Author manuscript, published in "Parallel Problem Solving from Nature (PPSN XI) (2010)"

inria-00534906, version 1 - 10 Nov 2010

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

Anne Auger[1] Johannes Bader[2] Dimo Brockhoff[1]

[1] TAO Team, INRIA Saclay, LRI, Paris Sud University, 91405 Orsay Cedex, France
firstname.lastname@inria.fr
[2] Computer Engineering and Networks Lab, ETH Zurich, 8092 Zurich, Switzerland
johannes.bader@tik.ee.ethz.ch

Abstract. Several indicator-based evolutionary multiobjective optimization algorithms have been proposed in the literature. The notion of *optimal μ-distributions* formalizes the optimization goal of such algorithms: find a set of μ solutions that maximizes the underlying indicator among all sets with μ solutions. In particular for the often used hypervolume indicator, optimal μ-distributions have been theoretically analyzed recently. All those results, however, cope with bi-objective problems only. It is the main goal of this paper to extend some of the results to the 3-objective case. This generalization is shown to be not straight-forward as a solution's hypervolume contribution has not a simple geometric shape anymore in opposition to the bi-objective case where it is always rectangular. In addition, we investigate the influence of the reference point on optimal μ-distributions and prove that also in the 3-objective case situations exist for which the Pareto front's extreme points cannot be guaranteed in optimal μ-distributions.

1 Introduction

Several evolutionary multiobjective optimization (EMO) algorithms have been proposed to tackle multiobjective optimization problems. Among them, the indicator-based algorithms are the most recent developments [17, 6, 13]. These algorithms often explicitly optimize a unary quality indicator which maps a set of solutions to a single real value. This not only allows to decouple preference articulation from the search algorithm [17] but also transforms the multiobjective problem into a single-objective one: the goal is no longer to find or approximate the so-called Pareto front, but to find a solution set of fixed size (typically the population size μ) that maximizes the indicator. Therefore, it is important to characterize these solution sets to understand the optimization goal implicitly defined by a given indicator. In particular when benchmarking algorithms on certain test functions, it is highly useful to know the largest possible indicator value achievable with μ points. Throughout the paper, and in line with [2], we use the term *optimal μ-distribution* for those sets of μ solutions optimizing a given indicator.

One of the most often used quality indicators within indicator-based EMO algorithms is the hypervolume indicator or $\mathcal{S}$-metric which maps a set of solutions to the *size of the objective space covered* [18]. It has the nice property of being a refinement of the

* This is an author version of the PPSN'2010 paper published by Springer Verlag. The final publication is available at www.springerlink.com.

Pareto dominance relation [19] which implies that the optimal μ-distributions contain only solutions that are mapped to the Pareto front [10]. The question of *how* the optimal μ-distributions are spread over the Pareto front, interestingly, has only gained attention recently. Besides specific results on optimal μ-distributions in the case of linear Pareto fronts [9, 5], optimal μ-distributions have been theoretically investigated in more detail in [2, 1] for bi-objective problems. The main results are an exact characterization of optimal μ-distributions for problems with arbitrary linear Pareto fronts and a limit result in terms of a density for general front shapes that can be described by a continuous and differentiable function f. The density result proves that the empirical density of points converges to a density proportional to the square root of the negative of the first derivative of the front. In other words, it is only the slope of the front which determines how the points that maximize the hypervolume indicator are distributed—independent of the second derivative, i.e., whether the front is convex or concave. It has also been proven in [2] that for certain types of fronts, no finite reference point of the hypervolume indicator allows to have the extreme points in the optimal μ-distribution; for the remaining cases, it has been shown where to place the reference point such that the extremes are included. Later, the relation between optimal μ-distributions for the hypervolume indicator and the approximation ratio has been investigated theoretically as well [11, 7]. However, also in these studies, the results are restricted to only two objectives. The main reason why almost no results about optimal μ-distributions for 3-objective problems are known[3] is that the geometry of the hypervolume becomes more complicated in higher dimensions. We will see later on that, e.g., the hypervolume contributions of single points are not anymore simple rectangles or cuboids if 3-objective problems are considered and that all solutions can have an influence on the optimal placement of one point—in comparison to the local property proven in [2] for bi-objective problems.

Contributions of this paper. In this paper, we present for the first time theoretical results about optimal μ-distributions for the hypervolume indicator for more than 2 objectives, in contrast to [2, 1, 11, 7] where only bi-objective problems were tackled. Besides fundamental results on the existence and the monotonicity of optimal μ-distributions (Sec. 3), we prove fundamental, yet often not obvious statements about the shape of the hypervolume contribution of a single solution (Sec. 4) and investigate their implications on optimal μ-distributions—in particular on the influence of the reference point (Sec. 5). More specifically, we prove that situations exist (and characterize them) for which the extreme points of the Pareto front will never be contained in an optimal μ-distribution for 3-objective problems which covers the results for the bi-objective case of [2]. The results show in particular, that the investigation of optimal μ-distributions is, indeed, more difficult for 3-objective problems than in the case of 2 objectives.

2 Preliminaries

Without loss of generality (w.l.o.g.), we consider minimization problems where the vector-valued objective function is defined as $\mathcal{F}\colon X \to \mathbb{R}^k$ and k is the number of objectives. In this paper, $k = 3$ most of the time. We say $\mathcal{F}$ maps a solution $x \in X$

[3] The only exception is a conjecture in [3] about the influence of the dimension on the density.

inria-00534906, version 1 - 10 Nov 2010

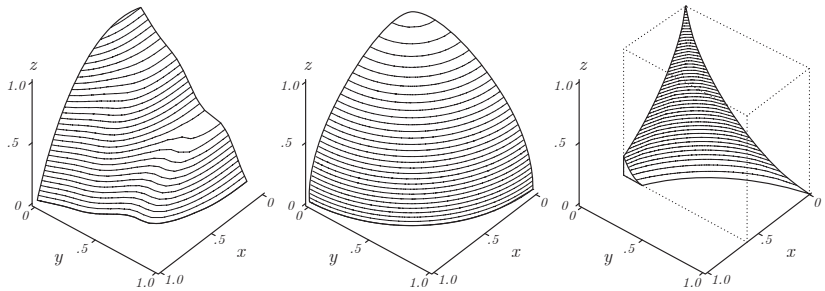

Fig. 1. 3-dimensional fronts implicitly described by $f_{3d}(x, y, z) = 0$ and restricted to the cube $[0, 1]^3$ (left and middle), and to $[0, 0.6] \times [0, 1] \times [0, 1]$ (right). **Left:** $f_{3d} = x^3 + (.1 \cdot (\sin(5\pi \cdot y) + 10)) \cdot y + z - 1$; **Middle:** $f_{3d} = x^2 + y^2 + z^2 - 1$; **Right:** $f_{3d} = x^{2/3} + y^{2/3} + z^{2/3} - 1$

from the decision space X to its objective vector $\mathcal{F}(x) = (\mathcal{F}_1(x), \ldots, \mathcal{F}_k(x)) \in \mathbb{R}^k$ within the objective space $\mathcal{F}(X) \subseteq \mathbb{R}^k$. As the single objective functions $\mathcal{F}_i$, in general, cannot be simultaneously minimized and therefore no single optimal solution exists, we denote the sought *set* of so-called *Pareto-optimal solutions* (or *Pareto set*) as the set $\{x \in X \mid \nexists y \in X\text{: } y \preceq x \text{ and } x \not\preceq y\}$. Thereby, the relation $\preceq$ is defined as: $x \preceq y$ if and only if $\mathcal{F}_i(x) \leq \mathcal{F}_i(y)$ for all $1 \leq i \leq k$ and we say, x is weakly dominating y if $x \preceq y$. The image of the Pareto set is called *Pareto front* or *front* for short. Note that in the remainder of this paper, we make an abuse of terminology and use the term *solution* both for a solution in the decision space and for its corresponding objective vector. Moreover, in order to increase readability, we also define $\preceq$ on objective vectors.

The Hypervolume Indicator. The hypervolume indicator of a solution set has been introduced as *the size of the objective space covered* [18]. Here, we formalize the hypervolume indicator $I_H(A, r)$ for sets of objective vectors $A \subseteq \mathbb{R}^k$ and a reference point $r \in \mathbb{R}^k$ according to [2] to ease readability compared to defining I_H for solutions in X as in the original paper: $I_H(A, r) = \lambda\big(\bigcup_{a \in A} C(a, r)\big)$ where $C(a, r) = \{z \in \mathbb{R}^k \mid a \preceq z \preceq r\}$ is the (hyper-)cuboid containing all objective vectors that are weakly dominated by a and themselves weakly dominate r. λ is the Lebesgue measure.

Notations for 3-objective problems. For the specific case of 3-objective problems, we assume the Pareto front to be implicitly describable as the points $(x, y, z) \in \mathbb{R}^3$ for which a function $f_{3d} : \mathbb{R}^3 \to \mathbb{R}$ is zero[4]: $f_{3d}(x, y, z) = 0$. W.l.o.g., we restrict the front to a cuboid $[x_{\min}, x_{\max}] \times [y_{\min}, y_{\max}] \times [z_{\max}, z_{\min}]$, see Fig. 1. Besides a few exceptions of disrupted fronts, the Pareto fronts of well-known test problems, e.g., from the DTLZ [8], IHR [13], or WFG [12] test suites, can be described as assumed. Furthermore, we denote the reference point of the hypervolume by $r = (r_1, r_2, r_3)$. For some proofs, we will need an explicit description of the front, i.e., in terms of $z = f(x, y)$ (resp. $y = f(x, z)$, $x = f(y, z)$). Note that it is not always possible to find an explicit representation of an implicit equation $f_{3d}(x, y, z) = 0$. However, it is possible locally assuming regularity of f_{3d} as stated by the implicit function theorem [14]. Therefore, assuming an explicit representation is not very restrictive.

[4] In addition, in order to describe a Pareto front, the partial derivatives of f_{3d} with respect to the first, second, and third variableare not supposed to change their sign, see for example [15].

inria-00534906, version 1 - 10 Nov 2010

inria-00534906, version 1 - 10 Nov 2010

3 General Results on Optimal μ-Distributions in 3-Objective Problems

In this section, we generalize some basic results of [2] about the existence of optimal μ-distributions and their monotonicity in μ to the 3-objective case. The proofs comprise the same ideas as in the bi-objective case though they are a bit more technical.

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.*

Proof. Assume w.l.o.g. that the front is described via $z = f(x, y)$. Let $p_1, \ldots, p_\mu$ be μ points of $\mathbb{R}^3$. A point p_i writes as $(x_i, y_i, f(x_i, y_i))$. Since f is continuous, the mapping $((x_1, y_1), \ldots, (x_\mu, y_\mu)) \rightarrow \lambda(\bigcup_i C((x_i, y_i, f(x_i, y_i)), r))$, where $C((x_i, y_i, f(x_i, y_i)), r)$ is the cuboid with space diagonal defined by the extremes p_i and r, is continuous according to the Lebesgue dominated convergence theorem [4]. Moreover I_H is upper bounded by the hypervolume of the entire front. From the Extreme Value Theorem, there exists a set of μ points maximizing the hypervolume indicator. □

Note that the previous theorem states the existence but not the uniqueness, which cannot be guaranteed in general and that, in principle, the result can be easily generalized to the weighted hypervolume of [16]. A set of points maximizing the hypervolume whose existence is proven in the previous theorem will be called *optimal μ-distribution*. The associated value of the hypervolume is denoted as $\overline{I_H^\mu}$.

The following proposition establishes that the hypervolume of optimal $(\mu + 1)$-distributions is strictly larger than the hypervolume of optimal μ-distributions in the case of 3-objective problems and when the Pareto front contains at least $\mu + 1$ distinct points. This result is a generalization of Lemma 1 in [2].

Proposition 1 (Strict monotonicity in μ of the optimal hypervolume value). *Let $x_{\min}$, $x_{\max}$, $y_{\min}$, $y_{\max}$, $z_{\max}$, $z_{\min} \in \mathbb{R}$, $f_{3d} : \mathbb{R}^3 \rightarrow \mathbb{R}$, 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.

Proof. To prove the proposition, it suffices to show the inequality for $\mu_2 = \mu_1 + 1$ where we denote μ_1 by μ for readability. Assume the optimal μ-distribution is $D_\mu = \{(x_1^\mu, y_1^\mu, z_1^\mu), \ldots, (x_\mu^\mu, y_\mu^\mu, z_\mu^\mu)\}$ with $x_i^\mu, y_i^\mu, z_i^\mu \in \mathbb{R}$ and $f_{3d}(x_i^\mu, y_i^\mu, z_i^\mu) = 0$ for all $1 \leq i \leq \mu$. Since P contains at least $\mu + 1$ elements, the set $P \backslash D_\mu$ is not empty and we can pick any $p_{\text{new}} = (x_{\text{new}}, y_{\text{new}}, z_{\text{new}}) \in P \backslash D_\mu$ to define a set $S = D_\mu \cup \{p_{\text{new}}\}$. As $I_H(D_{\mu_2}) \geq I_H(S)$ holds, it remains to prove that $I_H(S) > I_H(D_\mu)$. To this end, let us sort the points in D_μ with respect to each objective and pick the solution with the

smallest x- (y-, z-) value which is larger than x_{new} (y_{new}, z_{new}) and denote it by $\overline{x}$ ($\overline{y}$, $\overline{z}$). If such a solution does not exist in D_μ, we set $\overline{x}$ to r_1 ($\overline{y}$ to r_2, $\overline{z}$ to r_3):

$$\begin{aligned}
\overline{x} &= \min\left\{\left\{x_i^\mu \mid (x_i^\mu, y_i^\mu, z_i^\mu) \in D_\mu \wedge x_{\text{new}} < x_i^\mu < r_1\right\}, r_1\right\} \\
\overline{y} &= \min\left\{\left\{y_i^\mu \mid (x_i^\mu, y_i^\mu, z_i^\mu) \in D_\mu \wedge y_{\text{new}} < y_i^\mu < r_2\right\}, r_2\right\} \\
\overline{x} &= \min\left\{\left\{z_i^\mu \mid (x_i^\mu, y_i^\mu, z_i^\mu) \in D_\mu \wedge z_{\text{new}} < z_i^\mu < r_3\right\}, r_3\right\}
\end{aligned}$$

Then, all objective vectors within $H_{\text{new}} := [x_{\text{new}}, \overline{x}) \times [y_{\text{new}}, \overline{y}) \times [z_{\text{new}}, \overline{z})$ are weakly dominated by p_{new} but are not dominated by any vector in D_μ. Furthermore, H_{new} is not a null set (i.e. has a strictly positive Lebesgue measure) since $x_{\text{new}} < \overline{x}$, $y_{\text{new}} < \overline{y}$, and $z_{\text{new}} < \overline{z}$. This additional contribution makes $I_H(S)$ strictly larger than $I_H(D_\mu)$. □

Although the result is proven only for the 3-objective case, the generalization to an arbitrary number of objectives is straightforward though technical such that we refrain from presenting it here. Moreover, the same monotonicity directly follows for the weighted hypervolume indicator of [16] by replacing *Lebesgue* by *weighted Lebesgue*.

4 Geometrical Properties of the Hypervolume Contributions of Single Solutions in 3-Objective Problems

As we have seen so far, some basic results about optimal μ-distributions can be easily transferred to the 3-objective case. For some other results of [2], mainly regarding the exact distribution of μ solutions that maximize the hypervolume indicator, generalizations to higher dimensions are more difficult. The main reason is the fact that the optimal placement of a single solution is not determined by only two neighbors anymore as it is the case for bi-objective problems, see [2, Proposition 1]. As we will see in this section, the hypervolume contribution of a single solution in a 3-objective scenario can be influenced by all other solutions. The stated properties of the possible *shape* of a solution's hypervolume contribution will be used in the following section to generalize a non-trivial result of [2] about the absence of the extreme points in optimal μ-distributions to the 3-objective case.

Before we investigate the general shape of the hypervolume dominated by a single solution, let us define a geometrical object to be a *generalized cylinder* if there exists one coordinate axis for which all cross sections of the geometrical object, perpendicular to this axis, yield the same 2-dimensional shape and the corresponding projections of the cross sections along this axis on the coordinate system are the same. The usual cylinder with a circular cross section is one specific case of such a generalized cylinder when oriented along a coordinate axis. Figure 3 shows an example of a generalized cylinder with a steplike cross section. With this definition, we can state the first result about the volume solely dominated by a single solution in 3 objectives, see Fig. 2:

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.*

inria-00534906, version 1 - 10 Nov 2010

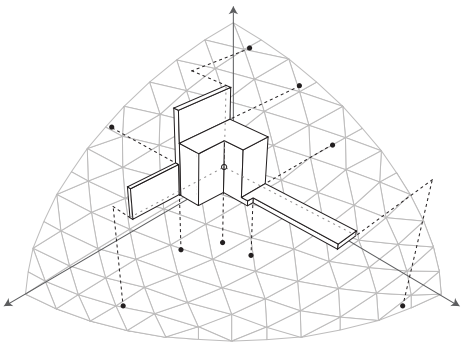

Fig. 2. Hypervolume contribution of a point (unfilled circle) on a three dimensional sphere function. The remaining nine points (black circles) all affect the shape of the contribution

Proof. The points that are weakly dominated by a specific 3-dimensional point $p = (x, y, z) \in \mathbb{R}^3$ and that weakly dominate the hypervolume's reference point form a cuboid $[x, r_1] \times [y, r_2] \times [z, r_3]$ with the point p as one corner and the reference point $r = (r_1, r_2, r_3)$ as the other end of its space diagonal. If we investigate now the points that are solely dominated by the point p, we have to subtract from this cuboid all solutions that are weakly dominated by other points $a = (a_1, a_2, a_3) \in A$, i.e., by the corresponding cuboids of which one corner is also the reference point. This gives the following set of points that are solely weakly dominated by p which can be obtained by deleting a general cylinder with steplike base area in each dimension from the cuboid associated to p: $([x, r_1] \times [y, r_2] \times [z, r_3]) \setminus \left(\bigcup_{\mathbf{a} \in A} [a_1, r_1] \times [a_2, r_2] \times [a_3, r_3]\right)$. $\square$

Note that the previous result does not only characterize the special shape of the hypervolume contribution of a single solution but also that this hypervolume contribution can be influenced by an arbitrary number of other solutions in a set A.

Interestingly, the shape of the space solely dominated by a single solution is becoming a generalized cylinder itself if we consider *extreme solutions* of a solution set A, see Fig. 3 (left). A solution a^i is thereby called extreme with respect to A (or extreme point of A) and objective $\mathcal{F}_i$, if no other solution in A has larger values in objective $\mathcal{F}_i$, i.e., $a^i \in \operatorname{argmax}\{a' \in A \mid \nexists a'' \in A \colon \mathcal{F}_i(a') < \mathcal{F}_i(a'')\}$. Note that extreme points are not unique in the 3-objective case in general and that their objective values do not always coincide with the values $x_{\max}$, $y_{\max}$, and $z_{\max}$, see, e.g., Fig. 1. We denote the obtained maximal values of extreme points in the three dimensions as $\overline{x}$, $\overline{y}$, and $\overline{z}$ respectively.

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. Let us consider w.l.o.g. only one extreme point in x-direction and its hypervolume contribution, i.e., a point $p_{\text{xmax}} = (\overline{x}, y, z)$ with the largest x-value $\overline{x}$ among μ solutions on the front. Without any other point, the hypervolume contribution of p_{xmax} would be again the cuboid from above with the point itself as one corner and the hypervolume's reference point as the other end of the cuboid's space diagonal starting

inria-00534906, version 1 - 10 Nov 2010

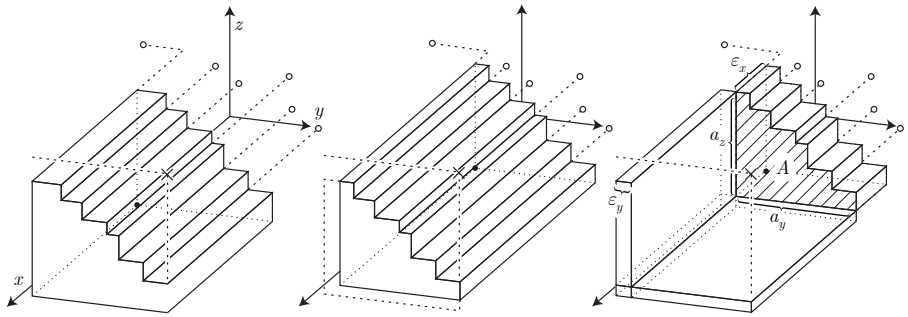

Fig. 3. Illustration of the hypervolume contribution of an extreme point in x-direction. **Left:** hypervolume solely dominated by the extreme point p_{xmax} which is depicted by a black circle; **Middle:** hypervolume solely dominated by the moved point $p_{\text{xmax}} - \varepsilon$; **Right:** illustration of benefit and deficit in hypervolume if we move the extreme point towards $p_{\text{xmax}} - \varepsilon$; in all three plots, the reference point is depicted by a cross and the remaining points influencing the extreme point's hypervolume contribution are depicted by unfilled circles

at p_{xmax}. Due to other incomparable solutions on the front, this cuboid is pruned in a specific way. To investigate how the hypervolume contribution of p_{xmax} is influenced by other points on the front, we consider the projection of all points to the y-z-plane, see Fig. 4. Two statements can be easily proven: (i) no point in the lower left region of p_{xmax} exists (otherwise it would be dominating p_{xmax} due to its better objective values in x-, y-, and z-direction) and (ii) all other solutions dominate a cuboid themselves and therefore cut this cuboid from the extreme point's cuboid (all solutions obviously dominate a volume that is a cuboid and the points are in addition not worse in x-direction and therefore their dominated volume is reaching in x-direction over the entire cuboid of p_{xmax}). This results in a volume solely dominated by p_{xmax}, that has a steplike projection and is a general cylinder in x-direction, see the leftmost plot of Fig. 3. □

5 Fronts For Which It Is Impossible to Obtain the Extreme Points

Given the knowledge about the shape of the hypervolume solely dominated by extreme solutions obtained above, we are able to generalize another result on optimal μ-distributions of [2] to the 3-objective case: There are cases where no finite reference point allows to have an extreme point of the Pareto front contained in optimal μ-distributions. In the 3-objective case, this corresponds to the cases where the (finite) partial derivative of the front at an extreme with respect to the first (second, third) axis is perpendicular to the first (second, third) axis:

Theorem 2. *Let f_{3d} be a continuous and differentiable function describing the front with gradient ∇f_{3d} continuous in $[x_{\min}, x_{\max}] \times [y_{\min}, y_{\max}] \times [z_{\min}, z_{\max}]$. If the gradient $\nabla f_{3d}(\overline{x}, y, z) = (\partial_1 f_{3d}(\overline{x}, y, z), \partial_2 f_{3d}(\overline{x}, y, z), \partial_3 f_{3d}(\overline{x}, y, z))$ of the front at an extreme point $(\overline{x}, y, z)$ (at an extreme point $(x, \overline{y}, z)$, or at $(x, y, \overline{z})$) is finite, i.e., the single components of the gradient are finite, and the gradient is perpendicular to the*

inria-00534906, version 1 - 10 Nov 2010

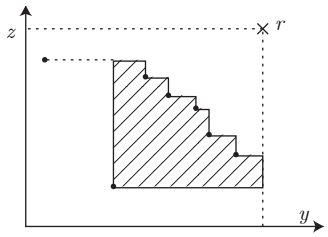

Fig. 4. Projection of all points to the y-z-plane

x-axis (y-, z-axis), i.e., if $\nabla f_{3d}(\overline{x}, y, z) \cdot (1,0,0) = 0$ ($\nabla f_{3d}(x, \overline{y}, z) \cdot (0,1,0) = 0$, $\nabla f_{3d}(x, y, \overline{z}) \cdot (0,0,1) = 0$), the corresponding extreme point is not included in any optimal μ-distribution with $\mu \geq 1$.

Proof. W.l.o.g., we consider only the case of the extreme point $p_{\text{xmax}} := (\overline{x}, y, z)$ where $\nabla f_{3d}(p_{\text{xmax}}) \cdot (1,0,0) = 0$ and therefore $\partial_1 f_{3d}(p_{\text{xmax}}) = 0$. The proof idea is similar to the bi-objective case in [2]: we consider the hypervolume that we gain and the hypervolume that we lose if we move the extreme point p_{xmax} towards larger values of y and z. To this end, we use the notations of Fig. 3. In particular, we move p_{xmax} by a small value $\varepsilon_x > 0$ in x direction towards smaller x-values and at the same time parallel to the plane defined by the x-axis and the gradient $\nabla f_{3d}(p_{\text{xmax}})$ along the front:

$$p_{\text{xmax}} - \varepsilon = p_{\text{xmax}} - \begin{pmatrix} \varepsilon_x \\ \varepsilon_y \\ \varepsilon_z \end{pmatrix} = p_{\text{xmax}} - \begin{pmatrix} \varepsilon_x \\ 0 \\ 0 \end{pmatrix} + \nu \begin{pmatrix} 0 \\ \partial_2 f(p_{\text{xmax}}) \\ \partial_3 f(p_{\text{xmax}}) \end{pmatrix} \tag{1}$$

where $\nu \in \mathbb{R}$ and the last equality follows from the assumption that the gradient $\nabla f(p_{\text{xmax}})$ is perpendicular to the x-axis. Note that at least one of the values ε_y and ε_z has to be negative since all points on the Pareto front are incomparable and a point with all three objectives smaller than p_{xmax} would therefore not lie on the front.

If we choose ε_x small enough, i.e., as long as there is no other point $p' = (x', y', z')$ with $x' > \overline{x} - \varepsilon_x$, $y' > y$, and $z' > z$ among the μ solutions under consideration, the hypervolume contribution of the moved extreme point keeps its shape, see Fig. 3. According to the rightmost plot of Fig. 3, we denote the area of the y-z-projection of the new point's hypervolume contribution as A, by a_z the height of this area in z-direction, and by a_y the length of this area in y-direction. Then, the benefit and deficit in hypervolume if we move the extreme point from p_{xmax} to $p_{\text{xmax}} - \varepsilon$ can be written as

$$\text{benefit: } \varepsilon_x \cdot A \qquad \text{deficit: } (\varepsilon_y \cdot a_z + \varepsilon_z \cdot a_y + \varepsilon_y \cdot \varepsilon_z) \cdot (r_1 - \overline{x}) \ .$$

Now, it remains to be shown that the ratio between deficit and benefit goes to zero when ε_x converges to zero. To this end, we decompose the ratio R between deficit and benefit as

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

Because A is lower bounded by a constant and a_x and a_z are upper bounded by a constant, showing that $\lim_{\varepsilon_x \to 0} \frac{\varepsilon_y}{\varepsilon_x} = 0$ and $\lim_{\varepsilon_x \to 0} \frac{\varepsilon_z}{\varepsilon_x} = 0$ will directly prove that

inria-00534906, version 1 - 10 Nov 2010

the ratio R converges to zero. W.l.o.g., we are only going to prove $\lim_{\varepsilon_x \to 0} \frac{\varepsilon_y}{\varepsilon_x} = 0$ in the following. The proof of $\lim_{\varepsilon_x \to 0} \frac{\varepsilon_z}{\varepsilon_x} = 0$ can be done in the same way by exchanging the roles of ε_y and ε_z.

Assuming $\partial_2 f_{3d}(p_{\text{xmax}}) \neq 0$ (otherwise, $\lim_{\varepsilon_x \to 0} \frac{\varepsilon_y}{\varepsilon_x} = 0$ follows directly), we know from Eq. 1 that $\nu = \frac{\varepsilon_y}{\partial_2 f_{3d}(p_{\text{xmax}})}$ and therefore that

$$\varepsilon_z = \varepsilon_y \frac{\partial_3 f_{3d}(p_{\text{xmax}})}{\partial_2 f_{3d}(p_{\text{xmax}})} \quad . \tag{2}$$

Since the gradient of f_{3d} is continuous within the cuboid restricting the front, we can expand $f_{3d}(p_{\text{xmax}} - \varepsilon)$ with the Taylor formula as $f_{3d}(p_{\text{xmax}} - \varepsilon) = f_{3d}(p_{\text{xmax}}) - \nabla f_{3d}(p_{\text{xmax}}) \cdot \varepsilon + \mathcal{O}(||\varepsilon||^2)$ which indicates that $\nabla f_{3d}(p_{\text{xmax}}) \cdot \varepsilon - \mathcal{O}(||\varepsilon||^2) = 0$ for any $\varepsilon \geq 0$ as f_{3d} equals zero for all points on the front by definition. From $\nabla f_{3d}(p_{\text{xmax}}) \cdot \varepsilon - \mathcal{O}(||\varepsilon||^2) = 0$ we can conclude that $\lim_{\varepsilon \to 0} \nabla f_{3d}(p_{\text{xmax}}) \cdot \varepsilon = \lim_{\varepsilon \to 0} \mathcal{O}(||\varepsilon||^2)$ and even $\lim_{\varepsilon \to 0} (\nabla f_{3d}(p_{\text{xmax}}) \cdot \varepsilon / ||\varepsilon||) = \lim_{\varepsilon \to 0} \mathcal{O}(||\varepsilon||) = 0$. Since $\partial_1 f_{3d}(p_{\text{xmax}}) = 0$ and the other partial derivatives $\partial_2 f_{3d}(p_{\text{xmax}})$ and $\partial_2 f_{3d}(p_{\text{xmax}})$ are finite and constant, the previous equation can be rewritten as

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

Using (2) in the previous equation and factorizing the numerator and denominator by ε_y we obtain

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

which implies that $\lim_{\varepsilon \to 0} \frac{\varepsilon_y}{\varepsilon_x} = 0$. □

Remark. Note that the case covered by the previous theorem is not an artificial case but obtained for some well-known test problems, e.g., DTLZ2–4 [8] and WFG4–9 [12]. It also covers the bi-objective case proven in [2] which, however, used a slightly different notation due to a simpler description of the front shape.

6 Conclusions

Obtaining optimal μ-distributions for a certain quality indicator I, i.e., a set of μ solutions maximizing I, is the optimization goal of several indicator-based multiobjective evolutionary algorithms [2]. In particular, the hypervolume indicator, among others employed in the SMS-EMOA [6] and the MO-CMA-ES [13], received interest as a selection criterion in multiobjective algorithms due to its property of being a refinement of the Pareto dominance relation. However, theoretical investigations of optimal μ-distributions for the hypervolume indicator are rare and limited to the bi-objective case so far [9, 5, 2, 1, 11, 7].

inria-00534906, version 1 - 10 Nov 2010

Here, we obtain first theoretical results on optimal μ-distributions for the hypervolume indicator in 3-objective scenarios. It turns out that the hypervolume contribution of a single point has a specific shape that is not as simple as in bi-objective problems anymore—indicating that *all* solutions have an influence on where to optimally place a solution instead of only *two* solutions in the bi-objective case. Besides generalizations of basic statements of [2] to the 3-objective case, we prove in particular that also in 3-objective problems there are situations where no finite reference point can ensure the extreme solutions of the Pareto front within an optimal μ-distribution.

Acknowledgments This work has been in part supported by the French national research agency (ANR) within the SYSCOMM project ANR-08-SYSC-017 and within the COSINUS project ANR-08-COSI-007-12.

References

1. A. Auger, J. Bader, D. Brockhoff, and E. Zitzler. Investigating and Exploiting the Bias of the Weighted Hypervolume to Articulate User Preferences. In *Genetic and Evolutionary Computation Conference (GECCO 2009)*, pages 563–570. ACM, 2009.
2. 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. ACM, 2009.
3. J. Bader. *Hypervolume-Based Search for Multiobjective Optimization: Theory and Methods*. PhD thesis, ETH Zurich, Switzerland, 2010.
4. R. G. Bartle. *The Elements of Integration and Lebesgue Measure*. Wiley, 1995.
5. N. Beume, C. M. Fonseca, M. Lopez-Ibanez, L. Paquete, and J. Vahrenhold. On the Complexity of Computing the Hypervolume Indicator. *IEEE T. Evolut. Comput.*, 13(5):1075–1082, 2009.
6. N. Beume, B. Naujoks, and M. Emmerich. SMS-EMOA: Multiobjective Selection Based on Dominated Hypervolume. *Eur. J. Oper. Res.*, 181:1653–1669, 2007.
7. K. Bringmann and T. Friedrich. The Maximum Hypervolume Set Yields Near-optimal Approximation. In *Genetic and Evolutionary Computation Conference (GECCO 2010)*. ACM, 2010. to appear.
8. K. Deb, L. Thiele, M. Laumanns, and E. Zitzler. Scalable Test Problems for Evolutionary Multi-Objective Optimization. In *Evolutionary Multiobjective Optimization: Theoretical Advances and Applications*, chapter 6, pages 105–145. Springer, 2005.
9. M. Emmerich, A. Deutz, and N. Beume. Gradient-Based/Evolutionary Relay Hybrid for Computing Pareto Front Approximations Maximizing the S-Metric. In *Hybrid Metaheuristics*, pages 140–156. Springer, 2007.
10. M. Fleischer. The Measure of Pareto Optima. Applications to Multi-Objective Metaheuristics. In *Conference on Evolutionary Multi-Criterion Optimization (EMO 2003)*, pages 519–533. Springer, 2003.
11. T. Friedrich, C. Horoba, and F. Neumann. Multiplicative Approximations and the Hypervolume Indicator. In *Genetic and Evolutionary Computation Conference (GECCO 2009)*, pages 571–578. ACM, 2009.
12. S. Huband, P. Hingston, L. Barone, and L. While. A Review of Multiobjective Test Problems and a Scalable Test Problem Toolkit. *IEEE T. Evolut. Comput.*, 10(5):477–506, 2006.
13. C. Igel, N. Hansen, and S. Roth. Covariance Matrix Adaptation for Multi-objective Optimization. *Evolutionary Computation*, 15(1):1–28, 2007.
14. W. Rudin. *Principles of Mathematical Analysis*. McGraw-Hill, third edition, 1976.

inria-00534906, version 1 - 10 Nov 2010

15. D. A. Van Veldhuizen and G. B. Lamont. Multiobjective evolutionary algorithm test suites. In *Symposium on Applied Computing*, pages 351–357. ACM, 1999.
16. E. Zitzler, D. Brockhoff, and L. Thiele. The Hypervolume Indicator Revisited: On the Design of Pareto-compliant Indicators Via Weighted Integration. In *Conference on Evolutionary Multi-Criterion Optimization (EMO 2007)*, pages 862–876. Springer, 2007.
17. E. Zitzler and S. Künzli. Indicator-Based Selection in Multiobjective Search. In *Conference on Parallel Problem Solving from Nature (PPSN VIII)*, pages 832–842. Springer, 2004.
18. E. Zitzler and L. Thiele. Multiobjective Evolutionary Algorithms: A Comparative Case Study and the Strength Pareto Approach. *IEEE T. Evolut. Comput.*, 3(4):257–271, 1999.
19. E. Zitzler, L. Thiele, and J. Bader. On Set-Based Multiobjective Optimization. *IEEE T. Evolut. Comput.*, 14(1):58–79, 2009.

inria-00534906, version 1 - 10 Nov 2010