m_i, M_i, c_i^-, c_i^+ := \texttt{samp\_approx}(c_i, \mathbf{K}, k, p)$ for $i \in \{1, 2\}$
  $t^-, t^+ := \texttt{compose\_bop}(c_1^-, c_1^+, c_2^-, c_2^+, \texttt{bop}, [m_2, M_2])$
end
return $\texttt{min\_sa}(t^-, \mathbf{K}, k), \texttt{max\_sa}(t^+, \mathbf{K}, k), t^-, t^+$

Minimax Approximation / For Comparison

- The precision is an integer d
- The best-uniform degree-d polynomial approximation of u:

$$\min_{h \in \mathbb{R}_d[x]} \|u - h\|_\infty = \min_{h \in \mathbb{R}_d[x]} \left(\sup_{x \in I} |u(x) - h(x)|\right)$$

- Implementation in `Sollya` [Chevillard-Joldes-Lauter 10]
- Interface of `NLCertify` with `Sollya`

High-degree Polynomial Approximation + SOS

SWF: $\min_{\mathbf{x} \in [1,500]^n} f(\mathbf{x}) = -\sum_{i=1}^{n} x_i \sin(\sqrt{x_i})$

- replace $\sin(\sqrt{\cdot})$ by a degree-d Chebyshev polynomial
- Hard to combine with SOS

High-degree Polynomial Approximation + SOS

Indeed:

- Small d: lack of accuracy $\Longrightarrow$ expensive Branch and Bound

- Large d: "No free lunch" rule with $\binom{n+d}{n}$ SDP variables

High-degree Polynomial Approximation + SOS

SWF with $n = 10, d = 4$:

- $38\,min$ to compute a lower bound of $-430n$

Comparison on Global Optimization Problems

$$\min_{\mathbf{x}\in[1,500]^n} f(\mathbf{x}) = -\sum_{i=1}^n x_i \sin(\sqrt{x_i})$$

$$f^* \lesssim -418.9n$$

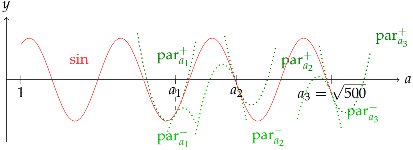
Interval Arithmetic for sin + SOS

n	lower bound	n_{lifting}	#boxes	time
10	$-430n$	0	3830	129 s
10	$-430n$	$2n$	16	40 s

Comparison on Global Optimization Problems

$$\min_{\mathbf{x} \in [1,500]^n} f(\mathbf{x}) = -\sum_{i=1}^{n} x_i \sin(\sqrt{x_i})$$

$$f^* \lesssim -418.9n$$



n	lower bound	n_{lifting}	#boxes	time
10	$-430n$	0	3830	$129\,s$
10	$-430n$	$2n$	16	$40\,s$

Comparison on Global Optimization Problems

$$\min_{\mathbf{x} \in [1,500]^n} f(\mathbf{x}) = -\sum_{i=1}^{n} x_i \sin(\sqrt{x_i})$$

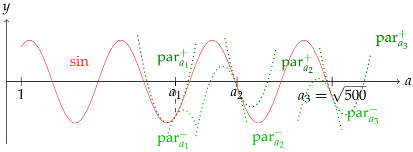
$$f^* \lesssim -418.9n$$

Interval Arithmetic for sin + SOS

n	lower bound	n_{lifting}	#boxes	time
100	$-440n$	0	> 10000	$> 10\,h$
100	$-440n$	$2n$	274	$1.9\,h$

Comparison on Global Optimization Problems

$$\min_{\mathbf{x} \in [1,500]^n} f(\mathbf{x}) = -\sum_{i=1}^{n} x_i \sin(\sqrt{x_i})$$

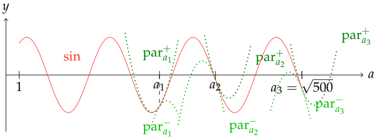


$$f^* \lesssim -418.9n$$

n	lower bound	n_{lifting}	#boxes	time
100	$-440n$	0	> 10000	$> 10\,h$
100	$-440n$	$2n$	274	$1.9\,h$

Comparison on Global Optimization Problems

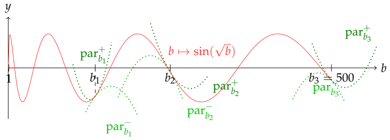
$$\min_{\mathbf{x}\in[1,500]^n} f(\mathbf{x}) = -\sum_{i=1}^{n-1} (x_i + x_{i+1}) \sin(\sqrt{x_i})$$



n	lower bound	n_{lifting}	#boxes	time
1000	$-967n$	$2n$	1	$543\,s$
1000	$-968n$	n	1	$272\,s$

Comparison on Global Optimization Problems

$$\min_{\mathbf{x}\in[1,500]^n} f(\mathbf{x}) = -\sum_{i=1}^{n-1} (x_i + x_{i+1}) \sin(\sqrt{x_i})$$



n	lower bound	n_{lifting}	#boxes	time
1000	$-967n$	$2n$	1	$543\,s$
1000	$-968n$	n	1	$272\,s$

Convergence of the Optimization Algorithm

- Let f be a multivariate transcendental function
- Let t_p^- be the underestimator of f, obtained at precision p
- Let $\mathbf{x}_{opt}^p$ be a minimizer of t_p^- over $\mathbf{K}$

Theorem [X. Allamigeon S. Gaubert VM B. Werner 13]

Every accumulation point of the sequence $(\mathbf{x}_{opt}^p)$ is a global minimizer of f on $\mathbf{K}$.

Ingredients of the proof:

- Convergence of Lasserre SOS hierarchy
- Uniform approximation schemes (Maxplus/Minimax)

Polynomial Approximations for Semialgebraic Functions

- Inspired from [Lasserre - Thanh 13]
- Let $f_{\text{sa}} \in \mathcal{A}$ defined on a box $K \subset \mathbb{R}^n$
- Let μ_n be the standard Lebesgue measure on $\mathbb{R}^n$
- Best polynomial underestimator $h \in \mathbb{R}_d[\mathbf{x}]$ of f_{sa} for the L_1 norm:

$$(P^{\text{sa}}) \left\{ \begin{array}{ll} \min\limits_{h \in \mathbb{R}_d[\mathbf{x}]} & \int_{\mathbf{K}} (f_{\text{sa}} - h) d\mu_n \\ \text{s.t.} & f_{\text{sa}} - h \geqslant 0 \text{ on } \mathbf{K} \end{array} \right. .$$

Lemma

Problem (P^{sa}) has a degree-d polynomial minimizer h_d.

Polynomial Approximations for Semialgebraic Functions

- b.s.a.l. $\hat{\mathbf{K}} := \{(\mathbf{x}, \mathbf{z}) \in \mathbb{R}^{n+p} : g_1(\mathbf{x}, \mathbf{z}) \geqslant 0, \dots, g_m(\mathbf{x}, \mathbf{z}) \geqslant 0\}$
- The quadratic module $\mathcal{M}(\hat{\mathbf{K}})$ is Archimedean

- The optimal solution h_d of (P^{sa}) is a maximizer of:

$$(P_d) \begin{cases} \max\limits_{h \in \mathbb{R}_d[\mathbf{x}]} & \displaystyle\int_{[0,1]^n} h \, d\mu_n \\ \text{s.t.} & (z_p - h) \in \mathcal{M}(\hat{\mathbf{K}}) \end{cases} .$$

Polynomial Approximations for Semialgebraic Functions

- Let m_d be the optimal value of Problem (P^{sa})
- Let h_{dk} be a maximizer of the SOS relaxation of (P_d)

Convergence of the SOS Hierarchy

The sequence $(\|f_{\mathrm{sa}} - h_{dk}\|_1)_{k \geqslant k_0}$ is non-increasing and converges to m_d. Each accumulation point of the sequence $(h_{dk})_{k \geqslant k_0}$ is an optimal solution of Problem (P^{sa}).

Polynomial Approximations for Semialgebraic Functions

$$f_{\mathrm{sa}}(\mathbf{x}) := \frac{\partial_4 \Delta \mathbf{x}}{\sqrt{4 x_1 \Delta \mathbf{x}}}$$

d	k	Upper bound of $\|f_{\mathrm{sa}} - h_{dk}\|_1$	Bound
2	2	0.8024	-1.171
	3	0.3709	-0.4479
4	2	1.617	-1.056
	3	0.1766	-0.4493

Polynomial Approximations for Semialgebraic Functions

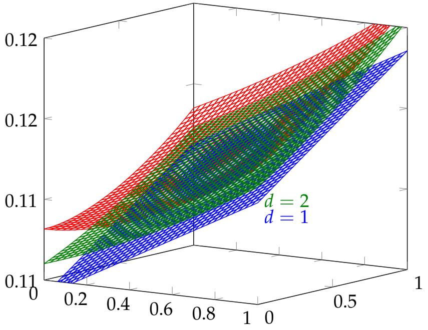
- rad$_2$: $(x_1, x_2) \mapsto \frac{-64x_1^2+128x_1x_2+1024x_1-64x_2^2+1024x_2-4096}{-8x_1^2+8x_1x_2+128x_1-8x_2^2+128x_2-512}$

- Linear and quadratic underestimators for rad$_2$ ($k = 3$):

Polynomial Approximations for Semialgebraic Functions

- $\text{rad}_2 : (x_1, x_2) \mapsto \frac{-64x_1^2+128x_1x_2+1024x_1-64x_2^2+1024x_2-4096}{-8x_1^2+8x_1x_2+128x_1-8x_2^2+128x_2-512}$

- Linear and quadratic underestimators for rad_2 ($k = 3$):



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.

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

Bounding the Polynomial Remainder

- Normalized POP ($\mathbf{x} \in [0,1]^n$)
- $\epsilon_{\text{pop}}(\mathbf{x}) := p(\mathbf{x}) - \lambda_k - \sum_{j=0}^{m} \sigma_j(\mathbf{x}) g_j(\mathbf{x})$
- $\forall \mathbf{x} \in [0,1]^n, \epsilon_{\text{pop}}(\mathbf{x}) \geqslant \epsilon^*_{\text{pop}} := \sum_{\epsilon_{\boldsymbol{\alpha}} \leqslant 0} \epsilon_{\boldsymbol{\alpha}}$

Formal SOS Results

- *POP1*: $\forall \mathbf{x} \in \mathbf{K}, \partial_4 \Delta \mathbf{x} \geqslant -41$.
- *POP2*: $\forall \mathbf{x} \in \mathbf{K}, \Delta \mathbf{x} \geqslant 0$.

Problem	n	`NLCertify`	`micromega` [Besson 07]
POP1	6	0.08 s	9.00 s
POP2	2	0.09 s	0.36 s
	3	0.39 s	—
	6	13.2 s	—

Sparse SOS relaxations $\Longrightarrow$ Speedup

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]

No free lunch: SDP informal bottleneck

22 times slower than SDP: $\boxed{q = q'}$ formal bottleneck

Contribution

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

Formal Nonlinear Optimization

- Formal nonlinear optimization: `NLCertify`
- Safe solutions for challenging problems, e.g. Flyspeck

Formal Nonlinear Optimization

Further research:

- OCAML API
- Alternative polynomial bounds using geometric programming [De Wolff and Iliman]
- Mixed LP/SOS certificates (trade-off CPU/precision)
- COQ tactic
- Improve formal polynomial checker

Formal Nonlinear Optimization

Further research:

Generalized problem of moments

$$
\begin{array}{llcll}
 & \text{(Moment)} & & & \text{(SOS)} \\
\inf & \int_{\mathbf{K}} p_0 \, d\mu & \geqslant & \sup & \lambda_0 + \sum_i \lambda_i b_i \\
\text{s.t.} & \int_{\mathbf{K}} p_i \, d\mu \leqslant b_i & & \text{s.t.} & \lambda_0, \lambda_i \leqslant 0 \ , \\
 & \mu \in \mathcal{M}_+(\mathbf{K}) & & & p_0 - \lambda_0 - \sum_i \lambda_i p_i \in \mathcal{Q}_k(\mathbf{K})
\end{array}
$$

- Formal bounds using SDP and `ring`

End

Thank you for your attention!

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