Optimal μ-Distributions for the Hypervolume Indicator for Problems With Linear Bi-Objective Fronts: Exact and Exhaustive Results*

Dimo Brockhoff**

TAO Team, INRIA Saclay, LRI, Paris Sud University, 91405 Orsay Cedex, France
dimo.brockhoff@inria.fr

Abstract. To simultaneously optimize multiple objective functions, several evolutionary multiobjective optimization (EMO) algorithms have been proposed. Nowadays, often set quality indicators are used when comparing the performance of those algorithms or when selecting "good" solutions during the algorithm run. Hence, characterizing the solution sets that maximize a certain indicator is crucial—complying with the optimization goal of many indicator-based EMO algorithms. If these optimal solution sets are upper bounded in size, e.g., by the population size μ, we call them *optimal μ-distributions*. Recently, optimal μ-distributions for the well-known hypervolume indicator have been theoretically analyzed, in particular, for bi-objective problems with a linear Pareto front. Although the exact optimal μ-distributions have been characterized in this case, not all possible choices of the hypervolume's reference point have been investigated. In this paper, we revisit the previous results and rigorously characterize the optimal μ-distributions also for all other reference point choices. In this sense, our characterization is now exhaustive as the result holds for any linear Pareto front and for any choice of the reference point and the optimal μ-distributions turn out to be always unique in those cases. We also prove a tight lower bound (depending on μ) such that choosing the reference point above this bound ensures the extremes of the Pareto front to be always included in optimal μ-distributions.

Keywords: multiobjective optimization, hypervolume indicator, optimal μ-distributions, theory

1 Introduction

Many evolutionary multiobjective optimization (EMO) algorithms have been proposed to tackle optimization problems with multiple objectives. The most recent ones employ quality indicators within their selection in order to (i) directly incorporate user preferences into the search [1, 16] and/or to (ii) avoid cyclic behavior of the current population

* This is an updated author version of the SEAL'2010 paper published by Springer Verlag which corrects the following errors. The final publication is available at www.springerlink.com.

ERRATUM: In Theorems 1, 2, and 3, it should read "the unique optimal μ-distribution $(x_1^\mu, \ldots, x_\mu^\mu)$".

** Currently at LIX, École Polytechnique, Palaiseau, France

[15, 18]. In particular the hypervolume indicator [17] is of interest here and due to its refinement property [18] employed in several EMO algorithms [4, 6, 14]. The hypervolume indicator assigns a set of solutions the "size of the objective value space which is covered" and at the same time is bounded by the indicator's reference point [17]. Although maximizing the hypervolume indicator, according to its refinement property, results in finding Pareto-optimal solutions only [10], the question arises which of these points are favored by hypervolume-based algorithms. In other words, we are interested in the optimization goal of hypervolume-based algorithms with a fixed population size μ, i.e., in finding a set of μ solutions with the highest hypervolume indicator value among all sets with μ solutions. Also in performance assessment, the hypervolume is used quite frequently [19]. Here, knowing the set of points maximizing the hypervolume is crucial as well. On the one hand, it allows to evaluate whether hypervolume-based algorithms really converge towards their optimization goal on certain test functions. On the other hand, only the knowledge of the best hypervolume value achievable with μ solutions allows to compare algorithms in an absolute manner similar to the state-of-the-art approach of benchmarking single-objective continuous optimization algorithms in the horizontal-cut view scenario, see [12, appendix] for details.

Theoretical investigations of the sets of μ points maximizing the hypervolume indicator—also known under the term of *optimal μ-distributions* [2]—have been started only recently. Although quite strong, i.e., very general, results on optimal μ-distributions are known [2, 7], most of them are approximation or limit results in order to study a wide range of problem classes. The only exact results consider problems with very specific Pareto fronts, namely linear fronts that can be described by a function $f : x \in [x_{\min}, x_{\max}] \mapsto \alpha x + \beta$ where $\alpha < 0$ and $\beta \in \mathbb{R}$ in the bi-objective case [2, 5, 9] or fronts that can be expressed as $f : x \in [1, c] \mapsto c/x$ with $c > 1$ [11].

The main scope of this paper is to revisit the results on optimal μ-distributions for bi-objective problems with linear Pareto fronts and to consider all conditions under which the exact optimal μ-distributions have not been characterized yet. The result is both exact and exhaustive, in the sense that a single formula is proven that characterizes the unique optimal μ-distribution for any choice of the hypervolume indicator's reference point and for any $\mu \geq 2$, covering also the previously known cases. It turns out that the specific case of $\mu = 2$ complies with a previous results of [2] and that for all linear front shapes, the optimal μ-distributions are always unique.

Before we present our results in Sec. 5–7, we introduce basic notations and definitions in Sec. 2, define and discuss the problem of finding optimal μ-distributions in Sec. 3 in more detail, and give an extensive overview of the known results in Sec. 4.

2 Preliminaries

Without loss of generality (w.l.o.g.), we consider bi-objective minimization problems where a vector-valued function $\mathcal{F} : X \to \mathbb{R}^2$ has to be minimized with respect to the weak Pareto dominance relation $\preceq$. We say a solution $x \in X$ is weakly dominating another solution $y \in X$ ($x \preceq y$) iff $\mathcal{F}_1(x) \leq \mathcal{F}_1(y)$ and $\mathcal{F}_2(x) \leq \mathcal{F}_2(y)$ where $\mathcal{F} = (\mathcal{F}_1, \mathcal{F}_2)$. We also say $x \in X$ is dominating $y \in X$ ($x \prec y$) if $x \preceq y$ but $y \not\preceq x$. The set of nondominated solutions is the so-called Pareto set $\mathcal{P}_s = \{x \in X \mid \nexists y \in X :$

$y \prec x\}$ and its image $\mathcal{F}(\mathcal{P}_s)$ in objective space is called Pareto front. Note that, to keep things simple, we make an abuse of terminology throughout the paper and use the term *solution* both for a point x in the decision space X and for its corresponding objective vector $\mathcal{F}(x) \in \mathbb{R}^2$. Moreover, we also define the orders $\preceq$ and $\prec$ on objective vectors.

In order to optimize multiobjective optimization problems like the bi-objective ones considered here, several recent EMO algorithms aim at optimizing the *hypervolume indicator* [17], a set quality indicator $I_H(A, r)$ that assigns a set A the Lebesgue measure λ of the set of solutions that are weakly dominated by solutions in A but that at the same time weakly dominate a given reference point $r \in \mathbb{R}^2$, see Fig. 1:

$$I_H(A, r) = \lambda\left(\{z \in \mathbb{R}^2 \,|\, \exists a \in A : f(a) \preceq z \preceq r\}\right) \tag{1}$$

The hypervolume indicator has the nice property of being a refinement of the Pareto dominance relation [18]. This means that maximizing the hypervolume indicator is equivalent to obtaining solutions in the Pareto set only [10]. However, it is more interesting to know *where* the solutions maximizing the hypervolume lie on the Pareto front if we restrict the size of the sets A to let us say, the population size μ. This set of μ points maximizing the hypervolume indicator among all sets of μ points is known under the term *optimal μ-distribution* [2] and finding an optimal μ-distribution coincides with the optimization goal of hypervolume-based algorithms with fixed population size.

To investigate optimal μ-distributions in this paper, we assume the Pareto front to be given by a function $f : \mathbb{R} \to \mathbb{R}$ and two values $x_{\min}, x_{\max} \in \mathbb{R}$ such that all points on the Pareto front have the form $(x, f(x))$ with $x \in [x_{\min}, x_{\max}]$. In case of a linear Pareto front, $f(x) = \alpha x + \beta$ for $\alpha, \beta \in \mathbb{R}$, see Fig. 1 for an example. W.l.o.g, we assume that $x_{\min} = 0$ and $\beta > 0$ in the remainder of the paper—otherwise, a simple linear transformation brings us back to this case. Moreover, $\alpha < 0$ follows from minimization. Note also that under not too strong assumptions on the Pareto front, and in particular for linear fronts, optimal μ-distributions always exist, see [2].

3 Problem Statement

In case of a linear Pareto front described by the function $f(x) = \alpha x + \beta$ ($\alpha < 0, \beta \in \mathbb{R}$), finding the optimal μ-distribution for the hypervolume indicator with reference point $r = (r_1, r_2)$ can be written as finding the minimum of the function

$$\begin{aligned} I_H(x_1, \ldots, x_\mu) &= \sum_{i=1}^{\mu} (x_{i+1} - x_i)\left(f(x_0) - f(x_i)\right) = \sum_{i=1}^{\mu} (x_{i+1} - x_i)\left(\alpha x_0 - \alpha x_i\right) \\ &= \alpha \sum_{i=1}^{\mu} \left[(x_i)^2 + x_0 x_{i+1} - x_0 x_i - x_i x_{i+1}\right] \\ &\text{with } x_{\min} \le x_i \le x_{\max} \text{ for all } 1 \le i \le \mu \end{aligned} \tag{2}$$

where we define $x_{\mu+1} = r_1$ and $x_0 = f^{-1}(r_2)$ [2], Fig. 1. According to [2], we denote the x-values of the optimal μ-distribution, maximizing (2), as $x_1^\mu \ldots x_\mu^\mu$. Although the term in (2) is quadratic in the variables $x_0, \ldots, x_{\mu+1}$, and therefore, in principle, solvable analytically, the restrictions of the variables to the interval $[x_{\min}, x_{\max}]$ makes it difficult to solve the problem. In the following, we therefore investigate the minima of (2) depending on the choice of r_1 and r_2 with another approach: we use the necessary condition for optimal μ-distributions of [2, Proposition 1] and apply it to linear fronts while the restriction of the variables to $[x_{\min}, x_{\max}]$ are handled "by hand".

4 Overview of Recent and New Results

Characterizing optimal μ-distributions for the hypervolume indicator has been started only recently but the number of results is already quite extensive, see for example [1–5, 7, 9–11]. Here, we restate, to the best of our knowledge, all previous results that relate to linear Pareto fronts and point out which problems are still open.

Besides the proof that maximizing the hypervolume indicator yields Pareto-optimal solutions [10], the authors of [5] and [9] were the first to investigate optimal μ-distributions for linear fronts. Under the assumption that the extreme points $(0, \beta)$ and $(x_{\max}, 0)$ are included in the optimal μ-distribution, it was shown for linear fronts with $\alpha = -1$ that neighbored points within a set maximizing the hypervolume are equally spaced. However, the result does not state where the leftmost and rightmost point of the optimal μ-distribution have to be placed in order to maximize the hypervolume and it has been shown later [2] that the assumption about the extreme points does not always hold.

The first results without assuming the positions of the leftmost and rightmost point have been proven in [2] where the result is based on a more general necessary condition about optimal μ-distributions for the hypervolume indicator. In particular, [2] presents the exact distribution of μ points maximizing the hypervolume indicator when the reference point is chosen close to the Pareto front (region I in Fig. 1, cp. [2, Theorem 5]) or far away from the front (region IX in Fig. 1, cp. [2, Theorem 6]). In the former case, both extreme points of the front do not dominate the reference point and the (in this case unique) optimal μ-distribution reads

$$x_i^\mu = f^{-1}(r_2) + \frac{i}{\mu+1} \cdot (r_1 - f^{-1}(r_2)) \quad . \tag{3}$$

In the latter case, the reference point is chosen far enough such that—independent of the reference point and μ—both extreme points are included in an optimal μ-distribution [1] and the (again unique) optimal μ-distributions can be expressed as

$$x_i^\mu = x_{\min} + \frac{i-1}{\mu-1}\left(x_{\max} - x_{\min}\right) \quad . \tag{4}$$

Note that the region IX, corresponding to choices of the reference point within Theorem 6 of [2] does not depend on μ but on a lower bound on the reference point to ensure that both extremes are included in the optimal μ-distribution. Recently, a limit result has been proven [3] which shows that the lower bound of [2, Theorem 6] converges to the nadir point[2] if μ goes to infinity but the result does not state how fast (in μ) the nadir point is approached. Clearly, choosing the reference point within the other regions

[1] Which is proven to be true for $r_1 > 2x_{\max}$ and $r_2 > 2\beta$ in another general theorem [2].

[2] In case of a linear front as defined above, the nadir point equals $n = (x_{\max}, f(x_{\min}))$.

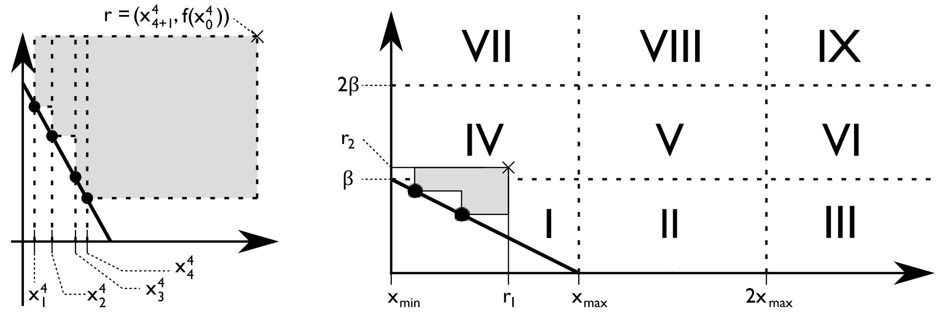

Fig. 1. Left: Illustration of the hypervolume indicator $I_H(A, r)$ (gray area). **Right:** Optimal μ-distributions and the choice of the reference point for linear fronts of shape $y = \alpha x + \beta$. Up-to-now, theoretical results are only known if the reference point is chosen within the regions I and IX [2]. Exemplary, the optimal 2-distribution (circles) is shown when choosing the reference point (cross) within region IV.

II–VIII in Fig. 1 is possible and the question arises how the reference point influences the optimal μ-distributions in these uninvestigated cases as well. The answer to this question is the main focus of this paper.

5 If the Reference Point is Dominated by Only the Right Extreme

As a first new result, we consider choosing the reference point within the regions II or III of Fig. 1. Here, the left extreme cannot be included in an optimal μ-distribution as it is never dominating the reference point and thereby always has a zero hypervolume contribution. Thus, the proof of the optimal μ-distribution has to consider only the restrictions of the μ points at the right extreme. Moreover, the uniqueness of the optimal μ-distribution in the cases II and III follows directly from case I.

Theorem 1. *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 $[0, x_{\max} = -\frac{\beta}{\alpha}]$. If $r_2 \leq \beta$ and $r_1 \geq x_{\max}$ (cases II and III), the unique optimal μ-distribution $(x_1^\mu, \ldots, x_\mu^\mu)$ for the hypervolume indicator I_H with reference point (r_1, r_2) can be described by*

$$x_i^\mu = f^{-1}(r_2) + \frac{i}{\mu+1}\left(\min\left\{r_1, \frac{\mu+1}{\mu}x_{\max} - \frac{f^{-1}(r_2)}{\mu}\right\} - f^{-1}(r_2)\right) \quad . \tag{5}$$

Proof. According to (3) and assuming no restrictions of the solutions on the linear front $\alpha x + \beta$ with $x \in \mathbb{R}$, the optimal μ-distribution would be given by $x_i^\mu = f^{-1}(r_2) + \frac{i}{\mu+1} \cdot (r_1 - f^{-1}(r_2))$ where the x_i^μ are possibly lying outside the interval $[0, x_{\max}]$. However, as long as r_1 is chosen such that $x_\mu^\mu \leq x_{\max}$, we can use (3) for describing

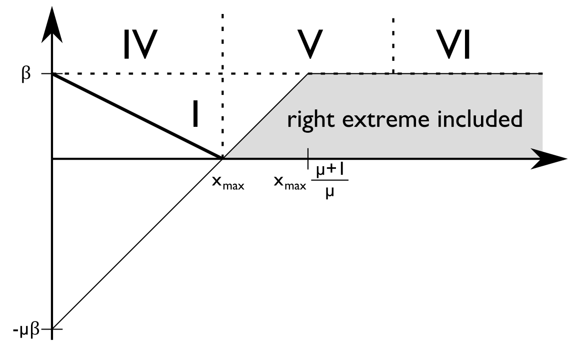

Fig. 2. When choosing the reference point within regions II and III, we prove that the right extreme is included in optimal μ-distributions if the reference point is chosen within the gray shaded area right of the line $y = -\alpha\mu x - \mu\beta$, see Corollary 1. The picture corresponds to $\mu = 2$.

the optimal μ-distributions, i.e., in the case that

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

$$\Leftrightarrow r_1 \leq \frac{\mu+1}{\mu} x_{\max} - \frac{f^{-1}(r_2)}{\mu} \left(= \frac{-r_2 - \beta\mu}{\alpha\mu} \right) \quad . \tag{6}$$

With larger r_1, the optimal μ-distribution does not change any further (only the hypervolume contribution of x_μ^μ increases linearly with r_1), i.e., we can rewrite (3) as (5). □

The previous theorem allows us also a more precise statement of when the right extreme is included in optimal μ-distributions than the statement in [2].

Corollary 1. *In case that $r_2 \leq \beta$ and $r_1 \geq \frac{\mu+1}{\mu} x_{\max} - \frac{f^{-1}(r_2)}{\mu}$, the right extreme point $(x_{\max}, 0)$ is included in all optimal μ-distributions for the front $\alpha x + \beta$.* □

Note that the choice of r_1 to guarantee the right extreme in optimal μ-distributions depends both on μ and r_2 here whereas the (not so tight) bound for r_1 to ensure the right extreme proven in [2] equals $2x_{\max}$. This is independent of μ and coincides with the new (tighter) result if $\mu = 2$ and $r_2 = \beta$. Figure 2 illustrates the region for which, if the reference point is chosen within, the right extreme is always included in an optimal μ-distribution. Compare also to the old result of [2] which states this inclusion of the right extreme only in case the reference point is chosen in region IX of Fig. 1. The description of the line $y = -\alpha\mu x - \mu\beta$ where choosing the reference point to the right of it ensures the right extreme in the optimal μ-distribution results from writing r_2 within $r_1 = \frac{\mu+1}{\mu} x_{\max} - \frac{f^{-1}(r_2)}{\mu}$ as a function of r_1.

6 If the Reference Point is Dominated by Only the Left Extreme

Obviously, the two cases IV and VII of Fig. 1 are symmetrical to the cases II and III where mainly the left extreme and the reference point's coordinate r_2 take the roles of the right extreme and the coordinate r_1 respectively from the previous proof.

Theorem 2. *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 $[0, x_{\max} = -\frac{\beta}{\alpha}]$. If $r_1 \leq x_{\max}$ and $r_2 \geq \beta$ (cases IV and VII), the unique optimal μ-distribution $(x_1^\mu, \ldots, x_\mu^\mu)$ for the hypervolume indicator I_H with reference point (r_1, r_2) can be described by*

$$x_i^\mu = f^{-1}\Big(\min\Big\{r_2, \tfrac{\mu+1}{\mu}\beta - \tfrac{f(r_1)}{\mu}\Big\}\Big) + \frac{i}{\mu+1}\Big(r_1 - f^{-1}\Big(\min\Big\{r_2, \tfrac{\mu+1}{\mu}\beta - \tfrac{f(r_1)}{\mu}\Big\}\Big)\Big)\ . \quad (7)$$

Proof. The proof is similar to the one of Theorem 1: As in case I, we can write the optimal μ-distribution according to (3) except that we have to ensure that $x_1^\mu \geq x_{\min} = 0$. This is equivalent to $f^{-1}(r_2) + \frac{1}{\mu+1}\left(r_1 - f^{-1}(r_2)\right) \geq 0$ or $\frac{r_2-\beta}{\alpha} + \frac{1}{\mu+1}\left(r_1 - \frac{r_2-\beta}{\alpha}\right) \geq 0$ or $\frac{r_2-\beta}{\alpha} + \frac{\alpha r_1 - r_2 + \beta}{(\mu+1)\alpha} \geq 0$. With $\alpha < 0$, this gives $(\mu+1)r_2 - (\mu+1)\beta + \alpha r_1 - r_2 + \beta \leq 0$ and finally $r_2 \leq \frac{(\mu+1)\beta - (\alpha r_1 + \beta)}{\mu} = \frac{\mu+1}{\mu}\beta - \frac{f(r_1)}{\mu}$ such that (3) becomes (7). □

7 General Result for All Cases I–IX

By combining the above results, we can now characterize the optimal μ-distributions also for the other cases V, VI, VII, and IX and give a general description of optimal μ-distributions for problems with bi-objective linear fronts, given any $\mu \geq 2$ and any meaningful choice of the reference point[3].

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 $(x_1^\mu, \ldots, x_\mu^\mu)$ for the hypervolume indicator I_H with reference point $(r_1, r_2) \in \mathbb{R}^2_{>0}$ can be described by*

$$x_i^\mu = f^{-1}(F_l) + \frac{i}{\mu+1}\left(F_r - f^{-1}(F_l)\right) \quad (8)$$

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

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

Proof. Again, the optimal μ-distribution would be given by (3) if we prolongate the front linearly outside the interval $[x_{\min}, x_{\max}]$ and therefore, no restrictions on the x_i^μ would hold. However, the points x_i^μ are restricted to $[x_{\min}, x_{\max}]$ and therefore (since we assume $x_i^\mu < x_{i+1}^\mu$) we have to ensure that both $x_1^\mu \geq x_{\min} = 0$ and $x_\mu^\mu \leq x_{\max} = -\beta/\alpha$ hold. According to the above proofs, the former is equivalent to

$$r_2 \leq \frac{\mu+1}{\mu}\beta - \frac{f(r_1)}{\mu} \quad (9)$$

[3] Choosing the reference point such that it weakly dominates a Pareto-optimal point does not make sense as no feasible solution would have a positive hypervolume.

and the latter is equivalent to

$$r_1 \leq \frac{\mu+1}{\mu} x_{\max} - \frac{f^{-1}(r_2)}{\mu} \tag{10}$$

however, with restrictions on r_1 ($r_1 \leq x_{\max}$) and $r_2 \leq \beta$ respectively which we do not have here. As long as both (9) and (10) hold as in the white area in Fig. 3, i.e., no constraint is violated, (3) can be used directly to describe the optimal μ-distribution as in region I. To cover all other cases, we could, at first sight, simply combine the results for the cases II, III, IV, and VII from above and use

$$F_l^* = \min\left\{r_2, \frac{\mu+1}{\mu}\beta - \frac{f(r_1)}{\mu}\right\} \quad \text{and} \quad F_r^* = \min\left\{r_1 \frac{\mu+1}{\mu} x_{\max} - \frac{f^{-1}(r_2)}{\mu}\right\}$$

as the extremes influencing the set $x_i^{\mu,*} = F_l^* + \frac{i}{\mu+1}(F_r^* - F_l^*)$. However, r_1 and r_2 are unrestricted and thus, F_l^* and F_r^* can become too large such that the points $x_i^{\mu,*}$ lie outside the feasible front part $[x_{\min}, x_{\max}]$. To this end, we compute where the two constraints (9) and (10) meet, i.e., what is the smallest possible reference point that results in having both extremes in the optimal μ-distribution. This point is depicted as the lower left point of the dark gray area in Fig. 3.

By combining the equalities in (9) and (10) which is equivalent to $r_2 = -\alpha\mu r_1 - \beta\mu$ (see end of Sec. 5), we obtain

$$r_2 = \tfrac{\mu+1}{\mu}\beta - \tfrac{f(r_1)}{\mu} = \tfrac{\mu+1}{\mu}\beta - \tfrac{\alpha r_1+\beta}{\mu} = -\alpha\mu r_1 - \beta\mu \ \text{ or } \ r_1 = -\tfrac{\beta}{\alpha}\tfrac{\mu}{\mu-1} = \tfrac{\mu}{\mu-1}x_{\max}$$

and thus $r_2 = \frac{\mu+1}{\mu}\beta - \frac{f(\frac{\mu}{\mu-1}x_{\max})}{\mu} = \frac{\mu}{\mu-1}\beta$. Hence, if we choose the reference point $r = (r_1, r_2)$ such that $r_1 \geq \frac{\mu}{\mu-1}x_{\max}$ and $r_2 \geq \frac{\mu}{\mu-1}\beta$, both extremes will be included in the optimal μ-distribution $x_i^{\mu} = F_l^{\text{extr}} + \frac{i}{\mu+1}(F_r^{\text{extr}} - F_l^{\text{extr}})$ with $F_l^{\text{extr}} = \frac{\mu}{\mu-1}\beta$ and $F_r^{\text{extr}} = \frac{\mu}{\mu-1}x_{\max}$. With this result, we know that, independent of r_2, the right extreme is included if $r_1 \geq \frac{\mu}{\mu-1}x_{\max}$ (if the leftmost extreme is not included, r_2 must be smaller than $\frac{\mu}{\mu-1}\beta$ and in this case $r_1 \geq \frac{\mu+1}{\mu}x_{\max}$ ensures that it is also greater or equal to $\frac{\mu+1}{\mu}x_{\max} - \frac{f^{-1}(r_2)}{\mu}$). The same can be said for the left extreme, which is included in an optimal μ-distribution whenever $r_2 \geq \frac{\mu}{\mu-1}\beta$. The optimal μ-distribution for those cases are the same than the optimal μ-distributions if we restrict r_1 and r_2 to be at most $\min\{\frac{\mu+1}{\mu}x_{\max} - \frac{1}{\mu}f^{-1}(r_2), \frac{\mu}{\mu-1}x_{\max}\}$, and $\min\{\frac{\mu+1}{\mu}\beta - \frac{1}{\mu}f(r_1), \frac{\mu}{\mu-1}\beta\}\}$ respectively, i.e., to the cases where the reference point is lying on the boundary of the white region of Fig. 3 and having one or even both extremes included in the optimal μ-distributions. In those cases, (3) can be used again for characterizing the optimal μ-distribution as the constraints on the x_i^{μ} are fulfilled. Using the mentioned restrictions on r_1 and r_2 results in the theorem. □

Note that the previous proof gives a tighter bound for how to choose the reference point $r = (r_1, r_2)$ in order to obtain the extremes in comparison to the old result in [2]: The former result states that whenever r_1 is chosen strictly larger than $2x_{\max}$ and r_2 is chosen strictly larger than 2β, both extremes are included in an optimal μ-distribution in the case of a linear Pareto front. This bound holds for every $\mu \geq 2$ but the previous

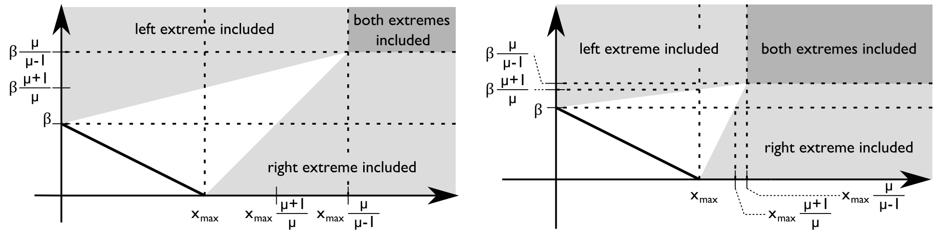

Fig. 3. How to choose the reference point to obtain the extremes in optimal μ-distributions: $\mu = 2$ (left) and $\mu = 4$ (right) for one and the same front $y = -x/2 + 1$.

theorem precises this bound to $r_1 \geq \frac{\mu+1}{\mu} x_{\max}$ and $r_2 \geq \frac{\mu+1}{\mu}\beta$ for a given μ which coincides with the old bound for $\mu = 2$ but is the closer to the nadir point $(x_{\max}, \beta)$, the larger μ gets—a result that has been previously shown as a limit result for arbitrary Pareto fronts [3].

Last, we want to note that, though the two equations (8) and (4) do not look the same at first sight, Theorem 3 complies with the characterization of optimal μ-distributions given in (4) [2, Theorem 6] for the case IX which can be shown by simple algebra.

8 Conclusions

Finding optimal μ-distributions, i.e., sets of μ points that have the highest quality indicator value among all sets of μ solutions coincides with the optimization goal of indicator-based multiobjective optimization algorithms and it is therefore important to characterize them. Here, we rigorously analyze optimal μ-distributions for the often used hypervolume indicator and for problems with linear Pareto fronts. The results are exhaustive in a sense that a single formula covers all possible choices of the hypervolume's reference point, including two previously proven cases. In addition to the newly covered cases, the new results show also how the choice of μ influences the fact that the extremes of the Pareto front are included in optimal μ-distributions for the case of linear fronts—a fact that has been only shown before by a lower bound result of choosing the reference point and not exact as here. The proofs also show that the optimal μ-distributions for problems with linear Pareto fronts are, given a $\mu \geq 2$ and a certain choice of the reference point, always unique.

Besides being the first exhaustive theoretical investigation of optimal μ-distributions for a specific front shape, the presented results are expected to have an impact in practical performance assessment as well. For the first time, it is now possible to use the exact optimal μ-distribution and its corresponding hypervolume when comparing algorithms on test problems with linear fronts such as DTLZ1 [8] or WFG3 [13] for any choice of the reference point[4]. It remains future work to theoretically characterize the optimal μ-distributions for test problems with other front shapes for which the optimal μ-distributions can only be approximated numerically at the moment [2].

[4] Theorem 3 can be applied directly with $\alpha = -1$ and $\beta = 0.5$ (DTLZ1) or $\beta = 1$ (WFG3).

Acknowledgments The author would like to thank Anne Auger for many fruitful discussions and comments as well as the French national research agency (via the SYSCOMM project ANR-08-SYSC-017) and the LIX laboratory for financial support. Many thanks also to Tobias Friedrich who pointed out the index errors within the Theorems 1–3.

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] A. Auger, J. Bader, D. Brockhoff, and E. Zitzler. Hypervolume-based Multiobjective Optimization: Theoretical Foundations and Practical Implications. *Theor. Comput. Sci.*, 2010. accepted for publication.

[4] J. Bader. *Hypervolume-Based Search For Multiobjective Optimization: Theory and Methods*. PhD thesis, ETH Zurich, 2010.

[5] N. Beume, C. M. Fonseca, M. Lopez-Ibanez, L. Paquete, and J. Vahrenhold. On the Complexity of Computing the Hypervolume Indicator. Technical Report CI-235/07, University of Dortmund, December 2007.

[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)*, pages 511–518. ACM, 2010.

[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*, 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)*, volume 2632 of *LNCS*, 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] N. Hansen, A. Auger, S. Finck, and R. Ros. Real-Parameter Black-Box Optimization Benchmarking 2009: Experimental Setup. INRIA Research Report RR-6828, INRIA Saclay—Ile-de-France, May 2009.

[13] 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.

[14] C. Igel, N. Hansen, and S. Roth. Covariance Matrix Adaptation for Multi-objective Optimization. *Evol. Comput.*, 15(1):1–28, 2007.

[15] T. Wagner, N. Beume, and B. Naujoks. Pareto-, Aggregation-, and Indicator-based Methods in Many-objective Optimization. In *Conference on Evolutionary Multi-Criterion Optimization (EMO 2007)*, volume 4403 of *LNCS*, pages 742–756. Springer, 2007.

[16] E. Zitzler and S. Künzli. Indicator-Based Selection in Multiobjective Search. In *Conference on Parallel Problem Solving from Nature (PPSN VIII)*, volume 3242 of *LNCS*, pages 832–842. Springer, 2004.

[17] E. Zitzler and L. Thiele. Multiobjective Optimization Using Evolutionary Algorithms - A Comparative Case Study. In *Conference on Parallel Problem Solving from Nature (PPSN V)*, pages 292–301, Amsterdam, 1998.

[18] E. Zitzler, L. Thiele, and J. Bader. On Set-Based Multiobjective Optimization. *IEEE T. Evolut. Comput.*, 14(1):58–79, 2010.

[19] E. Zitzler, L. Thiele, M. Laumanns, C. M. Fonseca, and V. Grunert da Fonseca. Performance Assessment of Multiobjective Optimizers: An Analysis and Review. *IEEE T. Evolut. Comput.*, 7(2):117–132, 2003.