Certification of inequalities involving transcendental functions using Semi-Definite Programming

Supervisor: Benjamin Werner (TypiCal)
Co-Supervisor: Stéphane Gaubert (Maxplus)

2^{nd} year PhD Victor MAGRON

LIX, École Polytechnique

Wednesday July 11 st 2012

Contents

Flyspeck-Like Problems

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 arctan, arccos, arcsin, exp, log, $|\cdot|$, $(\cdot)^{\frac{1}{p}} (p \in \mathbb{N}_0), +, -, \times, /, \sup, \inf$

Lemma$_{9922699028}$ from Flyspeck

$K := [4; 6.3504]^3 \times [6.3504; 8] \times [4; 6.3504]^2$
$\Delta x := x_1 x_4(-x_1 + x_2 + x_3 - x_4 + x_5 + x_6) + x_2 x_5(x_1 - x_2 + x_3 + x_4 - x_5 + x_6) + x_3 x_6(x_1 + x_2 - x_3 + x_4 + x_5 - x_6) - x_2 x_3 x_4 - x_1 x_3 x_5 - x_1 x_2 x_6 - x_4 x_5 x_6$
$\forall x \in K, -\frac{\pi}{2} - \text{arctan} \frac{-\partial_4 \Delta x}{\sqrt{4 x_1 \Delta x}} + 1.6294 - 0.2213 (\sqrt{x_2} + \sqrt{x_3} + \sqrt{x_5} + \sqrt{x_6} - 8.0) + 0.913 (\sqrt{x_4} - 2.52) + 0.728 (\sqrt{x_1} - 2.0) \geq 0.$

Flyspeck-Like Problems

Hales and Solovyev Method:

- Real numbers are represented by interval arithmetic
- Arithmetic is floating point with IEEE-754 directed rounding
- Analytic functions f are approximated with Taylor expansions with rigorously computed error terms:
  $|f(x) - f(x^0) - \nabla f(x^0)\,(x - x^0)| < \sum_{i,j} m_{ij}\, \epsilon_i\, \epsilon_j,$
  $\epsilon_i = |x_i - x_i^0|$
- The domain K is partitioned into smaller rectangles as needed until the Taylor approximations are accurate enough to yield the desired inequalities.
- The Taylor expansions are generated by symbolic differentiation using the chain rule, product rule, and so forth. A few primitive functions ($\sqrt{\cdot}$, $\frac{1}{\cdot}$, arctan and some common polynomials) are hand-coded.

General Framework

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

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

If the relaxations are accurate enough, $f^* \geq f_{sa}^* \geq f_{pop}^* \geq 0$.

SOS and SDP Relaxations

Polynomial Optimization Problem (POP):

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

SOS Assumption:

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) $u_0, u_1, \ldots, u_m \in \mathbb{R}[X]$

The SOS assumption is always verified if there exists $N \in \mathbb{N}$ such that $N - \sum_{i=1}^{n} X_i^2 = u_0 + \sum_{j=1}^{m} u_j \, g_j$. In our case, it as always verified since all the polynomial variables X_i are bounded.

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]:

$\exists L : \mathbb{R}[X] \to \mathbb{R}$ s.t. $(\exists \mu \in \mathcal{P}(K_{pop}), \forall p \in \mathbb{R}[X], L(p) = \int p \, d\mu) \iff (L(1) = 1$ and $L(s_0 + \sum_{j=1}^{m} s_j g_j) \geq 0$ for any SOS $s_0, \ldots, s_m \in \mathbb{R}[X])$.

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 psd $\}$, with $g_0 = 1, \mathcal{L}_{g_0}, \ldots, \mathcal{L}_{g_m}$ 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

- Let $\mathcal{B}$ the monomial basis $(X^\alpha)_{\alpha \in \mathbb{N}^n}$ and set $y_b = L(b)$ for $b \in \mathcal{B}$ identifies L with the infinite series $y = (y_b)_{b \in \mathcal{B}}$.
- 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 $k \geq k_0 := \max\{\lceil \deg f_{pop} \rceil / 2, \lceil \deg g_0 / 2 \rceil, ..., \lceil \deg g_m / 2 \rceil\}$. By truncating the previous matrices by considering only rows and columns indexed by elements in $\mathcal{B}$ of degree at most k, consider the hierarchy Q_k of semidefinite relaxations:

$$
Q_k : \left\{ \begin{array}{rll} & \inf_y L(f) & \\ M_{k - \lceil \deg g_j / 2 \rceil}(g_j y) & \succcurlyeq & 0, \quad 0 \leq j \leq m, \\ y_1 & = & 1 \end{array} \right.
$$

SOS and SDP Relaxations

Convergence Theorem [Lasserre]:

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

SDP relaxations:

Let $B = |\mathcal{B}|$. Many solvers (Sedumi, SDPA) solve the following standard form semidefinite program and its dual:

$$
(SDP) \begin{cases}
\mathcal{P}: & \min\limits_{y} & \sum\limits_{\alpha=1}^{B} c_\alpha y_\alpha \\
& \text{subject to} & \sum\limits_{\alpha=1}^{B} F_\alpha \, y_\alpha - F_0 \succcurlyeq 0 \\
\mathcal{D}: & \max\limits_{Y} & \text{Trace}\,(F_0 \, Y) \\
& \text{subject to} & \text{Trace}\,(F_\alpha \, Y) = c_\alpha \, (\alpha = 1, \ldots, B)
\end{cases}
$$

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 $K_{sa} := \{x \in \mathbb{R}^n \ : \ g_1(x) \geq 0, ..., g_m(x) \geq 0\}$ 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 (a b.s.a.l.), or f is basic semi-algebraic (b.s.a.) if $\exists p, s \in \mathbb{N}$, polynomials $(h_k)_{1 \leq k \leq s} \in \mathbb{R}[X, Z_1, ..., Z_p]$ and a b.s.a. set
$K_{pop} := \{(x, \ z) \in \mathbb{R}^{n+p} : x \in K_{sa}, \ h_k(x, z) \geq 0, \ k = 1, ..., s\}$
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 $z_1 := \sqrt{4 x_1 \Delta x}$, $m_1 = \inf_{x \in K_{sa}} z_1(x)$, $M_1 = \sup_{x \in K_{sa}} z_1(x)$.
- Define $h_1 := z_1 - m_1$, $h_2 := M_1 - z_1$, $h_3 := z_1^2 - \sqrt{4 x_1 \Delta x}$, $h_4 := -z_1^2 + \sqrt{4 x_1 \Delta x}$, $h_5 := z_1$, $h_6 := z_2 \, z_1 + \partial_4 \Delta x$, $h_7 := -z_2 \, z_1 - \partial_4 \Delta x$, $s = 7$, $p = 2$.
- $K_{pop} := \{(x, \, z) \in \mathbb{R}^{6+2} : x \in K_{sa}, \; h_k(x, z) \geq 0, \; k = 1, \ldots, 7\}$.
- $\Psi_{f_{sa}} := \{(x, \, f_{sa}(x)) : x \in K_{sa}\} = \{(x, \, z_2) : (x, \, z) \in K_{pop}\}$.

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

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

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

$$Q_k : \begin{array}{rcll} \inf\limits_{y} y_{0\dots01} & & & \\ M_{k-1}(g_j\, y) & \succcurlyeq & 0, & 1 \leq j \leq 12, \\ M_{k-\lceil \deg h_k/2 \rceil}(h_k\, y) & \succcurlyeq & 0, & 1 \leq k \leq 7, \\ y_{0\dots0} & = & 1 & \end{array}$$

b.s.a.l. Convergence:

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

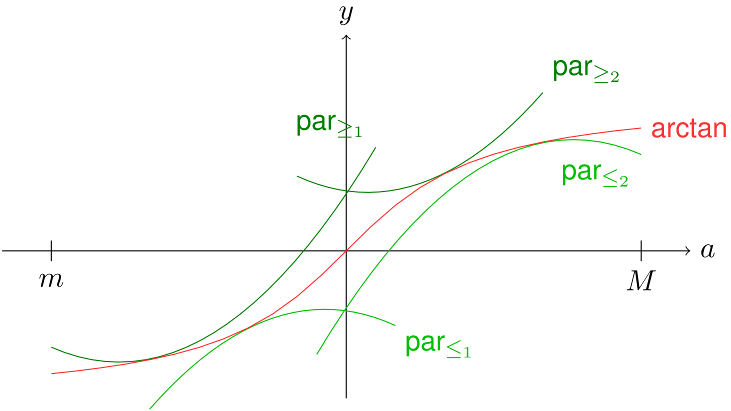
Transcendental Functions Underestimators

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

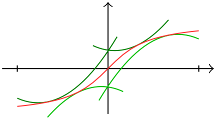
Example with arctan:

- $\forall a \in I = [m; M]$, $\text{arctan}(a) \geq \max_{p \in \mathcal{C}} \{ \text{par}_{\leq_p}(a) \}$ where $\mathcal{C}$ define an index collection of parabola tangent to the function curve and minoring f.
- $\text{par}_{\leq_p} := \frac{c_p}{2}(a - a_p)^2 + f'_{a_p}(a - a_p) + f(a_p)$, $f'_{a_p} = \frac{1}{1 + a_p^2}$, $f(a_p) = \text{arctan}(a_p)$.
- c_p depends on a_p and the curvature variations of arctan on the considered interval I. This is a consequence of the convexity of the $\text{arctan}(\cdot) - \frac{c_p}{2}(\cdot)^2$ function for a well-chosen c_p.

Example with arctan:



Transcendental Functions Underestimators



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

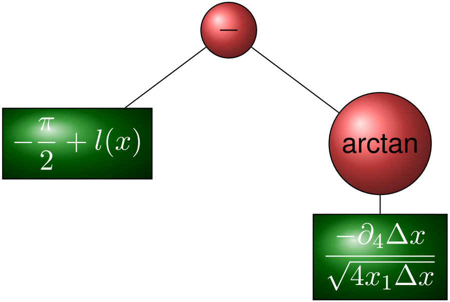
$$z = |p_1 - p_2| \iff z^2 = (p_1 - p_2)^2 \wedge z \geq 0$$

Lemma$_{9922699028}$ from Flyspeck:

- $K := [4; 6.3504]^3 \times [6.3504; 8] \times [4; 6.3504]^2$
  $f := -\frac{\pi}{2} - \text{arctan} \frac{-\partial_4 \Delta x}{\sqrt{4x_1 \Delta x}} + 1.6294 - 0.2213 (\sqrt{x_2} + \sqrt{x_3} + \sqrt{x_5} + \sqrt{x_6} - 8.0) + 0.913 (\sqrt{x_4} - 2.52) + 0.728 (\sqrt{x_1} - 2.0)$
- Using semi-algebraic optimization methods:
  $\forall x \in K, \; m \leq \frac{-\partial_4 \Delta x}{\sqrt{4x_1 \Delta x}} \leq M$
- Using the arctan properties: $\forall a \in I = [m; M]$,
  $\text{arctan}(a) \; \leq m_{sa}(a) = \min \{ \text{par}_{\geq_1}(a); \text{par}_{\geq_2}(a) \}$
- $f^* \geq f^*_{sa} = \min_{x \in K} \{ f_{sa}(x) =$
  $-\frac{\pi}{2} - m_{sa}(x) + 1.6294 - 0.2213 (\sqrt{x_2} + \sqrt{x_3} + \sqrt{x_5} + \sqrt{x_6} -$

Multi-Relaxations 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:



Multi-Relaxations Algorithm

$\texttt{algo}_{\mathcal{T}}$

Require: tree t, box K, sequence $s = (x_k)_{1 \leq k \leq r} \in K^r$
Ensure: lower bound m, upper bound M, lower tree $t_{\leq}$, upper tree $t_{\geq}$

```
if t is s.a. 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_≥ := algo_T (t, K, s)
        t_≤, t_≥ := relax (t, m_c, M_c, c_≤, 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 := algo_T (c_i, K, s)
            t_≤, t_≥ := bop (c_≤1, c_≥1, c_≤2, c_≥2)
            return min t_≤, max t_≥, t_≤, t_≥
        end if
    end if
end if
```

Multi-Relaxations Algorithm

$\texttt{algo}_{\texttt{newton}}$

Require: tree t , box K, tol

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

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

n := 0

m := -1

while $m < 0$ **or** $n \leq tol$ **do**

$m, M, t_{\leq}, t_{\geq}$:= $\texttt{algo}_{\mathcal{T}}$ (t, K, s)

x_{opt} := $\texttt{argmin}$ $t_{\leq}$ $\{t_{\leq}(x_{opt}) = m\}$

s := x_{opt} :: s

$n := n + 1$

end while

return m, x_{opt}

Multi-Relaxations Algorithm

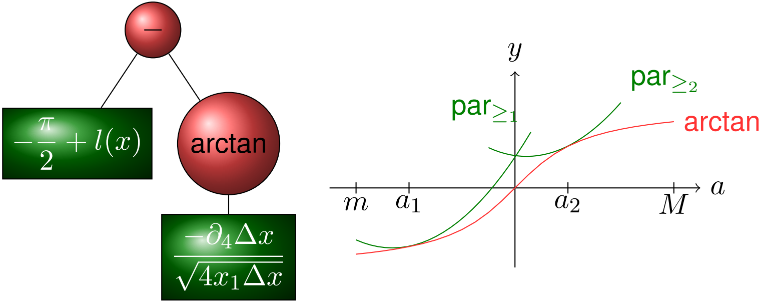
Example from Flyspeck:

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

- Here, $t = f_{sa}$, this is the first case of $\texttt{algo}_{\mathcal{T}}$, $\min f_{sa}$ is computed by rewritting the problem into a POP.
- Then solve the corresponding SDP problem Q_k for a given k. If the computed point is not a feasible solution, increase the relaxation order k.

Multi-Relaxations Algorithm

Semi-algebraic relaxations:



1. Compute $a_1 = f_{sa}(\texttt{argmin}(\texttt{randeval}\ t))$, the equation of par$_{\geq 1}$ and finally $\min t_{\leq 1}$. This is the first $\texttt{algo}_{\texttt{newton}}$ iteration.
2. Suppose that $\min t_{\leq 1} \leq 0$. Then, the POP solver returns a point $x_2 \in K$ with $\min t_{\leq 1} = t_{\leq 1}(x_2)$. Then compute $a_1 = f_{sa}(x_2)$, par$_{\geq 2}$, $\min t_{\leq 2}$
3. Repeat the procedure i times until $\min t_{\leq i} \geq 0$

Local Solutions to Global Issues

Two relaxation order types:

1. Semi-algebraic relaxation order which is the number of considered parabola, and the size of the sequence s in $\texttt{algo}_{\texttt{newton}}$
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 $\mathcal{O}(n^k)$

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 growing exponentially, $\texttt{algo}_{\texttt{newton}}$ fails to converge in a reasonable time

Local Solutions to Global Issues

Instead of increasing both relaxation orders, fix a tolerance for both and if $\texttt{algo}_{\texttt{newton}}$ fails to converge, cut the initial box K in several boxes $(K_i)_{1 \leq i \leq c}$ and solve the inequality on each K_i. But...

1. Where K should be cut?
2. How to partition K?

Multivariate Taylor-Models Underestimators:

- Let $x_{cut} \in K$ a point obtained by $\texttt{algo}_{\texttt{newton}}$ (f) after reaching the tolerance of both relaxation orders.
- Let f_{TM2} the quadratic form related to the second order Multivariate Taylor polynomial defined on a neighborhood $\mathcal{B}_{x_{cut}, r}$ of the point x_{cut}.
- Let $\lambda := \min_{x \in \mathcal{B}_{x_{cut}, r}} \{\lambda_{\min}(\mathcal{H}_f(x) - \mathcal{H}_f(x_{cut}))\}$
- $f_{TM2} := f(x) - f(x_{cut}) - \nabla_f(x_{cut})(x - x_{cut}) - \frac{1}{2}(x - x_{cut})^T \mathcal{H}_f(x_{cut})(x - x_{cut}) - \frac{1}{2}\lambda(x - x_{cut})^2$.

Theorem:

$\forall x \in \mathcal{B}_{x_{cut}, r}, \; f(x) \geq f_{TM2}(x)$

`algo`$_{\texttt{dicho}}$

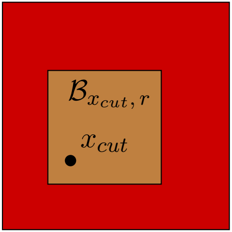
Require: tree t , K, x_{cut}, r_1, r_2, r_{tol}
Ensure: lower bound m

$r := \frac{r_1 + r_2}{2}$

Compute the infinite squared ball $\mathcal{B}_{x_{cut}, r}$ whose edges are parallel to the K ones and f_{TM2}

$m := \min_{x \in \mathcal{B}_{x_{cut}, r}} f_{TM2}$

if $m \geq 0$ **and** $|r_1 - r| \leq r_{tol}$ **then**
  return m
else
  if $m < 0$ **then**
    return `algo`$_{\texttt{dicho}}$ $(t, K, x_{cut}, r_1, r, r_{tol})$
  else
    return `algo`$_{\texttt{dicho}}$ $(t, K, x_{cut}, r, r_2, r_{tol})$
  end if
end if



algo$_{\texttt{bb}}$

Require: tree t , K, tol, r_{tol}

Ensure: lower bound m

m, x_{cut} := $\texttt{algo}_{\texttt{newton}}(\texttt{t, K, tol})$

if $m \leq 0$ **then**

$r_1 := 0$; $r_2 := \min\{\texttt{length}(\texttt{edges}\ (K))\}$

$r := \texttt{algo}_{\texttt{dicho}}\ (t,\ K,\ x_{cut},\ r_1,\ r_2,\ r_{tol})$

Compute the infinite squared ball $\mathcal{B}_{x_{cut}, r}$

Get a partition of $K \setminus \mathcal{B}_{x_{cut}, r} := (K_i)_{1 \leq i \leq c}$

$K_0 := \mathcal{B}_{x_{cut}, r}$

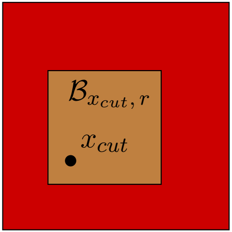
$m := \min_{0 \leq i \leq c} \{\ \texttt{algo}_{\texttt{bb}}\ (t\ ,\ K_i,\ tol,\ r_{tol})\ \}$

return m

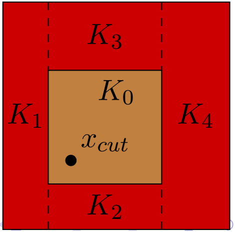
else

return m

end if



$\Downarrow$



- 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 polynomial variables correlation sparsity pattern maximal cliques.

End

Thank you for your attention!