Effects of Dominance Resistant Solutions on the Performance of Evolutionary Multi-Objective and Many-Objective Algorithms

Hisao Ishibuchi
Shenzhen Key Laboratory of Computational Intelligence, University Key Laboratory of Evolving Intelligent Systems of Guangdong Province, Department of Computer Science and Engineering, Southern University of Science and Technology, Shenzhen 518055, China
hisao@sustech.edu.cn

Takashi Matsumoto, Naoki Masuyama and Yusuke Nojima
Department of Computer Science and Intelligent Systems, Graduate School of Engineering, Osaka Prefecture University, Sakai, Osaka 599-8531, Japan
takashi.matsumoto@ci.cs.osakafu-u.ac.jp
masuyama@cs.osakafu-u.ac.jp
nojima@cs.osakafu-u.ac.jp

ABSTRACT

Dominance resistant solutions (DRSs) in multi-objective problems have very good values for some objectives and very bad values for other objectives. Whereas DRSs are far away from the Pareto front, they are hardly dominated by other solutions due to some very good objective values. It is well known that the existence of DRSs severely degrades the search ability of Pareto dominance-based algorithms such as NSGA-II and SPEA2. In this paper, we examine the effect of DRSs on the search ability of NSGA-II on the DTLZ test problems with many objectives. We slightly change their problem formulation to increase the size of the DRS region. Through computational experiments, we show that DRSs have a strong negative effect on the search ability of NSGA-II whereas they have almost no effect on MOEA/D with the PBI function. We also show that a slightly modified NSGA-II for decreasing the negative effect of DRSs works well on many-objective DTLZ test problems (its performance is similar to NSGA-III and MOEA/D). These results suggest that DTLZ is not an appropriate test suite for evaluating many-objective evolutionary algorithms. This issue is further addressed through computational experiments on newly formulated test problems with no distance function.

CCS CONCEPTS

• **Mathematics of computing → Optimization algorithms**

KEYWORDS

Evolutionary multi-objective optimization, evolutionary many-objective optimization, Pareto dominance-based algorithms, decomposition-based algorithms, NSGA-II, MOEA/D, NSGA-III.

GECCO '20, July 8–12, 2020, Cancún, Mexico

ACM ISBN 978-1-4503-7128-5/20/07…$15.00
https://doi.org/10.1145/3377930.3390166

ACM Reference format:

Hisao Ishibuchi, Takashi Matsumoto, Naoki Masuyama, and Yusuke Nojima. 2020. Effects of Dominance Resistant Solutions on the Performance of Evolutionary Multi-Objective and Many-Objective Algorithms. In *Proceedings of ACM GECCO Conference, Cancun, Mexico, July 2020 (GECCO'20)*, 9 pages.
https://doi.org/10.1145/3377930.3390166

1 INTRODUCTION

An m-objective minimization problem can be written as

$$\text{Minimize } f_i(\boldsymbol{x}),\ i = 1, 2, ..., m, \tag{1}$$

$$\text{subject to } \boldsymbol{x} \in \boldsymbol{X}, \tag{2}$$

where $\boldsymbol{x} = (x_1, x_2, ..., x_n)$ is an n-dimensional decision vector, $\boldsymbol{f}(\boldsymbol{x}) = (f_1(\boldsymbol{x}), f_2(\boldsymbol{x}), ..., f_m(\boldsymbol{x}))$ is an m-dimensional objective vector, and $\boldsymbol{X}$ is the feasible region of $\boldsymbol{x}$. Some or all of the m objectives are conflicting with each other. The multi-objective problem in (1)-(2) has an n-dimensional decision space and an m-dimensional objective space. When multi-objective problems have four or more objectives, they are referred to as many-objective problems.

In the evolutionary multi-objective optimization (EMO) field, solutions are compared based on the Pareto dominance relation: A solution $\boldsymbol{x}$ is referred to as being dominated by another solution $\boldsymbol{y}$ if and only if $f_i(\boldsymbol{y}) \le f_i(\boldsymbol{x})$ for $\forall i$ and $f_i(\boldsymbol{y}) < f_i(\boldsymbol{x})$ for $\exists i$. When $\boldsymbol{x}$ is not dominated by any other feasible solutions in $\boldsymbol{X}$, $\boldsymbol{x}$ is referred to as a Pareto optimal solution. The set of all Pareto optimal solutions is the Pareto optimal solution set. The projection of the Pareto optimal solution set to the objective space is often called the Pareto front. A number of EMO algorithms have been proposed to search for a set of non-dominated solutions which approximates the Pareto front in the last three decades [2].

A current hot topic in the EMO field is the handling of many-objective problems with more than three objectives. Various approaches have been proposed to evolutionary many-objective optimization in the literature [14], [15], [20]. In early studies on many-objective optimization, the DTLZ test problems [5] were frequently used to clearly demonstrate that Pareto dominance-based algorithms such as NSGA-II [4] and SPEA2 [26] did not work well on many-objective problems. For example, Wagner et al. [21] showed that NSGA-II and SPEA2 were clearly outperformed by a hypervolume-based algorithm SMS-EMOA [1] on DTLZ1-2 with 3-6 objectives. Mostaghim and Schmeck [17] demonstrated that NSGA-II was outperformed by random search on DTLZ2 with 10-20 objectives. Currently, it is widely accepted in the EMO community that Pareto dominance-based algorithms do not work well on many-objective problems. As a result, while NSGA-II has been a widely-used algorithm, it was not included in computational experiments to evaluate recently-proposed many-objective algorithms such as NSGA-III [3], MOEA/DD [16] and θ-DEA [23]. However, it has also been reported in some studies that NSGA-II works well on some many-objective problems such as distance minimization problems [19] and knapsack problems [8] with highly correlated objectives (whereas it does not work well when the objectives are not correlated).

The difficulty of many-objective problems can be explained by the size of the dominating region of a current solution (i.e., the better region than the current solution) in a high-dimensional objective space. Let us assume that we have a current solution $\boldsymbol{x}$ in an m-dimensional objective space. The region which dominates $\boldsymbol{x}$ is only $1/2^m$ of the neighborhood of $\boldsymbol{f}(\boldsymbol{x})$ in the objective space. For example, it is less than 0.1% of the neighborhood in a 10-dimensional objective space (i.e., $m = 10$). Almost all solutions in the neighborhood are non-dominated with the current solution. Thus, it is very difficult to push the population towards the Pareto front using the Pareto dominance relation.

This explanation is consistent with experimental results on many-objective knapsack problems in [8] where NSGA-II [4] and MOEA/D [24] with the weighted sum (WS), Tchebycheff and PBI functions were compared. In Fig. 1, the contour lines of each scalarizing function are shown for a two-objective knapsack problem (maximization problem) with the weight vector $\boldsymbol{w} = (0.5, 0.5)$. The contour lines of the Tchebycheff function are consistent with the dominating region of a point on the arrow in Fig. 1 for the two-objective maximization problem. The contour lines of the PBI function ($\theta = 5$) have a sharp angle. This means that the better region than a current solution is narrow (i.e., less than $1/2^m$ of the neighborhood of the current solution for the m-objective problem if the current solution is on the arrow). By decreasing the value of θ, the size of the better region of the PBI function increases. From Fig. 1 and the above discussions about the size of the dominating region, we can expect that the best and worst results will be obtained from the weighted sum and the PBI function ($\theta = 5$), respectively. We can also expect that the performance of MOEA/D-PBI will be improved by using a small θ value.

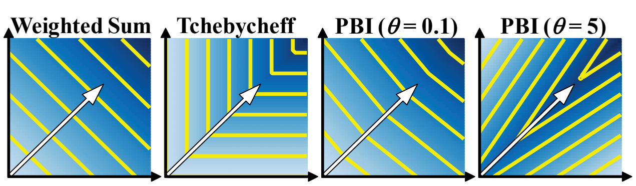


Figure 1: Contour lines of each scalarizing function for a two-objective maximization problem.

Reported results on many-objective knapsack problems in [8] are consistent with our expectation. The performance of NSGA-II and MOEA/D-Tchebycheff (both of which have the $1/2^m$ better region) was between the best results by MOEA/D-WS and the worst results by MOEA/D-PBI ($\theta = 5$). It was also shown in [8] that the performance of MOEA/D-PBI was improved by using a smaller value of θ (e.g., $\theta = 0.01$).

However, totally different results were reported in [13] on many-objective DTLZ test problems. Much better results were obtained by MOEA/D-PBI ($\theta = 5$) than NSGA-II and MOEA/D-Tchebycheff. This observation cannot be explained by the above discussions on the size of the better region in each algorithm. In this paper, we examine why NSGA-II does not work on many-objective DTLZ test problems whereas MOEA/D-PBI ($\theta = 5$) works very well. Through computational experiments, we explain the following:

(i) The reason for the poor performance of NSGA-II on many-objective DTLZ test problems is not the number of objectives but the existence of dominance resistant solutions (DRSs [7]).
(ii) Many-objective DTLZ test problems do not have any severe difficulty which is purely caused by the increase in the number of objectives. NSGA-II works well on those test problems after slight modification to decrease the negative effect of DRSs. Its performance is similar to that of NSGA-III and MOEA/D.
(iii) MOEA/D and NSGA-III can be outperformed by NSGA-II on many-objective test problems which are generated without using the frequently-used problem formulation with position and distance functions.

Our goal is to clearly show that reported experimental results on many-objective DTLZ test problems in the literature can be misleading. Our experimental results suggest that both very poor performance of Pareto dominance-based algorithms and very high performance of decomposition-based algorithms have been widely overemphasized. This paper is organized as follows. In Section 2, we explain the modification of DTLZ to widen the DRS region, the modification of NSGA-II to decrease the negative effect of DRSs, and the formulation of new many-objective test problems. In Section 3, we examine the effect of DRSs using the modified many-objective DTLZ test problems where the size of the DRSs region in the objective space can be arbitrarily specified. In Section 4, we compare the modified NSGA-II with MOEA/D and NSGA-III on the original many-objective DTLZ test problems. It is shown that the performance of NSGA-II can be significantly improved by removing the negative effect of DRSs. In Section 5, the modified NSGA-II algorithm is compared with MOEA/D and NSGA-III through computational experiments on our new many-objective test problems with no distance function. Better results are obtained from NSGA-II. This paper is concluded in Section 6.

2 MODIFICATIONS OF DTLZ AND NSGA-II

2.1 DTLZ with Adjustable DRS Regions

It is well known that the existence of DRSs severely degrades the performance of Pareto dominance-based algorithms [7], [22].

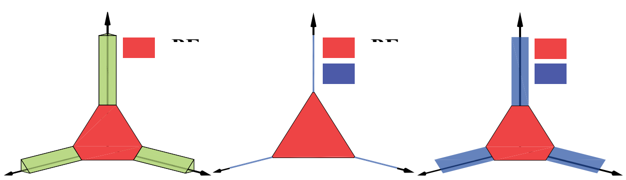
To examine the effect of DRSs on the performance of NSGA-II and MOEA/D on many-objective DTLZ test problems, we slightly modify their objectives so that we can arbitrarily specify the size of the DRS region in the objective space. Each objective of DTLZ is modified as follows:

$$\text{If } \sum_{\substack{i=1 \\ i\neq j}}^{m} f_i(\boldsymbol{x}) < \varepsilon \text{ for } \exists j \in \{1, 2, ..., m\}$$

$$\text{then } f_i(\boldsymbol{x}) = 10000 \text{ for } i = 1, 2, ..., m, \quad (3)$$

where ε is a non-negative small real number (e.g., $\varepsilon = 0.01, 0.001$).

By this modification, all solutions with very small values (i.e., very good values) of $(m-1)$ objectives are moved to the very bad point with $\boldsymbol{f}(\boldsymbol{x}) = (10000, ..., 10000)$ in the objective space. It is judged by ε whether $(m-1)$ objectives have very good values or not (i.e., whether their sum is smaller than ε or not). Fig. 2 (a) explains the modification of the objectives by (3) for the three-objective DTLZ1 problem. All solutions in each green triangular prism in Fig. 2 (a) are moved to the very bad point.



(a) Modified regions. (b) Original DRS. (c) Modified DRS.

Figure 2: Modification of the three-objective DTLZ1 problem.

As a result of the modification by (3), the objective space has the following boundary along the jth axis ($j = 1, 2, ..., m$):

$$f_1 + f_2 + ... + f_{j-1} + f_{j+1} + ... + f_m = \varepsilon, \quad (4)$$

where $\boldsymbol{f} = (f_1, f_2, ..., f_m)$ is an m-dimensional objective vector. This boundary is parallel to the jth axis since the left-hand side of (4) does not have the jth element f_j. In the case of three objectives, this boundary is a plane as shown by the three blue planes in Fig. 2 (c). All points on these three planes are hardly dominated by any other solutions (i.e., they are DRSs), which are explained in Fig. 3. Solution A in Fig. 3 is dominated by solutions on the white line (i.e., by those solutions with exactly the same f_1 and f_2 values as A and a better f_3 value than A). Any other solutions (except for solutions on the white line) cannot dominate solution A in Fig. 3.

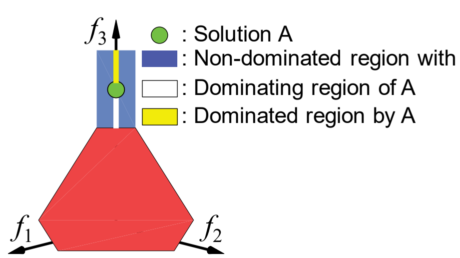


Figure 3: Pareto dominance relation on the DRS region.

When $\varepsilon = 0$, no solution is moved by (3) since all objective values are non-negative in DTLZ. By increasing the value of ε, the width of each DRS region is increased from Fig. 2 (b) to Fig. 2 (c). That is, we can arbitrarily specify the size of the DRS region in the modified DTLZ test problems. It should be noted that not only solutions on the DRS region but also very close solutions to the DRS region can survive as DRSs since they are also hardly dominated by other solutions. In Section 3, we examine the effect of DRSs on the performance of NSGA-II and MOEA/D using the modified DTLZ test problems.

2.2 Simple Modification of NSGA-II

It has been shown in some studies (e.g., Sato et al. [18]) that the convergence ability of Pareto dominance-based algorithms on many-objective problems was improved by modifying the Pareto dominance relation, i.e., by increasing the size of the dominating region of each solution. In this paper, we modify the objective functions (instead of modifying the Pareto dominance relation) to achieve the same effect to improve the convergence ability of Pareto dominance-based algorithms. Our idea is to simply modify each objective $f_i(\boldsymbol{x})$, $i = 1, 2, ..., m$, as follows:

$$u_i(\boldsymbol{x}) = (1-\alpha) f_i(\boldsymbol{x}) + \frac{\alpha}{m}\sum_{i=1}^{m} f_i(\boldsymbol{x}), \quad i = 1, 2, ..., m, \quad (5)$$

where α is a non-negative real number in $0 \le \alpha \le 1$. NSGA-II is applied to the multi-objective problem with $u_i(\boldsymbol{x})$.

In (5), $\alpha = 0$ means no modification: $u_i(\boldsymbol{x})$ is the same as $f_i(\boldsymbol{x})$. When $\alpha = 1$ in (5), each $u_i(\boldsymbol{x})$ has the same value, which is the average value of $f_i(\boldsymbol{x})$ over all objectives. The modification by (5) increases the correlation among the objectives $u_i(\boldsymbol{x})$, $i = 1, 2, ..., m$, which improves the convergence ability of Pareto dominance-based algorithms. It should be noted that the modification in (5) can be used in any EMO algorithm. All objective values of each solution are modified by (5) when the solution is evaluated. We use the objective value modification by (5) since this is much simpler than the modification of the Pareto dominance relation.

In Section 4, NSGA-II is applied to many-objective DTLZ test problems after the modification of their objectives by (5). Of course, the performance of NSGA-II (i.e., obtained solution sets by NSGA-II) is evaluated in the original objective space.

2.3 New Many-Objective Test Problems

In the DTLZ test suite [5], the first four test problems (i.e., DTLZ1-4) have been frequently used to evaluate many-objective algorithms in the literature. They have some special features. One is that all objectives have the same distance function $g(\boldsymbol{x}_D)$ as

$$f_i(\boldsymbol{x}) = (1 + g(\boldsymbol{x}_D))\, h_i(\boldsymbol{x}_P), \; i = 1, 2, ..., m, \quad (6)$$

where $h_i(\boldsymbol{x}_P)$ is the position function for the ith objective. In (6), the decision vector $\boldsymbol{x} = (x_1, x_2, ..., x_n)$ is divided into the position variable vector $\boldsymbol{x}_P = (x_1, x_2, ..., x_{m-1})$ and the distance variable vector $\boldsymbol{x}_D = (x_m, x_{m+1}, ..., x_n)$. The number of the position variables is $m-1$, and the number of the distance variables is $n-m+1$.

In DTLZ1-4, all feasible solutions with the minimum distance function value (i.e., $g(\boldsymbol{x}_\mathrm{D}) = 0$) are Pareto optimal. That is, Pareto optimal solutions can be obtained by simply optimizing $g(\boldsymbol{x}_\mathrm{D})$, and all feasible solutions with $g(\boldsymbol{x}_\mathrm{D}) = 0$ are Pareto optimal solutions. Thus, evolutionary multi-objective optimization for DTLZ1-4 can be viewed as having the following two tasks: (i) to find the optimal solution $\boldsymbol{x}^*_\mathrm{D}$ of $g(\boldsymbol{x}_\mathrm{D})$, and (ii) to maximize the diversity of solutions with $\boldsymbol{x}^*_\mathrm{D}$ in the objective space using the position variable vector $\boldsymbol{x}_\mathrm{P}$.

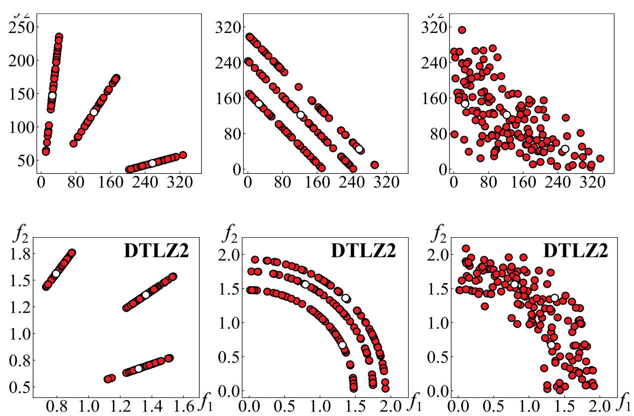
DTLZ1-4 have the following Pareto fronts [5]:

$$\text{DTLZ1: } \sum_{i=1}^{m} f_i(\boldsymbol{x}) = 0.5 \text{ and } f_i(x) \geq 0,\ i = 1, 2, ..., m, \quad (7)$$

$$\text{DTLZ2-4: } \sum_{i=1}^{m} f_i(\boldsymbol{x})^2 = 1 \text{ and } f_i(x) \geq 0,\ i = 1, 2, ..., m. \quad (8)$$

These Pareto fronts have triangular shape. This is another special (somewhat unrealistic) feature of the DTLZ1-4 test problems [9].

The special structure of DTLZ1-4 in (6) is visually illustrated in Fig. 4 for the two-objective DTLZ1 and DTLZ2 problems. In Fig. 4, 50 solutions (red circles) are generated from a randomly generated current solution (white circle). The value of a randomly selected distance variable of each current solution is randomly changed in Fig. 4 (a). A different variable is randomly selected in Fig. 4 (a) to generate a red solution. All the generated solutions from each solution are on a single line directed to the ideal point (0, 0). In Fig. 4 (b), the value of a randomly selected position variable of each current solution is randomly changed. A wide variety of solutions are generated in Fig. 4 (b) without changing the value of the distance function. In Fig. 4 (c), 50 solutions are generated by simultaneously performing these two changes.



(a) Distance variable. (b) Position variable. (c) Both variables.

Figure 4: Generated 50 solutions (red circles) from each current solution (white circle) by randomly changing the value of a single distance variable in (a), the value of a single position variable in (b), and the values of both variables in (c).

Fig. 4 (a) shows that we can push the population towards the Pareto front by adjusting only the distance variables. The diversity of the population can be improved by adjusting only the position variables as shown by Fig. 4 (b). If we have one Pareto optimal solution, we can generate all the other Pareto optimal solutions by changing the values of its position variables. The simultaneous adjustment of the two types of variables is not needed as shown in Fig. 4 (c). Fig. 4 and the formulation in (6) show that the convergence improvement for DTLZ1-4 is a single-objective optimization problem of the distance function $g(\boldsymbol{x}_\mathrm{D})$ independent of the number of objectives. The WFG4-9 test problems [6] also have a similar feature.

As shown by Fig. 4 and (6), DTLZ1-4 have an undesirable feature as many-objective test problems (i.e., the convergence of the population depends only on the distance function, which is single-objective optimization). In this paper, we propose the following simple problem formulation without this undesirable feature caused by the use of the distance function:

$$\text{Minimize} \quad f_i(\boldsymbol{x}) = \max\left\{0,\ x_i - \beta \sum_{\substack{j=1 \\ j \neq i}}^{m} x_j\right\},\ i = 1, 2, ..., m, \quad (9)$$

$$\text{subject to } 0 \leq x_i \leq 100, \quad i = 1, 2, ..., m, \quad (10)$$

where

$$f_i(\boldsymbol{x}) = 10000, \quad i = 1, 2, ..., m, \quad \text{if } \sum_{i=1}^{m} f_i(\boldsymbol{x}) < 1. \quad (11)$$

In this formulation, β is a small positive real number satisfying $\beta < 1/(m-1)$. As shown in (9), the ith objective $f_i(\boldsymbol{x})$ is basically the same as the ith decision variable x_i. The second term with β is added to introduce the correlation among the m objectives. If $\beta > 1/(m-1)$ in (9), the total effect of the other variables is larger than x_i on the ith objective $f_i(\boldsymbol{x})$. Thus, we use a small value of β satisfying $\beta < 1/(m-1)$. In our experiments in Section 5, the value of β is specified as $\beta = 0.1$ for the m-objective problem ($m \leq 10$).

Our test problem has the linear triangular Pareto front:

$$\sum_{i=1}^{m} f_i(\boldsymbol{x}) = 1 \text{ and } f_i(x) \geq 0,\ i = 1, 2, ..., m. \quad (12)$$

This is the same as the Pareto front of the normalized DTLZ1 test problem, and similar to the Pareto fronts of DTLZ2-4. It should be noted that the shape of the Pareto front of our test problem can be easily changed by changing the condition in (11). For example, we can formulate a concave Pareto front by changing $f_i(\boldsymbol{x})$ in (11) to $f_i(\boldsymbol{x})^2$, which is the same as the Pareto fronts of DTLZ2-4. It is also possible to formulate a convex Pareto front using $f_i(\boldsymbol{x})^{1/2}$.

In our test problem, the number of decision variables is the same as the number of objectives. For example, the 10-objective test problem has only 10 decision variables. No complicated nonlinear functions are used in the problem formulation in (9)-(11). The Pareto front is linear triangular as in DTLZ1. From these features, our test problem looks very easy. In Section 5, we demonstrate that our test problem is actually very difficult for MOEA/D and NSGA-III which have high search ability on many-objective DTLZ1-4 test problems.

3 EFFECTS OF INCREASED DRS REGIONS

In this section, we report experimental results by NSGA-II, NSGA-III, MOEA/D-Tchebycheff and MOEA/D-PBI ($\theta = 5$) on the modified DTLZ1-3 test problems with 3, 5, 8, 10 objectives. To examine the effect of increasing the size of the DRS region, we use four settings of ε: $\varepsilon = 0$, 0.1, 0.01, 0.001. When $\varepsilon = 0$, there is no modification on DTLZ1-3. A large value of ε means that a large (wide) DRS region is added. For example, the Pareto fronts of the m-objective DTLZ2-3 test problems are included in the unit hypercube $[0, 1]^m$. Along the axis of each objective, the region $[0, \varepsilon]$ of the other objectives is related to the DRS region. $[0, \varepsilon]$ is only 0.1% of [0, 1] when $\varepsilon = 0.001$, and it is 10% when $\varepsilon = 0.1$.

We perform computational experiments in the same manner as many other studies [3], [16], [23]. In each m-objective test problem, the number of position variables is $m - 1$, the number of distance variables is 5 (DTLZ1) and 10 (DTLZ2-3), and the range of each decision variable x_i is $0 \le x_i \le 1$ ($i = 1, 2, ..., n$). Each algorithm is executed under the following setting:

Initial population: Randomly generated solutions.
Population size: 91 ($m = 3$), 210 (5), 156 (8), 275(10).
Termination condition: Number of generations (Table 1).
Crossover: SBX (Probability = 1.0, Index = 30).
Mutation: PM (Probability = $1/n$, Index = 20).

Table 1: Termination condition (Number of generations).

	$m = 3$	$m = 5$	$m = 8$	$m = 10$
DTLZ1	400	600	750	1,000
DTLZ2	250	350	500	750
DTLZ3	1,000	1,000	1,000	1,500

Each algorithm is executed 51 times. The average value of the hypervolume indicator [27] is calculated over the 51 runs for each algorithm on each test problem. The reference point $\boldsymbol{r} = (r, ..., r)$ is specified for hypervolume calculation in the normalized objective space of the original DTLZ1-3 problems with the ideal point (0, ..., 0) and the nadir point (1, 1, 1) as $r = 13/12$ ($m = 3$), 7/6 ($m = 5$), 4/3 ($m = 8$), 4/3 ($m = 10$) from [10]. The hypervolume indicator is used since no other unary indicator is known as being Pareto compliant [25]. Whereas we also calculate the average value of the IGD indicator, we do not report IGD-based comparison results due to the page limitation of the conference paper. This is also because IGD is not Pareto compliant and the specification of reference points for IGD is not easy for many-objective problems [11].

The average hypervolume values are summarized in Tables 2-4. In each table, the best and worst results for each test problem are highlighted by red and blue. From these tables, we can see that the performance of NSGA-II is severely degraded by increasing the size of the DRS region even when the increase is very small ($\varepsilon = 0.001$) and the number of objectives is only three ($m = 3$). Moreover, even when $\varepsilon = 0$, NSGA-II does not work well on many-objective DTLZ test problems. However, NSGA-III and MOEA/D work well in Tables 2-4 even when the increase in the size of the DRS region is large ($\varepsilon = 0.1$). No clear negative effects of DRSs are observed for those algorithms in Tables 2-4.

Table 2: Average hypervolume for modified DTLZ1.

ε	m	NSGA-II	NSGA-III	PBI ($\theta = 5$)	Tchebycheff
0	3	1.028	1.058	**1.058**	**1.009**
	5	**1.761**	**2.129**	2.129	2.062
	8	**0.171**	9.980	**9.983**	9.817
	10	**0.000**	17.74	**17.76**	17.53
0.001	3	**0.046**	0.996	**1.057**	0.991
	5	**0.000**	2.117	**2.128**	2.038
	8	**0.000**	9.727	**9.982**	9.804
	10	**0.000**	17.73	**17.76**	17.51
0.01	3	**0.007**	1.014	**1.045**	0.980
	5	**0.000**	2.123	**2.125**	2.032
	8	**0.000**	9.629	**9.981**	9.798
	10	**0.000**	17.72	**17.76**	17.49
0.1	3	**0.000**	0.828	**0.864**	0.821
	5	**0.000**	2.016	**2.063**	1.891
	8	**0.000**	9.424	**9.952**	9.605
	10	**0.000**	17.50	**17.75**	17.24

Table 3: Average hypervolume for the modified DTLZ2.

ε	m	NSGA-II	NSGA-III	PBI ($\theta = 5$)	Tchebycheff
0	3	0.644	**0.684**	0.684	**0.641**
	5	**1.634**	1.859	**1.859**	1.697
	8	**0.028**	9.782	**9.825**	7.899
	10	**0.080**	17.63	**17.68**	13.59
0.001	3	**0.636**	0.684	**0.684**	0.641
	5	**1.394**	1.858	**1.859**	1.695
	8	**0.022**	9.801	**9.825**	7.819
	10	**0.028**	17.64	**17.68**	13.46
0.01	3	**0.614**	0.682	**0.684**	0.641
	5	**1.182**	1.857	**1.859**	1.692
	8	**0.005**	9.799	**9.825**	7.866
	10	**0.063**	17.63	**17.68**	13.58
0.1	3	**0.568**	0.670	**0.678**	0.633
	5	**0.568**	1.853	**1.856**	1.685
	8	**0.006**	9.785	**9.822**	7.813
	10	**0.049**	17.65	**17.68**	13.46

Table 4: Average hypervolume for the modified DTLZ3.

ε	m	NSGA-II	NSGA-III	PBI ($\theta = 5$)	Tchebycheff
0	3	0.641	0.678	**0.679**	**0.636**
	5	**1.643**	**1.856**	1.855	1.696
	8	**0.000**	8.968	**9.145**	8.084
	10	**0.000**	**17.56**	17.49	13.62
0.001	3	**0.000**	0.410	**0.678**	0.601
	5	**0.000**	1.813	**1.855**	1.530
	8	**0.000**	8.946	**9.096**	7.718
	10	**0.000**	**17.50**	16.95	13.58
0.01	3	**0.000**	0.422	**0.679**	0.5988
	5	**0.000**	1.836	**1.855**	1.536
	8	**0.000**	8.584	**8.721**	7.79
	10	**0.000**	**17.47**	16.67	13.49
0.1	3	**0.000**	0.375	**0.671**	0.5943
	5	**0.000**	1.837	**1.853**	1.527
	8	**0.000**	8.104	**9.276**	7.756
	10	**0.000**	17.35	**17.68**	13.59

To further examine the effect of the size of the DRS region, we use the sum of all objectives as a performance indicator:

$$f_{\mathrm{Sum}}(\boldsymbol{x}) = f_1(\boldsymbol{x}) + f_2(\boldsymbol{x}) + \ldots + f_m(\boldsymbol{x}). \quad (13)$$

In DTLZ1 and its variants with the increased DRS region by ε, $f_{\mathrm{Sum}}(\boldsymbol{x}) = 0.5$ always holds on the Pareto front. The range of $f_{\mathrm{Sum}}(\boldsymbol{x})$ is $0.5 \le f_{\mathrm{Sum}}(\boldsymbol{x}) \le 551.125$ [5]. Using the obtained solution set from each run in Table 2, we calculate the average value of $f_{\mathrm{Sum}}(\boldsymbol{x})$ over 51 runs. Results are summarized in Table 5. The average value of $f_{\mathrm{Sum}}(\boldsymbol{x})$ by NSGA-II is surprisingly large. For example, it is 432.1 for the 10-objective DTLZ1 with no modification (i.e., $\varepsilon = 0$). This value is much closer to its upper bound 551.125 than its lower bound 0.5 whereas all objectives $f_i(\boldsymbol{x})$ should be minimized. By slightly increasing the size of the DRS region to $\varepsilon = 0.001$, this value increases to 506.2 in Table 5. Only for the case of three objectives with $\varepsilon = 0$, all solutions are close to the Pareto front (i.e., $f_{\mathrm{Sum}}(\boldsymbol{x}) = 0.505$ in Table 5). The slight increase of ε to $\varepsilon = 0.001$ for the 3-objective DTLZ1 leads to the increase of $f_{\mathrm{Sum}}(\boldsymbol{x})$ to 54.34.

Table 5: Average objective sum for the modified DTLZ1.

ε	m	NSGA-II	NSGA-III	PBI (θ=5)	Tchebycheff
0	3	**0.505**	0.502	0.501	**0.500**
	5	**0.762**	0.500	0.500	**0.500**
	8	**170.6**	0.539	0.503	**0.500**
	10	**432.1**	0.584	0.501	**0.500**
0.001	3	**54.34**	0.865	**0.502**	7.957
	5	**217.6**	0.532	**0.500**	7.090
	8	**483.6**	2.044	**0.503**	0.520
	10	**506.2**	0.977	0.501	**0.500**
0.01	3	**74.28**	0.716	**0.502**	7.192
	5	**222.7**	0.506	**0.500**	8.217
	8	**484.2**	1.623	**0.503**	0.534
	10	**505.8**	1.564	0.501	**0.500**
0.1	3	**95.84**	0.896	**0.503**	6.760
	5	**233.1**	1.028	**0.501**	7.149
	8	**486.9**	2.532	0.503	**0.500**
	10	**508.5**	1.736	0.501	**0.500**

These observations suggest that the final solution sets obtained by NSGA-II include a large number of DRSs with some very bad objective values. This can be explained as a side-effect of the crowding distance, which is the secondary criterion of the fitness evaluation in NSGA-II. When all solutions are non-dominated with each other, the fitness evaluation of each solution depends only on its crowding distance. Roughly speaking, the crowding distance of a solution is the sum of the distance between its two adjacent solutions on each objective [4]. Thus DRSs with some very bad objective values are highly evaluated than solutions around the Pareto front by the crowding distance. In other words, the crowding distance drives the population toward DRSs with extremely bad objective values. As a result, the average value of $f_{\mathrm{Sum}}(\boldsymbol{x})$ of the obtained solutions by NSGA-II on many-objective problems in Table 5 is much closer to its upper bound 551.125 than its lower bound 0.5.

Since the other algorithms (NSGA-III and MOEA/D) have no fitness evaluation mechanisms which favor DRSs, large average values of $f_{\mathrm{Sum}}(\boldsymbol{x})$ are not obtained in Table 5. Especially, all average values of $f_{\mathrm{Sum}}(\boldsymbol{x})$ by MOEA/D-PBI (θ=5) in Table 5 are less than 0.504 (i.e., within 0.8% error from 0.5). This is because the better region by the PBI function includes a non-dominated region of the current solution when the weight vector is parallel (or close to parallel) to one axis of the objective space (e.g., (1, 0, 0), (0, 1, 0) or (0, 0, 1) in the case of three objectives). As a result, a DRS can be replaced with its non-dominated solution.

The convergence ability of NSGA-III is deteriorated by the increase in the size of the DRS region in Table 5 (e.g., by the increase of ε from $\varepsilon = 0$ to $\varepsilon = 0.001$). This is because the non-dominated sorting is the primary fitness evaluation criterion in NSGA-III (as in NSGA-II). However, in NSGA-III, since a decomposition-based mechanism is used instead of the crowding distance, many DRSs cannot survive. As a result, the effect of DRSs on NSGA-III is much smaller than that on NSGA-II (while it is larger than that on MOEA/D-PBI (θ=5)).

The effects of DRSs on MOEA/D-Tchebycheff in Table 5 look counter-intuitive. In Table 5, DRSs have large negative effects for the 3- and 5-objective test problems but no negative effects for the 8- and 10-objective test problems. Since the contour lines of the Tchebycheff function are consistent with the Pareto dominance relation, the current solution on a search direction cannot be replaced with its non-dominated solution. If a DRS is on a search direction which is (almost) parallel to one axis of the objective space, it is hardly replaced with a better solution. This is the reason for large negative effects of DRSs in Table 5 for the 3- and 5-objective test problems. However, in MOEA/D-Tchebycheff, the actual search direction for the weight vector (w_1, w_2, ..., w_m) is ($1/w_1$, $1/w_2$, ..., $1/w_m$). In the case of many-objective test problems with 8 or 10 objectives, all weight vectors have at least two zero elements (which is usually replaced with a very small value such as 10^{-6}). As a result, no search direction is close to any axis of the objective space. Thus, DRSs have no negative effects on the performance of MOEA/D-Tchebycheff for many-objective test problems with 8 or 10 objectives in Table 5. For the same reason, it is difficult to find a well-distributed solution set over the entire Pareto front by MOEA/D-Tchebycheff. As a result, its average hypervolume values are worse than those by MOEA/D-PBI (θ=5) in Tables 2-4 for many-objective problems (while its convergence ability on the 8- and 10-objective problems is the best in Table 5).

4 PERFORMANCE OF MODIFIED NSGA-II

Our experimental results in Section 3 demonstrated that DRSs severely degraded the performance of NSGA-II. As a remedy, we explained the modification of each objective by (5) in Section 2. Eq. (5) can be rewritten as follows:

$$u_i(\boldsymbol{x}) = (1-\alpha) f_i(\boldsymbol{x}) + \alpha\, f_{\mathrm{Ave}}(\boldsymbol{x}),\ \ i = 1, 2, ..., m, \quad (14)$$

where $f_{\mathrm{Ave}}(\boldsymbol{x})$ is the average value of all objectives $f_i(\boldsymbol{x})$, i = 1, 2, ..., m. Using the value of α, we can adjust the effect of $f_{\mathrm{Ave}}(\boldsymbol{x})$.

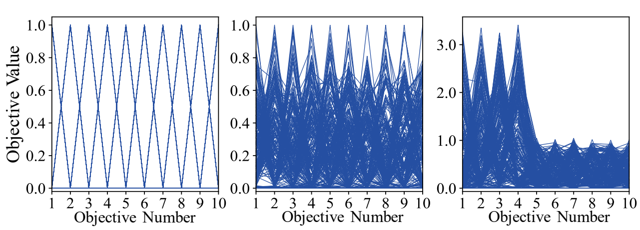
In the same manner as in Section 3, we applied NSGA-II with various values of α to the original DTLZ1-3 test problems. The average hypervolume values are summarized in Table 6. In this table, $\alpha = 0$ means no modification. In Table 6, we also show statistical test results. When the obtained results by the modified NSGA-II are significantly better (or worse) than those by the original NSGA-II with $\alpha = 0$, "+" (or "–") is assigned to the corresponding average hypervolume value. When they have no statistically significant difference, "=" is assigned. Our statistical test is performed at the 5% significance level using Shapiro-Wilk test, Bartlett's test, Wilcoxon's rank sum test, Student's t-test and Welch's t-test depending on the distributions of the hypervolume values obtained by the two algorithms to be compared.

In Table 6, the performance of NSGA-II is clearly improved on all test problems by a small modification with $\alpha = 0.01$ (which means only a 1% change of each objective value towards their average value). It seems that the best specification of α depends on the problem. However, similar results are obtained for each test problem from a wide range of α in [0.01, 0.2] in Table 6.

Table 6: Average hypervolume by the modified NSGA-II.

	m	$\alpha = 0.5$	$\alpha = 0.2$	$\alpha = 0.1$	$\alpha = 0.05$	$\alpha = 0.01$	$\alpha = 0$
DTLZ1	3	**1.035 +**	1.034 +	1.033 +	1.033 +	1.033 +	**1.028**
	5	**2.115 +**	2.112 +	2.110 +	2.106 +	2.100 +	**1.761**
	8	**9.977 +**	9.968 +	9.952 +	9.914 +	9.782 +	**0.170**
	10	**17.76 +**	17.75 +	17.74 +	17.71 +	17.40 +	**0.000**
DTLZ2	3	**0.271 –**	0.662 +	**0.665 +**	0.656 +	0.649 +	0.644
	5	**1.161 –**	**1.823 +**	1.782 +	1.750 +	1.706 +	1.634
	8	7.930 +	**9.748 +**	9.657 +	9.537 +	7.960 +	**0.028**
	10	15.56 +	**17.62 +**	17.52 +	17.36 +	14.75 +	**0.080**
DTLZ3	3	**0.268 –**	0.597 =	**0.659 +**	0.651 +	0.645 +	0.641
	5	**1.022 –**	**1.791 +**	1.788 +	1.766 +	1.729 +	1.643
	8	4.867 +	**9.716 +**	9.655 +	8.992 +	3.086 +	**0.000**
	10	7.220 +	**17.62 +**	17.52 +	17.01 +	8.700 +	**0.000**

In general, the convergence ability of NSGA-II is improved by increasing the value of α. However, the use of a too large value of α degrades the diversity of the obtained solutions. In Fig. 5, we show the obtained solution set by a single run with the median hypervolume value among 51 runs of the modified NSGA-II with $\alpha = 0.5$, 0.1, 0.01 on the 10-objective DTLZ2 test problem. It should be noted that the scale of the vertical axis in Fig. 5 (c) is more than three times larger than that in the other figures. The above-mentioned effects of α are clearly demonstrated in Fig. 5.



(a) $\alpha = 0.5$. (b) $\alpha = 0.1$. (c) $\alpha = 0.01$.

Figure 5: Results of a single run on the 10-objective DTLZ2.

In Table 7, NSGA-II with $\alpha = 0.1$ is compared with NSGA-III and MOEA/D on the DTLZ1-3 test problems. NSGA-II with $\alpha = 0.1$ is outperformed by NSGA-III and MOEA/D-PBI (θ=5) on almost all test problems. However, their average hypervolume values are very similar. As shown in Fig. 5 (b), good results are obtained by NSGA-II with $\alpha = 0.1$. From these observations, we can say that all of those algorithms work well on the DTLZ1-3 test problems with 3-10 objectives.

Table 7: Hypervolume-based comparison results.

	m	NSGA-II with $\alpha = 0.1$	NSGA-III	MOEA/D PBI (θ=5)	MOEA/D Tchebycheff
DTLZ1	3	1.033	1.058 +	**1.058 +**	**1.009 –**
	5	2.110	**2.129 +**	2.129 +	**2.062 –**
	8	9.952	9.980 +	**9.983 +**	**9.817 –**
	10	17.74	17.74 +	**17.76 +**	**17.53 –**
DTLZ2	3	0.665	**0.685 +**	0.685 +	**0.642 –**
	5	1.782	1.859 +	**1.859 +**	**1.697 –**
	8	9.657	9.782 +	**9.825 +**	**7.899 –**
	10	17.52	17.63 +	**17.68 +**	**13.59 –**
DTLZ3	3	0.660	0.679 +	**0.679 +**	**0.637 –**
	5	1.788	**1.856 +**	1.855 +	**1.696 –**
	8	**9.655**	8.968 –	9.145 –	**8.084 –**
	10	17.52	**17.56 +**	17.49 +	**13.62 –**

5 PERFORMANCE ON SIMPLE PROBLEMS

In Section 4, we demonstrated that NSGA-III, MOEA/D-PBI (θ=5) and the modified NSGA-II with $\alpha = 0.1$ work well on many-objective DTLZ1-3 test problems. However, this does not mean that they have high search ability. Using the simple test problem formulated in Section 2, we examine their performance (and the performance of some other algorithms). In our problem, the ith objective $f_i(\boldsymbol{x})$ is basically equal to the ith decision variable x_i. Thus, the m-objective problem has only m decision variables. The Pareto front is linear triangular: $f_1 + f_2 + \ldots + f_m = 1$. In the same manner as in Sections 3 and 4 (except that the termination condition is $1000m$ generations, which is about 10 times longer), we perform computational experiments. Experimental results are summarized in Table 8 with average hypervolume values and Table 9 with average sums of the objective values. Whereas the newly-formulated test problem looks very simple and we use much more computation load than the settings in Sections 3 and 4, non-zero hypervolume values are obtained only by NSGA-II with $\alpha = 0.1$ and MOEA/D with the weighted sum (WS) function in Table 8 for the 10-objective problem.

Table 8: Hypervolume-based comparison results on the newly formulated simple test problem.

m	NSGA-II $\alpha = 0.1$	NSGA-II (Original)	NSGA-III	MOEA/D PBI (θ=5)	MOEA/D Tcheby	MOEA/D WS
3	1.016	1.004 –	**1.058 +**	1.052 +	1.007 –	**0.271 –**
5	2.089	2.053 –	**2.129 +**	2.129 +	2.055 –	**1.161 –**
8	**9.862**	**0.000 –**	8.612 –	9.592 –	0.192 –	8.658 –
10	**17.70**	**0.000 –**	**0.000 –**	**0.000 –**	**0.000 –**	17.44 –

Table 9: Objective sum-based comparison results on the newly formulated simple test problem.

m	NSGA-II $\alpha = 0.1$	NSGA-II (Original)	NSGA-III	MOEA/D PBI (θ=5)	MOEA/D Tcheby	MOEA/D WS
3	1.044	**1.069 –**	1.005 +	1.016 +	1.003 +	**1.000 +**
5	1.175	**1.335 –**	1.001 +	1.003 +	1.006 +	**1.000 +**
8	1.749	**143.2 –**	4.888 –	2.153 –	48.56 –	**1.000 +**
10	1.856	**99.47 –**	25.80 –	51.57 –	55.47 –	**1.000 +**

In Table 9, very poor convergence results are obtained by NSGA-II, NSGA-III, and MOEA/D with PBI and Tchebycheff on the 10-objective test problem. The experimental results in Table 9 on the 10-objective problem are consistent with our discussions about the difficulty of many-objective problems based on the size of the better region. The best convergence results are obtained by MOEA/D-WS in Table 9. However, since a wide variety of solutions cannot be obtained by the weighted sum on the linear Pareto front, MOEA/D-WS is outperformed by the other algorithms on the test problems with 3-5 objectives in Table 8. NSGA-II with $\alpha = 0.1$ shows high performance on average over the four test problems in Table 8. This is because NSGA-II with $\alpha = 0.1$ has higher convergence ability than the other algorithms in Table 8 (except for MOEA/D-WS) thanks to the increase in the correlation among the objectives, which is a similar effect of the modification of the dominance relation. Since the increase in the correlation by $\alpha = 0.1$ is not large, the diversity deterioration is not serious in Table 8 in comparison with MOEA/D-WS.

6 CONCLUSIONS

In this paper, we tried to clearly explain that the main reason for the difficulty of many-objective DTLZ test problems is not the number of objectives but the existence of DRSs. For this purpose, we demonstrated that the performance of NSGA-II was severely degraded by increasing the size of the DRS region. We also demonstrated that the performance of NSGA-II was improved by removing the negative effect of DRSs. These observations show that the DTLZ test suite is not suitable for evaluating the performance of many-objective evolutionary algorithms. Its main drawback is the use of a single distance function for all objectives. Due to this drawback, the convergence improvement of solutions in many-objective evolutionary algorithms is single-objective optimization independent of the number of objectives. As a result, good results can be easily obtained without strong convergence ability even when the number of objectives is large. To further address this issue, we formulated a simple many-objective test problem with no distance function. Experimental results on our new test problem were totally different from those on the DTLZ test suite. Without strong convergence ability, good results were not obtained on our 10-objective test problem.

Our experimental results suggest the necessity of re-evaluating many-objective evolutionary algorithms. This is because widely-believed performance comparison results (i.e., poor performance of Pareto dominance-based algorithms and high performance of decomposition-based algorithms) were obtained from experiments on many-objective test problems with special features (i.e., DTLZ and WFG). Actually, it was reported in a recent study [12] that NSGA-II outperformed four well-known decomposition-based algorithms (MOEA/D, MOEA/DD, NSGA-III, θ-DEA) on 9 out of 18 many-objective test problems with various features (e.g., curvature and shape of Pareto fronts, shape of feasible regions) whereas NSGA-II was always outperformed on the DTLZ and WFG test suites. For fair performance comparison, it is needed to examine the suitability of many-objective test problems. This is an important future research topic since the examination of test problems in a high-dimensional objective space is very difficult. One related (and important) future research topic is the proposal of realistic many-objective test problems which have similar properties to real-world problems. Of course, the development of high-performance many-objective evolutionary algorithms for realistic test problems and real-world problems is the most exciting future research topic.

ACKNOWLEDGEMENT

This work was supported by National Natural Science Foundation of China (Grant No. 61876075), the Program for Guangdong Introducing Innovative and Enterpreneurial Teams (Grant No. 2017ZT07X386), Shenzhen Science and Technology Program (Grant No. KQTD2016112514355531), the Science and Technology Innovation Committee Foundation of Shenzhen (Grant No. ZDSYS201703031748284), the Program for University Key Laboratory of Guangdong Province (Grant No. 2017KSYS008).

REFERENCES

[1] N. Beume, B. Naujoks, and M. Emmerich. 2007. SMS-EMOA: Multiobjective selection based on dominated hypervolume. *European Journal of Operational Research* 181, 3 (2007), 1653-1669.

[2] K. Deb. 2001. *Multi-Objective Optimization Using Evolutionary Algorithms*. John Wiley & Sons, Chichester.

[3] K. Deb and H. Jain. 2014. An evolutionary many-objective optimization algorithm using reference-point-based non-dominated sorting approach, Part I: Solving problems with box constraints. *IEEE Trans. on Evolutionary Computation* 18, 4 (2014) 577-601.

[4] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan. 2002. A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Trans. on Evolutionary Computation* 6, 2 (2002) 182-197.

[5] K. Deb, L. Thiele, M. Laumanns, and E. Zitzler. 2002. Scalable multi-objective optimization test problems. *Proceedings of 2002 Congress on Evolutionary Computation (CEC 2002)*, IEEE, Honolulu, USA, 825-830.

[6] S. Huband, P. Hingston, L. Barone, and L. While. 2006. A review of multiobjective test problems and a scalable test problem toolkit. *IEEE Trans. on Evolutionary Computation* 10, 5 (2006), 477-506.

[7] K. Ikeda, H. Kita, and S. Kobayashi. 2001. Failure of Pareto-based MOEAs: Does non-dominated really mean near to optimal? *Proceedings of 2001 Congress on Evolutionary Computation (CEC 2001)*, IEEE, Seoul, Korea, 957-962.

[8] H. Ishibuchi, N. Akedo, and Y. Nojima. 2015. Behavior of multi-objective evolutionary algorithms on many-objective knapsack problems. *IEEE Trans. on Evolutionary Computation* 19, 2 (2015) 264-283.

[9] H. Ishibuchi, L. He, and K. Shang. 2019. Regular Pareto front shape is not realistic. *Proceedings of 2019 IEEE Congress on Evolutionary Computation (CEC 2019)*, IEEE, Wellington, New Zealand, 2035-2042.

[10] H. Ishibuchi, R. Imada, Y. Setoguchi, and Y. Nojima. 2018. How to specify a reference point in hypervolume calculation for fair performance comparison. *Evolutionary Computation* 26, 3 (2018) 411-440.

[11] H. Ishibuchi, R. Imada, Y. Setoguchi, and Y. Nojima. 2018. Reference point specification in inverted generational distance for triangular linear Pareto front. *IEEE Trans. on Evolutionary Computation* 22, 6 (2018) 961-975.

[12] H. Ishibuchi, T. Matsumoto, N. Masuyama, and Y. Nojima. 2020. Many-objective problems are not always difficult for Pareto dominance-based

evolutionary algorithms. *Proceedings of 24th European Conference on Artificial Intelligence (ECAI 2020)*, Santiago, Spain (Accepted).
[13] H. Ishibuchi, Y. Setoguchi, H. Masuda, and Y. Nojima. 2017. Performance of decomposition-based many-objective algorithms strongly depends on Pareto front shapes. *IEEE Trans. on Evolutionary Computation* 21, 2 (2017) 169-190.
[14] H. Ishibuchi, N. Tsukamoto, and Y. Nojima. 2008. Evolutionary many-objective optimization: A short review. *Proceedings of 2008 IEEE Congress on Evolutionary Computation (CEC 2008)*, IEEE, Hong Kong, China, 2419-2426.
[15] B. Li, J. Li, K. Tang, and X. Yao. 2015. Many-objective evolutionary algorithms: A survey. *ACM Computing Surveys* 48, 1 (2015) Article 13, 1-35.
[16] K. Li, K. Deb, Q. Zhang, and S. Kwong. 2015. An evolutionary many-objective optimization algorithm based on dominance and decomposition. *IEEE Trans. on Evolutionary Computation* 19, 5 (2015) 694-716.
[17] S. Mostaghim and H. Schmeck. 2008. Distance based ranking in many-objective particle swarm optimization. *Proceedings of PPSN X*, 753-762 (2008).
[18] H. Sato, H. E. Aguirre, and K. Tanaka. 2007. Controlling dominance area of solutions and its impact on the performance of MOEAs. *Proceedings of the 4th International Conference on Evolutionary Multi-Criterion Optimization (EMO 2007)*, Springer, 5-20.
[19] O. Schütze, A. Lara, and C. A. C. Coello. 2011. On the influence of the number of objectives on the hardness of a multiobjective optimization problem. *IEEE Trans. on Evolutionary Computation* 15, 4 (2011) 444-455.
[20] C. von Lücken, B. Barán, and C. Brizuela. 2014. A survey on multi-objective evolutionary algorithms for many-objective problems. *Computational Optimization and Applications* 58, 3 (2014) 707-756.
[21] T. Wagner, N. Beume, and B. Naujoks. 2007. Pareto-, aggregation-, and indicator-based methods in many-objective optimization. *Proceedings of the 4th International Conference on Evolutionary Multi-Criterion Optimization (EMO 2007)*, Springer, 742-756.
[22] Z. Wang, Y. S. Ong, and H. Ishibuchi. 2019. On scalable multiobjective test problems with hardly-dominated boundaries. *IEEE Trans. on Evolutionary Computation* 23, 2 (2019) 217-231.
[23] Y. Yuan, H. Xu, B. Wang, and X. Yao. 2016. A new dominance relation based evolutionary algorithm for many-objective optimization. *IEEE Trans. on Evolutionary Computation* 20, 1 (2016) 16-37.
[24] Q. Zhang and H. Li. 2007. MOEA/D: A multiobjective evolutionary algorithm based on decomposition. *IEEE Trans. on Evolutionary Computation* 11, 6 (2007) 712-731.
[25] E. Zitzler, D. Brockhoff, and L. Thiele. 2007. The hypervolume indicator revisited: On the design of Pareto-compliant indicators via weighted integration. *Proceedings of the 4th International Conference on Evolutionary Multi-Criterion Optimization (EMO 2007)*, Springer, 862-876.
[26] E. Zitzler, M. Laumanns, and L. Thiele. 2001. SPEA2: Improving the strength Pareto evolutionary algorithm. *TIK-Report 103*, Computer Engineering and Networks Laboratory (TIK), Department of Electrical Engineering, ETH, Zurich.
[27] E. Zitzler and L. Thiele. 1998. Multiobjective optimization using evolutionary algorithms - A comparative case study. *Proceedings of the 5th International Conference on Parallel Problem Solving from Nature (PPSN V)*, Amsterdam, Netherlands, 292-301.