technische universität
dortmund

Convergence Rates of (1+1) Evolutionary Multiobjective Optimization Algorithms

Nicola Beume
Marco Laumanns
Günter Rudolph

Algorithm Engineering Report
TR10-2-006
July 2010
ISSN 1864-4503

Faculty of Computer Science
Algorithm Engineering (Ls11)
44221 Dortmund / Germany
http://ls11-www.cs.uni-dortmund.de/

Convergence Rates of (1+1) Evolutionary Multiobjective Optimization Algorithms

Nicola Beume[1], Marco Laumanns[2], and Günter Rudolph[1]

[1] Fakultät für Informatik, Technische Universität Dortmund, Germany
{nicola.beume, guenter.rudolph}@tu-dortmund.de
[2] IBM Research – Zurich, Switzerland, mlm@zurich.ibm.com

Abstract. Convergence analyses of evolutionary multiobjective optimization algorithms typically deal with the convergence in limit (stochastic convergence) or the run time. Here, for the first time concrete results for convergence rates of several popular algorithms on certain classes of continuous functions are presented. We consider the algorithms in the version of using a (1+1) selection scheme. Then, SMS-EMOA and $\text{IBEA}_{\epsilon+}$ achieve linear convergence rate, proved by showing algorithmic equivalence to the single-objective (1+1)-EA with self-adaptation, whereas NSGA-II and SPEA2 have a sub-linear convergence rate, proved by reducing them to a multiobjective algorithm with known properties.

Keywords: multiobjective optimization, convergence rate, hypervolume, self-adaptation.

1 Introduction

Research on evolutionary algorithms is developed further for single-objective optimization than for multi-objective optimizers. A common hope is that the understanding of evolutionary multiobjective optimization algorithms (EMOA) can profit from the bases acquired for the single-objective case. Here, we transfer knowledge on the convergence of the single-objective (1+1)-EA to gain insights into the convergence behavior of complex EMOA.

Convergence properties of EMOA are yet not well understood. More recently, theory concentrated on the convergence or runtime of simple EMOA on special discrete problems, considering whether and how quickly the Pareto set is reached. For the case of a continuous search space $\mathbb{R}^n$ only a few results exist for specialized algorithms, the first obtained by Rudolph [1]. He showed that a multiobjective (1+1)-EA that accepts incomparable points with probability $\frac{1}{2}$ converges with probability 1 to the Pareto set if the step size is chosen proportional to the distance to the Pareto set, while two other step size concepts fail. Hanne [2] considered stochastic convergence of EMOA with different selection schemes, the possibilities of temporary fitness deterioration, and on problems with unattainable solutions. A recent subject of interest has been whether a certain distribution on the Pareto front can be obtained that is optimal regarding specified preferences.

Despite these advances, the convergence rate in continuous space remains a neglected topic. Teytaud [3] shows that the convergence rate scales badly with increasing number of objectives entailing that any comparison-based EMOA performs hardly better than random search for a large number of objectives. Also a general lower bound for the convergence time is given.

In this paper we consider popular EMOA in the simple version of using a (1+1) selection scheme. For (strongly) convex quadratic objective functions, the order of the convergence rate is calculated, whereas SMS-EMOA and $\text{IBEA}_{\epsilon+}$ reach a linear convergence rate. This is to the best of our knowledge the first time that a linear convergence rate is shown for a multiobjective evolutionary algorithms that does not use an explicit weighting of objectives.

The next section introduces the technical background of our topic. Section 3 shows the linear convergence rate for SMS-EMOA and IBEA, whereas Section 4 gives the negative results for NSGA-II and SPEA2. We summarize our findings in section 5 and give hints on future research.

2 Preliminaries

2.1 Single-objective Optimization with the (1+1)-EA

Let $f : \mathbb{R}^n \to \mathbb{R}$ be the objective function to be minimized. The $(1+1)$ Evolutionary Algorithm (EA) (cf. Alg. 1) minimizes $f(\cdot)$ by drawing an n-dimensional random vector from a multivariate standard normal distribution that is scaled by factor σ and then added to the current position. If the new point is better it is accepted, otherwise it is rejected. Then the scaling factor is adapted and this sequence is run again.

Algorithm 1: $(1+1)$-Evolutionary Algorithm with Self-Adaptation

choose $X^{(0)} \in \mathbb{R}^n$ and $\sigma^{(0)} > 0$, set $t = 0$ and $k = 0$
repeat
  draw $Z^{(t)}$ from a multivariate standard normal distribution
  $Y^{(t)} = X^{(t)} + \sigma^{(t)} Z^{(t)}$
  if $f(Y^{(t)}) \leq f(X^{(t)})$ **then**
    $X^{(t+1)} = Y^{(t)}$; increment k
  else $X^{(t+1)} = X^{(t)}$
  $\sigma^{(t+1)} = \texttt{adapt}(\sigma^{(t)}, t, k\,; \delta, p_s, \gamma)$
  increment t
until *termination criterion fulfilled*

Procedure adapt$(\sigma, t, k\,; \delta, p_s, \gamma)$

if $t \mod \delta \neq 0$ **then return** σ
$q_s = k/\delta$; $k = 0$
if $q_s \geq p_s$ **then return** $\sigma \times \gamma$ **else return** σ/γ

The self-adaptation mechanism considered here (Procedure `adapt`) is termed the $\frac{1}{5}$-success rule that has some parameters: the observation interval $\delta > 0$, the success probability $p_s = \frac{1}{5}$ and the adaptation factor $\gamma > 1$. If the self-adaptation procedure is properly parameterized a remarkable result has been proven by Jägersküpper [4, 5]:

Theorem 1. *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a quadratic function $f(x) = x'Ax + b'x + c$ with positive definite matrix A whose condition number is bounded. The $(1+1)$-EA with self-adaptation as in Procedure* `adapt` *using $\delta = \Theta(n)$, $p_s = \frac{1}{5}$ and $c \geq 2$ halves the distance to the optimum in $O(n)$ iterations in expectation, provided that $\sigma^{(0)} = \Theta(D/n)$ where D is the distance to the optimum after initialization.*

In other words: under the conditions of the theorem the (1+1)-EA with self-adaptation minimizes every strongly convex quadratic function with linear convergence rate, i.e., the approximation error decreases exponentially fast.

2.2 Multi-objective Optimization with the SMS-EMOA

We consider unconstrained multiobjective optimization problems $\min f(x) : \mathbb{R}^n \to \mathbb{R}^d$ where $f(x) = (f_1(x), \ldots, f_d(x))$ maps an n-dimensional vector of the search space to a d-dimensional vector of the objective space.

A strict partial order, called Pareto dominance, holds in the objective space based on the coordinate-wise total order: a point $p = (p_1, \ldots, p_d)$ weakly dominates a point q (written as $p \preceq q$) iff $p_i \leq q_i$ holds for all $1 \leq i \leq d$. A point p dominates q iff $p \preceq q$ and $p \neq q$. Two distinct points $p \neq q$ are incomparable ($p \parallel q$) iff neither point dominates the other. Considering a set $A \subseteq \mathbb{R}^d$, points of A that are not dominated by any other of A are referred to as non-dominated in A or the minima of A. Those points that are non-dominated regarding the whole objective space are Pareto-optimal and called the Pareto front. The set of their preimages in the search space is named the Pareto set.

In the continuous domain only an approximation of the Pareto set can be expected to be achieved. In order to compare the results of different EMOA, several quality measures exist, typically rewarding quantity, closeness to the Pareto set, and high diversity. Among these, the hypervolume indicator (or S-metric or Lebesque measure) by Zitzler and Thiele [6] is of outstanding importance due to its consistency with the Pareto dominance relation, cf. [7].

Definition 1. *Let $\{v^{(1)}, v^{(2)}, \ldots, v^{(\mu)}\} \subset \mathbb{R}^d$, $d \geq 2$ be a finite set of elements, which are mutually incomparable w. r. t. to the dominance relation $\preceq$. Let $r \in \mathbb{R}^d$ indicate the* reference point *with $v^{(i)} \prec r$ for all $i = 1, \ldots, \mu \in \mathbb{N}$. The quantity*

$$H(v^{(1)}, \ldots, v^{(\mu)}; r) = \mathsf{Leb}\left(\bigcup_{i=1}^{\mu} [v^{(i)}, r]\right) \tag{1}$$

is termed the dominated hypervolume *or* S-metric *where $\mathsf{Leb}(\cdot)$ denotes the Lebesgue measure in $\mathbb{R}^d$.*

If $d = 2$, provided the elements $v^{(1)}, \ldots, v^{(\mu)}$ have been labeled in ascending order of their first component, i.e., $v_1^{(1)} < v_1^{(2)} < \ldots < v_1^{(\mu)}$, equation (1) specializes to

$$H(v^{(1)}, \ldots, v^{(\mu)}; r) = (r_1 - v_1^{(1)})\,(r_2 - v_2^{(1)}) + \sum_{i=2}^{\mu} (r_1 - v_1^{(i)})\,(v_2^{(i-1)} - v_2^{(i)}). \quad (2)$$

The SMS-EMOA [8] is a steady-state, i.e. $(\mu+1)$, EMOA that aims to maximize the population's dominated hypervolume by incorporating it in the selection operator. The selection starts with non-dominated sorting in order to determine the worst front. Among these points, the one contributing least to the dominated hypervolume of the set is discarded. The hypervolume contribution of a point is defined as the dominated hypervolume that is exclusively dominated by the point and thus would get lost when the point was discarded. The calculation of the hypervolume requires the specification of a reference point r. Yet, it is no exogenous parameter of the SMS-EMOA but chosen automatically. For each objective function, the maximal value among the $\mu+1$ points is determined. The reference point is constructed by these maxima plus 1. The decisive properties that will be utilized to prove the linear convergence rate are: (1) The reference point is not static throughout the optimization process but dynamically adapted in each generation. (2) Those points that are the worst ones in the population regarding an objective function have a distance to the reference point of exactly 1 w.r.t. that worst objective (cf. Fig. 1, left).

The SMS-EMOA does not specify a certain variation operator. It has mainly been considered using SBX recombination and polynomial mutation (see e.g. [9]). Here, we consider Gaussian mutation with self-adaptation as detailed in the following section.

Section 3.3 contains the analysis of IBEA$_{\epsilon+}$ [10] that performs non-dominated sorting and afterwards selects among the worst points using the additive ϵ-indicator $I_{\epsilon+}$. Section 4 deals with NSGA-II [9] that firstly uses non-dominated sorting, and afterwards a particular density measure, the crowding distance, as the secondary selection criterion. SPEA2 [11] as well applies a selection criterion based on the Pareto dominance relation by counting dominated and dominating solutions for each point. Among the incomparable ones, again a kind of density measure comes into play, namely a k-nearest neighbor method.

3 Linear Convergence Rates

3.1 (1+1)-SMS-EMOA on 2-objective problems

If the $(\mu+1)$-SMS-EMOA is instantiated with $\mu = 1$ it reduces to the algorithm described below (see Alg. 2).

Theorem 2. *The* $(1+1)$*-SMS-EMOA with self-adaptation applied to applied to a bi-objective optimization problem* $\min\{f : \mathbb{R}^n \to \mathbb{R}^2\}$ *is algorithmically equivalent to a* $(1+1)$*-EA with self-adaptation applied to the minimization of the single-objective function* $f^s : \mathbb{R}^n \to \mathbb{R}$ *with* $f^s(x) = \frac{1}{2}\,(f_1(x) + f_2(x))$.

Algorithm 2: (1+1)-SMS-EMOA with Self-Adaptation

choose $X^{(0)} \in \mathbb{R}^n$ and $\sigma^{(0)} > 0$, set $t = 0$ and $k = 0$
repeat
  draw $Z^{(t)}$ from a multivariate standard normal distribution
  $Y^{(t)} = X^{(t)} + \sigma^{(t)} Z^{(t)}$
  $R^{(t)} = (\max\{f_1(X^{(t)}), f_1(Y^{(t)})\} + 1, \max\{f_2(X^{(t)}), f_2(Y^{(t)})\} + 1)'$
  if $f(Y^{(t)}) \prec f(X^{(t)})$ ***or***
    $\big(f(Y^{(t)}) \parallel f(X^{(t)})$ ***and*** $H(f(Y^{(t)}); R^{(t)}) > H(f(X^{(t)}); R^{(t)})\big)$ **then**
    $X^{(t+1)} = Y^{(t)}$; increment k
  else $X^{(t+1)} = X^{(t)}$
  $\sigma^{(t+1)} = \texttt{adapt}(\sigma^{(t)}, t, k\,; \delta, p_s, c)$
  increment t
until *termination criterion fulfilled*

Proof. The (1+1)-SMS-EMOA differs from the (1+1)-EA only in the additional determination of the reference point $R^{(t)}$ which is required in the seemingly more complex acceptance criterion. Evidently, it is sufficient to show that the $(1+1)$-SMS-EMOA accepts/rejects a new point if it would be accepted/rejected by the (1+1)-EA with the scalarized objective function (cf. Fig. 1 (right) for its regions of acceptance or rejection).

The mutated individual y is accepted if it dominates its parent x, i.e., $f(y) \prec f(x)$. This implies $f^s(y) < f^s(x)$ and the $(1+1)$-EA would accept y:

$$\begin{aligned} f(y) \prec f(x) &\Leftrightarrow f_1(y) < f_1(x) \;\wedge\; f_2(y) < f_2(x) \\ &\Rightarrow f_1(y) + f_2(y) < f_1(x) + f_2(x) \\ &\Leftrightarrow \frac{1}{2}\,f_1(y) + \frac{1}{2}\,f_2(y) < \frac{1}{2}\,f_1(y) + \frac{1}{2}\,f_2(y) \\ &\Leftrightarrow f^s(y) < f^s(x) \end{aligned}$$

Moreover, the mutated individual y is also accepted if it is incomparable to its parent x, i.e., $f(y) \parallel f(x)$, but has a larger dominated hypervolume. This condition also implies $f^s(y) < f^s(x)$ which is easily seen as follows: Since $f(y) \parallel f(x)$ we have to distinguish two cases.

1. $f_1(x) > f_1(y) \wedge f_2(x) < f_2(y)$
   According to Algorithm 2 the reference point is $r = (f_1(x) + 1, f_2(y) + 1)'$, here. Recall that $H(v; r) = (r_1 - v_1)\,(r_2 - v_2)$ for a single point. It follows:

   $$H_x = H(f(x); r) = [\, f_1(x) + 1 - f_1(x)\,]\,[\, f_2(y) + 1 - f_2(x)\,] = f_2(y) - f_2(x) + 1$$

   $$H_y = H(f(y); r) = [\, f_1(x) + 1 - f_1(y)\,]\,[\, f_2(y) + 1 - f_2(y)\,] = f_1(x) - f_1(y) + 1$$

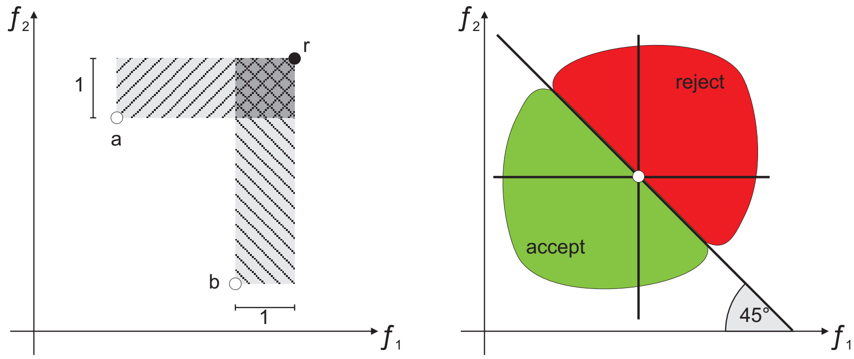

Fig. 1. Left: Dominated hypervolume of the points a and b w.r.t. the reference point r, whereas the hypervolume contribution is shaded in light gray. Right: Regions of acceptance or rejection for the substitute weighted sum function.

The new point y is accepted if

$$\begin{aligned} H_y \overset{!}{>} H_x &\Rightarrow f_1(x) - f_1(y) + 1 > f_2(y) - f_2(x) + 1 \\ &\Leftrightarrow f_1(x) - f_1(y) > f_2(y) - f_2(x) \\ &\Leftrightarrow f_1(x) + f_2(x) > f_1(y) + f_2(y) \\ &\Leftrightarrow f^s(x) > f^s(y)\,. \end{aligned}$$

Thus, in this particular situation the $(1+1)$-SMS-EMOA would accept y and so would do the $(1+1)$-EA.

2. $f_1(x) < f_1(y) \wedge f_2(x) > f_2(y)$
Now the reference point is $r = (f_1(y)+1, f_2(x)+1)'$ yielding a dominated hypervolume of
$H_x = H(f(x); r) = [\,f_1(y)+1-f_1(x)\,]\,[\,f_2(x)+1-f_2(x)\,] = f_1(y)-f_1(x)+1$
$H_y = H(f(y); r) = [\,f_1(y)+1-f_1(y)\,]\,[\,f_2(x)+1-f_2(y)\,] = f_2(x)-f_2(y)+1$
The new point y is accepted if

$$\begin{aligned} H_y \overset{!}{>} H_x &\Rightarrow f_2(x) - f_2(y) + 1 > f_1(y) - f_1(x) + 1 \\ &\Leftrightarrow f_2(x) - f_2(y) > f_1(y) - f_1(x) \\ &\Leftrightarrow f_1(x) + f_2(x) > f_1(y) + f_2(y) \\ &\Leftrightarrow f^s(x) > f^s(y) \end{aligned}$$

Again, the $(1+1)$-SMS-EMOA would accept the new point y and so would do the $(1+1)$-EA.

Putting all together we have shown that whenever the $(1+1)$-SMS-EMOA accepts a new element so does the $(1+1)$-EA. Finally we have to preclude that the $(1+1)$-EA accepts elements that are rejected by the SMS-EMOA. Or equivalently, if the $(1+1)$-SMS-EMOA rejects a new element then so must do the $(1+1)$-EA. Notice that the proof of this property is analogous to acceptance case above and therefore omitted here. □

The equivalence of $(1+1)$-SMS-EMOA and the specific $(1+1)$-EA as formulated in Th. 2 holds for all bi-objective problems. For a certain class of problems, we show a linear convergence rate:

Corollary 1. *The* $(1+1)$*-SMS-EMOA with self-adaptation applied to applied to a bi-objective optimization problem* $\min\{f : \mathbb{R}^n \to \mathbb{R}^2\}$ *approaches an element of the Pareto front with linear order of convergence if both objective functions are quadratically convex and at least one of them strongly convex.*

Proof. Since both objective functions are quadratically convex they are of form

$$f_1(x) = \tfrac{1}{2}x'Ax + b'x + c \qquad \text{and} \qquad f_2(x) = \tfrac{1}{2}x'\check{A}x + \check{b}'x + \check{c}$$

with positive semidefinite matrices A and $\check{A}$. Notice that at least one objective function is even strongly convex so that its Hessian matrix is positive definite. Suppose w. l. o. g. that A is positive definite. Since

$$f^s(x) = \frac{1}{2}(f_1(x) + f_2(x)) = \frac{1}{2}\left[\frac{1}{2}x'(A+\check{A})x + (b+\check{b})'x + (c+\check{c})\right]$$

$$\text{and} \qquad x'(A+\check{A})x \;=\; \underbrace{x'Ax}_{>0} + \underbrace{x'\check{A}x}_{\geq 0} \;>\; 0$$

for all $x \in \mathbb{R}^n \setminus \{0\}$, its Hessian matrix is positive definite ensuring that $f^s(x)$ is a strongly convex quadratic function. Now we can invoke Theorem 1 that guarantees linear convergence rate of the $(1+1)$-EA with self-adaptation for $f^s(x)$. Owing to Theorem 2 we know that the $(1+1)$-EA with self-adaptation is algorithmically equivalent to a $(1+1)$-SMS-EMOA for minimizing the bi-objective function $(f_1(x), f_2(x))'$. As a consequence, the $(1+1)$-SMS-EMOA must have linear convergence rate to an element of the Pareto front under the conditions of the corollary. □

3.2 (1+1)-SMS-EMOA Beyond Two Objectives

Expectedly, the result from the previous section does not generalize to more than two objectives, which we show by a simple counter-example. First notice that the reference point for two points $f(x)$ and $f(y)$ in objective space $\mathbb{R}^d$ is

$$r = \big(\max\{f_1(x), f_1(y)\} + 1, \max\{f_2(x), f_2(y)\} + 1, \ldots, \max\{f_d(x), f_d(y)\} + 1\big)'$$

where $f : \mathbb{R}^n \to \mathbb{R}^d$, $d \geq 2$. The dominated hypervolume $H(v; r)$ for a single point $v \in \mathbb{R}^d$ and the scalarized objective function $f^s(x)$ used by the single-objective $(1+1)$-EA are, respectively

$$H(v; r) = \prod_{i=1}^{d} [\, r_i - v_i \,] \qquad \text{and} \qquad f^s(x) = \frac{1}{d}\sum_{i=1}^{d} f_i(x)\,.$$

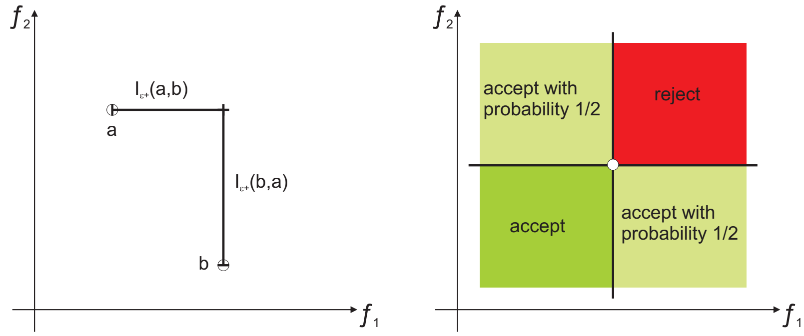

Fig. 2. Left: Values of the additive ϵ indicator $I_{\epsilon+}$ correspond to the hypervolume contributions in the (1+1)-SMS-EMOA. Right: Regions of acceptance, rejection, or random acceptance in case of incomparable points for (1+1)-NSGA-II, (1+1)-SPEA2, and the simple (1+1)-EA from [1].

Suppose there are two incomparable points $x, y \in \mathbb{R}^d$ with values $f(x) = (0, 0, \ldots, 0)'$ and $f(y) = (-1, -1, \ldots, -1, d-1+\epsilon)'$ where $\epsilon \in (0,1) \subset \mathbb{R}$. Insertion yields the reference point $r = (1, 1, \ldots, 1, d+\epsilon)$ leading to the dominated hypervolume $H_x = d + \epsilon$ and $H_y = 2^{d-1}$.

The (1+1)-SMS-EMOA would accept y if $H_y > H_x$. Notice that $H_y > H_x \Leftrightarrow 2^{d-1} > d + \epsilon$ is true for $d \geq 3$. But the (1+1)-EA would reject y since

$$f^s(y) = \frac{1}{d}\left[(d-1)\cdot(-1) + d - 1 + \epsilon\right] = \frac{\epsilon}{d} \;>\; 0 \;=\; f^s(x)\,.$$

As a consequence, both algorithms are not algorithmically equivalent for $d \geq 3$ in case of a uniformly weighted scalarized objective function. □

3.3 (1+1)-IBEA using the Additive ϵ-Indicator

We show that a (1+1)-IBEA [10] selecting according to the additive ϵ-indicator $I_{\epsilon+}$ [7] performs equal to the (1+1)-SMS-EMOA for two objectives. IBEA$_{\epsilon+}$ prefers non-dominated individuals over dominated ones, so for the case of two comparable individuals, the behavior of acceptance and rejection is clearly equal to the one of the SMS-EMOA. For incomparable individuals, the indicator $I_{\epsilon+}$ comes into play, which is a relative binary indicator, originally defined on two sets of points. For two points, $I_{\epsilon+}(a,b)$ calculates the minimal distance ϵ by which a can be moved in each direction until it is weakly dominated by b.

$$I_{\epsilon+}(a,b) = \min_{\epsilon}\{\forall i \in \{1,\ldots,d\} : f_i(a) + \epsilon \geq f_i(b)\} \qquad (3)$$

Obviously large values correspond to valuable individuals, analogously to the hypervolume (contribution). The hypervolume contribution has been shown to reduces to a distance for the (1+1)-SMS-EMOA. This distance is exactly equal to the value of the $I_{\epsilon+}$ (cf. Fig. 2, left). Thus it directly follows:

Corollary 2. *Theorem 2 and Corollary 1 hold as well for the (1+1)-IBEA$_{\epsilon+}$ which selects according to the additive ϵ-indicator $I_{\epsilon+}$.*

4 Sub-Linear Convergence Rates

We investigate whether the equivalence of (1+1)-SMS-EMOA and IBEA$_{\epsilon+}$ to the (1+1)-EA is outstanding. To this end, we consider further popular EMOA in the version of using a (1+1) selection scheme.

NSGA-II [9] has been developed with a $(\mu + \mu)$ selection, and is thus considered for $\mu = 1$. The selection starts by performing non-dominated sorting on the set of parent and offspring. If the individuals are comparable, the dominating one is kept and the dominated one discarded. In case of incomparable individuals the crowding distance is invoked. It rewards individuals with a large distance to their neighbors, and assigns a value of infinity to points at the boundary of the non-dominated front, i.e. those not having neighbors in one dimension. Here, both points are boundary points with equal crowding distance values. Thus, one is chosen to be discarded uniformly at random, so in case of incomparable points, each is accepted with probability 1/2 (cf. Fig. 2, right).

The same result holds for the (1+1)-SPEA2 [11]. For incomparable individuals, there are neither dominated nor dominating ones, thus the raw fitness of both individuals is zero. So, the secondary indicator based on a k-nearest neighbor method is used. The resulting values for the individuals are equal since they both are their only neighbors and distances are symmetrical.

We declare that both algorithms in their (1+1) version are equal to the EMOA considered by Rudolph [1]: Recall that this (1+1)-EA chooses uniformly at random one fitness function for selection. The better individual w.r.t. to the function is kept, the other one discarded. Two incomparable individuals have both worst and best values in interchanged functions. So, choosing a function is equivalent to choosing the preferred individual. Since [1] proves that convergence is given but only with a sub-linear rate for at least one instance from the problem class, we immediately get the following result.

Theorem 3. *The (1+1)-NSGA-II and the (1+1)-SPEA2 have sub-linear convergence rate under conditions for the step sizes given in [1].* □

It is still unclear how this step size rule can be realized in practice and, thus, whether NSGA-II and SPEA2 converge at all for any other known mutation operator. Nevertheless, our result indicates that sub-linear convergence might be the best one can hope for.

5 Conclusions

We showed that the (1+1) versions of SMS-EMOA and IBEA$_{\epsilon+}$ have linear convergence rate on the class of bi-objective problems whose functions all are quadratically convex with at least one being strongly convex. This is the first time that a linear convergence rate could be proved for evolutionary multiobjective optimization algorithms that do not require an explicit weighting of objectives. The convergence rate is proved by reduction to an already analyzed single-objective (1+1)-EA with self-adaptation. The equivalence of the algorithms holds for arbitrary bi-objective problems, and a result regarding the linear convergence rate on a certain class of functions is transferred to the EMOA. By a counter example it is shown that the selection behavior of the EMOA is no longer equal to the single-objective (1+1)-EA for more than two objectives.

(1+1)-NSGA-II and (1+1)-SPEA2 have a sub-linear convergence rate on the considered class of functions due to the fact that their selection operators degenerate to random choice among incomparable individuals.

Future research shall consider how population sizes greater than one influence the convergence properties of different evolutionary multiobjective optimization algorithms.

References

1. Rudolph, G.: On a multi–objective evolutionary algorithm and its convergence to the Pareto set. In: Proceedings of the 1998 IEEE International Conference on Evolutionary Computation. IEEE Press, Piscataway (NJ) (1998) 511–516
2. Hanne, T.: On the convergence of multiobjective evolutionary algorithms. European Journal of Operational Research **117**(3) (1999) 553–564
3. Teytaud, O.: On the hardness of offline multi-objective optimization. Evolutionary Computation **15**(4) (2007) 475–491
4. Jägersküpper, J.: How the (1+1) ES using isotropic mutations minimizes positive definite quadratic forms. Theoretical Computer Science **361**(1) (2006) 38–56
5. Jägersküpper, J.: Algorithmic analysis of a basic evolutionary algorithm for continuous optimization. Theoretical Computer Science **379**(3) (2007) 329–347
6. Zitzler, E., Thiele, L.: Multiobjective optimization using evolutionary algorithms – a comparative case study. In: Proc. of the 5th Intl. Conference on Parallel Problem Solving from Nature (PPSN V). LNCS 1498, Berlin, Springer (1998) 292–304
7. Zitzler, E., Thiele, L., Laumanns, M., Fonseca, C.M., Grunert da Fonseca, V.: Performance Assessment of Multiobjective Optimizers: An Analysis and Review. IEEE Transactions on Evolutionary Computation **7**(2) (2003) 117–132
8. Beume, N., Naujoks, B., Emmerich, M.: SMS-EMOA: Multiobjective selection based on dominated hypervolume. European Journal of Operational Research **181**(3) (2007) 1653–1669
9. Deb, K., Pratap, A., Agarwal, S., Meyarivan, T.: A Fast and Elitist Multiobjective Genetic Algorithm: NSGA-II. IEEE Transactions on Evolutionary Computation **6**(2) (2002) 182–197
10. Zitzler, E., Künzli, S.: Indicator-Based Selection in Multiobjective Search. In: Proc. of the 8th Intl. Conference on Parallel Problem Solving from Nature (PPSN VIII). LNCS 3242, Springer (2004) 832–842
11. Zitzler, E., Laumanns, M., Thiele, L.: SPEA2: Improving the Strength Pareto Evolutionary Algorithm for Multiobjective Optimization. In: Evolutionary Methods for Design, Optimisation and Control with Application to Industrial Problems (EUROGEN 2001), CIMNE (2002) 95–100