Approximating Pareto Curves using Semidefinite Relaxations

Victor MAGRON

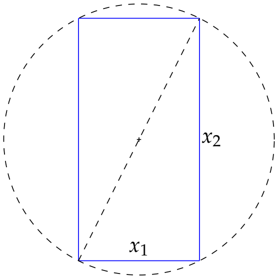
Postdoc LAAS-CNRS
(Joint work with Didier Henrion and Jean-Bernard Lasserre)

Optimisation Non Linéaire en Variables Continues et Discrètes
18 June 2014

LAAS

Multiobjective Polynomial Optimization

- Optimization Problems with several criteria in engineering, economics, applied mathematics.
- Design of a beam of length l, heigth x_1 and width x_2:
  1. **light** construction: minimize the volume lx_1x_2
  2. **cheap** construction: minimize the sectional area $\pi/4(x_1^2 + x_2^2)$
  3. under stress and nonnegativity constraints



Multiobjective Polynomial Optimization

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

$$(\mathbf{P}) \left\{ \min_{\mathbf{x} \in \mathbf{S}} (f_1(\mathbf{x})\ f_2(\mathbf{x}))^\top \right\}$$

Assumption

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

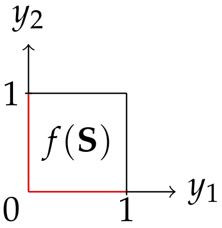
Multiobjective Polynomial Optimization

Definition

Let the previous assumption be satisfied.

- A point $\bar{\mathbf{x}} \in \mathbf{S}$ is called an *Edgeworth-Pareto (EP) optimal point* of Problem **P**, when there is no $\mathbf{x} \in \mathbf{S}$ such that $f_j(\mathbf{x}) \leqslant f_j(\bar{\mathbf{x}})$, $j = 1, 2$ and $f(\mathbf{x}) \neq f(\bar{\mathbf{x}})$.
- A point $\bar{\mathbf{x}} \in \mathbf{S}$ is called a *weakly (EP) optimal point* of Problem **P**, when there is no $\mathbf{x} \in \mathbf{S}$ such that $f_j(\mathbf{x}) < f_j(\bar{\mathbf{x}})$, $j = 1, 2$.

$$f_1(\mathbf{x}) := x_1 \ ,$$
$$f_2(\mathbf{x}) := x_2 \ ,$$
$$\mathbf{S} := \{\mathbf{x} \in \mathbb{R}^2 : 0 \leqslant x_1 \leqslant 1, \ 0 \leqslant x_2 \leqslant 1\} \ .$$



Pareto Curve

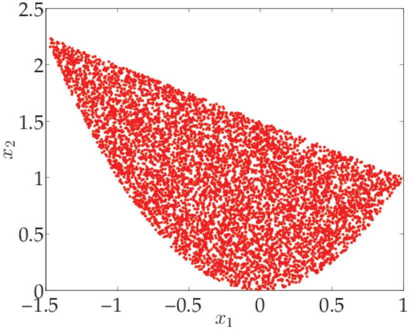
Definition

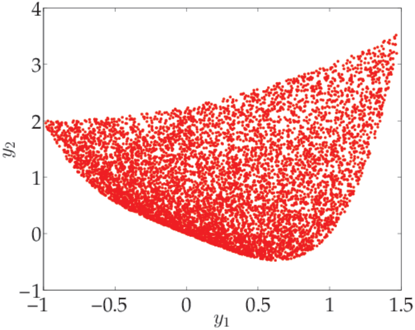
The image set of weakly Edgeworth-Pareto optimal points is called the *Pareto curve*.

Some Examples: $f(\mathbf{S}) + \mathbb{R}^2_+$ is convex

$$g_1 := -x_1^2 + x_2 \ ,$$
$$g_2 := -x_1 - 2x_2 + 3 \ ,$$
$$\mathbf{S} := \{\mathbf{x} \in \mathbb{R}^2 : g_1 \geqslant 0, g_2 \geqslant 0\} \ .$$

$$f_1 := -x_1 \ ,$$
$$f_2 := x_1 + x_2^2 \ .$$

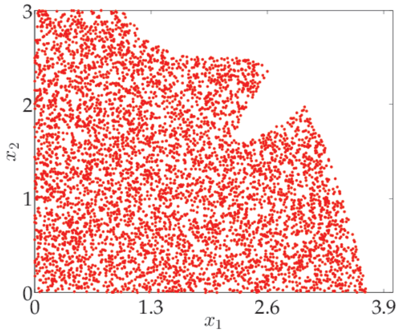


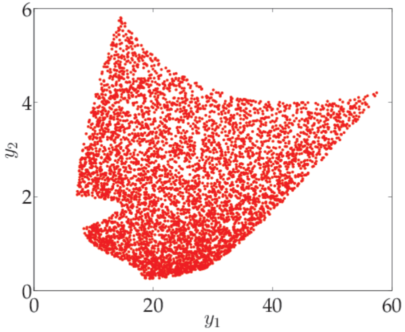


Some Examples: $f(\mathbf{S}) + \mathbb{R}^2_+$ is not convex

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

$$f_1 := (x_1 + x_2 - 7.5)^2/4 + (-x_1 + x_2 + 3)^2 \ ,$$
$$f_2 := (x_1 - 1)^2/4 + (x_2 - 4)^2/4 \ .$$





Scalarization Techniques

- Common workaround by reducing **P** to a scalar POP :

$$(\mathbf{P}_\lambda^p) \left\{ \min_{\mathbf{x}\in\mathbf{S}} f^p(\lambda, \mathbf{x}) := ((\lambda|f_1(\mathbf{x}) - \mu_1|)^p + ((1-\lambda)|f_2(\mathbf{x}) - \mu_2|)^p)^{\frac{1}{p}} \right\} ,$$

with the weight $\lambda \in [0,1]$ and the goals $\mu_1, \mu_2 \in \mathbb{R}$.

- Possible choice: $\mu_j < \min_{\mathbf{x}\in\mathbf{S}} f_j(\mathbf{x}), \ j = 1, 2.$

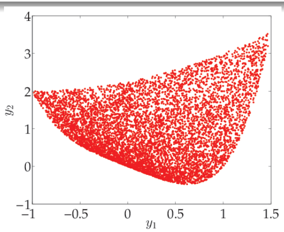
Weighted convex sum approximation: method (a)

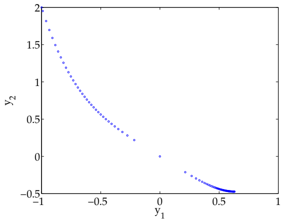
$$(\mathbf{P}_\lambda^1): \quad f^1(\lambda) := \min_{\mathbf{x}\in\mathbf{S}} f^1(\lambda, \mathbf{x})$$

$$f^1(\lambda, \mathbf{x}) := \lambda f_1(\mathbf{x}) + (1-\lambda) f_2(\mathbf{x})$$

Theorem ([Borwein 77], [Arrow-Barankin-Blackwell 53])

Assume that $f(\mathbf{S}) + \mathbb{R}_+^2$ is convex. A point $\bar{\mathbf{x}} \in \mathbf{S}$ is an EP optimal point of Problem $\mathbf{P}$ $\iff$ $\exists \lambda$ such that $\bar{\mathbf{x}}$ is an image unique solution of Problem $\mathbf{P}_\lambda^1$.





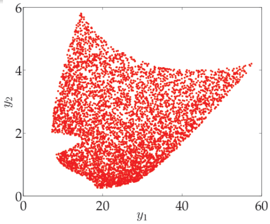
Weighted convex sum approximation: method (a)

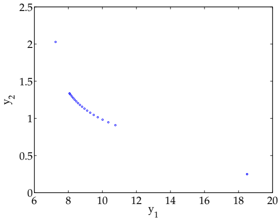
$$(\mathbf{P}^1_\lambda): \quad f^1(\lambda) := \min_{\mathbf{x}\in\mathbf{S}} f^1(\lambda, \mathbf{x})$$

$$f^1(\lambda, \mathbf{x}) := \lambda f_1(\mathbf{x}) + (1-\lambda) f_2(\mathbf{x})$$

Theorem ([Borwein 77], [Arrow-Barankin-Blackwell 53])

Assume that $f(\mathbf{S}) + \mathbb{R}^2_+$ is convex. A point $\bar{\mathbf{x}} \in \mathbf{S}$ is an EP optimal point of Problem $\mathbf{P} \iff \exists\lambda$ such that $\bar{\mathbf{x}}$ is an image unique solution of Problem $\mathbf{P}^1_\lambda$.





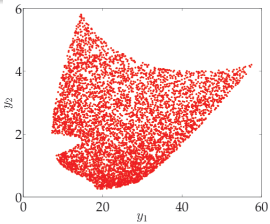
Weigthed Chebyshev approximation: method (b)

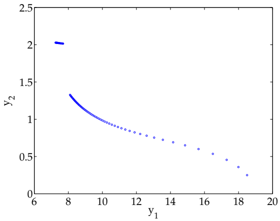
$$(\mathbf{P}^{\infty}_{\lambda}) : \quad f^{\infty}(\lambda) := \min_{\mathbf{x} \in \mathbf{S}} f^{\infty}(\lambda, \mathbf{x})$$

$$f^{\infty}(\lambda, \mathbf{x}) := \max\{\lambda(f_1(\mathbf{x}) - \mu_1), (1 - \lambda)(f_2(\mathbf{x}) - \mu_2)\}$$

Theorem ([Jahn 10, Corollary 11.21 (a)], [Bowman 76], [Steuer-Choo 83])

Suppose that $\forall \mathbf{x} \in \mathbf{S}, \mu_j < f_j(\mathbf{x}),\ j = 1, 2$. A point $\bar{\mathbf{x}} \in \mathbf{S}$ is an EP optimal point of Problem $\mathbf{P}$ $\iff$ $\exists \lambda \in (0, 1)$ such that $\bar{\mathbf{x}}$ is an image unique solution of Problem $\mathbf{P}^{\infty}_{\lambda}$.





Parametric sublevel set approximation: method (c)

Inspired by previous research on multiobjective linear optimization [1]

For each $\lambda \in [a_1, b_1]$, consider the following parametric POP

$$(\mathbf{P}_\lambda^u): \quad f^u(\lambda) := \min_{\mathbf{x} \in \mathbf{S}} \{ f_2(\mathbf{x}) : f_1(\mathbf{x}) \leqslant \lambda \} \ ,$$

with $a_1 := \min_{\mathbf{x} \in \mathbf{S}} f_1(\mathbf{x})$, $b_1 := f_1(\bar{\mathbf{x}})$ and $\bar{\mathbf{x}}$ a solution of $\min_{\mathbf{x} \in \mathbf{S}} f_2(\mathbf{x})$.

Lemma

Suppose that $\bar{\mathbf{x}} \in \mathbf{S}$ is an optimal solution of Problem $\mathbf{P}_\lambda^u$, with $\lambda \in [a_1, b_1]$. Then $\bar{\mathbf{x}}$ belongs to the set of weakly EP points of Problem $\mathbf{P}$.

[1] B. Gorissen, D. den Hertog. *Approximating the pareto set of multiobjective linear programs via robust optimization.* (2012)

Questions

- Is it mandatory to use discretization schemes?
- Can we approximate the Pareto curve in a relatively strong sense?

Contributions

Yes!

We provide two approaches together with numerical schemes that **avoid computing finitely many points**.

1. Parametric POP: for **methods (a) and (b)** (resp. **method (c)**), we approximate the Pareto curve with polynomials so that convergence in L_2-norm (resp. L_1-norm) holds
2. Hierarchy of outer approximation: we provide certified underestimators of the Pareto curve with strong convergence to $f(\mathbf{S})$ in L_1-norm

Contributions

Yes!

We provide two approaches together with numerical schemes that **avoid computing finitely many points**.

1. Parametric POP: for **methods (a) and (b)** (resp. **method (c)**), we approximate the Pareto curve with polynomials so that convergence in L_2-norm (resp. L_1-norm) holds
2. Hierarchy of outer approximation: we provide certified underestimators of the Pareto curve with strong convergence to $f(\mathbf{S})$ in L_1-norm

Contributions

Yes!

We provide two approaches together with numerical schemes that **avoid computing finitely many points**.

1. Parametric POP: for **methods (a) and (b)** (resp. **method (c)**), we approximate the Pareto curve with polynomials so that convergence in L_2-norm (resp. L_1-norm) holds
2. Hierarchy of outer approximation: we provide certified underestimators of the Pareto curve with strong convergence to $f(\mathbf{S})$ in L_1-norm

Outline

Preliminaries: method (a)

Parametric POP $(\mathbf{P}_\lambda^1)$: $f^*(\lambda) := f^1(\lambda) = \min_{\mathbf{x} \in \mathbf{S}} f(\lambda, \mathbf{x})$

Assumption

For almost all $\lambda \in [0,1]$, the solution $\mathbf{x}^*(\lambda)$ of the scalarized problem $(\mathbf{P}_\lambda^1)$ is unique.

Non-uniqueness may be tolerated on a Borel set $B \subset [0,1]$, in which case one assumes image uniqueness of the solution.

Preliminaries: method (a)

Parametric POP $(\mathbf{P}_\lambda^1)$: $f^*(\lambda) := f^1(\lambda) = \min_{\mathbf{x} \in \mathbf{S}} f(\lambda, \mathbf{x})$

- Let $\mathbf{K} := [0, 1] \times \mathbf{S}$
- Let $\mathcal{M}(\mathbf{K})$ the set of probability measures supported on $\mathbf{K}$

$$(P) \begin{cases} \rho := \min_{\mu \in \mathcal{M}(\mathbf{K})} & \int_{\mathbf{K}} f(\lambda, \mathbf{x}) d\mu(\lambda, \mathbf{x}) \\ \quad \text{s.t.} & \int_{\mathbf{K}} \lambda^k d\mu(\lambda, \mathbf{x}) = 1/(1+k), \; k \in \mathbb{N} \; . \end{cases}$$

Preliminaries: method (a)

Lemma (Corollary of [2, Theorem 2.2])

Problem (P) has an optimal solution $\mu^* \in \mathcal{M}(\mathbf{K})$. Then,

$$\rho = \int_{\mathbf{K}} f(\lambda, \mathbf{x}) d\mu^* = \int_0^1 f^*(\lambda) d\lambda \ .$$

Moreover, suppose that (P) has a unique (or image unique) global minimizer $\mathbf{x}^*(\lambda) \in \mathbf{S}$ and let $f_j^*(\lambda) := f_j(\mathbf{x}^*(\lambda)), \ j = 1, 2$. Then,

$$\rho = \int_0^1 [\lambda f_1^*(\lambda) + (1 - \lambda) f_2^*(\lambda)] d\lambda \ .$$

[2] J.B. Lasserre. A "joint + marginal" approach to parametric polynomial optimization (2010)

A hierarchy of semidefinite relaxations

- Let $g \in \mathbb{R}[\lambda, \mathbf{x}]$ with $g(\lambda, \mathbf{x}) := \sum_{k,\alpha} g_{k\alpha} \lambda^k \mathbf{x}^\alpha$.
- Consider the real sequence $\mathbf{z} = (z_{k\alpha}), \ (k, \alpha) \in \mathbb{N}_d^{n+1}$
- Consider the linear functional $L_z(g) := \sum_{k,\alpha} g_{k\alpha} z_{k\alpha}$

A hierarchy of semidefinite relaxations

- Let $g_0 := 1$.
- Let $d_0 := \max\{\deg f_1, \deg f_2, \deg g_1, \dots, \deg g_m\}$.

Consider the semidefinite relaxations of (P) for $d \geqslant d_0$:

$$(P_d) \begin{cases} \min_{\mathbf{z}} & L_{\mathbf{z}}(f) \\ \text{s.t.} & \mathbf{M}_{d-v_l}(g_l\, \mathbf{z}) \succcurlyeq 0,\ l = 0, \dots, m\ , \\ & L_{\mathbf{z}}(\lambda^k) = 1/(1+k),\ k = 0, \dots, 2d\ . \end{cases}$$

- $\mathbf{M}_d(\mathbf{z})$ is the moment matrix associated with $\mathbf{z}$
- $\mathbf{M}_{d-v_l}(g_l\, \mathbf{z})$ is the localizing matrix associated with $\mathbf{z}$ and g_l

Polynomial underestimators of $f^*(\lambda)$

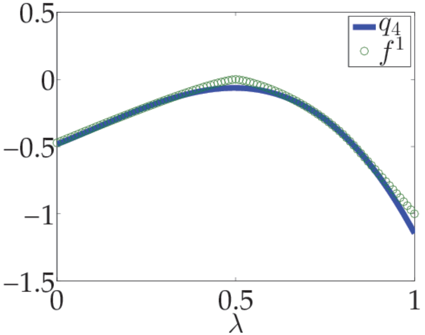
The dual SDP of (P_d) reads:

$$
(D_d) \begin{cases} \max\limits_{q,(\sigma_l)} & \sum\limits_{k=0}^{2d} q_k / (1+k) \\ \text{s.t.} & f(\lambda, \mathbf{x}) - q(\lambda) = \sum\limits_{l=0}^{m} \sigma_l(\lambda, \mathbf{x}) g_l(\mathbf{x}) \\ & q \in \mathbb{R}_{2d}[\lambda],\ \sigma_l \in \Sigma[\lambda, \mathbf{x}],\ l = 0, \dots, m\ , \\ & \deg(\sigma_l g_l) \leqslant 2d,\ l = 0, \dots, m\ . \end{cases}
$$

- The hierarchy (D_d) provides a sequence (q_d) of **polynomial underestimators** of $f^*(\lambda)$.
- $\lim\limits_{d \to \infty} \int_0^1 (f^*(\lambda) - q_d(\lambda)) d\lambda = 0$

Polynomial underestimators of $f^*(\lambda)$

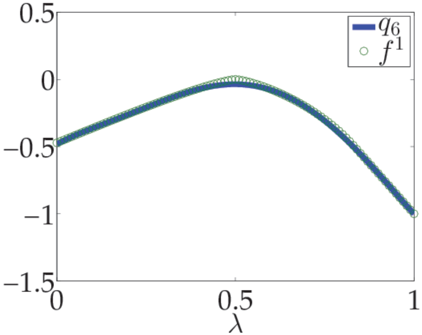
On the convex example:



Degree 4 underestimator

Polynomial underestimators of $f^*(\lambda)$

On the convex example:



Degree 6 underestimator

An inverse problem from generalized moments

Lemma (Corollary of [2, Theorem 3.3])

Assume that for a.a. $\lambda \in [0,1]$, Problem (P) has a unique global optimizer $\mathbf{x}^*(\lambda)$ and let $\mathbf{z}^d = (z^d_{k\alpha})$ be an optimal solution of (P_d). Then,

$$\lim_{d \to \infty} z^d_{k\alpha} = \int_0^1 \lambda^k (\mathbf{x}^*(\lambda))^\alpha d\lambda, \; k \in \mathbb{N} \; .$$

In particular, for $s \in \mathbb{N}$,

$$m^k_j := \lim_{d \to \infty} \sum_\alpha f_{j\alpha} z^d_{k\alpha} = \int_0^1 \lambda^k f^*_j(\lambda) d\lambda, \; j = 1, 2, \; k = 0, \ldots, s \; .$$

[2] J.B. Lasserre. A “joint + marginal” approach to parametric polynomial optimization (2010)

An inverse problem from generalized moments

For a fixed $s \in \mathbb{N}$, one can compute:

- Approximation $\mathbf{m}_j^{sd}$ of the vector $\mathbf{m}_j^s := (m_j^k)$
- Approximations of $f_j^*(\lambda)$, $j = 1, 2$, by solving:

$$\min_{h \in \mathbb{R}_s[\lambda]} \left\{ \int_0^1 (f_j^*(\lambda) - h(\lambda))^2 d\lambda \right\}, \; j = 1, 2 \; .$$

An inverse problem from generalized moments

Theorem

The Problem $\min_{h \in \mathbb{R}_s[\lambda]} \left\{ \int_0^1 (f_j^*(\lambda) - h(\lambda))^2 d\lambda \right\}$ has an optimal solution $h_{sj} \in \mathbb{R}_s[\lambda]$, whose vector of coefficients is $\mathbf{h}_{sj} = \mathbf{H}_s^{-1} \mathbf{m}_j$, $j = 1, 2$, where $\mathbf{H}_s \in \mathcal{S}^{2s+1}$ is the Hankel matrix, whose entries are defined by:

$$\mathbf{H}_s(a, b) := 1/(1 + a + b), \ \ a, b = 0, \ldots, 2s \ .$$

An inverse problem from generalized moments

Proof.

$$\int_0^1 (f_j^*(\lambda) - h(\lambda))^2 d\lambda = \underbrace{\int_0^1 f_j^*(\lambda)^2 d\lambda}_{A} - 2 \underbrace{\int_0^1 f_j^*(\lambda) h(\lambda) d\lambda}_{B} + \underbrace{\int_0^1 h(\lambda)^2 d\lambda}_{C},$$

$$B = \mathbf{h}'\mathbf{m}_j, \quad C = \mathbf{h}'\mathbf{H}_s\mathbf{h} \quad ,$$

thus the problem can be reformulated as:

$$\min_{\mathbf{h}} \{\mathbf{h}'\mathbf{H}_s\mathbf{h} - 2\mathbf{h}'\mathbf{m}_j\}, \quad j = 1, 2 \quad .$$

□

An inverse problem from generalized moments

Proof.

$$\int_0^1 (f_j^*(\lambda) - h(\lambda))^2 d\lambda = \underbrace{\int_0^1 f_j^*(\lambda)^2 d\lambda}_{A} - 2 \underbrace{\int_0^1 f_j^*(\lambda) h(\lambda) d\lambda}_{B} + \underbrace{\int_0^1 h(\lambda)^2 d\lambda}_{C},$$

$$B = \mathbf{h}' \mathbf{m}_j, \ \ C = \mathbf{h}' \mathbf{H}_s \mathbf{h} \ \ ,$$

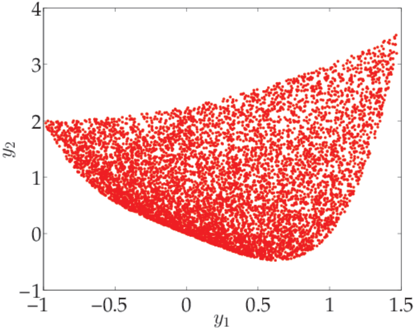
thus the problem can be reformulated as:

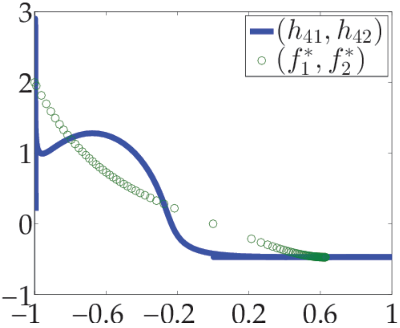
$$\min_{\mathbf{h}} \{ \mathbf{h}' \mathbf{H}_s \mathbf{h} - 2 \mathbf{h}' \mathbf{m}_j \}, \ \ j = 1, 2 \ \ .$$

□

Weighted convex sum approximation: method (a)

On the convex example:

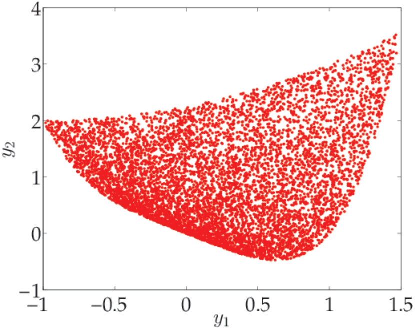


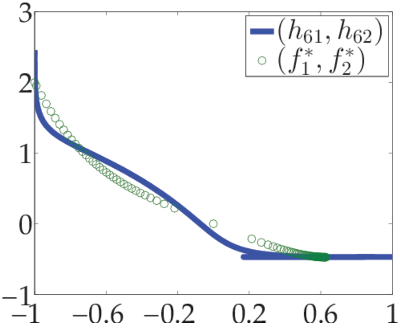


Degree 4

Weighted convex sum approximation: method (a)

On the convex example:

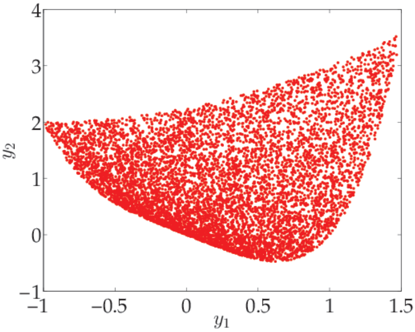


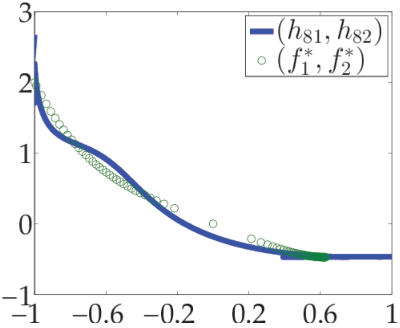


Degree 6

Weighted convex sum approximation: method (a)

On the convex example:

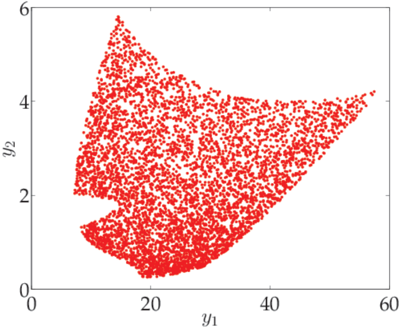


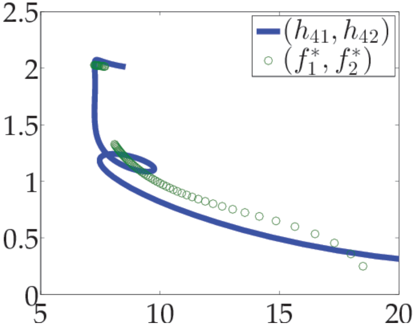


Degree 8

Weighted Chebyshev approximation: method (b)

On the non-convex example:

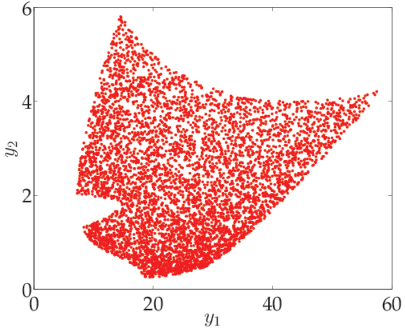


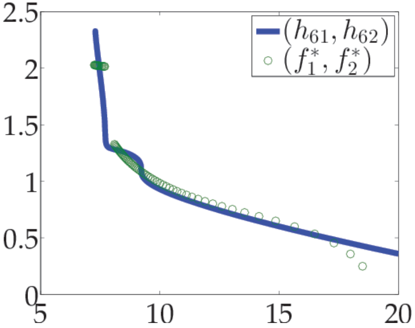


Degree 4

Weighted Chebyshev approximation: method (b)

On the non-convex example:

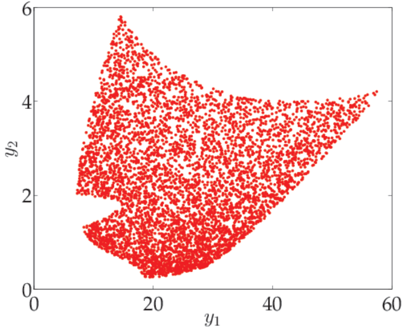


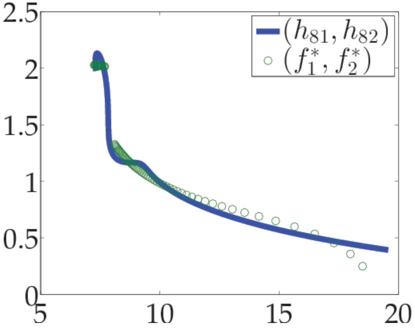


Degree 6

Weighted Chebyshev approximation: method (b)

On the non-convex example:





Degree 8

Parametric sublevel set approximation: method (c)

Scaling the problem:

$$\mathbf{K}^u := \{(\lambda, \mathbf{x}) \in [0,1] \times \mathbf{S} : (f_1(\mathbf{x}) - a_1)/(b_1 - a_1) \leqslant \lambda\},$$

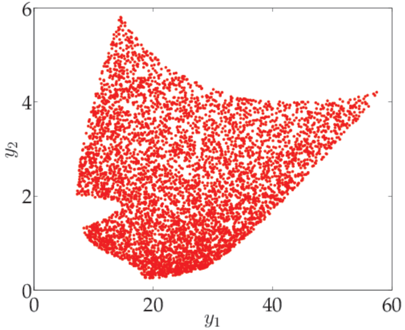
Parametric POP:

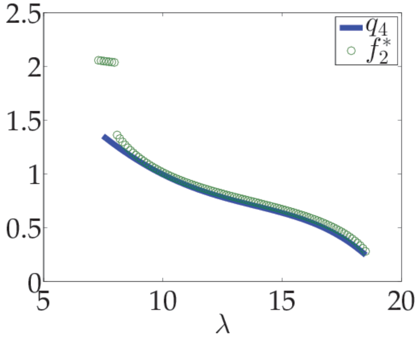
$$(\mathbf{P}^u_\lambda) : f^u(\lambda) = \min_{\mathbf{x} \in \mathbf{S}} \{f_2(\mathbf{x}) : (\lambda, \mathbf{x}) \in \mathbf{K}^u\}$$

Solving the dual SDP D_d yields underestimators for $\lambda \mapsto f^u(\lambda)$ over $[a_1, b_1]$. One can directly approximate the Pareto curve from below!

Parametric sublevel set approximation: method (c)

On the non-convex example:

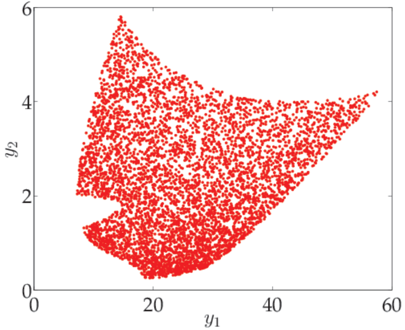


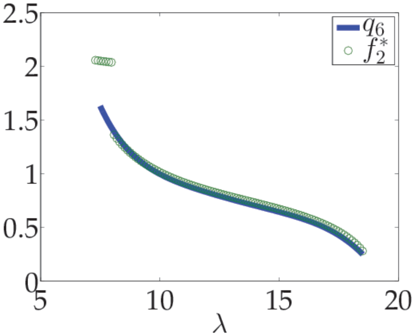


Degree 4

Parametric sublevel set approximation: method (c)

On the non-convex example:

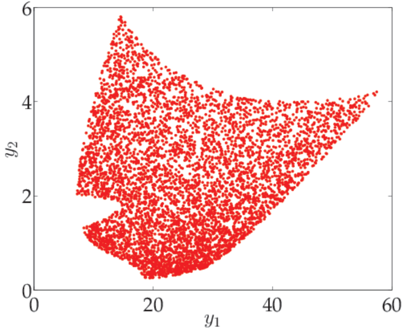


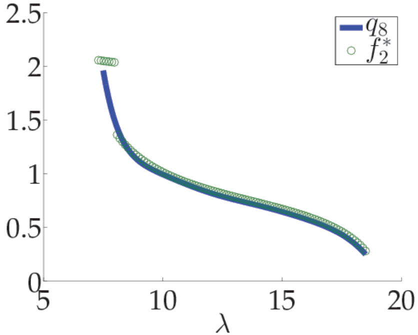


Degree 6

Parametric sublevel set approximation: method (c)

On the non-convex example:





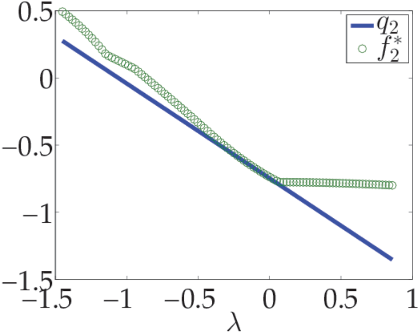
Degree 8

Parametric sublevel set approximation: method (c)

Medium size random bicriteria problem:

- $\mathbf{Q}_1, \mathbf{Q}_2 \in \mathbb{R}^{15\times 15}$, $\mathbf{q}_1, \mathbf{q}_2 \in \mathbb{R}^{15}$
- $\min_{\mathbf{x}\in[-1,1]^{15}} \{f_1(\mathbf{x}), f_2(\mathbf{x})\}$

  $f_j(\mathbf{x}) := \mathbf{x}^\top \mathbf{Q}_j \mathbf{x}/n^2 - \mathbf{q}_j^\top \mathbf{x}/n$



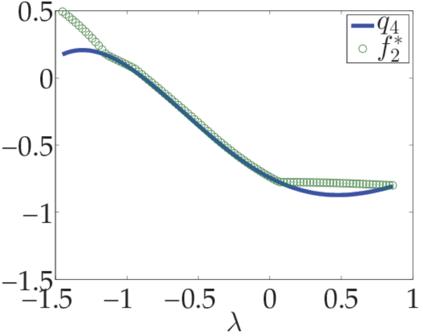
Degree 2

Parametric sublevel set approximation: method (c)

Medium size random bicriteria problem:

- $\mathbf{Q}_1, \mathbf{Q}_2 \in \mathbb{R}^{15\times 15}$, $\mathbf{q}_1, \mathbf{q}_2 \in \mathbb{R}^{15}$
- $\min_{\mathbf{x}\in[-1,1]^{15}} \{f_1(\mathbf{x}), f_2(\mathbf{x})\}$

  $f_j(\mathbf{x}) := \mathbf{x}^\top \mathbf{Q}_j \mathbf{x}/n^2 - \mathbf{q}_j^\top \mathbf{x}/n$



Degree 4

Outline

Approximation of sets defined with "∃"

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

- Another point of view:

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

  with

$$h(\mathbf{x}, \mathbf{y}) := (y_1 - f_1(\mathbf{x}))^2 + (y_2 - f_2(\mathbf{x}))^2 .$$

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

$$\Theta_d := \{\mathbf{y} \in \mathbf{B} : J_d(\mathbf{y}) \leqslant 0\} ,$$

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

Approximation of sets defined with "∃"

- Let $g_0 := 1$ and $\mathbf{Q}_d(\mathbf{S})$ be the d-truncated quadratic module generated by $g_0, \dots, g_m$:

$$\mathbf{Q}_d(\mathbf{S}) = \left\{ \sum_{l=0}^{m} \sigma_l(\mathbf{x}, \mathbf{y}) g_l(\mathbf{x}), \text{ with } \sigma_l \in \Sigma_{d-v_l}[\mathbf{x}, \mathbf{y}] \right\}$$

- Define $H(\mathbf{y}) := \min_{\mathbf{x} \in \mathbf{S}} h(\mathbf{x}, \mathbf{y})$
- Hierarchy of Semidefinite programs:

$$\rho_d := \min_{J \in \mathbb{R}_{2d}[\mathbf{y}], \sigma_l} \left\{ \int_{\mathbf{B}} (H - J) d\mathbf{y} : h - J \in \mathbf{Q}_d(\mathbf{S}) \right\} .$$

Yet another SOS program with an optimal solution $J_d \in \mathbb{R}_{2d}[\mathbf{y}]$!

A hierarchy of outer approximations of $f(\mathbf{S})$

From the definition of J_d, the sublevel sets

$$\Theta_d := \{\mathbf{y} \in \mathbf{B} : J_d(\mathbf{y}) \leqslant 0\} \supset f(\mathbf{S}), \;\; d \geqslant d_0 \;\; ,$$

provide a sequence of certified outer approximations of $f(\mathbf{S})$.

It comes from the following:

$$\forall (\mathbf{x}, \mathbf{y}) \in \mathbf{S} \times \mathbf{B}, J(\mathbf{y}) \leqslant h(\mathbf{x}, \mathbf{y}) \iff \forall \mathbf{y}, J(\mathbf{y}) \leqslant H(\mathbf{y}) \;\; .$$

Strong convergence property

Theorem

1. The sequence of underestimators $(J_d)_{d \geqslant d_0}$ converges to H w.r.t the $L_1(\mathbf{B})$-norm:
$$\lim_{d \to \infty} \int_{\mathbf{B}} |H - J_d| d\mathbf{y} = 0 \enspace .$$

2. $$\lim_{d \to \infty} V(\Theta_d \backslash f(\mathbf{S})) = 0 \enspace .$$

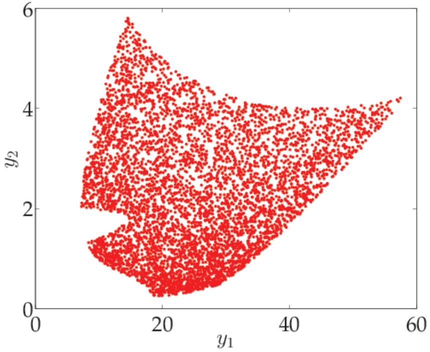
Strong convergence property

Theorem

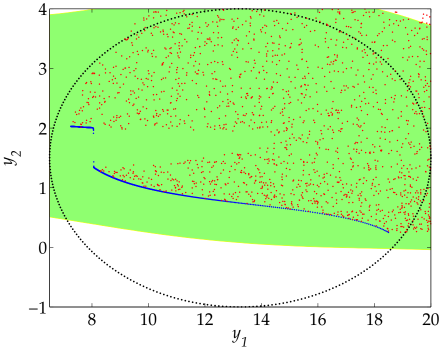
1. The sequence of underestimators $(J_d)_{d \geqslant d_0}$ converges to H w.r.t the $L_1(\mathbf{B})$-norm:
$$\lim_{d \to \infty} \int_{\mathbf{B}} |H - J_d| d\mathbf{y} = 0 \ .$$

2. $$\lim_{d \to \infty} V(\Theta_d \backslash f(\mathbf{S})) = 0 \ .$$

Back to the non-convex example

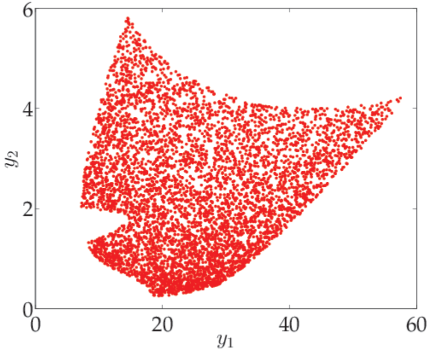


$f(\mathbf{S})$

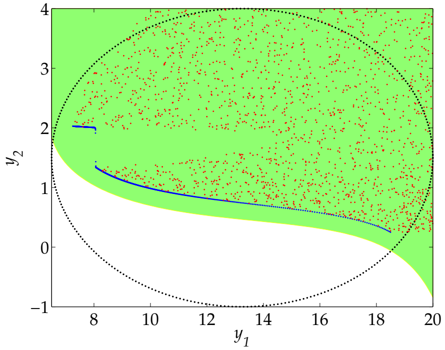


Θ_2

Back to the non-convex example

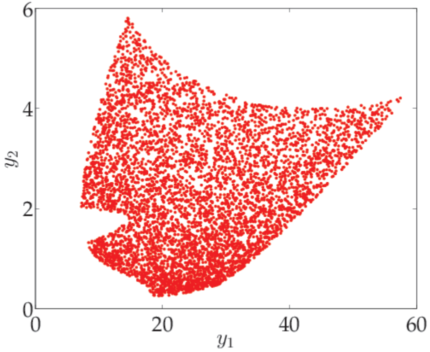


$f(\mathbf{S})$

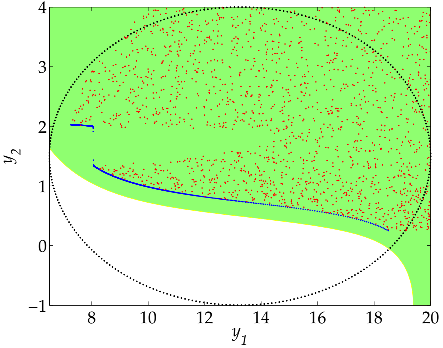


Θ_3

Back to the non-convex example

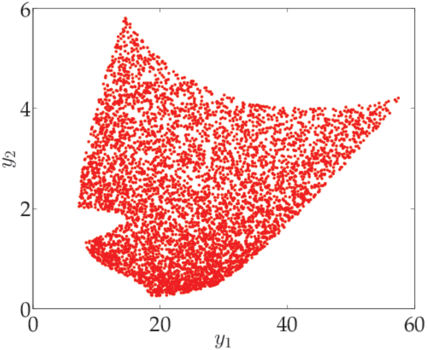


$f(\mathbf{S})$

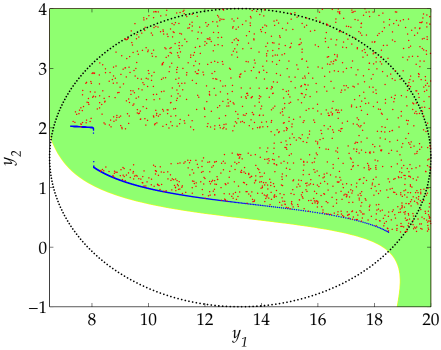


Θ_4

Back to the non-convex example

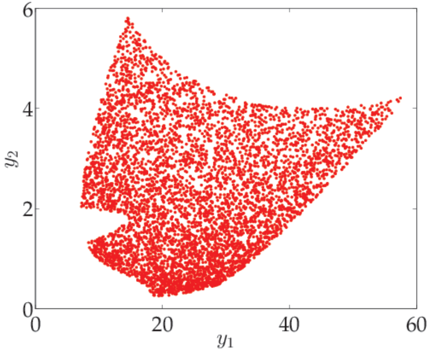


$f(\mathbf{S})$

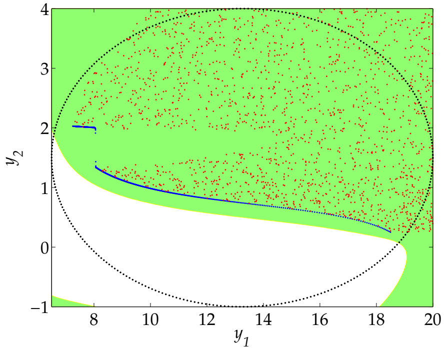


Θ_5

Back to the non-convex example

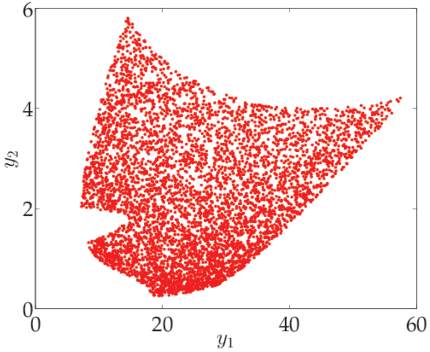


$f(\mathbf{S})$

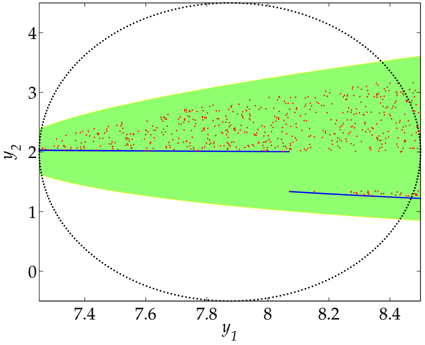


Θ_6

Branch and Bound: Zoom on the left

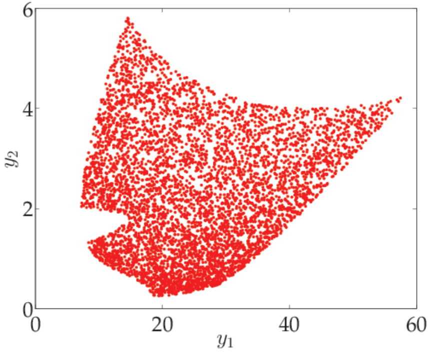


$f(\mathbf{S})$

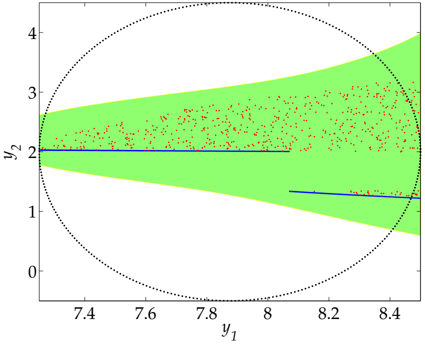


Θ_2

Branch and Bound: Zoom on the left

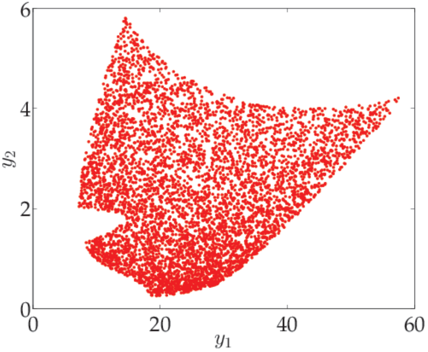


$f(\mathbf{S})$

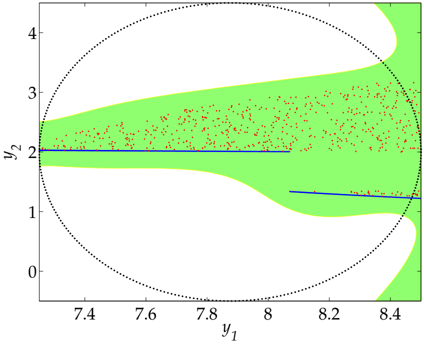


Θ_3

Branch and Bound: Zoom on the left

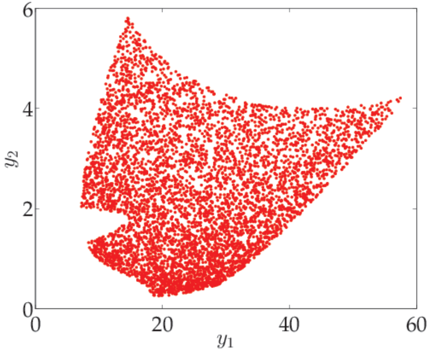


$f(\mathbf{S})$

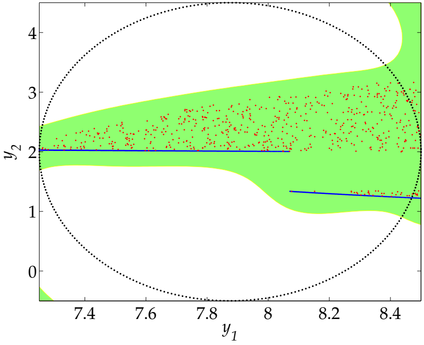


Θ_4

Branch and Bound: Zoom on the left

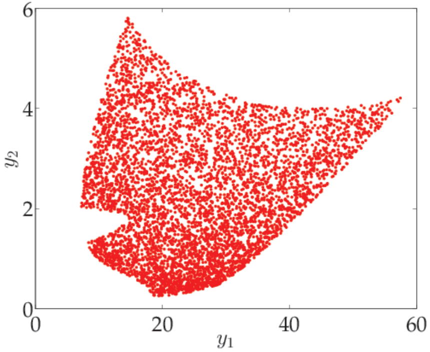


$f(\mathbf{S})$

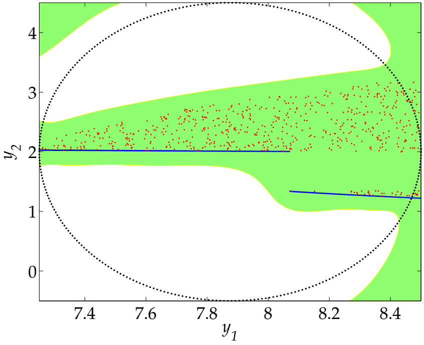


Θ_5

Branch and Bound: Zoom on the left



$f(\mathbf{S})$



Θ_6

Outline

Transcendental conflicting criteria

Now, consider the following Problem:

$$(\mathbf{P}) \left\{ \min_{\mathbf{x} \in \mathbf{S}} \quad (f_1(\mathbf{x})\, f_2(\mathbf{x}))^\top \quad . \right\}$$

with transcendental criteria f_1, f_2.

- Generalization of the single criterion problem $\min_{\mathbf{x} \in \mathbf{S}} f(\mathbf{x})$
- Hard to combine SOS hierarchies with Taylor/Chebyshev approximations [2]

[2] X. Allamigeon, S. Gaubert, V. Magron and B. Werner. *Certification of inequalities involving transcendental functions: combining SDP and max-plus approximation* (2013)

Transcendental conflicting criteria

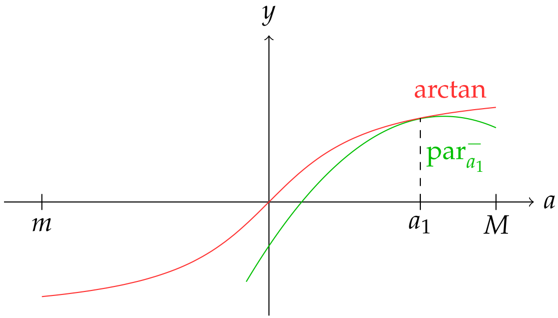
Definition: Semiconvex function

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.

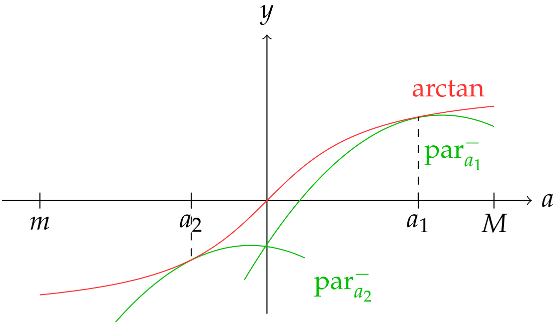
Proposition (by Legendre-Fenchel duality)

The set of functions $f : \mathbb{R}^n \to \mathbb{R}$ which can be written as the max-plus linear combination $f = \sup_{w \in \mathcal{B}}(a(w) + w)$ for some function $a : \mathcal{B} \to \mathbb{R} \cup \{-\infty\}$ is precisely the set of lower semicontinuous γ-semiconvex functions.

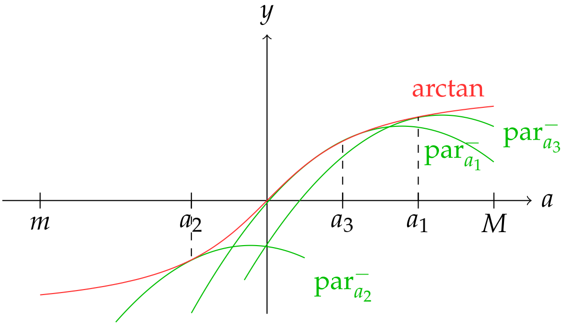
Transcendental conflicting criteria



Transcendental conflicting criteria



Transcendental conflicting criteria



Sublevel sets of semialgebraic underestimators

The sublevel sets

$$\Theta_d := \{\mathbf{y} \in \mathbf{B} : J_d(\mathbf{y}) \leqslant 0\} \supset f(\mathbf{S}), \;\; d \geqslant d_0 \;\; ,$$

provide a sequence of certified outer approximations of $f(\mathbf{S})$.

To avoid Branch and bound iterations (“Zooms”), one could underestimate H with a rational function

$$J := F/(1+\sigma) \;\; ,$$

with $F \in \mathbb{R}_{2d}[\mathbf{y}], \sigma \in \Sigma_{d_0}[\mathbf{y}]$.

Thank you for your attention!

Victor Magron, Didier Henrion, Jean-Bernard Lasserre. *Approximating Pareto Curves using Semidefinite Relaxations*. arxiv:1404.4772, 2014.

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