Theoretically Investigating Optimal μ-Distributions for the Hypervolume Indicator: First Results for Three Objectives

ETH Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

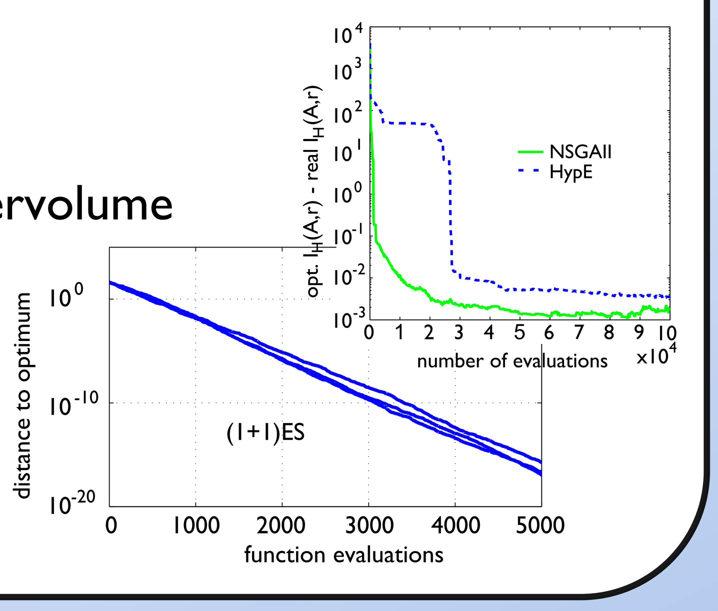
INRIA

Motivation

The **Hypervolume Indicator** is often used
- within indicator-based selection in evolutionary multiobjective optimization
- in performance assessment

Optimization Goal:
- find set of μ solutions that maximize the hypervolume indicator: the optimal μ-distribution

Why interesting?
- understanding indicator-based algorithms
  - convergence to the optimization goal?
  - convergence speed?
- performance assessment
  - absolute vs. relative interpretation of hypervolume
  - fixing target values (horizontal view)



Contribution

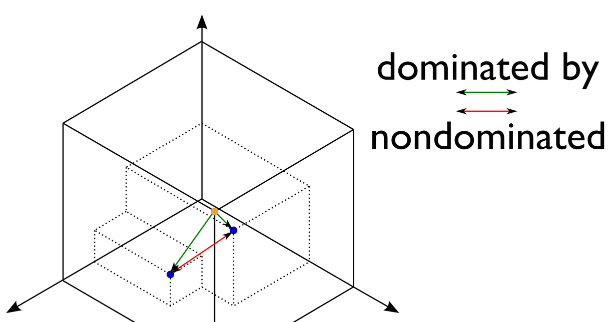
Previous work investigated optimal μ-distributions for biobjective problems only

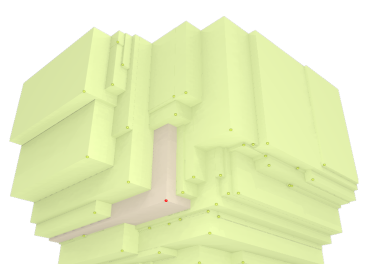
Here:
First results for 3 objectives

In particular:
- general results: existence and monotonicity
- geometric properties of hypervolume contributions
- fronts for which the extremes are never included in opt. μ-distributions

Definitions and Notations

Scenario:
- minimize k objective functions $\mathcal{F}: X \to \mathbb{R}^k$
- Pareto dominance
- Pareto set: {x∈X | ∄y∈X: y≼x and x⋠y}

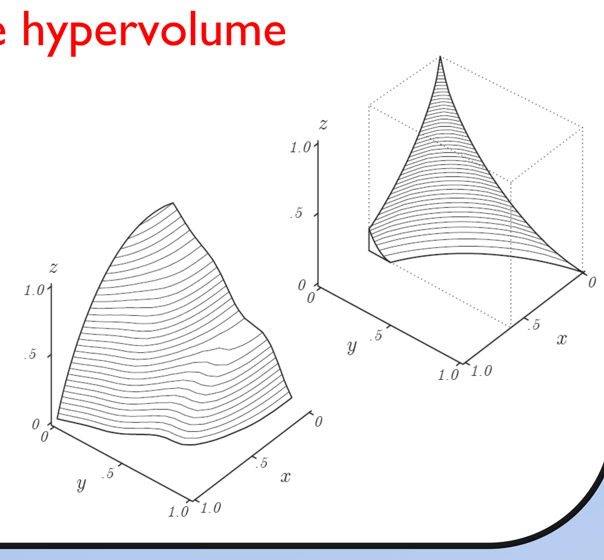




Hypervolume Indicator:
$I_H(A,r) = \lambda(\{z \mid \exists a \in A: a \preccurlyeq z \preccurlyeq r\})$

Optimal μ-Distribution:
set of μ points that maximize the hypervolume among all set of μ points

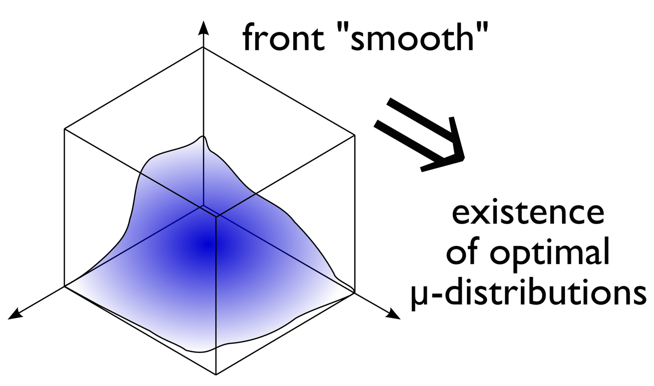
Specific for the 3-Objective Case:
- assume front to be described implicitly via f_{3d}
- i.e., for all points a on front $f_{3d}(a)=0$
- f_{3d} is restricted to [xmin,xmax]x[ymin,ymax]x[zmin,zmax]
- sometimes: explicit description: z = f(x,y)
- given for most DTLZ, IHR, and WFG problems



General Results

Theorem 1 (Existence of optimal μ-distributions for 3-objective problems). *Assume a 3-objective problem and assume that the front is described explicitly by a 2-dimensional function f, i.e., points of the Pareto front satisfy $z = f(x,y)$ (or $y = f(x,z)$ or $x = f(y,z)$). If the function f is continuous, there exists (at least) one set of μ points maximizing the hypervolume.*

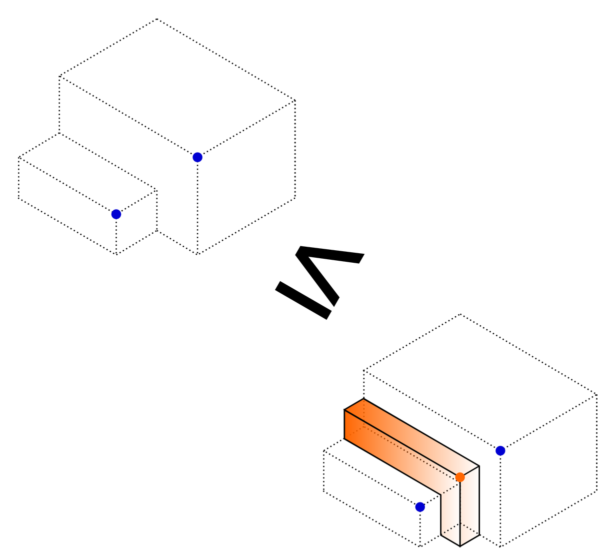
note: lower semi-continuity enough for existence



Proposition 1 (Strict monotonicity in μ of the optimal hypervolume value). *Let $x_{\min}, x_{\max}, y_{\min}, y_{\max}, z_{\max}, z_{\min} \in \mathbb{R}$, $f_{3d}: \mathbb{R}^3 \to \mathbb{R}$, and let $P = \{(x,y,z) \in \mathbb{R}^3 \mid f_{3d}(x,y,z) = 0 \wedge (x_{\min} \leq x \leq x_{\max}) \wedge (y_{\min} \leq y \leq y_{\max}) \wedge (z_{\min} \leq z \leq z_{\max})\}$ describe the corresponding Pareto front. Let μ_1 and $\mu_2 \in \mathbb{N}$ with $\mu_1 < \mu_2$, then*

$$\overline{I_H^{\mu_1}} < \overline{I_H^{\mu_2}}$$

holds if P contains at least $\mu_1 + 1$ elements (x_i, y_i, z_i) for which $x_i < r_1$, $y_i < r_2$, and $z_i < r_3$ holds where $r = (r_1, r_2, r_3)$ is the hypervolume's reference point.

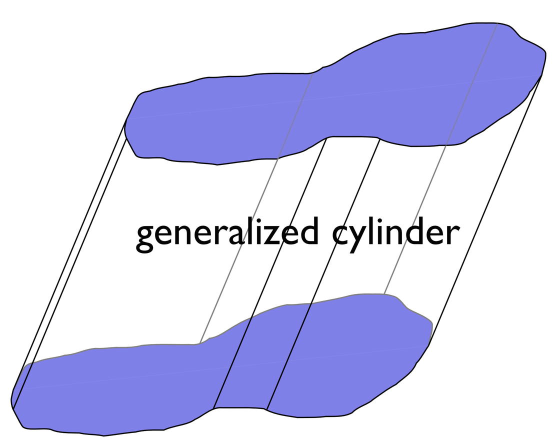


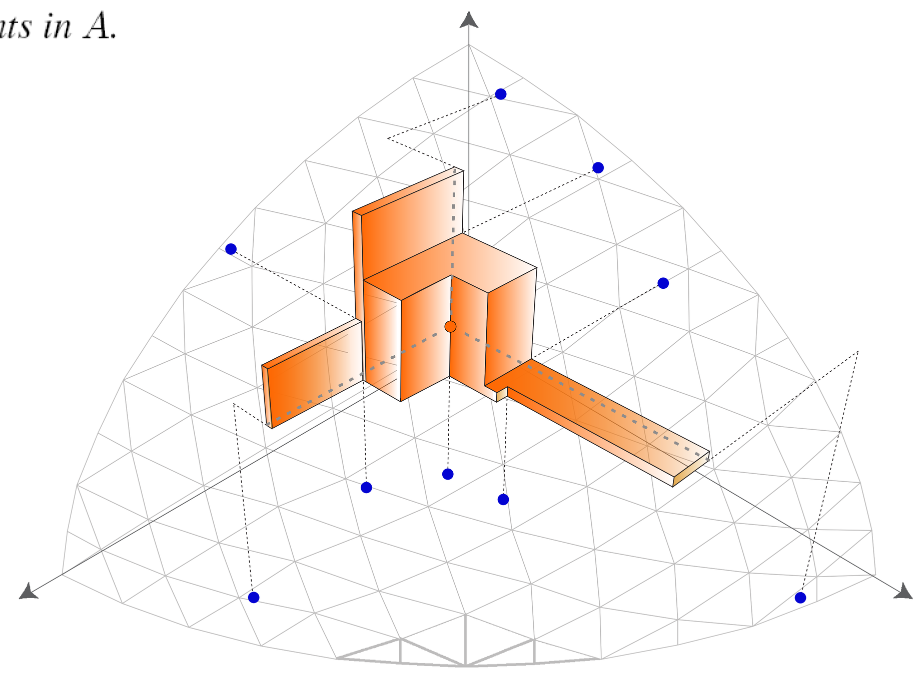
Geometric Properties

Lemma 1. *Given a set $A \subseteq \mathbb{R}^3$ of 3-dimensional objective vectors, the hypervolume solely dominated by a single point $a \in A$ is an axis-aligned cuboid, with the point itself and the reference point as the end points of one of the cuboid's space diagonals, from which three generalized cylinders are cut—one parallel to each coordinate axis with steplike base areas, which depend on the other points in A.*

Proof idea:

geometrical considerations

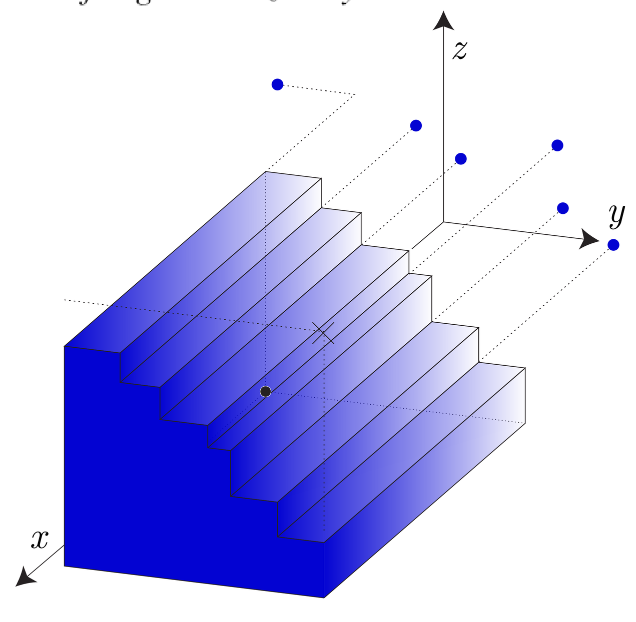




Lemma 2. *Given a set of 3-dimensional objective vectors $A \subseteq \mathbb{R}^3$. An extreme point of A, i.e., a point with the largest objective value among all points in A for (at least) one objective, solely dominates a region the shape of which is itself a generalized cylinder with a steplike base area.*

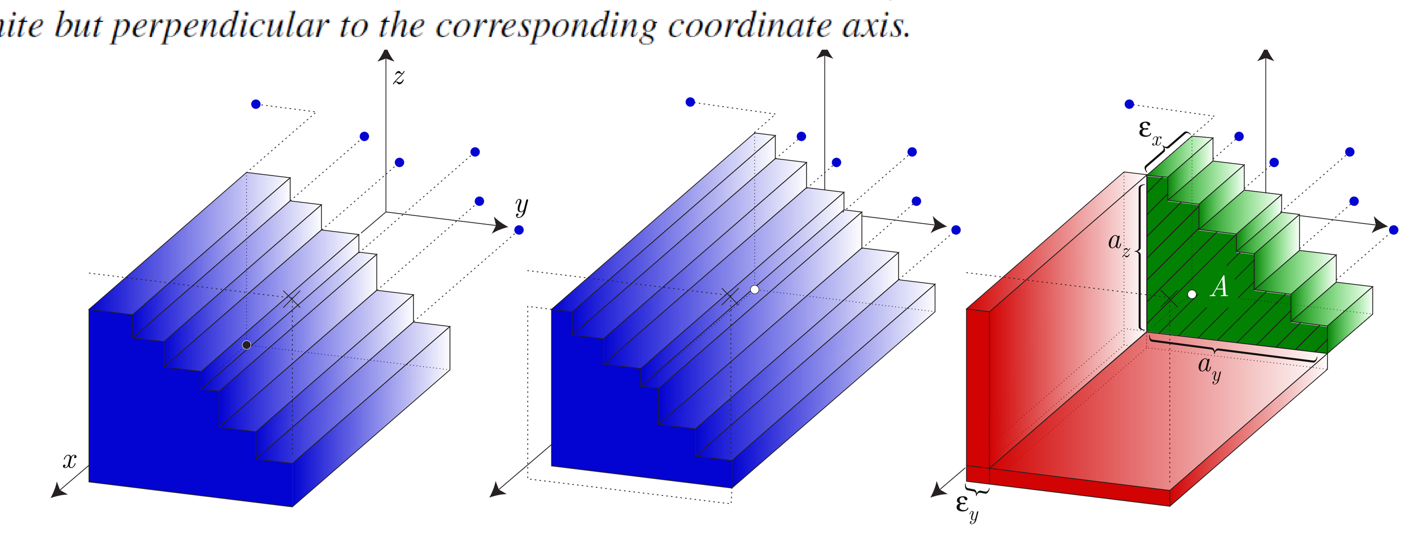
Proof idea:

geometrical considerations



Obtaining the Extremes

Theorem 2. *Extreme solutions are never included in optimal μ-distributions if the gradient at the extreme is finite but perpendicular to the corresponding coordinate axis.*



extreme point's contribution | contribution if shifted inside | losses and gains

Sketch of Proof:

Consider extreme $p_{\text{xmax}} = (\overline{x},y,z)$ in x-direction with $\nabla f_{3d}(p_{\text{xmax}}) \cdot (1,0,0) = 0$ and move it along the front to

$$p_{\text{xmax}} - \begin{pmatrix} \varepsilon_x \\ \varepsilon_y \\ \varepsilon_z \end{pmatrix} = p_{\text{xmax}} - \begin{pmatrix} \varepsilon_x \\ 0 \\ 0 \end{pmatrix} + \nu \begin{pmatrix} 0 \\ \partial_2 f(p_{\text{xmax}}) \\ \partial_3 f(p_{\text{xmax}}) \end{pmatrix}$$

then we get: benefit: $\varepsilon_x \cdot A$ deficit: $(\varepsilon_y \cdot a_z + \varepsilon_z \cdot a_y + \varepsilon_y \cdot \varepsilon_z) \cdot (r_1 - \overline{x})$ and

$$\lim_{\varepsilon_x \to 0} \frac{\text{deficit}}{\text{benefit}} = \lim_{\varepsilon_x \to 0} \frac{\varepsilon_y \cdot a_z \cdot (r_1 - \overline{x})}{\varepsilon_x \cdot A} + \frac{\varepsilon_z \cdot a_y \cdot (r_1 - \overline{x})}{\varepsilon_x \cdot A} + \frac{\varepsilon_y \cdot \varepsilon_z \cdot (r_1 - \overline{x})}{\varepsilon_x \cdot A} = 0$$

since $\lim_{\varepsilon_x \to 0} \frac{\varepsilon_y}{\varepsilon_x} = 0$ and $\lim_{\varepsilon_x \to 0} \frac{\varepsilon_z}{\varepsilon_x} = 0$ which is shown exemplary for the former:

Assuming $\partial_2 f_{3d}(p_{\text{xmax}}) \neq 0$ we get $\varepsilon_z = \varepsilon_y \frac{\partial_3 f_{3d}(p_{\text{xmax}})}{\partial_2 f_{3d}(p_{\text{xmax}})}$ and with Taylor expansion we have

$$f_{3d}(p_{\text{xmax}} - \varepsilon) = f_{3d}(p_{\text{xmax}}) - \nabla f_{3d}(p_{\text{xmax}}) \cdot \varepsilon + \mathcal{O}(||\varepsilon||^2)$$

and thus $\nabla f_{3d}(p_{\text{xmax}}) \cdot \varepsilon - \mathcal{O}(||\varepsilon||^2) = 0$ for any $\varepsilon \geq 0$ (definition of front). Hence $\lim_{\varepsilon \to 0} \nabla f_{3d}(p_{\text{xmax}}) \cdot \varepsilon = \lim_{\varepsilon \to 0} \mathcal{O}(||\varepsilon||^2)$ and even $\lim_{\varepsilon \to 0} (\nabla f_{3d}(p_{\text{xmax}}) \cdot \varepsilon / ||\varepsilon||) = \lim_{\varepsilon \to 0} \mathcal{O}(||\varepsilon||) = 0$ which is equivalent to

$$\lim_{\varepsilon \to 0} \frac{\partial_2 f_{3d}(p_{\text{xmax}}) \cdot \varepsilon_y + \partial_3 f_{3d}(p_{\text{xmax}}) \cdot \varepsilon_z}{\sqrt{\varepsilon_x^2 + \varepsilon_y^2 + \varepsilon_z^2}} = \lim_{\varepsilon \to 0} \frac{\partial_2 f_{3d}(p_{\text{xmax}}) + \frac{(\partial_3 f_{3d}(p_{\text{xmax}}))^2}{\partial_2 f_{3d}(p_{\text{xmax}})}}{\sqrt{\frac{\varepsilon_x^2}{\varepsilon_y^2} + 1 + \frac{(\partial_3 f_{3d}(p_{\text{xmax}}))^2}{(\partial_2 f_{3d}(p_{\text{xmax}}))^2}}} = 0$$

Anne Auger[1], Johannes Bader[2], and Dimo Brockhoff[1]
[1] INRIA Saclay Ile-de-France, TAO, LRI, Paris Sud University, Orsay, France, firstname.lastname@inria.fr
[2] ETH Zurich, 8092 Zurich, Switzerland, johannes.bader@tik.ee.ethz.ch