Artificial Intelligence 195 (2013) 265–290

Contents lists available at SciVerse ScienceDirect

Artificial Intelligence

www.elsevier.com/locate/artint

Approximation quality of the hypervolume indicator

Karl Bringmann [a], Tobias Friedrich [b,*]

[a] Max-Planck-Institut für Informatik, Saarbrücken, Germany
[b] Friedrich–Schiller-Universität, Jena, Germany

ARTICLE INFO

Article history:
Received 20 June 2011
Received in revised form 11 September 2012
Accepted 17 September 2012
Available online 19 September 2012

Keywords:
Evolutionary computation
Multi-objective optimization
Hypervolume indicator
Theory
Approximation
Selection

ABSTRACT

In order to allow a comparison of (otherwise incomparable) sets, many evolutionary multi-objective optimizers use indicator functions to guide the search and to evaluate the performance of search algorithms. The most widely used indicator is the hypervolume indicator. It measures the volume of the dominated portion of the objective space bounded from below by a reference point.

Though the hypervolume indicator is very popular, it has not been shown that maximizing the hypervolume indicator of sets of bounded size is indeed equivalent to the overall objective of finding a good approximation of the Pareto front. To address this question, we compare the optimal approximation ratio with the approximation ratio achieved by two-dimensional sets maximizing the hypervolume indicator. We bound the optimal multiplicative approximation ratio of n points by $1 + \Theta(1/n)$ for arbitrary Pareto fronts. Furthermore, we prove that the same asymptotic approximation ratio is achieved by sets of n points that maximize the hypervolume indicator. However, there is a provable gap between the two approximation ratios which is even exponential in the ratio between the largest and the smallest value of the front.

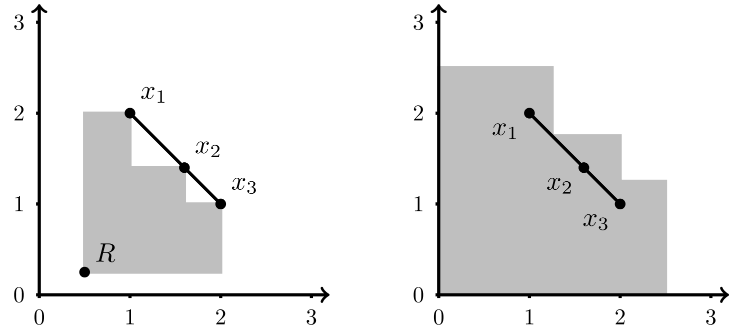
We also examine the additive approximation ratio of the hypervolume indicator in two dimensions and prove that it achieves the optimal additive approximation ratio apart from a small ratio $\leqslant n/(n-2)$, where n is the size of the population. Hence the hypervolume indicator can be used to achieve a good additive but not a good multiplicative approximation of a Pareto front. This motivates the introduction of a “logarithmic hypervolume indicator” which provably achieves a good multiplicative approximation ratio.

1. Introduction

Multi-objective problems are prevalent in many different fields like economics, engineering, management, and healthcare [15,17,25]. Such optimization problems with multiple objectives (like time vs. cost) often cannot be easily described by a single objective function. This implies that there is in general no unique optimum, but a possibly very large set of incomparable solutions which form the Pareto front. In the area of evolutionary computation, many different multi-objective evolutionary algorithms (MOEAs) have been developed to find a Pareto set of (small) size n which gives a *good approximation* of the Pareto front. A popular way to measure the quality of a Pareto set is the hypervolume indicator (HYP) which measures the volume of the dominated space bounded from below by a reference point [32]. For small numbers of objectives, MOEAs which directly use the hypervolume indicator to guide the search are the methods of choice. These include for example the generational MO-CMA-ES [19,29], SMS-EMOA [5,16], HypE [3], and variants of IBEA [31,34].

* Corresponding author.
E-mail address: friedrich@uni-jena.de (T. Friedrich).

0004-3702/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.artint.2012.09.005



(a) Hypervolume indicator. The shaded area shows the portion of the objective space which is dominated by the solution set $P = \{x_1, x_2, x_3\}$ up to the reference point R.

(b) Approximation ratio. The shaded area shows the portion of the objective space which is approximated by a factor of 1.25 by the solution set $P = \{x_1, x_2, x_3\}$.

Fig. 1. An illustration of the hypervolume indicator (cf. Section 3.2) compared to multiplicative approximation (cf. Section 2.1) for a linear front $f : [1, 2] \to [1, 2]$ with $f(x) = 3 - x$. The solution set $P = \{x_1, x_2, x_3\} = \{(1, 2), (1.6, 1.4), (2, 1)\}$ achieves a hypervolume of $\text{HYP}(P, R) = 1.865$ with respect to the reference point $R = (0.5, 0.25)$. The multiplicative approximation ratio of P is $\alpha^*(f, P) = 1.25$.

Despite its popularity, until recently there was not much rigorously known about the distribution of solution sets which maximize the hypervolume. Such solution sets have been described as empirically "well distributed" in [16,20,21]. In contrast to this, it was observed that "convex regions may be preferred to concave regions" [24,32] as well as that HYP is "biased towards the boundary solutions" [13]. It is known that some of these statements are invalid for the number of points $n \to \infty$ [2]. Auger et al. [2] proved that in this case the density of points depends only on the gradient.

We are interested in the approximation quality achieved by sets maximizing the hypervolume indicator. For this we have to formally define how to measure the approximation quality of solutions for multi-objective optimization problems. In the case of only one objective, the quality is typically measured by the (multiplicative) approximation ratio. For maximization problems this is the ratio between the optimal value and the best found value. This notion generalizes gracefully to our multi-objective setting. We say a Pareto set is an α-approximation if it approximately dominates the Pareto curve, that is, if for every point on the Pareto curve, the Pareto set contains a point that is at least as good approximately (within a factor α) in all objectives. For a sample of papers using this approach for classic (non-evolutionary) algorithms, see [11,12,14,26,27,30] and references therein. Fig. 1 gives an illustration of the hypervolume indicator and multiplicative approximation.

The advantage of the approximation ratio is that it gives a meaningful scalar value which allows us to compare the quality of solutions between different functions, different population sizes, and even different dimensions. In contrast to this, the hypervolume indicator always depends on the chosen reference point (cf. Section 3.1). A specific dominated volume does not give a priori any information on how well a front is approximated. This (often unwanted) freedom of choice not only changes the distribution of the points, but also makes the hypervolumes of different solutions measured relative to a reference point very hard to compare. This is even more true for algorithms (e.g. SMS-EMOA [5,16]) which dynamically change the reference point.

The choice by a decision maker between different Pareto fronts always remains subjective and there is no generally accepted optimization goal. However, if we are, for example, interested in a good multiplicative approximation, an "ideal" indicator would directly measure the approximation quality of a solution set P by returning the smallest $\alpha \in \mathbb{R}^+$ such that P is an α-approximation [23,33]. This corresponds to the unary multiplicative ε-indicator [35] where the reference set is the (possibly infinite) Pareto front. Unfortunately, such an indicator cannot be used in practice because the Pareto front is usually unknown.

This leads to the question of how close the approximations achieved by realistic indicators such as the hypervolume indicator come to those that could be obtained by such an "ideal" indicator. For this we consider the approximation ratio of a solution set maximizing the hypervolume.

At first glance, it is not obvious why maximizing the hypervolume indicator should yield a good approximation of the Pareto front. However, Friedrich, Horoba and Neumann [18] were the first to examine the approximation ratio of fronts maximizing the hypervolume. For linear and reciprocal functions, they were able to prove that maximizing HYP achieves an optimal approximation, while on other functions they showed empirically that the two might differ. In contrast, in this paper we provide a rigorous analysis of the approximation quality of hypervolume maximizing sets. So far this issue had been wide open even though it is crucial for understanding the implicit optimization goal when using the hypervolume indicator as a quality measure for populations.

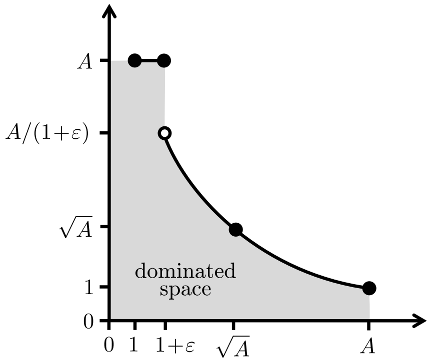


Fig. 2. Function $f_\varepsilon(x)$ as defined in Eq. (1.1) to show that depending on the scaling A, the approximation ratio of small solution sets with maximum hypervolume can be arbitrarily bad.

1.1. An illustrative example

We first give a simple example of why sets maximizing the hypervolume can be very bad approximations of the front. To define this properly, let us look at a maximization problem with a front that can be described by a monotonically decreasing function $f : [a, A] \to [b, B]$ with $0 < a < A$, $0 < b < B$. Then the approximation ratio (cf. Section 2.1) of a set of points $P := \{(x_1, y_1), \ldots, (x_n, y_n)\}$ (called the solution set) is the least $\alpha \geqslant 1$ such that for each $x \in [a, A]$ there is an $(x_i, y_i) \in P$ with

$$x \leqslant \alpha x_i \quad \text{and} \quad f(x) \leqslant \alpha y_i.$$

The approximation ratio does not depend on the scaling of $[a, A]$ and $[b, B]$. This can be seen by observing that for fixed constants $\mu, \nu > 0$, the function $f' : [\mu a, \mu A] \to [\nu b, \nu B]$ with $f'(x) = \nu f(x/\mu)$ achieves the same approximation ratio α with the solution set $P' := \{(\mu x_1, \nu y_1), \ldots, (\mu x_n, \nu y_n)\}$. However, the approximation ratio significantly depends on the proportions A/a and B/b. To see this, let us look at a function $f_\varepsilon : [1, A] \to [1, A]$ with $0 < \varepsilon < A - 1$ and

$$f_\varepsilon(x) := \begin{cases} A & \text{for } x \leqslant 1 + \varepsilon, \\ A/x & \text{for } x > 1 + \varepsilon. \end{cases} \tag{1.1}$$

A visualization of the function can be found in Fig. 2. Note that $A/a = B/b = A$ in this example. We want to see how well a single point from the set $\{(x, f_\varepsilon(x)) \mid x \in [1, A]\}$ can maximize the hypervolume and/or minimize the approximation ratio. By definition, the hypervolume[1] of a point $(x, f_\varepsilon(x))$ is

$$\mathrm{HYP}\big(\big\{\big(x, f_\varepsilon(x)\big)\big\}\big) = \begin{cases} xA & \text{for } x \leqslant 1 + \varepsilon, \\ A & \text{for } x > 1 + \varepsilon \end{cases}$$

while its multiplicative approximation ratio[2] is

$$\alpha^*\big(f_\varepsilon, \big\{\big(x, f_\varepsilon(x)\big)\big\}\big) = \max\big\{A/x, A/f_\varepsilon(x)\big\} = \begin{cases} A/x & \text{for } x \leqslant \sqrt{A}, \\ x & \text{for } x \geqslant \sqrt{A}. \end{cases}$$

The hypervolume is therefore maximized at exactly one point on the front, namely $(1+\varepsilon, 1)$. It achieves an approximation of $A/(1+\varepsilon)$. The best approximation of $\sqrt{A}$ is achieved by $(\sqrt{A}, \sqrt{A})$. Hence for $\varepsilon \to 0$, the approximation ratio of the solution set maximizing the hypervolume is off by a factor of $\sqrt{A}$ from the optimal ratio. This shows that the approximation ratio of sets maximizing the hypervolume can be very large for small numbers of points. However, this paper proves that this is *not* the case for sufficiently large solution sets.

This paper summarizes and extends our previous work presented in a sequence of three conference papers [7,8,10]. The first one was [8] which examined the multiplicative approximation factor of the hypervolume indicator. Afterwards, [7] studied the additive approximation factor and, finally, [10] proposed the logarithmic hypervolume indicator. The conference versions only discuss one particular aspect each and do not contain the full proofs. The majority of the material presented in Sections 4 to 6 is unpublished so far.

[1] We are assuming here that the size of the dominated space is measured relative to a common reference point $R = (0, 0)$. For the formal definition of HYP, see Eq. (3.2).

[2] For the formal definition of multiplicative approximation see Definition 2.1.

1.2. Our results

We are not interested in bounds for the approximation ratio on specific functions. Instead, we take a worst-case perspective for two objective optimization problems and look at all[3] functions $f : [a, A] \to [b, B]$ with $0 < a < A$, $0 < b < B$ and $f(a) = B$, $f(A) = b$.

We show that for all possible Pareto fronts the multiplicative approximation ratio achieved by a solution set of size n maximizing the hypervolume indicator is $1 + \Theta(1/n)$ (cf. Theorem 3.1).[4] This is shown to be *asymptotically equivalent* to the optimal multiplicative approximation ratio (cf. Corollary 2.4), which implies that the hypervolume indicator is guiding the search in the correct direction for sufficiently large n. However, the constant factor hidden by the Θ might be larger for the set maximizing hypervolume compared to the set with best possible approximation ratio. In fact, the multiplicative approximation ratio depends on the ratio A/a between the largest and smallest coordinate.[5] Using this notation, our precise result is the computation of the optimal multiplicative approximation ratio as $1 + \log(A/a)/n$ (cf. Corollary 2.4). We further show that the multiplicative approximation ratio for a set maximizing the hypervolume is strictly larger, namely on the order of at least $1 + \sqrt{A/a}/n$ (cf. Theorem 3.2). This implies that the multiplicative approximation ratio achieved by a set maximizing the hypervolume can be *exponentially worse* in the order of the ratio A/a. Hence for numerically very wide-spread fronts there are Pareto sets which give a much better multiplicative approximation than Pareto sets which maximize the hypervolume.

These results about the multiplicative approximation ratio can be seen as bad news for the hypervolume indicator. On the other hand, we examine the additive approximation ratio and observe that while the multiplicative approximation ratio is determined by the ratio A/a, the additive approximation ratio is determined by the width of the domain $A - a$. We prove that the optimal additive approximation ratio is $(A - a)/n$ (cf. Theorem 2.8) and upper bound the additive approximation ratio achieved by a set maximizing the hypervolume by $(A - a)/(n - 2)$ (cf. Theorem 3.3). This is a very strong statement, as apart from a small factor of $n/(n - 2)$, the additive approximation ratio achieved when maximizing the hypervolume is optimal. This shows that the hypervolume indicator yields a much *better additive than multiplicative* approximation.

It remains to find a natural indicator which provably achieves a good multiplicative approximation ratio. As this paper shows that the hypervolume gives a good additive approximation, we can use this to define an indicator which achieves a good multiplicative approximation: Logarithmize all axes before computing the classical hypervolume. We call this indicator the *logarithmic hypervolume indicator*. Note that in the setting of weighted hypervolume indicators [34] this corresponds to a reciprocal weight function (cf. Section 3.3). We prove that the logarithmic hypervolume indicator achieves a multiplicative approximation ratio of less than $1 + \log(A/a)/(n - 2)$ (cf. Corollary 3.6), which is again optimal apart from the factor $n/(n - 2)$. This indicates that as long as a multiplicative approximation is desired, the logarithmic hypervolume indicator should be preferred over the classic hypervolume indicator.

1.3. Outline

The outline of the paper is as follows. In Section 2 we define the notation used and the concepts of multiplicative and additive approximation ratios. Section 3 introduces the weighted, standard and logarithmic hypervolume indicator and presents our results on their approximation ratios. Afterwards, Section 4 justifies why we chose the definitions as they are. Most of the proofs of the paper are in the largest Section 5. We finally discuss how to translate our results to minimization problems in Section 6.

2. Preliminaries

We consider only the case of maximization problems on two objectives where there is a mapping from an arbitrary search space to an objective space which is a subset of $\mathbb{R}^2$. For minimization problems, see Section 6. Throughout this paper, we will work only on the objective space. For points from the objective space we define the following dominance relation:

$$(x_1, y_1) \preccurlyeq (x_2, y_2) \quad \text{iff} \quad x_1 \leqslant x_2 \text{ and } y_1 \leqslant y_2,$$

$$(x_1, y_1) \prec (x_2, y_2) \quad \text{iff} \quad (x_1, y_1) \preccurlyeq (x_2, y_2) \text{ and } (x_1, y_1) \neq (x_2, y_2).$$

We restrict ourselves to Pareto fronts that can be written as $\{(x, f(x)) \mid x \in [a, A]\}$ where $f : [a, A] \to [b, B]$ is a monotonically decreasing, upper semi-continuous[6] function with $f(a) = B$, $f(A) = b$ for some reals $a < A$, $b < B$. We write

[3] We restrict our attention to functions where there exists a set maximizing the hypervolume indicator. For technical details, see the definition of $\mathcal{F}$ at the beginning of Section 2.

[4] The precise statements of this and the following results are slightly more technical. For details see the respective theorems.

[5] The approximation ratio depends on the ratios in both dimensions. To simplify the presentation in the introduction, we assume that the ratio A/a in the first dimension is equal to the ratio B/b in the second dimension.

[6] Semi-continuity is a weaker property than normal continuity. A function f is said to be upper semi-continuous if for all points x of its domain, $\limsup_{y \to x} f(y) \leqslant f(x)$. Intuitively speaking this means that for all points x the function values for arguments near x are either close to $f(x)$ or less than $f(x)$. For more details see e.g. Rudin [28].

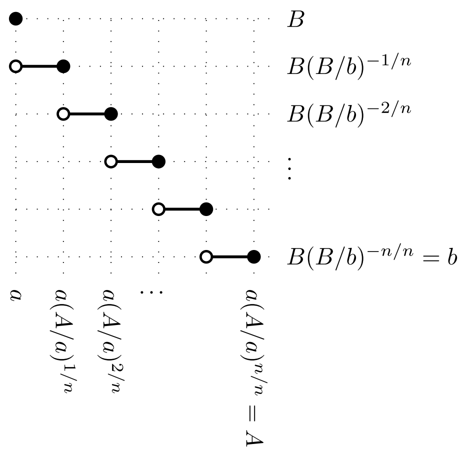


Fig. 3. Front f used for the lower bound construction in the proof of Theorem 2.3 in Section 5.1. Note that f is not only the black points, but the whole piecewise defined curve.

$\mathcal{F} = \mathcal{F}_{[a,A]\to[b,B]}$ for the set of all such functions f. We will use the term *front* for both the set of points $\{(x, f(x)) \mid x \in [a, A]\}$ and the function f.

Note that in contrast to the standard definition of a Pareto front, we consider a larger class of functions and do not require the functions f to be *strictly* monotonically decreasing. This has the advantage that we can handle step functions, as for example depicted in Fig. 3. Observe that such a function can be thought of as modeling a discrete front (namely the set of solid black points in Fig. 3). Moreover, sets maximizing the hypervolume indicator never contain points in the inner part of a constant interval of a function. Therefore, the sets maximizing the hypervolume on the discrete front are the same as the sets maximizing the hypervolume on the modeling step function. Since dominated points do not contribute to the hypervolume, our results carry over to discrete fronts.

The condition of f being upper semi-continuous cannot be relaxed further as without it the front lacks a certain symmetry in the two objectives: This condition is necessary and sufficient for the existence of the inverse function $f^{-1} : [b, B] \to [a, A]$ defined by setting

$$f^{-1}(y) := \max\bigl\{x \in [a, A] \bigm| f(x) \geqslant y\bigr\}.$$

Without upper semi-continuity, this maximum does not necessarily exist. Furthermore, this condition implies that there is a set maximizing the hypervolume indicator (see Section 4 for details).

Note that the set $\mathcal{F}$ of fronts we consider is a very general one. Most papers that theoretically examine the hypervolume indicator assume that the front is continuous and differentiable (e.g. [1,2,18]), and are thus not able to give results about step functions, which we can.

Let $n \in \mathbb{N}$. For fixed $[a, A], [b, B] \subset \mathbb{R}$ we call a set $P = \{p_1, \ldots, p_n\} \subset [a, A] \times [b, B]$ a *solution set* (of size[7] n) and write $\mathcal{P} := \mathcal{P}_n$ for the set of all such solution sets. A solution set P is said to be *feasible* for a front $f \in \mathcal{F}$, if $y \leqslant f(x)$ for all $p = (x, y) \in P$. We write $\mathcal{P}^f := \mathcal{P}_n^f \subseteq \mathcal{P}$ for the set of all solution sets (of size n) that are feasible for f.

To increase readability we occasionally write $P + r$ for $P \cup \{r\}$, where $P \subset \mathbb{R}^2$ and $r \in \mathbb{R}^2$, and similarly $P - r$ for $P \setminus \{r\}$.

A common approach to measure the quality of a solution set is to use unary indicator functions [35]. They assign to each solution set a real number that somehow reflects its quality, i.e., we have a function $Ind : \bigcup_{n=1}^{\infty} \mathcal{P}_n \to \mathbb{R}$. As throughout the paper $n \in \mathbb{N}$ is fixed, it is sufficient to define an indicator $Ind : \mathcal{P}_n \to \mathbb{R}$. Note that as we are only working in the objective space, we slightly deviate from the usual definition of an indicator function, where the domain is the search space, not the objective space.

In the following section we introduce notions of multiplicative and additive approximation quality.

2.1. Multiplicative approximation

When attempting to maximize an indicator function, we actually try to find a solution set $P \in \mathcal{P}_n^f$ that constitutes a good approximation of the front f. According to the custom for approximation algorithms, we measure the quality of a solution by its multiplicative approximation ratio. This can be transferred to the world of multi-objective optimization. For this we use the following definition of Papadimitriou and Yannakakis [26] which was also used in [7,8,10,18,22,23]. Note that it is crucial to require $a, b > 0$ here, as it is unclear what multiplicatively approximating a negative number should mean. We will always assume this when talking about multiplicative approximation throughout the paper.

[7] Note that the points $p_1, \ldots, p_n$ are not required to be pairwise different, so a solution set of size n has between 1 and n elements. This implies $\mathcal{P}_n \subset \mathcal{P}_m$ for $n < m$.

Definition 2.1. Let $f \in \mathcal{F}$ and $P \in \mathcal{P}_n^f$. The solution set P is a *multiplicative* α*-approximation* of f if for each $\hat{x} \in [a, A]$ there is a $p = (x, y) \in P$ with

$$\hat{x} \leqslant \alpha x \quad \text{and} \quad f(\hat{x}) \leqslant \alpha y \tag{2.1}$$

where $\alpha \in \mathbb{R}$, $\alpha \geqslant 1$. The *multiplicative approximation ratio* of P with respect to f is defined as

$$\alpha^*(f, P) := \inf\{\alpha \in \mathbb{R} \mid P \text{ is a multiplicative } \alpha\text{-approximation of } f\}.$$

The quality of an algorithm which calculates a solution set of size n for each Pareto front in $\mathcal{F}$ has to be compared with the respective optimal approximation ratio defined as follows.

Definition 2.2. For fixed $[a, A]$, $[b, B]$, and n, let

$$\alpha^*_{OPT} := \sup_{f \in \mathcal{F}} \inf_{P \in \mathcal{P}_n^f} \alpha^*(f, P).$$

The value α^*_{OPT} is chosen such that every front in $\mathcal{F}$ can be approximated by n points to a ratio of α^*_{OPT}, and there is a front which cannot be approximated better. In Section 5.1 we show the following two results.

Theorem 2.3. $\alpha^*_{OPT} = \min\{A/a, B/b\}^{1/n}$.

Corollary 2.4. *For all* $n \geqslant \log(\min\{A/a, B/b\})/\varepsilon$ *and* $\varepsilon \in (0, 1)$,

$$\alpha^*_{OPT} \geqslant 1 + \frac{\log(\min\{A/a, B/b\})}{n},$$
$$\alpha^*_{OPT} \leqslant 1 + (1+\varepsilon)\frac{\log(\min\{A/a, B/b\})}{n}.$$

We further want to measure the approximation of the solution set of size n maximizing an indicator *Ind*. As there might be several solution sets maximizing *Ind*, we consider the worst case and use the following definition.

Definition 2.5. For a unary indicator *Ind* and fixed $[a, A], [b, B], n$, and $f \in \mathcal{F}$ let

$$\mathcal{P}^f_{Ind} := \mathcal{P}^f_{Ind,n} := \left\{ P \in \mathcal{P}_n^f \;\middle|\; P \in \operatorname*{argmax}_{Q \in \mathcal{P}_n^f} Ind(Q) \right\} \quad \text{and}$$
$$\alpha^*_{Ind} := \alpha^*_{Ind,n} := \sup_{f \in \mathcal{F}} \sup_{P \in \mathcal{P}^f_{Ind,n}} \alpha^*(f, P).$$

The set $\mathcal{P}^f_{Ind}$ is the set of all feasible solution sets (of size n) that maximize *Ind* on f. The value α^*_{Ind} is chosen such that for every front f in $\mathcal{F}$ every solution set maximizing *Ind* approximates f by a ratio of at most α^*_{Ind}. Observe that we take a worst case viewpoint there, as we take the supremum over all solution sets maximizing the hypervolume indicator. This may seem unfair to the hypervolume indicator; however, Lemma 4.2 proves that it makes no difference whether we take the worst or best case perspective at this point, i.e., whether we take the supremum or infimum over $P \in \mathcal{P}^f_{Ind}$.

Note that we assume here that there exists a solution set that maximizes the indicator, i.e., we assume that the set $\mathcal{P}^f_{Ind}$ is non-empty. Since we restrict the fronts to be upper semi-continuous, this will be the case for all the indicators we consider, as shown in Lemma 4.1.

2.2. Additive approximation

Depending on the problem at hand, one can also consider an additive approximation ratio. We use the following definition, analogous to Definition 2.1.

Definition 2.6. Let $f \in \mathcal{F}$ and $P \in \mathcal{P}_n^f$. The solution set P is an *additive* α*-approximation* of f if for each $\hat{x} \in [a, A]$ there is a $p = (x, y) \in P$ with

$$\hat{x} \leqslant x + \alpha \quad \text{and} \quad f(\hat{x}) \leqslant y + \alpha \tag{2.2}$$

Table 1
Theoretical results for the optimal approximation ratio and upper bounds for the approximation ratios of HYP and LogHYP. See the cited theorems for the precise statements.

	Multiplicative approximation	Additive approximation
OPT	$1+\frac{\log(\min\{A/a,B/b\})}{n}$ (Corollary 2.4)	$\frac{\min\{A-a,B-b\}}{n}$ (Theorem 2.8)
HYP	$1+\frac{\sqrt{A/a}+\sqrt{B/b}}{n-4}$ (Theorem 3.1)	$\frac{\sqrt{(A-a)(B-b)}}{n-2}$ (Theorem 3.3)
LogHYP	$1+\frac{\sqrt{\log(A/a)\log(B/b)}}{n-2}$ (Corollary 3.6)	open

where $\alpha \in \mathbb{R}$, $\alpha \geqslant 0$. The *additive approximation ratio*[8] of P with respect to f is defined as

$$\alpha^{+}(f,P) := \inf\{\alpha \in \mathbb{R} \mid P \text{ is an additive } \alpha\text{-approximation of } f\}.$$

One thing that may come to mind when reading this definition is the following. It may be that the objectives are unbalanced, meaning that we would like to give them some kind of weight in the approximation. A possible definition for additive approximation incorporating this kind of weight uses weights $w_x, w_y > 0$ for the objectives and defines the point $(\hat{x}, f(\hat{x}))$ to be approximated by $(x,y) \in P$ by the ratio α iff $\hat{x} \leqslant x + w_x\alpha$ and $f(\hat{x}) \leqslant y + w_y\alpha$. This makes perfect sense and may be preferred over the standard unweighted definition in certain cases. However, it is already accounted for by the unweighted definition: After rescaling the x-axis by a factor of $1/w_x$ and the y-axis by a factor of $1/w_y$ we have $\hat{x}' \leqslant x' + \alpha$ iff $\hat{x} \leqslant x + w_x\alpha$ and $f'(\hat{x}') \leqslant y' + \alpha$ iff $f(\hat{x}) \leqslant y + w_y\alpha$ (where a primed variable denotes the variable after rescaling). Hence, all the results in this paper do apply to the weighted definition of additive approximation, one just has to rescale the axes correctly. Note that this kind of weight corresponds to a weighting of the form $\hat{x} \leqslant \alpha^{w_x} x$ and $f(\hat{x}) \leqslant \alpha^{w_y} y$ for multiplicative approximation.

Going on with the definitions, we are again interested in the optimal approximation ratio for Pareto fronts in $\mathcal{F}$. We use the following definition, analogous to Definition 2.2.

Definition 2.7. For fixed $[a,A]$, $[b,B]$, and n, let

$$\alpha^{+}_{OPT} := \sup_{f\in\mathcal{F}} \inf_{P\in\mathcal{P}_n^f} \alpha^{+}(f,P).$$

In Section 5.5 the following result will be proven using a relation between additive and multiplicative approximations and Theorem 2.3.

Theorem 2.8. $\alpha^{+}_{OPT} = \dfrac{\min\{A-a, B-b\}}{n}$.

Moreover, the analog for α^{*}_{Ind} is defined similarly to Definition 2.5.

Definition 2.9. For a unary indicator *Ind* and fixed $[a,A]$, $[b,B]$, n, and $f \in \mathcal{F}$ let

$$\alpha^{+}_{Ind} := \alpha^{+}_{Ind,n} := \sup_{f\in\mathcal{F}} \sup_{P\in\mathcal{P}^f_{Ind,n}} \alpha^{+}(f,P).$$

Again, Lemma 4.2 shows that it makes no difference whether we take a supremum or infimum over $P \in \mathcal{P}^f_{Ind}$.

3. Indicators and their approximation quality

This section presents the majority of the results of this paper. It is structured along the different indicators. First, we recap the general framework of the weighted hypervolume indicator. Afterwards, the standard, logarithmic, and hybrid hypervolume indicators are introduced and our respective results are presented. We also discuss briefly the well-known ε-indicator. The results are summarized in Table 1. Most proofs are deferred to Section 5.

3.1. Weighted hypervolume indicator

The classical definition of the hypervolume indicator is the volume of the dominated portion of the objective space relative to a fixed footpoint called the reference point $R = (R_x, R_y) \preccurlyeq (a,b)$. As a general framework for our two indicators

[8] To match the notation for multiplicative approximation, we call this value a “ratio”, although “difference” might be a more precise term.

we use the more general weighted hypervolume indicator of Zitzler et al. [34]. It weights points with a weight distribution $w : \mathbb{R}^2 \to \mathbb{R}_{>0}$ (or at least $w : [R_x, A] \times [R_y, B] \to \mathbb{R}_{>0}$), of which we require that the integral

$$\text{area}_w(x_1, y_1, x_2, y_2) := \int_{x_1}^{x_2} \int_{y_1}^{y_2} w(x, y)\, dy\, dx \tag{3.1}$$

exists. The *hypervolume* $\text{HYP}_w(P, R)$ (or $\text{HYP}_w(P)$ for short) of a solution set $P \in \mathcal{P}$ is then defined as

$$\begin{aligned} \text{HYP}_w(P) &:= \text{HYP}_w(P, R) \\ &:= \iint_{\mathbb{R}^2} A_{P,R}(x, y) w(x, y)\, dy\, dx \end{aligned} \tag{3.2}$$

where the attainment function $A_{P,R} : \mathbb{R}^2 \to \mathbb{R}$ is an indicator function on the objective space which describes the space above the reference point that is weakly dominated by P. Formally, $A_{P,R}(x, y) = 1$ if $(R_x, R_y) \preccurlyeq (x, y)$ and there is a $p = (p_x, p_y) \in P$ such that $(x, y) \preccurlyeq (p_x, p_y)$, and $A_{P,R}(x, y) = 0$ otherwise.

The original purpose of the weighted hypervolume indicator was to allow the decision maker to stress certain regions of the objective space. In this paper we unleash one of its hidden powers by showing that one gets a better multiplicative approximation by choosing the right weight distribution.

3.2. Standard hypervolume indicator

If w is the all-ones functions $\mathbf{1}$ with $\mathbf{1}(x, y) = 1$ for all $x, y \in \mathbb{R}$, the above definition matches the standard definition of the hypervolume indicator. In this case we write $\text{HYP} = \text{HYP}_{\mathbf{1}}$ for short. Bounds for this indicator are of particular interest. We prove in Section 5.2 an upper bound for α^*_{HYP}. As this is a key part of this paper, we give the precise result there in Theorem 5.4. Here we give only the following slightly weaker, but more readable bound, which immediately follows from Theorem 5.4.

Theorem 3.1. *Let $f \in \mathcal{F}$, $n > 4$, and let $R = (R_x, R_y)$ be the reference point. If we have*

$$\begin{aligned} &(n-2)(a - R_x) \geqslant \sqrt{Aa} \quad \text{and} \\ &(n-2)(b - R_y) \geqslant \sqrt{Bb} \end{aligned}$$

then

$$\alpha^*_{HYP} \leqslant 1 + \frac{\sqrt{A/a} + \sqrt{B/b}}{n-4}.$$

This shows that for sufficiently large n *or* a sufficiently far away reference point the hypervolume yields a multiplicative approximation with optimal asymptotic behavior in n. However, the constant factor is $\sqrt{A/a} + \sqrt{B/b}$ instead of the optimal $\log(\min\{A/a, B/b\})$ (see Corollary 2.4), so even for $A/a = B/b$ it is exponentially worse than the optimal constant. The following result shows that the above bound is more or less tight. Its proof is given in Section 5.3.

Theorem 3.2. *Let $n \geqslant 4$, $\frac{A}{a} = \frac{B}{b} \geqslant 13$, and $R = (R_x, R_y) \preccurlyeq (0, 0)$ be the reference point. Then*

$$\alpha^*_{HYP} \geqslant 1 + \frac{2\sqrt{A/a - 1}}{3(n-1)}.$$

This shows that the constant factor is indeed exponentially worse.

On the other hand, the following theorem (proven in Section 5.4) shows that HYP has a close-to-optimal additive approximation ratio.

Theorem 3.3. *If $n > 2$ and*

$$(n-2)\min\{a - R_x, b - R_y\} \geqslant \sqrt{(A-a)(B-b)}$$

we have

$$\alpha^+_{HYP} \leqslant \frac{\sqrt{(A-a)(B-b)}}{n-2}.$$

First note that the assumption is fulfilled if n is large enough *or* if the reference point is sufficiently far away from (a, b). Hence this is no real restriction. Moreover, compare this result to the bound for the optimal additive approximation ratio of Theorem 2.8. This shows that for $A - a \approx B - b$, α^{+}_{HYP} is very close to α^{+}_{OPT}. Further, for $A - a \ll B - b$ (or $A - a \gg B - b$) the constant in Theorem 3.3 is the geometric mean of $A - a$ and $B - b$ while in Theorem 2.8 it is instead the minimum of both. As there is a provable gap of log vs. square root of A/a for the multiplicative approximation ratio, this proves that HYP yields a much better additive approximation than a multiplicative one.

3.3. Logarithmic hypervolume indicator

Now we know an indicator yielding a good additive approximation, namely the (standard) hypervolume indicator HYP. For finding a good multiplicative approximation HYP turned out to be inapplicable, at least for large spreads A/a and B/b in the worst case. We propose the *logarithmic hypervolume indicator* to address this problem. For a solution set $P \in \mathcal{P}$ and reference point $R = (R_x, R_y)$ with $(R_x, R_y) \preccurlyeq (a, b)$, $R_x, R_y > 0$ we define

$$\text{LOGHYP}(P, R) := \text{HYP}_{\mathbf{1}}(\log P, \log R),$$

where $\log P := \{(\log x, \log y) \mid (x, y) \in P\}$ and $\log R := (\log R_x, \log R_y)$. Here, as in the standard case, the reference point is a parameter to be chosen by the user. Note that we do not really change the axes of the problem to logarithmic scale: We only change the calculation of the hypervolume, not the problem itself.

The above definition is nice in that it allows LOGHYP to be computed using existing implementations of algorithms for HYP, only wiring the input differently, i.e., logarithmizing everything beforehand.

It is very illustrative, though, to observe that the logarithmic hypervolume indicator fits very well in the weighted hypervolume framework: An equivalent definition of LOGHYP is

$$\text{LOGHYP}(P, R) := \text{HYP}_{\hat{w}}(P, R),$$

where $\hat{w}(x, y) = 1/(xy)$ is the appropriate weight distribution.

Lemma 3.4. $\text{HYP}_{\mathbf{1}}(\log P, \log R) = \text{HYP}_{\hat{w}}(P, R)$.

Proof. Let $\{(x_1, y_1), \dots, (x_k, y_k)\} \subseteq P$ be the points in P not dominated by any other point in P with $x_1 < \cdots < x_k$, $y_1 > \cdots > y_k$. With $x_0 := R_x$ we can then compute HYP as

$$\begin{aligned}
\text{HYP}_{\mathbf{1}}(\log P, \log R) &= \sum_{i=1}^{k} \int_{\log x_{i-1}}^{\log x_i} \int_{\log R_y}^{\log y_i} 1 \, dy \, dx \\
&= \sum_{i=1}^{k} \int_{x_{i-1}}^{x_i} \int_{R_y}^{y_i} \frac{1}{xy} \, dy \, dx \\
&= \text{HYP}_{\hat{w}}(P, R). \qquad \square
\end{aligned}$$

The next result, to be shown in Section 5.5, shows that the logarithmic hypervolume indicator yields a good multiplicative approximation, just as the standard hypervolume indicator yields a good additive approximation.

Theorem 3.5. *If $n > 2$ and*

$$(n-2) \log \min\{a/R_x, b/R_y\} \geqslant \sqrt{\log(A/a) \log(B/b)}$$

we have

$$\alpha^{*}_{logHYP} \leqslant \exp\left(\frac{\sqrt{\log(A/a) \log(B/b)}}{n-2}\right).$$

Note that the assumption is fulfilled if n is large enough *or* we choose the reference point near enough to $(0, 0)$. This is a very good upper bound compared to $\alpha^{*}_{OPT} = \exp(\min\{\log(A/a), \log(B/b)\}/n)$. Also compare the next corollary to Corollary 2.4. Its proof is analogous to the one of Corollary 2.4.

Corollary 3.6. *For $\varepsilon \in (0, 1)$ and all*

$$n \geqslant 2 + \sqrt{\log(A/a) \log(B/b)} / \min\{\varepsilon, \log(a/R_x), \log(b/R_y)\}$$

we have

$$\alpha^*_{logHYP} \leqslant 1 + (1+\varepsilon)\frac{\sqrt{\log(A/a)\log(B/b)}}{n-2}.$$

Hence we get a much better constant factor than in the bound of α^*_{HYP}.

3.4. Hybrid hypervolume indicator

The results of the preceding sections imply that guiding the search with the hypervolume indicator is an appropriate choice if we want an additive approximation. On the other hand, guiding the search with the logarithmic hypervolume indicator is preferable if we want a multiplicative approximation.

Of course, it may happen that one wants an additive approximation of some objectives and a multiplicative approximation of others. We propose a simple rule of thumb for this case: Logarithmize all objectives of the second type, i.e., those that should get multiplicatively approximated (leaving the objectives of the first type as they are) and then compute the hypervolume indicator. This hybrid indicator should work as intended, i.e., maximizing it should give a good additive approximation of the objectives of the first type and a good multiplicative approximation of the objectives of the second type.

As an illustration, assume we have two objectives, x and y, and want to approximate x additively and y multiplicatively. Then we use the hybrid indicator $Ind(P, R) := \mathrm{HYP}(P', (R_x, \log R_y))$, where $P' = \{(x_i, \log y_i) \mid (x_i, y_i) \in P\}$ and R is again a reference point. This indicator logarithmizes the y-axis and applies HYP afterwards. Along the lines of the proofs in this paper one can show that maximizing Ind on a front f yields a solution set P with the following property: For any $\hat{x} \in [a, A]$ there is a $p = (x, y) \in P$ with

$$\hat{x} \leqslant x + \alpha^+ \quad \text{and} \quad f(\hat{x}) \leqslant y\alpha^*,$$

where $\alpha^* = \exp \alpha^+$ and $\alpha^+ \leqslant \frac{\sqrt{(A-a)(\log(B)-\log(b))}}{n-2}$. This means, that we get an additive approximation of x and a multiplicative approximation of y, as desired.

3.5. ε-Indicator

Another important class of indicators which we want to discuss only briefly are the binary ε-indicators [23,33]. For two solution sets P and Q its additive version is defined as

$$I_{\varepsilon^+}(P, Q) := \max_{(x_1,y_1)\in P}\ \min_{(x_2,y_2)\in Q}\ \max\{x_1 - x_2, y_1 - y_2\}$$

which is the smallest value ε by which we have to shift Q along both axes such that it dominates P. This binary indicator favors P over Q if $I_{\varepsilon^+}(P, Q) > I_{\varepsilon^+}(Q, P)$. The multiplicative ε-indicator is defined analogously as

$$I_{\varepsilon^*}(P, Q) := \max_{(x_1,y_1)\in P}\ \min_{(x_2,y_2)\in Q}\ \max\left\{\frac{x_1}{x_2}, \frac{y_1}{y_2}\right\}.$$

This definition appears to be much closer to the definition of additive or multiplicative approximation (cf. Definitions 2.1 and 2.6) than the definition of the hypervolume indicator. The ε-indicator can even be seen as a relaxation of the "ideal" indicator noted in the introduction. In light of the above results regarding the hypervolume indicator it is natural to ask whether the ε^+-indicator also yields a good additive approximation and the ε^*-indicator also yields a good multiplicative approximation.

Unfortunately, this is not a well posed question as $\succ_{I_\varepsilon}$ is not a total order. The reason for this is that it lacks transitivity as it contains deteriorative cycles [4], i.e., an algorithm trying to maximize based on $\succ_{I_\varepsilon}$ may return to a set of search points that it has obtained before. This implies that there is, in general, no solution set P that is maximal for the relation $\succ_{I_\varepsilon}$. Hence statements on the approximation ratio of sets maximizing the ε-indicator are not meaningful. It is an open question how to describe the approximation quality achieved by the ε-indicator.

4. Two technicalities

Before we prove the claims from the previous section, we consider two details of the definitions which might look counterintuitive the first time encountered. These are (i) that we require the fronts to be upper semi-continuous and (ii) that the definition of α^*_{Ind} is the "worst case" approximation ratio over all sets maximizing the indicator and not, e.g., the "best case".

4.1. Why we need upper semi-continuity

We show that without upper semi-continuity there does not necessarily exist a solution set maximizing HYP. To see this, consider the front $f : [1,2] \to [1,2]$ with

$$f(x) := \begin{cases} 1 & \text{for } x = 2, \\ 2 & \text{for } 1 \leqslant x < 2 \end{cases} \tag{4.1}$$

and reference point $R = (0,0)$. The one-element solution set $P = \{(2-\varepsilon, 2)\}$ achieves $\mathrm{HYP}(P) = 4 - 2\varepsilon$ for each $\varepsilon > 0$. However, no solution set P' can have $\mathrm{HYP}(P') = 4$, as $f(2) = 1 < 2$. Thus, there exists no solution set maximizing HYP, as there is an infinite series of solution sets with larger and larger hypervolume indicator, but the limit $\sup_{P \in \mathcal{P}^f} \mathrm{HYP}(P) = 4$ is not taken by any solution set.

Next we prove that conditioning on fronts being upper semi-continuous implies that there are sets maximizing the weighted hypervolume indicator. In more detail, there is a solution set P of size n which maximizes HYP among all solution sets of size n.

Lemma 4.1. *Let $f \in \mathcal{F}$, $n \in \mathbb{N}$, and $w : \mathbb{R}^2 \to \mathbb{R}_{>0}$ be a weight function. Then there exists a (not necessarily unique) solution set $P \in \mathcal{P}_n^f$ that maximizes the weighted hypervolume indicator HYP_w on $\mathcal{P}_n^f$.*

Proof. Consider the sets $S := \{(x,y) \in [a,A] \times [b,B] \mid y \leqslant f(x)\}$ of feasible points for the front f, and S^n, the n-tuples of feasible points. Let us denote by π the direct mapping from S^n into $\mathcal{P}^f$ given by $((x_1,y_1),\dots,(x_n,y_n)) \mapsto \{(x_1,y_1),\dots,(x_n,y_n)\}$. Consider the map

$$\phi : S^n \xrightarrow{\ \pi\ } \mathcal{P}^f \xrightarrow{\ \mathrm{HYP}_w(\cdot,R)\ } \mathbb{R}.$$

Using the notion area_w from Section 3.1, we can explicitly express the map ϕ for a feasible tuple $((x_1,y_1),\dots,(x_n,y_n))$ using an inclusion–exclusion formula,

$$\phi\big((x_1,y_1),\dots,(x_n,y_n)\big) = \sum_{\emptyset \neq M \subseteq \{1,\dots,n\}} (-1)^{|M|+1} \mathrm{area}_w\Big(R_x, R_y, \min_{i\in M} x_i, \min_{i \in M} y_i\Big).$$

Since $\mathrm{area}_w(\cdot)$, as defined in Eq. (3.1), is continuous, this shows that the map ϕ is continuous too.

We now show that S is compact. To see this, take any sequence of points $(x_i,y_i)_{i\in\mathbb{N}}$ in S. In $[a,A]\times[b,B]$ (which is a superset of S) this sequence has a convergent subsequence which we again call $(x_i,y_i)_{i\in\mathbb{N}}$. For this convergent subsequence we have $\lim_{i\to\infty} y_i \leqslant \lim_{i\to\infty} f(x_i) \leqslant f(\lim_{i\to\infty} x_i)$, where we used $(x_i,y_i) \in S$ and the upper semi-continuity of f. This shows that the limit again lies in S, and thus the compactness of S. Compactness of S^n follows trivially.

This proves that ϕ takes its maximum as it is a continuous function on a compact set. Moreover, since π is surjective, $\mathrm{HYP}(\cdot, R)$ takes its maximum on $\mathcal{P}^f$, which is what was to be shown. □

Note that assuming upper semi-continuity is sufficient for the existence of solution sets which maximize HYP, but it is not necessary. There are fronts which are *not* upper semi-continuous in general, but still have a unique HYP-maximal solution set, since they are only not upper semi-continuous in parts where there are no HYP-maximizing points on the front (for fixed n). We need upper semi-continuity, however, for the existence of the inverse function f^{-1} as defined in Section 2, which implies symmetry of the two objectives.

4.2. Why we consider the worst-case approximation ratio

We show that in the definition of the approximation ratio of the hypervolume indicator we can replace "worst case" by "best case" and not change the value of α^*_{Ind} or α^+_{Ind}.

Before doing that, we confirm that the solution set maximizing the hypervolume indicator is indeed not unique in general. To show this, let us look again at the introductory example function f_ε from Eq. (1.1). By choosing $\varepsilon = 0$ we get a front $f_0 : [1,A] \to [1,A]$ with $f_0(x) = A/x$. With reference point $R = (0,0)$, we get $\mathrm{HYP}(\{(x, f(x))\}) = x(A/x) = A$ for all $x \in [1,A]$. Therefore the set of solution sets of size $n = 1$ which maximize HYP is far from unique as $\mathcal{P}^f_{\mathrm{HYP}} = \{\{(x, f(x))\} \mid x \in [1,A]\}$. Moreover, this example shows that the approximation ratios of two solution sets maximizing HYP can differ significantly as the solution set $\{(1,A)\}$ achieves an approximation ratio of A, while the solution set $\{(\sqrt{A}, \sqrt{A})\}$ achieves an approximation ratio of $\sqrt{A}$. However, by taking the supremum over all functions in $\mathcal{F}$ this difference is nullified as shown by the following lemma.

Here, we consider the definition of α^+_{HYP} (cf. Definition 2.9) as the worst or best case approximation ratio of the sets maximizing the hypervolume indicator and show that both values coincide.

Lemma 4.2. *In the definition of α^+_{HYP} it does not matter whether we take the best or worst case over the solution sets maximizing the hypervolume, that is,*

$$\sup_{f\in\mathcal{F}} \sup_{P\in\mathcal{P}^f_{\mathrm{HYP}}} \alpha^+(f,P) = \sup_{f\in\mathcal{F}} \inf_{P\in\mathcal{P}^f_{\mathrm{HYP}}} \alpha^+(f,P).$$

Proof. We show that for every front $f \in \mathcal{F}$ and solution set $P \in \mathcal{P}^f_{\mathrm{HYP}}$ with $\alpha^+(f,P) > 1$ and for each $\varepsilon > 0$, there is a front $f' \in \mathcal{F}$ that has only one solution set P' maximizing the hypervolume indicator on f', and we have $|\alpha^+(f,P) - \alpha^+(f',P')| < \varepsilon$. This means that when taking the supremum we can restrict our attention to fronts that have only one solution set maximizing the hypervolume, but for such fronts both definitions from above agree, which proves the claim.

Consider such f and P and a point $r = (x, f(x))$ that is not approximated by a ratio $\geqslant \alpha^+(f,P) - \varepsilon'$, $\varepsilon' > 0$. We know that such a point exists by definition of $\alpha^+(f,P)$. For sufficiently small ε', r is not dominated by any point in P, as $\alpha^+(f,P) - \varepsilon' > 1$.

Now, $P = \{p_1,\dots,p_n\}$ has n pairwise different points p_i (otherwise $P + r$ would have greater hypervolume than P) and there are no points $p_i, p_j \in P$ with $p_i \prec p_j$ (otherwise $P - p_i + r$ would have greater hypervolume), hence we can assume that $p_i = (x_i, y_i)$ with $a \leqslant x_1 < \cdots < x_n \leqslant A$ and $B \geqslant y_1 > \cdots > y_n \geqslant b$. Moreover, we have $y_i = f(x_i)$ (otherwise $P - p_i + (x_i, f(x_i))$ would have greater hypervolume).

Let us consider the step function defined by the points $p_i := (x_i, y_i)$, $i = 0,\dots,n+2$ where we set $(x_0, y_0) = (a, B)$, $(x_{n+1}, y_{n+1}) = (A, b)$ and $(x_{n+2}, y_{n+2}) = r$. Formally, this step function is

$$\widehat{f}(x) = \max\{y_i \mid i \in \{0,\dots,n+2\},\ x_i \geqslant x\}.$$

We have $\widehat{f}(x) \leqslant f(x)$ for all $x \in [a, A]$; therefore no solution set that is infeasible for f is feasible for $\widehat{f}$. Moreover, the solution set P is still feasible for $\widehat{f}$. Hence, P still maximizes the hypervolume indicator on $\widehat{f}$.

It is easy to see that the solution sets maximizing the hypervolume indicator on $\widehat{f}$ are among the sets $P_I := \{p_i \mid i \in I\}$, $I \subset \{0,\dots,n+2\}$, $|I| = n$, as any other solution set is dominated by some P_I. We need to make sure that the solution set $P = P_{\{1,\dots,n\}}$ has strictly greater hypervolume than any other P_I. For this we modify the front $\widehat{f}$ again, but do a case distinction.

Case 1: We have $y_1 < B$ or $x_n < A$. By symmetry we have to look at only one of these cases, so let $y_1 < B$. We change the points p_i $(1 \leqslant i \leqslant n)$ slightly by setting $y'_i := y_i + (n+1-i)\varepsilon'$, calling the resulting points $p'_i := (x'_i, y'_i) := (x_i, y'_i)$ (with $p'_i := p_i$ for $i = 0, n+1, n+2$) and the induced step function of $p'_0,\dots,p'_{n+2}$ by the name f'. Again, the solution sets maximizing the hypervolume indicator on f' are among the sets $P'_I := \{p'_i \mid i \in I\}$, $I \subset \{0,\dots,n+2\}$, $|I| = n$. Now, consider the space dominated by a solution set P_I. When going from P_I to P'_I this space increases in the y-direction. Observe that at a particular point x with $x_{i-1} < x \leqslant x_i$, the y-coordinate increases by at most

$$\varepsilon'\big(n+1-\min\{j \geqslant i \mid j \in I \cup \{n+1\}\}\big),$$

since the only increase can come from the point $p_j \in P_I$ with the next larger x-coordinate. Here we add $n+1$ to I so that the total term gets 0 if there is no point p_i $(1 \leqslant i \leqslant n)$ with the next larger x-coordinate. Note that the increase at a particular x can be smaller than this term, if the next larger point is $p_{n+2} = r$, which was not increased at all. Also note that for $P = P_{\{1,\dots,n\}}$ this upper bound is met with equality for all $x \in [a, A]$. Thus, we have for all I

$$\mathrm{HYP}(P'_I) - \mathrm{HYP}(P_I) \leqslant \sum_{i=1}^{n} \varepsilon'(x_i - x_{i-1})\big(n+1-\min\{j \geqslant i \mid j \in I \cup \{n+1\}\}\big),$$

with equality for $I = \{1,\dots,n\}$. Now, it is easy to see that this difference has a unique maximum for $I = \{1,\dots,n\}$, which is why $P' := P'_{\{1,\dots,n\}}$ is the single solution set maximizing the hypervolume indicator on f'. Also note that for ε' sufficiently small we changed the coordinates of the p_i $(1 \leqslant i \leqslant n)$ by less than $\varepsilon/2$, which implies that the additive approximation ratio $\alpha^+(f',P')$ differs from $\alpha^+(f,P)$ by at most ε (recall that r is approximated by the p_i $(1 \leqslant i \leqslant n)$ by a ratio of $\geqslant \alpha^+(f,P) - \varepsilon'$).

Case 2: We have $y_1 = B$ and $x_n = A$. Then either $p_0 = p_1$ or $p_0 \prec p_1$ and there is no solution set maximizing the hypervolume indicator that includes p_0. A similar statement holds for p_{n+1}, so we can discard p_0 and p_{n+1}, meaning that the solution sets maximizing the hypervolume indicator on $\widehat{f}$ are among the P_I with $I \subset \{1,\dots,n,n+2\}$, $|I| = n$. We make $r = p_{n+2}$ slightly worse by setting $y'_{n+2} := y_{n+2} - \varepsilon'$. For ε' sufficiently small r is still not dominated by any other point p_i. We call the resulting point again $r = p_{n+2}$ and the induced step function f'. Going from $\widehat{f}$ to f' the hypervolume decreases for solution sets containing r. Hence, $P_{\{1,\dots,n\}}$ is the single solution set maximizing the hypervolume indicator on f'. Moreover, for ε' sufficiently small we changed the coordinates of r by at most $\varepsilon/2$, which implies that the additive approximation ratio $\alpha^+(f',P')$ differs from $\alpha^+(f,P)$ by at most ε. □

Note that the same proof works in the multiplicative instead of the additive setting. Hence the same result holds for α^*_{HYP}. Also note that the relation between multiplicative and additive approximation from Section 5.5 carries the above result over to α^*_{logHYP}.

5. Proofs for the approximation ratio

5.1. Tight bound for α^*_{OPT}

In this section we examine the optimal approximation ratio α^*_{OPT}. Recall that no set of n points can achieve a better approximation ratio than α^*_{OPT}. This is the reason why bounds for α^*_{OPT} are important for comparison before examining α^*_{HYP} in Section 5.2.

Proof of Theorem 2.3. We want to show $\alpha^*_{OPT} = \min\{(A/a)^{1/n}, (B/b)^{1/n}\}$. For this, we first show $\alpha^*_{OPT} \leqslant (A/a)^{1/n}$. Let $\alpha := (A/a)^{1/n}$ and $x_i := a\alpha^{i-1}$ for $i \in \{1, \dots, n\}$. The solution set $\{(x_i, f(x_i)) \mid i \in \{1, \dots, n\}\}$ is an α-approximation of f as we have $x \leqslant \alpha x_i$, $f(x) \leqslant f(x_i)$ for any $x_i \leqslant x \leqslant \alpha x_i$. Hence, $\alpha^*_{OPT} \leqslant \alpha = (A/a)^{1/n}$.

To show that analogously $\alpha^*_{OPT} \leqslant (B/b)^{1/n}$, let $\alpha := (B/b)^{1/n}$ and $x_i := f^{-1}(B\alpha^{-i})$ for $i \in \{1, \dots, n\}$. Then $f(x_i) \geqslant B\alpha^{-i}$ and no point $(x, f(x))$ has $f(x_i) > f(x) > B\alpha^{-i}$. Hence, we have $x \leqslant x_i$, $f(x) \leqslant \alpha f(x_i)$ for any x with $B\alpha^{-i} \leqslant f(x) \leqslant B\alpha^{-i+1}$. Thus, we get $\alpha^*_{OPT} \leqslant \alpha = (B/b)^{1/n}$.

It remains to prove the lower bound $\alpha^*_{OPT} \geqslant \min\{A/a, B/b\}^{1/n}$. For this, we set $f(x) := B(B/b)^{-i/n}$ for $a(A/a)^{(i-1)/n} < x \leqslant a(A/a)^{i/n}$ and $i \in \{0, \dots, n\}$. Then f is a front which consists of $(n+1)$ levels. It is illustrated in Fig. 3. Let us now consider a solution set $(x_1, \dots, x_n)$ consisting of n points. As f has $n+1$ levels, the pigeonhole principle gives that there is at least one level having none of the n points. This implies that the rightmost point in this level is only approximated by a ratio of $\min\{(A/a)^{1/n}, (B/b)^{1/n}\}$. □

Proof of Corollary 2.4. Both inequalities follow directly from Theorem 2.3. For the first inequality note that $e^x \geqslant 1 + x$ for all $x \in \mathbb{R}$. For the second we upper bound e^x with $0 \leqslant x \leqslant \varepsilon$ by

$$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!} \leqslant 1 + \sum_{k=1}^{\infty} \frac{x^k}{2^{k-1}} \leqslant 1 + x \sum_{k=0}^{\infty} \frac{\varepsilon^k}{2^k}$$
$$= 1 + x \frac{1}{1 - \varepsilon/2} \leqslant 1 + (1+\varepsilon)x,$$

as $(1+\varepsilon)(1-\varepsilon/2) \geqslant 1$. □

5.2. Upper bound for α^*_{HYP}

In this section we give bounds on the multiplicative approximation ratio achieved by the sets maximizing the hypervolume indicator.

Let P be a solution set maximizing HYP on a front $f \in \mathcal{F}$, i.e., $P \in \mathcal{P}^f_{HYP}$, and let $n > 4$ be fixed. Assume that there are points $p, q \in P$ with $p \prec q$. Such a "redundant" set can maximize HYP only on degenerate fronts: If there is a point $r = (x, f(x))$ on the front which is not dominated by any point in P, then $P' := P + r - p$ would have $\text{HYP}(P') > \text{HYP}(P)$, as it dominates all the space P dominates united with the space r dominates. Thus, there is no such point r and P already dominates the whole front. In this case the approximation ratio $\alpha^*(f, P) = 1$ and the inequality we want to show holds trivially. This can only happen for f being a step function with less than n steps. In the same way we can exclude $P = \{p_1, \dots, p_n\}$ having less than n pairwise different points.

Hence, for the rest of the proof we can assume that there are no points $p, q \in P$ with $p \prec q$. Then we can write $P = \{p_1, \dots, p_n\}$, $p_i = (x_i, y_i)$ with $a \leqslant x_1 < \dots < x_n \leqslant A$ and $B \geqslant y_1 > \dots > y_n \geqslant b$. Furthermore, we can assume that $y_i = f(x_i)$ as otherwise $P - p_i + p'_i$ with $p'_i = (x_i, f(x_i))$ would have a larger hypervolume than P.

Now recall that the *contribution* of a point $p \in P$ to the hypervolume of a solution set $P \in \mathcal{P}$ is the volume dominated by p and no other element of P (see, e.g., [6]). More formally, the contribution of a point p is

$$\textsc{Con}_P(p) := \text{HYP}(P, R) - \text{HYP}(P - p, R).$$

In the following we mainly deal with the minimal contribution defined as

$$\textsc{MinCon}(P) := \min_{2 \leqslant i \leqslant n-1} \textsc{Con}_P(p_i)$$
$$= \min_{2 \leqslant i \leqslant n-1} (x_i - x_{i-1})\big(f(x_i) - f(x_{i+1})\big).$$

Fig. 4 gives an illustration of MinCon. Note that the above definition of $\textsc{MinCon}(P)$ is independent of the reference point R, as it only considers the minimal contribution of any of the points $p_2, \dots, p_{n-1}$. Restricted to these $(n-2)$ inner points, it corresponds to the definition of $\textsc{MinCon}(P)$ in [9].

We first show the following upper bound for $\textsc{MinCon}(P)$.

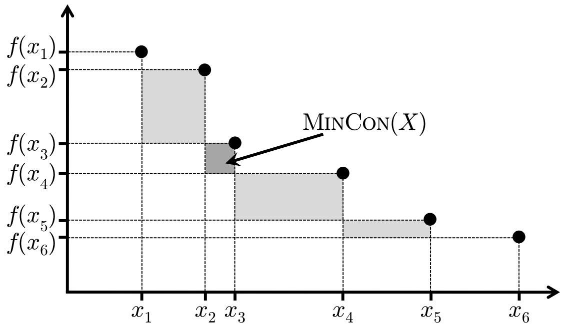


Fig. 4. The minimal contribution MinCon(X) of a solution set $X = \{p_1, p_2, \ldots, p_6\}$ is defined to be the least hypervolume contribution HYP$(P) -$ HYP$(P - p)$ for $p \in \{p_2, p_3, p_4, p_5\}$.

Lemma 5.1. *We have*

$$\text{MinCon}(P) \leqslant \frac{(x_n - x_1)\,(f(x_1) - f(x_n))}{(n-2)^2}.$$

Proof. Let $a_i := x_i - x_{i-1}$ for $2 \leqslant i \leqslant n$ and $b_i := f(x_i) - f(x_{i+1})$ for $1 \leqslant i \leqslant n-1$. This gives $\text{MinCon}(P) = \min_{2 \leqslant i \leqslant n-1} a_i b_i$ and

$$a_i \geqslant \text{MinCon}(P)/b_i \quad \text{for all } 2 \leqslant i \leqslant n-1.$$

This implies

$$\begin{aligned}\sum_{i=2}^{n-1} \text{MinCon}(P)/b_i &\leqslant \sum_{i=2}^{n-1} a_i \leqslant \sum_{i=2}^{n} a_i \\ &= \sum_{i=2}^{n} x_i - \sum_{i=1}^{n-1} x_i = x_n - x_1,\end{aligned}$$

and therefore

$$\text{MinCon}(P) \leqslant \frac{x_n - x_1}{\sum_{i=2}^{n-1} 1/b_i}.$$

We can now use the fact that the harmonic mean is less than the arithmetic mean, that is,

$$\frac{n-2}{\sum_{i=2}^{n-1} 1/b_i} \leqslant \frac{\sum_{i=2}^{n-1} b_i}{n-2}$$

to obtain

$$\begin{aligned}\text{MinCon}(P) &\leqslant \frac{(x_n - x_1)\sum_{i=2}^{n-1} b_i}{(n-2)^2} \\ &\leqslant \frac{(x_n - x_1)\,(f(x_1) - f(x_n))}{(n-2)^2}. \qquad \square\end{aligned}$$

To upper bound α^*_{HYP} we first calculate the approximation ratio of the "inner points", i.e., points $x \in [x_1, x_n]$. In a second step we determine how well the "outer points" x with $x < x_1$ or $x > x_n$ are approximated.

Lemma 5.2. *The solution set P achieves a*

$$1 + \frac{\sqrt{A/a} + \sqrt{B/b}}{n-4}$$

multiplicative approximation of all points $(x, f(x))$ with $x \in [x_1, x_n]$.

Proof. Assume there is a point $r = (x, f(x))$ which is not approximated by a ratio of

$$\alpha := 1 + \frac{\sqrt{A/a} + \sqrt{B/b}}{n-4}. \tag{5.1}$$

Let i be such that $x_i < x < x_{i+1}$ and therefore

$$x > \alpha x_i,$$
$$f(x) > \alpha f(x_{i+1}), \tag{5.2}$$

because r is approximated by neither p_i nor p_{i+1}.

Let p_j be a point contributing $\textsc{MinCon}(P)$ to P. As P maximizes the hypervolume indicator, we have

$$\begin{aligned}\text{HYP}(P) &\geqslant \text{HYP}(P - p_j + r)\\ &= \text{HYP}(P) - \textsc{Con}_P(p_j) + \textsc{Con}_{P-p_j+r}(r)\\ &\geqslant \text{HYP}(P) - \textsc{Con}_P(p_j) + \textsc{Con}_{P+r}(r),\end{aligned}$$

so we have

$$(x - x_i)\big(f(x) - f(x_{i+1})\big) = \textsc{Con}_{P+r}(r) \leqslant \textsc{MinCon}(P). \tag{5.3}$$

As Eq. (5.2) is equivalent with $x - x_i > (\alpha - 1)x_i$ and $f(x) - f(x_{i+1}) > (\alpha - 1) f(x_{i+1})$, Eq. (5.3) gives

$$\textsc{MinCon}(P) > (\alpha - 1)^2 x_i f(x_{i+1}). \tag{5.4}$$

For $3 \leqslant i \leqslant n-1$ we can upper bound the minimal contribution using Lemma 5.1 on the points $p_1, \dots, p_i$ by

$$\begin{aligned}\textsc{MinCon}(P) &\leqslant \textsc{MinCon}\big(\{p_1, \dots, p_i\}\big)\\ &\leqslant (x_i - x_1)\big(f(x_1) - f(x_i)\big)/(i-2)^2\\ &\leqslant x_i B/(i-2)^2.\end{aligned} \tag{5.5}$$

Analogously, for $1 \leqslant i \leqslant n-3$ we can upper bound the minimal contribution using Lemma 5.1 on the points $p_{i+1}, \dots, p_n$ by

$$\textsc{MinCon}(P) \leqslant A f(x_{i+1})/(n-i-2)^2. \tag{5.6}$$

Combining Eq. (5.4) with Eqs. (5.5) and (5.6), it follows for $3 \leqslant i \leqslant n-3$ that

$$(\alpha - 1)^2 x_i f(x_{i+1}) < \min\left\{\frac{x_i B}{(i-2)^2}, \frac{A f(x_{i+1})}{(n-i-2)^2}\right\} \tag{5.7}$$

or, equivalently,

$$\alpha < 1 + \min\left\{\frac{\sqrt{B/f(x_{i+1})}}{i-2}, \frac{\sqrt{A/x_i}}{n-i-2}\right\}$$

which yields with $x_i \geqslant a$ and $f(x_{i+1}) \geqslant b$ that

$$\alpha < 1 + \min\left\{\frac{\sqrt{B/b}}{i-2}, \frac{\sqrt{A/a}}{n-i-2}\right\} \tag{5.8}$$

for $3 \leqslant i \leqslant n-3$.

Now, the right hand side of Eq. (5.8) becomes maximal if the two terms are equal since one of them is monotonically increasing in i and the other one is monotonically decreasing in i. As this happens exactly for $i = 2 + \frac{(n-4)\sqrt{B/b}}{\sqrt{A/a}+\sqrt{B/b}}$, we get the upper bound

$$\alpha < 1 + \frac{\sqrt{A/a} + \sqrt{B/b}}{n-4}$$

for $3 \leqslant i \leqslant n-3$. This contradicts Eq. (5.1) and proves that every point $(x, f(x))$ with $x \in [x_3, x_{n-2}]$ is multiplicatively approximated by a ratio of α.

It remains to show a contradiction to Eq. (5.1) for $i = 1, 2$ and $i = n-2, n-1$. For $i = 1, 2$ we get from Eqs. (5.4) and (5.6)

$$\alpha < 1 + \frac{\sqrt{A/a}}{n-i-2} \leqslant 1 + \frac{\sqrt{A/a}}{n-4}$$

which contradicts Eq. (5.1).

Finally, for $i = n-2, n-1$ we get from Eqs. (5.4) and (5.5)

$$\alpha < 1 + \frac{\sqrt{B/b}}{i-2} \leqslant 1 + \frac{\sqrt{B/b}}{n-4}$$

which also contradicts Eq. (5.1) and finishes the proof. □

It remains to examine the approximation ratio of the "outer points" x with $x < x_1$ or $x > x_n$.

Lemma 5.3. *The solution set P achieves a*

$$1 + \frac{A}{(a-R_x)(n-2)^2}$$

multiplicative approximation of all points $(x, f(x))$ with $x < x_1$, and a

$$1 + \frac{B}{(b-R_y)(n-2)^2}$$

multiplicative approximation of all points $(x, f(x))$ with $x > x_n$.

Proof. We show only the statement for $x \leqslant x_1$. The case $x \geqslant x_n$ follows by symmetry in the two objectives.

The approximation ratio of any $x \leqslant x_1$ is exactly $\min\{x/x_1, f(x)/f(x_1)\} = f(x)/f(x_1)$. This is maximized for $x = a$, so that the approximation ratio of any $x \leqslant x_1$ is at most $B/f(x_1)$. We show that $B/f(x_1)$ is less than $1 + \frac{A}{(a-R_x)(n-2)^2}$.

Using Lemma 5.1 on the points $p_1, \ldots, p_n$ we get

$$\begin{aligned}\textsc{MinCon}(P) &\leqslant (x_n - x_1)\big(f(x_1) - f(x_n)\big)/(n-2)^2 \\ &\leqslant Af(x_1)/(n-2)^2. \end{aligned} \tag{5.9}$$

Let p_j be a point contributing $\textsc{MinCon}(P)$ to P and consider $P' := P - p_j + q$ with $q = (a, B)$. We have

$$\begin{aligned}\mathrm{HYP}\big(P'\big) &= \mathrm{HYP}(P) - \textsc{Con}_P(p_j) + \textsc{Con}_{P-p_j+q}(q) \\ &\geqslant \mathrm{HYP}(P) - \textsc{Con}_P(p_j) + \textsc{Con}_{P+q}(q) \\ &= \mathrm{HYP}(P) - \textsc{MinCon}(P) + (a - R_x)\big(B - f(x_1)\big).\end{aligned}$$

Together with $\mathrm{HYP}(P) \geqslant \mathrm{HYP}(P')$ this yields

$$\textsc{MinCon}(P) \geqslant (a - R_x)\big(B - f(x_1)\big). \tag{5.10}$$

Combining Eqs. (5.9) and (5.10), we finally get the desired

$$\frac{B}{f(x_1)} \leqslant 1 + \frac{A}{(a-R_x)(n-2)^2}. \qquad \square$$

Together Lemmas 5.2 and 5.3 directly imply the following theorem.

Theorem 5.4. *Let $f \in \mathcal{F}$, $n > 4$, and let $R = (R_x, R_y)$ be the reference point. Then*

$$\alpha^*_{HYP} \leqslant 1 + \max\left\{\frac{\sqrt{A/a} + \sqrt{B/b}}{n-4}, \frac{A}{(a-R_x)(n-2)^2}, \frac{B}{(b-R_y)(n-2)^2}\right\}.$$

For sufficiently large n *or* sufficiently small coordinates of the reference point, the two last terms in Theorem 5.4 are less than the first one. This proves the slightly simplified Theorem 3.1.

5.3. Lower bound for α^*_{HYP}

In this section we show that the upper bound of Theorem 3.1 is nearly tight. We restrict ourselves to the case of $\frac{A}{a} = \frac{B}{b}$. We show that in this situation we have $\alpha^*_{HYP} \geqslant 1 + \frac{\sqrt{A/a}}{cn}$ for some small constant c. Thus, the bounds are tight if $\frac{A}{a} \approx \frac{B}{b}$, except for the factor $2c$.

Proof of Theorem 3.2. As rescaling does not change any multiplicative approximation, we can assume w.l.o.g. that $a = b = 1$ and $A = B \geqslant 13$.

We set $k := \lceil n/2 \rceil$ and define $x_0, \dots, x_n$ as follows,

$$x_i := 1 + \frac{i}{2(k-1)} \quad \text{for } i = 0, \dots, k-1,$$
$$x_i := \frac{i-k}{n-k-1} \cdot \frac{A}{2} + \frac{n-i-1}{n-k-1}\left(\frac{3}{2} + \frac{\sqrt{A-1}}{n-2}\right) \quad \text{for } i = k, \dots, n-1,$$
$$x_i := A \quad \text{for } i = n.$$

To simplify the notation we further set $x_{-1} := 0$. With this we can calculate that

$$x_i - x_{i-1} = 1 \quad \text{for } i = 0,$$
$$x_i - x_{i-1} = \frac{1}{2(k-1)} \quad \text{for } i = 1, \dots, k-1,$$
$$x_i - x_{i-1} = \frac{\sqrt{A-1}}{n-2} \quad \text{for } i = k,$$
$$x_i - x_{i-1} = \frac{1}{n-k-1}\left(\frac{A}{2} - \frac{3}{2} - \frac{\sqrt{A-1}}{n-2}\right) \quad \text{for } i = k+1, \dots, n-1,$$
$$x_i - x_{i-1} = \frac{A}{2} \quad \text{for } i = n.$$

This implies $x_0 \leqslant \dots \leqslant x_n$. To confirm this for $i = k+1, \dots, n-1$, observe that

$$\frac{3}{2} + \frac{1}{n-2}\sqrt{A-1} \leqslant \frac{3}{2} + \frac{1}{2}\sqrt{A-1} \leqslant \frac{A}{4}, \tag{5.11}$$

as $n \geqslant 4$ and $A \geqslant 13$.

Moreover, we define, for $x_{i-1} < x \leqslant x_i$,

$$f(x) := \frac{\sum_{j=0}^{i-1} 1/(x_j - x_{j-1}) + A \sum_{j=i}^{n-1} 1/(x_j - x_{j-1})}{\sum_{j=0}^{n-1} 1/(x_j - x_{j-1})}.$$

This way, $f(x_0) = A = B$, $f(x_n) = 1 = b$ and all other function values are in between. We define solution sets $P_i := (p_0, \dots, p_{i-1}, p_{i+1}, \dots, p_n)$, $i = 0, \dots, n$, with n points in each solution set, $p_i = (x_i, f(x_i))$. We further define a solution set $P := (p_0, \dots, p_n)$ which has $n+1$ points and is therefore not among the possible solution sets maximizing HYP on f that we consider.

Let $Q = (q_1, \dots, q_n)$ be a solution set of n points which maximizes HYP. As P contains all non-dominated points, it also dominates Q. Moreover, by the pigeonhole principle there must be an $0 \leqslant i \leqslant n$ such that no q_j is contained in the set $(x_{i-1}, x_i] \times (f(x_{i+1}), f(x_i)]$ (where we set $f(x_{n+1})$ to a). But then P_i dominates Q, which implies $P_i = Q$, as otherwise $\text{HYP}(P_i)$ would be greater than $\text{HYP}(Q)$. Hence, the solution sets maximizing HYP are among the P_i.

We will determine the solution sets maximizing HYP by comparing $\textsc{Con}(p_i) := \textsc{Con}_P(p_i)$. The solution sets P_i minimizing $\textsc{Con}(p_i)$ are the sets maximizing HYP. We show that $\min_{0 \leqslant j \leqslant n} \textsc{Con}(p_j) = \textsc{Con}(p_i)$ for $1 \leqslant i < n$. To see this, we first examine $\textsc{Con}(p_i)$ for $1 \leqslant i < n$:

$$\begin{aligned}
\textsc{Con}(p_i) &= (x_i - x_{i-1})\big(f(x_i) - f(x_{i+1})\big) \\
&= \frac{A-1}{\sum_{j=0}^{n-1} 1/(x_j - x_{j-1})} \\
&< \frac{A-1}{\sum_{j=0}^{k-1} 1/(x_j - x_{j-1})} \\
&= \frac{A-1}{1 + 2(k-1)^2} < A/2.
\end{aligned} \tag{5.12}$$

Note that Eq. (5.12) is independent of i. We can further bound $\textsc{Con}(p_0)$ by

$$\begin{aligned}
\textsc{Con}(p_0) &= (x_0 - R_x)\big(f(x_0) - f(x_1)\big) \\
&\geqslant (x_0 - 0)\big(f(x_0) - f(x_1)\big) \\
&= (x_0 - x_{-1})\big(f(x_0) - f(x_1)\big) \\
&= \frac{A-1}{\sum_{j=0}^{n-1} 1/(x_j - x_{j-1})}.
\end{aligned}$$

Hence $\mathrm{Con}(p_0) \geqslant \mathrm{Con}(p_i)$ for $1 \leqslant i < n$. For $\mathrm{Con}(p_n)$ we get

$$\begin{aligned}\mathrm{Con}(p_n) &= (x_n - x_{n-1})\big(f(x_n) - R_y\big)\\ &= \left(A - \frac{A}{2}\right)(b - R_y) \geqslant A/2.\end{aligned}$$

Therefore also $\mathrm{Con}(p_i) < \mathrm{Con}(p_n)$ for $1 \leqslant i < n$.

This shows that the sets P_i, $1 \leqslant i < n$, and maybe P_0 maximize HYP on the front f. We now slightly adjust f such that the solution set maximizing HYP is P_k. For this we define $\widehat{f}$ to be

$$\widehat{f}(x) := \begin{cases} f(x) - \varepsilon & \text{for } x_{k-1} < x \leqslant x_k, \\ f(x) & \text{otherwise} \end{cases}$$

for a small $\varepsilon > 0$. If we go from f to $\widehat{f}$, then only $\mathrm{Con}(p_k)$ decreases and thus P_k is the solution set maximizing HYP for $\widehat{f}$.

Now we consider the approximation ratio $\alpha(\widehat{f}, X_k)$. It yields a lower bound for α^*_{HYP}. We have

$$\alpha(\widehat{f}, X_k) \geqslant \min\left\{\frac{x_k}{x_{k-1}}, \frac{\widehat{f}(x_k)}{\widehat{f}(x_{k+1})}\right\}.$$

The latter term goes to $\frac{f(x_k)}{f(x_{k+1})}$ for $\varepsilon \to 0$. Hence, we have

$$\alpha^*_{HYP} \geqslant \min\left\{\frac{x_k}{x_{k-1}}, \frac{f(x_k)}{f(x_{k+1})}\right\}. \tag{5.13}$$

By definition of x_i, the first term is

$$\frac{x_k}{x_{k-1}} = 1 + \frac{2\sqrt{A-1}}{3(n-2)}. \tag{5.14}$$

The second term of Eq. (5.13) is

$$\begin{aligned}\frac{f(x_k)}{f(x_{k+1})} &= \frac{\sum_{j=0}^{k-1} 1/(x_j - x_{j-1}) + A\sum_{j=k}^{n-1} 1/(x_j - x_{j-1})}{\sum_{j=0}^{k} 1/(x_j - x_{j-1}) + A\sum_{j=k+1}^{n-1} 1/(x_j - x_{j-1})}\\ &= 1 + \frac{(A-1)/(x_k - x_{k-1})}{\sum_{j=0}^{k} 1/(x_j - x_{j-1}) + A\sum_{j=k+1}^{n-1} 1/(x_j - x_{j-1})}.\end{aligned} \tag{5.15}$$

The nominator of the last fraction is $\sqrt{A-1}\,(n-2)$. The denominator can be bounded by

$$\begin{aligned}&1 + 2(k-1)^2 + \frac{n-2}{\sqrt{A-1}} + A\left(\frac{(n-k-1)^2}{A/2 - \frac{3}{2} - \frac{\sqrt{A-1}}{n-2}}\right)\\ &\quad\leqslant 1 + 2(k-1)^2 + (n-2)/\sqrt{A-1} + 4(n-k-1)^2\end{aligned} \tag{5.16}$$

where the last inequality is based on Eq. (5.11).

Note that $k = \lceil n/2 \rceil$ is either $n/2$, or $(n+1)/2$. In both cases,

$$2(k-1)^2 + 4(n-k-1)^2 \leqslant \frac{3}{2}(n-2)^2,$$

as $n \geqslant 4$.

With this we can upper bound Eq. (5.16) by

$$\begin{aligned}(5.16) &\leqslant 1 + (n-2)/\sqrt{A-1} + \frac{3}{2}(n-2)^2\\ &\leqslant \frac{3}{2}(n-1)(n-2)\end{aligned}$$

where the last inequality uses $n \geqslant 4$ and $A \geqslant 13$. Plugging these bounds into Eq. (5.15), we get

$$\frac{f(x_k)}{f(x_{k+1})} \geqslant 1 + \frac{\sqrt{A-1}(n-2)}{\frac{3}{2}(n-1)(n-2)} = 1 + \frac{2\sqrt{A-1}}{3(n-1)}.$$

Plugging this and Eq. (5.14) in Eq. (5.13) gives

$$\alpha^*_{HYP} \geqslant 1 + \frac{2\sqrt{A-1}}{3(n-1)},$$

which proves the claim. □

5.4. Upper bound for α^+_{HYP}

Having shown in the previous section that sets maximizing the hypervolume indicator have a suboptimal multiplicative approximation ratio in the worst case, we now analyze their additive approximation properties by proving an upper bound for α^+_{HYP}.

Proof of Theorem 3.3. We want to prove that $\alpha^+_{HYP} \leqslant \frac{\sqrt{(A-a)(B-b)}}{n-2}$ for $n > 2$ and $(n-2)\min\{a-R_x, b-R_y\} \geqslant \sqrt{(A-a)(B-b)}$.

Let $P \in \mathcal{P}^f_{HYP}$. As in the beginning of Section 5.2, we can assume that there are no points $p, q \in P$ with $p \prec q$ and that we can write $P = \{p_1, \dots, p_n\}$, $p_i = (x_i, f(x_i))$ with $a \leqslant x_1 < \cdots < x_n \leqslant A$ and $B \geqslant f(x_1) > \cdots > f(x_n) \geqslant b$, as otherwise we have $\alpha^+(f, P) = 1$.

Let $r = (x, f(x))$, $x \in [a, A]$ be an arbitrary point and let $\alpha > 0$ be such that r is not additively approximated by α. We make a case distinction depending on the position of r. Let us first assume that r is an "inner point", i.e., there is an $i \in \{1, \dots, n-1\}$ with $x_i \leqslant x < x_{i+1}$. As r is not additively approximated by α, we have

$$x > x_i + \alpha \quad \text{and} \quad f(x) > f(x_{i+1}) + \alpha. \tag{5.17}$$

As P maximizes the hypervolume indicator on f, replacing the point $p \in P$ contributing $\textsc{MinCon}(P)$ to P with the point r must not increase the hypervolume. Therefore,

$$\begin{aligned}\mathrm{HYP}(P) &\geqslant \mathrm{HYP}(P + r - p)\\ &= \mathrm{HYP}(P) - \textsc{Con}_P(p) + \textsc{Con}_{P+r-p}(r)\\ &\geqslant \mathrm{HYP}(P) - \textsc{Con}_P(p) + \textsc{Con}_{P+r}(r).\end{aligned}$$

This in turn implies

$$\begin{aligned}\textsc{MinCon}(P) &= \textsc{Con}_P(p) \geqslant \textsc{Con}_{P+r}(r)\\ &= (x - x_i)\big(f(x) - f(x_{i+1})\big) \overset{(5.17)}{>} \alpha^2.\end{aligned}$$

Using Lemma 5.1 and taking square roots on both sides gives the desired

$$\alpha < \frac{\sqrt{(A-a)(B-b)}}{n-2}.$$

It remains to study the case where $r = (x, f(x))$ is an "outer point" with $x \leqslant x_1$ or $x \geqslant x_n$. It suffices to examine $x \leqslant x_1$, as then the case $x \geqslant x_n$ follows by symmetry in the two objectives.

As r is not approximated by a ratio of α we have $f(x) > f(x_1) + \alpha$. Additionally, replacing the point $p \in P$ contributing $\textsc{MinCon}(P)$ to P by r must not increase the hypervolume, so we have

$$\begin{aligned}\textsc{MinCon}(P) &\geqslant \textsc{Con}_{P+r-p}(r) \geqslant \textsc{Con}_{P+r}(r)\\ &\geqslant (a - R_x)\big(f(x) - f(x_1)\big)\\ &\geqslant (a - R_x)\alpha.\end{aligned}$$

We use Lemma 5.1 again and get

$$\alpha \leqslant \frac{(A-a)(B-b)}{(a-R_x)(n-2)^2} \leqslant \sqrt{(A-a)(B-b)}/(n-2),$$

where the second inequality follows from the assumption of the theorem. □

Closely examining the above proof of Theorem 3.3, we see that it also gives an upper bound on the additive approximation ratio for solution sets P that are a *local maximum* for HYP, that is, where for all points $p \in P$ and $q = (x, y) \in [a, A] \times [b, B]$ with $y \leqslant f(x)$, we have $\mathrm{HYP}(P + q - p) \leqslant \mathrm{HYP}(P)$.

5.5. Tight bound for α^+_{OPT}

In this section we describe a relation that allows us to transfer results on multiplicative approximation into results on additive approximation and the other way around. This proves Theorems 2.8 and 3.5 and gives the intuition behind the logarithmic hypervolume indicator, as it is the standard hypervolume indicator transferred into the world of multiplicative approximation.

Consider a front[9] $f^* \in \mathcal{F}_{[a^*,A^*]\to[b^*,B^*]}$ and a solution set $P^* \in \mathcal{P}^{f^*}$ that is a multiplicative α^*-approximation of f^*. This means that we have for any $\hat{x}^* \in [a^*, A^*]$ a point $(x^*, y^*) \in P^*$ with

$$\hat{x}^* \leqslant \alpha^* x^* \quad \text{and} \quad f^*(\hat{x}^*) \leqslant \alpha^* y^*.$$

Logarithmizing both inequalities gives

$$\log \hat{x}^* \leqslant \log x^* + \log \alpha^* \quad \text{and} \quad \log f^*(\hat{x}^*) \leqslant \log y^* + \log \alpha^*.$$

This corresponds to an additive approximation. We set $x^+ := \log x^*$, $y^+ := \log y^*$, $\hat{x}^+ := \log \hat{x}^*$, $\alpha^+ := \log \alpha^*$ and $f^+ := \log \circ f^* \circ \exp$ and get

$$\hat{x}^+ \leqslant x^+ + \alpha^+ \quad \text{and} \quad f^+(\hat{x}^+) \leqslant y^+ + \alpha^+.$$

This means that $P^+ := \{(\log x, \log y) \mid (x, y) \in P^*\}$ is an additive α^+-approximation of the front $f^+ \in \mathcal{F}_{[a^+,A^+]\to[b^+,B^+]}$ with $a^+ = \log a^*$, $A^+ = \log A^*$, $b^+ = \log b^*$, $B^+ = \log B^*$. Observe that this corresponds to logarithmizing both axes.

All operations we used above are invertible, so that we can do the same thing the other way round: Having a solution set P^+ on a front f^+ achieving an additive α^+-approximation, we get a solution set $P^* = \{(\exp x, \exp y) \mid (x, y) \in P^+\}$ on a front $f^* = \exp \circ f^+ \circ \log$ achieving a multiplicative α^*-approximation, with $\alpha^* = \exp \alpha^+$. Thereby the interval bounds like a^+ are also exponentiated and we get $a^* = \exp a^+$.

Hence, we have a bijection[10] $\mathcal{F}^* \to \mathcal{F}^+$, $f^* \mapsto f^+$ and for any $f^* \in \mathcal{F}^*$ a bijection $\mathcal{P}^{f^*} \to \mathcal{P}^{f^+}$, $P^* \mapsto P^+$ that satisfies $\alpha^+(f^+, P^+) = \log \alpha^*(f^*, P^*)$.

This allows us to prove Theorem 2.8 by transferring Theorem 2.3 to the world of additive approximation:

Proof of Theorem 2.8. We want to prove $\alpha^+_{OPT} = \min\{A^+ - a^+, B^+ - b^+\}/n$. By definition and the above bijection $(*)$ we know that

$$\begin{aligned}
\alpha^+_{OPT} &= \sup_{f^+ \in \mathcal{F}^+} \inf_{P^+ \in \mathcal{P}^{f^+}} \alpha^+(f^+, P^+) \\
&\overset{(*)}{=} \sup_{f^+ \in \mathcal{F}^+} \inf_{P^+ \in \mathcal{P}^{f^+}} \log \alpha^*(f^*, P^*) \\
&\overset{(*)}{=} \sup_{f^* \in \mathcal{F}^*} \inf_{P^* \in \mathcal{P}^{f^*}} \log \alpha^*(f^*, P^*) \\
&= \log \sup_{f^* \in \mathcal{F}^*} \inf_{P^* \in \mathcal{P}^{f^*}} \alpha^*(f^*, P^*).
\end{aligned}$$

The last expression matches the definition of α^*_{OPT}. We replace α^*_{OPT} using Theorem 2.3 and a^* by $\exp a^+$ etc. and get

$$\begin{aligned}
\alpha^+_{OPT} &= \log \alpha^*_{OPT} \\
&= \log\left(\min\{A^*/a^*, B^*/b^*\}^{1/n}\right) \\
&= \min\{\log A^* - \log a^*, \log B^* - \log b^*\}/n \\
&= \min\{A^+ - a^+, B^+ - b^+\}/n. \qquad \square
\end{aligned}$$

5.6. Upper bound for $\alpha^*_{logHYP}\alpha^*_{logHYP}$

With similar reasoning we can now also prove Theorem 3.5.

Proof of Theorem 3.5. We want to show that

$$\alpha^*_{logHYP} \leqslant \exp\left(\frac{\sqrt{\log(A^*/a^*)\log(B^*/b^*)}}{n-2}\right).$$

For a solution set $P^* \in \mathcal{P}^*$ and a reference point $R^* = (R^*_x, R^*_y)$, $R^*_x, R^*_y > 0$ we defined LOGHYP by setting LOGHYP$(P^*, R^*) =$ HYP$_{\mathbf{1}}(\log P^*, \log R^*)$ with $\log P^* = \{(\log x, \log y) \mid (x, y) \in P^*\}$ and $\log R^* = (\log R^*_x, \log R^*_y)$. This $\log P^*$ is exactly P^+ as defined above. Writing $R^+ := \log R^*$ we thus have LOGHYP$(P^*, R^*) =$ HYP(P^+, R^+). Now, consider a solution set P^* maximizing LOGHYP(P^*, R^*), thus also maximizing HYP(P^+, R^+). We know that P^+ is an α^+_{HYP}-approximation of the front f^+, so using Theorem 3.3 and above bijections we get

[9] In this section we will mark every variable with a $+$ or $*$ depending on whether it belongs to the additive or multiplicative approximation.

[10] We write for short $\mathcal{F}^* = \mathcal{F}_{[a^*,A^*]\to[b^*,B^*]}$ and $\mathcal{F}^+ = \mathcal{F}_{[a^+,A^+]\to[b^+,B^+]}$.

$$\begin{aligned}\alpha^*(f^*, P^*) &= \exp \alpha^+(f^+, P^+)\\ &\leqslant \exp\left(\sqrt{(A^+ - a^+)(B^+ - b^+)}/(n-2)\right)\\ &= \exp\left(\sqrt{\log(A^*/a^*)\log(B^*/b^*)}/(n-2)\right).\end{aligned}$$

The observation that the assumption of Theorem 3.3 transforms directly into the assumption of Theorem 3.5 concludes the proof. □

Note that we could also have proceeded the other way round: proving a bound for α^*_{logHYP} and transforming it into a result for α^+_{HYP}. The above proof also makes clear why we defined LOGHYP as we did, as maximizing $\mathrm{HYP}(P^+, R^+)$ gives a good additive approximation which transforms into a good multiplicative approximation going back to P^*.

6. Minimization problems

All previous results a priori hold only for maximization problems. In this section we sketch how to adjust the definitions of Section 2 to minimization problems and what bounds hold in this case. Note that minimization and maximization are not isomorphic regarding multiplicative approximation.

6.1. Changes in the definitions and results

The main change in the definitions is reversing the direction of several inequalities. First, in Section 2 we have to require the fronts we consider to be lower semi-continuous instead of upper semi-continuous. We then set $f^{-1}(y) = \min\{x \in [a, A] \mid f(x) \leqslant y\}$. Moreover, a solution set P is called feasible for the front f if $y \geqslant f(x)$ for any $(x, y) \in P$.

For the definition of multiplicative and additive approximation in Sections 2.1 and 2.2 we have to change the inequalities (2.1) and (2.2) to

$$\hat{x} \geqslant x/\alpha \quad \text{and} \quad f(\hat{x}) \geqslant y/\alpha, \quad \text{and}$$
$$\hat{x} \geqslant x - \alpha \quad \text{and} \quad f(\hat{x}) \geqslant y - \alpha,$$

respectively. The remainder of Sections 2.1 and 2.2 also holds for minimization problems as it is written there. This is even the case for the results on α^*_{OPT} (Theorem 2.3 and Corollary 2.4) and α^+_{OPT} (Theorem 2.8). The following Section 6.2 describes how to derive these results.

Let us now go through the results presented in Section 3 for maximization problems and translate them to minimization. In Section 3.1 in the definition of the weighted hypervolume indicator, we now require the reference point to lie above all feasible points, i.e., $R \succcurlyeq (A, B)$ instead of $R \preccurlyeq (a, b)$. If we then define the attainment function $A_{P,R}$ by $A_{P,R}(x, y) = 1$, if $(R_x, R_y) \succcurlyeq (x, y)$ and there is a $p = (p_x, p_y) \in P$ such that $(x, y) \succcurlyeq (p_x, p_y)$, and $A_{P,R}(x, y) = 0$ otherwise, we get a meaningful definition of the weighted hypervolume indicator.

The results for the standard hypervolume indicator from Section 3.2 change slightly. In the upper bound for α^*_{HYP} the assumption changes and we get the following analog of Theorem 3.1. It is proven in Section 6.3.

Theorem 6.1. *Let $f \in \mathcal{F}$, $n > 4$, and let $R = (R_x, R_y)$ be the reference point. If we have*

$$(n-2)(R_y - B) \geqslant B\sqrt{\frac{A}{a}} \quad \textit{and}$$
$$(n-2)(R_x - A) \geqslant A\sqrt{\frac{B}{b}}$$

then

$$\alpha^*_{HYP} \leqslant 1 + \frac{\sqrt{A/a} + \sqrt{B/b}}{n-4}.$$

For the lower bound of α^*_{HYP} (Theorem 3.2) the proof has to be redone. We give a different construction than in the proof of Theorem 3.2, which gives a slightly better result (in terms of the constant factor). It has the additional assumption that n is even, which is not essential but simplifies the construction. The following theorem is proven in Section 6.4.

Theorem 6.2. *Let $n \geqslant 4$ be even, $A/a = B/b > 4$ and $R = (R_x, R_y)$ be the reference point satisfying $R_x \geqslant A + \frac{2A}{n-2}$ and $R_y \geqslant B + \frac{2B}{n-2}$. Then we have*

$$\alpha^*_{HYP} \geqslant \min\left\{\frac{A}{4a}, 1 + \frac{\sqrt{A/a}}{\sqrt{2}(n-2)}\right\}.$$

In the bound for α^+_{HYP} (Theorem 3.3) the assumption changes also. For minimization problems we get the following analog result. Its proof can be found in the following Section 6.2.

Theorem 6.3. *If $n > 2$ and*

$$(n-2)\min\{R_x - A, R_y - B\} \geqslant \sqrt{(A-a)(B-b)}$$

we have

$$\alpha^+_{HYP} \leqslant \frac{\sqrt{(A-a)(B-b)}}{n-2}.$$

It remains to translate the logarithmic hypervolume indicator as introduced in Section 3.3. In order to adjust the definition of the logarithmic hypervolume indicator to minimization, we require $R \succcurlyeq (A, B)$, but do not have to change anything besides that. Lemma 3.4 still holds; however, the proof has to be changed slightly, as we now have to integrate over the space above the solution set and not below. The upper bounds for α^*_{logHYP} (Theorem 3.5 and Corollary 3.6) still hold; we only have to change a/R_x to R_x/A and b/R_y to R_y/B in the assumptions.

We remark that the results of Section 4 on (i) why upper (lower) semi-continuity is important and (ii) why it does not matter if we take worst or best case in the definition of α^*_{HYP}, also translate to the minimization setting. However, for reasons of brevity, we omit these details.

6.2. Bounds for α^+_{OPT}, α^*_{OPT}, α^+_{HYP}, α^*_{logHYP} (minimization)

This section sketches how the analogous results stated above can be proven based on the corresponding maximization results.

Let us start with the results on additive approximation ratios. Consider the bijection $\mathcal{F}_{[a,A]\to[b,B]} \to \mathcal{F}_{[-A,-a]\to[-B,-b]}$ we get by negating both axes, i.e., $f \mapsto f^-$ with $f^-(x) := -f(-x)$. Moreover, let $P \mapsto P^-$ with $P^- = \{(-x, -y) \mid (x, y) \in P\}$ for $P \in \mathcal{P}$. Then P^- is a feasible solution set for f^- in the minimization setting iff P is feasible for f in the maximization setting. Additionally, P maximizes the hypervolume indicator on f iff P^- maximizes the (minimization) hypervolume indicator on f^-, and the additive approximation ratio of P equals the (minimization) additive approximation factor of P^-. This map gives the desired relation between maximization and minimization problems, as long as we are dealing with additive approximation only (since the requirement $a, b > 0$ does not hold for f or f^-). Using it we can easily show the analogous results on α^+_{OPT} (Theorem 2.8) and α^+_{HYP} (Theorem 3.3), similar to the way we used the relation between multiplicative and additive approximation in Section 5.5.

For the results on α^*_{OPT} and α^*_{logHYP} (Theorems 2.3 and 3.5) we use the relation between multiplicative and additive approximation again: The bijection works word by word the same way in the minimization setting and the proof of it works as in Section 5.5 with some minor changes like swapping inequality signs. Having this, we can proceed as in Section 5.5 to show the bounds for α^*_{OPT} and α^*_{logHYP} from the bounds for α^+_{OPT} and α^+_{HYP}.

These correspondences do not help to prove Theorems 6.1 and 6.2 (which are the minimization analogs of Theorems 3.1 and 3.2). The following two sections redo their proofs for minimization problems.

6.3. Upper bound for α^*_{HYP} (minimization)

Let $f \in \mathcal{F}$ and P be a solution set maximizing the hypervolume indicator on f. As in the proof of Theorem 3.1 we can assume that $P = \{p_1, \dots, p_n\}$ with $p_i = (x_i, f(x_i))$, $a \leqslant x_1 < \cdots < x_n \leqslant A$, $B \geqslant f(x_1) > \cdots > f(x_n) \geqslant b$.

We define Con and MinCon the same way as before. Lemma 5.1 still holds, i.e.,

$$\textsc{MinCon}(P) \leqslant \frac{(x_n - x_1)(f(x_1) - f(x_n))}{(n-2)^2}.$$

This can be proven by redoing the proof of Lemma 5.1 or by reducing the statement to Lemma 5.1 using the maximization–minimization bijection of the preceding Section 6.2 (mapping f to $f^-(x) := -f(-x)$).

We first calculate the approximation ratio of the "inner points" $x \in [x_1, x_n]$ analogously to Lemma 5.2.

Lemma 6.4. *The solution set P achieves a*

$$1 + \frac{\sqrt{A/a} + \sqrt{B/b}}{n-4}$$

multiplicative approximation of all points $(x, f(x))$ with $x \in [x_1, x_n]$.

Proof. We have to change only a few lines of the proof of Lemma 5.2. Eq. (5.2) now gives

$$x < x_i/\alpha,$$
$$f(x) < f(x_{i+1})/\alpha,$$

Eq. (5.3) gives

$$(x_{i+1} - x)\big(f(x_i) - f(x)\big) = \textsc{Con}_{P+r}(r) \leqslant \textsc{MinCon}(P),$$

Eq. (5.4) gives

$$\textsc{MinCon}(P) > (\alpha - 1)^2 x f(x),$$

and Eq. (5.7) gives

$$(\alpha - 1)^2 x f(x) < \min\left\{\frac{x_i B}{(i-2)^2}, \frac{A f(x_{i+1})}{(n-i-2)^2}\right\}.$$

As $x_i \leqslant x$ and $f(x_{i+1}) \leqslant f(x)$ this implies Eq. (5.7) and the rest of the proof works as before. □

For the outer points with $x < x_1$ or $x > x_n$ we proceed as in Lemma 5.3.

Lemma 6.5. *The solution set P achieves a*

$$1 + \frac{AB}{(R_y - B)a(n-2)^2}$$

multiplicative approximation of all points $(x, f(x))$ with $x < x_1$, and a

$$1 + \frac{AB}{(R_x - A)b(n-2)^2}$$

multiplicative approximation of all points $(x, f(x))$ with $x > x_n$.

Proof. We show only the statement for $x \leqslant x_1$. The case $x \geqslant x_n$ follows by symmetry in the two objectives.

The approximation ratio of any $x \leqslant x_1$ is exactly $\min\{x_1/x, f(x_1)/f(x)\} = x_1/x$. This is maximized for $x = a$, so that the approximation ratio of any $x \leqslant x_1$ is at most x_1/a. We show that x_1/a is less than $1 + \frac{AB}{(R_y - B)a(n-2)^2}$.

Using Lemma 5.1 on the points $p_1, \ldots, p_n$ we get

$$\begin{aligned} \textsc{MinCon}(P) &\leqslant (x_n - x_1)\big(f(x_1) - f(x_n)\big)/(n-2)^2 \\ &\leqslant AB/(n-2)^2. \end{aligned} \tag{6.1}$$

Let p_j be a point contributing $\textsc{MinCon}(P)$ to P and consider $P' := P - p_j + q$ with $q = (a, B)$. We have

$$\begin{aligned} \text{HYP}(P') &= \text{HYP}(P) - \textsc{Con}_P(p_j) + \textsc{Con}_{P-p_j+q}(q) \\ &\geqslant \text{HYP}(P) - \textsc{Con}_P(p_j) + \textsc{Con}_{P+q}(q) \\ &= \text{HYP}(P) - \textsc{MinCon}(P) + (x_1 - a)(R_y - B). \end{aligned}$$

Together with $\text{HYP}(P) \geqslant \text{HYP}(P')$ this yields

$$\textsc{MinCon}(P) \geqslant (x_1 - a)(R_y - B). \tag{6.2}$$

Combining Eqs. (6.1) and (6.2), we finally get the desired

$$x_1/a \leqslant 1 + \frac{AB}{(R_y - B)a(n-2)^2}. \qquad \square$$

After Lemmas 6.4 and 6.5, it is now an easy calculation that the approximation ratio for the outer points is smaller than $1 + \frac{\sqrt{A/a} + \sqrt{B/b}}{n-4}$ if we have

$$(n-2)(R_y - B) \geqslant B\sqrt{\frac{A}{a}}$$

and

$$(n-2)(R_x - A) \geqslant A\sqrt{\frac{B}{b}}.$$

This proves Theorem 6.1.

6.4. Lower bound for α^*_{HYP} (minimization)

Proof of Theorem 6.2. Let $n \geqslant 4$ even and set $m = n/2$. After rescaling we can assume that $A = B$ and $a = b$. We set $p_i := (x_i, y_i)$ with

$$\begin{aligned} x_i &:= a + (i-1)\varepsilon, \\ y_i &:= A - \frac{i-1}{m-1}(A - 2\alpha a), \quad \text{for } i = 1, \dots, m, \\ x_i &:= y_i := 2a, \quad \text{for } i = m+1, \\ x_i &:= y_{n+2-i}, \\ y_i &:= x_{n+2-i}, \quad \text{for } i = m+2, \dots, n+1. \end{aligned}$$

There, α is the approximation factor we will have in the end and we set $\varepsilon := \frac{4a^2(\alpha-1)^2(m-1)}{A-2\alpha a}$. Of these points we want $a \leqslant x_1 < \cdots < x_{n+1} \leqslant A$ and $A \geqslant y_1 > \cdots > y_{n+1} \geqslant a$. For this to hold, the following inequalities have to be fulfilled:

$$\begin{aligned} &\varepsilon > 0, \\ &2a > a + (m-1)\varepsilon, \\ &2\alpha a < A, \\ &\alpha > 1. \end{aligned}$$

For later use we require, additionally, $4\alpha a < A$. Plugging in the definition of ε and m these inequalities simplify to

$$\begin{aligned} &n > 2, \\ &1 < \alpha < \frac{A}{4a}, \\ &A - 2\alpha a > 4a(\alpha - 1)^2(m-1)^2. \end{aligned}$$

Using the second inequality, the third is fulfilled if we have

$$A/2 > 4a(\alpha-1)^2(m-1)^2,$$

which simplifies to

$$\alpha < 1 + \frac{\sqrt{A/a}}{\sqrt{2}(n-2)}.$$

Now we take the step function f defined by the points (x_i, y_i) as a front, i.e., we set

$$f(x) := \min\{y_i \mid i \in \{1, \dots, n+1\},\ x_i \leqslant x\}.$$

On this front the solution sets maximizing the hypervolume indicator are among the sets $P_i := \{p_j \mid 1 \leqslant j \leqslant n+1,\ \ j \neq i\}$. We set $P := \{p_j \mid 1 \leqslant j \leqslant n+1\}$ and compare the values $\text{Con}_P(p_i)$ to determine the set P_i maximizing the hypervolume indicator. We have $\text{Con}_P(p_i) = \varepsilon \cdot \frac{A-2\alpha a}{m-1}$ for $1 < i < n+1$, $i \neq m+1$ and $\text{Con}_P(p_{m+1}) = (2\alpha a - 2a)^2 = 4a^2(\alpha-1)^2$. By the definition of ε both values are equal. Moreover, by the choice of the reference point we have $\text{Con}_P(p_1) = \varepsilon \cdot (R_y - A) \leqslant \text{Con}_P(p_i)$ for any $1 < i < n+1$ and $\text{Con}_P(p_{n+1}) = \varepsilon \cdot (R_x - A) \leqslant \text{Con}_P(p_i)$ for any $1 < i < n+1$.

Hence, P_{m+1} maximizes the hypervolume indicator of f, as $\text{Con}_P(p_{m+1})$ is minimal. The approximation ratio of this solution set is $\min\{\frac{y_m}{y_{m+1}}, \frac{x_{m+2}}{x_{m+1}}\}$, which is exactly α.

By the above inequalities we can make α as large as $\min\{\frac{A}{4a}, 1 + \frac{\sqrt{A/a}}{\sqrt{2}(n-2)}\} - \varepsilon'$ for any $\varepsilon' > 0$, so by taking the supremum it follows that α^*_{HYP} is greater than equal to this value for $\varepsilon' = 0$, which proves the claim. □

7. Conclusion

We examined to what extent the goal of getting a "good approximation" of the Pareto front is reached when optimizing the hypervolume indicator. This has been done by theoretical considerations of the additive and multiplicative approximation ratio of sets of fixed size that maximize the hypervolume indicator in two dimensions on worst-case fronts. We proved that maximizing the hypervolume indicator gives a close-to-optimal additive, but no good multiplicative approximation ratio. Additionally, we introduced the logarithmic hypervolume indicator, which yields a close-to-optimal multiplicative approximation ratio.

Our results indicate for two dimensions that guiding the search with the hypervolume indicator is the right choice if one wants an additive approximation, while guiding the search with the logarithmic hypervolume indicator is the right choice if one wants a multiplicative approximation of the Pareto front. We expect similar results for higher dimensions, but a rigorous proof of this remains an open problem. The difficult part of the proof for higher dimensions will be controlling the outer points. However, the provable bounds on the approximation ratio will depend on the choice of the reference point R even more than in the two-dimensional case. Simple assumptions on R as in Theorems 3.1, 3.3 and 3.5 will not give general bounds on the approximation factor of sets maximizing the hypervolume independent of R. For an illustration on what sort of results we expect in higher dimensions, compare the assumptions and bounds of Theorems 3.1 and 5.4. We believe results like Theorem 5.4 are possible while more readable simplifications like Theorem 3.1 seem unlikely.

It is also an interesting open problem whether the approximation quality achieved by the ε-indicator can be measured in a similar manner (cf. Section 3.5). The same question can be asked for other indicator functions. This might allow a rigorous comparison between different indicators. For the weighted hypervolume indicator [34] it is obvious that regions with higher weights will be better approximated. A formal study how the weight function corresponds to the achieved approximation is another direction for future research.

References

[1] A. Auger, J. Bader, D. Brockhoff, E. Zitzler, Investigating and exploiting the bias of the weighted hypervolume to articulate user preferences, in: Proc. 11th Annual Conference on Genetic and Evolutionary Computation (GECCO'09), ACM Press, 2009, pp. 563–570.

[2] A. Auger, J. Bader, D. Brockhoff, E. Zitzler, Theory of the hypervolume indicator: Optimal μ-distributions and the choice of the reference point, in: Proc. 10th International Workshop on Foundations of Genetic Algorithms (FOGA'09), ACM Press, 2009, pp. 87–102.

[3] J. Bader, E. Zitzler, HypE: An algorithm for fast hypervolume-based many-objective optimization, Evolutionary Computation 19 (2011) 45–76.

[4] R. Berghammer, T. Friedrich, F. Neumann, Set-based multi-objective optimization, indicators, and deteriorative cycles, in: Proc. 12th Annual Conference on Genetic and Evolutionary Computation (GECCO'10), ACM Press, 2010, pp. 495–502.

[5] N. Beume, B. Naujoks, M.T.M. Emmerich, SMS-EMOA: Multiobjective selection based on dominated hypervolume, European Journal of Operational Research 181 (2007) 1653–1669.

[6] K. Bringmann, T. Friedrich, An efficient algorithm for computing hypervolume contributions, Evolutionary Computation 18 (2010) 383–402.

[7] K. Bringmann, T. Friedrich, Tight bounds for the approximation ratio of the hypervolume indicator, in: Proc. 11th International Conference on Parallel Problem Solving from Nature (PPSN XI), in: LNCS, vol. 6238, Springer, 2010, pp. 607–616.

[8] K. Bringmann, T. Friedrich, The maximum hypervolume set yields near-optimal approximation, in: Proc. 12th Annual Conference on Genetic and Evolutionary Computation (GECCO'10), ACM Press, 2010, pp. 511–518.

[9] K. Bringmann, T. Friedrich, Approximating the least hypervolume contributor: NP-hard in general, but fast in practice, Theoretical Computer Science 425 (2012) 104–116.

[10] K. Bringmann, T. Friedrich, C. Igel, T. Voß, The logarithmic hypervolume indicator, in: Proc. 11th International Workshop on Foundations of Genetic Algorithms (FOGA'11), ACM Press, 2011, pp. 81–92.

[11] T.C.E. Cheng, A. Janiak, M.Y. Kovalyov, Bicriterion single machine scheduling with resource dependent processing times, SIAM Journal on Optimization 8 (1998) 617–630.

[12] C. Daskalakis, I. Diakonikolas, M. Yannakakis, How good is the Chord algorithm? in: Proc. 21st Annual ACM–SIAM Symposium on Discrete Algorithms (SODA'10), 2010, pp. 978–991.

[13] K. Deb, M. Mohan, S. Mishra, Evaluating the ε-domination based multi-objective evolutionary algorithm for a quick computation of Pareto-optimal solutions, Evolutionary Computation 13 (2005) 501–525.

[14] I. Diakonikolas, M. Yannakakis, Small approximate Pareto sets for biobjective shortest paths and other problems, SIAM Journal on Computing 39 (2009) 1340–1371.

[15] M. Ehrgott, Multicriteria Optimization, 2nd edition, Springer, 2005.

[16] M.T.M. Emmerich, N. Beume, B. Naujoks, An EMO algorithm using the hypervolume measure as selection criterion, in: Proc. Third International Conference on Evolutionary Multi-Criterion Optimization (EMO'05), Springer, 2005, pp. 62–76.

[17] J. Figueira, S. Greco, M. Ehrgott, Multiple Criteria Decision Analysis: State of the Art Surveys, Springer Verlag, 2005.

[18] T. Friedrich, C. Horoba, F. Neumann, Multiplicative approximations and the hypervolume indicator, in: Proc. 11th Annual Conference on Genetic and Evolutionary Computation (GECCO'09), ACM Press, 2009, pp. 571–578.

[19] C. Igel, N. Hansen, S. Roth, Covariance matrix adaptation for multi-objective optimization, Evolutionary Computation 15 (2007) 1–28.

[20] J.D. Knowles, D. Corne, Properties of an adaptive archiving algorithm for storing nondominated vectors, IEEE Transactions on Evolutionary Computation 7 (2003) 100–116.

[21] J.D. Knowles, D.W. Corne, M. Fleischer, Bounded archiving using the Lebesgue measure, in: Proc. IEEE Congress on Evolutionary Computation (CEC'03), vol. 4, IEEE Press, 2003, pp. 2490–2497.

[22] R. Kumar, N. Banerjee, Running time analysis of a multiobjective evolutionary algorithm on simple and hard problems, in: Proc. 8th International Workshop on Foundations of Genetic Algorithms (FOGA'05), in: LNCS, vol. 3469, Springer, 2005, pp. 112–131.

[23] M. Laumanns, L. Thiele, K. Deb, E. Zitzler, Combining convergence and diversity in evolutionary multiobjective optimization, Evolutionary Computation 10 (3) (2002) 263–282.

[24] G. Lizarraga-Lizarraga, A. Hernandez-Aguirre, S. Botello-Rionda, G-metric: An m-ary quality indicator for the evaluation of non-dominated sets, in: Proc. 10th Annual Conference on Genetic and Evolutionary Computation (GECCO'08), ACM Press, 2008, pp. 665–672.

[25] K. Miettinen, Nonlinear Multiobjective Optimization, Springer, 1999.

[26] C.H. Papadimitriou, M. Yannakakis, On the approximability of trade-offs and optimal access of web sources, in: Proc. 41st Annual Symposium on Foundations of Computer Science (FOCS'00), IEEE Press, 2000, pp. 86–92.

[27] C.H. Papadimitriou, M. Yannakakis, Multiobjective query optimization, in: Proc. 20th ACM Symposium on Principles of Database Systems (PODS'01), 2001, pp. 52–59.

[28] W. Rudin, Real and Complex Analysis, 3rd edition, McGraw–Hill Book Co., New York, 1987.

[29] T. Suttorp, N. Hansen, C. Igel, Efficient covariance matrix update for variable metric evolution strategies, Machine Learning 75 (2009) 167–197.

[30] S. Vassilvitskii, M. Yannakakis, Efficiently computing succinct trade-off curves, Theoretical Computer Science 348 (2005) 334–356.

[31] E. Zitzler, S. Künzli, Indicator-based selection in multiobjective search, in: Proc. 8th International Conference on Parallel Problem Solving from Nature (PPSN VIII), in: LNCS, vol. 3242, Springer, 2004, pp. 832–842.

[32] E. Zitzler, L. Thiele, Multiobjective evolutionary algorithms: A comparative case study and the strength Pareto approach, IEEE Transactions on Evolutionary Computation 3 (1999) 257–271.
[33] E. Zitzler, L. Thiele, M. Laumanns, C.M. Fonseca, V. Grunert da Fonseca, Performance assessment of multiobjective optimizers: An analysis and review, IEEE Transactions on Evolutionary Computation 7 (2003) 117–132.
[34] E. Zitzler, D. Brockhoff, L. Thiele, The hypervolume indicator revisited: On the design of Pareto-compliant indicators via weighted integration, in: Proc. Fourth International Conference on Evolutionary Multi-Criterion Optimization (EMO'07), in: LNCS, vol. 4403, Springer, 2007, pp. 862–876.
[35] E. Zitzler, L. Thiele, J. Bader, On set-based multiobjective optimization, IEEE Transactions on Evolutionary Computation 14 (2010) 58–79.