

# The hypervolume based directed search method for multi-objective optimization problems

Oliver Schütze<sup>1</sup> · Víctor Adrián Sosa Hernández<sup>1</sup> ·  
Heike Trautmann<sup>2</sup> · Günter Rudolph<sup>3</sup>

Received: 28 September 2015 / Revised: 29 January 2016 / Accepted: 18 April 2016 /  
Published online: 30 April 2016  
© Springer Science+Business Media New York 2016

**Abstract** We present a new hybrid evolutionary algorithm for the effective hypervolume approximation of the Pareto front of a given differentiable multi-objective optimization problem. Starting point for the local search (LS) mechanism is a new division of the decision space as we will argue that in each of these regions a different LS strategy seems to be most promising. For the LS in two out of the three regions we will utilize and adapt the Directed Search method which is capable of steering the search into any direction given in objective space and which is thus well suited for the problem at hand. We further on integrate the resulting LS mechanism into SMS-EMOA, a state-of-the-art evolutionary algorithm for hypervolume approximations. Finally, we will present some numerical results on several benchmark problems with two and three objectives indicating the strength and competitiveness of the novel hybrid.

**Keywords** Multi-objective optimization · Evolutionary computation · Memetic algorithm · Directed search method · Hypervolume

---

✉ Oliver Schütze  
schuetze@cs.cinvestav.mx

<sup>1</sup> Computer Science Department, CINVESTAV-IPN, Av. IPN 2508, Col. San Pedro Zacatenco, Mexico C. P. 07360, Mexico

<sup>2</sup> Department of Information Systems, University of Münster, Leonardo-Campus 3, 48149 Münster, Germany

<sup>3</sup> Fakultät für Informatik, Technische Universität Dortmund, 44221 Dortmund, Germany

## 1 Introduction

In many applications one is faced with the problem that several objectives have to be optimized concurrently resulting in a *multi-objective optimization problem* (MOP). One important characteristic of MOPs is that their solution sets, the so-called Pareto sets, respectively their images, the Pareto fronts, do not consist of a single point as for 'classical' scalar optimization problems. Instead, these sets typically form  $(k - 1)$ -dimensional objects, where  $k$  is the number of objectives involved in the given MOP (Hillermeier 2001).

For the treatment of MOPs specialized evolutionary algorithms, multi-objective evolutionary algorithms (EMOAs), have caught the interest of many researchers (see, e.g., Coello et al. 2007; Deb 2001; Rozenberg et al. 2012; Zhang and Li 2007 and references therein). Reasons for this include that EMOAs are applicable to a wide range of problems, are of global nature and hence in principle not depending on the initial candidate set (i.e., the initial population). Further, due to their set based approach they allow to compute a finite size representation of the entire Pareto set in a single run of the algorithm.

On the other hand, it is widely accepted that EMOAs tend to converge slowly resulting in a relatively high number of function evaluations needed to obtain a suitable representation of the set of interest. As a possible remedy, researchers have proposed *memetic* or *hybrid* strategies (e.g., Bosman 2012; Ishibuchi et al. 2003; Jaskiewicz et al. 2012; Knowles and Corne 2005; Lara et al. 2010; Moscato 1989; Shukla 2007; Vasile and Zuiani 2011). Algorithms of that type hybridize local search strategies coming e.g. from mathematical programming techniques with EMOAs in order to obtain fast and reliable global search procedures.

The quality of the approximation of the Pareto front may be measured by the scalar-valued hypervolume indicator (HV) Zitzler (1999). The idea to integrate this indicator in the selection process of EMOAs was proposed in Knowles (2002) and Knowles and Corne (2003) and later generalized for arbitrary indicators (Zitzler and Künzli 2004). In general, indicator-based EMOAs (also known as IBEAs) aim at generating a set of solutions maximizing a specific performance indicator. The SMS-EMOA (Beume et al. 2006) is known as a powerful state-of-the-art EMOA internally maximizing the hypervolume indicator with particular merits in case of more than 2 objectives (Wagner et al. 2007). This performance measure has a number of appealing properties (see e.g. Bringmann and Friedrich 2010a,b; Knowles and Corne 2002; Zitzler et al. 2003) like strict monotonicity and Pareto compliance. The distributions obtained using the HV are biased toward the knee regions of the Pareto front (Auger et al. 2009). However, computational complexity increases exponentially with the problem dimension. Since the hypervolume indicator (as well as any other performance indicator) induces an auxiliary scalar optimization problem which is defined on the set of archives its gradient can be computed (Emmerich and Deutz 2014) which gives rise to set based mathematical programming techniques as done in Hernández et al. (2014).

In this work, we present a hybrid of a particular LS mechanism tailored for hypervolume approximations with the SMS-EMOA. The basis for the LS mechanism, the Hypervolume Directed Search (HVDS), is a certain region division which is possible

for MOPs with differentiable objectives. We will argue that a different local search strategy is most promising if the candidate solution  $x$  is either (i) 'far away' from the Pareto set, (ii) 'near', or (iii) 'in between'. In case  $x$  is 'far away' we propose to use a steepest descent like method in order to approach the rough location of the Pareto set as fast as possible, while for candidate solutions in the other two regions a particular steering is advisable. For this, we will use and adapt the recently proposed *Directed Search* (DS) method (Schütze et al. 2016) as this method is capable of steering the search from a given point into any direction  $d$  given in objective space. Thus, DS is well suited for the problem at hand as the hypervolume indicator is defined in objective space. We present HVDS both as standalone algorithm and as local search engine within SMS-EMOA leading to SMS-EMOA-HVDS which is the first hybrid EMOA for hypervolume approximations of the Pareto front. Numerical experiments on several benchmark MOPs with two and three objectives show the benefit of the new approach. A preliminary study of this work can be found in Hernández et al. (2013) which is restricted to bi-objective problems.

The idea to utilize steering features within hybrid evolutionary algorithms is not new. Brown and Smith (2005) use the multi-objective gradient (Fliege and Fux 2000) together with the so-called negative gradient simplex to obtain well-spread solutions along the Pareto front. Bosman (2012) uses a line search procedure as local search engine within an evolutionary algorithm such that the according movement in objective space (i.e.,  $F(x_{i+i}) - F(x_i)$ , where  $x_i$  is the current and  $x_{i+1}$  the novel iterate) is equal to a prescribed direction in objective space. Finally, the Directed Search method (Schütze et al. 2016) utilizes a similar idea than the one from Bosman (2012) but uses this to steer the search both toward and along the Pareto front. The main differences of these works to the current study are (i) the region division that is proposed here in order to perform a different local search in each of these regions, and (ii) that the method is tailored to the context of HV approximations which requires separate considerations.

The remainder of this paper is organized as follows: in Sect. 2, we briefly state the background required for the understanding of the sequel. In Sect. 3, we present the region division and the local searcher HVDS that is tailored for hypervolume maximization. In Sect. 4, we will present SMS-EMOA-HVDS which is a hybrid of SMS-EMOA and HVDS. In Sect. 5, we will show some numerical results and comparisons to demonstrate the strength of the novel hybrid. Finally, we conclude and show possible paths for future research in Sect. 6.

## 2 Background

In the following we consider continuous multi-objective optimization problems (MOPs) that can be expressed as

$$\min_{x \in Q} F(x), \quad (1)$$

where  $Q \subset \mathbb{R}^n$  is the domain and  $F : Q \rightarrow \mathbb{R}^k$  is defined as the vector of the objective functions

$$F(x) = (f_1(x), \dots, f_k(x))^T, \quad (2)$$

and where each objective  $f_i : Q \rightarrow \mathbb{R}$  is assumed to be sufficiently smooth. In this work, we will concentrate on unconstrained problems, i.e., problems of the form (1) where  $Q = \mathbb{R}^n$ .

The optimality of a MOP is defined by the concept of *dominance*.

- Definition 1** (a) Let  $v, w \in \mathbb{R}^k$ . Then the vector  $v$  is *less than*  $w$  ( $v <_p w$ ), if  $v_i < w_i$  for all  $i \in \{1, \dots, k\}$ . The relation  $\leq_p$  is defined analogously.  
 (b) A vector  $y \in Q$  is *dominated* by a vector  $x \in Q$  ( $x < y$ ) with respect to (1) if  $F(x) \leq_p F(y)$  and  $F(x) \neq F(y)$ , else  $y$  is called *non-dominated* by  $x$ .  
 (c) A point  $x \in Q$  is called (*Pareto*) *optimal* or a *Pareto point* if there is no  $y \in Q$  which dominates  $x$ .  
 (d) The set  $P_Q$  of all Pareto optimal solutions is called the *Pareto set* and its image  $F(P_Q)$  the *Pareto front*.

The Directed Search (DS) method is a point-wise iterative search procedure that is capable of steering the search from a given point into a desired direction  $d \in \mathbb{R}^k$  in objective space (Schütze et al. 2016). To be more precise, given a point  $x \in \mathbb{R}^n$ , a search direction  $v \in \mathbb{R}^n$  is sought such that

$$\lim_{t \searrow 0} \frac{f_i(x_0 + tv) - f_i(x_0)}{t} = d_i, \quad i = 1, \dots, k. \tag{3}$$

Such a direction vector  $v$  solves the following system of linear equations:

$$J(x_0)v = d. \tag{4}$$

Hereby,  $J(x)$  denotes the Jacobian of  $F$  at  $x$ , i.e.,

$$J(x) = \begin{pmatrix} \nabla f_1(x)^T \\ \vdots \\ \nabla f_k(x)^T \end{pmatrix} \in \mathbb{R}^{k \times n}, \tag{5}$$

where  $\nabla f_i(x)$  denotes the gradient of the  $i$ -th objective  $f_i$  at  $x$ . Since typically  $k \ll n$ , we can assume that the system in Eq. (4) is (highly) underdetermined. Among the solutions of Eq. (4), the one with the least 2-norm can be viewed as the greedy direction for the given context. This solution is given by

$$v_+ := J(x)^+ d, \tag{6}$$

where  $J(x)^+$  denotes the pseudo inverse of  $J(x)$ . If the rank of  $J := J(x)$  is  $k$  (i.e., maximal) the pseudo inverse is given by  $J^+ = J^T(JJ^T)^{-1}$ . Since there is no restriction on  $d$  the search can be steered in any direction, e.g., toward and along the Pareto set.

A commonly accepted measure for assessing the quality of an approximation is the so-called *hypervolume indicator* (Zitzler 1999).

- Definition 2** Let  $v^{(1)}, v^{(2)}, \dots, v^{(\mu)} \in \mathbb{R}^k$  be a non-dominated set and  $R \in \mathbb{R}^k$  such that  $v^{(i)} < R$  for all  $i = 1, \dots, \mu$ . The value

$$H(v^{(1)}, \dots, v^{(\mu)}; R) = \Lambda_k \left( \bigcup_{i=1}^{\mu} [v^{(i)}, R] \right) \tag{7}$$

is termed the *dominated hypervolume* or value of the hypervolume indicator with respect to *reference point*  $R = (r_1, \dots, r_k)^T$ , where  $\Lambda_k(\cdot)$  denotes the Lebesgue measure in  $\mathbb{R}^k$ .

This measure has a number of appealing properties, however, determining its value is getting the more tedious the larger the number of objectives is being considered (Beume 2009).<sup>1</sup> In case of two objectives ( $k = 2$ ) and a lexicographically ordered non-dominated set  $v^{(1)}, v^{(2)}, \dots, v^{(\mu)}$  the calculation of (7) reduces to

$$H(v^{(1)}, \dots, v^{(\mu)}; R) = [r_1 - v_1^{(1)}] \cdot [r_2 - v_2^{(1)}] + \sum_{i=2}^{\mu} [r_1 - v_1^{(i)}] \cdot [v_2^{(i-1)} - v_2^{(i)}].$$

### 3 HVDS: a local search algorithm for hypervolume maximization

In this section we present HVDS, a point-wise iterative local search method for hypervolume maximization. The underlying idea is to divide the search space into three different regions since points in each region demand a different treatment. We will propose such a region division and the according local search in each region.

#### 3.1 Motivation

Assume first that a point  $x \in Q$  assigned for local search is ‘far away’ from the Pareto set which is quite likely in early stages of the search process. It is thus desired for the new iterate  $x_{new}$  to approach the set of interest (as a whole) as fast as possible while the particular location of  $x_{new}$  with respect to the Pareto set is rather secondary. On the contrary, the following discussion shows that a steering of the search (e.g., via DS) might be even hindering the overall performance: in Brown and Smith (2005) it has been observed that the objectives’ gradients of a given MOP typically point toward similar directions if  $x$  is ‘far away’ from the Pareto set (if existing). Consider for simplicity the extreme case, namely that all gradients point into the same direction, i.e., let

$$\nabla f_i(x) = \lambda_i \nabla f_1(x), \quad i = 1, \dots, k, \tag{8}$$

where  $\lambda_i > 0, i = 1, \dots, k$ . If  $x_{new}$  is obtained via line search it can be written as

$$x_{new} = x + t v, \tag{9}$$

where  $t \in \mathbb{R}_+$  is the step size and  $v \in \mathbb{R}^n$  the chosen search direction. Note that  $v$  defines the movement in decision space, and that the related movement in objective

---

<sup>1</sup> An overview about recent results on the complexity of HV computation is provided in <http://simco.gforge.inria.fr/doku.php?id=openproblems>.

space is given by  $J(x)v$  (this can be seen from Eq. (3)). Thus, in case (8) holds, we have

$$J(x)v = \begin{pmatrix} \nabla f_1(x)^T v \\ \dots \\ \nabla f_k(x)^T v \end{pmatrix} = \nabla f_1(x)^T v \cdot \lambda, \tag{10}$$

where  $\lambda = (\lambda_1, \dots, \lambda_k)^T \in \mathbb{R}^k$ . Note that the vector in (10) is a multiple of  $\lambda$ , i.e., we obtain a one-dimensional movement regardless of the choice of  $v$  which is  $n$ -dimensional. Though (8) and thus (10) is apparently only true if the distance of  $x$  to the Pareto set is infinite, the range of directions  $d = J(x)v$  in which the search can be steered with a satisfactory step size is expected to be small for points  $x$  that are sufficiently far away from the Pareto set. We refer to Schütze et al. (2016) for a more thorough discussion. Thus, rather than using DS (or any other steering mechanism) it seems to be wise to perform a greedy search toward the Pareto set.

Second, we consider the other extreme situation, namely that  $x$  is already nearly optimal which happens in later stages of the search process. In that case no further significant improvements can be expected when searching along descent directions. Instead, it seems to be wise to perform the search along the Pareto front.

Finally, points do apparently not have to be ‘far away’ nor ‘near’ to the Pareto set. In that case we suggest to search along descent directions that (locally) improve the hypervolume value.

To summarize, we suggest to divide the decision space  $Q$  into three regions. These together with their suggested local search are as follows:

- **Region I**  $x$  is ‘far away’ from the Pareto set (denoted by ‘ $x \in I$ ’). In that case, a greedy search toward the rough location of the Pareto front is desired.
- **Region II**  $x$  is ‘in between’, i.e., neither ‘far away’ nor ‘near’ to the Pareto front. In that case, a descent direction has to be selected such that a movement in that direction maximizes the hypervolume.
- **Region III**  $x$  is ‘near’ to the Pareto set. In that case, a search *along* the Pareto front will be performed.

In the following we propose a particular way to realize the region division and will further on discuss the local search in each region.

### 3.2 Region division

Assume we are given a MOP whose objectives are differentiable. Let

$$R_Q := \{x \in Q : \nabla f_i(x) \neq 0, i = 1, \dots, k\} \tag{11}$$

and define  $\eta : R_Q \rightarrow \mathbb{R}$  via

$$\eta(x) = \left\| \sum_{i=1}^k \tilde{\alpha}_i \frac{\nabla f_i(x)}{\|\nabla f_i(x)\|_2} \right\|_2^2, \tag{12}$$

where  $\tilde{\alpha}$  is a solution of the following optimization problem

$$\min_{\alpha \in \mathbb{R}} \left\{ \left\| \sum_{i=1}^k \alpha_i \frac{\nabla f_i(x)}{\|\nabla f_i(x)\|_2} \right\|_2^2 : \alpha_i \geq 0, i = 1, \dots, k, \sum_{i=1}^k \alpha_i = 1 \right\}. \tag{13}$$

The following result is the basis for the region division.

**Proposition 1** *Let all objectives  $f_i, i = 1, \dots, k$ , be continuously differentiable in  $x \in R_Q, \eta$  be as in (12) and  $x \in R_Q$ . Then*

- (a)  $\eta(x) \in [0, 1]$
- (b)  $\eta(x) = 1 \Leftrightarrow \nabla f_i(x) = \lambda_i \nabla f_1(x)$  with  $\lambda_i > 0$  for  $i = 1, \dots, k$ .
- (c)  $\eta(x) = 0 \Leftrightarrow x \in R_Q$  is a Karush–Kuhn–Tucker (KKT) point.
- (d)  $\eta(x)$  is a continuous mapping.

*Proof* (a) Clearly,  $\eta(x) \geq 0$ . Further, it is

$$\eta(x) = \left\| \sum_{i=1}^k \tilde{\alpha}_i \frac{\nabla f_i(x)}{\|\nabla f_i(x)\|_2} \right\|_2^2 \leq \sum_{i=1}^k \tilde{\alpha}_i \left\| \frac{\nabla f_i(x)}{\|\nabla f_i(x)\|_2} \right\|_2^2 = \sum_{i=1}^k \tilde{\alpha}_i = 1. \tag{14}$$

(b) “ $\Rightarrow$ ”: by strict convexity of the function in (13) it follows directly that

$$\frac{\nabla f_1(x)}{\|\nabla f_1(x)\|_2} = \dots = \frac{\nabla f_k(x)}{\|\nabla f_k(x)\|_2} \tag{15}$$

which is equivalent to (8).

“ $\Leftarrow$ ”: let  $\alpha$  be a convex weight, then

$$\sum_{i=1}^k \alpha_i \frac{\nabla f_i(x)}{\|\nabla f_i(x)\|_2} = \sum_{i=1}^k \alpha_i \frac{\lambda_i \nabla f_1(x)}{\|\lambda_i \nabla f_1(x)\|_2} = \sum_{i=1}^k \alpha_i \frac{\nabla f_1(x)}{\|\nabla f_1(x)\|_2} = \frac{\nabla f_1(x)}{\|\nabla f_1(x)\|_2} \tag{16}$$

by which it follows that  $\eta(x) = 1$ .

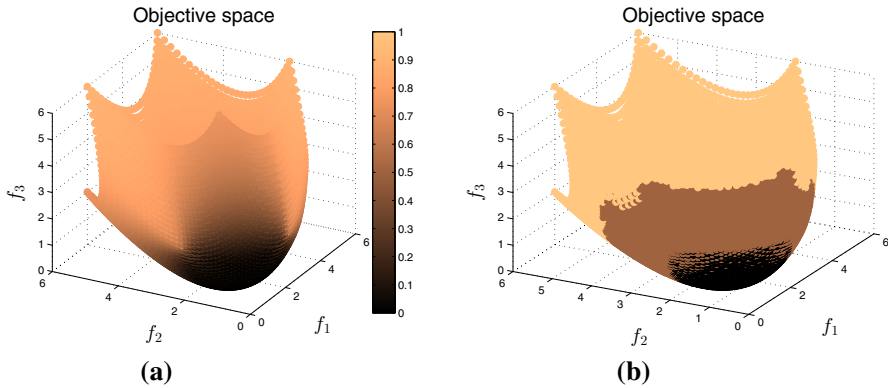
(c) “ $\Rightarrow$ ”: let  $\eta(x) = 0$ , then there exists a convex weight  $\alpha$  s.t.

$$\sum_{i=1}^k \alpha_i \frac{\nabla f_i(x)}{\|\nabla f_i(x)\|_2} = \sum_{i=1}^k \frac{\alpha_i}{\|\nabla f_i(x)\|_2} \nabla f_i(x) = 0. \tag{17}$$

Thus, choosing  $\tilde{w}_i := \alpha_i / \|\nabla f_i(x)\|_2, i = 1, \dots, k$ , the vector  $w := \tilde{w} / \|\tilde{w}\|_1$  is a convex weight with  $\sum_{i=1}^k w_i \nabla f_i(x) = 0$ , i.e.,  $x$  is a KKT point.

“ $\Leftarrow$ ”: now let  $x \in R_Q$  be a KKT point. That is, there exists a convex weight  $\alpha$  s.t.

$$\sum_{i=1}^k \alpha_i \nabla f_i(x) = \sum_{i=1}^k \alpha_i \|\nabla f_i(x)\|_2 \frac{\nabla f_i(x)}{\|\nabla f_i(x)\|_2}. \tag{18}$$



**Fig. 1** Values of  $\eta$  and possible region division for MOP (20). **a** Values of  $\eta$ . **b** Region division

Choosing  $\tilde{w}_i := \alpha_i \|\nabla f_i(x)\|_2, i = 1, \dots, k$ , then  $w := \tilde{w} / \|\tilde{w}\|_1$  is a convex weight such that  $\sum_{i=1}^k w_i \nabla f_i(x) / \|\nabla f_i(x)\|_2 = 0$ , i.e.,  $\eta(x) = 0$ .  
 (d) follows by construction of  $\eta$ . □

Thus, by the above result and using the observation made in [Brown and Smith \(2005\)](#) we can conclude that (i)  $x$  is far away from the Pareto set iff  $\eta(x)$  is close to 1, (ii)  $x$  is near to the Pareto set iff  $\eta(x)$  is close to 0, and (iii) else  $x$  is in between. Hence, by determining two values  $a, b \in (0, 1)$  with  $a < b$  we can make the region division via

$$\begin{aligned}
 x \in I &\Leftrightarrow \eta(x) \geq b \\
 x \in II &\Leftrightarrow \eta(x) \in (a, b) \\
 x \in III &\Leftrightarrow \eta(x) \leq a
 \end{aligned}
 \tag{19}$$

As example consider the three objective problem

$$\begin{aligned}
 f_1, f_2, f_3 &: \mathbb{R}^3 \rightarrow \mathbb{R} \\
 f_i(x) &= \sum_{j=1}^3 (x_j - a_j^i)^2,
 \end{aligned}
 \tag{20}$$

where  $a^1 = (1, 0, 0)^T, a^2 = (0, 1, 0)^T$  and  $a^3 = (0, 0, 1)^T$ . Figure 1a shows the value of  $\eta$  (in objective space for a better visualization) and Fig. 1b a possible division of the image space into the three regions. For the region division we have used  $a = 0.1$  and  $b = 0.5$ .

*Remark 1* For the important special case  $k = 2$  (i.e., for bi-objective problems) one can simplify the region division by using alternatively the angle between the two gradients:

$$\cos \alpha(x) = \frac{\langle \nabla f_1(x), \nabla f_2(x) \rangle}{\|\nabla f_1(x)\|_2 \|\nabla f_2(x)\|_2} \in [-1, 1].
 \tag{21}$$



Using (21) one can proceed analogously to (19) and set

$$\begin{aligned}
 x \in I &\Leftrightarrow \cos \alpha(x) \geq \tilde{b} \\
 x \in II &\Leftrightarrow \cos \alpha(x) \in (a, b) \\
 x \in III &\Leftrightarrow \cos \alpha(x) \leq \tilde{a}
 \end{aligned}
 \tag{22}$$

for  $\tilde{a}, \tilde{b} \in (-1, 1)$  with  $\tilde{a} < \tilde{b}$ . In the previous study (Hernández et al. 2013) the values  $\tilde{a} = -0.8$  and  $\tilde{b} = 0.8$  were proposed.

### 3.3 Performing the local search in each region

In the following we will propose a search strategy for each of the three regions. For this, we will first consider the extreme case that the given archive/population consists of only one element, and will later on consider the general case. Reason for this is that if a point  $x$  is assigned for local search in a general archive, this can be led back to the case of one element archives.

#### 3.3.1 One element archives

Assume we are given an archive  $A = \{x\}$ , i.e., a point  $x \in Q$  which is assigned for local search and a reference point  $R = (r_1, \dots, r_k)^T$  for the hypervolume indicator.

*Local search in Region I* As the above discussion shows the explicit steering of the search is a delicate problem. Thus, it seems to be wise to utilize established multi-objective descent methods in order to approach the Pareto set/front as fast as possible. Such methods can be found e.g. in Fliege et al. (2009), Fliege and Fux (2000), and Schäffler et al. (2002). A descent direction  $v$  at a point  $x$  satisfies  $\langle \nabla f_i(x), v \rangle < 0$ ,  $i = 1, \dots, k$ .

For the treatment of bi-objective problems we have used the descent direction proposed in Lara et al. (2010)

$$v = \frac{1}{2} \left( \frac{\nabla f_1(x)}{\|\nabla f_1(x)\|_2} + \frac{\nabla f_2(x)}{\|\nabla f_2(x)\|_2} \right)
 \tag{23}$$

as it does not require to solve an additional optimization problem coupled with an Armijo-like step size control as used in Lara et al. (2010). Note, however, that the descent direction in (23) is restricted to bi-objective problems and has no analogon for MOPs with more than two objectives. For the treatment of such problems we have used the descent method proposed in Schäffler et al. (2002) which requires to solve an additional optimization problem to compute the search direction. More precisely, we choose the vector  $-q(x)$  which is defined in the following.

**Theorem 1** (Schäffler et al. 2002) *Let problem (1) be given and the map  $q : \mathbb{R}^n \rightarrow \mathbb{R}^n$  be defined by*

$$q(x) = \sum_{i=1}^k \hat{\alpha}_i \nabla f_i(x), \tag{24}$$

where  $\hat{\alpha}$  is a solution of

$$\min_{\alpha \in \mathbb{R}^k} \left\{ \left\| \sum_{i=1}^k \alpha_i \nabla f_i(x) \right\|_2^2 ; \alpha_i \geq 0, i = 1, \dots, k, \sum_{i=1}^k \alpha_i = 1 \right\}. \tag{25}$$

Then the following statements hold.

- (a) Either  $q(x) = 0$  or  $-q(x)$  is a descent direction.
- (b) For each  $\hat{x} \in \mathbb{R}^n$ , there exists a neighborhood  $N(\hat{x})$  and a constant  $L_{\hat{x}} \in \mathbb{R}_0^+$  such that

$$\|q(x) - q(y)\|_2 \leq L_{\hat{x}} \|x - y\|_2, \quad \forall x, y \in N(\hat{x}). \tag{26}$$

Note that if  $q(x) = 0$ , then  $x$  is a KKT point. Thus, the computation of  $-q(x)$  includes a test for first order optimality.

*Local search in Region II* If  $x$  is in Region II, the task is to find a descent direction  $d_{II} <_p 0$  (in objective space) such that a movement in that direction maximizes the hypervolume. Since the new iterate  $x_{new}$  is performed via DS, we can (idealized) write its image  $y_{new} = F(x_{new})$  as

$$y_{new} = F(x) + td_{II}, \tag{27}$$

where  $t \in \mathbb{R}$  is a given (fixed) step size and  $d_{II}$  has to be chosen such that it solves the  $k$ -dimensional problem

$$\begin{aligned} \max_{d \in \mathbb{R}^k} \text{vol}(d) &= \prod_{i=1}^k (r_i - f_i(x) - td_i) \\ \text{s.t. } \|d\|_2^2 &= 1, \end{aligned} \tag{28}$$

which is related to the maximization of the hypervolume (compare to Fig. 2). If one replaces the 2-norm by the infinity norm in the constraint of problem (28) (which drops the assumption that the movement is done with an equal step in objective space) a straightforward computation shows that

$$d_{II,\infty} = F(x) - R \tag{29}$$

solves the problem (where we assume that  $F(x) <_p R$ ). For our implementations we have used the direction in Eq. (29) since it is easier to calculate and since we observed that it yields no difference in the performance of the algorithm compared to the solution of problem (28). Note, however, that the solution of (28) may come with additional CPU time but not with additional function evaluations.

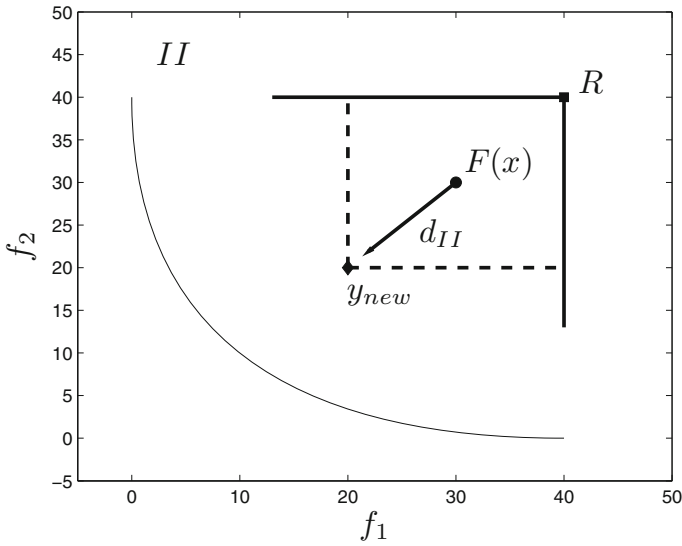


Fig. 2 Local search in Region II

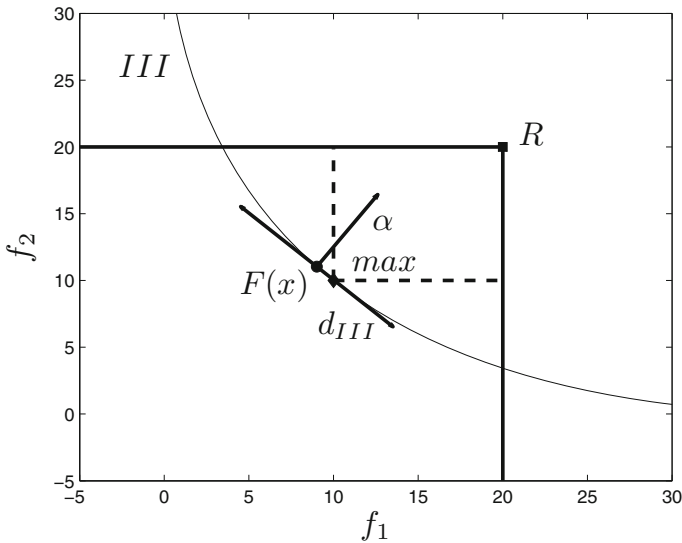


Fig. 3 Movement along linearized Pareto front in order to improve the hypervolume in Region III

*Local search in Region III* If  $x$  is nearly optimal, a search in a descent direction will not lead to significant improvements any more. Instead, it seems to be wise to continue the search along the Pareto set. Here, we suggest to perform a search via DS along the linearized Pareto front (see Fig. 3). The particular steps of the search are described in the following for KKT points, the steps for nearly optimal solutions are the same.

- (1) Compute the KKT weight  $\alpha$  associated to  $x$ . This can be done by solving (25).

(2) It is known that  $\alpha$  is orthogonal to the linearized Pareto front at  $F(x)$  (Hillermeier 2001). Thus, an orthonormal basis of the tangent space can be obtained as follows: compute a  $QU$ -factorization of  $\alpha$ ,

$$\alpha = QU = (q_1, q_2, \dots, q_k)U. \tag{30}$$

Since  $Q \in \mathbb{R}^{k \times k}$  is orthogonal, the vectors  $q_2, \dots, q_k$  span the desired tangent space.

(3) Now we can maximize the hypervolume in objective space similar to (28). The differences here are that (i) we are restricting the search to directions  $d \in \text{span}(q_2, \dots, q_k)$  and (ii) the optimization will not only yield the optimal direction but also the optimal step size (i.e.,  $y_{new}$  those corresponding decision vector  $x_{new}$  can be retrieved via DS). The  $(k - 1)$ -dimensional unconstrained optimization problem is given by

$$\max_{\lambda \in \mathbb{R}^{k-1}} \text{vol}(\lambda) = \prod_{i=1}^k (r_i - f_i(x) - (Q_2\lambda)_i), \tag{31}$$

where

$$Q_2 = (q_2, \dots, q_k) \in \mathbb{R}^{k \times k-1}. \tag{32}$$

The search direction is thus given by

$$d_{III} = Q_2\lambda^*, \tag{33}$$

where  $\lambda^*$  is a solution of (31).

(4) Using DS we compute  $x_{new}$  such that  $y_{new} \approx F(x) + Q_2\lambda^*$ . If the hypervolume value is not increased by  $x_{new}$  we perform the Armijo-like backtracking approach as in Lara et al. (2010) using the hypervolume value as objective function.

For the special case  $k = 2$  the solution of (31) can be expressed analytically if all the elements of the KKT weight are positive. Let  $\alpha = (\alpha_1, \alpha_2)^T$  be the KKT weight associated to  $x$ , then the tangent vector of the linearized Pareto front at  $F(x)$  is e.g. given by  $d = (-\alpha_2, \alpha_1)^T$  and problem (31) can be stated as

$$\max_{\lambda \in \mathbb{R}} \text{vol}_2(\lambda) = (r_1 - f_1(x) - \lambda d_1) \times (r_2 - f_2(x) - \lambda d_2). \tag{34}$$

**Proposition 2** *Let  $k = 2$  and  $\alpha >_p 0$ , then the global maximizer of Problem (34) is given by*

$$\lambda^* = \frac{d_1 r_2 + d_2 r_1 - d_1 f_2(x) - d_2 f_1(x)}{2d_1 d_2}. \tag{35}$$

*Proof* If  $\alpha_p >_p 0$ , then it follows that also  $d_1, d_2 \neq 0$ . The first derivative of  $\text{vol}_2$  is given by

$$\text{vol}_2'(\lambda) = 2\lambda d_1 d_2 + d_2 f_1(x) - r_1 d_2 + d_1 f_2(x) - d_1 r_2 \tag{36}$$

Setting this to zero leads to

$$\lambda^* = \frac{d_1 r_2 + d_2 r_1 - d_1 f_2(x) - d_2 f_1(x)}{2d_1 d_2}. \tag{37}$$

Further, the second derivative at  $\lambda^*$  is given by

$$vol_2''(\lambda^*) = 2d_1d_2 < 0. \tag{38}$$

The negativity holds since  $\alpha >_p 0$  and by construction of  $d$ , and the claim follows.  $\square$

*Remark 2* We stress that due to the linearization of the Pareto front at  $F(x)$  it can happen that the new iterate  $x_{new}$  is located in Region II leading to a kind of oscillating behavior in the sequence of iterates. We have, however, not observed any instabilities in our computations. Further, if  $x$  is already near to the optimal solution, such oscillations do typically not happen.

Algorithm 1 shows the pseudo code of the HVDS as standalone algorithm for one-element archives which puts together the above discussion. Figure 4 shows some exemplary iterations of HVDS in all three regions of MOP (20).

---

**Algorithm 1** HVDS as standalone algorithm for one element archives

---

**Require:**  $x_0$ : starting point,  $a, b$ : values for region assignment;  $R$ : reference point

**Ensure:** : sequence  $\{x_i\}$  of candidate solutions

$i := 0$

**repeat**

    Compute  $\eta(x_i)$  as in Eq. (12)

**if**  $\eta(x_i) \geq b$  **then**  $\triangleright x_i \in I$

        Compute  $v_I$  (e.g. as in Eq. (23) for  $k = 2$ )

        Compute  $t_I \in \mathbb{R}_+$

$x_{i+1} = x_i + t_I v_I$

**else if**  $\eta(x_i) \in (b, a)$  **then**  $\triangleright x_i \in II$

$d_{II} = F(x_i) - R$

$v_{II} = J(x_i)^+ d_{II}$

        Compute  $t_{II} \in \mathbb{R}_+$

$x_{i+1} = x_i + t_{II} v_{II}$

**else**  $\triangleright x_i \in III$

        Compute the KKT weight  $\alpha$  as in Eq. (25)

        Compute  $Q_2$  as in Eq. (32)

        Compute  $\lambda^*$  by solving (31)

$d_{III} = Q_2 \lambda^*$

        Compute  $t_{III} \in \mathbb{R}_+$

$x_{i+1} = x_i + t_{III} * v_{III}$   $\triangleright$  compute new iterate

**end if**

$i := i + 1$

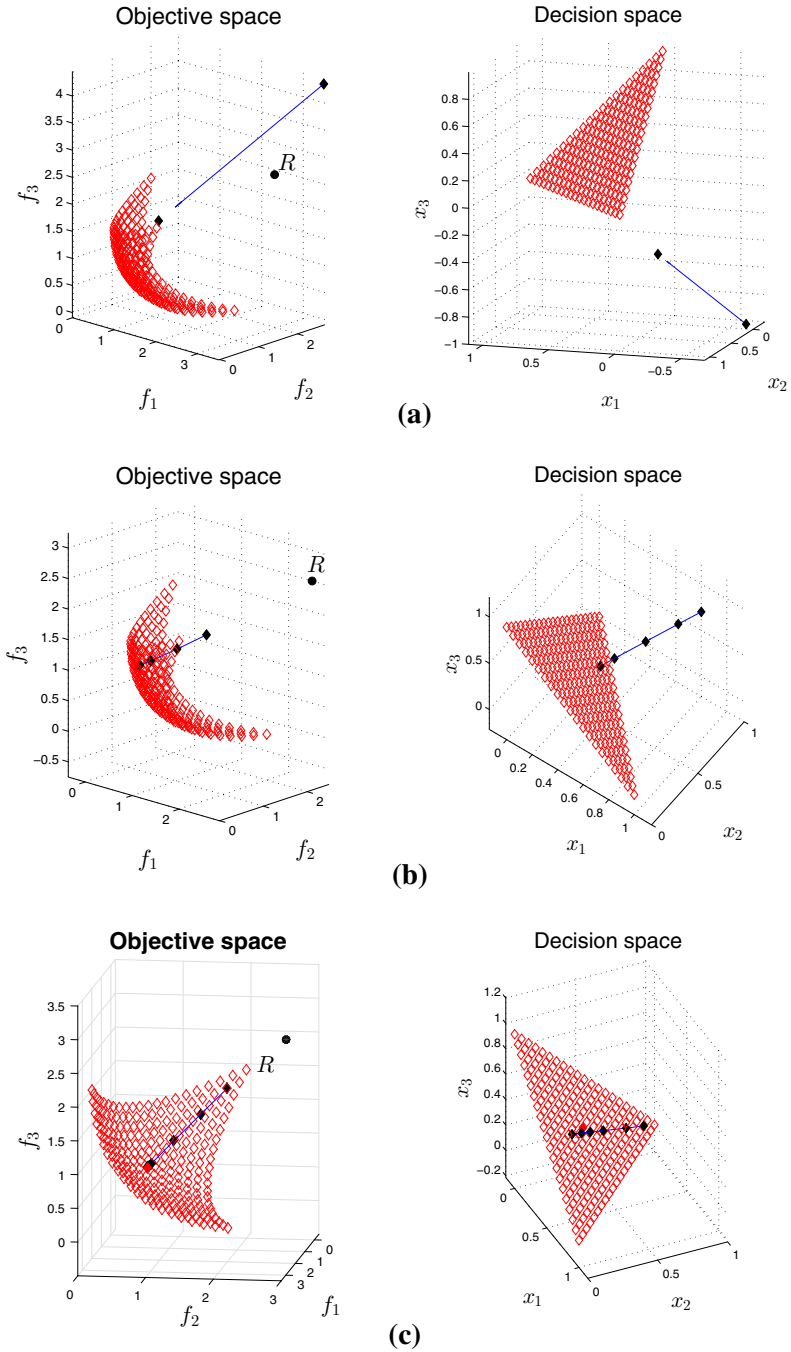
**until**  $t_{III} = 0$  or a maximum number of iterations is reached

---

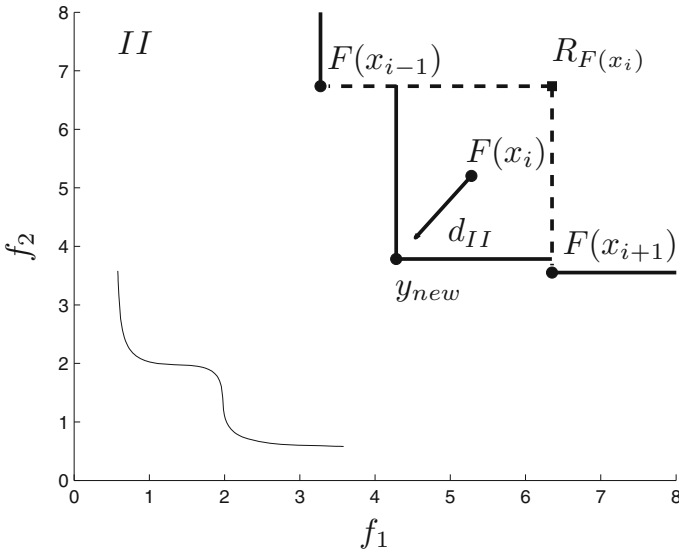
### 3.3.2 General archives

In the following we describe the adaption of the HVDS for archives with arbitrary sizes.

*Local search in Region I* If a point  $x \in A$  is chosen for local search that is according to the region division 'far away' from the Pareto front, an improvement toward the



**Fig. 4** Exemplary iterations of the HVDS in each region of MOP (20). **a** Region I. **b** Region II. **c** Region III



**Fig. 5** Local search in Region II for multiple archive entries

solution set is desired regardless of the location of the other elements of  $A$ . Hence, we recommend to proceed as the one element archive by choosing a descent direction from Eq. (23) in the case of  $k = 2$  or proceed as in Schäffler et al. (2002) for  $k > 2$ .

*Local search in Region II* If a point  $x \in A$  is located in the Region II it is necessary to consider its neighboring solutions in  $A$  since the chosen movement in objective space should ideally increase the contribution of the newly found solutions  $x_{new}$  computed from  $x$  without decreasing the contribution of the other elements in  $A$ . To perform this movement we have to look for the  $k$  neighboring entries of  $x$  in  $A$  in order to set a new reference point. For  $k = 2$  the adaption comes straightforward since the neighborhood from a point  $x$  is given by only two points for non-extreme points (where the archive entries are sorted according to one objective). For the extreme points we used the updated reference point given by the SMS-EMOA. Figure 5 shows the bi-objective case for setting the reference point  $R_{F(x)}$  for a non-extreme point.

For  $k > 2$ , the problem becomes more complicated since a neighborhood is not well defined (in the case of  $k = 2$  it is enough only ordering the points in descending order to define the neighbors, however for  $k > 2$  it is not the case). To overcome this problem we propose to choose such points that minimize the difference between each objective of  $F(x)$  from all non-dominated points in the population. Once having these points we are now in the position to construct a new  $R_F$  and use Equation (29) for obtaining the improvement direction. Algorithm 2 shows the pseudo code to obtain the reference point.

**Algorithm 2** Set reference point for Region II**Require:** Point  $x$  from  $A$  in Region II and the set of non-dominated points  $N$ .**Ensure:** Reference point  $R$ 

```

for  $i = 1, \dots, k$  do
  Set  $R_i = 0$ 
  Set  $D_i = \infty$ 
  for all points  $n$  in  $N$  do
     $aux = f_i(n) - f_i(x)$ 
    if  $f_i(n) > f_i(x)$  and  $aux < D_i$  then
       $D_i = aux$ 
       $R_i = f_i(n)$ 
    end if
  end for
end for

```

**Algorithm 3** SMS-EMOA-HVDS**Require:** MOP, archive size  $\mu$ , values  $a, b$  for region division**Ensure:** An approximation  $P$  of the given MOP.Initialize a population  $P \subset Q$  with  $\mu$  elements at random**repeat**generate offspring  $x \in Q$  from  $P$  by variationchoose  $p \in [0, 1]$  uniformly at random**if**  $p \leq p_{HVDS}$  **then**choose the generated offspring  $x$  for local search $\tilde{x} = HVDS(x, a, b, R)$ **else** $\tilde{x} = x$ **end if** $P := P \cup \{\tilde{x}\}$ build ranking  $S_1, \dots, S_h$  from  $P$ compute the hypervolume contribution for each  $x \in S_h$ denote by  $x^*$  the element with the least hypervolume contribution $P := P \setminus \{x^*\}$ **until** stopping criterion fulfilled**return**  $P$ 

▷ Apply HVDS

*Local search in Region III* For Region III we proceed to use the new reference point as stated for Region II. To be more precise, we propose to use the reference point  $R_{F(x_i)}$  for intermediate points (i.e.,  $i \in \{2, \dots, l-1\}$ ) and the original point  $R$  for the extreme archive entries. In this case we follow the same process as the one element archive using the generated reference point.

## 4 Integrating HVDS into SMS-EMOA

Here we make a first attempt to integrate HVDS into an EMOA. As base algorithm we have chosen to take SMS-EMOA since it is a state-of-the-art evolutionary algorithm that aims for optimal hypervolume approximations of the Pareto front. HVDS is integrated into SMS-EMOA as additional generational operator with a certain probability  $p_{HVDS}$  for application (for an empirical evaluation of the values of  $p_{HVDS}$ ,  $a$ , and  $b$  we refer to the next section). To be more precise, the new hybrid algorithm



**Table 1** Test function problems

Multi-objective optimization problems

DTLZ 1 (Linear front - 2D,3D) domain  $\mathbf{x} \in [0, 1]^n$

$$f_1(\mathbf{x}) = (1 + g(\mathbf{x}))0.5 \prod_{i=1}^{k-1} x_i$$

$$f_{m=2:k-1}(\mathbf{x}) = (1 + g(\mathbf{x}))0.5(\prod_{i=1}^{k-m} x_i)(1 - x_{k-m+1})$$

$$f_k(\mathbf{x}) = (1 + g(\mathbf{x}))0.5(1 - x_1)$$

$$g(\mathbf{x}) = 100 * (n + \sum_{i=1}^n ((x_i - 0.5)^2 - \cos(20 * \pi(x_i - 0.5))))$$

DTLZ 2 (Concave front - 2D,3D) domain  $\mathbf{x} \in [0, 1]^n$

$$f_1(\mathbf{x}) = (1 + g(\mathbf{x})) \prod_{i=1}^{k-1} \cos(x_i \frac{\pi}{2})$$

$$f_{m=2:k-1}(\mathbf{x}) = (1 + g(\mathbf{x}))(\prod_{i=1}^{k-m} \cos(x_i \frac{\pi}{2})) \sin(x_{k-m+1} \frac{\pi}{2})$$

$$f_k(\mathbf{x}) = (1 + g(\mathbf{x})) \sin(x_1 \frac{\pi}{2})$$

$$g(\mathbf{x}) = \sum_{i=1}^n (x_i - 0.5)^2$$

DTLZ 3 (Concave front - 2D,3D) domain  $\mathbf{x} \in [0, 1]^n$

$$f_1(\mathbf{x}) = (1 + g(\mathbf{x})) \prod_{i=1}^{k-1} \cos(x_i \frac{\pi}{2})$$

$$f_{m=2:k-1}(\mathbf{x}) = (1 + g(\mathbf{x}))(\prod_{i=1}^{k-m} \cos(x_i \frac{\pi}{2})) \sin(x_{k-m+1} \frac{\pi}{2})$$

$$f_k(\mathbf{x}) = (1 + g(\mathbf{x})) \sin(x_1 \frac{\pi}{2})$$

$$g(\mathbf{x}) = \sum_{i=1}^n (x_i - 0.5)^2$$

DTLZ 7 (Disconnected front - 3D) domain  $\mathbf{x} \in [0, 1]^n$

$$f_{m=1:k-1}(\mathbf{x}) = x_m$$

$$f_k(\mathbf{x}) = (1 + g(\mathbf{x})) \left( k - \sum_{i=1}^{k-1} \left( \frac{f_i}{1+g(\mathbf{x})} (1 + \sin(3 * \pi * f_i)) \right) \right)$$

$$g(\mathbf{x}) = 1 + (9 * \sum_{i=1}^n x_i) / k$$

P1-HVDS (Convex front - 2D) domain  $\mathbf{x} \in [-1, 1]^2$

$$f_1(\mathbf{x}) = (x_1 - 1)^4 + (x_2 - 1)^2$$

$$f_2(\mathbf{x}) = (x_1 + 1)^2 + (x_2 + 1)^2$$

P3-HVDS (Convex-Concave front - 2D) domain  $\mathbf{x} \in [-3, 3]^2$

$$f_1(\mathbf{x}) = \frac{1}{2} * (\sqrt{1 + (x_1 + x_2)^2} + \sqrt{1 + (x_1 - x_2)^2} + x_1 - x_2) + \lambda * e^{(-1*(x_1-x_2)^2)}$$

$$f_2(\mathbf{x}) = \frac{1}{2} * (\sqrt{1 + (x_1 + x_2)^2} + \sqrt{1 + (x_1 - x_2)^2} - x_1 + x_2) + \lambda * e^{(-1*(x_1-x_2)^2)}$$

Shifted ZDT1 (Convex front - 2D) domain  $\mathbf{x} \in [0, 1] \times [-1, 1]^{n-1}$

$$f_1(\mathbf{x}) = x_1$$

$$f_2(\mathbf{x}) = g(\mathbf{x}) * (2 - \sqrt{f_1(\mathbf{x})/g(\mathbf{x})})$$

$$g(\mathbf{x}) = 1 + \frac{9}{n-1} \sum_{i=1}^n x_i^2$$

Shifted ZDT2 (Concave front - 2D) domain  $\mathbf{x} \in [0, 1] \times [-1, 1]^{n-1}$

$$f_1(\mathbf{x}) = x_1$$

$$f_2(\mathbf{x}) = g(\mathbf{x}) * (2 - (f_1(\mathbf{x})/g(\mathbf{x}))^2)$$

$$g(\mathbf{x}) = 1 + \frac{9}{n-1} \sum_{i=1}^n x_i^2$$

Shifted ZDT3 (Disconnected front - 2D) domain  $\mathbf{x} \in [0, 1] \times [-1, 1]^{n-1}$

$$f_1(\mathbf{x}) = x_1$$

$$f_2(\mathbf{x}) = g(\mathbf{x}) * (2 - \sqrt{f_1(\mathbf{x})/g(\mathbf{x})}) - (f_1(\mathbf{x})/g(\mathbf{x})) \sin(10\pi * f_1(\mathbf{x}))$$

$$g(\mathbf{x}) = 1 + \frac{9}{n-1} \sum_{i=1}^n x_i^2$$

**Table 1** continued

Multi-objective optimization problems

Shifted ZDT6 (Concave front - 2D) domain  $\mathbf{x} \in [0, 1] \times [-1, 1]^{n-1}$

$$f_1(\mathbf{x}) = 1 - e^{-4 * x_1} * \sin(6 * \pi * x_1)$$

$$f_2(\mathbf{x}) = g(\mathbf{x}) * (2 - (f_1(\mathbf{x})/g(\mathbf{x}))^2)$$

$$g(\mathbf{x}) = 1 + 9 * \left( \frac{\sum_{i=1}^n x_i^2}{n-1} \right)^{0.25}$$

Convex2 (Convex front - 3D) domain  $\mathbf{x} \in [-5, 5]^n$

$$f_i(\mathbf{x}) = \sum_{\substack{j=1 \\ j \neq i}}^n (x_j - a_j^i)^2 + (x_i - a_i^i)^4$$

$$a^1 = (1, 1, 1, 1, \dots), a^2 = (-1, -1, -1, -1, \dots), a^3 = (1, -1, 1, -1, \dots)$$

**Table 2** Configuration values for SMS-EMOA-HVDSMO and SMS-EMOA, where *DI* means “distribution index” and *p* is the probability to apply the operator

Description	Parameter	2D values	3D values
Crossover operator	SBX	<i>DI</i> = 20.0, <i>p</i> = 0.9	<i>DI</i> = 20.0, <i>p</i> = 0.9
Mutation operator	PM	<i>DI</i> = 20.0, <i>p</i> = 1/ <i>n</i>	<i>DI</i> = 20.0, <i>p</i> = 1/ <i>n</i>
Population size	$\mu$	100	300
LS first parameter	<i>a</i>	0.2	0.1
LS second parameter	<i>b</i>	0.4	0.4
Probability for LS	<i>p<sub>HVDS</sub></i>	0.005	0.005

SBX stands for simulated binary crossover and PM for polynomial mutation

SMS-EMOA-HVDS starts by initializing a population *P* of  $\mu$  elements at random from the domain *Q*. Once having the initialized population the process begins by generating only one offspring *x* by genetic operators (crossover and mutation). HVDS is integrated as an additional generational operator for the produced new elements in each generation. That is, after producing a new element the algorithm selects a value  $p \in [0, 1]$  uniformly at random. If *p* is less or equal than the probability *p<sub>HVDS</sub>* the local search procedure will be applied, otherwise the new candidate solution is added to *P*. It is important to mention that the probability *p<sub>HVDS</sub>* balances the emphasis between global and local search during the whole run of the algorithm. The aim to apply HVDS is to push the entire population toward the solution set during the algorithm process. In this case the produced element can be everywhere within the objective space, however, thanks to the region division we can make the decision of which LS strategy should be chosen in order to improve the hypervolume. Each time the HVDS strategy is called, it executes only one iteration step in our current implementation. Further, we use the dynamic reference point described above to get *R*. Algorithm 3 shows the pseudocode of the resulting algorithm SMS-EMOA-HVDS.

**Table 3** HV results of SMS-EMOA with and without HVDS as operator (using the same number of function evaluations)

Problem	# Func	n	Average	SMS-EMOA-HVDSMO		
				Deviation	Median	Covered HV (%)
DTLZ1 (2D) M	30K	15	<b>118.25014</b>	3.65605	119.61943	97.82947
DTLZ2 (2D) U	10K	30	<b>120.20997</b>	0.00055	120.21009	99.99944
DTLZ3 (2D) M	30K	15	101.84429	23.49979	105.40498	84.72152
P1-HVDS (2D) U	10K	2	<b>3251.81142</b>	16.40663	3253.60993	99.81105
P3-HVDS (2D) U	10K	2	<b>17.34873</b>	0.00220	17.34852	99.88260
ZDT1 (2D) U	10K	30	<b>109.53759</b>	0.17000	109.60691	99.88643
ZDT2 (2D) U	10K	30	<b>108.61966</b>	1.95078	109.22509	99.35129
ZDT3 (2D) U	10K	30	116.73398	1.46083	117.67874	99.11522
ZDT6 (2D) M	10K	15	<b>104.81097</b>	0.97200	104.78108	98.40024
Convex2 (3D) U	10K	15	<b>21.14882</b>	0.03884	21.16804	99.99858
DTLZ1 (3D) M	50K	15	<b>1330.93727</b>	0.07304	1330.97379	99.99723
DTLZ2 (3D) U	10K	30	<b>1330.41646</b>	0.00501	1330.41637	99.99920
DTLZ3 (3D) M	50K	15	<b>1326.62757</b>	5.85269	1330.17426	99.71441
DTLZ7 (3D) M	30K	15	964.42986	57.13546	992.99807	97.10511
Problem	# Func	n	Average	SMS-EMOA		
				Deviation	Median	Covered HV (%)
DTLZ1 (2D) M	30K	15	117.40941	5.64543	119.24526	97.13393
DTLZ2 (2D) U	10K	30	120.20604	0.00130	120.20637	99.99617
DTLZ3 (2D) M	30K	15	<b>105.68880</b>	9.48479	107.17194	87.91967
P1-HVDS (2D) U	10K	2	3248.72641	22.57145	3253.23842	99.71636
P3-HVDS (2D) U	10K	2	17.34866	0.00270	17.34891	99.88220
ZDT1 (2D) U	10K	30	109.26564	0.63385	109.57209	99.63844
ZDT2 (2D) U	10K	30	106.42534	3.45765	109.13577	97.34421
ZDT3 (2D) U	10K	30	<b>117.05177</b>	1.27276	117.69188	99.38512
ZDT6 (2D) M	10K	15	97.60659	1.49610	97.41107	91.63651
Convex2 (3D) U	10K	15	21.10150	0.04299	21.09881	99.77483
DTLZ1 (3D) M	50K	15	1330.93127	0.07912	1330.97013	99.99678
DTLZ2 (3D) U	10K	30	1330.41083	0.00288	1330.41086	99.99878
DTLZ3 (3D) M	50K	15	1325.70533	8.01075	1330.19390	99.64509
DTLZ7 (3D) M	30K	15	<b>993.01561</b>	0.02094	993.01893	99.98331

The values are obtained over 20 test runs and  $k = 2, 3$

### 5 Experiments

Systematic experiments were conducted in order to investigate and validate the behavior of the proposed algorithm SMS-EMOA-HVDS. For this purpose well-established two-objective (2D) and three-objective (3D) test functions from the ZDT (shifted and

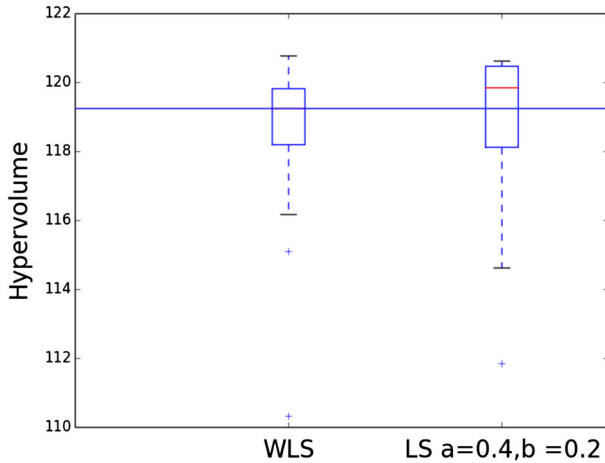


Fig. 6 DTLZ1 (2D)

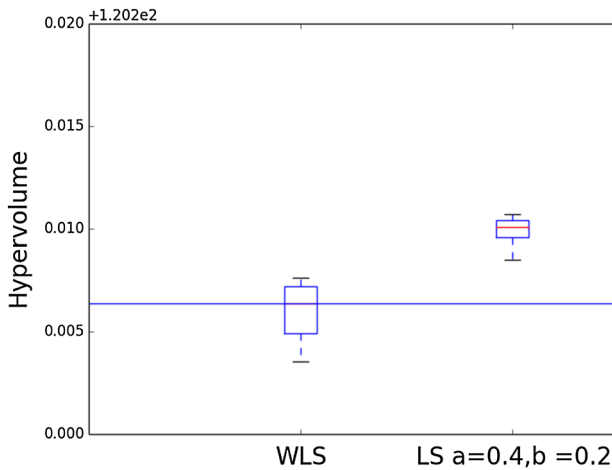


Fig. 7 DTLZ2 (2D) (★)

thus unconstrained versions, see [Lara et al. 2010](#), [Shukla 2007](#), and [Shukla 2007](#)) and DTLZ test suites were chosen supplemented by specific functions (see [Table 1](#) to look the description of the problems). The domains  $Q$  for all problems have been chosen such that the Pareto sets have an empty intersection with the boundary of  $Q$ . All the details regarding the use of SMS-EMOA and its hybrid variant can be found in [Table 2](#).

The algorithm was applied 20 times with population sizes 100 (2D) respectively 300 (3D). The value of the hypervolume indicator (HV) of the final result is used as basis of comparison to respective runs of the original SMS-EMOA. Based on extensive systematic investigations the settings  $a = 0.2, b = 0.4$  (2D) resp.  $a = 0.1$  (3D) together with  $p_{HVD_S} = 0.005$  turned out to be most stable over all settings while in

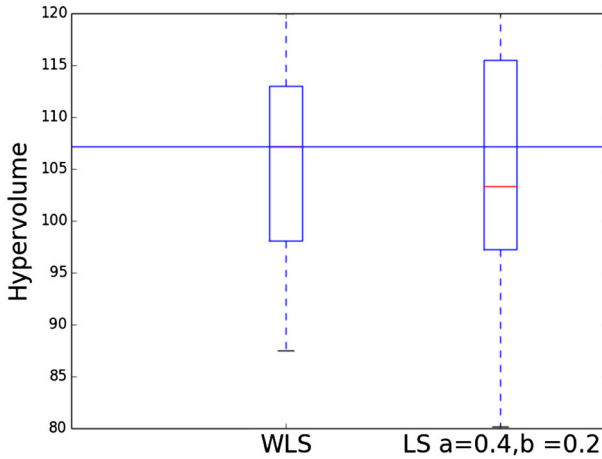


Fig. 8 DTLZ3 (2D)

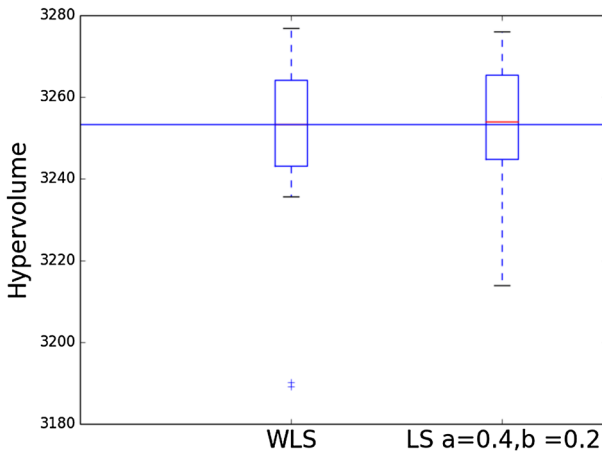


Fig. 9 P1-HVDS (2D)

general algorithm behavior is not extremely sensitive regarding the respective settings of  $a$  and  $b$ . A dynamic reference point was used for Region II (see Algorithm 2).

Table 3 provides detailed algorithm settings along with the experimental results and thereby allows for the assessment of the effect of the HVDS operator.

For all test functions, the superior algorithm regarding the average HV is highlighted in italic that the sophisticated local search strategy indeed has a positive effect on algorithm performance. The corresponding ordering regarding median performance only changes slightly. Moreover, on all considered problems the percentage of the covered optimal hypervolume is very high while keeping in mind that the number of function evaluations has not been specifically tuned to gain optimal hypervolume. The optimal HV values for all the problems have been computed by numerically solving the following  $(n \cdot \mu)$ -dimensional scalar optimization problem

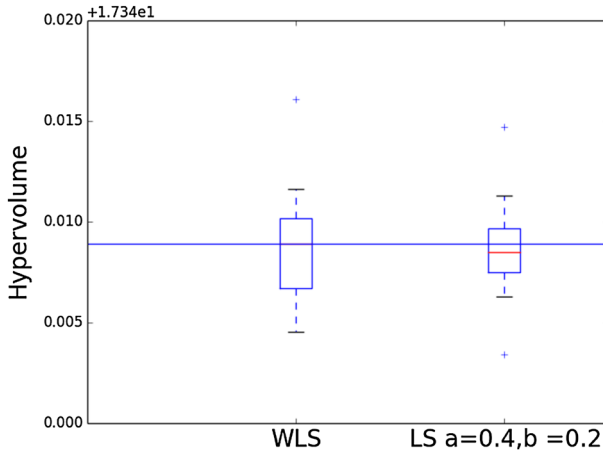


Fig. 10 P3-HVDS (2D)

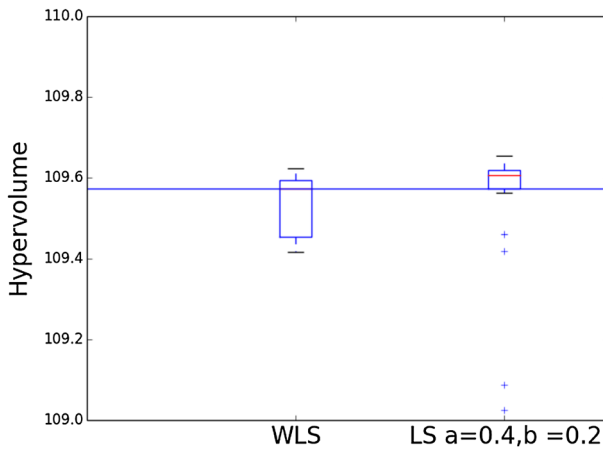


Fig. 11 ZDT1 (2D) (★)

$$\begin{aligned}
 &H : \mathbb{R}^{n \cdot \mu} \rightarrow \mathbb{R} \\
 &\max_{v^{(1)}, \dots, v^{(\mu)} \in \mathbb{R}^n} H(v^{(1)}, \dots, v^{(\mu)}; R), \tag{39}
 \end{aligned}$$

where  $\mu$  denotes the population sizes.

Robustness and statistical significance of the results can be assessed by observing boxplots of the obtained HV in the final generation over all runs in Figs. 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, and 19. The respective boxes range from the lower up to the upper quartile while the median is highlighted by a horizontal line inside the box. Points are plotted individually as outliers if they are larger than box boundaries plus resp. minus 1.5 times the interquartile range. The plotted whisker terminates at the adjacent point, which is the most extreme point that is not an outlier. Results with (LS)

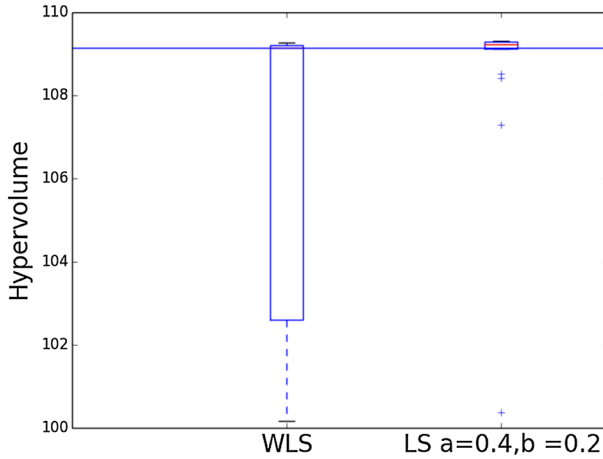


Fig. 12 ZDT2 (2D) (★)

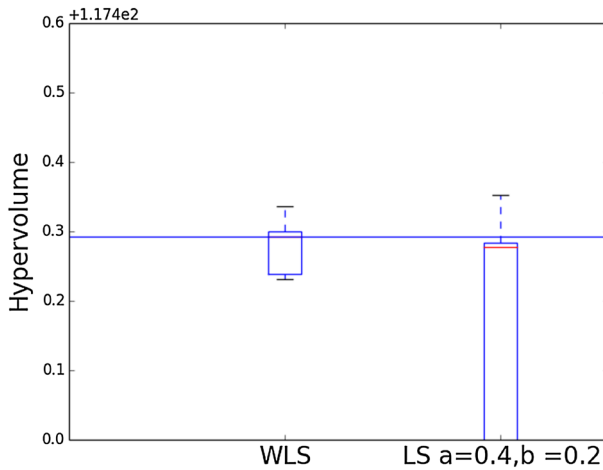


Fig. 13 ZDT3 (2D)

and without local search (WLS) are compared while the horizontal line visualizes the median HV of the original SMS-EMOA runs.

Figures marked with (★) reflect a statistically significant result based on the Wilcoxon Rank test using a significance level  $\alpha = 0.05$ . It becomes obvious that in 3D the SMS-EMOA supplemented with the HVDS strategy outperforms the original version for almost all test functions. Exceptions are DTLZ3, but in this case much more stable results were generated, and the disconnected problem DTLZ7. In 2D, results are comparable for all test functions and superiority of the local search variant is given for DTLZ3, ZDT1, ZDT2 and ZDT6. However, due to the sophisticated strategy in Region III which searches *along* the Pareto front, the algorithm in general faces special challenges for disconnected fronts such as ZDT3 and DTLZ7.

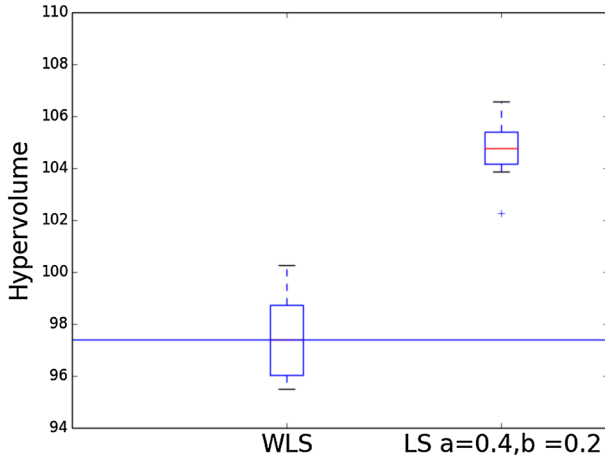


Fig. 14 ZDT6 (2D) (★)

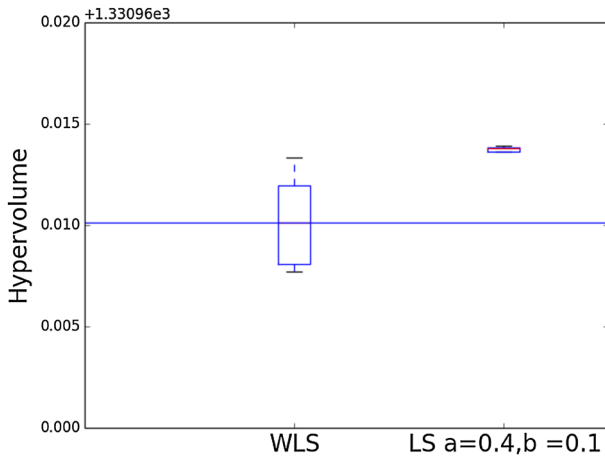


Fig. 15 DTLZ1 (3D) (★)

Moreover, observing the algorithms' behavior along the whole run, SMS-EMOA-HVDS often shows faster convergence in earlier phases of optimization, see the exemplary Figs. 20 and 21.

## 6 Conclusions and future work

In this paper, we have proposed a new hybrid evolutionary algorithm for the effective hypervolume approximation of the Pareto fronts of differentiable multi-objective optimization problems. The basis of the local search mechanism is a novel division of the decision space into three regions as we have argued that in each region another local search seems to be most promising. For the local search we have adapted the



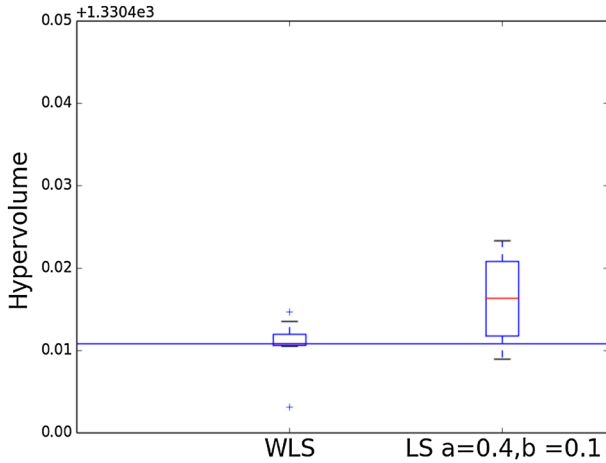


Fig. 16 DTLZ2 (3D) (★)

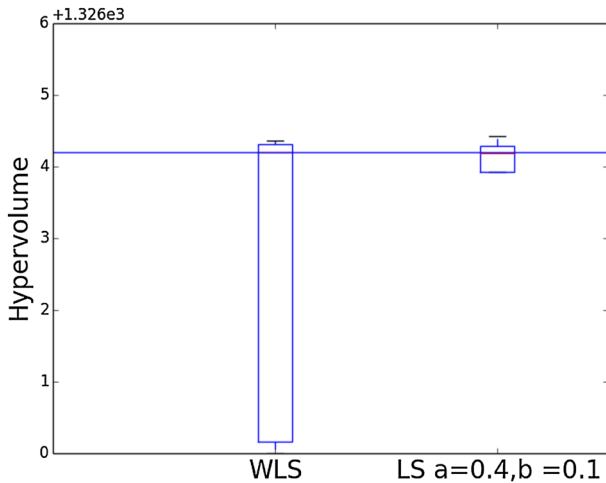


Fig. 17 DTLZ3 (3D) (★)

Directed Search method that is able to steer the search into any direction given in objective space and is thus well suited for the problem at hand as the hypervolume is defined in objective space. We have presented the resulting iterative method, HVDS, as standalone algorithm as well as local searcher within a hybrid variant of SMS-EMOA. Experimental results showed that especially in 3D, the integration of HVDS into the SMS-EMOA has a substantial positive effect on the algorithm performance while in 2D the results are at least comparable to the original SMS-EMOA but often superior as well. With respect to the considered test functions which cover a wide range of problem characteristics, robust parameter settings for the SMS-EMOA-HVDS could be provided. Therefore, we see promising perspectives for even higher dimensions and further research will focus on generalizing the introduced strategy in this direction.

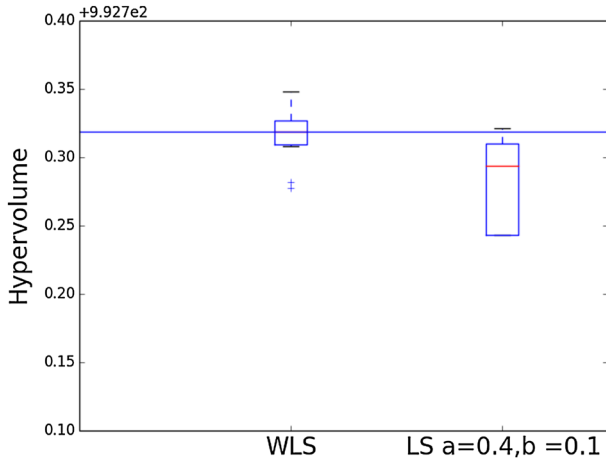


Fig. 18 DTLZ7 (3D) (★)

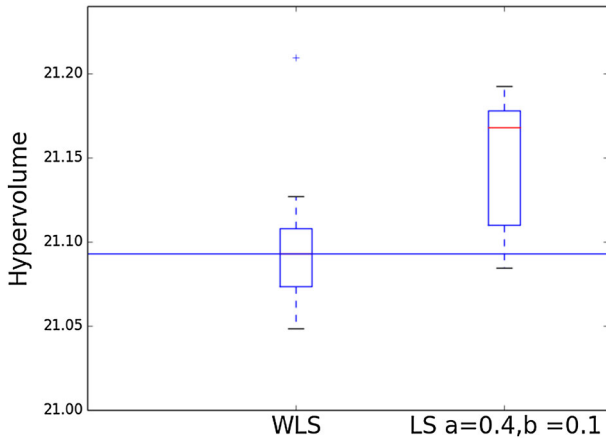


Fig. 19 Convex3 (3D) (★)

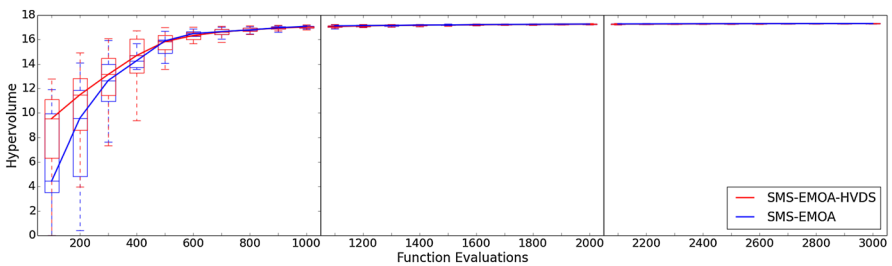
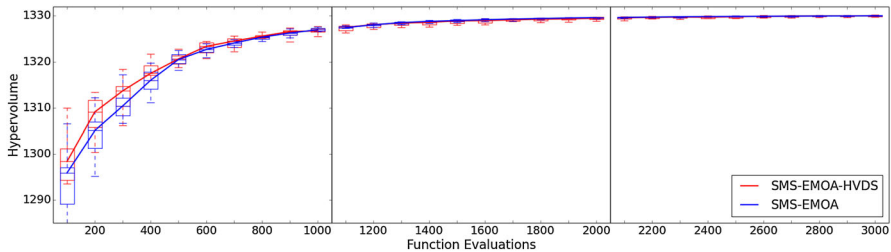


Fig. 20 Comparison using boxplots in P3 HVDS (2D) during 3000 function evaluations

Further paths of future research include the adaption of the method to constrained problems which needs separate considerations. Further, it is desirable to obtain a gradient-free realization of HVDS, including the region division, in order to increase



**Fig. 21** Comparison using boxplots in DTLZ2 (3D) during 3000 function evaluations

the applicability of the method to a broader class of methods. Finally, it would be interesting to investigate if and to which extent the chosen method can be applied to other performance indicators.

**Acknowledgments** Victor Adrian Sosa Hernández acknowledges support by the Consejo Nacional de Ciencia y Tecnología (CONACYT). Heike Trautmann acknowledges support by the European Center of Information Systems (ERCIS). All authors acknowledge support from CONACYT Project no. 174443 and DFG Project no. TR 891/5-1

## References

- Auger, A., Bader, J., Brockhoff, D., Zitzler, E.: Theory of the hypervolume indicator: optimal  $\mu$ -distributions and the choice of the reference point. In: Proceedings of the Tenth ACM SIGEVO Workshop on Foundations of Genetic Algorithms, FOGA '09, pp. 87–102. ACM, New York, NY, USA (2009)
- Beume, N.: S-Metric calculation by considering dominated hypervolume as Klee's measure problem. *Evolut. Comput.* **17**(4), 477–492 (2009)
- Beume, N., Naujoks, B., Emmerich, M.: SMS-EMOA: multiobjective selection based on dominated hypervolume. *Eur. J. Oper. Res.* **181**(3), 1653–1669 (2006)
- Bosman, P.A.N.: On gradients and hybrid evolutionary algorithms for real-valued multiobjective optimization. *IEEE Trans. Evolut. Comput.* **16**(1), 51–69 (2012)
- Bringmann, K., Friedrich, T.: The maximum hypervolume set yields near-optimal approximation. In: Proceedings of the ACM Genetic and Evolutionary Computation Conference (GECCO), pp. 511–518. ACM Press (2010)
- Bringmann, K., Friedrich, T.: Tight bounds for the approximation ratio of the hypervolume indicator. In: Proceedings of the 11th International Conference Parallel Problem Solving From Nature (PPSN). Lecture Notes in Computer Science, vol. 6238, pp. 607–616, Krakow, Poland, September. Springer (2010)
- Brown, M., Smith, R.E.: Directed multi-objective optimisation. *Int. J. Comput. Syst. Signals* **6**(1), 3–17 (2005)
- Coello, C.A., Lamont, G.B., Van Veldhuizen, D.A.: *Evolutionary Algorithms for Solving Multi-Objective Problems*, 2nd edn. Springer, New York (2007)
- Deb, K.: *Multi-Objective Optimization using Evolutionary Algorithms*. Wiley, Chichester (2001)
- Emmerich, M., Deutz, A.: Time complexity and zeros of the hypervolume indicator gradient field. In: Schütze, O., et al. (eds.) *EVOLVE - A Bridge between Probability, Set Oriented Numerics, and Evolutionary Computation III*. Studies in Computational Intelligence, vol. 500, pp. 169–193. Springer, Berlin (2014)
- Fliege, J., Graña Drummond, L.M., Svaiter, B.F.: Newton's method for multiobjective optimization. *SIAM J. Optim.* **20**, 602–626 (2009)
- Fliege, J., Fux Svaiter, B.: Steepest descent methods for multicriteria optimization. *Math. Methods Oper. Res.* **51**(3), 479–494 (2000)
- Hernández, V.A.S., Schütze, O., Emmerich, M.: Hypervolume maximization via set based newtons method. In: Tantar, A., et al. (eds.) *EVOLVE—A Bridge Between Probability, Set Oriented Numerics, and*

- Evolutionary Computation V. *Advances in Intelligent Systems and Computing*, vol. 288, pp. 15–28. Springer, Berlin (2014)
- Hernández, V.A.S., Schütze, O., Rudolph, G., Trautmann, H.: The directed search method for pareto front approximations with maximum dominated hypervolume. In: Emmerich, M., et al. (eds.) *EVOLVE—A Bridge Between Probability, Set Oriented Numerics, and Evolutionary Computation IV. Advances in Intelligent Systems and Computing*, vol. 227, pp. 189–205. Springer, Berlin (2013)
- Hilliermeier, C.: *Nonlinear Multiobjective Optimization: A Generalized Homotopy Approach*. Springer, Basel (2001)
- Ishibuchi, H., Yoshida, T., Murata, T.: Balance between genetic search and local search in memetic algorithms for multiobjective permutation flowshop scheduling. *IEEE Trans. Evolut. Comput.* **7**(2), 204–223 (2003)
- Jaszkiewicz, A., Ishibuchi, H., Zhang, Q.: Multiobjective memetic algorithms. *Handbook of Memetic Algorithms. Studies in Computational Intelligence*, vol. 379, pp. 201–217. Springer, Berlin (2012)
- Knowles, J., Corne, D.: Memetic algorithms for multiobjective optimization: issues, methods and prospects. In: Hart, W.E., Krasnogor, N., Smith, J.E. (eds.) *Recent Advances in Memetic Algorithms. Studies in Fuzziness and Soft Computing*, vol. 166, pp. 313–352. Springer, Berlin (2005)
- Knowles, J.D.: Local-search and hybrid evolutionary algorithms for pareto optimization. PhD thesis, University of Reading, UK (2002). (Section 4.3.4 “S metric Archiving”)
- Knowles, J.D., Corne, D.: Properties of an adaptive archiving algorithm for storing nondominated vectors. *IEEE Trans. Evolut. Comput.* **7**(2), 100–116 (2003)
- Knowles, J., Corne, D.: On metrics for comparing nondominated sets. In: *Proceedings of the 2002 Congress on Evolutionary Computation (CEC’02)*, vol. 1, pp. 711–716. IEEE (2002)
- Lara, A., Coello Coello, C. A., Schütze, O.: A painless gradient-assisted multi-objective memetic mechanism for solving continuous bi-objective optimization problems. In: *2010 IEEE Congress on Evolutionary Computation (CEC)*, pp. 1–8. IEEE Press, IEEE (2010)
- Lara, A., Sanchez, G., Coello Coello, C.A., Schütze, O.: HCS: a new local search strategy for memetic multiobjective evolutionary algorithms. *IEEE Trans. Evolut. Comput.* **14**(1), 112–132 (2010)
- Moscatto, P.: On evolution, search, optimization, genetic algorithms and martial arts: toward memetic algorithms. Technical Report C3P Report 826, Caltech Concurrent Computation Program (1989)
- Rozenberg, G., Bäck, T., Kok, J.N.: *Handbook of Natural Computing*. Springer, Berlin (2012)
- Schäffler, S., Schultz, R., Weinzierl, K.: A stochastic method for the solution of unconstrained vector optimization problems. *J. Optim. Theory Appl.* **114**(1), 209–222 (2002)
- Schütze, O., Martín, A., Lara, A., Alvarado, S., Salinas, E., Coello, C.A.: The directed search method for multiobjective memetic algorithms. *J. Comput. Optim. Appl.* **63**, 305–332 (2016)
- Shukla, P.: On gradient based local search methods in unconstrained evolutionary multi-objective optimization. In: Obayashi, S. et al. (ed), *EMO 2007*, pp. 96–110 (2007)
- Shukla, P.K.: Gradient based stochastic mutation operators in evolutionary multi-objective optimization. *Adaptive and Natural Computing Algorithms*, pp. 58–66. Springer, Berlin (2007)
- Vasile, M., Zuiani, F.: Multi-agent collaborative search: an agent-based memetic multi-objective optimization algorithm applied to space trajectory design. *Proc. Inst. Mech. Eng. G* **225**(11), 1211–1227 (2011)
- Wagner, T., Beume, N., Naujoks, B.: Pareto-, aggregation-, and indicator-based methods in many-objective optimization. In: *Proceedings of the 4th International Conference on Evolutionary Multi-criterion Optimization, EMO’07*, pp. 742–756 (2007). Springer, Berlin
- Zhang, Q., Li, H.: MOEA/D: a multiobjective evolutionary algorithm based on decomposition. *IEEE Trans. Evolut. Comput.* **11**(6), 712–731 (2007)
- Zitzler, E.: *Evolutionary algorithms for multiobjective optimization: methods and applications*. PhD thesis, ETH Zurich (1999)
- Zitzler, E., Künzli, S.: Indicator-based selection in multiobjective search. In: Yao, X. et al., (eds) *Proc. Int’l Conf. on Parallel Problem Solving from Nature (PPSN VIII)*, pp. 832–842. Springer, Berlin (2004)
- Zitzler, E., Thiele, L., Laumanns, M., Fonseca, C.M., Da Fonseca, V.G.: Performance assessment of multiobjective optimizers: an analysis and review. *IEEE Trans. Evolut. Comput.* **7**(2), 117–132 (2003)