

A New Two-stage Evolutionary Algorithm for Many-objective Optimization

Yanan Sun, *Member, IEEE*, Bing Xue, *Member, IEEE*, Mengjie Zhang, *Senior Member, IEEE*,
and Gary G. Yen, *Fellow, IEEE*

Abstract—Convergence and diversity are inter-dependently handled during the evolutionary process by most existing Many-objective Evolutionary Algorithms (MaOEAs). In such a design, the degraded performance of one would deteriorate the other, and only solutions with both are able to improve the performance of MaOEAs. Unfortunately, it is not easy to constantly maintain a population of solutions with both convergence and diversity. In this paper, an MaOEA based on two independent stages is proposed for effectively solving Many-objective Optimization Problems (MaOPs), where the convergence and diversity are addressed in two independent and sequential stages. To achieve this, we firstly propose a non-dominated dynamic weight aggregation method by using a genetic algorithm, which is capable of finding the Pareto-optimal solutions for MaOPs with concave, convex, linear and even mixed Pareto front shapes, and then these solutions are employed to learn the Pareto-optimal subspace for the convergence. Afterwards, the diversity is addressed by solving a set of single-objective optimization problems with reference lines within the learned Pareto-optimal subspace. To evaluate the performance of the proposed algorithm, a series of experiments are conducted against six state-of-the-art MaOEAs on benchmark test problems. The results show the significantly improved performance of the proposed algorithm over the peer competitors. In addition, the proposed algorithm can focus directly on a chosen part of the objective space if the preference area is known beforehand. Furthermore, the proposed algorithm can also be used to effectively find the nadir points.

Index Terms—Many-objective evolutionary optimization algorithm, two-stage method, nadir point, Pareto-optimal subspace.

I. INTRODUCTION

MANY-OBJECTIVE Optimization Problems (MaOPs) typically have more than three conflict objectives to be optimized concurrently [1]. Mathematically, an MaOP is represented by (1)

$$\begin{cases} \mathbf{f}(x) = (f_1(x), f_2(x), \dots, f_m(x)) \\ s.t. \quad x \in \Pi \end{cases} \quad (1)$$

where $m > 3$, $\Pi \subseteq R^n$ is the search space for the decision variable $x = (x_1, x_2, \dots, x_n)$, and \mathbf{f} is an objective vector. Without loss of generality, MaOPs discussed in this paper are to be minimized because the maximization problems can be transformed into the minimization ones by multiplying -1 .

Yanan Sun, Bing Xue, and Mengjie Zhang are with the School of Engineering and Computer Science, Victoria University of Wellington, PO Box 600, Wellington 6140, New Zealand (e-mail: yanan.sun@ecs.vuw.ac.nz; bing.xue@ecs.vuw.ac.nz; and mengjie.zhang@ecs.vuw.ac.nz).

Gary G. Yen is with the School of Electrical and Computer Engineering, Oklahoma State University, Stillwater, OK 74078 USA (e-mail: gyen@okstate.edu).

Due to the conflicting nature among $f_1(x), f_2(x), \dots, f_m(x)$, there is no single solution that satisfies all objectives simultaneously. Instead, a set of non-dominated solutions with promising *convergence* and *diversity* is sought for. Specifically, the convergence means that the solutions are close to the true Pareto Front (PF) as far as possible, and the diversity implies these solutions are uniformly distributed in the PF. Evolutionary Algorithms (EAs) are preferably used to solve these optimization problems with multiple conflicting objectives because of their population-based, meta-heuristic characteristics.

Numerous Multi-objective EAs (MOEAs), such as the elitist Non-dominated Sorting Genetic Algorithm (NSGA-II) [2] and the advanced version of Strength Pareto Evolutionary Algorithm (SPEA2) [3], have been designed for solving Multi-objective Optimization Problems (MOPs) satisfactorily where the number of conflicting objectives is no more than three. Typically, MOEAs employ a carefully crafted Pareto-domination-based elitism mechanism and a density-based diversity promotion mechanism to achieve the convergence and diversity, respectively. Regrettably, these MOEAs are not able to address MaOPs largely due to the loss of selection pressure [4]. Specifically, nearly all solutions from the population are non-dominated with respect to each other as the number of objectives increases because of the dominance resistance phenomenon [5]. In this regard, the Pareto-domination-based elitism mechanism cannot work effectively, and the density-based diversity promotion mechanism becomes overly emphasized. However, in MaOPs, the solutions selected by the density-based diversity promotion mechanism are typically far away from each other and also distant from the PF [6]. As a result, the parent solutions chosen by the density-based diversity promotion mechanism will often generate poor offspring in the next generation. It has been reported that the generated offspring in such a situation could be stagnant during the evolution or even far away from the PF [7]. To this end, various Many-objective EAs (MaOEAs) specifically designed to address MaOPs have been proposed in recent years.

Typically, MaOEAs can be classified into four different classes. The first refers to the ones based on the diversity enhancement using a set of reference lines that are uniformly distributed in the objective space. Examples include the reference-line-based non-dominated sorting approach (i.e., NSGA-III) [8], the new dominance relation-based EA (i.e., θ -DEA) [9], the reference line-based estimation of distribution algorithm (i.e., MaOEDA/RL) [10], the improved regularity model-based estimation of distribution algorithm (i.e.,

MaOEDA-IR) [11], and the Reference Vector guided EA (RVEA) [12]. The second is known as the decomposition-based algorithms; decomposing an MaOP into a set of Single-objective Optimization Problems (SOPs). The MOEA based on Decomposition (MOEA/D) [13] and its numerous variants, such as [14]–[16], belong to this category. The third covers the ones based on the convergence enhancement by modifying the Pareto-dominance-based comparison strategy. For instance, the ϵ -dominance method [17] employed a relaxed factor ϵ to compare the dominance relation between solutions; the preference order approach [18] used a ranking procedure to lower the adverse impact of the Pareto-dominance-based comparison; the fuzzy dominance method [19] was proposed with the fuzzy-based ranking scheme; the MaOEA in [20] was proposed with the grid-based approach to relax the Pareto-dominance-based method; and a Knee point-based EA (KnEA) [21] was recently presented to enhance the selection pressure. The fourth is the indicator-based approaches. Because the performance indicators extensively used in MaOEAs are the hypervolume (HV) and the Inverse Generational Distance (IGD), MaOEAs in this category are mainly based on these two indicators, such as [22]–[24]. Recently, Sun *et al.* [25] proposed an IGD-based MaOEA (MaOEA-IGD) that showed promising performance in solving MaOPs.

These MaOEAs mentioned above handle the convergence and diversity concurrently, i.e., the convergence and diversity are explicitly addressed at the same time of each generation, and then individuals with better convergence and diversity survive into the next generation to hopefully generate new solutions with both of such good characteristics. Because EAs are heuristic, it has no guarantee to constantly maintain a population of individuals that are with good convergence and diversity at every generation, especially in MaOPs, due to their large search space that solutions with good diversity are at times with poor convergence and vice versa [4]. In addition, the convergence and diversity are generally conflicting [26]–[28], which implies that the individuals with promising convergence are usually with poor diversity, and vice versa. To this end, it is hard to maintain a population of individuals with both promising convergence and diversity, but a compromise between them.

To address the issue, researchers have proposed the MaOEAs based on two-stage strategy where the convergence and diversity are handled in different phases, such as the objective space Reduction and Diversity improvement MaOEA (MaOEA-RD) [29], the MaOEA based on two-stage strategy and Parallel Cell Coordinate System (PCCS) [30], and the Two-phase EA with Penalty Based Adjustment for reference lines (TPEA-PBA) [31]. Typically, these MaOEAs in the first stage find the Pareto-optimal Space (PoS) that is spanned by the solutions from the Pareto Set (PS) (Pareto-optimal solutions form the PS in the decision space). Therefore, the convergence can be guaranteed if the solutions are found from the PoS. In the second stage, the diversity is processed within the PoS by using their claimed diversity promotion strategies. By dividing the evolution into such two stages, the convergence and diversity are obviously addressed without the deficiencies mentioned above. Unfortunately, there are two

major limitations when these algorithms are used in practice. Firstly, they only rely on the extreme points in finding the PoS; however, the same extreme points may give the different PoS. Secondly, they find the extreme points by obtaining the optimal solutions of each objective of the MaOP, which has been proven to be inaccurate [25], [32].

In order to make the best use of the advantages of the two-stage strategy, a new MaOEA based on an Independent Two-stage approach, in shortly named MaOEA-IT, is proposed in this paper, which has eliminated the above limitations often existed in two-stage MaOEAs. In summary, the contributions of MaOEA-IT are highlighted as follows:

- 1) The pursuit of convergence and diversity in MaOEA-IT is addressed independently through two sequential stages, which makes MaOEA-IT to be effective in solving MaOPs. To be specific, there is no requirement for introducing any extra mechanism, which is often used to simultaneously maintain the promising convergence and diversity at each generation during the evolution.
- 2) A Non-dominated Dynamic Weight Aggregation by using a Genetic Algorithm (NDWA-GA) method is proposed, which is capable of finding k Pareto-optimal solutions for an m -objective MaOP with convex, linear, convex and even mixed PF shapes, where $k \gg m$. The found solutions are with promising convergence but not necessarily good diversity, but are sufficient to be used to learn the PoS with dimension reduction techniques.
- 3) The diversity is addressed by converting an m -objective MaOEA into p SOPs with p reference lines that are uniformly distributed in the objective space. During the process of solving each SOP, the objective is to minimize the cosine value between the corresponding reference line and the solution via exploration and exploitation in the PoS. Theoretically, any single-objective optimization algorithm can be used in this stage.
- 4) The proposed algorithm can focus on only a small part of the search space when the preference area of the MaOP to be solved is known, which could save the computational resource at a large extent because the traditional MaOEAs must find all the solutions first, and then from which the preference one is selected. In addition, the proposed algorithm can also be used to effectively search for the nadir points, which is typically used as a preprocessing step in solving MaOPs.

The remainder of this paper is organized as follows. Section II reviews the limitations of related works to motive the proposed design in this paper. In Section III, the details of MaOEA-IT are documented. To evaluate the performance of MaOEA-IT, in Section IV, a series of experiments is conducted validating MaOEA-IT on benchmarks with different numbers of objectives against the state-of-the-art MaOEAs. In addition, the performance of the proposed critical component NDWA-GA is also investigated through the comparison to its seminal work. Finally, MaOEA-IT is concluded and future works are illustrated in Section V.

II. MOTIVATIONS

In this section, the limitations of the existing MaOEAs based on the two-stage approach are detailed, which form the main motivation of the proposed algorithm. In addition, the limitation of the existing dynamic weight aggregation strategy is also documented, which inspires the design of the proposed NDWA-GA method.

A. MaOEAs Based on Two-stage Approach

To the best of our knowledge, MaOEAs based on the two-stage approach include MaOEA-RD, PCCS and TPEA-PBA of which the two limitations have been pointed out in Section I. Because the extensive explanation regarding the second limitation can be found in [25], [32], we mainly discuss the first one at length in this subsection.

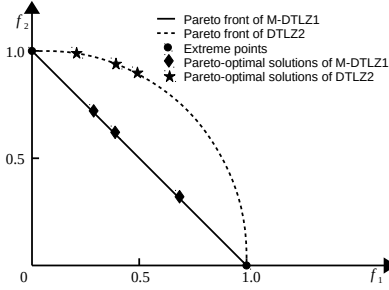


Fig. 1. An illustration of the extreme points, Pareto fronts and Pareto-optimal solutions of the 2-objective M-DTLZ1 and DTLZ2 test problems.

As mentioned above, MaOEA-RD, PCCS and TPEA-PBA employ the extreme points to obtain the PoS. In order to clearly describe this deficiency, an example is shown in Fig. 1 where the solid circles denote the extreme points of the 2-objective Modified-DTLZ1 (named M-DTLZ1) and DTLZ2 [33] test problems. Specifically, the M-DTLZ1 is implemented by enlarging each objective of DTLZ1 [33] twice. Therefore, the 2-objective M-DTLZ1 has the same extreme points as the 2-objective DTLZ2, i.e., $[0, 1]$ and $[1, 0]$. In addition, the solid line and the dash line in Fig. 1 refer to the PFs of the 2-objective M-DTLZ1 and DTLZ2 test problems, respectively. Obviously, the 2-objective M-DTLZ1 and DTLZ2 test problems have the same extreme points but different PFs. However, MaOEA-RD, PCCS and TPEA-PBA will treat them with the same PoS¹, which is obviously incorrect, or at least not accurate. In Fig. 1, the star and diamond symbols refer to the other Pareto-optimal solutions of the 2-objective M-DTLZ1 and DTLZ2 test problems, respectively. If these Pareto-optimal solutions are used together with the respective extreme points, the respective PoS for M-DTLZ1 and DTLZ2 will be found correctly. Unfortunately, obtaining a set of such Pareto-optimal solutions before solving the optimization problem is not an easy task especially for the problems involving non-convex PF shapes.

¹Noting that the PoS typically refers to the subspace in the decision space. This example is illustrated in the objective space of two concrete test problems which is very familiar to the researchers and easy to understand the limitation.

B. Dynamic Weight Aggregation Strategy

Weight Aggregation (WA) methods [34] have been widely used to find the Pareto-optimal solutions of MOPs at the very beginning of MOPs researches. Specifically, WA methods employ a weight, say $w = \{w_1, w_2, \dots, w_m\}$, to transform an m -objective MOP $\mathbf{f}(x) = (f_1(x), f_2(x), \dots, f_m(x))$ into a SOP $\mathbf{g}(\cdot)$ by $\mathbf{g}(x) = w_1 f_1(x) + w_2 f_2(x) + \dots + w_m f_m(x)$. By solving $\mathbf{g}(x)$, a Pareto-optimal solution of $\mathbf{f}(x)$ can be found. Consequently, a set of different weights is required to form multiple different SOPs, and then solving them for obtaining multiple Pareto-optimal solutions. Algorithms based on WA have also been proposed with different strategies for maintaining the diversity, which are beyond the scope here because we focus solely on the convergence when using WA-based algorithms in this paper.

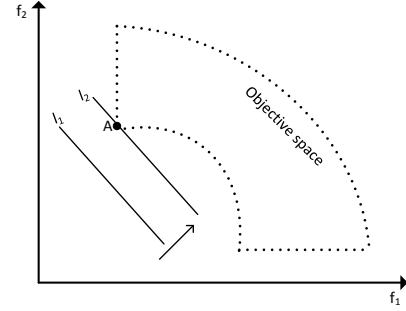


Fig. 2. Textbook explanation of the reason that weight aggregation-based methods cannot reach to the concave area.

Most researchers have known that WA methods are incapable of obtaining the Pareto-optimal solutions of the MOPs with concave PF shapes. The reason has been stated in a long time as that the contour of the transformed SOP $\mathbf{g}(x)$ cannot reach to the concave region of $\mathbf{f}(x)$ during solving $\mathbf{g}(x)$, i.e., $\mathbf{g}(x)$ converges when it has an intersection with the PF of $\mathbf{f}(x)$, and this intersection will not occur in the non-convex area. Fig. 2 shows such an example where the area closed by dash lines denotes the objective space of $\mathbf{f}(x)$, l_1 and l_2 refer to the contour of $\mathbf{g}(x)$ with different x , the arrow denotes the moving direction in solving $\mathbf{g}(x)$, A is a Pareto-optimal solution of $\mathbf{f}(x)$ and also the optimal solution of $\mathbf{g}(x)$. l_2 is also the stopping position of $\mathbf{g}(x)$ and it will not move toward up-right (i.e., the concave area). However, this explanation is incomplete because $\mathbf{g}(x)$ cannot stand beyond the PF [35], and the fact is that $\mathbf{g}(x)$ moves from up-right towards the PF until it has no intersection any more if it continues moving. To address this concerning issue, various remedies have also been proposed, such as [35]–[39], among which the Dynamic WA (DWA)-based method [35], [39] is most recognized one because it is capable of obtaining multiple Pareto-optimal solutions for MOPs with convex, linear and concave PF shapes in a single run. An example of DWA-based is shown in Fig. 3, where l_3 , l_4 and l_5 denote the different contours of $\mathbf{g}(x)$ with different x , and l_5 is the stopping position based on our explanation above. Obviously, the Pareto-optimal solutions have been found during solving $\mathbf{g}(x)$ (i.e., the points B and C). By changing the weights for transforming the original MOP into different $\mathbf{g}(x)$

and keeping an archive of the non-dominated solutions during solving $\mathbf{g}(x)$, a set of different Pareto-optimal solutions will be obtained. In addition, Jin in [40] has concluded that the DWA-based method is only suitable to the MOPs that satisfy the *neighborhoodness*, i.e., small variations in the decision space should keep the same scale of variation in the objective space, and only the Evolutionary Strategy (ES) is suitably used in such a situation, which induces the DWA-ES algorithm [35].

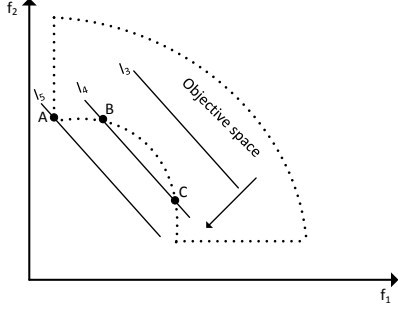


Fig. 3. An example to illustrate how DWA-based algorithms search for the Pareto-optimal solutions of the problems involving concave areas.

Experimental investigations indicate DWA-ES showing poor performance on MaOPs. The major reasons are: 1) Solutions in DWA-ES surviving into the next generation are selected based on the fitness of $\mathbf{g}(x)$. However, the goodness of solutions for $\mathbf{g}(x)$ does not promote the search towards the PF of $\mathbf{f}(x)$, i.e., if a solution can give a best fitness of $\mathbf{g}(x)$, it is not necessary to give a promising performance of $\mathbf{f}(x)$. 2) The MaOPs to be solved do not necessarily meet the *neighborhoodness* constraint, which causes the inappropriate using of ES. 3) Parameters in ES are usually difficult to be dynamically tuned, although they are optimized also through the evolution, which results into the poor performance of ES. In this paper, we will propose the NDWA-GA method by addressing the limitations above, which will still preserve the promising properties of the DWA-based strategy.

III. THE PROPOSED ALGORITHM

In this section, the proposed Independent Two-stage Many-objective Evolutionary Algorithm (i.e., MaOEA-IT) is presented. To be specific, the framework of MaOEA-IT is outlined in Subsection III-A. Then, the details of the main steps are documented in Subsections III-B to III-E. Finally, the computational complexity of MaOEA-IT is analyzed in Subsection III-F. Noting that the MaOEA-IT is described within the context formulated by (1).

A. Framework of MaOEA-IT

Algorithm 1 outlines the framework of the proposed MaOEA-IT algorithm that is composed of four steps. Firstly, the proposed NDWA-GA algorithm is developed to search for a set of Pareto-optimal solutions (denoted by P) of a given MaOP $\mathbf{f}(x) = \{f_1(x), f_2(x), \dots, f_m(x)\}$ (line 1). Secondly, the PoS (denoted by Ω) is learned from the found Pareto-optimal solutions P (line 2). Thirdly, the given reference lines

Algorithm 1: Framework of MaOEA-IT

Input: the MaOP to be optimized

$\mathbf{f}(x) = \{f_1(x), f_2(x), \dots, f_m(x)\}$, a set of uniformly distributed reference lines

$\mathbf{r} = \{r_1, r_2, \dots, r_N\}$.

Output: The solutions of $\mathbf{f}(x)$.

- 1 $P \leftarrow$ Find a set of Pareto-optimal solutions of $\mathbf{f}(x)$ by using the proposed NDWA-GA algorithm;
 - 2 $\Omega \leftarrow$ Learn the Pareto-optimal subspace based on P ;
 - 3 Update the given reference lines $\mathbf{r} = \{r_1, r_2, \dots, r_N\}$;
 - 4 $Q \leftarrow$ Diversity maintaining with \mathbf{r} in Ω ;
 - 5 **Return** Q .
-

are updated (line 3) where N is predefined by the end-users based on how many solutions they finally would like to have. Fourthly, the final solutions Q with promising convergence and diversity are obtained (line 4).

Noting that the given reference lines \mathbf{r} are generated by using Das and Dennis's method [41] that is extensively used in reference lines-based MaOEAs, such as NSGA-III. Because the reference lines will be used in the first and the third steps of the proposed algorithm, the details of Das and Dennis's method are introduced priori to presenting the details of each step. To be specific, r_1, r_2, \dots, r_N are generated from the $(m-1)$ -dimensional unit hyperplane located in the first quadrant of R^m , i.e. R_{+}^{m-1} , and the summation of the components in each reference line r_i ($1 \leq i \leq N$) satisfies $\sum_{k=1}^{m-1} r_{i,k} = 1$ and $r_{i,k} \in \{0/h, 1/h, \dots, h/h\}$, where $r_{i,k}$ refers to the k -th component of the i -th reference line, and h is the division size determining the total number of sampled reference lines by $N = C_{m+h-1}^h$. However, no inside reference lines will be generated if $h < m$, which leads to the loss of diversity, i.e., generated reference lines only pass through the boundary of the $(m-1)$ -dimensional unit hyperplane and, there is no generated reference lines passing through the inside part. Otherwise when $h \geq m$, too many reference lines are generated, which will lead to the impracticability especially when $m > 5$. To address this concerning issue, the two-layer sampling method [8] is used. Specifically, the two-layer method is composed of sampling reference lines through the boundary and the inner part of the unit hyperplane. For sampling reference lines through the boundary, just setting $h < m$. During sampling reference lines for the inside, the boundary sampling method is utilized again with the division size for the inside, and then the sampled reference lines are transformed to pass through the centers between the boundary reference lines and the center point $(1/m, \dots, 1/m)$.

Next, the details of each step shown in Algorithm 1, especially the motivations and the reasons of the designs, are specified in Subsections III-B, III-C, III-D and III-E, respectively.

B. NDWA-GA

Algorithm 2 shows the steps of the proposed NDWA-GA algorithm. Specifically, given an MaOP $\mathbf{f}(x)$ to be optimized, the maximal generation number T and the population size L ,

Algorithm 2: Non-dominated Dynamic Weight Aggregation by Using Genetic Algorithm

Input: The MaOP to be optimized

$\mathbf{f}(x) = \{f_1(x), f_2(x), \dots, f_m(x)\}$, the number of generations T , the population size L .

Output: The Pareto-optimal solutions P of $\mathbf{f}(x)$.

```

1  $archive \leftarrow \emptyset$ ;
2  $\mathbf{r} = \{r_1, r_2, \dots, r_k\} \leftarrow$  Generate a set of weight vectors
   by using Das and Dennis's method;
3  $P_0 \leftarrow$  Randomly initialize a population of individuals
   with the size of  $L$ ;
4  $t \leftarrow 0$ ;
5 while  $t < T$  do
6   for  $i \leftarrow 1$  to  $L$  do
7      $ind \leftarrow (t \bmod m) + 1$ ;
8     Let  $i$ -th individual in  $P_t$  model the minimization
       problem of  $\sum_{j=1}^m r_{ind,j} * f_j(x)$ ;
9   end
10  Evaluate the fitness of the individuals in  $P_t$ ;
11   $R \leftarrow$  Generate the offspring from  $P_t$ ;
12   $R \leftarrow R \cup P_t$ ;
13  Let individuals in  $R$  model the problem of  $\mathbf{f}(x)$ ;
14  Evaluate the fitness of the individuals in  $R$ ;
15  Non-dominate the solutions in  $R$  into the Pareto
     fronts  $\{F_1, F_2, \dots\}$ ;
16  Add individuals in  $F_1$  to  $archive$ ;
17   $t \leftarrow t + 1$ ;
18   $P_t \leftarrow \emptyset$ ;
19   $j \leftarrow 1$ ;
20  while  $|P_t| + |F_j| < L$  do
21    Add individuals in  $F_j$  to  $P_t$ ;
22     $j \leftarrow j + 1$ ;
23  end
24  Randomly select  $L - |P_t|$  individuals from  $F_j$  and
     add them to  $P_t$ ;
25 end
26  $P \leftarrow$  Select the non-dominated solutions from  $archive$ ;
27 Return  $P$ .
```

NDWA-GA begins to take effect, and finally a set of Pareto-optimal solutions of $\mathbf{f}(x)$ is obtained. During the execution of NDWA-GA, an archive denoted by *archive* is initialized firstly as an empty set. Then, a set of weight vectors are obtained by transforming the reference lines generated by Das and Dennis's method. This is followed by randomly initialization of the population P_0 . Next, NDWA-GA performs a series of T generations. Finally, the Pareto-optimal solutions are obtained by selecting the non-dominated solutions from the archive.

Furthermore, each repeated iteration in NDWA-GA is composed of four steps. The first is to periodically use the weight vector (line 7) to transform $\mathbf{f}(x)$ into a SOP, where the periodicity is based on the generation, and then let the individuals in the current population to encode the SOP (line 8). When the fitness of these individuals are evaluated (line 10), the offspring are generated using genetic operators (line 11), and the current population is updated by combining the parent

population and the generated offspring (line 12). Specifically, the binary tournament selection, Simulated binary crossover (SBX) [42] and the polynomial mutation [28] operators are used to generate the offspring. The second is to evaluate the fitness of individuals in the current population by encoding $\mathbf{f}(x)$, and divide these individuals into different PFs based on their domination relationship, which are shown in lines 14-15. The third is to copy the individuals in the first PF into the archive, which is shown in line 16. The fourth is to increase the generation number and select L individuals with promising convergence performance from the previous population as the parent solutions of the next generation. Specifically, $L - |P_t|$ individuals are randomly selected from the j -th non-dominated front F_j when $|P_t| + |F_j| \geq L$ and $|P_t| + |F_{j-1}| < L$, where $|\cdot|$ is a countable operator measuring the cardinality. The details of the fourth step are shown in lines 17-24.

As mentioned in Subsection II-B that the proposed NDWA-GA algorithm is inspired by DWA-ES. Next, we will mainly discuss the major differences between NDWA-GA and DWA-ES as well as the motivations why such new components are designed in NDWA-GA. Specifically, NDWA-GA differs from DWA-ES in two aspects: 1) NDWA-GA is designed based on GA while DWA-ES is on ES, and 2) the environmental selection in NDWA-GA is based on non-dominated selection while that in DWA-ES is based on the traditional selection upon SOP such as the roulette wheel selection. First, individuals in ES are generated through the variation by adding a $N(\mu, \lambda)$ to the parent solutions. Due to the unknown size of the search space, μ and λ cannot set to be constant values but to dynamic ones, which are problem-dependent and very sensitive. Although some advanced methods have been proposed to set the proper values, experiences on using ES are still required. However in GA, the promising performance can be achieved by simply setting the probabilities for crossover and mutation. In particular, genetic operator type and the parameters in the used genetic operators based on the convention from the EA community, such as the probabilities for crossover and mutation setting to be 0.9 and 0.1, respectively, can handle well the problems in most situations. Second, the individuals surviving into the next generation in DWA-ES are selected based on their fitness of the transformed SOP. However, an individual with good fitness regarding the SOP cannot show its promising convergence in the search space of the corresponding MaOP. For MOP, the deficiency of using such environmental selection is not significant due to the small search space, however, the performance becomes poor when addressing MaOPs. Because the non-dominated solutions are typically used as the convergence measurement for the individuals in solving MOPs and MaOPs, a revised approach adopting non-dominated sorting is expectedly used in such a situation. As a result, NDWA-GA is proposed. The experimental results in Section IV will show that our design in NDWA-GA against DWA-ES is significantly improved in solving not only MOPs but also MaOPs.

C. Pareto-optimal Subspace Learning

The steps of the PoS learning are shown in Algorithm 3 which is composed of two parts: finding the principal com-

Algorithm 3: Pareto-optimal Subspace Learning

Input: The Pareto-optimal solutions P generated by NDWA-GA, a threshold value ϵ , the lower bound $l_limit = [l_1, l_2, \dots, l_n]$ of x , and the upper bound $u_limit = [u_1, u_2, \dots, u_n]$ of x .

Output: The learned Pareto-optimal subspace Ω .

```

1 Let a matrix  $M \in R^{|P| \times n}$  denote the solutions in  $P$ ;
2  $mean \leftarrow$  Calculate the mean values of  $M$ ;
3  $M' \leftarrow$  Update  $M$  by subtracting  $mean$ ;
4  $[v_1, v_2, \dots, v_n] \leftarrow$  Calculate the eigenvalues of the
  covariance matrix of  $M'$ ;
5  $[ind_1, ind_2, \dots, ind_n] \leftarrow$  Record the positions of the
  eigenvalues with the descending order;
6  $j \leftarrow 0$ ;
7 while  $\sum_{i=1}^{ind_j} v_i / \sum_{i=1}^n v_i < \epsilon$  do
8   |  $j \leftarrow j + 1$ ;
9 end
10  $median \leftarrow$  Calculate the median value of each column
   in  $M$ ;
11 for  $i \leftarrow j + 1$  to  $n$  do
12   |  $l\_i \leftarrow$  Copy the  $ind_i$  component of  $median$ ;
13   |  $u\_i \leftarrow$  Copy the  $ind_i$  component of  $median$ ;
14 end
15  $\Omega \leftarrow$  All points in  $[l\_limit, u\_limit]$ ;
16 Return  $\Omega$ .
```

ponents of the Pareto-optimal space with the given threshold ϵ , and padding the non-principal components with the corresponding median values. These two steps are shown in lines 1-9 and 10-15, respectively. Specifically, the Principal Component Analysis (PCA) method is used in the first part. Firstly, the Pareto-optimal solutions in P are represented by a matrix $M \in R^{|P| \times n}$ where each row denotes a solution and each column denotes the same component of all solutions (line 1). Secondly, the mean values $mean$ of M are calculated, and M is updated to M' by subtracting $mean$ (lines 2-3). Thirdly, the eigenvalues $[v_1, v_2, \dots, v_n]$ of the covariance matrix of M' are calculated (line 4). Fourthly, the principal components are found based on the given threshold ϵ (lines 5-9). In the second part, the median values of each column in M are calculated first (line 10), and then the elements in the lower and upper bounds of x are set to be the corresponding median values if they belong to the non-principal components (lines 11-14). Finally, the PoS is generated by sampling all the points in the updated lower and upper bounds (line 15).

The reason that PCA is used in the first step is considered from two aspects: 1) PCA is a widely used dimension reduction technique, and its implementations with specifically optimized techniques are widely available in different libraries, and 2) PCA has been successfully used to reduce the search space in addressing MaOPs in our previous works [10], [11]. In addition, the median values are used here to pad the non-principal components instead of the mean values that is a standard operation in PCA. The reason is that not all the non-dominated solutions in P are precisely lying in the PF, and a small partition of them are beyond the PF especially

in the case of a larger number of objectives. These non-dominated solutions not standing at the PF are called “outliers” in dimension reduction. Removing the outliers from the entire dataset is not an easy task, while the outliers typically affect the quality of the resulted outcome. In this regard, using median values will generate a more accurate PoS than using mean values. For example, the element values of one non-principle component are $[0.3, 0.4, 0.5, 0.5, 0.5]$ and the ground-truth on this component is 0.5. Consequently, the mean value is $(0.3+0.4+0.5+0.5+0.5)/5 = 0.44$, while the median value is 0.5. Obviously, the resulted PoS by using the mean values is different from that by the median values. Although there is a slight difference between the mean value and the median value in one component just as this example shows, significantly inferior performance will be caused when such differences are accumulated to multiple non-principal components, which have been observed during the design of Algorithm 3.

D. Reference Lines Mapping

Algorithm 4: Reference Lines Mapping

Input: The MaOP to be optimized

$\mathbf{f}(x) = \{f_1(x), f_2(x), \dots, f_m(x)\}$, the Pareto-optimal solutions P generated by NDWA-GA, the uniformly distributed reference lines $\mathbf{r} = \{r_1, r_2, \dots, r_N\}$ generated by using Das and Dennis’s method.

Output: The updated reference lines \mathbf{r} .

```

1  $extreme\_points \leftarrow \emptyset$ ;
2 for  $i \leftarrow 1$  to  $m$  do
3   |  $s = [s_1, \dots, s_m] = [0, \dots, 1, \dots, 0] \leftarrow$  Initialize a
     | vector of  $m$  elements with zeros except for the  $i$ -th
     | element with one;
4   |  $x^* \leftarrow$  Solve the single-objective optimization
     | problem formulated by (2);
5   |  $extreme\_points \leftarrow extreme\_points \cup \mathbf{f}(x^*)$ ;
6 end
7  $nadir\_point \leftarrow \emptyset$ ;
8  $ideal\_point \leftarrow \emptyset$ ;
9 for  $i \leftarrow 1$  to  $m$  do
10  |  $nadir\_point \leftarrow nadir\_point \cup$  The maximal
    | element at the  $i$ -th component of  $extreme\_points$ ;
11  |  $ideal\_point \leftarrow ideal\_point \cup$  The minimal element
    | at the  $i$ -th component of  $extreme\_points$ ;
12 end
13 for each  $r$  in  $\mathbf{r} = \{r_1, r_2, \dots, r_N\}$  do
14   |  $r \leftarrow r \times (nadir\_point - ideal\_point) + ideal\_point$ ;
15 end
16 Return  $\mathbf{r}$ .
```

The steps of mapping the reference lines are shown in Algorithm 4 that is composed of three parts. They are finding the extreme points, extracting the nadir and ideal points, and mapping each reference line, which are detailed in lines 1-6, 7-12 and 13-15, respectively. In finding the extreme points, each extreme point is obtained by solving a SOP in the PoS Ω that is formulated by (2).

$$\begin{cases} \text{maximize} & \frac{\sum_{i=1}^m f_i(x)s_i}{\sqrt{\sum_{i=1}^m f_i^2(x)}\sqrt{\sum_{i=1}^m s_i^2}} \\ \text{s.t.} & x \in \Omega \end{cases} \quad (2)$$

Intuitively, it is to find the solution of which the objective vector has the closest included angle to the corresponding axis, i.e., the maximal cosine value. In addition, the nadir point and the ideal point is obtained by extracting the maximal value and minimal value from each dimension of the extreme points, and the mapping of each reference line is achieved by the algebraic operation shown in line 14 of Algorithm 4.

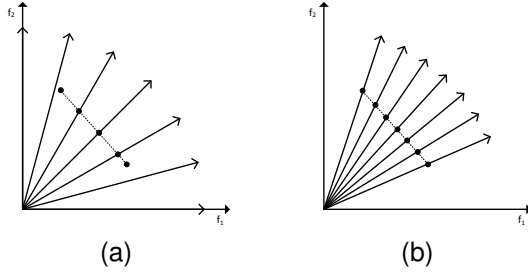


Fig. 4. The illustration regarding the intention of mapping the reference lines where the arrow lines, dash lines, and the solid circles denote the reference lines, Pareto front and the resulted Pareto-optimal solutions for maintaining the promising diversity. Specifically, reference lines in Fig. 4a refer to the ones generated by the Das and Dennis's method and in Fig. 4b refer to the ones after mapping.

The reason of mapping the reference lines is that such N reference lines generated by using Das and Dennis's method span the entire space of R_+^m . However, the PF of the MaOP does not necessarily occupy the entire space of R_+^m , i.e., only a part of such N reference lines pass through the PF. Because the promising diversity is achieved by finding the solution that has the minimal included angle to each reference line, the resulted final solutions will loss the diversity under such a situation. The aim of mapping the reference lines is to keep such predefined N reference lines within the PF for keeping the promising diversity. For a better understanding the intention of mapping the reference lines, an example is shown in Fig. 4 where the arrow lines, dash lines and solid circles denote the reference lines, PF and resulted Pareto-optimal solutions with diversity. It can be observed from Fig. 4a that the seven reference lines generated by Das and Dennis's method span the entire space of R_+ but the PF does not. In such situation, the resulted solutions are with poor diversity. Noting that there are two overlapping solutions at each end of the PF shown in Fig. 4a. The reference lines that have been mapped with Algorithm 4 are shown in Fig. 4b, where it can be investigated that the resulted Pareto-optimal solutions are with good diversity.

Noting that the extreme points in MaOEA-IT are obtained by repeatedly solving m times of the SOP formulated by (2) with m different v 's from the PoS, but choosing them from the Pareto-optimal solutions P generated by NDWA-GA in Subsection III-B. The reason is that all the solutions in P do not necessarily cover all the extreme points. If the extreme

points are forced to be obtained from P , the results would be with poor accuracy. In addition, there are also existing algorithms, such as NSGA-III, extracting the nadir points to perform the operation mapping the reference lines by using the intercepts, and the intercepts are from the hyperplane passing through the extreme points extracted from the Pareto-optimal solutions. However, the resulted nadir point is also inaccurate and not suitable for the proposed MaOEA-IT. NSGA-III can use such nadir points due to the updating strategy in its subsequent generations. As shown above, the proposed MaOEA-IT is capable of searching for the nadir points. To justify the fact, further discussion and experiments of MaOEA-IT on searching the nadir points are shown in the Supplemental Materials.

E. Diversity Maintenance

This is the final step of the proposed MaOEA-IT where a population of solutions with promising diversity is obtained from the PoS.

Algorithm 5: Diversity Maintaining

Input: The MaOP to be optimized

$\mathbf{f}(x) = \{f_1(x), f_2(x), \dots, f_m(x)\}$, the
Pareto-optimal subspace Ω found in Algorithm 3,
the updated reference lines $\mathbf{r} = \{r_1, r_2, \dots, r_N\}$
generated in Algorithm 4.

Output: The solutions with promising diversity Q .

```

1  $Q \leftarrow \emptyset$ ;
2 for  $i \leftarrow 1$  to  $N$  do
3    $s \leftarrow r_i$ ;
4    $x^* \leftarrow$  Solve the single-objective optimization
   problem formulated by (2);
5    $Q \leftarrow Q \cup x^*$ ;
6 end
7 Return  $Q$ .
```

The steps of diversity maintenance are shown in Algorithm 5, where each solution is obtained by solving the SOP (2) with a set of reference lines \mathbf{r} . Because reference lines in \mathbf{r} are uniformly distributed in the PF, and each solution in Q is with minimal cosine value to each reference line, solutions in Q are with promising diversity. In addition, the decision space of (2) is the PoS from which any solution sampled is with promising convergence. Therefore, the solutions in Q are also with promising convergence. In summary, solutions in Q are promising in terms of both convergence and diversity.

F. Computational Complexity

For the convenience of the analysis, it is assumed that the population size in Algorithm 2, the population size for solving (2), and the number of final solutions are set to be N . The number of generations for solving (2) is set to be T . Specifically, in Algorithm 2, the non-dominated sorting accounts for the most computational time. However, it is hard to determine the exact number of solutions in the archive. To this end, we could only calculate the computational

complexity under the worst case scenario, i.e., the generated N offspring are all non-dominated, and TN solutions in the archive. Consequently, the computational complexity of Algorithm 2 is $O((TNm)^2)$. In Algorithm 3, the calculation of the eigenvalues contributes to the major computational complexity. Calculating the eigenvalues involves the matrix factorization, of which the computational complexity is $O(n^3)$ where n denotes the number of decision variables. The computation time of Algorithm 4 is mainly from solving the SOP. Because the SBX and polynomial mutation are used for generating the offspring, the total computational complexity of Algorithm 4 is $O(m^2n)$. As a consequence, the computational complexity of Algorithm 5 is $O(Nmn)$. Because $n \gg m$ in MaOPs and $T \gg n$, the computational complexity of Algorithms 2 and 5 approaches $O((TNm)^2)$, and that of Algorithms 3 and 4 is $O(n^3)$. As a result, the worst computational complexity of the proposed MaOEA-IT algorithm is $O((TNm)^2)$ or $O(n^3)$ whichever is larger, while the best is $O(n^3)$. In summary, the non-dominated sorting accounts for the major computational complexity when the number of non-dominated solutions is large. Otherwise, the calculation of the eigenvalues for learning the PoS accounts for the major computational complexity.

IV. EXPERIMENTS

In order to examine the performance of the proposed MaOEA-IT algorithm in solving MaOPs, a series of experiments is performed against peer competitors. Particularly, NSGA-III [8] is selected firstly because it is typically used as a baseline in comparing the performance of MaOEAs. Additionally, the recently proposed or the most well-regarded ones are selected from each of the four categories mentioned in Section I as the corresponding representative peer competitors. They are RVEA [12], MOEA/D [13], KnEA [21] and MaOEA-IGD [25]. Furthermore, the MaOEA-RD [29] which is an MaOEA based on the two-stage strategy is also selected as the peer competitor.

The remaining of this section is organized as follows. Firstly, the benchmark test problems utilized in the comparisons are introduced in Subsection IV-A. Secondly, the chosen performance metric is given in Subsection IV-B to quantitatively evaluate the quality of the resulted solutions of all competing MaOEAs. Thirdly, the parameter settings used in all the algorithms considered are enumerated in Subsection IV-C, and experimental results measured by the chosen performance metric are presented and analyzed in Subsection IV-D. Fourthly, the performance of the proposed NDWA-GA algorithm is evaluated in Subsection IV-E by comparing to the pioneering work (i.e., the DWA-ES).

A. Benchmark Test Problems

The extensively used test problems, DTLZ1-DTLZ7 with configurable numbers of objectives from the DTLZ benchmark test suite [33], are used in the experiments. Particularly, each objective $f_i(x)$ of one given test problem from DTLZ is formulated by (3):

$$f_i(x) = g(x_1, \dots, x_{m-1}) \cdot h(x_m, \dots, x_{z+m-1}) \quad (3)$$

where $n = z + m - 1$ is the number of the decision variables, m is the number of objectives, and z is a configurable parameter that is set to be 5 for DTLZ1, 10 for DTLZ2-DTLZ6 and 20 for DTLZ7. Specifically, $g(\cdot)$ is controlled by the first $m - 1$ decision variables to measure the distribution of the solutions in the PF, and $h(\cdot)$ is controlled by the last $z + m - 1$ decision variables to determine the distance to the PF. Furthermore, diverse complicated forms of $g(\cdot)$ have been implemented, such as multi-modal, concave, biased, disconnected and deceptive, which raises the challenges for the MaOEAs and is sufficient to conclude whether the tested MaOEA is effective or not in solving MaOPs. Noting that the DTLZ benchmark test suite used in these experiments follows the version introduced in the original paper [33].

B. Performance Metric

The widely used Hypervolume (HV) [43], which is able to concurrently measure the convergence and diversity of the tested MaOEAs, is chosen as the performance metric in these experiments. Based on the suggestions in [9], [11], the resulted solutions by the competing MaOEAs are normalized prior to calculating their HV values. Because of the computational cost of HV significantly growing as the number of objectives increases, Monte Carlo simulation [44]² is used for the calculation when $m \geq 10$. Otherwise the precise approach proposed in [45] is utilized³. Furthermore, the HV values are normalized for the comparisons.

C. Parameter Settings

1) *Number of Objectives*: Because the proposed MaOEA-IT algorithm focuses on solving the MaOPs of which the number of objectives is larger than three, the 5-, 8-, 10- and 15-objective test problems are investigated. However, the experimental results are difficult to visually analyze when the number of objectives is larger than three. Therefore, the 3-objective test problems are also investigated in these experiments.

2) *Crossover and Mutation*: SBX [42] and polynomial mutation [28] are used as the crossover and mutation operators, respectively, in the competing MaOEAs. Based on the conventions, the probabilities of SBX and polynomial mutation are set to be 1.0 and $1/n$ (n is the number of decision variables), respectively. The distribution index of SBX and polynomial mutation are set to be 20, in addition to the NSGA-III whose distribution index needs to be specified as 30 based on its developers [8].

3) *Population Size*: The population sizes for all competing MaOEAs on each considered number of objectives are all set to be 240 for the reason of simplicity. However, the reference lines-based algorithms, such as NSGA-III, require the number of solutions to be the same as that of the generated reference lines. However, the reference lines generated by Das and Dennis's method is a binomial coefficient, which cannot be arbitrarily set. In order to do a fair comparison, the population

²The source code is available at: <http://www.tik.ee.ethz.ch/sop/download/supplementary/hype/>.

³The source code is available at: <http://www.wfg.csse.uwa.edu.au/hypervolume/>.

size of the reference lines-based algorithms is set to be the number that is the closest to 240 by choosing an appropriate division size. As mentioned above, when the division size is less than m when generating the reference lines, the resulted solutions will have a poor diversity. To this end, the two-layer approach [8] is employed when $m > 5$. As a result, the division sizes for 3-, 5-, 8-, 10- and 15-objective test problems are set to be [20], [6], [3, 3], [3, 2] and [2, 2], respectively. Consequently, the corresponding population sizes are set to be 231, 210, 240, 275 and 240, respectively.

4) *Number of Function Evaluations*: The maximal generation number for peer competitors are set to be with the order of 10^2 when $m < 10$ and 10^3 otherwise based on the conventions [8], [10]–[12]. Because of the two-stage nature of the proposed MaOEA-IT algorithm, the total numbers of function evaluation of MaOEA-IT in the two-stage are set to be the same to that of the peer competitors for a fair comparison. For this reason, the maximal generation number is transformed to the maximal numbers. Consequently, the maximal numbers of function evaluations are set to be 1.4×10^5 , 2.4×10^5 , 4.0×10^5 , 2.0×10^6 and 3.0×10^6 for 3-, 5-, 8-, 10- and 15-objective test problems, respectively. They are approximately 200 generations when $m < 10$ and 800 otherwise.

5) *Statistical Approach*: Because of the heuristic characteristic of MaOEAs, each competing MaOEA is independently performed 30 times with the same configurations, and the resulted HV values are statistically evaluated. Based on the conventions, the Mann-Whitney-Wilcoxon rank-sum test [46] with a 5% significance level is employed for this purpose.

D. Overall Results

The experimental results of the proposed MaOEA-IT algorithm against NSGA-III, RVEA, MOEA/D, KnEA, MaOEA-IGD and MaOEA-RD on 3-, 5-, 8-, 10- and 15-objective DTLZ1-DTLZ7 test problems are shown in Table I. In Table I, the numbers with bold face indicate the best mean values among all competing MaOEAs for the same numbers of objectives of each test problem. The symbols “+” and “-” denote whether the null hypothesis, i.e., the results of MaOEA-IT are better or worse than the peer competitors, is accepted by the Mann-Whitney-Wilcoxon rank-sum test at the significance level of 5%, while symbol “=” implies the corresponding null hypothesis is rejected. In addition, the numbers shown next to symbols “+,” “-,” and “=” in the last row summarize how many times MaOEA-IT performs significantly better than, worse than, and equal to, respectively, the corresponding peer competitors over each test problem on all considered numbers of objectives.

As can be seen in Table I, MaOEA-IT wins the best scores against all peer competitors on DTLZ1 with all considered numbers of objectives, which is the same to DTLZ2 and DTLZ3. For DTLZ4, MaOEA-IT achieves the best performance on the 5-, 8- and 15-objective test problems. Although the winner on 3-objective DTLZ4 is MOEA/D, MaOEA-IT shows the statistically equal performance to MOEA/D. On 10-objective DTLZ4, MaOEA-IT shows a little worse than RVEA which is credited as the winner. Although MOEA/D

wins the best performance on DTLZ5 with 8 and 10 objectives, MaOEA-IT shows the best performance on the 3-, 5- and 15-objective problems. In addition, MaOEA-IT also shows the best performance on DTLZ6 with 3, 10 and 15 objectives, but performs worse on 5 and 8 objectives on which RVEA and MOEA/D are the winners, respectively. Furthermore, MaOEA-IT is also the winner upon DTLZ7 except for its 10-objective test problem on which KnEA shows the best performance.

Based on the shapes of PF, DTLZ1-DTLZ7 can be divided into two categories. The first category includes DTLZ1-DTLZ4 of which the extreme points are located at the axes, while the other is composed of DTLZ5-DTLZ7 of which not all the extreme points are located at the axes. From Table I, it can also be clearly seen that MaOEA-IT performs better on DTLZ1-DTLZ4 than on DTLZ5-DTLZ7, which is caused by the reference line mapping in line 3 of Algorithm 1. To be specific, the nadir point estimation in MaOEA-IT is based on the extreme-point-to-nadir scheme, i.e., the extreme points are supposed to stand at the axes. However, the test problems are not necessary with such an assumption, which causes the inaccurate estimated nadir point from the extreme points. This can be viewed as the limitation of the nadir point estimation method based on the extreme-point-to-nadir scheme. However, the advantage of the extreme-point-to-nadir scheme remains in its significantly cheaper computational cost. Furthermore, MaOEA-IT still wins most of the test problems in DTLZ5-DTLZ7.

In summary, the proposed MaOEA-IT algorithm wins 164 out of the 210 competitions, which assure the superior performance of MaOEA-IT in addressing DTLZ1-DTLZ7 test problems with 3-, 5-, 8-, 10- and 15-objective problems. We believe the promising performance of the proposed algorithm is mainly owing to 1) its two-stage strategy where the converge and diversity are treated at different stages, and 2) the designed NWDA-GA method in finding the accurate PoS in the first stage. Specifically, the peer competitors used in this experiment can be classified into two different categories, i.e., the traditional MaOEAs which pursue the promising convergence and diversity in the same generation, and the two-stage strategy based MaOEAs which treat the convergence and diversity in different generations. Particularly, as discussed in Section I, the individuals with promising convergence are typically with poor diversity, and vice versa. For the peer competitors in the first category, the individuals selected by them are a compromise between the convergence and diversity, but not with both good convergence and diversity. As a result, the performance measured by HV may not be promising. For the peer competitors in the second category, two limitations exist for them; the first one is the incorrect PoS may be resulted and the second is the inaccurate extreme points they employed. For the benchmark problems, the first limitation does not exist, while the second limitation affects more about the performance, which is the main reason that peer competitors from this category show worse performance than the proposed algorithm. In addition, the proposed algorithm is composed of two components, i.e., the convergence component and the diversity component. The convergence component is viewed as the main contributor to the promising performance of the

TABLE I

HV RESULTS OF MAOEA-IT AGAINST NSGA-III, RVEA, MOEA/D, KNEA, MAOEA-IGD AND MAOEA-RD OVER DTLZ1-DTLZ7 TEST PROBLEMS WITH 3, 5, 8-, 10, AND 15 OBJECTIVES.

		3-objective	5-objective	8-objective	10-objective	15-objective
DTLZ1	MaOEA-IT	0.9756(1.773e-04)	0.9988(2.880e-05)	0.9996(9.485e-05)	0.9998(8.753e-05)	0.9996(1.963e-04)
	NSGA-III	0.9660(4.514e-05)(+)	0.9889(3.983e-05)(+)	0.9900(3.895e-06)(+)	0.9900(1.638e-06)(+)	0.9900(6.963e-06)(+)
	RVEA	0.9662(5.058e-07)(+)	0.9889(4.170e-05)(+)	0.9900(5.236e-06)(+)	0.9900(9.679e-07)(+)	0.9900(0.000e+00)(+)
	MOEA/D	0.9662(4.726e-07)(+)	0.9890(3.615e-05)(+)	0.9900(1.871e-06)(+)	0.9900(7.071e-07)(+)	0.9609(4.509e-03)(+)
	KnEA	0.9076(2.168e-02)(+)	0.6546(9.013e-02)(+)	0.8523(3.258e-02)(+)	0.0000(0.000e+00)(+)	0.0000(0.000e+00)(+)
	MaOEA-IGD	0.9565(3.713e-04)(+)	0.9936(3.171e-04)(+)	0.9746(4.264e-04)(+)	0.9982(4.003e-04)(=)	0.9815(3.425e-04)(+)
	MaOEA-RD	0.9739(2.681e-04)(=)	0.9974(9.890e-05)(=)	0.9651(9.191e-03)(+)	0.9096(2.183e-02)(+)	0.9107(2.932e-02)(+)
DTLZ2	MaOEA-IT	0.9295(8.891e-05)	0.9897(1.551e-04)	0.9991(6.549e-05)	0.9996(5.484e-05)	0.9997(6.834e-05)
	NSGA-III	0.9203(2.579e-05)(+)	0.9805(7.965e-05)(+)	0.9981(3.538e-04)(=)	0.9994(1.174e-04)(=)	0.9898(4.632e-05)(+)
	RVEA	0.9205(6.100e-07)(+)	0.9805(5.425e-05)(+)	0.9893(2.837e-05)(+)	0.9899(8.268e-06)(+)	0.9900(1.191e-06)(+)
	MOEA/D	0.9205(6.979e-08)(+)	0.9806(7.905e-05)(+)	0.9894(1.963e-05)(+)	0.9899(9.330e-06)(+)	0.9900(1.304e-06)(+)
	KnEA	0.9280(1.621e-04)(=)	0.9897(1.164e-04)(=)	0.9990(4.016e-05)(=)	0.9899(1.191e-05)(+)	0.9900(9.515e-07)(+)
	MaOEA-IGD	0.9216(3.300e-04)(+)	0.9771(1.869e-04)(+)	0.9875(5.212e-04)(+)	0.9768(2.759e-04)(+)	0.9870(3.737e-04)(+)
	MaOEA-RD	0.9174(1.711e-03)(+)	0.9698(3.733e-03)(+)	0.9575(5.797e-03)(+)	0.9069(2.260e-02)(+)	0.7854(7.653e-02)(+)
DTLZ3	MaOEA-IT	0.9293(1.733e-04)	0.9902(1.817e-04)	0.9991(6.494e-05)	0.9997(8.479e-05)	0.9998(5.811e-05)
	NSGA-III	0.9204(1.708e-05)(+)	0.9806(7.297e-05)(+)	0.9936(1.705e-03)(+)	0.9960(1.252e-03)(=)	0.9966(7.389e-04)(=)
	RVEA	0.9204(5.996e-06)(+)	0.9804(1.187e-04)(+)	0.9894(1.910e-05)(+)	0.9899(8.756e-06)(+)	0.9900(1.137e-06)(+)
	MOEA/D	0.9204(3.328e-05)(+)	0.9804(1.683e-04)(+)	0.9893(2.546e-05)(+)	0.9899(7.190e-06)(+)	0.9871(5.492e-03)(+)
	KnEA	0.8133(3.084e-02)(+)	0.5495(1.188e-01)(+)	0.0000(0.000e+00)(+)	0.0000(0.000e+00)(+)	0.0000(0.000e+00)(+)
	MaOEA-IGD	0.8924(3.334e-04)(+)	0.9772(1.073e-04)(+)	0.9683(1.776e-04)(+)	0.9875(3.113e-04)(+)	0.9773(3.158e-06)(+)
	MaOEA-RD	0.9244(7.885e-04)(=)	0.9769(2.351e-03)(+)	0.9527(1.026e-02)(+)	0.6618(1.073e-01)(+)	0.0046(2.671e-01)(+)
DTLZ4	MaOEA-IT	0.9186(6.153e-03)	0.9833(1.144e-03)	0.9912(2.333e-03)	0.9819(6.415e-03)	0.9975(1.794e-03)
	NSGA-III	0.9204(2.179e-05)(=)	0.9730(3.929e-03)(+)	0.9893(3.084e-05)(=)	0.9889(3.329e-04)(-)	0.9900(1.127e-05)(+)
	RVEA	0.9205(8.934e-07)(=)	0.9805(7.441e-05)(=)	0.9893(2.414e-05)(=)	0.9899(9.612e-06)(-)	0.9900(1.276e-06)(+)
	MOEA/D	0.9205(5.675e-08)(=)	0.9804(1.280e-04)(=)	0.9860(1.392e-03)(+)	0.9895(1.901e-04)(-)	0.9899(2.194e-05)(+)
	KnEA	0.9187(1.590e-04)(=)	0.9799(1.536e-04)(=)	0.9892(4.954e-05)(=)	0.9899(1.167e-05)(-)	0.9900(1.021e-06)(+)
	MaOEA-IGD	0.9115(2.088e-04)(+)	0.9725(3.592e-04)(+)	0.9803(3.515e-04)(+)	0.9725(5.241e-04)(+)	0.9871(3.205e-04)(+)
	MaOEA-RD	0.9128(3.018e-04)(+)	0.9808(7.628e-04)(=)	0.9889(1.552e-03)(=)	0.9732(3.111e-03)(+)	0.9647(8.225e-03)(+)
DTLZ5	MaOEA-IT	0.7578(5.256e-04)	0.7240(7.017e-04)	0.5686(1.996e-01)	0.5471(1.840e-01)	0.6606(8.474e-04)
	NSGA-III	0.7544(2.362e-04)(=)	0.6504(1.830e-02)(+)	0.5540(3.701e-02)(+)	0.5219(4.239e-02)(+)	0.5519(2.226e-02)(+)
	RVEA	0.7526(1.471e-04)(+)	0.6776(2.604e-03)(+)	0.5807(9.214e-03)(-)	0.5466(3.898e-02)(=)	0.5000(4.388e-02)(+)
	MOEA/D	0.7505(4.455e-07)(+)	0.6586(1.783e-02)(+)	0.6939(3.904e-04)(-)	0.6848(7.182e-04)(-)	0.5583(6.623e-02)(+)
	KnEA	0.7546(1.191e-04)(=)	0.5846(3.659e-02)(+)	0.6700(5.557e-03)(-)	0.6341(1.020e-02)(-)	0.0000(1.473e-01)(+)
	MaOEA-IGD	0.7228(2.736e-04)(+)	0.5697(5.903e-05)(+)	0.5753(4.039e-04)(-)	0.5308(8.704e-05)(+)	0.4322(3.353e-04)(+)
	MaOEA-RD	0.5000(0.000e+00)(+)	0.5000(0.000e+00)(+)	0.5000(0.000e+00)(+)	0.5000(0.000e+00)(+)	0.5000(0.000e+00)(+)
DTLZ6	MaOEA-IT	0.7578(4.034e-04)	0.6627(1.812e-02)	0.6229(4.146e-02)	0.6776(3.714e-04)	0.6604(1.535e-03)
	NSGA-III	0.7537(2.719e-04)(=)	0.6761(1.034e-02)(-)	0.3766(7.599e-02)(+)	0.0000(2.536e-01)(+)	0.0000(2.065e-01)(+)
	RVEA	0.7508(7.625e-04)(+)	0.6779(5.757e-03)(-)	0.6213(1.081e-02)(=)	0.6410(7.605e-03)(+)	0.4413(4.722e-02)(+)
	MOEA/D	0.7505(1.164e-07)(+)	0.7235(4.708e-04)(-)	0.6891(2.210e-03)(-)	0.5220(9.244e-03)(+)	0.4422(1.205e-02)(+)
	KnEA	0.7558(1.148e-05)(=)	0.6630(6.876e-03)(=)	0.5581(3.181e-02)(+)	0.0965(1.236e-01)(+)	0.0000(1.141e-01)(+)
	MaOEA-IGD	0.6651(4.619e-04)(+)	0.5944(2.693e-04)(+)	0.4293(3.525e-04)(+)	0.3208(7.166e-05)(+)	0.3305(2.501e-04)(+)
	MaOEA-RD	0.5000(0.000e+00)(+)	0.5000(0.000e+00)(+)	0.5000(0.000e+00)(+)	0.5000(0.000e+00)(+)	0.4950(0.000e+00)(-)
DTLZ7	MaOEA-IT	0.6694(4.010e-04)	0.6636(7.589e-04)	0.6482(1.098e-03)	0.6154(2.200e-03)	0.5507(3.113e-03)
	NSGA-III	0.5991(5.854e-04)(+)	0.6138(8.956e-03)(+)	0.5704(1.327e-02)(+)	0.5452(1.466e-02)(+)	0.4820(1.055e-02)(+)
	RVEA	0.5985(3.989e-04)(+)	0.6168(1.394e-03)(+)	0.6162(5.578e-03)(+)	0.5909(1.062e-02)(+)	0.5447(2.500e-02)(+)
	MOEA/D	0.6007(6.764e-06)(+)	0.6417(2.737e-04)(+)	0.4194(4.503e-02)(+)	0.4089(3.813e-02)(+)	0.3774(4.266e-02)(+)
	KnEA	0.5979(8.629e-04)(+)	0.6123(9.458e-03)(+)	0.6408(7.469e-03)(+)	0.6430(1.070e-02)(-)	0.5280(1.017e-02)(+)
	MaOEA-IGD	0.3961(3.574e-06)(+)	0.4087(3.282e-05)(+)	0.4005(2.532e-04)(+)	0.5041(3.837e-06)(+)	0.4324(2.061e-04)(+)
	MaOEA-RD	0.4252(4.062e-02)(+)	0.4002(1.490e-02)(+)	0.5280(1.997e-02)(+)	0.5134(2.326e-02)(+)	0.5239(8.403e-03)(+)
+/-		31/11/0	32/7/3	30/7/5	31/4/7	40/1/1

proposed algorithm. The main reason is that the diversity component used in the proposed algorithm is also used by the peer competitors, such as NSGA-III and RVEA. Nevertheless, the proposed algorithm shows the promising performance among these peer competitors.

E. Investigation Between NDWA-GA and DWA-ES

In this subsection, the performance of the proposed NDWA-GA in finding the Pareto-optimal solutions is investigated through the comparison against its preceding work, i.e., DWA-ES. To evaluate the quality of the generated Pareto-optimal so-

lutions, two performance indicators, the Generational Distance (GD) [28] and the Maximum Pareto Front Error (MPFE) [47] which are specifically designed to measure the convergence of the solutions, are applied in this investigation.

Specifically, GD and MPFE are mathematically represented by (4) and (5), respectively,

$$GD = \sqrt{\frac{\sum_{p \in P} \min_{pf \in PF} \|p - pf\|_2}{|P|}} \quad (4)$$

$$MPFE = \max_{p \in P} \min_{pf \in PF} \|p - pf\|_2 \quad (5)$$

TABLE II
THE CONVERGENCE OF NDWA-GA AND DWA-ES MEASURED BY GD AND MPFE.

		GD Indicator		MPFE Indicator	
		NDWA-GA	DWA-ES	NDWA-GA	DWA-ES
DTLZ1	3	1.116378E-05(3.801440E-06)	2.011676E-03(4.260983E-03)(+)	4.996066E-04(2.904490E-04)	2.739488E-01(3.456044E-01)(+)
	5	7.429154E-05(3.633830E-05)	2.164111E-02(9.207933E-04)(+)	2.033603E-03(4.772295E-03)	2.309088E-01(4.215690E-01)(+)
	8	1.745770E-04(7.215127E-04)	4.514194E-02(2.984266E-04)(+)	4.935391E-02(4.812665E-03)	6.737321E-01(1.060792E-01)(+)
	10	1.712473E-04(6.079962E-05)	5.367918E-02(2.384873E-04)(+)	3.048288E-01(2.429143E-04)	1.267158E-01(9.051564E-01)(+)
	15	2.738109E-04(4.173079E-05)	8.154291E-02(6.843897E-02)(+)	9.184138E-01(6.704740E-04)	3.083838E-01(4.878440E-01)(+)
DTLZ2	3	2.254637E-05(9.106223E-05)	6.828068E-03(6.964152E-05)(+)	8.246289E-04(4.004944E-05)	6.047725E-01(2.623298E-01)(+)
	5	1.737555E-04(4.703591E-04)	7.633826E-02(8.141903E-03)(+)	1.327162E-03(2.990603E-03)	9.310400E-01(3.674385E-01)(+)
	8	3.608976E-04(4.060436E-05)	1.422377E-01(5.481838E-02)(+)	7.725854E-02(9.548772E-02)	8.073665E-01(7.051854E-01)(+)
	10	3.316103E-04(4.174040E-04)	1.740351E-01(8.800398E-02)(+)	4.619234E-01(6.320157E-03)	6.620679E-01(8.816138E-01)(+)
	15	4.836233E-04(7.855027E-04)	3.111430E-01(2.251463E-03)(+)	9.139047E-01(6.931652E-02)	6.362358E-01(9.921861E-01)(+)
DTLZ3	3	2.203975E-05(4.319371E-04)	6.738774E-03(2.620434E-04)(+)	4.248374E-03(4.736167E-04)	2.585487E-01(9.340812E-01)(+)
	5	1.394320E-04(2.164811E-06)	6.445229E-02(7.243770E-02)(+)	1.603100E-02(8.976379E-03)	7.061242E-01(4.132775E-02)(+)
	8	2.454213E-04(9.153132E-04)	1.924174E-01(6.238226E-02)(+)	6.683038E-01(1.455276E-03)	1.191811E-01(6.425888E-02)(+)
	10	3.954295E-04(9.997008E-05)	1.840918E-01(3.309933E-02)(+)	9.343235E-01(7.210347E-02)	8.654473E-01(6.243297E-01)(+)
	15	4.084966E-04(6.931363E-05)	2.708953E-01(7.918663E-03)(+)	9.249660E-01(1.326772E-01)	5.588227E-01(3.930256E-01)(+)
DTLZ4	3	2.770528E-05(3.051793E-05)	5.777959E-03(4.167487E-02)(+)	9.826194E-05(5.442619E-03)	6.644040E-02(9.916023E-01)(+)
	5	1.204319E-04(7.474354E-04)	6.600408E-02(9.735102E-04)(+)	6.761374E-03(8.745897E-02)	9.557402E-01(1.413901E-01)(+)
	8	2.855726E-04(9.948585E-05)	2.529676E-01(6.881755E-02)(+)	6.424232E-01(1.592495E-03)	7.305326E-01(6.196558E-01)(+)
	10	3.434044E-04(2.329267E-06)	1.798878E-01(5.188940E-04)(+)	6.515113E-01(2.705081E-02)	7.967081E-01(9.963083E-01)(+)
	15	5.211931E-04(3.316237E-04)	4.282656E-01(8.293367E-02)(+)	4.109187E-01(7.646295E-02)	2.857676E-01(7.086567E-01)(+)

where P and PF denote the generated Pareto-optimal solutions and the solutions sampled from the PF, $\|\cdot\|_2$ is to measure the Euclidean distance, and $|\cdot|$ is a cardinality operator. It is obvious that the calculation of GD and MPFE requires a set of solutions sampled from the PF of the corresponding problem to be solved. Because only DTLZ1-DTLZ4 have the analytical form of their PFs, DTLZ5-DTLZ7 are excluded from this investigation [11].

The performance measured by GD and MPFE are shown in Table II where the numbers in bold face refer to the best results, and the symbol “+” denotes that the null hypothesis of the results regarding NDWA-GA is accepted by the Mann-Whitney-Wilcoxon rank-sum test at the significance level of 5%. It is clearly observed from Table II that NDWA-GA outperforms DWA-ES on all the test problems in the experiments.

To intuitively investigate the effectiveness of NDWA-GA, the true PFs, and the resulted approximated PFs of NDWA-GA and DWA-ES of the 3-objective DTLZ1-DTLZ4 are plotted in Figs. 5-8, where Figs. 5a, 6a, 7a and 8a denote the true PFs, Figs. 5b, 6b, 7b and 8b denote the approximate Pareto-optimal fronts generated by NDWA-GA, and Figs. 5c, 6c, 7c and 8c denote the approximate Pareto-optimal fronts generated by DWA-ES. Specifically, the ones, which are from the solutions generated by NDWA-GA and DWA-ES but not the true Pareto-optimal, are plotted in red, while the true Pareto-optimal solutions are in blue. For a fair comparison, we set the number of function evaluations to be 2×10^4 , and the same numbers of offspring at each generation throughout the evolutionary process. In addition, the CMA-ES [48], which is a promising implementation of ES and can achieve promising performance without parameter tuning, is adopted in DWA-ES for maximizing its performance.

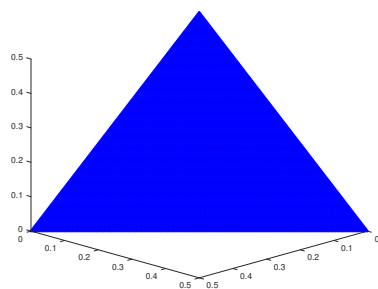
As can be seen clearly from these figures, all the solutions generated by NDWA-GA are located exactly on the PFs, while DWA-ES obtains promising solutions only on DTLZ4. In

addition, NDWA-GA also finds more Pareto-optimal solutions than DWA-ES does.

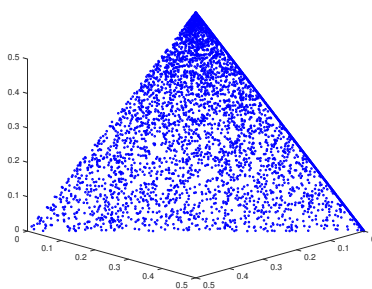
V. CONCLUSIONS

The aim of this paper is to propose an independent two-stage MaOEAs, i.e., MaOEA-IT, where the convergence and diversity are addressed in independent stages for effectively solving MaOPs. To achieve this, we firstly propose a non-dominated dynamic weight aggregation method based on genetic algorithm, i.e., NDWA-GA, which is capable of finding a population of Pareto-optimal solutions for MaOPs with concave, linear and convex Pareto front shapes in a single run. Further, these found Pareto-optimal solutions are used to learn the Pareto-optimal subspace to promote the convergence. Then, the diversity is handled by solving a set of single-objective optimization problems within the learned Pareto-optimal subspace. The solution of each single-objective optimization problem is with the closest distance to the corresponding reference line. There are totally N reference lines uniformly distributed in the objective space, where N denotes the desire number of final solutions that the end-users prefer. Through the quantitative comparisons on seven test problems with five different scales of objective numbers against six state-of-the-art MaOEAs, MaOEA-IT shows significantly better performance in addressing MaOPs. Moreover, the performance of NDWA-GA is also investigated against its seminal work, and NDWA-GA wins all of the comparisons. In addition, MaOEA-IT can also be used to effectively search for the nadir point, which had also been quantitatively and qualitatively proven by a series of experiments. Furthermore, MaOEA-IT is able to directly find the solutions within the preference area, while existing MaOEAs still need to find a population of solutions even when the preference area from end-users is known in advance.

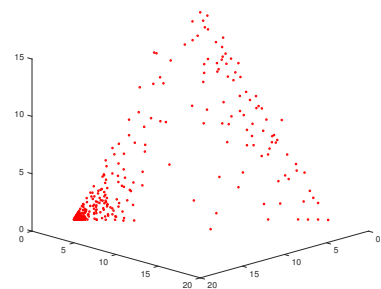
Based on the analysis on the computational complexity of MaOEA-IT, it is realized that non-dominated sorting on a large



(a) Pareto front

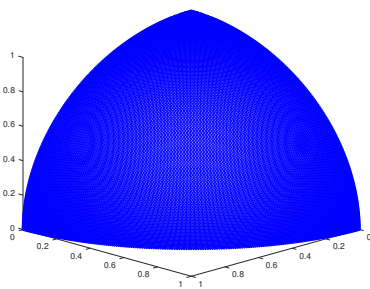


(b) NDWA-GA

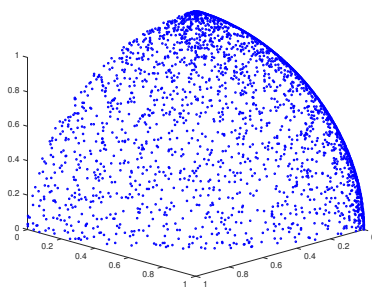


(c) DWA-ES

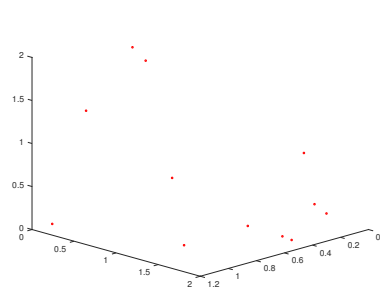
Fig. 5. The true Pareto front, the Pareto-optimal solutions generated by NDWA-GA and the Pareto-optimal solutions generated by DWA-ES of the 3-objective DTLZ1.



(a) Pareto front

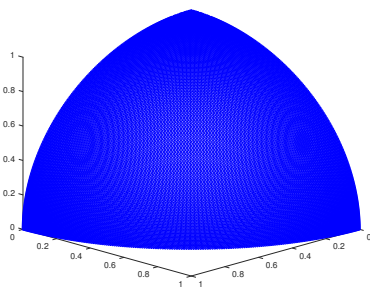


(b) NDWA-GA

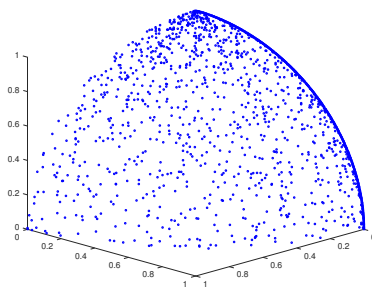


(c) DWA-ES

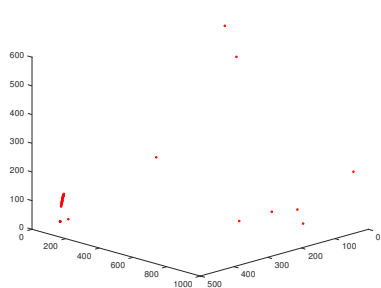
Fig. 6. The true Pareto front, the Pareto-optimal solutions generated by NDWA-GA and the Pareto-optimal solutions generated by DWA-ES of the 3-objective DTLZ2. Fig. 6c can be seen clearly on the computer screen.



(a) Pareto front



(b) NDWA-GA



(c) DWA-ES

Fig. 7. The true Pareto front, the Pareto-optimal solutions generated by NDWA-GA and the Pareto-optimal solutions generated by DWA-ES of the 3-objective DTLZ3. Fig. 7c can be seen clearly on the computer screen.

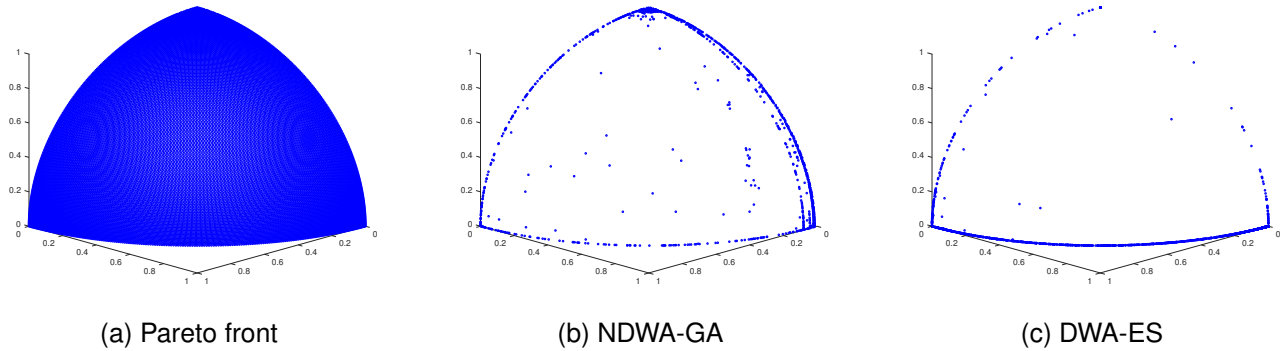


Fig. 8. The true Pareto front, the Pareto-optimal solutions generated by NDWA-GA and the Pareto-optimal solutions generated by DWA-ES of the 3-objective DTLZ4.

number of population size consumes extensive computational resources. Therefore, efficient non-dominated algorithm could be developed in our future work. In addition, MaOEA-IT can be used to find the nadir points, however, it works under the assumption that the extreme points stand at the axes due to its extreme-point-to-nadir scheme. As can be seen from the DTLZ5 and DTLZ6 test problems, a number of their extreme points do not locate at the axes. In this regard, generalizing MaOEA-IT for estimating the nadir points could also be part of our future work. Moreover, MaOEA-IT will be extended to solve constrained MaOPs that frequently exist in real-world applications.

REFERENCES

- [1] V. Khare, X. Yao, and K. Deb, "Performance scaling of multi-objective evolutionary algorithms," in *Proceedings of International Conference on Evolutionary Multi-Criterion Optimization*. Faro, Portugal: Springer, 2003, pp. 376–390.
- [2] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: NSGA-II," *IEEE Transactions on Evolutionary Computation*, vol. 6, no. 2, pp. 182–197, 2002.
- [3] E. Zitzler, M. Laumanns, L. Thiele, E. Zitzler, E. Zitzler, L. Thiele, and L. Thiele, "SPEA2: Improving the strength pareto evolutionary algorithm," 2001.
- [4] B. Li, J. Li, K. Tang, and X. Yao, "Many-objective evolutionary algorithms: a survey," *ACM Computing Surveys*, vol. 48, no. 1, pp. 13–47, 2015.
- [5] C. M. Fonseca and P. J. Fleming, "Multiobjective optimization and multiple constraint handling with evolutionary algorithms. I: a unified formulation," *IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans*, vol. 28, no. 1, pp. 26–37, 1998.
- [6] Y. Liu, D. Gong, J. Sun, and Y. Jin, "A many-objective evolutionary algorithm using a one-by-one selection strategy," *IEEE transactions on cybernetics*, vol. 47, no. 9, pp. 2689–2702, 2017.
- [7] T. Wagner, N. Beume, and B. Naujoks, "Pareto-, aggregation-, and indicator-based methods in many-objective optimization," in *Proceedings of International Conference on Evolutionary Multi-Criterion Optimization*. Matsushima, Japan: Springer, 2007, pp. 742–756.
- [8] 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 Transactions on Evolutionary Computation*, vol. 18, no. 4, pp. 577–601, 2014.
- [9] Y. Yuan, H. Xu, B. Wang, and X. Yao, "A new dominance relation-based evolutionary algorithm for many-objective optimization," *IEEE Transactions on Evolutionary Computation*, vol. 20, no. 1, pp. 16–37, 2016.
- [10] Y. Sun, G. G. Yen, and Z. Yi, "Reference line-based estimation of distribution algorithm for many-objective optimization," *Knowledge-Based Systems*, vol. 132, pp. 129–143, 2017.
- [11] —, "Improved regularity model-based EDA for many-objective optimization," *IEEE Transactions on Evolutionary Computation*, 2018, DOI:10.1109/TEVC.2018.2794319.
- [12] R. Cheng, Y. Jin, M. Olhofer, and B. Sendhoff, "A reference vector guided evolutionary algorithm for many-objective optimization," *IEEE Transactions on Evolutionary Computation*, vol. 20, no. 5, pp. 773–791, 2016.
- [13] Q. Zhang and H. Li, "MOEA/D: a multiobjective evolutionary algorithm based on decomposition," *IEEE Transactions on Evolutionary Computation*, vol. 11, no. 6, pp. 712–731, 2007.
- [14] Y. Yuan, H. Xu, B. Wang, B. Zhang, and X. Yao, "Balancing convergence and diversity in decomposition-based many-objective optimizers," *IEEE Transactions on Evolutionary Computation*, vol. 20, no. 2, pp. 180–198, 2016.
- [15] Z. Wang, Q. Zhang, M. Gong, and A. Zhou, "A replacement strategy for balancing convergence and diversity in MOEA/D," in *Proceedings of IEEE Congress on Evolutionary Computation*. Xi'an, China: IEEE, 2014, pp. 2132–2139.
- [16] K. Li, Q. Zhang, S. Kwong, M. Li, and R. Wang, "Stable matching-based selection in evolutionary multiobjective optimization," *IEEE Transactions on Evolutionary Computation*, vol. 18, no. 6, pp. 909–923, 2014.
- [17] M. Laumanns, L. Thiele, K. Deb, and E. Zitzler, "Combining convergence and diversity in evolutionary multiobjective optimization," *Evolutionary computation*, vol. 10, no. 3, pp. 263–282, 2002.
- [18] F. di Pietro, S.-T. Khu, and D. A. Savic, "An investigation on preference order ranking scheme for multiobjective evolutionary optimization," *IEEE Transactions on Evolutionary Computation*, vol. 11, no. 1, pp. 17–45, 2007.
- [19] Z. He, G. G. Yen, and J. Zhang, "Fuzzy-based Pareto optimality for many-objective evolutionary algorithms," *IEEE Transactions on Evolutionary Computation*, vol. 18, no. 2, pp. 269–285, 2014.
- [20] S. Yang, M. Li, X. Liu, and J. Zheng, "A grid-based evolutionary algorithm for many-objective optimization," *IEEE Transactions on Evolutionary Computation*, vol. 17, no. 5, pp. 721–736, 2013.
- [21] X. Zhang, Y. Tian, and Y. Jin, "A knee point-driven evolutionary algorithm for many-objective optimization," *IEEE Transactions on Evolutionary Computation*, vol. 19, no. 6, pp. 761–776, 2015.
- [22] R. Storn and K. Price, *Differential evolution—a simple and efficient adaptive scheme for global optimization over continuous spaces*. ICSI Berkeley, 1995, vol. 3.
- [23] H. Ishibuchi, H. Masuda, Y. Tanigaki, and Y. Nojima, "Modified distance calculation in generational distance and inverted generational distance," in *Proceedings of International Conference on Evolutionary Multi-Criterion Optimization*. Guimaraes, Portugal: Springer, 2015, pp. 110–125.
- [24] E. M. Lopez and C. A. C. Coello, "IGD+EMOA: a multi-objective evolutionary algorithm based on IGD+," in *Proceedings of IEE Congress on Evolutionary Computation*. Vancouver, Canada: IEEE, 2016, pp. 999–1006.
- [25] Y. Sun, G. G. Yen, and Z. Yi, "IGD indicator-based evolutionary algorithm for many-objective optimization problems," *IEEE Transactions on Evolutionary Computation*, 2018, DOI:10.1109/TEVC.2018.2791283.
- [26] S. Jiang and S. Yang, "Convergence versus diversity in multiobjective optimization," in *Proceedings of International Conference on Parallel Problem Solving from Nature*, Edinburgh, Scotland, 2016, pp. 984–993.

- [27] E. Zitzler and S. Künzli, "Indicator-based selection in multiobjective search," in *Proceedings of International Conference on Parallel Problem Solving from Nature*, Birmingham, UK, 2004, pp. 832–842.
- [28] K. Deb, *Multi-Objective Optimization Using Evolutionary Algorithms*. John Wiley & Sons, 2001, vol. 16.
- [29] Z. He and G. G. Yen, "Many-objective evolutionary algorithm: objective space reduction and diversity improvement," *IEEE Transactions on Evolutionary Computation*, vol. 20, no. 1, pp. 145–160, 2016.
- [30] W. Hu, G. G. Yen, and G. Luo, "Many-objective particle swarm optimization using two-stage strategy and parallel cell coordinate system," *IEEE Transactions on cybernetics*, vol. 47, no. 6, pp. 1446–1459, 2017.
- [31] C. Zhu, X. Cai, Z. Fan, and M. Sulaman, "A two-phase many-objective evolutionary algorithm with penalty based adjustment for reference lines," in *Proceedings of IEEE Congress on Evolutionary Computation*. Vancouver, Canada: IEEE, 2016, pp. 2161–2168.
- [32] H. Wang, S. He, and X. Yao, "Nadir point estimation for many-objective optimization problems based on emphasized critical regions," *Soft Computing*, pp. 1–13, 2015.
- [33] K. Deb, L. Thiele, M. Laumanns, and E. Zitzler, *Scalable Test Problems for Evolutionary Multiobjective Optimization*. London, UK: Springer, 2005.
- [34] P. Hajela and C.-Y. Lin, "Genetic search strategies in multicriterion optimal design," *Structural Optimization*, vol. 4, no. 2, pp. 99–107, 1992.
- [35] Y. Jin, M. Olhofer, and B. Sendhoff, "Dynamic weighted aggregation for evolutionary multi-objective optimization: why does it work and how?" in *Proceedings of the 3rd Annual Conference on Genetic and Evolutionary Computation*. San Francisco, USA: Morgan Kaufmann Publishers Inc., 2001, pp. 1042–1049.
- [36] I. Das and J. E. Dennis, "A closer look at drawbacks of minimizing weighted sums of objectives for Pareto set generation in multicriteria optimization problems," *Structural Optimization*, vol. 14, no. 1, pp. 63–69, 1997.
- [37] A. Messac, C. Puemi-Sukam, and E. Melachrinoudis, "Aggregate objective functions and Pareto frontiers: required relationships and practical implications," *Optimization and Engineering*, vol. 1, no. 2, pp. 171–188, 2000.
- [38] V. J. Bowman, "On the relationship of the Tchebycheff norm and the efficient frontier of multiple-criteria objectives," in *Multiple criteria decision making*. Springer, 1976, pp. 76–86.
- [39] Y. Jin, T. Okabe, and B. Sendhoff, "Adapting weighted aggregation for multiobjective evolution strategies," in *Proceedings of International Conference on Evolutionary Multi-Criterion Optimization*. Zurich, Switzerland: Springer, 2001, pp. 96–110.
- [40] Y. Jin, "Effectiveness of weighted sum of the objectives for evolutionary multi-objective optimization: Methods, analysis and applications," unpublished.
- [41] I. Das and J. E. Dennis, "Normal-boundary intersection: a new method for generating the Pareto surface in nonlinear multicriteria optimization problems," *SIAM Journal on Optimization*, vol. 8, no. 3, pp. 631–657, 1998.
- [42] K. Deb and R. B. Agrawal, "Simulated binary crossover for continuous search space," *Complex Systems*, vol. 9, no. 3, pp. 1–15, 1994.
- [43] E. Zitzler and L. Thiele, "Multiobjective evolutionary algorithms: a comparative case study and the strength Pareto approach," *IEEE Transactions on Evolutionary Computation*, vol. 3, no. 4, pp. 257–271, 1999.
- [44] J. Bader and E. Zitzler, "Hype: an algorithm for fast hypervolume-based many-objective optimization," *Evolutionary Computation*, vol. 19, no. 1, pp. 45–76, 2011.
- [45] L. While, L. Bradstreet, and L. Barone, "A fast way of calculating exact hypervolumes," *IEEE Transactions on Evolutionary Computation*, vol. 16, no. 1, pp. 86–95, 2012.
- [46] R. G. Steel, D. J. H. Dickey *et al.*, *Principles and Procedures of Statistics a Biometrical Approach*. WCB/McGraw-Hill, 1997, no. 519.5 S813 1997.
- [47] D. A. V. Veldhuizen and D. A. V. Veldhuizen, "Multiobjective evolutionary algorithms: Classifications, analyses, and new innovations," *Evolutionary Computation*, Tech. Rep., 1999.
- [48] N. Hansen, "The CMA evolution strategy: a comparing review," in *Towards a new evolutionary computation*. Springer, 2006, pp. 75–102.



Yanan Sun (S'15-M'18) received a Ph.D. degree in engineering from the Sichuan University, Chengdu, China, in 2017. He is currently a Research Fellow in the School of Engineering and Computer Science, Victoria University of Wellington, Wellington, New Zealand. His research topics are evolutionary algorithms, deep learning, and evolutionary deep learning. He is the leading organizer of the First Workshop on Evolutionary Deep Learning, the leading organizer of the Special Session on Evolutionary Deep Learning and Applications in CEC19, and the founding chair of the IEEE CIS Task Force on Evolutionary Deep Learning and Applications.



Bing Xue (M'10) received the B.Sc. degree from the Henan University of Economics and Law, Zhengzhou, China, in 2007, the M.Sc. degree in management from Shenzhen University, Shenzhen, China, in 2010, and the PhD degree in computer science in 2014 at Victoria University of Wellington, New Zealand. She is currently an Associate Professor in School of Engineering and Computer Science at Victoria University of Wellington. Her research focuses mainly on evolutionary computation, feature selection, feature construction, multi-objective optimization, image analysis, transfer learning, data mining, and machine learning. She has over 100 papers published in fully refereed international journals and conferences and most of them are on evolutionary feature selection and construction.

Dr Xue is currently the Chair of the IEEE Task Force on Evolutionary Feature Selection and Construction, IEEE Computational Intelligence Society (CIS), Vice-Chair of the IEEE CIS Data Mining and Big Data Analytics Technical Committee, and Vice-Chair of IEEE CIS Task Force on Transfer Learning & Transfer Optimization. She is also an Associate Editor/member of Editorial Board for five international journals and a reviewer of over 50 international journals. Dr Xue is the Finance Chair of IEEE Congress on Evolutionary Computation (CEC) 2019.



Mengjie Zhang (M'04-SM'10) received the B.E. and M.E. degrees from Artificial Intelligence Research Center, Agricultural University of Hebei, Hebei, China, and the Ph.D. degree in computer science from RMIT University, Melbourne, VIC, Australia, in 1989, 1992, and 2000, respectively. He is currently Professor of Computer Science, Head of the Evolutionary Computation Research Group, and the Associate Dean (Research and Innovation) in the Faculty of Engineering. His current research interests include evolutionary computation, particularly genetic programming, particle swarm optimization, and learning classifier systems with application areas of image analysis, multi-objective optimization, feature selection and reduction, job shop scheduling, and transfer learning. He has published over 350 research papers in refereed international journals and conferences. Prof. Zhang is a Fellow of Royal Society of New Zealand and have been a Panel member of the Marsden Fund (New Zealand Government Funding). He is also a senior member of IEEE and a member of ACM. He is currently chairing the IEEE CIS Intelligent Systems and Applications Technical Committee, and the immediate Past Chair for the IEEE CIS Emergent Technologies Technical Committee and the Evolutionary Computation Technical Committee, and a member of the IEEE CIS Award Committee. He is a vice-chair of the IEEE CIS Task Force on Evolutionary Feature Selection and Construction, a vice-chair of the Task Force on Evolutionary Computer Vision and Image Processing, and the founding chair of the IEEE Computational Intelligence Chapter in New Zealand. He is also a committee member of the IEEE NZ Central Section.



Gary G. Yen (S'87-M'88-SM'97-F'09) received a Ph.D. degree in electrical and computer engineering from the University of Notre Dame in 1992. Currently he is a Regents Professor in the School of Electrical and Computer Engineering, Oklahoma State University (OSU). Before joined OSU in 1997, he was with the Structure Control Division, U.S. Air Force Research Laboratory in Albuquerque. His research interest includes intelligent control, computational intelligence, conditional health monitoring, signal processing and their industrial/defense applications.

cations.

Dr. Yen was an associate editor of the *IEEE Control Systems Magazine*, *IEEE Transactions on Control Systems Technology*, *Automatica*, *Mechantronics*, *IEEE Transactions on Systems, Man and Cybernetics, Parts A and B* and *IEEE Transactions on Neural Networks*. He is currently serving as an associate editor for the *IEEE Transactions on Evolutionary Computation* and the *IEEE Transactions on Cybernetics*. He served as the General Chair for the 2003 *IEEE International Symposium on Intelligent Control* held in Houston, TX and 2006 *IEEE World Congress on Computational Intelligence* held in Vancouver, Canada. Dr. Yen served as Vice President for the Technical Activities in 2005-2006 and then President in 2010-2011 of the IEEE Computational intelligence Society. He was the founding editor-in-chief of the *IEEE Computational Intelligence Magazine*, 2006-2009. In 2011, he received Andrew P Sage Best Transactions Paper award from *IEEE Systems, Man and Cybernetics Society* and in 2014, he received Meritorious Service award from *IEEE Computational Intelligence Society*.