Partially Degenerate Multi-Objective Test Problems

Lie Meng Pang, Yang Nan, and Hisao Ishibuchi*

Guangdong Provincial Key Laboratory of Brain-inspired Intelligent Computation
Department of Computer Science and Engineering
Southern University of Science and Technology, Shenzhen 518055, China
{panglm@, nany@mail., hisao@}sustech.edu.cn

Abstract. Degenerate multi-objective test problems are included in test suites to evaluate EMO algorithms on a wide variety of test problems. However, it was pointed out in some studies that the frequently-used degenerate DTLZ5, DTLZ6 and WFG3 test problems do not have degenerate Pareto fronts. Their Pareto fronts are different from the originally intended degenerate shapes. Actually, they are partially degenerate test problems. Modified formulations of DTLZ5 and DTLZ6 were proposed to remove the non-degenerate parts of their Pareto fronts. However, the original formulations of DTLZ5, DTLZ6 and WFG3 continue to be used as degenerate test problems in many studies whereas they are not degenerate test problems. One issue in their use as degenerate test problems is that reference point sets for IGD calculation are sampled from the originally intended degenerate Pareto fronts whereas they are not the true Pareto fronts. Nevertheless, the original DTLZ5, DTLZ6 and WFG3 formulations are useful for performance evaluation of EMO algorithms since their Pareto front shapes are similar to some real-world problems and much more complicated than other test problems. That is, their use helps us to evaluate the performance of EMO algorithms on a wide variety of test problems including realistic and challenging test problems. In this paper, we clearly demonstrate the usefulness of the original DTLZ5, DTLZ6 and WFG3 formulations. Then, after pointing out the difficulty in their use in computational experiments, we explain how we can obtain reliable experimental results on those test problems.

Keywords: Evolutionary multi-objective optimization, test problems, degenerate Pareto fronts, partially degenerate Pareto fronts, IGD indicator.

1 Introduction

In the field of evolutionary multi-objective optimization (EMO), the performance of EMO algorithms is usually evaluated through computational experiments on benchmark test suites. Thus, it is highly desirable that a benchmark test suite consists of a wide variety of test problems with diverse characteristics including realistic test problems. In recent two decades, several benchmark test suites (e.g., ZDT [1], DTLZ [2], WFG [3], MaF [4], UF [5]) have been proposed to facilitate the growth of the EMO

* Corresponding Author: Hisao Ishibuchi

field. These test suites cover various problem characteristics. For example, these test suites include test problems with various fitness landscapes such as unimodal, multi-modal, biased, and deceptive. They also include test problems with various Pareto front shapes such as linear, convex, concave, and disconnected. In some test suites, multi-objective test problems with degenerate Pareto fronts are included to increase the diversity of test suites. For example, in the DTLZ test suite [2], DTLZ5 and DTLZ6 were designed as degenerate test problems. In the WFG test suite [3], WFG3 was designed as a degenerate test problem.

An M-objective problem is generally considered degenerate if the dimension of its Pareto front is smaller than $(M-1)$ [12], which can be a result of the existence of redundant objectives in its problem formulation [27]. Examples of degenerate Pareto fronts are illustrated in Fig. 1, in which the degenerate Pareto fronts of DTLZ5, DTLZ6 and WFG3 with three objectives are shown. These three test problems have been frequently used to demonstrate the ability of EMO algorithms to handle multi-objective problems with degenerate Pareto fronts (e.g., see [6]-[10]). If a problem contains both degenerate and non-degenerate parts of the Pareto front, it is referred to as a partially degenerate problem in this paper. In [27], it was demonstrated that partially redundant objectives can lead to a partially degenerate problem.

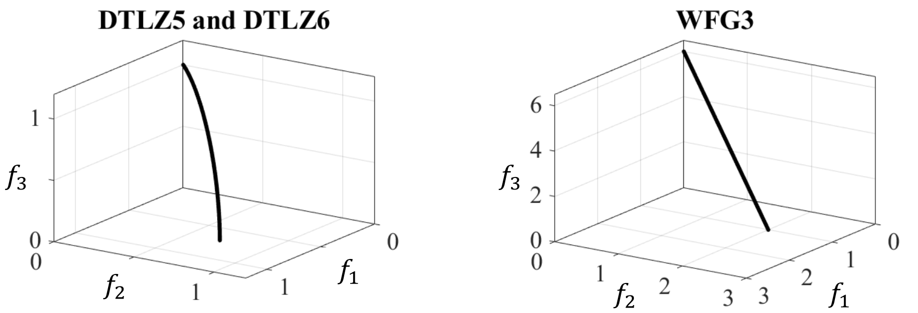

Fig. 1. The intended degenerate Pareto fronts for the three-objective DTLZ5, DTLZ6 and WFG3 test problems.

While the DTLZ5, DTLZ6 and WFG3 test problems have been frequently used to evaluate the performance of EMO algorithms on degenerate problems, it was pointed out in some studies that these three test problems are not degenerate test problems [3], [11]-[12]. Their Pareto fronts are different from the originally intended shapes. Actually, they are partially degenerate test problems [12]. The true Pareto fronts for DTLZ5 and DTLZ6 have non-degenerate parts when they have more than three objectives [3], [11]-[12]. WFG3 has a non-degenerate part of the Pareto front when it has three or more objectives [12]. In order to remove the non-degenerate parts, modified formulations of DTLZ5 and DTLZ6 were proposed in [11]. In [12], constraint conditions were derived to remove the non-degenerate part of the Pareto front of WFG3. Despite these efforts, the original formulations of DTLZ5, DTLZ6 and WFG3 are still used as degenerate test problems in many studies.

In this paper, we point out that the original formulations of DTLZ5, DTLZ6 and WFG3 with the partially degenerate Pareto fronts are good test problems. This is because their Pareto front shapes are similar to those of some real-world problems. This is also because their Pareto front shapes are much more complicated than those of the other DTLZ and WFG test problems. That is, the original formulations of DTLZ5, DTLZ6 and WFG3 are more realistic and challenging in performance evaluation of EMO algorithms than the other DTLZ and WFG test problems. One issue in their use is that the originally intended Pareto front of each test problem is often used to sample reference point sets for the inverted generational distance (IGD) [13] calculation. That is, the reference point sets and the test problems are not consistent. In other words, the original formulations of DTLZ5, DTLZ6 and WFG3 are not appropriately used for evaluating the performance of EMO algorithms. In this paper, we demonstrate the usefulness of the original DTLZ5, DTLZ6 and WFG3 test problems. We also provide suggestions on how to use them for performance evaluation of EMO algorithms.

The organization of this paper is as follows. Section 2 provides brief discussions on the Pareto fronts of DTLZ5, DTLZ6 and WFG3 with the original problem formulations. In Section 2, we also review the availability of these three test problems and their reference point sets for IGD calculation in frequently-used EMO experimental platforms: jMetal [22], PlatEMO [19] and pymoo [23]. Section 3 presents our experimental results for IGD-based performance evaluation. Section 4 concludes this paper.

2 DTLZ5, DTLZ6 and WFG3 Test Problems

2.1 Pareto Fronts of DTLZ5, DTLZ6 and WFG3

As shown in Fig. 1, the originally intended Pareto front shapes of the DTLZ5 and DTLZ6 test problems are one-dimensional curves independent of the number of objectives [2], [3], [12]. However, it was pointed out in [3], [11]-[12] that the true Pareto fronts of DTLZ5 and DTLZ6 are not degenerate when the number of objectives is larger than three. The true Pareto front shapes of DTLZ5 and DTLZ6 are unknown for the case of four or more objectives. For WFG3, the originally intended Pareto front shape is a line as shown in Fig. 1. However, the true Pareto front of WFG3 includes the line part and other solutions [12], which gives rise to a flag-like shape in the three-objective case (see Fig. 2 (a)). In Fig. 2 (b), we show an approximated Pareto front of a real-world three-objective “reactive power optimization” problem called DDMOP5 in [24]. We can see that the two Pareto fronts in Fig. 2 have similar shapes. A similar partially degenerate flag-shaped Pareto front is also shown in [20] for a real-world three-objective “two-bar truss design” problem called RE3-3-1. For the case of four or more objectives, the true Pareto front shape of WFG3 is unknown.

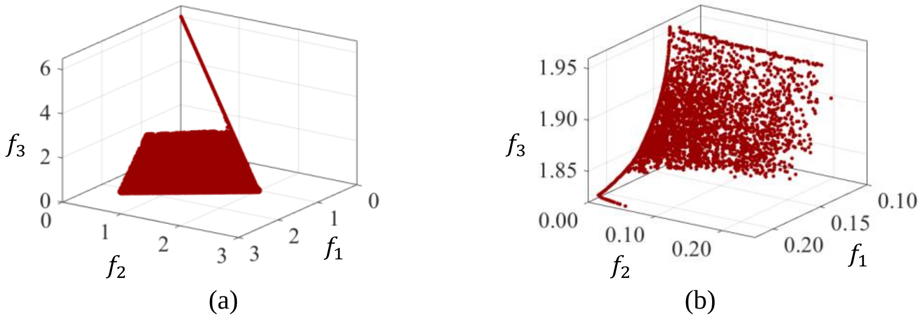

Fig. 2. The partially degenerate Pareto front of the three-objective WFG3 test problem in (a) and an approximated Pareto front of the real-world DDMOP5 problem [24] in (b).

To obtain clear pictures of the true Pareto front shapes of DTLZ5, DTLZ6 and WFG3 in a high-dimensional objective space, we use five EMO algorithms, i.e., MOEA/D with the PBI function [14], NSGA-III [15], θ-DEA [16], NSGA-II/SDR [17] and PREA [18] to approximate the Pareto front of each test problem in the five-objective space. These five algorithms are chosen based on the following considerations. MOEA/D and NSGA-III are frequently-used classic EMO algorithms. The other three are recently-proposed EMO algorithms which have shown promising performance on many-objective problems. We use PlatEMO [19] for our experiments in this paper. The population size in each algorithm is specified as 210. Each algorithm is terminated after 1,000 generations. For other specifications in each algorithm, the default settings in PlatEMO are used. Each algorithm is executed 31 times on each test problem. To approximate the true Pareto front of each test problem, we use all non-dominated solutions among obtained solutions by 31 runs of the five algorithms (i.e., $31\times 5 = 155$ runs in total).

Fig. 3 shows an approximated Pareto front for the five-objective DTLZ5. Due to the paper length limitation, approximated Pareto fronts for the other two test problems are shown in the supplementary file (which is available from https://github.com/HisaoLabSUSTC/EMO2023). The approximated Pareto front in Fig. 2 (b) was created in the same manner as in Fig. 3 whereas the population size was 91 in Fig. 2 (b). As shown in Fig. 3 (and Figs. S1-S2 in the supplementary file), the approximated Pareto fronts of DTLZ5, DTLZ6 and WFG3 are highly irregular in the high-dimensional objective space. They are clearly different from the other test problems in the DTLZ and WFG test suites. Thus, their use increases the diversity of the test problems in these test suites.

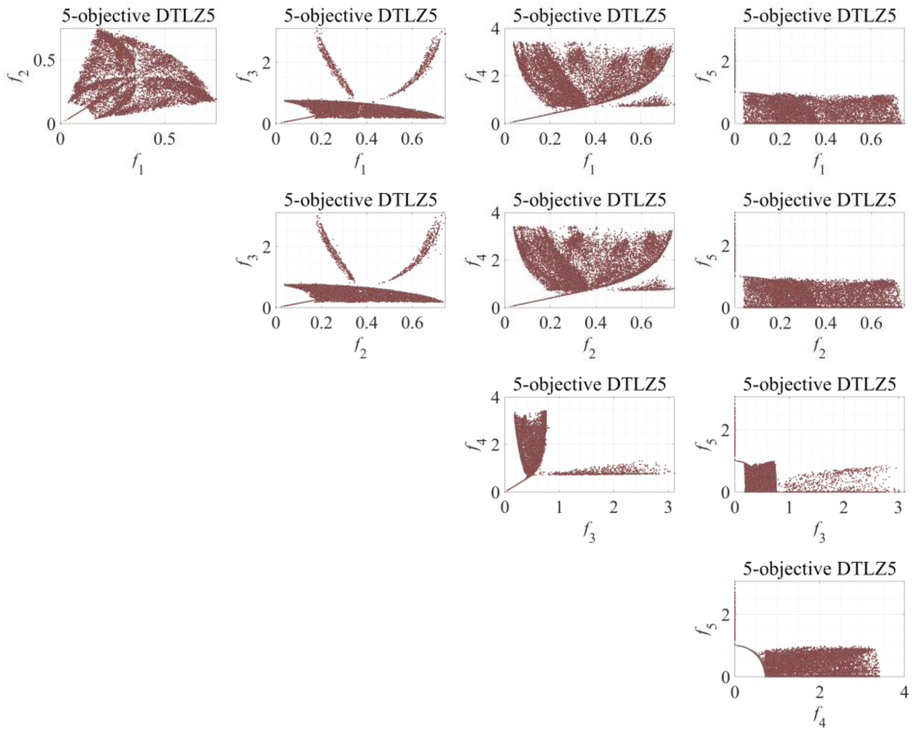

Fig. 3. An approximated Pareto front for the five-objective DTLZ5 test problem. Solutions are projected to the two-dimensional subspace.

A major challenge posed by DTLZ5, DTLZ6 and WFG3 is to find their entire Pareto fronts including the non-degenerate parts. Clearly different solution sets are often obtained by different EMO algorithms on these test problems even when almost the same results are obtained on other more standard test problems such as DTLZ1-4 and WFG4-9 with regular triangular Pareto fronts [21]. As an example, Fig. 4 shows the final population of a single run with the median IGD value among 31 runs of each algorithm on the five-objective DTLZ5 in the f_1-f_4 subspace (see Section 3 for IGD calculation). The f_1-f_4 subspace is shown here because it provides a clear demonstration of the search performance of the five EMO algorithms on the five-objective DTLZ5 problem with a partially degenerate Pareto front. The upper left subfigure shows the approximated Pareto front in the f_1-f_4 subspace, which is a copy from Fig. 3. In Fig. 4, clearly different solution sets are obtained from the five algorithms. No algorithms find a well-distributed solution set over the entire Pareto front. NSGA-III and PREA seem to find more diverse solutions on the non-degenerate part of the Pareto front than the other three algorithms. The main difference among the obtained solution sets in Fig. 4 is the diversity of solutions over the non-degenerate part (see also Fig. 7 for the ten-objective DTLZ5 and Fig. 8 for the ten-objective WFG3 in Section 3). Thus, the three partially degenerate test problems are useful for evaluating the diversification ability of EMO algorithms.

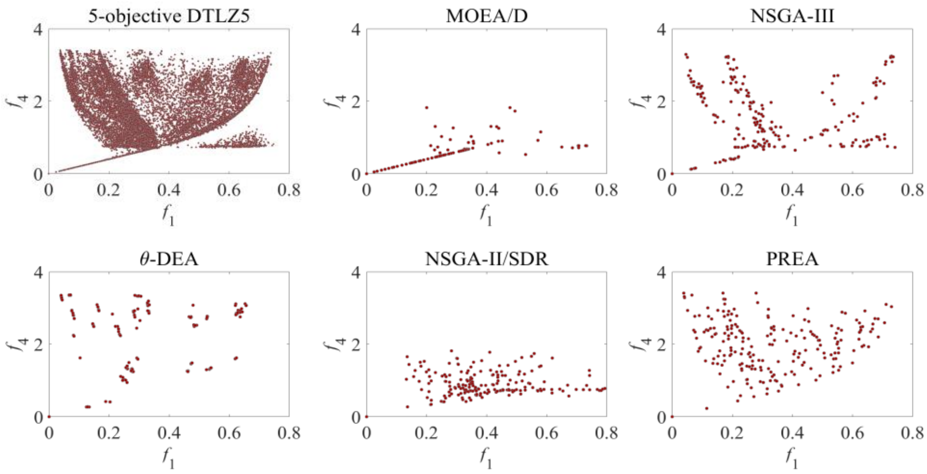

Fig. 4. A solution set obtained by a single run of each algorithm on the five-objective DTLZ5. The final population of a single run is projected to the two-dimensional subspace with f_1 and f_4.

2.2 Availability of the Test Problems

In the previous subsection, we have discussed the usefulness of the original problem formulations of DTLZ5, DTLZ6 and WFG3 for performance evaluation of EMO algorithms. Whereas the three test problems are useful, one critical issue is that the originally intended degenerate Pareto fronts have been used to sample reference point sets for IGD calculation. For DTLZ5, DTLZ6 and WFG3 with the original problem formulations, IGD-based evaluation results are unreliable and misleading if reference point sets are sampled from the originally intended degenerate Pareto fronts. Under this reference point sampling mechanism, the calculated IGD values evaluate the approximation quality of the obtained solution sets only for the degenerate parts of the partially degenerate Pareto fronts.

In many EMO experimental platforms, the original formulations of DTLZ5, DTLZ6 and WFG3 are available. It is therefore necessary to check whether the reference point set for IGD calculation is sampled from the entire partially degenerate Pareto front of each test problem. In this subsection, we review the problem formulations of the three test problems and the corresponding reference point sets for IGD calculation used in three commonly-used EMO experimental platforms: jMetal [22], PlatEMO [19] and pymoo [23].

All the jMetal, PlatEMO and pymoo platforms use the original problem formulations of DTLZ5, DTLZ6 and WFG3. Table 1 lists the reference point sets for IGD calculation for the three test problems in each platform. In jMetal and pymoo, the reference point sets for IGD calculation for DTLZ5, DTLZ6 and WFG3 are sampled from the originally intended degenerate Pareto fronts when the number of objectives (i.e., M) is three. This setting is appropriate for DTLZ5 and DTLZ6 since they have degenerate Pareto fronts when $M = 3$. However, this setting is not appropriate for WFG3 since its Pareto front is not degenerate when $M \geq 3$. For $M > 3$, the reference point sets for the three test problems are not provided in jMetal and pymoo. When reference point sets are not available in jMetal and pymoo, they can be constructed by combining the results of all runs of compared algorithms. This is a widely-used practice in the EMO field for unknown Pareto fronts (whereas this does not always lead to reliable comparison results [25]).

In PlatEMO, the provided reference point sets for IGD calculation for DTLZ5, DTLZ6 and WFG3 are sampled from the originally intended degenerate Pareto fronts regardless of the number of objectives. Thus, the reference points are not appropriate for DTLZ5 and DTLZ6 for $M > 3$ and WFG3 for $M \geq 3$. When the IGD indicator is used to evaluate the performance of EMO algorithms in PlatEMO for the three test problems, misleading results are likely to be obtained. It is therefore necessary to update the reference point sets for the three test problems in order to avoid creating unreliable IGD-based evaluation results. Moreover, it is important for users to be aware that the original problem formulations of DTLZ5, DTLZ6 and WFG3 are partially degenerate problems. When these three test problems are used for performance evaluation of EMO algorithms, we should always ensure that an appropriate reference point set for IGD calculation is used for each test problem.

Table 1. Reference point sets used for IGD calculation in jMetal, PlatEMO and pymoo.

Test problem	M	Reference point sets used for IGD calculation		
		jMetal [22]	PlatEMO [19]	pymoo [23]
DTLZ5, DTLZ6, WFG3	3	Sampled from the intended degenerate Pareto front.	Sampled from the intended degenerate Pareto front.	Sampled from the intended degenerate Pareto front.
DTLZ5, DTLZ6, WFG3	> 3	Not provided.	Sampled from the intended degenerate Pareto front.	Not provided.

3 Performance Evaluation Results

In this section, we examine the performance of the five EMO algorithms (MOEA/D [14], NSGA-III [15], θ-DEA [16], NSGA-II/SDR [17] and PREA [18]) on DTLZ5, DTLZ6 and WFG3 with the original formulations. The population size for each algorithm is specified as 91 for three-objective problems, 210 for five-objective problems, and 275 for ten-objective problems. The termination condition of each algorithm is 1,000 generations. Each algorithm is executed 31 times on each test problem.

We report two types of IGD-based performance evaluation results for DTLZ5, DTLZ6 and WFG3. One is based on the reference point set sampled from the originally intended degenerate Pareto front (i.e., the reference point set provided in PlatEMO) for each test problem, and the other is based on the reference point set consisting of all non-dominated solutions among obtained solutions by 31 runs of the five algorithms. In the latter setting, the reference point set for each test problem is an approximation of the partially degenerate (i.e., true) Pareto front. In order to examine the reliability of the constructed reference point sets using the obtained solutions by the five algorithms, two different termination conditions are used to construct the reference point sets. One is 1,000 generations (which is the same as the termination condition for performance evaluation of the five algorithms), and the other is 10,000 generations. That is, we perform IGD-based comparison of the five algorithms using the three reference point sets for each test problem. Figs. 5-6 show the three reference point sets for the ten-objective DTLZ5 and WFG3, respectively. The reference point sets provided in PlatEMO (i.e., the left figures in Figs. 5-6) are clearly different from the reference point sets obtained by the five algorithms with the two termination conditions (i.e., the center and right figures in Figs. 5-6). The difference in the reference point sets between the two termination conditions is small especially in Fig. 6 on the ten-objective WFG3. Reference point sets for the other many-objective test problems (i.e., the five-objective DTLZ5, DTLZ6 and WFG3, and the ten-objective DTLZ6) are shown in the supplementary file.

Experimental results using the three reference point sets are summarized in Tables 3-5 using the ranking of the five algorithms (“1” is the best and “5” is the worst).

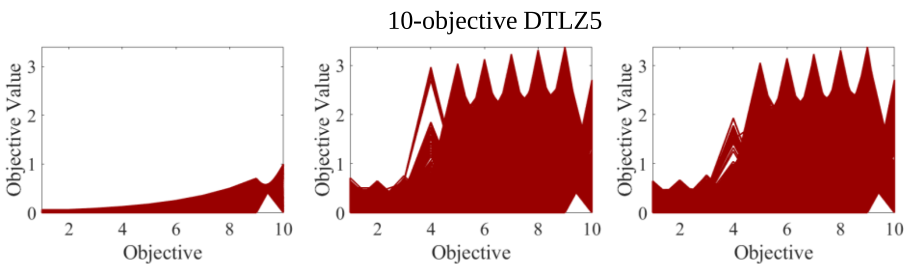

Fig. 5. Reference point sets used for IGD calculation for the ten-objective DTLZ5: (left) provided by PlatEMO, (middle) all non-dominated solutions obtained by the five algorithms with the termination condition of 1,000 generations, (right) all non-dominated solutions obtained by the five algorithms with the termination condition of 10,000 generations.

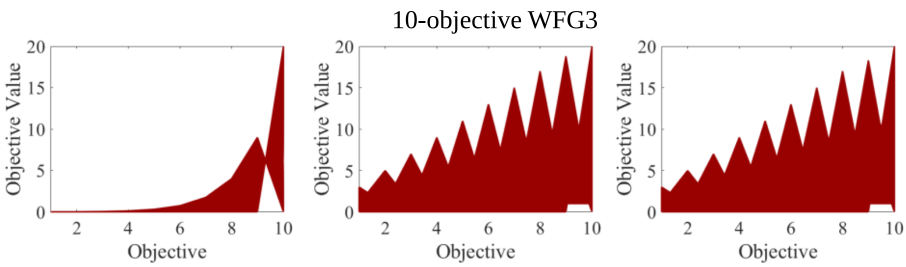

Fig. 6. Reference point sets used for IGD calculation for the ten-objective WFG3: (left) provided by PlatEMO, (middle) all non-dominated solutions obtained by the five algorithms with the termination condition of 1,000 generations, (right) all non-dominated solutions obtained by the five algorithms with the termination condition of 10,000 generations.

Table 2 is based on the reference point sets in PlatEMO. Tables 3-4 are based on the reference point sets obtained by the five algorithms (after 1,000 generations in Table 3 and 10,000 generations in Table 4). In each table, the best rank “1” is highlighted in bold. The average IGD value of each algorithm on each test problem is shown in Tables S1-S3 in the supplementary file for the three reference point sets.

In Table 2 with the PlatEMO reference point sets, NSGA-II/SDR has the best average rank over the three test problems (see the bottom line of Table 2). However, the difference in the average ranks among the five algorithms is small. A different algorithm has the best rank for a different test problem. For example, PREA has the best rank on the three-objective DTLZ5, DTLZ6 and WFG3 test problems whereas MOEA/D has the best rank on DTLZ5 and DTLZ6 with five and ten objectives.

In Tables 3-4, almost the same results are obtained. That is, Table 3 is almost the same as Table 4. For example, PREA always has the best rank for all test problems in these two tables. This is because similar reference point sets are obtained after 1,000 generations (in Table 3) and 10,000 generations (in Table 4) for each test problem as demonstrated in Figs. 5-6 (i.e., the center and right figures).

Table 2. The rank of each algorithm based on the average IGD value calculated using the reference point sets provided in PlatEMO.

Problem	M	MOEA/D	NSGA-III	θ-DEA	NSGA-II/SDR	PREA
DTLZ5	3	5	2	3	4	**1**
	5	**1**	3	5	2	4
	10	**1**	4	3	2	5
DTLZ6	3	3	2	4	5	**1**
	5	**1**	3	4	2	5
	10	**1**	5	2	3	4
WFG3	3	5	3	4	2	**1**
	5	5	3	4	**1**	2
	10	5	4	**1**	3	2
Average		3.00	3.22	3.33	**2.67**	2.78

Table 3. The rank of each algorithm based on the average IGD value calculated using the reference point set obtained by the five algorithms after 1,000 generations.

Problem	M	MOEA/D	NSGA-III	θ-DEA	NSGA-II/SDR	PREA
DTLZ5	3	4	2	3	5	**1**
	5	5	2	3	4	**1**
	10	5	2	3	4	**1**
DTLZ6	3	3	2	4	5	**1**
	5	4	2	3	5	**1**
	10	5	3	2	4	**1**
WFG3	3	2	3	5	4	**1**
	5	2	3	4	5	**1**
	10	4	2	5	3	**1**
Average		3.78	2.33	3.56	4.33	**1.00**

Table 4. The rank of each algorithm based on the average IGD value calculated using the reference point set obtained by the five algorithms after 10,000 generations.

Problem	M	MOEA/D	NSGA-III	θ-DEA	NSGA-II/SDR	PREA
DTLZ5	3	4	2	3	5	**1**
	5	5	2	3	4	**1**
	10	5	2	3	4	**1**
DTLZ6	3	3	2	4	5	**1**
	5	4	2	3	5	**1**
	10	5	3	2	4	**1**
WFG3	3	2	3	5	4	**1**
	5	2	3	4	5	**1**
	10	3	2	5	4	**1**
Average		3.67	2.33	3.56	4.44	**1.00**

One clear observation from Tables 2-4 is that totally different results are obtained between Table 2 and Tables 3-4. For the five test problem instances shaded in Tables 3-4, the best algorithm on each test problem instance in Table 2 shows the worst performance in Table 3-4. Especially, on the ten-objective DTLZ5, the totally opposite rankings of the five algorithms are obtained between Table 2 (i.e., 1, 4, 3, 2, 5) and Tables 3-4 (i.e., 5, 2, 3, 4, 1). These different results are obtained since Table 2 is based on only the degenerated parts whereas Tables 3-4 are based on the entire Pareto fronts. For example, in Fig. 4, the degenerate part of the five-objective DTLZ5 is well covered by the solution set obtained by MOEA/D. Thus, MOEA/D is evaluated as the best algorithm for the five-objective DTLZ5 in Table 2. However, the same solution set covers only a small region of the non-degenerate part in Fig. 4. Thus, MOEA/D is evaluated as the worst algorithm for the five-objective DTLZ5 in Tables 3-4.

To further examine the experimental results in Tables 2-4, the solution sets obtained by the five algorithms on the ten-objective DTLZ5 and WFG3 are shown as parallel coordinate plots in Figs. 7 and 8, respectively. For each algorithm on each test problem, a single run with the median IGD value among 31 runs is used in these figures. The reference point sets obtained after 10,000 generations in Table 4 are used for IGD calculation to choose a single run in Figs. 7 and 8 (and also in Fig. 4).

In Fig. 7, the solution set obtained by MOEA/D on the ten-objective DTLZ5 is similar to the PlatEMO reference point set in Fig. 5 (the left figure). Thus, MOEA/D is evaluated as the best algorithm on the ten-objective DTLZ5 in Table 2. However, the solution set obtained by MOEA/D is clearly different from the reference point sets obtained by the five algorithms after 1,000 and 10,000 generations in Fig. 5 (the center and right figures). Thus, MOEA/D is evaluated as the worst algorithm in Tables 3-4. Similar observations can be obtained for the solution sets of the other algorithms in Figs. 7-8 (e.g., the solution set by θ-DEA in Fig. 8 is similar to the PlatEMO reference point set in Fig. 6).

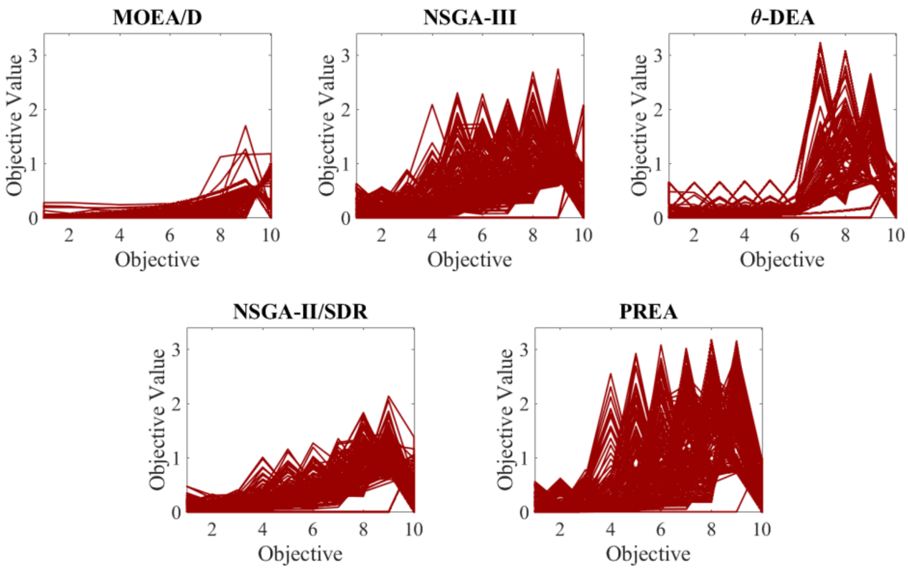

Fig. 7. The solution sets obtained by the five algorithms on the ten-objective DTLZ5 test problem. A single run with the median IGD value is selected from 31 runs of each algorithm.

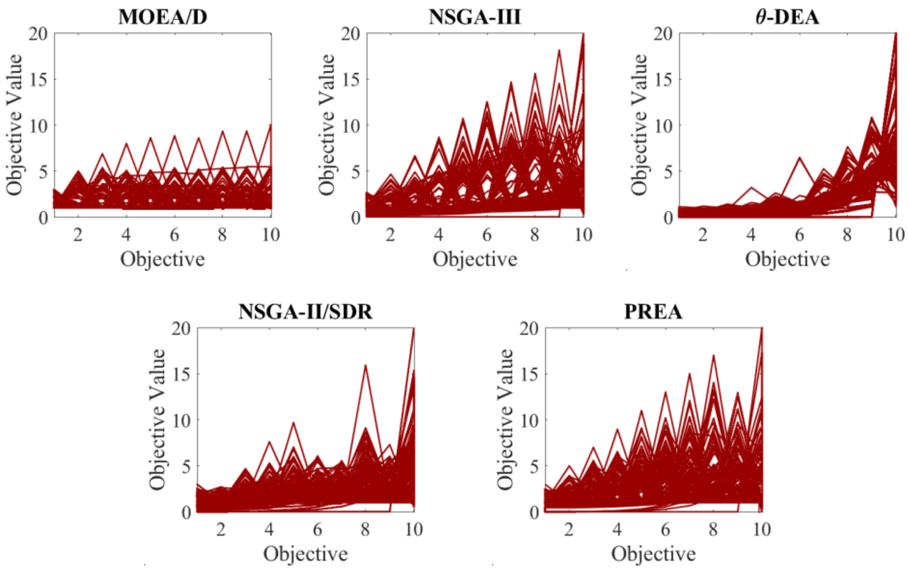

Fig. 8. The solution sets obtained by the five algorithms on the ten-objective WFG3 test problem. A single run with the median IGD value is selected from 31 runs of each algorithm.

Our experimental results in Tables 2-4 and Figs. 7-8 demonstrate that the reference point sets sampled from the originally intended degenerate Pareto front are not appropriate for DTLZ5, DTLZ6 and WFG3 with the original problem formulations (i.e., with the partially degenerate Pareto fronts). That is, IGD-based evaluation results on these test problems can be misleading when the reference point sets for IGD calculation are sampled from the originally intended degenerate Pareto front. Our suggestion is to use all non-dominated solutions among obtained solutions by all runs of all the examined algorithms as a reference point set for IGD calculation. Moreover, it is advisable to use an additional performance indicator (e.g., the hypervolume indicator) together with the IGD indicator for fair comparison of EMO algorithms. This is because performance comparison results based on a single indicator are not always reliable [26].

4 Conclusions

In this paper, we showed that the partially degenerate Pareto fronts of the DTLZ5, DTLZ6 and WFG3 test problems with the original problem formulations are highly irregular in a high-dimensional objective space, which are clearly different from the originally intended degenerate Pareto fronts. Their Pareto fronts are similar to those of some real-world problems. Hence, the original formulations of the three test problems can be used to increase the diversity of test problems in the DTLZ and WFG test suites. That is, their original formulations are good test problems to evaluate the performance of EMO algorithms. One critical issue in their use for performance evaluation of EMO algorithms is that the originally intended degenerate Pareto fronts have been used to sample reference point sets for IGD calculation. That is, these three test problems have not been used appropriately in IGD-based performance evaluation. Our computational experiments in this paper demonstrated that IGD-based evaluation results based on reference point sets from the originally intended degenerate Pareto fronts are not reliable. Thus, it is always necessary to ensure that an appropriate reference point set for each test problem is used for IGD calculation in IGD-based performance evaluation of EMO algorithms on these test problems.

Since degenerate and partially degenerate problems are common in real-world applications [27], an interesting future research study would be to investigate the possibility of quantifying or measuring degeneracy through exploratory landscape analysis [28].

Acknowledgement

This work was supported by National Natural Science Foundation of China (Grant No. 61876075), Guangdong Provincial Key Laboratory (Grant No. 2020B121201001), the Program for Guangdong Introducing Innovative and Enterpreneurial Teams (Grant No. 2017ZT07X386), The Stable Support Plan Program of Shenzhen Natural Science Fund (Grant No. 20200925174447003), Shenzhen Science and Technology Program (Grant No. KQTD2016112514355531).

References

1. Zitzler, E., Deb, K., Thiele, L.: Comparison of multiobjective evolutionary algorithms: Empirical results. Evol. Comput. 8(2), 173-195 (2000).
2. Deb, K., Thiele, L., Laumanns, M., Zitzler, E.: Scalable test problems for evolutionary multiobjective optimization. In: Evol. Multi. Optim., pp. 105-145. A. Abraham, L. Jain, and R. Goldberg, Eds. London: Springer-Verlag (2005).
3. Huband, S., Hingston, P., Barone, L., While, L.: A review of multiobjective test problems and a scalable test problem toolkit. IEEE Trans. Evol. Comput. 10(5), 477-506 (2006).
4. Cheng, R., Li, M., Tian, Y., Zhang, X., Yang, S., Jin, Y., Yao, X.: A benchmark test suite for evolutionary many-objective optimization. Complex & Intelligent Systems 3(1), 67-81 (2017).
5. Zhang, Q., Zhou, A., Zhao, S. Z., Suganthan, P. N., Liu, W., Tiwari, S.: Multiobjective optimization test instances for the CEC-2009 special session and competition. Nanyang Technol. Univ., Singapore, Tech. Rep. 1-30 (2008).
6. Liu, Y., Gong, D., Sun, J., Jin, Y.: A many-objective evolutionary algorithm using a one-by-one selection strategy. IEEE Trans. Cybn. 47(9), 2689-2702 (2017).
7. Pan, L., Li, L., He, C., Tan, K.: A subregion division-based evolutionary algorithm with effective mating selection for many-objective optimization. IEEE Trans. Cybn. 50(8), 3477-3490, 2020.
8. Gong, D., Liu, Y., Yen, G. G.: A meta-objective approach for many-objective evolutionary optimization. Evol. Comput. 28(1), 1-25 (2020).
9. Hua, Y., Jin, Y., Hao, K., Cao, Y.: Generating multiple reference vectors for a class of many-objective optimization problems with degenerate Pareto fronts. Complex & Intelligent Systems 6(1), 275-285 (2020).
10. Zhang, K., Xu, Z., Xie, S., Yen, G. G.: Evolution strategy-based many-objective evolutionary algorithm through vector equilibrium. IEEE Trans. Cybn. 51(11), 5455-5467 (2021).
11. Saxena, D. K., Duro, J. A., Tiwari, A., Deb, K., Zhang, Q.: Objective reduction in many-objective optimization: linear and nonlinear algorithms. IEEE Trans. Evol. Comput. 17(1), 77-99 (2013).
12. Ishibuchi, H., Masuda, H., Nojima, Y.: Pareto fronts of many-objective degenerate test problems. IEEE Trans. Evol. Comput. 20(5), 807-813 (2016).
13. Coello, C. A. C., Sierra, M. R.: A study of the parallelization of a coevolutionary multi-objective evolutionary algorithm. In: MICAI 2004: Advances in Artificial Intelligence, pp. 688-697. Mexico City, Mexico (2004).
14. Zhang, Q., Li, H.: MOEA/D: A multiobjective evolutionary algorithm based on decomposition. IEEE Trans. Evol. Comput. 11(6), 712-731 (2007).
15. Deb, K., Jain, H.: An evolutionary many-objective optimization algorithm using reference-point-based non-dominated sorting approach, Part I: Solving problems with box constraints. IEEE Trans. Evol. Comput. 18 (4), 577-601 (2014).
16. Yuan, Y., Xu, H., Wang, B., Yao, X.: A new dominance relation-based evolutionary algorithm for many-objective optimization. IEEE Trans. Evol. Comput. 20(1), 16-37 (2016).
17. Tian, Y., Cheng, R., Zhang, X., Su, Y., Jin, Y.: A strengthened dominance relation considering convergence and diversity for evolutionary many-objective optimization. IEEE Trans. Evol. Comput. 23(2), 331-345 (2019).
18. Yuan, J., Liu, H.-L., Gu, F., Zhang, Q., He, Z.: Investigating the properties of indicators and an evolutionary many-objective algorithm using promising regions. IEEE Trans. Evol. Comput. 25(1), 75-86 (2021).

19. Tian, Y., Cheng, R., Zhang, X., Jin, Y.: PlatEMO: A MATLAB platform for evolutionary multi-objective optimization. IEEE Computational Intelligence Magazine 12(4), 73-87 (2017).
20. Tanabe, R., Ishibuchi, H.: An easy-to-use real-world multi-objective optimization problem suite. Applied Soft Computing, Article 106078 (2020).
21. Ishibuchi, H., He, L., Shang, K.: Regular Pareto front shape is not realistic. In: Proc. IEEE Congr. Evol. Comput., pp. 2034-2041. Wellington, New Zealand (2019).
22. Nebro, A. J., Durillo, J. J., Vergne, M.: Redesigning the jMetal multi-objective optimization framework. In: Proc. Conf. Genet. Evol. Comput., pp. 1093-1100. Madrid Spain (2015).
23. Blank, J., Deb, K.: pymoo: Multi-objective optimization in Python. IEEE Access 8, 89497-89509 (2020).
24. He, C., Tian, Y., Wang, H., Jin, Y.: A repository of real-world datasets for data-driven evolutionary multiobjective optimization. Complex & Intelligent Systems 6(1), 189-197 (2020).
25. Ishibuchi, H., Imada, R., Setoguchi, Y., Nojima, Y.: Reference point specification in inverted generational distance for triangular linear Pareto front. IEEE Trans. Evol. Comput. 22(6), 961-975 (2018).
26. Ishibuchi, H., Pang, L. M., Shang, K.: Difficulties in fair performance comparison of multi-objective evolutionary algorithms. IEEE Computational Intelligence Magazine 17(1), 86-101 (2022).
27. Zhen, L., Li, M., Cheng, R., Peng, D., Yao, X.: Multiobjective test problems with degenerate Pareto fronts. Available at https://arxiv.org/abs/1806.02706
28. Kerschke, P., Trautmann, H.: The R-Package FLACCO for exploratory landscape analysis with applications to multi-objective optimization problems. In: Proc. IEEE Congr. Evol. Comput., pp. 5262-5269. Vancouver, BC, Canada (2016).