Certification of inequalities involving transcendental functions using Semi-Definite Programming

Joint Work with B. Werner, S. Gaubert and X. Allamigeon

Second year PhD Victor MAGRON

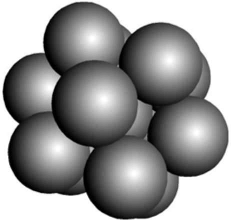
LIX/INRIA, École Polytechnique

ISMP 2012 Friday August 24 th

Kepler Conjecture (1611):

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

- It corresponds to the way people would intuitively stack oranges, as a pyramid shape
- The proof of T. Hales (1998) consists of thousands of non-linear inequalities
- Many recent efforts have been done to give a formal proof of these inequalities: Flyspeck Project
- Motivation: get positivity certificates and check them with Proof assistants like COQ



Inequalities issued from Flyspeck non-linear part involve:

1. Semi-Algebraic functions algebra $\mathcal{A}$: composition of polynomials with $|\cdot|, (\cdot)^{\frac{1}{p}} (p \in \mathbb{N}_0), +, -, \times, /, \sup, \inf$
2. Transcendental functions $\mathcal{T}$: composition of semi-algebraic functions with $\mathtt{arctan}$, $\mathtt{arcos}$, $\mathtt{arcsin}$, $\mathtt{exp}$, $\mathtt{log}$, $|\cdot|$, $(\cdot)^{\frac{1}{p}} (p \in \mathbb{N}_0), +, -, \times, /, \sup, \inf$

Lemma$_{9922699028}$ from Flyspeck

$K := [4; 6.3504]^3 \times [6.3504; 8] \times [4; 6.3504]^2$

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

$\forall x \in K, -\frac{\pi}{2} + \mathtt{arctan}\, \frac{\partial_4 \Delta x}{\sqrt{4 x_1 \Delta x}} + 1.6294 - 0.2213\, (\sqrt{x_2} + \sqrt{x_3} + \sqrt{x_5} + \sqrt{x_6} - 8.0) + 0.913\, (\sqrt{x_4} - 2.52) + 0.728\, (\sqrt{x_1} - 2.0) \geq 0.$

Tight inequality: global optimum $= 1.7 \times 10^{-4}$

Contents

Flyspeck-Like Problems [Hales and Solovyev Method]

- Real numbers are represented by interval arithmetic
- Analytic functions f (e.g. $\sqrt{\cdot}$, $\frac{1}{\cdot}$, $arctan$) are approximated with Taylor expansions and the error terms are bounded:
  $$|f(x) - f(x_0) - \mathcal{D}_f(x_0)\,(x - x_0)| < \sum_{i,j} m_{ij}\, \epsilon_i\, \epsilon_j$$
  $$\epsilon_i := |x^i - x_0^i|$$
- To satisfy the inequalities, the initial box K is partitioned into smaller boxes until the Taylor approximations are accurate enough (the error terms become small enough)
- The Taylor expansions are generated by symbolic differentiation using the chain rule, product rule

General Framework

We consider the same problem: given K a compact set, and f a transcendental function, bound from below $f^* = \inf_{x \in K} f(x)$ and prove $f^* \geq 0$

1. f is underestimated by a semi-algebraic function f_{sa} on a compact set $K_{sa} \supset K$
2. Reduce the problem $\inf_{x \in K_{sa}} f_{sa}(x)$ to a polynomial optimization problem (POP) in a lifted space K_{pop}
3. Solve classicaly the POP problem $\inf_{x \in K_{pop}} f_{pop}(x)$ using a hierarchy of SDP relaxations by Lasserre

$$f^* \geq f_{sa}^* \geq f_{pop}^* \underbrace{\geq 0}_{\text{If the relaxations are accurate enough}}$$

General Framework

We consider the same problem: given K a compact set, and f a transcendental function, bound from below $f^* = \inf_{x \in K} f(x)$ and prove $f^* \geq 0$

1. f is underestimated by a semi-algebraic function f_{sa} on a compact set $K_{sa} \supset K$
2. Reduce the problem $\inf_{x \in K_{sa}} f_{sa}(x)$ to a polynomial optimization problem (POP) in a lifted space K_{pop}
3. Solve classicaly the POP problem $\inf_{x \in K_{pop}} f_{pop}(x)$ using a hierarchy of SDP relaxations by Lasserre

$$f^* \geq f_{sa}^* \geq f_{pop}^* \underbrace{\geq 0}_{\text{If the relaxations are accurate enough}}$$

General Framework

We consider the same problem: given K a compact set, and f a transcendental function, bound from below $f^* = \inf_{x \in K} f(x)$ and prove $f^* \geq 0$

1. f is underestimated by a semi-algebraic function f_{sa} on a compact set $K_{sa} \supset K$
2. Reduce the problem $\inf_{x \in K_{sa}} f_{sa}(x)$ to a polynomial optimization problem (POP) in a lifted space K_{pop}
3. Solve classicaly the POP problem $\inf_{x \in K_{pop}} f_{pop}(x)$ using a hierarchy of SDP relaxations by Lasserre

$$f^* \geq f^*_{sa} \geq f^*_{pop} \underbrace{\geq 0}$$

If the relaxations are accurate enough

General Framework

We consider the same problem: given K a compact set, and f a transcendental function, bound from below $f^* = \inf_{x \in K} f(x)$ and prove $f^* \geq 0$

1. f is underestimated by a semi-algebraic function f_{sa} on a compact set $K_{sa} \supset K$
2. Reduce the problem $\inf_{x \in K_{sa}} f_{sa}(x)$ to a polynomial optimization problem (POP) in a lifted space K_{pop}
3. Solve classicaly the POP problem $\inf_{x \in K_{pop}} f_{pop}(x)$ using a hierarchy of SDP relaxations by Lasserre

$$f^* \geq f_{sa}^* \geq f_{pop}^* \underbrace{\geq 0}_{\text{If the relaxations are accurate enough}}$$

SOS and SDP Relaxations

Polynomial Optimization Problem (POP):

Let $f, g_1, \cdots, g_m \in \mathbb{R}[X_1, \cdots, X_n]$
$K_{pop} := \{x \in \mathbb{R}^n : g_1(x) \geq 0, \cdots, g_m(x) \geq 0\}$ is the feasible set
General POP: compute $f^*_{pop} = \inf_{x \in K_{pop}} f(x)$

SOS Assumption: [e.g. Lasserre]

K is compact, $\exists u \in \mathbb{R}[X]$ s.t. the level set $\{x \in \mathbb{R}^n : u(x) \geq 0\}$ is compact and $u = u_0 + \sum_{j=1}^{m} u_j \, g_j$ for some sum of squares (SOS) $u_0, u_1, \cdots, u_m \in \Sigma[X]$

- Normalize the feasibility set to get $K' := [-1; 1]^n$
  $K' := \{x \in \mathbb{R}^n : g_1 := 1 - x_1^2 \geq 0, \cdots, g_n := 1 - x_n^2 \geq 0\}$
- The polynomial $u(x) := n - \sum_{j=1}^{n} x_j^2$ satisfies the assumption

SOS and SDP Relaxations

To convexify the problem, use the equivalent formulation:

$f^*_{pop} = \inf_{x \in K_{pop}} f_{pop}(x) = \inf_{\mu \in \mathcal{P}(K_{pop})} \int f_{pop}\, d\mu$, where $\mathcal{P}(K_{pop})$ is the set of all probability measures μ supported on the set K_{pop}.

Theorem [Putinar]:

Given $L : \mathbb{R}[X] \to \mathbb{R}$, the following are equivalent:

1. $\exists \mu \in \mathcal{P}(K_{pop}), \forall p \in \mathbb{R}[X], L(p) = \int p\, d\mu$
2. $L(1) = 1,\ L(s_0 + \sum_{j=1}^{m} s_j g_j) \geq 0$ for any $s_0, \cdots, s_m \in \Sigma[X]$

Equivalent formulation:

$f^*_{pop} = \min \{ L(f) : L : \mathbb{R}[X] \to \mathbb{R}$ linear, $L(1) = 1$ and each $\mathcal{L}_{g_j}$ is SDP $\}$, with $g_0 = 1$, $\mathcal{L}_{g_0}, \cdots, \mathcal{L}_{g_m}$ defined by:

$$\begin{aligned} \mathcal{L}_{g_j} : \mathbb{R}[X] \times \mathbb{R}[X] &\to \mathbb{R} \\ (p, q) &\mapsto L(p \cdot q \cdot g_j) \end{aligned}$$

SOS and SDP Relaxations: Lasserre Hierarchy

- Let $\mathcal{B} := (X^\alpha)_{\alpha \in \mathbb{N}^n}$ denote the monomial basis and set $y_\alpha = L(X^\alpha)$, this identifies L with the infinite series $y = (y_\alpha)_{\alpha \in \mathbb{N}^n}$.
- The infinite moment matrix M associated to y indexed by $\mathcal{B}$ is: $M(y)_{u,v} := L(u \cdot v), \ u, v \in \mathcal{B}$.
- The localizing matrix $M(g_j y)$ is: $M(g_j y)_{u,v} := L(u \cdot v \cdot g_j), \ u, v \in \mathcal{B}$.
- Let $k \geq k_0 := \max\{\lceil \deg f_{pop} \rceil / 2, \lceil \deg g_0 / 2 \rceil, \cdots, \lceil \deg g_m / 2 \rceil\}$. Truncate the previous matrices by considering only rows and columns indexed by elements in $\mathcal{B}$ of degree at most k, and consider the hierarchy Q_k of semidefinite relaxations:

$$Q_k : \begin{cases} \inf_y L(f) & = \int f_\alpha \, x^\alpha \, d\mu(x) = \sum_\alpha f_\alpha \, y_\alpha \\ M_{k - \lceil \deg g_j / 2 \rceil}(g_j y) & \succcurlyeq 0, \qquad 0 \leq j \leq m, \\ y_1 & = 1 \end{cases}$$

SOS and SDP Relaxations

Convergence Theorem [Lasserre]:

The sequence $\inf(Q_k)_{k \geq k_0}$ is non-decreasing and under the SOS assumption converges to f^*_{pop}.

SDP relaxations:

Many solvers (e.g. Sedumi [?], SDPA) solve the pair of (standard form) semidefinite programs:

$$(SDP) \begin{cases} \mathcal{P}: & \min\limits_{y} & \sum\limits_{\alpha} c_\alpha y_\alpha \\ & \text{subject to} & \sum\limits_{\alpha} F_\alpha \, y_\alpha - F_0 \succcurlyeq 0 \\ \mathcal{D}: & \max\limits_{Y} & \text{Trace}\,(F_0\, Y) \\ & \text{subject to} & \text{Trace}\,(F_\alpha\, Y) = c_\alpha \end{cases}$$

Basic Semi-Algebraic Relaxations

- Let $\mathcal{A}$ be a set of semi-algebraic functions and $f_{sa} \in \mathcal{A}$.
- We consider the problem $f_{sa}^* = \inf_{x \in K_{sa}} f_{sa}(x)$ with $K_{sa} := \{x \in \mathbb{R}^n : g_1(x) \geq 0, \cdots, g_m(x) \geq 0\}$ a basic semi-algebraic set

Basic Semi-Algebraic Lifting:

A function $f_{sa} \in \mathcal{A}$ is said to have a basic semi-algebraic lifting (b.s.a.l.) if $\exists p, s \in \mathbb{N}$, polynomials $h_1, \cdots h_s \in \mathbb{R}[X, Z_1, \cdots, Z_p]$ and a b.s.a. set K_{pop} defined by:

$$
\begin{aligned}
K_{pop} := \; & \{(x, z_1, \cdots, z_p) \in \mathbb{R}^{n+p} : x \in K_{sa}, \\
& h_1(x, z_1, \cdots, z_p) \geq 0, \cdots, h_s(x, z_1, \cdots, z_p) \geq 0\}
\end{aligned}
$$

such that the graph of f_{sa} (denoted $\Psi_{f_{sa}}$) satisfies:
$\Psi_{f_{sa}} := \{(x, f_{sa}(x)) : x \in K_{sa}\} = \{(x, z_p) : (x, z) \in K_{pop}\}$

Basic Semi-Algebraic Relaxations

b.s.a.l. lemma [Lasserre, Putinar] :

Let $\mathcal{A}$ be the semi-algebraic functions algebra obtained by composition of polynomials with $|\cdot|$, $(\cdot)^{\frac{1}{p}} (p \in \mathbb{N}_0)$, $+, -, \times, /, \sup, \inf$. Then every well-defined $f_{sa} \in \mathcal{A}$ has a basic semi-algebraic lifting.

Example from Flyspeck:

$$f_{sa} := \frac{\partial_4 \Delta x}{\sqrt{4 x_1 \Delta x}}, K_{sa} := [4; 6.3504]^3 \times [6.3504; 8] \times [4; 6.3504]^2.$$

- Define $z_1 := \sqrt{4 x_1 \Delta x}$, $m_1 \leq \inf_{x \in K_{sa}} z_1(x)$, $M_1 \geq \sup_{x \in K_{sa}} z_1(x)$,

  $z_2 := \dfrac{\partial_4 \Delta x}{\sqrt{4 x_1 \Delta x}}$ How to compute $m_2 \leq \inf_{x \in K_{sa}} z_2(x)$?
- Define $h_1 := z_1 - m_1$, $h_2 := M_1 - z_1$, $h_3 := z_1^2 - 4 x_1 \Delta x$, $h_4 := -z_1^2 + 4 x_1 \Delta x$, $h_5 := z_1$, $h_6 := z_2 \, z_1 - \partial_4 \Delta x$, $h_7 := -z_2 \, z_1 + \partial_4 \Delta x$
- $K_{pop} := \{(x, \, z) \in \mathbb{R}^{6+2} : x \in K_{sa}, \, h_k(x, z) \geq 0, \, k = 1, \cdots, 7\}$.
- $\Psi_{f_{sa}} := \{(x, \, f_{sa}(x)) : x \in K_{sa}\} = \{(x, \, z_2) : (x, \, z) \in K_{pop}\}$.

Basic Semi-Algebraic Relaxations

Example from Flyspeck:

$$f_{sa} := \frac{\partial_4 \Delta x}{\sqrt{4 x_1 \Delta x}}, \; K_{sa} := [4; 6.3504]^3 \times [6.3504; 8] \times [4; 6.3504]^2.$$

- Define $g_1 := x_1 - 4$, $g_2 := 6.3504 - x_1, \cdots, g_{11} := x_6 - 4$, $g_{12} := 6.3504 - x_6$. Solve:

$$Q_k : \begin{cases} \inf_y L(f_{pop}) = \inf_y y_{0\ldots01} = \int z_2 \, d\mu \\ M_{k - \lceil \deg g_j / 2 \rceil}(g_j \, y) \succcurlyeq 0, \quad 1 \leq j \leq 12, \\ M_{k - \lceil \deg h_k / 2 \rceil}(h_k \, y) \succcurlyeq 0, \quad 1 \leq k \leq 7, \\ y_{0\ldots0} = 1 \end{cases}$$

Basic Semi-Algebraic Relaxations

Example from Flyspeck:

$$f_{sa} := \frac{\partial_4 \Delta x}{\sqrt{4 x_1 \Delta x}}, \; K_{sa} := [4; 6.3504]^3 \times [6.3504; 8] \times [4; 6.3504]^2.$$

- Define $g_1 := x_1 - 4,\ g_2 := 6.3504 - x_1 \cdots g_{11} := x_6 - 4,\ g_{12} := 6.3504 - x_6$. Solve:

$$Q_k : \left\{ \begin{array}{rcll} \inf\limits_{y} L(f_{pop}) & = & \inf\limits_{y} y_{0\ldots01} & = \int z_2 \, d\mu \\ M_{k-1}(g_j \, y) & \succcurlyeq & 0, & 1 \leq j \leq 12, \\ M_{k-\lceil \deg h_k/2 \rceil}(h_k \, y) & \succcurlyeq & 0, & 1 \leq k \leq 7, \\ y_{0\ldots0} & = & 1 & \end{array} \right.$$

b.s.a.l. Convergence: (Special case of Convergence Theorem)

- Let $k \geq k_0 := \max\{1, \lceil \deg h_1/2 \rceil, \cdots, \lceil \deg h_7/2 \rceil\}$.
- The sequence $\inf(Q_k)_{k \geq k_0}$ is monotically non-decreasing and converges to f_{sa}^*.

Decrease the SDP Problems Size

- Exploiting symmetries in SDP-relaxations for POP [Riener, Theobald, Andren, Lasserre] to replace one SDP problem Q_k of size $\mathcal{O}(n^k)$ by several smaller SDPS of size $\mathcal{O}(\eta_i^k)$.
- SOS and SDP Relaxations for Polynomial Optimization Problems with Structured Sparsity [Waki, Kim, Kojima, Muramatsu] to replace one SDP problem Q_k of size $\mathcal{O}(n^k)$ by a SDP problem of size $\mathcal{O}(\kappa^k)$ where κ is the average size of the maximal cliques correlation pattern of the polynomial variables.

Decrease the SDP Problems Size

- Exploiting symmetries in SDP-relaxations for POP [Riener, Theobald, Andren, Lasserre] to replace one SDP problem Q_k of size $\mathcal{O}(n^k)$ by several smaller SDPS of size $\mathcal{O}(\eta_i^k)$.
- SOS and SDP Relaxations for Polynomial Optimization Problems with Structured Sparsity [Waki, Kim, Kojima, Muramatsu] to replace one SDP problem Q_k of size $\mathcal{O}(n^k)$ by a SDP problem of size $\mathcal{O}(\kappa^k)$ where κ is the average size of the maximal cliques correlation pattern of the polynomial variables.

Issues and Solutions

Issues:

1. How to deal with transcendant functions?
2. Even when exploiting sparsity and symmetries, a direct implementation of basic-semialgebraic relaxation is not enough to prove Hales's lemmas (inequalities are too tight, requiring high order relaxations, and so a high execution time)

Solutions:

An adaptative basic-semialgebraic relaxation, with a max-plus semi-convex approximation (lower approximate a transcendant functions by a sup of quadratic forms)

Issues and Solutions

Issues:

1. How to deal with transcendant functions?
2. Even when exploiting sparsity and symmetries, a direct implementation of basic-semialgebraic relaxation is not enough to prove Hales's lemmas (inequalities are too tight, requiring high order relaxations, and so a high execution time)

Solutions:

An adaptative basic-semialgebraic relaxation, with a max-plus semi-convex approximation (lower approximate a transcendant functions by a sup of quadratic forms)

Issues and Solutions

Issues:

1. How to deal with transcendant functions?
2. Even when exploiting sparsity and symmetries, a direct implementation of basic-semialgebraic relaxation is not enough to prove Hales's lemmas (inequalities are too tight, requiring high order relaxations, and so a high execution time)

Solutions:

An adaptative basic-semialgebraic relaxation, with a max-plus semi-convex approximation (lower approximate a transcendant functions by a sup of quadratic forms)

Transcendental Functions Underestimators

- Let $f \in \mathcal{T}$ be a transcendental univariate elementary function such as $\mathtt{arctan}$, $\mathtt{exp}$, ..., defined on a real interval I.
- Basic convexity/quasiconvexity properties and monotonicity of f are used to find lower and upper semi-algebraic bounds.

Example with $\mathtt{arctan}$:

- $\mathtt{arctan}$ is quasiconvex on I:
  $\exists\, c < 0$ such that $\mathtt{arctan} - \frac{c}{2}(\cdot)^2$ is convex on I
- $\forall a \in I = [m; M]$, $\mathtt{arctan}(a) \geq \max_{i \in \mathcal{C}} \{\mathtt{par}^-_{a_i}(a)\}$ where $\mathcal{C}$ define an index collection of parabola tangent to the function curve and underestimating f.
- $\mathtt{par}^-_{a_i} := \frac{c_i}{2}(a - a_i)^2 + f'_{a_i}(a - a_i) + f(a_i)$,
  $f'_{a_i} = \frac{1}{1 + a_i^2}$, $f(a_i) = \mathtt{arctan}(a_i)$.
- c_p depends on a_p and the curvature variations of $\mathtt{arctan}$ on the considered interval I.

Transcendental Functions Underestimators

- Let $f \in \mathcal{T}$ be a transcendental univariate elementary function such as $\mathtt{arctan}$, exp, ..., defined on a real interval I.
- Basic convexity/quasiconvexity properties and monotonicity of f are used to find lower and upper semi-algebraic bounds.

Example with $\mathtt{arctan}$:

- $\mathtt{arctan}$ is quasiconvex on I:
  $\exists\, c < 0$ such that $\mathtt{arctan} - \frac{c}{2}(\cdot)^2$ is convex on I
- $\forall a \in I = [m; M]$, $\mathtt{arctan}(a) \geq \max_{i \in \mathcal{C}} \{par^-_{a_i}(a)\}$ where $\mathcal{C}$ define an index collection of parabola tangent to the function curve and underestimating f.
- $par^-_{a_i} := \frac{c_i}{2}(a - a_i)^2 + f'_{a_i}(a - a_i) + f(a_i)$,
  $f'_{a_i} = \frac{1}{1 + a_i^2}$, $f(a_i) = \mathtt{arctan}(a_i)$.
- c_p depends on a_p and the curvature variations of $\mathtt{arctan}$ on the considered interval I.

Transcendental Functions Underestimators

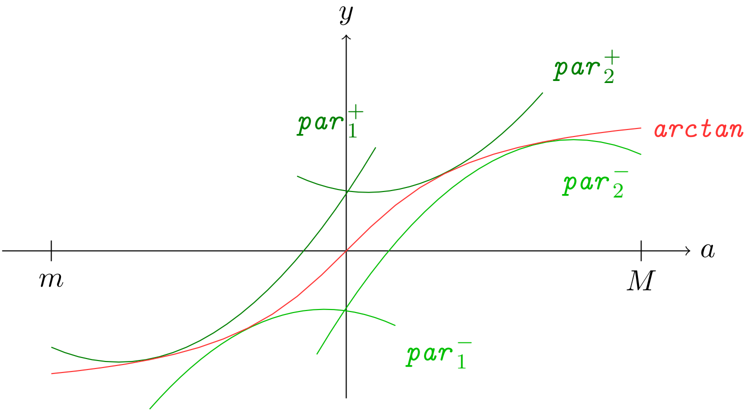
- Let $f \in \mathcal{T}$ be a transcendental univariate elementary function such as $\texttt{arctan}$, exp, ..., defined on a real interval I.
- Basic convexity/quasiconvexity properties and monotonicity of f are used to find lower and upper semi-algebraic bounds.

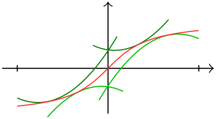
Example with $\texttt{arctan}$:

- $\texttt{arctan}$ is quasiconvex on I:
  $\exists\, c < 0$ such that $\texttt{arctan} - \frac{c}{2}(\cdot)^2$ is convex on I
- $\forall a \in I = [m; M],\ \texttt{arctan}(a) \geq \max\limits_{i \in \mathcal{C}} \{\texttt{par}^-_{a_i}(a)\}$ where $\mathcal{C}$ define an index collection of parabola tangent to the function curve and underestimating f.
- $\texttt{par}^-_{a_i} := \frac{c_i}{2}(a - a_i)^2 + f'_{a_i}(a - a_i) + f(a_i)$,
  $f'_{a_i} = \frac{1}{1 + a_i^2}$, $f(a_i) = \texttt{arctan}(a_i)$.
- c_p depends on a_p and the curvature variations of $\texttt{arctan}$ on the considered interval I.

Transcendental Functions Underestimators

Example with arctan:





$$\max(p_1, p_2) = \frac{p_1 + p_2 + |p_1 - p_2|}{2}$$

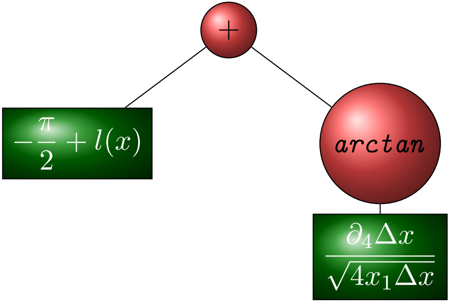
$$z = |p_1 - p_2| \iff z^2 = (p_1 - p_2)^2 \wedge z \geq 0$$

Lemma$_{9922699028}$ from Flyspeck:

- $K := [4; 6.3504]^3 \times [6.3504; 8] \times [4; 6.3504]^2$
  $f := -\frac{\pi}{2} + l(x) + \texttt{arctan} \frac{\partial_4 \Delta x}{\sqrt{4 x_1 \Delta x}}$
- Using semi-algebraic optimization methods:
  $\forall x \in K, \, m \leq \frac{\partial_4 \Delta x}{\sqrt{4 x_1 \Delta x}} \leq M$
- Using the $\texttt{arctan}$ properties: $\forall a \in I = [m; M]$,
  $\texttt{arctan}(a) \geq m_{sa}(a) = \max\{ \texttt{par}^-_{a_1}(a); \texttt{par}^-_{a_2}(a)\}$
- $f^* \geq f^*_{sa} = \min_{x \in K}\{f_{sa}(x) = -\frac{\pi}{2} + l(x) + m_{sa}(x)\}$

Adaptative Semi-algebraic Approximations Algorithm

- The first step is to build the abstract syntax tree from an inequality, where leaves are semi-algebraic functions and nodes are univariate transcendental functions (arctan, exp, ...) or basic operations ($+$, $\times$, $-$, $/$).
- With $l := 1.6294 - 0.2213\,(\sqrt{x_2} + \sqrt{x_3} + \sqrt{x_5} + \sqrt{x_6} - 8.0) + 0.913\,(\sqrt{x_4} - 2.52) + 0.728\,(\sqrt{x_1} - 2.0)$, the tree for the flyspeck example is:



Adaptative Semi-algebraic Approximations Algorithm

algo$_{\mathcal{T}}$

Input: tree t, box K, control points finite sequence $s = x_1, \cdots, x_r \in K$
Output: lower bound m, upper bound M, lower tree t^-, upper tree t^+

if t is semialgebraic **then**
  return min t, max t, t, t
else if t is a transcendental node with a child c **then**
  $m_c, M_c, c^-, c^+ := \texttt{algo}_{\mathcal{T}}\ (t,\ K,\ s)$
  $par^-, par^+ := \texttt{build}_{\texttt{par}}\ (\texttt{t}, \texttt{m}_\texttt{c}, \texttt{M}_\texttt{c}, \texttt{s})$
  $t^-, t^+ := \texttt{compose}\ (par^-, par^+, \texttt{c}^-, \texttt{c}^+)$
  return min t^-, max t^+, t^-, t^+
else if t is a dyadic operation node bop parent of c_1 and c_2 **then**
  $m_{c_i}, M_{c_i}, c_i^-, c_i^+ := \texttt{algo}_{\mathcal{T}}\ (c_i,\ K,\ s)$
  $t^-, t^+ := \texttt{compose}_{\texttt{bop}}\ (\texttt{c}_1^-, \texttt{c}_1^+, \texttt{c}_2^-, \texttt{c}_2^+)$
  return min t^-, max t^+, t^-, t^+
end

Adaptative Semi-algebraic Approximations Algorithm

```
algo_iter
```

Input: tree t , box K, $iter_{\max}$

Output: lower bound m, feasible solution x_{opt}

$s := [\, \texttt{argmin}\ (\texttt{randeval}\ t)\,]$ $\quad \triangleright\ s \in K$

$n := 0$

$m := -\infty$

while $m < 0$ **or** $n \leq iter_{\max}$ **do**

  $m, M, t^{-}, t^{+} := \texttt{algo}_{\mathcal{T}}\ (t,\ K,\ s)$

  $x_{opt} := \texttt{guess\_of\_argmin}\ (t^{-})$ $\quad \triangleright\ t^{-}\ (x_{opt}) = m$

  $s := s \cup \{\, x_{opt}\, \}$

  $n := n + 1$

done

return $m,\ x_{opt}$

Adaptative Semi-algebraic Approximations

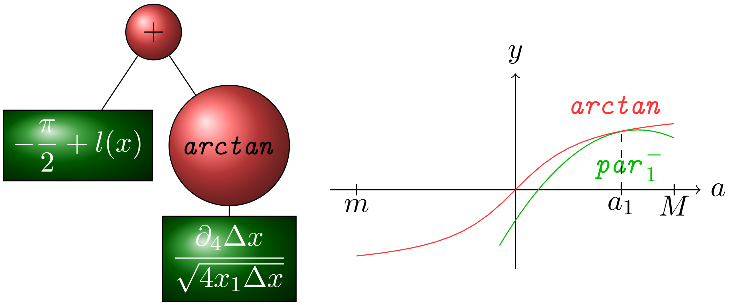
Example from Flyspeck:

$$f_{sa} := \frac{\partial_4 \Delta x}{\sqrt{4 x_1 \Delta x}}, K_{sa} := [4; 6.3504]^3 \times [6.3504; 8] \times [4; 6.3504]^2.$$

- Here, $t = f_{sa}$, this is the first cell of $\texttt{algo}_{\mathcal{T}}$, a lower bound m of $\min f_{sa}$ is computed by rewritting the problem into a POP.
- For a relaxation order $k = 2$, we find $m_2 = -0.618$ and $M_2 = 0.891$. The feasibility error is too big.
- For a relaxation order $k = 3$, we find $m_3 = -0.445$ and $M_3 = 0.87$ with a low feasibility error.
- The argument of *arctan* lies in $[m_3; M_3]$. Notice that it lies also in $[m_2; M_2]$ but the parabola approximations would be less accurate.

Adaptative Semi-algebraic Approximations

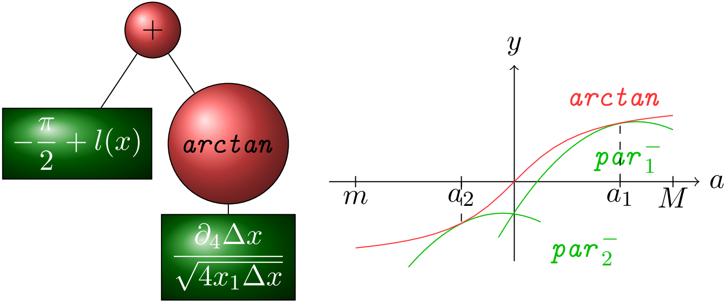
$\mathtt{algo}_{\mathtt{iter}}$ First iteration:



1. Evaluate f with `randeval` and obtain a minimizer guess x_1.
   Compute $a_1 := \frac{\partial_4 \Delta x}{\sqrt{4 x_1 \Delta x}}(x_1) = f_{sa}(x_1) = 0.84460$
2. Get the equation of par_1^- with $\mathtt{build}_{\mathtt{par}}$
3. Compute $m_1 \leq \min_{x \in K} \{-\frac{\pi}{2} + l(x) + par_1^- (f_{sa}(x))\}$

Adaptative Semi-algebraic Approximations

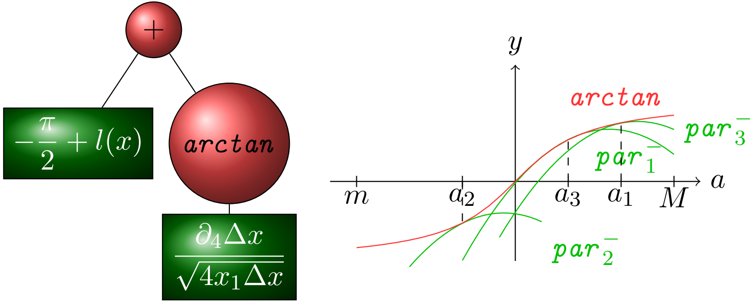
algo$_{\texttt{iter}}$ Second iteration:



1. For $k = 3, m_1 = -0.746 < 0$, obtain a new minimizer x_2.
2. Compute $a_2 := f_{sa}(x_2) = -0.374$ and par_2^-
3. Compute $m_2 \leq \min_{x \in K} \{ -\frac{\pi}{2} + l(x) + \max_{i \in \{1,2\}} \{ par_i^- (f_{sa}(x)) \} \}$

Adaptative Semi-algebraic Approximations

$\texttt{algo}_{\texttt{iter}}$ Third iteration:



1. For $k = 3, m_2 = -0.112 < 0$, obtain a new minimizer x_3.
2. Compute $a_3 := f_{sa}(x_3) = 0.357$ and par_3^-
3. Compute $m_3 \leq \min_{x \in K} \{-\frac{\pi}{2} + l(x) + \max_{i \in \{1,2,3\}} \{par_i^- (f_{sa}(x))\}\}$

Adaptative Semi-algebraic Approximations

- For $k = 3, m_3 = -0.0333 < 0$, obtain a new minimizer x_4 and iterate again...
- Actually, many iterations are needed and if we take $k = 3$ then that is not enough to ensure convergence of $\mathtt{algo}_{\mathtt{iter}}$.
- But the following convergence theorem holds:

Theorem: Convergence of Semi-algebraic underestimators

Let $f \in \mathcal{T}$ and $(x_p^{opt})_{p \in \mathbb{N}}$ be a sequence of control points obtained to define the hierarchy of f-underestimators in the previous algorithm $\mathtt{algo}_{\mathtt{iter}}$ and x^* be an accumulation point of $(x_p^{opt})_{p \in \mathbb{N}}$. Then, x^* is a global minimizer of f on K.

Proof

It comes from the convergence of Lasserre' hierarchy of SDP (the SOS assumption holds) and the properties of the accumulation point.

Adaptative Semi-algebraic Approximations

- For $k = 3, m_3 = -0.0333 < 0$, obtain a new minimizer x_4 and iterate again...
- Actually, many iterations are needed and if we take $k = 3$ then that is not enough to ensure convergence of $\texttt{algo}_{\texttt{iter}}$.
- But the following convergence theorem holds:

Theorem: Convergence of Semi-algebraic underestimators

Let $f \in \mathcal{T}$ and $(x_p^{opt})_{p \in \mathbb{N}}$ be a sequence of control points obtained to define the hierarchy of f-underestimators in the previous algorithm $\texttt{algo}_{\texttt{iter}}$ and x^* be an accumulation point of $(x_p^{opt})_{p \in \mathbb{N}}$. Then, x^* is a global minimizer of f on K.

Proof

It comes from the convergence of Lasserre' hierarchy of SDP (the SOS assumption holds) and the properties of the accumulation point.

Adaptative Semi-algebraic Approximations

- For $k = 3, m_3 = -0.0333 < 0$, obtain a new minimizer x_4 and iterate again...
- Actually, many iterations are needed and if we take $k = 3$ then that is not enough to ensure convergence of $\texttt{algo}_{\texttt{iter}}$.
- But the following convergence theorem holds:

Theorem: Convergence of Semi-algebraic underestimators

Let $f \in \mathcal{T}$ and $(x_p^{opt})_{p \in \mathbb{N}}$ be a sequence of control points obtained to define the hierarchy of f-underestimators in the previous algorithm $\texttt{algo}_{\texttt{iter}}$ and x^* be an accumulation point of $(x_p^{opt})_{p \in \mathbb{N}}$. Then, x^* is a global minimizer of f on K.

Proof

It comes from the convergence of Lasserre' hierarchy of SDP (the SOS assumption holds) and the properties of the accumulation point.

Local Solutions to Global Issues

Two relaxation parameters:

1. Semi-algebraic relaxation order which is the number of parabola, and the size of the sequence s in $\texttt{algo}_{\texttt{iter}}$
2. SDP relaxation order $k \geq \max\{\lceil \deg f_{pop} \rceil / 2, \lceil \deg g_j / 2 \rceil\}$. The size of the moment SDP matrices grows with the SDP-relaxation order and the number of lifting variables: $\mathcal{O}((n+p)^{2k})$ variables and linear matrix inequalities (LMIs) of size $\mathcal{O}((n+p)^{k})$: polynomial in p, exponential in k

The number of parabola increases

$\Downarrow$

The number p of lifting variables increases:
2 by argument of the $\max$)

$\Downarrow$

The size of the SDP problems grows exponentially with the SDP relaxation order

$\texttt{algo}_{\texttt{iter}}$ may not converge in a reasonable time

Local Solutions to Global Issues

Two relaxation parameters:

1. Semi-algebraic relaxation order which is the number of parabola, and the size of the sequence s in $\texttt{algo}_{\texttt{iter}}$
2. SDP relaxation order $k \geq \max\{\lceil \deg f_{pop} \rceil / 2, \lceil \deg g_j / 2 \rceil\}$. The size of the moment SDP matrices grows with the SDP-relaxation order and the number of lifting variables: $\mathcal{O}((n+p)^{2k})$ variables and linear matrix inequalities (LMIs) of size $\mathcal{O}((n+p)^k)$: polynomial in p, exponential in k

The number of parabola increases

$\Downarrow$

The number p of lifting variables increases: 2 by argument of the $\max$)

$\Downarrow$

The size of the SDP problems grows exponentially with the SDP relaxation order

$\texttt{algo}_{\texttt{iter}}$ may not converge in a reasonable time

Local Solutions to Global Issues

Two relaxation parameters:

1. Semi-algebraic relaxation order which is the number of parabola, and the size of the sequence s in $\texttt{algo}_{\texttt{iter}}$
2. SDP relaxation order $k \geq \max\{\lceil \deg f_{pop} \rceil/2, \lceil \deg g_j/2 \rceil\}$.
   The size of the moment SDP matrices grows with the SDP-relaxation order and the number of lifting variables: $\mathcal{O}((n+p)^{2k})$ variables and linear matrix inequalities (LMIs) of size $\mathcal{O}((n+p)^k)$: polynomial in p, exponential in k

The number of parabola increases

$\Downarrow$

The number p of lifting variables increases:
2 by argument of the $\max$)

$\Downarrow$

The size of the SDP problems grows exponentially with the SDP relaxation order

$\texttt{algo}_{\texttt{iter}}$ may not converge in a reasonable time

Two relaxation parameters:

1. Semi-algebraic relaxation order which is the number of parabola, and the size of the sequence s in $\texttt{algo}_{\texttt{iter}}$
2. SDP relaxation order $k \geq \max\{\lceil \deg f_{pop} \rceil / 2, \lceil \deg g_j / 2 \rceil\}$. The size of the moment SDP matrices grows with the SDP-relaxation order and the number of lifting variables: $\mathcal{O}((n+p)^{2k})$ variables and linear matrix inequalities (LMIs) of size $\mathcal{O}((n+p)^k)$: polynomial in p, exponential in k

The number of parabola increases

$\Downarrow$

The number p of lifting variables increases:

2 by argument of the $\max$)

$\Downarrow$

The size of the SDP problems grows exponentially with the SDP relaxation order

$\texttt{algo}_{\texttt{iter}}$ may not converge in a reasonable time

Local Solutions to Global Issues

Two relaxation parameters:

1. Semi-algebraic relaxation order which is the number of parabola, and the size of the sequence s in $\texttt{algo}_{\texttt{iter}}$
2. SDP relaxation order $k \geq \max\{\lceil \deg f_{pop} \rceil / 2, \lceil \deg g_j / 2 \rceil\}$.
   The size of the moment SDP matrices grows with the SDP-relaxation order and the number of lifting variables: $\mathcal{O}((n+p)^{2k})$ variables and linear matrix inequalities (LMIs) of size $\mathcal{O}((n+p)^k)$: polynomial in p, exponential in k

The number of parabola increases

$$\Downarrow$$

The number p of lifting variables increases:
2 by argument of the $\max$)

$$\Downarrow$$

The size of the SDP problems grows exponentially with the SDP relaxation order

$\texttt{algo}_{\texttt{iter}}$ may not converge in a reasonable time

Local Solutions to Global Issues

Two relaxation parameters:

1. Semi-algebraic relaxation order which is the number of parabola, and the size of the sequence s in $\texttt{algo}_{\texttt{iter}}$
2. SDP relaxation order $k \geq \max\{\lceil \deg f_{pop} \rceil/2, \lceil \deg g_j/2 \rceil\}$.
   The size of the moment SDP matrices grows with the SDP-relaxation order and the number of lifting variables: $\mathcal{O}((n+p)^{2k})$ variables and linear matrix inequalities (LMIs) of size $\mathcal{O}((n+p)^{k})$: polynomial in p, exponential in k

The number of parabola increases

$\Downarrow$

The number p of lifting variables increases:

2 by argument of the $\max$)

$\Downarrow$

The size of the SDP problems grows exponentially with the SDP relaxation order

$\texttt{algo}_{\texttt{iter}}$ may not converge in a reasonable time

Local Solutions to Global Issues

- Instead of increasing both relaxation orders, fix the SDP relaxation order $k \leq 3$ (computable SDP in practice) and the number of control points (the number of lifting variables p).
- If $\texttt{algo}_{\texttt{iter}}$ returns a negative lower bound then cut the initial box K in several boxes $(K_i)_{1 \leq i \leq c}$ and solve the inequality on each K_i. But...
- How to partition K?

Multivariate Taylor-Models Underestimators:

- Consider a global minimizer x_c^* candidate obtained after $\texttt{algo}_{\texttt{iter}}$ returned a negative value m_k. For a given r, define the L_∞-ball $\mathcal{B}_{x_c^*, r} := \{x \in K : ||x - x_c^*|| \leq r\}$.
- Then, let $f_{x_c^*, r}$ be the quadratic form defined by:

$$\begin{aligned} f_{x_c^*, r} : \mathcal{B}_{x_c^*, r} & \longrightarrow \mathbb{R} \\ x & \longmapsto f(x_c^*) + \mathcal{D}_f(x_c^*)\,(x - x_c^*) \\ & \quad + \frac{1}{2}(x - x_c^*)^T \mathcal{D}_f^2(x_c^*)\,(x - x_c^*) + \lambda (x - x_c^*)^2 \end{aligned}$$

with $\lambda := \min\limits_{x \in \mathcal{B}_{x_c^*, r}} \{\lambda_{\min}(\mathcal{D}_f^2(x) - \mathcal{D}_f^2(x_c^*))\}$

Theorem:

$\forall x \in \mathcal{B}_{x_c^*, r}, f(x) \geq f_{x_c^*, r}$, that is $f_{x_c^*, r}$ underestimates f on $\mathcal{B}_{x_c^*, r}$.

- $\lambda := \min_{x \in \mathcal{B}_{x_c^*, r}} \{\lambda_{\min}(\mathcal{D}_f^2(x) - \mathcal{D}_f^2(x_c^*))\}$
- Bound the hessian on $\mathcal{B}_{x_c^*, r}$ by interval artithmetic or SDP relaxations to get $\bar{\mathcal{D}}_f^2$:
- Define the symmetric matrix B containing the bounds on the entries of $\bar{\mathcal{D}}_f^2$.
- Let $\mathcal{S}^n$ be the set of diagonal matrices of sign.
  $\mathcal{S}^n := \{\text{diag}\,(s_1, \cdots, s_n),\ s_1 = \pm 1, \cdots\ s_n = \pm 1\}$
  $\lambda := \lambda_{\min}(\bar{\mathcal{D}}_f^2 - \mathcal{D}_f^2(x_c^*))$: minimal eigenvalue of an interval matrix

Robut Optimization with Reduced Vertex Set [Calafiore, Dabbene]

The robust interval SDP problem $\lambda_{\min}(\bar{\mathcal{D}}_f^2 - \mathcal{D}_f^2(x_c^*))$ is equivalent to the following SDP in the single variable $t \in \mathbb{R}$:

$$\begin{cases} \min & -t \\ \text{s.t.} & -t\, I - \mathcal{D}_f^2(x_c^*) - S\, B\, S \succeq 0,\ S = \text{diag}\,(1,\ \tilde{S}),\ \forall \tilde{S} \in \mathcal{S}^{n-1} \end{cases}$$

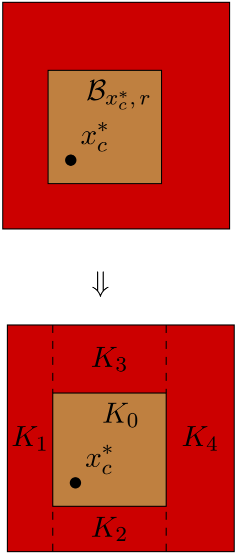
Local Solutions to Global Issues
Branch and Bound Algorithm

$\texttt{algo}_{\texttt{dicho}}$ returns the L_∞-ball $\mathcal{B}_{x_c^*, r}$ of maximal radius r (by dichotomy) such that the underestimator $f_{x_c^*, r}$ is positive on $\mathcal{B}_{x_c^*, r}$

$\texttt{algo}_{\texttt{bb}}$

Input: tree t , K, $iter_{\max}$
Output: lower bound m

$m,\ x_c^* := \texttt{algo}_{\texttt{iter}}\ (t,\ K,\ iter_{\max})$
if $m < 0$ **then**
  $\mathcal{B}_{x_c^*, r} := \texttt{algo}_{\texttt{dicho}}\ (t,\ K,\ x_c^*)$
  Partition $K \setminus \mathcal{B}_{x_c^*, r} := (K_i)_{1 \leq i \leq c}$
  $K_0 := \mathcal{B}_{x_c^*, r}$
  $m := \min\limits_{0 \leq i \leq c} \{\ \texttt{algo}_{\texttt{bb}}\ (t\ ,\ K_i,\ iter_{\max})\ \}$
  return m
else
  return m
end



Ineq. id	$n_{\mathcal{T}}$	n_{vars}	$k_{\max}$	n_{pop}	n_{cuts}	m	cpu time
9922	1	6	2	222	27	3.07×10^{-5}	$20\,min$
3526	1	6	2	156	17	4.89×10^{-6}	$13.4\,min$
6836	1	6	2	173	22	4.68×10^{-5}	$14\,min$
6619	1	6	2	163	21	4.57×10^{-5}	$13.4\,min$
3872	1	6	2	250	30	7.72×10^{-5}	$20.3\,min$
3139	1	6	2	162	17	1.03×10^{-5}	$13.2\,min$
4841	1	6	2	624	73	2.34×10^{-6}	$50.4\,min$
3020	1	5	3	80	9	2.96×10^{-5}	$31\,min$
3318	1	6	3	26	2	3.12×10^{-5}	$1.2\,h$

Lemma$_{7394_6}$ from Flyspeck

$K := [4; 6.3504]^6$

$$\forall x \in K, \frac{\pi}{2} + \sum_{i=1}^{3} \mathtt{arctan}_i \frac{-\partial_4 \Delta x}{\sqrt{4 x_1 \Delta x}} - 0.55125 - 0.196 (\sqrt{x_4} + \sqrt{x_5} + \sqrt{x_6} - 6.0) + 0.38 (\sqrt{x_1} + \sqrt{x_2} + \sqrt{x_3} - 6.0) \geq 0.$$

Ineq. id	$n_{\mathcal{T}}$	n_{vars}	$k_{\max}$	n_{pop}	n_{cuts}	m	cpu time
7726	3	6	2	450	70	1.22×10^{-6}	$3.4\, h$
7394_3	3	3	3	1	0	3.44×10^{-5}	$11\, s$
7394_4	3	4	3	47	10	3.55×10^{-5}	$26\, min$
7394_5	3	5	3	290	55	3.55×10^{-5}	$12\, h$

Conclusion and Further Work

- Results are encouraging for the easiest inequalities even if disjunctions occur.
- We could reduce the computation time by computing underestimators for some semi-algebraic functions like $\frac{\partial_4 \Delta x}{\sqrt{4 x_1 \Delta x}}$ by using SDP again [e.g. Lasserre, Tanh].
- Obtain good feasible points is necessary to get fast convergence of $\texttt{algo}_{\texttt{iter}}$, "joint+marginal" algorithms are available for POP [e.g. Lasserre, Tanh]. Randomization methods could also work out.
- Maybe a hybrid method using both SDP certificates and Solovyev method (interval arithmetic with Taylor-Models) on appropriate subsets of K would be more performant.
- It is possible to perform exact certification for polynomials with rational coefficients [e.g. Kaltofen, Parrilo] in order to verify the positivity certificates with the formal proof assistant COQ.

Conclusion and Further Work

- Results are encouraging for the easiest inequalities even if disjunctions occur.
- We could reduce the computation time by computing underestimators for some semi-algebraic functions like $\frac{\partial_4 \Delta x}{\sqrt{4 x_1 \Delta x}}$ by using SDP again [e.g. Lasserre, Tanh].
- Obtain good feasible points is necessary to get fast convergence of $\mathtt{algo}_{\mathtt{iter}}$, "joint+marginal" algorithms are available for POP [e.g. Lasserre, Tanh]. Randomization methods could also work out.
- Maybe a hybrid method using both SDP certificates and Solovyev method (interval arithmetic with Taylor-Models) on appropriate subsets of K would be more performant.
- It is possible to perform exact certification for polynomials with rational coefficients [e.g. Kaltofen, Parrilo] in order to verify the positivity certificates with the formal proof assistant COQ.

Conclusion and Further Work

- Results are encouraging for the easiest inequalities even if disjunctions occur.
- We could reduce the computation time by computing underestimators for some semi-algebraic functions like $\frac{\partial_4 \Delta x}{\sqrt{4 x_1 \Delta x}}$ by using SDP again [e.g. Lasserre, Tanh].
- Obtain good feasible points is necessary to get fast convergence of $\texttt{algo}_{\texttt{iter}}$, "joint+marginal" algorithms are available for POP [e.g. Lasserre, Tanh]. Randomization methods could also work out.
- Maybe a hybrid method using both SDP certificates and Solovyev method (interval arithmetic with Taylor-Models) on appropriate subsets of K would be more performant.
- It is possible to perform exact certification for polynomials with rational coefficients [e.g. Kaltofen, Parrilo] in order to verify the positivity certificates with the formal proof assistant COQ.

Conclusion and Further Work

- Results are encouraging for the easiest inequalities even if disjunctions occur.
- We could reduce the computation time by computing underestimators for some semi-algebraic functions like $\dfrac{\partial_4 \Delta x}{\sqrt{4 x_1 \Delta x}}$ by using SDP again [e.g. Lasserre, Tanh].
- Obtain good feasible points is necessary to get fast convergence of $\texttt{algo}_{\texttt{iter}}$, "joint+marginal" algorithms are available for POP [e.g. Lasserre, Tanh]. Randomization methods could also work out.
- Maybe a hybrid method using both SDP certificates and Solovyev method (interval arithmetic with Taylor-Models) on appropriate subsets of K would be more performant.
- It is possible to perform exact certification for polynomials with rational coefficients [e.g. Kaltofen, Parrilo] in order to verify the positivity certificates with the formal proof assistant COQ.

Conclusion and Further Work

- Results are encouraging for the easiest inequalities even if disjunctions occur.
- We could reduce the computation time by computing underestimators for some semi-algebraic functions like $\dfrac{\partial_4 \Delta x}{\sqrt{4 x_1 \Delta x}}$ by using SDP again [e.g. Lasserre, Tanh].
- Obtain good feasible points is necessary to get fast convergence of $\texttt{algo}_{\texttt{iter}}$, "joint+marginal" algorithms are available for POP [e.g. Lasserre, Tanh]. Randomization methods could also work out.
- Maybe a hybrid method using both SDP certificates and Solovyev method (interval arithmetic with Taylor-Models) on appropriate subsets of K would be more performant.
- It is possible to perform exact certification for polynomials with rational coefficients [e.g. Kaltofen, Parrilo] in order to verify the positivity certificates with the formal proof assistant COQ.

End

Thank you for your attention!