A Strength Pareto Evolutionary Algorithm Based on Reference Direction for Multi-objective and Many-objective Optimization

Shouyong Jiang and Shengxiang Yang, *Senior Member, IEEE*

***Abstract*—While Pareto-based multi-objective optimization algorithms continue to show effectiveness for a wide range of practical problems that involve mostly two or three objectives, their limited application for many-objective problems, due to the increasing proportion of nondominated solutions and the lack of sufficient selection pressure, has also been gradually recognized. In this paper, we revive an early-developed and computationally expensive strength Pareto based evolutionary algorithm by introducing an efficient reference direction based density estimator, a new fitness assignment scheme, and a new environmental selection strategy, for handling both multi- and many-objective problems. The performance of the proposed algorithm is validated and compared with some state-of-the-art algorithms on a number of test problems. Experimental studies demonstrate that the proposed method shows very competitive performance on both multi- and many-objective problems considered in this study. Besides, our extensive investigations and discussions reveal an interesting finding, that is, diversity-first-and-convergence-second selection strategies may have great potential to deal with many-objective optimization.**

***Index Terms*—Multi-objective optimization, many-objective optimization, strength Pareto evolutionary algorithm, reference direction, computational complexity.**

I. Introduction

MULTIOBJECTIVE evolutionary algorithms (MOEAs), such as nondominated sorting genetic algorithm II (NSGA-II) [10], strength Pareto based evolutionary algorithm 2 (SPEA2) [53], and Pareto archived evolution strategy (PESA) [28], have shown their tremendous potential to handle multi-objective optimization problems (MOPs) with two or three objectives [11]. However, in many real-world applications, optimization problems often involve four or more objectives [23]. Recent studies have suggested that conventional MOEAs are subjected to the scalability challenge, i.e., the performance of these MOEAs degrades dramatically with the increase of the number of objectives. This fact gives rise to a new term, known as many-objective optimization problems (MaOPs), to better refer to those MOPs that have four or more objectives. Note that some communities, such as the multi-criterion decision making (MCDM) [3], do not differentiate between multiobjective and many-objective optimization.

Pareto-based MOEAs employ the (weak) Pareto-dominance relation (denoted as "$\succeq$") [11], a kind of notion that defines a partial order in the objective space, to discriminate individuals in the population. For two individuals x and y, if x is not worse on all objectives and better on at least one objective than y, then the "$\succeq$" relation induces a partial order as $x \succeq y$, which means that individual x dominates y. Despite its great success for dealing with MOPs, the Pareto-dominance relation becomes less discriminating for MaOPs as most solutions become incomparable or nondominated, and for over ten objectives, almost all the solutions are nondominated [23]. For a geometrical interpretation, the reader is referred to [26], [39]. As a consequence, the Pareto-dominance relation becomes of limited use for MaOPs, since it cannot induce sufficient selection pressure toward a set of tradeoff solutions, known as the Pareto-optimal set (POS) in the decision objective or the Pareto-opimal front (POF) in the objective space.

There have been a number of attempts to improve the Pareto-based MOEAs for MaOPs. The first and foremost approach is to modify or develop the definition of the Pareto-dominance relation. In an early attempt, a relaxed version of Pareto-dominance, known as ϵ-dominance, was proposed by Laumanns *et al.* [30] to combine both the convergence and diversity of solutions in a compact form. This modification makes it possible for Pareto-based MOEAs to strengthen selection pressure among solutions and has shown to be very promising for MaOPs [18], [27], [41]. Other studies along this direction, such as cone ϵ-dominance [5], k-optimality [16], preference order ranking [36], fuzzy-dominance [16], [19], θ-dominance [47], and generalized Pareto-optimality [50], have also been shown to provide competitive results.

Another feasible way is to replace the Pareto-dominance relation with an indicator function intended to evaluate the quality of solutions, which is called an indicator-based approach [22]. The hypervolume indicator [54] possesses some nice properties and is often used as the indicator function. Hypervolume-based MOEAs do not require any explicit diversity preservation strategy to maintain population diversity, instead, they promote diversity by the hypervolume indicator itself. The indicator-based evolutionary algorithm (IBEA) [51]

Manuscript received January 14, 2016; revised April 23, 2016; accepted June 27, 2016. This work was supported by the Engineering and Physical Sciences Research Council (EPSRC) of UK under Grant EP/K001310/1 (Corresponding author: Shengxiang Yang).

The authors are with the Centre for Computational Intelligence (CCI), School of Computer Science and Informatics, De Montfort University, Leicester LE1 9BH, U.K. (email: shouyong.jiang@email.dmu.ac.uk, syang@dmu.ac.uk).

This paper has supplementary downloadable material provided by the authors, which is available at http://ieeexplore.org. This supplementary material provides some test problems and genetic operators used in the paper and supplementary experimental results. This material is 205KB in size.

is an early implementation among hypervolume-based MOEAs and can provide good results for MaOPs [41]. However, A potential drawback of hypervolume-based methods is the high computational burden for computing the hypervolume measure on high-dimensional problems, thus reducing of its efficiency for MaOPs. A recent study [4] has reported a new hypervolume-based algorithm, called HypE, which fulfils hypervolume calculations by a Monte Carlo simulation technique. This technique aims to reduce the computational complexity of HypE, thus rendering it competitive for handling MaOPs.

Decomposition-based MOEAs, which convert an MOP to a number of subproblems and simultaneously solve them in a collaborative manner, are another promising method for MaOPs. The MOEA based on decomposition (MOEA/D) [48] is a representative of this class of metaheuristics. MOEA/D employs three possible decomposition functions, i.e., the weighted sum function, the Chebyshev function, and the penalty-based boundary intersection function, to decompose a high-dimensional problem into a set of scalarizing subprolems. It maintains population diversity by a set of evenly-distributed weight vectors. This way, MOEA/D is capable of solving different types of optimization problems with varying degrees of success [17], [24], [25], [33], [48]. Since its introduction, MOEA/D has been regarded as a benchmark for new MOEAs by winning the unconstrained MOEA competition in the 2009 IEEE Congress on Evolutionary Computation (IEEE CEC 2009) [49]. Besides, the decomposition-based idea has also been exploited in some recently-developed MOEAs, e.g., NSGA-III [13], MOEA based on dominance and decomposition (MOEA/DD) [31] and MOEA/D with a distance-based updating strategy [46], to maintain population diversity or control convergence for many-objective optimization.

Another method for MaOPs is to alleviate the loss of selection pressure by enhancing diversity management [2], [32], [42]. In [2], a diversity management operator was introduced to manage the activation/deactivation of diversity promotion on the crowding distance of NSGA-II [10]. Wagner *et al.* [41] reported a significant improvement on the convergence performance of NSGA-II after modifying the assignment of crowding distance values for boundary solutions. Recently, Li *et al.* [32] proposed a shift-based density estimation (SDE) strategy to increase selection pressure for MaOPs. For fitness assignment, SDE takes into account both the distribution and convergence information of solutions, and nondominated solutions with poor convergence are penalized. The empirical study in [32] showed a clear improvement for MOEAs incorporating this strategy.

On the other hand, there has been a large amount of contribution to preference-based approaches and objective reduction. The former attempts to interactively introduce preferences and thus produce tradeoff surfaces in objective subspaces of interest to decision makers. In [14], the interactive use of preferences is implemented by first modelling a strictly monotone value function based on the accepted preference information, and then using the resulting value function to redefine the Pareto-dominance relation, directing the search to more preferred areas. In a recent work, Wang *et al.* [43] proposed to coevolve a family of preferences simultaneously with a population of candidate solutions, which leads to preference-inspired coevolutionary algorithms (PICEAs). Following this idea, they suggested a realization of PICEAs, called PICEA-g, and demonstrated that this method provides highly competitive performance for MaOPs. The latter (i.e., objective reduction) focusses on the reduction of the number of objectives [6], [8], [37], [38], which attempts to circumvent the problems of MaOPs by means of identification and removal of redundant objectives. As a result, the reduced lower-dimensional problems can be solved effectively using existing MOEAs.

Most existing MOEAs adopt a convergence-first-and-diversity-second selection strategy [34] to balance convergence and diversity. This strategy generally works well in multiobjecitve optimization, where the proportion of nondominated solutions in the population is not very high. Despite that, it may fail if the search environments of multi-objective optimization are very complex, which has been observed in the study [34]. In many-objective optimization, the convergence-first-and-diversity-second strategy can be of limited use because the proportion of nondominated solutions is very high and diversity preservation is very likely to be carried out only on nondominated solutions. The population is at the risk of losing diversity and preserved solutions may be far from each other if nondominated solutions are not well distributed. Correspondingly, reproduction operators struggle to generate promising solutions for unexplored regions as distant parents are not very effective to generate good offspring solutions in many-objective optimization [13], [29]. In fact, some dominated but promising solutions can contribute to population diversity, and proper use of them can increase the selection pressure in high-dimensional optimization. In this sense, diversity outweighs convergence and should be emphasized for many-objective optimization. Bearing this in mind, in this paper, we propose a new strength Pareto evolutionary algorithm (SPEA) based on reference direction, denoted SPEA/R, for both multi- and many-objective optimization. SPEA/R is a substantial extension of early-developed prominent SPEA methods [53], [54]. It inherits the advantage of fitness assignment of SPEA2 [53] in quantifying solutions' diversity and convergence in a compact form, but replaces the most time-consuming density estimator by a reference direction based one. Our proposed fitness assignment also takes into account both local and global convergence. More importantly, unlike most MOEAs, we adopts a diversity-first-convergence-second selection strategy, which can soundly balance diversity and convergence. SPEA/R is examined on difficult multi- and many-objective test suites, showing very competitive and even better performance compared with several popular algorithms. Furthermore, we extensively investigate possible reasons for the high performance of SPEA/R and reveal some interesting findings.

The rest of this paper is organized as follows. Section II describes the framework of the proposed algorithm, together with detailed descriptions of its components. Section III presents experimental studies on multi-objective optimization, followed by studies on many-objective optimization described in Section IV. Extensive investigations and discussions regarding the

Algorithm 1 Framework of SPEA/R

Input: N (population size)
Output: approximated Pareto-optimal front
Generate a diverse reference direction set W:
  $W :=$ `Reference_Generation`();
Create an initial parent population P;
while *stopping criterion not met* **do**
  Apply genetic operators on P to generate offspring population $\overline{P}$;
  $Q := P \cup \overline{P}$;
  Normalize objectives of members in Q:
  $\overline{Q} :=$ `Objective_Normalization`(Q);
  for each reference direction $i \in W$ **do**
    Identify members of $\overline{Q}$ close to i:
    $E(i) :=$ `Associate`($\overline{Q}, W, i$);
    Calculate fitness values of members in $E(i)$:
    `Fitness_Assignment`($E(i)$);
  end for
  $P :=$ `Environment_Selection`($\overline{Q}, W$);
end while

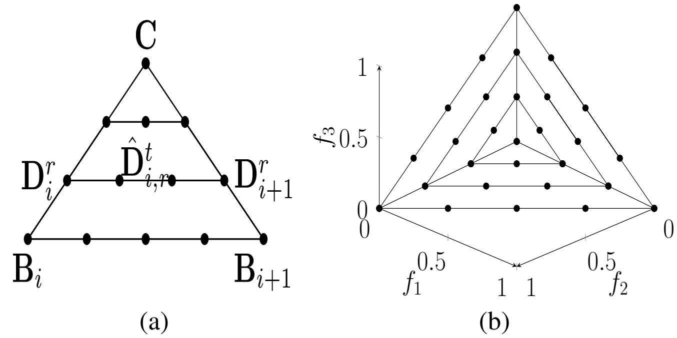


Fig. 1: Intersections of reference directions and a unit simplex: (a) reference directions on the subsimplex $Simp(i)$; (b) reference directions (with 28 directions generated by 3 layers) in three-dimensional space.

proposed algorithm are provided in Section V. Section VI concludes this paper.

II. Proposed SPEA/R Algorithm

The basic framework of the proposed SPEA/R algorithm is presented in Algorithm 1. SPEA/R starts with an initial population and the construction of a predefined set of reference directions, which splits the objective space into a number of independent subregions, helping guide the search toward the whole POF with a good guarantee of population diversity in the objective space. For each generational cycle, on the basis of the preserved parent population, SPEA/R applies genetic operator to reproduce an offspring population, followed by a union of the parent and offspring populations. Then, to make it capable of handling problems with disparately scaled objectives, SPEA/R introduces an objective normalization strategy after the merging of the two populations. Afterwards, each member in the combined population is associated with a reference direction (or a subregion). This way, the combined population members are distributed to different subregions. A novel fitness assignment technique is applied on individuals residing in each subregion. Thereafter, a diversity-first and convergence-second selection strategy is adopted to construct a new parent population for the next generation. In the following subsections, the implementation of each component of SPEA/R will be detailed step by step.

A. Generation of the Reference Direction Set

Any reference-direction-based MOEA cannot ignore the importance of the setting of reference directions (or weight vectors in [48]). Early MOEA/D algorithms employ a systematic approach, developed by Das and Dennis [9], to generate $H = \binom{p+M-1}{M-1}$ reference directions on a unit simplex for M objectives, where p is the number of divisions considered along each objective coordinate. The systematic approach works well for low-dimensional problems, especially for bi-objective problems, where the number of reference directions can be arbitrarily designated. For high-dimensional problems, however, this approach will generate a large amount of reference directions if intermediate reference directions (which require $p \geq M$) within the simplex are pursued [13]. This inevitably pushes up the computational burden of MOEAs. To avoid such a situation, a two-layer (boundary and inside layers) approach for objectives over seven was proposed in [13], [31], which uses the systematic approach to generate two reference direction sets: one set on the boundary layer and the other on the inside layer. Despite that the two-layer approach improves the generation of reference directions, it still produces a large number of reference directions for high-dimensional problems, which will be illustrated later.

To reduce these drawbacks, we present a k-layer reference direction generation approach. Since any reference direction should be sampled from a unit simplex, we can partition the unit simplex into a number of subsimplexes and then generate a set of diverse reference directions for each subsimplex. First, we denote the central reference direction as $\mathbf{C} = (1/M, \cdots, 1/M)$, and the ith extreme reference direction (the intercept on the ith axis) as $\mathbf{B}_i = (b_1, \cdots, b_M)$ where $b_i = 1$ and $b_j = 0$ for all $j \neq i$, $1 \leq j \leq M$, $1 \leq i \leq M$. Thus, the unit simplex can be partitioned into M subsimplexes, each of which (denoted as $Simp(i)$) is bounded by points $\mathbf{C}$, $\mathbf{B}_i$ and $\mathbf{B}_{i+1}$. In the following, we explain how to use our proposed k-layer approach to generate reference directions for the subsimplex $Simp(i)$, and reference directions in other segments can be constructed in the same way.

For the subsimplex $Simp(i)$, we first generate points on sides $\mathbf{CB}_i$ and $\mathbf{CB}_{i+1}$. As illustrated in Fig. 1(a), the rth reference direction (denoted as $\mathbf{D}_i^r$) within the line $\mathbf{CB}_i$ can be calculated as follows:

$$\mathbf{D}_i^r = \mathbf{C} + \frac{r}{k}(\mathbf{B}_i - \mathbf{C}) \tag{1}$$

where $r \in \{1, \cdots, k\}$. This generates k reference directions (actually k layers from vertex $\mathbf{C}$ to the base $\mathbf{B}_i\mathbf{B}_{i+1}$) from the central point to the ith extreme point. After that, we focus on calculating reference directions within the rth layer. Likewise, the tth reference direction within the line $\mathbf{D}_i^r\mathbf{D}_{i+1}^r$ on the rth

Algorithm 2 `Reference_Generation()`

Input: K (number of layers), M (number of objectives), $\overline{N}$ (archive size)
Output: W (reference direction set)
if $M < 3$ **then**
  Use Das and Dennis's method [9] to generate W;
else
  Generate extreme points $\mathbf{B}_i$ for $i = 1, \cdots, M$, and the central point $\mathbf{C}$;
  for $i := 1$ to M **do**
    for $r := 1$ to K **do**
      Calculate all points on the r-th layer by Eq. (2);
    end for
  end for
end if

layer is computed by

$$\hat{\mathbf{D}}_{i,r}^{t} = \mathbf{D}_i^t + \frac{t}{r+1}(\mathbf{D}_{i+1}^r - \mathbf{D}_i^t) \tag{2}$$

where $t \in \{1, \cdots, r\}$. This generates r reference directions for the rth layer of $Simp(i)$. Similarly, diverse reference directions in the rest of $M-1$ subsimplexes can be produced by the above method. At last, the constructed reference direction set W comprises the central reference direction $\mathbf{C}$, and reference directions of each layer in each subsimplex.

It is easy to see that, for k layers, the total number (H_M^k) of reference directions for an M-objective problem is given by

$$H_M^k = \sum_{i=1}^{M}\{\sum_{r=1}^{k} r + k\} + 1 = \frac{Mk(k+3)}{2} + 1 \tag{3}$$

For example, for $M = 3$ and $k = 1$, the reference directions are created on a triangle with vertices at (1,0,0), (0,1,0), and (0,0,1), including three midpoints of the sides of the triangle and an intermediate point at (1/3,1/3,1/3). Fig. 1(b) presents a simple example of reference direction set generated by three layers. In this paper, for bi-objective problems, we use Das and Dennis's systematic approach to predefine a set of uniform reference directions, while for $M > 2$, the k-layer approach is used. The generation of a predefined reference direction set is described in Algorithm 2.

B. Offspring Reproduction and Objective Normalization

Reproduction (line 6 in Algorithm 1) is a step to create a new offspring population to update the parent population $\overline{P}$ (which is actually regarded as the archive). Here, mating selection plays a important role in reproduction. Each parent individual $P_1 \in \overline{P}$ needs a mate $P_2 \in \overline{P}$ to do reproduction. SPEA/R employs a restricted mating scheme to select the mate P_2 for P_1. Specifically, K candidates different from P_1 are randomly chosen from the parent population. Then, the candidate minimizing the Euclidean distance (in objective space) to P_1 can be screened as P_2. $K = 20$ is recommended in this paper based on some preliminary experiments. The restricted mating scheme may help alleviate recombination issues in many-objective optimization, where recombining two distant or very different parents is too disruptive and not likely to generate good children [13], [29].

Algorithm 3 `Objective_Normalization(`Q`)`

Input: Q (combined population)
Output: $\overline{Q}$ (normalized population)
for $i := 1$ to M **do**
  Compute the ideal point $z_{min}^i := \min_{q \in Q} f_i(q)$;
  Compute the worst point $z_{max}^i := \max_{q \in Q} f_i(q)$;
end for
for each member $q \in Q$ **do**
  Computed the normalized objective vector by Eq. (4);
  Save the normalized q to $\overline{Q}$;
end for

After the production of the new offspring population P, SPEA/R then combines it and the parent population to form a population Q (line 7 in Algorithm 1), which is used later to normalize the objectives of individuals (line 8 in Algorithm 1). The normalization procedure is described in Algorithm 3. First, the ideal point $z_{min} = (z_{min}^1, \cdots, z_{min}^M)$ and the worst point $z_{max} = (z_{max}^1, \cdots, z_{max}^M)$ are constructed from the nondominated set of the combined $\overline{P}$ and P, where $z_{min}^i = min(f_i(q))$ and $z_{max}^i = max(f_i(q))$, $q \in Q$, $i = 1, \cdots, M$. Then, the objectives of member q are translated as follows:

$$\hat{f}_i(q) = \frac{f_i(q) - z_{min}^i}{z_{max}^i - z_{min}^i} \tag{4}$$

where $i \in \{1, \cdots, M\}$ and $\hat{f}_i(q)$ denotes the ith normalized objective of member q.

C. Member Association and Fitness Assignment

After mapping the objectives of members of Q into a unit hypercube, next we need to associate each member in the normalized population $\overline{Q}$ with a reference direction (line 10 in Algorithm 1). The member association procedure is presented in Algorithm 4. For each reference direction $w^i \in W$, $i \in \{1, \cdots, H_M^k\}$, we define a subregion, denoted as Ψ^i, in the objective space, as follows:

$$\Psi^i = \{\hat{F}(x) \in \Omega_f | \langle \hat{F}(x), w^i \rangle \leq \langle \hat{F}(x), w^j \rangle\} \tag{5}$$

where $j \in \{1, \cdots, H_M^k\}$, $x \in \Omega_x$, $\hat{F}(x)$ is the normalized objective vector of x, and $\langle \hat{F}(x), w^j \rangle$ is the acute angle between vectors $\hat{F}(x)$ and w^j. Using this definition can easily identify a number of members residing in Ψ^i, denoted as $E(i)$, from the normalized population $\overline{Q}$.

The idea of decomposing the objective space has also been employed in [7], [31], [34]. In both [7] and [34], the objective space decomposition provides a way to approximate a small segment of the POF, while in [31], it is used for local density estimation and diversity maintenance.

The decomposition of objective space can facilitate fitness assignment, as shown in Algorithm 5. In detail, each member

Algorithm 4 $\texttt{Associate}(\overline{Q}, W, i)$

Input: $\overline{Q}$ (combined population), W (reference direction set)
Output: $E(i)$ (individuals in the ith subregion)
for each $q \in \overline{Q}$ **do**
  for each $w \in W$ **do**
    Compute the acute angle $\langle \hat{F}(q), w \rangle$
  end for
  Assign $\hat{w} = w : \text{argmin}_{w \in W} \langle \hat{F}(q), w \rangle$;
  Assign $\theta_q = \langle \hat{F}(q), \hat{w} \rangle$;
  Save q in $E(\hat{w})$
end for

a in $E(i)$ is assigned a "local" [1] strength value $S(a)$, representing the number of solutions it dominates in $E(i)$:

$$S(a) = C(\{a \in E(i) | a \succeq b\}) \tag{6}$$

where $b \in E(i)$ and $C(\cdot)$ denotes the cardinality of a set. The "local" strength value is then used to calculate the "local" raw fitness $R(a)$ of a member a in $E(i)$, as follows:

$$R(a) = \sum_{b \in E(i), b \succeq a} S(b) \tag{7}$$

where the "local" raw fitness depends on the strengths of its dominators in the same subregion. Note that, similar to SPEA2, the fitness is to be minimized here.

In the case where individuals in $E(i)$ do not dominate each other, their raw fitness values will be zero and the above fitness assignment will make no sense. Fig. 2 presents such a situation, where both a and b are in the same subregion and they are nondominated individuals. Intuitively, a is better than b because it is closer to the associated reference direction (y-axis). Thus, individuals' other information should be considered. We adopt an angle-based density estimation technique to discriminate between individuals having identical raw fitness values. Each individual $a \in E(i)$ has a unique angle value $\theta_a = \langle \hat{F}(a), w^i \rangle$, which is actually the acute angle between $\hat{F}(a)$ and the associated reference direction w^i. Then, the density $D(a)$ of individual a is estimated by

$$D(a) = \frac{\theta_a}{\theta_a + \theta_m} \tag{8}$$

where $\theta_m = \underset{1 \leq i \leq H_M^k}{max} \; \underset{j \neq i}{min}(w^i, w^j)$, i.e., the largest acute angle between two neighbouring reference directions, is added to ensure that $D(a)$ is smaller than one. The "local" fitness value of individual a, denoted as $FV_l(a)$, is composed of its raw fitness and density value, combined in the following form:

$$FV_l(a) = R(a) + D(a) \tag{9}$$

This way, individuals with better local diversity and convergence will have higher final fitness. Thus, a is better than b in the case illustrated in Fig. 2.

[1]The term "local" used here is to clarify the difference between our fitness assignment and that in SPEA2 [53].

Algorithm 5 $\texttt{Fitness\_Assignment}(E(i))$

Input: $E(i)$ (individuals in the ith subregion), $\overline{Q}$ (combined population), W (reference direction set)
Output: FV (fitness values of members in $E(i)$)
for each $a \in E(i)$ **do**
  Compute the "local" raw fitness $R(a)$ using Eq. (7);
  Estimate the density value $D(a)$ using Eq. (8);
  Compute the "local" fitness value $FV_l(a) := R(a) + D(a)$;
  Assign the final fitness value $FV(a)$ using Eq. (10);
end for

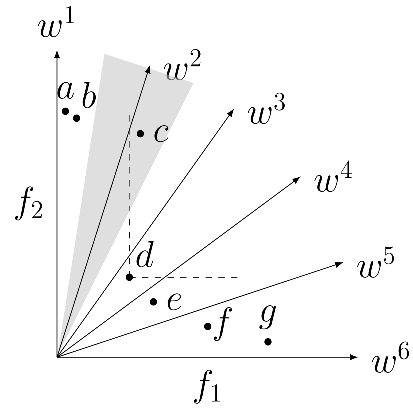


Fig. 2: Influence of decomposed subregions on environmental selection. The grey area represents the subregion occupied by w^2, i.e., Ψ^2, and the dashed lines are used to indicate d dominates c.

Despite great benefit for subregion diversity and local convergence, the "local" fitness assignment may impair global convergence if all individuals in Ψ^i are dominated by individuals in other subregions. To avoid this situation, individual a is also assigned a "global" fitness value, denoted as $FV_g(a)$, which is actually the raw fitness in SPEA2 (see [53]). Besides, if a is the only member in Ψ^i and dominated by individuals in other subregions, it should be given a chance to survive to the next generation. Thus, the final fitness of a, or $FV(a)$, is calculated as:

$$FV(a) = \begin{cases} FV_l(a) & \text{if } |\Psi^i| = 1; \\ FV_l(a) + FV_g(a) & \text{otherwise .} \end{cases} \tag{10}$$

Considering again the example in Fig. 2, individual c is the only member in the associated subregion Ψ^2, but, it is dominated by d in another subregion. This means Ψ^2 might be an underexploited area in the objective space and the search in this area should be enhanced. Conventional Pareto-dominance based techniques, e.g., NASG-II and SPEA2, however, are likely to ignore or even simply abandon important individuals like c in this area. In contrast, the proposed fitness assignment rewards the isolated c at an attempt to attract other individuals toward the underexploited area. This way, the fitness assignment hopefully provides a good approximation to each region of the POF.

Algorithm 6 `Environment_Selection`$(\overline{Q}, W)$

1: **Input**: N (population size), $\overline{Q}$ (combined population), W (reference direction set)
2: **Output**: P (new parent population).
3: Set $P = \emptyset$;
4: **while** $C(P) < N$ **do**
5: Set $H = \emptyset$;
6: **for** each reference direction $i \in W$ **do**
7: **if** $E(i) \neq \emptyset$ **then**
8: Assign $\hat{q} = q : \text{argmin}_{q \in E(i)} FV(q)$;
9: Save $\hat{q}$ in H and remove it from $E(i)$;
10: **end if**
11: **end for**
12: **if** $C(P \cup H) <= N$ **then**
13: $P = P \cup H$;
14: **else**
15: Fill up P with the best $N - C(P)$ individuals in terms of fitness from H;
16: **end if**
17: **end while**

D. Environmental Selection

In the environmental selection (line 13 in Algorithm 1), the best N individuals that can balance diversity and convergence should be preserved. Here, we present a new environmental selection strategy, which is shown in Algorithm 6. The strategy repeatedly selects an array (H) of individuals coming from each subregion, and copy the selected individuals to the new population P if the population size $C(P)$ is not larger than N. Otherwise, $N - C(P)$ individuals from the last considered array are required to exactly fill up the new population. In this situation, the last array is sorted according individuals' fitness and then the best $N - C(P)$ individuals are copied to P.

It should be noted that SPEA/R adopts a diversity-first-and-convergence-second strategy to perform the environmental selection, which is different from most existing MOEAs. SPEA/R repeatedly gives each subregion priority to preserve the most promising individual in the subregion, so individuals in the first loop (lines 6-11 in Algorithm 6) of selection have the highest diversity, and those in the second loop has the second highest diversity, and so on. This way, population diversity can be well maintained. Besides, promising individuals chosen from each subregion often have good fitness, so convergence is also soundly considered. The selection strategy can be further enhanced by elaborating niche count of each subregion when performing convergence selection (line 15 of Algorithm 6) for filling up the population P.

E. Computational Complexity of SPEA/R

The objective normalization (line 8 in Algorithm 1) requires $O(MN)$ computations. In line 10 of Algorithm 1, associating a combined population of $2N$ individuals to H_M^k reference directions takes $O(MNH_M^k)$. Suppose that $L_i = C(E(i))$, the number of individuals in the subregion Ψ^i, then $\sum_{i=1}^{H_M^k} L_i = N$. Thus, fitness assignment for $E(i)$ (line 11 in Algorithm 1) requires $O(ML_i^2)$ operations. For environmental selection, computational resources are mainly consumed by convergence selection. In Algorithm 6, lines 6-11 require $O(H_M^k)$ comparisons and sorting (line 15) spends $O(N \log N)$. In this paper, the population size N depends on H_M^k, as $N \approx H_M^k$. On average, the number of individuals in the ith subregion will be $L_i = 2N/H_M^k \approx 2$. Thus, the average complexity of one generational cycle of SPEA/R is $O(MN^2)$. In the worst case, that is, all the $2N$ individuals get trapped into one subregion and other subregions do not contain any member, the computational complexity reaches $O(MN^2)$, which is the same as the average complexity.

III. Experiments on Multiobjective Optimization

As a starting point, SPEA/R is studied on multi-objective problems. The test problems used here are the MOP [34] test suite, which is a modification of ZDT [52] and DTLZ [15] but more difficult than its predecessors. Since SPEA/R uses a framework similar to MOEA/D-M2M [34], it will be interesting to make a comparison between them. Additionally, we also compared SPEA/R with a subproblem-constrained MOEA/D, i.e., MOEA/D-ACD [44], which is a recently-developed algorithm and has shown great promise for the MOP test problems. These three algorithms[2] employ the recombination operator [35], as suggested in MOEA/D-M2M [34]. For fairness, MOEA/D-ACD uses our reference direction initialization method. The population size was set to 100 (by the systematic approach [9]) and 313 (by our k-layer approach with $k = 13$) for bi- and three-objective problems, respectively. The maximum number of generation was set to 5000 for all the problems, and each algorithm was executed 30 independent runs for each problem. A detailed description of the MOP [34] test suite and the recombination operator [35] is provided in the supplementary material.

A. Performance Metrics

In our experimental studies, we adopt the following widely used performance metrics.

1) Inverted generational distance (IGD) [34]: IGD can provide reliable information on both the diversity and convergence of obtained solutions. Let PF be a set of solutions uniformly sampled from the true POF, and PF^* be the approximated solutions in the objective space, the metric measures the gap between PF^* and PF, calculated as follows:

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

where $d(p, PF^*)$ is the distance between the member p of PF and the nearest member of PF^*. The sizes of the uniformly-sampled PF are 5000 and 5050 for two and three objectives, respectively.

[2]The source codes of MOEA/D-M2M and MOEA/D-ACD are from http://www.cs.cityu.edu.hk/∼qzhang/publications.html.

2) Hypervolume (HV) [54]: The HV metric measures the size of the objective space dominated by the approximated solutions S and bounded by a reference point $R = (R_1, \ldots, R_M)^T$ that is dominated by all points on the POF, and is computed by:

$$HV(S) = Leb(\bigcup_{x \in S} [f_1(x), R_1] \times \cdots \times [f_M(x), R_M]) \quad (12)$$

where $Leb(A)$ is the Lebesgue measure of a set A. In our experiments, R is set to $(2.0, \ldots, 2.0)^T$ unless otherwise stated, and the exact HV metric (calculated by the WFG algorithm[3] [40]) is considered for comparison.

B. Results on Multiobjective Optimization

Table I presents the results of SPEA/R, MOEA/D-M2M, and MOEA/D-ACD, where mean and stand deviation values of IGD and HV are reported and the best value for each problem is marked in boldface. The differences between the approximations are assessed by the Wilcoxon rank-sum test [45] at the 0.05 significance level, with the standard Bonferroni correction [1] to deal with the problem of the higher probability of Type I errors in multiple comparisons.

The MOP [34] test suite contains seven hard-to-converge problems. In this suite, MOP4 is the only disconnected problem, and MOP6 and MOP7 are two three-objective problems. Besides, MOP4 to MOP7 are also diversity-resistant, which may be a big challenge to approximating well-distributed POFs if population diversity is not well maintained. Table I shows that, SPEA/R performs significantly better than MOEA/D-M2M on most of the test problems, in terms of IGD and HV. SPEA/R competes well with MOEA/D-ACD in terms of HV on these problems. Generally, SPEA/R mainly loses on the three-objective MOP6. On another three-objective MOP7, however, SPEA/R wins the comparison by a clear margin.

To have a better understanding of these algorithms' performance, approximated POFs over 30 runs for the seven MOP problems are displayed in Fig. 3. As can be seen from the figure, SPEA/R, MOEA/D-M2M, and MOEA/D-ACD are all able to approximate the POF for the seven problems, but they perform differently in terms of convergence and diversity. Specifically, SPEA/R converges better than the other two algorithms on the first four bi-objective problems. On the two three-objective problems, i.e., MOP6 and MOP7, MOEA/D-M2M cannot achieve uniformly-distributed approximations and misses some boundary regions of the POF. This means that MOEA/D-M2M may not be able to cover the whole POF in higher-dimensional problems. MOEA/D-ACD performs poorly in terms of diversity for MOP7, implying that adding constraints to subproblems is not enough to deal with hard-to-converge and diversity-resistant problems like MOP7. In contrast, SPEA/R maintains diversity well on both MOP6 and MOP7, although it does not fully converge to the POF in some runs.

The experiment on the MOP test suite shows that, SPEA/R and MOEA/D-M2M perform distinctly although both share some similar properties, e.g., decomposition of the objective space. The high performance of SPEA/R may be attributed to its good balance between diversity and convergence, which is achieved by our new proposed fitness assignment and environmental selection.

[3]The latest implementation of WFG can be downloaded from http://www.wfg.csse.uwa.edu.au/hypervolume/.

C. Comparison of Evolution Behaviour with MOEA/D-M2M

Experimental results in the previous subsection have validated the performance of SPEA/R, but it is still not clear why SPEA/R performs better than MOEA/D-M2M on the MOP test suite despite their similar framework. To answer this question, we further compare the evolution behaviour of these two algorithms on MOP2 and MOP3. To be more specific, the obtained approximations of three stages, i.e., the 50th (early stage), 500th (middle stage), and 1000th generation (late stage), are recorded, which are plotted in Fig. 4. It is clear to see from the figure that, SPEA/R maintains good population diversity all the time, whereas MOEA/D-M2M tends to partition population into several subpopulations far away from each other before the late stage, which means diversity between neighbouring subpopulations is poorly controlled. As a consequence, MOEA/D-M2M takes more effort than SPEA/R to search unexplored regions before converging toward the POF and providing a good distribution of population, as illustrated by the 1000th-generation approximation for MOP2. This reason can be also used to explain the poor distribution of MOEA/D-M2M on the three-objective MOP6 and MOP7 in the previous experiment. The figure also indicates that the use of diversity-first-and-convergence-second selection strategy can help SPEA/R to manage diversity and convergence well during the search.

IV. Experiments on Many-objective Optimization

Having had a good start on multi-objective optimization, SPEA/R is now examined on many-objective optimization. The section contributes to making a comparison of SPEA/R with state-of-the-art algorithms on many-objective problems.

A. Test Problems

The test problems used for algorithm comparison come from the WFG toolkit [21]. These problems contains a number of challenging characteristics, i.e., nonseparability, deception, multimodality, biased attributes, and various POF geometries. For each WFG test problem, the number of objectives varies from two to twelve, which considers both multi-objective and many-objective optimization. As recommended by the developers [21], the number of decision variables of all test instances is $n = k + l$, where k and l are the number of position-related variables and distance-related variables, respectively. $k = 2 \times (M - 1)$ and $l = 10$ are used in this paper.

B. Compared Algorithms

Five popular or newly-developed MOEAs are used for comparison in our experimental studies. They are MOEA/D

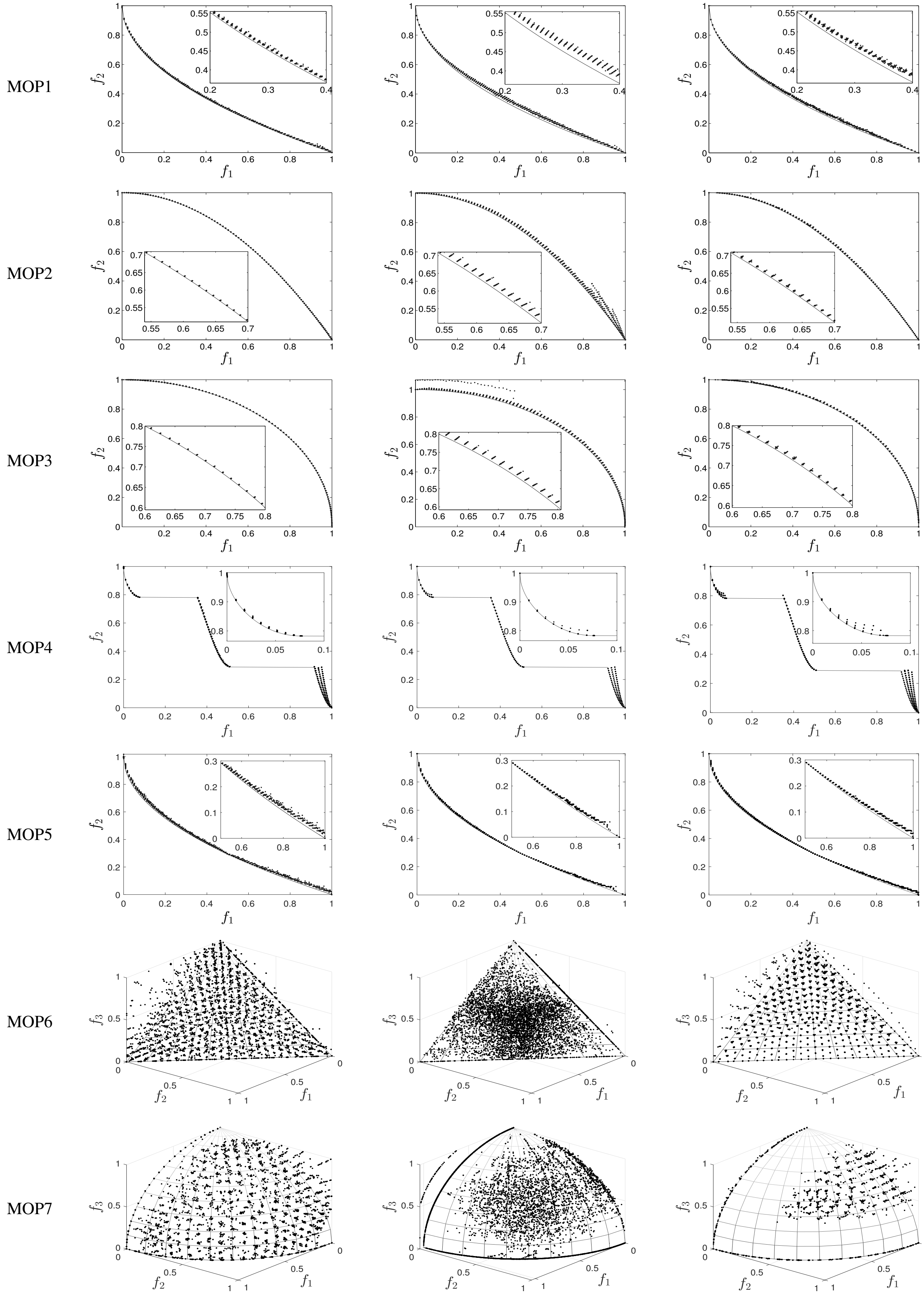


Fig. 3: Approximated POFs for MOP test problems over 30 runs. Left column: SPEA/R; middle column: MOEA/D-M2M; right column: MOEA/D-ACD.

TABLE I: Mean and stand deviation IGD and HV values on MOP problems

	IGD			HV		
Prob.	SPEA/R	MOEA/D-M2M	MOEA/D-ACD	SPEA/R	MOEA/D-M2M	MOEA/D-ACD
MOP1	**8.7805E-3(1.9373E-4)**	9.4133E-3(8.4998E-4)‡	9.0088E-3(1.6739E-4)†	**3.6522E+0(2.8854E-4)**	3.6514E+0(1.0269E-3)‡	3.6520E+0(2.4682E-4)†
MOP2	**4.2374E-3(3.8551E-5)**	8.2719E-3(1.6819E-2)‡	4.4633E-3(5.8251E-5)†	**3.3264E+0(1.0877E-4)**	3.3226E+0(1.8367E-2)‡	3.3207E+0(3.8048E-4)‡
MOP3	**4.8235E-3(1.4936E-4)**	1.0236E-2(1.9945E-2)‡	4.9031E-3(1.5359E-4)†	**3.2101E+0(1.1819E-4)**	3.1825E+0(1.1641E-1)‡	3.2084E+0(1.1626E-3)†
MOP4	**5.8664E-3(1.5107E-3)**	6.5855E-3(1.6268E-3)‡	7.7672E-3(1.5285E-3)‡	**3.5128E+0(2.3030E-3)**	3.5109E+0(2.9883E-3)†	3.5071E+0(3.5365E-3)‡
MOP5	1.2053E-2(7.0952E-4)	9.3834E-3(5.1483E-4)	**8.6467E-3(2.2862E-4)**	3.6457E+0(1.1503E-3)	**3.6502E+0(1.4470E-3)**	3.6414E+0(5.7291E-3)†
MOP6	4.0020E-2(2.6624E-3)	3.8164E-2(1.6047E-3)†	**2.5999E-2(3.5385E-4)**	7.7687E+0(3.9912E-3)	7.7356E+0(1.4549E-2)‡	**7.7956E+0(1.6302E-3)**
MOP7	**5.7604E-2(2.3640E-3)**	8.7838E-2(2.9091E-2)‡	1.0901E-1(3.9980E-3)‡	**7.3919E+0(3.2270E-3)**	7.3659E+0(3.0804E-2)‡	7.3730E+0(1.6802E-2)‡

‡ and † indicate SPEA/R performs significantly better than and equivalently to the corresponding algorithm, respectively.

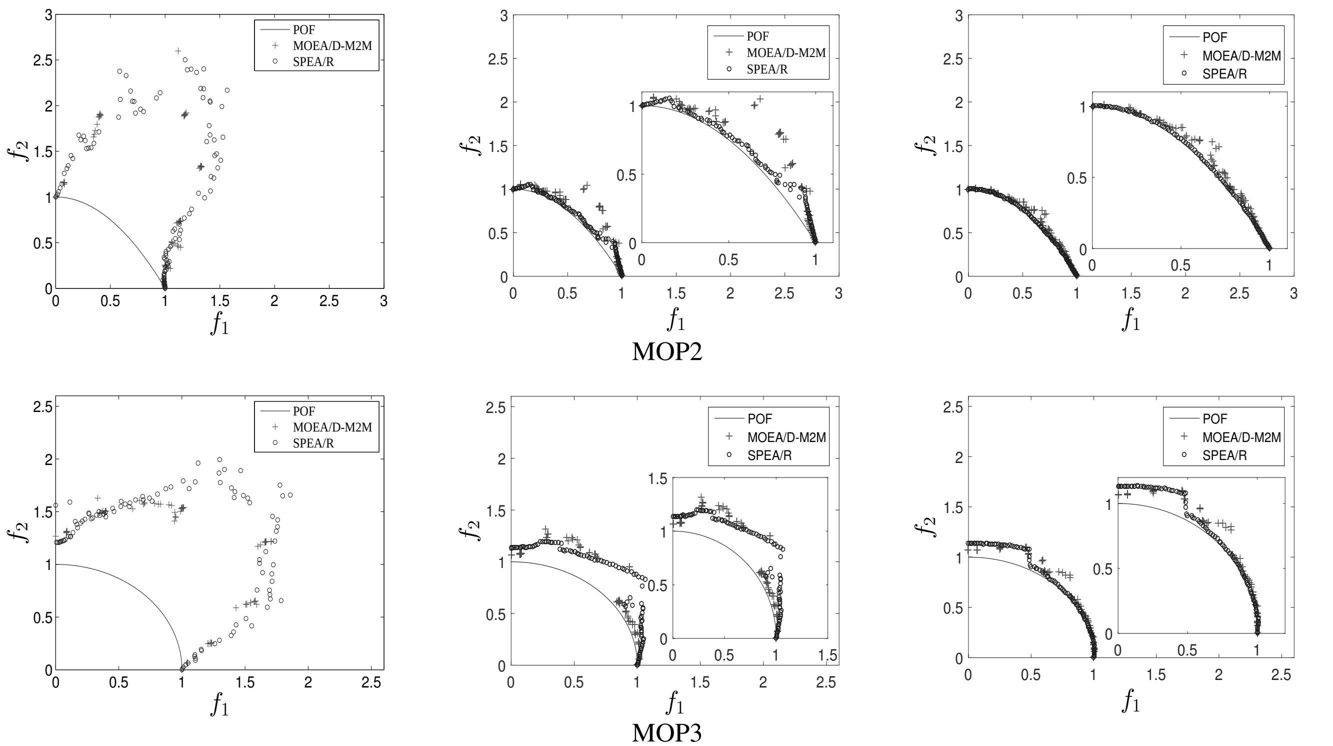


Fig. 4: Evolution behaviour comparison between SPEA/R and MOEA/D-M2M for three stages on MOP2 and MOP3. Left: 50th generation; middle: 500th generation; right: 1000th generation.

[48], HypE [4], SPEA2+SDE [32], PICEA-g [43], and NSGA-III [13], and represent different classes of metaheuristics. A brief description of each compared algorithm is given below.

1) MOEA/D [4] [48]: it is a representative of decomposition-based algorithms. In this paper, PBI is adopted as the aggregation function for MOEA/D because it is empirically proved to be more effective than other decomposition methods for many-objective optimization in a recent study [13], and normalization [48] is used for scaled problems.
2) HypE[5] [4]: it is a representative of indicator-based MOEAs, which employs the hypervolume metric as an indicator in the environmental selection. In HypE, the fitness value of a solution is determined by not only its own hypervolume contribution but also the hypervolume contribution shared with others. Additionally, for the sake of computational complexity, HypE uses Monte Carlo simulation to approximate the exact hypervolume values.
3) SPEA2+SDE[6] [32]: this method introduces a density estimator that considers both the distribution and convergence information of individuals to increase the selection pressure in many-objective optimization. SPEA2+SDE has shown to be very promising for MaOPs [32].
4) PICEA-g[7] [43]: it introduces a new concept of preference-based coevolutionary algorithm (PICEA), which coevolves a family of decision-maker preferences together with a population of candidate solutions, for many-objective optimization. PICEA-g is an implementation of such a concept, where preferences gain higher fitness if it is satisfied by fewer solutions, and solutions gain fitness by meeting as many preferences as possible.

[4]The code of MOEA/D is from http://dces.essex.ac.uk/staff/qzhang/.

[5]The code of HypE is from http://www.tik.ee.ethz.ch/pisa/.

[6]The source code of SPEA2+SDE can be downloaded from http://www.brunel.ac.uk/ cspgmml1/home.html.

[7]The source code of the PICEA-g algorithm can be downloaded from http://www.sheffield.ac.uk/acse/staff/rstu/ruiwang/index.

TABLE II: Population size for different algorithms using the k-layer approach

M	k	H_M^k	MOEA/D	NSGA-III	SPEA/R
2	-	100	100	100	100
3	7	105	105	108	108
5	7	175	175	176	176
8	6	217	217	220	220
12	5	241	241	244	244

5) NSGA-III[8] [13]: it is an upgraded version of the most popular dominance-based NSGA-II algorithm, where a number of supplied reference points is used as a guideline for handling MaOPs. The basic framework of NSGA-III remains similar to NSGA-II, except that it maintains population diversity by niche preservation.

C. Parameter Settings

The parameters of the six MOEAs considered in the experiments are referenced from their original papers. Some key parameters in these algorithms were set as follows:

1) Reproduction parameters: All the algorithms used the simulated binary crossover (SBX) and polynomial mutation [13] as their genetic operators. The crossover probability was $p_c = 1.0$ and its distribution index was $\eta_c = 20$. The mutation probability was $p_m = 1/n$ and its distribution $\eta_m = 20$.
2) Population size: The population sizes (N) of different algorithms are presented in Table II. The population sizes of all the algorithms except MOEA/D were set as $4\lceil H_M^k/4\rceil$, which is the smallest multiple of four not smaller than H_M^k, according to the suggestion in [13]. In other words, HypE, PICEA-g, and SPEA2+SDE use the same population size settings as NSGA-III and SPEA/R.
3) Stopping criterion and the number of executions: Each algorithm was terminated after a pre-specified number of generations. To be specific, for WFG problems, each algorithm stops after 300, 600, 1000, 1500, and 2000 generations for 2-, 3-, 5-, 8-, 12-objective cases, respectively. Additionally, each algorithm was executed 30 independent times on each test instance.

D. Experimental Results and Analysis

The performance measures for quantifying the performance of the compared algorithms in this section are IGD [13] and HV [54]. Note that, the POF points used for computing IGD here are a set of target points on the POF associated with reference directions, as suggested in [13]. For HV computation, the ith objective of the reference point used is $2i+2$ for all the WFG problems, and the HV values presented in this paper are all normalized to $[0,1]$ by dividing $\prod_{i=1}^{M}(2i+2)$. The IGD and HV values of six algorithms on nine WFG test problems are presented in Tables III and IV, respectively.

The WFG1 problem mainly examines whether an MOEA can handle bias and mixed POF shapes. Both IGD and HV metrics indicate that PICEA-g is more suitable for this kind of problem than the other compared algorithms. SPEA/R competes well and even outperforms the others for relatively low-dimensional cases. But, it is defeated by NSGA-III on the 12-objective WFG1 in terms of the HV metric.

WFG2 challenges algorithms' ability to locate all disconnected POF segments and handle nonseparable variable dependencies. For this problem, all the algorithms can achieve impressive performance in low-dimensional cases, and SPEA/R wins by a clear margin. However, when the number of objectives is over five, the performance of MOEA/D and NSGA-III degrades sharply whereas SPEA/R continues to yield good results. SPEA/R wins in the 8-objective case and can compete with HypE and SPEA2+SDE in the 12-objective case, as indicated by both IGD and HV metrics. This means SPEA/R can deal with disconnectivity.

WFG3 features a degenerated and linear POF shape and its variable is nonseparable as well. For this problem, while SPEA/R performs best for the 2-objective case, its performance degrades sharply when the number of redundant objectives increases, which is also the case for the other algorithms except PICEA-g. PCIEA-g is roughly the best performer for this problem because it generates nondominated reference points in the objective space to guide the search in every generation. The other algorithms to a certain extent try to spread population over the whole objective space for the sake of diversity, leading to a very limited number of points on the degenerated POF. Despite that, SPEA/R outperforms MOEA/D and NSGA-III and performs competitively with SPEA2+SDE for the 8- and 12-objective cases. This may be because fitness assignment in SPEA/R favours nondominated solutions.

The problems WFG4 to WFG9 have an identical hyperellipse surface, but they differ in some other characteristics. To be specific, WFG4 introduces multimodality to test algorithms' ability to escape from local optima, and WFG5 is a deceptive problem, and the difficulty lies in the large "aperture" size of the well/basin leading to the global minimum. WFG6 has a significant nonseparable reduction, and WFG7-9 all introduce some bias to challenge algorithms' diversity, but WFG8-9 are nonseparable. Also, variable linkages in WFG8 are much more difficult than that in WFG9.

For WFG4-WFG9, SPEA/R wins nearly all the tested cases in terms of IGD and HV, showing high ability to deal with a number of considered characteristics in these problems. Considering the HV metric, PICEA-g also achieves very competitive results on these problems, and outperforms or compares well with SPEA/R in some cases. However, none of the other algorithms can compete with SPEA/R.

The above experimental studies show that the tested algorithms' performance can be influenced by at least two factors, i.e., problem characteristics and the number of objectives. Clearly, degeneration in WFG3 poses a big challenge to reference-based algorithms, i.e., MOEA/D, NSGA-III, and SPEA/R, as they roughly pursue diversified population over the whole objective space. On the other hand, an increase in the number of objectives to some extent influences all the tested algorithms. MOEA/D is the most influenced one among six algorithms, which experiences a sharp drop when the num-

[8]The source code of NSGA-III (version 1.1) can be downloaded from http://web.ntnu.edu.tw/~tcchiang/publications/nsga3cpp/nsga3cpp.htm.

TABLE III: Mean and standard deviation IGD values obtained by six algorithms for WFG problems

Prob.	M	HypE	PICEA-g	MOEA/D	NSGA-III	SPEA2+SDE	SPEA/R
WFG1	2	8.8317E-1(3.0615E-2)†	8.4573E-1(7.7703E-2)†	1.0781E+0(4.1187E-2)‡	1.0260E+0(5.5762E-2)‡	**8.2124E-1(3.9450E-2)**	9.0085E-1(2.0778E-2)
	3	1.2748E+0(4.3651E-2)‡	**7.8534E-1(8.9205E-2)**	1.2041E+0(3.1711E-2)†	1.3784E+0(4.0202E-2)‡	1.2648E+0(2.3129E-2)‡	1.1794E+0(2.4394E-2)
	5	1.8676E+0(6.2653E-2)‡	**5.3311E-1(1.3060E-1)**	1.4877E+0(3.0123E-2)†	2.0084E+0(1.0479E-1)‡	1.8980E+0(1.6956E-2)‡	1.4780E+0(3.5911E-2)
	8	2.5670E+0(6.2118E-2)†	**1.4327E+0(1.4821E-1)**†	2.7002E+0(3.2555E-1)‡	2.7334E+0(9.6093E-2)‡	2.7327E+0(6.3839E-2)‡	1.5879E+0(6.9790E-2)
	12	3.6640E+0(4.1488E-1)	**3.2660E+0(6.1898E-1)**	4.9319E+0(7.5027E-1)†	4.8169E+0(2.8691E-1)†	3.6292E+0(8.4763E-2)	4.7653E+0(7.7595E-1)
WFG2	2	1.3134E-1(5.5902E-2)‡	8.6982E-2(6.4363E-2)‡	6.5756E-1(2.4753E-1)‡	9.8399E-2(7.0368E-2)‡	8.6488E-2(6.8054E-2)‡	**5.3339E-2(5.9669E-2)**
	3	3.9965E-1(9.4628E-2)‡	3.0184E-1(1.3370E-1)‡	1.6004E+0(5.1826E-1)‡	2.7108E-1(1.9464E-1)‡	2.9113E-1(1.3529E-1)‡	**1.9273E-1(1.4337E-1)**
	5	1.0282E+0(3.1081E-1)‡	6.9710E-1(3.7222E-1)‡	3.4876E+0(9.7999E-1)‡	6.1589E-1(5.8567E-1)‡	5.2034E-1(1.8118E-1)‡	**4.1324E-1(1.9367E-1)**
	8	1.6356E+0(2.0123E-1)‡	1.0825E+0(3.2538E-1)‡	6.9688E+0(3.6616E+0)‡	2.5990E+0(5.5451E-1)‡	1.0398E+0(8.0346E-2)†	**9.2918E-1(1.5704E-2)**
	12	3.6569E+0(1.0009E+0)‡	1.9909E+0(1.6166E-1)†	1.4420E+1(7.0122E+0)‡	5.3146E+0(2.1231E+0)‡	2.0416E+0(2.9444E-1)†	**1.9406E+0(4.8101E-1)**
WFG3	2	6.1055E-2(2.1969E-2)‡	1.5970E-2(1.0060E-3)‡	2.1120E-2(1.0954E-2)‡	1.7672E-2(3.9260E-3)‡	1.4922E-2(1.3055E-3)‡	**9.5981E-3(1.4972E-3)**
	3	3.6647E-1(7.0170E-2)‡	1.2092E-1(8.5003E-3)†	**8.4610E-2(2.0841E-2)**	1.1033E-1(1.9565E-2)‡	1.0415E-1(9.6999E-3)	1.2706E-1(1.9111E-2)
	5	8.7070E-1(3.4817E-1)‡	4.2029E-1(2.6257E-2)†	**2.1410E-1(9.2308E-3)**	3.9046E-1(5.6144E-2)	6.8446E-1(7.4234E-2)‡	4.7455E-1(5.0952E-2)
	8	1.2895E+0(3.8251E-1)	**1.0091E+0(1.1303E-1)**	8.6567E+0(2.5278E-2)‡	2.4852E+0(1.3923E+0)‡	2.3216E+0(2.1709E-1)‡	1.8096E+0(7.9589E-1)
	12	2.1678E+0(6.4471E-1)†	**1.5164E+0(4.9334E-1)**	1.3202E+1(5.0194E-2)‡	6.1449E+0(1.0145E+0)‡	4.3313E+0(2.2150E-1)‡	2.2071E+0(8.9759E-1)
WFG4	2	3.1877E-2(6.5145E-3)‡	1.6304E-2(1.2351E-3)‡	2.7940E-2(6.4054E-3)‡	1.2596E-2(3.0353E-3)‡	3.0674E-2(6.3801E-3)‡	**4.0642E-3(6.3104E-4)**
	3	5.6412E-1(8.5813E-2)‡	2.0043E-1(7.9213E-3)‡	6.9098E-2(1.0745E-2)‡	6.3007E-2(5.4682E-3)‡	2.9956E-1(1.5741E-2)‡	**2.8864E-2(2.1139E-3)**
	5	2.0444E+0(2.2951E-1)‡	1.1045E+0(4.7317E-1)‡	1.4182E-1(1.4965E-2)‡	3.6864E-1(4.4117E-1)‡	1.2366E+0(7.4328E-2)‡	**1.0932E-1(7.3118E-3)**
	8	6.0523E+0(1.5295E+0)‡	6.6026E+0(7.9339E-1)‡	1.4683E+1(1.1304E+0)‡	2.3447E+0(8.3237E-1)‡	3.7387E+0(2.1384E-1)‡	**3.0963E-1(4.4228E-2)**
	12	1.1022E+1(1.4125E+0)‡	1.4136E+1(9.6910E-1)‡	2.4085E+1(3.8119E-7)‡	9.3126E+0(9.3937E-1)‡	7.8294E+0(2.7645E-1)‡	**5.545E-1(5.8877E-2)**
WFG5	2	1.4442E-1(2.9026E-2)‡	6.9351E-2(1.8443E-3)†	7.2942E-2(1.6355E-3)†	6.9201E-2(1.9862E-3)‡	8.1770E-2(4.0625E-3)‡	**6.8656E-2(6.0334E-4)**
	3	7.9362E-1(1.4324E-1)‡	2.1530E-1(6.1536E-3)‡	1.0567E-1(4.0651E-3)†	2.1164E-1(1.2089E-2)‡	2.8447E-1(1.1878E-2)‡	**1.0007E-1(2.7774E-3)**
	5	2.3647E+0(5.0181E-1)‡	9.3128E-1(1.7072E-2)‡	1.7300E-1(2.0329E-2)‡	3.6011E-1(2.7441E-2)‡	1.1152E+0(5.7730E-2)‡	**1.5221E-1(3.3155E-3)**
	8	5.0708E+0(9.3870E-1)‡	3.6695E+0(6.6358E-1)‡	1.4665E+1(1.6616E-1)‡	1.0424E+0(1.2969E+0)‡	3.0894E+0(1.6408E-1)‡	**2.9294E-1(7.8991E-3)**
	12	1.1590E+1(3.7303E+0)‡	1.1189E+1(6.7757E-1)‡	2.3809E+1(3.6910E-2)‡	1.1152E+1(1.7458E+0)‡	6.8783E+0(3.5983E-1)‡	**5.9182E-1(5.6574E-2)**
WFG6	2	9.5505E-2(2.7439E-2)‡	8.7771E-2(1.6443E-2)†	1.2231E-1(2.9457E-2)‡	**6.2542E-2(8.0865E-3)**	8.3074E-2(2.2805E-2)†	8.2235E-2(1.7636E-2)
	3	5.0657E-1(5.9845E-2)‡	2.2653E-1(1.1940E-2)‡	1.6990E-1(4.3933E-2)‡	1.3889E-1(1.6059E-2)†	3.0489E-1(2.5190E-2)‡	**1.2494E-1(1.8689E-2)**
	5	1.7161E+0(1.7421E-1)‡	9.3252E-1(2.8199E-2)‡	2.4944E-1(5.9006E-2)‡	2.5876E-1(1.8224E-2)‡	1.1083E+0(4.5895E-2)‡	**1.9652E-1(2.1632E-2)**
	8	3.5742E+0(2.2169E-1)‡	2.4527E+0(1.0243E-1)‡	1.3441E+1(3.4359E+0)‡	3.5306E-1(3.9311E-2)‡	2.8712E+0(1.5114E-1)‡	**3.1442E-1(4.5861E-2)**
	12	8.5555E+0(1.2719E+0)‡	8.3868E+0(1.5120E+0)‡	2.4086E+1(1.0499E-3)‡	1.2026E+0(1.4127E+0)‡	6.5637E+0(3.4155E-1)‡	**5.6185E-1(6.4777E-2)**
WFG7	2	8.2165E-2(2.8926E-2)‡	1.5910E-2(6.6738E-4)‡	2.3036E-2(5.5223E-3)‡	6.1770E-3(1.2478E-3)‡	2.9225E-2(4.7514E-3)‡	**3.0428E-3(6.4407E-4)**
	3	6.6418E-1(9.6773E-2)‡	1.9958E-1(6.0806E-3)‡	1.0757E-1(6.5050E-2)‡	4.5169E-2(5.3488E-3)‡	2.6732E-1(1.8452E-2)‡	**1.7752E-2(2.2178E-3)**
	5	2.1347E+0(2.2676E-1)‡	9.3521E-1(2.3560E-2)‡	1.3512E-1(2.0759E-2)‡	1.9524E-1(3.9223E-2)‡	1.2314E+0(7.5941E-2)‡	**7.6146E-2(6.0145E-3)**
	8	5.7576E+0(1.2137E+0)‡	4.7858E+0(1.3194E+0)‡	4.1326E+0(4.8119E+0)‡	1.9059E+0(5.2245E-1)‡	3.3239E+0(2.3293E-1)‡	**4.4555E-2(1.4906E-3)**
	12	1.3634E+1(2.8675E+0)‡	1.1684E+1(1.1888E+0)‡	1.8854E+1(7.2573E+0)‡	9.0098E+0(1.1672E+0)‡	7.2319E+0(2.0432E-1)‡	**1.2239E+0(1.7913E-1)**
WFG8	2	1.1761E-1(1.3492E-2)‡	1.7830E-1(8.4903E-3)‡	1.9949E-1(7.7770E-2)‡	1.0440E-1(4.4972E-3)‡	9.6227E-2(7.0488E-3)‡	**6.5348E-2(8.3518E-3)**
	3	6.4260E-1(9.6987E-2)‡	3.6302E-1(6.6066E-3)‡	2.9721E-1(1.8361E-2)‡	2.6280E-1(1.1173E-2)‡	3.9859E-1(1.2363E-2)‡	**1.8791E-1(1.2203E-2)**
	5	3.0822E+0(3.6658E-1)‡	1.1359E+0(1.5299E-1)‡	6.2743E-1(2.4932E-2)‡	6.2983E-1(3.4558E-2)‡	1.3975E+0(7.1725E-2)‡	**4.4959E-1(6.3168E-2)**
	8	6.6256E+0(6.9786E-1)‡	4.8593E+0(6.4988E-1)‡	1.4786E+1(5.2920E-1)‡	3.9291E+0(8.5793E-1)‡	3.6385E+0(1.3014E-1)‡	**7.7559E-1(1.7947E-1)**
	12	1.2621E+1(1.3872E+0)‡	1.2125E+1(8.1783E-1)‡	2.4081E+1(1.6969E-2)‡	9.5723E+0(7.4981E-1)‡	7.5282E+0(2.7831E-1)‡	**25918E+0(1.5246E+0)**
WFG9	2	1.1042E-1(2.1126E-1)‡	2.5073E-2(2.0117E-3)‡	1.1382E-1(9.2790E-2)‡	4.1288E-2(3.6557E-2)‡	4.6615E-2(4.5168E-2)‡	**2.3333E-2(1.8028E-3)**
	3	8.5309E-1(3.9767E-1)‡	2.0280E-1(5.3020E-3)‡	3.6735E-1(8.0129E-2)‡	2.1455E-1(2.8924E-2)‡	2.8436E-1(2.9944E-2)‡	**1.3565E-1(6.0332E-2)**
	5	2.2225E+0(5.9275E-1)‡	9.0468E-1(2.3965E-2)‡	5.6825E-1(5.1271E-2)‡	4.7457E-1(2.0885E-2)†	1.2446E+0(1.0233E-1)‡	**4.4527E-1(1.0612E-1)**
	8	5.6313E+0(2.0539E+0)‡	2.3506E+0(1.8201E-1)‡	1.4075E+1(3.1523E+0)‡	1.4547E+0(7.5551E-1)‡	3.2874E+0(2.1696E-1)‡	**1.0596E+0(2.1459E-1)**
	12	1.1425E+1(4.1445E+0)‡	9.0759E+0(1.1149E+0)‡	2.3886E+1(1.0426E-1)‡	7.0250E+0(2.2799E+0)‡	6.8976E+0(3.8740E-1)‡	**1.9587E+0(4.1090E-1)**

‡ and † indicate SPEA/R performs significantly better than and equivalently to the corresponding algorithm, respectively.

ber of objectives increases from five to twelve, as indicated by the deterioration of IGD and HV. This is consistent with some recent studies [31], [47]. This observation shows MOEA/D struggles to solve difficult many-objective WFG problems.

To understand why SPEA/R generally performs better than the other algorithms, we graphically plot the parallel coordinates (normalized by the nadir point) of final solutions obtained by each algorithm for the 12-objective WFG4 in Fig. 5. For the inspection of parallel coordinates for several other WFG instances, the interested readers are referred to the supplementary material. The figure clearly shows that SPEA/R is able to obtain a good spread of solutions in the entire range of the POF ($f_i \in [0, 2i]$, for all i), whereas HypE, PICEA-g, MOEA/D, and NSGA-III miss some part of the POF. Due to effective density estimation, SPEA2+SDE shows very competitive diversity performance, but it does not cover well the entire POF. Thus, we can conclude that the outperformance of SPEA/R over the other algorithms results largely from its sound diversity maintenance and its effective fitness assignment capable of driving population toward the POF.

V. Investigations and Discussions

A. Comparison of Different Reference Direction Generation Approaches

As the population size ($Popsize$) of MOEAs is closely associated with the amount of reference directions, we compare our proposed k-layer approach with the systematic approach used in MOEA/D [48] and the two-layer approach used in NSGA-III [13] in terms of required $Popsize$ for different numbers of objectives. Since the two-layer approach is an improved version of the systematic approach for generating reference directions in the case of 7 or more objectives, we just need to compare our k-layer approach with the former and the latter in low-dimensional cases and high-dimensional cases, respectively.

TABLE IV: Mean and standard deviation HV values obtained by six algorithms for WFG problems

Prob.	M	HypE	PICEA-g	MOEA/D	NSGA-III	SPEA2+SDE	SPEA/R
WFG1	2	6.1349E-1(8.4520E-3)†	6.2383E-1(2.1342E-2)	5.2963E-1(2.6156E-2)‡	5.2796E-1(2.1870E-2)‡	**6.3186E-1(1.1067E-2)**	6.0736E-1(5.9395E-3)
	3	5.9929E-1(8.2784E-3)‡	**7.3526E-1(2.6053E-2)**	5.8131E-1(2.7399E-2)‡	5.4419E-1(2.4954E-2)‡	6.0147E-1(6.0920E-3)‡	6.2420E-1(6.0584E-3)
	5	5.2153E-1(7.6354E-3)‡	**9.1200E-1(5.4121E-2)**	5.7746E-1(7.5357E-3)‡	5.4737E-1(1.7873E-2)‡	5.0145E-1(3.6774E-3)‡	5.7919E-1(5.4070E-3)
	8	4.3738E-1(1.3688E-2)‡	**9.3707E-1(2.1198E-2)**	4.3174E-1(4.4287E-2)‡	4.8199E-1(1.7796E-2)‡	4.2294E-1(2.3498E-3)‡	6.2128E-1(7.8831E-2)
	12	3.7561E-1(4.6241E-3)‡	**9.2162E-1(2.9084E-2)**	3.2474E-1(7.3978E-2)‡	7.4386E-1(4.0646E-2)	3.6416E-1(3.7633E-3)‡	4.2884E-1(5.9670E-2)
WFG2	2	7.9746E-1(3.0072E-2)‡	8.1726E-1(3.3376E-2)‡	6.5750E-1(5.0778E-2)‡	8.0801E-1(3.4064E-2)‡	8.1659E-1(3.4444E-2)‡	**8.3092E-1(2.9447E-2)**
	3	9.2195E-1(5.6280E-2)‡	9.0282E-1(6.9396E-2)‡	6.4363E-1(7.0612E-2)‡	9.1538E-1(6.8576E-2)‡	9.0856E-1(6.6674E-2)‡	**9.4106E-1(6.0781E-2)**
	5	9.2808E-1(7.4718E-2)‡	9.4948E-1(7.9765E-2)‡	6.2190E-1(8.7394E-2)‡	9.6131E-1(6.1941E-2)‡	9.7887E-1(3.5391E-2)‡	**9.8674E-1(3.7587E-2)**
	8	9.7630E-1(3.8039E-3)‡	9.8194E-1(5.5797E-2)‡	5.5675E-1(2.5975E-1)‡	9.1862E-1(1.0630E-1)‡	9.8544E-1(2.1334E-3)‡	**9.9759E-1(6.6128E-4)**
	12	9.6801E-1(1.1392E-2)‡	**9.9860E-1(8.5921E-4)**	3.3908E-1(3.1366E-1)‡	7.9068E-1(1.6581E-1)‡	9.7314E-1(3.9702E-2)‡	9.8832E-1(9.2631E-2)
WFG3	2	8.1688E-1(4.4403E-3)‡	8.2899E-1(4.8416E-4)†	8.2486E-1(4.3417E-3)‡	8.2634E-1(1.1226E-3)‡	8.2739E-1(1.0559E-3)‡	**8.2945E-1(1.3184E-3)**
	3	7.5060E-1(9.0712E-3)‡	**8.9256E-1(7.5315E-4)**	7.9356E-1(5.8079E-3)	7.8197E-1(4.1079E-3)†	7.8383E-1(4.3288E-3)‡	7.7947E-1(8.0418E-3)
	5	6.6263E-1(2.6816E-2)‡	**9.5989E-1(1.2055E-3)**	7.3844E-1(5.3374E-3)	7.1730E-1(7.5465E-3)†	6.8788E-1(1.3525E-2)†	6.8696E-1(1.1909E-2)
	8	6.0613E-1(2.3359E-2)	**9.8966E-1(5.0622E-4)**	1.2210E-1(1.2553E-3)‡	4.6929E-1(1.0987E-1)‡	4.8493E-1(3.3216E-2)†	4.9757E-1(4.4357E-2)
	12	5.6477E-1(2.7196E-2)	**9.5252E-1(6.7478E-2)**	8.6302E-2(1.5356E-3)‡	3.3546E-1(3.7684E-2)‡	3.2742E-1(6.1689E-3)‡	4.8913E-1(4.4723E-2)
WFG4	2	7.2511E-1(6.1498E-3)‡	7.3428E-1(5.4716E-4)†	7.2972E-1(1.3665E-3)‡	7.3314E-1(9.9494E-4)‡	7.1096E-1(1.1254E-2)‡	**7.3497E-1(6.0901E-4)**
	3	8.2289E-1(2.1555E-2)‡	8.4966E-1(1.1654E-3)†	8.3709E-1(2.2772E-3)‡	8.3960E-1(1.2736E-3)‡	7.6916E-1(1.1329E-2)‡	**8.5842E-1(9.0221E-4)**
	5	7.4786E-1(3.9699E-2)‡	9.1105E-1(7.2895E-2)‡	9.1690E-1(1.2970E-3)‡	8.7615E-1(4.9747E-2)‡	7.5415E-1(1.2288E-2)‡	**9.2169E-1(1.2998E-3)**
	8	5.8468E-1(6.8400E-2)‡	7.0094E-1(7.4721E-2)‡	1.3762E-1(8.1625E-2)‡	8.0583E-1(4.5507E-2)‡	5.8432E-1(2.5894E-2)‡	**9.5797E-1(3.9059E-3)**
	12	5.3389E-1(5.8040E-2)‡	6.6260E-1(7.2208E-2)‡	7.6923E-2(5.9845E-9)‡	7.1332E-1(3.0720E-2)‡	5.3058E-1(3.7951E-2)‡	**9.1609E-1(7.8544E-2)**
WFG5	2	6.8762E-1(4.8597E-3)‡	7.0799E-1(2.7043E-3)‡	7.0280E-1(1.5455E-3)†	7.0558E-1(3.3458E-3)†	6.8987E-1(7.5992E-3)‡	**7.1227E-1(9.3626E-4)**
	3	7.6888E-1(2.2719E-2)‡	8.2305E-1(1.5867E-3)†	8.1368E-1(2.3212E-3)†	8.0667E-1(2.8385E-3)‡	7.7206E-1(7.9184E-3)‡	**8.2437E-1(2.0573E-3)**
	5	7.4369E-1(5.0270E-2)‡	**9.0566E-1(1.5576E-3)**	8.7778E-1(2.3524E-3)†	8.4922E-1(3.6613E-3)‡	7.8108E-1(1.0898E-2)‡	8.8792E-1(1.1292E-3)
	8	6.0049E-1(7.9653E-2)‡	8.4677E-1(4.8498E-2)‡	1.0334E-1(5.3172E-3)‡	8.5495E-1(6.6536E-2)‡	6.9019E-1(1.6596E-2)‡	**9.0023E-1(2.9488E-3)**
	12	4.3256E-1(6.7488E-2)‡	7.3490E-1(4.4608E-2)‡	6.8603E-2(8.1182E-4)‡	6.3626E-1(5.0428E-2)‡	5.9168E-1(2.0913E-2)‡	**9.0236E-1(6.6030E-3)**
WFG6	2	6.9937E-1(1.1071E-2)†	7.0821E-1(5.5304E-3)†	6.9412E-1(1.0969E-2)‡	**7.1557E-1(3.0357E-3)**	6.9758E-1(1.5288E-2)†	7.1001E-1(6.5120E-3)
	3	7.8656E-1(1.1842E-2)‡	8.1866E-1(6.8386E-3)‡	8.0201E-1(1.2884E-2)‡	8.1457E-1(4.2668E-3)‡	7.7216E-1(1.2142E-2)‡	**8.2552E-1(6.3065E-3)**
	5	8.0206E-1(2.2080E-2)‡	8.9790E-1(1.0352E-2)†	8.6441E-1(1.4974E-2)‡	8.6594E-1(5.3327E-3)‡	7.7777E-1(1.3556E-2)‡	**8.9930E-1(7.7577E-3)**
	8	7.0169E-1(5.5025E-2)‡	9.2290E-1(7.9111E-3)†	1.9017E-1(1.8471E-1)‡	9.0534E-1(9.3956E-3)‡	7.2350E-1(1.6221E-2)‡	**9.2855E-1(1.5045E-2)**
	12	5.1133E-1(4.6498E-2)‡	8.2963E-1(5.2905E-2)‡	6.8017E-2(2.6487E-3)‡	9.0514E-1(3.8475E-2)†	6.3134E-1(2.2436E-2)‡	**9.1604E-1(1.2921E-2)**
WFG7	2	7.1926E-1(4.5563E-3)‡	7.3505E-1(2.4614E-4)†	7.3136E-1(1.0965E-3)‡	7.3510E-1(3.1131E-4)†	7.2082E-1(8.9974E-3)‡	**7.3615E-1(1.1625E-3)**
	3	8.1856E-1(1.4867E-2)‡	8.5110E-1(8.3476E-4)†	8.3278E-1(1.5401E-2)‡	8.4263E-1(1.1488E-3)‡	7.9338E-1(1.3433E-2)‡	**8.5171E-1(3.4081E-4)**
	5	7.4819E-1(5.6439E-2)‡	**9.4257E-1(9.5013E-4)**	9.2071E-1(1.3121E-3)‡	8.9358E-1(4.7463E-3)‡	7.4066E-1(1.4528E-2)‡	9.2528E-1(6.1088E-4)
	8	5.8792E-1(7.2407E-2)‡	8.2636E-1(8.6834E-2)‡	7.4490E-1(2.7588E-1)‡	8.4280E-1(2.5436E-2)‡	6.9531E-1(1.9407E-2)‡	**9.5110E-1(1.7669E-2)**
	12	4.6000E-1(8.2486E-2)‡	7.6346E-1(6.1408E-2)‡	2.7676E-1(2.8129E-1)‡	7.5706E-1(3.6351E-2)‡	6.3525E-1(2.3553E-2)‡	**9.4407E-1(9.1515E-3)**
WFG8	2	6.7904E-1(1.0189E-2)‡	6.7029E-1(3.8045E-3)‡	6.6299E-1(2.3559E-2)‡	6.9584E-1(1.4755E-3)‡	6.8948E-1(4.3227E-3)‡	**7.0318E-1(3.3615E-3)**
	3	7.2386E-1(2.5175E-2)‡	7.7446E-1(2.1990E-3)‡	7.6918E-1(4.6432E-3)‡	7.9129E-1(2.5827E-3)‡	7.2355E-1(8.3819E-3)‡	**8.0651E-1(5.3416E-3)**
	5	6.2664E-1(4.0359E-2)‡	8.3085E-1(1.5580E-2)‡	7.9897E-1(1.7626E-3)‡	7.9816E-1(5.4098E-3)‡	7.0110E-1(1.1835E-2)‡	**8.5983E-1(1.8504E-2)**
	8	5.6633E-1(4.5007E-2)‡	8.0925E-1(4.2507E-2)‡	1.2037E-1(2.3436E-2)‡	7.3449E-1(1.9296E-2)‡	6.3416E-1(1.7580E-2)‡	**8.8805E-1(3.8120E-2)**
	12	5.4577E-1(6.2990E-2)‡	7.4233E-1(4.8267E-2)‡	7.6453E-2(2.1549E-3)‡	6.9534E-1(3.5933E-2)‡	5.5797E-1(1.9437E-2)‡	**8.7084E-1(2.2243E-1)**
WFG9	2	6.9556E-1(3.9680E-2)‡	7.1483E-1(8.6702E-4)†	6.8383E-1(2.8965E-2)‡	7.0797E-1(1.1834E-2)‡	7.0116E-1(1.4582E-2)‡	**7.1499E-1(1.4646E-3)**
	3	7.4158E-1(7.1225E-2)‡	8.1680E-1(2.8524E-3)†	7.2575E-1(2.5631E-2)‡	7.8864E-1(1.2055E-2)‡	7.6490E-1(1.6052E-2)‡	**8.2240E-1(2.2316E-2)**
	5	6.4243E-1(1.0678E-1)‡	8.5030E-1(5.3667E-2)†	7.5157E-1(2.2857E-2)‡	7.9189E-1(3.9893E-3)‡	7.1535E-1(2.8081E-2)‡	**8.6833E-1(4.2512E-2)**
	8	4.4778E-1(1.0543E-1)‡	**8.4756E-1(5.2963E-2)**	1.2548E-1(1.6679E-1)‡	7.5797E-1(2.7755E-2)‡	6.3961E-1(1.9066E-2)‡	8.1250E-1(4.0473E-2)
	12	3.8782E-1(1.3885E-1)‡	7.7356E-1(4.2001E-2)‡	5.9415E-2(6.0946E-3)‡	6.6920E-1(3.6199E-2)‡	5.9417E-1(2.9346E-2)‡	**7.8937E-1(4.7711E-2)**

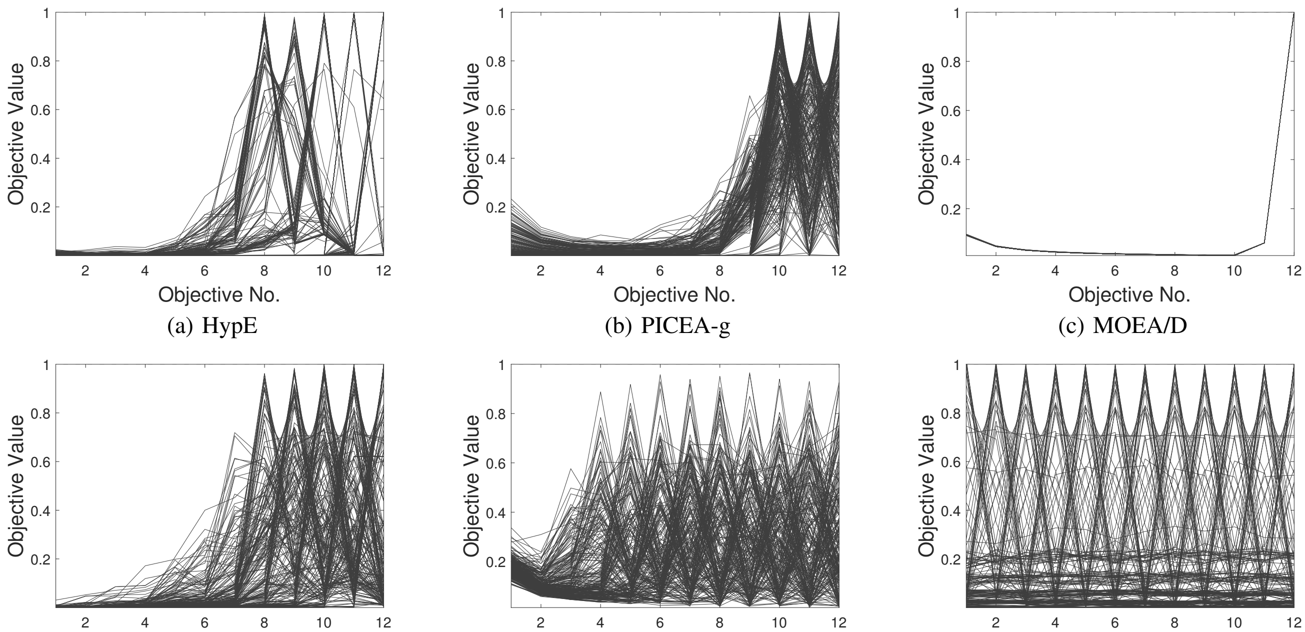
‡ and † indicate SPEA/R performs significantly better than and equivalently to the corresponding algorithm, respectively.

The results are given in Fig. 6, where for each approach, 20 different levels of $Popsize$ are continuously sampled. Clearly, for three objectives, the systematic approach shows better $Popsize$ settings than the k-layer approach. However, when M is increased from 5 to 7, the k-layer approach has more choice to set the population size in the range of 10 to 1000. For 8-objective problems, the two-layer approach works slightly better than the k-layer method, but for much higher objectives, the $Popsize$ generated by the two-layer approach grows very fast, particularly for 30 objectives. In this case, the k-layer approach gives more options to configure a desirable population size. Thus, in comparison with the other two approaches, the k-layer approach appears more suitable for generating a reasonable size of population for MaOPs with a large number of objectives.

The uniformity of reference sets generated by different approaches is investigated on different levels of population size and various number of objectives. The most widely used discrepancy measure, i.e., centred L_2-discrepancy (CD_2) [20], is employed to measure the uniformity of generated reference sets. Comparisons between the simplex-lattice design [13], [48] and our k-layer approach is provided in the supplementary material. The comparisons indicate that the k-layer approach can generate a reference set that covers the whole reference vector space in a good manner.

B. SPEA/R vs NSGA-III

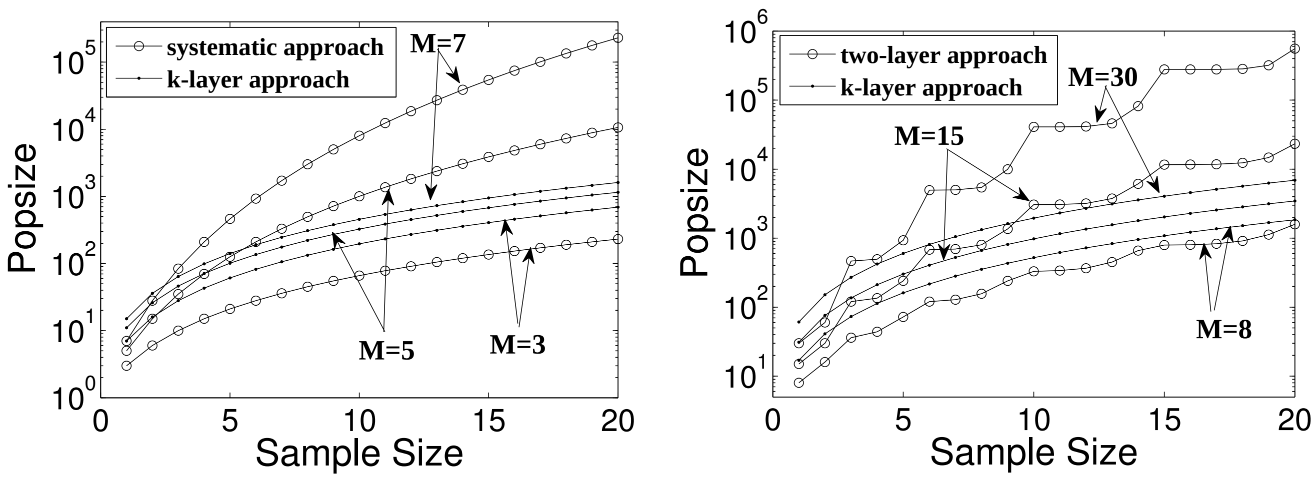
SPEA/R and NSGA-III share some similarities in the way that they keep diversity with the aid of reference directions. Besides different methods for constructing reference directions, there are several key differences between them, resulting in distinct search behaviours. First, SPEA/R introduces a restricted mating selection to enhance reproduction instead of random selection, which is very helpful for many-objective optimization where distant parents are not likely to generate good solutions. Second, SPEA/R uses a simple normalization



(a) HypE (b) PICEA-g (c) MOEA/D

(d) NSGA-III (e) SPEA2+SDE (f) SPEA/R

Fig. 5: Parallel coordinates of final solutions obtained by six algorithms for the 12-objective WFG4 instance.

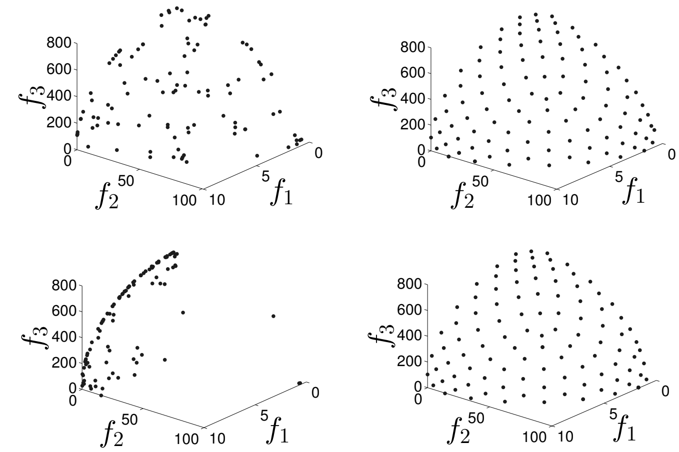


(a) (b)

Fig. 6: Comparison of the population size required by different methods: (a) the systematic approach and k-layer approach for low-dimensional cases; (b) the two-layer approach and k-layer approach for high-dimensional cases.

method based on the worst value of each objective, whereas normalization in NSGA-III requires intercept computation and hyperplane construction, which are computationally expensive, particularly for many-objective problems, and NSGA-III may also have difficulty in hyperplane construction due to duplicate extreme points. Third, niche preservation strategies are different in SPEA/R and NSGA-III. Whenever preserving a member from the last front considered, NSGA-III tries to repeatedly identify reference directions having the worst niche count, which is the number of members associated with these reference directions that has been preserved in higher fronts (the higher, the better). This procedure is computationally inefficient. Furthermore, this strategy may result in some isolated but promising members in lower fronts being abandoned if nondominated sorting terminates before considering these lower fronts, implying that population diversity in NSGA-III is still not well maintained. In contrast, as illustrated in Section III-C, SPEA/R intentionally gives higher priority to diversity than convergence when performing environmental selection, leading to impressive performance on MOP problems, and the niche preservation strategy in SPEA/R has also been further validated on multi- and many-objective WFG problems.

Generally, normalization and niche preservation are all aimed to help keep diversity. To understand the second and third differences, we tested SPEA/R and NSGA-III on disparately scaled three-objective WFG5. That is, the objectives f_1, f_2, and f_3 are multiplied with 5, 5^2, and 5^3, respectively. Fig. 7 plots approximated POFs of the median and worst IGD

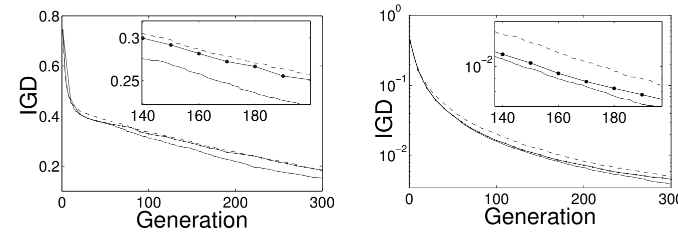


NSGA-III SPEA/R

Fig. 7: Approximated POFs for scaled WFG5. Top: the median approximation; bottom: the worst approximation.

TABLE V: Statistical difference between SPEA/R and two variants

	Sign	SPEA/R-A	SPEA/R-B	SPEA/R-C
SPEA/R vs.	B	12	4	51
	E	40	47	1
	W	0	1	0



(a) MOP3 (b) WFG4

Fig. 8: The IGD metric against the number of generations for two instances: SPEA/R (solid); SPEA/R-A (dashed); SPEA/R-B (dotted).

values over 31 runs, showing that the simple normalization method used in SPEA/R can deal with scaling objectives and the new diversity-first-and-convergence-second strategy is capable of providing a uniform distribution of solutions. NSGA-III, however, struggles to solve the scaled WFG5. In the median case of NSGA-III, the intercept-based normalization is able to diversify points over the whole POF but the niche preservation cannot provide a good distribution. In the worst case, NSGA-III drives the majority of points toward the $f_1 f_3$ plane, and misses a large part of the POF. One reason for this is that NSGA-III tends to preserve members in higher fronts that have better convergence, and less-converged isolated ones in lower fronts are likely to leave unconsidered, leading to poor diversity during the search. Therefore, NSGA-III cannot compete with SPEA/R in terms of diversity.

C. *Influence of Fitness Assignment and Niche Preservation*

Although Section III-C has revealed that good population diversity contributes to the performance of SPEA/R, in this subsection we try to unveil more reasons behind the high performance of SPEA/R.

Fitness assignment and diversity preservation are the core of SPEA/R, which control the balance between convergence and diversity. To understand why our strategy yields high performance, we further design three other SPEA/R variants that use different strategies. The first one, called SPEA/R-A, removes global fitness from Eq. (9) when calculating individuals' fitness. Instead of removing global fitness, the second variant, i.e., SPEA/R-B, removes local fitness, so an individual's final fitness is composed of global fitness and density. The third variant (named SPEA/R-C) allows individuals with the highest fitness to enter into the next generation in environmental selection. Thus, SPEA/R-A favours diversity whereas SPEA/R-C favours convergence, and SPEA/R-B does not consider local convergence. The variants are compared with the original SPEA/R on 7 MOP problems and 9 WFG problems with 2, 3, 5, 8, and 12 objectives. The HV results obtained by each algorithm for a total of 52 instances are presented in our supplementary material, and only statistical testing result is given in Table V, which is based on the Wilcoxon signed-rank test [45] at the 0.05 significance level with Bonferroni correction. In the table, the signs 'B', 'E', and 'W' represent SPEA/R is significantly better, equivalent to, and worse than the compared variant, respectively.

It is clear that, SPEA/R generally performs better than the other variants. Specifically, SPEA/R outperforms SPEA/R-A in a number of cases, indicating that the use of global fitness is a good choice for SPEA/R in some situations. The comparison between SPEA/R and SPEA/R-B shows that the use of local fitness does not make a big difference but may help SPEA/R achieve slightly better performance for a few instances. For SPEA/R-C, the lack of diversity maintenance induces a significant lag behind SPEA/R. This observation further confirms that the high performance of SPEA/R is mainly due to sound diversity preservation.

Since SPEA/R, SPEA/R-A, and SPEA/R-B differ only in fitness assignment, one would wonder why SPEA/R works better (though not significantly better in most cases) than the other two variants. To investigate this, we plot the mean IGD curves of these variants against the first 300 generations on MOP3 and 2-objective WFG4, as shown in Fig. 8. It can be observed that SPEA/R converges fastest, followed by SPEA/R-B, and SPEA/R-A ranks last. SPEA/R is better than SPEA/R-A because adding local fitness can strengthen discrimination between individuals so that more-converged individuals can be preserved. In contrast, without the use of global fitness, SPEA/R-A converges relatively slower than SPEA/R and SPEA/R-B. This illustrates that the joint use of global fitness and local fitness can speed up the search process, although not very significantly. However, we should point out that there is no much difference between SPEA/R and SPEA/R-B in high-dimensional problems. This is because the local fitness value will be zero when the majority of individuals are nondominated in high dimensions. In other words, SPEA/R may degenerate to SPEA/R-B in this situation.

TABLE VI: Mean and standard deviation HV values obtained by SPEA/R with different K values for four WFG problems

Prob.	M	K=2	K=5	K=10	K=20	K=40	K=50
WFG5	2	7.0444E-1(2.6807E-3)	7.0331E-1(1.4893E-3)	7.0290E-1(1.5121E-3)	**7.1227E-1(9.3626E-4)**	7.0240E-1(1.0190E-3)	7.0225E-1(1.2918E-3)
	3	8.1491E-1(2.2212E-3)	8.1506E-1(2.5168E-3)	8.1509E-1(2.6994E-3)	**8.2437E-1(2.0573E-3)**	8.1398E-1(1.3143E-3)	8.1376E-1(1.1227E-3)
	5	8.7474E-1(2.0417E-3)	8.7699E-1(2.8228E-3)	8.7754E-1(1.4904E-3)	**8.8792E-1(1.1292E-3)**	8.7778E-1(1.3177E-3)	8.7786E-1(1.0863E-3)
	8	8.8822E-1(8.3131E-3)	8.9728E-1(3.1274E-3)	8.9833E-1(1.2579E-2)	9.0023E-1(2.9488E-3)	**9.0064E-1(3.1272E-3)**	8.9896E-1(3.0760E-3)
	12	8.2542E-1(4.0708E-2)	8.3170E-1(6.4925E-2)	8.6648E-1(9.6071E-2)	9.0236E-1(6.6030E-3)	**9.0357E-1(5.3041E-3)**	8.9903E-1(1.3462E-2)
WFG6	2	7.0956E-1(7.4317E-3)	**7.1207E-1(4.9637E-3)**	7.1049E-1(6.8942E-3)	7.1001E-1(6.5120E-3)	7.1125E-1(4.4444E-3)	7.1112E-1(6.4365E-3)
	3	8.1414E-1(8.6380E-3)	8.1048E-1(6.0997E-3)	8.1501E-1(5.5185E-3)	**8.2552E-1(6.3065E-3)**	8.1704E-1(6.4877E-3)	8.1612E-1(6.9906E-3)
	5	8.7020E-1(1.2008E-2)	8.7366E-1(1.2462E-2)	8.7589E-1(8.0459E-3)	**8.9930E-1(7.7577E-3)**	8.7429E-1(1.3280E-2)	8.7756E-1(8.3509E-3)
	8	8.8848E-1(2.9622E-2)	8.9990E-1(1.9014E-2)	9.0585E-1(1.0637E-2)	**9.2855E-1(1.5045E-2)**	9.0146E-1(1.2872E-2)	9.0041E-1(1.2033E-2)
	12	8.6720E-1(8.0397E-2)	8.8930E-1(4.9023E-2)	8.9676E-1(5.2199E-3)	**9.1604E-1(1.2921E-2)**	9.0072E-1(2.5975E-2)	8.9359E-1(6.1651E-2)
WFG7	2	7.3554E-1(5.8256E-4)	7.3573E-1(4.3877E-4)	7.3570E-1(4.2235E-4)	**7.3615E-1(1.1625E-3)**	7.3531E-1(5.4513E-4)	7.3455E-1(1.5055E-3)
	3	8.5109E-1(3.6231E-4)	8.5135E-1(2.8815E-4)	8.5120E-1(3.8447E-4)	**8.5171E-1(3.4081E-4)**	8.5022E-1(3.7511E-4)	8.5007E-1(3.9649E-4)
	5	9.2148E-1(1.0175E-3)	9.2385E-1(5.7867E-4)	9.2498E-1(6.6708E-4)	**9.2528E-1(6.1088E-4)**	9.2504E-1(8.4886E-4)	9.2451E-1(1.7558E-3)
	8	8.1636E-1(7.3633E-2)	8.8675E-1(5.2499E-2)	9.1107E-1(6.1033E-2)	**9.5110E-1(1.7669E-2)**	9.0279E-1(3.2266E-2)	8.4598E-1(6.7261E-2)
	12	9.0203E-1(1.1885E-1)	9.2156E-1(8.3763E-2)	9.3326E-1(3.5592E-2)	**9.4407E-1(9.1515E-3)**	8.8898E-1(8.1680E-2)	8.5109E-1(1.4778E-1)
WFG8	2	6.9842E-1(2.5016E-3)	7.0146E-1(3.4153E-3)	7.0138E-1(3.2326E-3)	**7.0318E-1(3.3615E-3)**	7.0262E-1(3.5139E-3)	7.0264E-1(3.2564E-3)
	3	7.9341E-1(3.4538E-3)	8.0085E-1(6.2528E-3)	8.0478E-1(7.1091E-3)	**8.0651E-1(5.3416E-3)**	8.0817E-1(5.9836E-3)	8.0802E-1(5.7856E-3)
	5	8.0213E-1(3.0860E-3)	8.0994E-1(3.9769E-3)	8.2573E-1(2.0662E-2)	**8.5983E-1(1.8504E-2)**	8.5576E-1(3.6934E-2)	8.4240E-1(1.7857E-2)
	8	8.6003E-1(6.8351E-2)	8.6155E-1(4.5106E-2)	8.7114E-1(4.3606E-2)	**8.8805E-1(3.8120E-2)**	8.6770E-1(5.7081E-2)	8.5679E-1(6.3412E-2)
	12	8.4546E-1(5.2716E-2)	8.5492E-1(2.0422E-1)	8.5888E-1(2.2243E-1)	**8.7084E-1(2.2243E-1)**	8.5445E-1(1.6750E-1)	7.4837E-1(1.6586E-1)

D. *Influence of Restricted Mating*

Restricted mating selection is somewhat similar to the concept of neighbourhood used in MOEA/D, in which close parents are likely to generate good offspring. It has a key parameter K, i.e., the number of parent candidates, and the influence of this parameter is investigated on four WFG problems. Table VI reports the HV values obtained by SPEA/R with different settings of K. It can be observed that, $K = 20$ (10%–20% of population size) yields better results than the other settings for all the cases except the 8- and 12-objective WFG5 and the 2-objective WFG6. Particularly, for many-objective problems, e.g., the 8- and 12-objective cases, there is a noticeable improvement on the HV metric when K is increased from 2 to 20. This means proper restricted mating can benefit population reproduction, thereby promoting algorithms' performance for many-objective optimization.

The above experiment has shown that proper restricted mating is good for population reproduction. However, we should point out that, restricted mating can be used only when population diversity is well maintained. This is because, if population individuals are not well distributed, then restricted mating can cause overexploitation in overcrowded regions so that isolated regions may be left under-explored or even unexplored, resulting in a further deterioration of population diversity. This has been illustrated in Section V-C, where the overlook of diversity maintenance makes SPEA/R-C significantly worse than SPEA/R although restricted mating has been employed there.

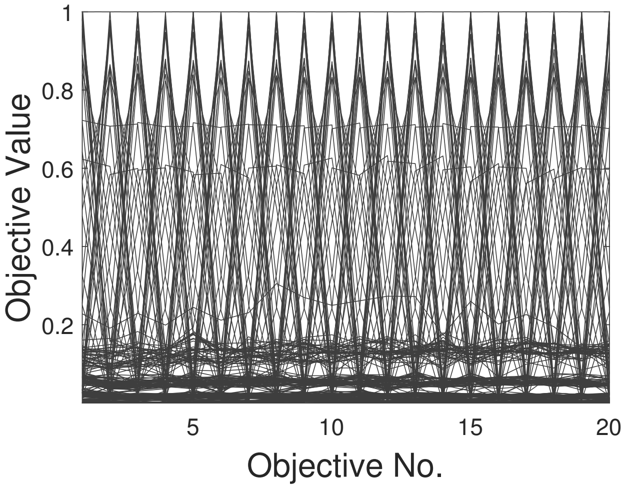
E. *Peformance of SPEA/R on Problems with More Objectives*

SPEA/R has the advantage of population diversity maintenance so that it can handle high-dimensional problems. To further investigate whether this advantage can deal with problems with more objectives, we tested SPEA/R on WFG4 with 20 and 40 objectives. This means the difficulty of the problem is massively increased as nearly all population members are nondominated with respect to each other. The population size was set to 280 and 560 for 20 and 40 objectives, respectively. Due to the increase of the difficulty of the problem, SPEA/R should be given more computational resources. Hence, the maximum number of generations was set to 3000 and 5000 for 20 and 40 objectives, respectively.

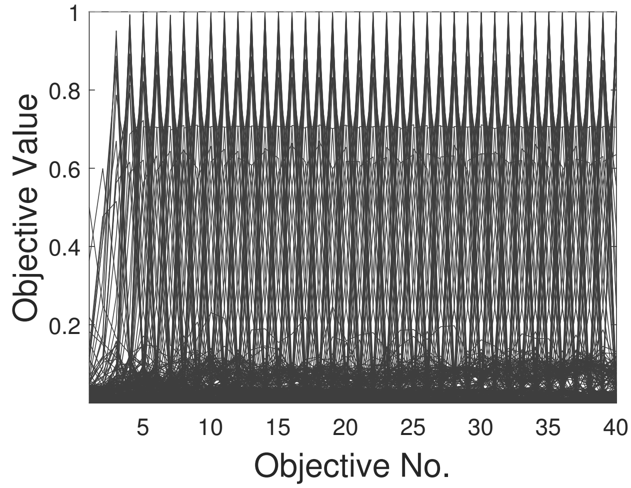
Fig. 9 shows the normalized parallel coordinates of final solutions obtained by SPEA/R for two instances. Clearly, on both 20 and 40 objectives, SPEA/R can still obtain a set of diverse solutions in the entire range of the POF. Thus, the proposed diversity-first-and-convergence-second selection strategy in SPEA/R is very promising for solving many-objective problems.

F. *More Discussions*

It has been well recognized that convergence and diversity are two main but hard-to-balance goals in designing MOEAs. Any bias toward one goal will inevitably aggravate the other. In many-objective optimization the balance between them is still of great importance. However, when handling MaOPs, most MOEAs inherit elitist preservation from their counterparts of low-dimensional optimization that emphasizes nondominated solutions in the population, resulting in very little room left for diversity maintenance. Even if these MOEAs did not intentionally emphasize convergence, they could not elude the fact that an increasingly large fraction of population becomes nondominated with an increase in the number of objectives. In other words, they perform environmental selection in a convergence-first-and-diversity-second manner. As a result, when the MOEAs are applied to high-dimensional optimization, there will be a large number of nondominated individuals after the convergence-first selection, and diversity preservation will be performed only on the nondominated individuals. Correspondingly, some regions occupied by dominated individuals will be scarcely explored, and diversity preservation becomes of limited use in this case. In contrast, SPEA/R adopts a diversity-first-and-convergence-second strategy to perform environmental selection, at an attempt to maximize population diversity and strengthen exploitation in



(a)



(b)

Fig. 9: Parallel coordinates of final solutions obtained by SPEA/R for WFG with 20 (a) and 40 (b) objectives.

less-converged regions during the search. Our experiments have shown its promise for both multi- and many-objective optimization.

However, we may wonder why SPEA/R can work well on problems with over 12 objectives, where nearly all individuals (over 95% of population) are nondominated [23]. This means, in this situation, the diversity-first-and-convergence-second strategy in SPEA/R has no advantage over other MOEAs in diversity preservation because there is hardly any region that can be occupied by very few dominated individuals. There is no doubt that, when population is randomly generated, the fraction of dominated individuals is close to zero for 10 objectives and over [23]. But, what if the population is a combination of parent and offspring populations, which is the case with MOEAs? To investigate this, we consider the search behaviour of SPEA/R on the 12-objective WFG5 over 2000 generations. In every generation, SPEA/R distributes a combined population toward H_M^k subregions (which equals the total number of reference directions) of the objective space, and the number (N_d) of subregions in which only dominated solutions reside is recorded. Fig. 10 shows the relative frequency of different N_d values over 2000 generations. Clearly, in the majority of generations dominated solutions do not solely occupy any subregions. In this situation, dominated solutions make little contribution to diversity as nondominated solutions covers all subregions of the evolving population. However, there are also over 20% generations in which some subregions are occupied by dominated but not nondominated solutions. In this case, dominated solutions make a difference to population diversity.

Additionally, we also compute the percentage of dominated solutions in the combined population of every generation of SPEA/R for a single run, as shown in Fig. 11. It can be observed from the figure that, there is still a noticeable proportion of dominated solutions in the combined population before the population converges to the POF. All these observations clearly confirm that preservation of dominated solutions for diversity promotion through the diversity-first-and-convergence-second

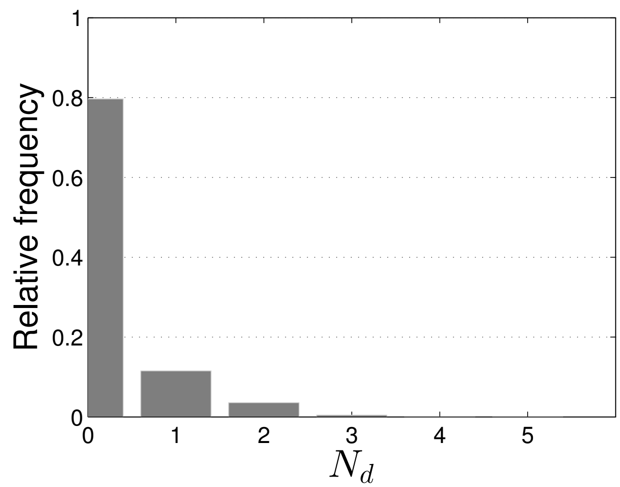


Fig. 10: The relative frequency of the number of subregions occupied only by dominated solutions.

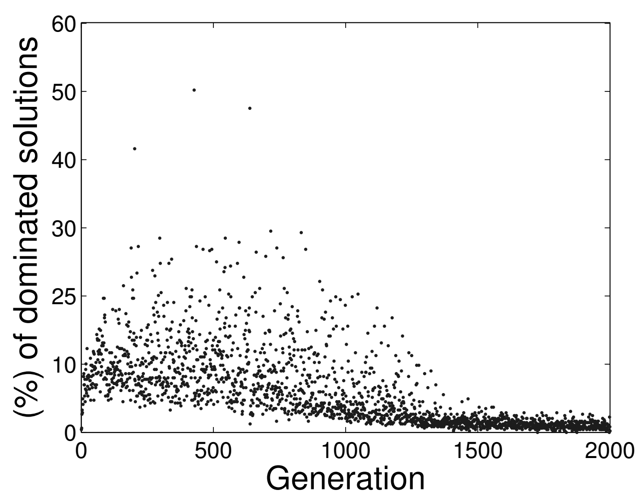


Fig. 11: The percentage of dominated solutions in every generation of SPEA/R for 12-objective WFG5.

strategy is still beneficial to SPEA/R when handing high-dimensional problems.

On the other hand, Fig. 11 can also be used to explain why the compared MOEAs in this paper cannot compete with SPEA/R. As shown in this figure, there are at least 50% non-dominated solutions in the combined population nearly every generation. Since HypE, MOEA/D, PICEA-g, NSGA-III, and SPEA2+SDE all prefer nondominated solutions, there is no room for them to preserve dominated but diverse solutions when selecting only half of the combined population for next generation. As a consequence, regions occupied by dominated solutions will be left unexplored, which can cause diversity loss in the population.

The reason why our observation is inconsistent with the study of [23] in terms of the proportion of dominated solutions is that, the combined population comprises parent and offspring members, and there is a close relationship between them. Thus, there are more dominated members than expected. However, we should be aware that there might be very few or even no dominated solutions if the number of objectives is considerably large, e.g., 100. In this case, the diversity-first-and-convergence-second selection strategy may be of limited use.

VI. Conclusions

It has been repeatedly reported that conventional Pareto-dominance based MOEAs may be unsuitable for many-objective optimization, although they can successfully solve two- or three-objective problems. In this paper, we have suggested a reference-direction based method to revive an early SPEA algorithm for handling both MOPs and MaOPs. Through incorporating a set of predefined reference directions, the proposed algorithm, i.e., SPEA/R, partitions the objective space into a number of subregions of interest, and individuals in each subregion are guided toward predefined search directions. Unlike most existing MOEAs prefering nondominatd solutions, SPEA/R adopts a diversity-first-and-convergence-second selection strategy, which can increase the selection pressure for many-objective optimization where a large fraction of population is nondominated. SPEA/R also employs a restricted mating scheme to improve reproduction efficiency. Besides, the proposed framework has significantly reduced the computational effort of SPEA-based methods, providing the overall computational complexity bounded by $O(MN^2)$.

Our experimental study has demonstrated the efficacy of SPEA/R on a number of MOP and WFG test problems with 2 to 40 objectives and various optimization difficulties. A fair comparison with several state-of-the-art MOEAs suggests that SPEA/R is very comparative for both multi- and many-objective optimization. This implies that giving high priority to diversity over convergence can be another effective way to handle many-objective optimization.

Although SPEA/R has provided encouraging performance on the test problems considered in this paper, it needs to be examined on a wider range of problems (e.g., complicated POS and POF shapes). Also, as the research on many-objective optimization is still in its infancy, there are some open issues remaining to be solved, such as the computationally expensive calculation of performance metrics and visualization of a higher-dimensional trade-off front. Therefore, these should be very interesting topics for our future work.

References

[1] H. Abdi, "Bonferroni and Šidák corrections for multiple comparison," in *Encyclopedia of Measurement and Statistics*, N. J. Salkind, Eds. Thousand Oaks, CA: Sage, 2007.

[2] S. F. Adra and P. J. Fleming, "Diversity management in evolutionary many-objective optimization," *IEEE Trans. Evol. Comput.*, vol. 15, no. 2, pp. 183–195, 2011.

[3] M. Aruldoss, T. M. Lakshmi, and V. P. Venkatesan, "A survey on multi criteria decision making methods and its applications," *American Journal of Information Systems*, vol. 1, no. 1, pp. 31–43, 2013.

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

[5] L. S. Batista, F. Campelo, F. G. Guimaraes, and J. A. Ramirez, "Pareto cone ϵ-dominance: Improving convergence and diversityin multiobjective evolutionary algorithms," in *Proc. 6th Int. Conf. Evol. Multi-criterion Optimiz.*, LNCS, vol. 6576, 2011, pp. 76–90.

[6] D. Brockhoff, E. Zitzler, "Dimensionality reduction in multiobjective optimization: The minimum objective subset problem," in *Oper. Res. Proc. 2006*. Berlin: Springer, vol. 2006, 2007, pp. 423–429.

[7] D. Cai and Y. Wang, "A new multiobjective evolutionary algorithm based on decomposition of the objective space for multiobjective optimization," *J. Appl. Math.*, vol. 2014, pp. 1–9, 2014.

[8] Y. Cheung, F. Gu, and H. Liu, "Objective extraction for many-objective optimization problems: Algorithm and test problems," *IEEE Trans. Evol. Comput.*, 2015, in press. DOI: 10.1109/TEVC.2016.2519758.

[9] I. Das and J. Dennis, "Normal-boundary intersection: A new method for generating the Pareto surface in nonlinear multicriteria optimization problems," *SIAM J. Optimiz.*, vol. 8, no. 3, pp. 631–657, 1998.

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

[11] K. Deb, *Multi-Objective Optimization Using Evolutionary Algorithms*. Chichester: John Wiley and Sons, 2001.

[12] K. Deb and M. Goyal, "A combined genetic adaptive search (GeneAS) for engineering design," *Computer Science and Informatics*, vol. 26, pp. 30–45, 1996.

[13] K. Deb and H. Jain, "An evolutionary many-Objective optimization algorithm using reference-point based non-dominated sorting approach, Part I: Solving problems with box constraints," *IEEE Trans. Evol. Comput.*, vol. 18, no. 4, pp. 577–601, 2014.

[14] K. Deb, A. Sinha, P. J. Korhonen, and J. Wallenius, "An interactive evolutionary multiobjective optimization method based on progressively approximated value functions," *IEEE Trans. Evol. Comput.*, vol. 14, no. 5, pp. 723–739, Oct. 2010.

[15] K. Deb, L. Thiele, M. Laumanns, and E. Zitzler, "Scable test problems for evolutionary multi-objective optimization," in *Evolutionary Multiobjective Optimization*. London: Springer, 2005, pp. 105–145.

[16] M. Farina and P. Amato, "A fuzzy definition of optimality for many-criteria problems," *IEEE Trans. Syst., Man and Cybern. Part A: Syst. and Humans*, vol. 34, no. 3, pp. 315–326, 2004.

[17] I. Giagkiozis, R. C. Purshouse, and P. J. Fleming, "Generalized decomposition and cross entropy methods for many-objective optimization," *Inform. Sci.*, vol. 282, pp. 363–387, 2014.

[18] D. Hadka and P. Reed, "Diagnostic assessment of search controls and failure modes in many-objective optimization," *Evol. Comput.*, vol. 20, no. 3, pp. 423–453, 2012.

[19] Z. He, G. G. Yen, and J. Zhang, "Fuzzy-based Pareto optimality for many-objective evolutionary algorithms," *IEEE Trans. Evol. Comput.*, vol. 18, no. 2, pp. 269–285, 2014.

[20] F. J. Hickernell, "A generalized discrepancy and quadrature error bound," *Math. Comput.*, vol. 67, no. 221, pp. 299–322, 1998.

[21] 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. 2, pp. 477–506, 2006.

[22] H. Ishibuchi, N. Tsukamoto, Y. Hitotsuyanagi, and Y. Nojima, "Indicator-based evolutionary algorithm with hypervolume approximation by achievement scalarizing function," in *Proc. 12th Ann. Conf. Genetic Evol. Comput. (GECCO 2010)*, 2010, pp. 527–534.

[23] H. Ishibuchi, N. Tsukamoto, and Y. Nojima, "Evolutionary many-objective optimization: A short review," in *Proc. 2008 IEEE Congr. Evol. Comput.*, 2008, pp. 2424–2431.

[24] H. Ishibuchi, Y. Sakane, N. Tsukamoto, and Y. Nojima, "Evolutionary many-objective optimization by NSGA-II and MOEA/D with large populations," in *Proc. 2009 IEEE Int. Conf. Syst. Man, Cybern.*, vol. 1, 2009, pp. 1820–1825.

[25] S. Jiang and S. Yang, "An improved multi-objective optimization evolutionary algorithm based on decomposition for complex Pareto fronts," *IEEE Trans. Cybern.*, vol. 46, no. 2, pp. 421–437, 2016.

[26] R. Joshi, B. Deshpande, "Empirical and analytical study of many-objective optimization problems: Analysing distribution of nondominated solutions and population size for scalability of randomized heuristics," *Memetic Comput.*, vol. 6, no. 2, pp. 133–145, 2014.

[27] J. R. Kasprzyk, P. M. Reed, B. R. Kirsch, G. W. Characklis, "Managing population and drought risks using many-objective water portfolio planning under uncertainty," *Water Resour. Res.*, vol. 45, no. 12, Dec. 2009.

[28] J. D. Knowles and D. W. Corne, "The Pareto archived evolution strategy: A new baseline algorithm for multiobjective optimisation," in *Proc. 1999 IEEE Congr. Evol. Comput.*, 1999, pp. 98–105.

[29] N. Kowatari, A. Oyama, H. E. Aguirre, and K. Tanaka, "A study on large population MOEA using adaptive ϵ-box dominance and neighborhood recombination for many-objective optimization," in *Learning and Intelligent Optimization*, LNCS, vol. 7219, 2012, pp. 86–100.

[30] 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, Jan. 2002.

[31] 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, 2015.

[32] M. Li, S. Yang, X. Liu, "Shift-based density estimation for Pareto-based algorithms in many-objective optimization," *IEEE Trans. Evol. Comput.*, vol. 18, no. 3, pp. 348–365, Jun. 2014.

[33] H. Li and Q. Zhang, "Multiobjective optimization problems with complicated Pareto sets, MOEA/D and NSGA-II," *IEEE Trans. Evol. Comput.*, vol. 13, no. 2, pp. 284–302, 2009.

[34] H. 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, 2014.

[35] H. Liu and X. Li, "The multiobjective evolutionary algorithm based on determined weight and sub-regional search," in *Proc. Congr. Evol. Comput.*, vol. 15, 2009, pp. 361–384.

[36] F. di Pierro, S.-T. Khu, and D. A. Savic, "An investigation on preference order ranking scheme for multiobjective evolutionary optimization," *IEEE Trans. Evol. Comput.*, vol. 11, no. 1, pp. 17–45, 2007.

[37] D. K. Saxena, J. A. Duro, A. Tiwari, K. Deb, and Q. Zhang, "Objective reduction in many-objective optimization: Linear and nonlinear algorithms," *IEEE Trans. Evol. Comput.*, no. 17, no. 1, pp. 77–99, Feb. 2013.

[38] H. Singh, A. Isaacs, T. Ray, "A Pareto corner search evolutionary algorithm and dimensionality reduction in many-objective problems," *IEEE Trans. Evol. Comput.*, vol. 15, no. 4, pp. 539–556, 2011.

[39] R. Takahashi, E. Carrano, and E. Wanner, "On a stochastic differential equation approach for multiobjective optimization up to Pareto-criticality," in *Proc. 6th Int. Conf. Evol. Multi-criterion Optimiz. (EMO 2011)*, LNCS, vol. 6576, 2011, pp. 61–75.

[40] R. 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, 2012.

[41] T. Wagner, N. Beume, and B. Naujoks, "Pareto-, aggregation-, and indicator-based methods in many-objective optimization," in *Proc. 4th Int. Conf. Evol. Multi-criterion Optimiz.*, 2007, pp. 742–756.

[42] H. Wang, L. Jiao, and X. Yao,"Two arch2: An improved two-archive algorithm for many-objective optimization," *IEEE Trans. Evol. Comput.*, vol. 19, no. 4, pp. 524âĂŞ-541, 2015.

[43] R. Wang, R. C. Purshouse, and P. J. Fleming, "Preference-inspired coevolutionary algorithms for many-objective optimization," *IEEE Trans. Evol. Comput.*, vol. 17, no. 4, pp. 474–494, Aug. 2013.

[44] L. Wang, Q. Zhang, A. Zhou, M. Gong, and, L. Jiao, "Constrained subproblems in decomposition based multiobjective evolutionary algorithm," *IEEE Trans. Evol. Comput.*, vol. 20, no. 3, pp. 475–480, 2015.

[45] F. Wilcoxon, "Individual comparisons by ranking methods," *Biometrics*, vol. 1, no. 6, pp. 80–83, 1945.

[46] Y. Yuan, H. Xu, B. Wang, B. Zhang, and X. Yao, "Balancing convergence and diversity in decomposition-based many-objective optimizers," *IEEE Trans. Evol. Comput.*, vol. 20, no. 2, pp. 180–198, 2016.

[47] 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, 2016.

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

[49] Q. Zhang, W. Liu, and H. Li, "The performance of a new version of MOEA/D on CEC09 unconstrained MOP test instances," Work. Rep. CES-491, School of CS and EE, University of Essex, Colchester, UK, 2009.

[50] C. Zhu, L. Xu, and E. D. Goodman, "Generalization of Pareto-optimality for many-objective evolutionary optimization," *IEEE Trans. Evol. Comput.*, vol. 20, no. 2, pp. 299âĂŞ-315, 2015.

[51] E. Zitzler, S. Kunzli, "Indicator-based selection in multiobjective search," in *Proc. 8th Int. Conf. Parallel Problem Solving from Nature (PPSN VIII)*, LNCS, vol. 3242, 2004, pp. 832–842.

[52] E. Zitzler, K. Deb, and L. Thiele, "Comparison of multiobjective evolutionary algorithms: Empirical results," *Evol. Comput.*, vol. 8, no. 2, pp. 173–195, 2000.

[53] E. Zitzler, M. Laumanns, and L. Thiele, "SPEA2: Improving the strength Pareto evolutionary algorithm for multiobjective optimization," in *Proc. Evol. Methods Des. Optimiz. Contr. Applicat. Indust. Problems*, EUROGEN, vol. 3242, no. 103, 2002, pp. 95-100.

[54] E. Zitzler and L. Thiele, " Multiobjective evolutionary algorithms: A comparative case study and the strength Pareto approach," *IEEE Trans. Evol. Comput.*, vol. 3, no. 4, pp. 257–271, 1999.

Shouyong Jiang received the B.Sc. degree in information and computation science and the M.Sc. degree in control theory and control engineering from Northeastern University, Shenyang, China in 2011 and 2013, respectively.

He is currently pursuing the Ph.D. degree in the School of Computer Science and Informatics, De Montfort University, Leicester, U.K. His current research interests include evolutionary computation, multi-objective optimization, and dynamic optimization.

Shengxiang Yang (M'00–SM'14) received the B.Sc. and M.Sc. degrees in automatic control and the Ph.D. degree in systems engineering from Northeastern University, Shenyang, China in 1993, 1996, and 1999, respectively.

He is currently a Professor in Computational Intelligence and Director of the Centre for Computational Intelligence, School of Computer Science and Informatics, De Montfort University, Leicester, U.K. He has over 210 publications. His current research interests include evolutionary and genetic algorithms, swarm intelligence, computational intelligence in dynamic and uncertain environments, artificial neural networks for scheduling, and relevant real-world applications.

Prof. Yang is the Chair of the Task Force on Evolutionary Computation in Dynamic and Uncertain Environments, under the Evolutionary Computation Technical Committee of the IEEE Computational Intelligence Society and the Founding Chair of the Task Force on Intelligent Network Systems, under the Intelligent Systems Applications Technical Committee of the IEEE Computational Intelligence Society.

A Strength Pareto Evolutionary Algorithm Based on Reference Direction for Multi-objective and Many-objective Optimization

Shouyong Jiang and Shengxiang Yang, *Senior Member, IEEE*

This is the supplementary material to the paper entitled "A Strength Pareto Evolutionary Algorithm Based on Reference Direction for Multi-objective and Many-objective Optimization", written by Shouyong Jiang and Shengxiang Yang. This material provides a detailed description of a multi-objective test suite and and a recombination operator used in the paper. After that, there are discussions on the performance of reference direction set used in the proposed algorithm, followed by supplementary results for the MOP test suite and detailed comparisons between different fitness assignments and niche preservation schemes. Some additional figures of parallel coordinates indicating algorithms' final performance on many-objective optimization are also presented in this material.

VII. MOP Test Suite

- MOP1

$$f_1(x) = (1+g(x))x_1$$
$$f_2(x) = (1+g(x))(1-\sqrt{x_1})$$

where $g(x) = 2\sin(\pi x_1)\sum_{i=2}^{n}(-0.9t_i^2 + |t_i|^{0.6})$, $t_i = x_i - \sin(0.5\pi x_1)$, $i = 2,\dots,n$. The search space is $[0,1]^n$, $n = 10$. Its POS is $\{x \in R^{(n-1)} | x_i = \sin(0.5\pi x_1), n = 2,\dots,n, x_1 \in [0,1]\}$. Its POF is $\{(f_1,f_2)|f_2 = 1-\sqrt{f_1}, f_1 \in [0,1]\}$.

- MOP2

$$f_1(x) = (1+g(x))x_1$$
$$f_2(x) = (1+g(x))(1-x_1^2)$$

where $g(x) = 10\sin(\pi x_1)\sum_{i=2}^{n}\frac{|t_i|}{1+\exp(5|t_i|)}$, $t_i = x_i - \sin(0.5\pi x_1)$, $i = 2,\dots,n$. The search space is $[0,1]^n$, $n = 10$. Its POS is the same as that of MOP1. Its POF is $\{(f_1,f_2)|f_2 = 1-f_1^2, f_1 \in [0,1]\}$.

- MOP3

$$f_1(x) = (1+g(x))\cos(0.5\pi x_1)$$
$$f_2(x) = (1+g(x))\sin(0.5\pi x_1)$$

where $g(x) = 10\sin(\pi x_1)\sum_{i=2}^{n}\frac{|t_i|}{1+\exp(5|t_i|)}$, $t_i = x_i - \sin(0.5\pi x_1)$, $i = 2,\dots,n$. The search space is $[0,1]^n$, $n = 10$. Its POS is the same as that of MOP1. Its POF is $\{(f_1,f_2)|f_2 = \sqrt{1-f_1^2}, f_1 \in [0,1]\}$.

- MOP4

$$f_1(x) = (1+g(x))x_1$$
$$f_2(x) = (1+g(x))(1-\sqrt{x_1}\cos^2(2\pi x_1))$$

where $g(x) = 10\sin(\pi x_1)\sum_{i=2}^{n}\frac{|t_i|}{1+\exp(5|t_i|)}$, $t_i = x_i - \sin(0.5\pi x_1)$, $i = 2,\dots,n$. The search space is $[0,1]^n$, $n = 10$. Its POS is the same as that of MOP1. Its POF is $\{(f_1,f_2)|f_2 = 1-\sqrt{f_1}\cos^2(2\pi f_1), f_1 \in [0,1]\}$.

- MOP5

$$f_1(x) = (1+g(x))x_1$$
$$f_2(x) = (1+g(x))(1-\sqrt{x_1})$$

where $g(x) = 2\cos(\pi x_1)\sum_{i=2}^{n}(-0.9t_i^2 + |t_i|^{0.6})$, $t_i = x_i - \sin(0.5\pi x_1)$, $i = 2,\dots,n$. The search space is $[0,1]^n$, $n = 10$. Its POS is the same as that of MOP1. Its POF is a part of $\{(f_1,f_2)|f_2 = 1-\sqrt{f_1}, f_1 \in [0,1]\}$.

- MOP6

$$f_1(x) = (1+g(x))x_1x_2$$
$$f_2(x) = (1+g(x))x_1(1-x_2)$$
$$f_3(x) = (1+g(x))(1-x_1)$$

where $g(x) = 2\sin(\pi x_1)\sum_{i=3}^{n}(-0.9t_i^2 + |t_i|^{0.6})$, $t_i = x_i - x_1x_2$, $i = 3,\dots,n$. The search space is $[0,1]^n$, $n = 10$. Its POS is $\{x \in R^{(n-2)} | x_i = \sin(0.5\pi x_1), i = 3,\dots,n, x_j \in [0,1], j = 1,2\}$. Its POF is $\{(f_1,f_2,f_3)|f_1+f_2+f_3 = 1, f_i \in [0,1], i = 1,2,3\}$.

- MOP7

$$f_1(x) = (1+g(x))\cos(0.5\pi x_1)\cos(0.5\pi x_2)$$
$$f_2(x) = (1+g(x))\cos(0.5\pi x_1)\sin(0.5\pi x_2)$$
$$f_2(x) = (1+g(x))\sin(0.5\pi x_1)$$

where $g(x) = 2\sin(\pi x_1)\sum_{i=3}^{n}(-0.9t_i^2 + |t_i|^{0.6})$, $t_i = x_i - x_1x_2$, $i = 3,\dots,n$. The search space is $[0,1]^n$, $n = 10$. Its POS is $\{x \in R^{(n-2)} | x_i = \sin(0.5\pi x_1), n = 3,\dots,n, x_j \in [0,1], j = 1,2\}$. Its POF is $\{(f_1,f_2,f_3)|f_1^2+f_2^2+f_3^2 = 1, f_i \in [0,1], i = 1,2,3\}$.

VIII. Liu and Li's Recombination Operator

A. Crossover

Suppose x^i and x^j are mating parents, a child y can be generated by

$$y = x^i + rc(x^i - x^j) \quad (1)$$

with

$$rc = (2rnd-1)(1-rnd^{-(1-g/Mg)^{0.7}}) \quad (2)$$

The authors are with the Centre for Computational Intelligence (CCI), School of Computer Science and Informatics, De Montfort University, The Gateway, Leicester LE1 9BH, U.K. (email: shouyong.jiang@email.dmu.ac.uk, syang@dmu.ac.uk).

TABLE VII
POPULATION SIZE SETTINGS FOR SLD AND K-LAYER FOR DIFFERENT NUMBERS OF OBJECTIVES

M	SLD				k-layer			
3	45	105	231	528	43	106	232	511
4	56	120	286	560	57	109	261	541
5	70	126	210	495	71	101	221	451
6	56	126	252	462	55	121	211	463
7	28	84	210	462	36	99	190	456
8	36	120	240	660	41	113	281	521
9	45	165	210	495	46	127	244	487
10	55	110	230	440	51	141	201	441
15	30	135	240	680	31	136	211	526

TABLE VIII
CENTRED L_2-DISCREPANCY VALUES OF SLD AND K-LAYER FOR DIFFERENT POPULATION SIZES

M	SLD				k-layer			
3	1.13E-01	**1.11E-01**	**1.10E-01**	**1.11E-01**	**1.12E-01**	1.12E-01	1.12E-01	1.12E-01
4	4.87E-01	5.17E-01	5.52E-01	5.89E-01	**4.20E-01**	**4.14E-01**	**4.14E-01**	**4.14E-01**
5	1.18E+00	1.28E+00	1.32E+00	1.43E+00	**9.93E-01**	**9.79E-01**	**9.78E-01**	**9.78E-01**
6	2.39E+00	2.44E+00	2.70E+00	2.80E+00	**1.92E+00**	**1.92E+00**	**1.93E+00**	**1.93E+00**
7	3.78E+00	4.43E+00	4.58E+00	5.05E+00	**3.61E+00**	**3.50E+00**	**3.47E+00**	**3.45E+00**
8	6.46E+00	7.72E+00	**5.46E+00**	6.21E+00	**6.09E+00**	**5.92E+00**	5.84E+00	**5.86E+00**
9	1.26E+01	1.27E+01	1.08E+01	1.43E+01	**9.91E+00**	**9.67E+00**	**9.61E+00**	**9.55E+00**
10	1.98E+01	1.45E+01	1.89E+01	**1.36E+01**	**1.58E+01**	**1.54E+01**	**1.52E+01**	1.53E+01
15	1.58E+02	1.62E+02	**1.31E+02**	1.61E+02	**1.39E+02**	**1.32E+02**	1.33E+02	**1.32E+02**

where rnd is a random number in [0,1]. g and Mg are the current generation and the maximum number of generations, respectively. After that, y is repaired by

$$y_k = \begin{cases} y_k & \text{if } a_k \leq \bar{y_k} \leq b_k \\ a_k + \frac{rnd}{2}(y_k - a_k) & \text{if } \bar{y_k} < a_k \\ b_k + \frac{rnd}{2}(b_k - a_k) & \text{if } \bar{y_k} > b_k \end{cases} \tag{3}$$

where y_k is the k-th component of y, and a_k and b_k are the lower and upper bounds of the k-th variable.

B. Mutation

Each component of y is mutated by

$$y_k = y_k + rm(b_k - a_k) \tag{4}$$

with

$$rm = 0.5(rnd - 0.5)(1 - rnd^{-(1-g/Mg)^{0.7}}) \tag{5}$$

After mutation, y is repaired by Eq. (3).

IX. DISCUSSIONS ON REFERENCE DIRECTION SET

A. Uniformity

Unlike the simplex-lattice design (SLD) method used in MOEA/D and NSGA-III, our k-layer method generates reference direction set from the subsimplex's point of view. As a result, the k-layer method can give more options for setting population size when the number of objectives is very large. Here, we investigate the uniformity of reference direction set generated by different methods. Since the coupling between population size and the number of objectives (M) exists in both SLD and the k-layer method, it is very unlikely to generate the same population size using these two methods. A feasible way to compare the uniformity between SLD and the k-layer method is to generate similar population size. Also, we consider four levels of population size, that is, population size approximately equals 50, 100, 250, and 500, which is shown in Table VII. Note that, SLD can generate each considered population size of points by choosing different settings for the two-layered method of NSGA-III. The most widely used discrepancy measure, i.e., centred L_2-discrepancy (CD_2) is adopted to evaluate the uniformity of a reference direction set, and the smaller CD_2 is, the more uniform a reference direction set will be.

Table VIII presents the CD_2 values of SLD and k-layers on different population sizes, where the best value for each population size is marked in boldface. It can be observed that, in the 3-objective case, SLD has slightly better distribution than the k-layer method, whereas for $M > 3$, the k-layer method gives much better CD_2 values than SLD in most cases. It is understandable that the k-layer method generates more uniform reference direction set in higher-dimension cases, because it decomposes the whole simplex into M subsimplexes, and each subsimplex can get a number of uniformly-distributed points. In contrast, SLD use a two-layered structure to distribute points in high-dimensional cases. It generates points only on the layers, and cannot generate intermediate or near intermediate points on the simplex. So, the distribution of SLD is not very uniform in high-dimensional cases.

The k-layer method is just an example of using subsimplexes to design the reference direction set. The whole simplex consists of M subsimplexes for the M-objective case, and each subsimplex is a two-dimensional simplex. So, it is easier to distributing points on these subsimplexes than on the whole M-dimensional simplex. However, it is far from being satisfying because there is still a coupling between the population size and the number of objectives. Subsimplex-based design

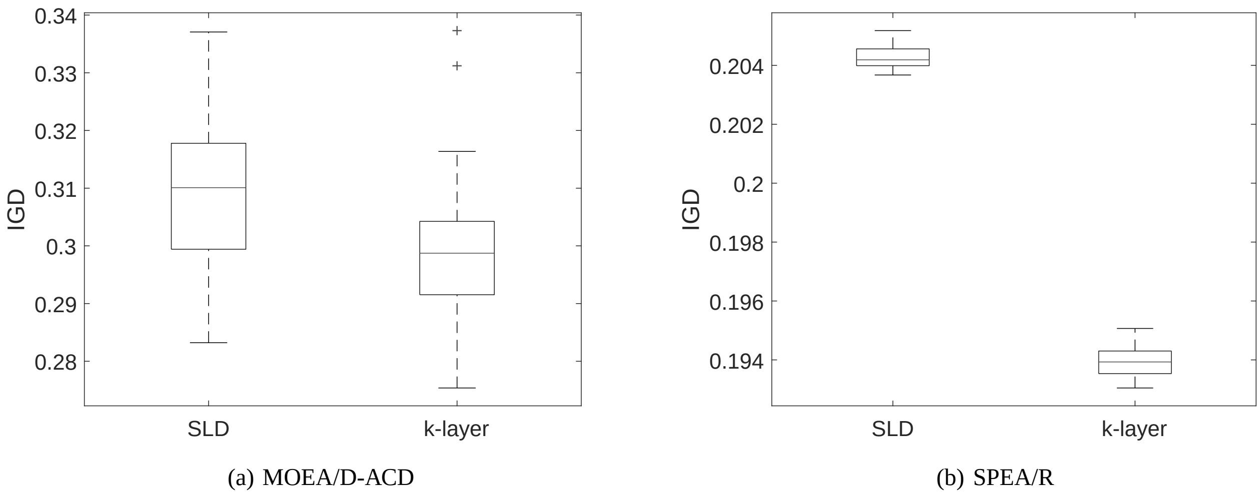


Fig. 12. Boxplots of the IGD results obtained by algorithms using the SLD and k-layer approaches for the 3-objective WFG4 instance.

methods are promising for reducing the coupling, and more work is required to generate an arbitrary population size free from the influence of the number of objectives. It should be noted that a uniformly-distributed reference direction set does not necessarily produce a uniform distribution of solutions on the Pareto-optimal front (POF). This is because the task to uniform a reference set on the simplex is not equivalent to that to well distribute solutions on the POF that has various geometries.

B. *Influence on Search Performance*

Next, we analyze the search performance of algorithms with different reference direction generation methods, i.e., the SLD method and the k-layer method. The two methods are tested on MOEA/D-ACD and SPEA/R. Since both SLD and k-layer methods have a close coupling with population size, it is desirable to choose similar population sizes to make a fair comparison between these two methods. For this reason, we use a population size of 105 and 106 for SLD and k-layer, respectively, in 3-objective cases, which has been shown in Table VII.

The influence of SLD and k-layer is examined on 3-objective WFG4, and the IGD results obtained MOEA/D-ACD and SPEA/R are presented in Fig. 12. Note that, 5050 uniformly-sampled points from the true POF of the 3-objective WFG4 are used to calculate the IGD metric. Two interesting observations can be obtained from the figure. First, SPEA/R outperforms MOEA/D-ACD in terms of IGD, regardless of the reference direction generation methods. Second, for both algorithms, the k-layer method helps yield better IGD values than the SLD method. Thus, the proposed reference direction generation method is very effective to guide the search toward the 3-dimensional POF.

Afterwards, we examine the effectiveness of the k-layer method on high-dimensional cases. To do this, we compare the k-layer method with SLD on 8-objective WFG4, and the population size of SPEA/R is set to 161. Note that, here SLD represents the two-layered approach proposed in NSGA-III. Fig. 13 presents boxplots of the IGD values obtained SPEA/R with different reference direction generation methods, showing the k-layer method help achieve better IGD values than SLD. Additionally, the normalized parallel coordinate plot of SPEA/D with the two different methods is displayed in Fig. 14. We can observe from the figure that both methods are able to provide a set of solutions residing in the entire POF. While the k-layer method covers well the POF, SLD misses a small part of the POF, e.g., there is no solution in the region where $f_i \in [0.1, 0.4]$ for all i. Therefore, the proposed reference direction generation method can guide the search well in high-dimensional cases.

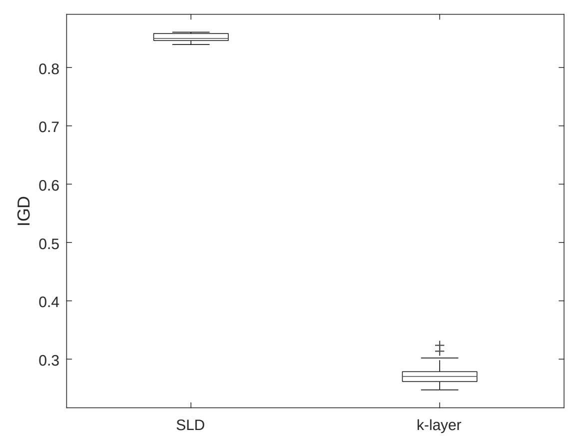
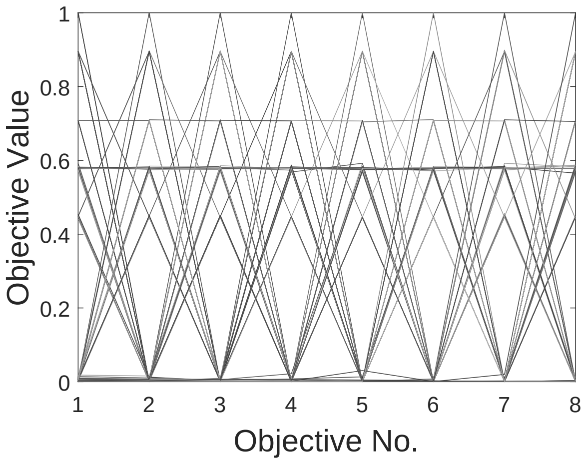


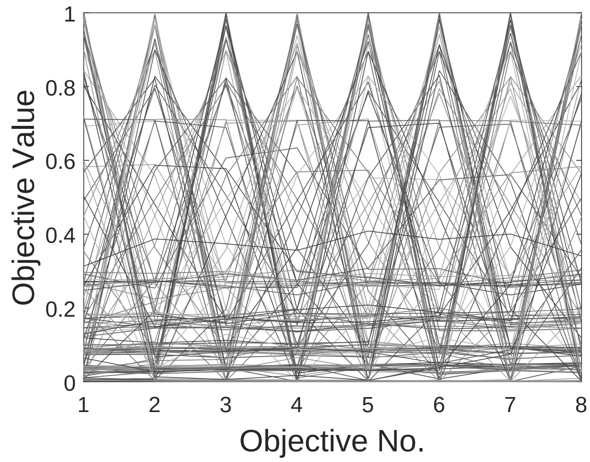
Fig. 13. Boxplots of the IGD results obtained by SPEA/R using the SLD and k-layer approaches for the 8-objective WFG4 instance.

X. Supplementary Results on The MOP Test Suite

To study the evolution difference among SPEA/R, MOEA/D-ACD and MOEA/D-M2M, we provide the box-plot of the obtained HV values after different number of generations for each test MOP problem, which is shown in Fig. 15. It is clear to see that the HV value roughly increases monotonically with the number of generations while the variance of the HV decreases. In MOP1, MOEA/D-ACD



(a) SLD



(b) k-layer

Fig. 14. Parallel coordinates of final solutions obtained by different reference direction generation methods for the 8-objective WFG4 instance.

performs better than the others at the early stage (before the 1000th generation), but the difference between the algorithms becomes small when the generation proceeds. In MOP2 and MOP4, SPEA/R performs the best after 500 generations, as indicated by the variance of HV. The three algorithms perform similarly on MOP5, MOP6, and MOP7, but SPEA/R seems to have less outliers than MOEA/D-M2M and MOEA/D-ACD during the evolution, particularly on MOP6 and MOP7. The boxplots of the HV clearly show that SPEA/R can evolve in a good behavior and achieve competitive performance against the other algorithms.

XI. Supplementary Results on The Influence of Fitness Assignment and Niche Preservation

Table IX provides some supplementary results of four SPEA/R variants for the WFG and MOP test suites in terms of the HV metric. Note that, the reported HV values for the WFG instances have been normalized whereas those for the MOP instances do not. The reference point for the computation of HV remains the same as in Section III and Section IV for the WFG and MOP test suites, respectively. The Wilcoxon signed-rank test at the 0.05 significance level with Bonferroni correction is used to clarify the difference between the original SPEA/R and the other compared variants.

We can observe from the table that, generally, SPEA/R obtains the best HV values on the majority of the 52 test instances. Specifically, SPEA/R outperforms SPEA/R-A, SPEA/R-B, and SPEA/R-C in 12, 4, and 51 out of 52 instances, respectively. Compared with SPEA/R, SPEA/R-A loses on some instances, indicating that the use of global fitness is helpful. In most cases, SPEA/R performs similarly to SPEA/R-B. This means that the use of local fitness does not significantly enhance the final performance of SPEA/R. On the other hand, there is a significant difference between SPEA/R and SPEA/R-C. This shows that the lack of proper diversity maintenance is detrimental to the algorithm's performance.

The comparison between SPEA/R and the other variants clearly demonstrates that the proposed diversity-first-and-convergence-second environmental selection strategy contributes most to solving various problems considered in this paper. But, this does not mean that fitness assignment is not important. The outperformance of SPEA/R over SPEA/R-A illustrates that proper use of fitness assignment can boost algorithms' performance. Therefore, we can conclude that both diversity and convergence are of great importance to SPEA/R, but diversity outweighs convergence in our study.

XII. Solution Representation for Many-objective Problems

To further show the performance of different algorithms on many-objective optimization, we graphically plot the parallel coordinates of final solutions obtained by each algorithm for 12-objective WFG5 and WFG6 in Figs. 16 and 17, respectively. For WFG5, SPEA2+SDE and SPEA/R are the best performers in terms of diversity, but the former does not cover as well as the latter. For WFG6, NSGA-III and SPEA/R perform similarly, and SPEA2+SDE also achieves very diverse solutions, but its coverage is again far from being satisfying. The figures clearly show that SPEA/R is very capable of balancing diversity and convergence for many-objective optimization.

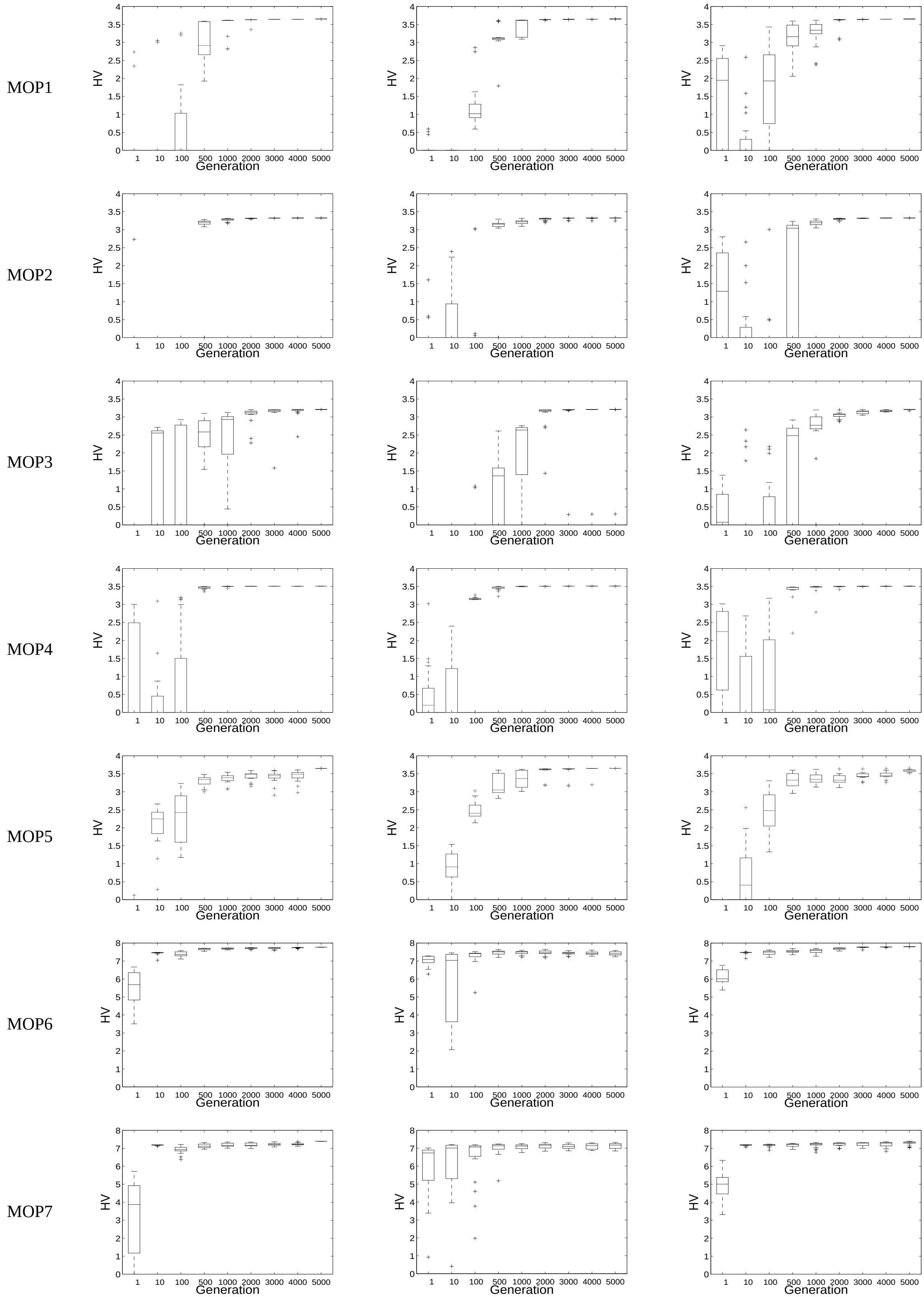
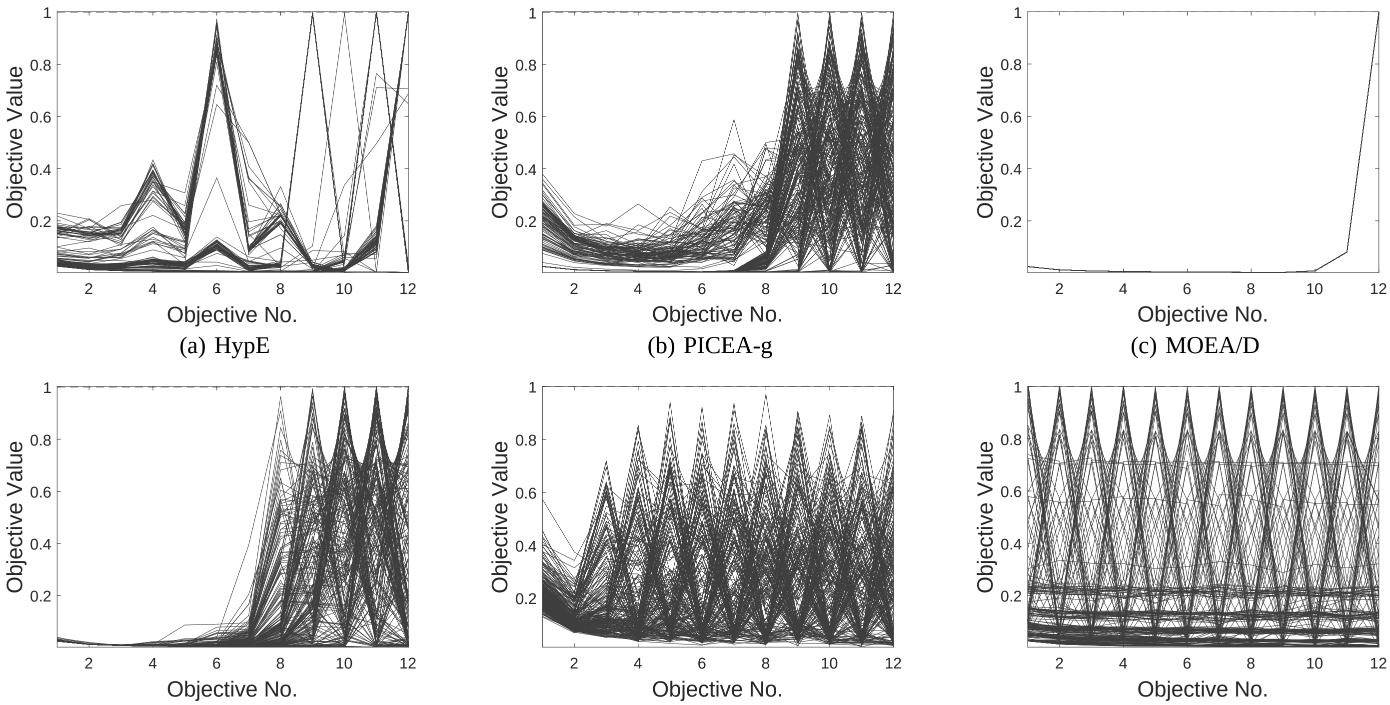


Fig. 15. Boxplots of obtained HV values at different search stages. Left column: SPEA/R; middle column: MOEA/D-M2M; right column: MOEA/D-ACD.

TABLE IX
MEAN AND STAND DEVIATION HV VALUES OBTAINED BY FOUR SPEA/R VARIANTS ON A TOTAL OF 52 TEST INSTANCES

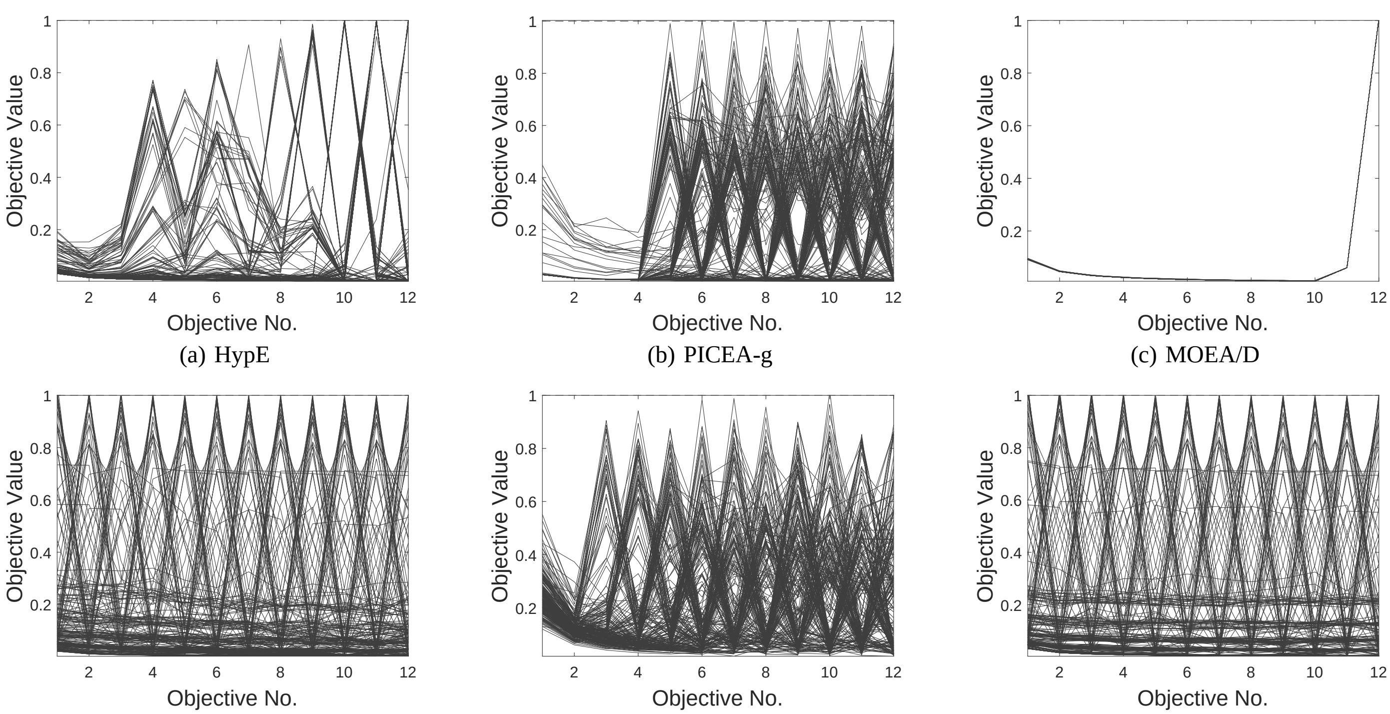
Prob.	M	SPEA/R	SPEA/R-A	SPEA/R-B	SPEA/R-C
WFG1	2	6.0736E-1(5.9395E-3)	6.0544E-1(3.9530E-3)†	**6.0978E-1(5.3469E-3)**†	5.0624E-1(3.1928E-2)‡
	3	6.2420E-1(6.0584E-3)	6.2393E-1(6.2900E-3)†	**6.2529E-1(4.8780E-3)**†	3.7656E-1(7.3531E-2)‡
	5	5.7919E-1(5.4070E-3)	5.7709E-1(4.4688E-3)‡	**5.8382E-1(1.2427E-2)**	4.0328E-1(8.2468E-2)‡
	8	6.2128E-1(7.8831E-2)	5.9634E-1(9.2591E-2)‡	**5.1144E-1(8.2269E-2)**†	3.3294E-1(7.1250E-2)‡
	12	4.2884E-1(5.9670E-2)	3.3142E-1(5.9896E-2)†	3.1374E-1(1.2585E-2)†	3.0224E-1(6.4692E-2)‡
WFG2	2	8.3092E-1(2.9447E-2)	**8.3745E-1(2.4617E-2)**†	8.3397E-1(2.7915E-2)†	6.4052E-1(1.3187E-1)‡
	3	9.4106E-1(6.0781E-2)	9.4029E-1(6.1166E-2)†	**9.4759E-1(5.6287E-2)**†	3.9638E-1(1.4135E-1)‡
	5	9.8674E-1(3.7587E-2)	9.4430E-1(7.9392E-2)‡	**9.8694E-1(3.7738E-2)**†	3.0529E-1(9.9552E-2)‡
	8	**9.9759E-1(6.6128E-4)**	9.8321E-1(7.8581E-2)‡	9.9127E-1(9.0229E-2)†	2.9570E-1(1.0492E-1)†
	12	**9.8832E-1(9.2631E-2)**	8.8805E-1(9.2739E-2)†	9.8023E-1(5.6782E-2)‡	2.8799E-1(1.8007E-1)‡
WFG3	2	**8.2945E-1(1.3184E-3)**	8.2688E-1(1.3610E-3)†	8.2767E-1(1.6579E-3)†	6.2481E-1(1.2766E-1)‡
	3	**7.7947E-1(8.0418E-3)**	7.7253E-1(8.6399E-3)†	7.7651E-1(1.0792E-2)†	3.5801E-1(1.1030E-1)‡
	5	**6.8696E-1(1.1909E-1)**	3.3371E-1(1.1242E-1)†	3.3844E-1(9.9643E-2)†	2.6090E-1(9.0781E-2)‡
	8	4.9757E-1(4.4357E-2)	**2.9351E-1(5.1645E-2)**†	2.6538E-1(6.2345E-2)†	2.1330E-1(3.4329E-2)‡
	12	**4.8913E-1(4.4723E-2)**	4.7754E-1(4.6565E-2)†	4.7406E-1(5.4434E-2)†	2.0211E-1(7.7641E-3)‡
WFG4	2	7.3497E-1(6.0901E-4)	7.3488E-1(7.1729E-4)†	**7.3517E-1(4.5024E-4)**†	6.8561E-1(1.6689E-2)‡
	3	**8.5842E-1(9.0221E-4)**	8.4828E-1(7.1718E-4)†	8.4854E-1(5.9400E-4)†	2.6268E-1(1.6449E-2)‡
	5	9.2169E-1(1.2998E-3)	9.2103E-1(9.0865E-4)†	**9.2190E-1(1.0292E-3)**†	1.6953E-1(7.4838E-3)‡
	8	**9.5797E-1(3.9059E-3)**	9.5641E-1(4.6414E-3)†	9.5349E-1(1.5037E-2)†	1.1496E-1(9.1398E-3)‡
	12	**9.1209E-1(7.8544E-2)**	8.7905E-1(1.1934E-1)‡	9.0516E-1(3.2617E-2)†	7.7403E-2(2.4276E-3)‡
WFG5	2	**7.1227E-1(9.3626E-4)**	7.0231E-1(1.3487E-3)†	7.0263E-1(7.5781E-4)†	3.4370E-1(6.8752E-2)‡
	3	**8.2437E-1(2.0573E-3)**	8.1446E-1(1.8781E-3)†	8.1405E-1(1.5404E-3)†	2.8645E-1(5.6845E-2)‡
	5	**8.8792E-1(1.1292E-3)**	8.7773E-1(1.6310E-3)†	8.7815E-1(1.6021E-3)†	1.5839E-1(1.0153E-2)‡
	8	9.0023E-1(2.9488E-3)	**9.0208E-1(3.1050E-3)**†	9.0094E-1(3.5224E-3)†	1.1891E-1(1.4041E-2)‡
	12	**9.0236E-1(6.6030E-3)**	9.0210E-1(8.5838E-3)†	8.9487E-1(2.6142E-2)†	8.3108E-2(1.6591E-2)‡
WFG6	2	7.1001E-1(6.5120E-3)	7.1076E-1(5.5471E-3)†	**7.1264E-1(6.1166E-3)**†	6.3429E-1(4.0134E-2)‡
	3	**8.2552E-1(6.3065E-3)**	8.1690E-1(5.5111E-3)†	8.1921E-1(7.6680E-3)†	2.8573E-1(6.0441E-2)‡
	5	**8.9930E-1(7.7577E-3)**	8.7553E-1(1.0191E-2)‡	8.7243E-1(1.1251E-2)†	1.9397E-1(3.3503E-2)‡
	8	**9.2855E-1(1.5045E-2)**	9.0496E-1(1.2911E-2)†	8.9911E-1(1.1880E-2)‡	1.1849E-1(1.3896E-2)‡
	12	**9.1604E-1(1.2921E-2)**	9.1268E-1(1.3382E-2)†	9.0929E-1(1.8871E-2)‡	4.9899E-2(1.5898E-2)‡
WFG7	2	**7.3615E-1(1.1625E-3)**	7.3509E-1(8.5502E-4)‡	7.3520E-1(8.2879E-4)‡	5.0023E-1(1.3808E-1)‡
	3	**8.5171E-1(3.4081E-4)**	8.5084E-1(3.2037E-4)†	8.5097E-1(4.1257E-4)†	2.6602E-1(2.0120E-2)‡
	5	9.2528E-1(6.1088E-4)	9.2461E-1(5.0415E-4)†	9.2505E-1(7.8590E-4)†	1.7514E-1(1.3465E-2)‡
	8	**9.5110E-1(2.5113E-2)**	9.1126E-1(3.7057E-2)‡	9.4247E-1(3.9013E-2)†	1.0843E-1(1.8472E-2)‡
	12	9.4407E-1(9.1515E-3)	9.4246E-1(1.2369E-1)†	**9.4488E-1(2.5453E-2)**†	8.7954E-2(2.4103E-2)‡
WFG8	2	**7.0318E-1(3.3615E-3)**	7.0184E-1(3.4096E-3)†	7.0316E-1(3.1633E-3)†	6.1993E-1(4.5839E-2)‡
	3	8.0651E-1(5.3416E-3)	**8.0749E-1(6.1788E-3)**†	8.0554E-1(4.5734E-3)†	3.2625E-1(5.2069E-2)‡
	5	**8.5983E-1(1.8504E-2)**	8.2695E-1(1.1548E-2)†	8.3747E-1(2.4503E-2)†	1.8475E-1(1.3918E-2)‡
	8	**8.8805E-1(3.8120E-2)**	8.8089E-1(4.5522E-2)†	8.8219E-1(4.3581E-2)†	1.1907E-1(1.6269E-2)‡
	12	**8.7084E-1(2.2243E-1)**	8.6295E-1(2.3946E-1)†	8.6963E-1(2.5606E-1)†	6.4869E-2(4.8458E-1)‡
WFG9	2	**7.1499E-1(1.4646E-3)**	7.1358E-1(9.7322E-4)‡	7.1404E-1(1.1510E-3)†	5.5476E-1(1.4902E-1)‡
	3	**8.2240E-1(2.2316E-2)**	8.0276E-1(1.7211E-2)†	8.0298E-1(1.7383E-2)†	2.7150E-1(9.4910E-2)‡
	5	**8.6833E-1(4.2512E-2)**	7.7071E-1(3.7155E-2)†	7.9251E-1(3.6113E-2)†	1.8923E-1(8.1110E-2)‡
	8	**8.1250E-1(4.0473E-2)**	7.3161E-1(4.0893E-2)†	7.3690E-1(3.6129E-2)†	1.0369E-1(2.5244E-2)‡
	12	**7.8937E-1(4.7711E-2)**	7.2725E-1(3.3145E-2)‡	7.8130E-1(4.0693E-2)†	8.0192E-2(2.2657E-3)‡
MOP1	2	3.6546E+0(2.0423E-4)	3.6547E+0(1.9115E-4)†	3.6549E+0(2.4230E-4)†	2.0087E+0(1.1919E-2)‡
MOP2	2	**3.3264E+0(5.3229E-5)**	3.3264E+0(6.1860E-5)†	3.3264E+0(1.0981E-4)†	2.0000E+0(1.0649E-1)‡
MOP3	2	3.2081E+0(1.1321E-4)	3.2081E+0(1.4979E-4)†	**3.2081E+0(7.6603E-5)**†	2.0864E+0(2.1585E-1)‡
MOP4	2	**3.5109E+0(6.4177E-4)**	3.5082E+0(2.4014E-3)†	3.5089E+0(2.7698E-3)†	2.0091E+0(1.4614E-2)‡
MOP5	2	**3.6464E+0(8.5389E-4)**	3.6448E+0(9.9301E-4)‡	3.6460E+0(9.3730E-4)†	4.2000E-1(5.5782E-1)‡
MOP6	3	7.7792E+0(7.6112E-3)	7.7700E+0(6.6929E-3)‡	**7.7793E+0(5.4549E-3)**†	4.0004E+0(6.8309E-4)‡
MOP7	3	**7.3932E+0(7.0526E-3)**	7.3930E+0(5.3230E-3)†	7.3930E+0(7.4332E-3)†	4.0000E+0(0.0000E+0)‡

‡ and † indicate SPEA/R performs significantly better than and equivalently to the corresponding algorithm, respectively.



(a) HypE (b) PICEA-g (c) MOEA/D
(d) NSGA-III (e) SPEA2+SDE (f) SPEA/R

Fig. 16. Parallel coordinates of final solutions obtained by six algorithms for the 12-objective WFG5 instance.



(a) HypE (b) PICEA-g (c) MOEA/D
(d) NSGA-III (e) SPEA2+SDE (f) SPEA/R

Fig. 17. Parallel coordinates of final solutions obtained by six algorithms for the 12-objective WFG6 instance.