

Hypervolume-Optimal μ -Distributions on Line/Plane-based Pareto Fronts in Three Dimensions

Ke Shang, *Member, IEEE*, Hisao Ishibuchi, *Fellow, IEEE*, Weiyu Chen, Yang Nan, and Weiduo Liao

Abstract—Hypervolume is widely used in the evolutionary multi-objective optimization (EMO) field to evaluate the quality of a solution set. For a solution set with μ solutions on a Pareto front, a larger hypervolume means a better solution set. Investigating the distribution of the solution set with the largest hypervolume is an important topic in EMO, which is the so-called hypervolume optimal μ -distribution. Theoretical results have shown that the μ solutions are uniformly distributed on a linear Pareto front in two dimensions. However, the μ solutions are not always uniformly distributed on a single-line Pareto front in three dimensions. They are only uniform when the single-line Pareto front has one constant objective. In this paper, we further investigate the hypervolume optimal μ -distribution in three dimensions. We consider the line- and plane-based Pareto fronts. For the line-based Pareto fronts, we extend the single-line Pareto front to two-line and three-line Pareto fronts, where each line has one constant objective. For the plane-based Pareto fronts, the linear triangular and inverted triangular Pareto fronts are considered. First, we show that the μ solutions are not always uniformly distributed on the line-based Pareto fronts. The uniformity depends on how the lines are combined. Then, we show that a uniform solution set on the plane-based Pareto front is not always optimal for hypervolume maximization. It is locally optimal with respect to a $(\mu + 1)$ selection scheme. Our results can help researchers in the community to better understand and utilize the hypervolume indicator.

Index Terms—Hypervolume indicator, Evolutionary multi-objective optimization, Optimal μ -distribution.

I. INTRODUCTION

THE hypervolume indicator is a popular performance indicator in the field of evolutionary multi-objective optimization (EMO). Informally, the hypervolume of a solution set is the volume of the objective space dominated by the solution set and dominating

a reference point. A well recognized fact in the EMO community is that the hypervolume indicator is able to evaluate the convergence and the diversity of a solution set simultaneously. The hypervolume indicator is strictly Pareto compliant [1], which guarantees that a solution set maximizing the hypervolume indicator are all Pareto optimal [2]. Thus, the hypervolume indicator is adopted in some EMO algorithms (EMOAs) to guide the population converge to the Pareto front. We call these algorithms hypervolume-based EMOAs. Some representative algorithms are SMS-EMOA [3], [4], FV-MOEA [5], HypE [6], and R2HCA-EMOA [7]. For a comprehensive survey of the hypervolume indicator, please refer to [8].

For a hypervolume-based EMOA, a well-converged and widely-distributed solution set on the Pareto front is expected. The convergence of the solution set can be achieved due to the Pareto compliance property of the hypervolume indicator. For the diversity of the solution set, a widely-distributed solution set on the Pareto front is expected by maximizing the hypervolume indicator. Many researchers analyze how the solution set is distributed on the Pareto front when the hypervolume is maximized. This research issue is usually called the hypervolume optimal μ -distribution.

Theoretical studies have proved that for a linear Pareto front in two dimensions, the μ solutions are uniformly distributed on the Pareto front in order to maximize the hypervolume indicator [9], [10]. To the best of our knowledge, this conclusion is the only one to precisely describe the hypervolume optimal μ -distribution on the Pareto front in two dimensions. Auger et al. [9] and Friedrich et al. [11] theoretically investigated the hypervolume optimal μ -distribution of solutions on the two-objective nonlinear Pareto fronts. However, in general, it is not easy to precisely describe the location of the μ solutions on the nonlinear Pareto fronts. In three dimensional cases, Shukla et al. [12] theoretically investigated the hypervolume optimal μ -distribution on a single-line Pareto front (i.e., a degenerated Pareto front). It shows that the uniformity property of the optimal distribution for a linear Pareto front in two dimensions cannot be generalized to three dimensions. The solutions are only uniform on a single-line Pareto front when one objective of the Pareto front is constant (i.e., this Pareto front can be degenerated to a linear Pareto front in two dimensions). Auger et al. [13] theoretically studied the hy-

This work was supported by National Natural Science Foundation of China (Grant No. 62002152, 61876075), Guangdong Provincial Key Laboratory (Grant No. 2020B121201001), the Program for Guangdong Introducing Innovative and Entrepreneurial Teams (Grant No. 2017ZT07X386), The Stable Support Plan Program of Shenzhen Natural Science Fund (Grant No. 20200925174447003), Shenzhen Science and Technology Program (Grant No. KQTD2016112514355531). (*Corresponding Author: Hisao Ishibuchi.*)

K. Shang, H. Ishibuchi, W. Chen, Y. Nan and W. Liao are with Guangdong Provincial Key Laboratory of Brain-inspired Intelligent Computation, Department of Computer Science and Engineering, Southern University of Science and Technology, Shenzhen 518055, China (e-mail: kshang@foxmail.com; hisao@sustech.edu.cn; 11711904@mail.sustech.edu.cn; nany@mail.sustech.edu.cn; 11849249@mail.sustech.edu.cn).

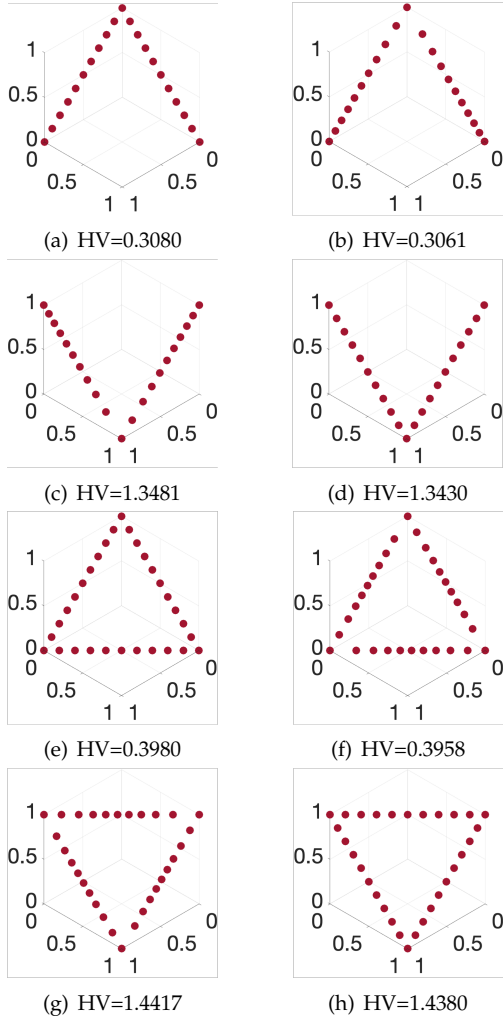


Fig. 1. Uniform and non-uniform solution sets on four types of line-based Pareto fronts. The reference point is specified as $\mathbf{r} = (-0.2, -0.2, -0.2)^1$. The hypervolume of each solution set is shown under each figure.

pervolume optimal μ -distribution in three dimensions. However, the exact hypervolume optimal μ -distribution is not obtained in [13]. Singh [14] theoretically studied the hypervolume behavior when μ is infinite (i.e., the whole Pareto front) whereas the hypervolume optimal μ -distribution is not described. Due to the difficulty of the theoretical study, many studies only investigate the hypervolume optimal μ -distribution empirically [15]–[19], i.e., the hypervolume optimal μ -distribution is approximated on the Pareto front in these studies.

In this paper, we further study the hypervolume optimal μ -distribution in three dimensions, both theoretically and empirically. We consider the line- and plane-based Pareto fronts. The line-based Pareto fronts have two and three lines, where each line has one constant objective. The plane-based Pareto fronts are with triangular and inverted triangular shapes.

For the line-based Pareto fronts, we show that the solutions are not always uniformly distributed. The uniformity depends on how the lines are combined. For

example, Fig. 1 shows four types of the line-based Pareto fronts with uniform and non-uniform solution sets. The hypervolume value of each solution set is also shown in Fig. 1 for the reference point $\mathbf{r} = (-0.2, -0.2, -0.2)^1$. In some cases (i.e., (a) and (e)), the uniform solution set has a larger hypervolume than the non-uniform solution set (i.e., (a) > (b) and (e) > (f)). However, in the other cases (i.e., (c) and (g)), the non-uniform solution set has a larger hypervolume than the uniform solution set (i.e., (c) > (d) and (g) > (h)). In this paper, we investigate this issue and explain why the uniform solution set is not always optimal in Fig. 1.

For the plane-based Pareto fronts, we show that the uniform solution set is not always optimal for hypervolume maximization. It is locally optimal with respect to a $(\mu + 1)$ selection scheme. For example, Fig. 2 shows two types of the plane-based Pareto fronts with uniform and non-uniform solution sets. The hypervolume value of each solution set is shown in Fig. 2 for the reference point $\mathbf{r} = (-1/8, -1/8, -1/8)$. For each Pareto front, the uniform solution set has a smaller hypervolume than the non-uniform solution set (e.g., (a) < (b) and (c) < (d)). In this paper, we investigate this issue and prove the local optimality of the uniform solution set on the plane-based Pareto fronts.

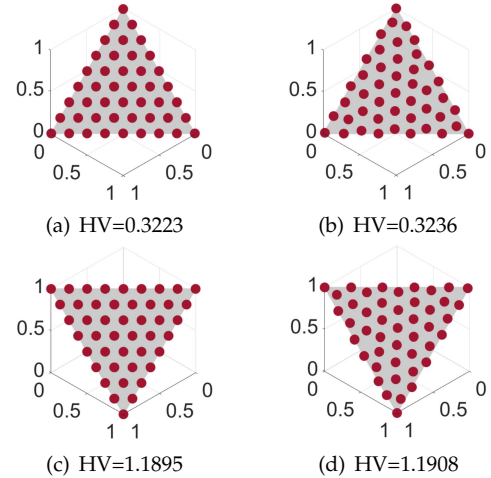


Fig. 2. Uniform and non-uniform solution sets on two types of plane-based Pareto fronts. The reference point is specified as $\mathbf{r} = (-1/8, -1/8, -1/8)$. The hypervolume of each solution set is shown under each figure.

The main contributions of this paper are summarized as follows:

- 1) We extend the single-line Pareto fronts to line-based Pareto fronts with more than one line in three dimensions, and reveal that a uniform solution set cannot always be obtained by maximizing the hypervolume indicator. The solution set is uniform when the lines are the boundaries of a linear triangular Pareto front, whereas it is non-uniform when

¹In this paper, maximization of each objective is assumed in multi-objective optimization problems. Thus, the reference point is in the negative orthant.

the lines are the boundaries of a linear inverted triangular Pareto front.

- 2) In addition to the line-based Pareto fronts, we also investigate the plane-based Pareto fronts in three dimensions. The main result is that a uniform solution set (i.e., a solution set generated by the DAS method [20]) is not always optimal for hypervolume maximization. It is locally optimal with respect to a $(\mu + 1)$ selection scheme.

The rest of the paper is structured as follows. In Section II, the preliminaries of the paper is presented. In Section III, the line-based Pareto fronts are investigated. In Section IV, the plane-based Pareto fronts are investigated. Finally, the conclusions are drawn in Section V.

This paper is an extended version of our conference paper [21]. In our conference paper, we only investigated the line-based Pareto fronts. In this paper, in addition to the line-based Pareto fronts, we also investigate the plane-based Pareto fronts. More discussions on the line- and plane-based Pareto fronts are provided in this paper.

II. PRELIMINARIES

A. Basic Definitions

First, the definitions of the hypervolume indicator, the hypervolume contribution, and the hypervolume optimal μ -distributions are presented.

The hypervolume of a solution set is defined as follows. For a solution set $A \subset \mathbb{R}^m$ and a reference point $\mathbf{r} \in \mathbb{R}^m$, the hypervolume of the solution set A is defined as

$$\text{HV}(A, \mathbf{r}) = \mathcal{L} \left(\bigcup_{\mathbf{a} \in A} \{\mathbf{b} | \mathbf{a} \succeq \mathbf{b} \succeq \mathbf{r}\} \right), \quad (1)$$

where $\mathcal{L}(\cdot)$ is the Lebesgue measure of a set, and $\mathbf{a} \succeq \mathbf{b}$ means that \mathbf{a} Pareto dominates \mathbf{b} (i.e., $a_i \geq b_i$ for all $i = 1, \dots, m$ and $a_j > b_j$ for at least one $j = 1, \dots, m$ in the maximization case).

The hypervolume contribution is an important concept based on the hypervolume indicator. It describes the amount of the hypervolume value contributed by a solution to the solution set. Formally, for a solution $\mathbf{s} \in A$, the hypervolume contribution of \mathbf{s} to A is defined as

$$\text{HVC}(\mathbf{s}, A, \mathbf{r}) = \text{HV}(A, \mathbf{r}) - \text{HV}(A \setminus \{\mathbf{s}\}, \mathbf{r}). \quad (2)$$

Based on the hypervolume indicator, the hypervolume optimal μ -distribution is defined as follows [22]. For a Pareto front $\mathcal{F} \subset \mathbb{R}^m$ and a reference point $\mathbf{r} \in \mathbb{R}^m$, the hypervolume optimal μ -distribution is $\mu \in \mathbb{N}$ points on the Pareto front which maximize the hypervolume of the μ points. The set A containing the optimal μ points is

$$A = \arg \max_{|A'|=\mu, A' \subset \mathcal{F}} \text{HV}(A', \mathbf{r}). \quad (3)$$

B. Hypervolume Optimal μ -Distribution in Two Dimensions

For a linear Pareto front in two dimensions, Emmerich et al. [10] and Auger et al. [9] theoretically show that the μ solutions which maximize the hypervolume value are uniformly distributed on the Pareto front. When the reference point is sufficiently far away from the Pareto front, the optimal μ -distribution includes the two extreme points of the Pareto front. For example, consider the linear Pareto front $f_1 + f_2 = 1$ and $f_1, f_2 \geq 0$, if the reference point $\mathbf{r} = (r, r)$ satisfies $r \leq -1/(\mu - 1)$ [23]², the two extreme points $(0, 1)$ and $(1, 0)$ of the Pareto front are included in the hypervolume optimal μ -distribution.

Fig. 3 gives an illustration of the hypervolume optimal μ -distribution in two dimensions. In Fig. 3, four solutions (i.e., $\mu = 4$) are uniformly distributed on the Pareto front. If we set the reference point to $\mathbf{r} = -1/(\mu - 1)$ (i.e., $r = -1/3$), each solution has the same hypervolume contribution (i.e., the colored square).

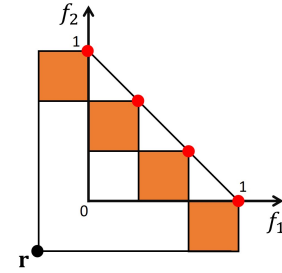


Fig. 3. An illustration of the hypervolume optimal μ -distribution on a linear Pareto front in two dimensions.

C. Hypervolume Optimal μ -Distribution in Three Dimensions

In this paper, all the Pareto fronts are studied in the normalized objective space (i.e., $[0, 1]^3$). The reference point is specified as $\mathbf{r} = (r, r, r)$ (i.e., each element of \mathbf{r} is the same).

Shukla et al. [12] theoretically studied the hypervolume optimal μ -distribution on a single-line Pareto front in three dimensions and showed that the solutions are not always uniformly distributed for hypervolume maximization. A single-line Pareto front in three dimensions is a degenerated Pareto front. For such a single-line Pareto front, it has the following two types:

- 1) **Type I:** Two objectives are conflicting with each other, and the other objective is constant. Fig. 4 (a) shows an example of Type I Pareto front.
- 2) **Type II:** Two objectives are consistent (i.e., not conflicting) with each other, and the other objective is conflicting with the two objectives. Fig. 4 (b) shows an example of Type II Pareto front.

For the Type I Pareto front $f_1 + f_3 = 1, f_1, f_3 \geq 0, f_2 = 0$ in Fig. 4 (a), when the reference point is specified as

²In the case of minimization, this condition is rewritten as $r \geq 1 + 1/(\mu - 1)$.

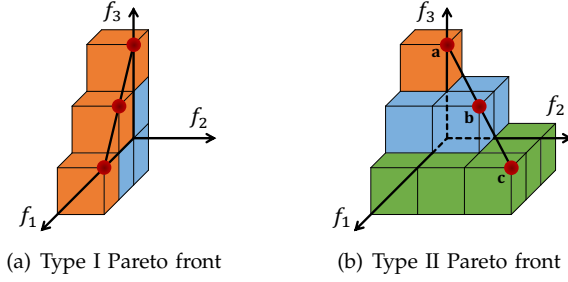


Fig. 4. Single-line Pareto fronts in three dimensions.

$r = -1/(\mu - 1)$, each solution of the uniform solution set has the same hypervolume contribution. We can see that the hypervolume of the three solutions on Type I Pareto front can be calculated as the two-dimensional hypervolume of the three solutions in f_1 - f_3 space multiplying $|r|$. Thus, maximizing the hypervolume in three dimensions is equivalent to maximizing the hypervolume in two dimensions (i.e., f_1 - f_3 space). Therefore, the solutions are uniformly distributed on the Type I Pareto front.

For the Type II Pareto front $f_1 + f_3 = 1, f_1, f_3 \geq 0, f_1 = f_2$ in Fig. 4 (b), when the reference point is specified as $r = -1/(\mu - 1)$, each solution of the uniform solution set has a different hypervolume contribution. For example, the hypervolume contribution of each solution $\mathbf{a} = (0, 0, 1)$, $\mathbf{b} = (0.5, 0.5, 0.5)$, and $\mathbf{c} = (1, 1, 0)$ in Fig. 4 (b) is 0.5^3 , 3×0.5^3 , and 5×0.5^3 , respectively. \mathbf{c} has a larger hypervolume contribution than \mathbf{b} , and \mathbf{b} has a larger hypervolume contribution than \mathbf{a} . If we move \mathbf{b} towards \mathbf{c} to a new location $\mathbf{b}' = (0.6, 0.6, 0.4)$ as illustrated in Fig. 5, then the hypervolume contribution of \mathbf{b}' is $(0.5^2 \times 3 + 0.1^2 + 0.1 \times 0.5 \times 4) \times 0.4 = 0.384 > 0.375 = 3 \times 0.5^3$. This means that the whole hypervolume in Fig. 4 (b) can be improved by moving \mathbf{b} to \mathbf{b}' . Therefore, the uniform solution set is not optimal for hypervolume maximization on the Type II Pareto front.

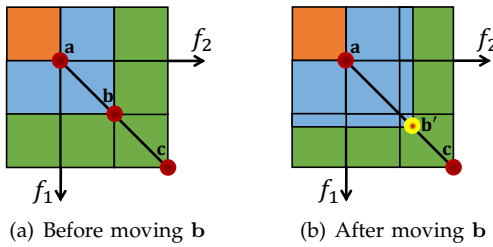


Fig. 5. An illustration of moving solution \mathbf{b} in Fig. 4 (b). Only f_1 - f_2 subspace is shown. The hypervolume contributions of \mathbf{b} and \mathbf{b}' are the blue areas in (a) and (b) times their f_3 values, respectively.

D. DAS Method and Reference Point Specification for Hypervolume Calculation

The DAS method [20] is a widely used method for generating uniformly distributed points on a simplex. DAS is adopted in many popular EMOAs (e.g., MOEA/D [24], NSGA-III [25]) for generating weight vectors or reference points.

DAS generates all points $\mathbf{w} = (w_1, w_2, \dots, w_m)$ satisfying the following relations:

$$\sum_{i=1}^m w_i = 1 \text{ and } w_i \in \left\{0, \frac{1}{H}, \frac{2}{H}, \dots, 1\right\} \text{ for } i = 1, 2, \dots, m, \quad (4)$$

where m is the number of objectives and H is a positive integer.

The total number of generated points is $\binom{H+m-1}{m-1}$. Fig. 6 (a) illustrates the points generated by DAS in three dimensions with $H = 3$ (i.e., $\mu = 10$). In the EMO community, a DAS solution set is implicitly assumed to be a perfectly uniform solution set on the unit simplex (i.e., a triangular plane Pareto front). Thus, a uniform solution set on a triangular plane Pareto front means a DAS solution set in this paper.

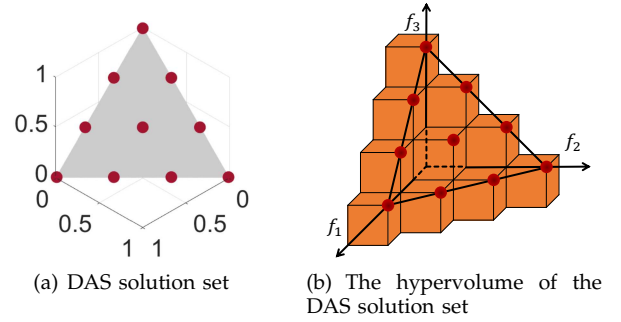


Fig. 6. A DAS solution set in three dimensions and its hypervolume.

Ishibuchi et al. [18] suggested a reference point specification method for hypervolume calculation. The general idea of the specification method is to set the reference point so that all solutions of the DAS solution set on the triangular linear Pareto front have the same hypervolume contribution.

Let us consider the linear Pareto front $f_1 + f_2 + \dots + f_m = 1$ and $f_1, f_2, \dots, f_m \geq 0$. Given a DAS solution set on the Pareto front, the reference point $\mathbf{r} = (r, \dots, r)$ is specified as follows [18]³:

$$r = -\frac{1}{H}, \quad (5)$$

where H is the parameter in (4).

Using this reference point specification, all the solutions have the equal hypervolume contribution as illustrated in Fig. 6 (b). Eq. (5) is a generalization of the reference point specification $r = -1/(\mu - 1)$ in two dimensions (see Fig. 3) since $H = \mu - 1$ when $m = 2$.

III. LINE-BASED PARETO FRONTS

In this section, we extend the Type I Pareto front, and investigate the hypervolume optimal μ -distribution on a Pareto front which consists of several joint Type I Pareto fronts.

We consider the following four types of Pareto fronts:

- 1) **Type III:** The two lines are the boundary of a triangular front. Fig. 7 (a) shows a Type III Pareto

³In the case of minimization, Eq. (5) is rewritten as $r = 1 + 1/H$.

front, where the triangular front is $f_1 + f_2 + f_3 = 1, f_1, f_2, f_3 \geq 0$.

- 2) **Type IV:** The two lines are the boundary of an inverted triangular front. Fig. 7 (b) shows a Type IV Pareto front where the inverted triangular front is $f_1 + f_2 + f_3 = 2, 0 \leq f_1, f_2, f_3 \leq 1$.
- 3) **Type V:** The three lines are the boundary of a triangular front. Fig. 7 (c) shows a Type V Pareto front where the triangular front is $f_1 + f_2 + f_3 = 1, f_1, f_2, f_3 \geq 0$.
- 4) **Type VI:** The three lines are the boundary of an inverted triangular front. Fig. 7 (d) shows a Type VI Pareto front where the inverted triangular front is $f_1 + f_2 + f_3 = 2, 0 \leq f_1, f_2, f_3 \leq 1$.

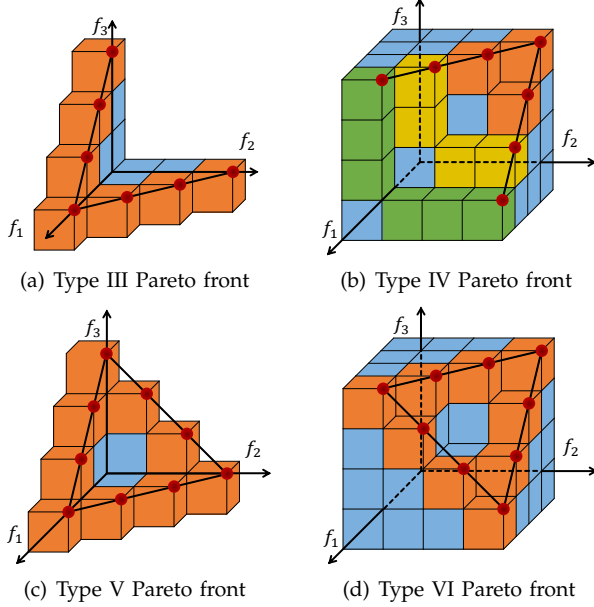


Fig. 7. Line-based Pareto fronts in three dimensions.

The four line-based Pareto fronts considered above are not overly realistic. They can be hardly found in real-world applications. However, as shown in [17], when the reference point is sufficiently far away, the solutions for hypervolume maximization are distributed on the boundary of some realistic Pareto fronts (e.g., the Pareto fronts of DTLZ1⁻¹ [26] and WFG3 [27]). In this case, it is meaningful to investigate the hypervolume optimal μ -distribution on the line-based Pareto fronts since the line-based Pareto fronts can be viewed as the boundary of some realistic Pareto fronts. This is the main motivation to investigate the line-based Pareto fronts.

Next, the hypervolume optimal μ -distribution on each of these four types of Pareto fronts is investigated.

A. Type III Pareto Front

Let us consider the Type III Pareto front with two lines where the first line is $f_1 + f_3 = 1, f_1, f_3 \geq 0$ and $f_2 = 0$, and the second line is $f_1 + f_2 = 1, f_1, f_2 \geq 0$ and $f_3 = 0$ as in Fig. 7 (a). Before investigating the hypervolume optimal μ -distribution on this type of Pareto front, we

show the following lemma which will be used in our analysis.

Lemma 1 (Reproduced from Theorem 3 in [23]). *Given $\mu \in \mathbb{N}_{\geq 2}$, and a linear Pareto front $f(x) = 1 - x$ where $x \in [0, 1]$, the unique optimal μ -distribution $(x_1^\mu, \dots, x_\mu^\mu)$ for the hypervolume indicator with reference point $\mathbf{r} = (r_1, r_2) \in \mathbb{R}_{<1}^2$ can be described by*

$$x_i^\mu = f^{-1}(F_l) + \frac{i}{\mu+1}(F_r - f^{-1}(F_l)) \quad (6)$$

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

$$\begin{aligned} F_l &= \min \left\{ 1 - r_1, \frac{\mu+1}{\mu} - \frac{1}{\mu}f(1 - r_2), \frac{\mu}{\mu-1} \right\}, \\ F_r &= \min \left\{ 1 - r_2, \frac{\mu+1}{\mu} - \frac{1}{\mu}f^{-1}(1 - r_1), \frac{\mu}{\mu-1} \right\}. \end{aligned} \quad (7)$$

Lemma 1 provides the following two main conclusions:

- 1) The optimal μ solutions are equispaced on the Pareto front, since $x_i^\mu - x_{i-1}^\mu$ is constant for all i .
- 2) The inclusion of the extreme points $(0, 1)$ and $(1, 0)$ in the optimal μ -distribution depends on the location of the reference point. If the reference point is $\mathbf{r} = (r, r)$, then $r \leq -1/(\mu - 1)$ guarantees the inclusion of the two extreme points. If the reference point is $\mathbf{r} = (r, 0)$, then $r \leq -1/\mu$ guarantees the inclusion of one extreme point $(0, 1)$.

Based on Lemma 1, we present the following theorem describing the hypervolume optimal μ -distribution on the Type III Pareto front.

Theorem 1. *For $\mu > 3$ solutions on the Type III Pareto front,*

- 1) *if μ is odd, the optimal μ -distribution is that the same number of solutions lie uniformly on each of the two lines where the three extreme points (i.e., $(1, 0, 0)$, $(0, 1, 0)$, $(0, 0, 1)$) are included, under the condition of $r \leq -2/(\mu - 1)$,*
- 2) *if μ is even, the optimal μ -distribution is that a different number of solutions lie uniformly on each of the two lines where the difference in the number of solutions on each line is one, and the three extreme points (i.e., $(1, 0, 0)$, $(0, 1, 0)$, $(0, 0, 1)$) are included, under the condition of $r \leq -1/\lfloor \frac{\mu-1}{2} \rfloor$.*

Proof. First, we assume that μ_1 solutions lie on the first line (from $(1, 0, 0)$ to $(0, 0, 1)$) and μ_2 solutions lie on the second line (from $(1, 0, 0)$ to $(0, 1, 0)$), where $\mu_1 + \mu_2 = \mu$ if $(1, 0, 0)$ is not included, and $\mu_1 + \mu_2 = \mu + 1$ if $(1, 0, 0)$ is included (since $(1, 0, 0)$ belongs to both lines). Denote the solutions on the first line from $(1, 0, 0)$ to $(0, 0, 1)$ as $\mathbf{a}_1, \dots, \mathbf{a}_{\mu_1}$, and the solutions on the second line from $(1, 0, 0)$ to $(0, 1, 0)$ as $\mathbf{b}_1, \dots, \mathbf{b}_{\mu_2}$. Without loss of generality, we assume that $f_1(\mathbf{a}_1) \geq f_1(\mathbf{b}_1)$. Then the whole hypervolume can be decomposed into two parts as illustrated in Fig. 8. That is, the whole hypervolume is sliced according to the plane $f_2 = 0$.

One part (denoted as HV_1) is the hypervolume determined by the solutions on the first line with the reference

point $\mathbf{r}_1 = (r, r, r)$. The other part (denoted as HV_2) is the hypervolume determined by the solutions on the second line with the reference point $\mathbf{r}_2 = (r, 0, r)$.

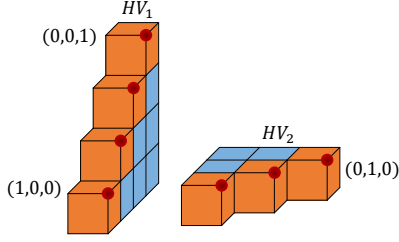


Fig. 8. An illustration of the decomposition of the hypervolume into two parts.

Then, the hypervolume can be calculated as $HV = HV_1 + HV_2$. Thus, maximizing HV is equivalent to maximizing $HV_1 + HV_2$. We can observe that HV_1 is the two-dimensional hypervolume in the f_1 - f_3 space (denoted as $HV_1^{f_1-f_3}$) multiplying $|r|$, and HV_2 is the two-dimensional hypervolume in the f_1 - f_2 space (denoted as $HV_2^{f_1-f_2}$) multiplying $|r|$. Therefore, maximizing $HV_1 + HV_2$ is equivalent to maximizing $HV_1^{f_1-f_3} + HV_2^{f_1-f_2}$.

Based on Lemma 1, we can get that the μ_1 solutions are equispaced on the first line and the μ_2 solutions are equispaced on the second line, in order to maximize $HV_1^{f_1-f_3}$ and $HV_2^{f_1-f_2}$ respectively. Note that the reference points for $HV_1^{f_1-f_3}$ and $HV_2^{f_1-f_2}$ are $\mathbf{r}_1^{f_1-f_3} = (r, r)$ and $\mathbf{r}_2^{f_1-f_2} = (r, 0)$, respectively. Based on Lemma 1, if $r \leq -1/(\mu_1 - 1)$, the two extreme points $(1, 0, 0)$ and $(0, 0, 1)$ of the first line are included in the μ_1 solutions. Since $(1, 0, 0)$ is included in this situation, we use $\mu_1 + \mu_2 = \mu + 1$ in the following analysis. If $r \leq -1/(\mu_2 - 1)$, the extreme point $(0, 1, 0)$ of the second line is included in the μ_2 solutions. Thus, if $r \leq \min\{-1/(\mu_1 - 1), -1/(\mu_2 - 1)\}$, the three extreme points are all included.

Next, based on the condition $r \leq \min\{-1/(\mu_1 - 1), -1/(\mu_2 - 1)\}$, we determine the values of μ_1 and μ_2 . First, $HV_1^{f_1-f_3}$ and $HV_2^{f_1-f_2}$ are calculated as illustrated in Fig. 9. $HV_1^{f_1-f_3}$ is the sum of the areas of the triangle A and the three rectangles B , C and D , minus the areas of the grey triangles in Fig. 9 (a). $HV_2^{f_1-f_2}$ is the sum of the areas of the triangle A and the rectangle B , minus the areas of the grey triangles in Fig. 9 (b). More specifically, they are calculated as follows:

$$\begin{aligned} HV_1^{f_1-f_3} &= \frac{1}{2} - r - r + r^2 - \frac{1}{2(\mu_1 - 1)^2} \times (\mu_1 - 1), \\ &= \frac{1}{2} + r(r - 2) - \frac{1}{2(\mu_1 - 1)}, \\ HV_2^{f_1-f_2} &= \frac{1}{2} - r - \frac{1}{2(\mu_2 - 1)^2} \times (\mu_2 - 1), \\ &= \frac{1}{2} - r - \frac{1}{2(\mu_2 - 1)}. \end{aligned} \quad (8)$$

For hypervolume maximization, we can formulate the

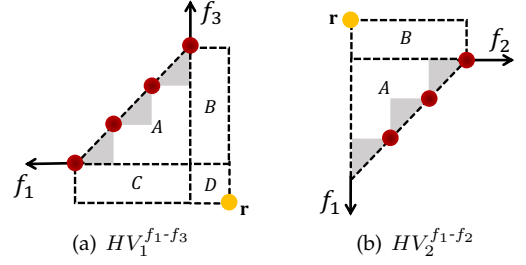


Fig. 9. An illustration of calculating $HV_1^{f_1-f_3}$ and $HV_2^{f_1-f_2}$.

following optimization problem:

$$\begin{aligned} \text{maximize: } & HV_1^{f_1-f_3} + HV_2^{f_1-f_2} \\ &= 1 - \frac{1}{2(\mu_1 - 1)} - \frac{1}{2(\mu_2 - 1)} + r(r - 3), \quad (9) \\ \text{subject to: } & \mu_1 + \mu_2 = \mu + 1 \text{ and } \mu_1, \mu_2 \in \mathbb{Z}_+. \end{aligned}$$

Let $\mu_2 = \mu + 1 - \mu_1$, optimization problem (9) can be rewritten as

$$\begin{aligned} \text{maximize: } & HV_1^{f_1-f_3} + HV_2^{f_1-f_2} \\ &= 1 - \frac{1}{2(\mu_1 - 1)} - \frac{1}{2(\mu - \mu_1)} + r(r - 3), \\ &= 1 - \frac{\mu - 1}{2(\mu_1 - 1)(\mu - \mu_1)} + r(r - 3), \\ \text{subject to: } & \mu_1 \in \mathbb{Z}_+. \end{aligned} \quad (10)$$

Thus, (10) is equivalent to maximizing $(\mu_1 - 1)(\mu - \mu_1)$ where $\mu_1 \in \mathbb{Z}_+$. By relaxing μ_1, μ_2 to real values, we can get the optimal solution of (10) as

$$\mu_1 = \mu_2 = \frac{\mu + 1}{2}. \quad (11)$$

If μ is odd, Eq. (11) is the optimal solution of (9). In this case, $\mu_1 = \mu_2$, which means that the same number of solutions lie on each of the two lines.

If μ is even, then the optimal solution of (9) is $\mu_1 = \lceil \frac{\mu+1}{2} \rceil$, $\mu_2 = \lfloor \frac{\mu+1}{2} \rfloor$, or $\mu_1 = \lfloor \frac{\mu+1}{2} \rfloor$, $\mu_2 = \lceil \frac{\mu+1}{2} \rceil$. In this case, $\mu_1 - \mu_2 = 1$ or $\mu_1 - \mu_2 = -1$, which means that one more or less solution lie on the first line than on the second line.

We have shown the condition for the reference point $r \leq \min\{-1/(\mu_1 - 1), -1/(\mu_2 - 1)\}$ in order to guarantee the inclusion of the three extreme points. Based on the above results, if μ is odd, $r \leq \min\{-1/(\mu_1 - 1), -1/(\mu_2 - 1)\} = -2/(\mu - 1)$. If μ is even, $r \leq \min\{-1/(\mu_1 - 1), -1/(\mu_2 - 1)\} = -1/\lfloor \frac{\mu-1}{2} \rfloor$. \square

B. Type IV Pareto Front

Let us consider the Type IV Pareto front with two lines where the first line is $f_1 + f_2 = 1, f_1, f_2 \geq 0$ and $f_3 = 1$, and the second line is $f_1 + f_3 = 1, f_1, f_3 \geq 0$ and $f_2 = 1$ as in Fig. 7 (b). Different from the analysis on the Type III Pareto front, we focus on investigating a uniform solution set on the Type IV Pareto front as illustrated in Fig. 7 (b), and analyze whether a uniform solution set is optimal for hypervolume maximization.

Here a uniform solution set means the same number of solutions are uniformly distributed on each of the two lines.

We have the following theorem for this type of Pareto front.

Theorem 2. For $\mu > 3$ solutions on the Type IV Pareto front, a uniform solution set is not optimal for hypervolume maximization.

Proof. First, we assume that there are μ' solutions uniformly distributed on each of the two lines, then $\mu' = (\mu + 1)/2$. Let $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_{\mu'}$ denote the solutions on the line from $(0, 1, 1)$ to $(1, 1, 0)$ where $\mathbf{a}_1 = (0, 1, 1)$ and $\mathbf{a}_{\mu'} = (1, 1, 0)$ are the two extreme solutions of this line. Fig. 7 (b) illustrates the hypervolume contributions of the uniformly distributed solutions on the Type IV Pareto front. We can see that different solutions have different hypervolume contributions, and each hypervolume contribution is a cuboid. For the hypervolume contribution of solution \mathbf{a}_i , the height of the cuboid is $\frac{1}{\mu'-1}$ and the basal area of the cuboid is $(i-1)(\frac{1}{\mu'-1})^2$. Thus, the hypervolume contribution of \mathbf{a}_i is

$$HVC(\mathbf{a}_i) = (i-1) \left(\frac{1}{\mu'-1} \right)^3, \quad i = 2, \dots, \mu'-1. \quad (12)$$

If we move \mathbf{a}_i towards \mathbf{a}_{i+1} for any $i \in \{2, \dots, \mu'-1\}$ (i.e., only one solution is moved and all the others are fixed as illustrated in Fig. 10), a new solution $\mathbf{a}'_i = \mathbf{a}_i + \alpha(\mathbf{a}_{i+1} - \mathbf{a}_i)$ where $\alpha \in [0, 1]$ is obtained. We can see that the hypervolume contribution of \mathbf{a}'_i is composed by two cuboids A and B . The height of the two cuboids is $(1-\alpha)(\frac{1}{\mu'-1})$. The basal area of A is $\alpha(\frac{1}{\mu'-1})^2$, and the basal area of B is $(1+\alpha)(\frac{1}{\mu'-1}) \times (i-1)(\frac{1}{\mu'-1}) = (1+\alpha)(i-1)(\frac{1}{\mu'-1})^2$. Thus, the hypervolume contribution of \mathbf{a}'_i is

$$\begin{aligned} HVC(\mathbf{a}'_i) &= \left(\alpha \left(\frac{1}{\mu'-1} \right)^2 + (1+\alpha)(i-1) \left(\frac{1}{\mu'-1} \right)^2 \right) \\ &\quad \times (1-\alpha) \left(\frac{1}{\mu'-1} \right) \\ &= [(i-1)(1-\alpha)(1+\alpha) + \alpha(1-\alpha)] \left(\frac{1}{\mu'-1} \right)^3. \end{aligned} \quad (13)$$

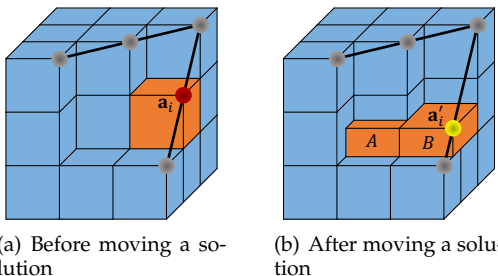


Fig. 10. An illustration of moving a solution.

The hypervolume contribution difference between \mathbf{a}'_i and \mathbf{a}_i is

$$HVC(\mathbf{a}'_i) - HVC(\mathbf{a}_i) = (-i\alpha^2 + \alpha) \left(\frac{1}{\mu'-1} \right)^3. \quad (14)$$

It is easy to obtain that if $0 < \alpha < 1/i$, $HVC(\mathbf{a}'_i) - HVC(\mathbf{a}_i) > 0$. This indicates that \mathbf{a}'_i has a larger hypervolume contribution than \mathbf{a}_i . Thus, we can improve the overall hypervolume by moving \mathbf{a}_i to \mathbf{a}'_i . Therefore, the original uniform solution set is not optimal for hypervolume maximization. \square

C. Type V Pareto Front

Let us consider the Type V Pareto front with a triangular shape specified by $f_1 + f_2 + f_3 = 1, f_1, f_2, f_3 \geq 0$ as shown in Fig. 7 (c).

We present the following theorem describing the hypervolume optimal μ -distribution on the Type V Pareto front.

Theorem 3. For $\mu > 3$ solutions on the Type V Pareto front,

- 1) if $(\mu \bmod 3) = 0$, the optimal μ -distribution is that the same number of solutions lie uniformly on each of the three lines where the three extreme points (i.e., $(1,0,0)$, $(0,1,0)$, $(0,0,1)$) are included, under the condition of $r \leq -3/\mu$,
- 2) if $(\mu \bmod 3) \neq 0$, the optimal μ -distribution is that a different number of solutions lie uniformly on each of the three lines where two lines have the same number of solutions and the other line has one more/less solution, and the three extreme points (i.e., $(1,0,0)$, $(0,1,0)$, $(0,0,1)$) are included, under the condition of $r \leq -1/\lfloor \frac{\mu}{3} \rfloor$.

Remark 1. The proof of Theorem 3 is similar to that of Theorem 1. We provide the detailed proof of Theorem 3 in Section I of the supplementary material.

D. Type VI Pareto Front

Let us consider the Type VI Pareto front with an inverted triangular shape specified by $f_1 + f_2 + f_3 = 2, 0 \leq f_1, f_2, f_3 \leq 1$ as shown in Fig. 7 (d). Similar to the analysis on the Type IV Pareto front, we investigate whether a uniform solution set is optimal for hypervolume maximization. Here a uniform solution set means the same number of solutions are uniformly distributed on each of the three lines.

We have the following theorem for this type of Pareto front.

Theorem 4. For $\mu > 6$ solutions on the Type VI Pareto front, a uniform solution set is not optimal for hypervolume maximization.

Remark 2. The proof of Theorem 4 is similar to that of Theorem 2. That is, the hypervolume of a uniform solution set can be improved by moving one solution. We provide the detailed proof of Theorem 4 in Section II of the supplementary material.

Remark 3. Here we need to emphasize that Theorem 4 only holds for $\mu > 6$. For the case of $\mu \leq 6$ (i.e., at most three solutions lie on each line), we cannot move one solution to get a better overall hypervolume. For example, if there are three uniformly distributed solutions on each line (i.e., one midpoint and two extreme points), the overall hypervolume will be decreased (instead of increased) by moving any of the three midpoints. As will be shown in Section IV, when $\mu = 3$ and 6 the uniform solution set on the Type VI Pareto front is optimal for hypervolume maximization.

E. Further Empirical Investigations

In the previous subsections, we have theoretically investigated the hypervolume optimal μ -distribution on four types of line-based Pareto fronts. Through theoretical analysis, we showed that the uniform solution set is optimal for hypervolume maximization on the Type III and Type V Pareto fronts, whereas it is not optimal for hypervolume maximization on the Type IV and Type VI Pareto fronts. In this subsection, we further empirically investigate the hypervolume optimal μ -distribution on the line-based Pareto fronts.

We use SMS-EMOA to search for the hypervolume optimal μ -distribution on the four types of line-based Pareto fronts. In order to find the optimal (or near-optimal) solution distribution, we perform SMS-EMOA 100 times with 10,000 generations and choose the best run (i.e., the solution set with the best hypervolume value) as the obtained hypervolume optimal μ -distribution. The reference point is specified as $\mathbf{r} = (-0.2, -0.2, -0.2)$. For each Pareto front, we specify the number of solutions (i.e., population size) as $\mu = 21$ for the Types III and IV Pareto fronts, and $\mu = 30$ for the Types V and VI Pareto fronts. The obtained optimal μ -distributions on the line-based Pareto fronts are shown in Fig. 1 (a), (c), (e) and (g) (i.e., the left column in Fig. 1). We can see that the empirical results are consistent with our theoretical results.

Fig. 1 (c) and (g) show the obtained optimal μ -distributions on the Types IV and VI Pareto fronts, respectively. For the Type IV Pareto front, we can see that solutions are sparsely distributed around the joint extreme point of the two lines. From the joint extreme point to the other two extreme points, solutions become denser and denser. For the Type VI Pareto front, we can see that solutions are sparsely distributed around the three extreme points. More solutions are distributed in the middle of each line.

To further analyze the optimal μ -distributions in Fig. 1, we show the hypervolume contribution information (i.e., the relative size of the contribution of each solution) and the shape of the hypervolume for each solution set in Fig. 11. In the left column in Fig. 11, the size of each circle is proportional to the hypervolume contribution of the corresponding solution. We can observe that for the Types III and V Pareto fronts (Fig. 11 (a) and

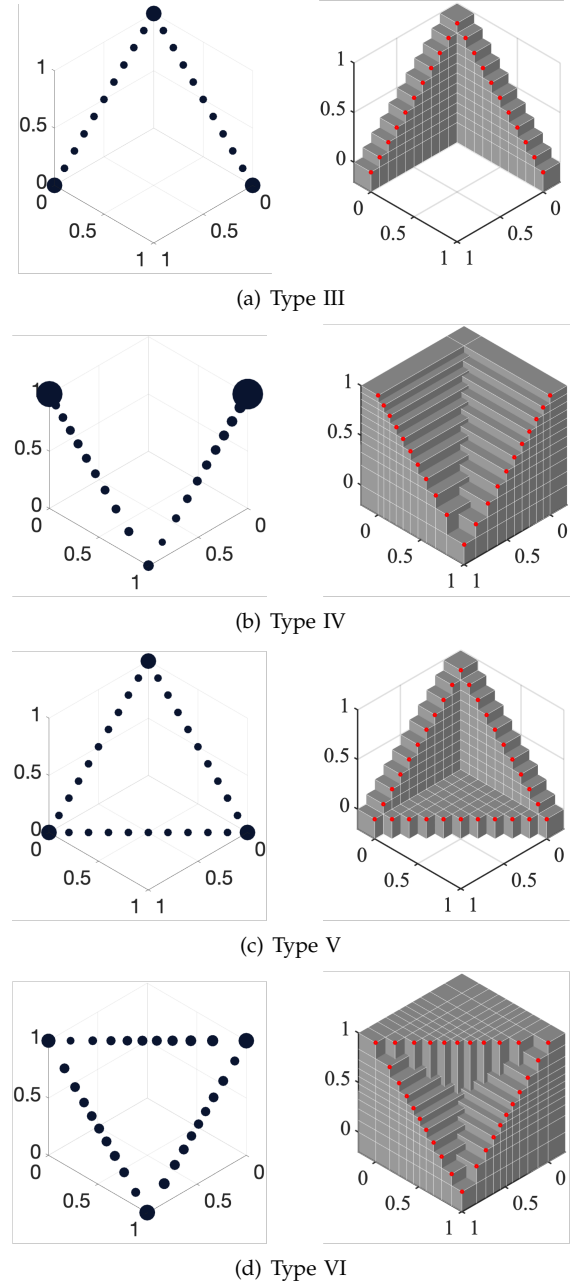


Fig. 11. The best solution sets obtained by SMS-EMOA. The reference point is specified as $\mathbf{r} = (-0.2, -0.2, -0.2)$. The left column shows the hypervolume contribution information where the size of the circle represents the relative hypervolume contribution of each solution in each solution set. The right column shows the shape of the hypervolume for each solution set.

(c)), the three extreme points have equal hypervolume contributions, and all the other points also have equal hypervolume contributions. The location of the reference point influences the hypervolume contributions of all points in Fig. 11 (a) and (c). For the Types IV and VI Pareto fronts (Fig. 11 (b) and (d)), each point tends to have a different hypervolume contribution. The location of the reference point only influences the hypervolume contribution of the three extreme points.

The right column in Fig. 11 shows the shape of the

hypervolume for each solution set. These figures intuitively show the shape of the hypervolume contribution for each solution. We can observe that for the Types III and V Pareto fronts (Fig. 11 (a) and (c)), each solution's hypervolume contribution is a regular cuboid. However, for the Types IV and VI Pareto fronts (Fig. 11 (b) and (d)), most solutions' hypervolume contribution is irregular. These results can help us to more intuitively and deeply understand the optimal distribution of solutions on the line-based Pareto fronts.

F. Discussions

1) *The difference between maximization and minimization:* We need to note that maximization of each objective is assumed in this paper. All the conclusions are based on this assumption. If the minimization case is considered, then the conclusions for the Type III/V and the Type IV/VI Pareto fronts should be exchanged. That is, in the minimization case, the uniform solution set is optimal for hypervolume maximization on the Type IV and Type VI Pareto fronts whereas it is not optimal on the Type III and Type V Pareto fronts.

2) *What can we learn from line-based Pareto fronts:* From Fig. 7 we can see that the line-based Pareto fronts are constructed by the Type I Pareto front. As already discussed in Section II-C, the hypervolume optimal μ -distribution is uniform on the Type I Pareto front. However, based on our results in this section, we can see that the hypervolume optimal μ -distribution is not always uniform on the line-based Pareto fronts constructed by the Type I Pareto front. This means that if we simply use the hypervolume indicator to compare two solution sets as shown in Fig. 1, the non-uniform solution set can be evaluated better than the uniform solution set, which is counter-intuitive. Therefore, the behavior of the hypervolume indicator is still not well understood in the EMO community.

For the Type III and V Pareto fronts, the hypervolume can be decomposed into different parts where each part is determined only by the solutions on each line (e.g., Fig. 8 illustrates such a decomposition). However, for the Type IV and VI Pareto fronts, the hypervolume cannot be decomposed in such a way. This observation may lead to further investigation of the hypervolume indicator.

3) *Construction of Type III-VI Pareto fronts:* We can also construct test problems with Type III-VI Pareto fronts. For example, let us consider a multi-line distance minimization problem [28] with constraints as shown in the left figure of Fig. 12, where the objectives are to minimize the distance to each of the three lines: AB , AC and BC . In the left figure of Fig. 12, the inside of the triangle ABC and the line BC are infeasible. Then the Pareto set of this problem is line AB and line AC . The Pareto front (i.e., the projection of these two lines into the objective space) is Type III. If we consider to minimize the minus distance to each line (which is equivalent to maximize the distance to each line) as shown in the right

figure of Fig. 12, given the infeasible region defined, the Pareto set is the union of the two red lines. In this case, the Pareto front is Type IV (the objective is the minus distance to each line).

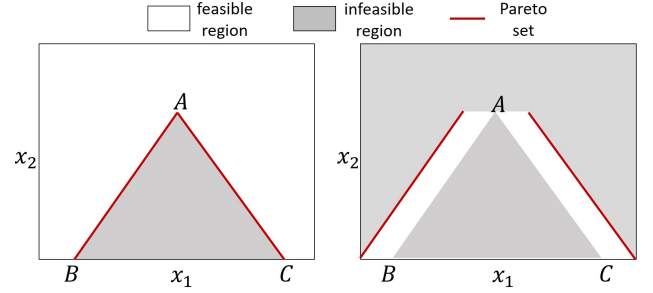


Fig. 12. An illustration of the multi-line distance minimization problem with constraints. (a) The inside of the triangle ABC and the line BC are infeasible. The union of the line AB and the line AC is the Pareto set. (b) The objective is the minus distance to each line. The union of the two line segments is the Pareto set.

Similarly, we can use the multi-line distance minimization problem to construct a Type V or Type VI Pareto front. The details are provided in Section IV of the supplementary material.

IV. PLANE-BASED PARETO FRONTS

In the last section, we only considered the line-based Pareto fronts in three dimensions and showed that the hypervolume optimal μ -distribution is not always uniform on these Pareto fronts. In this section, we investigate plane-based Pareto fronts, which are more commonly seen in multi-objective optimization test problems (e.g., three-objective DTLZ1 [29] and DTLZ1⁻¹ [26]).

The following two types of plane-based Pareto fronts were considered by Ishibuchi et al. [17]:

- 1) **Type VII:** The linear triangular Pareto front specified by $f_1 + f_2 + f_3 = 1, f_1, f_2, f_3 \geq 0$ (see Fig. 13 (a)).
- 2) **Type VIII:** The linear inverted triangular Pareto front specified by $f_1 + f_2 + f_3 = 2, 0 \leq f_1, f_2, f_3 \leq 1$ (see Fig. 13 (b)).

Next, the hypervolume optimal μ -distribution on each of these two types of Pareto fronts is investigated.

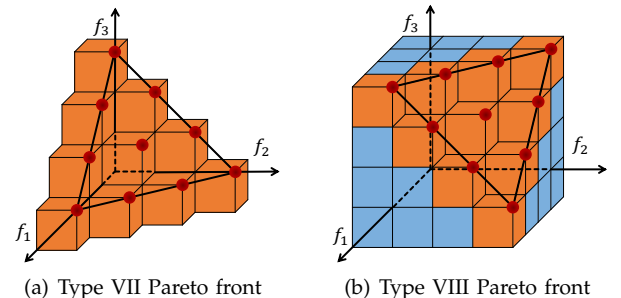


Fig. 13. Plane-based Pareto fronts in three dimensions.

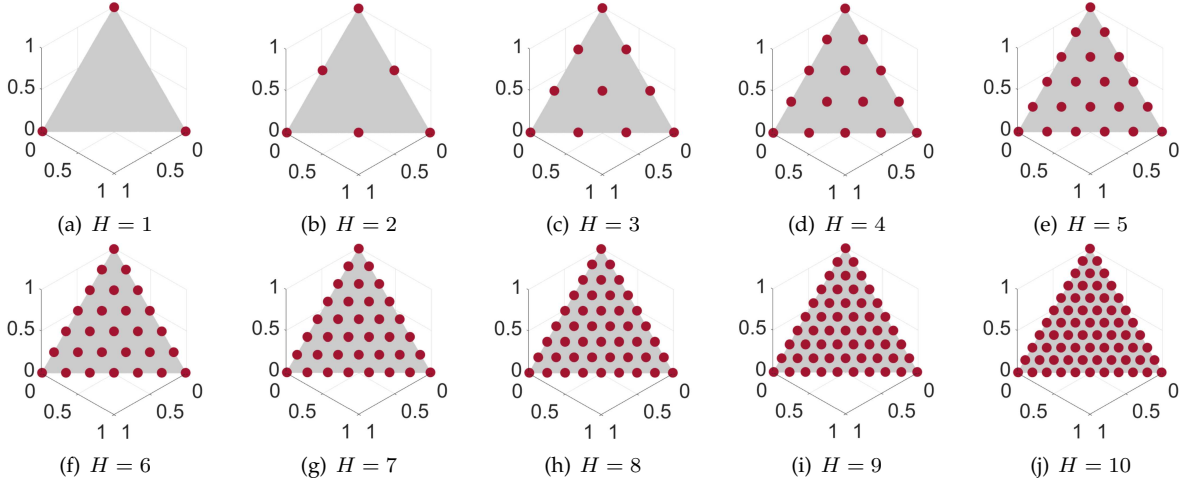


Fig. 14. The DAS solution sets generated under different specifications of H (see Eq. (4)).

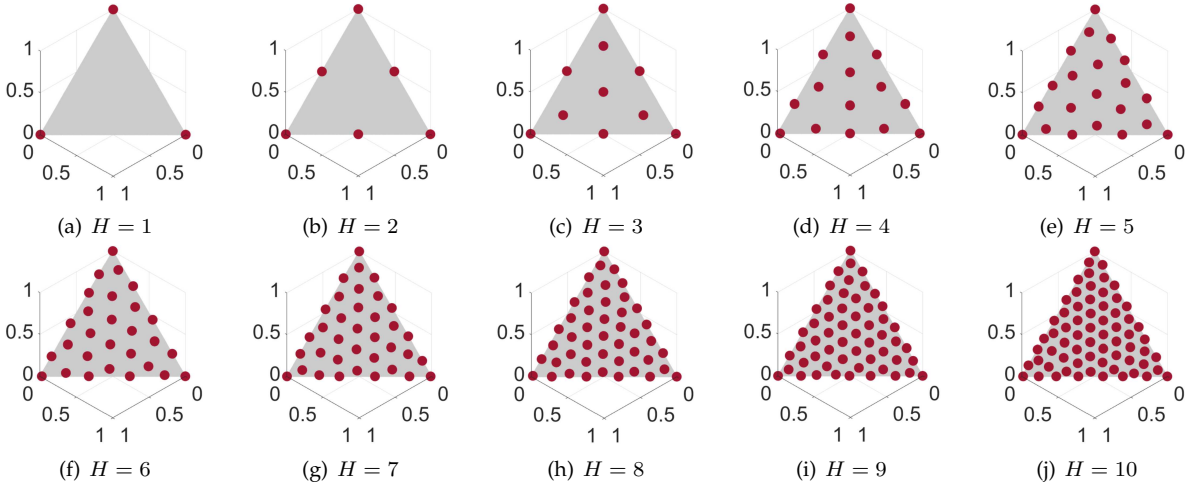


Fig. 15. The best solution sets obtained by SMS-EMOA. The number of solutions in each solution set is the same as the number of solutions in the DAS solution set under the corresponding H .

A. Type VII Pareto Front

For the DAS solution set on the Type VII Pareto front, if the reference point is specified as $r = -1/H$, all the solutions have the same hypervolume contribution as shown in Fig. 13 (a). As already shown in the previous sections, a uniformly distributed solution set is optimal for hypervolume maximization when each solution has the same hypervolume contribution (e.g., Types I, III and V Pareto fronts). Thus, intuitively we may think that the DAS solution set is also optimal for hypervolume maximization.

In order to examine whether this intuition is correct or not, we perform SMS-EMOA on the Type VII Pareto front to search for the optimal μ -distribution. Here we use SMS-EMOA to do the search since it is proved by Beume et al. [30] that SMS-EMOA is able to find the optimal μ -distribution on a two-objective linear Pareto front. Whereas there is no theoretical guarantee on the three-objective linear Pareto front, SMS-EMOA is powerful for searching for the optimal μ -distribution in the three-

objective case. We set $H = 1, 2, \dots, 10$ and the reference point $r = -1/H$, and use SMS-EMOA to search for the optimal μ -distribution on the Type VII Pareto front. In order to make sure that the obtained solution set is optimal (or very close to optimal), we use a large number of generations 10,000 and a large number of independent runs 100 for each H . The best solution set among 100 runs is used as the best optimal μ -distribution for each case.

Table I shows the hypervolume of the best solution sets obtained by SMS-EMOA, and the DAS solution sets. From Table I, we can observe that for each H , the hypervolume of the DAS solution set is smaller than or equal to the hypervolume of the best solution set obtained by SMS-EMOA. Specifically, the DAS solution set has the same hypervolume value as the best solution set obtained by SMS-EMOA when $H = 1, 2$. Moreover, all the 100 solution sets obtained by SMS-EMOA for each $H = 1, 2$ have the same hypervolume value. This means that the DAS solution set is optimal for hypervolume maximization when $H = 1, 2$, which is consistent with

our intuition. However, it is not optimal for hypervolume maximization when $H = 3, \dots, 10$, which is contrary to our intuition. Thus, the intuition mentioned above is not always correct.

TABLE I

HYPERVOLUME OF THE SOLUTION SETS GENERATED BY DAS AND OBTAINED BY SMS-EMOA. THE REFERENCE POINT IS SPECIFIED AS $r = -1/H$. THE BETTER HYPERVOLUME VALUE IS HIGHLIGHTED IN BOLD FOR EACH H .

| H (μ) | DAS | SMS-EMOA |
|---------------|---------------|---------------|
| 1 (3) | 4.0000 | 4.0000 |
| 2 (6) | 1.2500 | 1.2500 |
| 3 (10) | 0.7407 | 0.7422 |
| 4 (15) | 0.5469 | 0.5483 |
| 5 (21) | 0.4480 | 0.4497 |
| 6 (28) | 0.3889 | 0.3905 |
| 7 (36) | 0.3499 | 0.3515 |
| 8 (45) | 0.3223 | 0.3236 |
| 9 (55) | 0.3018 | 0.3031 |
| 10 (66) | 0.2860 | 0.2872 |

Figs. 14-15 show the DAS solution sets and the best solution sets obtained by SMS-EMOA, respectively. We can see that when $H = 1, 2$, the DAS solution sets are the same as the best solution sets obtained by SMS-EMOA. However, when $H = 3, \dots, 10$, they are totally different. The solution sets obtained by SMS-EMOA are not as uniform as the DAS solution sets for $H = 3, \dots, 10$.

In the above experiments, we use the reference point suggested in [18] for hypervolume comparison (i.e., $r = -1/H$). In the case of the Type VII Pareto front, the optimal μ -distribution strongly depends on the reference point specification [18]⁴. When the reference point is far away from the Pareto front (e.g., $r = -100$), all solutions in the optimal μ -distribution are on the sides of the Pareto front. When the nadir point is used as the reference point (i.e., $r = 0$), no solutions in the optimal μ -distribution are on the sides of the Pareto front.

Now we want to know whether we can improve the DAS solution set for $H = 3, \dots, 10$ using SMS-EMOA. Through experiments, we found that the DAS solution set cannot be further improved by SMS-EMOA. This interesting observation motivates us to examine the property of the DAS solution set. The following theorem shows that the DAS solution set is locally optimal for hypervolume maximization with respect to the $(\mu + 1)$ selection scheme, i.e., the DAS solution set cannot be further improved by replacing one solution.

Theorem 5. *Given the reference point $r = -1/H$, the DAS solution set is locally optimal for hypervolume maximization with respect to a $(\mu + 1)$ selection scheme.*

Proof. We will prove this theorem by adding one arbitrary solution on the Pareto front and showing that this

solution has the least hypervolume contribution, which implies that this solution is removed and the original set cannot be improved.

Firstly, the Pareto front is divided into different triangular regions as illustrated in Fig. 16. We can see that there are two types of regions: the triangular regions and the inverted triangular regions. Next, the two types of regions are considered separately.

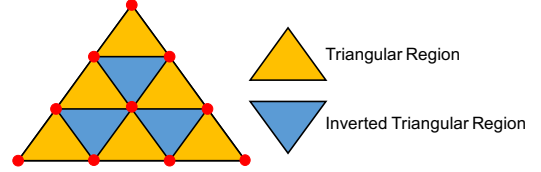


Fig. 16. The Pareto front is divided into triangular and inverted triangular regions.

If a solution \mathbf{p} is added in an inverted triangular region as illustrated in Fig. 17 (a), the hypervolume contributions of three solutions $\mathbf{a}, \mathbf{b}, \mathbf{c}$ (i.e., the three vertices of the triangle) are influenced by \mathbf{p} .

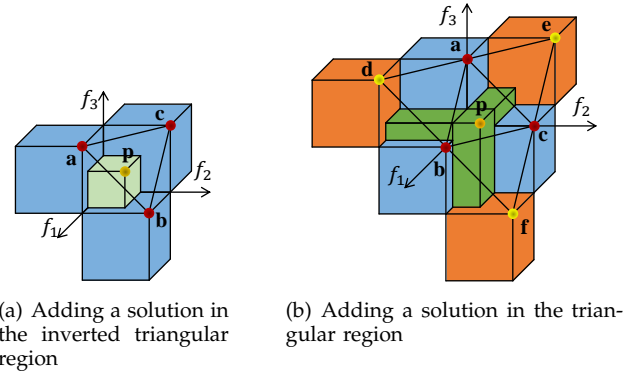


Fig. 17. Adding a solution on the Pareto front.

Suppose $\mathbf{p} = (x, y, z)$, $\mathbf{a} = (1, 0, 1)$, $\mathbf{b} = (0, 1, 1)$ and $\mathbf{c} = (1, 1, 0)$, then the hypervolume contribution of each of these solutions can be calculated as follows:

$$\begin{aligned} HVC(\mathbf{p}) &= xyz, \\ HVC(\mathbf{a}) &= 1 - xz, \quad HVC(\mathbf{b}) = 1 - yz, \quad HVC(\mathbf{c}) = 1 - xy. \end{aligned} \quad (15)$$

Given the conditions $x + y + z = 2$ and $0 \leq x, y, z \leq 1$, we prove that $HVC(\mathbf{p}) < HVC(\mathbf{c})$ as follows:

$$\begin{aligned} HVC(\mathbf{p}) &= xyz = xy(2 - x - y) = xy(1 - x) + xy(1 - y) \\ &< y(1 - x) + x(1 - y) = (x + y - xy) - xy. \end{aligned} \quad (16)$$

Let $f(x, y) = x + y - xy$, we have $f(0, 0) = 0$ and $f(1, 1) = 1$. Since $\nabla f(x, y) = (1 - y, 1 - x) > 0$, we have $f(x, y) < f(1, 1) = 1$. Then, based on Eq. (16), we have $HVC(\mathbf{p}) < 1 - xy = HVC(\mathbf{c})$. Similarly, we can also prove that $HVC(\mathbf{p}) < HVC(\mathbf{a})$ and $HVC(\mathbf{p}) < HVC(\mathbf{b})$. Thus, \mathbf{p} is the least hypervolume contributor and \mathbf{p} is removed.

If a solution \mathbf{p} is added in a triangular region as illustrated in Fig. 17 (b), the hypervolume contributions of six solutions $\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d}, \mathbf{e}, \mathbf{f}$ are influenced by \mathbf{p} .

⁴This conclusion holds for the inverted triangular Pareto front in [18] since the minimization case is considered in [18]. However, this conclusion holds for the triangular Pareto front in this paper since we consider the maximization case.

Suppose $\mathbf{p} = (x, y, z)$, $\mathbf{a} = (0, 0, 1)$, $\mathbf{b} = (1, 0, 0)$, $\mathbf{c} = (0, 1, 0)$, $\mathbf{d} = (1, -1, 1)$, $\mathbf{e} = (-1, 1, 1)$ and $\mathbf{f} = (1, 1, -1)$, then the hypervolume contribution of each of these solutions can be calculated as follows:

$$\begin{aligned} HVC(\mathbf{p}) &= xyz + xy + xz + yz, \\ HVC(\mathbf{a}) &= 1 - z, \quad HVC(\mathbf{b}) = 1 - x, \quad HVC(\mathbf{c}) = 1 - y, \\ HVC(\mathbf{d}) &= 1 - xz, \quad HVC(\mathbf{e}) = 1 - yz, \quad HVC(\mathbf{f}) = 1 - xy. \end{aligned} \quad (17)$$

Given the conditions $x + y + z = 1$ and $0 \leq x, y, z \leq 1$, we prove that $HVC(\mathbf{p}) < HVC(\mathbf{a})$ as follows:

$$\begin{aligned} HVC(\mathbf{p}) &= xyz + xy + xz + yz \\ &= xy(1 - x - y) + xy \\ &\quad + x(1 - x - y) + y(1 - x - y) \quad (18) \\ &= x + y - x^2(1 + y) - y^2(1 + x) \\ &< x + y = 1 - z = HVC(\mathbf{a}). \end{aligned}$$

Similarly, we can prove that $HVC(\mathbf{p}) < HVC(\mathbf{b})$ and $HVC(\mathbf{p}) < HVC(\mathbf{c})$. It is easy to observe that $HVC(\mathbf{a}) < HVC(\mathbf{d})$ and $HVC(\mathbf{a}) < HVC(\mathbf{e})$. Similarly, $HVC(\mathbf{b}) < HVC(\mathbf{d})$ and $HVC(\mathbf{b}) < HVC(\mathbf{f})$, $HVC(\mathbf{c}) < HVC(\mathbf{e})$ and $HVC(\mathbf{c}) < HVC(\mathbf{f})$. Based on these relations, we can conclude that \mathbf{p} is the least hypervolume contributor and \mathbf{p} is removed. \square

Remark 4. In Theorem 5, the reference point is specified as $r = -1/H$. This is the suggested reference point specification for the hypervolume indicator as discussed in Section II-D. If the reference point is specified as $r < -1/H$, Theorem 5 may not hold anymore since the solutions on the boundary of the Pareto front have larger hypervolume contributions than the inner solutions. In this case, replacing an inner solution by a boundary solution can improve the hypervolume of the whole solution set. When $r > -1/H$, Theorem 5 does not hold since the reference point is too close to the Pareto front.

B. Type VIII Pareto Front

If we invert the DAS solution set, we can obtain a uniformly distributed solution set on the inverted triangular Pareto front as shown in Fig. 13 (b). We call this uniform solution set as the inverted DAS solution set. If the reference point is specified as $r = -1/H$, all the solutions have the same hypervolume contribution as shown in Fig. 13 (b). Similarly, we use SMS-EMOA to search for the optimal μ -distribution on the Type VIII Pareto front. All the experimental settings are the same as in the previous subsection. Table II shows the hypervolume of the inverted DAS solution sets and the best solution sets obtained by SMS-EMOA. From Table II, we can observe that for each H , the hypervolume of the inverted DAS solution set is smaller than or equal to the hypervolume of the best solution set obtained by SMS-EMOA, which means that the inverted DAS solution set is not always optimal for hypervolume maximization. In Section V of the supplementary material, we visually show the inverted DAS solution sets and the best solution sets

obtained by SMS-EMOA. Similar observations can be obtained to the case of the Type VII Pareto front.

TABLE II
HYPERVOLUME OF THE INVERTED DAS SOLUTION SETS AND THE BEST SOLUTION SETS OBTAINED BY SMS-EMOA. THE REFERENCE POINT IS SPECIFIED AS $r = -1/H$. THE BETTER HYPERVOLUME VALUE IS HIGHLIGHTED IN BOLD FOR EACH H .

| H (μ) | Inverted DAS | SMS-EMOA |
|---------------|---------------|---------------|
| 1 (3) | 7.0000 | 7.0000 |
| 2 (6) | 2.8750 | 2.8750 |
| 3 (10) | 2.0000 | 2.0019 |
| 4 (15) | 1.6406 | 1.6421 |
| 5 (21) | 1.4480 | 1.4496 |
| 6 (28) | 1.3287 | 1.3303 |
| 7 (36) | 1.2478 | 1.2493 |
| 8 (45) | 1.1895 | 1.1908 |
| 9 (55) | 1.1454 | 1.1466 |
| 10 (66) | 1.1110 | 1.1122 |

We use the suggested reference point specification in [18] for hypervolume comparison (i.e., $r = -1/H$). In the case of the Type VIII Pareto front, the optimal μ -distribution does not change when $r \leq -1/H$ [18]. That is, the same distribution is always optimal for any specification of the reference point $\mathbf{r} = (r, r, r)$ satisfying $r \leq -1/H$. This is totally different from the case of the Type VII Pareto front.

In a similar manner to the case of the triangular Pareto front, we have the following theorem which shows that the inverted DAS solution set is locally optimal for hypervolume maximization with respect to the $(\mu + 1)$ selection scheme.

Theorem 6. Given the reference point $r = -1/H$, the inverted DAS solution set is locally optimal for hypervolume maximization with respect to the $(\mu + 1)$ selection scheme.

Proof. Follow the proof way of Theorem 5, we can easily get this conclusion. \square

Remark 5. In Theorem 6, the reference point is also specified as $r = -1/H$. Different from Theorem 5, Theorem 6 still holds for $r < -1/H$. This is because the change of the reference point only influences the hypervolume contribution of the three extreme points of the Type VIII Pareto front. This is very different from the Type VII Pareto front. When $r > -1/H$, Theorem 6 does not hold since the reference point is too close to the Pareto front.

C. Further Empirical Investigations

In the previous subsections, we have investigated the hypervolume optimal μ -distributions on the plane-based Pareto fronts, both empirically and theoretically. We have shown that each solution of the DAS solution set (or the inverted DAS solution set) has the same hypervolume contribution when the reference point is specified as $r = -1/H$. The counter-intuitive fact is that the whole solution set is not optimal for hypervolume

maximization when $H = 3, \dots, 10$. Thus, it is interesting to further investigate the hypervolume contribution of each solution for the optimal solution set on the plane-based Pareto fronts. In this subsection, we investigate this issue empirically.

We take the Type VII Pareto front and $H = 3, 4$ for illustration. The left column in Fig. 18 shows the hypervolume contribution information of each solution for the DAS solution set and the best solution set obtained by SMS-EMOA. We can observe that each solution has the same hypervolume contribution in the DAS solution sets ((a) and (c)) whereas each solution has a different hypervolume contribution in the best solution set obtained by SMS-EMOA ((b) and (d)). However, in the best solution sets ((b) and (d)) obtained by SMS-EMOA, an inner solution (a solution inside the Pareto front) tends to have a smaller hypervolume contribution than an outer solution (a solution on the boundary of the Pareto front).

The right column in Fig. 18 intuitively shows the shape of the hypervolume for different solution sets. We can more intuitively understand the hypervolume contribution of each solution. For the DAS solution sets ((a) and (c)), we can see that each solution's hypervolume contribution is a regular cuboid. However, for the best solution sets obtained by SMS-EMOA ((b) and (d)), most solutions' hypervolume contribution is irregular. These observations may motivate the EMO researchers to further investigate the hypervolume optimal μ -distribution in three dimensions.

D. Discussions

1) *The uniqueness of the hypervolume optimal μ -distribution:* It is clear that a DAS solution set is unique on the Pareto front. That is, it is impossible to have two different DAS solution sets with the same number of solutions. However, it is not always the case for a solution set maximizing the hypervolume indicator. As shown in Fig. 15, it is clear that the solution sets with $H = 1, 2, 3$ are unique. However, for the solution set with $H = 4$, we can rotate the solution set $\pm 120^\circ$ to get a different solution set. Thus, the hypervolume optimal μ -distribution does not necessarily need to be unique.

2) *The solution distribution for hypervolume maximization:* Although the theoretical hypervolume optimal μ -distributions on the plane-based Pareto fronts are not derived in this paper, we can use the empirical results in Fig. 15 to get a sense of how the solutions are distributed on the Pareto fronts. From Fig. 15 (a)–(d) we can observe that the solutions are symmetrically distributed. In our previous work [31], we empirically investigated the hypervolume optimal μ -distributions on the plane-based Pareto fronts for $\mu = 2, \dots, 10$. We have similar observations, i.e., in most cases the solutions are symmetrically distributed. Thus, this may be the law of the hypervolume optimal μ -distribution on the plane-based Pareto fronts. However, proving this is a

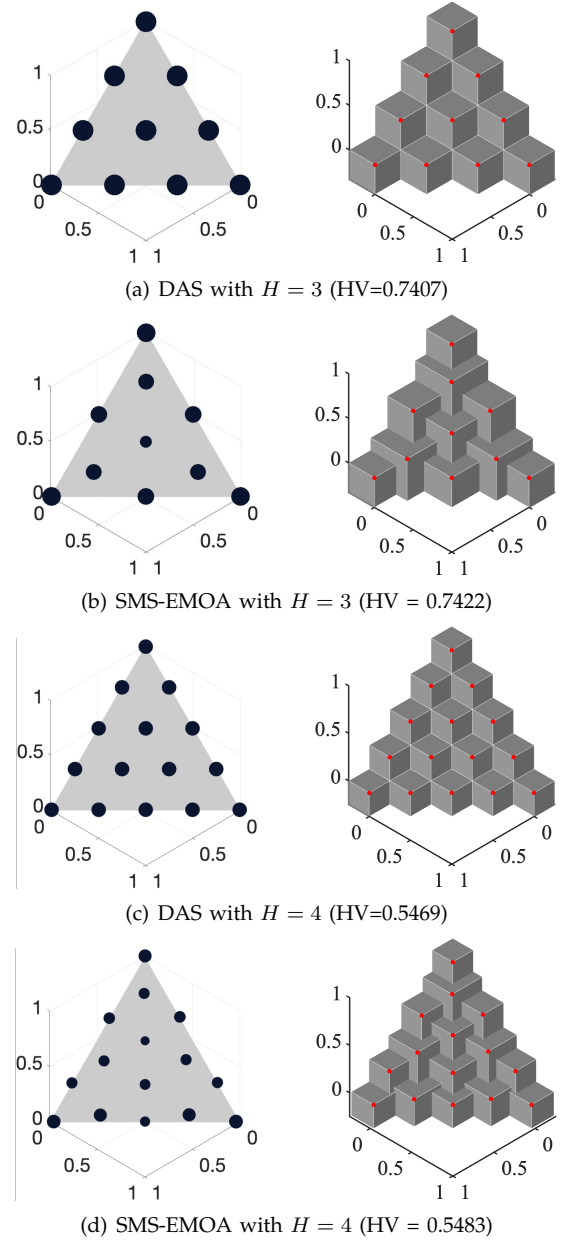


Fig. 18. Illustrations of the hypervolume contribution of each solution in (a) and (c) the DAS solution set, (b) and (d) the best solution set obtained by SMS-EMOA. The size of the circle represents the relative hypervolume contribution of each solution. The reference point is specified as $\mathbf{r} = (-1/H, -1/H, -1/H)$ where $H = 3$ in (a) and (b), and $H = 4$ in (c) and (d).

challenging task. We leave it as an open question for future research.

3) *Inspirations to EMO algorithm design:* In the EMO literature, most decomposition-based EMOs use the DAS method to generate weight vectors or reference points. The discussions in this section suggest the necessity of reconsidering the generation methods for weight vectors or reference points, in order to obtain better hypervolume results for these decomposition-based EMOs. For the hypervolume-based EMOs (e.g., SMS-EMOA), the results in this section reveal that a $(\mu + 1)$ selection scheme may fall into local optimum. Thus, a mechanism

to jump out of local optimum is needed in this type of algorithms.

V. CONCLUSIONS

In this paper, we investigated the hypervolume optimal μ -distributions on line- and plane-based Pareto fronts in three dimensions. First, we investigated the optimal μ -distributions on the line-based Pareto fronts with two and three lines. We showed that the solution set is not always uniform on these Pareto fronts for hypervolume maximization. The solutions are only uniform on the Type III and Type V Pareto fronts, whereas the solutions are non-uniform on the Type IV and Type VI Pareto fronts. Then, we investigated the optimal μ -distributions for the plane-based Pareto fronts. We showed that the DAS solution set on the Type VII Pareto front and the inverted DAS solution set on the Type VIII Pareto front are not always optimal for hypervolume maximization, which is contrary to our intuition. Table III summarizes all the conclusions mentioned in this paper. These conclusions can provide more knowledge for the EMO researchers to better understand the hypervolume indicator. A uniform solution set cannot always be obtained by maximizing the hypervolume indicator, which reminds us to utilize the hypervolume indicator in three dimensions more carefully.

TABLE III

A SUMMARY OF THE CONCLUSIONS FOR THE HYPERVOLUME OPTIMAL μ -DISTRIBUTION ON DIFFERENT PARETO FRONTS. MAXIMIZATION OF EACH OBJECTIVE IS CONSIDERED.

| Pareto front | Optimal μ -distribution | Reference point location |
|--------------|--|--------------------------|
| Type I | Uniform | $r \leq -1/(\mu - 1)$ |
| Type II | Nonuniform | Sufficiently far |
| Type III | Uniform | $r \leq -2/(\mu - 1)$ |
| Type IV | Nonuniform | Sufficiently far |
| Type V | Uniform | $r \leq -3/\mu$ |
| Type VI | Nonuniform | Sufficiently far |
| Type VII | Uniform ($H = 1, 2$) Nonuniform ($H > 2$) | $r = -1/H$ |
| Type VIII | Uniform ($H = 1, 2$) Nonuniform ($H > 2$) | $r = -1/H$ |

For our future work, we have the following research directions. 1) The exact hypervolume optimal μ -distributions on the Type IV, Type VI, Type VII and Type VIII Pareto fronts were not derived in this paper. One future direction is to precisely obtain the hypervolume optimal μ -distributions for these Pareto fronts. The hypervolume Newton method [32] may be a promising method for this task. 2) We only considered three dimensions in this paper. It is interesting to extend our research to higher dimensions. Some preliminary results are provided in Section VI of the supplementary material. 3) The non-uniformity caused by maximizing the hypervolume indicator can be seen as a undesired property of the

hypervolume indicator. How to solve this issue is one research direction. Another research direction is to utilize this property of hypervolume maximization for decision making (e.g., for the search of knee regions of nonlinear Pareto fronts).

REFERENCES

- [1] E. Zitzler, D. Brockhoff, and L. Thiele, "The hypervolume indicator revisited: On the design of pareto-compliant indicators via weighted integration," in *2007 International Conference on Evolutionary Multi-Criterion Optimization*. Springer, 2007, pp. 862–876.
- [2] M. Fleischer, "The measure of pareto optima applications to multi-objective metaheuristics," in *2003 International Conference on Evolutionary Multi-Criterion Optimization*. Springer, 2003, pp. 519–533.
- [3] N. Beume, B. Naujoks, and M. Emmerich, "SMS-EMOA: Multi-objective selection based on dominated hypervolume," *European Journal of Operational Research*, vol. 181, no. 3, pp. 1653–1669, 2007.
- [4] M. Emmerich, N. Beume, and B. Naujoks, "An EMO algorithm using the hypervolume measure as selection criterion," in *2005 International Conference on Evolutionary Multi-Criterion Optimization*, 2005, pp. 62–76.
- [5] S. Jiang, J. Zhang, Y.-S. Ong, A. N. Zhang, and P. S. Tan, "A simple and fast hypervolume indicator-based multiobjective evolutionary algorithm," *IEEE Transactions on Cybernetics*, vol. 45, no. 10, pp. 2202–2213, 2015.
- [6] J. Bader and E. Zitzler, "HypE: An algorithm for fast hypervolume-based many-objective optimization," *Evolutionary Computation*, vol. 19, no. 1, pp. 45–76, 2011.
- [7] K. Shang and H. Ishibuchi, "A new hypervolume-based evolutionary algorithm for many-objective optimization," *IEEE Transactions on Evolutionary Computation (Early access)*, 2020.
- [8] K. Shang, H. Ishibuchi, L. He, and L. M. Pang, "A survey on the hypervolume indicator in evolutionary multi-objective optimization," *IEEE Transactions on Evolutionary Computation*, 2020.
- [9] 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 *Proceedings of the 10th ACM SIGEVO Workshop on Foundations of Genetic Algorithms*. ACM, 2009, pp. 87–102.
- [10] M. Emmerich, A. Deutz, and N. Beume, "Gradient-based/evolutionary relay hybrid for computing pareto front approximations maximizing the s-metric," in *2007 International Workshop on Hybrid Metaheuristics*. Springer, 2007, pp. 140–156.
- [11] T. Friedrich, F. Neumann, and C. Thyssen, "Multiplicative approximations, optimal hypervolume distributions, and the choice of the reference point," *Evolutionary computation*, vol. 23, no. 1, pp. 131–159, 2015.
- [12] P. K. Shukla, N. Doll, and H. Schmeck, "A theoretical analysis of volume based Pareto front approximations," in *Proceedings of the 2014 Annual Conference on Genetic and Evolutionary Computation*, 2014, pp. 1415–1422.
- [13] A. Auger, J. Bader, and D. Brockhoff, "Theoretically investigating optimal μ -distributions for the hypervolume indicator: First results for three objectives," in *2010 International Conference on Parallel Problem Solving from Nature*. Springer, 2010, pp. 586–596.
- [14] H. K. Singh, "Understanding hypervolume behavior theoretically for benchmarking in evolutionary multi-many-objective optimization," *IEEE Transactions on Evolutionary Computation*, 2019.
- [15] T. Glasmachers, "Optimized approximation sets for low-dimensional benchmark pareto fronts," in *2014 International Conference on Parallel Problem Solving from Nature*. Springer, 2014, pp. 569–578.
- [16] H. Ishibuchi, R. Imada, N. Masuyama, and Y. Nojima, "Comparison of hypervolume, IGD and IGD+ from the viewpoint of optimal distributions of solutions," in *2019 International Conference on Evolutionary Multi-Criterion Optimization*. Springer, 2019, pp. 332–345.
- [17] H. Ishibuchi, R. Imada, Y. Setoguchi, and Y. Nojima, "Hypervolume subset selection for triangular and inverted triangular pareto fronts of three-objective problems," in *Proceedings of the 14th ACM/SIGEVO Conference on Foundations of Genetic Algorithms*. ACM, 2017, pp. 95–110.

- [18] —, "Reference point specification in hypervolume calculation for fair comparison and efficient search," in *Proceedings of the 2017 Genetic and Evolutionary Computation Conference*. ACM, 2017, pp. 585–592.
- [19] R. Tanabe and H. Ishibuchi, "An analysis of quality indicators using approximated optimal distributions in a three-dimensional objective space," *IEEE Transactions on Evolutionary Computation* (Early access), 2020.
- [20] I. Das and J. E. Dennis, "Normal-boundary intersection: A new method for generating the pareto surface in nonlinear multicriteria optimization problems," *SIAM Journal on Optimization*, vol. 8, no. 3, pp. 631–657, 1998.
- [21] K. Shang, H. Ishibuchi, W. Chen, and L. Adam, "Hypervolume optimal μ -distributions on line-based pareto fronts in three dimensions," in *Parallel Problem Solving from Nature – PPSN XVI*, T. Bäck, M. Preuss, A. Deutz, H. Wang, C. Doerr, M. Emmerich, and H. Trautmann, Eds. Cham: Springer International Publishing, 2020, pp. 257–270.
- [22] A. Auger, J. Bader, D. Brockhoff, and E. Zitzler, "Hypervolume-based multiobjective optimization: Theoretical foundations and practical implications," *Theoretical Computer Science*, vol. 425, pp. 75–103, 2012.
- [23] D. Brockhoff, "Optimal μ -distributions for the hypervolume indicator for problems with linear bi-objective fronts: Exact and exhaustive results," in *2010 Asia-Pacific Conference on Simulated Evolution and Learning*. Springer, 2010, pp. 24–34.
- [24] Q. Zhang and H. Li, "MOEA/D: A multiobjective evolutionary algorithm based on decomposition," *IEEE Transactions on Evolutionary Computation*, vol. 11, no. 6, pp. 712–731, 2007.
- [25] K. Deb and H. Jain, "An evolutionary many-objective optimization algorithm using reference-point-based nondominated sorting approach, part i: solving problems with box constraints," *IEEE Transactions on Evolutionary Computation*, vol. 18, no. 4, pp. 577–601, 2013.
- [26] H. Ishibuchi, Y. Setoguchi, H. Masuda, and Y. Nojima, "Performance of decomposition-based many-objective algorithms strongly depends on pareto front shapes," *IEEE Transactions on Evolutionary Computation*, vol. 21, no. 2, pp. 169–190, 2017.
- [27] H. Ishibuchi, H. Masuda, and Y. Nojima, "Pareto fronts of many-objective degenerate test problems," *IEEE Transactions on Evolutionary Computation*, vol. 20, no. 5, pp. 807–813, 2015.
- [28] M. Li, C. Grosan, S. Yang, X. Liu, and X. Yao, "Multiline distance minimization: A visualized many-objective test problem suite," *IEEE Transactions on Evolutionary Computation*, vol. 22, no. 1, pp. 61–78, 2017.
- [29] K. Deb, L. Thiele, M. Laumanns, and E. Zitzler, "Scalable test problems for evolutionary multiobjective optimization," in *Evolutionary Multiobjective Optimization*. Springer, 2005, pp. 105–145.
- [30] N. Beume, B. Naujoks, M. Preuss, G. Rudolph, and T. Wagner, "Effects of 1-greedy S -metric-selection on innumerable large pareto fronts," in *2009 International Conference on Evolutionary Multi-Criterion Optimization*. Springer, 2009, pp. 21–35.
- [31] H. Ishibuchi, L. M. Pang, and K. Shang, "Numerical analysis on optimal distributions of solutions for hypervolume maximization," in *2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. IEEE, 2020, pp. 1103–1110.
- [32] V. A. S. Hernández, O. Schütze, H. Wang, A. Deutz, and M. Emmerich, "The set-based hypervolume Newton method for bi-objective optimization," *IEEE Transactions on Cybernetics*, vol. 50, no. 5, pp. 2186–2196, 2018.



Ke Shang (M'19) received the B.S. and Ph.D. degrees from Xi'an Jiaotong University, China, in 2009 and 2016, respectively. He is currently a research assistant professor at Southern University of Science and Technology, China. His current research interests include evolutionary multi-objective optimization and its applications. He received GECCO 2018 Best Paper Award and CEC 2019 First Runner-up Conference Paper Award, and best paper nominations at PPSN 2020 and GECCO 2021.



Hisao Ishibuchi (M'93–SM'10–F'14) received the B.S. and M.S. degrees from Kyoto University in 1985 and 1987, respectively, and the Ph.D. degree from Osaka Prefecture University in 1992. He was with Osaka Prefecture University in 1987–2017. Since 2017, he is a Chair Professor at Southern University of Science and Technology, China. His research interests include fuzzy rule-based classifiers, evolutionary multi-objective and many-objective optimization, memetic algorithms, and evolutionary games. Dr. Ishibuchi was the IEEE CIS Vice-President for Technical Activities in 2010–2013 and the Editor-in-Chief of the IEEE COMPUTATIONAL INTELLIGENCE MAGAZINE in 2014–2019. Currently he is an AdCom member of the IEEE CIS (2014–2019, 2021–2023).



Weiyu Chen received his B.Eng. degree in computer science and technology from Southern University of Science and Technology, Shenzhen, China, in 2021. He will be a Ph.D. student at Hong Kong University of Science and Technology, Hong Kong, China. His current research interests include multi-/many-objective optimization, neural architecture search, and graph neural networks.



Yang Nan received his B.E. degree in Electrical and Electronic Engineering with the Southern University of Science and Technology (SUSTech), Shenzhen, China, in July 2018. He is currently a research assistant at the Department of Computer Science and Engineering in SUSTech. His current research interests include evolutionary multi-objective optimization and its applications.



Weiduo Liao received his B.E. degree in Computer Science and Engineering with the Southern University of Science and Technology (SUSTech), China, in July 2018 and M.E. degree in Computer Science and Engineering with the Harbin Institute of Technology, in July 2020. He is currently a PhD student at the Computer Science Department in City University of Hong Kong. His current research interests include multi-objective optimization and machine learning.