Tight Bounds for the Approximation Ratio of the Hypervolume Indicator

Karl Bringmann[1] and Tobias Friedrich[2]

[1] Universität des Saarlandes, Saarbrücken, Germany
[2] Max-Planck-Institut für Informatik, Saarbrücken, Germany

Abstract The hypervolume indicator is widely used to guide the search and to evaluate the performance of evolutionary multi-objective optimization algorithms. It measures the volume of the dominated portion of the objective space which is considered to give a good approximation of the Pareto front. There is surprisingly little theoretically known about the quality of this approximation. We examine the multiplicative approximation ratio achieved by two-dimensional sets maximizing the hypervolume indicator and prove that it deviates significantly from the optimal approximation ratio. This provable gap 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 and prove that it achieves the optimal additive approximation ratio apart from a small factor $\leqslant n/(n-2)$, where n is the size of the population. Hence the hypervolume indicator can be used to achieve a very good additive but not a good multiplicative approximation of a Pareto front.

1 Introduction

Most real-world optimization problems have to deal with multiple objectives (like time vs. cost) and 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 forms a Pareto front. Many different multi-objective evolutionary algorithms (MOEAs) have been developed to find a Pareto set of (preferably small) size n which gives a *good approximation* of the Pareto front. A popular way to measure the quality the approximation is the hypervolume indicator. It measures the volume of the dominated space [18]. For a small number of objective, 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 [8, 16], the SMS-EMOA [3, 6], and variants of IBEA [17, 19].

One of the reasons why the hypervolume indicator is so popular is that it matches very well with our *intuition* how a good approximation of a Pareto front should look like. However, there is only little known whether maximizing the hypervolume also gives a good approximation of the Pareto front in *strictly mathematical sense*. Considering the wide use of the hypervolume indicator, the question whether it achieves a good approximation appears to be fundamental. The distribution of the points maximizing the hypervolume indicator has been examined by several authors. It was observed that “convex regions may be preferred to concave regions” [13, 18] as well as that HYP is “biased towards the boundary solutions” [5]. In contrast to this, such sets are empirically “well distributed” according to [6, 9, 10] and it was also proven that for the number of points $n \to \infty$ the density of points only depends on the gradient [2].

However, the question whether sets maximizing the hypervolume give an approximation of the Pareto front in the mathematical sense remained open besides two preliminary papers [4, 7]. We follow up on this and study the approximation quality of the hypervolume indicator by classic approximation theory. Which concept from approximation theory is the right measure depends on the problem at hand. As a general rule of thumb, for linear axes this is the additive approximation ratio while for exponential axes this is the multiplicative approximation ratio.

To illustrate this with a small example, consider a knapsack problem (see e.g. [18]) with linearly distributed weights and exponentially distributed profits. In this case a good approximation of the front should be an additive approximation of the weights and a multiplicative approximation of the profits. Within this example the result of this paper is that compared to the optimal set with best possible approximation, sets maximizing the hypervolume only achieve the first aim, not necessarily the latter.

In our previous paper [4], we proved that for all possible Pareto fronts the multiplicative approximation factor achieved by a set of n solutions maximizing the hypervolume indicator is $1+\Theta(1/n)$ (cf. Theorem 3.6)[3]. As this was shown to be *asymptotically equivalent* to the optimal multiplicative approximation factor (cf. Corollary 3.4), we concluded that the hypervolume indicator is guiding the search in the correct direction for sufficiently large n. However, the size n of a population is usually not large. Also, the constant factors hidden by the Θ might still be larger for the set maximizing hypervolume compared to the set with best possible approximation factor.

Our results

We significantly extend the results of [4]. First, we are now able to give tight bounds on the multiplicative approximation ratio depending on the ratio A/a between the largest and smallest coordinate[4]. Using this notation, the precise result of [4] is the computation of the optimal multiplicative approximation ratio as $1 + \log(A/a)/n$ (cf. Corollary 3.4). We are now able to show that the multiplicative approximation ratio for a set maximizing the hypervolume is strictly

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

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

larger, namely of the order of *at least* $1+\sqrt{A/a}/n$ (cf. Theorem 3.7). This implies that the dependence of this multiplicative approximation ratio on the ratio A/a can be exponentially worse than in the optimal case. Hence for numerically very wide spread fronts (that is, large A/a) there are Pareto sets which give a much better multiplicative approximation than the Pareto sets which maximize the hypervolume.

Second, we now also analyze the additive approximation ratio of the hypervolume indicator. While the multiplicative approximation factor is determined by the ratio A/a, the additive approximation factor is determined by the width of the domain $A-a$. We prove that the optimal additive approximation ratio is $(A-a)/n$ (cf. Theorem 4.3) and upper bound the additive approximation ratio achieved by a set maximizing the hypervolume by $(A-a)/(n-2)$ (cf. Theorem 4.5). This is a very strong statement, as apart from the small factor $n/(n-2)$ 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.

2 Preliminaries

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

$$
\begin{aligned}
&(x_1, y_1) \preceq (x_2, y_2) \text{ iff } x_1 \leqslant x_2 \text{ and } y_1 \leqslant y_2,\\
&(x_1, y_1) \prec (x_2, y_2) \text{ iff } (x_1, y_1) \preceq (x_2, y_2) \text{ and } (x_1, y_1) \neq (x_2, y_2).
\end{aligned}
$$

We restrict ourselves to Pareto fronts that can be written as $\{(x, f(x)) \mid x \in [a, A]\}$ where $f\colon [a, A] \to [b, B]$ is a (not necessarily strictly) monotonically decreasing, upper semi-continuous[5] function with $f(a) = B$, $f(A) = b$ for $a < A$, $b < B$. We write $\mathcal{F} = \mathcal{F}_{[a,A]\to[b,B]}$ 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.

The condition of f being upper semi-continuous cannot be relaxed further as without it the f lacks symmetry in the two objectives in the following sense: Being upper semi-continuous is necessary and sufficient for the existence of the inverse function $f^{-1}\colon [b, B] \to [a, A]$ defined by setting $f^{-1}(y) := \max\{x \in [a, A] \mid f(x) \geqslant y\}$. Without upper semi-continuity this maximum does not exist in general. Furthermore, this condition implies that there is a set maximizing the hypervolume indicator.

[5] 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 [15].

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, 7]), and are thus not able to give results about discrete fronts, 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 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 $\mathcal{P}^f \subseteq \mathcal{P}$ for the set of all solution sets that are feasible for f.

We are now ready to formally define the hypervolume indicator. It was first introduced for performance assessment in multiobjective optimization by Zitzler and Thiele [18], but since then also has become a very popular way to guide the search in multi-objective evolutionary optimizers. On a two-dimensional objective space it is defined as follows.

Definition 2.1 *The hypervolume indicator* HYP(P) *of a solution set* $P \in \mathcal{P}$ *relative to a reference point* $R = (R_x, R_y)$ *is*

$$\mathrm{HYP}(P) := \mathrm{VOL}\left(\bigcup_{(x,y) \in P} [R_x, x] \times [R_y, y] \right).$$

with $\mathrm{VOL}(\,\cdot\,)$ *being the usual Lebesgue measure.*

3 Multiplicative Approximation

The standard measure of approximation quality in approximation theory is the multiplicative approximation ratio. We use the multi-objective definition for the multiplicative approximation ratio by Papadimitriou and Yannakakis [14] which was also used in [4, 7, 11, 12]. Note that here and in the rest of the paper when talking about multiplicative approximation we require $a, b > 0$ as this ratio only makes sense for positive values.

Definition 3.1 *Let* $f \in \mathcal{F}$ *and* $P \in \mathcal{P}^f$. *The solution set* P *is a* multiplicative α-approximation *of* f *if for each* $\hat{x} \in [a, A]$ *there is a point* $p = (x, y) \in P$ *with*

$$\hat{x} \leqslant \alpha\, x \quad \textit{and} \quad f(\hat{x}) \leqslant \alpha\, y$$

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

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

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 3.2 *For fixed* $[a, A]$, $[b, B]$, *and* n, *let*

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

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 [4] the authors showed the following two results.

Theorem 3.3 (from [4]) $\alpha^*_{OPT} = \min\{A/a, B/b\}^{1/n}$.

Corollary 3.4 (from [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 HYP. As there might be several solution sets maximizing HYP, we consider the worst case and use the following definition.

Definition 3.5 *For fixed* $[a, A]$, $[b, B]$, *and* n, *let*

$$\mathcal{P}^f_{HYP} := \left\{P \in \mathcal{P}^f \;\middle|\; \mathrm{HYP}(P) = \max_{Q \in \mathcal{P}^f} \mathrm{HYP}(Q)\right\} \text{ for } f \in \mathcal{F}, \text{ and}$$
$$\alpha^*_{HYP} := \sup_{f \in \mathcal{F}} \sup_{P \in \mathcal{P}^f_{HYP}} \alpha^*(f, P).$$

The set $\mathcal{P}^f_{HYP}$ is the set of all feasible solution sets that maximize HYP on f. The value α^*_{HYP} is chosen such that for every front f in $\mathcal{F}$ every solution set maximizing HYP approximates f by a ratio of at most α^*_{HYP}. Note that this assumes that there is at least one solution set which maximizes the indicator, i.e., the set $\mathcal{P}^f_{HYP}$ is non-empty. That this is indeed the case was proven in [4].

In [4] the authors also examined α^*_{HYP} and showed an upper bound that has the same asymptotic behavior as α^*_{OPT}, but a much larger constant factor.

Theorem 3.6 (from [4]) *Let* $f \in \mathcal{F}$, $n > 4$, *and let* $R = (R_x, R_y) \preceq (0, 0)$ *be the reference point. If we have*

- $n \geqslant 2 + \max\left\{\sqrt{A/a}, \sqrt{B/b}\right\}$ or
- $R_x \leqslant -\sqrt{A\,a}/n$, $R_y \leqslant -\sqrt{B\,b}/n$,

then

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

The previous paper [4] left open whether (i) the upper bound of Theorem 3.6 is not tight and α^*_{HYP} is actually much closer to the bounds for α^*_{OPT} given in Corollary 3.4 or (ii) α^*_{HYP} is indeed significantly larger than α^*_{OPT}. By giving a lower bound for α^*_{HYP} we can now prove the latter. In the following theorem we restrict ourselves to the case of $A/a = B/b$. We show that in this situation the bound of Theorem 3.6 is tight except for a small constant factor.

Theorem 3.7 *Let $n \geqslant 4$, $A/a = B/b \geqslant 13$, and $R = (R_x, R_y) \preceq (0,0)$ be the reference point. Then*

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

The proof of this theorem will be provided in the full version of the paper.

4 Additive Approximation

After the previous section showed that sets maximizing the hypervolume have sub-optimal multiplicative approximation ratio we now analyze their additive approximation properties. Analogous to Definition 3.1 we use the following definition.

Definition 4.1 *Let $f \in \mathcal{F}$ and $P \in \mathcal{P}^f$. The solution set P is an* additive α-approximation *of f if for each $\hat{x} \in [a, A]$ there is a point $p = (x, y) \in P$ with*

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

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

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

Again, we are interested in the optimal approximation ratio for Pareto fronts in $\mathcal{F}$. Analogous to Definition 3.2 we give the following definition.

Definition 4.2 *For fixed $[a, A]$, $[b, B]$, and n, let*

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

Analogously to the precise bound $\alpha^*_{OPT} = \min\{A/a, B/b\}^{1/n}$ of Theorem 3.3 for the optimal multiplicative approximation ratio, we can prove the following for the optimal additive approximation ratio α^+_{OPT}.

Theorem 4.3 $\alpha^+_{OPT} = \dfrac{\min\{A - a, B - b\}}{n}$.

The proof of Theorem 4.3 will be provided in the full version of the paper.

In order to compare the optimal additive approximation ratio with the approximation ratio achieved by the hypervolume, we give the following definition analogously to the definition of α^*_{HYP} in Definition 3.5.

Definition 4.4 *For fixed $[a, A], [b, B], n$, and $f \in \mathcal{F}$ let*

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

We can now state the main result of this paper that α^+_{HYP} is very close to α^+_{OPT}. Similar to the proof of the upper bound for α^*_{HYP} of Theorem 3.6 we can prove the following upper bound for α^+_{HYP}.

Theorem 4.5 *For all $n > 2$ and $(n-2)\min\{a - R_x, b - R_y\} \geqslant \sqrt{(A-a)\,(B-b)}$,*

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

Let us briefly discuss the result before the theorem will be proven in the remainder of this section. First note that the precondition 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 4.3. This shows that for $A-a \approx B-b$ and moderately sized n, α^+_{HYP} is very close to α^+_{OPT}. More precisely, for $A - a \ll B - b$ (or $A - a \gg B - b$) the constant in Theorem 4.5 is the geometric mean of $A - a$ and $B - b$ while in Theorem 4.3 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.

Proof of Theorem 4.5. Let P be a solution set maximizing HYP on a front $f \in \mathcal{F}$, i.e., $P \in \mathcal{P}^f_{HYP}$. 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[6] $P' := P + r - p$ would have $\mathrm{HYP}(P') > \mathrm{HYP}(P)$, as it dominates all the space P dominates united with the space r dominates. Thus, there is no such point r and P dominates already 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.

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 $P = \{p_1, \ldots, p_n\}$, $p_i = (x_i, y_i)$ with $a \leqslant x_1 < \ldots < x_n \leqslant A$ and $B \geqslant y_1 > \ldots > y_n \geqslant b$. 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 (this uses that the points in P are mutually non-dominating).

We want to argue about the contribution of a point p to the hypervolume of a solution set P, namely $\mathrm{CON}_P(p) := \mathrm{HYP}(P) - \mathrm{HYP}(P - p)$. In particular we need the minimal contribution of any of the points $p_2, \ldots, p_{n-1}$:

$$\begin{aligned}\textsc{MinCon}(P) &:= \min_{1<i<n} \mathrm{CON}_P(p_i)\\ &= \min_{1<i<n} (x_i - x_{i-1})\,(f(x_i) - f(x_{i+1})).\end{aligned}$$

[6] To increase the readability, for a set $P \subset \mathbb{R}^2$ and a point $r \in \mathbb{R}^2$ we define $P + r := P \cup \{r\}$ and $P - r := P \setminus \{r\}$ here and in the remainder of this section.

This value has been (with slightly different notation) examined in [4]. In particular, the authors showed that for $n > 2$

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

This implies

$$\text{MinCon}(P) \leqslant \frac{(A-a)\,(B-b)}{(n-2)^2}. \tag{1}$$

Let $r = (x, f(x))$, $x \in [a, A]$ be an arbitrary point and let $\alpha > 0$ be such that r it 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 $i \in \{1, \ldots, n-1\}$ with $x_i \leqslant x < x_{i+1}$. As r is not additively approximated by α, we have

$$x > x_i + \alpha \quad \text{and} \quad f(x) > f(x_{i+1}) + \alpha. \tag{2}$$

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

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

which in turn implies

$$\text{MinCon}(P) = \text{CON}_P(p) \geqslant \text{CON}_{P+r}(r) = (x - x_i)\,(f(x) - f(x_{i+1})) \overset{(2)}{>} \alpha^2.$$

Using equation (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 $\text{MinCon}(P)$ to P by r must not increase the hypervolume, so we have

$$\begin{aligned}\text{MinCon}(P) \geqslant \text{CON}_{P+r-p}(r) \geqslant \text{CON}_{P+r}(r) &= (a - R_x)\,(f(x) - f(x_1))\\ &\geqslant (a - R_x)\alpha.\end{aligned}$$

We use equation (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 precondition of the theorem. $\square$

	Multiplicative approximation		Additive approximation	
OPT	$1+\dfrac{\log(\min\{A/a,B/b\})}{n}$	(Cor. 3.4)	$\dfrac{\min\{A-a,B-b\}}{n}$	(Thm. 4.3)
HYP	$1+\dfrac{\sqrt{A/a}+\sqrt{B/b}}{n-4}$	(Thm. 3.6)	$\dfrac{\sqrt{(A-a)\,(B-b)}}{n-2}$	(Thm. 4.5)

Table 1: Results for the optimal approximation ratio and upper bounds for the approximation ratios of HYP. See the cited theorems for the precise statements.

5 Conclusion

Many modern MOEA use the hypervolume indicator to guide the search process. We presented a mathematically rigorous framework to analyze the approximation ratio achieved by sets maximizing the hypervolume. We prove that sets maximizing HYP do *not* give a perfect multiplicative approximation. The proven bounds can be found in Table 1. The multiplicative approximation ratio of HYP is getting large for numerically wide spread fronts with large A/a. On the other hand, we can prove that maximizing HYP gives a close-to-optimal additive approximation.

Bibliography

[1] A. Auger, J. Bader, D. Brockhoff, and 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)*, pp. 563–570, 2009.

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

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

[4] K. Bringmann and T. Friedrich. The maximum hypervolume set yields near-optimal approximation. In *Proc. 12th annual Conference on Genetic and Evolutionary Computation (GECCO '10)*, 2010.

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

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

[7] T. Friedrich, C. Horoba, and F. Neumann. Multiplicative approximations and the hypervolume indicator. In *Proc. 11th annual Conference on Genetic and Evolutionary Computation (GECCO '09)*, pp. 571–578, 2009.

[8] C. Igel, N. Hansen, and S. Roth. Covariance matrix adaptation for multi-objective optimization. *Evolutionary Computation*, 15:1–28, 2007.

[9] J. D. Knowles and D. Corne. Properties of an adaptive archiving algorithm for storing nondominated vectors. *IEEE Trans. Evolutionary Computation*, 7:100–116, 2003.

[10] J. D. Knowles, D. W. Corne, and M. Fleischer. Bounded archiving using the Lebesgue measure. In *Proc. IEEE Congress on Evolutionary Computation (CEC '03)*, Vol. 4, pp. 2490–2497, 2003.

[11] R. Kumar and 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)*, Vol. 3469 of *Lecture Notes in Computer Science*, pp. 112–131. Springer, 2005.

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

[13] G. Lizarraga-Lizarraga, A. Hernandez-Aguirre, and 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)*, pp. 665–672, 2008.

[14] C. H. Papadimitriou and 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)*, pp. 86–92, 2000.

[15] W. Rudin. *Real and complex analysis.* McGraw-Hill Book Co., New York, 3rd edition, 1987.

[16] T. Suttorp, N. Hansen, and C. Igel. Efficient covariance matrix update for variable metric evolution strategies. *Machine Learning*, 75:167–197, 2009.

[17] E. Zitzler and S. Künzli. Indicator-based selection in multiobjective search. In *Proc. 8th International Conference Parallel Problem Solving from Nature (PPSN VIII)*, Vol. 3242 of *LNCS*, pp. 832–842, 2004.

[18] E. Zitzler and L. Thiele. Multiobjective evolutionary algorithms: a comparative case study and the strength Pareto approach. *IEEE Trans. Evolutionary Computation*, 3:257–271, 1999.

[19] E. Zitzler, D. Brockhoff, and 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)*, Vol. 4403 of *LNCS*, pp. 862–876, 2007.