Large-Scale Evolutionary Multiobjective Optimization Assisted by Directed Sampling

Shufen Qin, Chaoli Sun, *Member, IEEE*, Yaochu Jin , *Fellow, IEEE*,
Ying Tan, and Jonathan Fieldsend , *Member, IEEE*

Abstract—It is particularly challenging for evolutionary algorithms to quickly converge to the Pareto front in large-scale multiobjective optimization. To tackle this problem, this article proposes a large-scale multiobjective evolutionary algorithm assisted by some selected individuals generated by directed sampling (DS). At each generation, a set of individuals closer to the ideal point is chosen for performing a DS in the decision space, and those nondominated ones of the sampled solutions are used to assist the reproduction to improve the convergence in evolutionary large-scale multiobjective optimization. In addition, elitist nondominated sorting is adopted complementarily for environmental selection with a reference vector-based method in order to maintain diversity of the population. Our experimental results show that the proposed algorithm is highly competitive on large-scale multiobjective optimization test problems with up to 5000 decision variables compared to five state-of-the-art multiobjective evolutionary algorithms.

Index Terms—Directed sampling (DS), evolutionary multiobjective optimization, large-scale multiobjective problems (LSMOPs), nondominated sorting, reference vectors.

I. INTRODUCTION

A LARGE number of multiobjective evolutionary algorithms (MOEAs) [1]–[8] have been proposed in the past few decades for solving multiobjective optimization problems (MOPs). They can roughly be divided into four classes [9], [10]: 1) Pareto dominance-based MOEAs [11]–[22]; 2) decomposition-based MOEAs [23]–[28];

3) performance indicator-based MOEAs [29]–[38]; and finally, 4) those MOEAs that do not fall in the previous three categories [39]–[42]. In recent years, most research on MOEAs [7], [17]–[21], [28], [36], [39], [40], [43]–[47] has focused on the scalability of the performance to the increase in the number of objectives, and little attention has been paid to the scalability of the performance in the decision space. However, many real-world MOPs may have hundreds or even thousands of decision variables, which are known as large-scale optimization problems [48]. For example, a flight safety system has over 1500 parameters [49], and in the search of optimal structure of deep neural networks, the numbers of weight and structure parameters dramatically increase as the number of hidden nodes and the number of hidden layers increase [50], [51].

The performance of classic metaheuristic algorithms often deteriorates as the dimension of the decision space increases [52], which is called curse of dimensionality [53]. Although large-scale single-objective optimization has been a popular research topic for many years [54]–[58], research on large-scale multiobjective optimization is still in its early stage [52], [59]–[68]. Generally, existing approaches to large-scale multiobjective problems (LSMOPs) can be roughly classified into four categories.

1) *Cooperative Coevolution (CC) Framework-Based MOEAs:* Antonio and Coello Coello applied a differential evolution algorithm (GDE3) in the CC framework for solving large-scale many-objective problems [59] and then they further proposed to combine MOEA/D with coevolutionary techniques for decomposition in both objective and decision spaces [61]. Li and Wei [64] also proposed a cooperative coevolutionary algorithm, in which a fast interdependency identification grouping method was utilized, for LSMOPs.

2) *Decision Variable Clustering-Based MOEAs:* Ma *et al.* [62] divided the decision variables into three groups according to their properties related to convergence and diversity and proposed a multiobjective evolutionary algorithm based on decision variable analyses (MOEA/DVA). Zhang *et al.* [52] extended the idea of MOEA/DVA [62] and proposed an evolutionary algorithm for large-scale many-objective optimization (LMEA) by dividing the decision variables more precisely and introducing different search and selection strategies for different groups. Liu *et al.* [69] proposed to group large-scale decision variables into

two categories, i.e., convergence-related and diversity-related variables. Then, principal components analysis (PCA) is performed to reduce the dimension of the convergence-related variables, which will further be decomposed into a number of subproblems based on interdependence analysis for optimization. Chen *et al.* also separated the variables into convergence-related and diversity-related clusters and proposed to optimize subproblems in parallel [63] and sequentially [70], respectively. Note, however, that such algorithms often require a large number of fitness evaluations to achieve adequately precise clustering of the decision variables.

3) *Problem Transformation-Based MOEAs:* Zille *et al.* [65] proposed a framework for large-scale multiobjective optimization (WOF) using a problem transformation technique to reduce the search space, which can serve as a genetic method for any population-based meta-heuristic algorithms. Following that, He *et al.* [66] implemented a dimension reduction method by optimizing a set of weight variables along different directions in the decision space. Most recently, He *et al.* [71] further proposed an adaptive offspring generation method for large-scale optimization problems. Liu *et al.* [72] proposed to use a random dynamic grouping instead of ordered grouping to improve the performance of the WOF framework [65].

4) *New Learning Strategies-Based MOEAs:* Not so many methods falling in this category have been proposed. Tian *et al.* [67] proposed a modified competitive swarm optimizer for LSMOPs, called LMOCSO, by generating promising offspring to accelerate the search of the global optimum. Zhang *et al.* [73] utilized the information of the previous positions of the population and built an information feedback model (IFM) to generate an offspring population.

Although the existing large-scale optimization algorithms have showed promising performances, each category of the algorithms has its own drawback. For example, some CC and clustering techniques-based MOEAs need to classify the decision variables into a set of groups, which will cost a great number of objective evaluations in order to detect interactive decision variables. Furthermore, the performance of the CC framework-based MOEAs will seriously deteriorate on inappropriate groupings. In addition, the assumption of separability between groups of decision variables does not always hold. Therefore, the grouping strategies, such as linear grouping, random grouping, or ordered grouping, which do not cost additional objective evaluations to detect interactive decision variables, are not applicable for solving large-scale MOPs with interactions between all decision variables. The new learning strategy-based MOEAs also need to cost a large number of evaluations to obtain better solutions, for example, $15\,000 \times D$ evaluations are required in LMOCSO [67] to obtain better solutions. The problem transformation-based MOEAs are very competitive in enhancing the convergence capability for large-scale optimization, however, they are susceptible to local optima, as pointed out in [65]–[67]. Moreover, the grouping scheme severely restricts the flexibility of the method given in [65].

Based on the above discussions, we can see that it is vital for an MOEA to be able to search in the right direction to achieve a set of good solutions with a limited computational budget for large-scale optimization problems. For this purpose, this work proposes a large-scale evolutionary multiobjective algorithm assisted by directed sampling (DS), termed LMOEA-DS for short. The main idea is to estimate the promising search directions in the decision space, which is similar to the idea presented in [66]. Different to the method in [66] that uses nondominated solutions of the current parent population to generate search directions in the decision space, in this work, by contrast, we use solutions closest to the ideal point in subpopulations divided based on a set of reference vectors. Furthermore, promising solutions are generated by random sampling along the estimated search directions, instead of performing an optimization as done in [66]. The nondominated solutions of the sampled solutions are then used for reproduction. This above process of identifying promising solutions using DS is repeated in each generation. In [66], by contrast, the promising solutions are used to seed the initial population and then an arbitrary existing MOEA can be used to search for a set of nondominated solutions. The main contributions of this work are summarized as follows.

1) At each generation of the evolutionary search, a DS strategy is adopted for generating promising solutions to guide the search. For this purpose, a set of reference vectors is adopted to group the population into a number of clusters. Then, the individual having the minimum distance to the ideal point in each cluster will be used to generate two promising search directions in the decision space. A certain number of solutions are then randomly sampled along the two search directions. After that, the sampled solutions are combined and the nondominated solutions among them are chosen for reproduction. These nondominated solutions are called guiding solutions.

2) To strike a balance between convergence and diversity, two reproduction processes, each of which is followed by the proposed environmental selection, are introduced. In the first reproduction process, the guiding solutions are used to perform crossover with the current parent individuals, followed by the second reproduction, in which no guiding solutions will take part in the generation of offspring individuals.

3) Reference vectors and elitist nondominated sorting are complementarily used in the environmental selection. By default, a set of reference vectors is used to decompose the MOP and both the Euclidean distance and angle-based metric are employed for selection for each subproblem. However, in case the population diversity is lost, i.e., when a large number of reference vectors are empty, the elitist nondominated sorting strategy [12] will replace the reference-based selection, aiming to promote the diversity of the population.
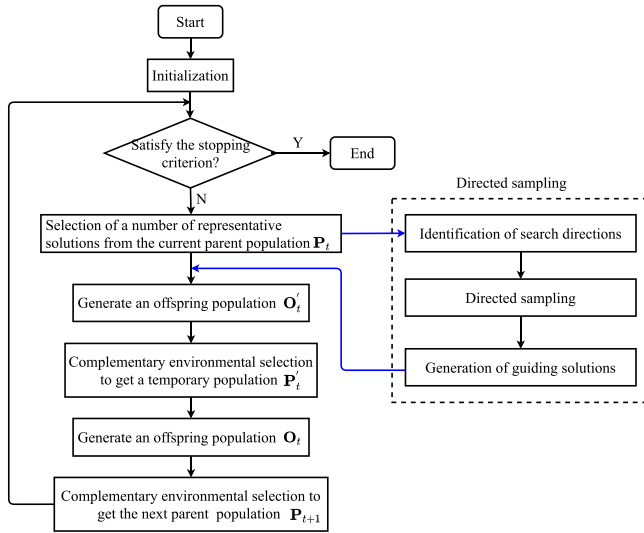
Fig. 1. Diagram of the overall framework of the proposed LMOEA-DS, which consists of three main components, namely, generation of the guiding solutions, double reproduction, and complementary environmental selection.

Section II describes the main framework and details of the proposed LMOEA-DS algorithm. In Section III, the test problems used in our experiments, the performance indicator for measuring the performance of the obtained nondominated solution sets are presented, followed by a description of the experimental settings and the comparative results. Finally, conclusions are drawn and future work is outlined in Section IV.

II. PROPOSED ALGORITHM

Without loss of generality, a multiobjective minimization problem can be formulated as follows:

$$\min \ \mathbf{F}(\mathbf{x}) = (f_1(\mathbf{x}), f_2(\mathbf{x}), \ldots, f_m(\mathbf{x}))$$
$$\text{s.t.} \ \mathbf{x} \in \mathbb{R}^D \tag{1}$$

where m is the number of objectives and $\mathbf{x} = (x_1, x_2, \ldots, x_D)$ is the decision vector, where D is the dimension of the decision space. In this work, we consider biobjective and three-objective problems with a large number of decision variables, which may range from hundreds to thousands [66].

The search space exponentially increases when the number of decision variables increases, dramatically worsening the search performance, in particular, the convergence capability of optimization algorithms. In order to address this issue, we propose a DS strategy in combination with a double-reproduction strategy to accelerate the convergence, together with a new environmental selection method to promote population diversity. Fig. 1 depicts the flowchart of the overall framework of the proposed LMOEA-DS, which is composed of three main components: 1) identification of the guiding solutions; 2) guided double reproduction; and 3) complementary environmental selection. Note that two complementary environmental selection strategies are provided to accelerate convergence while maintaining the diversity of the population. From Fig. 1, we can see that guiding solutions are embedded in the reproduction process in each generation,

which is substantially different from [66], where the promising solutions found are used only for population initialization. In addition, the method for identifying the promising search directions is also different. In the following, we will detail the three main components of LMOEA-DS.

A. Initialization

The population is randomly initialized, like in most MOEAs. In addition, two sets of reference vectors $\mathbf{W}$ and $\mathbf{W}'$ are created, which partition the objective space into subspaces for environmental selection and for generation of the search directions, respectively. Note that the number of reference vectors for generating the search directions should usually be considerably smaller than that for environmental selection to reduce additional computational costs introduced by the DS. Suppose that $N_{w'}$ solutions will be selected for generating the search directions, we simply group $\mathbf{W}$ reference vectors for environmental selection into $N_{w'} - m$ clusters using k-means [74] and then add the m boundary reference vectors in $\mathbf{W}$.

B. Directed Sampling and Generation of Guiding Solutions

Guiding solutions are meant to assist the MOEA to find the right search direction to accelerate the convergence. Here, we hypothesize that solutions that are closest to the ideal point in the current population are able to provide information about the desired search directions. However, it may be harmful to the diversity if only the closest solution for the whole population is used to guide the search. Therefore, we divide the overall search space into a number of subspaces using the reference vector set $\mathbf{W}'$. The method for determining the search directions and guiding solutions will be elaborated as follows.

1) Identification of Search Directions: Recall that the search is performed in the decision space and, consequently, we aim to identify promising search directions in the decision space. To this end, we identify the solutions closest to the ideal point of the objective space at first and then define the promising search directions by setting out a line from the lower or upper bound points to these solutions in the decision space. Note that the objective of using two directions is to increase the chance for the direction vectors to intersect with the Pareto set in the decision space. The experimental results of the methods that only use the direction setting out from $\mathbf{L}$ and that only use the direction setting out from $\mathbf{U}$, respectively, are given and analyzed in Section A of the supplementary material. As previously mentioned, we first group the current parent population into $N_{w'}$ subpopulations by assigning each solution to its closest reference vector of $\mathbf{W}'$. Then, for each reference vector $\mathbf{w}'_i$, where $i = 1, 2, \ldots, N_{w'}$, to which at least one solution is assigned, the solution that has the minimum projected distance to the ideal point is chosen to define two search directions. To illustrate how the search directions are defined, an example is given in Fig. 2(a), where the current population has six solutions denoted by black dots, $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_6$ and $\mathbf{W}'$ has four reference vectors. The blue dotted ellipse in Fig. 2(a) shows that the solutions in each ellipse are assigned to the same reference vector. The green dotted line indicates the distance of
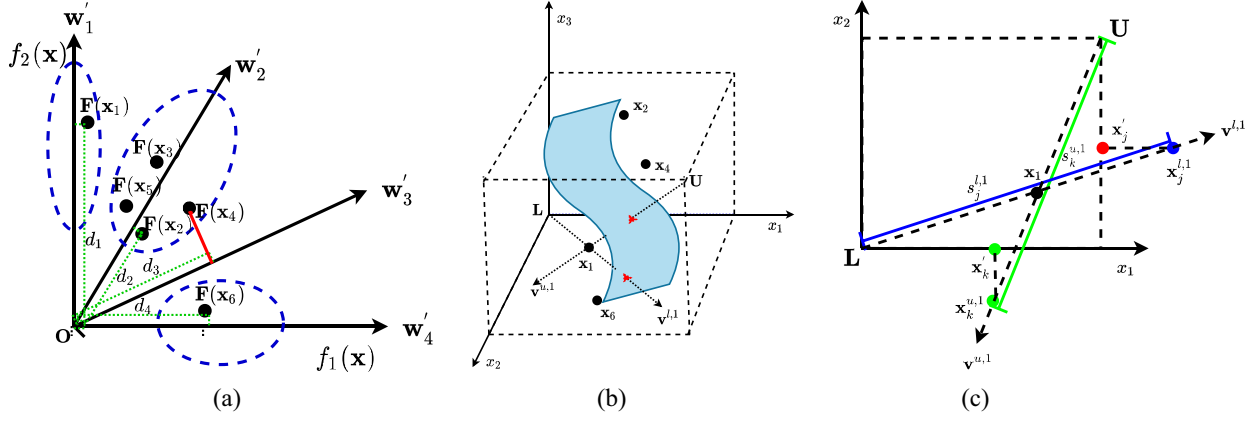
Fig. 2. (a) Solution selection of a solution for each reference vector that has the minimum projected distance to the ideal point in the objective space and (b) define two promising search directions based on the selected solution as well as the lower and upper bound solutions in the decision space. (c) Illustration of repairing randomly sampled infeasible solutions, in which any decision variables that exceed the upper or lower bound are projected back to the feasible solutions.
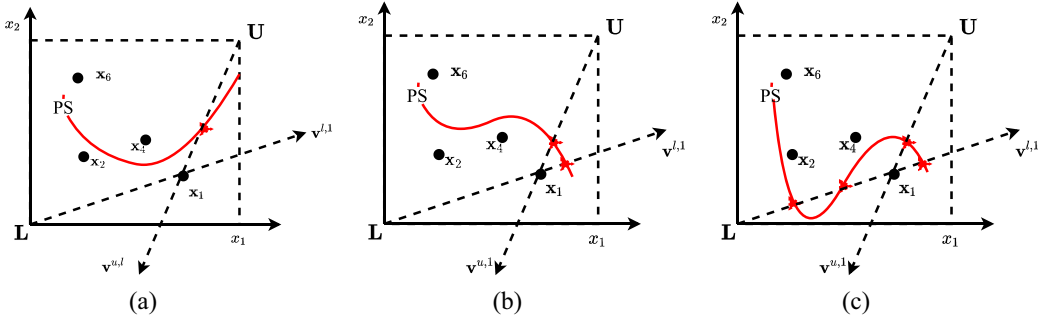

Fig. 3. Three cases of the intersection between a search direction $\mathbf{v}^{l,1}$ and the PS. (a) No solution on the search direction is in the Pareto set. (b) One and only one solution on the search direction is in the Pareto set. (c) More than one solutions on the search direction are in the Pareto set.

a solution to the origin along the related reference vector in the objective space. Note that reference vector $\mathbf{w}'_3$ has no solutions assigned to it in this example. To address this issue, the solution, which is closest to the reference vector $\mathbf{w}'_3$ among all individuals that have not been chosen to define the search directions, will be assigned to $\mathbf{w}'_3$ so that each reference vector has at least one solution assigned to. For example, in Fig. 2(a), vector $\mathbf{w}'_3$ has no solutions assigned to and $\mathbf{x}_4$ is the closest solution to it among $\mathbf{x}_3$, $\mathbf{x}_4$, and $\mathbf{x}_5$, which are not the closest solutions in their own clusters. In this case, solution $\mathbf{x}_4$ will be chosen to define the search directions for vector $\mathbf{w}'_3$.

As a result, $2 \times N_{w'}$ search directions, namely, $\mathbf{v}^{l,1}, \mathbf{v}^{u,1}, \ldots, \mathbf{v}^{l,i}, \mathbf{v}^{u,i}, \ldots, \mathbf{v}^{l,N_{w'}}, \mathbf{v}^{u,N_{w'}}$ are defined, where $\mathbf{v}^{l,i}, i = 1, 2, \ldots, N_{w'}$ and $\mathbf{v}^{u,i}, i = 1, 2, \ldots, N_{w'}$ represent the search directions setting out from the lower and upper bound points $\mathbf{L}$ and $\mathbf{U}$, respectively. Note that the search direction is always a line setting out from the lower or upper boundary point through the chosen solutions regardless the dimension of the search space, which can be defined as follows:

$$\mathbf{v}^{l,i} = \mathbf{x}_j - \mathbf{L}$$
$$\mathbf{v}^{u,i} = \mathbf{x}_j - \mathbf{U}. \tag{2}$$

In (2), $\mathbf{x}_j$ represents the solution that has the minimum distance to the ideal point along the reference vector $\mathbf{w}'_i$ for the ith vector. For example, in Fig. 2(b), $\mathbf{v}^{l,1}$ and $\mathbf{v}^{u,1}$ are two promising search directions defined by solution $\mathbf{x}_1$ in a 3-D decision

space. The blue surface is the true PS and two red solid triangles are the interaction points between the two direction vectors and the PS in a 3-D decision space.

Algorithm 1 presents the pseudocode of the method for defining the search directions. The angles between each individual $i, i = 1, 2, \ldots, |P|$ of the current parent population and all reference vectors in $\mathbf{W}'$ will be calculated, and individual i will be assigned to its closest reference vector that has the minimal angle (lines 2–5). Then, for each reference vector $\mathbf{w}'_i$ that has been assigned at least one individual, the solution having the minimum projected distance to the ideal point in the objective space will be stored in set S (lines 6–11). For a reference vector without any individual being assigned to, its closest individual assigned to another reference vector but is not included in S, will be assigned to this reference vector and then is stored in S (lines 12–17). As a result, a total of $N_{w'}$ solutions will be chosen and stored in S, each of which will be used for defining two search directions.

2) Directed Solution Sampling: Given a set of search directions $\mathbf{V} = \{\mathbf{v}^{l,1}, \mathbf{v}^{u,1}, \ldots, \mathbf{v}^{l,i}, \mathbf{v}^{u,i}, \ldots, \mathbf{v}^{l,N_{w'}}, \mathbf{v}^{u,N_{w'}}\}$, we are now going to generate a set of solutions along each of the search directions to identify solutions that can accelerate convergence. Each search direction is a line setting out from the lower or upper boundary point and, consequently, there will be three different situations for the positional relationship between the line and the Pareto set, i.e., no intersection, one intersection, or more than one intersection. Fig. 3 gives an

Algorithm 1: Identification of Promising Search Directions

Input:
    $\mathbf{W}'$: a set of reference vectors for identifying search directions;
    P_t: the current parent population of LMOEA-DS;
    $N_{w'}$: the number of solutions provided by LMOEA-DS;
Output:
    $\mathbf{V}$: a set of search directions;

1 $S = \emptyset$;
2 **for** $i = 1 : |P_t|$ **do**
3   Calculate the angle between individual i and each reference vector of $\mathbf{W}'$;
4   Assign individual i to its closest reference vector that has the minimum angle between this individual and all reference vectors;
5 **end for**
6 **for** $i = 1 : N_{w'}$ **do**
7   **if** *the reference vector $\mathbf{w}'_i$ has been assigned at least one individual* **then**
8   Calculate the distance of each individual assigned this reference vector to the ideal point along its reference vector;
9   Select the individual that has the minimum distance and saved to the set S;
10   **end if**
11 **end for**
12 **for** $i = 1 : N_{w'}$ **do**
13   **if** *the reference vector $\mathbf{w}'_i$ has not been assigned any individual* **then**
14   Sort the individuals of the population except those in S in an ascending order according to the angles to the reference vector $\mathbf{w}'_i$;
15   Select the first individual to be saved to S;
16   **end if**
17 **end for**
18 Define two search directions according to (2) for each individual in the set S;
19 Output the search directions;

example of a 2-D problem to show three cases. Theoretically, we can use an optimization algorithm to find the most promising solution along each search direction. For example, in Fig. 3(a) and (b), we can utilize an optimization method to find a solution that is closest to the Pareto set on each search direction so that it can be used to guide offspring generation in LMOEA-DS. However, when there is more than one solution on a search direction, such as the case in Fig. 3(c), a search algorithm may fail to find the most promising guiding solutions it can easily end up with one solution that is closet to the initial search point. Therefore, in this work, we decide to randomly sample multiple solutions along the search directions and then find the nondominated solutions from all sampled solutions. The benefit of DS over optimization is that DS is able to maintain a higher degree of diversity of the guiding solutions for speeding up the convergence of LMOEA-DS.

Assume solutions $\mathbf{x}^{l,i}_j$ and $\mathbf{x}^{u,i}_k$ are generated along the search directions $\mathbf{v}^{l,i}$ and $\mathbf{v}^{u,i}$, respectively, which can be formulated as follows:

$$\mathbf{x}^{l,i}_j = \mathbf{L} + s^{l,i}_j \cdot \frac{\mathbf{v}^{l,i}}{\|\mathbf{v}^{l,i}\|}$$
$$\mathbf{x}^{u,i}_k = \mathbf{U} + s^{u,i}_k \cdot \frac{\mathbf{v}^{u,i}}{\|\mathbf{v}^{u,i}\|} \qquad (3)$$

where $s^{l,i}_j$ and $s^{u,i}_k$ are non-negative coefficients that determine which solutions along the directions $\mathbf{v}^{l,i}$ and $\mathbf{v}^{u,i}$, respectively, will be generated, j and k represent the jth and kth sampling

on $\mathbf{v}^{l,i}$ and $\mathbf{v}^{u,i}$, respectively. Note that all solutions generated along these two directions are bounded by $\|\mathbf{U} - \mathbf{L}\|$, where $\| \cdot \|$ is the Euclidean distance. In order to ensure the diversity of the generated solutions, the value of $s^{l,i}_j$ and $s^{u,i}_k$ are random numbers uniformly distributed in the interval $[0, \|\mathbf{U} - \mathbf{L}\|]$. The influence of the upper bound of the coefficient on the performance of the proposed algorithm is analyzed in Section B of the supplementary material.

Note, however, that some solutions generated by the sampled coefficients will be outside the upper and lower bounds using $\|\mathbf{U} - \mathbf{L}\|$ defined above. To tackle this problem, all decision variables whose values are outside the lower or upper bound are projected back to the feasible region. Fig. 2(c) gives an illustrative example showing how two solutions generated outside the lower and upper bounds, respectively, are projected back to the feasible region. In Fig. 2(c), $s^{l,1}_j$ and $s^{u,1}_k$ are two coefficients randomly sampled between $[0, \|\mathbf{U} - \mathbf{L}\|]$, resulting two solutions that are outside the feasible region. More specifically, solution $\mathbf{x}^{l,1}_j$ exceeds the upper bound of x_1 and solution $\mathbf{x}^{u,1}_k$ exceeds the lower bound of x_2. Therefore, as shown in Fig. 2(c), solutions $\mathbf{x}^{l,i}_j$ and $\mathbf{x}^{u,1}_k$ will be projected to points $\mathbf{x}'_j$ and $\mathbf{x}'_k$, respectively, in the following way:

$$x'_{j,d} = \min\left\{x^{l,i}_{j,d}, U_d\right\}$$
$$x'_{k,d} = \max\left\{x^{u,i}_{k,d}, L_d\right\} \qquad (4)$$

where i indicates the ith reference vector, j and k represent the jth and kth individuals sampled on the search directions $\mathbf{v}^{l,i}$ and $\mathbf{v}^{u,i}$, respectively, and $d = 1, 2, \ldots, D$ is the dth decision variable.

3) Generation of Guiding Solutions: As described above, a number of solutions will be generated along the two search directions for each reference vector by randomly generating a number of coefficients uniformly distributed between 0 and $\|\mathbf{U} - \mathbf{L}\|$. Assume N_s solutions will be sampled for each search direction of each vector, then altogether $2 \times N_s \times N_{w'}$ solutions will be sampled in total. These solutions are put together and nondominated sorting is performed to find out the nondominated solutions. These nondominated solutions stored in archive S_G will be eventually used as the guiding solutions in the reproduction of the LMOEA-DS for speeding up the convergence.

C. Double Reproduction and Complementary Environmental Selection

In the proposed method, the guiding solutions are used in the reproduction of LMOEA-DS so as to speed up the convergence. However, it is often important in an MOEA to maintain a good balance between convergence and diversity. For this reason, a double reproduction process and a complementary environmental selection strategy are introduced in LMOEA-DS, the details of which are described in the following.

In the first reproduction, each individual in the parent population P_t will be selected to do a crossover with a randomly selected guiding solution, thereby generating $|P_t|$ candidate

TABLE I
CHARACTERISTICS OF LSMOP [75] TEST PROBLEMS

Problem	Characteristics		
	Modality	Shape	Separability
LSMOP1	Unimodal	Linear	Fully Separable
LSMOP2	Mixed	Linear	Partially Separable
LSMOP3	Multi-modal	Linear	Mixed
LSMOP4	Mixed	Linear	Mixed
LSMOP5	Unimodal	Convex	Fully Separable
LSMOP6	Mixed	Convex	Partially Separable
LSMOP7	Multi-modal	Convex	Mixed
LSMOP8	Mixed	Convex	Mixed
LSMOP9	Mixed	Disconnected	Fully Separable

solutions, where $|P_t|$ is the population size. A mutation operator will further be performed on these solutions to get intermediate offspring population O'_t. P_t is then combined with O'_t and the guiding solution set S_G to form a combined population $C'_t = P_t \cup O'_t \cup S_G$, on which the first environmental selection is applied, resulting in an intermediate parent population P'_t. The environmental selection applied here is a complementary selection strategy, which is to be detailed later on. Once the first environmental selection is conducted, a second reproduction operation is implemented by performing crossover on the intermediate parent population P'_t and mutation sequentially to produce the offspring O_t and finally, the parent population P_{t+1} for the next generation will be obtained by performing the same environmental selection on the combined population $C_t = P'_t \cup O_t$.

The pseudocode of the environmental selection adopted in this work is given in Algorithm 2. From Algorithm 2, we can see that the environmental selection strategy complementarily uses the decomposition-based method (lines 4–7) and the dominance-based method (lines 10–12). The objective values of each individual in the combined population C_t will be normalized at first and assigned to their closest reference vector for decomposition-based selection. The decomposition-based environmental selection will be utilized if the number of reference vectors that have been assigned at least one solution to is not less than a threshold N_ϵ. Otherwise, the elitist nondominated sorting proposed in NSGA-II [12] will be employed for environmental selection. The complementary selection strategy ensures that the number of selected offspring will not be too small and, therefore, a certain degree of population diversity can be maintained.

The decomposition-based selection strategy is performed according to the performance measure described in

$$\Psi_i = \frac{\cos\theta_{i,j}}{d_i} \qquad (5)$$

where $\cos\theta_{i,j} = ([\mathbf{F}(\mathbf{x_i}) \cdot \mathbf{w}_j]/[\|\mathbf{F}(\mathbf{x_i})\| \cdot \|\mathbf{w}_j\|])$ represents the cosine value of the angle $\theta_{i,j}$ between individual i and its associated reference vector $\mathbf{w}_j$, and d_i is the Euclidean distance between individual i and the ideal point of the objective space. The dominance-based selection method is the same as the one presented in [12]. Nondominated sorting is performed on the combined population so that a front number is assigned to each solution. Then, a crowding distance is calculated for all solutions having the same front number. All solutions are ranked

---

Algorithm 2: Complementary Environmental Selection

Input:
  C'_t in the first selection and C_t in the second selection;
  W: a set of reference vectors;
  N: the size of initial population for LMOEA-DS;
  N_ϵ: the threshold to determine which method is used in the environmental selection;
Output:
  P_{t+1}: a parent population for the next generation
1 Normalization of the objective values of the combined population;
2 Assign each individual in the combined population to its closest reference vectors in **W**;
3 **if** *the number of reference vectors that have been assigned to at least one individual is not less than N_ϵ* **then**
4   **for** *each reference vector* $\mathbf{w}_i$ **do**
5   Calculate the performance Ψ using (5) for each individual assigned to $\mathbf{w}_i$;
6   Select the individual with the best performance among all individuals assigned to $\mathbf{w}_i$ as one individual in the next parent population P;
7   **end for**
8 **end if**
9 **else**
10   Nondominated sorting is performed on the combined population;
11   Select individuals layer by layer until reaching the Lth layer that the total number of selected individuals will exceed N if all individuals at this layer is added;
12   Calculate the crowding degree of each individual at Lth layer and select individuals in an descending order on the crowding degree until the total number of individuals reaches N;
13 **end if**

---

according to their front number in an ascending order and solutions having the same front number are ranked according to their crowding distance in descending order. A truncation selection is then performed to select N best solutions. The reader is referred to [12] for more details.

III. EXPERIMENTAL STUDIES

To investigate the performance of the proposed method, a set of empirical studies are conducted on nine large-scale multiobjective benchmark problems, namely, LSMOP1–9 [75], with 500, 1000, 2000, and 5000 decision variables, respectively. The characteristics of the nine test problems are given in Table I. A commonly used performance indicator, i.e., the inverted generational distance (IGD) [76], is adopted to assess the quality of the solution sets obtained by the algorithms, since IGD is able to account for both accuracy and diversity of a solution set in approximating the true Pareto front [77]. Given a set of solutions **P**, the IGD value is calculated as follows:

$$\text{IGD}(\mathbf{P}, \mathbf{P}^*) = \frac{\sum_{\mathbf{x}^* \in \mathbf{P}^*} d(\mathbf{x}^*, \mathbf{P})}{|\mathbf{P}^*|} \qquad (6)$$

where $\mathbf{P}^*$ is a set of uniformly distributed reference points and $d(\mathbf{x}^*, \mathbf{P})$ is the minimum Euclidean distance between a reference solution $\mathbf{x}^*$ in $\mathbf{P}^*$ and all solutions in **P**. $|\mathbf{P}^*|$ is the size of $\mathbf{P}^*$. Note that the lower the IGD value, the better the quality of the obtained optimal solutions.

In the remainder of this section, we will first describe the settings used in the experiments. Then, some empirical analysis on the contributions of different components in LMOEA-DS are given. Finally, the method is compared to
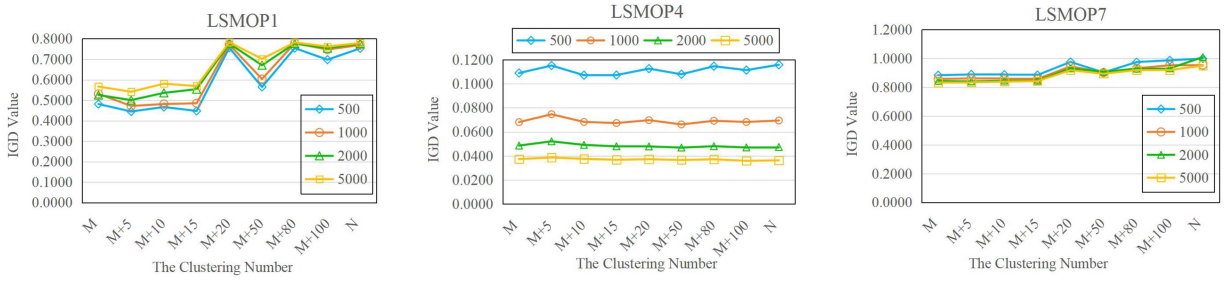
Fig. 4.   Parameter analysis of the clustering number $N_{w'}$.

five state-of-the-art algorithms on nine bi- and 3-D LSMOPs with up to 5000 decision variables.

A. Experimental Settings

Each algorithm under comparison is independently run 20 times on each test instance, and the terminal condition, i.e., the allowed maximum number of objective evaluations (denoted as *FEmax*), is set to 80 000. To calculate the IGD value of the solution set obtained by each algorithm, 10 000 uniformly distributed reference points are sampled on the Pareto front of each test problem [39], which is the same as given in [78]. The Wilcoxon rank-sum test [79] with Bonferroni correction is applied at a significance level of 0.05 to assess whether the performance of a solution set obtained by one of the two compared algorithms is statistically different to another [44], [80]. Symbols "+," "≈," and "−" indicate that the results obtained by the proposed LMOEA-DS is statistically better than, comparable to, or worse than the compared algorithm, respectively.

In our method, a real-coded genetic algorithm (GA) is adopted to optimize the large-scale multiobjective problem. The initial population size of the GA is set to $N = 153$. The simulated binary crossover (SBX) [1] and polynomial mutation (PM) [1], [81] are used as the reproduction operators. For the SBX, the distribution index η_c and the crossover probability p_c are set to 20 and 0.9, respectively. The distribution index η_m and mutation probability p_m for the PM are set to 20 and $1/D$, respectively, where D is the number of decision variables. The number of solutions to be randomly generated along each search direction is set to 30. The threshold N_ϵ is set to $2N/3$ to determine which selection strategy is used in the environmental selection.

Note that the number of clusters used to cluster the reference vectors **W** plays an important role in the proposed method. The larger the cluster number, the more search directions will be defined and, consequently, a larger number of objective evaluations will be consumed. Therefore, we conduct some pilot studies with different settings for the number of clusters in order to find out a proper setting of this parameter. Fig. 4 gives the mean IGD values of the solution sets obtained by LMOEA-DS on 20 independently runs using different cluster numbers on LSMOP1, LSMOP4, and LSMOP7 with up to 5000 decision variables. From Fig. 4, we can see that LMOEA-DS can obtain better results when the number of clusters is not more than $M + 15$, and there is no clear differences when the cluster number is between $M+5$ and $M+15$. Thus, $M+10$

is adopted in all experiments in the subsequent comparative studies.

The parameter settings used in all compared algorithms are the same to those given in their papers except that for WOF-NSGA-II [65], the number of evaluations are set to 800 and 400, respectively, for the optimization of the original problem and of each transformed problem to meet the maximum number of objective evaluations.

B. Analysis of the Effects of the Individual Components

In this section, we conduct a set of experiments on nine LSMOP test problems given in Table I with 500, 1000, 2000, and 5000 dimensions to show the efficiency and effectiveness of different components designed for LMOEA-DS.

1) Effects of the Directed Sampling Method: In our method, a number of solutions will be randomly generated along two defined search directions for each reference vector. The non-dominated solutions in the set of the randomly solutions will be used as the guiding solutions to accelerate the convergence to the Pareto front. In order to examine the effectiveness of the DS strategy, we compare the DS method with one of its variants, called DS-inbound, and the method proposed in LSMOF [66], called hypervolume optimization (HO), respectively, on three-objective LSMOP test problems with 500, 1000, 2000, and 5000 decision variables. The main difference between DS and DS-inbound lies in their repairing strategy for handling infeasible solutions generated by the DS. In DS-inbound, we first determine the bound along the given search direction and then uniformly sample solutions within the bound. Different to random sampling along the search direction, HO formulates a single-objective optimization problem to obtain a set of guiding solutions. The optimization problem aims to optimize the coefficients for all defined search directions to maximize the hypervolume (HV) of the solutions set generated by the optimized coefficients. In order to make a fair comparison, the population is set to 10 and the optimization is conducted for two generations so that the total number of fitness evaluations consumed by HO, DS, and DS-inbound is same at each generation.

Tables II and III present the statistical results on nine LSMOP test problems with up to 5000 dimensions obtained by the proposed LMOEA-DS, LMOEA-DS-inbound, and LMOEA-HO, respectively. From Table II, we can see that compared to LMOEA-DS-inbound, in which all solutions are limited within the bound along the search direction, the proposed method can obtain better or comparable results on all

TABLE II
STATISTICAL RESULTS (MEDIAN AND MEDIAN ABSOLUTE DEVIATION)
OBTAINED BY LMOEA-DS AND LMOEA-DS-INBOUND ON
THREE-OBJECTIVE LSMOP TEST PROBLEMS WITH 500, 1000, 2000,
AND 5000 DIMENSIONS. THE BEST MEDIAN RESULT
IN EACH ROW IS HIGHLIGHTED

Problem	D	LMOEA-DS-inbound	LMOEA-DS
LSMOP1	500	6.4115e-01(1.99e-02) [+]	4.5498e-01(2.41e-02)
	1000	6.8246e-01(4.13e-02) [+]	4.7643e-01(4.11e-02)
	2000	7.3085e-01(5.18e-02) [+]	4.7684e-01(8.15e-02)
	5000	7.4239e-01(2.10e-02) [+]	5.5356e-01(8.04e-02)
LSMOP2	500	6.8072e-02(8.28e-03) [+]	4.4929e-02(8.38e-04)
	1000	4.5114e-02(7.68e-03) [+]	3.8613e-02(7.89e-04)
	2000	3.6419e-02(1.58e-03) [+]	3.4450e-02(8.04e-04)
	5000	3.4356e-02(6.10e-04) [+]	3.3239e-02(9.80e-04)
LSMOP3	500	8.6056e-01(4.53e-05) [+]	8.6048e-01(3.62e-03)
	1000	8.6059e-01(8.34e-01) [+]	8.6049e-01(1.21e-03)
	2000	8.6062e-01(2.51e-03) [≈]	8.6057e-01(3.11e-04)
	5000	8.6088e-01(1.71e-03) [+]	8.6064e-01(1.45e-03)
LSMOP4	500	1.4157e-01(3.21e-03) [+]	1.0818e-01(1.78e-03)
	1000	9.3838e-02(3.94e-03) [+]	6.8001e-02(8.82e-04)
	2000	6.5975e-02(6.85e-03) [+]	4.8670e-02(9.49e-04)
	5000	3.8798e-02(2.61e-03) [+]	3.7368e-02(1.04e-03)
LSMOP5	500	9.7699e-01(1.80e+00) [+]	5.3450e-01(1.73e-02)
	1000	8.5876e-01(1.18e+00) [+]	5.3833e-01(1.24e-02)
	2000	9.0060e-01(1.20e+00) [+]	5.3963e-01(1.66e-03)
	5000	1.6300e+00(7.42e-01) [+]	5.3883e-01(1.61e-02)
LSMOP6	500	1.1803e+00(6.02e+03) [+]	7.6982e-01(2.93e-02)
	1000	1.5130e+00(4.13e+03) [+]	7.6694e-01(2.70e-02)
	2000	1.4335e+00(1.65e+03) [+]	7.6152e-01(2.72e-02)
	5000	1.6980e+00(7.28e+02) [+]	7.4373e-01(2.71e-02)
LSMOP7	500	1.0084e+00(1.07e+04) [+]	8.8589e-01(4.07e-03)
	1000	9.2927e-01(6.15e+03) [+]	8.5813e-01(1.55e-03)
	2000	9.3152e-01(1.05e+04) [+]	8.4561e-01(5.58e-04)
	5000	9.1965e-01(1.65e+03) [+]	8.3922e-01(1.76e-03)
LSMOP8	500	9.6354e-01(3.98e+02) [+]	2.9319e-01(4.09e-02)
	1000	1.4595e+00(2.28e+00) [+]	2.1473e-01(2.97e-02)
	2000	1.4905e+00(5.28e-01) [+]	2.2651e-01(1.29e-02)
	5000	1.7374e+00(4.62e-01) [+]	2.1994e-01(1.67e-02)
LSMOP9	500	5.8796e-01(8.04e-03) [≈]	5.8575e-01(1.18e-03)
	1000	5.8680e-01(6.59e-03) [+]	5.8175e-01(1.28e-03)
	2000	5.8292e-01(1.24e-02) [≈]	5.8036e-01(1.57e-03)
	5000	5.5922e-01(1.22e-02) [-]	5.7991e-01(1.84e-03)
+/≈/-		32/3/1	

TABLE III
STATISTICAL RESULTS (MEDIAN AND MEDIAN ABSOLUTE DEVIATION)
OBTAINED BY LMOEA-DS AND LMOEA-HO ON THREE-OBJECTIVE
LSMOP TEST PROBLEMS WITH 500, 1000, 2000, AND 5000 DIMENSIONS.
THE BEST MEDIAN RESULT IN EACH ROW IS HIGHLIGHTED

Problem	D	LMOEA-HO	LMOEA-DS
LSMOP1	500	4.5505e-01(4.45e-02) [≈]	4.5498e-01(2.41e-02)
	1000	4.8421e-01(4.95e-02) [≈]	4.7643e-01(4.11e-02)
	2000	5.2131e-01(7.53e-02) [≈]	4.7684e-01(8.15e-02)
	5000	5.7129e-01(7.95e-02) [≈]	5.5356e-01(8.04e-02)
LSMOP2	500	4.6390e-02(1.04e-03) [+]	4.4929e-02(8.38e-04)
	1000	3.8432e-02(7.05e-04) [≈]	3.8613e-02(7.89e-04)
	2000	3.4713e-02(4.43e-04) [≈]	3.4450e-02(8.04e-04)
	5000	3.3212e-02(7.88e-04) [≈]	3.3239e-02(9.80e-04)
LSMOP3	500	8.6061e-01(4.96e-04) [+]	8.6048e-01(3.62e-03)
	1000	8.6055e-01(3.57e-03) [≈]	8.6049e-01(1.21e-03)
	2000	8.6056e-01(1.20e-03) [≈]	8.6057e-01(3.11e-04)
	5000	8.6060e-01(1.04e-04) [≈]	8.6064e-01(1.45e-03)
LSMOP4	500	1.1235e-01(1.43e-03) [+]	1.0818e-01(1.78e-03)
	1000	6.8730e-02(1.12e-03) [+]	6.8001e-02(8.82e-04)
	2000	4.9041e-02(7.68e-04) [≈]	4.8670e-02(9.49e-04)
	5000	3.7602e-02(7.66e-04) [≈]	3.7368e-02(1.04e-03)
LSMOP5	500	5.3985e-01(1.19e-02) [≈]	5.3450e-01(1.73e-02)
	1000	5.4045e-01(1.03e-02) [+]	5.3833e-01(1.24e-02)
	2000	5.4059e-01(1.02e-02) [+]	5.3963e-01(1.66e-03)
	5000	5.4057e-01(9.99e-03) [+]	5.3883e-01(1.61e-02)
LSMOP6	500	7.7099e-01(2.07e-02) [≈]	7.6982e-01(2.93e-02)
	1000	7.8703e-01(2.74e-02) [≈]	7.6694e-01(2.70e-02)
	2000	7.8589e-01(2.52e-02) [≈]	7.6152e-01(2.72e-02)
	5000	7.9285e-01(2.47e-02) [+]	7.4373e-01(2.71e-02)
LSMOP7	500	8.8412e-01(3.12e-03) [≈]	8.8589e-01(4.07e-03)
	1000	8.5616e-01(2.52e-03) [-]	8.5813e-01(1.55e-03)
	2000	8.4546e-01(2.78e-03) [≈]	8.4561e-01(5.58e-04)
	5000	8.3901e-01(2.83e-04) [-]	8.3922e-01(1.76e-03)
LSMOP8	500	2.9837e-01(2.58e-02) [≈]	2.9319e-01(4.09e-02)
	1000	3.0210e-01(3.54e-02) [+]	2.1473e-01(2.97e-02)
	2000	2.5855e-01(3.55e-02) [+]	2.2651e-01(1.29e-02)
	5000	2.3600e-01(1.31e-02) [+]	2.1994e-01(1.67e-02)
LSMOP9	500	5.8698e-01(8.50e-04) [+]	5.8575e-01(1.18e-03)
	1000	5.8525e-01(1.31e-03) [+]	5.8175e-01(1.28e-03)
	2000	5.8256e-01(1.64e-03) [+]	5.8036e-01(1.57e-03)
	5000	5.8212e-01(1.67e-03) [+]	5.7991e-01(1.84e-03)
+/≈/-		15/19/2	

36 problems except LSMOP9 of dimension 5000. The results show that the method adopted in LMOEA-DS to repair infeasible solutions is more effective. Compared to LMOEA-HO, we can see that LMOEA-DS is worse only on LSMOP7 with 2000 and 5000 decision variables. Based on the results in both Tables II and III, we can conclude that the proposed DS is effective for finding guiding solutions.

2) Effectiveness of Double Regeneration: In order to evaluate the performance of the double reproduction strategy, we compare it with a variant with the first reproduction operation only, which is denoted as LMOEA-DS-one. Table IV lists the statistical results obtained by LMOEA-DS and LMOEA-DS-one on LSMOP1, LSMOP4, and LSMOP7 with up to 5000 dimensions, which are unimodal, mixed, and multimodel
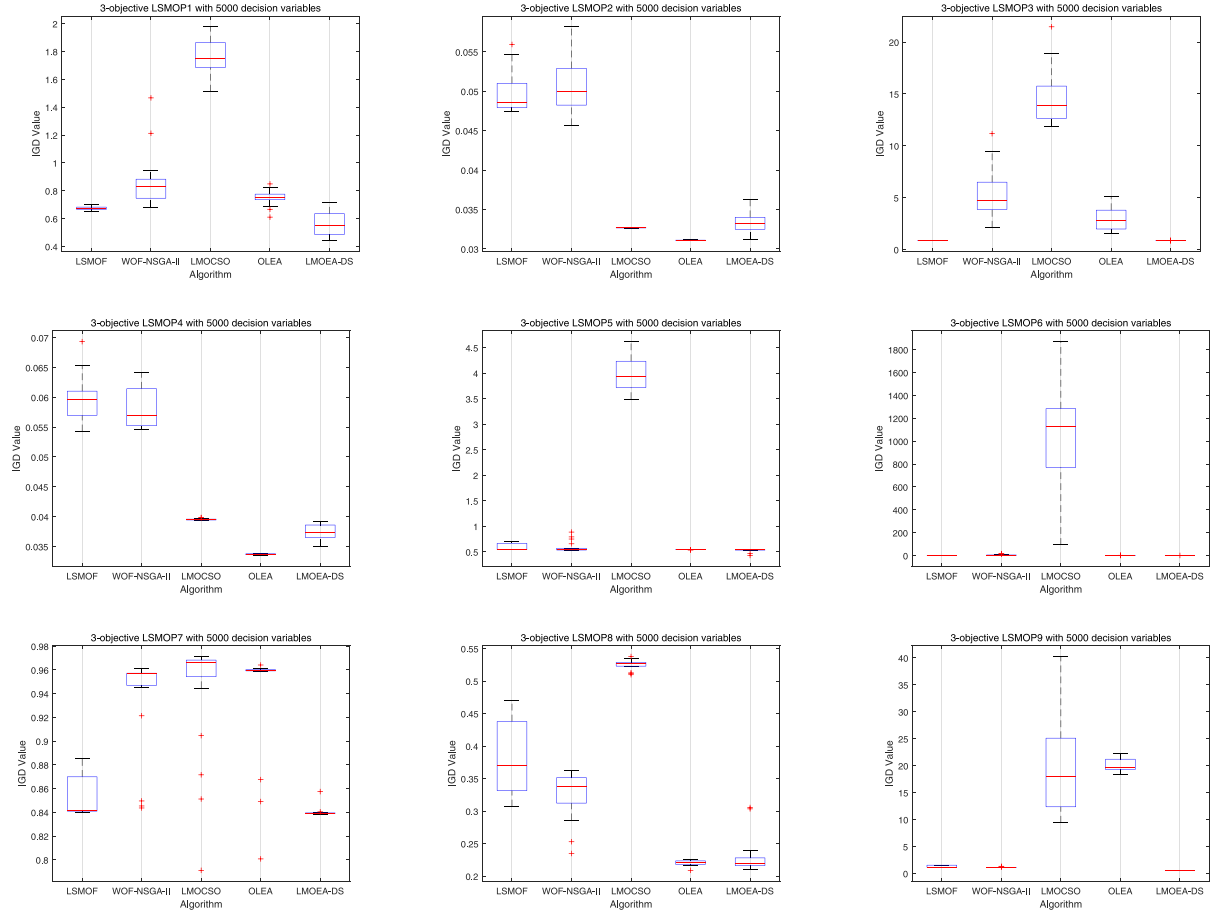
Fig. 5. Boxplots of LMOEA-DS and other four algorithms on three-objective LSMOP1-9 with 5000 decision variables.

TABLE IV
STATISTICAL RESULTS (MEDIAN AND MEDIAN ABSOLUTE DEVIATION) OBTAINED BY LMOEA-DS AND LMOEA-DS-ONE ON THREE-OBJECTIVE LSMOP1, LSMOP4, AND LSMOP7 WITH 500, 1000, 2000, AND 5000 DIMENSIONS. THE BEST MEDIAN RESULT IN EACH ROW IS HIGHLIGHTED

Problem	D	LMOEA-DS-one	LMOEA-DS
LSMOP1	500	6.0554e-01(6.26e-02) [+]	4.5498e-01(2.41e-02)
	1000	6.2847e-01(5.74e-02) [+]	4.7643e-01(4.11e-02)
	2000	6.6456e-01(6.13e-02) [+]	4.7684e-01(8.15e-02)
	5000	6.4850e-01(5.56e-02) [+]	5.5356e-01(8.04e-02)
LSMOP4	500	1.0783e-01(1.43e-03) [≈]	1.0818e-01(1.78e-03)
	1000	6.7159e-02(8.59e-04) [≈]	6.8001e-02(8.82e-04)
	2000	4.6646e-02(8.88e-04) [-]	4.8670e-02(9.49e-04)
	5000	3.5242e-02(4.97e-04) [-]	3.7368e-02(1.04e-03)
LSMOP7	500	8.8812e-01(4.87e-03) [≈]	8.8589e-01(4.07e-03)
	1000	8.5982e-01(2.42e-03) [+]	8.5813e-01(1.55e-03)
	2000	8.4934e-01(1.49e-03) [+]	8.4561e-01(5.58e-04)
	5000	8.4245e-01(1.29e-03) [+]	8.3922e-01(1.76e-03)
+/≈/-		7/3/2	-

problems, respectively. From Table IV, we can see that LMOEA-DS wins on 7/12 problems and loses only on 2/12 problems compared to the LMOEA-DS-one, indicating that

the second reproduction is helpful in improving the quality of the obtained solution set. Furthermore, the IGD results obtained by LMOEA-DS-one and LMOEA-DS on all three-objective LSMOP1–9 test problems are given in Table III of the supplementary material.

3) Effects of Complementary Environmental Selection: The complementary environmental selection strategy adopted in our work aims to ensure that a sufficient number of individuals will be kept in the parent population so that a high degree of diversity of the next parent population will be maintained. In order to verify the effectiveness of the complementary environmental selection, we compare it with two LMOEA-DS variants, one utilizing the decomposition-based selection only and the other using the nondominated sorting-based selection only, which are denoted as LMOEA-DS-decomposition and LMOEA-DS-domination, respectively. The statistical results of LMOEA-DS-decomposition, LMOEA-DS-domination and LMOEA-DS on LSMOP1, LSMOP4, and LSMOP7 with 500, 1000, 2000, and 5000 dimensions are given in Table V. The IGD results on the three-objective LSMOP1–9 problems with 500, 1000, 2000 and 5000 decision variables are also given in Table IV of the supplementary material. From Table V, we can see that LMOEA-DS obtains eight competitive and four better results compared to LMOEA-DS-decomposition on 12 problems with up to 5000 dimensions, indicating that LMOEA-DS with decomposition-based selection only cannot

TABLE V
STATISTICAL RESULTS (MEDIAN AND MEDIAN ABSOLUTE DEVIATION) OBTAINED BY LMOEA-DS-DECOMPOSITION, LMOEA-DS-DOMINATION AND LMOEA-DS ON THREE-OBJECTIVE LSMOP1, LSMOP4, AND LSMOP7 WITH 500, 1000, 2000, AND 5000 DECISION VARIABLES. THE BEST MEDIAN RESULT IN EACH ROW IS HIGHLIGHTED

Problem	D	LMOEA-DS-decomposition	LMOEA-DS-domination	LMOEA-DS
LSMOP1	500	4.4066e-01(2.59e-02) [≈]	5.5297e-01(2.29e-02) [+]	4.5498e-01(2.41e-02)
	1000	4.6985e-01(2.94e-02) [≈]	6.0421e-01(2.30e-02) [+]	4.7643e-01(4.11e-02)
	2000	4.8324e-01(3.62e-02) [≈]	6.5004e-01(2.28e-02) [+]	4.7684e-01(8.15e-02)
	5000	4.9964e-01(4.11e-02) [≈]	6.6624e-01(2.45e-02) [+]	5.5356e-01(8.04e-02)
LSMOP4	500	1.0766e-01(1.37e-03) [≈]	1.3539e-01(2.28e-03) [+]	1.0818e-01(1.78e-03)
	1000	6.8479e-02(8.16e-04) [≈]	9.2675e-02(2.15e-03) [+]	6.8001e-02(8.82e-04)
	2000	4.8911e-02(8.58e-04) [≈]	6.7097e-02(1.45e-03) [+]	4.8670e-02(9.49e-04)
	5000	3.7381e-02(6.88e-04) [≈]	5.2832e-02(1.85e-03) [+]	3.7368e-02(1.04e-03)
LSMOP7	500	1.3250e+00(1.81e-02) [+]	8.8454e-01(4.68e-03) [≈]	8.8589e-01(4.07e-03)
	1000	1.3366e+00(2.68e-02) [+]	8.5835e-01(1.58e-03) [≈]	8.5813e-01(1.55e-03)
	2000	1.3406e+00(3.01e-02) [+]	8.4522e-01(7.08e-04) [≈]	8.4561e-01(5.58e-04)
	5000	1.3441e+00(4.02e-02) [+]	8.3902e-01(4.72e-04) [≈]	8.3922e-01(1.76e-03)
+/≈/-		4/8/0	8/4/0	-

ensure a sufficient degree of population diversity for many problems. Furthermore, we can see from Table V that the complementary selection strategy in LMOEA-DS can make good use of the decomposition-based selection and nondominated sorting-based selection and, therefore, achieve a good balance between maintaining population diversity and accelerating convergence.

C. Comparisons With State-of-the-Art Large-Scale Multiobjective Methods

In order to rigorously evaluate the performance of LMOEA-DS, we compare LMOEA-DS with several recently proposed large-scale optimization algorithms, namely, LSMOF [66], WOF-NSGA-II [65], LMOCSO [67], and MOEA/DVA [62], on LSMOP test problems. Furthermore, an orthogonal crossover algorithm with learning automata (OLEA), proposed in [82], is also used for comparison. In OLEA, the decision variables are divided into groups, aiming to generate fewer promising offspring solutions to reduce the number of objective evaluations. The method was proposed for lower-dimensional MOPs. However, it is also efficient for LSMOPs because in OLEA, a set of solutions, which are located close to the Pareto optimal solutions, are chosen for the quantization orthogonal crossover with learning automata to guide the search for large-scale optimization.

Tables VI and VII give the statistical results obtained by the compared algorithms on biobjective and three-objective LSMOP1–9 problems with up to 5000 decision variables, respectively. From Tables VI and VII, we can see that compared to MOEA/DVA using the decision variable clustering technique and LMOCSO that uses the new learning strategy, our proposed LMOEA-DS has achieved better results on 36 and 34, respectively, out of 36 biobjective problems. Both MOEA/DVA and LMOCSO have obtained better results on one problem only out of all three-objective LSMOP test problems. LMOEA-DS has outperformed WOF-NSGA-II on 25

out of 36 biobjective LSMOP problems. Furthermore, WOF-NSGA-II was not able to obtain a better result on any of the three-objective LSMOP problems. The LSMOF method utilizes a problem transformation technique to guide the search for large-scale multiobjective optimization. From Tables VI and VII, we can see that our LMOEA-DS has obtained better results on 27 and 29 out of 36 biobjective and 36 three-objective problems, respectively. Especially, as can be observed from Table VII, LMOEA-DS can achieve better or competitive performance on all three-objective LSMOP benchmark problems when the number of decision variables is larger than 500. Compared to OLEA, which does not fall in any of the four categories of large-scale multiobjective optimization, we can see that LMOEA-DS is underperformed on five biobjective LSMOP test problems and 11 three-objective LSMOP test problems, respectively, with up to 5000 dimensions. However, LMOEA-DS has outperfomed OLEA on 26 and 22 out of 36 biobjective and three-objective problems. The statistical results of the HV values obtained by the compared algorithms on biobjective and three-objective LSMOP1–9 problems with up to 5000 decision variables are given in Tables VII and VIII, respectively, of the supplementary material. From these results, we can conclude that the proposed LMOEA-DS can obtain better results than MOEA/DVA, LSMOF, WOF-NSGA-II, LMOCSO, and OLEA.

The boxplots of the IGD values obtained by LSMOF, LMOCSO, MOEA/DVA, OLEA, and the proposed LMOEA-DS on three-objective LSMOP1–9 test problems with 5000 dimensions are presented in Fig. 5. Note that the boxplot of MOEA/DVA is not included in order to display the results of the other algorithms more clearly. The median values of each algorithm have been summarized in Table VII, from which, and from Fig. 5, we can observe that LMOEA-DS can obtain the best median results among the five algorithms under comparison on LSMOP1, LSMOP3, LSMOP5, LSMOP6, LSMOP7, LSMOP8, and LSMOP9 with

TABLE VI
Statistical Results (Median and Median Absolute Deviation) Obtained by LMOEA-DS and Five Compared Algorithms on 500-D, 1000-D, 2000-D, and 5000-D 2-Objective LSMOP1–9 Problems. The Best Median Result in Each Row Is Highlighted

Problem	M	D	MOEA/DVA	LSMOF	WOF-NSGA-II	LMOCSO	OLEA	LMOEA-DS
LSMOP1	2	500	8.7860e+00(1.78e-01) [+]	6.0526e-01(2.81e-02) [+]	4.7532e-01(7.91e-02) [+]	7.8942e-01(1.09e-01) [+]	2.5706e-01(2.16e-02) [-]	3.1311e-01(1.03e-02)
		1000	1.0281e+01(2.19e-01) [+]	6.4223e-01(2.07e-02) [+]	4.4680e-01(8.46e-02) [+]	1.5818e+00(1.30e-01) [+]	3.3678e-01(2.46e-02) [≈]	3.2838e-01(7.99e-03)
		2000	1.1047e+01(1.23e-01) [+]	6.4302e-01(9.72e-03) [+]	6.1641e-01(6.56e-02) [+]	2.0286e+00(1.35e-01) [+]	4.0885e-01(3.89e-02) [+]	3.3388e-01(6.14e-03)
		5000	1.1458e+01(9.82e-02) [+]	6.5744e-01(7.74e-03) [+]	5.2010e-01(9.69e-02) [+]	2.3083e+00(1.36e-01) [+]	4.8361e-01(5.11e-02) [+]	3.3982e-01(5.00e-03)
LSMOP2	2	500	7.2875e-02(1.22e-04) [+]	1.9443e-02(3.83e-04) [+]	1.5334e-02(5.18e-04) [+]	4.3181e-02(6.69e-04) [+]	1.6252e-02(1.11e-03) [+]	1.3724e-02(4.01e-04)
		1000	3.9697e-02(1.92e-04) [+]	1.3987e-02(2.71e-04) [+]	1.2195e-02(4.22e-04) [+]	2.3962e-02(3.78e-04) [+]	8.2990e-03(3.74e-04) [≈]	8.0297e-03(2.13e-04)
		2000	2.1937e-02(1.51e-04) [+]	1.0537e-02(2.57e-04) [+]	1.1354e-02(2.53e-04) [+]	1.3406e-02(1.66e-04) [+]	4.7441e-03(1.52e-04) [-]	5.0935e-03(1.23e-04)
		5000	1.1257e-02(2.64e-04) [+]	7.6264e-03(2.04e-04) [+]	7.3807e-03(3.30e-04) [+]	6.5956e-03(5.19e-05) [+]	3.0850e-03(6.39e-05) [-]	3.5386e-03(8.05e-05)
LSMOP3	2	500	1.0666e+03(6.91e+02) [+]	1.5649e+00(7.79e-04) [-]	1.4971e+00(1.23e-01) [≈]	1.2704e+00(5.45e-01) [-]	1.1303e+01(1.81e+00) [+]	1.5669e+00(3.60e-04)
		1000	8.2434e+02(5.53e+02) [+]	1.5737e+00(3.63e-04) [≈]	1.5725e+00(8.52e-02) [≈]	3.8277e+00(1.64e+00) [+]	1.5351e+01(1.64e+00) [+]	1.5736e+00(7.99e-04)
		2000	3.3261e+02(4.69e+02) [+]	1.5772e+00(2.31e-04) [+]	1.5861e+00(1.76e+00) [+]	1.0422e+01(2.41e+00) [+]	1.6861e+01(1.03e+00) [+]	1.5767e+00(8.75e-04)
		5000	7.2178e+02(6.44e+02) [+]	1.5792e+00(9.04e-05) [≈]	1.3898e+00(8.76e-02) [-]	1.9829e+01(2.68e+00) [+]	1.8158e+01(1.45e+00) [+]	1.5789e+00(1.62e-03)
LSMOP4	2	500	1.2705e-01(3.27e-04) [+]	4.8133e-02(6.47e-04) [+]	4.4211e-02(5.39e-04) [+]	8.7136e-02(8.00e-04) [+]	4.1694e-02(1.41e-03) [≈]	4.0862e-02(5.85e-04)
		1000	7.5277e-02(1.88e-04) [+]	2.8534e-02(3.87e-04) [+]	2.6086e-02(6.01e-04) [+]	5.1722e-02(3.28e-04) [+]	2.4541e-02(9.09e-04) [+]	2.2166e-02(3.45e-04)
		2000	4.3467e-02(1.24e-04) [+]	1.8825e-02(5.24e-04) [+]	2.4730e-02(2.01e-03) [+]	2.9576e-02(1.00e-04) [+]	1.3421e-02(4.43e-04) [+]	1.2274e-02(1.60e-04)
		5000	2.1755e-02(1.20e-04) [+]	1.2809e-02(4.27e-04) [+]	1.1724e-02(3.33e-04) [+]	1.4318e-02(4.87e-05) [+]	6.1598e-03(1.68e-04) [≈]	6.1763e-03(1.07e-04)
LSMOP5	2	500	1.9324e+01(3.46e-01) [+]	7.4209e-01(1.11e-16) [+]	4.5847e-01(8.91e-02) [-]	2.0255e+00(2.59e-01) [+]	5.0112e-01(3.15e-02) [-]	7.1467e-01(5.26e-03)
		1000	2.2142e+01(2.96e-01) [+]	7.4209e-01(1.11e-16) [+]	6.0693e-01(7.53e-02) [-]	3.7939e+00(3.55e-01) [+]	7.4259e-01(5.79e-02) [+]	7.1436e-01(5.06e-03)
		2000	2.3666e+01(3.37e-01) [+]	7.4209e-01(1.11e-16) [+]	7.4209e-01(3.02e-03) [+]	5.0262e+00(4.15e-01) [+]	8.0350e-01(3.73e-02) [+]	7.1493e-01(7.79e-03)
		5000	2.4395e+01(2.06e-01) [+]	7.4209e-01(1.11e-16) [+]	7.3429e-01(3.19e-02) [+]	5.6360e+00(4.24e-01) [+]	7.7274e-01(2.32e-02) [+]	7.1410e-01(7.19e-03)
LSMOP6	2	500	1.4834e+03(1.37e+03) [+]	3.2011e-01(4.40e-04) [+]	6.6740e-01(3.52e-02) [+]	7.8431e-01(6.90e-03) [+]	7.7864e-01(1.48e-03) [+]	3.1896e-01(6.67e-04)
		1000	1.1952e+03(1.00e+03) [+]	3.1246e-01(2.11e-04) [≈]	6.6937e-01(8.15e-02) [+]	7.6739e-01(1.44e-03) [+]	7.6141e-01(5.25e-03) [+]	3.1232e-01(2.25e-04)
		2000	1.1512e+03(1.38e+03) [+]	3.0879e-01(1.21e-04) [≈]	6.7175e-01(3.31e-02) [+]	7.5558e-01(1.13e-03) [+]	7.5176e-01(5.15e-02) [+]	3.0878e-01(7.37e-05)
		5000	1.1498e+03(1.88e+03) [+]	3.0701e-01(2.13e-03) [+]	6.7004e-01(6.38e-02) [+]	7.4745e-01(1.22e-04) [+]	7.4619e-01(8.97e-02) [+]	3.0699e-01(2.89e-05)
LSMOP7	2	500	6.3996e+04(3.13e+03) [+]	1.5019e+00(1.17e-03) [-]	1.5039e+00(5.62e-02) [-]	1.6290e+02(7.84e+01) [+]	1.5601e+00(6.26e-02) [+]	1.5067e+00(1.19e-03)
		1000	7.8286e+04(2.80e+03) [+]	1.5101e+00(3.23e-04) [-]	1.5128e+00(9.56e-02) [≈]	2.3653e+03(7.26e+02) [+]	1.5468e+00(3.88e-02) [+]	1.5119e+00(1.21e-03)
		2000	8.7518e+04(1.39e+03) [+]	1.5138e+00(3.16e-04) [-]	1.5192e+00(1.37e+00) [+]	4.9912e+03(1.11e+03) [+]	1.5450e+00(1.45e-02) [+]	1.5147e+00(6.13e-04)
		5000	9.1438e+04(1.35e+03) [+]	1.5157e+00(3.57e-04) [-]	1.5191e+00(7.88e-02) [≈]	6.7201e+03(1.62e+03) [+]	1.5419e+00(8.13e-01) [+]	1.5164e+00(4.99e-04)
LSMOP8	2	500	1.6204e+01(4.11e-01) [+]	7.4209e-01(1.11e-16) [+]	4.0305e-01(1.38e-01) [-]	1.5896e+00(2.46e-01) [+]	4.2698e-01(8.40e-02) [-]	7.0130e-01(1.67e-02)
		1000	1.8883e+01(3.16e-01) [+]	7.4209e-01(1.11e-16) [+]	6.1110e-01(1.56e-01) [≈]	3.2331e+00(3.25e-01) [+]	7.4614e-01(1.57e-02) [+]	7.0728e-01(1.00e-02)
		2000	2.0118e+01(2.11e-01) [+]	7.4209e-01(1.11e-16) [+]	7.4209e-01(2.08e-02) [+]	4.0539e+00(4.17e-01) [+]	7.5238e-01(1.74e-02) [+]	7.0942e-01(1.65e-02)
		5000	2.0839e+01(8.24e-02) [+]	7.4209e-01(1.11e-16) [+]	6.7999e-01(8.86e-02) [≈]	4.2317e+00(2.69e-01) [+]	7.4544e-01(6.95e-03) [+]	7.0954e-01(1.02e-02)
LSMOP9	2	500	4.3027e+01(1.55e+00) [+]	8.0878e-01(6.89e-04) [+]	8.0997e-01(1.76e-04) [+]	7.7112e-01(8.66e-02) [≈]	7.4164e-01(1.72e-01) [≈]	7.9250e-01(4.19e-02)
		1000	5.2526e+01(1.27e+00) [+]	8.0752e-01(1.36e-03) [+]	8.0913e-01(6.82e-04) [+]	2.9507e+00(7.39e-01) [+]	8.4399e-01(3.84e-01) [+]	6.4566e-01(3.11e-02)
		2000	5.8631e+01(9.35e-01) [+]	8.0538e-01(1.48e-03) [+]	8.0694e-01(4.18e-03) [+]	7.0281e+00(1.94e+00) [+]	1.2944e+00(7.48e-01) [+]	6.1974e-01(1.99e-02)
		5000	6.1803e+01(5.47e-01) [+]	8.0324e-01(1.95e-03) [+]	8.0512e-01(2.25e-03) [+]	1.0530e+01(1.98e+00) [+]	2.6056e+00(2.61e+00) [+]	6.0717e-01(1.93e-02)
+/≈/-			36/0/0	27/4/5	25/6/5	34/1/1	26/5/5	-

5000 decision variables. Furthermore, we can see from Fig. 5 that the interquartile ranges of the IGD values obtained by LMOEA-DS are relatively small, meaning that it performs robustly on all the nine 5000-D LSMOPs.

Table VIII summarizes the comparative results on all 72 benchmark problems. From Table VIII, we can see that the proposed method has obtained better results on 72, 56, 57, 69, and 48 out of 72 LSMOP test problems than MOEA/DVA, LSMOF, WOF-NSGA-II, LMOCSO, and OLEA, respectively, which shows the promising performance of our proposed method.

IV. CONCLUSION

To enhance the convergence performance of evolutionary algorithm in solving LSMOPs, this article proposed a method for generating guiding solutions along search directions defined by a few chosen individuals in the current parent population. The guiding solutions are then utilized in a double reproduction process in combination with a complementary selection strategy in an effort to generate well converged and diverse solutions. The experimental results confirmed that our proposed method is highly competitive in comparison with five state-of-the-art algorithms on biobjective and three-objective LSMOPs with up to 5000 dimensions. Notably, the proposed algorithm scales particularly well to the increase in the number of decision variables.

The proposed algorithm also suffers from one common weakness of decomposition-based algorithms is that their performance heavily depends on the degree of match between the distribution of the reference vectors and the shape of the Pareto fronts of the problem to be solved. In the future,

TABLE VII
STATISTICAL RESULTS (MEDIAN AND MEDIAN ABSOLUTE DEVIATION) OBTAINED BY LMOEA-DS AND FIVE COMPARED ALGORITHMS ON 500-D, 1000-D, 2000-D, AND 5000-D 3-OBJECTIVE LSMOP1–9 PROBLEMS. THE BEST MEDIAN RESULT IN EACH ROW IS HIGHLIGHTED

Problem	M	D	MOEA/DVA	LSMOF	WOF-NSGA-II	LMOCSO	OLEA	LMOEA-DS
LSMOP1	3	500	9.0575e+00(2.86e-01) [+]	5.7949e-01(6.54e-03) [+]	5.5391e-01(4.18e-02) [+]	1.4212e+00(1.02e-01) [+]	5.3261e-01(4.28e-02) [+]	4.5498e-01(2.41e-02)
		1000	1.0452e+01(2.39e-01) [+]	6.0963e-01(6.94e-03) [+]	5.8839e-01(2.45e-02) [+]	1.6175e+00(6.74e-02) [+]	6.5391e-01(3.40e-02) [+]	4.7643e-01(4.11e-02)
		2000	1.1130e+01(1.87e-01) [+]	6.3059e-01(7.28e-03) [+]	6.4489e-01(2.18e-02) [+]	1.6884e+00(8.12e-02) [+]	7.2100e-01(4.34e-02) [+]	4.7684e-01(8.15e-02)
		5000	1.1570e+01(1.58e-01) [+]	6.7556e-01(9.43e-03) [+]	8.2925e-01(1.17e-01) [+]	1.7494e+00(8.86e-02) [+]	7.5408e-01(3.65e-02) [+]	5.5356e-01(8.04e-02)
LSMOP2	3	500	7.1208e-02(1.74e-03) [+]	7.9365e-02(1.85e-03) [+]	8.0370e-02(2.11e-03) [+]	5.0817e-02(4.32e-04) [+]	4.3960e-02(4.21e-04) [-]	4.4929e-02(8.38e-04)
		1000	5.5339e-02(1.90e-03) [+]	6.1492e-02(1.92e-03) [+]	6.2031e-02(2.06e-03) [+]	4.0061e-02(2.26e-04) [+]	3.5460e-02(1.60e-04) [-]	3.8613e-02(7.89e-04)
		2000	4.8338e-02(1.67e-03) [+]	5.3557e-02(1.74e-03) [+]	5.2751e-02(1.71e-03) [+]	3.5307e-02(1.06e-04) [+]	3.2396e-02(5.95e-05) [-]	3.4450e-02(8.04e-04)
		5000	4.6236e-02(1.64e-03) [+]	4.8623e-02(2.07e-03) [+]	4.9913e-02(2.69e-03) [+]	3.2692e-02(2.85e-05) [≈]	3.1101e-02(1.93e-05) [-]	3.3239e-02(9.80e-04)
LSMOP3	3	500	1.8441e+02(1.52e+02) [+]	8.6049e-01(5.89e-03) [≈]	8.6072e-01(2.74e-04) [+]	1.0544e+01(1.87e+00) [+]	2.3092e+00(7.78e-01) [+]	8.6048e-01(3.62e-03)
		1000	1.2616e+02(2.86e+02) [+]	8.6064e-01(6.54e-05) [+]	8.6072e-01(7.26e-02) [+]	1.3054e+00(1.29e+00) [+]	3.4660e+00(1.02e+00) [+]	8.6049e-01(1.21e-03)
		2000	2.3955e+02(1.49e+02) [+]	8.6069e-01(6.31e-05) [+]	8.6072e-01(1.35e-01) [+]	1.3251e+00(1.39e+00) [+]	3.6046e+00(8.66e-01) [+]	8.6057e-01(3.11e-04)
		5000	4.4864e+02(2.65e+02) [+]	8.6072e-01(6.30e-05) [≈]	4.7664e+00(1.94e+00) [+]	1.3893e+01(1.93e+00) [+]	2.7917e+00(8.99e-01) [+]	8.6064e-01(1.45e-03)
LSMOP4	3	500	1.9972e-01(2.38e-03) [+]	2.0397e-01(3.32e-03) [+]	2.0556e-01(3.67e-03) [+]	1.4828e-01(8.24e-04) [+]	1.0086e-01(2.00e-03) [-]	1.0818e-01(1.78e-03)
		1000	1.2439e-01(1.36e-03) [+]	1.3318e-01(3.08e-03) [+]	1.3437e-01(2.62e-03) [+]	9.1873e-02(3.97e-04) [+]	6.1469e-02(1.00e-03) [-]	6.8001e-02(8.82e-04)
		2000	7.9424e-02(1.97e-03) [+]	8.6416e-02(2.65e-03) [+]	8.5148e-02(1.78e-03) [+]	5.8623e-02(1.64e-04) [+]	4.2840e-02(3.09e-04) [-]	4.8670e-02(9.49e-04)
		5000	5.3296e-02(1.74e-03) [+]	5.9626e-02(2.59e-03) [+]	5.6995e-02(3.03e-03) [+]	3.9503e-02(1.10e-04) [+]	3.3681e-02(7.83e-05) [-]	3.7368e-02(1.04e-03)
LSMOP5	3	500	1.5973e+01(6.48e-01) [+]	5.4027e-01(4.23e-02) [≈]	5.3951e-01(1.74e-02) [≈]	3.0430e+00(1.54e-01) [+]	3.7158e-01(1.87e-02) [-]	5.3450e-01(1.73e-02)
		1000	1.8618e+01(4.23e-01) [+]	6.5452e-01(7.02e-02) [+]	5.3927e-01(2.98e-03) [+]	3.5928e+00(1.99e-01) [+]	4.5737e-01(3.27e-02) [-]	5.3833e-01(1.24e-02)
		2000	1.9689e+01(3.00e-01) [+]	5.8513e-01(6.93e-02) [+]	5.4088e-01(2.87e-03) [+]	3.9550e+00(2.62e-01) [+]	5.4197e-01(2.17e-03) [+]	5.3963e-01(1.66e-03)
		5000	2.0607e+01(2.35e-01) [+]	5.4498e-01(5.76e-02) [+]	5.4039e-01(7.35e-02) [≈]	3.9418e+00(2.41e-01) [+]	5.4295e-01(1.21e-03) [+]	5.3883e-01(1.61e-02)
LSMOP6	3	500	2.6519e+04(3.33e+03) [+]	7.3246e-01(1.02e-02) [-]	1.2904e+00(3.94e-02) [+]	1.4435e+02(1.06e+02) [+]	1.3383e+00(8.44e-02) [+]	7.6982e-01(2.93e-02)
		1000	3.2665e+04(2.78e+03) [+]	7.5434e-01(1.33e-02) [≈]	1.3898e+00(8.37e-02) [+]	4.9111e+02(1.63e+02) [+]	1.4357e+00(6.18e-02) [+]	7.6694e-01(2.70e-02)
		2000	3.5515e+04(4.29e+03) [+]	7.5886e-01(1.39e-02) [≈]	1.4068e+00(1.22e-01) [+]	8.1687e+02(4.26e+02) [+]	1.4381e+00(2.36e-02) [+]	7.6152e-01(2.72e-02)
		5000	3.7730e+04(4.50e+03) [+]	7.5805e-01(1.49e-02) [≈]	2.5774e+00(3.51e+00) [+]	1.1276e+03(3.27e+02) [+]	1.4375e+00(8.38e-02) [+]	7.4373e-01(2.71e-02)
LSMOP7	3	500	3.4148e+02(8.14e+02) [+]	8.9275e-01(1.09e-02) [+]	8.7809e-01(2.83e-02) [≈]	9.4821e-01(8.75e-02) [+]	1.0456e+00(4.17e-02) [+]	8.8589e-01(4.07e-03)
		1000	7.1323e+02(5.59e+02) [+]	8.6558e-01(2.91e-02) [+]	9.0019e-01(3.00e-02) [≈]	1.0254e+00(4.99e-02) [+]	1.0074e+00(1.13e-02) [+]	8.5813e-01(1.55e-03)
		2000	1.3412e+03(9.33e+02) [+]	8.4927e-01(1.81e-02) [+]	9.2012e-01(3.27e-02) [+]	9.8468e-01(1.86e-02) [+]	9.7898e-01(1.63e-02) [+]	8.4561e-01(5.58e-04)
		5000	1.4717e+03(8.82e+02) [+]	8.4160e-01(1.55e-02) [+]	9.5660e-01(2.90e-02) [+]	9.6594e-01(3.55e-02) [+]	9.5983e-01(3.09e-02) [+]	8.3922e-01(1.76e-03)
LSMOP8	3	500	6.7181e-01(2.89e-02) [+]	3.6024e-01(6.10e-02) [+]	3.2020e-01(2.74e-02) [+]	5.4967e-01(5.46e-03) [+]	2.3505e-01(2.63e-03) [-]	2.9319e-01(4.09e-02)
		1000	6.4024e-01(4.86e-02) [+]	4.0204e-01(5.59e-02) [+]	2.9967e-01(3.91e-02) [+]	5.3382e-01(3.85e-03) [+]	2.2887e-01(3.76e-03) [≈]	2.1473e-01(2.97e-02)
		2000	6.5943e-01(3.93e-02) [+]	3.4037e-01(4.47e-02) [+]	2.9675e-01(3.86e-02) [+]	5.3069e-01(6.01e-03) [+]	2.2557e-01(2.19e-03) [≈]	2.2651e-01(1.29e-02)
		5000	6.3529e-01(4.36e-02) [+]	3.6954e-01(4.90e-02) [+]	3.3795e-01(2.72e-02) [+]	5.2626e-01(5.09e-03) [+]	2.2167e-01(2.93e-03) [≈]	2.1994e-01(1.67e-02)
LSMOP9	3	500	1.1061e+02(4.37e+00) [+]	1.5379e+00(7.08e-02) [+]	1.1450e+00(2.07e-04) [+]	9.3374e-01(1.04e-01) [+]	2.6282e+00(5.61e-01) [+]	5.8575e-01(1.18e-03)
		1000	1.3103e+02(3.29e+00) [+]	1.5379e+00(1.83e-01) [+]	1.1447e+00(2.37e-04) [+]	1.0165e+00(1.80e-01) [+]	5.4831e+00(3.80e-01) [+]	5.8175e-01(1.28e-03)
		2000	1.4292e+02(2.31e+00) [+]	1.1446e+00(1.65e-01) [+]	1.1446e+00(7.16e-02) [+]	2.1782e+00(1.44e+00) [+]	9.5920e+00(3.81e-01) [+]	5.8036e-01(1.57e-03)
		5000	1.5146e+02(1.08e+00) [+]	1.1499e+00(1.67e-01) [+]	1.1459e+00(2.53e-02) [+]	1.7971e+01(7.17e+00) [+]	1.9637e+01(1.12e+00) [+]	5.7991e-01(1.84e-03)
+/≈/-			36/0/0	29/6/1	32/4/0	35/1/0	22/3/11	-

TABLE VIII
SUMMARY OF THE RESULTS ON LSMOP1–LSMOP9 TEST PROBLEMS, WHERE LMOEA-DS IS BETTER THAN (+), WORST THAN (-), AND COMPARABLE TO (≈) EACH OF FIVE COMPARED ALGORITHMS ACCORDING TO THE WILCOXON RANK SUM TEST WITH BONFERRONI CORRECTION

LMOEA-DS $v.s.$	MOEA/DVA	LSMOF	WOF-NSGA-II	LMOCSO	OLEA
+	72/72	56/72	57/72	69/72	48/72
≈	0/72	10/72	10/72	2/72	8/72
-	0/72	6/72	5/72	1/72	16/72

we will make efforts to develop strategies for adapting the reference vectors to achieve more powerful environmental selection during the optimization. In addition, since maintaining a sufficient degree of population diversity is extremely important for large-scale MOPs, we will design new regeneration strategies to enhance the population diversity to further improve the search performance of the proposed algorithm on LSMOPs.

REFERENCES

[1] K. Deb, *Multi-Objective Optimization Using Evolutionary Algorithms*. New York, NY, USA: Wiley, 2001.
[2] R. C. Purshouse and P. J. Fleming, "On the evolutionary optimization of many conflicting objectives," *IEEE Trans. Evol. Comput.*, vol. 11, no. 6, pp. 770–784, Dec. 2007.
[3] D. Brockhoff, T. Friedrich, N. Hebbinghaus, C. Klein, F. Neumann, and E. Zitzler, "On the effects of adding objectives to plateau functions," *IEEE Trans. Evol. Comput.*, vol. 13, no. 3, pp. 591–603, Jun. 2009.
[4] A. Zhou, B.-Y. Qu, H. Li, S.-Z. Zhao, P. N. Suganthan, and Q. Zhang, "Multiobjective evolutionary algorithms: A survey of the state of the art," *Swarm Evol. Comput.*, vol. 1, no. 1, pp. 32–49, 2011.
[5] Y. Zhou, Z. Chen, and J. Zhang, "Ranking vectors by means of the dominance degree matrix," *IEEE Trans. Evol. Comput.*, vol. 21, no. 1, pp. 34–51, Feb. 2017.

[6] Y. Zhang, D.-W. Gong, J.-Y. Sun, and B.-Y. Qu, "A decomposition-based archiving approach for multi-objective evolutionary optimization," *Inf. Sci.*, vols. 430–431, pp. 397–413, Mar. 2018.

[7] J. Maltese, B. M. Ombuki-Berman, and A. P. Engelbrecht, "A scalability study of many-objective optimization algorithms," *IEEE Trans. Evol. Comput.*, vol. 22, no. 1, pp. 79–96, Feb. 2018.

[8] D. Gong, Y. Han, and J. Sun, "A novel hybrid multi-objective artificial bee colony algorithm for blocking lot-streaming flow shop scheduling problems," *Knowl. Based Syst.*, vol. 148, pp. 115–130, May 2018.

[9] H. Ishibuchi, N. Tsukamoto, and Y. Nojima, "Behavior of evolutionary many-objective optimization," in *Proc. 10th Int. Conf. Comput. Model. Simulat.*, Cambridge, U.K., 2008, pp. 266–271.

[10] Z. He and G. G. Yen, "Ranking many-objective evolutionary algorithms using performance metrics ensemble," in *Proc. IEEE Congr. Evol. Comput.*, Cancun, Mexico, 2013, pp. 2480–2487.

[11] D. W. Corne, J. D. Knowles, and M. J. Oates, "The Pareto envelope-based selection algorithm for multiobjective optimization," in *Parallel Problem Solving From Nature PPSN VI*, vol. 1917. Berlin, Germany: Springer, 2000, pp. 839–884.

[12] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: NSGA-II," *IEEE Trans. Evol. Comput.*, vol. 6, no. 2, pp. 182–197, Apr. 2002.

[13] M. Laumanns, L. Thiele, K. Deb, and E. Zitzler, "Combining convergence and diversity in evolutionary multiobjective optimization," *Evol. Comput.*, vol. 10, no. 3, pp. 263–282, 2002.

[14] L. Raisanen and R. M. Whitaker, "Comparison and evaluation of multiple objective genetic algorithms for the antenna placement problem," *Mobile Netw. Appl.*, vol. 10, nos. 1–2, pp. 79–88, 2005.

[15] D. W. Corne and J. D. Knowles, "Techniques for highly multiobjective optimisation: Some nondominated points are better than others," in *Proc. Conf. Genet. Evol. Comput.*, 2007, pp. 773–780.

[16] L. S. Batista, F. Campelo, F. G. Guimaraes, and J. A. Ramirez, "A comparison of dominance criteria in many-objective optimization problems," in *Proc. Congr. Evol. Comput.*, New Orleans, LA, USA, 2011, pp. 2359–2366.

[17] D. Hadka and P. Reed, "Borg: An auto-adaptive many-objective evolutionary computing framework," *Evol. Comput.*, vol. 21, no. 2, pp. 231–259, May 2013.

[18] 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.

[19] M. Li, S. Yang, and X. Liu, "Bi-goal evolution for many-objective optimization problems," *Artif. Intell.*, vol. 228, pp. 45–65, Nov. 2015.

[20] X. Zhang, Y. Tian, and Y. Jin, "A knee point-driven evolutionary algorithm for many-objective optimization," *IEEE Trans. Evol. Comput.*, vol. 19, no. 6, pp. 761–776, Dec. 2015.

[21] 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.

[22] D. Gong, B. Xu, Y. Zhang, Y. Guo, and S. Yang, "A similarity-based cooperative co-evolutionary algorithm for dynamic interval multiobjective optimization problems," *IEEE Trans. Evol. Comput.*, vol. 24, no. 1, pp. 142–156, Feb. 2020.

[23] H. Ishibuchi and T. Murata, "A multi-objective genetic local search algorithm and its application to flowshop scheduling," *Comput. Ind. Eng.*, vol. 30, no. 4, pp. 957–968, 1996.

[24] Y. Jin, T. Okabe, and B. Sendho, "Adapting weighted aggregation for multiobjective evolution strategies," in *Proc. Int. Conf. Evol. Multi-Criterion Optim.*, 2001, pp. 96–110.

[25] T. Murata, H. Ishibuchi, and M. Gen, "Specification of genetic search directions in cellular multi-objective genetic algorithms," in *Proc. Int. Conf. Evol. Multi-Criterion Optim.*, 2001, pp. 82–95.

[26] Q. Zhang and L. Hui, "MOEA/D: A multiobjective evolutionary algorithm based on decomposition," *IEEE Trans. Evol. Comput.*, vol. 11, no. 6, pp. 712–731, Dec. 2007.

[27] 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.

[28] R. Cheng, Y. Jin, 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.

[29] N. Beumea, B. Naujoks, and M. Emmerich, "SMS-EMOA: Multiobjective selection based on dominated hypervolume," *Eur. J. Oper. Res.*, vol. 181, no. 3, pp. 1653–1669, 2007.

[30] E. Zitzler and S. Künzli, "Indicator-based selection in multiobjective search," in *Parallel Problem Solving From Nature PPSN VIII*, vol. 3242. Berlin, Germany: Springer, 2004, pp. 832–842.

[31] K. Bringmann and T. Friedrich, "An efficient algorithm for computing hypervolume contributions," *Evol. Comput.*, vol. 18, no. 3, pp. 383–402, 2010.

[32] J. Bader and E. Zitzler, "HypE: An algorithm for fast hypervolume-based many-objective optimization," *Evol. Comput.*, vol. 19, no. 1, pp. 45–76, Mar. 2011.

[33] L. While, L. Bradstreet, and L. Barone, "A fast way of calculating exact hypervolumes," *IEEE Trans. Evol. Comput.*, vol. 16, no. 1, pp. 86–95, Feb. 2012.

[34] L. M. S. Russo and A. P. Francisco, "Quick hypervolume," *IEEE Trans. Evol. Comput.*, vol. 18, no. 4, pp. 481–502, Aug. 2014.

[35] D. Brockhoff, T. Wagner, and H. Trautmann, "On the properties of the R2 indicator," in *Proc. Conf. Genet. Evol. Comput.*, 2012, pp. 465–472.

[36] R. H. Gómez and C. A. C. Coello, "MOMBI: A new metaheuristic for many-objective optimization based on the R2 indicator," in *Proc. Congr. Evol. Comput.*, 2013, pp. 2488–2495.

[37] H. Trautmann, T. Wagner, and D. Brockhoff, "R2-EMOA: Focused multiobjective search using R2-indicator-based selection," in *Proc. Int. Conf. Learn. Intell. Optim.*, vol. 7997, 2013, pp. 70–74.

[38] Y. Tian, R. Cheng, X. Zhang, F. Cheng, and Y. Jin, "An Indicator-based multiobjective evolutionary algorithm with reference point adaptation for better versatility," *IEEE Trans. Evol. Comput.*, vol. 22, no. 4, pp. 609–622, Aug. 2018.

[39] 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 Trans. Evol. Comput.*, vol. 18, no. 4, pp. 577–601, Aug. 2014.

[40] 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.

[41] D. Gong, J. Sun, and X. Ji, "Evolutionary algorithms with preference polyhedron for interval multi-objective optimization problems," *Inf. Sci.*, vol. 233, pp. 141–161, Jun. 2013.

[42] M. Rong, D. Gong, Y. Zhang, Y. Jin, and W. Pedrycz, "Multidirectional prediction approach for dynamic multiobjective optimization problems," *IEEE Trans. Cybern.*, vol. 49, no. 9, pp. 3362–3374, Sep. 2019.

[43] B. Li, J. Li, K. Tang, and X. Yao, "Many-objective evolutionary algorithms: A survey," *ACM Comput. Surveys*, vol. 48, no. 1, pp. 1–35, 2015.

[44] Z. He and G. G. Yen, "Many-objective evolutionary algorithms based on coordinated selection strategy," *IEEE Trans. Evol. Comput.*, vol. 21, no. 2, pp. 220–233, Apr. 2017.

[45] X. Cai, Z. Yang, Z. Fan, and Q. Zhang, "Decomposition-based-sorting and angle-based-selection for evolutionary multiobjective and many-objective optimization," *IEEE Trans. Cybern.*, vol. 47, no. 9, pp. 2824–2837, Sep. 2017.

[46] C. Liu, Q. Zhao, B. Yan, S. Elsayed, T. Ray, and R. Sarker, "Adaptive sorting-based evolutionary algorithm for many-objective optimization," *IEEE Trans. Evol. Comput.*, vol. 23, no. 2, pp. 247–257, Apr. 2019.

[47] S. Qin, C. Sun, G. Zhang, X. He, and Y. Tan, "A modified particle swarm optimization based on decomposition with different ideal points for many-objective optimization problems," *Complex Intell. Syst.*, vol. 6, no. 2, pp. 263–274, 2020.

[48] W. W. Hager, D. W. Hearn, and P. M. Pardalos, *Large Scale Optimization: State of the Art*. New York, NY, USA: Springer, 2013.

[49] R. M. Everson and J. E. Fieldsend, "Multiobjective optimization of safety related systems: An application to short-term conflict alert," *IEEE Trans. Evol. Comput.*, vol. 10, no. 2, pp. 187–198, Apr. 2006.

[50] Y. Tian, C. Lu, X. Zhang, K. C. Tan, and Y. Jin, "Solving large-scale multiobjective optimization problems with sparse optimal solutions via unsupervised neural networks," *IEEE Trans. Cybern.*, early access, Mar. 26, 2020, doi: 10.1109/TCYB.2020.2979930.

[51] Y. Tian, X. Zhang, C. Wang, and Y. Jin, "An evolutionary algorithm for large-scale sparse multiobjective optimization problems," *IEEE Trans. Evol. Comput.*, vol. 24, no. 2, pp. 380–393, Apr. 2020.

[52] X. Zhang, Y. Tian, R. Cheng, and Y. Jin, "A decision variable clustering-based evolutionary algorithm for large-scale many-objective optimization," *IEEE Trans. Evol. Comput.*, vol. 22, no. 1, pp. 97–112, Feb. 2018.

[53] L. Parsons, E. Haque, and H. Liu, "Subspace clustering for high dimensional data: A review," *ACM SIGKDD Explorations Newslett.*, vol. 6, no. 1, pp. 90–105, 2004.

[54] R. Cheng and Y. Jin, "A competitive swarm optimizer for large scale optimization," *IEEE Trans. Cybern.*, vol. 45, no. 2, pp. 191–204, Feb. 2015.

[55] R. Cheng and Y. Jin, "A social learning particle swarm optimization algorithm for scalable optimization," *Inf. Sci.*, vol. 291, pp. 43–60, Jan. 2015.

[56] Q. Yang *et al.*, "Segment-based predominant learning swarm optimizer for large-scale optimization," *IEEE Trans. Cybern.*, vol. 47, no. 9, pp. 2896–2910, Sep. 2017.

[57] M. N. Omidvar, X. Li, Z. Yang, and X. Yao, "Cooperative co-evolution for large scale optimization through more frequent random grouping," in *Proc. IEEE Congr. Evol. Comput.*, Barcelona, Spain, 2010, pp. 1–8.

[58] H. Wang, M. Liang, C. Sun, G. Zhang, and L. Xie, "Multiple-strategy learning particle swarm optimization for large-scale optimization problems," *Complex Intell. Syst.*, vol. 7, pp. 1–16, May 2020.

[59] L. M. Antonio and C. A. Coello Coello, "Use of cooperative coevolution for solving large scale multiobjective optimization problems," in *Proc. IEEE Congr. Evol. Comput.*, 2013, pp. 2758–2765.

[60] A. Song, Q. Yang, W.-N. Chen, and J. Zhang, "A random-based dynamic grouping strategy for large scale multi-objective optimization," in *Proc. IEEE Congr. Evol. Comput. (CEC)*, Vancouver, BC, Canada, 2016, pp. 468–475.

[61] L. M. Antonio and C. A. Coello Coello, "Decomposition-based approach for solving large scale multi-objective problems," in *Proc. Int. Conf. Parallel Problem Solving Nat.*, 2016, pp. 525–534.

[62] X. Ma *et al.*, "A multiobjective evolutionary algorithm based on decision variable analyses for multiobjective optimization problems with large-scale variables," *IEEE Trans. Evol. Comput.*, vol. 20, no. 2, pp. 275–298, Apr. 2016.

[63] H. Chen, X. Zhu, W. Pedrycz, S. Yin, G. Wu, and H. Yan, "PEA: Parallel evolutionary algorithm by separating convergence and diversity for large-scale multi-objective optimization," in *Proc. IEEE 38th Int. Conf. Distrib. Comput. Syst. (ICDCS)*, Vienna, Austria, 2018, pp. 223–232.

[64] M. Li and J. Wei, "A cooperative co-evolutionary algorithm for large-scale multi-objective optimization problems," in *Proc. Genet. Evol. Comput. Conf. Companion*, 2018, pp. 1716–1721.

[65] H. Zille, H. Ishibuchi, S. Mostaghim, and Y. Nojima, "A framework for large-scale multiobjective optimization based on problem transformation," *IEEE Trans. Evol. Comput.*, vol. 22, no. 2, pp. 260–275, Apr. 2018.

[66] C. He *et al.*, "Accelerating large-scale multiobjective optimization via problem reformulation," *IEEE Trans. Evol. Comput.*, vol. 23, no. 6, pp. 949–961, Dec. 2019.

[67] Y. Tian, X. Zheng, X. Zhang, and Y. Jin, "Efficient large-scale multiobjective optimization based on a competitive swarm optimizer," *IEEE Trans. Cybern.*, vol. 50, no. 8, pp. 3696–3708, Aug. 2020.

[68] C. He, R. Cheng, C. Zhang, Y. Tian, Q. Chen, and X. Yao, "Evolutionary large-scale multiobjective optimization for ratio error estimation of voltage transformers," *IEEE Trans. Evol. Comput.*, vol. 24, no. 5, pp. 868–881, Oct. 2020.

[69] R. Liu, R. Ren, J. Liu, and J. Liu, "A clustering and dimensionality reduction based evolutionary algorithm for large-scale multi-objective problems," *Appl. Soft Comput.*, vol. 89, Apr. 2020, Art. no. 106120.

[70] H. Chen, R. Cheng, J. Wen, H. Li, and J. Weng, "Solving large-scale many-objective optimization problems by covariance matrix adaptation evolution strategy with scalable small subpopulations," *Inf. Sci.*, vol. 509, pp. 457–469, Jan. 2020.

[71] C. He, R. Cheng, and D. Yazdani, "Adaptive offspring generation for evolutionary large-scale multiobjective optimization," *IEEE Trans. Syst., Man, Cybern., Syst.*, early access, Jul. 10, 2020, doi: 10.1109/TSMC.2020.3003926.

[72] R. Liu, J. Liu, Y. Li, and J. Liu, "A random dynamic grouping based weight optimization framework for large-scale multi-objective optimization problems," *Swarm Evol. Comput.*, vol. 55, Jun. 2020, Art. no. 100684.

[73] Y. Zhang, G.-G. Wang, K. Li, W.-C. Yeh, M. Jian, and J. Dong, "Enhancing MOEA/D with information feedback models for large-scale many-objective optimization," *Inf. Sci.*, vol. 522, pp. 1–16, Jun. 2020.

[74] K. Krishna and M. N. Murty, "Genetic *K*-means algorithm," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 29, no. 3, pp. 433–439, Jun. 1999.

[75] R. Cheng, Y. Jin, M. Olhofer, and B. Sendhoff, "Test problems for large-scale multiobjective and many-objective optimization," *IEEE Trans. Cybern.*, vol. 47, no. 12, pp. 4108–4121, Dec. 2017.

[76] A. Zhou, Y. Jin, Q. Zhang, B. Sendhoff, and E. P. K. Tsang, "Combining model-based and genetics-based offspring generation for multi-objective optimization using a convergence criterion," in *Proc. IEEE Congr. Evol. Comput.*, 2006, pp. 892–899.

[77] P. Czyzżak and A. Jaszkiewicz, "Pareto simulated annealing—A meta-heuristic technique for multiple-objective combinatorial optimization," *J. Multi-Criteria Decis. Anal.*, vol. 7, no. 1, pp. 34–47, 1998.

[78] Y. Tian, R. Cheng, X. Zhang, and Y. Jin, "PlatEMO: A MATLAB platform for evolutionary multi-objective optimization [Educational Forum]," *IEEE Comput. Intell. Mag.*, vol. 12, no. 4, pp. 73–87, Nov. 2017.

[79] R. G. D. Steel and J. H. Torrie, *Principles and Procedures of Statistics: A Biometrical Approach*. New York, NY, USA: McGraw-Hill Kogakusha, Ltd., 1980.

[80] J. Dufner, U. Jensen, and E. Schumacher, *Statistik MIT SAS*. Heidelberg, Germany: Springer, 2013.

[81] K. Deb and M. Goyal, "A combined genetic adaptive search (GeneAS) for engineering design," *Comput. Sci. Informat.*, vol. 26, no. 4, pp. 30–45, 1996.

[82] C. Dai, Y. Wang, M. Ye, X. Xue, and H. Liu, "An orthogonal evolutionary algorithm with learning automata for multiobjective optimization," *IEEE Trans. Cybern.*, vol. 46, no. 12, pp. 3306–3319, Dec. 2016.

Shufen Qin received the B.Sc. degree in computer science and technology from Jinzhong University, Jinzhong, China, in 2017. She is currently pursuing the Ph.D. degree with the Department of Electronic Information Engineering, Taiyuan University of Science and Technology, Taiyuan, China.

Her current research interests include particle swarm optimization, multi/many-objective optimization, large-scale optimization, surrogate-assisted evolutionary optimization, and their applications to practical engineering problems.

Chaoli Sun (Member, IEEE) received the B.Sc. and M.Sc. degrees from Hohai University, Nanjing, China, in 2000 and 2003, respectively, and the Ph.D. degree from the Taiyuan University of Science and Technology, Taiyuan, China, in 2011.

She is a Professor with the Department of Computer Science and Technology, Taiyuan University of Science and Technology. Her research interests include swarm intelligence, particle swarm optimization, surrogate-assisted evolutionary optimization, and their applications to practical engineering problems.

Prof. Sun is an Associate Editor of *Soft Computing and Complex & Intelligent Systems*, and was the Chair of the Task Force on Data-Driven Evolutionary Optimization of Expensive Problems within the Evolutionary Computation Technical Committee of IEEE Computational Intelligence Society from 2015 to 2020. She has also served as a Reviewer for journals, including IEEE TRANSACTIONS ON EVOLUTIONARY COMPUTATION, IEEE TRANSACTIONS ON CYBERNETICS, IEEE TRANSACTIONS ON NEURAL NETWORKS AND LEARNING SYSTEMS, and *IEEE Computational Intelligence Magazine*.

Yaochu Jin (Fellow, IEEE) received the B.Sc., M.Sc., and Ph.D. degrees from Zhejiang University, Hangzhou, China, in 1988, 1991, and 1996, respectively, and the Dr.-Ing. degree from Ruhr University Bochum, Bochum, Germany, in 2001.

He is a Professor of Computational Intelligence with the Department of Computer Science, University of Surrey, Guildford, U.K., where he heads the Nature Inspired Computing and Engineering Group. He is also a Finland Distinguished Professor funded by the Finnish Funding Agency for Innovation (Tekes) and a Changjiang Distinguished Visiting Professor appointed by the Ministry of Education, China. He has coauthored over 300 peer-reviewed journal and conference papers and been granted eight patents on evolutionary optimization. He is also particularly interested in the application of nature-inspired algorithms to solving real-world optimization, learning and self-organization problems. His research interests lie primarily in the crossdisciplinary areas of computational intelligence, computational neuroscience, and computational systems biology.

Prof. Jin was a recipient of the 2014 and 2016 *IEEE Computational Intelligence Magazine* Outstanding Paper Award, and the 2017 IEEE TRANSACTIONS ON EVOLUTIONARY COMPUTATION Outstanding Paper Award. He is the Editor-in-Chief of the IEEE TRANSACTIONS ON COGNITIVE AND DEVELOPMENTAL SYSTEMS and *Complex & Intelligent Systems*. He is an IEEE Distinguished Lecturer from 2017 to 2019 and was the Vice President for Technical Activities of the IEEE Computational Intelligence Society from 2014 to 2015.

Ying Tan received the B.Sc. degree in industrial automation from the Taiyuan University of Science and Technology, Taiyuan, China, in 1986, and the M.Sc. degree in control theory and application from Tsinghua University, Beijing, China, in 1994.

She is a Professor with the Department of Computer Science and Technology, Taiyuan University of Science and Technology. Her current research interests include intelligent computing, cloud computing, and distributed network security.

Jonathan Fieldsend (Member, IEEE) received the B.A. degree (Hons.) in economics from Durham University, Durham, U.K., in 1998, the M.Sc. degree in computational intelligence from Plymouth University, Plymouth, U.K., in 1999, and the Ph.D. degree in computer science from the University of Exeter, Exeter, U.K., in 2003.

He is currently a Professor of Computational Intelligence with the University of Exeter. His research interests include multiobjective optimization, optimization with uncertainty, robust optimization, multimodal optimization, expensive optimization, machine learning, and data visualization.

Prof. Fieldsend is a Member of the IEEE Computational Intelligence Society (CIS) and the Vice Chair of the IEEE CIS Task Force on Multimodal Optimization and the IEEE CIS Task Force on Data-Driven Evolutionary Optimization of Expensive Problems. He is an Associate Editor of IEEE TRANSACTIONS ON EVOLUTIONARY COMPUTATION.