IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received March 26, 2018, accepted April 18, 2018, date of publication May 7, 2018, date of current version June 5, 2018.

Digital Object Identifier 10.1109/ACCESS.2018.2832181

Evolutionary Many-Objective Optimization: A Comparative Study of the State-of-the-Art

KAIWEN LI[1], RUI WANG[1], TAO ZHANG[1], AND HISAO ISHIBUCHI[2], (Fellow, IEEE)
[1]College of System Engineering, National University of Defense Technology, Changsha 410073, China
[2]Department of Computer Science and Engineering, Southern University of Science and Technology, Shenzhen 518055, China

Corresponding author: Rui Wang (ruiwangnudt@gmail.com)

This work was supported in part by the National Natural Science Foundation of China under Grant 61403404, Grant 61773390, Grant 71571187, and Grant 71371181, and in part by the Distinguished Natural Science Foundation of Hunan Province under Grant 2017JJ1001.

ABSTRACT With the increasing attention paid to many-objective optimization in the evolutionary multi-objective optimization community, various approaches have been proposed to solve many-objective problems. However, existing experimental comparative studies are usually restricted to a few methods. Few studies have encompassed most of the recently proposed state-of-the-art approaches and made an experimental comparison. To this end, this paper offers a systematic comparison of 13 algorithms covering various categories to solve many-objective problems. The experimental comparison is conducted on three groups of test functions by using two performance metrics and a visual observation in the decision space. The experimental results demonstrate that different approaches have different search abilities. None of the test approaches outperform the others on all types of problems. However, some of the approaches are competitive on a large number of test problems. Moreover, inconsistent results from the hypervolume and the inverted generational distance metrics are revealed in this paper. Based on these comparative results, researchers can obtain useful suggestions for choosing appropriate algorithms for different problems.

INDEX TERMS Evolutionary computation, experimental comparison, HV, IGD, multi-objective optimization.

I. INTRODUCTION

Multi-objective optimization problems (MOPs) frequently appear in many fields such as engineering [1], business [2], mathematics [3] and physics [4], when two or more potentially conflicting objectives are required to be minimized simultaneously. Problems with multiple objectives are much common in real world, such as minimizing cost while maximizing comfort while buying a car. Advanced methods are especially required for solving the more complicated problems in real world application. For instance, a unit commitment problem [5] can be solved as a multi-objective optimization problem considering minimizing cost and emission as the multiple objectives. In medical aid distribution route problems [6], the travel time and priorities for the shelters with higher demand are both supposed to be considered simultaneously. Moreover, in financial field, a central bank must choose a stance for monetary policy that balances competing objectives: low inflation, low unemployment and low balance of trade deficit. Overall, there are a number of problems that are characterized by more than one goal and the trade-offs are required in order to satisfy the different objectives. To model such problems, a MOP can be formulated as follows:

$$\min_{\mathbf{x}} \ \mathbf{f}(\mathbf{x}) = (f_1(\mathbf{x}), f_2(\mathbf{x}), \ldots, f_M(\mathbf{x}))$$
$$\mathbf{s.t.} \ \mathbf{x} \in X$$

where $\mathbf{f}(\mathbf{x})$ is an objective function vector that consists of M objective functions and $X \subseteq R_D$ is the decision space. $\mathbf{x} = (x_1, x_2, \ldots, x_D) \in X$ is the decision vector with D dimensions.

Since optimizing one objective often leads to a deterioration in at least one other objective, a set of trade-off solutions, termed Pareto optimal solutions, are expected to be found for MOPs. A solution is called nondominated or Pareto optimal, if none of the objective functions can be improved in value without degrading some of the other objective values. All Pareto optimal solutions are considered equally good if there is no additional subjective preference information. More formally, let $\mathbf{x}^1, \mathbf{x}^2 \in X$; $\mathbf{x}^1$ is said to dominate $\mathbf{x}^2$ (termed $\mathbf{x}^1 \prec \mathbf{x}^2$) if and only if $f_i(\mathbf{x}^1) \leq f_i(\mathbf{x}^2)$ for all $i \in 1, \ldots, M$ and

$f_i(\mathbf{x}^1) < f_i(\mathbf{x}^2)$ for at least one index i. Here, a point $\mathbf{x}^* \in X$ is a Pareto optimal solution if there is no $\mathbf{x} \in X$ such that $\mathbf{x} \prec \mathbf{x}^*$. $\mathbf{f}(\mathbf{x}^*)$ is then called a Pareto optimal (objective) vector. The set of all Pareto optimal points in X is called the Pareto set (PS), and the set of all the Pareto optimal objective vectors is the Pareto front (PF).

Over the past two decades, evolutionary multi-objective optimization (EMO) has been demonstrated to be an efficient approach for solving MOPs that can obtain a set of solutions in a single run. A number of EMO algorithms have been proposed to solve MOPs, e.g., NSGA-II [7], SPEA2 [8], IBEA [9], and MOEA/D [10]. Among these EMO algorithms, Pareto-based approaches are the most popular class of approaches in the EMO community. In these approaches, solutions with a better Pareto rank are selected according to what is known as the dominance-based selection criterion. To achieve a diversity of solutions, a secondary diversity-related criterion is also often adopted. NSGA-II and SPEA2 are two representative Pareto-based approaches. Such approaches have shown very effective performance in tackling MOPs with two or three objectives. However, for MOPs with more than three objectives, known as many-objective optimization problems (MaOPs), the efficiency of Pareto-based approaches decreases significantly.

In real-world applications, such as water distribution systems [11] and land use management problems [12], there often exist MaOPs that involve more than three objectives and even reach 10 to 15 objectives. Strategies for solving MaOPs have drawn increasing attention in recent years because conventional EMO algorithms developed for solving MOPs with two or three objectives cannot tackle MaOPs efficiently. The main reason for this failure is that the percentage of non-dominated solutions increases dramatically as the number of objectives increases, and thus, the dominance-based selection criterion is unable to distinguish the generated solutions in a population. When the primary criterion fails to provide sufficient selection pressure to discriminate solutions, the diversity-related secondary criterion is activated to select solutions, which is known as the Active Diversity Promotion (ADP) phenomenon [13]. Many experimental observations have shown that the ADP phenomenon may cause poor convergence of the final solutions. As a result, the final set of solutions may present good diversity over the objective space without converging to the Pareto front. Moreover, the calculation of performance metrics, such as the hypervolume measure, may be too computationally expensive when dealing with MaOPs.

To enhance the performance of EMO algorithms to tackle MaOPs effectively, a variety of approaches have been developed in recent years. These approaches can be roughly divided into five categories.

As the Pareto-based approaches fail to provide sufficient selection pressure towards the PF when tackling MaOPs, the first category covers various non-Pareto-based approaches, such as decomposition-based and indicator-based approaches. The decomposition-based type of methods aggregate the objectives into a scalar function, and each objective has a weighting coefficient that indicates its importance, known as the weight vector. By employing a set of weighting vectors, an MaOP is decomposed into a number of single-objective sub-problems that can be optimized simultaneously. The multiobjective evolutionary algorithm based on decomposition (MOEA/D) and multiple single objective Pareto sampling (MSOPS) [14] are two representative algorithms. In MOEA/D, the fitness of each solution is evaluated by a unique weight vector, whereas each solution in MSOPS is evaluated by a set of weight vectors. In [15], an improved version of MSOPS, termed MSOPS-II, is proposed with two comprehensive extensions to MSOPS. Moreover, there is another type of decomposition-based approach in which an MOP is decomposed into a set of sub-MOPs, such as MOEA/D-M2M [16]. A reference vector-guided evolutionary algorithm (RVEA) [17] belongs to this category of methods. In RVEA, the objective space is partitioned into a number of subspaces using the reference vectors, which helps balance the convergence and diversity. Other algorithms (e.g., NSGA-III [18] and MOEA/DD [19]) can also be viewed as decomposition-based approaches that use hybrid mechanisms of the Pareto dominance-based fitness evaluation and the MOEA/D framework.

The indicator-based approaches use the value of the performance indicator to guide the search process, e.g., the indicator based evolutionary algorithm (IBEA) [9], the S-metric selection based evolutionary multi-objective algorithm (SMS-EMOA) [20], and the fast hypervolume based evolutionary algorithm (HypE) [21]. IBEA was the first indicator-based approach to solve MaOPs using a predefined optimization goal to measure the contribution of each solution. However, IBEA has bad performance with respect to diversity due to the lack of diversity maintenance in its indicator. As the hypervolume (HV) value can balance both convergence and diversity effectively, SMS-EMOA was developed based on the HV value. Moreover, to reduce the large computational cost of calculating the HV values, HypE was developed by applying Monte Carlo simulation to approximate the exact HV values.

Due to the ineffectiveness of the Pareto dominance for solving MaOPs, the second category of approaches modifies the traditional dominance relation to enhance the selection pressure towards the PF. Many approaches have been developed, such as ε-dominance [22], [23], L-optimality [24], fuzzy dominance [25], and preference order ranking [26]. Yang *et al.* [27] developed a grid-based evolutionary algorithm (GrEA) in which a grid dominance is introduced to strengthen the selection pressure toward the optimal direction. A new dominance relation, termed θ dominance, is introduced for many-objective optimization in [28]. Based on the θ dominance, an effective θ dominance-based evolutionary algorithm (θ-DEA) is proposed to tackle MaOPs effectively.

Although extensive work has focused on enhancing the selection pressure towards PF, the third category of methods aims to modify the secondary diversity-related criterion to maintain a good balance between convergence and diversity for tackling MaOPs. Adra and Fleming [29] developed a diversity management mechanism named DM1 to determine whether or not to activate the diversity promotion according to the convergence condition. In GrEA, in addition to introducing a grid-dominance to enhance the selection pressure, three grid-based criteria (i.e., grid ranking, grid crowding distance, and grid coordinate point distance) are applied to maintain the diversity of solutions. Moreover, Li *et al.* [30] propose a general modification of the diversity-related criterion for dominance-based EMO algorithms, termed the shift-based density estimation (SDE) strategy, to reduce the detrimental impact caused by the ADP phenomenon. In the recently proposed knee point-driven evolutionary algorithm (KnEA) [31], a knee point-based secondary selection is designed in addition to the dominance-based selection criterion to enhance the convergence pressure. In addition, NSGA-III replaces the crowding distance-based diversity-maintenance operator in NSGA-II with a reference point-based strategy to solve MaOPs effectively.

As the objective space dimensions in MaOPs increase, the population size should increase exponentially to approximate the entire PF [32]. However, the population size in real-world MaOPs is often too limited compared to the large objective space [33]. Thus, it is rational to focus on a subset of the PF according to the user's preference. This category of approaches is known as the preference-based approach and mainly includes three classes: *a priori* algorithms [34], [35], interactive algorithms [36]–[38], and *a posteriori* algorithms [39]–[42]. In *a priori* algorithms, the preference information is specified before the search. In interactive algorithms, the decision maker is required to provide preference information interactively. Similarly, in *a posteriori* algorithms, the preference information is introduced after the search. PICEA-g [41], which is a highly cited *a posteriori* algorithm, harnesses the power of coevolution of candidate solutions and preferences for the purposes of optimization. It achieves competitive performance as the number of objectives increases.

A few other types of recently proposed approaches also show competitive performance in maintaining a good balance between convergence and diversity for tackling MaOPs. For example, the recently proposed MOEA/DD algorithm combines the advantages of both the dominance and decomposition-based approaches to achieve satisfactory performance for MaOPs. In [43] and [44], a novel two-archive algorithm (TAA) and its improved version (Two_Arch2) are proposed. These algorithms apply two archives focusing on convergence and diversity separately. Most recently, other advanced methods have been proposed, such as MnRP-BILDE [45], which was developed based on the concept of objective function space reduction. Jiang and Yang [46] propose a new SPEA based on the reference direction, denoted SPEA/R, for both multiobjective and many-objective optimization. Moreover, RPEA [47] was proposed to exploit the potential of the reference points-based approach. To solve unconstrained many-objective problems, a vector angle-based evolutionary algorithm [48] was developed.

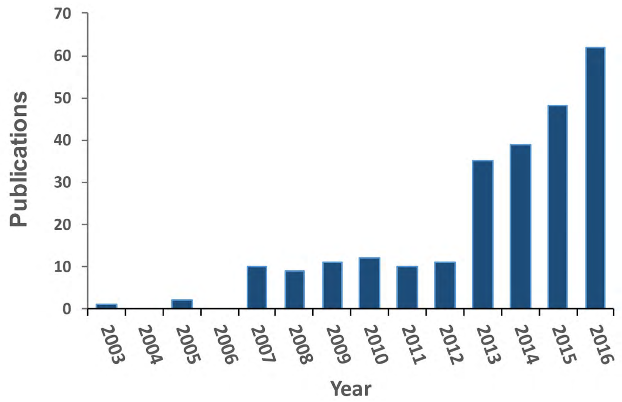

FIGURE 1. Number of publications about many-objective optimization in the Web of Science Core Collection from 2003 to 2016.

Overall, a variety of algorithms have been proposed to handle many-objective optimization problems. However, questions remain, including which algorithm is suitable for a specific type of problems and if any algorithm has a clear advantage over the others. To this end, Li *et al.* [49] compared the performance of eight different representative algorithms with a total of eleven test functions in 2013. However, studies on solving MaOPs have increased greatly since 2013. Fig. 1 shows the research trends of many-objective optimization, as indicated by the number of publications in this field from 2003 to 2016 in the Core Collection of Web of Science. Many approaches have been proposed in the most recent four years. In 2016, Ma *et al.* [50] conducted a comparative study only on decomposition-based approaches for many-objective optimization. In addition, a large number of state-of-the-art many-objective algorithms (21 algorithms in total) are examined in [51]. However, only WFG1-9 test problems are used for performance comparison. Thus, this work provides a systematic comparison of 13 algorithms covering different popular categories of approaches and recently proposed novel algorithms and applies three groups of test functions (17 test functions in total) to evaluate the above algorithms.

II. THE THIRTEEN INVESTIGATED ALGORITHMS

In this part, the 13 algorithms investigated in this paper are introduced. These algorithms are introduced in chronological order.

- **MSOPS-II** is an improved version of MSOPS for many-objective optimization that uses a set of target vectors to direct and control the search process based on the chosen aggregation method. MSOPS-II applies two extensions to MSOPS. The first extension is the redefinition of the fitness assignment process to simplify analysis and enable more comprehensive constraint handling. The second extension permits automatic generation of the target vector, with no need to specify a search direction *a priori*.

The operation of MSOPS is to generate a set of T target vectors. For every target vector, the performance of every individual in the population (size P) is evaluated based on the chosen aggregation/target vector method(s). Thus each of the P members of the population has a set of scores, held in a score matrix S with dimensions $P \times T$. Each column of the matrix S is now ranked, with the best performing population member being given rank one and the worst a rank P. MSOPS sorts the individual of population based on these aggregated rank instead of objective functions.
In MSOPS-II, the aggregate fitness of the $i^t h$ member of P has been redefined as follows:
1) for each column (target vector), find the minimum and second smallest metric value (from valid population members only),
2) scale each column by the minimum value found, except for the row which gave the minimum value: use the second lowest to scale this result,
3) for each row (population member), find the minimum scaled value to represent fitness, resulting in a column vector as the final aggregate fitness,
4) sort column vector to rank population.
Moreover, MSOPS-II designs an automatic target vector generation system, that uses the current population to provide the source for the new target vectors instead of specifying them a-priori. Those modifications make MSOPS-II a general-purpose many objective optimization algorithm, requiring minimal initial configuration.

- **MOEA/D** decomposes a multi/many-objective problem into a set of single-objective problems with uniformly distributed weight vectors and optimizes them simultaneously. Each problem is optimized and updated by the information from its neighbouring problems. The neighbourhood relations are defined based on the distances among their weight vectors. Several aggregation functions, weighted sum (WS), Tchebycheff and penalty-based boundary intersection (PBI), can be used in MOEA/D. The brief process of MOEA/D is introduced as follows:
    1) give a uniform spread of N weight vectors: $\lambda^1, \ldots, \lambda^N$, while generating a N-sized initial population; compute the Euclidean distances between any two weight vectors and then work out the T closest weight vectors to each weight vector i, forming the neighbourhood set $B(i)$.
    2) for each i in N, perform genetic operators to generate a new solution y from two randomly chosen solutions in neighbourhood $B(i)$ in a probability, and then update neighbouring solutions if y is better, evaluated by the chosen aggregation function.
    3) repeat 2) if stopping criteria is not satisfied.

  In this work, PBI is selected as the aggregation function since it shows more competitive performance than other functions when solving problems with a high-dimensional objective space [49]. More advanced variants of MOEA/D, e.g., MOEA/D-STM [52] and MOEA/D-PaS [53], are discussed in [54].
- **HypE** is a hypervolume-based algorithm designed for many-objective optimization. Monte Carlo simulation is adopted to approximate the exact hypervolume values, thereby significantly reducing the computational cost. Moreover, HypE applies the rankings of solutions induced by HV values instead of the actual HV values. In HypE, the hypervolume-based fitness of a solution is not only calculated based on its own hypervolume contribution but also the hypervolume contribution associated with other solutions. This is a more refined approach than that adopted in the other hypervolume-based approaches, such as the S metric selection-based evolutionary multi-objective optimization algorithm (SMS-EMOA), in which contribution calculations are limited to single solutions, without consideration of the wider population context. Experimental results indicate that HypE is highly effective for many-objective problems and represents a trade-off between the accuracy of hypervolume estimation and computing cost.
- **PICEA-g** is a realization of the preference-inspired coevolutionary algorithm (PICEA). In PICEAs, a family of decision-maker preferences are coevolved together with the candidate solutions. The use of coevolution as the adaptation mechanism is particularly notable as it is very challenging to harness the power of coevolution for the purposes of optimization (rather than simply to explore coevolutionary dynamics). It focuses on the a posteriori optimization to provide decision makers with the approximation of the entire Pareto front.
    1) In PICEA-g, first, a population of candidate solutions S of fixed size N and a population of preference sets G of fixed size $NGoal$ are initialized.
    2) In each generation t, N offspring $Sc(t)$ are produced by applying genetic variation operators to parents $S(t)$. Simultaneously, $NGoal$ new preference sets $Gc(t)$, are randomly regenerated based on the initial bounds.
    3) The combined populations, $S(t)+Sc(t)$ and $G(t)+Gc(t)$, are sorted according to the fitness. The method of calculating fitness of a candidate solution and a preference is based on the fitness assignment method of [55].
    4) Accordingly, the best N solutions are selected as new parent population $S(t+1)$, and the best $NGoal$ solutions are selected as new preference population $G(t+1)$.

  In systematic comparisons of PICEA-g with other representative methods, PICEA-g shows competitive performance as the number of objectives increases.
- **SDE** was proposed as a general modification of the diversity maintenance mechanism to make Pareto-based algorithms suitable for many-objective optimization. The idea of SDE is to put individuals with poor convergence into crowded regions. Thus, individuals that are

poorly converged will be eliminated easily during the evolution as they will be assigned a higher density value. More specifically, when estimating the density of an individual p, SDE shifts the positions of other individuals in the population according to the convergence comparison between these individuals and p on each objective. If an individual performs better than p for an objective, it will be shifted to the same position of p on this objective; otherwise, it remains unchanged. This process would assign p a higher density value if it has a poor convergence. Thus, only the individual with both good convergence and good diversity has a low crowding degree in SDE.
The application of SDE in three popular Pareto-based algorithms validates its efficiency in many-objective optimization. Here, the application of SDE in the SPEA2 algorithm (SPEA2+SDE) is selected to conduct experiments.

- **GrEA** is a grid-based evolutionary algorithm for solving many-objective problems aimed at balancing the convergence and diversity of solutions. To strengthen the selection pressure towards the optimal direction, the grid dominance and grid difference are introduced to determine the mutual relationships of individuals in a grid environment. Moreover, the fitness assignment process is modified by three grid-based criteria (grid ranking(GR), grid crowding distance(GCD), and grid coordinate point distance(GCPD)). The basic procedure of GrEA is similar to NSGA-II. The initial population P and the grid environment for the current population P are set at first. Then the fitness of individuals in P is assigned according to their location in the grid, evaluated by GR, GCD, and GCPD. Mating selection is performed to pick out promising solutions for variation. Finally, the environmental selection procedure is implemented according to the grid-based fitness of individuals.
- **NSGA-III** uses the framework of the NSGA-II procedure with significant modifications in its crowding distance operator to tackle MaOPs. It is an elitist approach and the basic procedure is similar to NSGA-II: the parent and offspring population are combined and evaluated using a fast nondominated sorting approach and an efficient crowding scheme. When more than N population members of the combined population belong to the nondominated set, only those with a better crowding measure are chosen.
In NSGA-III, the crowding distance operator is replaced with the following approaches. First, a number of well-spread reference points, similar to the weight vectors in MOEA/D, are applied to maintain the diversity among population members. Each population member is associated with a reference point based on its perpendicular distance to the reference line. Then, non-dominated solutions close to the reference points are prioritized. After the generation of offspring solutions, the non-dominated sorting method and elitism mechanism are employed following the procedure in NSGA-II.
- **KnEA** is a knee point-driven evolutionary algorithm for solving many-objective problems. In KnEA, solutions for the next generation are first selected based on the non-dominance selection criterion, and then knee points are used as the secondary selection criterion. The work shows that preference over knee points can be approximately viewed as a bias towards larger HV, thus maintaining good convergence and diversity.
The basic procedure of KnEA is similar to NSGA-II. First, an initial parent population of size N is randomly generated. Second, N offspring individuals are selected from the parent population by a binary tournament strategy using a variation method. In the binary tournament selection, three tournament metrics are adopted, namely, the dominance relationship, the knee point criterion, and a weighted distance measure. Third, nondominated sorting is performed on the combination of the parent and offspring population, followed by an adaptive strategy to identify solutions located in the knee regions of each nondominated front in the combined population. Fourth, an environmental selection is conducted to select N individuals as the parent population of the next generation.
- **RVEA** applies a framework similar to that of the NSGA-II algorithm, from which RVEA adopts an elitism strategy, where the offspring population is generated using traditional genetic operations, and then the offspring population is combined with the parent population to undergo an elitism selection. The main new contributions in the RVEA lie in the two other components, i.e., the reference vector guided selection and the reference vector adaptation. RVEA uses the reference vectors to partition the objective space into a number of subspaces, where the selection is performed separately. To obtain a uniform distribution of Pareto optimal solutions, an adaptive strategy of reference vectors is proposed to deal with objective functions that are not well normalized. In addition, RVEA applies a scalarization approach called the angle penalized distance (APD) to dynamically balance the convergence and diversity of solutions in high-dimensional objective space.
- **Two_Arch2** is a significantly improved version of the two-archive algorithm (Two_Arch), which is a low-complexity algorithm with two archives that focus on convergence (CA) and diversity (DA) separately. The basic framework of Two_Arch2 is similar to general MOEAs (reproduction and iteration). The main difference is that the nondominated solution set is divided into two archives (CA and DA). After the generation of children, the environmental selection is performed by CA and DA, separately. Different from Two_Arch, Two_Arch2 makes crossover between CA and DA but mutation on CA only during the process of reproduction. Moreover, the $I_{\epsilon+}$ indicator and a L_p norm based crowding distance metric are applied to the individuals

in CA and DA, respectively. As the diversity of CA is too poor, Two_Arch2 use DA as the final output, which is different from the union set of CA and DA in Two_Arch. Two_Arch2 is reported to solve many-objective problems with satisfactory convergence, diversity, and complexity.

- **$\theta-$DEA** introduces a new dominance relation, termed θ dominance, to improve the convergence of NSGA-III by referring to the procedure of the MOEA/D algorithm, while preserving the strength of NSGA-III in its diversity maintenance. In θ dominance, given a set of well-spread reference points, solutions are clustered into different groups, with each group represented by a reference point. Competitive relationships between solutions only exist in the same cluster, where a fitness function similar to PBI is defined. Based on the θ-dominance relation, the combined parent-offspring population is sorted into different non-domination levels, and a new population is constructed as in NSGA-III.
The framework of $\theta-$DEA is briefly described as follows. First, a set of N reference points, the initial population P_0 of size N, the ideal point $\mathbf{z}^*$ and the nadir point $\mathbf{z}^{nad}$ are generated. $\mathbf{z}^*$ and $\mathbf{z}^{nad}$ are updated during the search. The offspring population is produced by using the recombination operator, which is then combined with the current population to form a new population. Next, a population S_t is obtained according to the Pareto non-domination level. The normalization procedure is executed to S_t assisted by $\mathbf{z}^*$ and $\mathbf{z}^{nad}$. After normalization, the clustering operator is used to split the members in S_t into a set of N clusters, where each cluster is represented by a reference point. Then, the nondominated sorting based on θ dominance (not Pareto dominance) is employed to classify S_t into different θ-nondomination levels. Once θ-nondominated sorting has been finished, the remaining steps fill the population slots in P_{t+1} using one level at a time.
- **MOEA/DD** is a combination of MOEA/D and NSGA-III for many-objective optimization, exploiting the merits offered by both dominance and decomposition-based approaches. Each weight vector in MOEA/DD defines a subproblem, while also specifying a unique subregion in the objective space. To maintain the diversity in a high-dimensional objective space, the density of a population is estimated by the local niche count of a subregion. In the population updating process of MOEA/DD, only one offspring solution is considered each time. That is to say, multiple rounds of this update procedure will be implemented if more than one offspring solution have been generated. Upon updating, the associated subregion of the offspring solution $\mathbf{x}^c$ is identified at first. Then, $\mathbf{x}^c$ is combined with the parent population P to form a hybrid population P'. The method proposed in [56] is used to update the nondomination level structure of the population after introducing an offspring solution.
- **AnD** is a very simple and efficient approach which is free from the use of the Pareto relationship, the reference vectors or points and the performance indicators. Instead, it only makes use of two strategies (i.e., the angle-based selection strategy and the shift-based density estimation strategy) to delete the poor individuals one by one in the environmental selection process. More specifically, the angle based selection strategy is employed to maintain the diversity of the search directions and it can be used to identify a pair of solutions with the minimum vector, which means these two solutions search in the most similar directions. Subsequently, the shift-based density estimation strategy is conducted to differentiate these two solutions considering both the diversity and convergence performance, and then delete the inferior one. The experimental results indicate that AnD can achieve a highly competitive performance when comparing with seven state-of-the-art MaOEAs. In addition, AnD can also be conveniently extended (C-AnD) for solving constrained MaOPs.

TABLE 1. The categories of thirteen selected algorithms. C1: decomposition-based approaches. C2: indicator-based approaches. C3: approaches that modify the secondary diversity-based selection criterion. C4: approaches that modify the traditional dominance relation. C5: preference-based approaches.

Algorithms	Categories	Algorithms	Categories
MSOPS-II	C1	KnEA	C3
MOEA/D	C1	RVEA	C1,C5
HypE	C2	Two_Arch2	C5
PICEA-g	C5	$\theta-$DEA	C1,C4
SDE	C3	MOEA/DD	C1,C3
GrEA	C4	AnD	C4
NSGA-III	C1,C3		

Overall, the main differences and similarities of the thirteen algorithms above can be summarized by the their categories as discussed in Introduction, as shown in Table 1. In this sense, MSOPS-II and MOEA/D are decomposition-based methods (let this category be Category C1), which does not apply the Pareto dominance relation. HypE is non-Pareto-based approach as well, which belongs to the indicator-based approach (Category C2). SDE and KnEA are still Pareto-based approaches but they modify the secondary diversity-based selection criterion to better solve MaOPs (Category C3). Moreover, GrEA, $\theta-$DEA and AnD modify the traditional dominance relation to enhance the selection pressure (Category C4). PICEA-g is a preference-based approach (Category C5) and Two_Arch2 can be considered as this category to some extent. In terms of RVEA, when a set of evenly distributed reference vectors are generated for achieving representative solutions of the whole PF, the proposed RVEA can be considered as one of the decomposition-based approaches. However, if user preferences are available and a set of specific reference vectors are generated for achieving only a preferred section of the PF, RVEA can also be seen as a preference based approach. In addition, MOEA/DD and NSGA-III can be categorized into both the C1 and C3 category, as they use hybrid mechanisms of the Pareto dominance-based fitness evaluation and the decomposition-based techniques.

TABLE 2. The properties of test instances in WFG problem suite.

Problem	Properties
WFG1	Convex, Mixed, Biased, Unimodal
WFG2	Convex, Disconnected, Non-separable
WFG3	Linear, Partially degenerate, Unimodal, Non-separable
WFG4	Concave, Multimodal
WFG5	Concave, Deceptive
WFG6	Concave, Unimodal, Non-separable
WFG7	Concave, Biased, Unimodal
WFG8	Concave, Biased, Unimodal, Non-separable
WFG9	Concave, Biased, Multimodal, Deceptive, Non-separable

III. EXPERIMENTAL DESIGN

A. TEST PROBLEMS

Three problem suites are used in this study: WFG [57], MaF [58], and Pareto-Box [59]. Problem attributes of separability or nonseparability, unimodality or multimodality, unbiased or biased parameters, and convex or concave geometries are all covered.

- WFG is a commonly used continuous benchmark function and is scalable to any number of objectives. To achieve a thorough comparison, all test instances in WFG (9 test functions) are considered in this paper. Table 2 lists the properties of WFG problems.
- MaF is a benchmark function suite proposed in 2017 to promote research on evolutionary many-objective optimization. Fifteen benchmark functions are designed in MaF with diverse properties that provide a good representation of various real-world scenarios. Due to space limitations, seven representative problems with different characteristics are selected to evaluate the algorithms. The properties of the seven test functions are summarized in Table 3.
- The Pareto-Box problem is a simple and interesting many-objective test function developed by Köppen and Yoshida [59] and extended by Ishibuchi *et al.* [60]. There are two characteristics of the Pareto-Box problem. One is that its Pareto optimal set in the decision space is one (or several) two-dimensional closure(s). The second important characteristic is that the crowding in its decision space is closely related to the crowding in its objective space. Thus, we can view the distribution of solutions in the decision space while inferring the behaviour of algorithms in the objective.

B. PERFORMANCE METRICS

The widely used performance metrics of inverted generational distance (IGD) [10] and hypervolume (HV) [61] are used in this work to evaluate the performance of all compared algorithms. IGD and HV can measure both convergence and diversity.

Inverted Generational Distance (IGD): Let P^* be a set of uniformly distributed reference points on the Pareto front and P be the set of solutions obtained by an EMO algorithm. The IGD value of P can be defined as follows:

$$IGD(P, P^*) = \frac{\sum_{v \in P^*} d(v, P)}{|P^*|} \tag{1}$$

where $d(v, P)$ is the Euclidean distance from point $v \in P^*$ to its nearest point in P. $|P^*|$ is the cardinality of P^*. Here, the number of reference points is taken as 10,000 in the computational experiments. The smaller the IGD, the better the quality of P for approximating the whole PF.

Hypervolume (HV): Let $y^* = (y_1^*, \ldots, y_m^*)$ be a reference point in the objective space dominated by all Pareto optimal solutions. The HV metric measures the size of the region that is dominated by P and dominates y^*. Given a reference point y^*, a larger HV value indicates better performance. The selection of the reference point is an important issue. It is better for y^* to be slightly larger than the maximum value of each objective to emphasize the balance between convergence and diversity. Here, the reference point is set as 1.1 times the upper bound of the Pareto front, as recommended by [60].

C. GENERAL EXPERIMENTAL SETTINGS

In our experimental study, all test results are obtained from a recently developed software platform, PlatEMO[1] [62], which has more than 50 representative algorithms and more than 100 benchmark functions. Assume that D is the number of decision variables and M is the number of objectives. For each selected test problem, M is taken as 2, 3 or 7.

- For MaF test instances, the number of decision variables is set as $D = M + K - 1$, where K is set as 10 for problems MaF1-MaF6 and 20 for problem MaF7.
- For WFG test instances, the number of decision variables is set as $D = M + 9$. The number of position-related variables $K = M - 1$, and the number of distance-related variables $L = D - K$.
- The simulated binary crossover (SBX) operator is applied for crossover, and polynomial mutation is used for mutation, with distribution indexes both set to 20.
- The population size N is set to 100. The final results are obtained by executing 31 independent runs of each algorithm after 250*100 evaluations.
- We set the parameters of the algorithms as recommended in their original papers or some relevant literatures. The specific parameter settings in each algorithm are shown in Table 4. Here, N is the population size and M is the number of objectives. It is noted that the best parameters recommended by the authors of GrEA and KnEA are also associated with the specific problems, e.g., $div = 10$ for 3-objective WFG1. However, we use the same set of

[1] PlatEMO can be downloaded at http://bimk.ahu.edu.cn/index.php?s=/Index/Software/index.html

TABLE 3. The properties of test instances in MaF problem suite.

Problem	Properties	Note
MaF1	Linear	No single optimal solution in any subset of objectives
MaF2	Concave	No single optimal solution in any subset of objectives
MaF3	Convex, Multimodal	
MaF4	Concave, Multimodal	Badly-scaled and no single optimal solution in any sub-set of objectives
MaF5	Convex, Biased	Badly-scaled
MaF6	Concave, Degenerate	
MaF7	Mixed, Disconnected, Multimodal	

TABLE 4. The parameters of the test algorithms. ***N*** **is the population size and** ***M*** **is the number of objectives.**

Algorithms	Parameters
MOEA/D	The penalty parameter θ of the PBI function: 5, neighbourhood size T: $N/10$
HypE	The number of sampling points: 10,000
PICEA-g	The number of preferences $NGoal$: $100 * M$
SDE	Archive size: N
GrEA	The grid division $div = 45$ for 2 objectives, $div = 15$ for 3 objectives and $div = 8$ for 7 objectives
KnEA	The rate of knee points in population $K = 0.6$ for 2 objectives and $K = 0.5$ for other conditions.
RVEA	The index α of penalty function: 2, the frequency f_r of reference vector adaptation: 0.1
Two_Arch2	The sizes of CA and DA: N, the p for L_p-norm-based distances: $1/M$
θ-DEA	The penalty parameter θ: 5
MOEA/DD	The penalty parameter θ of the PBI function: 5, neighbourhood size T: $N/10$, neighbourhood selection probability δ: 0.9

TABLE 5. HV results (Mean and SD) of the 13 algorithms on the 2-objective WFG problems. The top-ranked algorithms for each problem instance are highlighted in grey background, the worst algorithms highlighted in boldface.

Problem	M	D	MSOPSII	MOEAD	HypE	PICEA-g	SPEA2SDE	GrEA	NSGAIII	KnEA	RVEA	Two_Arch2	θ-DEA	MOEADD	AnD
WFG1	2	11	5.7269e+0	**4.0578e+0**	4.5281e+0	5.6664e+0	5.9717e+0	6.0119e+0	5.5877e+0	5.7855e+0	**3.8278e+0**	5.5287e+0	5.5124e+0	5.2285e+0	5.2936e+0
			(5.19e-1)	(2.72e-1)	(5.96e-1)	(1.74e-1)	(3.07e-1)	(4.33e-1)	(1.76e-1)	(4.06e-1)	(2.02e-1)	(5.16e-1)	(3.76e-1)	(1.94e-1)	(1.74e-1)
WFG2	2	11	6.1004e+0	5.8777e+0	6.1297e+0	6.0947e+0	6.1210e+0	6.1007e+0	6.1103e+0	**4.6332e+0**	5.8352e+0	6.1048e+0	6.0923e+0	6.0902e+0	6.0793e+0
			(8.43e-3)	(1.43e-1)	(4.46e-3)	(4.86e-2)	(6.20e-3)	(4.35e-3)	(6.23e-3)	(5.53e-1)	(5.60e-2)	(1.66e-2)	(4.76e-2)	(1.25e-2)	(1.44e-2)
WFG3	2	11	5.6060e+0	5.5267e+0	5.6236e+0	5.5776e+0	5.6163e+0	5.5844e+0	5.6104e+0	5.6023e+0	**5.3659e+0**	5.6057e+0	5.6150e+0	5.5988e+0	5.5898e+0
			(7.10e-3)	(3.03e-2)	(4.52e-3)	(1.08e-2)	(2.27e-3)	(4.35e-3)	(7.33e-3)	(9.70e-3)	(5.61e-2)	(1.28e-2)	(5.01e-3)	(1.17e-2)	(1.10e-2)
WFG4	2	11	3.3410e+0	3.2323e+0	3.3636e+0	3.3228e+0	3.3532e+0	3.3285e+0	3.3456e+0	3.3009e+0	**3.0510e+0**	3.3572e+0	3.3462e+0	3.3390e+0	3.3415e+0
			(3.06e-3)	(2.70e-2)	(1.43e-3)	(7.24e-3)	(2.31e-3)	(2.05e-3)	(4.70e-3)	(2.20e-2)	(4.68e-2)	(3.10e-3)	(3.23e-3)	(7.53e-3)	(3.92e-3)
WFG5	2	11	3.0189e+0	2.9619e+0	3.0333e+0	3.0133e+0	3.0125e+0	2.9844e+0	3.0224e+0	2.9405e+0	**2.8835e+0**	3.0231e+0	3.0151e+0	2.9949e+0	3.0200e+0
			(4.59e-3)	(6.29e-3)	(1.39e-2)	(2.07e-2)	(2.00e-2)	(2.09e-2)	(1.21e-2)	(5.02e-2)	(3.58e-2)	(2.15e-2)	(1.99e-2)	(2.21e-2)	(1.42e-2)
WFG6	2	11	2.9624e+0	2.8820e+0	2.9595e+0	2.8711e+0	2.9387e+0	2.9528e+0	2.9127e+0	**2.0673e+0**	2.6033e+0	2.9906e+0	2.9406e+0	2.9177e+0	2.9404e+0
			(1.18e-1)	(1.21e-1)	(1.12e-1)	(1.05e-1)	(1.00e-1)	(1.50e-1)	(1.29e-1)	(2.33e-1)	(1.08e-1)	(1.11e-1)	(7.60e-2)	(1.02e-1)	(9.97e-2)
WFG7	2	11	3.3394e+0	3.2300e+0	3.3662e+0	3.3327e+0	3.3563e+0	3.3291e+0	3.3531e+0	**2.8652e+0**	3.1496e+0	3.3623e+0	3.3534e+0	3.3456e+0	3.3437e+0
			(2.28e-3)	(1.68e-2)	(5.20e-4)	(3.87e-3)	(1.61e-3)	(1.20e-3)	(1.57e-3)	(1.84e-1)	(2.64e-2)	(6.93e-4)	(1.62e-3)	(1.79e-3)	(2.72e-3)
WFG8	2	11	2.7762e+0	2.7076e+0	2.7973e+0	2.7488e+0	2.7865e+0	2.7758e+0	2.7766e+0	**1.4156e+0**	2.3644e+0	2.7890e+0	2.7702e+0	2.7789e+0	2.7592e+0
			(1.33e-2)	(3.44e-2)	(1.09e-2)	(1.17e-2)	(8.82e-3)	(4.35e-3)	(8.53e-3)	(1.21e-1)	(5.62e-2)	(2.09e-2)	(1.28e-2)	(9.40e-3)	(9.98e-3)
WFG9	2	11	3.1560e+0	**2.9861e+0**	3.3157e+0	3.1493e+0	3.2928e+0	3.2665e+0	3.2641e+0	3.1901e+0	3.1106e+0	3.3016e+0	3.2674e+0	3.2642e+0	3.2562e+0
			(3.93e-1)	(2.72e-1)	(1.06e-2)	(2.66e-1)	(1.75e-2)	(1.93e-2)	(1.60e-2)	(2.35e-1)	(2.90e-2)	(2.07e-2)	(2.05e-2)	(1.76e-2)	(1.84e-2)

parameters for all problem instances, which is a little different from the best one. This difference should not lead to a large deterioration of the results if this algorithm is robust to minor changes in parameters, which should also be an assessment of the algorithm.

IV. RESULTS AND DISCUSSION

To ensure clarity in the result comparisons, the results are presented in two stages. First, each of the 16 test problems on which the performances of the different algorithms are compared is listed and introduced separately. Due to the large number of algorithms and test instances, the results for the HV metric are first illustrated in this comparison stage. Then, all of the results are summarized in terms of both HV and IGD. A variety of conclusions and differences between the behaviours of the HV and IGD metrics are obtained.

The HV values obtained by the 13 algorithms (mean and standard deviation) on 2-, 3- and 7-objective WFG and MaF problems are presented in Tables 5, 6, 7, 8, 9, and 10.

TABLE 6. HV results (Mean and SD) of the 13 algorithms on the 3-objective WFG problems. The top-ranked algorithms for each problem instance are highlighted in grey background, the worst algorithms highlighted in boldface.

Problem	M	D	MSOPSII	MOEAD	HypE	PICEA-g	SPEA2SDE	GrEA	NSGAIII	KnEA	RVEA	Two_Arch2	θ-DEA	MOEADD	AnD
WFG1	3	12	5.1126e+1	4.0333e+1	4.0727e+1	**3.2010e+1**	5.5223e+1	5.2279e+1	4.5021e+1	5.2412e+1	4.2492e+1	4.7383e+1	4.7641e+1	**3.1662e+1**	4.6596e+1
			(2.46e+0)	(3.12e+0)	(2.21e+0)	(3.25e+0)	(2.21e+0)	(1.91e+0)	(2.52e+0)	(2.10e+0)	(2.01e+0)	(3.89e+0)	(2.66e+0)	(3.51e+0)	(1.68e+0)
WFG2	3	12	5.8898e+1	**5.6393e+1**	5.9721e+1	5.8787e+1	5.9408e+1	5.8985e+1	5.9174e+1	5.9183e+1	5.8634e+1	5.9525e+1	5.9383e+1	5.8858e+1	5.9015e+1
			(1.95e-1)	(1.06e+0)	(6.12e-2)	(3.56e-1)	(1.60e-1)	(1.71e-1)	(8.72e-2)	(1.27e-1)	(2.46e-1)	(1.61e-1)	(1.36e-1)	(2.22e-1)	(1.58e-1)
WFG3	3	12	6.1101e+0	**5.4032e+0**	6.5857e+0	5.8623e+0	6.3197e+0	6.2468e+0	6.0228e+0	5.7085e+0	**5.2051e+0**	6.2051e+0	6.0672e+0	**5.1317e+0**	5.7264e+0
			(2.35e-1)	(3.69e-1)	(1.19e-2)	(9.11e-2)	(7.17e-2)	(8.13e-2)	(8.34e-2)	(4.77e-1)	(1.64e-1)	(7.67e-2)	(1.13e-1)	(5.19e-1)	(1.25e-1)
WFG4	3	12	3.4142e+1	**3.3536e+1**	3.5640e+1	**3.3608e+1**	3.5203e+1	3.5396e+1	3.5064e+1	3.4245e+1	3.4541e+1	3.5139e+1	3.5121e+1	3.4594e+1	3.4800e+1
			(2.24e-1)	(2.24e-1)	(7.08e-2)	(1.64e-1)	(1.25e-1)	(7.02e-2)	(1.40e-1)	(2.29e-1)	(1.71e-1)	(1.72e-1)	(9.78e-2)	(1.52e-1)	(1.37e-1)
WFG5	3	12	3.2028e+1	**3.1740e+1**	3.3039e+1	3.1966e+1	3.2967e+1	3.2939e+1	3.3030e+1	3.2204e+1	3.2778e+1	3.2595e+1	3.3013e+1	3.2485e+1	3.2893e+1
			(2.66e-1)	(1.61e-1)	(4.79e-2)	(2.72e-1)	(9.75e-2)	(1.68e-1)	(4.11e-2)	(2.88e-1)	(1.57e-1)	(2.06e-1)	(2.07e-1)	(8.05e-2)	(6.75e-2)
WFG6	3	12	**3.0432e+1**	**3.0040e+1**	3.2511e+1	**3.0473e+1**	3.2028e+1	3.2043e+1	3.1592e+1	**3.0655e+1**	3.1352e+1	3.2319e+1	3.1857e+1	3.1424e+1	3.2232e+1
			(8.62e-1)	(6.90e-1)	(1.10e+0)	(1.18e+0)	(8.81e-1)	(7.94e-1)	(9.19e-1)	(5.77e-1)	(1.29e+0)	(1.03e+0)	(8.56e-1)	(1.11e+0)	(1.03e+0)
WFG7	3	12	3.4205e+1	**3.0991e+1**	3.5463e+1	3.4086e+1	3.5521e+1	3.5657e+1	3.5219e+1	3.4676e+1	3.4629e+1	3.5493e+1	3.5327e+1	3.4589e+1	3.5215e+1
			(2.80e-1)	(1.44e+0)	(8.35e-2)	(1.80e-1)	(9.79e-2)	(5.65e-2)	(7.07e-2)	(2.75e-1)	(1.74e-1)	(7.00e-2)	(5.71e-2)	(1.02e-1)	(1.16e-1)
WFG8	3	12	**2.7389e+1**	2.8476e+1	2.9716e+1	2.8473e+1	3.0265e+1	3.0312e+1	2.9502e+1	2.8588e+1	2.8932e+1	2.9843e+1	2.9529e+1	2.9255e+1	2.9305e+1
			(2.18e-1)	(4.93e-1)	(1.23e-1)	(2.67e-1)	(1.28e-1)	(9.64e-2)	(1.96e-1)	(3.18e-1)	(4.74e-1)	(1.66e-1)	(2.04e-1)	(2.28e-1)	(1.71e-1)
WFG9	3	12	3.2996e+1	**2.9486e+1**	3.4826e+1	3.2241e+1	3.4098e+1	3.4302e+1	3.2742e+1	3.3413e+1	3.2922e+1	3.3809e+1	3.3040e+1	3.3016e+1	3.3589e+1
			(2.09e+0)	(2.05e+0)	(1.61e-1)	(7.77e-1)	(2.63e-1)	(1.19e-1)	(1.99e+0)	(2.49e-1)	(4.23e-1)	(3.35e-1)	(1.95e+0)	(2.09e-1)	(6.52e-1)

TABLE 7. HV results (Mean and SD) of the 13 algorithms on the 7-objective WFG problems. The top-ranked algorithms for each problem instance are highlighted in grey background, the worst algorithms highlighted in boldface.

Problem	M	D	MSOPSII	MOEAD	HypE	PICEA-g	SPEA2SDE	GrEA	NSGAIII	KnEA	RVEA	Two_Arch2	θ-DEA	MOEADD	AnD
WFG1	7	16	1.0597e+6	6.7814e+5	1.0152e+6	**3.7052e+5**	9.4067e+5	9.0173e+5	6.7005e+5	7.6370e+5	7.1587e+5	6.2475e+5	8.6348e+5	6.2374e+5	7.3361e+5
			(6.98e+4)	(1.03e+5)	(1.09e+5)	(2.31e+4)	(7.45e+4)	(9.08e+4)	(5.83e+4)	(4.87e+4)	(6.61e+4)	(7.16e+4)	(1.18e+5)	(7.10e+4)	(5.01e+4)
WFG2	7	16	1.2459e+6	**1.0982e+6**	1.2533e+6	1.2052e+6	1.2208e+6	1.2188e+6	1.2357e+6	1.2340e+6	1.2027e+6	1.2353e+6	1.2136e+6	1.1823e+6	1.2460e+6
			(7.89e+3)	(5.91e+4)	(1.60e+3)	(1.07e+4)	(7.96e+3)	(6.92e+3)	(6.87e+3)	(5.22e+3)	(1.38e+4)	(4.80e+3)	(1.60e+4)	(1.03e+4)	(3.18e+3)
WFG3	7	16	2.2280e-1	**0.0000e+0**	2.5137e-1	**0.0000e+0**	**2.5654e-3**	1.4401e-2	2.5049e-2	**0.0000e+0**	**0.0000e+0**	**8.9220e-3**	1.0394e-1	**0.0000e+0**	3.6690e-2
			(1.27e-2)	(0.00e+0)	(9.41e-3)	(0.00e+0)	(5.17e-3)	(2.28e-2)	(2.41e-2)	(0.00e+0)	(0.00e+0)	(1.69e-2)	(2.67e-2)	(0.00e+0)	(1.70e-2)
WFG4	7	16	1.0543e+6	**4.5993e+5**	8.5398e+5	9.0674e+5	1.0003e+6	1.0330e+6	1.0497e+6	1.0455e+6	1.0332e+6	8.8404e+5	1.0583e+6	1.0007e+6	1.0295e+6
			(8.39e+3)	(7.66e+4)	(6.86e+4)	(1.88e+4)	(8.53e+3)	(7.16e+3)	(1.98e+4)	(7.76e+3)	(1.77e+4)	(1.13e+4)	(5.46e+3)	(1.23e+4)	(1.30e+4)
WFG5	7	16	9.5088e+5	**4.9829e+5**	1.0147e+6	8.4682e+5	9.6742e+5	9.9047e+5	1.0101e+6	9.6967e+5	1.0090e+6	8.4639e+5	1.0131e+6	9.2436e+5	9.8610e+5
			(9.74e+3)	(4.00e+4)	(2.30e+4)	(2.05e+4)	(8.44e+3)	(5.08e+3)	(3.23e+3)	(7.67e+3)	(5.70e+3)	(1.22e+4)	(2.24e+3)	(1.50e+4)	(6.57e+3)
WFG6	7	16	9.4253e+5	**2.9428e+5**	9.7930e+5	8.1007e+5	9.4228e+5	9.5978e+5	9.7440e+5	9.0745e+5	9.4762e+5	7.9656e+5	9.7788e+5	8.9230e+5	9.6920e+5
			(3.23e+4)	(3.25e+4)	(3.90e+4)	(2.65e+4)	(2.28e+4)	(2.65e+4)	(3.01e+4)	(3.28e+4)	(3.21e+4)	(2.60e+4)	(2.25e+4)	(3.34e+4)	(2.67e+4)
WFG7	7	16	1.0421e+6	**4.2584e+5**	1.0239e+6	8.7506e+5	1.0408e+6	1.0714e+6	1.0574e+6	1.0419e+6	1.0357e+6	9.0888e+5	1.0708e+6	1.0062e+6	1.0640e+6
			(1.07e+4)	(4.81e+4)	(4.79e+4)	(3.25e+4)	(6.79e+3)	(7.19e+3)	(8.72e+3)	(1.21e+4)	(2.09e+4)	(1.62e+4)	(5.99e+3)	(1.88e+4)	(5.96e+3)
WFG8	7	16	7.3295e+5	**4.3223e+4**	9.3137e+5	7.7498e+5	9.1649e+5	8.3782e+5	8.7260e+5	8.0024e+5	7.1463e+5	6.8366e+5	8.7966e+5	8.5141e+5	8.5580e+5
			(2.25e+4)	(4.46e+4)	(1.50e+4)	(2.35e+4)	(9.21e+3)	(1.96e+4)	(2.01e+4)	(3.20e+4)	(9.72e+4)	(1.60e+4)	(1.09e+4)	(1.94e+4)	(3.14e+4)
WFG9	7	16	8.7803e+5	**3.7184e+5**	9.7818e+5	7.3898e+5	9.4424e+5	9.6204e+5	9.3590e+5	9.6448e+5	9.1356e+5	8.1836e+5	9.6162e+5	7.8588e+5	9.4747e+5
			(7.97e+4)	(1.33e+5)	(9.01e+4)	(5.50e+4)	(1.58e+4)	(1.12e+4)	(4.23e+4)	(4.99e+4)	(3.61e+4)	(3.38e+4)	(4.37e+4)	(5.30e+4)	(4.48e+4)

TABLE 8. HV results (Mean and SD) of the 13 algorithms on the 2-objective MaF problems. The top-ranked algorithms for each problem instance are highlighted in grey background, the worst algorithms highlighted in boldface.

Problem	M	D	MSOPSII	MOEAD	HypE	PICEA-g	SPEA2SDE	GrEA	NSGAIII	KnEA	RVEA	Two_Arch2	θ-DEA	MOEADD	AnD
MaF1	2	11	7.0098e-1	7.0492e-1	7.0476e-1	7.0398e-1	7.0398e-1	**6.9890e-1**	7.0492e-1	7.0239e-1	7.0481e-1	7.0432e-1	7.0493e-1	7.0493e-1	7.0226e-1
			(6.87e-4)	(3.94e-6)	(2.38e-5)	(1.31e-4)	(3.08e-4)	(7.96e-5)	(1.12e-5)	(4.54e-4)	(6.94e-5)	(6.43e-6)	(5.27e-6)	(4.83e-6)	(4.66e-4)
MaF2	2	11	2.1348e-1	2.1342e-1	2.1435e-1	2.1394e-1	2.1414e-1	2.1298e-1	2.1425e-1	**1.9767e-1**	2.1312e-1	2.1423e-1	2.1430e-1	2.1203e-1	2.1379e-1
			(1.19e-4)	(1.84e-4)	(6.43e-6)	(4.65e-5)	(3.85e-5)	(5.52e-5)	(7.08e-5)	(5.44e-3)	(1.75e-4)	(6.94e-6)	(2.64e-5)	(3.24e-4)	(7.32e-5)
MaF3	2	11	7.1015e-2	5.2653e-1	4.5060e-1	**0.0000e+0**	5.6362e-1	5.1389e-1	4.3498e-1	4.8506e-1	**1.6746e-2**	**3.4074e-2**	2.7362e-1	**0.0000e+0**	**0.0000e+0**
			(7.40e-2)	(2.44e-1)	(3.57e-1)	(0.00e+0)	(3.59e-1)	(3.12e-1)	(3.49e-1)	(3.70e-1)	(6.49e-2)	(9.43e-2)	(3.48e-1)	(0.00e+0)	(0.00e+0)
MaF4	2	11	6.7087e+0	5.2863e+0	6.9701e+0	4.3490e+0	6.4118e+0	5.2783e+0	5.0135e+0	6.2956e+0	2.2686e+0	3.6919e+0	6.9843e+0	**1.6247e+0**	4.8977e-1
			(1.76e+0)	(2.59e+0)	(1.70e+0)	(3.07e+0)	(2.00e+0)	(2.64e+0)	(3.07e+0)	(1.85e+0)	(3.22e+0)	(3.46e+0)	(1.60e+0)	(2.32e+0)	(8.19e-1)
MaF5	2	11	2.6856e+0	3.1907e+0	2.5931e+0	3.1930e+0	3.1933e+0	3.3298e+0	3.1955e+0	2.5882e+0	3.3533e+0	**1.3769e+0**	2.8646e+0	3.3560e+0	3.3543e+0
			(1.13e+0)	(6.39e-1)	(1.11e+0)	(6.40e-1)	(6.40e-1)	(1.64e-3)	(6.41e-1)	(1.07e+0)	(6.14e-3)	(1.03e+0)	(1.03e+0)	(4.88e-5)	(2.17e-3)
MaF6	2	11	3.4747e-1	4.1955e-1	4.1206e-1	4.1890e-1	4.1988e-1	4.1628e-1	4.1979e-1	**2.6272e-1**	4.1510e-1	4.2037e-1	4.1986e-1	4.1979e-1	4.1796e-1
			(5.38e-2)	(2.52e-4)	(6.92e-3)	(6.74e-4)	(1.85e-4)	(1.66e-4)	(1.41e-4)	(1.13e-2)	(1.22e-3)	(1.40e-4)	(1.79e-4)	(1.19e-4)	(1.14e-3)
MaF7	2	21	9.5766e-1	9.5821e-1	**8.0646e-1**	9.8818e-1	1.0088e+0	9.9604e-1	1.0083e+0	9.9579e-1	9.5293e-1	9.7138e-1	1.0098e+0	9.9575e-1	1.0026e+0
			(9.49e-2)	(9.27e-2)	(1.27e-1)	(7.13e-2)	(7.58e-4)	(3.52e-3)	(1.44e-4)	(7.31e-3)	(1.08e-2)	(9.82e-2)	(9.46e-5)	(1.07e-3)	(1.22e-3)

As a supplement, the IGD results are presented in Tables 13, 14, 15, 16, 17, and 18.

Based on the HV and IGD results, a statistical comparison between different algorithms on a particular test problem can be designed as follows. Because the Kruskal-Wallis test rejects the hypothesis that all algorithms are equivalent at the 95% confidence level, pairwise performance comparisons between algorithms are first conducted (according to

TABLE 9. HV results (Mean and SD) of the 13 algorithms on the 3-objective MaF problems. The top-ranked algorithms for each problem instance are highlighted in grey background, the worst algorithms highlighted in boldface.

Problem	M	D	MSOPSII	MOEAD	HypE	PICEA-g	SPEA2SDE	GrEA	NSGAIII	KnEA	RVEA	Two_Arch2	θ-DEA	MOEADD	AnD
MaF1	3	12	2.7358e-1 (4.26e-3)	2.6376e-1 (1.19e-4)	**2.4915e-1** (6.82e-3)	2.8988e-1 (1.30e-3)	2.9341e-1 (9.03e-4)	2.9174e-1 (4.67e-4)	2.7202e-1 (2.19e-3)	2.8429e-1 (5.52e-3)	**2.4774e-1** (1.24e-3)	2.9375e-1 (3.78e-4)	**2.5059e-1** (1.11e-3)	**2.5315e-1** (2.32e-3)	2.8973e-1 (7.03e-4)
MaF2	3	12	2.1199e-1 (1.04e-3)	2.0895e-1 (1.01e-3)	2.1554e-1 (4.62e-4)	2.1232e-1 (7.02e-4)	2.1751e-1 (5.49e-4)	2.1781e-1 (4.56e-4)	2.1000e-1 (9.38e-4)	2.1628e-1 (7.27e-4)	2.0314e-1 (1.34e-3)	2.1771e-1 (3.77e-4)	2.1128e-1 (8.73e-4)	**1.7853e-1** (1.72e-3)	2.1643e-1 (6.27e-4)
MaF3	3	12	2.4073e-1 (4.09e-1)	7.6963e-1 (5.79e-1)	2.9862e-1 (5.13e-1)	**0.0000e+0** (0.00e+0)	9.3834e-1 (5.23e-1)	5.9616e-1 (5.89e-1)	1.8691e-1 (3.89e-1)	6.7333e-1 (5.22e-1)	**0.0000e+0** (0.00e+0)	**2.7673e-2** (1.07e-1)	5.5516e-1 (5.63e-1)	**0.0000e+0** (0.00e+0)	**0.0000e+0** (0.00e+0)
MaF4	3	12	1.4408e+1 (1.79e+1)	1.7611e+1 (1.29e+1)	1.5399e+1 (1.78e+1)	7.6143e+0 (1.35e+1)	2.6429e+1 (2.03e+1)	1.8114e+1 (2.05e+1)	**2.5122e+0** (5.83e+0)	2.5561e+1 (1.76e+1)	1.7710e+1 (1.92e+1)	7.0120e+0 (1.47e+1)	2.4121e+1 (1.89e+1)	2.4145e+1 (2.01e+1)	**2.7831e+0** (1.08e+1)
MaF5	3	12	4.0906e+1 (8.44e+0)	**3.2277e+1** (1.15e+1)	**3.2154e+1** (1.04e+1)	4.0069e+1 (8.90e+0)	4.3915e+1 (7.45e+0)	4.3686e+1 (8.04e+0)	4.5138e+1 (6.59e+0)	4.6275e+1 (2.60e-1)	4.7612e+1 (2.44e-2)	4.7625e+1 (1.45e-1)	4.2355e+1 (9.09e+0)	4.6504e+1 (1.19e-2)	4.7343e+1 (1.22e-1)
MaF6	3	12	1.2476e-1 (1.84e-2)	1.0613e-1 (2.34e-2)	**7.4042e-2** (3.15e-3)	1.3247e-1 (2.26e-4)	1.3275e-1 (1.01e-4)	1.2523e-1 (2.33e-4)	1.2855e-1 (8.30e-4)	1.1236e-1 (1.41e-2)	1.1403e-1 (7.16e-3)	1.3315e-1 (5.46e-5)	1.2172e-1 (5.60e-4)	1.2132e-1 (3.79e-4)	**8.8436e-2** (1.05e-2)
MaF7	3	22	1.4970e+0 (1.69e-2)	1.4889e+0 (1.27e-2)	**1.1556e+0** (5.75e-3)	1.3939e+0 (1.44e-1)	1.6380e+0 (4.35e-3)	1.6001e+0 (1.51e-2)	1.6055e+0 (5.30e-3)	1.6293e+0 (9.42e-3)	1.5392e+0 (1.34e-2)	1.6004e+0 (7.11e-2)	1.5598e+0 (4.50e-2)	1.2402e+0 (1.27e-1)	1.5880e+0 (9.49e-3)

TABLE 10. HV results (Mean and SD) of the 13 algorithms on the 7-objective MaF problems. The top-ranked algorithms for each problem instance are highlighted in grey background, the worst algorithms highlighted in boldface.

Problem	M	D	MSOPSII	MOEAD	HypE	PICEA-g	SPEA2SDE	GrEA	NSGAIII	KnEA	RVEA	Two_Arch2	θ-DEA	MOEADD	AnD
MaF1	7	16	2.5898e-4 (3.46e-5)	**3.8959e-5** (1.53e-5)	1.0025e-4 (1.33e-5)	1.6472e-4 (1.85e-5)	3.3003e-4 (2.25e-5)	2.7123e-4 (2.50e-5)	2.0416e-4 (4.13e-5)	3.1912e-4 (1.71e-5)	**2.7286e-5** (1.34e-5)	3.1119e-4 (4.04e-5)	2.7087e-4 (1.44e-5)	1.5618e-4 (4.11e-5)	2.5775e-4 (2.64e-5)
MaF2	7	16	4.4388e-2 (1.60e-3)	4.5893e-2 (6.54e-4)	5.4584e-2 (6.83e-4)	4.6531e-2 (1.02e-3)	5.0658e-2 (6.50e-4)	5.2130e-2 (6.07e-4)	4.2899e-2 (1.60e-3)	4.2868e-2 (3.45e-3)	**3.0243e-2** (9.28e-3)	4.8095e-2 (4.99e-4)	4.3466e-2 (1.74e-3)	3.6499e-2 (1.94e-3)	5.0503e-2 (5.04e-4)
MaF3	7	16	6.5777e-1 (7.12e-1)	1.2914e+0 (7.69e-1)	**0.0000e+0** (0.00e+0)	**0.0000e+0** (0.00e+0)	1.2253e+0 (8.74e-1)	**0.0000e+0** (0.00e+0)	**0.0000e+0** (0.00e+0)	**0.0000e+0** (0.00e+0)	**0.0000e+0** (0.00e+0)	**0.0000e+0** (0.00e+0)	**0.0000e+0** (0.00e+0)	**0.0000e+0** (0.00e+0)	**0.0000e+0** (0.00e+0)
MaF4	7	16	**1.3169e+0** (5.10e+0)	1.8759e+3 (1.94e+3)	1.4601e+4 (3.92e+4)	2.2905e+4 (8.87e+4)	1.1058e+5 (1.37e+5)	4.5907e+5 (1.08e+6)	1.2877e+5 (4.63e+5)	2.0110e+4 (7.29e+4)	1.4287e+4 (2.73e+4)	1.0859e+4 (4.19e+4)	1.2419e+6 (1.74e+6)	1.6827e+4 (2.27e+4)	**0.0000e+0** (0.00e+0)
MaF5	7	16	2.9597e+8 (7.23e+7)	**1.2230e+8** (4.49e+7)	3.4229e+8 (4.15e+7)	3.7557e+8 (1.67e+7)	3.6820e+8 (3.14e+7)	4.5862e+8 (7.60e+5)	4.5107e+8 (2.08e+7)	4.4164e+8 (4.66e+6)	4.2163e+8 (3.91e+7)	3.6133e+8 (1.04e+7)	4.5697e+8 (6.29e+5)	2.6007e+8 (4.71e+7)	4.5134e+8 (1.52e+6)
MaF6	7	16	2.0576e-4 (1.92e-6)	9.4096e-5 (9.57e-5)	1.8629e-4 (5.24e-6)	2.0882e-4 (4.90e-7)	2.0755e-4 (1.06e-6)	1.8234e-4 (5.05e-5)	1.3925e-4 (8.71e-5)	1.9188e-4 (5.32e-5)	1.6360e-4 (2.41e-5)	2.1007e-4 (5.45e-7)	1.8337e-4 (2.14e-6)	1.6249e-4 (3.06e-5)	**7.3615e-7** (1.27e-6)
MaF7	7	26	2.0200e+0 (1.06e-1)	**9.0376e-2** (1.72e-1)	1.6748e+0 (1.08e-1)	6.4972e-1 (3.59e-1)	1.7873e+0 (1.51e-1)	2.3327e+0 (1.15e-1)	1.8372e+0 (1.08e-1)	1.4256e+0 (1.23e-1)	1.5702e+0 (1.85e-1)	1.5455e+0 (9.46e-2)	1.8822e+0 (1.95e-1)	4.2413e-1 (3.81e-1)	2.0293e+0 (6.28e-2)

TABLE 11. The average ranks of the 13 algorithms for the different types of problems based on HV metric. A smaller rank indicates a better performance.

M	MSOPSII	MOEAD	HypE	PICEA-g	SPEA2SDE	GrEA	NSGAIII	KnEA	RVEA	Two_Arch2	θ-DEA	MOEADD	AnD
2	6	11	2	10	1	7	4	12	13	5	3	8	9
3	9	13	4	12	1	2	7	6	10	3	5	11	8
7	6	13	3	11	4	1	5	8	10	9	2	12	7

TABLE 12. The average ranks of the 13 algorithms for the different types of problems based on IGD metric. A smaller rank indicates a better performance.

M	MSOPSII	MOEAD	HypE	PICEA-g	SPEA2SDE	GrEA	NSGAIII	KnEA	RVEA	Two_Arch2	θ-DEA	MOEADD	AnD
2	8	11	4	7	6	9	2	12	13	5	1	3	10
3	11	12	13	4	7	5	3	6	9	1	2	10	8
7	9	13	12	5	4	1	7	10	8	3	6	11	2

the Wilcoxon-rank sum test). Then, partial orderings of the algorithms can be constructed. For example, if A and B are both better than C, then A and B are assigned a better rank than C. Similarly, if A is comparable to B, then A has the same rank as B. Hence, the algorithms with the best rank are identified as the *best* algorithms (highlighted in grey background in the Tables). The algorithms with the worst rank are identified as the *worst* algorithms (highlighted in bold). A more comprehensive comparison can be obtained in this way than by simply highlighting the algorithm with the highest HV value.

In addition, box plots are used to visualize the HV values of the 31 runs. The bottom and top of the box are the first and third quartiles, whereas the line inside the box indicates the median. The ends of the whiskers represent the lowest datum that is still within 1.5 times the interquartile

TABLE 13. IGD Results (Mean and SD) of the 13 algorithms on the 2-objective WFG problems. The top-ranked algorithms for each problem instance are highlighted in grey background, the worst algorithms highlighted in boldface.

Problem	M	D	MSOPSII	MOEAD	HypE	PICEA-g	SPEA2SDE	GrEA	NSGAIII	KnEA	RVEA	Two_Arch2	θ-DEA	MOEADD	AnD
WFG1	2	11	2.1134e-1	5.2779e-1	**7.2651e-1**	2.0358e-1	1.9256e-1	1.9163e-1	2.7023e-1	2.8776e-1	5.8172e-1	2.5722e-1	2.7203e-1	3.1137e-1	2.8482e-1
			(9.19e-2)	(5.95e-2)	(1.53e-1)	(3.63e-2)	(4.81e-2)	(9.01e-2)	(5.03e-2)	(1.39e-1)	(4.75e-2)	(9.01e-2)	(8.29e-2)	(3.60e-2)	(2.60e-2)
WFG2	2	11	2.6824e-2	1.0926e-1	1.0763e-2	2.5927e-2	1.2562e-2	3.1444e-2	1.5249e-2	**9.1879e-1**	7.7188e-2	1.2996e-2	3.0173e-2	2.4768e-2	3.0270e-2
			(3.18e-3)	(6.86e-2)	(2.78e-4)	(4.83e-2)	(7.44e-4)	(2.06e-3)	(6.28e-4)	(2.70e-1)	(1.07e-2)	(1.97e-3)	(4.77e-2)	(1.68e-3)	(3.96e-3)
WFG3	2	11	1.6767e-2	2.6843e-2	1.2341e-2	1.8146e-2	1.3715e-2	2.3784e-2	1.3469e-2	1.7967e-2	**5.7999e-2**	1.4775e-2	1.2904e-2	1.5832e-2	1.9044e-2
			(9.57e-4)	(4.89e-3)	(3.48e-4)	(1.67e-3)	(3.66e-4)	(3.33e-4)	(8.46e-4)	(8.33e-4)	(1.04e-2)	(1.46e-3)	(5.23e-4)	(1.21e-3)	(1.64e-3)
WFG4	2	11	1.8001e-2	3.6040e-2	1.7771e-2	1.8532e-2	3.1811e-2	2.6055e-2	1.3953e-2	2.5249e-2	**9.4425e-2**	1.6245e-2	1.4012e-2	1.4325e-2	1.9852e-2
			(1.16e-3)	(4.78e-3)	(1.24e-3)	(1.97e-3)	(5.11e-3)	(1.42e-3)	(1.13e-3)	(4.50e-3)	(1.59e-2)	(8.82e-4)	(1.21e-3)	(4.79e-4)	(2.00e-3)
WFG5	2	11	6.6120e-2	7.2322e-2	6.6914e-2	6.5859e-2	7.8220e-2	7.4116e-2	6.4440e-2	7.8605e-2	**1.0122e-1**	6.5940e-2	6.5213e-2	6.7091e-2	6.6924e-2
			(4.58e-4)	(1.45e-3)	(1.48e-3)	(2.22e-3)	(4.69e-3)	(2.13e-3)	(1.01e-3)	(9.41e-3)	(1.40e-2)	(2.36e-3)	(2.14e-3)	(2.75e-3)	(1.91e-3)
WFG6	2	11	7.6567e-2	9.5519e-2	8.1120e-2	9.5516e-2	9.4419e-2	8.0767e-2	8.6392e-2	**3.0893e-1**	1.6972e-1	7.3811e-2	8.0992e-2	8.3543e-2	8.4227e-2
			(2.14e-2)	(2.32e-2)	(2.09e-2)	(1.99e-2)	(1.88e-2)	(2.56e-2)	(2.42e-2)	(6.94e-2)	(2.42e-2)	(2.00e-2)	(1.42e-2)	(1.87e-2)	(1.87e-2)
WFG7	2	11	1.9124e-2	3.3514e-2	1.7886e-2	1.6001e-2	3.5441e-2	2.9776e-2	1.2700e-2	**1.4060e-1**	6.6515e-2	1.6222e-2	1.2699e-2	1.3963e-2	1.9752e-2
			(1.42e-3)	(3.34e-3)	(8.14e-4)	(9.18e-4)	(5.56e-3)	(1.90e-3)	(2.35e-4)	(5.41e-2)	(1.26e-2)	(2.69e-4)	(1.98e-4)	(3.45e-4)	(1.68e-3)
WFG8	2	11	1.1263e-1	1.2663e-1	1.1143e-1	1.2033e-1	1.1775e-1	1.1217e-1	1.1293e-1	**5.0156e-1**	2.0629e-1	1.1728e-1	1.1495e-1	1.1042e-1	1.1835e-1
			(2.64e-3)	(5.63e-3)	(3.79e-3)	(4.01e-3)	(3.46e-3)	(9.73e-4)	(1.70e-3)	(7.25e-2)	(1.51e-2)	(8.79e-3)	(3.42e-3)	(1.98e-3)	(2.65e-3)
WFG9	2	11	4.7621e-2	**7.4091e-2**	2.0663e-2	4.2529e-2	3.5560e-2	3.0308e-2	2.3035e-2	3.7442e-2	6.0373e-2	2.0218e-2	2.1812e-2	2.2160e-2	2.8940e-2
			(7.28e-2)	(5.48e-2)	(1.00e-3)	(5.09e-2)	(5.90e-3)	(2.65e-3)	(1.97e-3)	(4.19e-2)	(5.71e-3)	(2.34e-3)	(2.34e-3)	(2.23e-3)	(3.46e-3)

TABLE 14. IGD results (Mean and SD) of the 13 algorithms on the 3-objective WFG problems. The top-ranked algorithms for each problem instance are highlighted in grey background, the worst algorithms highlighted in boldface.

Problem	M	D	MSOPSII	MOEAD	HypE	PICEA-g	SPEA2SDE	GrEA	NSGAIII	KnEA	RVEA	Two_Arch2	θ-DEA	MOEADD	AnD
WFG1	3	12	3.8619e-1	6.5207e-1	**1.3305e+0**	9.7828e-1	2.9430e-1	3.0458e-1	5.5535e-1	3.7930e-1	6.5424e-1	4.5853e-1	4.7516e-1	1.0232e+0	4.7946e-1
			(7.13e-2)	(9.52e-2)	(1.22e-1)	(1.07e-1)	(5.17e-2)	(4.41e-2)	(7.70e-2)	(5.38e-2)	(6.52e-2)	(1.14e-1)	(6.57e-2)	(1.49e-1)	(4.94e-2)
WFG2	3	12	2.7288e-1	**1.0224e+0**	2.7156e-1	1.5379e-1	2.4798e-1	2.6076e-1	1.8288e-1	2.3662e-1	2.1757e-1	1.5356e-1	2.1064e-1	4.8589e-1	2.4939e-1
			(3.40e-2)	(3.31e-2)	(4.30e-2)	(9.67e-3)	(5.53e-2)	(2.64e-2)	(5.38e-3)	(4.36e-2)	(2.05e-2)	(3.41e-3)	(2.24e-2)	(1.11e-1)	(2.26e-2)
WFG3	3	12	9.7974e-2	**2.0491e-1**	3.7216e-2	1.2526e-1	6.6408e-2	9.1062e-2	1.1934e-1	1.3620e-1	**2.3089e-1**	8.7444e-2	1.3446e-1	**2.6050e-1**	1.5499e-1
			(2.44e-2)	(5.78e-2)	(3.53e-3)	(1.04e-2)	(5.31e-3)	(8.78e-3)	(8.98e-3)	(5.76e-2)	(1.96e-2)	(6.22e-3)	(1.84e-2)	(1.01e-1)	(1.87e-2)
WFG4	3	12	2.6046e-1	2.6322e-1	**3.3366e-1**	2.2341e-1	3.2817e-1	2.4107e-1	2.2294e-1	2.5496e-1	2.4346e-1	2.2780e-1	2.2194e-1	2.4151e-1	2.2851e-1
			(9.62e-3)	(5.94e-3)	(1.48e-2)	(3.05e-3)	(1.37e-2)	(2.99e-3)	(9.79e-4)	(1.04e-2)	(5.85e-3)	(5.49e-3)	(5.44e-4)	(9.34e-4)	(6.08e-3)
WFG5	3	12	2.8036e-1	2.5416e-1	**3.6200e-1**	2.2784e-1	3.3427e-1	2.6071e-1	2.3121e-1	2.6812e-1	2.3700e-1	2.3775e-1	2.3047e-1	2.4583e-1	2.3805e-1
			(9.31e-3)	(3.69e-3)	(1.19e-2)	(3.29e-3)	(1.60e-2)	(4.44e-3)	(4.39e-4)	(1.54e-2)	(2.78e-3)	(3.98e-3)	(8.14e-4)	(1.68e-3)	(4.09e-3)
WFG6	3	12	3.2058e-1	2.9976e-1	**3.7273e-1**	2.6303e-1	3.5553e-1	2.7211e-1	2.5119e-1	3.0266e-1	2.7282e-1	2.5355e-1	2.4670e-1	2.6101e-1	2.5513e-1
			(1.71e-2)	(8.25e-3)	(2.28e-2)	(2.14e-2)	(1.97e-2)	(9.51e-3)	(1.27e-2)	(1.45e-2)	(1.72e-2)	(1.38e-2)	(1.13e-2)	(1.29e-2)	(1.69e-2)
WFG7	3	12	2.7116e-1	3.7391e-1	**3.8358e-1**	2.1863e-1	3.2703e-1	2.5514e-1	2.2287e-1	2.5276e-1	2.3948e-1	2.2493e-1	2.2219e-1	2.4466e-1	2.2943e-1
			(1.35e-2)	(4.54e-2)	(1.44e-2)	(3.58e-3)	(1.41e-2)	(9.13e-3)	(4.19e-4)	(1.36e-2)	(5.28e-3)	(4.51e-3)	(4.77e-4)	(1.85e-3)	(4.27e-3)
WFG8	3	12	**3.9153e-1**	3.2510e-1	3.7156e-1	3.0942e-1	3.6055e-1	3.0201e-1	2.9562e-1	3.3820e-1	3.2874e-1	3.1132e-1	2.9379e-1	3.0529e-1	3.2855e-1
			(1.25e-2)	(1.10e-2)	(1.41e-2)	(4.52e-3)	(1.11e-2)	(8.89e-3)	(5.07e-3)	(1.29e-2)	(1.62e-2)	(5.61e-3)	(4.59e-3)	(3.66e-3)	(9.38e-3)
WFG9	3	12	2.5643e-1	3.0339e-1	**3.6226e-1**	2.2134e-1	3.1156e-1	2.3984e-1	2.3560e-1	2.2941e-1	2.3645e-1	2.2152e-1	2.3258e-1	2.3941e-1	2.2678e-1
			(3.01e-2)	(3.76e-2)	(1.31e-2)	(1.10e-2)	(1.39e-2)	(5.39e-3)	(3.10e-2)	(5.63e-3)	(6.77e-3)	(3.75e-3)	(3.04e-2)	(1.92e-3)	(9.73e-3)

TABLE 15. IGD results (Mean and SD) of the 13 algorithms on the 7-objective WFG problems. The top-ranked algorithms for each problem instance are highlighted in grey background, the worst algorithms highlighted in boldface.

Problem	M	D	MSOPSII	MOEAD	HypE	PICEA-g	SPEA2SDE	GrEA	NSGAIII	KnEA	RVEA	Two_Arch2	θ-DEA	MOEADD	AnD
WFG1	7	16	1.1701e+0	2.1200e+0	**2.5278e+0**	**2.4666e+0**	1.1335e+0	1.3179e+0	1.5770e+0	1.2925e+0	1.3695e+0	1.6191e+0	1.2951e+0	1.8633e+0	1.2592e+0
			(9.67e-2)	(2.52e-1)	(1.31e-1)	(5.13e-2)	(9.70e-2)	(1.90e-1)	(1.43e-1)	(1.32e-1)	(1.27e-1)	(1.57e-1)	(2.93e-1)	(1.17e-1)	(1.07e-1)
WFG2	7	16	3.1401e+0	**1.0638e+1**	3.8990e+0	2.0458e+0	6.1501e+0	3.0014e+0	3.3498e+0	2.1641e+0	5.3945e+0	2.0238e+0	3.6040e+0	8.9116e+0	3.2923e+0
			(7.86e-1)	(1.18e-1)	(6.24e-1)	(3.88e-1)	(1.27e+0)	(6.88e-1)	(2.17e+0)	(3.84e-1)	(1.05e+0)	(4.16e-1)	(1.65e+0)	(2.73e-1)	(1.03e+0)
WFG3	7	16	1.8591e-1	**3.0516e+0**	9.2976e-2	8.7637e-1	1.2527e+0	8.8233e-1	1.2086e+0	1.5650e+0	1.9303e+0	9.5599e-1	1.2243e+0	1.7802e+0	1.1449e+0
			(4.01e-2)	(1.70e-1)	(9.16e-3)	(8.43e-2)	(4.51e-1)	(1.37e-1)	(2.95e-1)	(5.84e-1)	(5.32e-1)	(1.15e-1)	(2.03e-1)	(1.35e-1)	(1.66e-1)
WFG4	7	16	2.7841e+0	**6.0011e+0**	4.4275e+0	2.5209e+0	2.7668e+0	2.4768e+0	2.6653e+0	2.8293e+0	2.6378e+0	2.5980e+0	2.6539e+0	2.9280e+0	2.5652e+0
			(3.14e-2)	(1.65e-1)	(5.97e-1)	(1.51e-1)	(3.53e-2)	(1.66e-2)	(4.57e-2)	(3.81e-2)	(1.09e-2)	(2.16e-2)	(1.08e-2)	(7.84e-2)	(2.35e-2)
WFG5	7	16	2.9268e+0	**5.7344e+0**	2.9075e+0	2.4616e+0	2.7503e+0	2.4734e+0	2.6097e+0	2.8500e+0	2.6167e+0	2.5379e+0	2.6104e+0	3.0551e+0	2.5644e+0
			(8.08e-2)	(1.30e-1)	(7.70e-2)	(1.90e-2)	(4.16e-2)	(2.12e-2)	(7.85e-3)	(3.74e-2)	(8.80e-3)	(2.43e-2)	(7.90e-3)	(1.05e-1)	(1.97e-2)
WFG6	7	16	2.9209e+0	**6.1898e+0**	2.9564e+0	2.5041e+0	2.8708e+0	2.5246e+0	2.6625e+0	3.0413e+0	2.6422e+0	2.6144e+0	2.6520e+0	2.9873e+0	2.6353e+0
			(5.61e-2)	(1.35e-1)	(1.32e-1)	(1.66e-2)	(4.93e-2)	(2.33e-2)	(1.85e-2)	(8.66e-2)	(3.43e-2)	(3.18e-2)	(1.55e-2)	(9.44e-2)	(3.29e-2)
WFG7	7	16	2.9244e+0	**6.1043e+0**	3.2130e+0	2.4754e+0	2.7981e+0	2.5062e+0	2.6623e+0	2.9069e+0	2.6454e+0	2.5451e+0	2.6609e+0	2.9292e+0	2.6260e+0
			(6.22e-2)	(1.34e-1)	(2.53e-1)	(1.69e-2)	(4.18e-2)	(1.22e-2)	(9.73e-3)	(5.22e-2)	(1.85e-2)	(2.17e-2)	(1.14e-2)	(1.13e-1)	(3.41e-2)
WFG8	7	16	2.9926e+0	**5.3740e+0**	3.3965e+0	2.7070e+0	2.8302e+0	2.5975e+0	2.6761e+0	2.8775e+0	2.6752e+0	2.8671e+0	2.6073e+0	2.8237e+0	2.6307e+0
			(4.02e-2)	(1.53e-1)	(1.90e-1)	(1.28e-1)	(3.60e-2)	(2.52e-2)	(1.67e-1)	(7.35e-2)	(4.91e-2)	(3.88e-2)	(1.21e-2)	(2.01e-2)	(2.59e-2)
WFG9	7	16	2.7560e+0	**5.5685e+0**	2.8930e+0	2.5480e+0	2.6709e+0	2.4286e+0	2.5417e+0	2.6597e+0	2.5686e+0	2.5832e+0	2.5362e+0	3.1163e+0	2.4784e+0
			(3.94e-2)	(4.40e-1)	(2.03e-1)	(4.65e-2)	(3.09e-2)	(1.46e-2)	(1.99e-2)	(4.73e-2)	(3.13e-2)	(3.74e-2)	(7.77e-3)	(1.15e-1)	(2.39e-2)

range (IQR) of the lower quartile and the highest datum that is still within 1.5 IQR of the upper quartile [63]. Moreover, outlying values are marked as "+". Taking the WFG problems as an example, box plots of the HV values are shown

TABLE 16. IGD results (Mean and SD) of the 13 algorithms on the 2-objective MaF problems. The top-ranked algorithms for each problem instance are highlighted in grey background, the worst algorithms highlighted in boldface.

Problem	M	D	MSOPSII	MOEAD	HypE	PICEA-g	SPEA2SDE	GrEA	NSGAIII	KnEA	RVEA	Two_Arch2	θ-DEA	MOEADD	AnD
MaF1	2	11	5.6810e-3	3.5713e-3	3.6811e-3	3.8228e-3	3.9873e-3	**7.8323e-3**	3.5730e-3	5.0649e-3	3.6031e-3	4.0003e-3	3.5717e-3	3.5711e-3	4.8193e-3
			(2.82e-4)	(1.21e-7)	(1.28e-5)	(3.90e-5)	(6.27e-5)	(6.05e-5)	(2.33e-6)	(2.14e-4)	(3.29e-5)	(3.10e-6)	(4.88e-7)	(9.24e-8)	(2.25e-4)
MaF2	2	11	3.0294e-3	2.5850e-3	2.0589e-3	2.2134e-3	2.3771e-3	3.9900e-3	2.0506e-3	**2.4517e-2**	2.7615e-3	2.2410e-3	2.0178e-3	4.4431e-3	2.5245e-3
			(1.52e-4)	(1.54e-4)	(1.27e-5)	(3.00e-5)	(6.38e-5)	(8.78e-5)	(7.13e-5)	(7.69e-3)	(1.61e-4)	(2.06e-5)	(7.51e-6)	(1.38e-4)	(9.47e-5)
MaF3	2	11	1.8291e+0	6.6597e-1	1.6633e+0	**2.1880e+2**	7.5481e-1	1.0699e+0	3.2158e+0	9.2754e-1	**1.9819e+2**	2.0891e+1	1.0912e+1	4.1478e+1	**1.2351e+4**
			(2.43e+0)	(1.32e+0)	(2.68e+0)	(1.38e+2)	(1.20e+0)	(2.08e+0)	(7.66e+0)	(1.16e+0)	(2.81e+2)	(2.46e+1)	(3.26e+1)	(3.78e+1)	(1.91e+4)
MaF4	2	11	2.9184e-1	7.4127e-1	2.1209e-1	9.7959e-1	3.4737e-1	7.0168e-1	7.5005e-1	4.6349e-1	**2.3290e+0**	**2.1026e+0**	2.1240e-1	**2.1883e+0**	**2.9545e+0**
			(3.55e-1)	(6.20e-1)	(3.69e-1)	(1.05e+0)	(4.34e-1)	(6.75e-1)	(9.42e-1)	(3.83e-1)	(1.80e+0)	(2.61e+0)	(3.12e-1)	(1.83e+0)	(1.51e+0)
MaF5	2	11	**5.4905e-1**	1.4614e-1	**6.2759e-1**	1.4596e-1	1.6317e-1	3.2002e-2	1.4533e-1	**5.7085e-1**	1.3733e-2	1.6099e+0	**4.1147e-1**	1.3105e-2	1.5941e-2
			(9.11e-1)	(5.15e-1)	(8.89e-1)	(5.15e-1)	(5.10e-1)	(1.18e-3)	(5.15e-1)	(8.97e-1)	(1.25e-3)	(8.25e-1)	(8.26e-1)	(1.20e-6)	(1.23e-3)
MaF6	2	11	7.5531e-2	4.1426e-3	9.5642e-3	4.5703e-3	1.0576e-2	9.8516e-3	4.0193e-3	**1.4833e-1**	7.5484e-3	5.1792e-3	4.0186e-3	4.0359e-3	6.2367e-3
			(5.84e-2)	(1.52e-4)	(3.88e-3)	(3.10e-4)	(1.32e-3)	(6.53e-4)	(3.97e-5)	(1.80e-2)	(8.55e-4)	(1.66e-4)	(6.45e-5)	(4.73e-5)	(7.01e-4)
MaF7	2	21	7.7581e-2	7.0297e-2	**3.2599e-1**	3.4671e-2	5.2226e-3	2.9774e-2	6.8521e-3	3.4033e-2	2.9131e-2	6.3284e-2	5.1101e-3	2.0118e-2	9.3801e-3
			(1.52e-1)	(1.53e-1)	(2.01e-1)	(1.14e-1)	(2.02e-4)	(7.39e-3)	(1.74e-4)	(1.48e-2)	(4.64e-3)	(1.54e-1)	(7.25e-5)	(1.15e-3)	(8.54e-4)

TABLE 17. IGD results (Mean and SD) of the 13 algorithms on the 3-objective MaF problems. The top-ranked algorithms for each problem instance are highlighted in grey background, the worst algorithms highlighted in boldface.

Problem	M	D	MSOPSII	MOEAD	HypE	PICEA-g	SPEA2SDE	GrEA	NSGAIII	KnEA	RVEA	Two_Arch2	θ-DEA	MOEADD	AnD
MaF1	3	12	5.4087e-2	7.0482e-2	**8.5055e-2**	4.1616e-2	4.2015e-2	4.0369e-2	6.1660e-2	4.8449e-2	**8.2252e-2**	4.1494e-2	**8.0433e-2**	**7.8236e-2**	4.3822e-2
			(5.31e-3)	(1.70e-5)	(5.82e-3)	(4.72e-4)	(6.20e-4)	(9.01e-4)	(1.91e-3)	(7.05e-3)	(2.56e-4)	(4.28e-4)	(9.64e-4)	(1.97e-3)	(6.16e-4)
MaF2	3	12	3.7213e-2	4.1389e-2	4.5576e-2	3.0403e-2	3.0958e-2	3.1316e-2	3.6691e-2	3.4179e-2	4.2260e-2	2.9120e-2	3.6560e-2	**5.5815e-2**	3.0384e-2
			(1.18e-3)	(1.20e-3)	(1.71e-3)	(7.07e-4)	(7.86e-4)	(6.91e-4)	(8.54e-4)	(1.66e-3)	(1.36e-3)	(4.68e-4)	(4.26e-4)	(2.08e-3)	(4.94e-4)
MaF3	3	12	5.4802e+0	9.0333e-1	3.7068e+0	2.6300e+1	5.5117e-1	2.3723e+0	4.0919e+0	2.3301e+0	**8.0373e+2**	1.2461e+1	2.8153e+0	2.6562e+1	**7.9681e+3**
			(7.03e+0)	(1.18e+0)	(5.44e+0)	(1.81e+1)	(9.40e-1)	(3.09e+0)	(3.51e+0)	(5.10e+0)	(1.33e+3)	(1.32e+1)	(4.86e+0)	(2.73e+1)	(9.75e+3)
MaF4	3	12	2.1426e+0	2.2734e+0	2.6851e+0	**6.2649e+0**	1.8903e+0	3.0299e+0	4.6438e+0	1.5476e+0	3.7054e+0	**5.4989e+0**	1.7999e+0	2.0123e+0	**8.6721e+0**
			(1.56e+0)	(1.05e+0)	(2.48e+0)	(5.48e+0)	(1.98e+0)	(3.21e+0)	(3.28e+0)	(1.60e+0)	(5.88e+0)	(3.61e+0)	(2.72e+0)	(2.42e+0)	(5.21e+0)
MaF5	3	12	7.7765e-1	**1.5838e+0**	**1.6413e+0**	7.7898e-1	6.6023e-1	5.3953e-1	4.9540e-1	3.1111e-1	2.6045e-1	2.5404e-1	6.6639e-1	2.9684e-1	2.6257e-1
			(6.97e-1)	(1.22e+0)	(1.09e+0)	(6.90e-1)	(6.28e-1)	(5.10e-1)	(6.22e-1)	(9.91e-3)	(9.96e-4)	(6.40e-3)	(7.11e-1)	(1.89e-4)	(5.32e-3)
MaF6	3	12	2.7059e-2	9.9411e-2	**1.9634e-1**	4.5807e-3	9.5724e-3	2.0980e-2	1.4992e-2	4.7401e-2	5.1103e-2	5.8424e-3	3.3400e-2	3.0559e-2	6.2634e-2
			(5.22e-2)	(1.48e-1)	(2.54e-2)	(3.02e-4)	(1.19e-3)	(5.31e-4)	(1.56e-3)	(3.94e-2)	(2.43e-2)	(2.83e-4)	(2.28e-3)	(1.45e-3)	(1.53e-2)
MaF7	3	22	1.4040e-1	1.5432e-1	**8.2187e-1**	3.7759e-1	5.8598e-2	8.2876e-2	7.6744e-2	6.8147e-2	1.0807e-1	9.6961e-2	1.1040e-1	5.0578e-1	8.5394e-2
			(1.21e-2)	(1.74e-3)	(5.03e-3)	(2.65e-1)	(2.56e-3)	(4.07e-3)	(2.96e-3)	(6.14e-3)	(2.83e-3)	(1.03e-1)	(6.96e-2)	(2.54e-1)	(3.08e-3)

TABLE 18. IGD results (Mean and SD) of the 13 algorithms on the 7-objective MaF problems. The top-ranked algorithms for each problem instance are highlighted in grey background, the worst algorithms highlighted in boldface.

Problem	M	D	MSOPSII	MOEAD	HypE	PICEA-g	SPEA2SDE	GrEA	NSGAIII	KnEA	RVEA	Two_Arch2	θ-DEA	MOEADD	AnD
MaF1	7	16	2.3936e-1	**4.7689e-1**	2.9999e-1	2.1617e-1	2.0528e-1	2.2143e-1	2.5592e-1	2.0477e-1	**4.9956e-1**	2.0759e-1	2.6282e-1	3.3403e-1	2.1735e-1
			(1.83e-2)	(3.74e-2)	(6.63e-3)	(3.34e-3)	(2.51e-3)	(4.66e-3)	(1.70e-2)	(5.09e-3)	(7.00e-2)	(4.22e-3)	(5.71e-3)	(3.01e-2)	(1.24e-3)
MaF2	7	16	1.7082e-1	2.0985e-1	**4.2959e-1**	2.1610e-1	1.6039e-1	1.6573e-1	1.9632e-1	1.6342e-1	**4.5886e-1**	1.6222e-1	2.0272e-1	2.2763e-1	1.5759e-1
			(4.50e-3)	(2.95e-3)	(2.29e-2)	(4.08e-2)	(7.08e-3)	(3.06e-3)	(1.60e-2)	(7.03e-3)	(1.47e-1)	(3.64e-3)	(1.32e-2)	(2.34e-2)	(5.08e-3)
MaF3	7	16	2.3979e+1	4.6374e-1	1.0999e+5	**1.3124e+9**	2.0896e+0	2.7940e+5	4.4695e+2	6.3674e+6	5.7476e+1	2.0884e+5	2.0947e+1	2.0362e+2	7.4302e+3
			(5.05e+1)	(4.54e-1)	(1.62e+5)	(1.32e+9)	(5.61e+0)	(6.58e+5)	(5.91e+2)	(1.69e+7)	(5.02e+1)	(7.07e+5)	(1.80e+1)	(1.63e+2)	(1.38e+4)
MaF4	7	16	1.5613e+2	7.2359e+1	7.5916e+1	**2.9488e+2**	2.4424e+1	6.4112e+1	1.3026e+2	**3.0335e+2**	4.0914e+1	1.5007e+2	4.9970e+1	4.5604e+1	9.2275e+1
			(1.33e+2)	(5.27e+0)	(5.92e+1)	(3.36e+2)	(1.37e+1)	(6.69e+1)	(1.42e+2)	(2.66e+2)	(1.97e+1)	(1.35e+2)	(4.01e+1)	(1.04e+1)	(8.02e+1)
MaF5	7	16	2.0851e+1	**4.4755e+1**	1.9667e+1	1.0756e+1	1.0998e+1	8.6030e+0	1.2110e+1	1.2917e+1	1.5123e+1	8.9530e+0	1.1825e+1	**3.9432e+1**	9.5665e+0
			(8.71e+0)	(2.70e+0)	(3.79e+0)	(3.50e+0)	(2.47e+0)	(2.44e-1)	(9.47e-1)	(5.11e-1)	(3.26e+0)	(2.39e-1)	(5.89e-1)	(2.30e+0)	(4.03e-1)
MaF6	7	16	1.4533e-2	**4.1364e-1**	2.0197e-1	4.3817e-3	1.0291e-2	8.2390e-2	2.0767e-1	**4.5005e-1**	1.3148e-1	7.6470e-3	1.4603e-1	1.2960e-1	**3.1191e-1**
			(2.29e-3)	(2.02e-1)	(3.26e-2)	(1.77e-4)	(1.33e-3)	(1.54e-1)	(2.67e-1)	(1.62e+0)	(2.53e-2)	(7.45e-4)	(5.93e-2)	(1.10e-2)	(6.39e-2)
MaF7	7	26	9.2744e-1	1.3566e+0	**3.2770e+0**	**3.0301e+0**	5.4030e-1	7.1338e-1	7.2494e-1	5.0694e-1	1.2352e+0	5.5575e-1	7.1364e-1	2.0982e+0	5.6616e-1
			(2.52e-1)	(1.75e-1)	(2.32e-1)	(8.23e-1)	(9.29e-3)	(7.15e-2)	(3.74e-2)	(1.38e-2)	(2.41e-1)	(3.59e-2)	(7.91e-2)	(5.77e-1)	(2.63e-2)

in Fig. 2. From the box plots, the ranks of different algorithms can be graphically inferred, consistent with the statistical results obtained by the method above.

A. PERFORMANCE COMPARISONS ON WFG PROBLEMS

The WFG1 problem is designed with a flat bias and a PF containing both convex and concave segments. This test problem is used to assess whether EMO algorithms are capable of dealing with PFs of complicated mixed geometries. GrEA and SPEA2+SDE are jointly ranked in the first class on the 2-objective WFG1. SPEA2+SDE also outperforms the other algorithms on the 3-objective WFG1. For the 7-objective WFG1, MSOPS-II and HypE perform the best. However, MOEA/D, together with RVEA, exhibits the worst performance on the 2-objective WFG1. Moreover, PICEA-g has the lowest HV value for the 3- and 7-objective WFG1.

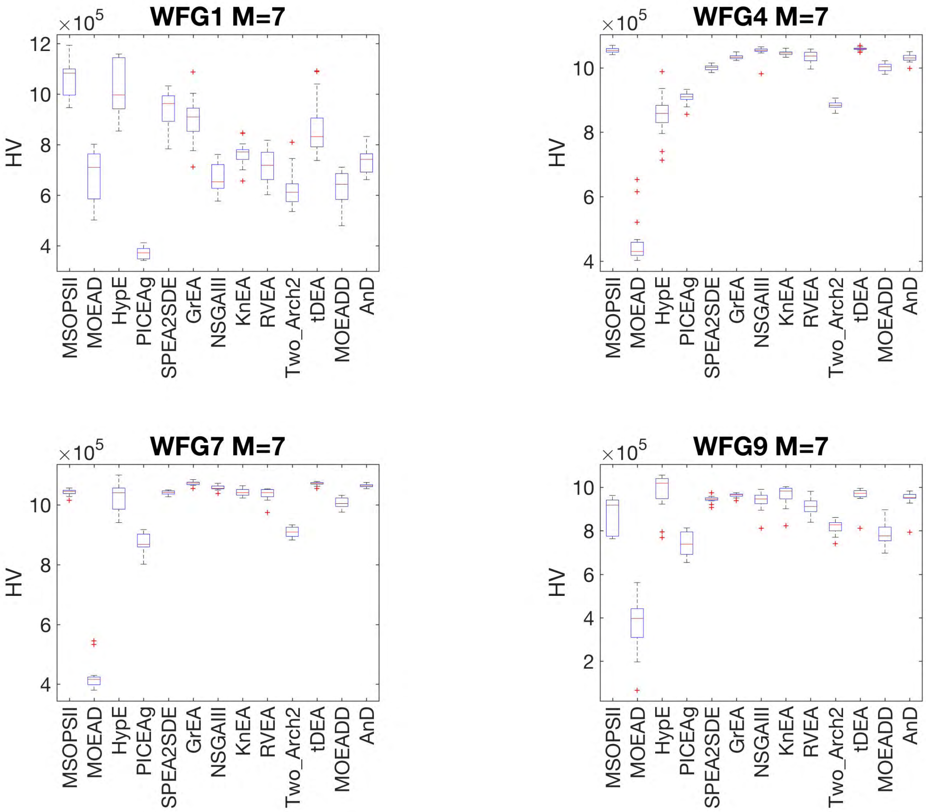

FIGURE 2. Box plots of the distribution of the 31 HV values for the WFG1, WFG4, WFG7, and WFG9 problems with 7 objectives of the 13 test algorithms.

TABLE 19. HV results of MOEA/D with normalization on 3- and 7-objective WFG problems. The top-ranked algorithms for each problem instance are highlighted in grey background, the worst algorithms highlighted in boldface.

Problem	M	D	MOEA/D-normalization	Problem	M	D	MOEA/D-normaliztion
WFG1	3	12	5.4485e+1 (2.42e+0)	WFG1	7	16	**4.5272e+5 (6.49e+4)**
WFG2	3	12	**5.6872e+1 (4.28e-1)**	WFG1	7	16	**1.1056e+6 (1.10e+5)**
WFG3	3	12	6.4578e+0 (3.10e-2)	WFG1	7	16	**1.6112e-1 (4.02e-2)**
WFG4	3	12	**3.3519e+1 (3.00e-1)**	WFG1	7	16	**6.4908e+5 (5.19e+4)**
WFG5	3	12	**3.0441e+1 (2.00e-1)**	WFG1	7	16	**5.9850e+5 (3.83e+4)**
WFG6	3	12	**2.9770e+1 (1.33e+0)**	WFG1	7	16	**5.4729e+5 (8.61e+4)**
WFG7	3	12	**3.3413e+1 (2.59e-1)**	WFG1	7	16	**6.0718e+5 (6.16e+4)**
WFG8	3	12	2.8746e+1 (1.91e-1)	WFG1	7	16	**5.3131e+5 (1.73e+4)**
WFG9	3	12	**3.1458e+1 (2.15e+0)**	WFG1	7	16	**4.9883e+5 (6.85e+4)**

The WFG2 problem has a scaled disconnected PF, which enables an assessment of whether EMO algorithms are capable of dealing with scaled disconnected PFs. Most of the test algorithms work well on the 2-objective WFG2, and KnEA performs the worst. In terms of the 3-objective WFG2, HypE, SPEA2+SDE, and Two_Arch2 are the top-performing algorithms, whereas MOEA/D performs the worst. Notably, MOEA/D shows the worst performances on most of the WFG test problems with 3 and 7 objectives. In addition, HypE, AnD, and MSOPS-II are ranked first on the 7-objective WFG2.

WFG3 is a difficult problem with a partially degenerate front and non-separable decision variables. Whereas WFG3 was proposed as a degenerate test problem, its Pareto front is not a degenerate Pareto front as studied in [64]. In terms of the 2-objective WFG3, most algorithms have comparable performance. However, seven of the testing algorithms, including MOEA/D and SPEA2+SDE, fail to solve the 7-objective WFG3, and HypE performs the best. HypE is also ranked first for the 3-objective WFG3.

The WFG4-WFG9 problems have an identical PF shape with different characteristics. For example, the characteristic of multimodality is introduced in the WFG4 problem, which can make MOEAs trap in local optima easily. WFG5 is a deceptive problem with the global optimum in an unlikely place, which increases the difficulty of searching the true optimum. WFG6 is a non-separable problem with decision variables that are linked with other variables; this type of problem can be more difficult to tackle than separable problems. WFG7-9 adopt some bias, which increases the challenge of maintaining diversity. WFG7 is a unimodal and separable problem. By contrast, WFG8 and WFG9 are non-separable problems. WFG9 is a non-separable, multimodal, deceptive and bias problem and thus is the most difficult among the WFG problems.

For problems WFG4 to WFG9 with 2 objectives, the testing algorithms show similar performance: most of the algorithms show comparable performance, except MOEA/D, KnEA and RVEA. For the 3-objective WFG4-9, HypE is always among the top-performing algorithms, except for WFG8, for which SPEA2+SDE and GrEA perform the best.

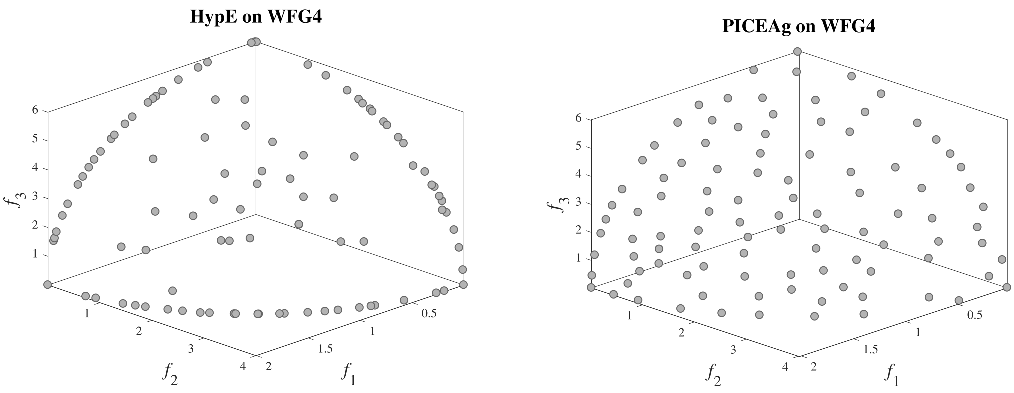

FIGURE 3. Distribution of solutions obtained by HypE and PICEA-g on a 3-objective WFG problem.

SPEA2+SDE and GrEA are also ranked in the first class for the 3-objective WFG7. In addition, both NSGA-III and θ-DEA are ranked first for the 3-objective WFG5 and WFG7. Moreover, Two_Arch2 and AnD are both ranked first for the 3-objective WFG6 and WFG7. For the 7-objective WFG4-9, HypE, NSGA-III and θ-DEA exhibit comparable performance. NSGA-III and θ-DEA show similar behaviour, and both work well on WFG4-7. For the most difficult WFG9 problem, HypE outperforms the other algorithms, whereas MOEA/D performs the worst.

Overall, several interesting and insightful conclusions on the performance of the 13 test algorithms on the WFG problem suite can be drawn, and the IGD metric is also taken into consideration. Here, we assume the PF of WFG3 is a degenerate line to facilitate the calculation of the IGD value, although this assumption may not be true, as pointed out by Ishibuchi *et al.* [64].

1) For some test problems, the results for the IGD metric show a different behaviour compared with the HV metric. For instance, in terms of the HV value, HypE is among the top-performing algorithms on most 3-objective WFG problems (WFG2-7 and WFG9). However, HypE obtains the worst IGD values for those problems, e.g., WFG4-7 and WFG9. Moreover, PICEA-g achieves the best IGD values for the 3-objective WFG4 and WFG6 but the worst HV values for those problems.

   The dissimilar performances of the HV and IGD metrics can be explained as follows. If the convergence to the Pareto front of a test problem is not difficult, the performance comparison results are mainly based on the diversity of solutions over the Pareto front. However, the diversity assessment mechanism of the HV metric differs from that of the IGD metric, as can be inferred from their calculation formulas. To provide a more intuitive explanation, Fig. 3 shows the solutions of HypE and PICEA-g for the 3-objective WFG4 problem. For this problem, HypE achieves the best HV value and the worst IGD value, whereas PICEA-g achieves the best IGD value and the worst HV value. The results show that the solutions of the two algorithms both converge well (with values of GD [65], a convergence metric, of 0.00215 and 0.00220, respectively). However, they show obviously different behaviours in the distribution of solutions, with the HV metric preferring solutions located on the boundary. In this case, the IGD metric shows a more conceivable behaviour of the diversity measure. Therefore, for problems in which the convergence of solutions to the Pareto front is relatively easy (like most WFGs), the difference in the diversity assessment mechanisms of HV and IGD is the key reason for the dissimilar performances of the HV and IGD metrics.

2) Most of the test algorithms work well and show comparable performance on the **2-objective WFG problems**. However, KnEA and RVEA exhibit slightly lower performance than the other algorithms.

   This poorer performance might occur because the knee points concept used during mating selection and environmental selection in KnEA and the angle-penalized distance adopted in RVEA are specifically designed to solve MaOPs in the high-dimensional objective space. Consequently, these mechanisms may not work well for easy problems, such as a WFG with 2 objectives. Moreover, the behaviours of the HV and IGD values are nearly identical for the 2-objective WFG problems.

3) In terms of the HV metric, HypE performs the best for the **3-objective WFG problem suite** (ranked first for 7 problems). SPEA2+SDE shows good performance as well (ranked first for 4 problems). In terms of the IGD metric, PICEA-g and NSGA-III perform the best on the 3-objective WFG problem suite (both ranked first for 6 problems). Two_Arch2 and θ-DEA also work well and are ranked first for four and five of the nine 3-objective WFG problems, respectively.

4) In terms of the HV metric, HypE exhibits the best performance for the **7-objective WFG problem suite** (ranked first for 7 problems). NSGA-III and θ-DEA are also among the top-performing algorithms for WFG4-7. By contrast, PICEA-g and GrEA perform the best for the 7-objective WFG problem suite in terms of the IGD metric. In terms of both the HV and

IGD metrics, MOEA/D performs the worst among the 13 algorithms on most 7-objective WFG problems. The lack of a normalization mechanism may be the reason for the poor performance of MOEA/D. However, after conducting experiments to test the performance of MOEA/D with normalization, the results in Table 19 show that MOEA/D with normalization still performs the worst on most problem instances among the 13 algorithms. It even shows a worse performance on some problems than MOEA/D without normalization. This phenomenon is also pointed out by [66], that the incorporation of a normalization mechanism into MOEA/D possibly has unexpected negative and positive effects on the performance of MOEA/D. Note that the normalization is performed by using the estimated nadir and ideal point that are constructed by Pareto optimal solutions in the offline archive, which is a naive normalization procedure. We think the main reason for the poor performance of MOEA/D may be the inappropriate setting of the number of neighbourhood size T, as explained in [41]. It could also be the use of the naive normalization procedure.

5) In conclusion, HypE achieves the top-ranked HV values for most 2-, 3- and 7-objective WFG problems. HypE may have an unfair advantage over other algorithms as it was developed based on the HV indicator. By comparison, PICEA-g always achieves the top-ranked IGD values on the WFG2 and WFG4-7 problems with all 2, 3 and 7 objectives. For PICEA-g, the advanced idea of coevolving a family of preferences simultaneously with the population of candidate solutions is responsible for its excellent ability to generate a diverse distribution of solutions. As the convergence of solutions to the Pareto front is not difficult for most of the WFG1-9 problems, the superiority of PICEA-g in maintaining diversity makes it a top algorithm for most WFG problems.

6) In addition, some of the recently proposed approaches, such as MOEA/DD, did not show outstanding performance, in contrast to their initial reports. A likely reason is the setting of the maximum number of function evaluations ($FEvals^{max}$) [51]. For example, for HypE, a value of $FEvals^{max}$ of 1×10^4 was applied in its paper published in [21]. However, in more recent papers, such as those for MOEA/DD, $FEvals^{max}$ is set to 5.5×10^5, 55 times larger. As users cannot always set $FEvals^{max}$ to a large number for real-world problems due to the large computational cost, $FEvals^{max}$ is set to 2.5×10^4 for all algorithms in this work. This difference in the value of $FEvals^{max}$ might be responsible for the poor performance of some recently proposed algorithms.

B. PERFORMANCE COMPARISONS ON MAF PROBLEMS

MaF1 is a linear problem with an inverted PF, which means that there is no single optimal solution in any subset of objectives. Most algorithms perform comparably on the 2-objective MaF1, for which GrEA exhibits slightly worse performance. For the 3-objective MaF1, the top-performing algorithms are PICEA-g, SPEA2+SDE, GrEA and Two_Arch2. SPEA2+SDE and Two_Arch2 are also ranked in the first class for the 7-objective MaF1.

MaF2 is also a problem with no single optimal solution in any subset of objectives. In contrast to MaF1, MaF2 is a concave problem. Most of the test algorithms have comparable performance for the 2-objective MaF2, for which KnEA performs slightly worse. The behaviour of the test algorithms on the 3-objective MaF2 is similar to that for MaF1, with SPEA2+SDE, GrEA and Two_Arch2 showing the best performance. In contrast to the results for MaF1, HypE outperforms the other algorithms on the 7-objective MaF2, whereas RVEA performs the worst.

MaF3 is a multimodal problem with a convex PF. In general, the chosen test algorithms show worse performance on MaF3 than on the other MaFs. SPEA2+SDE performs the best on all of the 2-, 3- and 7-objective MaF3 problems. MSOPS-II and MOEA/D are also ranked in the first class for the 7-objective MaF3, whereas the other algorithms fail to achieve a converged solution for this problem. Moreover, AnD fails to converge for all of the 2-, 3- and 7-objective MaF3 problems.

MaF4 is a multimodal problem with a badly scaled PF. It also has no single optimal solution in any subset of objectives, similar to MaF1 and MaF2. In terms of the 2-objective MaF4, HypE and θ-DEA show the best performance. θ-DEA is also ranked first on the 3- and 7-objective MaF4. Interestingly, MOEA/DD is among the top-performing algorithms on the 3-objective MaF4 but performs the worst on the 2-objective MaF4. SPEA2+SDE and KnEA also work well on the 3-objective MaF4. In addition, AnD performs the worst for all of the 2-, 3- and 7-objective MaF4 problems.

MaF5 is a biased problem with a convex and badly scaled PF. When the objective number is two or three, most algorithms have comparable performance. Two_Arch2 shows the worst performance on the 2-objective MaF5, whereas MOEA/D and HypE exhibit poor performance on the 3-objective MaF5. In terms of the 7-objective MaF5, the top-performing algorithms are GrEA, NSGA-III, θ-DEA, and AnD.

MaF6 has a degenerate PF. Most algorithms exhibit similar performance on the 2-objective MaF6, for which KnEA performs the worst. For both the 3- and 7-objective MaF6, MSOPS-II, PICEA-g, SPEA2+SDE and Two_Arch2 are always the top-performing algorithms. KnEA is also ranked in the first class on the 7-objective MaF6.

MaF7 has a mixed disconnected Pareto front and is also characterized by a multimodal property. As shown in Tables 6 and 7, there are respectively five top-performing algorithms for the 2- and 3-objective MaF7, with SPEA2+SDE, NSGA-III and Two_Arch2 working well for both the 2- and 3-objective MaF7. In terms of the 7-objective MaF7, only GrEA and AnD exhibit superior performance compared to the other algorithms. AnD, which shows a poor performance on MaF3 and MaF4, is always among the top-performing algorithms on MaF7 problem.

In summary, several interesting and insightful conclusions on the 13 test algorithms can be drawn based on the results for the MaF problem suite:

1) In contrast to the WFG problems, there is no severe difference between the behaviours of the HV metric and the IGD metric on the MaF problems, primarily because the convergence of the solutions to the Pareto front is more difficult for most MaFs compared to the WFG problems. In this case, the performance comparison results are mainly based on the convergence ability of each algorithm. As a result, the difference in the performance comparison results between HV and IGD is small since the distance of the solutions to the Pareto front may have much larger effects on the performance comparison results than the distribution of solutions over the Pareto front.
2) AnD shows a poor performance on MaF3 and MaF4 problems, both having a large number of local Pareto-optimal fronts. However, it can always perform the best on MaF7 problem, which has a disconnected PF. We can see that the combination of angle-based selection strategy and SDE strategy makes AnD a top algorithm for solving problems with disconnected PFs. However, it is not a good choice for highly multimodal problems.
3) Although MOEA/D performs poorly on the WFG problem suite due to the lack of objective space normalization, it is ranked first for four problems of the 2-objective MaF in terms of both HV and IGD.
4) In contrast to the WFG suite, for which KnEA and RVEA perform poorly in most instances, RVEA achieves the top-ranked performance for most of the 2-objective MaF problems, whereas KnEA is ranked in the first class for nearly half of the 3- and 7-objective MaF problems.
5) Compared to the WFG suite, HypE shows much worse performance for the 3- and 7-objective MaF problems even when evaluated by the HV metric, even though HypE was developed based on the HV value. The reason for this poor performance may be that the convergence ability of HypE is not as strong as those of the other algorithms.
6) For the 3-objective MaFs, SPEA2+SDE ranks first for all of the problems in terms of the HV metric, followed by Two_Arch2, which performs the best for five of the problems. PICEA-g and GrEA are also ranked in the first class for approximately half of the 3-objective MaFs based on both the HV and IGD metrics.
7) In terms of MaF problems with 7 objectives, SPEA2+SDE achieves the top-ranked performance for nearly half of the problems considering the HV metric and for most problems considering the IGD metric, whereas most of the other algorithms perform the best on at least one problem. Two_Arch2 and KnEA also perform well and are ranked first on approximately half of the 7-objective MaFs.
   SPEA2+SDE exhibits strong convergence ability due to its SDE strategy, which virtually places individuals with poor convergence in crowded regions, thus easily eliminating individuals that are poorly converged during the evolution. As the convergence of solutions to the Pareto front can be difficult for most MaFs, SPEA2+SDE consequently shows outstanding performance.
8) In conclusion, even though SPEA2+SDE exhibits top-ranked performance for most of the MaF test instances, other algorithms also work well for some of the MaFs. None of the test algorithms outperforms the others on all MaFs.

C. PERFORMANCE COMPARISONS ON PARETO-BOX PROBLEM

Due to the relation of the crowding between the objective space and the decision space, the behaviour of algorithms in the objective space can be directly inferred by viewing the distribution of solutions in the decision space.

Fig. 4 depicts the solutions obtained by the 13 test algorithms in the decision space in a single run. It is clear that the 13 test algorithms show different behaviour characteristics. As presented in Fig 4, MOEA/D, RVEA, θ-DEA, MOEA/DD, and NSGA-III concentrate on only one or several subregion(s) of the decision space, resulting in poor diversity maintenance. Notably, these five algorithms all belong to decomposition-based approaches with weight vectors. Therefore, their poor diversity maintenance is attributable to the inverted triangular shape of the Pareto front of the Pareto-Box problem. Due to this shape, only a small number of weight vectors intersect with the Pareto front [67]. In MOEA/D, when a weight vector is outside the Pareto front, only the best solution for the weight vector is obtained. Thus, many solutions are obtained in a small region of the Pareto front, as shown in Fig. 4. In MOEA/DD and θ-DEA, when a weight vector is outside the Pareto front, no solution is assigned to this weight vector. The second best solution to another weight vector inside the Pareto front is included in the population. Since the second best solution is almost (or exactly) the same as the best solution for each weight vector, the diversity of the obtained solutions becomes very small (i.e., the number of obtained solutions appears very small). In NSGA-III, when a weight vector is outside the Pareto front, no solution is assigned to this weight vector. A solution from non-dominated solutions is randomly selected and included in the population. Consequently, the obtained solution set has greater diversity. However, the uniformity or regularity of solutions of MOEA/D decreases due to the random selection mechanism.

As a decomposition-based method, MSOPS-II shows greater diversity maintenance than the above five algorithms. In addition, the distribution of solutions obtained by GrEA and HypE is not satisfactory. Although some solutions

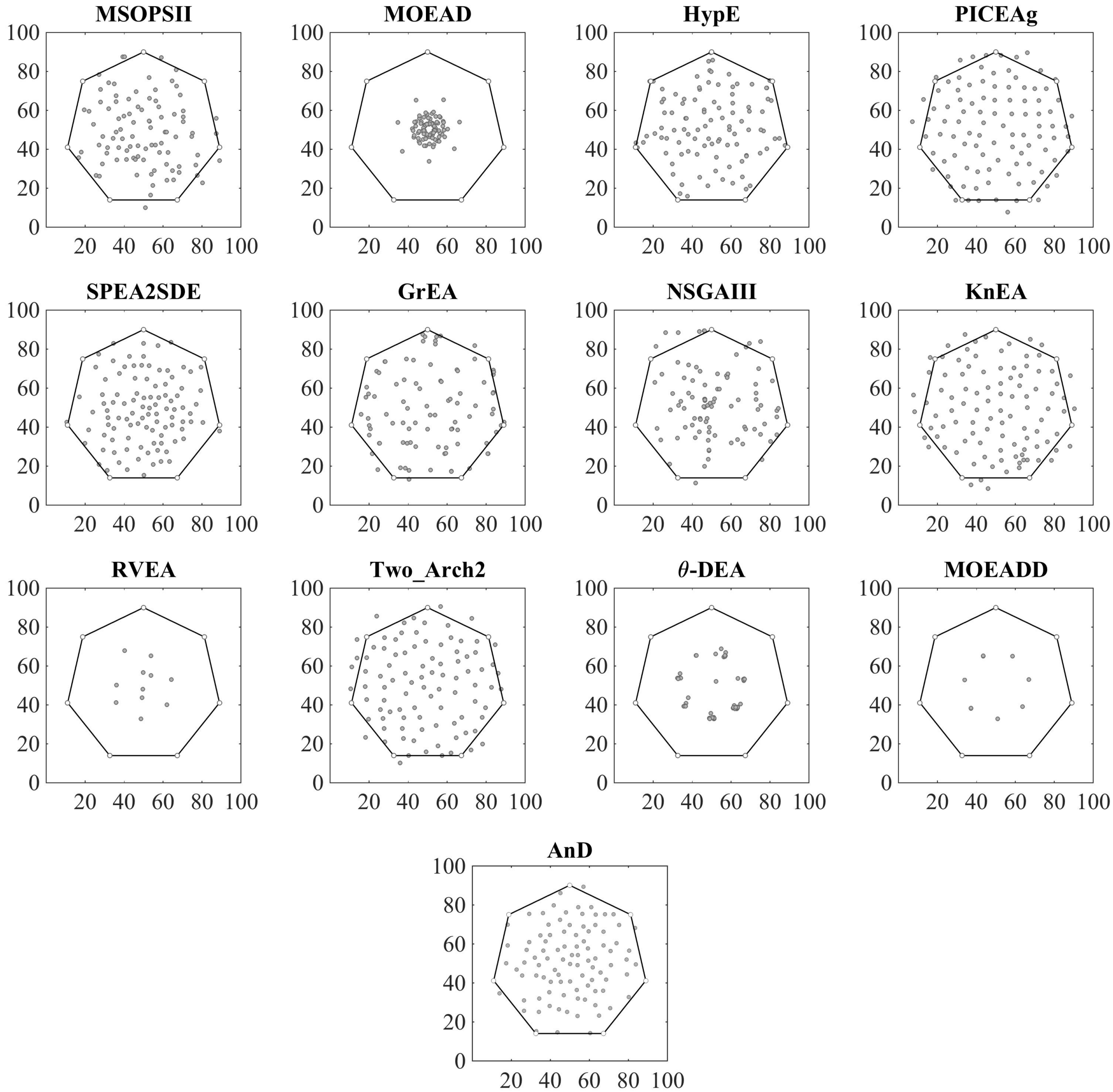

FIGURE 4. The distribution of solutions obtained by the 13 test algorithms in the decision space on the Pareto-Box problem.

obtained by HypE can reach each angle of the heptagon, many vacant places remain.

In this single run, PICEA-g, SPEA2+SDE, Two_Arch2, KnEA, and AnD show better performance compared with the other algorithms, with much better diversity maintenance. However, these algorithms have their own drawbacks. For Two_Arch2, PICEA-g, KnEA, and AnD, several solutions fail to converge into the Pareto optimal region. Moreover, SPEA2+SDE shows difficulty in keeping the boundary solutions.

Although the Pareto-Box is a relatively "easy" problem, none of the test algorithms provides an acceptable balance of convergence and diversity. Producing a set of solutions that are both well-converged and well-distributed remains a challenging problem. In this respect, perhaps a localized concept-based decomposition [68] or a constrained decomposition [69] can help balance convergence and diversity.

D. OVERALL DISCUSSION ON THE PERFORMANCE OF THE 13 ALGORITHMS

Unsurprisingly, the results described above show that the performance of the 13 algorithms significantly depends on the problem type and the number of objectives. Thus, further analysis of the performance of the algorithms can be conducted in connection to their characteristics and also in connection to the features of the test problems. For example, MaF3 has a convex PF with a large number of local fronts. Most algorithms fail to achieve a converged solution for this problem with 7 objectives. However, the two decomposition-based algorithms, MSOPS-II and MOEA/D, show a good performance on it. The idea of decomposing a MaOP into a set of scalar optimization subproblems may work well for such problems with a high dimension. Moreover, MSOPS-II also show a competitive behaviour on WFG1, which has a PF of complicated mixed geometries. This finding is similar with the conclusion in [14], that complex objective spaces can be analyzed easily and efficiently using multiple target vector approaches. Moreover, WFG3 is a difficult problem with a partially degenerate front and non-separable decision variables. Most algorithms show a poor performance on this problem especially when the dimension increases. However, HypE performs the best on WFG3 with all 2, 3 and 7 objectives, evaluated by both the IGD and HV metrics. To some extent, the indicator-based approach may have a competitive advantage on such difficult problems with a degenerate front and non-separable decision variables.

The advanced idea of coevolving a family of preferences simultaneously with the population of candidate solutions gives PICEA-g an excellent ability to maintain diversity. Such superiority of PICEA-g makes it a top algorithm for different types of problems, such as most WFG problems and the Pareto-Box problem. However, PICEA-g shows a relatively worse performance on WFG1, WFG3, WFG8, and WFG9 when the dimension of problems increases. It is interesting that WFG3, WFG8, and WFG9 are all non-separable problems, with their decision variables linked with other variables. In this sense, it should be careful to select the preference-based approach to solve such non-separable problems.

From the results, we can see that SPEA2+SDE shows a better behaviour on MaF problems compared with the WFG problems. It is primarily because the convergence of solutions to the Pareto front is more difficult for most MaFs compared to the WFG problems. The SDE strategy, which virtually places individuals with poor convergence in crowded regions, can easily eliminate individuals that are poorly converged during the evolution. Thus SPEA2+SDE exhibits a strong convergence ability due to its SDE strategy and it is suitable for those problems which are hard to converge.

Moreover, it can be concluded from the results that decomposition-based approaches have a poor ability of maintaining diversity for problems with inverted PFs. In contrast, solutions obtained by Pareto-based approaches are better distributed. For instance, regarding the Pareto-Box problem, MOEA/D, RVEA, θ-DEA, MOEA/DD, and NSGA-III concentrate on only one or several subregion(s) of the decision space, resulting in poor diversity maintenance. In addition, those decomposition-based approaches, MOEA/D, RVEA, θ-DEA, MOEA/DD, and NSGA-III show a poor performance on MaF1 and MaF2, which have inverted PFs. It is primarily because that, for problems with inverted PFs, only a small number of weight vectors intersect with the Pareto front, thus leading to a bad diversity maintenance. Hence, the decomposition-based approach is not a good choice for problems with inverted PFs.

In order to show the relative performance of the compared algorithms in an overall aspect, the average ranks of the algorithms for the different types of problems are shown in Table 11 and 12, based on HV and IGD respectively. Here, a smaller rank indicates a better performance.

V. CONCLUSION AND FUTURE DIRECTIONS

Thirteen state-of-the-art evolutionary algorithms of various categories for many-objective optimization are compared in this work. Most of these algorithms have been published in top journals in the last four years, a time period during which studies on the solution of MaOPs have increased greatly. Seventeen test instances involving the popular WFG, MaF and Pareto-Box problem suites, which encompass different types of problem properties, are used in this study to evaluate the investigated algorithms. The performances of the algorithms are evaluated by the HV and IGD metrics.

The experimental studies demonstrate that none of the compared algorithms outperforms the others for all types of test problems. Different algorithms have advantages for different problems. Interestingly, the experimental studies also show that the behaviours of the IGD and HV metrics may differ in some cases, particularly for the HypE and PICEA-g algorithms on the WFG test suite. HypE has an unfair advantage over other algorithms when using the HV metric, as this algorithm was developed based on the HV indicator. However, HypE fails to show such overwhelming performance when using the HV metric on MaF problems.

There is no *champion* algorithm that works well for different problems. Several algorithms show competitiveness for a large number of test instances. For WFG problems, HypE, SPEA2+SDE, NSGA-III, and θ-DEA perform well in terms of the HV metric. PICEA-g, GrEA, Two_Arch2, θ-DEA and NSGA-III perform well in terms of the IGD metric. For the MaF problems, SPEA2+SDE, Two_Arch2, KnEA, and GrEA perform well in terms of both the HV and IGD metrics.

In addition, even for some easy problems, such as the Pareto-Box problem, most of the algorithms fail to generate a set of well-converged and well-distributed solutions. Compared to other algorithms, PICEA-g and Two_Arch2 show better diversity maintenance but are still unable to converge all solutions into the Pareto optimal region.

Evolutionary many-objective algorithms have been developed and evaluated mainly for the DTLZ [70] and WFG [57] test problems with special characteristic features [67]. As a result, as demonstrated in this paper, in computational experiments (e.g., Figure 4), these algorithms do not always work well on other test problems. One important future research direction is to create a set of test problems with various characteristic features. Some attempts have been proposed recently [71]–[73]. The creation of such a set of test problems will lead to the development of more versatile many-objective algorithms. The performance evaluation of many-objective algorithms on real-world problems is also an important future research topic. Moreover, another important future direction for MaOP study is the search for the knee front, instead of just searching for the whole Pareto front, as the approximation of the whole front always requires a large number of points. With respect to algorithm development, diversity maintenance in the decision space has not been discussed in the context of many-objective optimization in the literature. This is an interesting and promising future research topic since no algorithms work well on a simple Pareto-Box problem in Figure 4. This research direction includes automated adaptation of weight vectors in MOEA/D and its variants.

APPENDIX

IGD RESULTS

IGD results of the 13 algorithms on WFG and MaF test suites are listed in Tables 13, 14, 15, 16, 17, and 18. The HV results of MOEA/D with normalization is shown in Table 19.

REFERENCES

[1] W.-Z. Zhao, C.-Y. Wang, L.-Y. Yu, and T. Chen, "Performance optimization of electric power steering based on multi-objective genetic algorithm," *J. Central South Univ.*, vol. 20, pp. 98–104, Jan. 2013.
[2] R. Chelouah and C. Baron, "Ant colony algorithm hybridized with tabu and greedy searches as applied to multi-objective optimization in project management," *J. Heuristics*, vol. 13, p. 640, Dec. 2007.
[3] S. Kitayama, J. Srirat, M. Arakawa, and K. Yamazaki, "Sequential approximate multi-objective optimization using radial basis function network," *Struct. Multidisciplinary Optim.*, vol. 48, pp. 501–515, Sep. 2013.
[4] F. Djeffal and N. Lakhdar, "An improved analog electrical performance of submicron dual-material gate (DM) GaAs-MESFETs using multi-objective computation," *J. Comput. Electron.*, vol. 12, pp. 29–35, Mar. 2013.
[5] A. Trivedi, D. Srinivasan, K. Pal, C. Saha, and T. Reindl, "Enhanced multiobjective evolutionary algorithm based on decomposition for solving the unit commitment problem," *IEEE Trans. Ind. Informat.*, vol. 11, no. 6, pp. 1346–1357, Dec. 2014.
[6] S. R. Sulistyo and R. I. Ratnasari, "Determining medical aid distribution route using multi-objective optimization," in *Proc. IEEE Int. Conf. Ind. Eng. Eng. Manage. (IEEM)*, Dec. 2017, pp. 417–421.
[7] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: NSGA-II," *IEEE Trans. Evol. Comput.*, vol. 6, no. 2, pp. 182–197, Apr. 2002.
[8] E. Zitzler, M. Laumanns, and L. Thiele, "SPEA2: Improving the strength Pareto evolutionary algorithm for multiobjective optimization," in *Proc. Evol. Methods Design, Optim. Control Appl. Ind. Prob. (EUROGEN)*, Athens, Greece, Sep. 2001, pp. 95–100.
[9] E. Zitzler and S. Künzli, "Indicator-based selection in multiobjective search," in *Proc. Int. Conf. Parallel Problem Solving From Nature*, vol. 3242. Berlin, Germany: Springer, 2004, pp. 832–842.
[10] 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, Dec. 2007.
[11] G. Fu, Z. Kapelan, J. R. Kasprzyk, and P. Reed, "Optimal design of water distribution systems using many-objective visual analytics," *J. Water Resour. Planning Manage.*, vol. 139, pp. 624–633, Nov. 2013.
[12] O. Chikumbo, E. Goodman, and K. Deb, "Approximating a multi-dimensional Pareto front for a land use management problem: A modified MOEA with an epigenetic silencing metaphor," in *Proc. IEEE Congr. Evol. Comput.*, Jun. 2012, pp. 1–9.
[13] R. C. Purshouse and P. J. Fleming, "On the evolutionary optimization of many conflicting objectives," *IEEE Trans. Evol. Comput.*, vol. 11, no. 6, pp. 770–784, Dec. 2007.
[14] E. J. Hughes, "Multiple single objective Pareto sampling," in *Proc. IEEE Congr. Evol. Comput. (CEC)*, vol. 4. Dec. 2003, pp. 2678–2684.
[15] E. J. Hughes, "MSOPS-II: A general-purpose many-objective optimiser," in *Proc. IEEE Congr. Evol. Comput. (CEC)*, Sep. 2007, pp. 3944–3951.
[16] H.-L. Liu, F. Gu, and Q. Zhang, "Decomposition of a multiobjective optimization problem into a number of simple multiobjective subproblems," *IEEE Trans. Evol. Comput.*, vol. 18, no. 3, pp. 450–455, Jun. 2014.
[17] R. Cheng, Y. Jin, M. Olhofer, and B. Sendhoff, "A reference vector guided evolutionary algorithm for many-objective optimization," *IEEE Trans. Evol. Comput.*, vol. 20, no. 5, pp. 773–791, Oct. 2016.
[18] K. Deb and H. Jain, "An evolutionary many-objective optimization algorithm using reference-point-based nondominated sorting approach—Part I: Solving problems with box constraints," *IEEE Trans. Evol. Comput.*, vol. 18, no. 4, pp. 577–601, Apr. 2013.
[19] K. Li, K. Deb, Q. Zhang, and S. Kwong, "An evolutionary many-objective optimization algorithm based on dominance and decomposition," *IEEE Trans. Evol. Comput.*, vol. 19, no. 5, pp. 694–716, Oct. 2015.
[20] N. Beume, B. Naujoks, and M. Emmerich, "SMS-EMOA: Multiobjective selection based on dominated hypervolume," *Eur. J. Oper. Res.*, vol. 181, pp. 1653–1669, Sep. 2007.
[21] J. Bader and E. Zitzler, *Hype: An Algorithm for Fast Hypervolume-Based Many-Objective Optimization*. Cambridge, MA, USA: MIT Press, 2011.
[22] M. Laumanns, L. Thiele, K. Deb, and E. Zitzler, "Combining convergence and diversity in evolutionary multiobjective optimization," *Evol. Comput.*, vol. 10, no. 3, 263–282, 2002.
[23] D. Hadka and P. Reed, "Borg: An auto-adaptive many-objective evolutionary computing framework," *Evol. Comput.*, vol. 21, no. 2, pp. 231–259, 2013.
[24] X. Zou, Y. Chen, M. Liu, and L. Kang, "A new evolutionary algorithm for solving many-objective optimization problems," *IEEE Trans. Syst., Man, Cybern, B, Cybern.*, vol. 38, no. 5, pp. 1402–1412, Oct. 2008.
[25] G. Wang and H. Jiang, "Fuzzy-dominance and its application in evolutionary many objective optimization," in *Proc. Int. Conf. Comput. Intell. Secur. Workshops*, 2007, pp. 195–198.
[26] F. D. 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, Feb. 2007.
[27] S. Yang, M. Li, X. Liu, and J. Zheng, "A grid-based evolutionary algorithm for many-objective optimization," *IEEE Trans. Evol. Comput.*, vol. 17, no. 5, pp. 721–736, Oct. 2013.
[28] Y. Yuan, H. Xu, B. Wang, and X. Yao, "A new dominance relation-based evolutionary algorithm for many-objective optimization," *IEEE Trans. Evol. Comput.*, vol. 20, no. 1, pp. 16–37, Feb. 2016.
[29] 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, Apr. 2011.
[30] M. Li, S. Yang, and 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.
[31] X. Zhang, Y. Tian, and Y. Jin, "A knee point-driven evolutionary algorithm for many-objective optimization," *IEEE Trans. Evol. Comput.*, vol. 19, no. 6, pp. 761–776, Dec. 2015.
[32] H. Ishibuchi, N. Tsukamoto, and Y. Nojima, "Evolutionary many-objective optimization: A short review," in *Proc. IEEE Congr. Evol. Comput. (CEC), IEEE World Congr. Comput. Intell.*, Jun. 2008, pp. 2419–2426.
[33] L. Rachmawati and D. Srinivasan, "Preference incorporation in multi-objective evolutionary algorithms: A survey," in *Proc. IEEE Congr. Evol. Comput. (CEC)*, Jul. 2006, pp. 962–968.
[34] J.-H. Kim, J.-H. Han, Y.-H. Kim, S.-H. Choi, and E.-S. Kim, "Preference-based solution selection algorithm for evolutionary multiobjective optimization," *IEEE Trans. Evol. Comput.*, vol. 16, no. 1, pp. 20–34, Feb. 2012.
[35] R. Landa, C. A. C. Coello, and G. Toscano-Pulido, "Goal-constraint: Incorporating preferences through an evolutionary ε-constraint based method," in *Proc. IEEE Congr. Evol.*, Jun. 2013, pp. 741–747.
[36] A. L. Jaimes, A. A. Montaño, and C. A. C. Coello, "Preference incorporation to solve many-objective airfoil design problems," in *Proc. IEEE Congr. Evol.*, Jun. 2011, pp. 1605–1612.
[37] D. Gong, J. Sun, and X. Ji, "Evolutionary algorithms with preference polyhedron for interval multi-objective optimization problems," *Inf. Sci.*, vol. 233, pp. 141–161, 2013.
[38] R. Wang, R. C. Purshouse, I. Giagkiozis, and P. J. Fleming, "The iPICEA-g: As new hybrid evolutionary multi-criteria decision making approach using the brushing technique," *Eur. J. Oper. Res.*, vol. 243, pp. 442–453, Jun. 2014.
[39] R. C. Purshouse, C. Jalbă, and P. J. Fleming, "Preference-driven co-evolutionary algorithms show promise for many-objective optimisation," in *Evolutionary Multi-Criterion Optimization*. Berlin, Germany: Springer, 2011, pp. 136–150.

[40] R. Wang, R. C. Purshouse, and P. J. Fleming, "On finding well-spread Pareto optimal solutions by preference-inspired co-evolutionary algorithm," in *Proc. 15th Annu. Conf. Genetic Evol. Comput.*, 2013, pp. 695–702.
[41] 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.
[42] R. Wang, R. C. Purshouse, and P. J. Fleming, "Preference-inspired co-evolutionary algorithms using weight vectors," *Eur. J. Oper. Res.*, vol. 243, pp. 423–441, Jun. 2015.
[43] K. Praditwong and X. Yao, "A new multi-objective evolutionary optimisation algorithm: The two-archive algorithm," in *Proc. Int. Conf. Comput. Intell. Secur.*, vol. 1. 2006, pp. 286–291.
[44] B. Li, J. Li, K. Tang, and X. Yao, "An improved two archive algorithm for many-objective optimization," in *Proc. IEEE Congr. Evol. Comput. (CEC)*, Jul. 2014, pp. 2869–2876.
[45] N. Pholdee, S. Bureerat, and A. R. Yildiz, "Hybrid real-code population-based incremental learning and differential evolution for many-objective optimisation of an automotive floor-frame," *Int. J. Vehicle Des.*, vol. 73, nos. 1–3, pp. 20–53, 2017.
[46] S. Jiang and S. Yang, "A strength Pareto evolutionary algorithm based on reference direction for multiobjective and many-objective optimization," *IEEE Trans. Evol. Comput.*, vol. 21, no. 3, pp. 329–346, Jun. 2017.
[47] Y. Liu, D. Gong, X. Sun, and Y. Zhang, "Many-objective evolutionary optimization based on reference points," *Appl. Soft Comput.*, vol. 50, pp. 344–355, Jan. 2017.
[48] Y. Xiang, Y. Zhou, M. Li, and Z. Chen, "A vector angle-based evolutionary algorithm for unconstrained many-objective optimization," *IEEE Trans. Evol. Comput.*, vol. 21, no. 1, pp. 131–152, Feb. 2017.
[49] M. Li, S. Yang, X. Liu, and R. Shen, "A comparative study on evolutionary algorithms for many-objective optimization," in *Proc. EMO*, 2013, pp. 261–275.
[50] X. Ma, J. Yang, N. Wu, Z. Ji, and Z. Zhu, "A comparative study on decomposition-based multi-objective evolutionary algorithms for many-objective optimization," in *Proc. IEEE Congr. Evol. Comput.*, Jul. 2016, pp. 2477–2483.
[51] R. Tanabe, H. Ishibuchi, and A. Oyama, "Benchmarking multi- and many-objective evolutionary algorithms under two optimization scenarios," *IEEE Access*, vol. 5, pp. 19597–19619, Sep. 2017.
[52] K. Li, Q. Zhang, S. Kwong, M. Li, and R. Wang, "Stable matching-based selection in evolutionary multiobjective optimization," *IEEE Trans. Evol. Comput.*, vol. 18, no. 6, pp. 909–923, Dec. 2014.
[53] R. Wang, Q. Zhang, and T. Zhang, "Decomposition-based algorithms using Pareto adaptive scalarizing methods," *IEEE Trans. Evol. Comput.*, vol. 20, no. 6, pp. 821–837, Dec. 2016.
[54] A. Trivedi, D. Srinivasan, K. Sanyal, and A. Ghosh, "A survey of multiobjective evolutionary algorithms based on decomposition," *IEEE Trans. Evol. Comput.*, vol. 21, no. 3, pp. 440–462, Jun. 2017.
[55] J. D. Lohn, W. F. Kraus, and G. L. Haith, "Comparing a coevolutionary genetic algorithm for multiobjective optimization," in *Proc. IEEE Congr. Evol. Comput. (CEC)*, May 2002, pp. 1157–1162.
[56] K. Li, K. Deb, Q. Zhang, and Q. Zhang, "Efficient nondomination level update method for steady-state evolutionary multiobjective optimization," *IEEE Trans. Cybern.* , vol. 47, no. 9, pp. 2838–2849, Sep. 2017.
[57] S. Huband, P. Hingston, L. Barone, and L. While, "A review of multiobjective test problems and a scalable test problem toolkit," *IEEE Trans. Evol. Comput.*, vol. 10, no. 5, pp. 477–506, Oct. 2006.
[58] R. Cheng *et al.*, "A benchmark test suite for evolutionary many-objective optimization," *Complex Intell. Syst.*, vol. 3, no. 1, pp. 67–81, Mar. 2017.
[59] M. Köppen, K. Yoshida, "Substitute distance assignments in NSGA-II for handling many-objective optimization problems," in *Evolutionary Multi-Criterion Optimization*. Berlin, Germany: Springer, 2007, pp. 727–741.
[60] H. Ishibuchi, Y. Hitotsuyanagi, N. Tsukamoto, and Y. Nojima, "Many-objective test problems to visually examine the behavior of multiobjective evolution in a decision space," in *Proc. 11th Int. Conf. Parallel Problem Solving Nature (PPSN)*, 2010, pp. 91–100.
[61] E. Zitzler and L. Thiele, "Multiobjective evolutionary algorithms: A comparative case study and the strength Pareto approach," *IEEE Trans. Evol. Comput.*, vol. 3, no. 4, pp. 257–271, Nov. 1999.
[62] Y. Tian, R. Cheng, X. Zhang, and Y. Jin, "PlatEMO: A MATLAB platform for evolutionary multi-objective optimization [Educational Forum]," *IEEE Comput. Intell. Mag.*, vol. 12, no. 4, pp. 73–87, Sep. 2017.
[63] M. Frigge, D. C. Hoaglin, and B. Iglewicz, "Some implementations of the boxplot," *Amer. Statistician*, vol. 43, no. 1, pp. 50–54, 1989.
[64] H. Ishibuchi, H. Masuda, and Y. Nojima, "Pareto fronts of many-objective degenerate test problems," *IEEE Trans. Evol. Comput.*, vol. 20, no. 5, pp. 807–813, Oct. 2016.
[65] D. A. Van Veldhuizen, and G. B. Lamont, "Evolutionary computation and convergence to a pareto front," in *Proc. Late Breaking Papers Genetic Program. Conf.*, 1998, pp. 221–228.
[66] H. Ishibuchi, K. Doi, and Y. Nojima, "On the effect of normalization in MOEA/D for multi-objective and many-objective optimization," *Complex Intell. Syst.*, vol. 3, pp. 279–294, Dec. 2017.
[67] H. Ishibuchi, S. Yu, H. Masuda, and Y. Nojima, "Performance of decomposition-based many-objective algorithms strongly depends on Pareto front shapes," *IEEE Trans. Evol. Comput.*, vol. 21, no. 2, pp. 169–190, Apr. 2016.
[68] R. Wang, Z. Zhou, H. Ishibuchi, T. Liao, and T. Zhang, "Localized weighted sum method for many-objective optimization," *IEEE Trans. Evol. Comput.*, vol. 22, no. 1, pp. 3–18, Feb. 2016.
[69] L. Wang, Q. Zhang, A. Zhou, M. Gong, and L. Jiao, "Constrained subproblems in a decomposition-based multiobjective evolutionary algorithm," *IEEE Trans. Evol. Comput.*, vol. 20, no. 3, pp. 475–480, Jun. 2016.
[70] K. Deb *et al*, "Scalable test problems for evolutionary multiobjective optimization," in *Evolutionary Multiobjective Optimization*. London, U.K.: Springer, 2005, pp. 105–145.
[71] H. Masuda, Y. Nojima, and H. Ishibuchi, "Common properties of scalable multiobjective problems and a new framework of test problems," in *Proc. IEEE Congr. Evol. Comput.*, Jul. 2016, pp. 3011–3018.
[72] R. Cheng *et al.*, "Benchmark functions for the CEC'2017 competition on many-objective optimization," School Comput. Sci., Univ. Birmingham, Birmingham, U.K., Tech. Rep. CSR-17-01, 2017.
[73] M. Li, C. Grosan, S. Yang, X. Liu, and X. Yao, "Multiline distance minimization: A visualized many-objective test problem suite," *IEEE Trans. Evol. Comput.*, vol. 22, no. 1, pp. 61–78, Feb. 2017.

KAIWEN LI received the B.S. degree from the National University of Defense Technology (NUDT), Changsha, China, in 2016.

He is currently pursuing the degree with the College of System Engineering, NUDT. His research interests include prediction technique, multicriteria decision making, optimal scheduling, data mining, and optimization methods on energy Internet network.

RUI WANG received the B.S. degree from the National University of Defense Technology, Changsha, China, in 2008, and the Ph.D. degree from The University of Sheffield, Sheffield, U.K., in 2013.

He is currently a Lecturer with the National University of Defense Technology. His research interests include evolutionary computation, multiobjective optimization, machine learning, and various applications using evolutionary algorithms.

TAO ZHANG received the B.S., M.S., Ph.D. degrees from the National University of Defense Technology (NUDT), Changsha, China, in 1998, 2001, and 2004, respectively.

He is currently a Professor with the College of Information System and Management, NUDT. His research interests include multicriteria decision making, optimal scheduling, data mining, and optimization methods on energy Internet network.

HISAO ISHIBUCHI (M'93–SM'10–F'14) received the B.S. and M.S. degrees in precision mechanics from Kyoto University, Kyoto, Japan, in 1985 and 1987, respectively, and the Ph.D. degree in computer science from Osaka Prefecture University, Sakai, Japan, in 1992.

Since 1987, he has been with Osaka Prefecture University for 30 years. He is currently a Chair Professor with the Department of Computer Science and Engineering, Southern University of Science and Technology, Shenzhen, China. His current research interests include fuzzy rule-based classifiers, evolutionary multiobjective optimization, many-objective optimization, and memetic algorithms.

Dr. Ishibuchi was the IEEE Computational Intelligence Society (CIS) Vice-President for Technical Activities from 2010 to 2013. He is an IEEE CIS AdCom Member from 2014 to 2019, an IEEE CIS Distinguished Lecturer from 2015 to 2017, and an Editor-in-Chief of the IEEE *Computational Intelligence Magazine* from 2014 to 2019. He is also an Associate Editor of the IEEE Transactions on Evolutionary Computation, the IEEE Transactions on Cybernetics, and the IEEE Access.