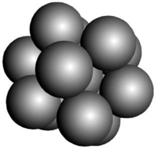
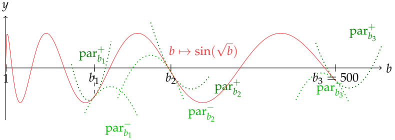
Formal Proofs for Nonlinear Optimization

Victor Magron, RA Imperial College

3 March 2015

QUADS Seminar
Dept of Computing, Imperial College





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 Jan-Sept: Postdoc at LAAS-CNRS
  MOMENT-SOS APPLICATIONS
  (D. Henrion and J.B. Lasserre)

- From 2014 Oct: Postdoc at Imperial
  SDP FOR AUTOMATED HARDWARE TUNING
  (G. Constantinides and A. Donaldson)

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)

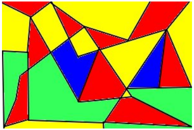
  HOL-LIGHT (Harrison, Gordon 1980)

  Built in top of OCAML

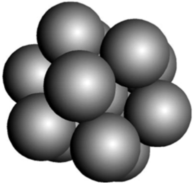
Complex Proofs

- Complex mathematical proofs / mandatory computation

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



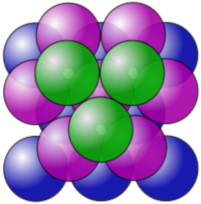
T. Hales, A Proof of the Kepler Conjecture, 1994.



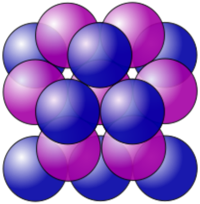
From Oranges Stack...

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."
- **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!!**

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

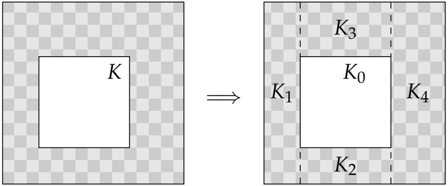
- Certified interval arithmetic in COQ [Melquiond 12]
- Taylor methods in HOL Light [Solovyev thesis 13]
  - Formal verification of floating-point operations
- robust but subject to the **Curse of Dimensionality**

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 (SOS) techniques

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

but

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

Existing Formal Frameworks

The "No Free Lunch" Rule:

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

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 : Dual approach

- Semialgebraic set $\mathbf{K} := \{\mathbf{x} \in \mathbb{R}^n : g_1(\mathbf{x}) \geqslant 0, \ldots, g_m(\mathbf{x}) \geqslant 0\}$
- $p^* := \min_{\mathbf{x} \in \mathbf{K}} p(\mathbf{x})$: 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 : Dual approach

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 $\mathcal{Q}(\mathbf{K}')$

Polynomial Optimization : Primal approach

- 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. $\mu : \mathcal{B} \to [0, 1]$ $\quad$ $\mu(\emptyset) = 0$
  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 $\mu(\mathbb{R}^n \backslash \mathbf{K}) = 0$

Polynomial Optimization : Primal approach

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

Polynomial Optimization : Primal approach

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

Polynomial Optimization : Primal approach

- 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

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

Polynomial Optimization : Primal approach

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

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:

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

with column labels $1 \mid x_1 \; x_2 \mid x_1^2 \; x_1x_2 \; x_2^2$.

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:

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

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 $\mathbf{K}$
- $\mathcal{Q}_k(\mathbf{K})$: truncated quadratic module

Polynomial Optimization Problems (POP)

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

Primal-dual Moment-SOS

- For large enough k, **zero duality gap** [Lasserre 01]:

Polynomial Optimization Problems (POP)

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

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:

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

  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 on $[0, 1]^4$
- SOS of degree at most 4
- 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

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)

General informal Framework

Approximation theory: Chebyshev/Taylor models

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

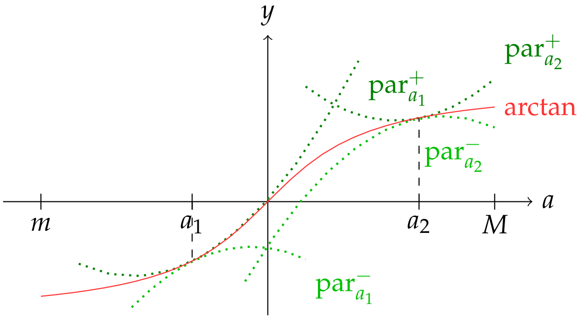
Maxplus Approximation

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



Maxplus Approximation Error

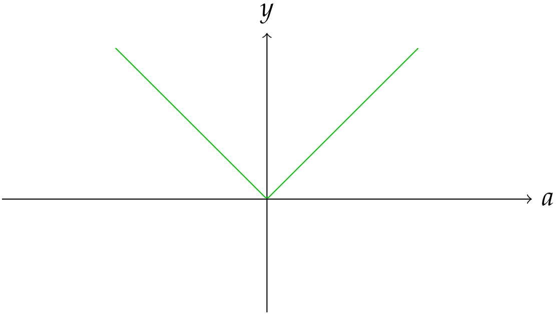
Theorem [Akian-Gaubert-Lakhoua 08]

Let $\gamma \in \mathbb{R}$, $\eta > 0$. Let ϕ be $(\gamma - \eta)$-semiconvex and Lipschitz-continuous on compact convex $K \subset \mathbb{R}^n$. Let ϕ_N denote the best maxplus approximation by N quadratic forms of Hessian $-\gamma I$. Then $\|\boldsymbol{\phi} - \boldsymbol{\phi}_\mathbf{N}\|_\infty = \mathbf{O}(\mathbf{1}/\mathbf{N}^{\mathbf{2}/\mathbf{n}})$.

- **Differentiability not mandatory** by contrast with Taylor
- In our case $n = 1$, one needs $O(1/\sqrt{\epsilon})$ basis functions

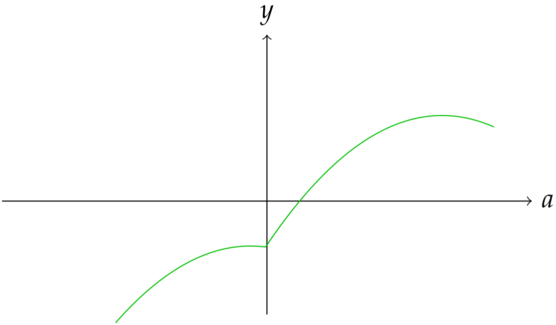
Nonlinear Function Representation

Exact parsimonious maxplus representations



Nonlinear Function Representation

Exact parsimonious maxplus representations



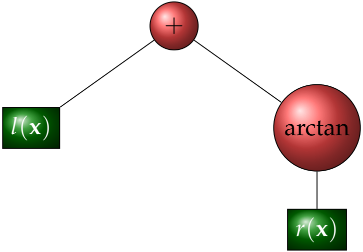
Nonlinear Function Representation

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

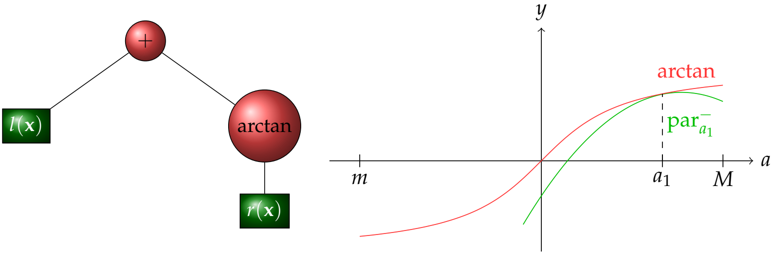
Nonlinear Function Representation

- For the "Simple" Example from Flyspeck:



Maxplus Optimization Algorithm

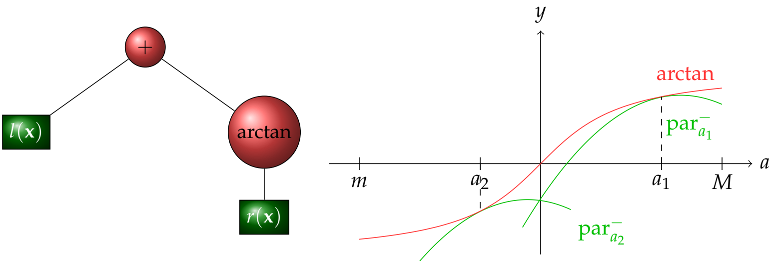
First iteration:



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

Maxplus Optimization Algorithm

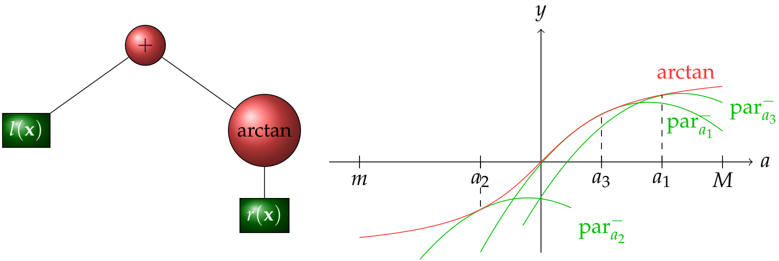
Second iteration:



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!

Convergence of the Optimization Algorithm

- f under-approximated by t_p^- (precision p)
- Let $\mathbf{x}_{opt}^p$ be a minimizer of t_p^- over $\mathbf{K}$

Theorem

Each limit point of $(\mathbf{x}_{opt}^p)_p$ is a **global minimizer** of f on $\mathbf{K}$.

Ingredients of the proof:

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

Comparison Results

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

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

Best uniform Approximation
+ SOS

- $d = 4, n = 10$
- 38 *min* to certify a lower bound of $-430n$
- Poor accuracy of Minimax Estimators

Comparison Results

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

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

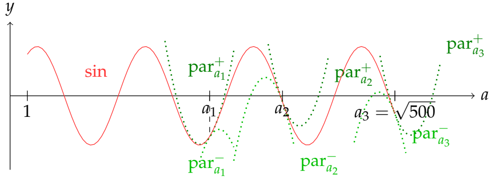
Inteval Arithmetic for sin + SOS

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

Comparison Results

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

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



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

Comparison Results

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

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

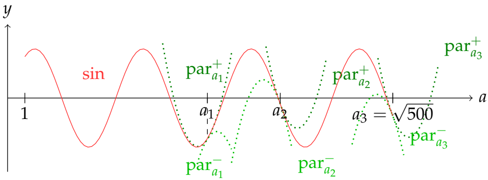
Inteval Arithmetic for sin + SOS

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

Comparison Results

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

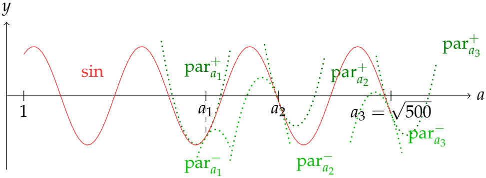
$$f^* \lesssim -418.9n$$



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

Comparison Results

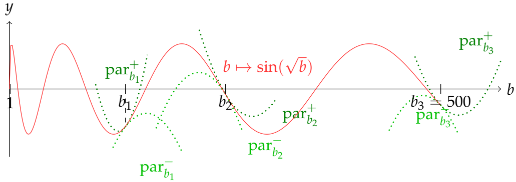
$$\min_{\mathbf{x}\in[1,500]^n} -\sum_{i=1}^{n-1}(x_i + x_{i+1})\sin(\sqrt{x_i})$$



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

Comparison Results

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



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

Contributions

V. Magron, X. Allamigeon, S. Gaubert, and B. Werner.
Certification of Real Inequalities – Templates and Sums of Squares, arxiv:1403.5899, 2014. Accepted for publication in *Mathematical Programming SERIES B, volume on Polynomial Optimization*.

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 Polynomial Optimization

- Alternative to the projection and rounding by [Parrilo-Peyrl 08]:
  - Normalized POP ($\mathbf{x} \in [0,1]^n$)
  - Conversion into rationals: SOS $\sigma_0, \ldots, \sigma_m$, lower bound μ_k
  - $\epsilon_{\text{pop}}(\mathbf{x}) := p(\mathbf{x}) - \mu_k - \sum_{j=0}^{m} \sigma_j(\mathbf{x}) g_j(\mathbf{x})$
  - Bounding: $\forall \mathbf{x} \in [0,1]^n, \epsilon_{\text{pop}}(\mathbf{x}) \geqslant \epsilon_{\text{pop}}^* := \sum_{\epsilon_{\boldsymbol{\alpha}} \leqslant 0} \epsilon_{\boldsymbol{\alpha}}$
- More concise SOS certificates / Simpler rounding

Formal Polynomial Optimization

When $q \in \mathcal{Q}(\mathbf{K}), \sigma_0, \ldots, \sigma_m$ 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

Formal Semialgebraic Optimization

```
Inductive cert_sa : Type :=
| Poly  : PExpr → itv → cert_pop_itv → cert_sa
| Fdiv  : cert_sa → cert_sa → itv →
   cert_pop_itv → cert_sa
| Fsqrt : cert_sa → itv → cert_sa
| ...
```

Formal Semialgebraic Optimization

```
Inductive cert_sa : Type :=
| Poly  : PExpr → itv → cert_pop_itv → cert_sa
| Fdiv  : cert_sa → cert_sa → itv →
   cert_pop_itv → cert_sa
| Fsqrt : cert_sa → itv → cert_sa
| ...
```

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

Definition p := Poly p i_p c_p.
Definition q := Poly q i_q c_q.
Definition sqrtq := Fsqrt q $i_{\sqrt{}}$.
Definition r := Fdiv p sqrtq i_r c_r.

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

For more details on the formal side:

- V. M., X. Allamigeon, S. Gaubert and B. Werner. Formal Proofs for Nonlinear Optimization, arxiv:1404.7282, 2015. *Journal of Formalized Reasoning*.

- Victor Magron. NLCertify: A Tool for Formal Nonlinear Optimization, arxiv:1405.5668, *Proceedings of ICMS 2014*.

Conclusion

With **SUMS OF SQUARES**, you can

- Optimize nonlinear (transcendental) functions
- Formal nonlinear optimization: `NLCertify`

Conclusion

Further research:

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

End

Thank you for your attention!

`cas.ee.ic.ac.uk/people/vmagron`