Recent Results on Optimal μ-Distributions for the Hypervolume Indicator

Dimo Brockhoff

[joint work with Anne Auger, Johannes Bader, and Eckart Zitzler]

Schloss Dagstuhl, Wadern, September 6, 2010

INSTITUT NATIONAL
DE RECHERCHE
EN INFORMATIQUE
ET EN AUTOMATIQUE

INRIA

Multiobjective Optimization

Almost all problems are multiobjective in nature…

Scenario

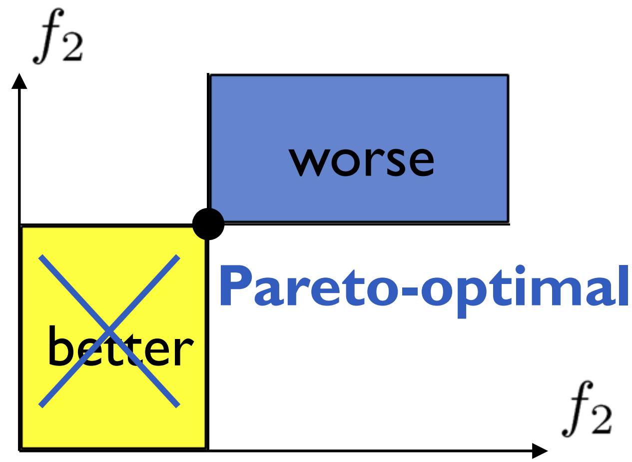
- Minimize >1 objective functions simultaneously

$$\min_{x \in X} \mathcal{F}(x) = (f_1(x), \ldots, f_k(x))$$

where $x \in X \to \mathcal{F}(x) \in \mathbb{R}^k$

- Pareto dominance

$x \preceq y$ iff $\forall 1 \leq i \leq k : f_i(x) \leq f_i(y)$



- Pareto set/Pareto front
- $\Rightarrow$ set problem: generalize Pareto dominance on sets

Multiobjective Optimization

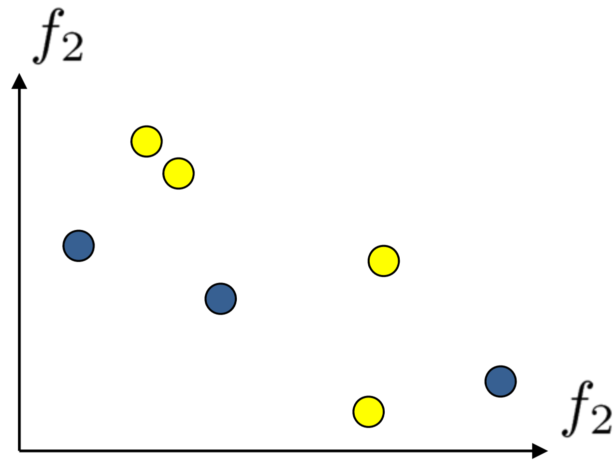
Almost all problems are multiobjective in nature…

Scenario

- Minimize >1 objective functions simultaneously

$$\min_{x \in X} f(x) = (f_1(x), \ldots, f_k(x))$$

$$\text{where } x \in X \to f(x) \in \mathbb{R}^k$$



- Pareto dominance

$$x \preceq y \text{ iff } \forall 1 \leq i \leq k : f_i(x) \leq f_i(y)$$

- Pareto set/Pareto front
- ⇒ set problem: generalize Pareto dominance on sets

$$A \preceq B \text{ iff } \forall b \in B : \exists a \in A : a \preceq b$$

Unary Quality Indicators

Take set A ⇒ assign real value I(A) ∈ $\mathbb{R}$

Why?

Performance assessment (provides a total order on sets via ≥)

Nowadays also used in selection of EAs (explicit preference articulation)

Result

Transforms multiobjective problem to single-objective one: max. I(A)
s.t. |A| ≤ μ

! Not any indicator interesting: we need also a refinement

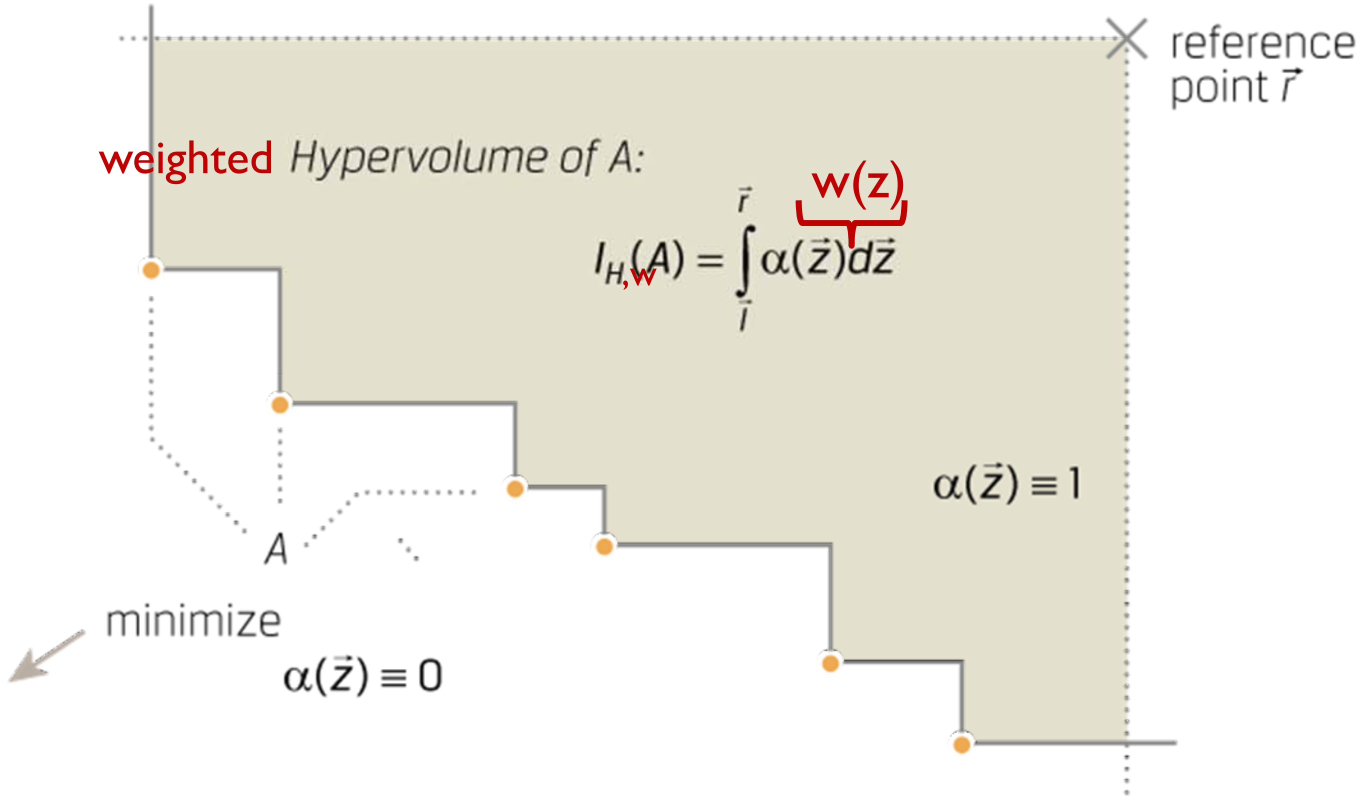
Definition 2.4: Given a set Ψ. Then the preference relation $\preccurlyeq_{\text{ref}}$ refines $\preccurlyeq$ if for all $A, B \in \Psi$ we have

$$(A \preccurlyeq B) \wedge (B \not\preccurlyeq A) \Rightarrow (A \preccurlyeq_{\text{ref}} B) \wedge (B \not\preccurlyeq_{\text{ref}} A).$$

from [Zitzler et al. in IEEETEC'10]

The (Weighted) Hypervolume Indicator

The only unary indicators that are refinements are the (weighted) hypervolume indicators



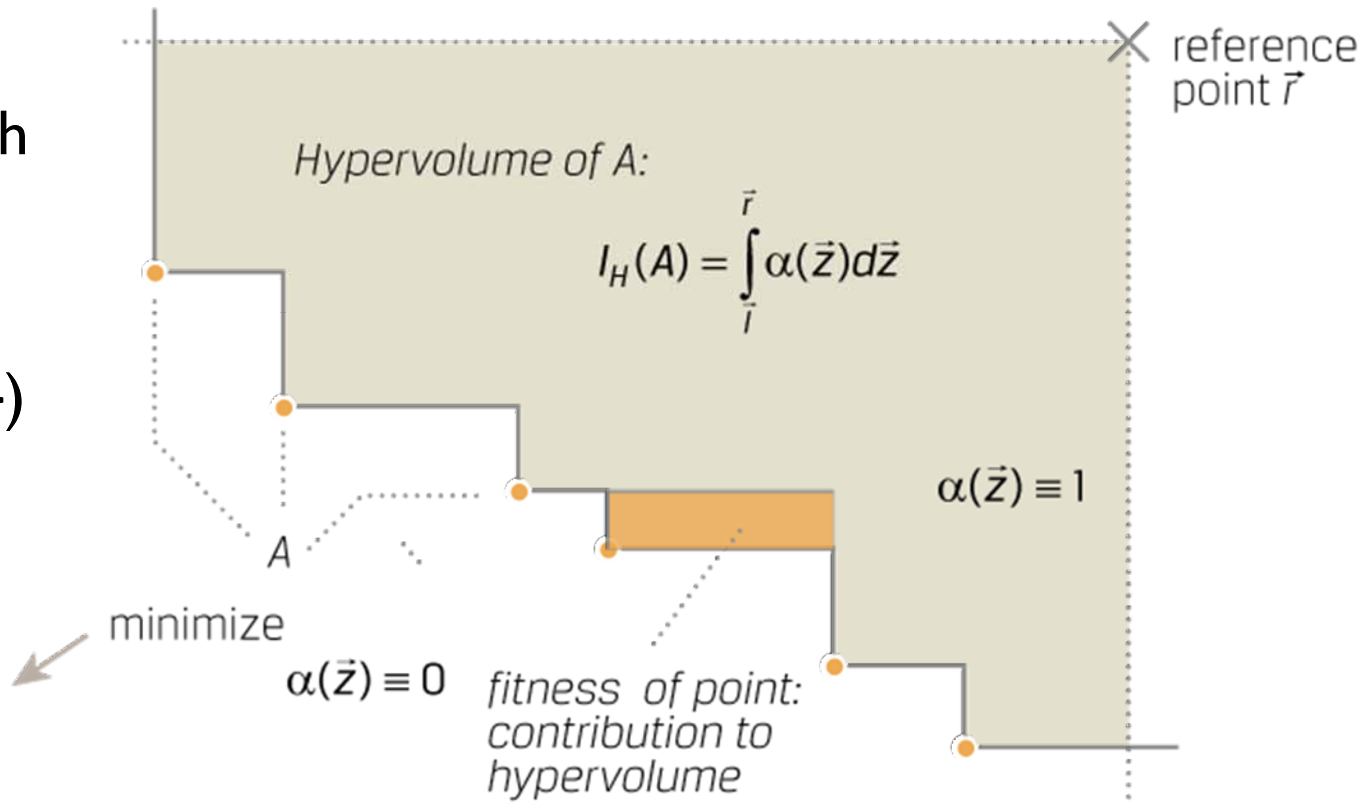
Hypervolume-Based Evolutionary Algorithms

State-of-the-art algorithms (SMS-EMOA, MO-CMA-ES, HypE, …)

use hypervolume indicator as 2nd selection criterion: refinement!

Main idea

Delete solutions with the smallest hypervolume loss $d(s) = I_H(P)-I_H(P / \{s\})$ iteratively



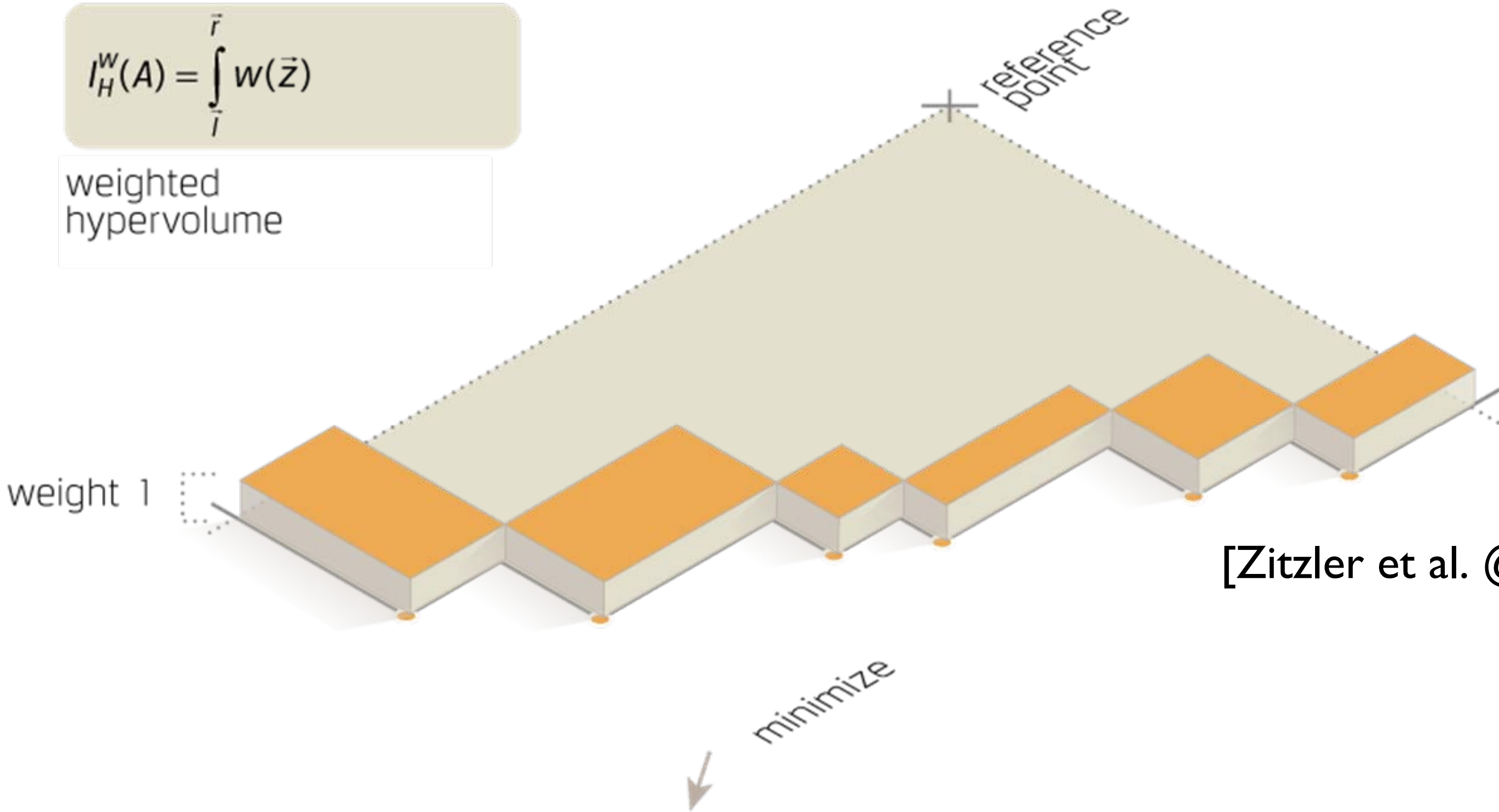
Weighted Hypervolume Selection

$$I_H(A) = \int_{\vec{l}}^{\vec{r}}$$

normal hypervolume

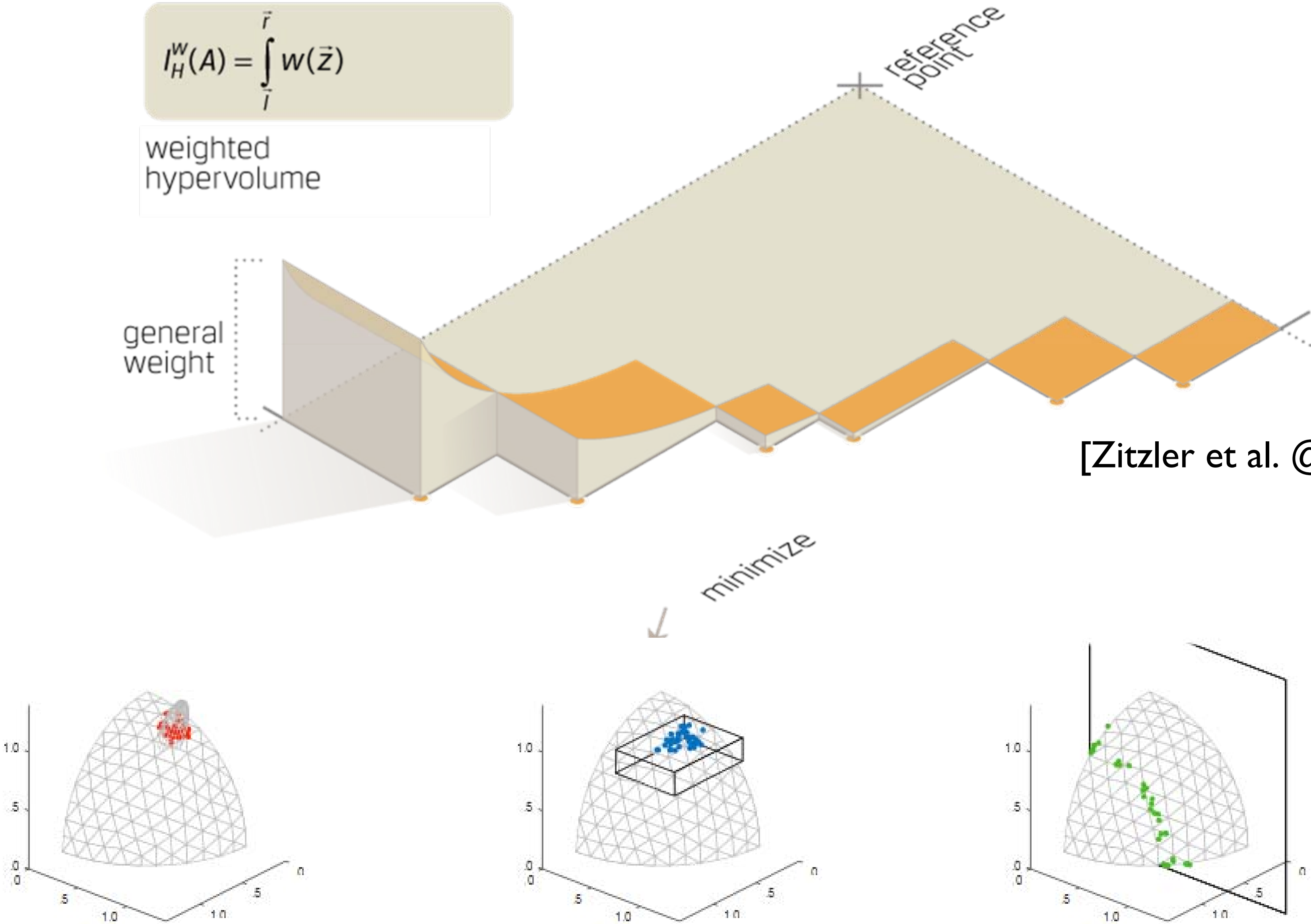
[Zitzler et al. @ EMO'07]

Weighted Hypervolume Selection



[Zitzler et al. @ EMO'07]

Weighted Hypervolume Selection



[Zitzler et al. @ EMO'07]

Optimal μ-Distributions

When the goal is to maximize the hypervolume…

- this yields sets with only Pareto-optimal solutions [Fleischer @ EMO'03]
- those sets, if unrestricted in size, cover the entire Pareto front
- many hypervolume-based evolutionary algorithms have a population size μ

Optimal μ-Distribution:

A set of μ solutions that maximizes the hypervolume indicator among all sets of μ solutions is called optimal μ-distribution.

Questions:

- How distributed? ⇒ performance assessment
- Do algorithms converge to it?
- And even before: do optimal μ-distributions exist?

Overview

Today:

❶ New existence results [submitted to TCS]

❷ An exact and exhaustive result for linear bi-objective fronts [Brockhoff @ SEAL'10]

Overview

Today:

❶ New existence results [submitted to TCS]

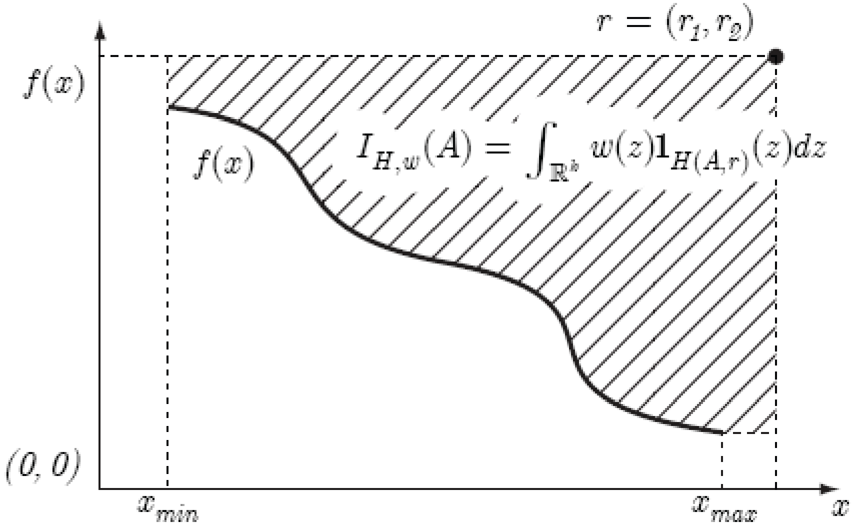
❷ An exact and exhaustive result for linear bi-objective fronts [Brockhoff @ SEAL'10]

Existence of Optimal μ-Distributions

Scenario

f: $[x_{min}, x_{max}] \rightarrow \mathbb{R}$ strictly monotone

reference point $r = (r_1, r_2)$



Known results on Existence

- if f continuous [Auger et al. @ FOGA'09]
- if f upper semi-continuous for maximization

[Bringmann and Friedrich @ GECCO'10]

Upper and Lower Semi-Continuity

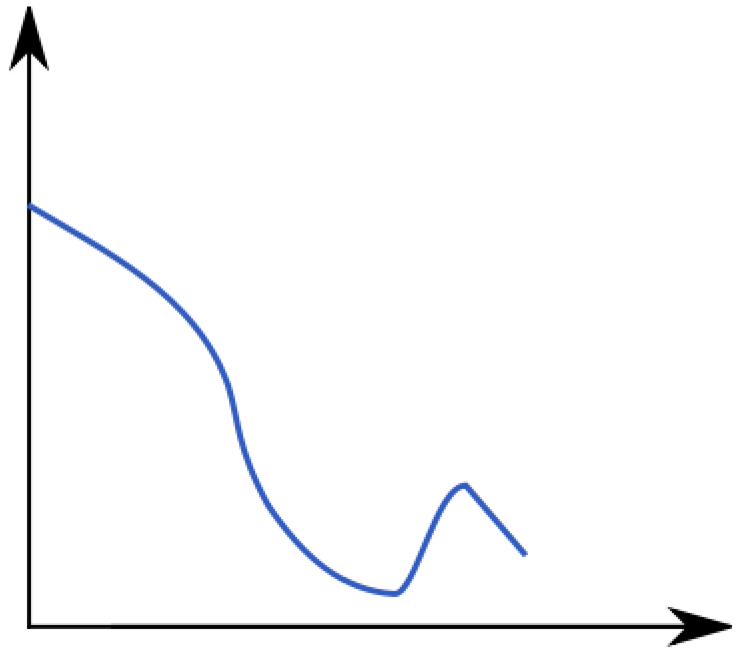
A function f is upper (lower) semi-continuous if for all x_0

$$\limsup_{x \rightarrow x_0} f(x) \leq f(x_0) \quad \left(\liminf_{x \rightarrow x_0} f(x) \geq f(x_0) \right)$$

holds.

Upper and Lower Semi-Continuity

Definition not overly intuitive at first sight: $\forall x_0$: $\liminf_{x \to x_0} f(x) \geq f(x_0)$

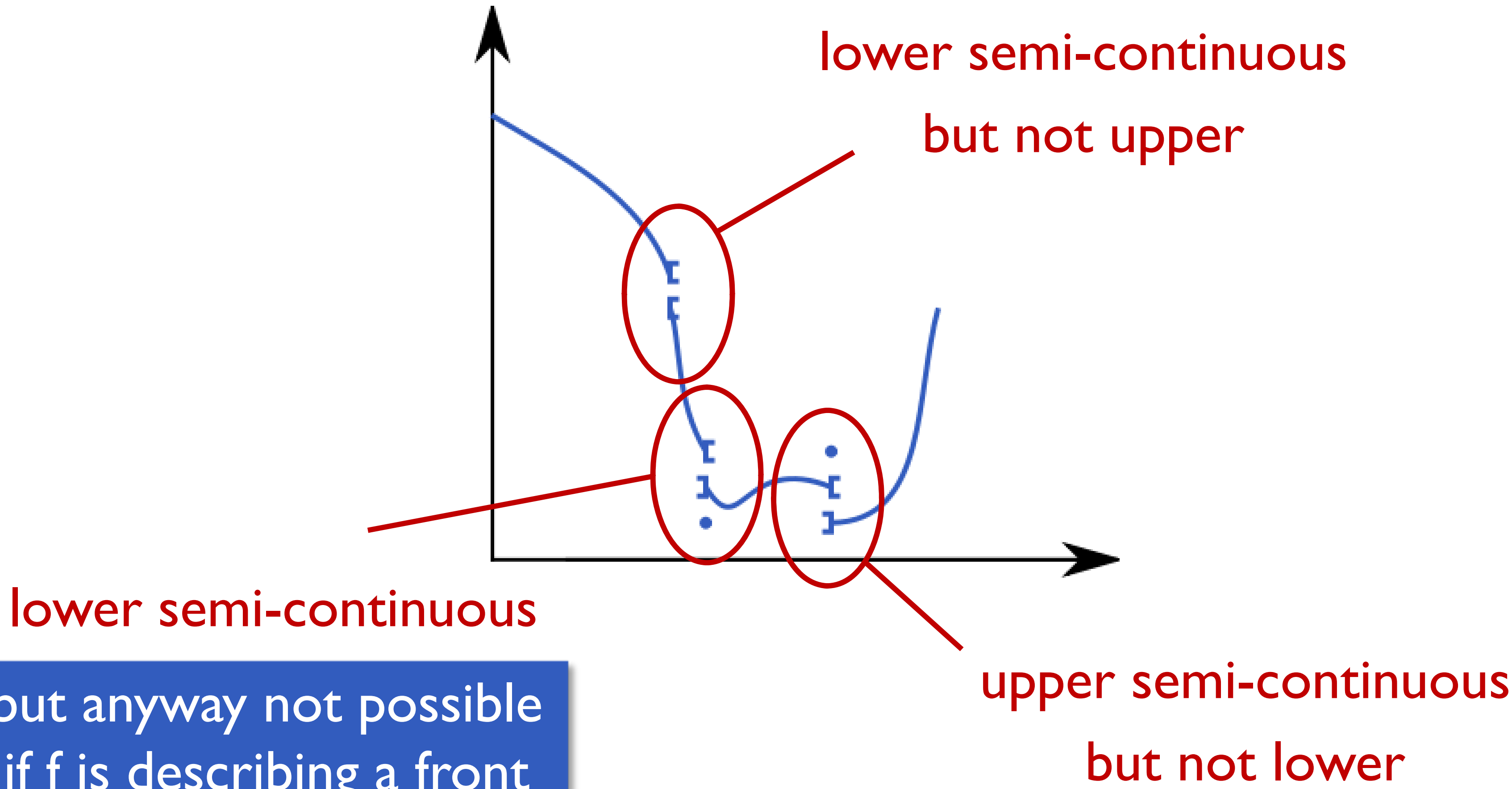


f continuous

$\Rightarrow$ f upper and lower semi-continuous

Upper and Lower Semi-Continuity

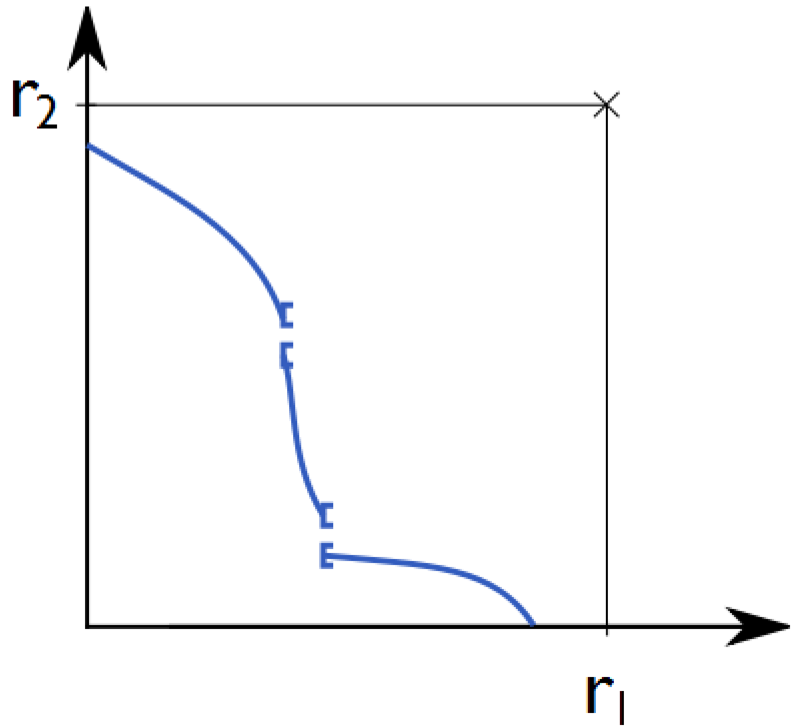
Definition not overly intuitive at first sight: $\forall x_0$: $\liminf_{x \to x_0} f(x) \geq f(x_0)$



but anyway not possible if f is describing a front

Upper and Lower Semi-Continuity

Definition not overly intuitive at first sight: $\forall x_0$: $\liminf_{x \to x_0} f(x) \geq f(x_0)$



if f is describing a front:

lower semi-continuity $\Leftrightarrow$ continuous from the right

Interesting Questions

- Why upper semi-continuous? What about minimization?
- Existence for the weighted hypervolume indicator?
- Sufficent and/or necessary criteria?

Today:

- Lower semi-continuity for minimization is a sufficient criterion in the weighted hypervolume case
- There are fronts for which no optimal 1-distribution exists
- Lower semi-continuity is not a necessary condition for existence

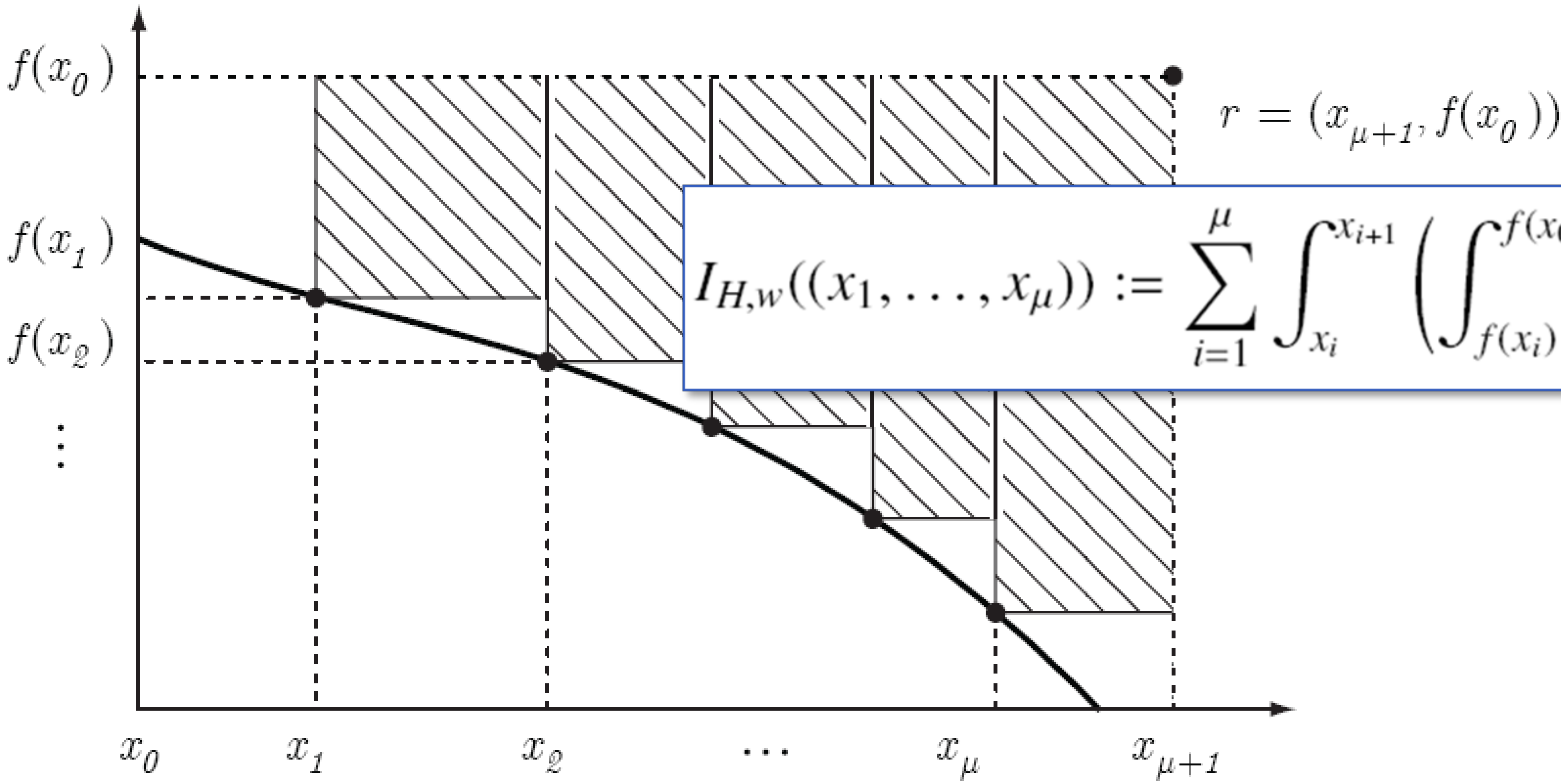
Again: The Scenario

Given

f: $[x_{min}, x_{max}] \rightarrow \mathbb{R}$ strictly monotone

reference point $r=(r_1,r_2)$

μ x-values enough to characterize optimal μ -distribution



$$I_{H,w}((x_1,\dots,x_\mu)) := \sum_{i=1}^{\mu} \int_{x_i}^{x_{i+1}} \left(\int_{f(x_i)}^{f(x_0)} w(x,y)dy \right) dx$$

General Result on Existence for Weighted Case

Theorem Let μ ∈ ℕ. If f is lower semi-continuous, there exists (at least) one set of μ points maximizing the hypervolume.

Proof. We are going to prove that $I_{H,w}$ is upper semi-continuous if f is lower semi-continuous, and then apply the $\boxed{\text{Extreme Value Theorem.}}$ Since $I_{H,w}$ is the sum of μ functions $g(x_i, x_{i+1})$ where $g(\alpha,\beta) = \int_\alpha^\beta \left(\int_{f(\alpha)}^{f(x_0)} w(x,y)dy \right) dx$, we will prove the upper semi-continuity of $g(x_i, x_{i+1})$ for $(x_i, x_{i+1}) \in [x_{\min}, x_{\max}]$. This will imply the upper semi-continuity of $I_{H,w}$ (Bourbaki, 1989, p 362). Let $(x_i, x_{i+1}) \in [x_{\min}, x_{\max}]$ and let $(x_i^n, x_{i+1}^n)_{n\in\mathbb{N}}$ converging to (x_i, x_{i+1}). We will now prove that $\limsup g(x_i^n, x_{i+1}^n) \le g(x_i, x_{i+1})$ (see Knapp, 2005, p 481). Since

$$\limsup_{n\to\infty} g(x_i^n, x_{i+1}^n) = \limsup_{n\to\infty} \int\int \mathbf{1}_{[x_i^n, x_{i+1}^n]}(x)\mathbf{1}_{[f(x_i^n), f(x_0)]}(y) w(x,y)dydx \ ,$$

and $\mathbf{1}_{[x_i^n, x_{i+1}^n]}(x)\mathbf{1}_{[f(x_i), f(x_0)]}(x)w(x,y) \le \mathbf{1}_{[x_{\min}, x_{\max}]}(x)\mathbf{1}_{[f(x_{\max}), f(x_0)]}(x)w(x,y)$ we can use the (Reverse) Fatou Lemma (Knapp, 2005, p 252) that implies $\limsup g(x_i^n, x_{i+1}^n) \le \int\int \limsup \mathbf{1}_{[x_i^n, x_{i+1}^n]}(x)\mathbf{1}_{[f(x_i^n), f(x_0)]}(y)w(x,y)dydx$. Since f is lower semi-continuous, $\liminf f(x_i^n) \ge f(x_i)$ holds which is equivalent to $\limsup(f(x_0) - f(x_i^n)) = f(x_0) - \liminf f(x_i^n) \le f(x_0) - f(x_i)$. Hence, $\limsup \mathbf{1}_{[f(x_i^n), f(x_0)]}(y) \le \mathbf{1}_{[f(x_i), f(x_0)]}(y)$ and thus

$$\limsup_{n\to\infty} g(x_i^n, x_{i+1}^n) \le \int\int \mathbf{1}_{[x_i, x_{i+1}]}(x)\mathbf{1}_{[f(x_i), f(x_0)]}(y)w(x,y)dydx = g(x_i, x_{i+1}) \ .$$

We have proven the upper semi-continuity of g which implies the upper semi-continuity of $I_{H,w}$. Since in addition $I_{H,w}$ is upper bounded by the hypervolume contribution of the entire front which is finite, we can imply from the Extreme Value Theorem that there exists a set of μ points maximizing the hypervolume indicator. □

General Result on Existence for Weighted Case

Theorem Let $\mu \in \mathbb{N}$. If f is lower semi-continuous, there exists (at least) one set of μ points maximizing the hypervolume.

Sketch of Proof

Extreme Value Theorem / Weierstrass Theorem:

If $G: K \rightarrow \mathbb{R}$ is upper semi continuous and K compact then

$\exists\, x^* \in K$ s.t. $\forall\, x \in K$: $G(x) \leq G(x^*)$

Here: $G = I_{H,w}(A,r)$

- ✓ $K=[x_{min},x_{max}]^\mu$ compact
- ✓ f lower semi-continuous $\Rightarrow$ $I_{H,w}$ upper semi-continuous (technical: tools from Lebesgue integration theory, in particular "Fatou Lemma")

Are There Fronts Without Existence?

Shown:

f lower semi-continuous $\Rightarrow$ existence (sufficient criterion)

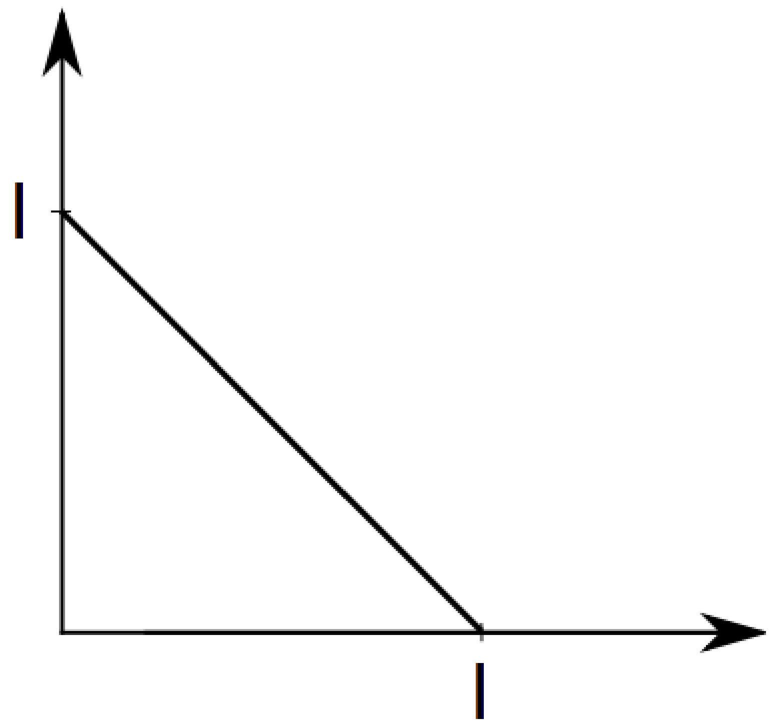
Still unclear:

does an example with no existence really exist?

Are There Fronts Without Existence?

Theorem There are problems with no optimal μ-distributions.

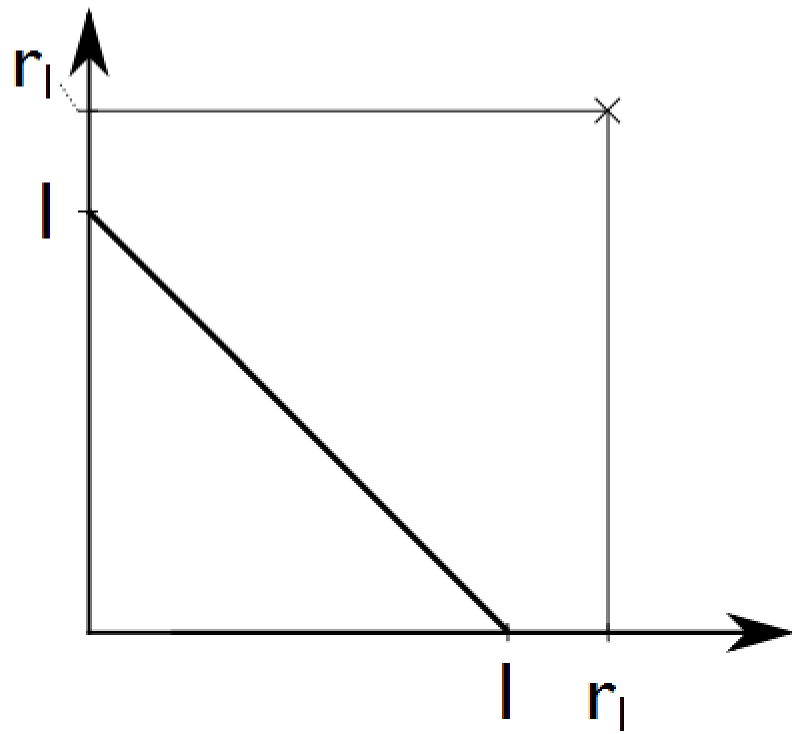
Proof



Are There Fronts Without Existence?

Theorem There are problems with no optimal μ-distributions.

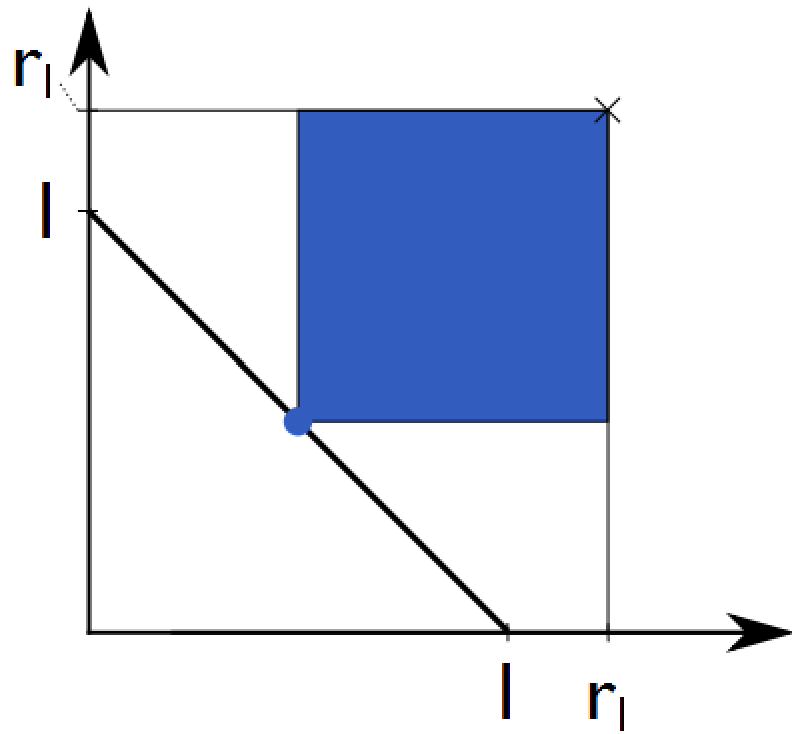
Proof



Are There Fronts Without Existence?

Theorem There are problems with no optimal μ-distributions.

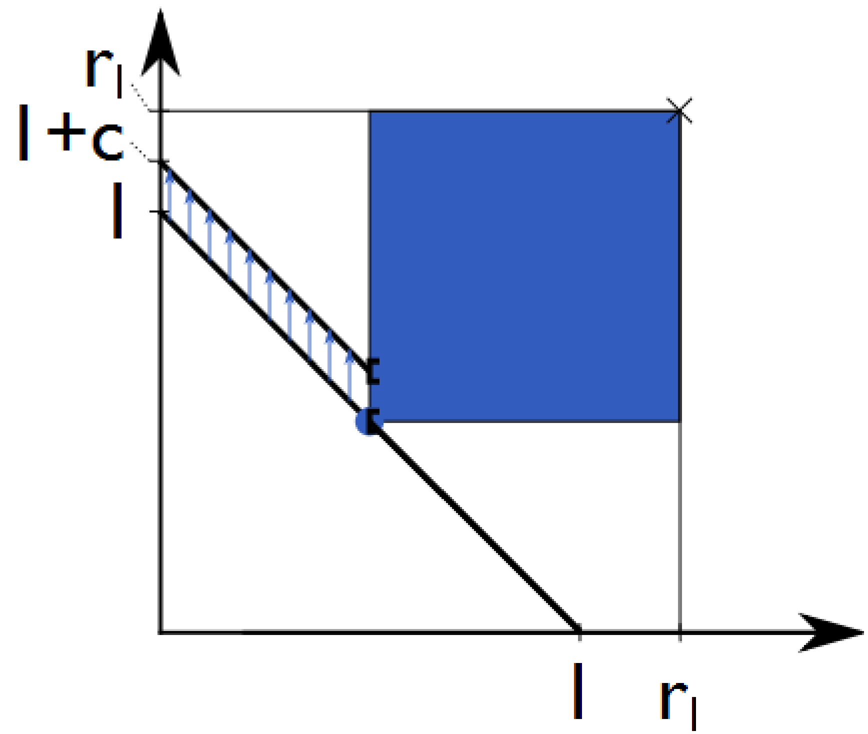
Proof



Are There Fronts Without Existence?

Theorem There are problems with no optimal μ-distributions.

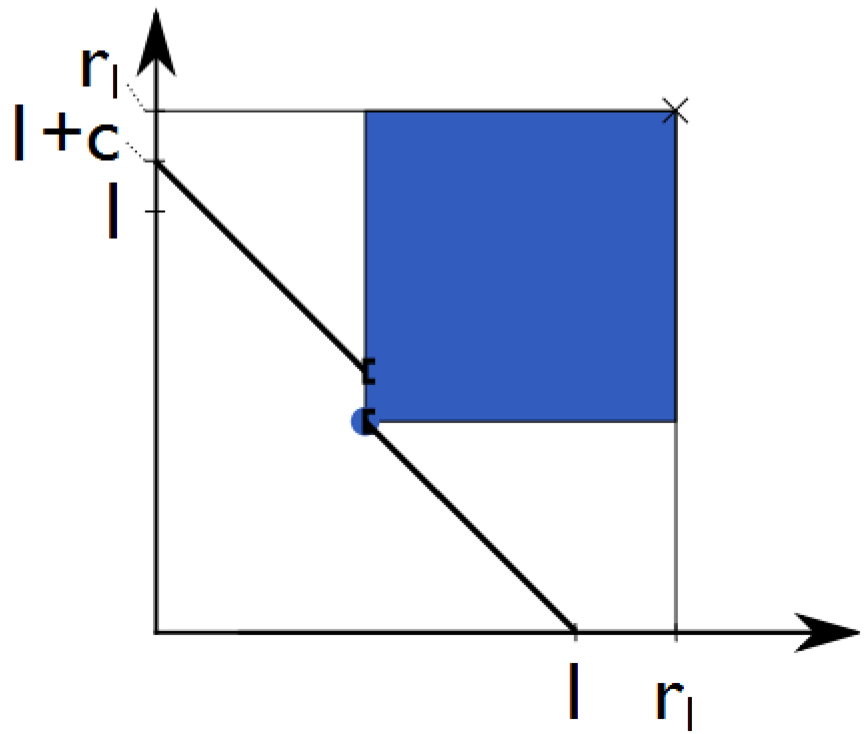
Proof



Are There Fronts Without Existence?

Theorem There are problems with no optimal μ-distributions.

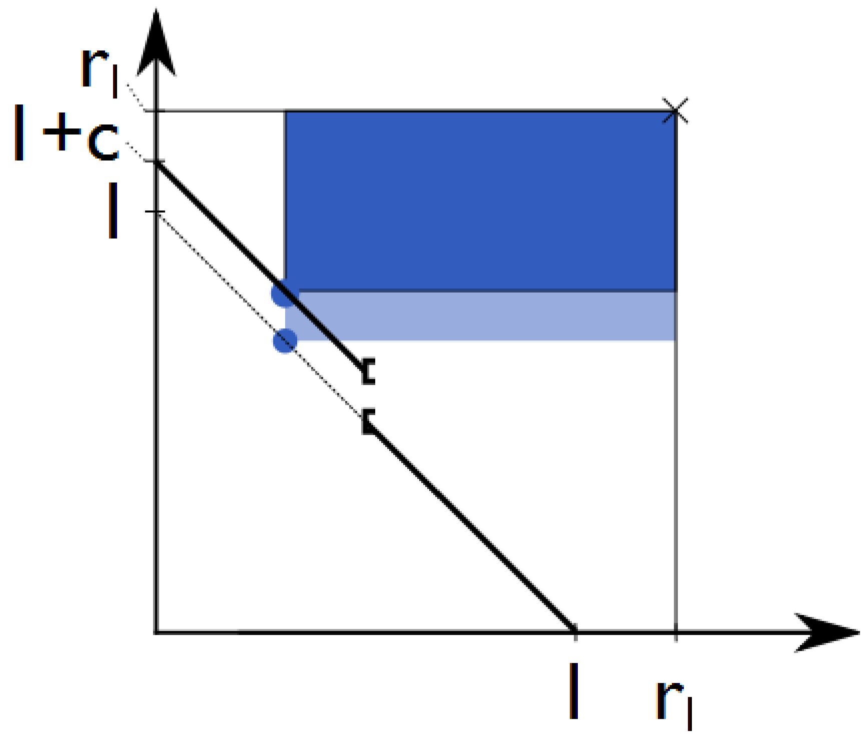
Proof



Are There Fronts Without Existence?

Theorem There are problems with no optimal μ-distributions.

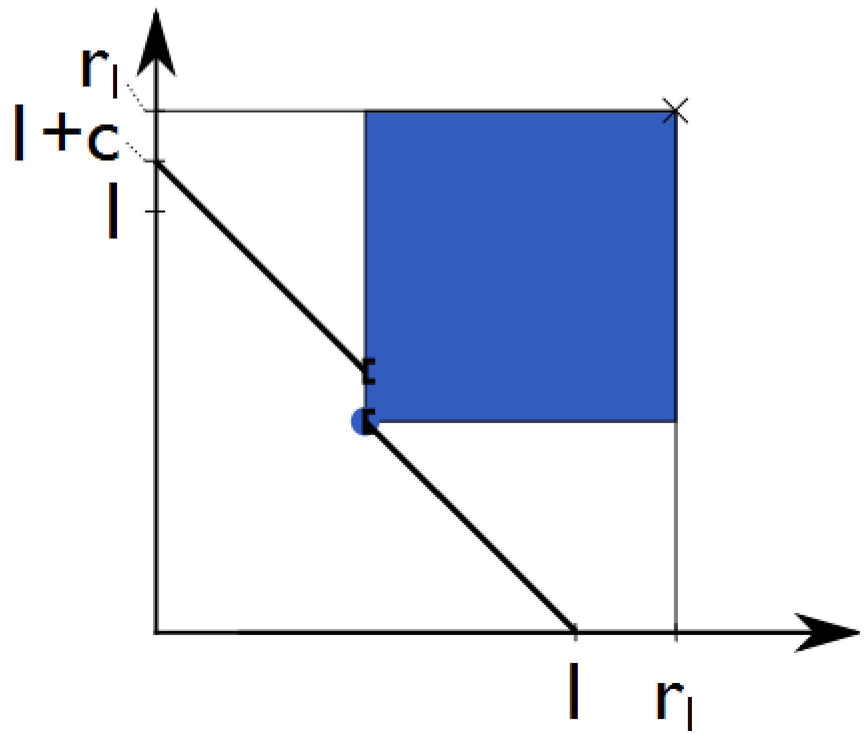
Proof



Are There Fronts Without Existence?

Theorem There are problems with no optimal μ-distributions.

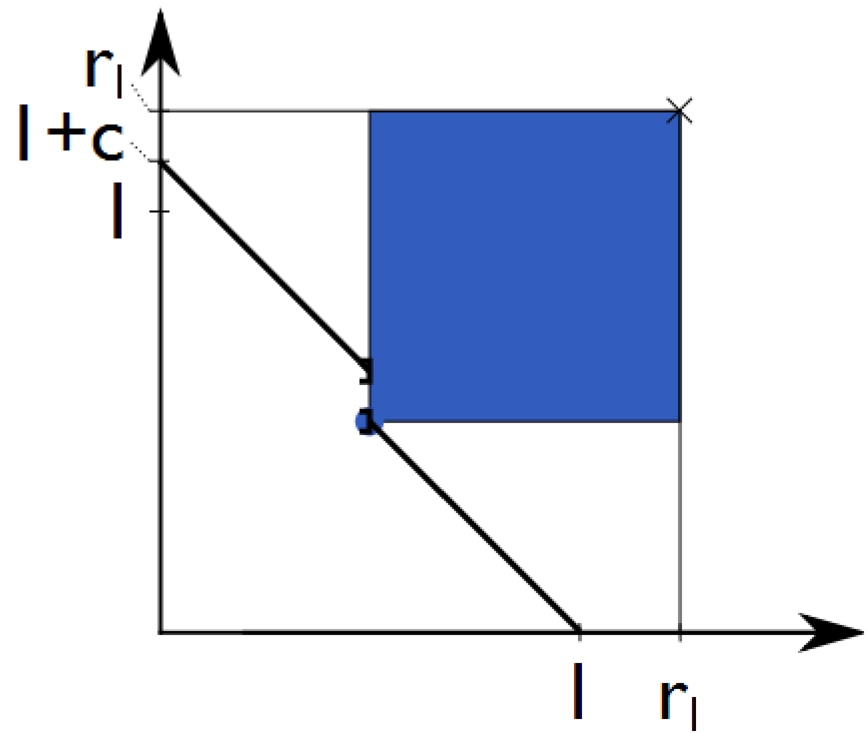
Proof



Are There Fronts Without Existence?

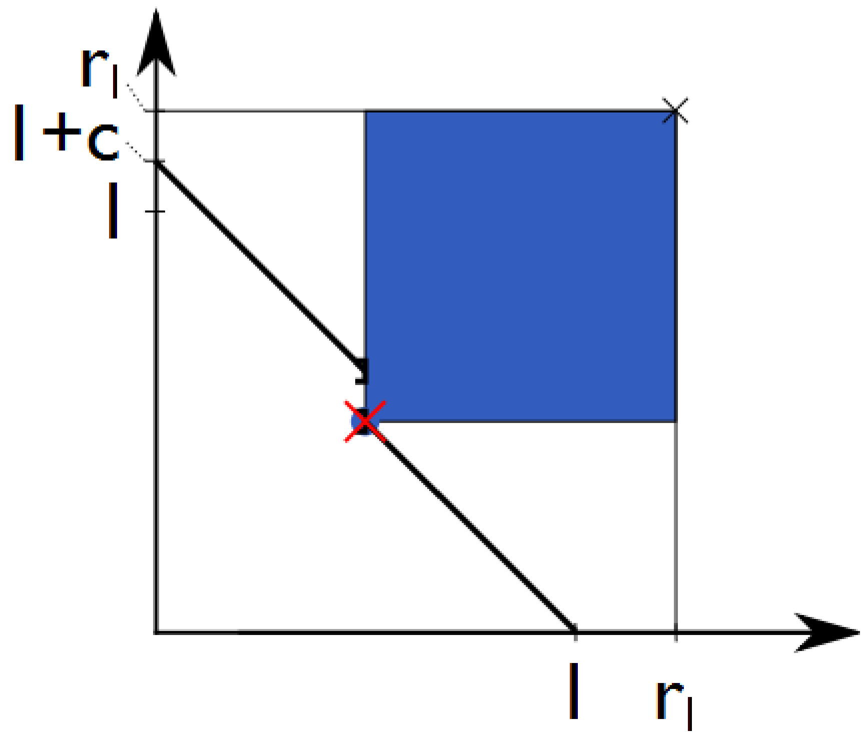
Theorem There are problems with no optimal μ-distributions.

Proof



Theorem There are problems with no optimal μ-distributions.

Proof



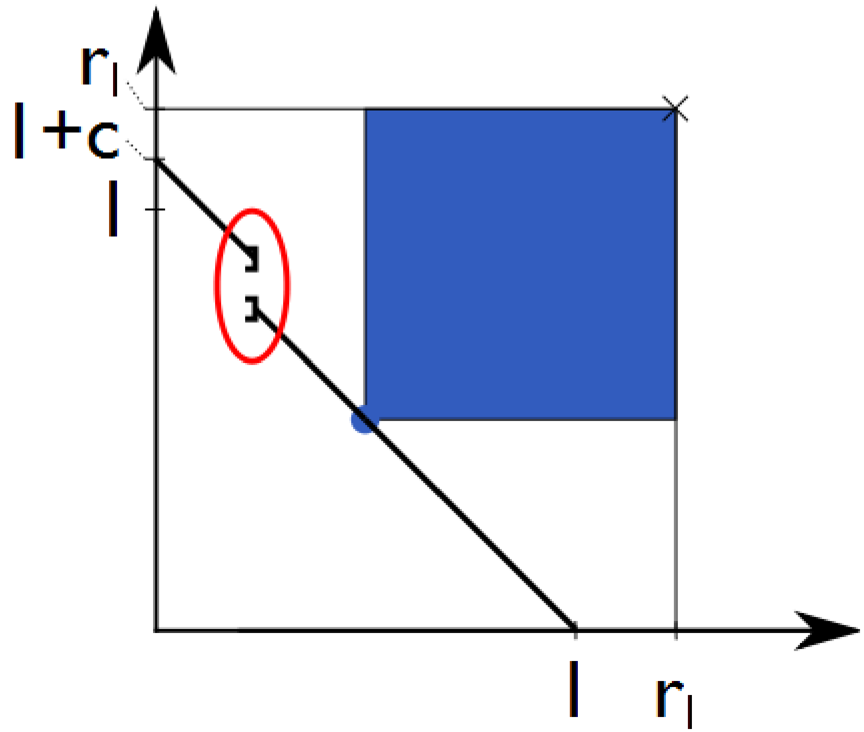
⇒ no optimal 1-distribution exists!

A Necessary Condition for Existence?

Theorem Lower semi-continuity is not a necessary criterion for existence.

Proof

To show: f not lower semi-continuous but optimal μ-distribution exists



Open Questions…

…to be discussed here at Dagstuhl (or later)

- Existence results for higher dimensional optimal μ-distributions?
  - if continuous, no problem but is simply lower/upper semi-continuity enough?
  - are there differences to the 2-objective case?
- Are there criteria for the uniqueness?
  - cf. results of Beume et al. [Beume et al. @ EMO'09]
- Constructive results: how do optimal μ-distributions look like?

Overview

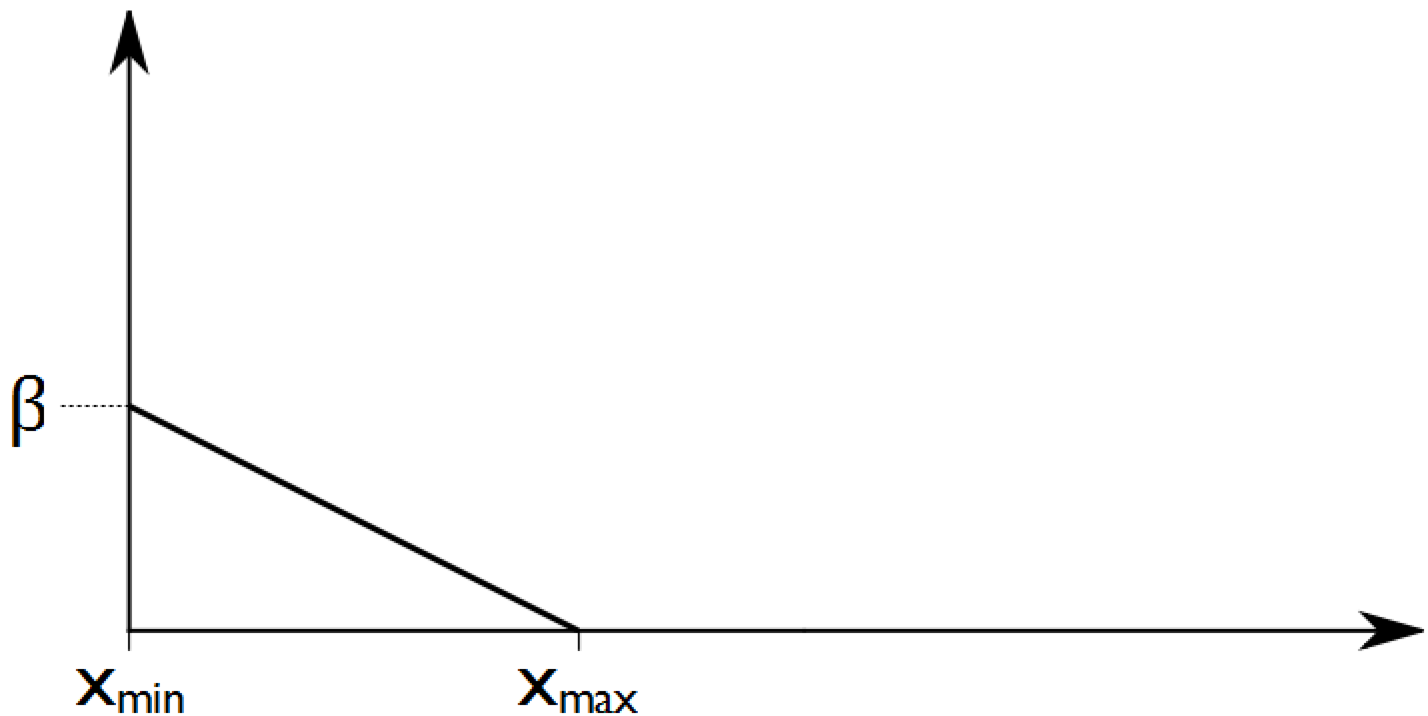
❶ New existence results [submitted to TCS]

❷ An exact and exhaustive result for linear bi-objective fronts

[Brockhoff @ SEAL'10]

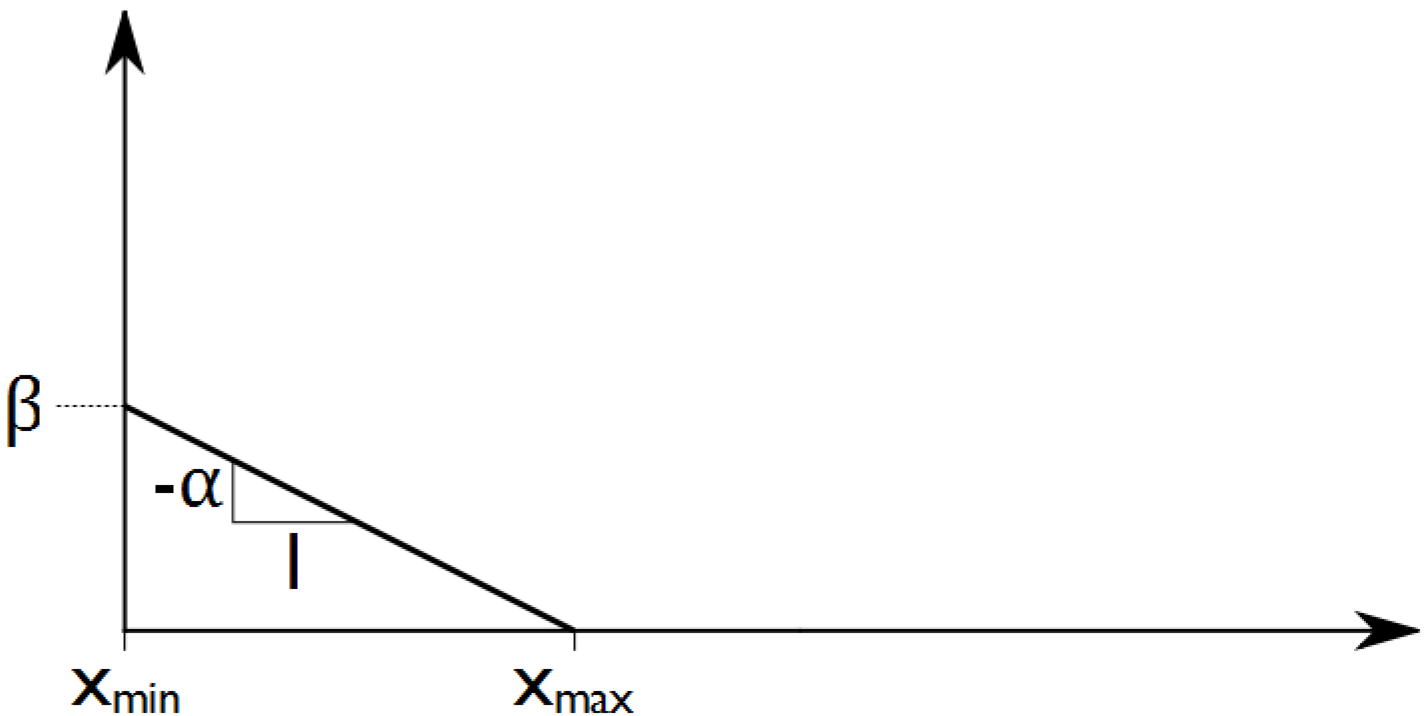
Starting Point

Assume problem with linear bi-objective front: $f(x)=\alpha x+\beta$ in $[x_{min}, x_{max}]$



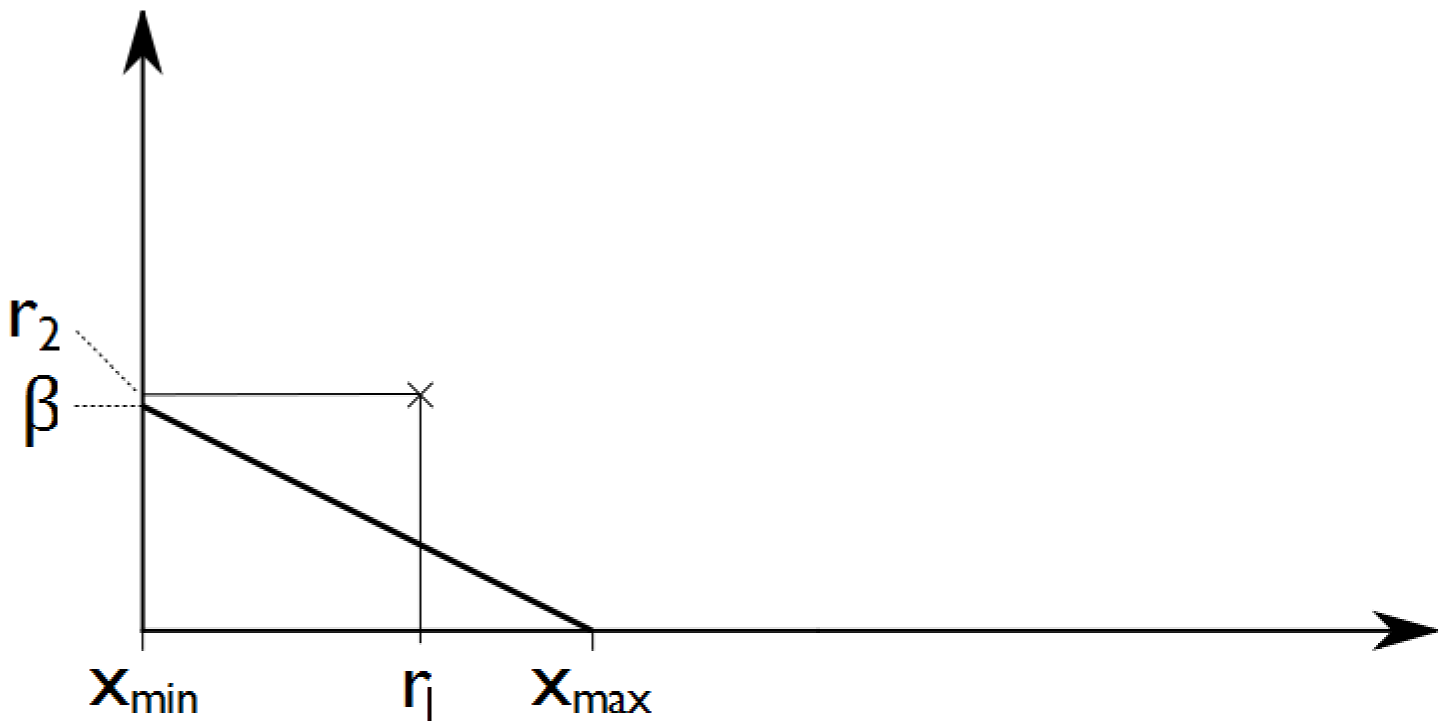
Starting Point

Assume problem with linear bi-objective front: $f(x)=\alpha x+\beta$ in $[x_{min}, x_{max}]$



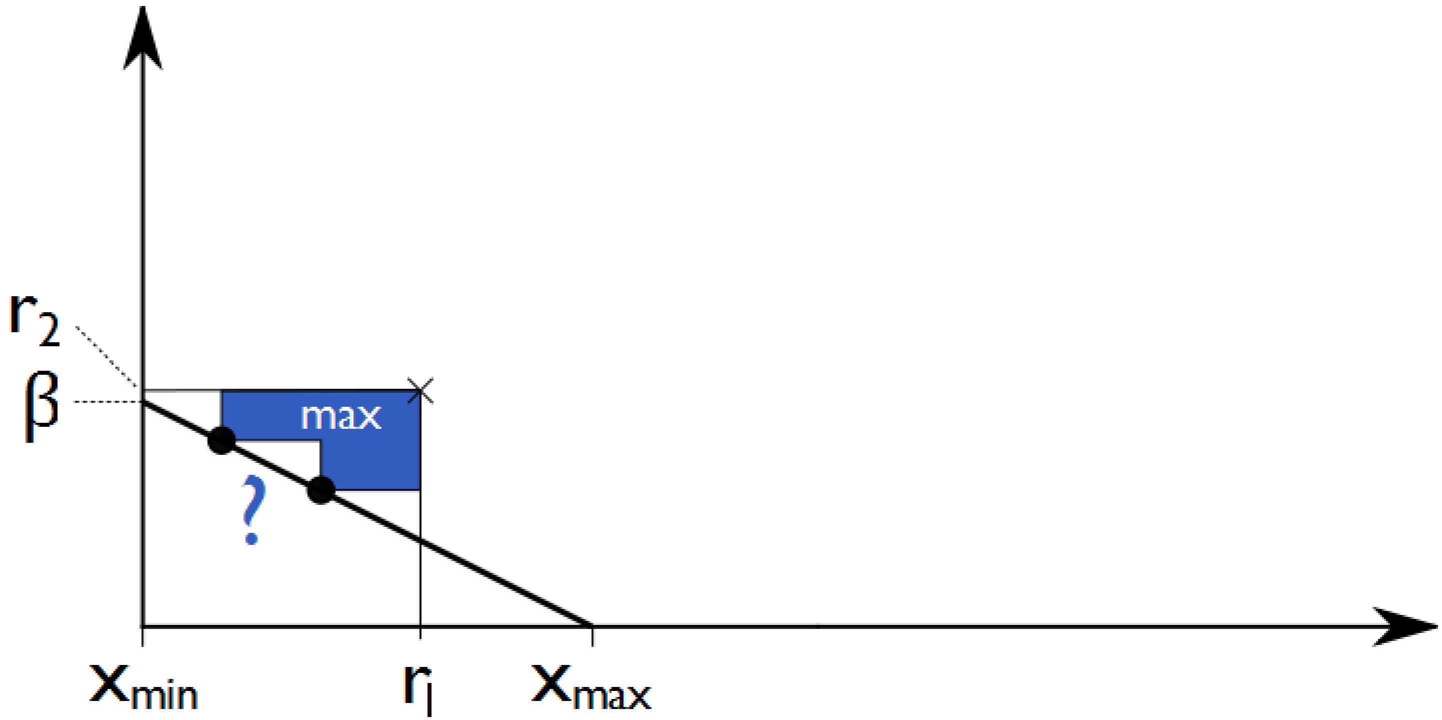
Starting Point

Assume problem with linear bi-objective front: $f(x)=\alpha x+\beta$ in $[x_{min}, x_{max}]$



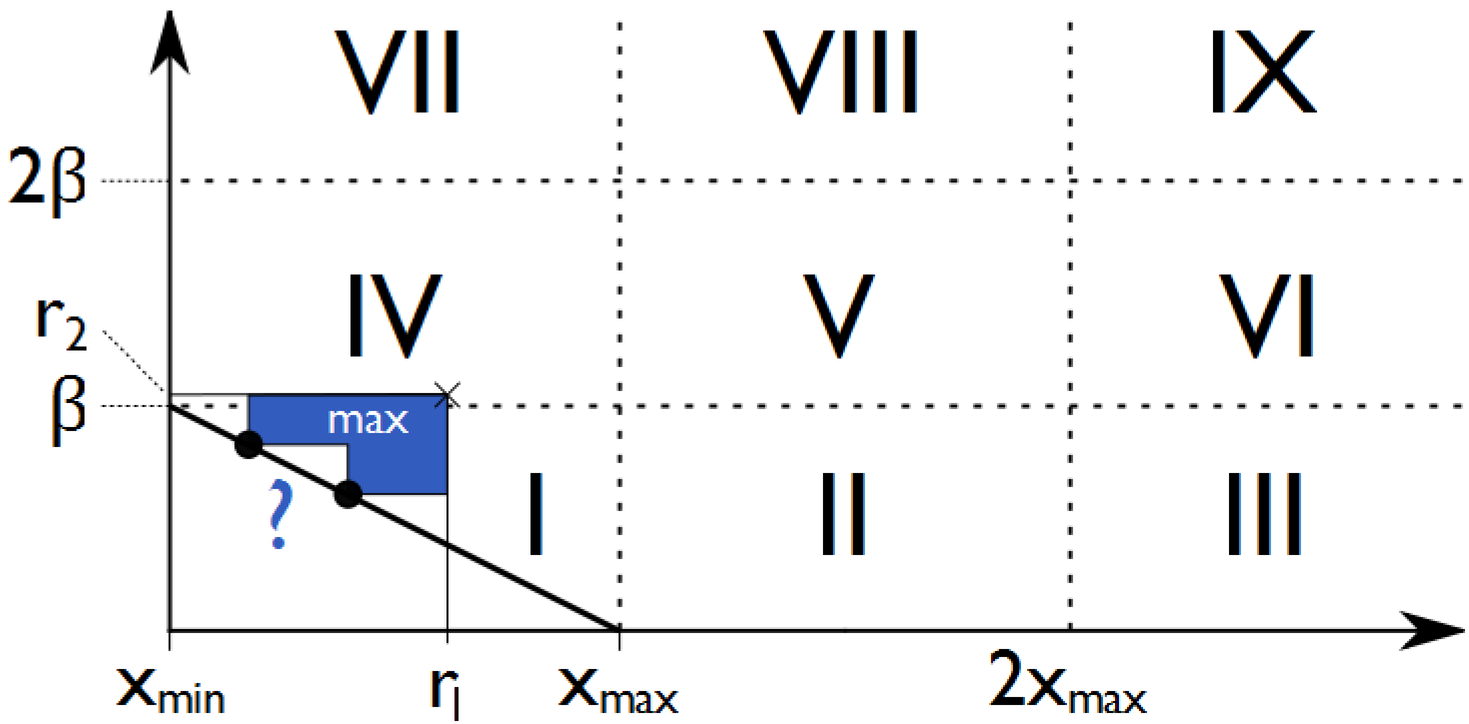
Starting Point

Assume problem with linear bi-objective front: $f(x)=\alpha x+\beta$ in $[x_{min}, x_{max}]$



Starting Point

Assume problem with linear bi-objective front: $f(x)=\alpha x+\beta$ in $[x_{min}, x_{max}]$



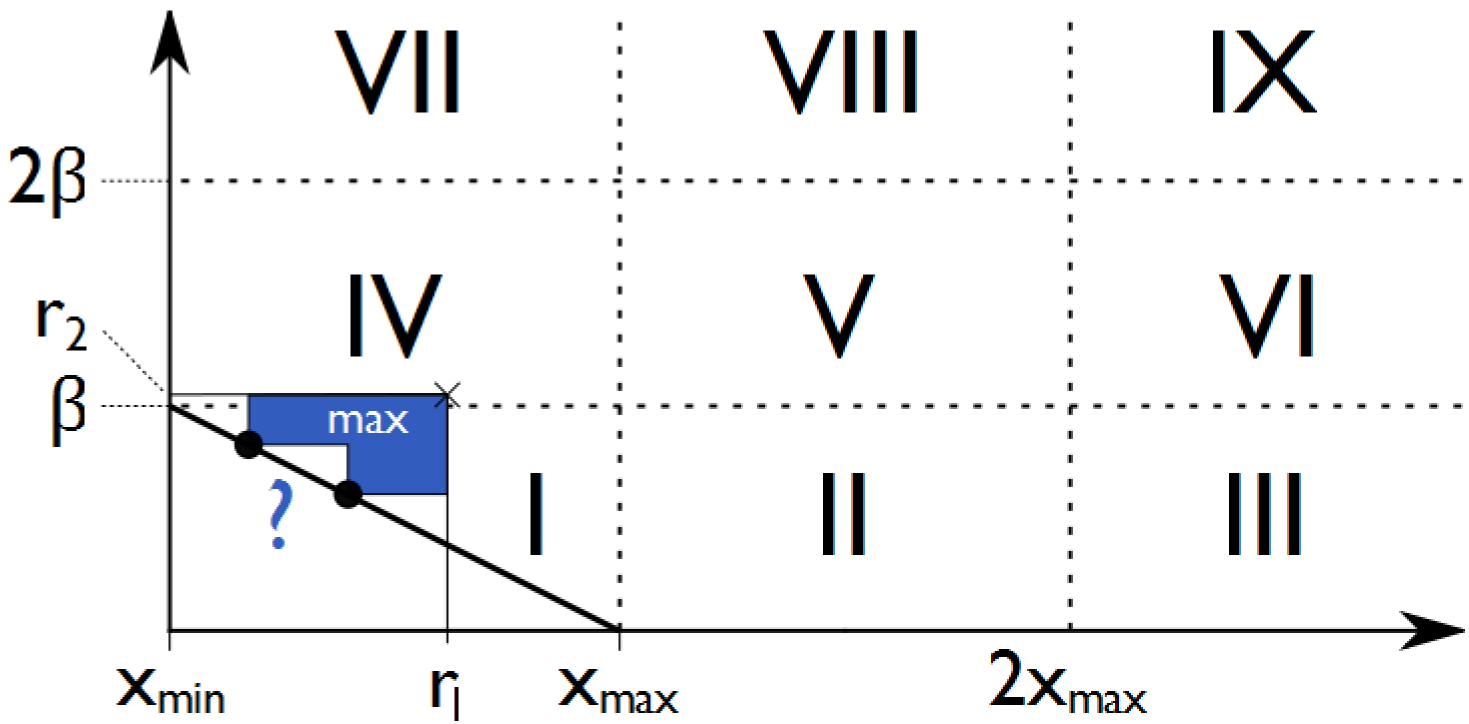
Problems:

❶ exact results only known when choosing r in I and IX [Auger et al.@FOGA'09]

❷ Case IX independent of μ

Starting Point

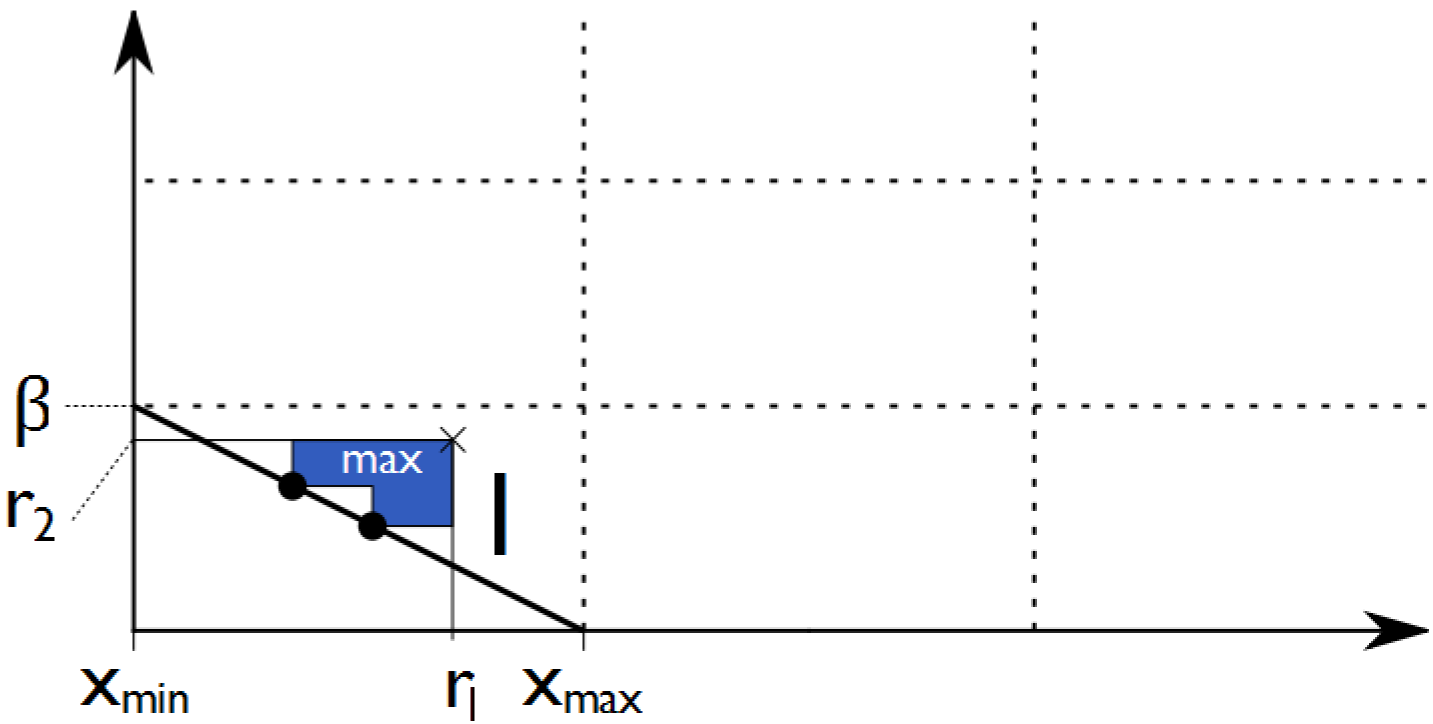
Assume problem with linear bi-objective front: $f(x)=\alpha x+\beta$ in $[x_{min}, x_{max}]$



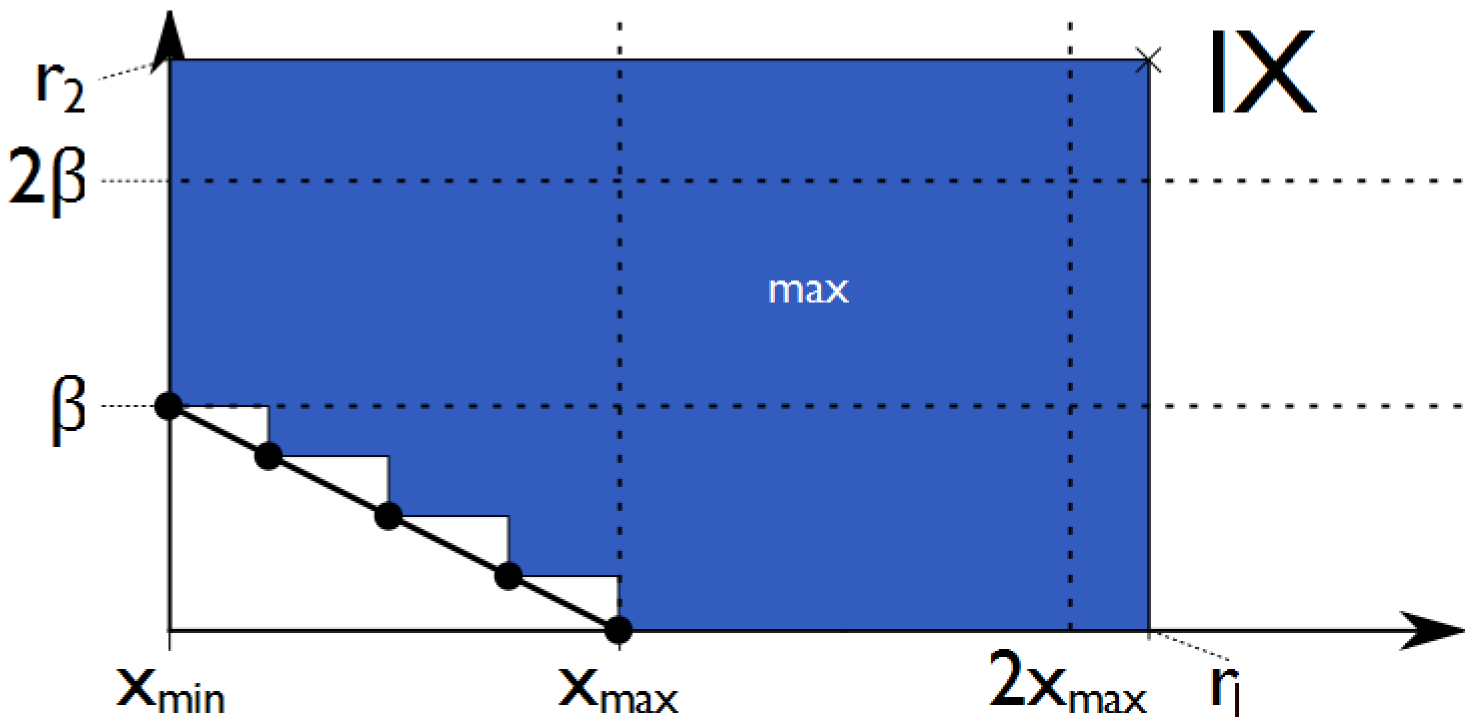
Goal:

Have exact and exhaustive results for any reasonable r and μ>1

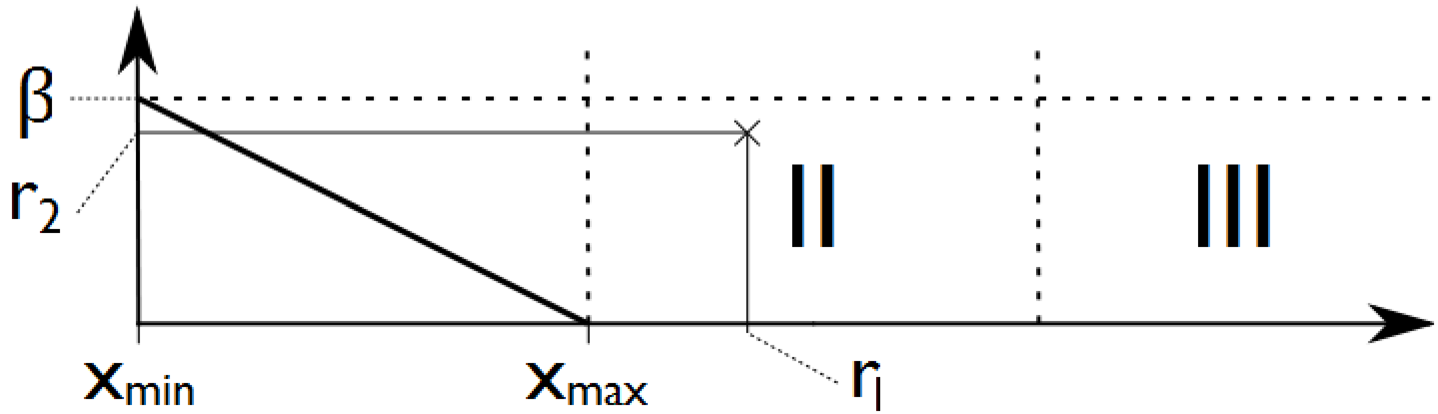
Revisiting Results for Cases I and IX



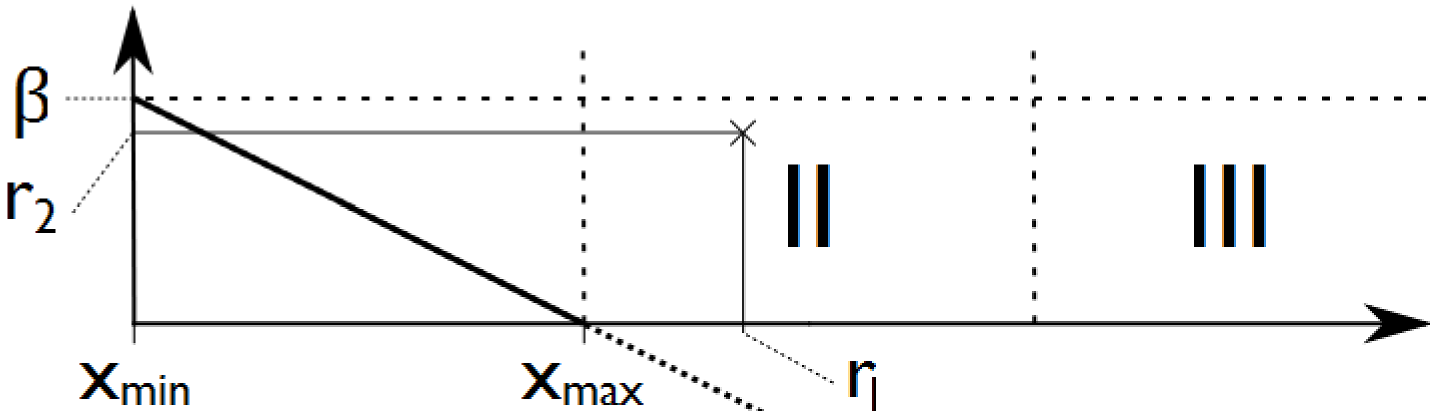
Revisiting Results for Cases I and IX



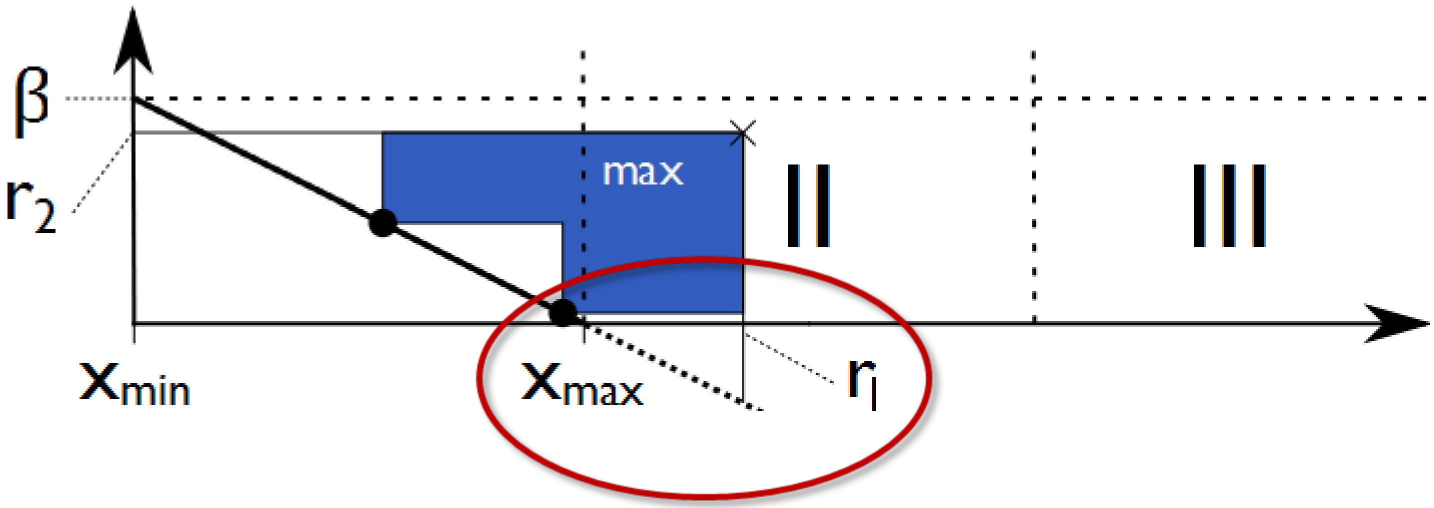
Cases II and III



Cases II and III



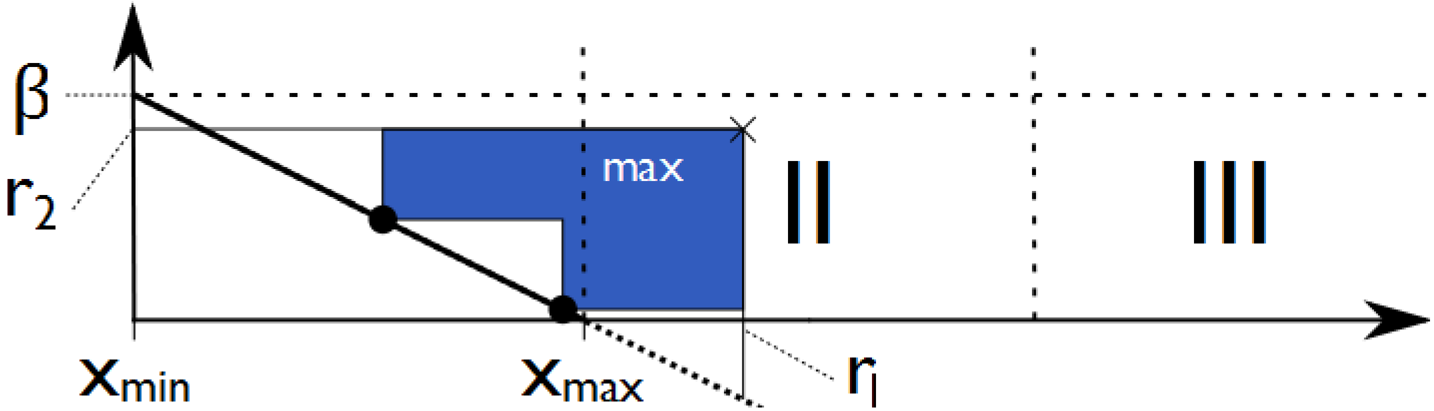
Cases II and III



What if rightmost point outside $[x_{min}, x_{max}]$?

⇒ rightmost point simply at x_{max}!

Cases II and III



simply take result of case I $\quad x_i^\mu = f^{-1}(r_2) + \frac{i}{\mu+1} \cdot (r_1 - f^{-1}(r_2))$

and restrict x_μ to x_{max}:

$$x_\mu^\mu = f^{-1}(r_2) + \frac{\mu}{\mu+1} \cdot (r_1 - f^{-1}(r_2)) \leq x_{\max} \Leftrightarrow \frac{f^{-1}(r_2)}{\mu+1} + \frac{\mu}{\mu+1} r_1 \leq x_{\max}$$

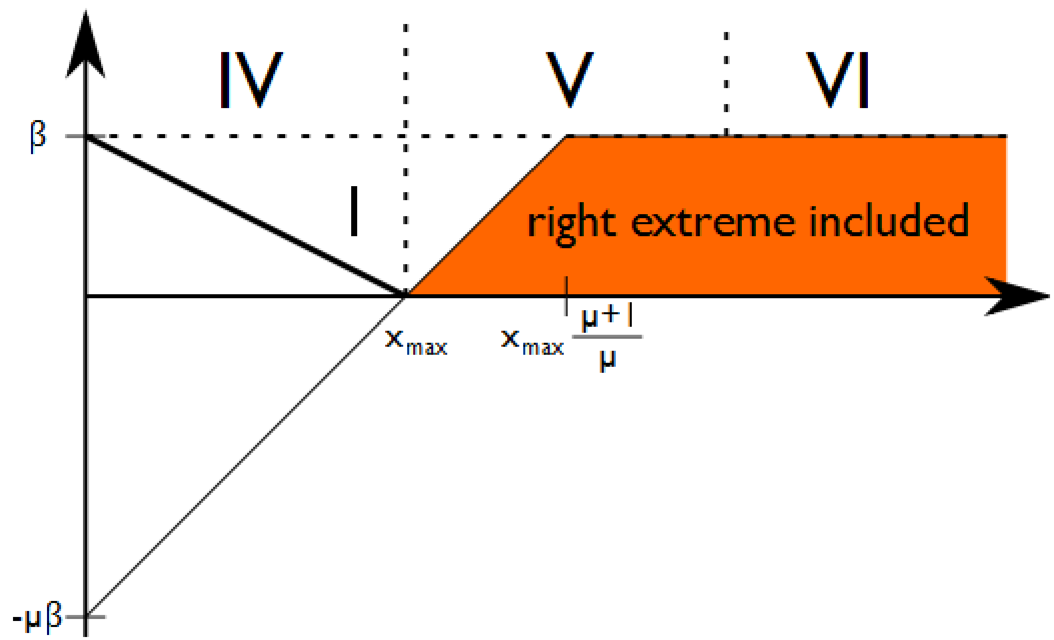
$$\Leftrightarrow r_1 \leq \frac{\mu+1}{\mu} x_{\max} - \frac{f^{-1}(r_2)}{\mu} \quad .$$

Result for Cases II and III

Optimal μ-distribution

$$x_i^\mu = f^{-1}(r_2) + \frac{i}{\mu+1}\left(\min\left\{r_1, \frac{\mu+1}{\mu}x_{\max} - \frac{f^{-1}(r_2)}{\mu}\right\} - f^{-1}(r_2)\right)$$

Right extreme included



General Result for All Cases

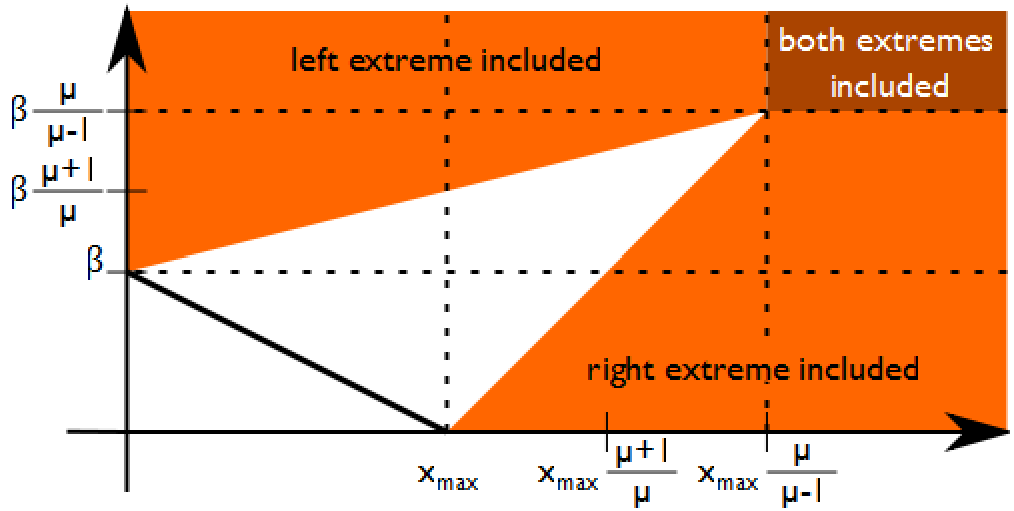
Theorem 3. *Given $\mu \in \mathbb{N}_{\geq 2}$, $\alpha \in \mathbb{R}_{<0}$, $\beta \in \mathbb{R}_{>0}$, and a linear Pareto front $f(x) = \alpha x + \beta$ within $[0, x_{\max} = -\frac{\beta}{\alpha}]$, the unique optimal μ-distribution $(x_0^\mu, \ldots, x_\mu^\mu)$ for the hypervolume indicator I_H with reference point $(r_1, r_2) \in \mathbb{R}^2_{>0}$ can be described by*

$$x_i^\mu = f^{-1}(F_l) + \frac{i}{\mu+1}\left(F_r - f^{-1}(F_l)\right)$$

for all $1 \leq i \leq \mu$ where

$$F_l = \min\{r_2, \frac{\mu+1}{\mu}\beta - \frac{1}{\mu}f(r_1), \frac{\mu}{\mu-1}\beta\} \text{ and}$$

$$F_r = \min\{r_1, \frac{\mu+1}{\mu}x_{\max} - \frac{1}{\mu}f^{-1}(r_2), \frac{\mu}{\mu-1}x_{\max}\} \ .$$



General Result for All Cases

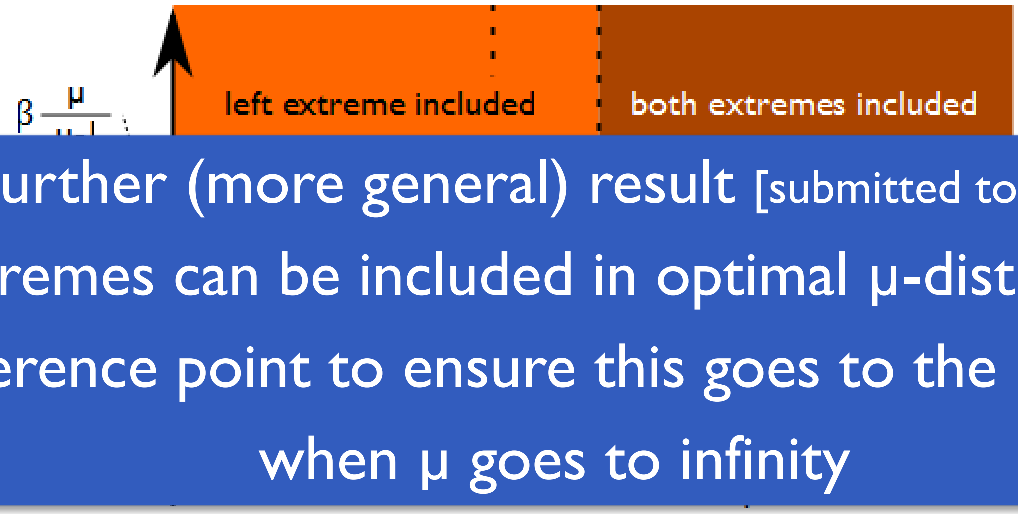
Theorem 3. *Given $\mu \in \mathbb{N}_{\geq 2}$, $\alpha \in \mathbb{R}_{<0}$, $\beta \in \mathbb{R}_{>0}$, and a linear Pareto front $f(x) = \alpha x + \beta$ within $[0, x_{\max} = -\frac{\beta}{\alpha}]$, the unique optimal μ-distribution $(x_0^\mu, \ldots, x_\mu^\mu)$ for the hypervolume indicator I_H with reference point $(r_1, r_2) \in \mathbb{R}^2_{>0}$ can be described by*

$$x_i^\mu = f^{-1}(F_l) + \frac{i}{\mu+1}\left(F_r - f^{-1}(F_l)\right)$$

for all $1 \leq i \leq \mu$ where

$$F_l = \min\{r_2, \frac{\mu+1}{\mu}\beta - \frac{1}{\mu}f(r_1), \frac{\mu}{\mu-1}\beta\} \text{ and}$$

$$F_r = \min\{r_1, \frac{\mu+1}{\mu}x_{\max} - \frac{1}{\mu}f^{-1}(r_2), \frac{\mu}{\mu-1}x_{\max}\} \ .$$



Further (more general) result [submitted to TCS]:

If extremes can be included in optimal μ-distributions the reference point to ensure this goes to the nadir point when μ goes to infinity

Open Questions…

…to be discussed here at Dagstuhl (or later)

- optimal μ-distributions for other front shapes
    - fronts forZDF, DTLZ, WFG test problems quite simple, but…
- more results for >2 objectives
    - difficult since no recurrence relations known
- optimal μ-distributions for other indicators
    - in particular some that are refinements (“binary indicators”)

Conclusions

Optimal μ-distributions for the hypervolume indicator

❶ New existence results

- difference maximization/minimization
- sufficient but not necessary criterion: lower semi-continuity
- example of no existence

❷ An exact and exhaustive result for linear bi-objective fronts

- exact description of unique optimal μ-distributions for linear case
- for all reasonable reference points and all μ>1

Several open questions that can be discussed this week:
uniqueness, other front types, other indicators, …
and their implications for performance assessment

Announcement

EMO session @ MCDM'2011 in Jyväskylä, Finland

organizers: Dimo Brockhoff and Kalyanmoy Deb

tentative deadlines: **Nov. 15, 2010** (full papers) & **Jan. 31, 2011** (abstracts)

JYVÄSKYLÄN YLIOPISTO
UNIVERSITY OF JYVÄSKYLÄ

Log in Site Map Accessibility

Search Site Search

only in current section Contact Search

Home Campus Map Introduction Contact Information Feedback

Home → In English → Congresses → The 21st International Conference on Multiple Criteria Decision Making

The 21st International Conference on Multiple Criteria Decision Making

Navigation

- The 21st International Conference on Multiple Criteria Decision Making
- Latest updates
- Registration
- Important dates
- Call for sessions
- Call for papers
- Submission procedure
- Preliminary programme
- Publications
- Venue and travel
- Accommodation
- Social Program
- Committees
- Further Information
- Media
- Contacts
- Sponsors

The 21st International Conference on Multiple Criteria Decision Making

Welcome to the land of Midnight Sun - June 13-17, 2011!

Welcome to the 21st International Conference on Multiple Criteria Decision Making! The theme for the Conference is Environment and Policy, which are active areas of application of MCDM tools in Finland (as well as around the world). The Conference will be held in the Agora Building, located on the Mattilanniemi campus of the University of Jyväskylä. The Agora Building offers an excellent venue for this conference next to lake Jyväsjärvi.

Contacts

MCDM JYVÄSKYLÄ 2011

P.O. Box 35 (Agora)
FI-40014 University of Jyvaskyla, Finland

mcdm2011@mcdmsociety.org
fax: +358 14 2602771

Visit Jyväskylä

- **Multimedia show of JYU**

Questions?

References

[Auger et al. @ FOGA'09] A. Auger, J. Bader, D. Brockhoff, and E. Zitzler. **Theory of the Hypervolume Indicator: Optimal μ-Distributions and the Choice of the Reference Point**. In *Foundations of Genetic Algorithms (FOGA 2009)*, pages 87–102, New York, NY, USA, 2009. ACM

[Beume et al. @ EMO'09] N. Beume, B. Naujoks, M. Preuss, G. Rudolph, and T. Wagner. **Effects of 1-Greedy *S*-Metric-Selection on Innumerably Large Pareto Fronts**. In M. Ehrgott et al., editors, *Conference on Evolutionary Multi-Criterion Optimization (EMO 2009)*, volume 5467 of *LNCS*, pages 21–35. Springer, 2009

[Bringmann and Friedrich @ GECCO'10] K. Bringmann and T. Friedrich. **The Maximum Hypervolume Set Yields Near-optimal Approximation**. In J. Branke et al., editors, *Genetic and Evolutionary Computation Conference (GECCO 2010)*, pages 511–518. ACM, 2010

[Brockhoff @ SEAL'10] D. Brockhoff. **Optimal μ-Distributions for the Hypervolume Indicator for Problems With Linear Bi-Objective Fronts: Exact and Exhaustive Results**. In *Simulated Evolution and Learning (SEAL 2010)*. Springer, 2010. to appear

[Fleischer @ EMO'03] M. Fleischer. **The Measure of Pareto Optima. Applications to Multi-Objective Metaheuristics**. In C. M. Fonseca et al., editors, *Conference on Evolutionary Multi-Criterion Optimization (EMO 2003)*, volume 2632 of *LNCS*, pages 519–533, Faro, Portugal, 2003. Springer

[Zitzler et al. @ EMO'07] E. Zitzler, D. Brockhoff, and L. Thiele. **The Hypervolume Indicator Revisited: On the Design of Pareto-compliant Indicators Via Weighted Integration**. In S. Obayashi et al., editors, *Conference on Evolutionary Multi-Criterion Optimization (EMO 2007)*, volume 4403 of *LNCS*, pages 862–876, Berlin, 2007. Springer