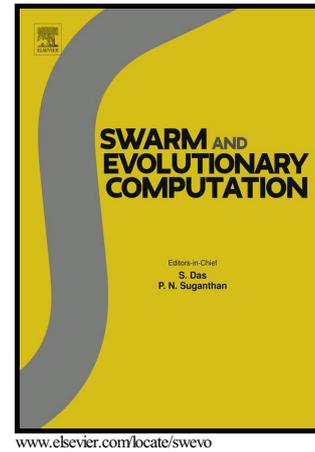


Author's Accepted Manuscript

On The Use of Two Reference Points in
Decomposition Based Multiobjective Evolutionary
Algorithms

Zhenkun Wang, Qingfu Zhang, Hui Li, Hisao
Ishibuchi, Licheng Jiao



PII: S2210-6502(16)30243-7
DOI: <http://dx.doi.org/10.1016/j.swevo.2017.01.002>
Reference: SWEVO249

To appear in: *Swarm and Evolutionary Computation*

Received date: 4 September 2016
Revised date: 2 December 2016
Accepted date: 19 January 2017

Cite this article as: Zhenkun Wang, Qingfu Zhang, Hui Li, Hisao Ishibuchi and Licheng Jiao, On The Use of Two Reference Points in Decomposition Based Multiobjective Evolutionary Algorithms, *Swarm and Evolutionary Computation*, <http://dx.doi.org/10.1016/j.swevo.2017.01.002>

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and a review of the resulting galley proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

On The Use of Two Reference Points in Decomposition Based Multiobjective Evolutionary Algorithms

Zhenkun Wang, Qingfu Zhang, Hui Li, Hisao Ishibuchi and Licheng Jiao

Abstract

Decomposition based multiobjective evolutionary algorithms approximate the Pareto front of a multiobjective optimization problem by optimizing a set of subproblems in a collaborative manner. Often, each subproblem is associated with a direction vector and a reference point. The settings of these parameters have a very critical impact on convergence and diversity of the algorithm. Some work has been done to study how to set and adjust direction vectors to enhance algorithm performance for particular problems. In contrast, little effort has been made to study how to use reference points for controlling diversity in decomposition based algorithms. In this paper, we first study the impact of the reference point setting on selection in decomposition based algorithms. To balance the diversity and convergence, a new variant of the multiobjective evolutionary algorithm based on decomposition with both the ideal point and the nadir point is then proposed. This new variant also employs an improved global replacement strategy for performance enhancement. Comparison of our proposed algorithm with some other state-of-the-art algorithms is conducted on a set of multiobjective test problems. Experimental results show that our proposed algorithm is promising.

Index Terms

Multiobjective optimization, decomposition, reference points, test instance

I. INTRODUCTION

Recently decomposition based multiobjective evolutionary algorithms (MOEA/D) have received much attention [1]–[5]. These algorithms decompose a multiobjective optimization problem (MOP) into a number of simple subproblems and optimize them in a collaborative manner. How to decompose a MOP is a major issue in the design of a MOEA/D algorithm.

In principle, any decomposition approach can be used for decomposing a MOP. Although some decomposition approaches generate a set of multiobjective subproblems [6], most approaches are single-objective decomposition which this paper focuses on. These approaches divide a MOP into a number of single-objective optimization subproblems. The objective function in each subproblem is defined by aggregating all the objectives in the MOP in question into a scalar objective by some parameters such as a direction vector and reference points [3], [7]–[9]. The optimal solutions of these subproblems will be sought by MOEA/D. Hopefully, all the obtained (nearly) optimal solutions collectively approximate the Pareto front of the MOP. The distribution of the obtained solutions, which is critical to the approximation quality, can often be determined by the used aggregation method, the setting of direction vectors and reference points. Some efforts have been made to study how to dynamically adjust aggregation functions [10], [11] and direction vectors [12]–[14] for obtaining solutions with a desirable distribution on the Pareto front. However, little work has been done on the effect of referent points in aggregation functions on the distribution of the final optimal solutions.

The setting of reference points in aggregation functions plays a key role in the performance of MOEA/D. Actually, different types of reference points could have different impacts on the search behaviors of MOEA/D. Most MOEA/D variants adopt the ideal point as its reference point. The use of the ideal point would be effective when the population's diversity is easy to maintain as argued in [15]. In contrast, a utopian point is helpful for approximating the PF boundary. In [16], a set of reference points uniformly distributed along the convex hull of PF are employed in MOEA/D for ensuring good diversity. Moreover, some attempts on the use of the nadir point in MOEA/D can be found in [17], [18]. In [18], the reference point changes from the ideal point to the nadir point if fewer solutions are obtained in the boundary regions than in the middle region of the PF after several generations. Very recently, both the utopian point and a generalized nadir point were suggested to be used simultaneously as the reference points in [19].

This work was supported by the National Natural Science Foundation of China (Grant Nos. 61473241, and 61573279).

Z. Wang and L. Jiao are with the Key Laboratory of Intelligent Perception and Image Understanding of Ministry of Education of China, Xidian University, Xi'an, Shaanxi Province 710071, China.(e-mail: wangzhenkun90@gmail.com, jlc1023@163.com).

Q. Zhang is with the Department of Computer Science, City University of Hong Kong, Kowloon, Hong Kong SAR (e-mail: qingfu.zhang@cityu.edu.hk), and the School of Computer Science and Electronic Engineering, University of Essex, Colchester CO4 3SQ, U.K. (e-mail: qzhang@essex.ac.uk).

H. Li is with the School of Mathematics and Statistics, Xi'an Jiaotong University Xi'an, Shaanxi Province 710049, China.(e-mail: lihui10@mail.xjtu.edu.cn).

H. Ishibuchi is with Department of Computer Science and Intelligent Systems, Osaka Prefecture University, Japan.(phone: +81-72-254-9350; fax:+81-72-254-9915; e-mail: hisaoi@cs.osakafu-u.ac.jp).

In this paper, we first investigate the difference between the ideal point and the nadir point on their impacts on the algorithm performance and show that they can complement each other. Then we propose a new variant of MOEA/D using both the ideal point and nadir points. A new global replacement strategy (GR) [20] is designed and used in the proposed MOEA/D variant to achieve this purpose. The proposed algorithm is referred to as the MOEA/D-MR (MOEA/D with multiple reference points). Several state-of-the-art MOEAs are compared with MOEA/D-MR on our constructed test problems.

This paper is organized as follows. In Section II, we show some basic definitions in multiobjective optimization. In Section III, we explain the motivation for using both the ideal point and the nadir points in the MOEA/D. The proposed algorithm MOEA/D-MR is explained in detail in Section IV. Performance comparison results are reported in Section V. This paper is concluded in Sections VI where some future research directions are also suggested.

II. BACKGROUND

A. Basic Definitions

A MOP can be defined as follows:

$$\begin{aligned} & \text{minimize} && F(x) = (f_1(x), \dots, f_m(x))^T \\ & \text{subject to} && x \in \Omega \end{aligned} \quad (1)$$

where $x = (x_1, \dots, x_n)^T \in R^n$ is a vector of decision variables, Ω is the feasible search region, and $F : \Omega \rightarrow R^m$ is an objective vector consisting of m objective functions, i.e., $f_i(x), i = 1, \dots, m$. If Ω is a closed and connected region in R^n and all objectives are continuous functions of x , then the MOP in (1) is called a *continuous MOP*.

In the following, some basic definitions in multiobjective optimization are given.

- Let $x, y \in \Omega$, x is said to **dominate** y , denoted by $x \prec y$, if and only if $f_i(x) \leq f_i(y)$ for all $i \in I = \{1, \dots, m\}$, and $f_j(x) < f_j(y)$ for at least one index $j \in I$.
- Let $x^* \in \Omega$, it is called **Pareto optimal** if there does not exist $x \in \Omega$ satisfying $x \prec x^*$.
- The set of all Pareto optimal solutions in the decision space is called the **Pareto set (PS)**, which can be defined as follows:

$$PS = \{x^* \in \Omega \mid \nexists x \in \Omega \text{ such that } x \prec x^*\}$$

The **Pareto front (PF)** is the set of the images of all solutions in the PS in the objective space, i.e., $PF = \{F(x) \mid x \in PS\}$.

- A point $z^{\text{id}} = (z_1^{\text{id}}, \dots, z_m^{\text{id}})^T$ is called the **ideal point** if $z_i^{\text{id}} = \inf_{x \in \Omega} f_i(x), i = 1, \dots, m$.
- A point $z^{\text{na}} = (z_1^{\text{na}}, \dots, z_m^{\text{na}})^T$ is called the **nadir point** if $z_i^{\text{na}} = \sup_{x \in PS} f_i(x), i = 1, \dots, m$.

B. Tchebycheff Decomposition

The weighted Tchebycheff approach is used for decomposition in this paper. As shown in Fig. 1, it defines a subproblem as follows:

$$\begin{aligned} & \text{minimize} && g(x|w, z^*) = \max_{1 \leq i \leq m} \{w_i |f_i(x) - z_i^*|\} \\ & \text{subject to} && x \in \Omega \end{aligned} \quad (2)$$

where

- $w = (w_1, \dots, w_m)^T$ is the weight vector with $w_i \geq 0, i = 1, \dots, m$ and $\sum_{i=1}^m w_i = 1$.
- $z^* = (z_1^*, \dots, z_m^*)^T$ is the reference point.

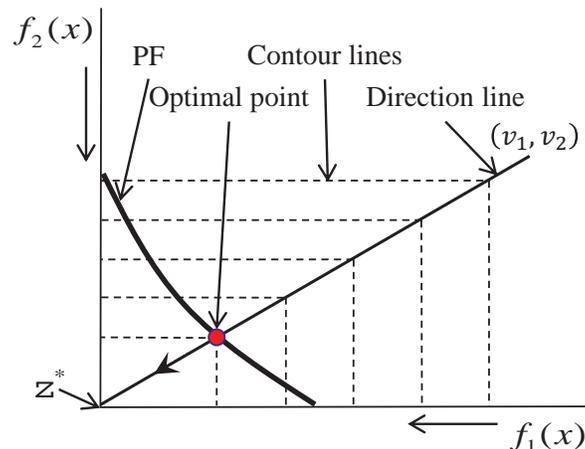


Fig. 1. Illustration of the Tchebycheff decomposition approach

As shown in Fig. 1, we can define a direction line which goes through the reference point and the optimal solution of subproblem (2). It can be represented as:

$$z = \phi v + z^* \quad (3)$$

where ϕ is the scale coefficient and $v = (v_1, \dots, v_m)^T$ is called the direction vector. v is as follows [9], [12]:

$$v_k = \frac{\frac{1}{w_k}}{\sum_{i=1}^m \frac{1}{w_i}}, k = 1, \dots, m \quad (4)$$

In practice, w_i is changed to 10^{-6} if $w_i = 0$. The intersection between the direction line and the PF is the optimal solution of subproblem (2).

III. MOTIVATIONS

In MOEA/D, the optimal solutions of the subproblems collectively approximate the PF. Thus, it is desirable that these optimal solutions are distributed uniformly along the PF. If a direction line and the PF intersect, the optimal solution of the corresponding subproblem is the intersection point as discussed in Section II. Otherwise, its optimal solution should be in the boundary of the PF. Very often, we have no good prior knowledge of the PF shape and position, it is impossible to set these direction lines properly in advance. Therefore, a fixed setting of direction lines may not efficiently solve the problem. A popular and commonly used approach for dealing with this issue in the literature is to dynamically adjust the direction vectors so that the resultant optimal solutions can approximate the PF well [12]–[14], [21], [22]. However, adjusting or resetting all the direction vectors at the same time is not always an easy task. Moreover, multiobjective search with adaptive direction vectors is less efficient than the case of fixed direction vectors with respect to the convergence to the Pareto front as discussed in [23]. In this paper, we argue that a good distribution of the optimal solutions can be obtained by appropriately specifying a single or multiple reference points.

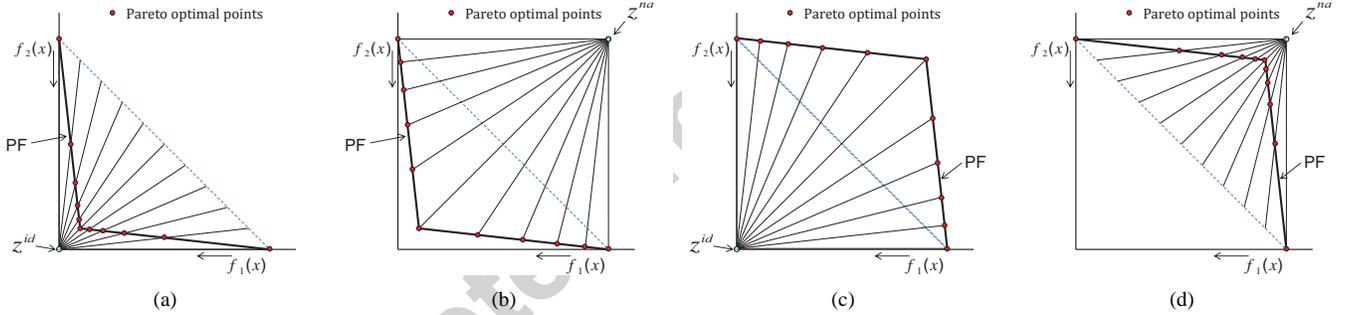


Fig. 2. Pareto optimal points obtained by the search with the ideal point z^{id} and the nadir point z^{na} respectively for the convex and concave PFs.

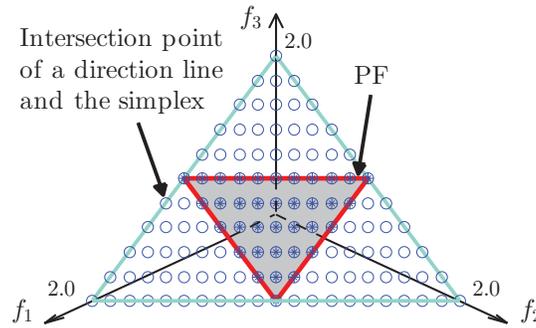


Fig. 3. Direction lines of subproblems defined using z^{id} and the inverted triangular shaped PF.

In the following, we illustrate our idea using the following two examples:

- In the case of a convex PF, the optimal solutions of subproblems are shown in Fig. 2(a) and Fig. 2(b), where z^{id} and z^{na} are used as reference points respectively. In Fig. 2(a), it is clear that the density of solutions in the central part is much larger than that near the boundaries. In Fig. 2(b), the opposite situation can be observed. In the case of a concave PF, the optimal solutions of subproblems with z^{id} are plotted in Fig. 2(c) while those with z^{na} are in Fig. 2(d). It is obvious that the optimization of subproblems defined by z^{na} will find more solutions in the central part. Since the final population

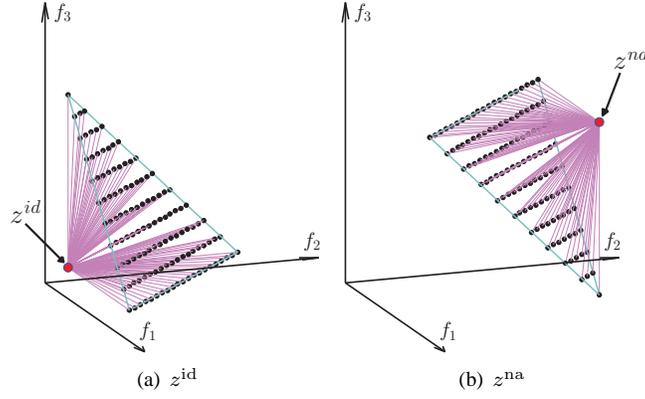


Fig. 4. Distribution of direction lines of subproblems defined using z^{id} and z^{na} in the three-objective space.

distributions by using z^{id} and z^{na} are complementary, their simultaneous use as the reference points in MOEA/D is likely to improve its performance to approximate both convex and concave PFs.

- MOEA/D with z^{id} does not perform well on a MOP with an inverted triangular shaped PF as reported in [24], [25]. As shown in Fig. 3 ([24], [25]), many direction lines do not intersect with the PF (i.e., the shaded area) whereas they are uniformly distributed in the simplex. This may lead to an undesirable situation where many subproblems can only find the solutions on the boundary of the PF. It can be observed from Fig. 4 ([25]), on one hand, the direction lines of subproblems passing z^{id} are uniformly distributed along the triangular shaped PF. On the other hand, the z^{na} based subproblems' direction lines have a even distribution on the converted triangular shaped PF. Therefore, MOEA/D with both z^{id} and z^{na} can handle the problem of mismatch in Fig. 3 in a very simple manner.

IV. MOEA/D-MR

Based on the discussion in the previous section, we propose to use two sets of subproblems simultaneously in MOEA/D, i.e., $S = S^1 \cup S^2$. Subproblems in S^1 use z^{id} as the reference point, while subproblems in S^2 adopts z^{na} as the reference point. The proposed algorithm is called MOEA/D-MR, and its pseudo-code is presented in Algorithm 1. Its some important components, such as initialization (line 3 of Algorithm 1), reproduction (line 5 to line 14 of Algorithm 1), update and replacement (line 17 to line 19 of Algorithm 1) are further illustrated in the following.

Algorithm 1: MOEA/D-MR

```

1 Input: MOP, a stopping criterion,  $N$ ,  $T_m$ ,  $T_r$ ,  $\delta$ ;
2 Output:  $P^e$ ;
3 Generate the direction vectors  $V = V^1 + V^2 = \{v^1, \dots, v^N\}$  for  $S^1$  and  $S^2$ . Work out the neighborhood relations
  between direction vectors within  $V^1$  and  $V^2$  respectively. Initialize the population  $P \leftarrow \{x^1, \dots, x^N\}$ ,  $z^{id}$  and  $z^{na}$ ;
4 Set  $P^e \leftarrow \emptyset$ ;
5 while the stopping criterion is not satisfied do
6    $Q \leftarrow \emptyset$ ;
7   for  $i \leftarrow 1$  to  $N$  do
8     if  $rand < \delta$  then
9        $E \leftarrow B(i)$ ;
10    else
11       $E \leftarrow \{1, \dots, N\}$ ;
12    end
13    Set  $x^{r_1} \leftarrow x^i$  and randomly select  $r_2$  and  $r_3$  from  $E$ , then generate a new solution  $\bar{x}^i$  by applying the
    reproduction operators on them, and  $Q \leftarrow Q \cup \{\bar{x}^i\}$ ;
14  end
15  Evaluate the function values of each solution in  $Q$ ;
16   $R \leftarrow P \cup Q$ ;
17  Update  $z^{id}$  and  $z^{na}$ ;
18   $P \leftarrow GR(R, V, T_r, z^{id}, z^{na})$ ;
19  Update  $P^e$  with  $P$ ;
20 end

```

A. Initialization

In MOEA/D-MR, we use two subproblem sets of the same size for simplicity, each set has $\frac{N}{2}$ subproblems. The direction vectors of subproblems can be represented as $V = V^1 + V^2 = \{v^1, \dots, v^N\}$. $V^1 = \{v^1, \dots, v^{\frac{N}{2}}\}$ is the direction vector set corresponding to S^1 , and $V^2 = \{v^{\frac{N}{2}+1}, \dots, v^N\}$ is the direction vector set of S^2 .

In MOEA/D-MR, the generation of two sets of direction vectors are the same, i.e., $V^1 = V^2$. Direction vectors in V^1 and V^2 are uniformly sampled from the $(m-1)$ -dimensional unit simplex. That is, for every $v \in V^1$ and V^2 , each of its elements takes a value from $\{\frac{0}{H}, \frac{1}{H}, \dots, \frac{H}{H}\}$ under the condition $\sum_{i=1}^m v_i = 1$. H is a user-defined integer. Then the weight vector of each subproblem can be calculated by using Eq. (4).

The neighborhood relationships between subproblems are obtained within each subproblem set by computing the Euclidean distances between their corresponding direction vectors. For one subproblem in S^1 or S^2 , its neighborhood is composed by T subproblems from the same subproblems set, whose direction vectors are closest to its own direction vector in terms of Euclidean distance. As suggested in [26], [27], two neighborhoods, the mating neighborhood and the replacement neighborhood are used in this paper. The former neighborhood is adopted for the mating parents selection and the latter one is used for determining the replacement range, T_m and T_r are their sizes, respectively.

The initial population $P = \{x^1, \dots, x^N\}$ is randomly sampled from the decision space. A randomly generated solution x^i is randomly assigned to a subproblem s^i . To initialize z^{id} and z^{na} , we first find all the nondominated solutions in P and let them form P^{nd} , then set $z_i^{\text{id}} = \min\{f_i(x)|x \in P^{\text{nd}}\}$ and $z_i^{\text{na}} = \max\{f_i(x)|x \in P^{\text{nd}}\}$, for all $i \in \{1, \dots, m\}$.

B. Reproduction

We use the same reproduction operators as in [4]. That is, the differential evolution (DE) operator and polynomial mutation operator are employed for producing new solutions in this paper. For the each subproblem s^i , $B(i)$ consists of the indexes of T_m neighboring subproblems. As show in the line 8 to line 13 of Algorithm 1, three mating parents x^{r_1} , x^{r_2} and x^{r_3} are first selected. Then, an intermediate solution $y^i = (y_1^i, \dots, y_n^i)$ is generated by:

$$y_k^i = \begin{cases} x_k^{r_1} + F \times (x_k^{r_2} - x_k^{r_3}) & \text{if } \text{rand} \leq CR \\ x_k^{r_1} & \text{otherwise} \end{cases} \quad (5)$$

Then, a new solution $\bar{x}^i = (\bar{x}_1^i, \bar{x}_2^i, \dots, \bar{x}_n^i)$ is generated by polynomial mutation as follows:

$$\bar{x}_k^i = \begin{cases} y_k^i + \sigma_k \times (b_k - a_k) & \text{if } \text{rand} \leq p_m \\ y_k^i & \text{otherwise} \end{cases} \quad (6)$$

with

$$\sigma_k = \begin{cases} (2 \times \text{rand})^{\frac{1}{\eta+1}} - 1 & \text{if } \text{rand} < 0.5 \\ 1 - (2 - 2 \times \text{rand})^{\frac{1}{\eta+1}} & \text{otherwise} \end{cases} \quad (7)$$

where CR and F are two parameters in the DE operator, the distribution index η and the mutation rate p_m are control parameters in the polynomial mutation operator. rand is a uniform random number from $[0, 1]$. a_k and b_k are the lower and upper bounds of the k -th decision variable.

After the whole reproduction operation, the offspring population Q is evaluated (line 15 of Algorithm 1) and combined with the parent population P to form R (line 16 of Algorithm 1).

C. Global Replacement

Before performing the replacement strategies for subproblems, the reference points, z^{id} and z^{na} , are updated firstly (line 17 of Algorithm 1). To do so, just as in Initialization, we first find all the nondominated solutions in R and let them form P^{nd} , then set $z_i^{\text{id}} = \min\{f_i(x)|x \in P^{\text{nd}}\}$ and $z_i^{\text{na}} = \max\{f_i(x)|x \in P^{\text{nd}}\}$, for all $i \in \{1, \dots, m\}$.

In the following, the replacement operator is conducted to update the current solutions of subproblems with the whole population R . To achieve this, a new GR strategy with mutual selection between solutions and subproblems is proposed in this paper (line 18 of Algorithm 1). The proposed replacement strategy can reduce the mismatch between subproblems and solutions, and thus improve the algorithm performance [20], [28]–[31]. Its implementation is illustrated in Algorithm 2, it mainly works as the following steps:

- Objective Normalization (line 4 of Algorithm 2): To tackle MOPs with dissimilar PF ranges, the objective values of all solutions in the population are normalized before the replacement in MOEA/D. For one solution x , the vector $(\bar{f}_1, \dots, \bar{f}_m)^T$ of its normalized objective function values is calculated by:

$$\bar{f}_i(x|z^*) = \frac{f_i(x) - z_i^*}{z_i^{\text{na}} - z_i^{\text{id}}} \quad (8)$$

where, the setting of z^* depends on the type of subproblems. $z^* = z^{\text{id}}$ if $i = 1, \dots, N/2$; otherwise, $z^* = z^{\text{na}}$.

Algorithm 2: GR ($R, V, T_r, z^{\text{id}}, z^{\text{na}}$)

```

1 Input:  $R, V, T_r, z^{\text{id}}$  and  $z^{\text{na}}$ ;
2 Output:  $P$ ;
3  $P \leftarrow \emptyset$ ;
4 Normalize ( $R$ );
5  $\pi \leftarrow \text{Associate} (R, V, z^{\text{id}}, z^{\text{na}}, T_r)$ ;
6 for  $i \leftarrow 1$  to  $N$  do
7   if  $\pi(i) \neq \emptyset$  then
8      $x^j \leftarrow \arg \min_{x^k \in \pi(i)} \{g(x^k|v^i, z^*)\}$ ;
9     if  $g(x^j|v^i, z^*) < g(x^i|v^i, z^*)$  then
10       $P \leftarrow P \cup \{x^j\}, V \leftarrow V \setminus \{v^i\}$ ;
11      if  $x^j \in R$  then
12         $R \leftarrow R \setminus \{x^j\}$ ;
13      end
14    end
15  end
16 end
17 while  $V \neq \emptyset$  do
18   Randomly choose a direction vector  $v^i$  from  $V$ ;
19   Find  $T_r$  closest solutions to  $v^i$  among  $R$ ;
20    $x^j \leftarrow \arg \min_{x^k \in T_r(i)} \{g(x^k|v^i, z^*)\}$ ;
21   if  $g(x^j|v^i, z^*) < g(x^i|v^i, z^*)$  then
22      $P \leftarrow P \cup \{x^j\}, V \leftarrow V \setminus \{v^i\}, R \leftarrow R \setminus \{x^j\}$ ;
23   else
24      $P \leftarrow P \cup \{x^i\}, V \leftarrow V \setminus \{v^i\}$ ;
25   end
26 end

```

Algorithm 3: Associate ($R, V, z^{\text{id}}, z^{\text{na}}, T_r$)

```

1 Input:  $R, V, z^{\text{id}}, z^{\text{na}}$  and  $T_r$ ;
2 Output:  $\pi$ ;
3  $d \leftarrow \emptyset, \pi \leftarrow \emptyset$ ;
4 for each  $x^i \in R$  do
5   for each  $v^j \in V$  do
6     Compute  $d_{ij}$  by Eq. (9) and  $d(i) \leftarrow d(i) \cup d_{ij}$ ;
7   end
8   Sort  $d(i)$  in ascending order, and allocate  $x^i$  to  $T_r$  closest direction vectors,
9   i.e.,  $\pi(k) \leftarrow \pi(k) \cup \{x^i\}$ , for every  $v^k \in T_r(i)$ ; /*  $T_r(i)$  is the set of  $T_r$  direction vectors which
   are closest to  $x^i$  */;
10 end

```

- Assignment of Subproblems for Solutions (line 5 of Algorithm 2): The details of this process are shown in Algorithm 3. At the end of this operation, each member of the population is associated with T_r nearest subproblems. For each solution x^i and each direction vector v^j , the distance between them is calculated as follows:

$$d_{ij} = \left\| \bar{F}(x^i|z^*) - \frac{v^j T \bar{F}(x^i|z^*) v^j}{v^j T v^j} \right\| \quad (9)$$

where the setting of z^* is the same as above. Let $T_r(i)$ is the set of T_r direction vectors which are closet to x^i . For each subproblem $v^k \in T_r(i)$, we update $\pi(k) \leftarrow \pi(k) \cup \{x^i\}$.

- Assignment of Solutions for Subproblems (line 6 to line 16 of Algorithm 2): Since the normalization has been done all solutions, the aggregation function of each subproblem $s^i \in S$ can be defined as:

$$g(x|w^i, z^*) = \max_{1 \leq k \leq m} \{w_k^i \bar{f}_k(x|z^*)\} \quad (10)$$

where the setting of z^* is also the same as above. Let x^j is the best solution from $\pi(i)$ in terms of Eq. (10). If x^j is better than x^i , then match subproblem i with x^j and remove x^j from the population if x^j is still its member.

After the loop, there might be some subproblems without being updated. To deal with this issue, we repeat the following procedure (line 17 to line 26 of Algorithm 2):

- Randomly chose one subproblem from these subproblems and remove it from them. Find T_r closest solutions to its direction vectors from the population.
- Determine the best one from these solutions. If it is better than the current solution in terms of the objective function (10), we replace the current solution with it and remove it from the population, otherwise keep the current solution.

Note that the parameter T_r has a great influence on the balance between convergence and diversity. On one hand, a hand, a small value of T_r encourages the diversity of population. This is because any new solution can only replace the solutions in a small neighborhood. On the other hand, a large value of T_r emphasizes the convergence

D. External Population

In MOEA/D-MR, an external population (P^e) is employed to store the non-dominated solutions found during the search. If the maximum size of P^e is exceeded, the overcrowded ones are removed. To do this, the k -th nearest neighbor method is used to measure the density of each solution in P^e . Its details can be found in [32].

V. EXPERIMENTAL STUDIES

A. Performance Metrics

In our experimental studies, two widely-used performance metrics, i.e., the inverted generational distance metric (IGD) and the Hypervolume indicator (I_H) [33], [34] are adopted in assessing the performance of the compared algorithms.

- 1) IGD-metric [33]: Let P^* be a set of uniformly distributed Pareto optimal points along the PF in the objective space. Let P be an approximate set to the PF obtained by an algorithm. The inverted generational distance from P^* to P is defined as:

$$IGD(P^*, P) = \frac{\sum_{v \in P^*} d(v, P)}{|P^*|} \quad (11)$$

where $d(v, P)$ is the minimum Euclidean distance between v and P . If $|P^*|$ is large enough, $IGD(P^*, P)$ can measure the quality of P both in convergence and in diversity. When the PFs are known in advance, we choose 500 representative points to approximate P^* for all bi-objective test problems, and 1,000 for the test problems with three objectives.

- 2) I_H -metric [35]: Let $y^* = (y_1^*, y_2^*, \dots, y_m^*)$ be a reference point satisfying $y_i^* \geq \max_{x \in PS} f_i(x)$. The I_H value of the approximation S of PF against y^* is the volume of the region between S and y^* , which is computed by:

$$I_H(S, y^*) = \text{vol}(\bigcup_{y \in S} [y_1, y_1^*] \times \dots \times [y_m, y_m^*]) \quad (12)$$

where $\text{vol}(\cdot)$ is the Lebesgue measure. In our experiments, each component y_i^* of the reference point is set to be $z_i^{\text{na}} + 0.2$.

B. Experimental Settings

To verify the efficiency of our proposed method in diversity, a set of nine new test instances formulated in Table I are exclusively constructed. The major features of these test problems are summarized as follows:

- F1-F3 are three unimodal MOPs with linear variable linkages. The whole or part PF shapes of them are either highly convex or highly concave. As stated in Section III, F1-F3 can cause difficulties for the decomposition based MOEAs [14], [36].
- F4 and F5 are two unimodal MOPs with complicated PS shapes. The boundary parts of their PFs are more difficult to approximate than other parts. F6 and F7 are two multimodal MOPs with complicated PS shapes.
- F8 and F9 are two unimodal triple-objective problems with complicated PS shapes. These two test problems have the inverted triangle-like shapes. The similar test problems can refer to [24], [37].

We compare our proposed algorithm MOEA/D-MR with four other state-of-the-art MOEAs, i.e., NSGA-II [38], SMS-EMOA [39], MOEA/D-DE [4] and NSGA-III [40]. NSGA-II is a popular domination based MOEA while SMS-EMOA is a well-known indicator based MOEA. MOEA/D-DE is a representative steady-state decomposition based MOEA, which is good at solving the MOPs with complicated PS shapes. NSGA-III is an extension of NSGA-II for dealing with many-objective optimization, which maintains the diversity of population via decomposition. The detailed parameter settings are summarized as follows:

- 1) *The setting of population size*: N is set to 200 for the bi-objective test problems and 600 for the ones with three objectives. In MOEA/D-MR, the two subsets S^1 and S^2 have the same size, i.e., $N/2$.

TABLE I
TEST INSTANCES

Instance	n	Variable domains	Objective functions	Characteristics
F1	30	$x_1 \in [0, 1]$, $x_i \in [-1, 1]$, $2 \leq i \leq n$	$f_1 = 1 - \cos(0.5\pi x_1) + \frac{2}{ J_1 } \sum_{j \in J_1} y_j ^{0.7}$ $f_2 = 10 - 10\sin(0.5\pi x_1) + \frac{2}{ J_2 } \sum_{j \in J_2} y_j ^{0.7}$ $y_j = x_j - 0.9\sin(\frac{j\pi}{n})$, $j \in 2, \dots, n$. where $J_1 = \{j j \text{ is odd and } 2 \leq j \leq n\}$ and $J_2 = \{j j \text{ is even and } 2 \leq j \leq n\}$	Dissimilar PF ranges, convex PF, linear PS, unimodal, separability
F2	30	$x_1 \in [0, 1]$, $x_i \in [-1, 1]$, $2 \leq i \leq n$	$f_1 = x_1 + \frac{2}{ J_1 } \sum_{j \in J_1} y_j + \frac{\sin(\pi y_j)}{\pi} $ $f_2 = \begin{cases} 1 - 19x_1 + \frac{2}{ J_2 } \sum_{j \in J_2} y_j + \frac{\sin(\pi y_j)}{\pi} & \text{if } x_1 \leq 0.005 \\ \frac{1}{19} - \frac{x}{19} + \frac{2}{ J_2 } \sum_{j \in J_2} y_j + \frac{\sin(\pi y_j)}{\pi} & \text{otherwise} \end{cases}$ $y_j = x_j - 0.9\sin(\frac{j\pi}{n})$, $j \in 2, \dots, n$. where J_1 and J_2 are the same with as those of F1.	Convex PF, piecewise linear PF, linear PS, unimodal, separability
F3	30	$x_1 \in [0, 1]$, $x_i \in [-1, 1]$, $2 \leq i \leq n$	$f_1 = x_1 + \frac{2}{ J_1 } \sum_{j \in J_1} (1 - e^{- y_j })$ $f_2 = \begin{cases} 1 - 8x_1^4 + \frac{2}{ J_2 } \sum_{j \in J_2} (1 - e^{- y_j }) & \text{if } x_1 \leq 0.5 \\ 8(1 - x_1)^4 + \frac{2}{ J_2 } \sum_{j \in J_2} (1 - e^{- y_j }) & \text{otherwise} \end{cases}$ $y_j = x_j - 0.9\sin(\frac{j\pi}{n})$, $j \in 2, \dots, n$. where J_1 and J_2 are the same with as those of F1.	Convex PF, concave PF, Mixed PF, linear PS, unimodal, separability
F4	30	$x_1 \in [0, 1]$, $x_i \in [-1, 1]$, $2 \leq i \leq n$	$f_1 = x_1 + \frac{2}{ J_1 } \sum_{j \in J_1} (y_j)^2$ $f_2 = 1 - \sqrt{x_1} + \frac{2}{ J_2 } \sum_{j \in J_2} (y_j)^2$ $y_j = \begin{cases} x_j - 1.8 x_1 - 0.55 \cos(5\pi x_1 + \frac{j\pi}{n}) & j \in J_1 \\ x_j - 1.8 x_1 - 0.55 \sin(5\pi x_1 + \frac{j\pi}{n}) & j \in J_2 \end{cases}$ where J_1 and J_2 are the same with as those of F1.	Convex PF, nonlinear PS, unimodal, nonseparability
F5	30	$x_1 \in [0, 1]$, $x_i \in [-1, 1]$, $2 \leq i \leq n$	$f_1 = x_1 + \frac{2}{ J_1 } \sum_{j \in J_1} (y_j)^2$ $f_2 = 1 - \sqrt{x_1} + \frac{2}{ J_2 } \sum_{j \in J_2} (y_j)^2$ $y_j = \begin{cases} x_j - [0.75t^2 \cos(16\pi x_1 + \frac{4j\pi}{n}) + 1.4 t] \cos(4\pi x_1 + \frac{j\pi}{n}) & j \in J_1 \\ x_j - [0.75t^2 \cos(16\pi x_1 + \frac{4j\pi}{n}) + 1.4 t] \sin(4\pi x_1 + \frac{j\pi}{n}) & j \in J_2 \end{cases}$ where J_1 and J_2 are the same with as those of F1 and $t = x_1 - 0.55$.	Convex PF, nonlinear PS, nonseparability, unimodal
F6	30	$x_1 \in [0, 1]$, $x_i \in [-1, 1]$, $2 \leq i \leq n$	$f_1 = x_1 + \frac{0.02}{ J_1 } \sum_{j \in J_1} (100y_j^2 - \cos(10\pi y_j) + 1)$ $f_2 = 1 - \sqrt{x_1} + \frac{0.02}{ J_2 } \sum_{j \in J_2} (100y_j^2 - \cos(10\pi y_j) + 1)$ $y_j = x_j - 0.9\sin(2\pi x_1 + \frac{j\pi}{n})$, $j \in 2, \dots, n$. where J_1 and J_2 are the same with as those of F1.	Convex PF, nonlinear PS, nonseparability, multimodal
F7	30	$x_1 \in [0, 1]$, $x_i \in [-1, 1]$, $2 \leq i \leq n$	$f_1 = x_1 + \frac{0.05}{ J_1 } (100 \sum_{j \in J_1} y_j^2 - \prod_{j \in J_1} \cos(\frac{20\pi y_j}{\sqrt{j}}) + 1)$ $f_2 = 1 - \sqrt{x_1} + \frac{0.05}{ J_2 } (100 \sum_{j \in J_2} y_j^2 - \prod_{j \in J_2} \cos(\frac{20\pi y_j}{\sqrt{j}}) + 1)$ $y_j = x_j - 0.9\sin(2\pi x_1 + \frac{j\pi}{n})$, $j \in 2, \dots, n$. where J_1 and J_2 are the same with as those of F1.	Convex PF, nonlinear PS, nonseparability, multimodal
F8	10	$x_1, x_2 \in [0, 1]$, $x_i \in [-1, 1]$, $3 \leq i \leq n$	$f_1 = 1 - x_1 x_2 + \sum_{j=3}^n y_j^2$ $f_2 = 1 - x_1 + x_1 x_2 + \sum_{j=3}^n y_j^2$ $f_3 = x_1 + \sum_{j=3}^n y_j^2$ $y_j = x_j - 0.9 \prod_{k=1}^2 \sin(2\pi x_k + \frac{j\pi}{n})$, $j \in 3, \dots, n$.	Linear PF, nonlinear PS, nonseparability, unimodal
F9	10	$x_1, x_2 \in [0, 1]$, $x_i \in [-1, 1]$, $3 \leq i \leq n$	$f_1 = 1 - \cos(\frac{\pi x_1}{2}) \cos(\frac{\pi x_2}{2}) + 0.01 \sum_{j=3}^n (100y_j^2 - \cos(4\pi y_j) + 1)$ $f_2 = 1 - \cos(\frac{\pi x_1}{2}) \sin(\frac{\pi x_2}{2}) + 0.01 \sum_{j=3}^n (100y_j^2 - \cos(4\pi y_j) + 1)$ $f_3 = 1 - \sin(\frac{\pi x_1}{2}) + 0.01 \sum_{j=3}^n (100y_j^2 - \cos(4\pi y_j) + 1)$ $y_j = x_j - 0.9 \prod_{k=1}^2 \sin(2\pi x_k + \frac{j\pi}{n})$, $j \in 3, \dots, n$.	Convex PF, nonlinear PS, nonseparability, multimodal

- 2) *Number of runs and stopping condition*: Each algorithm is run 30 times independently for each test problem. In each run, the maximal number of function evaluations is set to 150,000 for terminating every algorithm.
- 3) *Control parameters in reproduction operators*: As suggested in [4], the crossover rate CR and the scaling factor F in the DE operator are set to 1.0 and 0.5 respectively. In SBX, the crossover probability p_c is set to 1.0, and the distribution index e_c is set to 20. In polynomial mutation [41], the mutation rate p_m is set to $1/n$ while the distribution index η_m is set to 20.
- 4) *The neighborhood sizes*: In MOEA/D-DE, the mating neighborhood size T_m is set to $0.1 \times N$. In MOEA/D-MR, two mating neighborhood sizes, T_{m1} and T_{m2} , are set to $0.05 \times N$. In both MOEA/D variants, the replacement neighborhood size T_r is set to 5 for F1-F7 and 10 for F8 and F9.
- 5) *Other control parameters*: In both MOEA/D-DE and MOEA/D-MR, the probability δ of selecting mating parents from the whole population is set to 0.8. In MOEA/D-DE, the maximal number n_r of solutions replaced by one offspring solutions is limited to 2.

Other exclusive parameters of these four algorithms are the same as in [4], [38]–[40].

TABLE II
IGD-METRIC VALUES OBTAINED BY MOEA/D-MR AND FOUR OTHER ALGORITHMS ON F1-F9

IGD Instance	NSGA-II		SMS-EMOA		MOEA/D-DE		NSGA-III		MOEA/D-MR					
		median	best		median	best		median	best	median	best			
F1	+	1.67E-2	1.49E-2	+	1.42E-2	1.39E-2	-	1.67E-1	1.44E-1	-	3.07E-2	2.35E-2	2.03E-2	1.85E-2
F2	+	3.00E-3	2.90E-3	+	2.40E-3	2.39E-3	-	1.12E-2	9.92E-3	-	1.17E-2	7.68E-3	3.51E-3	3.02E-3
F3	+	2.89E-3	2.54E-3	-	4.77E-3	3.86E-3	-	1.02E-2	8.29E-3	-	8.20E-3	6.31E-3	3.75E-3	3.43E-3
F4	-	3.09E-2	1.98E-2	-	3.45E-2	2.38E-2	-	4.78E-2	3.83E-2	-	6.42E-2	3.61E-2	3.85E-3	2.93E-3
F5	-	1.78E-2	1.44E-2	-	1.83E-2	1.39E-2	-	1.03E-2	5.77E-3	-	3.92E-2	2.83E-2	5.91E-3	4.06E-3
F6	-	1.04E-1	8.46E-2	-	6.06E-2	3.32E-2	-	7.58E-3	6.20E-3	-	1.20E-1	8.91E-2	3.70E-3	3.07E-3
F7	-	1.28E-1	1.10E-1	-	9.85E-2	5.23E-2	-	8.49E-3	6.13E-3	-	2.12E-1	1.04E-1	3.91E-3	3.48E-3
F8	-	8.69E-2	6.71E-2	-	5.87E-2	5.37E-2	-	4.62E-2	4.25E-2	-	8.42E-2	7.28E-2	2.82E-2	2.60E-2
F9	-	1.41E-1	1.13E-1	-	8.09E-2	7.40E-2	-	5.06E-2	4.51E-2	-	1.25E-1	1.03E-1	4.04E-2	3.39E-2

+, - and \approx denote that the performance of the corresponding algorithm is significantly better than, worse than, and similar to MOEA/D-MR respectively by Wilcoxon's rank sum test with $\alpha = 0.05$.

TABLE III
 I_H -METRIC VALUES OBTAINED BY MOEA/D-MR AND FOUR OTHER ALGORITHMS ON F1-F9

I_H Instance	NSGA-II		SMS-EMOA		MOEA/D-DE		NSGA-III		MOEA/D-MR					
		median	best		median	best		median	best	median	best			
F1	+	10.0550	10.0585	+	10.0754	10.0755	-	9.7898	9.8370	+	10.0566	10.0612	9.9727	9.9876
F2	+	1.3891	1.3892	+	1.3895	1.3895	-	1.3794	1.3815	\approx	1.3864	1.3882	1.3858	1.3869
F3	+	0.9370	0.9373	+	0.9384	0.9384	-	0.9258	0.9279	+	0.9369	0.9373	0.9332	0.9344
F4	-	1.0134	1.0447	-	0.9969	1.0564	-	0.9468	0.9691	-	0.9457	0.9979	1.0972	1.1003
F5	-	1.0454	1.0599	-	1.0630	1.0746	-	1.0625	1.0864	-	0.9898	1.0325	1.0873	1.0978
F6	-	0.9512	0.9772	-	1.0109	1.0433	-	1.0870	1.0917	-	0.8982	0.9386	1.0995	1.1008
F7	-	0.9100	0.9275	-	0.9289	1.0146	-	1.0829	1.0927	-	0.7870	0.9234	1.0989	1.0998
F8	-	0.4318	0.4672	-	0.4816	0.4882	-	0.5095	0.5148	-	0.4204	0.4539	0.5373	0.5411
F9	-	0.9057	0.9405	-	0.9261	0.9405	-	1.0220	1.0290	-	0.8991	0.9298	1.0568	1.0608

+, - and \approx denote that the performance of the corresponding algorithm is significantly better than, worse than, and similar to MOEA/D-MR respectively by Wilcoxon's rank sum test with $\alpha = 0.05$.

C. Experimental Results

The quantitative results on the IGD-metric values and the I_H -metric values obtained by five algorithms on F1-F9 are presented in Table II and Table III respectively. From the IGD-metric values and the Wilcoxon's rank sum tests [42] in Table II, it is evident that MOEA/D-MR significantly outperforms other four MOEAs on F4-F9 in minimizing IGD-metric values. Moreover, MOEA/D-MR is only slightly worse than NSGA-II on F3. For F1-F2, although MOEA/D-MR is not as good as NSGA-II and SMS-EMOA, it still performs better than MOEA/D-DE and NSGA-III. From the results in Table III, it is clear that MOEA/D-MR has better performance than other four algorithms on the majority of test problems. On F4-F9, MOEA/D-MR performs best among five algorithms. On F1-F3, MOEA/D-MR cannot beat NSGA-II, SMS-EMOA and NSGA-III in achieving better I_H -metric values. However, it should be pointed out that MOEA/D-MR performs much better than MOEA/D-DE on F1-F3 according to the I_H -metric values and the Wilcoxon's rank sum tests. This indicates that MOEA/D-MR has better ability than MOEA/D-DE for handling the MOPs with highly convex PFs.

Apart from the quantitative results on two indicators, the PF graphical plotting is used to visualize the diversity of the final solutions. Fig. 5 shows the distributions of the final populations with the median IGD-metric values found by five algorithms in 30 runs on F1-F3. It can be observed in Fig 5(a) that MOEA/D-DE is the only MOEA that fails to approximate the whole PF of F1 well. Precisely speaking, MOEA/D-DE can not find a good approximation along the left part of the PF. Moreover, the reason that both NSGA-III and MOEA/D-MR are superior to MOEA/D-DE is due to the fact that both of them adopt objective normalization for dealing with dissimilar objective ranges. In Fig 5(b), it is easy to see that NSGA-II, SMS-EMOA and MOEA/D-MR can obtain a set of solutions with satisfactory distribution along the PF of F2. In contrast, both NSGA-III and MOEA/D-DE find fewer solutions on the boundary parts of the PF. As shown in Fig 5(c), NSGA-III, SMS-EMOA and MOEA/D-DE are unable to approximate the boundary parts of the PF of F3, while other two MOEAs can approximate the whole PF well. To compare the convergence speed of five algorithms, the mean IGD values obtained by them on F1-F3 at different generations are shown in Fig 6. It can be observed that NAGA-II, SMS-EMOA and MOEA/D-MR converge faster than MOEA/D-DE and NSGA-III. Among three decomposition based algorithms, MOEA/D-MR performs best while MOEA/D-DE performs worst.

On F4 and F5, the final solutions with the median IGD-metric values obtained by five algorithms in 30 runs are plotted in Fig. 7 and Fig. 8. It is evident that MOEA/D-MR obtains the best approximation of the PFs on these two problems. NSGA-II, SMS-EMOA and NSGA-III can only find the middle part of the PFs. MOEA/D-DE is unable to approximate the boundary parts well. In Fig. 9, the evolution of the mean IGD-metric values found by five algorithms is plotted. These results clearly show that MOEA/D-MR is superior to all others in convergence towards the PFs of F4 and F5 with complicated PS shapes.

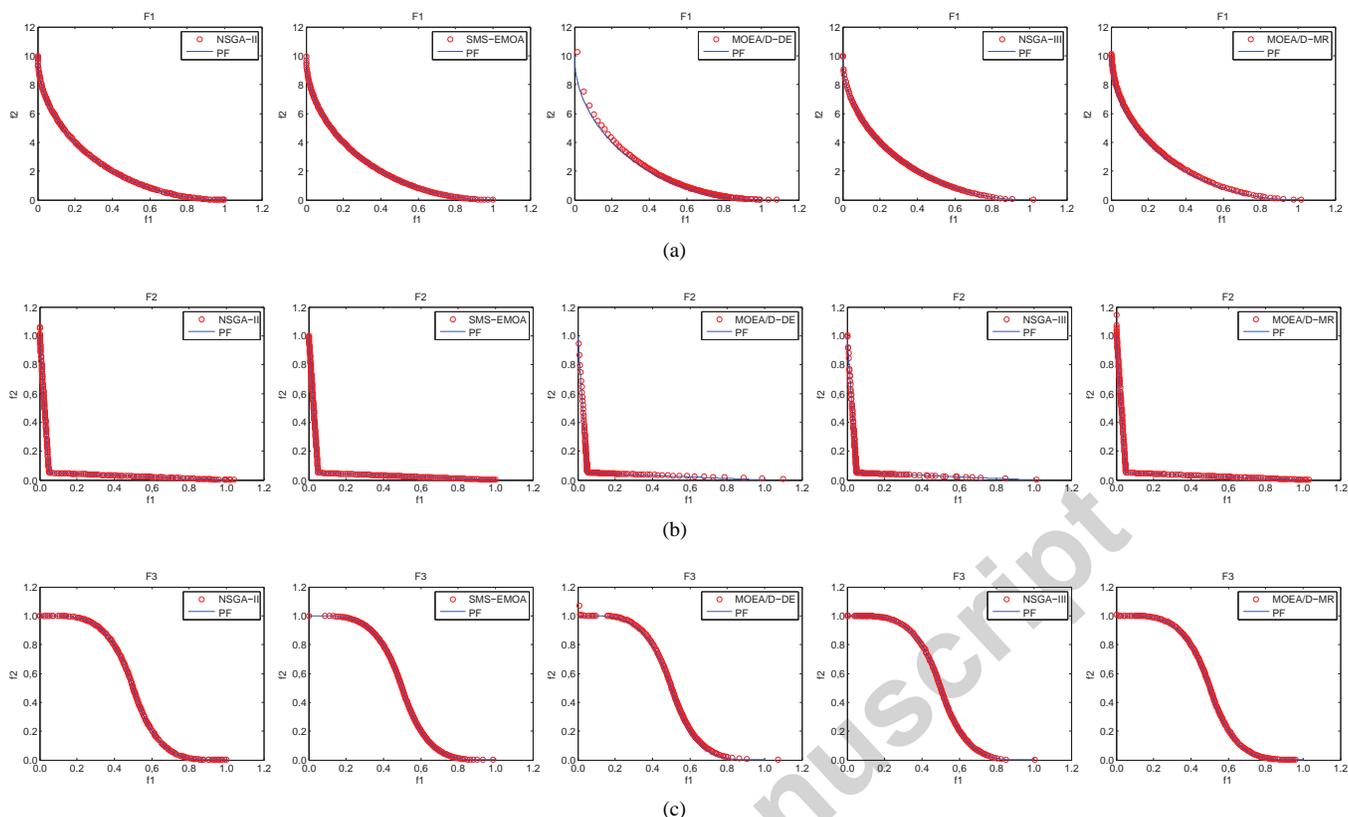


Fig. 5. Plots of the final solutions with the median IGD-metric values found by MOEA/D-MR and four other MOEAs in 30 runs in the objective space on F1-F3.

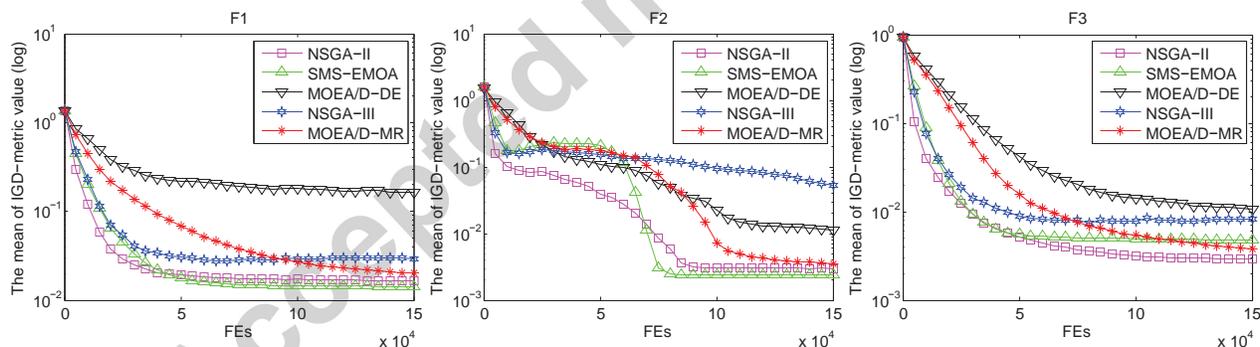


Fig. 6. Evolution of the mean IGD-metric values of MOEA/D-MR and four other MOEAs on F1-F3 during the evolutionary process.

On F6 and F7, the final solutions found by five algorithms in their median runs are shown in Fig. 10 and Fig. 11. Obviously, MOEA/D-MR approximates the whole PFs well while MOEA/D-DE performs slightly worse in approximating the boundary parts of their PFs. In contrast, other three algorithms perform very badly on these two problems. NSGA-II and NSGA-III can only find several solutions close to the PFs while SMS-EMOA fail to approximate the whole PF well. Fig. 12 shows the evolution of the mean IGD values of the populations obtained by five algorithms. It can be observed that MOEA/D-MR and MOEA/D-DE perform remarkably better than three other algorithms on F6 and F7. Compared with the results on F4 and F5, MOEA/D-MR is more advantageous than other MOEAs when dealing multimodal MOPs like F6 and F7.

F8 and F9 are two three-objective problems with inverted triangle-like PF shapes. Fig. 13 and Fig. 14 plot the final solutions with the median IGD-metric values obtained by different algorithms. It is easy to see that both NSGA-II and NSGA-III find several points near the PFs. The solutions obtained by SMS-EMOA and MOEA/D-DE are unevenly distributed along the PFs (Fig. 3 explains the MOEA/D-DE results). In contrast, the solutions found by MOEA/D-MR cover the whole PFs very well. Fig. 15 shows the evolution of the mean IGD values obtained by five algorithms. Again, MOEA/D-MR outperforms four other algorithms on F8 and F9.

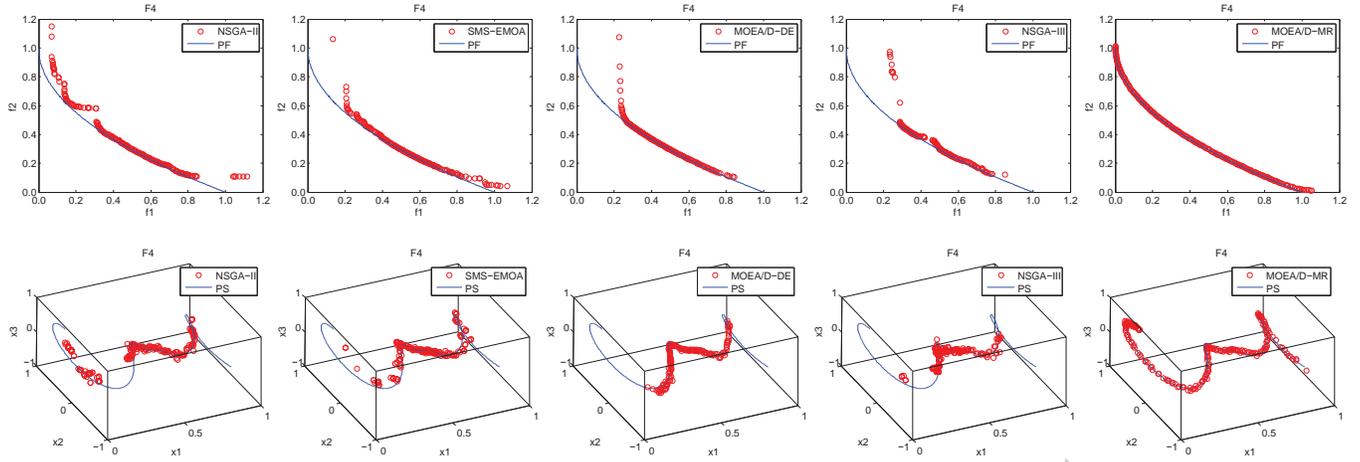


Fig. 7. Plots of the final solutions with the median IGD-metric values found by MOEA/D-MR and four other MOEAs in 30 runs in the objective space (the top row) and the $x_1 - x_2 - x_3$ space (the bottom row) on F4.

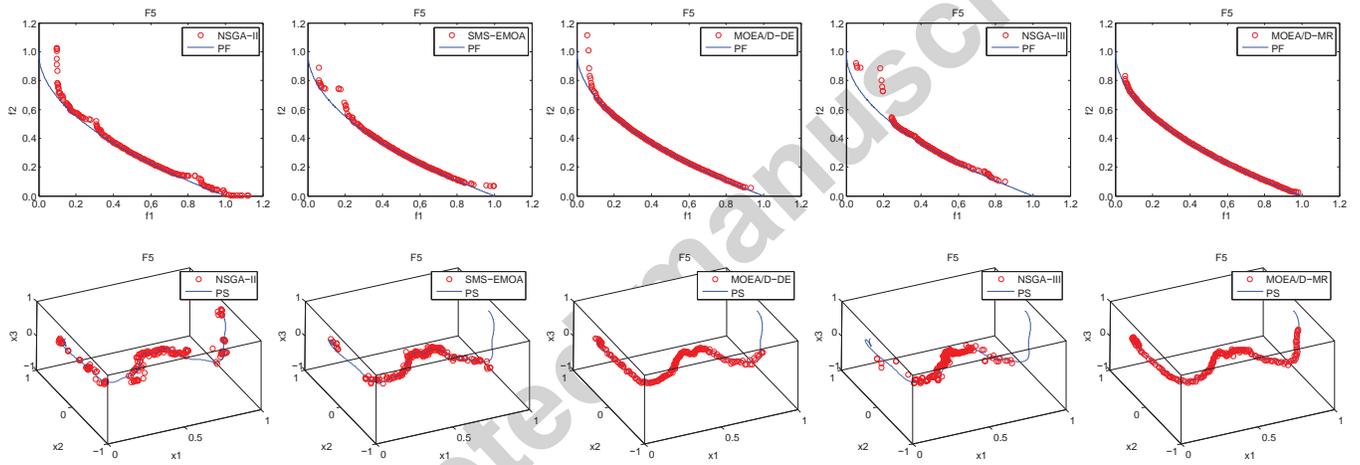


Fig. 8. Plots of the final solutions with the median IGD-metric values found by MOEA/D-MR and four other MOEAs in 30 runs in the objective space (the top row) and the $x_1 - x_2 - x_3$ space (the bottom row) on F5.

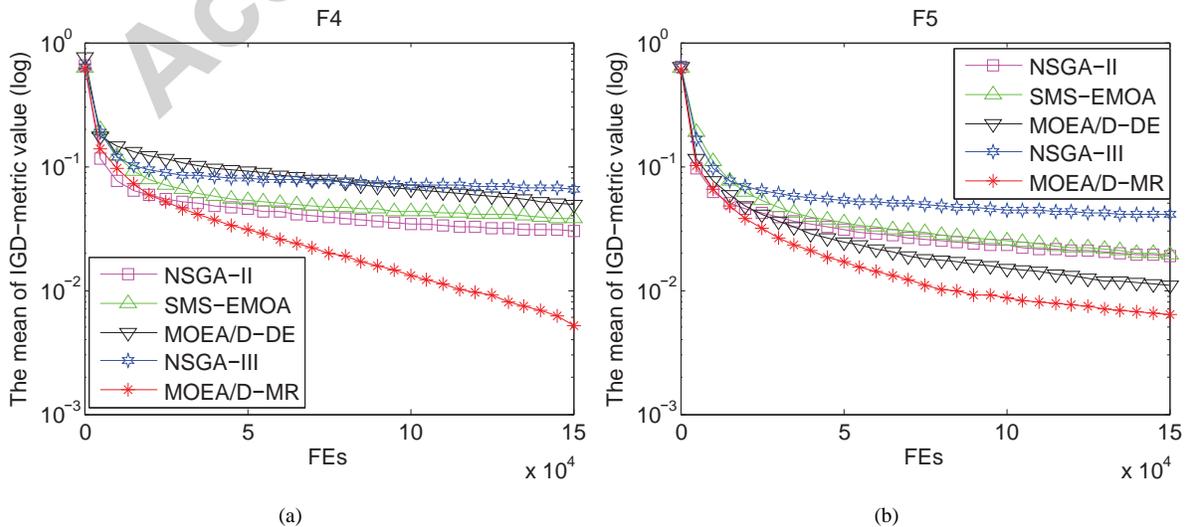


Fig. 9. Evolution of the mean IGD-metric values of MOEA/D-MR and four other MOEAs on F4 and F5 during the evolutionary process.

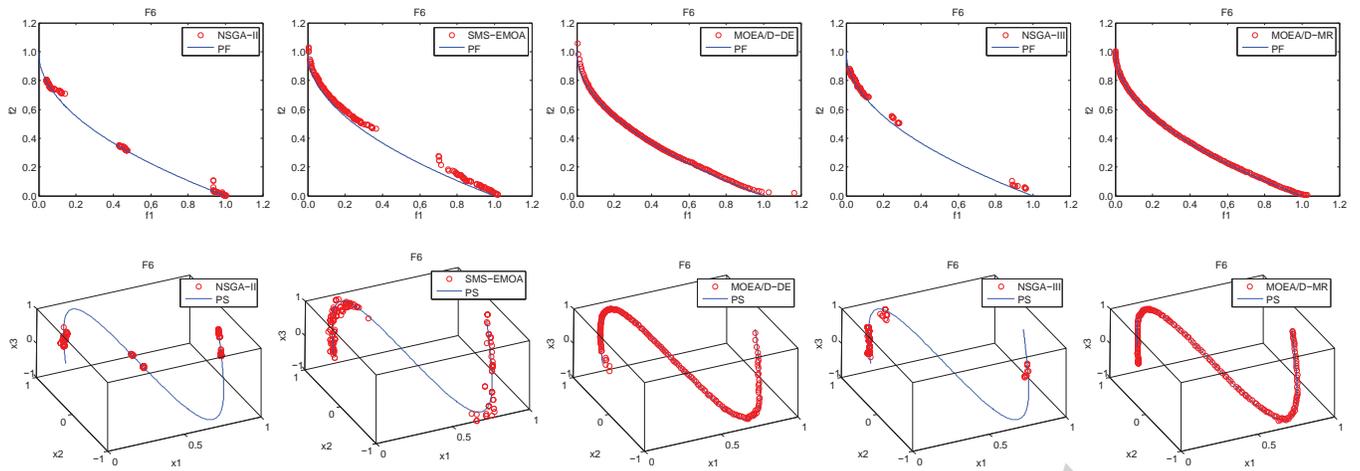


Fig. 10. Plots of the final solutions with the median IGD-metric values found by MOEA/D-MR and four other MOEAs in 30 runs in the objective space (the top row) and the $x_1 - x_2 - x_3$ space (the bottom row) on F6.

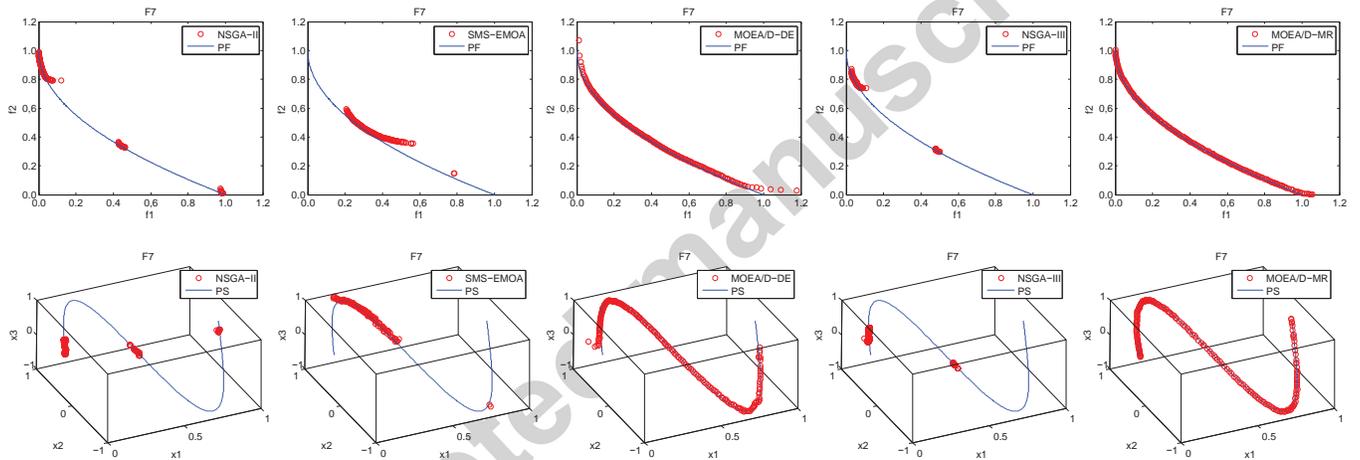


Fig. 11. Plots of the final solutions with the median IGD-metric values found by MOEA/D-MR and four other MOEAs in 30 runs in the objective space (the top row) and the $x_1 - x_2 - x_3$ space (the bottom row) on F7.

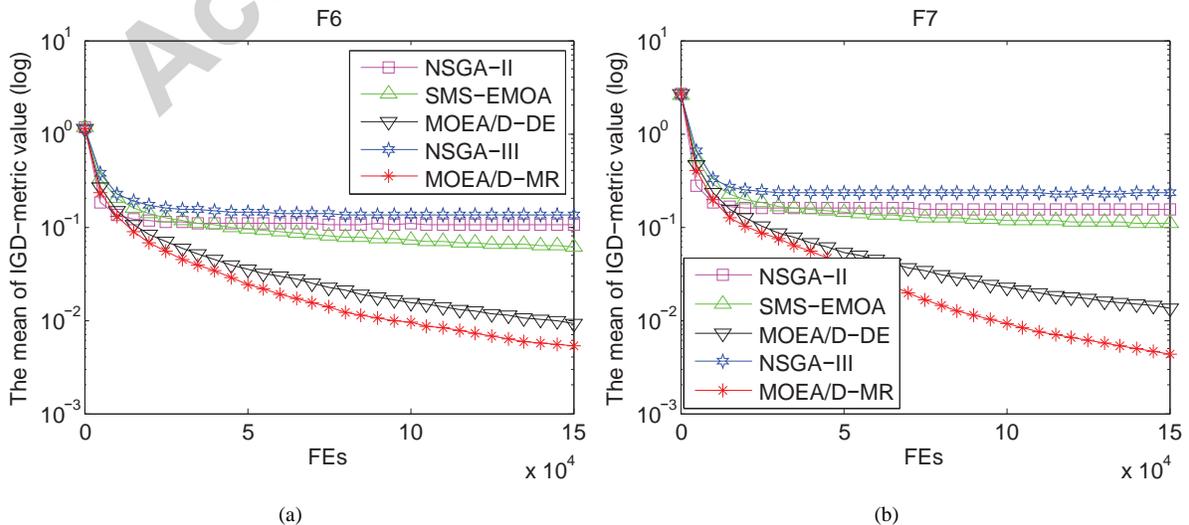


Fig. 12. Evolution of the mean IGD-metric values of MOEA/D-MR and four other MOEAs on F6 and F7 during the evolutionary process.

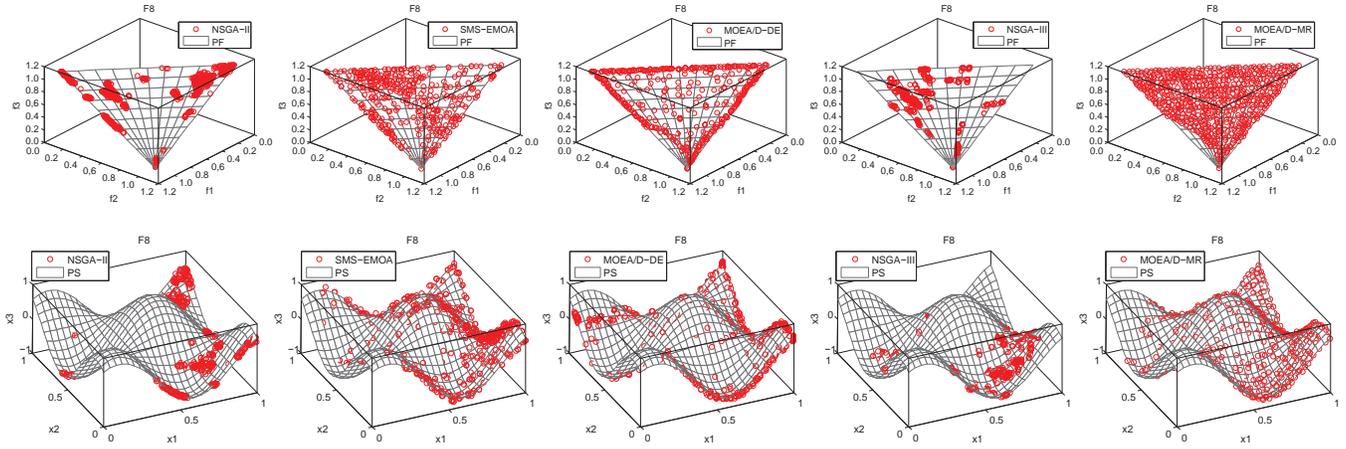


Fig. 13. Plots of the final solutions with the median IGD-metric values found by MOEA/D-MR and four other MOEAs in 30 runs in the objective space (the top row) and the $x_1 - x_2 - x_3$ space (the bottom row) on F8.

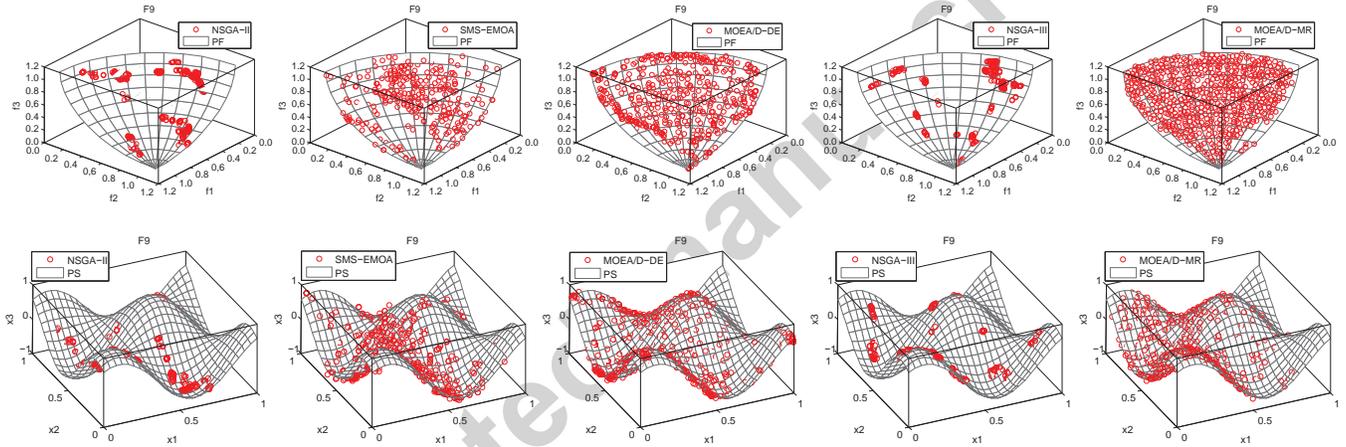


Fig. 14. Plots of the final solutions with the median IGD-metric values found by MOEA/D-MR and four other MOEAs in 30 runs in the objective space (the top row) and the $x_1 - x_2 - x_3$ space (the bottom row) on F9.

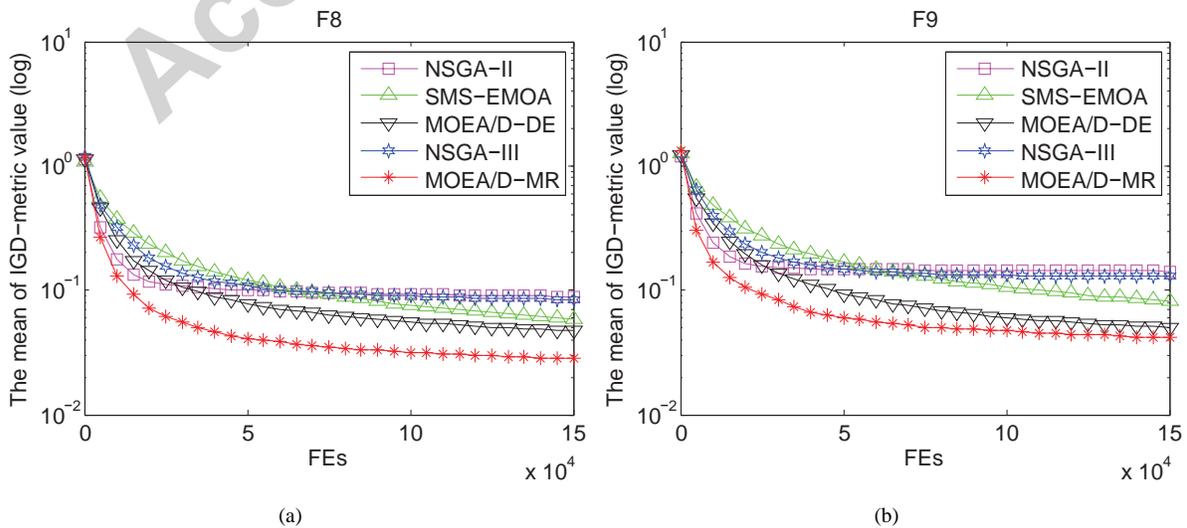


Fig. 15. Evolution of the mean IGD-metric values of MOEA/D-MR and four other MOEAs on F8 and F9 during the evolutionary process.

TABLE IV
PERFORMANCE COMPARISONS OF MOEA/D-MR AND ITS TWO VARIANTS

Instance	IGD						I_H					
	MOEA/D-MR		Variant-ideal		Variant-nadir		MOEA/D-MR		Variant-ideal		Variant-nadir	
	median	best	median	best	median	best	median	best	median	best	median	best
F1	2.03E-2	1.85E-2	2.35E-2	1.99E-2	1.84E-2	1.77E-2	9.9727	9.9876	9.9937	10.0165	9.9768	9.9843
F2	3.51E-3	3.02E-3	4.11E-3	3.51E-3	3.10E-3	2.93E-3	1.3858	1.3869	1.3864	1.3871	1.3863	1.3870
F3	3.75E-3	3.43E-3	4.68E-3	3.63E-3	4.56E-3	3.66E-3	0.9332	0.9344	0.9330	0.9345	0.9310	0.9329
F4	3.85E-3	2.93E-3	1.77E-2	3.89E-3	4.86E-2	2.52E-2	1.0972	1.1003	1.0324	1.0967	0.9824	1.0640
F5	5.91E-3	4.06E-3	7.60E-3	5.32E-3	2.99E-2	1.64E-2	1.0873	1.0978	1.0653	1.0876	1.0148	1.0714
F6	3.70E-3	3.07E-3	5.29E-3	3.60E-3	8.68E-2	2.04E-2	1.0995	1.1008	1.0895	1.0968	1.0086	1.0644
F7	3.91E-3	3.48E-3	7.55E-3	3.42E-3	8.93E-2	2.25E-2	1.0989	1.0998	1.0840	1.0978	0.9815	1.0483
F8	2.82E-2	2.60E-2	3.50E-2	3.06E-2	4.46E-2	2.67E-2	0.5373	0.5411	0.5303	0.5394	0.5152	0.5305
F9	4.04E-2	3.39E-2	4.89E-2	3.93E-2	1.46E-1	1.04E-1	1.0568	1.0608	1.0420	1.0505	1.0061	1.0342

TABLE V
MEAN VALUE RANKS AND STATISTICAL SIGNIFICANCE TESTS RESULTS OF MOEA/D-MR AND ITS TWO VARIANTS

Instance	IGD			I_H		
	MOEA/D-MR	Variant-ideal	Variant-nadir	MOEA/D-MR	Variant-ideal	Variant-nadir
F1	2	3-	1+	3	1+	2+
F2	2	3-	1+	3	1+	2+
F3	1	3-	2-	1	2 \approx	3-
F4	1	2-	3-	1	2-	3-
F5	1	2-	3-	1	2-	3-
F6	1	2-	3-	1	2-	3-
F7	1	2-	3-	1	2-	3-
F8	1	2-	3-	1	2-	3-
F9	1	2-	3-	1	2-	3-

+ , - and \approx denote that the performance of the corresponding algorithm is significantly better than, worse than, and similar to MOEA/D-MR respectively by Wilcoxon's rank sum test with $\alpha = 0.05$.

D. Multiple Reference Points vs Single Reference Point

To further demonstrate the effectiveness on the use of multiple reference points in MOEA/D, we compare MOEA/D-MR with its two variants: Variant-ideal and Variant-nadir, which only use z^{id} and z^{na} as its reference point, respectively. The parameters settings in both variants are the same as in MOEA/D-MR described in Section V-B. The experimental results on the IGD-metric and the I_H -metric are given in Table IV and Table V. It is evident that MOEA/D-MR significantly outperforms both Variant-ideal and Variant-nadir.

- On F3-F9, MOEA/D-MR performs remarkably better than two variants in terms of the median of the IGD-metric values and the I_H -metric values. As for the results in the best run with the minimal IGD-metric values or the I_H -metric values on F3-F9, MOEA/D-MR are also better than other two variants except F3 and F7. On F3, MOEA/D-MR has the best IGD value and the second best I_H value. On F7, MOEA/D-MR has the second best IGD value and the best I_H value. The the Wilcoxon's rank sum tests in Table V also demonstrates that MOEA/D-MR is significantly better than its two variants on F3-F9.
- On F1 and F2, the performances of three algorithms are similar. Variant-nadir achieves the best IGD values while Variant-ideal gains the best I_H values. MOEA/D-MR is slightly worse than them in terms of I_H metric values. But it still can yield the second best IGD results.

Overall, the performance of MOEA/D-MR is remarkably better than its two variants. From these results, we can conclude that the use of multiple reference points can improve the performance of MOEA/D with single reference point.

E. MOEA/D-MR With the PBI Decomposition

Apart from the Tchebycheff approach, the PBI decomposition can also serve in MOEA/D-MR. With the normalization of Eq. (8), the subproblem defined by the PBI decomposition can be formulated as follows:

$$\begin{aligned} & \text{minimize} && \mathbf{g}(x|w, z^*) = d_1 + \theta d_2 \\ & \text{subject to} && x \in \Omega \end{aligned} \quad (13)$$

where

$$d_1 = \bar{F}(x|z^*)^T w / \|w\|, \quad d_2 = \|\bar{F}(x|z^*) - d_1 \frac{w}{\|w\|}\|.$$

z^* can be either z^{id} or z^{na} . When z^* is set to be z^{id} , Eq. (13) is the same as the original PBI decomposition [3]. In [17], z^* is set to z^{na} . In this case, Eq. (13) is called the IPBI decomposition. θ is a penalty parameter. It should be noted that the

TABLE VI
IGD-METRIC VALUES OBTAINED BY MOEA/D-MR AND ITS FOUR PBI VARIANTS ON F1-F9

IGD	MOEA/D-MR		MOEA/D-MR-PBI-0		MOEA/D-MR-PBI-5		MOEA/D-MR-PBI-20		MOEA/D-MR-PBI-100					
	median	best	median	best	median	best	median	best	median	best				
F1	2.03E-2	1.85E-2	—	1.57	7.64E-1	—	8.32E-2	4.15E-2	—	2.91E-2	2.41E-2	—	3.68E-2	2.71E-2
F2	3.51E-3	3.02E-3	—	2.49E-1	2.09E-1	—	1.43E-1	9.19E-2	—	1.44E-1	1.14E-1	—	1.00E-1	8.51E-3
F3	3.75E-3	3.43E-3	—	1.09E-1	8.60E-2	—	2.07E-2	1.83E-2	—	1.01E-2	6.63E-3	—	6.66E-3	4.71E-3
F4	3.85E-3	2.93E-3	—	2.24E-1	1.51E-1	—	5.59E-2	3.07E-2	—	3.64E-2	1.33E-2	—	3.56E-2	1.56E-2
F5	5.91E-3	4.06E-3	—	1.88E-1	1.18E-1	—	2.50E-2	1.07E-2	—	1.69E-2	8.42E-3	—	1.89E-2	9.32E-3
F6	3.70E-3	3.07E-3	—	1.84E-1	9.65E-2	—	3.79E-2	1.31E-2	—	1.23E-2	5.62E-3	—	1.38E-2	6.68E-3
F7	3.91E-3	3.48E-3	—	3.14E-1	1.57E-1	—	4.99E-2	7.23E-3	—	2.08E-2	6.34E-3	—	3.47E-2	6.60E-3
F8	2.82E-2	2.60E-2	—	1.66E-1	8.59E-2	—	8.47E-2	5.37E-2	—	7.34E-2	4.96E-2	—	6.62E-2	4.71E-2
F9	4.04E-2	3.39E-2	—	2.85E-1	2.07E-1	—	2.32E-1	1.58E-1	—	2.31E-1	1.41E-1	—	2.01E-1	1.18E-1

+ , - and \approx denote that the performance of the corresponding algorithm is significantly better than, worse than, and similar to MOEA/D-MR respectively by Wilcoxon's rank sum test with $\alpha = 0.05$.

TABLE VII
 I_H -METRIC VALUES OBTAINED BY MOEA/D-MR AND ITS FOUR PBI VARIANTS ON F1-F9

IGD	MOEA/D-MR		MOEA/D-MR-PBI-0		MOEA/D-MR-PBI-5		MOEA/D-MR-PBI-20		MOEA/D-MR-PBI-100					
	median	best	median	best	median	best	median	best	median	best				
F1	9.9727	9.9876	—	8.0156	8.7567	—	9.9016	9.9333	—	9.9672	9.9874	—	9.9640	9.9854
F2	1.3858	1.3869	—	1.3348	1.3505	—	1.3584	1.3612	—	1.3581	1.3634	—	1.3612	1.3816
F3	0.9331	0.9344	—	0.8244	0.8463	—	0.9269	0.9289	—	0.9302	0.9320	—	0.9298	0.9323
F4	1.0972	1.1003	—	0.7485	0.8119	—	0.9626	1.0528	—	0.9977	1.0793	—	0.9981	1.0697
F5	1.0873	1.0978	—	0.7716	0.8367	—	1.0254	1.0680	—	1.0402	1.0790	—	1.0294	1.0780
F6	1.0995	1.1008	—	0.8225	0.9223	—	1.0214	1.0633	—	1.0734	1.0892	—	1.0700	1.0851
F7	1.0989	1.0998	—	0.6666	0.9050	—	1.0183	1.0844	—	1.0474	1.0928	—	1.0328	1.0888
F8	0.5373	0.5411	—	0.3316	0.4051	—	0.4391	0.4977	—	0.4591	0.5009	—	0.4650	0.5093
F9	1.0568	1.0608	—	0.7204	0.8128	—	0.8268	0.8704	—	0.8274	0.9030	—	0.8534	0.9314

+ , - and \approx denote that the performance of the corresponding algorithm is significantly better than, worse than, and similar to MOEA/D-MR respectively by Wilcoxon's rank sum test with $\alpha = 0.05$.

definitions of w and v for each subproblem is different from those in Eq. (4). Actually, both vectors are the same in the PBI decomposition, i.e., $w = v$.

Since the performance of the PBI decomposition based algorithm is sensitive to the values of θ as pointed out in [43], we test four PBI versions of MOEA/D-MR with different θ values, MOEA/D-MR-PBI-0, MOEA/D-MR-PBI-5, MOEA/D-MR-PBI-20 and MOEA/D-MR-PBI-100, and their θ values are set to be 0, 5, 20 and 100, respectively. The IGD and I_H comparison results of MOEA/D-MR and its four PBI versions are shown in Table VI and Table VII, respectively. It is clear that all the PBI decomposition variants are worse than the Tchebycheff decomposition one on all nine test problems. These results also illustrate that a larger θ value can make MOEA/D-MR-PBI performs better on F1-F9. The main reason is that some subproblems in MOEA/D-MR-PBI will have the same optimal solution, and a larger θ value can reduce this overlap. Although the PBI decomposition works not as well as the Tchebycheff decomposition in MOEA/D-MR on F1-F9, it is able to solve MOPs with other characteristics [17], [25].

VI. CONCLUSION

In this paper, we studied and analyzed the possible benefits of MOEA/D with two reference points. We observed and argued that the ideal point and the nadir point can complement each other. A new MOEA/D version denoted as MOEA/D-MR, which uses both the ideal point and the nadir point as the reference points, was designed. MOEA/D-MR was compared with four other state-of-the-art MOEAs on a set of newly designed difficult test problems. The comparison results showed that our proposed MOEA/D-MR works well and is competitive.

Actually, there are still some challenging and promising research directions on MOEA/D with multiple reference points.

- Use of global and local reference points: It may be useful on many-objective optimization. Global reference points can help to give a global picture of the PF and local ones can help to refine the search for some areas preferred by the decision maker. In general, the use of multiple reference points (i.e., more than two) is an interesting future research direction.
- The size of the subproblem set for each reference point: It could be adaptively adjusted for improving the algorithm performance.
- The search effort: Reference points can be used to guide how to distribute the search effort.

REFERENCES

- [1] H. Ishibuchi and T. Murata, "A multi-objective genetic local search algorithm and its application to flowshop scheduling," *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews*, vol. 28, no. 3, pp. 392-403, 1998.

- [2] T. Murata, H. Ishibuchi, and M. Gen, "Specification of genetic search directions in cellular multi-objective genetic algorithms," in *Evolutionary Multi-Criterion Optimization (EMO)*, vol. LNCS 1993. Springer, 2001, pp. 82–95.
- [3] 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.
- [4] H. Li and Q. Zhang, "Comparison between NSGA-II and MOEA/D on a set of multiobjective optimization problems with complicated Pareto sets," *IEEE Transactions on Evolutionary Computation*, vol. 13, no. 2, pp. 284–302, 2009.
- [5] G. Yu, J. Zheng, R. Shen, and M. Li, "Decomposing the user-preference in multiobjective optimization," *Soft Computing*, pp. 1–17, 2015.
- [6] H.-L. Liu, F. Gu, and Q. Zhang, "Decomposition of a multiobjective optimization problem into a number of simple multiobjective subproblems," *IEEE Transactions on Evolutionary Computation*, vol. 18, no. 3, pp. 450–455, 2014.
- [7] J. Jahn, *Mathematical vector optimization in partially ordered linear spaces*. Lang Frankfurt/Main Bern, 1986.
- [8] R. T. Marler and J. S. Arora, "Survey of multi-objective optimization methods for engineering," *Structural and multidisciplinary optimization*, vol. 26, no. 6, pp. 369–395, 2004.
- [9] I. Giagkiozis, R. C. Purshouse, and P. J. Fleming, "Generalized decomposition," in *Evolutionary Multi-Criterion Optimization (EMO)*, vol. LNCS 7811. Springer, 2013, pp. 428–442.
- [10] H. Ishibuchi, Y. Sakane, N. Tsukamoto, and Y. Nojima, "Adaptation of scalarizing functions in MOEA/D: An adaptive scalarizing function-based multiobjective evolutionary algorithm," in *Evolutionary Multi-Criterion Optimization (EMO)*, 2009. Springer, 2009, pp. 438–452.
- [11] —, "Simultaneous use of different scalarizing functions in MOEA/D," in *Proceedings of the 12th annual conference on Genetic and evolutionary computation*. ACM, 2010, pp. 519–526.
- [12] Y. Qi, X. Ma, F. Liu, L. Jiao, J. Sun, and J. Wu, "MOEA/D with adaptive weight adjustment," *Evolutionary computation*, vol. 22, no. 2, pp. 231–264, 2014.
- [13] H. Li and D. Landa-Silva, "An adaptive evolutionary multi-objective approach based on simulated annealing," *Evolutionary Computation*, vol. 19, no. 4, pp. 561–595, 2011.
- [14] F. Gu, H.-L. Liu, and K. C. Tan, "A multiobjective evolutionary algorithm using dynamic weight design method," *International Journal of Innovative Computing, Information and Control*, vol. 8, no. 5(B), pp. 3677–3688, 2012.
- [15] H. Ishibuchi, K. Doi, H. Masuda, and Y. Nojima, "Relation between weight vectors and solutions in MOEA/D," in *Proc. of 2015 IEEE Symposium Series on Computational Intelligence in Multi-Criteria Decision-Making*. IEEE, 2015, pp. 861–868.
- [16] Q. Zhang, H. Li, D. Maringer, and E. Tsang, "MOEA/D with NBI-style tchebycheff approach for portfolio management," in *IEEE Congress on Evolutionary Computation (CEC) 2010*. IEEE, 2010, pp. 1–8.
- [17] H. Sato, "Inverted PBI in MOEA/D and its impact on the search performance on multi and many-objective optimization," in *Proceedings of the 2014 conference on Genetic and evolutionary computation*. ACM, 2014, pp. 645–652.
- [18] S. Jiang and S. Yang, "An improved multiobjective optimization evolutionary algorithm based on decomposition for complex pareto fronts," *IEEE Transactions on Cybernetics*, vol. 46, no. 2, pp. 421–437, 2016.
- [19] R. Saborido, A. B. Ruiz, and M. Luque, "Global WASF-GA: An evolutionary algorithm in multiobjective optimization to approximate the whole Pareto optimal front," *Evolutionary Computation*, (in press).
- [20] Z. Wang, Q. Zhang, M. Gong, and A. Zhou, "A replacement strategy for balancing convergence and diversity in MOEA/D," in *IEEE Congress on Evolutionary Computation (CEC), 2014*. IEEE, 2014, pp. 2132–2139.
- [21] A. Mohammadi, M. N. Omidvar, and X. Li, "Reference point based multi-objective optimization through decomposition," in *IEEE Congress on Evolutionary Computation (CEC), 2012*. IEEE, 2012, pp. 1–8.
- [22] A. Mohammadi, M. N. Omidvar, X. Li, and K. Deb, "Integrating user preferences and decomposition methods for many-objective optimization," in *IEEE Congress on Evolutionary Computation (CEC), 2014*. IEEE, 2014, pp. 421–428.
- [23] I. Giagkiozis, R. C. Purshouse, and P. J. Fleming, "Towards understanding the cost of adaptation in decomposition-based optimization algorithms," in *Proc. of 2013 IEEE International Conference on Systems, Man, and Cybernetics (SMC 2013)*. IEEE, 2013, pp. 615–620.
- [24] H. Jain and K. Deb, "An evolutionary many-objective optimization algorithm using reference-point based nondominated sorting approach, Part II: Handling constraints and extending to an adaptive approach," *IEEE Transactions on Evolutionary Computation*, vol. 18, no. 4, pp. 602–622, 2014.
- [25] 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*, 2016, in press.
- [26] H. Ishibuchi, Y. Sakane, N. Tsukamoto, and Y. Nojima, "Effects of using two neighborhood structures on the performance of cellular evolutionary algorithms for many-objective optimization," in *IEEE Congress on Evolutionary Computation, 2009. (CEC)*. IEEE, 2009, pp. 2508–2515.
- [27] H. Ishibuchi, N. Akedo, and Y. Nojima, "Relation between neighborhood size and MOEA/D performance on many-objective problems," in *Evolutionary Multi-Criterion Optimization (EMO)*, vol. LNCS 7811. Springer, 2013, pp. 459–474.
- [28] K. Li, Q. Zhang, S. Kwong, M. Li, and R. Wang, "Stable matching based selection in evolutionary multiobjective optimization," *IEEE Transactions on Evolutionary Computation*, vol. 18, no. 6, pp. 909–923, 2014.
- [29] Z. Wang, Q. Zhang, A. Zhou, M. Gong, and L. Jiao, "Adaptive replacement strategies for MOEA/D," *IEEE Transactions on Cybernetics*, vol. 46, no. 2, pp. 474–486, 2016.
- [30] Y. Yuan, H. Xu, B. Wang, B. Zhang, and X. Yao, "Balancing convergence and diversity in decomposition-based many-objective optimizers," *IEEE Transactions on Evolutionary Computation*, vol. 20, no. 2, pp. 180–198, 2016.
- [31] K. Li, S. Kwong, Q. Zhang, and K. Deb, "Interrelationship-based selection for decomposition multiobjective optimization," *IEEE Transactions on Cybernetics*, vol. 45, no. 10, pp. 2076–2088, 2015.
- [32] E. Zitzler, M. Laumanns, and L. Thiele, "SPEA2: Improving the Strength Pareto Evolutionary Algorithm for Multiobjective Optimization," in *Evolutionary Methods for Design, Optimization and Control with Applications to Industrial Problems (EUROGEN 2001)*. K. Giannakoglou, et al., Eds. International Center for Numerical Methods in Engineering (CIMNE), 2001, pp. 95–100.
- [33] E. Zitzler, L. Thiele, M. Laumanns, C. M. Fonseca, and V. G. Da Fonseca, "Performance assessment of multiobjective optimizers: An analysis and review," *IEEE Transactions on Evolutionary Computation*, vol. 7, no. 2, pp. 117–132, 2003.
- [34] M. R. Sierra and C. A. C. Coello, "Improving PSO-based multi-objective optimization using crowding, mutation and ϵ -dominance," in *International Conference on Evolutionary Multi-Criterion Optimization*. Springer, 2005, pp. 505–519.
- [35] E. Zitzler and L. Thiele, "Multiobjective evolutionary algorithms: A comparative case study and the strength pareto approach," *IEEE Transactions on Evolutionary Computation*, vol. 3, no. 4, pp. 257–271, 1999.
- [36] H.-L. Liu, F. Gu, and Y. Cheung, "T-MOEA/D: MOEA/D with objective transform in multi-objective problems," in *2010 International Conference of Information Science and Management Engineering (ISME)*, vol. 2. IEEE, 2010, pp. 282–285.
- [37] S. Huband, P. Hingston, L. Barone, and L. While, "A review of multiobjective test problems and a scalable test problem toolkit," *IEEE Transactions on Evolutionary Computation*, vol. 10, no. 5, pp. 477–506, 2006.
- [38] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: NSGA-II," *IEEE Transactions on Evolutionary Computation*, vol. 6, no. 2, pp. 182–197, 2002.
- [39] N. Beume, B. Naujoks, and M. Emmerich, "SMS-EMOA: Multiobjective selection based on dominated hypervolume," *European Journal of Operational Research*, vol. 181, no. 3, pp. 1653–1669, 2007.
- [40] 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, 2014.

- [41] K. Deb and M. Goyal, "A combined genetic adaptive search (GeneAS) for engineering design," *Computer Science and Informatics*, vol. 26, pp. 30–45, 1996.
- [42] F. Wilcoxon, "Individual comparisons by ranking methods," *Biometrics bulletin*, vol. 1, no. 6, pp. 80–83, 1945.
- [43] A. Mohammadi, M. N. Omidvar, X. Li, and K. Deb, "Sensitivity analysis of penalty-based boundary intersection on aggregation-based emo algorithms," in *IEEE Congress on Evolutionary Computation (CEC), 2015*. IEEE, 2015, pp. 2891–2898.

Accepted manuscript