

Received 8 September 2023, accepted 24 September 2023, date of publication 2 October 2023, date of current version 10 October 2023. *Digital Object Identifier 10.1109/ACCESS.2023.3321313*

RESEARCH ARTICLE

Large-Scale Evolutionary Multi-Objective Optimization Based on Direction Vector Sampling

YIN[G](https://orcid.org/0000-0001-8646-3900)J[I](https://orcid.org/0000-0002-3012-7913)AN XIONG[®] AND XUHUA SHI[®]

Faculty of Electrical Engineering and Computer Science, Ningbo University, Ningbo 315211, China

Corresponding author: Xuhua Shi (shixuhua@nbu.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 61773225 and Grant 61803214; in part by the K. C. Wong Magna Fund in Ningbo University, China; and in part by the Ningbo Key Research and Development Program under Grant 2023Z067.

ABSTRACT The large-scale multi-objective optimization problem is characterized by a large decision space. How to design an efficient optimization algorithm that can search a large decision space and find the global optimum in the objective space is a very challenging problem at present. In order to solve this problem, this paper proposes a sampling strategy based on direction vectors, which takes into account both convergence and diversity. First, select some excellent individuals who are close to the ideal point based on the reference vector. Secondly, construct a three-way search direction vector using the boundary point and an additional center point, and execute a directional sampling strategy called the convergence-related sampling strategy to improve the convergence of the algorithm. After, the direction vector is constructed among excellent individuals and executes a directional sampling strategy called the diversity-related sampling strategy to maintain the diversity of the population. Finally, the adjustment strategy of the reference vector in the Reference Vector Guidance Algorithm (RVEA) is adopted to adjust the reference vector. Numerical experiments are performed on large-scale multi-objective benchmark problem sets with 500, 1000, and 2000 decision variables and compared with the state-of-the-art algorithms. Experimental results show that the algorithm proposed in this paper is effective and can obtain solutions that are significantly better than those of the compared algorithms.

š **INDEX TERMS** Large-scale, multi-objective optimization, decision space, direction vector, directional sampling.

I. INTRODUCTION

In real-life production, there are problems involving simultaneous optimization of multiple objectives, which are called Multi-Objective Problems (MOPs) [\[1\]. In](#page-18-0) the past few decades, scholars have proposed a large number of multi-objective evolutionary algorithms (MOEAs) to solve such problems, which can be broadly categorized into three categories: Pareto dominance based algorithm [\[2\],](#page-18-1) [\[3\],](#page-18-2) [\[4\]; in](#page-18-3)dicator based algorithms [\[5\],](#page-18-4) [\[6\],](#page-18-5) [\[7\]; d](#page-18-6)ecomposition based algorithm [\[8\],](#page-18-7) [\[9\],](#page-18-8) [\[10\].T](#page-18-9)he aforementioned techniques, however, are only appropriate for multi-objective optimization problems with a limited set of decision variables because they only consider the scalability of the number of objectives. But for the real world, there exist large-scale multi-objective optimization problems (LSMOPs) that contain hundreds or thousands of decision variables, such as large-scale public transportation network design problems [\[11\]. T](#page-18-10)he search space of the algorithm grows exponentially as the number of decision variables rises [\[12\], r](#page-18-11)esulting in the dimensionality curse $[13]$, while the search efficiency of traditional multi-objective optimization algorithms rapidly declines [\[14\],](#page-19-0) [\[15\], a](#page-19-1)nd it becomes more challenging to produce promising offspring solutions in a larger decision space. In order to solve these problems, many large-scale multiobjective evolutionary algorithms (LSMOEAs) have been proposed, which can be roughly divided into the following categories.

The associate editor coordinating the review of this manuscript and approving it for publication was Jason Gu $^{\textcircled{\textsf{D}}}$ [.](https://orcid.org/0000-0002-7626-1077)

1) MOEAs based on cooperative coevolution (CC). Based on the principle of divide and conquer, this kind of

MOEAs divide the decision variables into numerous groups with various dimensions, optimize each group separately, and then achieve the optimization of the entire problem through the co-evolution between subpopulations. CCGDE3 [\[16\]](#page-19-2) divides decision variables into multiple subpopulations of equal length based on a random grouping strategy, and then uses GDE3 [\[17\]](#page-19-3) to optimize each subpopulation separately. MOEA/D2 [\[18\]](#page-19-4) further proposed to combine MOEA/D [8] [wi](#page-18-7)th co-evolutionary techniques for decomposition in both the decision space and the objective space at the same time.

- 2) MOEAs based on decision variable analysis. This kind of MOEAs is used to more precisely categorize decision variables by evaluating the link between them. Ma et al. [\[19\]](#page-19-5) proposed an MOEA based on decision variable analyses (MOEA/DVA). The algorithm observes the change of the dominant state by perturbing the value of the decision variable, divides the decision variable into convergence variables, diversity variables and mixed variables and optimizes them respectively. Zhang et al. [\[20\]](#page-19-6) proposed an evolutionary algorithm based on clustering of decision variables (LMEA). The algorithm divides the decision variables into convergence variables and diversity variables through a clustering algorithm, and then uses two different strategies to optimize the two groups of variables. Li and Wei [\[21\]](#page-19-7) proposed a coevolutionary algorithm to solve large-scale multi-objective problems by using a fast interdependence recognition grouping method. In addition, PEA [\[22\], F](#page-19-8)R [\[23\], e](#page-19-9)tc. also belong to the category based on decision variable analysis.
- 3) MOEAs based on problem transformation. This type of MOEAs transforms the original large-scale decision variable optimization problem into a small-scale weight optimization problem through the problem transformation function. Zille et al. [\[24\]](#page-19-10) proposed a weight optimization framework (WOF). The algorithm divides decision variables into multiple groups and optimizes the associated weight vector for each group, thereby transforming it into a small-scale MOP. To enhance the effectiveness of this framework, Liu et al [\[25\]](#page-19-11) proposed to use random dynamic grouping rather than ordered grouping. He et al. [\[26\]](#page-19-12) proposed a problem reconstruction framework (LSMOF). The algorithm constructs a bidirectional reference vector in the decision space and associates it with a set of weight variables. The weight variables are then optimized using the HV indicator to track the Pareto optimal set and speed up the search process.
- 4) The last category is MOEAs that do not belong to the above three categories. For example, Tian et al. [\[27\]](#page-19-13) proposed a modified competitive swarm optimizer for large-scale multi-objective problems (LMOCSO). The algorithm proposes a new particle update strategy,

in which inferior particles learn from superior particles to generate promising offspring to accelerate the search for the global optimal solution. He et al. [\[28\]](#page-19-14) propose an adaptive offspring generation method for large-scale multi-objective optimization (DGEA). The algorithm constructs convergence and diversity search directions through dominated and non-dominated solutions, and adaptively guides the generation of offspring. Qin et al. [\[29\]](#page-19-15) introduce a direct sampling method for large-scale problems (LMOEA-DS), which constructs a bidirectional search direction vector through boundary points and individuals closest to the origin. Directional sampling is then performed on the direction vector to generate offspring, while complementary non-dominated sorting and decomposition-based environmental selection strategies are used for selection. Yang et al. [\[30\]](#page-19-16) proposes a fuzzy decision variables framework for large-scale multi-objective optimization (FDV), which divides the evolution process into two main stages: fuzzy evolution and precise evolution by fuzzy operations on decision variables.

Although the existing large-scale multi-objective algorithms can improve the search efficiency to some level, each category of the algorithms has drawbacks. The performance of MOEAs algorithm based on cooperative coevolution depends on the grouping of decision variables. When the grouping of decision variables is not appropriate, the performance of the algorithm will significantly suffer. These algorithms, which use methods like random grouping [\[31\],](#page-19-17) linear grouping [\[32\],](#page-19-18) ordered grouping [\[33\],](#page-19-19) differential grouping [\[34\], a](#page-19-20)nd others, do not account for the interrelationships between decision variables, so they are not suitable for large-scale multi-objective optimization problems with interactions between decision variables. In order to divide the decision variables more accurately, MOEAs based on decision variable analysis often need to spend a lot of function evaluation, and the decision variable categories generated by the decision variable analysis are few, so the corresponding sub-problems may still be large-scale problems. The MOEAs based on problem transformation are prone to fall into local optima, and the final solution set has poor diversity. The LMOCSO [\[27\], D](#page-19-13)GEA [\[28\], L](#page-19-14)MOEA-DS [\[29\]](#page-19-15) algorithms directly search for the optimal solution in the original decision space. They still cannot effectively find promising offspring solutions to speed up search for some problems with high dimensional decision variables. The FDV [\[30\]](#page-19-16) algorithm introduces fuzzy operations to increase the complexity of the algorithm and increase the computational burden.

Based on the above discussion, a large-scale multiobjective evolutionary algorithm based on direction vector sampling (LSMOEA-DVS) is proposed. The main idea of the algorithm is to construct a search direction vector that is related to convergence and diversity through excellent individuals, boundary points and a center point. The direction vector is then directional sampling to produce the desired

offspring solutions in the decision space. Since the method proposed in this article directly generates offspring through direction vector sampling in the original decision space, it avoids the defects caused by the need to group decision variables in the first two types of algorithms. At the same time, it also avoids the problem of lack of diversity caused by the need for problem conversion in the third type of algorithm. In addition, the method proposed in this article further improves the performance of the last type of algorithm by enhancing the direction vector generation and implementing a sampling strategy that considers both convergence and diversity. The main contributions of this work are as follows:

- 1) A convergence-related sampling strategy is proposed: in each generation of the search, three promising search direction vectors are constructed through each excellent individual close to the ideal point, the boundary point and the additionally introduced center point. Then, directional sampling is performed along each search direction vector to generate individuals with good convergence.
- 2) A diversity-related sampling strategy is proposed: in each generation of the search, the strategy takes a set of diversity-related excellent individuals that is closed to the ideal point to construct a set of search directions. Directional sampling is then performed along each search direction. Additionally, non-dominated individuals perform crossover mutation operations to further improve diversity.
- 3) The algorithm proposed in this paper takes both convergence and diversity into account through two sampling strategies in the generation of offspring. In order to verify the effectiveness of the proposed LSMOEA-DVS algorithm in solving LSMOPs, it is compared with other five representative large-scale MOEAs. The experimental results show that LSMOEA-DVS is significantly better than other large-scale MOEAs evolutionary algorithm. In addition, an ablation experiment was conducted to confirm the efficacy of the proposed method, which verified the effectiveness of the method.

The rest of this article is organized as follows. In Section [II,](#page-2-0) we introduce the related work of the paper. Section [III](#page-3-0) introduces the LSMOEA-DVS algorithm framework and the principle of program implementation. Section [IV](#page-6-0) is the analysis of experimental results, which are compared with that of other algorithms. Section V is the conclusions of this paper.

II. RELATED WORK

A. MULTIOBJECTIVE OPTIMIZATION PROBLEM MOPs can be expressed mathematically as follows [\[35\]:](#page-19-21)

$$
\begin{cases} \min F(X) = (f_1(X), f_2(X), \dots, f_m(X)) \\ \text{subject to } X \in \Omega \end{cases} \tag{1}
$$

where $X = (x_1, x_2, \ldots, x_D)$ is a decision vector from a *D*-dimensional decision space Ω , and *F* (*X*) is the objective function vector, which contains m conflicting objective

functions. Due to the presence of conflicting objective functions, there is no single solution that can simultaneously minimize all objective functions. In the field of multi-objective evolution, a non-dominated solution set is defined as a trade-off between different objectives. Assuming X_1, X_2 ∈ Ω , we define X_1 Pareto dominate X_2 as $X_1 \prec X_2$, if and only if $f_i(X_1) \le f_i(X_2)$ ($\forall i \in \{1, 2, ..., m\}$), and at least on one objective $f_i(X_1) \leq f_i(X_2)$ ($j \in \{1, 2, ..., m\}$). All Pareto optimal solutions in the decision space are called Pareto optimal set (PS), and the projection of PS in the objective space is called Pareto optimal front (PF).

B. OBJECTIVE VALUE CONVERSION

Since the initial point of the reference vector is always the origin of the coordinates, in order to facilitate the division of the population individuals, the objective value of the population individual P_t needs to be transformed. Assuming F_t = $f_{t,1}, f_{t,2}, \ldots, f_{t,|p_t|}$, *t* represents the *t*-th iteration. Converted by the following formula:

$$
f'_{t,i} = f_{t,i} - z_t^{\min} \tag{2}
$$

where $i = 1, 2, \ldots, |p_t|, f_{t,i}$ and $f'_{t,i}$ are the objective vectors of individual *i* before and after the transformation. $z_t^{min} =$
 $\left(z_{min}^{min} - z_{min}\right)$ represents the minimum value of each $z_{t,1}^{min}$, $z_{t,2}^{min}$, ..., $z_{t,m}^{min}$ represents the minimum value of each objective in the current population. The purpose of this is to ensure that all objective values and reference vectors of the transformed population are in the first quadrant of the objective space. At this point, the ideal point is the coordinate origin.

C. SOLUTION ASSOCIATION

In order to achieve a population with good diversity, it is necessary to assign individuals to uniformly distributed reference vectors [\[36\], a](#page-19-22)nd allocate individuals based on the angle relationship between individuals and reference vectors in the objective space. The formula is as follows:

$$
cos\theta_{t, i, j} = \frac{f'_{t, i} \cdot v_{t, j}}{\left\|f'_{t, i}\right\| \cdot \left\|v_{t, j}\right\|}
$$
 (3)

where $\theta_{t,i,j}$ represents the angle between the objective vector $f'_{t,i}$ after the transformation of the population individual $P_{t,i}$ and the reference vector $v_{t,j}$, if and only When the angle between the population individual $P_{t,i}$ and the reference vector $v_{t,k}$ is the smallest (the cosine value is the largest), it is assigned to the reference vector $v_{t,k}$.

D. MAIN FRAMEWORK

The main framework of the LSMOEA-DVS algorithm proposed in this paper is listed in Algorithm [1.](#page-3-1) It can be seen from Algorithm [1](#page-3-1) that the LSMOEA-DVS algorithm mainly uses two sampling strategies to generate offspring population individuals, which is also the main contribution of this paper. In the following subsections, we will describe the main parts of Algorithm [1](#page-3-1) in detail.

Algorithm 1 LSMOEA-DVS

Input: *N* (population size), *FEmax* (the maximum number of function evaluation), *mu* (number of candidate solutions to construct the direction vector).

Output: *P* (final population).

- 1: V_0 , V^* ← generate-reference-vectors(N);
- 2: $V \leftarrow V_0 \cup V^*$;
- 3: $P \leftarrow$ Initialization(N);
- 4: **while** termination criterion is not fulfilled **do**
- 5: *coffspring* ← convergence-related sampling (P, V, mu) ;
- 6: *doffspring*, *offspring* ← diversity-related sampling($P \cup \text{coffspring}, V, mu, N$);
- 7: *P* ← Environmental Selection(*P* ∪ *coffspring*∪ *doffspring* ∪ *offspring*, *V*);
- 8: $V(1:N) \leftarrow$ reference-vector-adaptation(*P*, *V*₀);
- 9: **if** $FE > \frac{1}{2}FE_{max}$ then
- 10: $V(N + 1 : end) \leftarrow$ Reference-Vector-Regeneration $(P, V (N + 1 : end));$
- 11: **end if**
- 12: **end while**

III. THE PROPOSED ALGORITHM

A. PREPROCESSING

The purpose of preprocessing is to select excellent individuals from each generation of the population that are required by the subsequent sampling strategy to construct the direction vector. We first transform the individual objective value of the current population, and then assign it to the current reference vector V^t . The reference vector that assigns at least one individual is called the active reference vector V_{act}^t . The reference vector of unassigned population individuals is called the inactive reference vector V_{ina}^t . Use K-means clustering to cluster the active reference vector V_{act}^t into *mu* classes. If the number of active reference vectors V_{act}^t is less than the number of clusters *mu* required for clustering, then cluster the active reference vectors V_{act}^t into the number of active reference vectors *Nact* classes. Therefore, the population individuals are divided into *min* (*mu*,*Nact*) subpopulations. This method of clustering the active reference vectors is based on the K-RVEA algorithm [\[37\], s](#page-19-23)o that the selected individuals have the widest distribution. Finally, an individual who is closest to the ideal point is selected from each subpopulation. In this way, individuals with good convergence and diversity can be obtained to construct the direction vector. To illustrate how to find excellent individuals to construct the direction vector, an example is given in Figure [1.](#page-3-2) The four blue dots in Figure [1](#page-3-2) represent the current population individuals, and the five solid black lines represent the reference vector. It can be seen that the reference vector V_a has no population individuals assigned, so it is an inactive reference vector, while the remaining four assigned population individuals are active reference vectors. The four active reference vectors are clustered into two categories, Cluster1 and Cluster2. Then,

FIGURE 1. The selection of the best individual closest to the ideal point constituting the direction vector.

the individuals X_2 and X_4 , which are closest to the ideal point (original point), are selected from each category. *X*² and *X*⁴ are the excellent individuals required to construct the direction vector.

B. SAMPLING STRATEGY

1) CONVERGENCE-RELATED SAMPLING STRATEGY

With the increase of the dimension of decision variables, the search space of the algorithm increases exponentially. The traditional approach of generating offspring by cross-mutation makes the algorithm slow to converge in a large search space. Construct a promising direction vector to guide the search towards the Pareto optimal front, which can greatly speed up the convergence of the algorithm. Both the LSMOF [\[26\]](#page-19-12) and LMOEA-DS [\[29\]](#page-19-15) algorithms use the upper and lower boundary points of the decision space and excellent individuals to construct the direction vector of the bidirectional search. It is expected that the direction vector can intersect with the Pareto optimal set in the decision space, so as to follow the direction vector and generating population individuals on the Pareto optimal front to accelerate convergence. However, the effectiveness of this method is contingent upon the presence of an intersection between the direction vector and the Pareto optimal set. If the direction vector is far from the Pareto optimal set, it cannot guarantee the generation of population individuals with better convergence. Figure [2](#page-4-0) shows an example of constructing a direction vector in a two-dimensional decision variable space using only boundary points. In the figure, the blue curve represents PS in the decision variable space, while the four black solid circles represent the selected outstanding population individuals. Additionally, the red solid line represents the direction vector of the construct. As can be seen from Figure [2,](#page-4-0) the four selected population individuals exhibit a concentration in the lower half of the space, whereas the PS curves are predominantly located in the upper half of the space. The direction vector that has been constructed does not intersect with PS, which makes it difficult for the offspring individuals

FIGURE 2. In the two-dimensional decision variable space, use the boundary points to construct the direction vector.

FIGURE 3. In the two-dimensional decision variable space, use the boundary points and a center point to construct a direction vector.

generated along this vector to approach the vicinity of PS, so that the optimal solution cannot be directly found. Based on this problem, in this study, we decided to introduce an additional center point *C*, This point will be used to construct direction vectors with excellent individuals, allowing each excellent individual to generate three direction vectors. These vectors will then form a three-way search. Thereby increasing the possibility of intersection between the direction vector and PS. Figure [3](#page-4-1) shows an example of constructing the direction vector in the two-dimensional decision variable space using the boundary points and the center point. From Figure [2](#page-4-0) and Figure [3,](#page-4-1) it can be seen that the inclusion of the center point *C* enhances the likelihood of intersection between the direction vector and PS. Consequently, the distribution of offspring individuals generated along the direction vector becomes more extensive, which is beneficial to improve the exploration efficiency of the algorithm across the entire decision variable space.

After selecting *mu* excellent individuals $X = \{X_1, \ldots, X_{mu}\}$ from the parent population *P* through preprocessing, a direction vector is constructed using the upper and lower boundary points, as well as the center point, which is defined as follows:

$$
cdv^{l,i}=X_i-L
$$

$$
cdv^{u,i} = X_i - U
$$

\n
$$
cdv^{c,i} = X_i - C
$$
\n(4)

where *L* and *U* represent the lower and upper boundary points of the decision variable space, and the center point *C* is defined as follows:

$$
C = L + \frac{U - L}{2} \tag{5}
$$

Generate a set of direction vectors $c dv = \{cdv^{l,1}, cdv^{u,1}\}$, *cdvc*,¹ , . . . , *cdvl*,*mu* , *cdvu*,*mu* , *cdvc*,*mu*}. Each direction vector represents a search direction. The search starting point can be the upper boundary point, the lower boundary point, or the center point. We randomly sample multiple solutions in each search direction, the formula for this process is as follows:

$$
coffspring_j^{l,i} = L + \gamma_j^{l,i} \cdot \frac{cdv^{l,i}}{\|cdv^{l,i}\|}
$$

$$
coffspring_k^{u,i} = U + \gamma_k^{u,i} \cdot \frac{cdv^{u,i}}{\|cdv^{u,i}\|}
$$

$$
coffspring_i^{c,i} = C \pm \gamma_i^{c,i} \cdot \frac{cdv^{c,i}}{\|cdv^{c,i}\|}
$$
(6)

where the sampling coefficients $\gamma_j^{l,i}$ and $\gamma_k^{u,i}$ are random numbers uniformly distributed on [0, $||U - L||$], $\gamma_l^{c,i}$ is a random number uniformly distributed on [0, $\frac{\|U - L\|}{2}$ $\frac{-L\parallel}{2}$], $\|\cdot\|$ represents the Euclidean distance. This ensures that the individuals generated along the direction vector $cdv^{l,i}$, $cdv^{u,i}$ and $cdv^{c,i}$ cover the entire search space, *j*, *k*, and *l* represent the *j*-th, *k*-th and *l*-th samplings along the direction vector $cdv^{l,i}$, $cdv^{u,i}$ and $cdv^{c,i}$. Figure [4](#page-5-0) gives an illustrative example in the two-dimensional decision variable space. An excellent individual X_1 and boundary points L , U construct direction vectors $cdv^{l,1}$ and $cdv^{u,1}$. X_1 and the center point *C* construct a direction vector $cdv^{c,1}$. At the same time, random sampling is performed on $cdv^{l,1}$, $cdv^{u,1}$, and $cdv^{c,1}$ to generate $\text{coffspring}_i^{l,1}$, $\text{coffspring}_k^{u,1}$, and $\text{coffspring}_l^{c,1}$.

Algorithm [2](#page-5-1) gives the pseudocode for the convergencerelated sampling strategy method. First, the preprocessing selects *mu* excellent individuals (line 1). Then 3 direction vectors were constructed for each excellent individual (lines 2-4). Subsequently, random sampling was performed on each direction vector to generate offspring individuals (lines 5-7). Duplicate individuals are eliminated (line 8) and Out-of-bounds handling is performed on the resulting offspring individuals (lines 9-11). Output the offspring individuals generated through the convergence-related sampling strategy.

2) DIVERSITY-RELATED SAMPLING STRATEGY

Although the convergence-related sampling strategy can generate individuals with good convergence to accelerate the convergence process, it cannot guarantee that the generated individuals will have good diversity. Therefore, a diversityrelated sampling strategy is proposed to improve the diversity

Algorithm 2 Convergence-Related Sampling

Input: *P* (the parent population), *mu* (number of excellent individuals to construct the direction vector), *V* (a set of reference vectors for identifying search directions). **Output:** *coffspring* (sampling solution)

- 1: $gp \leftarrow$ The *mu* excellent individuals obtained by preprocessing for constructing the direction vector;
- 2: **for** $i = 1 : |gp|$ **do**
- 3: Define three direction vectors according to equation [\(4\)](#page-4-2) and individual *i*;

4: **end for**

- 5: **for** $i = 1$: |gp| **do**
- 6: Generate offspring individuals *coffspring* according to equation (6) ;

7: **end for**

8: delete the same individual;

9: **for** $i = 1$: $|coffspring|$ **do**

- 10: Out-of-bounds handling based on equation [\(9\)](#page-5-2) for each individual *i*;
- 11: **end for**
- 12: Output *coffspring*.

of each generation of populations. From the offspring individuals *coffspring* generated by the convergence-related sampling strategy and the parent population *P*, *mu* excellent individuals $X = \{X_1, \ldots, X_{mu}\}\$ are selected through preprocessing. The *mu* excellent individuals are selected in sequence to serve as the starting point of the construction direction vector. Then, one of the other *mu* − 1 candidate individuals is randomly selected to serve as the end point of the construction direction vector. The construction process is defined as follows:

$$
ddv^i = X_j - X_i \tag{7}
$$

where $i = \{1, 2, \ldots, mu\}$, *j* belongs to any number not equal to *i* in $\{1, 2, \ldots, mu\}$. Generate a set of direction vectors $ddv = \{ddv^1, ddv^2, \dots, ddv^{mu}\},$ each direction vector represents a search direction. The search starting point is *Xⁱ* , and multiple solutions are randomly sampled in each direction. The process formula is as follows:

$$
doffspring_j^i = X_i \pm \delta_j^i \cdot \frac{ddv^i}{\|ddv^i\|} \tag{8}
$$

The sampling coefficient δ_j^i is a random number uniformly distributed on [0, $\left\| ddv^i \right\|$], $\| \cdot \|$ represents the Euclidean distance, and *j* represents the *j*-th sampling along the direction vector *ddvⁱ* . After the offspring individual *doffspring* is generated, it is combined with the parent *P* and the offspring individual *coffspring* generated by the convergence sampling strategy to select *N* non-dominated individuals through nondominated sorting. Cross-mutating it to generate offspring *offspring* to further improve population diversity. Figure [5](#page-5-3) gives an illustrative example. In the two-dimensional decision variable space, a direction vector *ddv*¹ is constructed between

FIGURE 4. Construction and sampling of direction vectors in convergence-related sampling strategy.

FIGURE 5. Construction and sampling of direction vectors in diversity-related sampling strategy.

two excellent individuals X_1 and X_2 and at the same time, d *offspring*¹ is randomly sampled on ddv ¹.

Algorithm [3](#page-6-1) gives the pseudocode of the diversity-related sampling strategy method. First, the preprocessing selects *mu* excellent individuals (line 1). Then construct a direction vector for each excellent individual (lines 2-4) and randomly sample from each direction vector to generate offspring individuals (lines 5-7). Duplicate individuals are removed (line 8) and out-of-bounds handing is performed on the resulting offspring individuals (lines 9-11). Finally perform non-dominated sorting on the individuals of the current population, select the top *N* non-dominated solutions for cross-mutation (lines 12-13), and output the offspring individuals generated by the diversity-related sampling strategy.

3) OUT OF BOUNDS HANDLING

Some solutions generated by the sampling coefficients may exceed the boundary of the decision space, so here we use the value of the decision variable beyond the boundary position to reproject back to the feasible region. The specific process is defined as follows:

$$
x'_{i,d} = \min\left(x_{i,d}^{cd}, U_d\right)
$$

Algorithm 3 Diversity-Related Sampling

Input: *P* (the parent population), *coffspring* (convergencerelated sampling individuals), *mu* (number of excellent individuals to construct the direction vector), *V* (a set of reference vectors for identifying search directions), *N* (population size).

Output: *doffspring* (sampling solution), *offspring*

- 1: $gp \leftarrow$ The *mu* excellent individuals obtained by preprocessing for constructing the direction vector;
- 2: **for** $i = 1 : |gp|$ **do**
- 3: Define a direction vector according to equation [\(7\)](#page-5-4) and individual *i*;
- 4: **end for**
- 5: **for** $i = 1 : |gp|$ **do**
- 6: Generate offspring individuals *doffspring* according to equation (8) ;
- 7: **end for**
- 8: delete the same individual;
- 9: **for** $i = 1$: $\left| \frac{d \text{off} \text{spring}}{d \text{ of } \text{off}} \right|$ **do**
- 10: Out-of-bounds handling based on equation [\(9\)](#page-5-2) for each individual *i*;
- 11: **end for**
- 12: Sort the population *P* ∪ *coffspring* ∪ *doffspring* non-dominated and select the top *N* non-dominated individuals;
- 13: Cross mutation operation produces *offspring*;
- 14: Output *doffspring*, *offspring*.

$$
x'_{i,d} = \max\left(x_{i,d}^{cd}, L_d\right) \tag{9}
$$

where $x_{i,d}^{cd}$ represents the *d*-th decision variable of the *i*-th solution \hat{X}_i^{cd} . U_d and L_d are the values of the *d*-th decision variable at the upper and lower boundary points. Figure [6](#page-6-2) shows such an example. The x_1 value $x_{i,1}^{cd}$ of individual X_i^{cd} and the x_2 value $x_{k,2}^{cd}$ of individual X_k^{cd} sampled on the direction vector are beyond the decision variable space, and these individuals are projected back to positions X_i' and X_k' in the feasible region, respectively.

C. ENVIRONMENTAL SELECTION

Decomposition-based environmental selection strategy for populations *P*, *coffspring*, *doffspring* and *offspring*. First, the objective value of each individual in the population is transformed and then assigned to the reference vector *V*. From the individuals assigned by each reference vector, the individual with the closest Euclidean distance to the ideal point in the objective space is selected and retained for the next generation.

D. REFERENCE VECTOR ADJUSTMENT

1) REFERENCE-VECTOR-ADAPTATION

Since the individual objective values of the population are not normalized, there may be some MAOPs where different

FIGURE 6. Project out-of-bounds decision variables back to the feasible region.

objectives are scaled to different ranges. As a result, even if a uniformly distributed reference vector is used, it may not yield a uniformly distributed solution. We adopt the reference vector adaptation method in RVEA [9] [to a](#page-18-8)djust the reference vector according to the objective value, refer to [\[9\]](#page-18-8) for more details.

2) REFERENCE-VECTOR-REGENERATION

In some optimization problems characterized by irregular PF geometry, uniformly distributed reference vectors may result in the presence of invalid reference vectors, which lack non-dominated individuals associated with them. If we continue to rely on them for guiding environment selection, this practice will result in a decrease density for obtaining Pareto optimal solutions, as well as a wastage of computational resources on these invalid reference vectors. Therefore, in the second half of the optimization process, the reference vector regeneration strategy in RVEA [9] [is u](#page-18-8)sed to regenerate these invalid reference vectors, please refer to [\[9\]](#page-18-8) for more details.

IV. EXPERIMENTAL STUDY

The performance of the proposed LSMOEA-DVS is evaluated by comparing with five algorithms, including LMOEA-DS [\[29\], L](#page-19-15)MOCSO [\[27\], L](#page-19-13)SMOF [\[26\], W](#page-19-12)OF [\[24\]](#page-19-10) and FDV [\[30\]](#page-19-16) on the seven test problems of the DTLZ [\[38\]](#page-19-24) test suite, the nine test problems of the LSMOP [\[14\]](#page-19-0) test suite, and the nine test problems of the WFG [\[39\]te](#page-19-25)st suite. These problems have 500, 1000, 2000 decision variables and 2 and 3 objective dimensions, respectively. For fair comparison, all compared algorithms are implemented in PlatEMO [\[40\]](#page-19-26) and use an AMD Ryzen 7 5800H CPU, 3.20GHZ processor, and 16GB RAM. They are independently run 20 times on each benchmark problem.

A. PERFORMANCE INDICATOR

In the experiments, the inverted generational distance (IGD) [\[41\]](#page-19-27) is used as a performance indicator to evaluate the

comparison algorithms. This indicator can measure the convergence and diversity of the population obtained from the algorithm. The smaller the value of the IGD indicator, the better the obtained population. Defined as follows:

$$
IGD(P^*, P) = \frac{\sum_{x \in P^*} d(x, P)}{|P^*|}
$$
 (10)

where P^* is a set of uniformly distributed reference points on the true PF, P is the obtained solution set, $d(x, P)$ is the minimum Euclidean distance from the point *x* on the Pareto optimal front to the individual in the final solution set P , $|P^*|$ is the number of elements in P^* . The smaller the IGD value, the better the quality of the obtained solution set.

We also use the HV indicator to further evaluate the algorithm's performance. The calculation formula for the HV indicator is as follows:

$$
HV = \bigcup_{i} vol_i \tag{11}
$$

where $i \in PF$, *vol*_{*i*} corresponds to a super region bounded by a pre-specified reference point and a solution *i*. The HV indicator estimates the hypervolume of the area enclosed by a point in the population and a reference point. The HV indicator can measure both the convergence and diversity of the population obtained by the algorithm simultaneously. The larger the HV value, the better the algorithm performs.

B. EXPERIMENTAL SETTINGS

All algorithms are run independently 20 times. The maximum number of objective evaluations (denoted as *FEmax*) allowed is 200000 for test problems with three objectives and 150000 for test problems with two objectives. The initial population size is set to 100. Moreover, the Wilcoxon rank sum test [\[42\]](#page-19-28) is employed to assess whether the performance of the solution set obtained by one of the two comparison algorithms is statistically different from the other. The symbols "+," "-," and "=" indicate that the compared algorithms are significantly better, significantly worse, or statistically associated with LSMOEA-DVS.

In our method, the genetic operators used are simulated binary crossover (SBX) [\[43\]](#page-19-29) and polynomial mutation (PM) [\[44\]. F](#page-19-30)or SBX, the distribution index is set to 20 and the crossover probability is set to 1. For PM, the distribution index is set to 20 and the mutation probability is set to $\frac{1}{D}$, where *D* is the number of decision variables. The number *mu* of excellent individuals selected by preprocessing is set to 10. The number of randomly generated solutions along each direction vector is set to 30.

The parameters of all comparison algorithms are consistent with those given in their papers. For WOF and FDV, NSGA-II is used as the optimization algorithm for embedding.

C. ANALYSIS OF THE EFFECTS OF THE INDIVIDUAL **COMPONENTS**

In this section, a series of experiments are conducted on some DTLZ, LSMOP, and WFG test problems with decision

variable dimensions of 500, 1000, and 2000 to demonstrate the effectiveness of the proposed method.

1) THE EFFECTS OF PREPROCESSING

To evaluate the effectiveness of the preprocessing strategy, we conducted comparative experiments with a variant (LSMOEA-DVS-r). The variant algorithm LSMOEA-DVS-r selects excellent individuals by randomly choosing from the current population. As can be seen from Table [1,](#page-8-0) LSMOEA-DVS performs better on 15 out of the 27 test problems. It only performs worse than LSMOEA-DVS-r on the WFG2 problem. According to the results in Table [1,](#page-8-0) we can demonstrate the effectiveness of the preprocessing strategy. The number of excellent individuals selected by the preprocessing strategy *mu* is crucial for the subsequent sampling strategy. In theory, the larger the *mu* setting is, the more direction vectors are defined, and the more function fitness evaluations are consumed in one iteration. Therefore, setting appropriate values is essential. Figure [10](#page-17-0) shows the IGD values of different *mu* values for LSMOP2, LSMOP5, and LSMOP9 in the case of three objectives. It can be seen from Figure [10](#page-17-0) that setting *mu* to 10 results in better algorithm performance. Therefore, the value of *mu* is set to 10.

2) THE EFFECTS OF CONVERGENCE-RELATED SAMPLING **STRATEGY**

In our method, three points of the boundary point and center point are used as starting points to construct the search direction vector of the three-way for directional sampling. To evaluate the effectiveness of the additional center point, we compare it with a method (LSMOEA-DVS-two) that only takes two boundary points as starting points to construct the direction vector for the two-way search. As can be seen from Table [2,](#page-8-1) LSMOEA-DVS wins 18 out of 27 test problems compared to LSMOEA-DVS-two. On 9 test problems, both of them perform equally well. According to the results in Table [2,](#page-8-1) we can conclude that our proposed method is effective.

3) THE EFFECTS OF DIVERSITY-RELATED SAMPLING STRATEGY

To evaluate the performance of the diversity-related sampling strategy, we compare it with a variant that uses only the convergence-related sampling strategy (LSMOEA-DVS-c). Table [3](#page-9-0) shows the results of the experimental comparison. It can be seen that LSMOEA-DVS achieves better performance on all 27 test problems. This shows that the method significantly enhances the quality of the obtained solution set, thereby proving its effectiveness.

4) THE EFFECTS OF OUT-OF-BOUNDS HANDLING STRATEGY To demonstrate the effectiveness of the out-of-bounds handling strategy, we conducted comparative experiments with a variant (LSMOEA-DVS-i). The variant algorithm first determines the bounds of the search direction and then samples within the bounds, so it does not require out-of-bounds

TABLE 1. Statics of IGD results obtained by LSMOEA-DVS-R and LSMOEA-DVS on 500, 1000, and 2000 dimensional three-objective DTLZ LSMOP WFG some test questions. the best results in each row are highlighted.

Problem	D	LSMOEA-DVS-r	LSMOEA-DVS
DTLZ1	500	$2.0864e-2$ (3.04e-4) -	2.0545e-2 (2.11e-4)
	1000	$2.0899e-2(3.87e-4) =$	2.0966e-2 (1.09e-3)
	2000	2.9448e-2 (3.77e-2) -	2.0613e-2 (2.75e-4)
DTLZ5	500	$1.0320e-2$ (1.49e-3) -	8.5825e-3 (1.18e-3)
	1000	$1.0442e-2(1.85e-3)$ -	8.3597e-3 (1.02e-3)
	2000	$1.0787e-2$ (1.62e-3) -	7.7326e-3 (6.13e-4)
DTLZ7	500	$6.1925e-2(4.30e-3) =$	$6.2074e-2$ (3.30e-3)
	1000	$6.1040e-2(4.07e-3) =$	$6.0136e-2$ (2.05e-3)
	2000	6.0837e-2 $(2.92e-3)$ =	6.1873e-2 (3.21e-3)
LSMOP2	500	4.7229e-2 $(6.26e-4)$ +	4.7783e-2 (7.86e-4)
	1000	$4.3748e-2(4.38e-4) =$	4.4069e-2 (6.20e-4)
	2000	$4.2095e-2(5.39e-4) =$	4.2049e-2 (3.62e-4)
	500	$6.9905e-1$ $(8.10e-3)$ -	6.8951e-1 $(2.40e-3)$
LSMOP6	1000	$6.9940e-1$ (1.96e-2) -	$6.9518e-1$ (4.22e-3)
	2000	$7.0319e-1$ (5.71e-3) -	$6.9653e-1$ (4.89e-3)
LSMOP9	500	$5.8691e-1$ $(2.21e-3)$ -	5.8358e-1 (2.55e-3)
	1000	$5.8396e-1$ $(2.84e-3)$ -	5.7913e-1 (2.17e-3)
	2000	$5.8195e-1$ $(2.59e-3)$ -	5.7728e-1 (1.34e-3)
WFG2	500	$1.8639e-1(8.28e-3) +$	2.0497e-1 (1.38e-2)
	1000	$1.8983e-1(8.48e-3) +$	2.0401e-1 (1.37e-2)
	2000	$1.9089e-1(8.03e-3) +$	2.0834e-1 (1.16e-2)
WFG3	500	$1.5933e-1$ (2.39e-2) -	$1.0030e-1$ (2.31e-2)
	1000	$1.5106e-1$ (2.75e-2) -	$1.0016e-1(2.60e-2)$
	2000	$1.5122e-1$ $(2.31e-2)$ -	$1.0094e-1$ (1.85e-2)
WFG9	500	$2.4629e-1(5.97e-3) =$	2.4541e-1 (3.65e-3)
	1000	$2.4884e-1$ (4.85e-3) =	2.4796e-1 (5.69e-3)
	2000	$2.4833e-1$ (5.47e-3) -	2.4472e-1 (5.68e-3)
$+/-/=$		4/15/8	

handling. As can be seen from Table [4,](#page-9-1) LSMOEA-DVS outperformed LSMOEA-DVS-i on 17 out of the 27 test problems, and performed worse on only two problems. Experimental results prove the effectiveness of the out-ofbounds handling strategy.

D. COMPARISONS WITH STATE-OF-THE-ART LARGE-SCALE MULTI-OBJECTIVE METHODS

In order to evaluate the performance of the LSMOEA-DVS algorithm, we give the experimental results of the LSMOEA-DVS algorithm along with five other state-of-theart algorithms, namely LMOEA-DS, LMOCSO, LSMOF, WOF, and FDV. These algorithms were evaluated on DTLZ1-7, LSMOP1-9, and WFG1-9 test problems. The performance was compared based on the IGD values, with the best results are highlighted.

TABLE 2. Statics of IGD results obtained by LSMOEA-DVS-TWO and LSMOEA-DVS on 500, 1000, and 2000 dimensional two-objective DTLZ LSMOP WFG some test questions. the best results in each row are highlighted.

It can be seen from Table [5](#page-10-0) that in DTLZ1-7 problems, our proposed LSMOEA-DVS algorithm achieved the best results in 36 out of 42 cases. In 42 cases, LSMOEA-DVS outperformed the corresponding competing algorithms in 39, 42, 36, 42, and 41 cases. Our proposed LSMOEA-DVS algorithm performs the best on DTLZ1-4 and DTLZ7 problems, particularly on the bi-objective DTLZ problem, where our algorithm performs the best on all problems. The PF of the DTLZ5 and DTLZ6 test problems is degraded on three objectives. Our proposed algorithms do not achieve optimal results for these problems. LMOEA-DS only has good performance on DTLZ6-7 problems. Due to its complementary environment selection strategy, it has certain advantages in dealing with degenerate and discontinuous PF multi-objective problems like DTLZ6 and DTLZ7. on other remaining problems, the algorithm itself may not converge sufficiently, requiring

TABLE 3. Statics of IGD results obtained by LSMOEA-DVS-C and LSMOEA-DVS on 500, 1000, and 2000 dimensional three-objective DTLZ LSMOP WFG some test questions. the best results in each row are highlighted.

Problem	D	LSMOEA-DVS-c	LSMOEA-DVS
DTLZ1	500	$2.0880e-1$ (9.32e-4) -	2.0545e-2 (2.11e-4)
	1000	$2.0900e-1$ (1.07e-3) -	$2.0966e-2(1.09e-3)$
	2000	$2.1003e-1$ (4.18e-3) -	2.0613e-2 (2.75e-4)
DTLZ5	500	$4.3219e-1$ (1.11e-3) -	$8.5825e-3(1.18e-3)$
	1000	$4.3241e-1$ (7.01e-4) -	8.3597e-3 (1.02e-3)
	2000	$4.3250e-1$ $(3.15e-4)$ -	7.7326e-3 (6.13e-4)
DTLZ7	500	$1.3473e+0$ (3.42e-1) -	$6.2074e-2(3.30e-3)$
	1000	$1.4169e+0$ (2.73e-1) -	$6.0136e-2$ (2.05e-3)
	2000	$1.5381e+0$ (1.23e-4) -	$6.1873e-2(3.21e-3)$
LSMOP2	500	$5.9053e-2(1.04e-3)$ -	4.7783e-2 (7.86e-4)
	1000	$4.7893e-2$ (5.68e-4) -	4.4069e-2 $(6.20e-4)$
	2000	$4.3980e-2$ (4.37e-4) -	4.2049e-2 $(3.62e-4)$
LSMOP6	500	7.9920e-1 (6.97e-3) -	$6.8951e-1(2.40e-3)$
	1000	8.0996e-1 (8.47e-4) -	6.9518e-1 (4.22e-3)
	2000	$8.1140e-1$ (1.41e-3) -	$6.9653e-1$ (4.89e-3)
LSMOP9	500	$1.5364e+0$ (9.62e-4) -	$5.8358e-1$ (2.55e-3)
	1000	$1.5337e+0$ (1.28e-3) -	5.7913e-1 (2.17e-3)
	2000	$1.5315e+0$ (2.16e-3) -	5.7728e-1 (1.34e-3)
WFG2	500	6.2490e-1 $(3.10e-2)$ -	2.0497e-1 (1.38e-2)
	1000	6.2880e-1 $(2.67e-2)$ -	2.0401e-1 (1.37e-2)
	2000	$6.2992e-1$ $(3.59e-2)$ -	2.0834e-1 (1.16e-2)
WFG3	500	4.4572e-1 (4.48e-2) -	$1.0030e-1(2.31e-2)$
	1000	$4.5106e-1$ (5.44e-2) -	$1.0016e-1(2.60e-2)$
	2000	$4.3592e-1$ $(4.66e-2)$ -	1.0094e-1 (1.85e-2)
WFG9	500	$3.0981e-1$ (1.36e-2) -	$2.4541e-1$ (3.65e-3)
	1000	$3.0253e-1$ (1.12e-2) -	$2.4796e-1(5.69e-3)$
	2000	$3.0096e-1$ (1.44e-2) -	2.4472e-1 (5.68e-3)
$+/-/=$		0/27/0	

additional fitness evaluations to improve performance. Compared with other competing algorithms, LSMOF performs the best among the five competing algorithms and has relatively good performance on all DTLZ problems. However, due to the lack of diversity maintenance in the algorithm itself, it did not achieve the best results. Especially in the case of the multi-modal discontinuous PF problem of DTLZ7, LSMOF often only manages to find a portion of the optimal solution. This limitation hinders the algorithm's ability to further enhance its performance. The other three competing algorithms failed to achieve good results.

It can be seen from Table [6](#page-11-0) that in LSMOP1-9, our proposed algorithm achieves the best results in 34 out of 54 cases. In 54 cases, LSMOEA-DVS outperformed the corresponding competing algorithms in 33, 54, 43, 41, and 53 cases.

TABLE 4. Statics of IGD results obtained by LSMOEA-DVS-I and LSMOEA-DVS on 500, 1000, and 2000 dimensional three-objective DTLZ LSMOP WFG some test questions. the best results in each row are highlighted.

LMOCSO and FDV failed to achieve good results on the LSMOP problem, LSMOF and WOF only outperformed our proposed algorithm in 5 out of 54 problems, and mainly focus on the two-objective problem. On the 3-objective problem, the PF of the LSMOP2-4 and LSMOP6-9 test problems is mixed or multimodal, and algorithms that rely on problem transformation, such as LSMOF and WOF, do not perform well. LMOEA-DS is the best among the five competing algorithms. It outperforms LSMOF and WOF algorithms on problems with multimodal or mixed PF. LMOEA-DS outperforms our proposed algorithm in 14 out of 54 problems, especially on the LSMOP8 problem. It can be seen that our proposed algorithm can achieve the best results on LSMOP2, 4, 6, and 9 problems, regardless of whether they have 2 or 3 objectives. As for other LSMOP problems, our algorithm mainly fails to achieve excellent results on two-objective.

TABLE 5. Statistics of IGD results obtained by LSMOEA-DVS and five comparison algorithms on 500, 1000 and 2000 dimensional two- and three-objective DTLZ test problems. the best result in each row are highlighted.

As can be seen from Table [7,](#page-12-0) in WFG1-9, our proposed LSMOEA-DVS algorithm achieves the best results in 33 out

of 54 cases. Among the 54 cases, LSMOEA-DVS outperformed the corresponding competing algorithms in 31, 49,

TABLE 6. Statistics of IGD results obtained by LSMOEA-DVS and five comparison algorithms on 500, 1000 and 2000 dimensional two- and three-objective LSMOP test problems. the best result in each row are highlighted.

40, 46, and 46 cases, respectively. The PF of the WFG4-9 test problem is a concave function on the three objectives. The PF of WFG1 is mixed, that of WFG2 is discontinuous, and that of WFG3 is linear. LSMOF does not achieve

TABLE 7. Statistics of IGD results obtained by LSMOEA-DVS and five comparison algorithms on 500, 1000 and 2000 dimensional two- and three-objective WFG test problems. the best result in each row are highlighted.

the best results on the WFG problem. There are 2 cases in which LMOCSO and WOF achieve the best results. FDV

achieves the best results on the WFG1 problem. LMOEA-DS is the best performer among the 5 competing algorithms, as it

TABLE 8. Statistics of HV results obtained by LSMOEA-DVS and five comparison algorithms on 500, 1000 and 2000 dimensional two- and three-objective DTLZ test problems. the best result in each row are highlighted.

achieves the best results in 11 cases, specifically targeting the WFG4, WFG5, and WFG6 problems. It can be seen that our proposed algorithm has achieved the best results on WFG2, 3, 7, and 9 problems. Especially on two-objective problems,

TABLE 9. Statistics of HV results obtained by LSMOEA-DVS and five comparison algorithms on 500, 1000 and 2000 dimensional two- and three-objective LSMOP test problems. the best result in each row are highlighted.

our proposed algorithm performs well in the majority of cases.

Table [8,](#page-13-0) Table [9,](#page-14-0) and Table [10](#page-15-0) give the comparative experimental results of HV indicators. According to the comparison

TABLE 10. Statistics of HV results obtained by LSMOEA-DVS and five comparison algorithms on 500, 1000 and 2000 dimensional two- and three-objective WFG test problems. the best result in each row are highlighted.

results of HV indicators in Table [8,](#page-13-0) in 42 cases, our proposed algorithm LSMOEA-DVS outperforms the corresponding competing algorithms in 39, 42, 33, 42, and 39 cases. LSMOEA-DVS achieves the best performance on DTLZ2-4 and DTLZ7 problems. Especially for the bi-objective DTLZ problem, LSMOEA-DVS achieved the best performance.

FIGURE 7. The pareto front obtained by each algorithm in 3-objective DTLZ3 with 2000 decision variables. (a) LMOEA-DS; (b) LMOCSO; (c) LSMOF; (d) WOF; (e) FDV; (f) LSMOEA-DVS.

FIGURE 8. The pareto front obtained by each algorithm in 2-objective DTLZ7 with 1000 decision variables. (a) LMOEA-DS; (b) LMOCSO; (c) LSMOF; (d) WOF; (e) FDV; (f) LSMOEA-DVS.

This result is similar to the conclusion reached by the IGD indicator. According to the results in Table [9,](#page-14-0) out of 54 cases, LSMOEA-DVS outperforms the corresponding competing algorithms in 30, 42, 33, 37, and 45 cases. LSMOEA-DVS has the best performance on LSMOP4, 6, 8, and 9 problems. This result is similar to the conclusion reached by the IGD indicator. On certain bi-objective LSMOP problems, all algorithms perform poorly, resulting in an HV indicator score of 0. Therefore, it is impossible to intuitively compare the advantages and disadvantages using the HV indicator on these problems. However, it can be compared based on the IGD index. According to the results in Table [10,](#page-15-0) out of 54 cases, LSMOEA-DVS outperforms the corresponding competitive algorithms in 29, 49, 31, 39, and 43 cases. LSMOEA-DVS performs best on WFG3, 7, and 9 problems, especially for the bi-objective WFG problem. In most cases, LSMOEA-DVS has the best performance. This result is similar to the conclusion reached by the IGD indicator. It can be seen that the algorithm we proposed still outperforms the comparative algorithm in most problems according to the HV indicators. Therefore, the experimental results provide evidence of the efficiency and superiority of

IEEE Access

FIGURE 9. The convergence curves of each algorithm in 3-objective DTLZ1 with 2000 decision variables and 2-objective DTLZ5 with 1000 decision variables. (a) DTLZ1; (b) DTLZ5.

FIGURE 10. Parameter analysis of the mu value of the number of excellent individuals selected for preprocessing. (a) LSMOP2; (b) LSMOP5; (c) LSMOP9.

FIGURE 11. Experiment on time complexity.

LSMOEA-DVS in comparison to the other five competing algorithms.

In order to illustrate the superiority of the algorithm proposed in this paper more intuitively. Figures [7](#page-16-0) and [8](#page-16-1) present the distribution diagrams of the final Pareto front obtained

by the algorithm in this paper, as well as other comparative algorithms, on the DTLZ test problem. In these diagrams, the gray solid circle represents a solution, while the gray grid lines represent the real Pareto front. It can be seen from Figure [7](#page-16-0) that the LSMOEA-DVS algorithm proposed in this paper really converges to the real PF, which is consistent with the conclusion drawn from the IGD value. LSMOF is the best performing algorithm among the comparison algorithms, and it can also converge to the real PF. However, the obtained solution set distribution is not widely spread across the real PF. This also shows that algorithms based on problem transformation, such as LSMOF, have a tendency to converge towards local optima, leading to a limited diversity in the solution set. LMOEA-DS can also converge to the real PF. However, it determines whether the environment selection is based on decomposition or Pareto domination using a threshold, which introduces the possibility of retaining the dominate individual. The solution sets obtained by the remaining algorithms fail to converge to the real PF. It can be seen from Figure [8](#page-16-1) that all algorithms have good performance on 2 objectives. The algorithm LSMOEA-DVS proposed in this paper still has the best performance, and the obtained solution set achieves full coverage on the real PF. Algorithms based on problem transformation, such as LSMOF and WOF, only find half of the true PF. Although LMOEA-DS can find all the real PF, it retains the dominant solution in the discontinuous area, so it cannot handle irregular PF optimization problems. The LMOCSO algorithm also cannot deal with irregular PF optimization problems and still needs more fitness function evaluations to converge to the real PF. FDV can handle irregular PF optimization problems, but it also requires more fitness function evaluations to converge to the real PF.

In order to provide a more intuitive illustration of the convergence speed of the algorithm proposed in this paper, Figure [9](#page-17-1) shows the convergence curves of all algorithms. It can be seen from Figure [9](#page-17-1) that the algorithm proposed in this paper has the fastest convergence speed and can converge to a lower level before 50000 fitness function evaluations. Although LSMOF has the slowest convergence speed, it can achieve better performance among the compared algorithms. In contrast, the remaining algorithms have prematurely converged to local optima, leading to sluggish convergence in the later stages of the algorithm.

E. TIME COMPLEXITY EXPERIMENT AND ANALYSIS

As the number of decision variables increases, the complexity of multi-objective optimization problems also increases. Consequently, it becomes more time-consuming for the algorithm to find the optimal solution. We compare the time complexity of each algorithm by measuring its actual running time on the test problem. Figure [11](#page-17-2) shows the CPU running time of the comparison algorithm for the LSMOP1 problem. The abscissa represents the number of decision variables, while the ordinate represents the CPU running time.

The results show that the time complexity of our proposed algorithm, LSMOEA-DVS, is similar to that of LMOEA-DS and is lower compared to the other algorithms being compared. This shows that the time complexity of LSMOEA-DVS is lower than that of competing algorithms. It is worth noting that as the number of decision variables increases, the running time of the algorithm also increases. This is mainly because the cost of evaluating the function is closely related to the dimension of the decision variables. As the number of decision variables increases, the time complexity will inevitably increase. Since the FDV algorithm utilizes fuzzy search and is heavily influenced by the spatial dimension of decision variables, the time complexity of the FDV algorithm becomes more sensitive to the quantity of decision variables.

V. CONCLUSION

In this paper, a large-scale multi-objective optimization algorithm named LSMOEA-DVS is proposed which is based on direction vector sampling. The algorithm first selects multiple excellent individuals with good convergence and diversity to construct direction vectors by clustering. The directional sampling is guided by the boundary points, center

points, and excellent individuals to construct a three-way search direction vector in order to generate a promising offspring solution. To promote diversity in the next generation population, two individuals are selected from the excellent individuals to construct a direction vector for directional sampling, and then cross-mutate is implemented. At the same time, the reference vector is regenerated to solve the irregular PF optimization problem.

Ablation experiments verify the effectiveness of the two sampling strategies proposed in this paper. Compared with other five large-scale multi-objective evolutionary algorithms, the experimental results show the superiority of the proposed LSMOEA-DVS algorithm. The distribution and convergence curve of the solution set prove that the LSMOEA-DVS algorithm has better convergence and diversity than other algorithms. In future work, we will improve the selection of excellent individuals for constructing direction vectors to further improve the efficiency of directional sampling. At the same time, the proposed LSMOEA-DVS is extended to higher dimensions of objectives and practical problems.

REFERENCES

- [\[1\] C](#page-0-0). A. C. Coello, G. B. Lamont, and D. A. Van Veldhuizen, *Evolutionary Algorithms for Solving Multi-Objective Problems*. New York, NY, USA: Springer, 2007, pp. 79–104.
- [\[2\] X](#page-0-1). 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.
- [\[3\] K](#page-0-2). 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.
- [\[4\] S](#page-0-3). Z. Qasim and M. A. Ismail, ''RODE: Ranking-dominance-based algorithm for many-objective optimization with opposition-based differential evolution,'' *Arabian J. Sci. Eng.*, vol. 45, no. 12, pp. 10079–10096, May 2020.
- [\[5\] L](#page-0-4). 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.
- [\[6\] Y](#page-0-5). Sun, G. G. Yen, and Z. Yi, ''IGD indicator-based evolutionary algorithm for many-objective optimization problems,'' *IEEE Trans. Evol. Comput.*, vol. 23, no. 2, pp. 173–187, Apr. 2019.
- [\[7\] K](#page-0-6). Shang and H. Ishibuchi, ''A new hypervolume-based evolutionary algorithm for many-objective optimization,'' *IEEE Trans. Evol. Comput.*, vol. 24, no. 5, pp. 839–852, Oct. 2020.
- [\[8\] Q](#page-0-7). Zhang and H. Li, "MOEA/D: A multiobjective evolutionary algorithm based on decomposition,'' *IEEE Trans. Evol. Comput.*, vol. 11, no. 6, pp. 712–731, Dec. 2007.
- [\[9\] R](#page-0-8). 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.
- [\[10\]](#page-0-9) X. He, Y. Zhou, Z. Chen, and Q. Zhang, "Evolutionary many-objective optimization based on dynamical decomposition,'' *IEEE Trans. Evol. Comput.*, vol. 23, no. 3, pp. 361–375, Jun. 2019.
- [\[11\]](#page-0-10) I. M. Cooper, M. P. John, R. Lewis, C. L. Mumford, and A. Olden, "Optimising large scale public transport network design problems using mixedmode parallel multi-objective evolutionary algorithms,'' in *Proc. IEEE Congr. Evol. Comput. (CEC)*, Beijing, China, Jul. 2014, pp. 2841–2848.
- [\[12\]](#page-0-11) H. Wang, L. Jiao, R. Shang, S. He, and F. Liu, "A memetic optimization strategy based on dimension reduction in decision space,'' *Evol. Comput.*, vol. 23, no. 1, pp. 69–100, Mar. 2015.
- [\[13\]](#page-0-12) L. Parsons, E. Haque, and H. Liu, "Subspace clustering for high dimensional data: A review,'' *ACM SIGKDD Explor. Newslett.*, vol. 6, no. 1, pp. 90–105, 2004.
- [\[14\]](#page-0-13) 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.
- [\[15\]](#page-0-14) 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.
- [\[16\]](#page-1-0) L. M. Antonio and C. A. C. Coello, "Use of cooperative coevolution for solving large scale multiobjective optimization problems,'' in *Proc. IEEE Congr. Evol. Comput. (CEC)*, Cancun, Mexico, Jun. 2013, pp. 2758–2765.
- [\[17\]](#page-1-1) S. Kukkonen and J. Lampinen, "GDE3: The third evolution step of generalized differential evolution,'' in *Proc. IEEE Congr. Evol. Comput. (CEC)*, Edinburgh, U.K., Sep. 2005, pp. 443–450.
- [\[18\]](#page-1-2) L. M. Antonio and C. A. C. Coello, "Decomposition-based approach for solving large scale multi-objective problems,'' in *Proc. Int. Conf. Parallel Probl. Solving Nat. (PPSNXI)*, 2016, pp. 525–534.
- [\[19\]](#page-1-3) X. Ma, F. Liu, Y. Qi, X. Wang, L. Li, L. Jiao, M. Yin, and M. Gong, ''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.
- [\[20\]](#page-1-4) X. Zhang, Y. Tian, R. Cheng, and Y. Jin, "A decision variable clusteringbased evolutionary algorithm for large-scale many-objective optimization,'' *IEEE Trans. Evol. Comput.*, vol. 22, no. 1, pp. 97–112, Feb. 2018.
- [\[21\]](#page-1-5) M. Li and J. Wei, "A cooperative co-evolutionary algorithm for large-scale multi-objective optimization problems,'' in *Proc. Genet. Evol. Comput. Conf. Companion (GECCO)*, Kyoto, Japan, Jul. 2018, pp. 1716–1721.
- [\[22\]](#page-1-6) H. Chen, X. Zhu, W. Pedrycz, S. Yin, G. Wu, and H. Yan, ''PEA: Parallel evolutionary algorithm by separating convergence and diversity for largescale multi-objective optimization,'' in *Proc. IEEE 38th Int. Conf. Distrib. Comput. Syst. (ICDCS)*, Vienna, Austria, Jul. 2018, pp. 223–232.
- [\[23\]](#page-1-7) W. Du, L. Tong, and Y. Tang, "A framework for high-dimensional robust evolutionary multi-objective optimization,'' in *Proc. Genet. Evol. Comput. Conf. Companion (GECCO)*, Kyoto, Japan, Jul. 2018, pp. 1791–1796.
- [\[24\]](#page-1-8) 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.
- [\[25\]](#page-1-9) 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.
- [\[26\]](#page-1-10) C. He, L. Li, Y. Tian, X. Zhang, R. Cheng, Y. Jin, and X. Yao, "Accelerating large-scale multiobjective optimization via problem reformulation,'' *IEEE Trans. Evol. Comput.*, vol. 23, no. 6, pp. 949–961, Dec. 2019.
- [\[27\]](#page-1-11) 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.
- [\[28\]](#page-1-12) C. He, R. Cheng, and D. Yazdani, "Adaptive offspring generation for evolutionary large-scale multiobjective optimization,'' *IEEE Trans. Syst., Man, Cybern. Syst.*, vol. 52, no. 2, pp. 786–798, Feb. 2022.
- [\[29\]](#page-1-13) S. Qin, C. Sun, Y. Jin, Y. Tan, and J. Fieldsend, "Large-scale evolutionary multiobjective optimization assisted by directed sampling,'' *IEEE Trans. Evol. Comput.*, vol. 25, no. 4, pp. 724–738, Aug. 2021.
- [\[30\]](#page-1-14) X. Yang, J. Zou, S. Yang, J. Zheng, and Y. Liu, ''A fuzzy decision variables framework for large-scale multiobjective optimization,'' *IEEE Trans. Evol. Comput.*, vol. 27, no. 3, pp. 445–459, Jun. 2023.
- [\[31\]](#page-1-15) 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. (CEC)*, Barcelona, Spain, Jul. 2010, pp. 1–8.
- [\[32\]](#page-1-16) S. Van Aelst, X. S. Wang, R. H. Zamar, and R. Zhu, ''Linear grouping using orthogonal regression,'' *Comput. Statist. Data Anal.*, vol. 50, no. 5, pp. 1287–1312, Mar. 2006.
- [\[33\]](#page-1-17) W. Chen, T. Weise, Z. Yang, and K. Tang, "Large-scale global optimization using cooperative coevolution with variable interaction learning,'' in *Parallel Problem Solving From Nature, PPSN XI* (Lecture Notes in Computer Science). Berlin, Germany: Springer, 2010, pp. 300–309.
- [\[34\]](#page-1-18) M. N. Omidvar, M. Yang, Y. Mei, X. Li, and X. Yao, "DG2: A faster and more accurate differential grouping for large-scale black-box optimization,'' *IEEE Trans. Evol. Comput.*, vol. 21, no. 6, pp. 929–942, Dec. 2017.
- [\[35\]](#page-2-1) I. Giagkiozis, R. C. Purshouse, and P. J. Fleming, "An overview of population-based algorithms for multi-objective optimisation,'' *Int. J. Syst. Sci.*, vol. 46, no. 9, pp. 1572–1599, Jul. 2015.
- [\[36\]](#page-2-2) I. Das and J. E. 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, Aug. 1998.
- [\[38\]](#page-6-3) K. Deb, L. Thiele, M. Laumanns, and E. Zitzler, ''Scalable test problems for evolutionary multiobjective optimization,'' in *Evolutionary Multiobjective Optimization*. London, U.K.: Springer, 2005, pp. 105–145.
- [\[39\]](#page-6-4) 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.
- [\[40\]](#page-6-5) 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.
- [\[41\]](#page-6-6) K. Shang, H. Ishibuchi, L. He, and L. M. Pang, "A survey on the hypervolume indicator in evolutionary multiobjective optimization,'' *IEEE Trans. Evol. Comput.*, vol. 25, no. 1, pp. 1–20, Feb. 2021.
- [\[42\]](#page-7-0) W. Haynes, ''Wilcoxon rank sum test,'' in *Encyclopedia of Systems Biology*, W. Dubitzky et al., Eds. New York, NY, USA: Springer, 2013, pp. 2354–2355.
- [\[43\]](#page-7-1) K. Deb and R. B. Agrawal, "Simulated binary crossover for continuous search space,'' *Complex Syst.*, vol. 9, no. 2, pp. 115–148, Sep. 1995.
- [\[44\]](#page-7-2) T. Si, N. D. Jana, and J. Sil, "Particle swarm optimization with adaptive polynomial mutation,'' in *Proc. World Congr. Inf. Commun. Technol. (WICT)*, Mumbai, India, Dec. 2011, pp. 143–147.

YINGJIAN XIONG is currently pursuing the degree with the Faculty of Electrical Engineering and Computer Science, Ningbo University, Zhejiang, China.

His research interests include pattern recognition, intelligent systems, and machine learning.

XUHUA SHI received the Ph.D. degree in control science and engineering from the East China University of Science and Technology, Shanghai, China, in 2011.

She is currently a Professor with the Faculty of Electrical Engineering and Computer Science, Ningbo University, Zhejiang, China. Her research interests include process modeling, optimization and control, intelligent information, and production scheduling.