

Adaptively Allocating Search Effort in Challenging Many-Objective Optimization Problems

Hai-Lin Liu, Lei Chen, Qingfu Zhang, *Fellow, IEEE* and Kalyanmoy Deb, *Fellow, IEEE*

Abstract—An effective allocation of search effort is important in multi-objective optimization, particularly in many-objective optimization problems. This paper presents a new adaptive search effort allocation strategy for MOEA/D-M2M, a recent MOEA/D algorithm for challenging Many-Objective Optimization Problems (MaOPs). This proposed method adaptively adjusts the subregions of its subproblems by detecting the importance of different objectives in an adaptive manner. More specifically, it periodically resets the subregion setting based on the distribution of the current solutions in the objective space such that the search effort is not wasted on unpromising regions. The basic idea is that the current population can be regarded as an approximation to the Pareto front (PF) and thus one can implicitly estimate the shape of the PF and such estimation can be used for adjusting the search focus. The performance of proposed algorithm has been verified by comparing it with eight representative and competitive algorithms on a set of degenerated many-objective optimization problems with disconnected and connected PFs. Performances of the proposed algorithm on a number of non-degenerated test instances with connected and disconnected PFs are also studied.

Index Terms—Many-objective optimization, MOEA/D, evolutionary algorithm, adaptive allocation.

I. INTRODUCTION

Multiobjective objective optimization problems with more than four objectives, often referred to as many-objective optimization problems (MaOPs), have recently attracted growing research interest from the evolutionary computation community due to its wide applications and importance [1]–[5]. Decomposition based evolutionary multiobjective optimization algorithms have been regarded as a very promising approach for dealing with many objectives. Multiobjective evolutionary algorithm based on decomposition (MOEA/D) is a popular decomposition multiobjective optimization algorithmic framework [6]–[9]. It decomposes a multiobjective optimization problem into a number of subproblems, and each subproblem can be a single objective optimization problem [6], [7] or a

This work was supported in part by the National Natural Science Foundation of China under Grant 61673121, in part by the Projects of Science and Technology of Guangzhou under Grant 201508010008, in part by the China Scholarship Council, and in part by a grant from ANR/RCC Joint Research Scheme sponsored by the Research Grants Council of the Hong Kong Special Administrative Region, China and France National Research Agency (Project No. A-CityU101/16).

H.-L. Liu and L. Chen are with Guangdong University of Technology, Guangzhou, China, e-mail: hliu@gdut.edu.cn, chenaction@126.com.

Q. Zhang is with the Department of Computer Science, City University of Hong Kong, Hong Kong, email: qingfu.zhang@cityu.edu.hk

K. Deb is with Department of Electrical and Computer Engineering and BEACON Center for the Study of Evolution in Action, Michigan State University, 428 S. Shaw Lane, 2120 EB, East Lansing, MI 48824, USA, e-mail: kdeb@msu.edu.

multiobjective optimization one [8], [9]. MOEA/D solves these subproblems in a single run and exploits correlation among different subproblems for promoting its search efficiency.

Under regularity conditions, the dimensionality of the PF is one less than the number of objectives in a continuous MaOP [10], [11]. Objectives in many real-life MaOPs, however, can be highly correlated with one another and thus some objectives are redundant. The PFs of those MaOPs can be of very low dimensionality, and we call such MaOPs degenerated in this paper. For a degenerated MaOP, one can find its essential objectives in an online or offline manner and then turn it into a problem with few of objectives as in [12]–[20]. Many machine learning algorithms, such as feature selection [12], [13], Maximum Variance Unfolding (MVU) [14] and Principal Component Analysis (PCA) [15]–[17] have been used to reduce the number of objectives. However, when there is no redundant objectives, these objective reduction strategies may waste a lot of search effort.

MOEA/D algorithms [21], [22] often use a set of weight vectors for their decomposition. In the case when no prior knowledge of the PF is available, most MOEA/D algorithms choose a set of uniform distributed weight vectors. However, PF shapes may vary from problem to problem, some effort has been made to try to address this issue by adaptively adjusting weight vectors [23]–[27]. To the best of our knowledge, the existing adaptive strategies in MOEA/D framework cannot handle degenerated problems, in which each subproblem considers a single scalarized objective function.

In MOEA/D-M2M [8], a very recent variant of MOEA/D, each subproblem is multiobjective and has a different feasible region (subregion), and its Pareto solution set is part of the Pareto solution set of the original problem. The feasible region of each subproblem is defined by a direction vector (or reference vector) in the objective space and thus different objectives are not of equal importance. In order to make the best use of the search effort, we propose a new variant of MOEA/D-M2M, called MOEA/D-AM2M. To do search effort allocation effectively, this algorithm adaptively adjusts the feasible region of each subproblem, using information collected during the search. More specifically, it periodically resets the subregion setting based on the distribution of the current solutions in the objective space such that the search effort will not be wasted on unpromising regions. The basic idea is that the current population can be regarded as an approximation to the PF and thus one can implicitly estimate the shape of the PF and such estimation can be used for adjusting the search focus. To verify the effectiveness of our proposed algorithm, a set of degenerated MaOPs with disconnected

PFs and non-degenerated MaOPs with disconnected PFs are constructed for experimental studies. The proposed algorithm has been compared with MOEA/D-DE [7], NSGA-III [28], MOEA/D-AWA [27], MOEA/D-M2M [8], GrEA [29], NL-MVU-PCA [14], HypE [30] and RVEA [23]. In addition, the performances of the suggested algorithm on non-degenerated MaOPs are also studied.

The remainder of this paper is organized as follows: Section II gives the definition of the discussed MaOPs and describes how MOEA/D-M2M works; Section III discusses the adaptive subregion division and adaptive weight setting in detail, and then presents the proposed MOEA/D-AM2M framework. In Section IV, we introduce how the challenging MaOPs are constructed. A series of experiments are conducted, and the simulation results are analyzed in Section V. Section VI concludes this paper and presents some future work.

II. PRELIMINARIES

In this section, we first introduce some basic definitions related to multi-objective optimization problems, and then give the details of MOEA/D-M2M algorithm, which decomposes an MOP (MaOP) into a set of simple multi-objective (many-objective) optimization subproblems [8].

A. Problem Definition

A continuous multi-objective optimization problem can be stated as follows:

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

where \mathbf{D} , the *decision (variable) space*, is a closed and connected region in \mathbf{R}^n , $F : \mathbf{D} \rightarrow \mathbf{R}^m$ consists of m real-valued connected objective functions f_1, \dots, f_m . \mathbf{R}^m is called the objective space. When $m \leq 4$, it is called a many-objective optimization problem (MaOP) [1].

Let $\mathbf{u} = (u_1, \dots, u_m)$ and $\mathbf{v} = (v_1, \dots, v_m)$, \mathbf{u} is said to dominate \mathbf{v} if $u_i \leq v_i$ for all $i = 1, \dots, m$, and $\mathbf{u} \neq \mathbf{v}$. A point \mathbf{x}^* is called Pareto optimal if there is no $\mathbf{x} \in \mathbf{D}$ such that $F(\mathbf{x})$ dominates $F(\mathbf{x}^*)$. The set of all the Pareto optimal points, denoted by PS, is called the Pareto set. The set of all the Pareto objective vectors, $\text{PF} = \{F(\mathbf{x}) \in \mathbf{R}^m | \mathbf{x} \in \text{PS}\}$, is called the Pareto Front (PF) [11]. $\mathbf{u}^* = (u_1^*, \dots, u_m^*)$ is called the ideal point if u_i^* is the minimal value of $f_i(\mathbf{x})$ over the decision space. $\mathbf{z}^{nad} = (z_1^{nad}, \dots, z_m^{nad})$ is called the nadir point if z_i^{nad} is the maximal value of $f_i(\mathbf{x})$ over the PS.

B. MOEA/D-M2M

Without loss of generality, we assume that all the objective functions are positive. MOEA/D-M2M [8] requires K unit direction vectors $\mathbf{v}^1, \dots, \mathbf{v}^K$ in \mathbf{R}_+^m . It divides \mathbf{R}_+^m into K subregions $\Omega_1, \dots, \Omega_K$, where $\Omega_k (k = 1, \dots, K)$ is:

$$\Omega_k = \{\mathbf{u} \in \mathbf{R}_+^m | \langle \mathbf{u}, \mathbf{v}^k \rangle \leq \langle \mathbf{u}, \mathbf{v}^j \rangle \text{ for any } j = 1, \dots, K\}, \quad (2)$$

where $\langle \mathbf{u}, \mathbf{v}^j \rangle$ is the acute angle between \mathbf{u} and \mathbf{v}^j . In other words, \mathbf{u} belongs to Ω_k if and only if \mathbf{v}^k has the smallest angle to \mathbf{u} among all the K direction vectors. Based on this division,

(1) can be transformed into K constrained multiobjective optimization subproblems. Subproblem k is:

$$\begin{aligned} & \text{minimize } F(\mathbf{x}) = (f_1(\mathbf{x}), \dots, f_m(\mathbf{x})), \\ & \text{subject to } F(\mathbf{x}) \in \Omega_k. \end{aligned} \quad (3)$$

It is obvious that a Pareto optimal solution to (1) is Pareto optimal to one subproblem. However, some Pareto optimal solutions to a subproblem may not be Pareto optimal to (1). MOEA/D-M2M optimizes these K subproblems in a collaborative way. During its search process, it maintains and evolves K subpopulations: $\mathbf{P}_1, \dots, \mathbf{P}_K$, where $\mathbf{P}_k (k = 1, \dots, K)$ is to approximate the PF of subproblem k .

$$\mathbf{P} = \bigcup_{k=1}^K \mathbf{P}_k$$

is the population in MOEA/D-M2M. The size of \mathbf{P}_k is denoted as N_k . Therefore, the whole population size

$$N = \sum_{k=1}^K N_k.$$

At each generation, MOEA/D-M2M updates \mathbf{P}_k as follows:

- 1) New Solution Generation: Generate \mathbf{Q} , a set of N new solutions.
- 2) Division: Divide \mathbf{Q} into $\mathbf{Q}_1, \dots, \mathbf{Q}_K$, where \mathbf{Q}_k contains all the new solutions in Ω_k .
- 3) Update: For each k , select N_k elements from $\mathbf{Q}_k \cup \mathbf{P}_k$ to replace \mathbf{P}_k .

In the above framework, the value of N_k determines how the search effort is distributed among different multiobjective subproblems. Roughly speaking, the algorithm allocates N_k/N of search effort to the k th multiobjective subproblem.

There are many ways for implementing Step 1 and Step 3. In this paper, we adopt the following schemes for doing so.

C. New Solution Generation

For each element $\mathbf{x} \in \mathbf{P}_k$, we generate its child solution \mathbf{z} as follows:

Step 1: Randomly select a solution \mathbf{y} with probability α from \mathbf{P}_k , and with probability $1 - \alpha$ from any other subpopulations.

Step 2: Conduct genetic operators on \mathbf{x} and \mathbf{y} to generate \mathbf{z} .

This scheme needs a control parameter α to specify genetic operators in Step 2. If another selected parent solution \mathbf{y} is from \mathbf{P}_k where \mathbf{x} is, \mathbf{y} should be close to \mathbf{x} since both \mathbf{x} and \mathbf{y} are from the same subpopulation, therefore, \mathbf{z} may be near to \mathbf{x} too. In this case, the solution generation operation mainly do exploitation. When \mathbf{y} is from another subpopulation, \mathbf{y} should not be very close to \mathbf{x} and neither is \mathbf{z} . Therefore, the operation focuses on exploration. For this reason, α may control the balance between exploration and exploitation. This is our motivation for introducing α . This parameter enables us to make fully use of subpopulation for local search, and, at the same time, maintains population's global search ability.

In Step 1 of MOEA/D-M2M, we can generate one child solution for each solution in \mathbf{P} and all these child solutions constitute \mathbf{Q} .

D. Update

Update is to do selection. Therefore, any selection methods such as Pareto dominance, indicator and decomposition based approaches can be used for this purpose. In this paper, we use the Tchebycheff decomposition approach [11]. More precisely, for each subproblem k that \mathbf{P}_k is for, we decompose it into N_k (i.e., the size of \mathbf{P}_k) single optimization subproblems. Each single objective subproblem (k, j) is defined as:

$$g^{k,j}(\mathbf{x}) = \max_{1 \leq i \leq m} \left(\frac{f_i(\mathbf{x}) - z_i}{w_i^{k,j}} \right), \quad (4)$$

subject to $\mathbf{x} \in \Omega_k$,

where $\mathbf{w}^{k,j} = (w_1^{k,j}, \dots, w_m^{k,j}) > 0$, $\sum_{i=1}^m w_i^{k,j} = 1$, $\mathbf{w}^{k,j}$ are in Ω_k , and z_i is the minimum value of the whole population \mathbf{P} for f_i . The distribution of N_k weight vectors plays a major role in guiding the search in Ω_k . In our approach, each solution $\mathbf{x}^{k,j}$ in \mathbf{P}_k is associated with single objective optimization subproblem (k, j) . It updates \mathbf{P}_k by using solutions in \mathbf{Q}_k as follows:

- 1) For each solution $\mathbf{z} \in \mathbf{Q}_k$. Do:

If there are solutions in \mathbf{P}_k that are poorer than \mathbf{z} in terms of their respective single objective functions, replace a randomly selected such one solution by \mathbf{z} .

- 2) For each subpopulation \mathbf{P}_k . Do:

If $|\mathbf{P}_k| < N_k$, randomly select $N_k - |\mathbf{P}_k|$ individuals from \mathbf{P} .

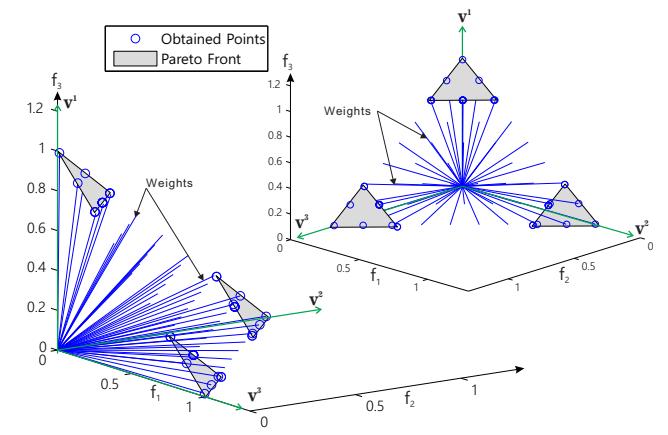
III. ADAPTIVE SUBREGION DIVISION AND WEIGHT VECTOR SETTING

In the case when there is no good a priori knowledge of an MOP (MaOP) with disconnected PF, a fixed subregion division (i.e., the setting of \mathbf{v}^k) and the weight setting for the update (i.e., $\mathbf{w}^{k,j}$) are not suitable [33] and much search effort will be unavoidably wasted since many single objective subproblems (4) will share the same optimal solutions as shown in Fig. 1(a). If the weight and direction vectors are designed according to the PF shape, we can obtain a set of well distributed Pareto optimal points on the PF (see Fig. 1(b)). Based on this consideration, we propose an adaptive subregion division and weight vector setting strategy. Our basic idea is that the population at each generation can be regarded as an approximation to the PF, and one can explicitly or implicitly estimate the PF shape and adjust subregion division and weight vectors such that their settings match the PF shape well.

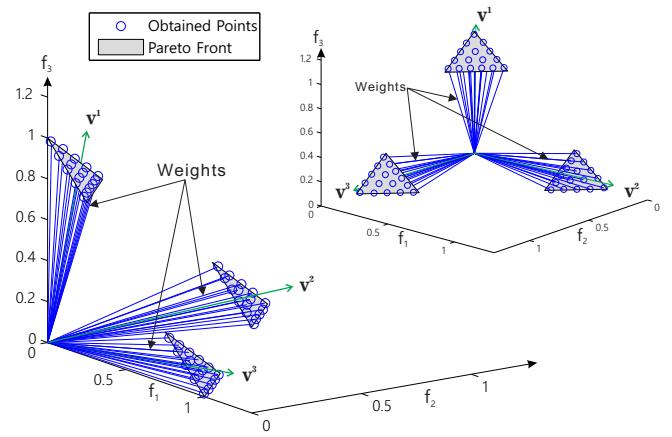
A. Adaptive Subregion Division

To reset the subregion division at a generation, we treat $\tilde{\mathbf{P}} = \mathbf{P} \cup \mathbf{Q}$ as an approximation to the PF. We reset the set of unit direction vectors as follows:

- 1) Set $\mathbf{V} = \emptyset$.
- 2) Randomly select a vector in $\tilde{\mathbf{P}}$, find a vector \mathbf{y} in $\tilde{\mathbf{P}}$ with the largest angle to the selected vector, delete \mathbf{y} from $\tilde{\mathbf{P}}$, and then add $\frac{\mathbf{y}}{\|\mathbf{y}\|}$ to \mathbf{V} .
- 3) Find a vector \mathbf{v} in $\tilde{\mathbf{P}}$ with the largest angle to \mathbf{V} , delete \mathbf{v} from $\tilde{\mathbf{P}}$, and then add $\frac{\mathbf{v}}{\|\mathbf{v}\|}$ to \mathbf{V} .



(a) Simulation results obtained by MOEA/D-M2M with evenly distributed direction/weight vectors



(b) Simulation results obtained by MOEA/D-M2M with designed direction/weight vectors

Fig. 1. Illustration of the direction/weight vectors' influence to the final obtained solutions.

- 4) If the size of \mathbf{V} is K , stop and return \mathbf{V} . Otherwise, go to 3).

The angle of a vector to a set is defined as the smallest one among the angles between this vector and all the vectors in the set.

This method uses the angle to measure the similarity among vectors. Other similarity measures can also be used. This method is able to select a set of representative vectors from $\tilde{\mathbf{P}}$. Therefore, the subregion division using their normalized vectors can focus the search on more promising areas (regions containing Pareto optimal solutions).

Fig. 2 illustrates how the proposed subregion division method generates three direction vectors. We first randomly select a vector \mathbf{x}^r from $\tilde{\mathbf{P}}$. The first direction vector \mathbf{v}^1 has the largest angle with \mathbf{x}^r . The second one \mathbf{v}^2 has the largest angle with \mathbf{v}^1 . \mathbf{v}^3 has the largest angle to $\{\mathbf{v}^1, \mathbf{v}^2\}$ among all the normalized vectors in $\tilde{\mathbf{P}}$.

To better understand how the adaptive subregions division method works, we illustrate the process of adaptive subregions design using a disconnected PF shown in Fig. 3. Suppose that the population is far from the PF at generation G , and

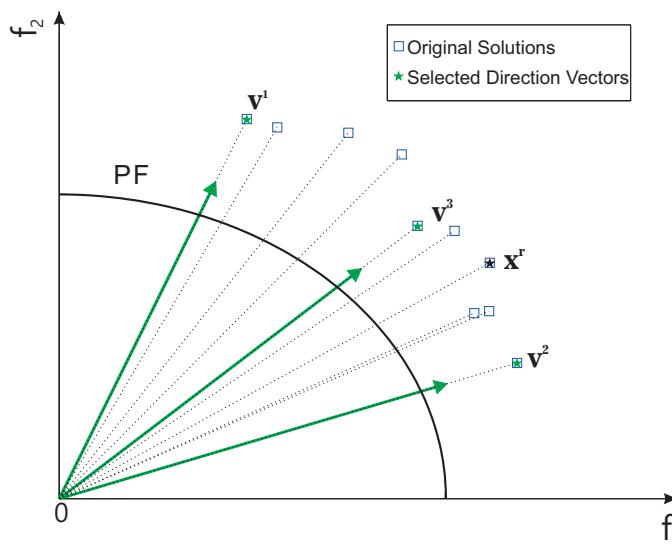


Fig. 2. Illustration of direction vectors design.

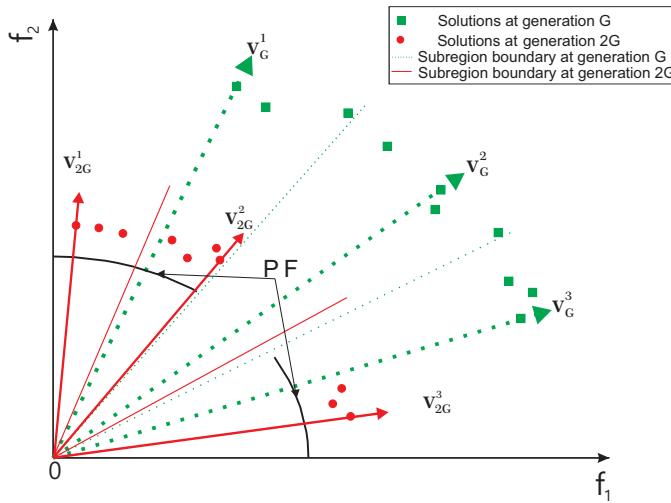


Fig. 3. Illustration of adaptive direction vectors design.

subregion Ω_2 centered by direction vector v_G^2 does not cover any part of PF. With the evolution of population, the proposed method will adaptively adjust the subregions according to the distribution of the new population. Fig. 3 shows how those subregions change from generation G to generation $2G$, where subregion Ω_2 at generation G (centered by direction vector v_G^2) is reset to the more promising area with part of PF (centered by direction vector v_{2G}^2) at generation $2G$.

B. Adaptive Weight Setting

$\tilde{\mathbf{P}}_k = \mathbf{P}_k \cup \mathbf{Q}_k$ is treated as an approximation to the PF in subregion $\Omega_k(k = 1, \dots, K)$. To reset the weight vectors, we first need to determine S_k , the number of weight vectors for subregion $\Omega_k(k = 1, \dots, K)$. In order to select the next generation solutions from the doubled population, we can averagely set S_k as $|\tilde{\mathbf{P}}_k|/2$ in each subregion. However, special care needs to be taken to handle subregions with an number of solutions. The detailed procedure of computing S_k is as follows:

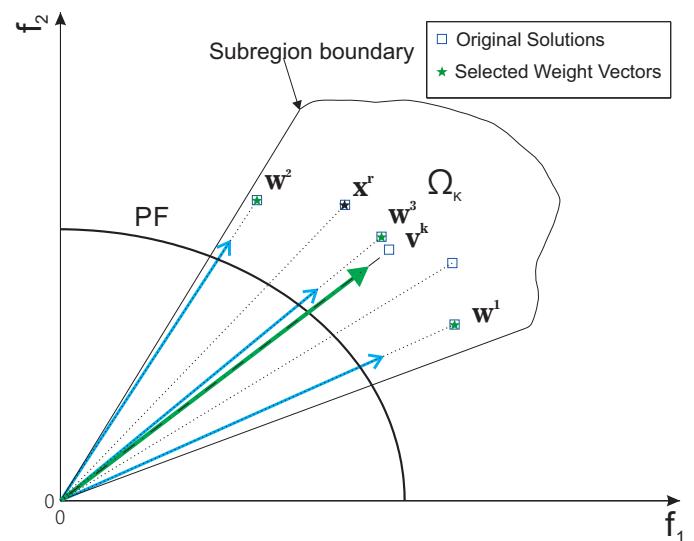


Fig. 4. Illustration of adaptive weight vectors design.

- 1) Set $\tilde{\mathbf{P}}_k = \{\mathbf{x} \in \Omega_k | \mathbf{x} \in \tilde{\mathbf{P}}\}$, $k = 1, \dots, K$.
- 2) If $|\tilde{\mathbf{P}}_k| = 1$, $S_k = 1$; Else $S_k = |\tilde{\mathbf{P}}_k|/2$.
- 3) If $\sum_{k=1}^K S_k = |\mathbf{P}|$, stop and return $S_k(k = 1, \dots, K)$.
- 4) Sort $S_k(k = 1, \dots, K)$ such that $0 < S_{k_t} \leq S_{k_{t+1}} \dots \leq S_{k_K}$. Do:
 $i = t$;
repeat
 $S_{k_i} = S_{k_i} + 1$;
 $i = i + 1$.
until $\sum_{k=1}^K S_k = |\mathbf{P}|$.
Stop and return $S_k(k = 1, \dots, K)$.
In 4), $S_{k_1} = S_{k_2} = \dots = S_{k_{t-1}} = 0$.

Once the number of weight vectors in each subregion is determined, we can reset the weights in each subregion using the same method in the above subsection.

Fig. 4 shows an example of how we generate weight vectors in subregion Ω_k . Just like the generation of direction vectors described above, we first randomly select a vector \mathbf{x}^r from $\tilde{\mathbf{P}}_k$. The first weight vector \mathbf{w}^1 has the largest angle with \mathbf{x}^r . The second one \mathbf{w}^2 has the largest angle with \mathbf{w}^1 . \mathbf{w}^3 has the largest angle to $\{\mathbf{w}^1, \mathbf{w}^2\}$ among all the normalized vectors in $\tilde{\mathbf{P}}_k$.

C. Main Framework of MOEA/D-AM2M

MOEA/D-AM2M mainly includes adaptive subregion division and adaptive weight setting. Based on the discussion above, we present the basic framework of MOEA/D-AM2M as follows (**Algorithm 1**):

We would like to make the following comments on the proposed algorithm:

- 1) The suggested method embeds MOEA/D-M2M with the proposed adaptive search effort allocation strategy, which can focus the search effort on the promising regions. The adaptive subregion division in MOEA/D-AM2M can quickly extend the population to the unexplored areas, and at the same time withdraw the search

Algorithm 1: MOEA/D-AM2M

Input :

- N : Population size;
- K : Number of the subproblems;
- G : Parameter update frequency;
- $MaxGen$: Maximum number of generations.

Output: a set of solutions.

Initialization: Uniformly initialize direction/weight vectors and population, use them to set subpopulation $\mathbf{P}_k(k = 1, \dots, K)$, and set $gen = 1$.

while $MaxGen$ is not exceeded **do**

 Generation of new solution set \mathbf{Q} using Section II-C;

if $mod(gen, G) == 0$ **then**

 Reset direction vectors using Section III-A and weight vectors using Section III-B;
 Update subpopulation \mathbf{P}_k by $\mathbf{Q} \cup \mathbf{P}$
 $(k = 1, \dots, K)$.

else

 Divide \mathbf{Q} to $\mathbf{Q}_1, \dots, \mathbf{Q}_K$ using Section II-B;
 Update subpopulation \mathbf{P}_k by $\mathbf{Q}_k(k = 1, \dots, K)$ using Section II-D.

end

$gen = gen + 1$;

 output $\cup_{k=1}^K \mathbf{P}_k$.

end

effort from the explored areas which are not worth for further exploitation. Therefore, search effort can be more effectively utilized in MOEA/D-AM2M framework. During the subregions adjustment process, the weights in each subregion will also be reset according to the new division of the subregions.

- 2) Although the adaptive subregion division strategy is applied, MOEA/D-AM2M still inherits MOEA/D-M2M's strong ability in population diversity maintenance. In the implementation of MOEA/D-AM2M, we adjust the subregions once every G generations, and thus the population will have enough time to fully exploit each subregion. In this way, some solutions will be protected from being eliminated by the easy-to-find Pareto optimal solutions during the G generations, which means MOEA/D-AM2M can also be very effective in dealing with the imbalanced MOPs [9]. Moreover, this method can avoid the waste of precious computing resource brought by the frequent subregion adjustment and make the proposed algorithm more stable.
- 3) It is true that when the population has no representative from one patch of PF, the algorithm will lose weight vectors from there, but as soon as a single point in the lost patch emerges the weight vectors will come back to the patch. We believe this is a reasonable approach, as losing a patch in a population may also mean that there is no PF point there and in this case our algorithm will not spend computations in the patch.

IV. CONSTRUCTION OF CHALLENGING MAOPs

In this section, we introduce two types of challenging MaOPs. The first type is degenerated MaOPs with disconnected PFs, and the second type is non-degenerated MaOPs with disconnected PFs.

A. Degenerated MaOPs with disconnected PFs

We first show how to construct the new class of degenerated MaOPs with disconnected PFs, and it is of the following form:

$$\begin{aligned} \text{minimize: } & \begin{cases} f_{r_1}(\mathbf{x}) = (1 + g(\mathbf{x}_{II}) + h(\mathbf{x}_I))t_{r_1}(\mathbf{x}_I) \\ f_{r_2}(\mathbf{x}) = (1 + g(\mathbf{x}_{II}) + h(\mathbf{x}_I))t_{r_2}(\mathbf{x}_I) \\ \dots \\ f_{r_k}(\mathbf{x}) = (1 + g(\mathbf{x}_{II}) + h(\mathbf{x}_I))t_{r_k}(\mathbf{x}_I) \\ f_{r_i}(\mathbf{x}) = q_{r_i}(\mathbf{x}_{II})t_{r_i}(\mathbf{x}_I) \quad (i = k+1, \dots, m) \end{cases} \\ \text{subject to } & \mathbf{x} \in \prod_{i=1}^n [a_i, b_i] \subset R^n. \end{aligned} \quad (5)$$

where $\mathbf{x} = (x_1, x_2, \dots, x_n)$. $\mathbf{x}_I = (x_1, x_2, \dots, x_d)$, $\mathbf{x}_{II} = (x_{d+1}, x_{d+2}, \dots, x_n)$ are two subvectors of \mathbf{x} , m is the number of objectives, and n is the number of variables ($m \leq n$). $\{r_1, r_2, \dots, r_k\}$ is a subset of $\{1, 2, \dots, m\}$, which represents the k random non-degenerated objectives. $\{r_{k+1}, r_{k+2}, \dots, r_m\}$ is the complement set of $\{r_1, r_2, \dots, r_k\}$, which represents the $m - k$ degenerated objectives. We assume:

- 1) $g(\mathbf{x}_{II})$ and $q_{r_i}(\mathbf{x}_{II})$ ($i = k+1, \dots, m$) are non-negative functions from $\prod_{i=d+1}^n [a_i, b_i]$ to R^+ , and their minimal values are zero.
- 2) $t_{r_i}(\mathbf{x}_I)$ ($i = 1, \dots, m$) are functions from $\prod_{i=1}^d [a_i, b_i]$ to R^+ , and $t_{r_1}(\mathbf{x}_I), \dots, t_{r_k}(\mathbf{x}_I)$ cannot be zero at the same time.
- 3) $h(\mathbf{x}_I)$ is a non-negative function from $\prod_{i=1}^d [a_i, b_i]$ to R^+ and its minimal values is zero. For any \mathbf{x}_I , if $h(\mathbf{x}_I) \neq 0$, then we can find \mathbf{x}'_I satisfying $h(\mathbf{x}'_I) = 0$ and $(t_{r_1}(\mathbf{x}'_I), \dots, t_{r_k}(\mathbf{x}'_I))$ dominates $(t_{r_1}(\mathbf{x}_I), \dots, t_{r_k}(\mathbf{x}_I))$.
- 4) If $g(\mathbf{x}_{II}) = 0$, then $q_{r_i}(\mathbf{x}_{II}) = 0$ for any $i = k+1, \dots, m$.

Thus, the manifold of PF is totally determined by $t_{r_1}(\mathbf{x}_I), \dots, t_{r_k}(\mathbf{x}_I)$ and $h(\mathbf{x}_I)$, where $h(\mathbf{x}_I)$ is used to further shape the PF. Actually, we have the following theorem:

Theorem 1. For degenerated MaOP (5), if $t_{r_i}(\mathbf{x}_I)$ ($i = 1, \dots, k$) cannot be further minimized, then $\mathbf{x} = (\mathbf{x}_I, \mathbf{x}_{II}) \in \text{PS}$ if and only if $g(\mathbf{x}_{II}) = 0$ and $h(\mathbf{x}_I) = 0$.

Proof: We first give the proof of the ‘if’ part. Since $g(\mathbf{x}_{II}) = 0$, it follows that $q_{r_i}(\mathbf{x}_{II}) = 0$ ($i = k+1, \dots, m$) by condition 4). Substituting $h(\mathbf{x}_I) = 0$, we can further obtain $f_i(\mathbf{x}) = t_i(\mathbf{x}_I)$ ($i = r_1, \dots, r_k$) and $f_i(\mathbf{x}) = 0$ ($i = r_{k+1}, \dots, r_m$). Thus, each $f_i(\mathbf{x}) = 0$ ($i = r_{k+1}, \dots, r_m$) cannot be further minimized, and every $f_i(\mathbf{x}) = t_i(\mathbf{x}_I)$ ($i = r_1, \dots, r_k$) can also not be further minimized according to the given condition. Therefore, we can deduce that $\mathbf{x} = (\mathbf{x}_I, \mathbf{x}_{II}) \in \text{PS}$.

Then, we prove the ‘only if’ part. For any given $\mathbf{x} = (\mathbf{x}_I, \mathbf{x}_{II})$, if $g(\mathbf{x}_{II}) > 0$, we can find $\mathbf{x}'_{II} \in \prod_{i=d+1}^n [a_i, b_i]$ such that $g(\mathbf{x}'_{II}) = 0$. By condition 4), we can get

$q_i(\mathbf{x}'_{II}) = 0$ ($i = r_{k+1}, \dots, r_m$) from $g(\mathbf{x}'_{II}) = 0$, and further $f_i(\mathbf{x}') = 0$ ($i = r_{k+1}, \dots, r_m$). Let $\mathbf{x}' = (\mathbf{x}_I, \mathbf{x}'_{II})$, we have $f_i(\mathbf{x}') = (1 + h(\mathbf{x}_I))t_i(\mathbf{x}_I) \leq f_i(\mathbf{x}) = (1 + g(\mathbf{x}_{II}) + h(\mathbf{x}_I))t_i(\mathbf{x}_I)$ ($i = r_1, \dots, r_k$), and $f_i(\mathbf{x}') = 0 \leq f_i(\mathbf{x}) = q_i(\mathbf{x}_{II})t_i(\mathbf{x}_I)$ ($i = r_{k+1}, \dots, r_m$). Note that for any $\mathbf{x} = (\mathbf{x}_I, \mathbf{x}_{II})$, $t_{r_1}(\mathbf{x}_I), \dots, t_{r_k}(\mathbf{x}_I)$ cannot be zero at the same time, and thus there is at least one index $j \in \{r_1, \dots, r_k\}$ such that $f_j(\mathbf{x}') < f_j(\mathbf{x})$. Now we can see that \mathbf{x}' dominates \mathbf{x} , which is contradict with the condition that $\mathbf{x} \in \text{PS}$.

If $g(\mathbf{x}_{II}) = 0$ and $h(\mathbf{x}_I) > 0$, we can find $\mathbf{x}'_I \in \prod_{i=1}^d [a_i, b_i]$ such that $h(\mathbf{x}'_I) = 0$ and $(t_{r_1}(\mathbf{x}'_I), \dots, t_{r_k}(\mathbf{x}'_I))$ dominates $(t_{r_1}(\mathbf{x}_I), \dots, t_{r_k}(\mathbf{x}_I))$ by condition 3). Let $\mathbf{x}' = (\mathbf{x}'_I, \mathbf{x}_{II})$, we have $f_i(\mathbf{x}') = t_i(\mathbf{x}'_I) \leq t_i(\mathbf{x}_I) \leq f_i(\mathbf{x}) = (1 + h(\mathbf{x}_I))t_i(\mathbf{x}_I)$ ($i = r_1, \dots, r_k$), and $f_i(\mathbf{x}') = f_i(\mathbf{x}) = 0$ ($i = r_{k+1}, \dots, r_m$). Note that $(t_{r_1}(\mathbf{x}'_I), \dots, t_{r_k}(\mathbf{x}'_I))$ dominates $(t_{r_1}(\mathbf{x}_I), \dots, t_{r_k}(\mathbf{x}_I))$, and hence there is at least one index $j \in \{r_1, \dots, r_k\}$ such that $f_j(\mathbf{x}') < f_j(\mathbf{x})$. That is to say, \mathbf{x}' dominates \mathbf{x} , which is also contradict with the condition that $\mathbf{x} \in \text{PS}$. ■

Considering that PF shapes have a great influence on the algorithm performance [33], we construct a set of seven degenerated test problems with disconnected PF of various shapes based on the method introduced above. Their search space is $[0, 1]^n$, $n = 10$, the number of objectives is $m = 10$, and all these instances are for minimization. In our study, each test instance has its own $t_{r_1}(\mathbf{x}_I)$, $t_{r_2}(\mathbf{x}_I)$ and $t_{r_3}(\mathbf{x}_I)$, and we simply set $t_{r_4}(\mathbf{x}_I)$ to $t_{r_{10}}(\mathbf{x}_I)$ as functions of $t_{r_1}(\mathbf{x}_I)$, $t_{r_2}(\mathbf{x}_I)$ and $t_{r_3}(\mathbf{x}_I)$:

$$\begin{aligned} t_{r_4}(\mathbf{x}_I) &= 0.1t_{r_1}(\mathbf{x}_I) + 0.1t_{r_2}(\mathbf{x}_I) + 0.3t_{r_3}(\mathbf{x}_I); \\ t_{r_5}(\mathbf{x}_I) &= 0.4t_{r_1}(\mathbf{x}_I) + 0.1t_{r_2}(\mathbf{x}_I) + 0.3t_{r_3}(\mathbf{x}_I); \\ t_{r_6}(\mathbf{x}_I) &= 0.3t_{r_1}(\mathbf{x}_I) + 0.2t_{r_2}(\mathbf{x}_I) + 0.4t_{r_3}(\mathbf{x}_I); \\ t_{r_7}(\mathbf{x}_I) &= 0.3t_{r_1}(\mathbf{x}_I) + 0.1t_{r_2}(\mathbf{x}_I) + 0.1t_{r_3}(\mathbf{x}_I); \\ t_{r_8}(\mathbf{x}_I) &= 0.4t_{r_1}(\mathbf{x}_I) + 0.3t_{r_2}(\mathbf{x}_I) + 0.1t_{r_3}(\mathbf{x}_I); \\ t_{r_9}(\mathbf{x}_I) &= 0.3t_{r_1}(\mathbf{x}_I) + 0.2t_{r_2}(\mathbf{x}_I) + 0.3t_{r_3}(\mathbf{x}_I); \\ t_{r_{10}}(\mathbf{x}_I) &= 0.1t_{r_1}(\mathbf{x}_I) + 0.3t_{r_2}(\mathbf{x}_I) + 0.1t_{r_3}(\mathbf{x}_I). \end{aligned}$$

$$\text{MaOP1: } \begin{cases} f_1(\mathbf{x}) = (1 + g(\mathbf{x}_{II}) + h(x_1))\frac{\sqrt{2}x_1}{2} \\ f_2(\mathbf{x}) = (1 + g(\mathbf{x}_{II}) + h(x_1))\frac{\sqrt{2}x_1}{2} \\ f_3(\mathbf{x}) = (1 + g(\mathbf{x}_{II}) + h(x_1))3(1 - x_1^2) \\ f_{r_i}(\mathbf{x}) = q_{r_i}(\mathbf{x}_{II})t_{r_i}(\mathbf{x}_I) \quad (i = 4, \dots, 10) \end{cases}$$

where

$$\begin{aligned} \mathbf{x} &= (\mathbf{x}_I, \mathbf{x}_{II}), \mathbf{x}_I = (x_1), \mathbf{x}_{II} = (x_2, x_3, \dots, x_n), \\ r_1 &= 1, r_2 = 2, r_3 = 3, \\ g(\mathbf{x}_{II}) &= \sum_{j=2}^n |x_j - \sin(\frac{\pi x_1}{2})|^2, \\ h(x_1) &= \max(0, 2\sin(4\pi x_1)), \\ q_i(\mathbf{x}_{II}) &= \exp(|x_i - \sin(\frac{\pi x_1}{2})|^2) - 1, \quad i = 4, \dots, 10. \end{aligned}$$

Its PF is a disconnected curve in the reduced objective space of f_1 , f_2 , and f_3 satisfying $f_1^2 + f_2^2 + f_3/3 = 1$ and $f_1 = f_2$ when $\sqrt{2}/8 \leq f_1 \leq \sqrt{2}/4$ or $3\sqrt{2}/8 \leq f_1 \leq \sqrt{2}/2$. Its PS is $x_i = \sin(0.5\pi x_1)$ for $0.25 \leq x_1 \leq 0.5$ or $0.75 \leq x_1 \leq 1$. and for $i = 2, \dots, n$. Its ideal point is $(0, 0, 0, \dots, 0)$, and its nadir point is $(1, 1, 3, 0, \dots, 0)$.

MaOP2 to MaOP7 problems are described in the Appendix.

The seven proposed many-objective optimization test problems have disconnected and three-dimension PFs in a reduced objective space. When optimizing this type of problems, the EMO algorithm needs to focus its search effort on the reduced space to find the disconnected PFs. All these char-

acteristics make the constructed MaOPs suitable for testing EMO algorithms' capability to adjust the search effort during the evolutionary process, which can effectively facilitate the design, testing and applications of EMO algorithms for degenerated MaOPs. In addition, we also have developed two non-degenerated test problems (MaOP8 and MaOP9) with disconnected PFs and they are described in the Appendix.

V. EXPERIMENTAL STUDY

In this section, we execute extensive experimental studies to investigate the performance of MOEA/D-AM2M. We first briefly introduce the comparing EMO algorithms, and then compare MOEA/D-AM2M with them on a set of many-objective test problems. The test problems include the constructed challenging MaOP1-MaOP9 with disconnected PFs and the classic MaOP benchmark problems from the DTLZ [34] and WFG [35] series.

A. EMO Algorithms in Comparison

We compare MOEA/D-AM2M with eight many-objective optimization algorithms, i.e., MOEA/D-DE [7], NSGA-III [28], MOEA/D-AWA [27], MOEA/D-M2M [8], GrEA [29], NL-MVU-PCA [14], HypE [30] and RVEA [23]. Since we have already introduced the general principle of MOEA/D-M2M in Section II-B, we only briefly introduce the other seven comparing algorithms.

1) *MOEA/D-DE* [7]: MOEA/D is a decomposition based EMO algorithm, and various decomposition methods [11], [32] can be applied for decomposition. MOEA/D decomposes an MaOP into a number of single objective optimization subproblems with the aid of a set of predefined weights. MOEA/D-DE is an efficient and effective version of MOEA/D based on differential evolution (DE). In our study, MOEA/D-DE with the Tchebycheff decomposition is used for comparing. In addition, a niching parameter T that is used to define the neighboring weight vectors for crossover and mutation, and the maximal number of solutions replaced by each child solution n_r are also needed in MOEA/D-DE.

2) *NSGA-III* [28]: NSGA-III works under the similar framework to the original NSGA-II except for the operation on the last selected non-domination level. In NSGA-III, reference points based niching technique is proposed to enhance the population convergence and maintain the population diversity, and Das and Dennis's [31] systematic approach is used to generate a set of reference points with a good distribution. When selecting the next generation population, the nondominated sorting selection is firstly executed, and the selected solutions are then normalized. After normalization, the niching technique based selection works to fill the slots of the next generation population.

3) *MOEA/D-AWA* [27]: MOEA/D-AWA uses MOEA/D with the adaptive weight vector adjustment strategy to address the MOPs with complex PFs. In MOEA/D-AWA, a certain kind of computing resources are assigned to adjust the weight vectors periodically, and the procedures are controlled by two parameters *rate_evol* and *wag* respectively. The weight vectors are initialized evenly at first, and then MOEA/D-AWA

adjusts the weights according to the sparsity level defined by the vicinity distance. With the help of an external elite population, MOEA/D-AWA adjusts the weights by deleting the overcrowded weight vectors and adding new weight vectors into the so-called real sparse regions. Numerical experiments have shown that MOEA/D-AWA can achieve good results in solving MOPs with complex PFs. It is worth noting that the calculation of vicinity distance is very time consuming, and vicinity distance of each solution need to be recalculated after each deleting or adding operation.

4) *GrEA* [29]: GrEA is a grid-based evolutionary algorithm for many-objective optimization, and its basic procedure is similar to NSGA-II. In order to increase the selection pressure of Pareto dominance based selection and maintain a good diversity among solutions, the grid strategy is introduced by modifying the Pareto dominance to grid dominance and the crowding distance to grid crowding distance. Grid dominance can better distinguish solutions in both mating and environmental selection processes when optimizing MaOPs, which makes GrEA an effective many-objective optimizer.

5) *NL-MVU-PCA* [14]: NL-MVU-PCA is a nonlinear (NL) objective reduction framework for MaOPs with redundant objectives. In NL-MVU-PCA, both MVU and PCA are utilized to remove the dependencies in the nondominated solutions, and then the essential objectives can be identified by eigenvalue analysis. Once the essential objectives of an MaOP are revealed, the MaOP can be optimized easier by the existing EMO algorithms. In our study, the classic NSGA-II is applied to the NL-MVU-PCA framework and the objective reduction procedure will be performed periodically during the NSGA-II run.

6) *HypE* [30]: HypE (Hypervolume Estimation algorithm) is a hypervolume indicator-based EMO algorithm for many-objective optimization. HypE adopts the Monte-Carlo simulation to approximate the exact hypervolume value instead of calculating it directly. Monte-Carlo simulation can effectively reduce the computational effort required for hypervolume indicator calculation, and thus make the hypervolume-based search to be easily applied to many-objective optimization. Simulation results indicate that HypE is highly competitive for MaOPs in comparison to many existing EMO algorithms.

7) *RVEA* [23]: Reference vector-guided evolutionary algorithm (RVEA) is a recent decomposition based algorithm for many-objective optimization, where the reference vectors are used for decomposition and representing the user preferences to a preferred subset of PF. A new scalarization approach, called angle-penalized distance, is proposed for selection. RVEA employs an adaptation strategy to dynamically adjust the distribution of the reference vectors according to the scales of the objective functions. A reference vector regeneration strategy is proposed for handling MaOPs with irregular PFs. Experiments on a set of many-objective test problems have shown that the RVEA is very competitive.

B. Performance Metrics

Two performance metrics are used to measure the quality of obtained solutions by those algorithms in our experimental studies.

1) *IGD-metric* [36]: Let \mathbf{P}^* be a set of points uniformly distributed over the PF in objective space, and \mathbf{P} be the set of obtained solutions of an EMO algorithm. The IGD value of \mathbf{P} can be calculated as:

$$IGD(\mathbf{P}, \mathbf{P}^*) = \frac{\sum_{\mathbf{x}^* \in \mathbf{P}^*} dist(\mathbf{x}^*, \mathbf{P})}{|\mathbf{P}^*|} \quad (6)$$

where $dist(\mathbf{x}^*, \mathbf{P})$ is the Euclidean distance between a point $\mathbf{x}^* \in \mathbf{P}^*$ and its nearest neighbor in \mathbf{P} , and $|\mathbf{P}^*|$ is the cardinality of \mathbf{P}^* , all computed in the objective space. Obviously, the lower the IGD value is, the better the algorithm performs. Considering that the generation the evenly distributed points across the PF in high-dimension objective space is not an easy task, we only use IGD-metric to measure the quality of obtained solutions for the seven constructed test MaOPs with degenerated PFs. The non-degenerated objective values of all the reference sets can be easily generated by uniformly sampling from the degenerated PFs and their degenerated objective values are all zero.

C. Experimental Setting

The experimental settings of MOEA/D-AM2M, MOEA/D-DE, NSGA-III, MOEA/D-AWA, MOEA/D-M2M, GrEA, NL-MVU-PCA, HypE and RVEA are as follows:

- Population size of the nine testing algorithms is set as $N = 110$ for MaOP1, and $N = 275$ for MaOP2-MaOP9, DTLZ2 and DTLZ4.
- The initial weight vectors (reference points or vectors) of MOEA/D-AM2M, MOEA/D-DE, NSGA-III, MOEA/D-AWA and RVEA are all generated by Das and Dennis's systematic approach with two layers strategy [28] for a fair comparison. For MaOP1, 55 points are generated on the boundary layer ($p_1 = 2$) and 55 points on the inside level ($p_2 = 2$); For the rest, 220 points are generated on the boundary layer ($p_1 = 3$) and 55 points on the inside level ($p_2 = 2$).
- The crossover and mutation operators with the same control parameters in MOEA/D-M2M [8] and MOEA/D [6] are used in MOEA/D-AM2M, MOEA/D-M2M and MOEA/D-DE respectively. The simulated binary crossover (SBX) operator [38] with $p_c = 1$ and $\eta_c = 30$, and polynomial mutation [39] with $p_m = 1/n$ and $\eta_m = 20$ are used in NSGA-III, MOEA/D-AWA, GrEA, NL-MVU-PCA, HypE and RVEA.
- Other control parameters in MOEA/D-AM2M and MOEA/D-M2M: $K = 10$ for MaOP1, $K = 20$ for the remaining, and $G = 100$ in MOEA/D-AM2M.
- Other control parameters in NL-MVU-PCA: the objective reduction procedure will be conducted every 100 generations, and the variance threshold $\theta = 0.997$.
- Other control parameters in MOEA/D-DE: $T = 20$, $\delta = 0.9$, and $n_r = 2$, which are the same as in [6].
- Other control parameters in MOEA/D-AWA: the size of external elite is set as $1.5N$, the parameter *rate_evol* is set as 0.8, and *wag* is set as 300, which are the same as in [27].
- Other control parameters in GrEA: the grid division is set as 8.

- Other control parameters in HypE: the number of sampling points in HypE is set as 10,000, which are the same as in [30].
- Other control parameters in RVEA: the rate to change the penalty function is set as $\alpha = 2$, the frequency to employ the reference vector adaptation is set as $f_r = 0.1$, which are the same as in [23].
- Stopping condition: all algorithms stop after $1,000 \times N$ function evaluations for MaOP1, MaOP8-MaOP9, DTLZ2 and DTLZ4; Stop after $2,000 \times N$ function evaluations for MaOP2-MaOP7.

D. Experimental Study on Degenerated MaOPs with Disconnected PFs

In this subsection, the seven degenerated test instances with disconnected PFs are tested to investigate the efficiency of MOEA/D-AM2M by comparing it with MOEA/D-DE, NSGA-III, MOEA/D-AWA, MOEA/D-M2M, GrEA, NL-MVU-PCA, HypE and RVEA. Considering that PFs of these test instances are all located in a reduced objective space, the HV-metric [37] is not directly calculated in the whole objective space but in the reduced objective space for simplicity, and we call it the approximate HV-metric (AHV-metric) in this paper. When calculating the AHV-metric, the setting of reference point is very crucial. Since the nadir point (\mathbf{z}^{nad}) of each test problem can be easily obtained and normalized to 1, we simply set the reference point as $\mathbf{z}^{nad} + 0.001\vec{\mathbf{e}}$, where $\vec{\mathbf{e}} = (1, \dots, 1)$. Table I shows the best and mean of the obtained AHV-metric value of the nine algorithms for each test instance in 20 independent runs. IGD values are presented in the supplementary document. In the tables, the data of the winner of the nine algorithms are bolded to stand out from other data for the seven test problems. It shows that MOEA/D-AM2M almost outperforms all the eight comparing algorithms in terms of both AHV-metric and IGD-metric value.

The respective solutions for a few of the algorithms are shown in Fig. 5. Performance of remaining algorithms are shown in the supplementary document. It is amply clear that MOEA/D-AM2M algorithm performs the best.

The reason why MOEA/D-AM2M outperforms MOEA/D-DE and NSGA-III is obvious. The adaptive subregion division in MOEA/D-AM2M can effectively adjust the subregions and weights according to the population distribution, which can make the search effort focus on the more promising areas. The search effort represented by the evenly distributed weight vectors/reference points in MOEA/D-DE and NSGA-III is inappropriate for those degenerated MaOPs with disconnected PFs, and it cannot guarantee a good population distribution along the PF. Moreover, the unreasonable allocation of search effort can also lead to the deterioration of NSGA-III's convergence. Since the true PFs of those MaOPs are all located in a reduced objective space, most of the initial reference points are not effective to guide the population search. The nondominated sorting and niching selection cannot provide enough selection pressure to push the population to the reduced objective space, let alone converge to the disconnected PFs. The normalization procedure conducted by the niching selection is based on

the assumption that PFs are distributed among the whole objective space. In this case, it is more likely that many solutions associate with one reference point and thus more than one solutions associated with the same reference point need to be selected to fill the next population slots. The niche based selection tends to first select the solutions with less perpendicular distance to the associated reference line and then push the associated solutions along their reference line to the PFs. Therefore, the population may not be able to approximate the PF very well.

MOEA/D-AWA periodically adjusts the weights by using an external archive set which stores the nondominated solutions found by now. However, the dominated solutions may also contain some useful PF information especially for those degenerated MaOPs, and this information can have some benefits to guide the population jump out of over searched areas and explore some more unexplored areas. The ignorance of those dominated solutions may lead to the loss of the opportunity to quickly extend population search to more promising areas, and eventually waste a lot of search effort. In those degenerated MaOPs optimization, consequence brought by this ignorance has made MOEA/D-AWA fail to adjust the weight vectors effectively.

In GrEA, grid ranking and grid crowding distance are applied to select the next generation population. Although the grid dominance ranking can provide more selection pressure than the Pareto dominance ranking does, and the grid crowding distance can distinguish the overcrowded solutions more effectively than the crowding distance does, it still fails to push the population to the degenerated PFs. The reason is that GrEA cannot focus its search effort on the specific reduced objective space where the degenerated PFs are located. NL-MVU-PCA objective reduction based NSGA-II is very competitive among all these comparing algorithms, but it may fail to find the essential objectives for some MaOPs. And even when the essential objectives are revealed, its population distribution along those degenerated PFs is still not as good as that of MOEA/D-AM2M. It is because that the diversity management strategy of MOEA/D-AM2M is much better than the crowding distance method in NSGA-II. HypE is considered to be very promising for many-objective optimization, but its performance can be greatly restricted by the accuracy of Monte-Carlo simulation. Moreover, the setting of the reference point for hypervolume calculation may also have an effect on the performance of HypE. Consequently, the performance of HypE on those degenerated MaOPs cannot catch up with MOEA/D-AM2M.

Similar to MOEA/D-AM2M, RVEA also adaptively adjusts the reference vectors, but in an entirely different way. RVEA periodically adjusts the distribution of the reference vectors according to the scales of the objective functions. During the selection, RVEA first associates the population with the reference vectors and then utilizes the angle-penalized distance to determine which one to be selected from the set of solutions associated with the same reference vector. When solving those degenerated MaOPs, it is highly possible that many reference vectors do not have any solutions to associate with, and under this condition only a small number of solutions can be

TABLE I
THE BEST AND MEAN OF AHV-METRIC VALUES OF MOEA/D-AM2M, MOEA/D-DE, NSGA-III MOEA/D-AWA, GRE, NL-MVU-PCA, HYPE AND RVEA IN 20 INDEPENDENT RUNS FOR EACH 10-OBJECTIVE TEST INSTANCE. BEST PERFORMANCE IS HIGHLIGHTED IN BOLD

AHV-metric		MOEA/D -AM2M	MOEA/D -DE	NSGA-III	MOEA/D -AWA	MOEA/D -M2M	GrEA	NL-MVU -PCA	HypE	RVEA
MaOP1	Best	0.819926	0.652716	0.489734	0.66525	0.743976	0.348232	0.959969	0.538851	0.438923
	Mean	0.817634	0.647993	0.403754	0.661352	0.690142	0.319444	0.795649	0.533429	0.379753
MaOP2	Best	1.277999	0.812332	0.777752	0.879183	0.258203	0.660127	1.108237	0.189388	0.270449
	Mean	1.261981	0.792226	0.643808	0.82233	0.248125	0.539335	0.518808	0.123797	0.183959
MaOP3	Best	9.369566	7.371739	7.188816	7.847038	1.018329	3.861009	8.967739	7.243972	2.804674
	Mean	9.344276	7.236861	6.866672	7.710055	0.887693	1.810892	7.088994	7.017994	2.158043
MaOP4	Best	0.380116	0.285377	0.195129	0.255330	0.313981	0.195899	0.408792	0.325505	0.035116
	Mean	0.377943	0.271247	0.148368	0.245111	0.292559	0.085772	0.402454	0.319095	0.003512
MaOP5	Best	0.968094	0.93667	0.931094	0.966614	0.944085	0.861348	0.960846	0.966614	0.323897
	Mean	0.967833	0.93471	0.92623	0.966306	0.940865	0.829418	0.958894	0.966306	0.099898
MaOP6	Best	0.742152	0.680867	0.573799	0.650539	0.684177	0.343088	0.727454	0.697626	0.505059
	Mean	0.737105	0.679094	0.540439	0.645879	0.664930	0.179983	0.724435	0.582382	0.368046
MaOP7	Best	0.382664	0.298496	0.211224	0.270925	0.310110	0.002988	0.151594	0.328421	0.135698
	Mean	0.380375	0.295207	0.167972	0.261672	0.299639	0.002918	0.031119	0.306248	0.045722

selected at each generation. RVEA tries to ease the situation by regenerating a set of population size reference vectors at each generation, but the quality of the newly generated reference vector cannot be guaranteed. In fact, the regeneration of reference vectors is highly dependent on the distribution of population, whereas the selection strategy adopted by RVEA may not be able to select a set of high-quality solutions. This will in turn affect the selection of next generation population, and thus RVEA could not effectively approximate those degenerated PFs.

E. Further Performance Study of MOEA/D-AM2M on Degenerated MaOPs with Connected PFs

DTLZ5(M, I) and WFG3(M) are two widely-used degenerated many-objective test problems. Their connected PFs differentiate them from the constructed degenerated MaOPs. NL-MVU-PCA is specially designed for the degenerated MaOPs, and it has been reported to be very effective in dealing with the DTLZ5(M, I) and WFG3(M) series test problems [14]. Thus, we do further performance study of MOEA/D-AM2M by comparing it with NL-MVU-PCA on a set of DTLZ5(M, I) and WFG3(M) test problems. DTLZ5(M, I) and WFG3(M) both have M objectives, where DTLZ5(M, I) has I scalable essential objectives and WFG3(M) has two essential objectives. In our study, we extend the original WFG3(M) to WFG3(M, I) with I scalable essential objectives (see Table III), and then compare the two algorithms on a set of original DTLZ5(M, I) and extended WFG(M, I) test problems with three essential objectives. The decision variables of DTLZ5 and WFG3 are set as $M + 9$ and $M + 19$. The population size of MOEA/D-AM2M and NL-MVU-PCA are both set as 210 ($p_1 = 6$), 156 ($p_1 = 3, p_2 = 2$), 275 ($p_1 = 3, p_2 = 2$) and 135 ($p_1 = 3, p_2 = 1$) for 5-, 8-, 10- and 15-objective test problems. Simulation results of MOEA/D-AM2M and NL-MVU-PCA on the original DTLZ5(M, I) and extended WFG3(M, I) test

problems are shown in Table II in terms of AHV-metric, where the reference point is set as $\mathbf{z}^{nad} + 0.01\vec{\mathbf{e}}$. It indicates that MOEA/D-AM2M outperforms NL-MVU-PCA on the original DTLZ5(M, I) test problems. As for the extended WFG3(M, I) test problems, the performances of MOEA/D-AM2M are not as good as that of NL-MVU-PCA, but still very competitive.

NL-MVU-PCA tends to select the objectives with higher magnitudes as the essential objectives during the objective reduction process, while both the original DTLZ5(M, I) and extended WFG3(M, I) have essential objectives with relatively higher magnitudes. Therefore, we modify the original DTLZ5(M, I) and extended WFG(M, I) to make them have the same objective magnitude. The parameter setting of the modified WFG3(M, I) are shown on Table III, and that of the modified DTLZ5(M, I) are shown in the supplementary document. We compare the two algorithms on a set of modified DTLZ5(M, I) and WFG(M, I) test problems with three essential objectives. The simulation results show that the advantages of MOEA/D-AM2M stand out when it comes to the modified test problems. Table IV shows that MOEA/D-AM2M outperforms NL-MVU-PCA both on the modified DTLZ5(M, I) and WFG3(M, I) test problems. We can also see that the mean AHV-metric values of 15 independent NL-MVU-PCA runs are much worse than that of MOEA/D-AM2M. It indicates that most of the NL-MVU-PCA runs cannot achieve comparable results to MOEA/D-AM2M on the modified DTLZ5(M, I) and WFG3(M, I) test problems and its performance is unstable. When the essential objectives are revealed, NL-MVU-PCA performs well. But if the essential objectives are mistaken, its performances cannot be guaranteed. The experimental results suggest that MOEA/D-AM2M can effectively deal with the two kinds of DTLZ5(M, I) and WFG3(M, I) test problems, which also means that MOEA/D-AM2M is more robust and has wider practical application scope than NL-MVU-PCA.

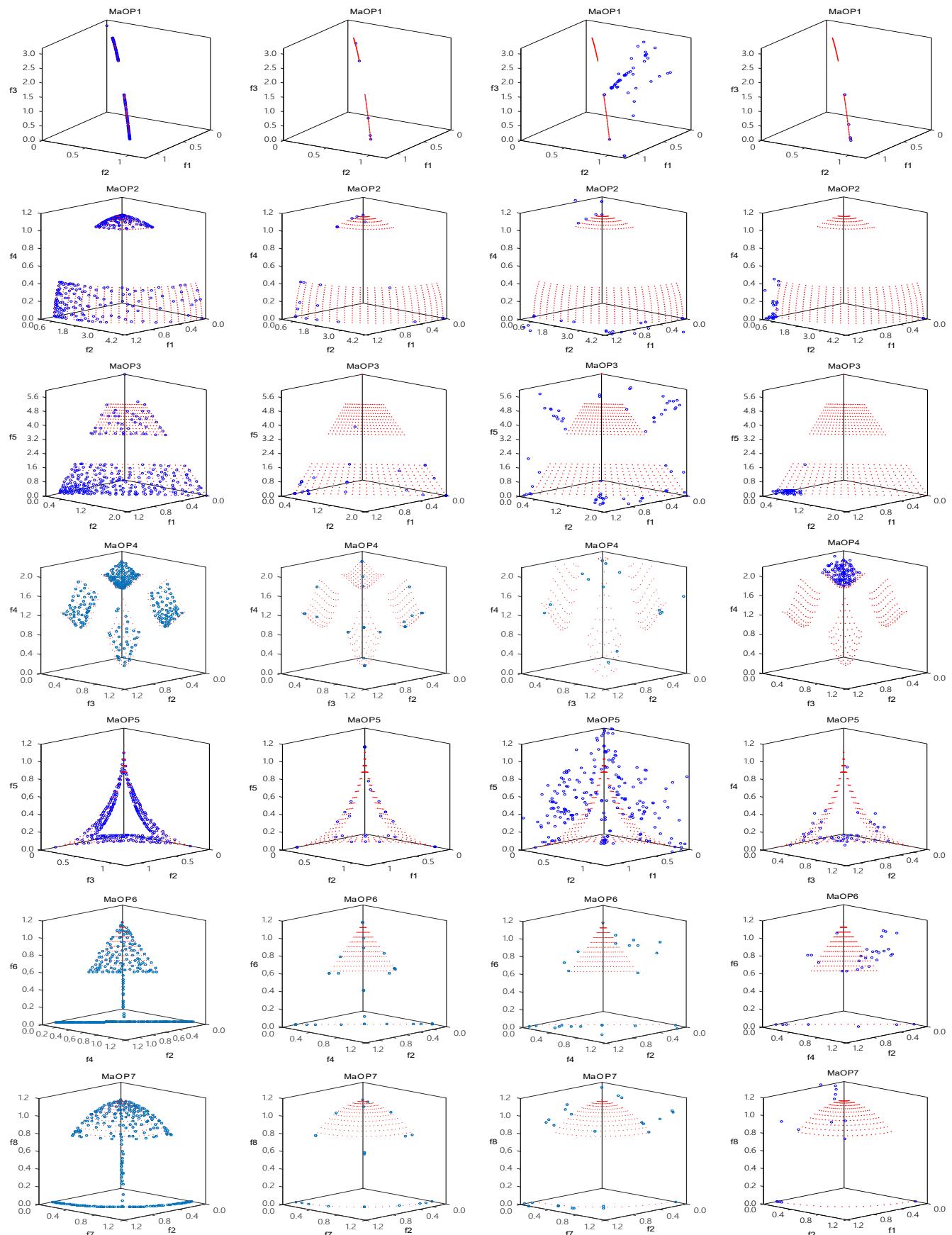


Fig. 5. Plot of the nondominated front in the reduced objective space with the median AHV-metric value found by MOEA/D-AM2M, MOEA/D-DE, NSGA-III and RVEA for test problems MaOP1-MaOP7.

TABLE II
THE BEST AND MEAN OF AHV-METRIC VALUES OF MOEA/D-AM2M AND NL-MVU-PCA IN 15 INDEPENDENT RUNS FOR EACH *Original*, *Extended* TEST INSTANCE. BETTER PERFORMANCE IS HIGHLIGHTED IN BOLD

Test instance	MOEA/D-AM2M		NL-MVU-PCA	
	Best	Mean	Best	Mean
DTLZ5(5,3)	0.561828	0.553378	0.550664	0.545266
DTLZ5(8,3)	0.555296	0.547446	0.547560	0.353382
DTLZ5(10,3)	0.555296	0.547446	0.556739	0.503888
DTLZ5(15,3)	0.545611	0.517906	0.543214	0.480756
WFG3(5,3)	1.225095	1.215011	1.222915	1.219923
WFG3(8,3)	1.225632	1.213020	1.254357	1.229149
WFG3(10,3)	1.240932	1.215174	1.274100	1.159448
WFG3(15,3)	1.265618	1.237599	1.263666	0.702839

TABLE III
PARAMETER SETTING OF ORIGINAL WFG3(M), EXTENDED WFG3(M, I) AND MODIFIED WFG3(M, I)

Original WFG3(M)	Extended WFG3(M, I)	Modified WFG3(M, I)
$S_1 = 2$	$S_1 = 2$	$S_1 = 2^{M-I}$
$S_{m=2:M} = 2m$	$S_{m=2:M} = 2m$	$S_{m=2:M} = 2^{M-I+2-m}$
$A_1 = 1$	$A_1 = 1$	$A_1 = 1$
$A_{2:M-1} = 0$	$A_{I:M-1} = 0$	$A_{I:M-1} = 0$

F. Experimental Study on Non-degenerated MaOPs

In this subsection, the widely-used DTLZ2 and DTLZ4 together with the constructed MaOP8 and MaOP9 are tested to investigate the performance of MOEA/D-AM2M in solving non-degenerated MaOPs. DTLZ2 and DTLZ4 have connected PFs, whereas MaOP8 and MaOP9 have disconnected PFs. The number of objective is $m = 10$ for the four test problems. The number of decision variables is set as $n = 19$ for DTLZ2 and DTLZ4 [34], and the number of decision variables is set as $n = 20$ for MaOP8 and MaOP9 in our study. For all those problems, we use a population size of 275 ($p_1 = 3$, $p_2 = 2$). MOEA/D-AM2M and all the comparing EMO algorithms independently run 15 times for each test problem. Since those test instances are non-degenerated, the simulation results are presented in Table V in terms of HV-metric. To visually view those algorithms' performance, we plot the parallel coordinates of the nondominated fronts with the median HV-metric value found by the comparing algorithms for DTLZ2 in Fig. 6, and that of the remaining test problems are shown in the supplementary document.

The simulation results manifest that MOEA/D-AM2M is

TABLE IV
THE BEST AND MEAN OF AHV-METRIC VALUES OF MOEA/D-AM2M AND NL-MVU-PCA IN 15 INDEPENDENT RUNS FOR EACH *Modified* TEST INSTANCE. BETTER PERFORMANCE IS HIGHLIGHTED IN BOLD

Test instance	MOEA/D-AM2M		NL-MVU-PCA	
	Best	Mean	Best	Mean
DTLZ5(5,3)	0.734448	0.713756	0.731789	0.361913
DTLZ5(8,3)	0.721911	0.664545	0.685292	0.372478
DTLZ5(10,3)	0.728259	0.692928	0.728070	0.440681
DTLZ5(15,3)	0.663554	0.583470	0.554742	0.322677
WFG3(5,3)	1.124336	1.118949	1.106239	0.635537
WFG3(8,3)	1.055720	1.020944	0.865680	0.588942
WFG3(10,3)	1.070010	1.040848	0.841642	0.406114
WFG3(15,3)	0.989431	0.954825	1.050141	0.626440

also good and robust in solving those non-degenerated MaOPs. As shown in Table V, although MOEA/D-AM2M does not perform best for DTLZ2 and DTLZ4 among the nine comparing algorithms, the differences in their performances are not significant. It is reasonable that NSGA-III, MOEA/D, MOEA/D-M2M and RVEA have better performances. The PFs of DTLZ2 and DTLZ4 are distributed in the whole objective space, and thus the well-distributed weight vectors (reference points or vectors) are good enough to guide the population search. The adaptive strategy in MOEA/D-AM2M will, in turn, disorganizes those well-distributed weight vectors, which leads to the MOEA/D-AM2M's worse performance than MOEA/D-DE, NSGA-III, MOEA/D-M2M and RVEA. For MaOP8 and MaOP9 with disconnected PFs, MOEA/D-AM2M can effectively focus the search on more promising areas due to the adaptive search effort allocation, and thus achieves better results.

G. Experimental Study on Imbalanced MOPs

To show the generality of the proposed MOEA/D-AM2M, we also investigate its performance on the imbalanced MOPs. Imbalanced MOPs are first studied in [8], [9], where the 'imbalance' means part of the PF are much more easily to be obtained and the PS of easy-to-get PF dominates a significantly larger part of the feasible variable space than those of the others. MOEA/D-M2M has been reported to be very effective in solving the imbalanced MOPs due to its strong ability in diversity maintenance. In our study, we compare MOEA/D-AM2M with MOEA/D-M2M on the tri-objective imbalanced MOP6 [8]. For any solutions satisfying $x_1 = 0$ or $x_1 = 1$, they are automatically the Pareto optimal solutions of MOP6. Those easy-to-get Pareto optimal solutions will emerge early, and they also dominates a large part of the variable space. As has been reported in [8], [9], if no other diversity management strategies are adopted, the whole PF of MOP6 cannot be effectively approximated by MOEA/D and NSGA-II.

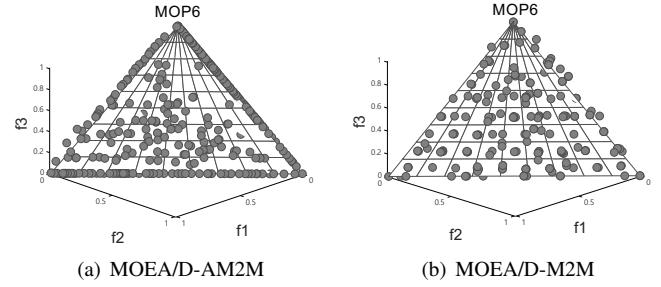


Fig. 7. Plot of the nondominated front with the median HV-metric value found by MOEA/D-AM2M and MOEA/D-M2M.

Fig. 7 gives an intuitive illustration of the obtained solutions on the PF of MOP6 with the median HV-metric value of the 15 independent runs found by MOEA/D-AM2M and MOEA/D-M2M. We can see that MOEA/D-AM2M does inherit MOEA/D-M2M's strong ability in diversity maintenance. In fact, the adaptive subregion division of MOEA/D-AM2M is conducted every G generations, and the corresponding population decomposition and selection are kept the same as MOEA/D-M2M during the G generations.

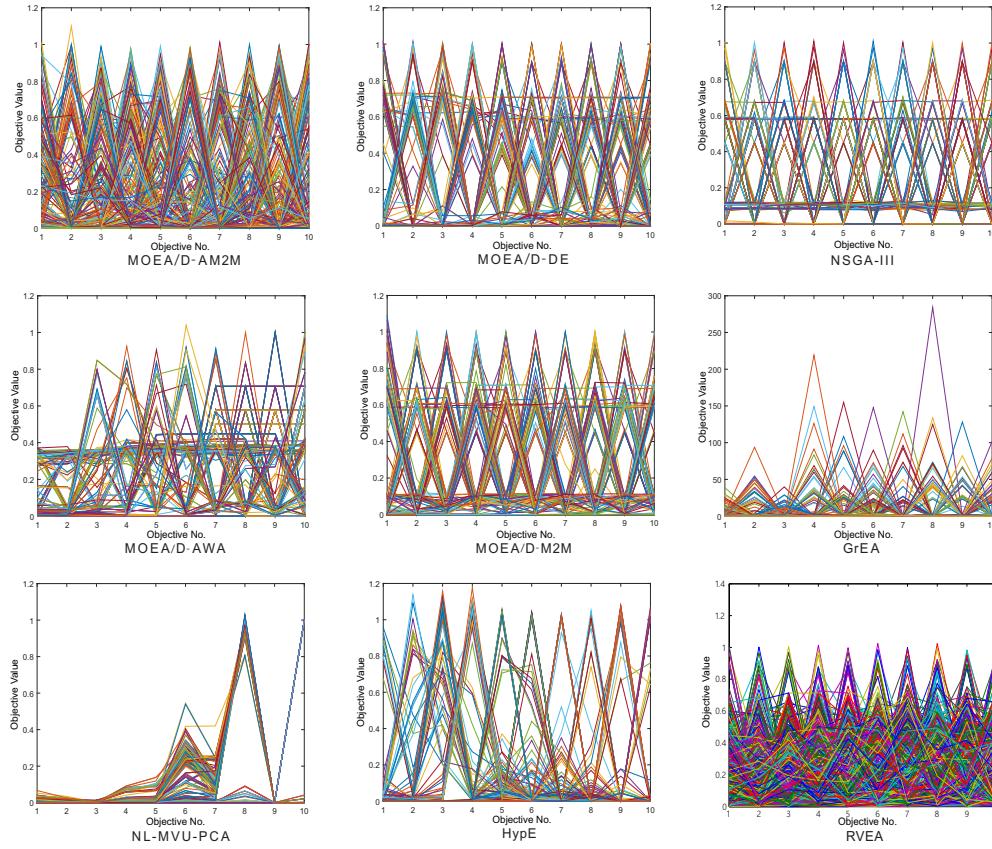


Fig. 6. Parallel coordinates of the nondominated front with the median HV-metric value found by MOEA/D-AM2M, MOEA/D-DE, NSGA-III, MOEA/D-AWA, MOEA/D-M2M, GrEA, NL-MVU-PCA, HypE and RVEA for the 10-objective DTLZ2.

TABLE V
THE BEST AND MEAN OF HV-METRIC VALUES OF MOEA/D-AM2M, MOEA/D-DE, NSGA-III MOEA/D-AWA, GREA, NL-MVU-PCA, HYPE AND RVEA IN 15 INDEPENDENT RUNS FOR EACH 10-OBJECTIVE TEST INSTANCE. BEST PERFORMANCE IS HIGHLIGHTED IN BOLD

HV-metric		MOEA/D -AM2M	MOEA/D -DE	NSGA-III	MOEA/D -AWA	MOEA/D -M2M	GrEA	NL-MVU -PCA	HypE	RVEA
DTLZ2	Best	2.482340	2.489943	2.512184	2.016623	2.498579	2.377166	1.550355	2.08867	2.494919
	Mean	2.472182	2.432555	2.511509	1.983044	2.491823	1.958930	0.731335	1.973589	2.473874
DTLZ4	Best	2.499119	2.506492	2.505612	1.915372	2.511029	2.215632	0	1.745092	2.507191
	Mean	2.492412	2.497723	2.500588	1.835551	2.508396	2.172000	0	1.631444	2.501037
MaOP8	Best	2.583155	2.545327	2.461032	2.556791	2.560456	0	0.606839	2.560822	2.501357
	Mean	2.580421	2.515927	2.417337	2.550973	2.557242	0	0.237302	2.557662	2.461335
MaOP9	Best	2.298554	2.216781	2.211247	1.817575	1.555534	0.000595	1.582893	2.016583	2.193324
	Mean	2.274088	2.071834	2.149641	1.812865	1.457524	0.000272	0.773845	1.922927	1.824980

H. The Setting of Update Parameter (G) in MOEA/D-AM2M

Generally, the population's distribution will change at every generation, but it may not change dramatically within very few generations. Therefore, extracting PF information from the population to adaptively adjust the subregions and weights at every generation is unnecessary, and frequent adjustments also cost the precious computing resources. MOEA/D-AM2M adjusts the subregions and weights every G generations. To investigate the effect of this parameter setting on the algorithm performance, we have conducted a series of experiments with different values of G on problem MaOP2. The AHV-metric

and IGD-metric of experimental results obtained by MOEA/D-AM2M with different setting of G from 1 to 400 are plotted in Fig. 8. It is clear from this figure that the performance of MOEA/D-AM2M is relatively stable on a wide range of G : $60 \leq G \leq 180$. These results indicate that MOEA/D-AM2M is not particularly sensitive to this parameter. We suggest that G is set among [60, 180] for MOEA/D-AM2M in our study.

VI. CONCLUSION AND FUTURE WORK

In this paper, we have proposed an adaptive search effort allocation framework called MOEA/D-AM2M for MaOPs. In

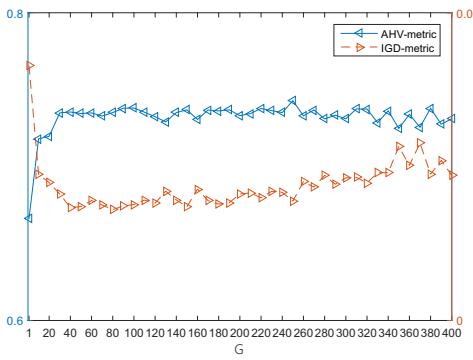


Fig. 8. Variation of AHV-metric and IGD-metric for MOEA/D-AM2M with different setting of G for solving MaOP2.

MOEA/D-AM2M, search effort can be adaptively allocated by adaptive subregion division and weight setting. A set of degenerated and non-degenerated MaOPs with disconnected PFs are constructed for numerical experiments. The performances of MOEA/D-AM2M on the classic degenerated DTLZ5 and WFG3 with connected PFs are also studied. Additionally, we also compare MOEA/D-AM2M with the eight comparing algorithms on non-degenerated MaOPs with connected and disconnected PFs. Simulation experiments show that the proposed algorithm is good and robust in performance.

MOEA/D-AM2M is an improved version of MOEA/D-M2M, and it inherits MOEA/D-M2M's advantages in high-efficiency selection and population diversity maintenance. As a limitation, like MOEA/D-M2M, MOEA/D-AM2M also has less global competition in its working of finding multiple regions on the PF, as they both work for each subregion separately. Nevertheless, as depicted through our simulation results, MOEA/D-AM2M is better able to spread population members on the entire PF compared to MOEA/D-M2M and other methods in solving difficult multi-objective optimization problems. Our further research will focus on the practical application of MOEA/D-AM2M on real-world problems.

REFERENCES

- [1] H. Ishibuchi, N. Tsukamoto, and Y. Nojima, "Evolutionary many-objective optimization: a short review," in *Proc. IEEE CEC*, Hong Kong, 2008, pp. 2424-2431.
- [2] N. Beume, B. Naujoks, and M. Emmerich, "SMS-EMOA: Multiobjective selection based on dominated hypervolume," *Eur. J. Oper. Res.*, vol. 181, no. 3, pp. 1653-1669, Sep. 2007.
- [3] M. Farina and P. Amato, "A fuzzy definition of 'optimality' for many-criteria optimization problems," *IEEE Trans. Syst. Man. Cybern., Part A: Syst. Hum.*, vol. 34, no. 3, pp. 315-326, May 2004.
- [4] H. Wang, L. Jiao, and X. Yao, "Two_Arch2: An improved two-archive algorithm for many-objective optimization," *IEEE Trans. Evol. Comput.*, vol. 19, no. 4, pp. 524-541, Aug. 2015.
- [5] M. Asafuddoula, T. Ray, and R. Sarker, "A decomposition-based evolutionary algorithm for many objective optimization," *IEEE Trans. Evol. Comput.*, vol. 19, no. 3, pp. 445-460, Jun. 2015.
- [6] Q. Zhang and H. Li, "MOEA/D: A multiobjective evolutionary algorithm based on decomposition," *IEEE Trans. Evol. Comput.*, vol. 11, pp. 712-731, Dec. 2007.
- [7] H. Li and Q. Zhang, "Multiobjective optimization problems with complicated Pareto sets, MOEA/D and NSGA-II," *IEEE Trans. Evol. Comput.*, vol. 13, no. 2, pp. 284-302, Apr. 2009.
- [8] H.-L. Liu, F. Gu, and Q. Zhang, "Decomposition of a multiobjective optimization problem into a number of simple multiobjective subproblems," *IEEE Trans. Evol. Comput.*, vol. 18, no. 3, pp. 450-455, Jun. 2014.
- [9] H.-L. Liu, L. Chen, K. Deb, and E. D. Goodman, "Investigating the effect of imbalance between convergence and diversity in evolutionary multiobjective algorithms," *IEEE Trans. Evol. Comput.*, vol. 21, no. 3, pp. 408-425, Jun. 2017.
- [10] Q. Zhang, A. Zhou, and Y. Jin, "RM-MEDA: A regularity model-based multiobjective estimation of distribution algorithm," *IEEE Trans. Evol. Comput.*, vol. 12, no. 1, pp. 41-63, Feb. 2008.
- [11] K. Miettinen, *Nonlinear Multiobjective Optimization*. Boston, MA: Kluwer, 1999.
- [12] A. L. Jaimes, C. A. C. Coello, and D. Chakraborty, "Objective reduction using a feature selection technique," in *Proc. GECCO 2008*, Atlanta Ga USA, pp. 673-680, 2008.
- [13] A. L. Jaimes, C. A. C. Coello, and J. E. U. Barrientos, "Online objective reduction to deal with many-objective problems," in *Proc. EMO*, 2009, pp. 423-437.
- [14] D. K. Saxena, J. A. Duro, A. Tiwari, K. Deb, and Q. Zhang, "Objective reduction in many-objective optimization: Linear and nonlinear algorithms," *IEEE Trans. Evol. Comput.*, vol. 17, no. 1, pp. 77-99, Feb. 2013.
- [15] Y. M. Cheung, F. Gu, and H.-L. Liu, "Objective extraction for many-objective optimization problems: Algorithm and test problems," *IEEE Trans. Evol. Comput.*, In press, 2016.
- [16] K. Deb and D. K. Saxena, "Searching for Pareto-optimal solutions through dimensionality reduction for certain large-dimensional multi-objective optimization problems," in *Proc. IEEE CEC*, 2006, pp. 3353-3360.
- [17] D. K. Saxena and K. Deb, "Non-linear dimensionality reduction procedures for certain large-dimensional multiobjective optimization problems: Employing correntropy and a novel maximum variance unfolding," in *Proc. EMO*, 2007, pp. 772-787.
- [18] H. K. Singh, A. Isaacs, and T. Ray, "A pareto corner search evolutionary algorithm and dimensionality reduction in many-objective optimization problems," *IEEE Trans. Evol. Comput.*, vol. 15, no. 4, pp. 539-556, Aug. 2011.
- [19] D. Brockhoff and E. Zitzler, "Are all objectives necessary? On dimensionality reduction in evolutionary multiobjective optimization," in *Proc. PPSN*, vol. 9, 2006, pp. 533-542.
- [20] D. Brockhoff and E. Zitzler, "Objective reduction in evolutionary multiobjective optimization: Theory and applications," in *Evol. Comput.*, vol. 17, no. 2, pp. 135-166, Summer 2009.
- [21] K. Li, K. Deb, Q. Zhang, and S. Kwong, "An evolutionary many-objective optimization algorithm based on dominance and decomposition," *IEEE Trans. Evol. Comput.*, vol. 19, no. 5, pp. 694-716, Oct. 2015.
- [22] Y. Yuan, H. Xu, B. Wang, and X. Yao, "A new dominance relation based evolutionary algorithm for many-objective optimization," *IEEE Trans. Evol. Comput.*, vol. 20, no. 1, pp. 16-37, Feb. 2016.
- [23] R. Cheng, Y. Jing, M. Olhofer and B. Sendhoff, "A reference vector guided evolutionary algorithm for many-objective optimization," *IEEE Trans. Evol. Comput.*, vol. 20, no. 5, pp. 773-791, Oct. 2016.
- [24] K. Deb and H. Jain, "An evolutionary many-objective optimization algorithm using reference-point based non-dominated sorting approach, part II: handling constraints and extending to an adaptive approach," *IEEE Trans. Evol. Comput.*, vol. 18, no. 4, pp. 601-622, Aug. 2015.
- [25] H. Jain and K. Deb, "An improved adaptive approach for elitist non-dominated sorting genetic algorithm for many-objective optimization," in *proc. EMO*, 2013, pp. 307-321.
- [26] H.-L. Liu, L. Chen, Q. Zhang and K. Deb, "An evolutionary many-objective optimisation algorithm with adaptive region decomposition," in *Proc. IEEE CEC*, 2016, pp. 4763-4769.
- [27] Y. Qi, X. Ma, F. Liu, L. Jiao, J. Sun, and J. Wu, "MOEA/D with adaptive weight adjustment," *Evol. Comput.*, vol. 22, no. 2, pp. 231-264, Summer 2014.
- [28] K. Deb and H. Jain, "An evolutionary many-objective optimization algorithm using reference-point based non-dominated sorting approach, Part I: Solving problems with box constraints," *IEEE Trans. Evol. Comput.*, vol. 18, no. 4, pp. 577-601, Aug. 2015.
- [29] S. Yang, M. Li, X. Liu, and J. Zheng, "A grid-based evolutionary algorithm for many-objective optimization," *IEEE Trans. Evol. Comput.*, vol. 17, no. 5, pp. 721-736, Oct. 2013.
- [30] J. Bader and E. Zitzler, "HypE: An algorithm for fast hypervolume-based many-objective optimization," *Evol. Comput.*, vol. 19, no. 1, pp. 45-76, Spring 2011.

- [31] I. Das and J. Dennis, "Normal-boundary intersection: A new method for generating the Pareto surface in nonlinear multicriteria optimization problems," *SIAM J. Optim.*, vol. 8, no. 3, pp. 631-657, Jul. 1998.
- [32] H. Sato, "Inverted PBI in MOEA/D and its impact on the search performance on multi and many-objective optimization," in *Proc. ACM GECCO*, 2014, pp. 645-652.
- [33] H. Ishibuchi, Y. Setoguchi, H. Masuda , and Y. Nojima, " Performance of decomposition-based many-objective algorithms strongly depends on Pareto front shapes," *IEEE Trans. Evol. Comput.*, In press, 2016.
- [34] K. Deb, L. Thiele, M. Laumanns, and E. Zitzler, "Scalable test problems for evolutionary multi-objective optimization," Kanpur, India: Kanpur Genetic Algorithms Lab. (KanGAL), Indian Inst. Technol., 2001.
- [35] S. Huband, P. Hingston, L. Barone, and L. While, "A review of multiobjective test problems and a scalable test problem toolkit," *IEEE Trans. Evol. Comput.*, vol. 10, no. 5, pp. 477-506, Oct. 2006.
- [36] E. Zitzler, L. Thiele, M. Laumanns, C. M. Fonseca, and V. G. da Fonseca, "Performance assessment of multiobjective optimizers: An analysis and review," *IEEE Trans. Evol. Comput.*, vol. 7, no. 2, pp. 117-132, Apr. 2003.
- [37] E. Zitzler and L. Thiele, "Multiobjective evolutionary algorithms: A comparative case study and the strength Pareto approach," *IEEE Trans. Evol. Comput.*, vol. 3, no. 4, pp. 257-271, Nov. 1999.
- [38] K. Deb and R. B. Agrawal, "Simulated binary crossover for continuous search space," *Complex Syst.*, vol. 9, no. 2, pp. 115-148, Nov. 1995.
- [39] K. Deb, "Multi-objective optimization using evolutionary algorithms," Chichester, London: Wiley.
- [40] H. Seada and K. Deb, "A unified evolutionary optimization procedure for single, multiple, and many objectives," *IEEE Trans. Evol. Comput.*, vol. 20, no. 3, pp. 358-369, Jun. 2016.

APPENDIX TEST PROBLEMS

A. Degenerated MaOPs with disconnected PFs

In addition to MaOP1 described in the text, we suggest six more test problems.

$$\text{MaOP2: } \begin{cases} f_1(\mathbf{x}) = (1 + g(\mathbf{x}_{II})) \cos\left(\frac{\pi x_1}{2}\right) \cos\left(\frac{\pi x_2}{2}\right) \\ f_2(\mathbf{x}) = (1 + g(\mathbf{x}_{II})) + h(x_2) \\ f_3(\mathbf{x}) = (1 + g(\mathbf{x}_{II})) 4 \cos\left(\frac{\pi x_1}{2}\right) \sin\left(\frac{\pi x_2}{2}\right) \\ f_4(\mathbf{x}) = (1 + g(\mathbf{x}_{II})) \sin\left(\frac{\pi x_1}{2}\right) \\ f_{r_i}(\mathbf{x}) = q_{r_i}(\mathbf{x}_{II}) t_{r_i}(\mathbf{x}_I) \quad (i = 4, \dots, 10) \end{cases}$$

where

$$\begin{aligned} \mathbf{x} &= (\mathbf{x}_I, \mathbf{x}_{II}), \mathbf{x}_I = (x_1, x_2), \mathbf{x}_{II} = (x_3, x_4, \dots, x_n), \\ r_1 &= 1, r_2 = 2, r_3 = 4, \\ g(\mathbf{x}_{II}) &= \sum_{i=3}^n |x_i - x_1 x_2|^2, \\ h(x_1) &= \max(0, -1.4(\cos(2\pi x_1))), \\ q_i(\mathbf{x}_{II}) &= \exp(|x_i - x_1 x_2|^2) - 1, \quad i = 4, \dots, 10. \end{aligned}$$

Its PF is a disconnected spherical surface in the reduced objective space of f_1 , f_2 , and f_4 satisfying $f_1^2 + f_2^2/16 + f_4^2 = 1$ when $0 \leq f_2 \leq 4 \sin(\pi/8)$ or $4 \sin(3\pi/8) \leq f_2 \leq 4$. Its PS is $x_i = x_1 x_2$ for $0 \leq x_1 \leq 0.25$ or $0.75 \leq x_1 \leq 1$, $0 \leq x_2 \leq 1$ and for $i = 3, \dots, 10$. Its ideal point is $(0, 0, 0, \dots, 0)$, and its nadir point is $(1, 4, 0, 1, 0, \dots, 0)$.

$$\text{MaOP3: } \begin{cases} f_1(\mathbf{x}) = (1 + g(\mathbf{x}_{II})) + h(x_1)) x_1 x_2 \\ f_2(\mathbf{x}) = (1 + g(\mathbf{x}_{II})) + h(x_1)) 2(1 - x_1)x_2 \\ f_3(\mathbf{x}) = (1 + g(\mathbf{x}_{II})) + h(x_1)) 6(1 - x_2) \\ f_{r_i}(\mathbf{x}) = q_{r_i}(\mathbf{x}_{II}) t_{r_i}(\mathbf{x}_I) \quad (i = 4, \dots, 10) \end{cases}$$

where

$$\begin{aligned} \mathbf{x} &= (\mathbf{x}_I, \mathbf{x}_{II}), \mathbf{x}_I = (x_1, x_2), \mathbf{x}_{II} = (x_3, x_4, \dots, x_n), \\ r_1 &= 1, r_2 = 2, r_3 = 5, \\ g(\mathbf{x}_{II}) &= \sum_{i=3}^n |x_i - x_1 x_2|^2, \\ h(x_1) &= \max(0, 1.4 \sin(4\pi x_1)), \\ q_i(\mathbf{x}_{II}) &= \exp(|x_i - x_1 x_2|^2) - 1, \quad i = 4, \dots, 10. \end{aligned}$$

Its PF is a disconnected plane surface in the reduced objective space of f_1 , f_2 , and f_5 satisfying $f_1 + f_2/2 + f_5/6 = 1$

when $0 \leq f_5 \leq 1.5$ $3 \leq f_5 \leq 4.5$ or $f_5 = 6$. Its PS is $x_i = x_1 x_2$ for $x_1 = 0$, $0.25 \leq x_1 \leq 0.5$ or $0.75 \leq x_1 \leq 1$, $0 \leq x_2 \leq 1$ and for $i = 3, \dots, 10$. Its ideal point is $(0, 0, 0, \dots, 0)$, and its nadir point is $(1, 2, 0, 0, 6, 0, \dots, 0)$.

$$\text{MaOP4: } \begin{cases} f_2(\mathbf{x}) = (1 + g(\mathbf{x}_{II})) x_1 \\ f_3(\mathbf{x}) = (1 + g(\mathbf{x}_{II})) x_2 \\ f_4(\mathbf{x}) = (1 + g(\mathbf{x}_{II})) t_{r_3}(\mathbf{x}_I) \\ f_{r_i}(\mathbf{x}) = q_{r_i}(\mathbf{x}_{II}) t_{r_i}(\mathbf{x}_I) \quad (i = 4, \dots, 10) \end{cases}$$

where

$$t_3(\mathbf{x}_I) = 2 - \sum_{i=1}^2 \frac{x_i}{2} (1 + \sin(3\pi x_i)),$$

$$r_1 = 2, r_2 = 3, r_3 = 4,$$

$$\mathbf{x} = (\mathbf{x}_I, \mathbf{x}_{II}), \mathbf{x}_I = (x_1, x_2), \mathbf{x}_{II} = (x_3, x_4, \dots, x_n),$$

$$g(\mathbf{x}_{II}) = \sum_{i=3}^n |x_i - x_1 x_2|^2,$$

$$q_i(\mathbf{x}_{II}) = \exp(|x_i - x_1 x_2|^2) - 1, \quad i = 4, \dots, 10.$$

Its PF is composed by four disconnected surfaces in the reduced objective space of f_2 , f_3 , and f_4 . Its PS is $x_i = x_1 x_2$ for $0 \leq x_1, x_2 \leq 1$ and for $i = 3, \dots, 10$. Its ideal point is $(0, 0, 0, \dots, 0)$, and its nadir point is $(0, 1, 1, 1, 0, \dots, 0)$.

MaOP5:

$$\begin{cases} f_2(\mathbf{x}) = (1 + g(\mathbf{x}_{II})) 2t_{r_1}(\mathbf{x}_I) \\ f_3(\mathbf{x}) = (1 + g(\mathbf{x}_{II})) 2t_{r_2}(\mathbf{x}_I) \\ f_5(\mathbf{x}) = (1 + g(\mathbf{x}_{II})) 2t_{r_3}(\mathbf{x}_I) \\ f_2(\mathbf{x}) = (1 + g(\mathbf{x}_{II})) t_{r_1}(\mathbf{x}_I) \\ f_3(\mathbf{x}) = (1 + g(\mathbf{x}_{II})) t_{r_2}(\mathbf{x}_I) \\ f_5(\mathbf{x}) = (1 + g(\mathbf{x}_{II})) t_{r_3}(\mathbf{x}_I) \\ f_i(\mathbf{x}) = q_i(\mathbf{x}_{II}) t_i(\mathbf{x}_I) \quad (i = 4, \dots, 10) \end{cases} \begin{cases} \text{if } x_1 \in [0.2, 0.8] \\ x_2 \in [0.2, 0.8] \end{cases}$$

where

$$\mathbf{x} = (\mathbf{x}_I, \mathbf{x}_{II}), \mathbf{x}_I = (x_1, x_2), \mathbf{x}_{II} = (x_3, x_4, \dots, x_n),$$

$$r_1 = 2, r_2 = 3, r_3 = 5,$$

$$g(\mathbf{x}_{II}) = \sum_{i=3}^n |x_i - 0.5|^2,$$

$$q_i(\mathbf{x}_{II}) = \exp(|x_i - 0.5|^2) - 1, \quad i = 4, \dots, 10,$$

$$t_1(\mathbf{x}_I) = (1 - \cos(\frac{\pi x_1}{2}))(1 - \cos(\frac{\pi x_2}{2})),$$

$$t_2(\mathbf{x}_I) = (1 - \cos(\frac{\pi x_1}{2}))(1 - \sin(\frac{\pi x_2}{2})),$$

$$t_3(\mathbf{x}_I) = (1 - \sin(\frac{\pi x_1}{2})).$$

Its PF is a disconnected spherical surface in the reduced objective space of f_2 , f_3 , and f_5 satisfying $f_2^2 + f_3^2 + f_5^2 = 1$ when $0 < f_3 < 1 - \cos(\pi/10)$ or $(1 - \sin(\pi/10)) < f_3 < (1 - \cos(2\pi/5))$ and $0 < f_5 < 1 - \sin(2\pi/5)$ or $1 - \sin(\pi/10) < f_5 < 1$. Its PS is $x_i = 0.5$ for $x_1 < 0.2$ or $x_1 > 0.8$, $x_2 < 0.2$ or $x_2 > 0.8$ and for $i = 3, \dots, 10$. Its ideal point is $(0, 0, 0, \dots, 0)$, and its nadir point is $(0, 1, 1, 0, 1, 0, \dots, 0)$.

$$\text{MaOP6: } \begin{cases} f_2(\mathbf{x}) = (1 + g(\mathbf{x}_{II})) + h(x_1, x_2)) x_1 x_2 \\ f_4(\mathbf{x}) = (1 + g(\mathbf{x}_{II})) + h(x_1, x_2)) x_1 (1 - x_2) \\ f_6(\mathbf{x}) = (1 + g(\mathbf{x}_{II})) + h(x_1, x_2)) (1 - x_1) \\ f_{r_i}(\mathbf{x}) = q_{r_i}(\mathbf{x}_{II}) t_{r_i}(\mathbf{x}_I) \quad (i = 4, \dots, 10) \end{cases}$$

where

$$\mathbf{x} = (\mathbf{x}_I, \mathbf{x}_{II}), \mathbf{x}_I = (x_1, x_2), \mathbf{x}_{II} = (x_3, x_4, \dots, x_n),$$

$$r_1 = 2, r_2 = 4, r_3 = 6,$$

$$g(\mathbf{x}_{II}) = \sum_{i=3}^n |x_i - \frac{x_1 + x_2}{2}|^2,$$

$$h(x_1, x_2) = \begin{cases} 0 & \text{if } |x_1(x_2 - 0.5)| < 0.0005 \\ \max(0, -1.4 \sin(2\pi x_1)) & \text{else} \end{cases},$$

$$q_i(\mathbf{x}_{II}) = \exp(|x_i - \frac{x_1 + x_2}{2}|^2) - 1, \quad i = 4, \dots, 10.$$

Its PF is composed by a plane surface and a curve which are approximately connected in the reduced objective space of f_2 , f_4 , and f_6 satisfying $f_2 + f_4 + f_6 = 1$. Its PS is $x_i = \frac{x_1 + x_2}{2}$ and $h(x_1, x_2) = 0$ for $0 \leq x_1, x_2 \leq 1$ and for $i =$

3, ..., 10. Its ideal point is $(0, 0, 0, \dots, 0)$, and its nadir point is $(0, 1, 0, 1, 0, 1, 0, \dots, 0)$.

MaOP7:

$$\begin{cases} f_2(\mathbf{x}) = (1 + g(\mathbf{x}_{II}) + h(x_1, x_2)) \cos\left(\frac{\pi x_1}{2}\right) \cos\left(\frac{\pi x_2}{2}\right) \\ f_7(\mathbf{x}) = (1 + g(\mathbf{x}_{II}) + h(x_1, x_2)) \cos\left(\frac{\pi x_1}{2}\right) \sin\left(\frac{\pi x_2}{2}\right) \\ f_8(\mathbf{x}) = (1 + g(\mathbf{x}_{II}) + h(x_1, x_2)) \sin\left(\frac{\pi x_1}{2}\right) \\ f_{r_i}(\mathbf{x}) = q_{r_i}(\mathbf{x}_{II}) t_{r_i}(\mathbf{x}_I) \quad (i = 4, \dots, 10) \end{cases}$$

where

$$\mathbf{x} = (\mathbf{x}_I, \mathbf{x}_{II}), \mathbf{x}_I = (x_1, x_2), \mathbf{x}_{II} = (x_3, x_4, \dots, x_n),$$

$$r_1 = 2, r_2 = 7, r_3 = 8,$$

$$g(\mathbf{x}_{II}) = \sum_{i=3}^n |x_i - \frac{x_1+x_2}{2}|^2,$$

$$h(x_1, x_2) = \begin{cases} 0 & \text{if } |\cos\left(\frac{\pi x_1}{2}\right)(\cos\left(\frac{\pi x_2}{2}\right) - \sin\left(\frac{\pi x_2}{2}\right))| < 0.01 \\ \max(0, 1.4\sin(2\pi x_1)) & \text{else} \end{cases},$$

$$q_i(\mathbf{x}_{II}) = \exp(|x_i - \frac{x_1+x_2}{2}|^2) - 1, \quad i = 4, \dots, 10.$$

Its PF is composed by two approximately connected spherical surfaces in the reduced objective space of f_2 , f_7 , and f_8 satisfying $f_2^2 + f_7^2 + f_8^2 = 1$. Its PS is $x_i = \frac{x_1+x_2}{2}$ and $h(x_1, x_2) = 0$ for $0 \leq x_1, x_2 \leq 1$ and for $i = 3, \dots, 10$. Its ideal point is $(0, 0, 0, \dots, 0)$, and its nadir point is $(0, 1, 0, \dots, 0, 1, 1, 0, 0)$.

B. Non-degenerated MaOPs with disconnected PFs

Non-degenerated MaOPs can also have disconnected PFs, and thus we construct the following two non-degenerated MaOPs with disconnected PFs for experimental study. Their search space is $[0, 1]^n$ ($n = 20$), the number of objectives is $m = 10$, and the two instances are for minimization.

$$\text{MaOP8: } \begin{cases} f_1(\mathbf{x}) = (1 + g(\mathbf{x}_{II}) + h(\mathbf{x}_I)) \prod_{i=1}^{m-1} x_i \\ f_2(\mathbf{x}) = (1 + g(\mathbf{x}_{II}) + h(\mathbf{x}_I)) \prod_{i=1}^{m-2} x_i (1 - x_{m-1}) \\ f_3(\mathbf{x}) = (1 + g(\mathbf{x}_{II}) + h(\mathbf{x}_I)) \prod_{i=1}^{m-3} x_i (1 - x_{m-2}) \\ \dots \\ f_m(\mathbf{x}) = (1 + g(\mathbf{x}_{II}) + h(\mathbf{x}_I)) (1 - x_1) \end{cases}$$

where

$$\mathbf{x} = (\mathbf{x}_I, \mathbf{x}_{II}), \mathbf{x}_I = (x_1, \dots, x_{m-1}), \mathbf{x}_{II} = (x_m, \dots, x_n),$$

$$g(\mathbf{x}_{II}) = \sum_{j=m}^n |x_j - 0.5|^2,$$

$$h(\mathbf{x}_I) = \sum_{i=1}^{m-2} \max(0, -(\cos(2\pi x_i) - 0.5)).$$

Its PF is a disconnected hyperplane surface satisfying $f_1 + f_2 + \dots + f_m = 1$, where $0 \leq f_i \leq 1/6$ or $(5/6)^{m-i} \leq f_i \leq 1$, $i = 1, \dots, m$. Its PS is $x_j = 0.5$ for $j = m, \dots, n$. Its ideal point is $(0, 0, 0, \dots, 0)$, and its nadir point is $(1, 1, 1, \dots, 1)$.

MaOP9:

$$\begin{cases} f_1(\mathbf{x}) = (1 + g(\mathbf{x}_{II}) + h(\mathbf{x}_I)) \prod_{i=1}^{m-2} \cos\left(\frac{\pi x_{m-1}}{2}\right) \\ f_2(\mathbf{x}) = (1 + g(\mathbf{x}_{II}) + h(\mathbf{x}_I)) \prod_{i=1}^{m-2} \cos\left(\frac{\pi x_i}{2}\right) \sin\left(\frac{\pi x_{m-1}}{2}\right) \\ f_3(\mathbf{x}) = (1 + g(\mathbf{x}_{II}) + h(\mathbf{x}_I)) \prod_{i=1}^{m-3} \cos\left(\frac{\pi x_i}{2}\right) \sin\left(\frac{\pi x_{m-2}}{2}\right) \\ \dots \\ f_m(\mathbf{x}) = (1 + g(\mathbf{x}_{II}) + h(\mathbf{x}_I)) \sin\left(\frac{\pi x_1}{2}\right) \end{cases}$$

where

$$\mathbf{x} = (\mathbf{x}_I, \mathbf{x}_{II}), \mathbf{x}_I = (x_1, \dots, x_{m-1}), \mathbf{x}_{II} = (x_m, \dots, x_n),$$

$$g(\mathbf{x}_{II}) = \sum_{j=m}^n |x_j - 0.5|^2,$$

$$h(\mathbf{x}_I) = \sum_{i=1}^{m-2} \max(0, -\cos(2\pi x_i)).$$

Its PF is a disconnected hyperspherical surface satisfying $f_1^2 + f_2^2 + \dots + f_m^2 = 1$, where $0 \leq f_i \leq \sin(\pi/8)$ or $(\sin(3\pi/8))^{m-i} \leq f_i \leq 1$, $i = 1, \dots, m$. Its PS is $x_j = 0.5$ for

$j = m, \dots, n$. Its ideal point is $(0, 0, 0, \dots, 0)$, and its nadir point is $(1, 1, 1, \dots, 1)$.



Hai-Lin Liu is currently a Professor of School of applied mathematics at the Guangdong University of Technology. He received the B.S. degree in mathematics from Henan Normal University, Xinxiang, China, the M.S degree in applied mathematics from Xidian University, Xi'an, China, the Ph.D. degree in control theory and engineering from South China University of Technology, Guangzhou, China, and Post-doctor in the Institute of Electronic and Information, South China University of Technology, Guangzhou, China. His research interests include evolutionary computation and optimization, wireless network planning and optimization, and their applications.



Lei Chen received the B.S. degree in mathematics and applied mathematics from Henan University, Kaifeng, China, in 2011. He is currently pursuing the Ph.D. degree in the School of Automation, Guangdong University of Technology, Guangzhou, China. He is now a visiting Ph.D. student in the Department of Electrical and Computer Engineering, Michigan State University, USA. His current research interests include evolutionary multi-objective optimization and its applications.



Qingfu Zhang (M'01-SM'06-F'17) received the BSc degree in mathematics from Shanxi University, China in 1984, the MSc degree in applied mathematics and the Ph.D degree in information engineering from Xidian University, China, in 1991 and 1994, respectively. He is a Professor at the Department of Computer Science, City University of Hong Kong, Hong Kong, and a Changjiang Visiting Chair Professor in Xidian University, China. His main research interests include evolutionary computation, optimization, neural networks, data analysis, and their applications. He is currently leading the Metaheuristic Optimization Research (MOP) Group in City University of Hong Kong.

Dr. Zhang is an Associate Editor of the IEEE Transactions on Evolutionary Computation and the IEEE Transactions on Systems, Man, and Cybernetics-Part B. He was awarded the 2010 IEEE Transactions on Evolutionary Computation Outstanding Paper Award. He is a Thomson Reuters' 2016 highly cited researcher in Computer Science.



Kalyanmoy Deb (F'13) is Koenig Endowed Chair Professor at Department of Electrical and Computer Engineering in Michigan State University, USA. Prof. Deb's research interests are in evolutionary optimization and their application in optimization, modeling, and machine learning. He was awarded Infosys Prize, TWAS Prize in Engineering Sciences, CajAsturMamdanji Prize, Distinguished Alumni Award from IIT Kharagpur, Edgeworth-Pareto award, Bhatnagar Prize in Engineering Sciences, and Bessel Research award from Germany. He is fellow of IEEE, ASME, and three Indian science and engineering academies. He has published over 435 research papers with Google Scholar citation of 102,000 with h-index 102. He is in the editorial board on 20 major international journals. More information about his research contribution can be found from <http://www.coin-laboratory.com>.