Pareto Fronts of Many-Objective Degenerate Test Problems

Hisao Ishibuchi, *Fellow, IEEE*, Hiroyuki Masuda,
and Yusuke Nojima, *Member, IEEE*

***Abstract*—In general, an M-objective continuous optimization problem has an $(M-1)$-dimensional Pareto front in the objective space. If its dimension is smaller than $(M-1)$, it is called a degenerate Pareto front. DTLZ5 and WFG3 have often been used as many-objective continuous test problems with degenerate Pareto fronts. However, it was pointed out that DTLZ5 has a non-degenerate part of the Pareto front. Constraints have been proposed to remove the non-degenerate part. In this paper, first we show that WFG3 also has a non-degenerate part. Then we derive constraints to remove the non-degenerate part. Finally we show that the existence of the non-degenerate part makes WFG3 an interesting test problem through computational experiments.**

***Index Terms*—Evolutionary many-objective optimization, WFG test problems, DTLZ test problems, degenerate Pareto front.**

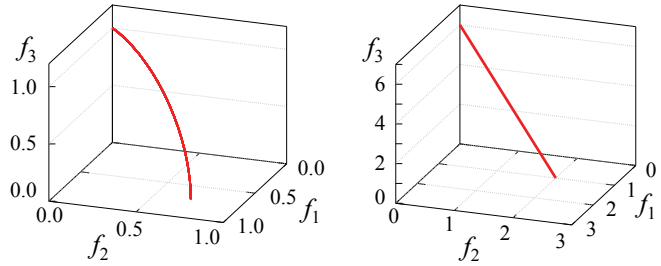
I. Introduction

Evolutionary many-objective optimization has been a hot topic in the field of evolutionary multiobjective optimization (EMO). A variety of approaches have been proposed to handle many objectives in the framework of EMO algorithms [1], [2]. Many-objective continuous test problems called DTLZ [3], [4] and WFG [5], [6] have been frequently used in the literature. The original DTLZ family in [3] had seven test problems. Two problems were added in [4]. In this paper, the nine problems in [4] are referred to as the DTLZ family. We focus our attention on many-objective continuous test problems (i.e., we do not discuss discrete problems).

In general, an M-objective continuous optimization problem has an $(M-1)$-dimensional Pareto front in the objective space. For example, the Pareto front of a bi-objective (two-objective) problem is a tradeoff curve, and a tri-objective (three-objective) problem has a tradeoff surface. However, this is not always the case. For example, if some objectives of an M-objective problem are not conflicting, the dimension of its Pareto front is smaller than $(M-1)$. The Pareto front of an M-objective problem is referred to as being degenerate when its dimension is smaller than $(M-1)$.

Among various test problems in the DTLZ [4] and WFG [6] families, it is known that DTLZ5, DTLZ6 and WFG3 were originally intended to be test problems with degenerate Pareto fronts. Their intended Pareto fronts are one-dimensional curves independent of the number of objectives (see Fig. 1). The intended Pareto fronts of DTLZ5 and DTLZ6 are the same (see Fig. 1 (a)). DTLZ5 was generalized to DTLZ5(I, M) with an $(I-1)$-dimensional intended Pareto front in [7] where $2 \le I \le M$.

While DTLZ5 and DTLZ6 were originally intended to be multi-objective test problems with degenerate Pareto fronts, it was pointed out in the literature [5], [6], [8] that their true Pareto fronts have non-degenerate parts when they have four or more objectives. The true Pareto front of DTLZ5(I, M) also has a non-degenerate part. To remove the non-degenerate parts, constraints were introduced to their original formulations in [8]. In this paper, we show that the Pareto front of WFG3 also has a non-degenerate part as well as the degenerate part (i.e., the line in Fig. 1 (b)) when it has three or more objectives. For other test problems with degenerate Pareto fronts, see [9] and [10] where test problems with high flexibility were proposed.



(a) DTLZ5 (DTLZ6). (b) WFG3.

Fig. 1. Intended degenerate Pareto fronts of DTLZ5 (DTLZ6) and WFG3. Each plot is depicted for the case of three objectives.

Recently, the performance of EMO algorithms on many-objective problems has been frequently evaluated using the generational distance (GD [11]) and the inverted generational distance (IGD [12]). This is because the quality evaluation of a solution set by the hypervolume indicator [13] is not easy when the number of objectives is large (e.g., more than 10) and/or the size of solution sets is large (e.g., more than 1000).

Continuous test problems are usually constructed in such a way that their Pareto fronts are known, which facilitates their use as benchmark problems for performance evaluation. In this case, a number of uniformly sampled reference points from the Pareto front are used for GD and IGD calculations. If reference points are sampled only from a small part of the Pareto front, misleading results are likely to be obtained. We examine this issue through computational experiments on WFG3 using NSGA-II [14], MOEA/D [15], and HypE [16].

In this paper, first we review two families of many-objective test problems called DTLZ and WFG in Section II and Section III, respectively. In Section IV, we show that WFG3 has a non-degenerate part of the Pareto front. We also derive constraints to remove the non-degenerate part. In Section V, we report experimental results on WFG3 using two settings of reference point sampling: One is the sampling only from the degenerate part and the other is from the entire Pareto front. Much smaller GD values are obtained by the reference point sampling from the entire Pareto front. When we use IGD, different performance comparison results of EMO algorithms are obtained depending on the reference point sampling. One interesting observation is that hypervolume-based comparison results are consistent with IGD-based results when reference points are sampled only from the degenerate part (not from the entire Pareto front). This observation is discussed in detail in Section V. Finally, we conclude this paper in Section VI.

Manuscript received ???, ??, 2015; revised ??? ????, 2015; accepted ????, 2015. Date of publication ????, ????.

The authors are with the Department of Computer Science and Intelligent Systems, Osaka Prefecture University, Osaka 599-8531, Japan (e-mail: hisaoi@cs.osakafu-u.ac.jp; hiroyuki.masuda@ci.cs.osakafu-u.ac.jp; nojima@cs.osakafu-u.ac.jp).

II. DTLZ Test Problems

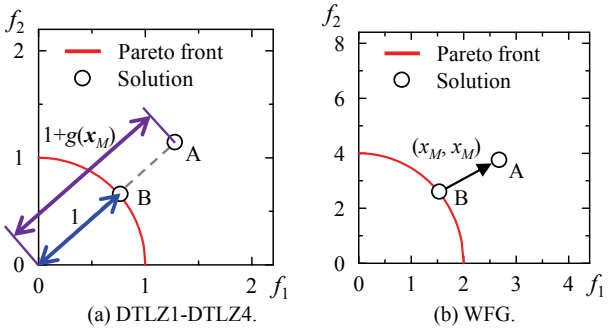
A. DTLZ1-DTLZ4

The first four test problems of the DTLZ family (i.e., DTLZ1-DTLZ4) have M minimization objectives of the following form [4]:

$$f_j(\boldsymbol{x}) = (1+g(\boldsymbol{x}_M))h_j(\boldsymbol{x}_{\text{pos}}),\ j = 1, 2, ..., M, \quad (1)$$

where $\boldsymbol{x}=(x_1, x_2, \ldots, x_{M+k-1})$ is a decision vector with $0 \le x_i \le 1$, $h_j(\boldsymbol{x}_{\text{pos}})$ specifies the shape of the Pareto front ($h_j(\boldsymbol{x}_{\text{pos}}) \ge 0$), and $g(\boldsymbol{x}_M)$ specifies the distance from the Pareto front ($g(\boldsymbol{x}_M) \ge 0$). The M functions $h_j(\boldsymbol{x}_{\text{pos}})$ are fully conflicting with each other. The decision vector $\boldsymbol{x}$ is divided into the first $M-1$ variables $\boldsymbol{x}_{\text{pos}}=(x_1, x_2, ..., x_{M-1})$ in $h_j(\boldsymbol{x}_{\text{pos}})$ and the other k variables $\boldsymbol{x}_M = (x_M, x_{M+1}, ..., x_{M+k-1})$ in $g(\boldsymbol{x}_M)$ where k is a parameter to specify the number of the decision variables.

Since all objectives have the form of $(1+g(\boldsymbol{x}_M))h_j(\boldsymbol{x}_{\text{pos}})$ with $h_j(\boldsymbol{x}_{\text{pos}}) \ge 0$ and $g(\boldsymbol{x}_M) \ge 0$, they can be simultaneously improved by minimizing $g(\boldsymbol{x}_M)$. That is, the M objectives are not conflicting with respect to the minimization of $g(\boldsymbol{x}_M)$. However, they are conflicting since the M functions $h_j(\boldsymbol{x}_{\text{pos}})$ are fully conflicting. The decrease of $h_j(\boldsymbol{x}_{\text{pos}})$ always leads to the increase of at least one $h_i(\boldsymbol{x}_{\text{pos}})$ with $i \ne j$. Thus all solutions with $g(\boldsymbol{x}_M)=0$ are Pareto optimal while no solutions with $g(\boldsymbol{x}_M)>0$ are Pareto optimal. This structure of DTLZ1-DTLZ4 is illustrated in Fig. 2 (a) where A is a dominated solution and B is a Pareto optimal solution.



(a) DTLZ1-DTLZ4. (b) WFG.

Fig. 2. Illustrations of test problems for the case of two objectives. In (a), the objective vector of A is obtained by multiplying the objective vector of B by $(1 + g(\boldsymbol{x}_M))$. In (b), the objective vector of A is obtained by adding a vector (x_M, x_M) to the objective vector of B.

B. DTLZ5

Objectives of DTLZ5 have the same form as (1) except that the $h_j(.)$ part of (1) depends on $g(\boldsymbol{x}_M)$ as follows:

$$f_j(\boldsymbol{x}) = \begin{cases} (1+g(\boldsymbol{x}_M))\cos(\theta_1)\cdots\cos(\theta_{M-2})\cos(\theta_{M-1}) & \text{for } j=1, \\ (1+g(\boldsymbol{x}_M))\cos(\theta_1)\cdots\cos(\theta_{M-j})\sin(\theta_{M-j+1}) & \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{for } 1<j<M, & \\ (1+g(\boldsymbol{x}_M))\sin(\theta_1) & \text{for } j=M, \end{cases} \quad (2)$$

where

$$\theta_i = \begin{cases} (\pi/2)x_i & \text{for } i=1, \\ \dfrac{\pi}{4(1+g(\boldsymbol{x}_M))}(1+2g(\boldsymbol{x}_M)x_i) & \text{for } i=2,3,...,M-1, \end{cases} \quad (3)$$

$$g(\boldsymbol{x}_M) = (x_M - 0.5)^2 + \cdots + (x_{M+k-1} - 0.5)^2, \quad (4)$$

$$0 \le x_i \le 1 \text{ for } i = 1,2,...,M+k-1. \quad (5)$$

When $g(\boldsymbol{x}_M)=0$ in (3), $\theta_1=(\pi/2)x_1$ and $\theta_i=\pi/4$ for $i>1$. Thus (2) can be rewritten as follows since $\cos(\pi/4)=\sin(\pi/4)$:

$$f_j(\boldsymbol{x}) = \begin{cases} \cos((\pi/2)x_1)(\cos(\pi/4))^{M-2} & \text{for } j=1, \\ \cos((\pi/2)x_1)(\cos(\pi/4))^{M-j} & \text{for } 1<j<M, \\ \sin((\pi/2)x_1) & \text{for } j=M, \end{cases} \quad (6)$$

where $0 \le x_1 \le 1$. This is the intended degenerate Pareto front of DTLZ5, which is shown in Fig. 1 (a).

In DTLZ1-DTLZ4, the Pareto front is characterized by $g(\boldsymbol{x}_M)=0$ because $h_j(\boldsymbol{x}_{\text{pos}})$ and $g(\boldsymbol{x}_M)$ are separable. However, in DTLZ5, θ_i depends on $g(\boldsymbol{x}_M)$ in (3). Thus we cannot say that no solutions with $g(\boldsymbol{x}_M) \ne 0$ are Pareto optimal. As pointed out in [5], [6], [8], some solutions with $g(\boldsymbol{x}_M) \ne 0$ are Pareto optimal when $M>3$. That is, the true Pareto front of DTLZ5 is not a degenerate curve in the case of four or more objectives.

To remove the non-degenerate part of the Pareto front of DTLZ5, the following constraints were introduced in [8]:

$$2^{M-2} f_1^2 + f_M^2 \ge 1, \quad (7)$$

$$2^{M-j} f_j^2 + f_M^2 \ge 1 \text{ for } j = 2,3,..., M-1. \quad (8)$$

DTLZ6 in [4], which is DTLZ5 in [3], is the same as DTLZ5 in (2)-(5) except that $g(\boldsymbol{x}_M)$ in (4) is defined as follows:

$$g(\boldsymbol{x}_M) = x_M^{0.1} + x_{M+1}^{0.1} + \cdots + x_{M+k-1}^{0.1}. \quad (9)$$

DTLZ5 and DTLZ6 have the same intended degenerate Pareto front in (6), which was derived from $g(\boldsymbol{x}_M) = 0$.

C. DTLZ5(I, M)

The dimension of the intended degenerate Pareto front of DTLZ5 is always one independent of the number of objectives. As an extension of DTLZ5, DTLZ5(I, M) was proposed in [7]. This extension was intended to design an M-objective problem with an $(I-1)$-dimensional degenerate Pareto front by changing the definition of θ_i in (3) as follows:

$$\theta_i = \begin{cases} (\pi/2)x_i & \text{for } i=1,2,...,I-1, \\ \dfrac{\pi}{4(1+g(\boldsymbol{x}_M))}(1+2g(\boldsymbol{x}_M)x_i) & \text{for } i=I,...,M-1. \end{cases} \quad (10)$$

The intended degenerate Pareto front is the set of solutions with $g(\boldsymbol{x}_M) = 0$ in (10). However, for the same reason as in Subsection II.B, the Pareto front of DTLZ5(I, M) has a non-degenerate part [8]. Constraint conditions similar to (7)-(8) were introduced in [8] to remove the non-degenerate part.

III. WFG Test Problems

A toolkit called WFG was proposed to generate a variety of multiobjective test problems in [5], [6]. Nine test problems WFG1-WGF9 have often been used in the EMO community.

Objectives of WFG1-WFG9 can be written as follows [6]:

$$f_j(\boldsymbol{x}) = Dx_M + S_j h_j(x_1, x_2, ..., x_{M-1}),\ j = 1,2,...,M, \quad (11)$$

$$x_i = \begin{cases} \max\{t_M^p, A_i\} \cdot (t_i^p - 0.5) + 0.5 & \text{for } i=1,2,...,M-1, \\ t_M^p & \text{for } i=M, \end{cases} \quad (12)$$

$$\boldsymbol{t}^p \leftarrow \boldsymbol{t}^{p-1} \leftarrow \boldsymbol{t}^{p-2} \leftarrow \cdots \leftarrow \boldsymbol{t}^1 \leftarrow \boldsymbol{z}. \quad (13)$$

In (13), a transition vector $\boldsymbol{t}^p$ is calculated from a decision vector $\boldsymbol{z}$ through a p-step transformation. Then, an underlying vector $\boldsymbol{x}$ is calculated from $\boldsymbol{t}^p$ by (12). In WFG, both $\boldsymbol{t}^p$ and $\boldsymbol{x}$ are M-dimensional vectors such that $\boldsymbol{t}^p \in [0, 1]^M$ and $\boldsymbol{x} \in [0, 1]^M$. In (12), A_i is a degeneracy constant, which is 0 or 1. Each objective function in (11) is calculated from the underlying vector $\boldsymbol{x}$. While the number of decision variables in $\boldsymbol{z}$ can be arbitrarily specified, WFG always has M transition variables in $\boldsymbol{t}^p$ and M underlying variables in $\boldsymbol{x}$. The domain of $\boldsymbol{t}^p$ is the entire M-dimensional hyper-cube $[0, 1]^M$. Using this property of the domain of $\boldsymbol{t}^p$, we discuss the shape of the Pareto front by viewing the objective functions in (11)-(13) as a mapping from $\boldsymbol{t}^p$ in $[0, 1]^M$ to an M-dimensional objective vector. In (11), D and S_j are specified as $D = 1$ and $S_j = 2j$ for $j = 1, 2, ..., M$.

As in DTLZ1-DTLZ4 in (1), $h_j(.)$ in (11) with the $M-1$ underlying position variables $(x_1, x_2, \ldots, x_{M-1})$ determines the shape of the Pareto front. The distance from the Pareto front is determined by x_M, which is called an underlying distance variable. This structure of WFG is illustrated in Fig. 1 (b). Fig. 1 (a) and Fig. 1 (b) show that $g(\boldsymbol{x}_M)$ in DTLZ and x_M in WFG play the same role. When x_M is separable from $(x_1, x_2, \ldots, x_{M-1})$, the Pareto front is characterized by $x_M = 0$. In this case, no solutions with $x_M > 0$ are Pareto optimal. This is because solutions with $x_M > 0$ can be further improved by decreasing x_M.

Let us discuss the role of the degeneracy constant A_i in (12). When A_i is specified as $A_i = 1$, $\max\{t_M^p, A_i\} = 1$ always holds since $0 \le t_M^p \le 1$. Thus we have $x_i = t_i^p$ from (12). When all A_i's are specified as $A_i = 1$, the underlying vector $\boldsymbol{x}$ is the same as the transition vector $\boldsymbol{t}^p$. In this case, x_M is separable from $(x_1, x_2, ..., x_{M-1})$ in each objective function in (11). Thus x_M can be minimized to $x_M = 0$ independent of the other underlying variables $(x_1, x_2, ..., x_{M-1})$. When the shape functions $h_j(.)$ in (11) are conflicting with each other and all A_i's are specified as $A_i = 1$, the Pareto front obtained from $x_M = 0$ is not degenerate.

Let us assume that A_i is specified as $A_i = 0$ for a particular i. We also assume that $x_M = 0$. In this case, $x_i = 0.5$ in (12) since $x_M = t_M^p = 0$. As a result, the dimension of the Pareto front is decreased to $(M-2)$. By increasing the number of A_i with $A_i = 0$, the dimension of the Pareto front can be decreased from $(M-1)$ to 1. This is the basic idea of formulating a multiobjective test problem with a degenerate Pareto front in WFG. However, this idea does not always work as intended. This is because x_i depends on x_M when $A_i = 0$. That is, x_i is calculated from (12) as follows when $A_i = 0$ since $x_M = t_M^p \ge 0$:

$$x_i = x_M(t_i^p - 0.5) + 0.5. \tag{14}$$

Due to the dependence of x_i on x_M, we cannot say that no solutions with $x_M > 0$ are Pareto optimal. There is a possibility that some solutions with $x_M > 0$ are Pareto optimal. This means that we cannot always generate a multiobjective test problem with a degenerate Pareto front by specifying A_i as $A_i = 0$. This issue is further discussed in the next section.

IV. Pareto Front of WFG3

A. WFG3 Test Problem

In WFG3, A_i is specified as $A_i = 1$ for $i = 1$ and $A_i = 0$ for $i = 2, 3, ..., M-1$. Thus the underlying vector $\boldsymbol{x}$ is specified as

$$x_i = \begin{cases} t_1^p & \text{for } i = 1, \\ t_M^p(t_i^p - 0.5) + 0.5 & \text{for } 1 < i < M, \\ t_M^p & \text{for } i = M. \end{cases} \tag{15}$$

WFG3 has the following M objective functions:

$$f_j(\boldsymbol{x}) = \begin{cases} x_M + S_1 x_1 x_2 \cdots x_{M-1} & \text{for } j = 1, \\ x_M + S_j(1 - x_{M-j+1}) x_1 x_2 \cdots x_{M-j} & \text{for } 1 < j < M, \\ x_M + S_M(1 - x_1) & \text{for } j = M. \end{cases} \tag{16}$$

The intended degenerate Pareto front is derived from the assumption that no solutions with $x_M > 0$ are Pareto optimal (whereas this is not always the case). For deriving the intended degenerate Pareto front, it is assumed that $x_M = 0$ holds for all Pareto optimal solutions of WFG3. From $x_M = 0$, the underlying vector $\boldsymbol{x}$ in (15) is calculated as follows: $x_1 = t_1^p$, $x_i = 0.5$ for $1 < i < M$, and $x_M = 0$: $\boldsymbol{x} = (x_1, 0.5, ..., 0.5, 0)$. In this case, the objective functions in (16) can be rewritten as follows:

$$f_j(x_1) = \begin{cases} S_1 x_1 / 2^{(M-2)} & \text{for } j = 1, \\ S_j x_1 / 2^{(M-j)} & \text{for } 1 < j < M, \\ S_M(1 - x_1) & \text{for } j = M, \end{cases} \tag{17}$$

where $0 \le x_1 \le 1$.

By changing the value of x_1 from 0 to 1, the M-dimensional objective vector $\boldsymbol{f}(x_1) = (f_1(x_1), f_2(x_1), ..., f_M(x_1))$ in (17) forms a straight line from $(0, 0, 0, ..., 0, S_M)$ to $(S_1/2^{M-2}, S_2/2^{M-2}, S_3/2^{M-3}, ..., S_{M-1}/2, 0)$. This is the intended degenerate Pareto front of WFG3. We denote this Pareto front by P^*, which is the straight line specified by (17). In the following, we examine whether the degenerate Pareto front P^* is the true Pareto front or not.

B. Degenerate Part of the Pareto Front

In this subsection, we show that all solutions in P^* are Pareto optimal. In WFG3 [6], the shape functions $h_j(.)$ are specified as "linear" as follows:

$$\sum_{j=1}^{M} h_j(\boldsymbol{x}) = 1. \tag{18}$$

Since D in (11) is always specified as $D = 1$ in WFG, the following relation is derived from (11) and (18):

$$\sum_{j=1}^{M} \frac{f_j(\boldsymbol{x})}{S_j} = 1 + \sum_{j=1}^{M} \frac{x_M}{S_j}. \tag{19}$$

From the definition of P^*, $x_M = 0$ holds for all solutions in P^*. Thus, the right-hand side of (19) is 1 for all solutions in P^*. Let us assume that a solution $\boldsymbol{x}$ dominates $\boldsymbol{x}^*$ in P^* (i.e., $\boldsymbol{f}(\boldsymbol{x})$ dominates $\boldsymbol{f}(\boldsymbol{x}^*)$ in P^*). In this case, the following relation in (20) holds from (19) since $x_M^* = 0$, $S_j > 0$ and $f_j(\boldsymbol{x}) \le f_j(\boldsymbol{x}^*)$ for all j's, and $f_j(\boldsymbol{x}) < f_j(\boldsymbol{x}^*)$ for at least one j:

$$1 + \sum_{j=1}^{M} \frac{x_M}{S_j} = \sum_{j=1}^{M} \frac{f_j(\boldsymbol{x})}{S_j} < \sum_{j=1}^{M} \frac{f_j(\boldsymbol{x}^*)}{S_j} = 1 + \sum_{j=1}^{M} \frac{x_M^*}{S_j} = 1. \tag{20}$$

This implies that x_M is negative (since $S_j > 0$ for all j's). However, there is no solution with a negative x_M since the range of x_M is the unit interval [0, 1]. Thus we can say that no solution in P^* is dominated by any other solution. From these discussions, we have the following property:

Property 1. All solutions in P^* are Pareto optimal.

We can show from (19) that $\boldsymbol{f}(\boldsymbol{x})$ with $x_M > 0$ is not the same as any point in P^* with $x_M = 0$. Let us assume that $\boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{f}(\boldsymbol{x}^*)$ holds for $\boldsymbol{x}$ with $x_M > 0$ and $\boldsymbol{x}^*$ with $x^*_M = 0$. In this case, we have the following relation from (19) and $\boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{f}(\boldsymbol{x}^*)$:

$$1 < 1 + \sum_{j=1}^{M} \frac{x_M}{S_j} = \sum_{j=1}^{M} \frac{f_j(\boldsymbol{x})}{S_j} = \sum_{j=1}^{M} \frac{f_j(\boldsymbol{x}^*)}{S_j} = 1 + \sum_{j=1}^{M} \frac{x^*_M}{S_j} = 1. \quad (21)$$

This inconsistency shows that $\boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{f}(\boldsymbol{x}^*)$ never holds for $\boldsymbol{x}$ with $x_M > 0$ and $\boldsymbol{x}^*$ with $x^*_M = 0$. This means that the solution set with $x_M > 0$ has no overlap with P^* with $x^*_M = 0$.

C. Dominance Test

In this subsection, we explain how we can check whether a solution $\boldsymbol{x}$ is dominated by at least one solution in P^* (i.e., whether $\boldsymbol{f}(\boldsymbol{x})$ is dominated by at least one point in P^*). This is to examine whether WFG3 has other Pareto optimal solutions.

The degenerate Pareto front P^* of WFG3 is a straight line from (0, 0, 0, ..., 0, S_M) to ($S_1/2^{M-2}$, $S_2/2^{M-2}$, $S_3/2^{M-3}$, ..., $S_{M-1}/2$, 0). Thus we can specify a single point in P^* using the value of one objective. We use the value of the M-th objective to specify a single point in P^*. Let u be a value of the M-th objective where $0 \le u \le S_M$. We denote the corresponding point in P^* as $\boldsymbol{p}^*(u) = (p^*_1(u), p^*_2(u), ..., p^*_M(u))$ where $p^*_M(u) = u$ and $\boldsymbol{p}^*(u) \in P^*$. This point can be calculated using x_1 in (17) as

$$\boldsymbol{p}^*(u) = (f_1(x_1), f_2(x_1), ..., f_M(x_1)) \text{ with } x_1 = 1 - u/S_M. \quad (22)$$

It should be noted that all points in P^* can be represented in the form of $\boldsymbol{p}^*(u)$ with $0 \le u \le S_M$: $P^* = \{\boldsymbol{p}^*(u) \mid 0 \le u \le S_M\}$.

We can check whether a solution $\boldsymbol{x}$ is dominated by P^* or not using the following property.

PROPERTY 2. A solution $\boldsymbol{x}$ is not dominated by any solution in P^* if and only if $f_M(\boldsymbol{x}) \le S_M$ holds and $\boldsymbol{f}(\boldsymbol{x})$ is not dominated by $\boldsymbol{p}^*(f_M(\boldsymbol{x}))$.

PROOF. First, let us assume that $f_M(\boldsymbol{x}) \le S_M$ holds and $\boldsymbol{f}(\boldsymbol{x})$ is not dominated by $\boldsymbol{p}^*(f_M(\boldsymbol{x}))$. When $u > f_M(\boldsymbol{x})$, it is clear that $\boldsymbol{f}(\boldsymbol{x})$ is not dominated by $\boldsymbol{p}^*(u)$ since $f_M(\boldsymbol{x}) < u = p^*_M(u)$. When $u = f_M(\boldsymbol{x})$, $\boldsymbol{f}(\boldsymbol{x})$ is not dominated by $\boldsymbol{p}^*(u)$ from the assumption. Therefore, we can focus our attention on solutions $\boldsymbol{p}^*(u)$ in P^* with $0 \le u < f_M(\boldsymbol{x})$. From (17), we can see that the decrease in the M-th objective increases all the other objectives in P^*. Thus $p^*_j(f_M(\boldsymbol{x})) < p^*_j(u)$ holds for $j = 1, 2, ..., M-1$ when $u < f_M(\boldsymbol{x})$. Since $\boldsymbol{f}(\boldsymbol{x})$ is not dominated by $\boldsymbol{p}^*(f_M(\boldsymbol{x}))$ and $f_M(\boldsymbol{x}) = p^*_M(f_M(\boldsymbol{x}))$, there exists at least one objective f_j ($j \ne M$) such that $f_j(\boldsymbol{x}) \le p^*_j(f_M(\boldsymbol{x}))$ holds (including the case of $\boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{p}^*(f_M(\boldsymbol{x}))$). Thus $f_j(\boldsymbol{x}) \le p^*_j(f_M(\boldsymbol{x})) < p^*_j(u)$ holds for at least one objective f_j ($j \ne M$) when $u < f_M(\boldsymbol{x})$. This means that $\boldsymbol{f}(\boldsymbol{x})$ is not dominated by $\boldsymbol{p}^*(u)$ when $0 \le u < f_M(\boldsymbol{x})$. Thus we can say that $\boldsymbol{f}(\boldsymbol{x})$ is not dominated by $\boldsymbol{p}^*(u)$ for $0 \le u \le S_M$ (since we have already shown that $\boldsymbol{f}(\boldsymbol{x})$ is not dominated by $\boldsymbol{p}^*(u)$ for $f_M(\boldsymbol{x}) \le u \le S_M$).

Next, let us assume that a solution $\boldsymbol{x}$ is not dominated by any solution in P^*. From this assumption, it is clear that $\boldsymbol{f}(\boldsymbol{x})$ is not dominated by $\boldsymbol{p}^*(f_M(\boldsymbol{x}))$. If $f_M(\boldsymbol{x}) > S_M$ holds, $\boldsymbol{f}(\boldsymbol{x})$ is dominated by $\boldsymbol{p}^*(S_M) = (0, 0, ..., 0, S_M)$ since the value of each objective in WFG3 is always non-negative (i.e., $f_j(\boldsymbol{x}) \ge 0$ for $j = 1, 2, ..., M$). This contradicts the assumption that $\boldsymbol{x}$ is not dominated by any solution in P^*. Thus we can say that $f_M(\boldsymbol{x}) \le S_M$. This completes the proof of Property 2.

D. Non-Degenerate Pareto Front

In this subsection, we show that WFG3 with three or more objectives has other Pareto optimal solutions in addition to the degenerate Pareto front P^* using Property 2. Let us consider a solution $\boldsymbol{x}$ with $x_1 = 1$, $x_2 = 1$ and $x_M = 1$. In this case, we have $f_M(\boldsymbol{x}) = 1$ and $f_{M-1}(\boldsymbol{x}) = 1$ from (16). Since $x_M \ne 0$, this solution is not included in the degenerate Pareto front P^*. Let us check whether this solution is dominated by the degenerate Pareto front P^* or not using Property 2.

Since $f_M(\boldsymbol{x}) = 1$ and $S_M = 2M$, $f_M(\boldsymbol{x}) \le S_M$ in Property 2 holds. Let us examine the Pareto dominance relation between $\boldsymbol{f}(\boldsymbol{x})$ and $\boldsymbol{p}^*(f_M(\boldsymbol{x}))$. From $f_M(\boldsymbol{x}) = 1$ and $S_j = 2j$, $p^*_{M-1}(f_M(\boldsymbol{x}))$ is calculated from (17) and (22) as:

$$p^*_{M-1}(f_M(\boldsymbol{x})) = p^*_{M-1}(1) = \frac{(M-1)(2M-1)}{2M}. \quad (23)$$

If the following relation holds, we can say that $\boldsymbol{f}(\boldsymbol{x})$ is not dominated by $\boldsymbol{p}^*(f_M(\boldsymbol{x}))$:

$$p^*_{M-1}(f_M(\boldsymbol{x})) = \frac{(M-1)(2M-1)}{2M} > f_{M-1}(\boldsymbol{x}) = 1. \quad (24)$$

This inequality holds when $M < 0.22$ or $2.28 < M$. Since M is the number of objectives, we can see that this inequality holds for three or more objectives. This means that a solution $\boldsymbol{x}$ with $x_1 = 1$, $x_2 = 1$ and $x_M = 1$ satisfies the conditions of Property 2. That is, such a solution is not dominated by the degenerate Pareto front P^*.

From these discussions, we can see that there exists a Pareto optimal solution in addition to the degenerate Pareto front P^*. As an example, let us consider a solution $\boldsymbol{x} = (1, 1, 1)$ of a three-objective WFG3 problem. The objective vector $\boldsymbol{f}(\boldsymbol{x})$ is calculated from (16) as $\boldsymbol{f}(\boldsymbol{x}) = (3, 1, 1)$ for $\boldsymbol{x} = (1, 1, 1)$. Using (17) and (22), $\boldsymbol{p}^*(f_M(\boldsymbol{x}))$ is calculated as $\boldsymbol{p}^*(f_M(\boldsymbol{x})) = (5/6, 5/3, 1)$ from $f_M(\boldsymbol{x}) = 1$ and $M = 3$. Since $f_M(\boldsymbol{x}) \le S_M$ holds from $S_M = 2M = 6$ and $\boldsymbol{f}(\boldsymbol{x})$ is not dominated by $\boldsymbol{p}^*(f_M(\boldsymbol{x}))$, $\boldsymbol{f}(\boldsymbol{x})$ is not dominated by any point in the degenerate Pareto front P^*.

From these discussions, we can see that WFG3 with three or more objectives has some other Pareto optimal solutions in addition to the intended degenerate Pareto front P^*. The true Pareto front consists of the intended degenerate Pareto front P^* and the non-dominated solutions among all solutions satisfying Property 2. However, it is not easy to derive the shape of those non-dominated solutions in the objective space.

Even for the case of three objectives, the derivation of the non-degenerate part of the Pareto front of WFG3 needs a lot of calculations. After a long derivation process, we obtain the following characterization of the non-degenerate part of the Pareto front of the three-objective WFG3 problem.

$$t^p_2 = 1, 0 \le f_3(\boldsymbol{x}) < 3, \max\{0, f_3(\boldsymbol{x}) - 2\} < x_3 \le \min\{1, f_3(\boldsymbol{x})\}. \quad (25)$$

The non-degenerate part consists of all solutions satisfying this formulation (i.e., all equality and inequality conditions in (25)). The outline of the derivation of (25) is as follows. First it is shown that solutions with $t^p_2 \ne 1$ are dominated by those with $t^p_2 = 1$. Next the Pareto optimality of solutions with $t^p_2 = 1$ is examined for the three cases: $0 \le f_3(\boldsymbol{x}) < 2$, $2 \le f_3(\boldsymbol{x}) < 3$ and $3 \le f_3(\boldsymbol{x})$. Then the obtained results are summarized. Instead of explaining such a long derivation process in detail using a number of pages, we will compare this formulation and the obtained solutions by EMO algorithms in Section V.

E. Introduction of Constraints

In this subsection, we derive constraints from Property 2 to remove the non-degenerate part of the Pareto front. Property 2 can be rewritten as follows: A solution $\boldsymbol{x}$ is dominated by at least one solution in P^* if and only if $f_M(\boldsymbol{x}) > S_M$ holds or $\boldsymbol{f}(\boldsymbol{x})$ is dominated by $\boldsymbol{p}^*(f_M(\boldsymbol{x}))$. Thus it is clear that $\boldsymbol{x}$ is not a Pareto optimal solution when $f_M(\boldsymbol{x}) > S_M$ holds.

Let us assume that $f_M(\boldsymbol{x}) \le S_M$ holds. From Property 2, $\boldsymbol{x}$ is dominated by at least one solution in P^* if and only if $\boldsymbol{f}(\boldsymbol{x})$ is dominated by $\boldsymbol{p}^*(f_M(\boldsymbol{x}))$. In addition to this case, $\boldsymbol{f}(\boldsymbol{x})$ is not a point in the non-degenerate part when $\boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{p}^*(f_M(\boldsymbol{x}))$ holds. Thus we can see that $\boldsymbol{f}(\boldsymbol{x})$ is dominated by at least one point in P^* or $\boldsymbol{f}(\boldsymbol{x})$ is in P^* if $p_i^*(f_M(\boldsymbol{x})) \le f_i(\boldsymbol{x})$ holds for $i = 1, 2, ..., M$.

From the definition of $\boldsymbol{p}^*(u)$ in Subsection IV.C, $p_M^*(f_M(\boldsymbol{x})) = f_M(\boldsymbol{x})$. Thus $p_i^*(f_M(\boldsymbol{x})) \le f_i(\boldsymbol{x})$ always holds for $i = M$. Using (22) with $u = f_M(\boldsymbol{x})$ in (17), we have the following formulation of $\boldsymbol{p}^*(f_M(\boldsymbol{x}))$:

$$p_j^*(f_M(\boldsymbol{x})) = \begin{cases} S_1(1 - f_M(\boldsymbol{x})/S_M)/2^{(M-2)} & \text{for } j = 1, \\ S_j(1 - f_M(\boldsymbol{x})/S_M)/2^{(M-j)} & \text{for } 1 < j < M, \\ f_M(\boldsymbol{x}) & \text{for } j = M. \end{cases} \quad (26)$$

Using (26), the condition $p_j^*(f_M(\boldsymbol{x})) \le f_j(\boldsymbol{x})$ for $j = 1, 2, ..., M-1$ can be written as follows (for $j = M$, this relation always holds since $p_j^*(f_M(\boldsymbol{x})) = f_j(\boldsymbol{x})$):

$$\frac{2^{M-2}}{S_1} f_1(\boldsymbol{x}) + \frac{f_M(\boldsymbol{x})}{S_M} \ge 1, \quad (27)$$

$$\frac{2^{M-j}}{S_j} f_j(\boldsymbol{x}) + \frac{f_M(\boldsymbol{x})}{S_M} \ge 1 \text{ for } j = 2, 3, \ldots, M-1. \quad (28)$$

Thus we can see that a solution $\boldsymbol{x}$ is dominated by at least one solution in P^* if and only if $f_M(\boldsymbol{x}) > S_M$ holds or both (27) and (28) hold. Since $0 \le S_j$ and $0 \le f_j(\boldsymbol{x})$ for all j's in (27) and (28), all solutions satisfying $f_M(\boldsymbol{x}) > S_M$ always satisfy both (27) and (28). Thus a solution $\boldsymbol{x}$ is dominated by at least one solution in P^* if and only if both (27) and (28) hold. This means that the non-degenerate part of the Pareto front of WFG3 can be removed by adding (27) and (28). It should be noted that $p_i^*(f_M(\boldsymbol{x})) \le f_i(\boldsymbol{x})$ always holds for all j's when $\boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{p}^*(f_M(\boldsymbol{x}))$. This means that the addition of (27) and (28) does not change the intended degenerate Pareto front of WFG3 while it removes all the other Pareto optimal solutions.

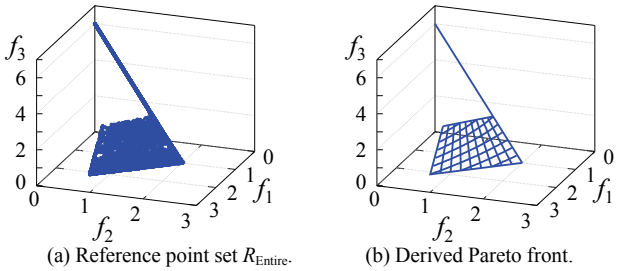
V. Computational Experiments

A. Settings of Computational Experiments

We examine the performance of NSGA-II [14], MOEA/D [15] with no archive and HypE [16] using GD and IGD. In GD and IGD, we specify the parameter p as $p = 1$ (see [17]). Each algorithm is applied to WFG3 under the following setting:

Number of objectives (WFG3): $M = 3, 4, 5, 6$,
Population size: 105, 84, 126, 126 (for 3, 4, 5, 6 objectives),
Number of position parameters (WFG3): $k = 2(M-1)$,
Number of distance parameters (WFG3): $l = 20$,
Number of decision variables (WFG3): $n = k + l$,
Stopping condition: 1000 generations,
Crossover: SBX ($\mu_C = 15$) with a crossover probability of 0.8,
Mutation: PM ($\mu_M = 20$) with a mutation probability of $1/n$,
Scalarizing function (MOEA/D): Tchebycheff function,
Neighbor size (MOEA/D): 10,
Number of sampling points (HypE): 10000,
Reference point (HypE): (3, 5, 7, …, $2M+1$).

Each algorithm is applied to each test problem 100 times. One reference point set R_{Part} for the calculation of GD and IGD is created by evenly sampling 1000 points from the degenerate part of the Pareto front. The size of R_{Part} is always 1000. Another reference point set R_{Entire} is constructed by selecting non-dominated solutions from all the obtained solutions by 100 runs of each algorithm (300 runs in total). The number of obtained reference points in R_{Entire} for each problem is 11,759 ($M = 3$), 12,283 ($M = 4$), 20,371 ($M = 5$), and 21,649 ($M = 6$). Fig. 3 (a) shows the reference points in R_{Entire} for the three-objective problem. The derived non-degenerate Pareto front in (25) is depicted in Fig. 3 (b) together with the degenerate line defined in (17). We can see that the reference point set R_{Entire} in Fig. 3 (a) is similar to the entire Pareto front in Fig. 3 (b).



(a) Reference point set R_{Entire}. (b) Derived Pareto front.

Fig. 3. Reference point set R_{Entire} and the Pareto front in (17) and (25).

B. Results of Computational Experiments

For examining the effect of the non-degenerate part on the behavior of each EMO algorithm, we monitor the number of solutions which are not dominated by the degenerate part P^* using Property 2 (i.e., solutions in the non-degenerate part). In Table I, we show the average percentage of those solutions (and the standard deviation) in the final generation over 100 runs of each EMO algorithm. The increase in the number of objectives clearly increases the percentage of those solutions in NSGA-II and HypE. That is, more solutions move to the non-degenerate part by increasing the number of objectives. This is because the dimension of the non-degenerate part increases with the number of objectives whereas the degenerate part is always one-dimensional. However, when MOEA/D is used, the number of objectives does not have a clear effect in Table I. Moreover, the standard deviation is very small in this case.

TABLE I
PERCENTAGE OF SOLUTIONS WHICH ARE NOT DOMINATED BY THE DEGENERATE PART P^* IN THE FINAL GENERATION OF EACH ALGORITHM.

Number of objectives	NSGA-II	MOEA/D	HypE
$M = 3$	43.0% (0.3)	37.1% (0.0)	20.0% (0.9)
$M = 4$	66.6% (0.4)	53.7% (0.0)	39.7% (0.8)
$M = 5$	77.4% (0.3)	49.8% (0.1)	48.6% (0.5)
$M = 6$	81.8% (0.3)	37.6% (0.2)	62.1% (0.5)

In Fig. 4 and Fig. 5, we show the average percentage of solutions which are not dominated by the degenerate part at each generation for WFG3 with three and six objectives, respectively. In these figures, the percentage of those solutions in MOEA/D is almost constant after 100 generations.

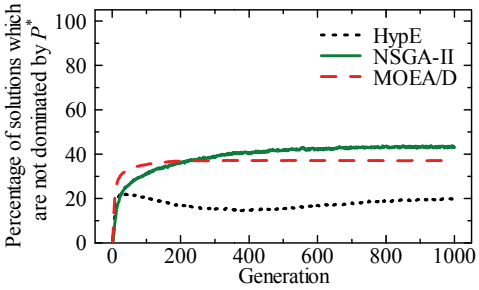


Fig. 4. Percentage of solutions which are not dominated by the degenerate part P^* (M=3).

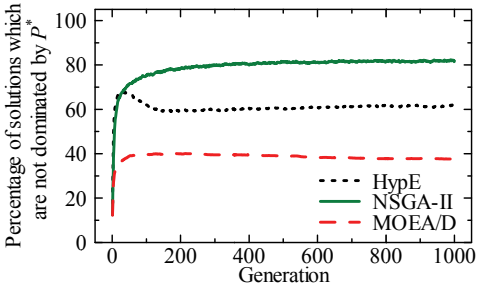
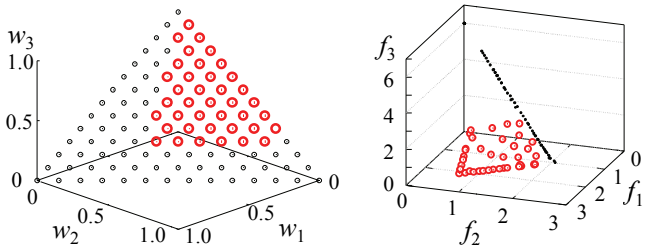


Fig. 5. Percentage of solutions which are not dominated by the degenerate part P^* (M=6).

When MOEA/D is used, the number of objectives does not have a clear effect on the percentage of solutions in the non-degenerate part. For further discussing this observation, we show the relation between the weight vectors and the obtained solutions by a single run of MOEA/D on WFG3 with three objectives in Fig. 6. Large circles in Fig. 6 (a) show weight vectors from which solutions are obtained in the non-degenerate part in Fig. 6 (b). Among 100 runs in Table I and Fig. 4, a single run is selected so that the percentage of the obtained solutions in the non-degenerate part is closest to its average value in Table I. The number of those solutions in Fig. 6 (a) is 39, which is 37.1% of the population size 105. For each test problem, we select a single run of MOEA/D using the same criteria based on the average percentage in Table I. Weight vectors corresponding to those solutions are shown in Fig. 7 by bold lines in each parallel coordinate plot. We can observe a similarity of bold lines among the four plots in Fig. 7. This observation explains why the average percentage of solutions in the non-degenerate part in Table I is not sensitive to the number of objectives when MOEA/D is used.

In the same manner, we also select a single run of the other algorithms for WFG with three objectives. Fig. 8 shows the obtained solutions by a single run in the f_1-f_3 subspace. Fig. 8 clearly demonstrates the characteristics of each algorithm.



(a) Weight vectors. (b) Obtained solutions.

Fig. 6. Relation between the weight vectors and the obtained solutions by MOEA/D (M=3).

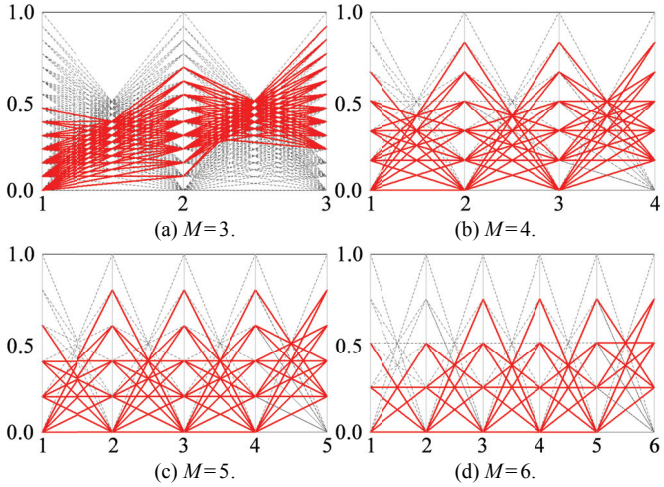


Fig. 7. Weight vectors from which solutions in the non-degenerate part are obtained by MOEA/D for WFG3 with M=3, 4, 5, 6.

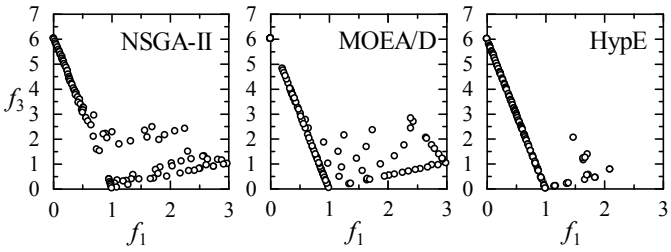


Fig. 8. Obtained solutions by each algorithm (M=3).

In Tables II-V, we show the average value and the standard deviation of GD and IGD over 100 runs for each reference point set. In Table VI, we show the results of the hypervolume indicator. The reference point for the hypervolume calculation is specified as (3, 5, 7, …, 2M+1) for the M-objective problem, which is the same specification as in the execution of HypE.

In each table, the statistically significant worst and best results are shown by bold face and "*", respectively. We use the Wilcoxon signed-rank test with the Bonferroni correction at the significance level of 5%. In Table II and Table III with R_{Part}, HypE is clearly the best among the three algorithms. However, in Table IV and V with R_{Entire}, NSGA-II is best for IGD while HypE is still the best for GD. These observations show that the performance comparison results by GD and IGD depend on the choice of the reference point set. Large GD values in Table II can be viewed as misleading results due to the use of R_{Part}. This is because the actual convergence performance of each algorithm is much better in Table IV. Large differences in the IGD values among the three algorithms in Table III can be also viewed as misleading results because the actual search ability of each algorithm evaluated by IGD is not so different in Table V.

An interesting observation is that the hypervolume-based performance comparison in Table VI is totally different from Table V. Especially for the three-objective WFG3 problem, HypE is evaluated as being the worst in Table V and the best in Table VI. This observation looks strange since the reference points from the entire Pareto front are used in Table V (and these two measures are usually consistent [18]). The reason for such a strange observation is a special structure of the Pareto front of WFG3. In Fig. 3 (a), R_{Entire} includes a large number of reference points on the widely spread non-degenerate part.

Thus many solutions are needed in this part for good results in Table V. However, the contribution of those solutions to the hypervolume is very small. The increase of the hypervolume from the degenerate part in Fig. 1 (b) to the entire Pareto front in Fig. 3 (b) is only 1.44% (76 vs. 77.0926). This also explains why only 20% solutions of HypE move to the non-degenerate part in Fig. 4 (also see Fig. 8). These observations suggest that WFG3 is a very interesting test problem to be further examined.

Our experimental results show different characteristics of each performance indicator with respect to the sensitivity to the specification of reference points. GD is sensitive because solutions in the non-degenerate part (e.g., large circles in Fig. 6 (b)) have large GD values when the reference points are only in the degenerate part as in Table II. IGD is less sensitive because those solutions are not used in its calculation in Table III. The hypervolume does not need any information about the true Pareto front. This is a clear advantage over the other indicators.

TABLE II
GD WITH THE REFERENCE POINT SET R_{PART}.

Obj.	NSGA-II	MOEA/D	HypE
$M=3$	0.5005 (0.0331)	**0.5494 (0.0185)**	0.1140* (0.0145)
$M=4$	**1.4833 (0.0674)**	1.4629 (0.0519)	0.4785* (0.0388)
$M=5$	**2.4423 (0.0811)**	2.3659 (0.0735)	0.9904* (0.0665)
$M=6$	**3.4228 (0.0963)**	3.1522 (0.1753)	1.4601* (0.0936)

*: Best results, Bold: Worst results

TABLE III
IGD WITH THE REFERENCE POINT SET R_{PART}.

Obj.	NSGA-II	MOEA/D	HypE
$M=3$	0.1005 (0.0161)	0.0990 (0.0059)	0.0369* (0.0029)
$M=4$	0.3141 (0.0543)	**0.9944 (0.0542)**	0.0698* (0.0053)
$M=5$	0.4637 (0.0737)	**1.9406 (0.0642)**	0.0840* (0.0076)
$M=6$	0.6684 (0.1172)	**2.9076 (0.0564)**	0.1150* (0.0103)

TABLE IV
GD WITH THE REFERENCE POINT SET R_{ENTIRE}.

Obj.	NSGA-II	MOEA/D	HypE
$M=3$	**0.0282 (0.0090)**	0.0058 (0.0060)	0.0021* (0.0015)
$M=4$	**0.0673 (0.0186)**	0.0185 (0.0099)	0.0030* (0.0023)
$M=5$	**0.0865 (0.0207)**	0.0093 (0.0047)	0.0024* (0.0016)
$M=6$	**0.0951 (0.0250)**	0.0072 (0.0058)	0.0023* (0.0014)

TABLE V
IGD WITH THE REFERENCE POINT SET R_{ENTIRE}.

Obj.	NSGA-II	MOEA/D	HypE
$M=3$	0.1444 (0.0112)	0.0810* (0.0038)	**0.3130 (0.0224)**
$M=4$	0.3971*(0.0220)	**0.5498 (0.0213)**	0.5261 (0.0302)
$M=5$	0.5628*(0.0234)	**1.0462 (0.0344)**	0.6527 (0.0401)
$M=6$	0.7714*(0.0319)	**1.8077 (0.0713)**	0.8942 (0.0691)

TABLE VI
HYPERVOLUME.

Obj.	NSGA-II	MOEA/D	HypE
$M=3$	7.477e1 (0.026e1)	**7.468e1 (0.021e1)**	7.562e1* (0.008e1)
$M=4$	6.413e2 (0.092e2)	**5.745e2 (0.059e2)**	6.706e2* (0.011e2)
$M=5$	6.930e3 (0.125e3)	**6.010e3 (0.060e3)**	7.381e3* (0.014e3)
$M=6$	8.772e4 (0.238e4)	**7.227e4 (0.093e4)**	9.567e4* (0.022e4)

VI. CONCLUSION

In this paper, we showed that the Pareto front of WFG3 has a non-degenerate part as well as the intended degenerate part. We also derived constraints to remove the non-degenerate part. Through computational experiments, we demonstrated that the use of reference points only from the degenerate part can lead to misleading performance evaluation results. We also obtained the following interesting observations about WFG3:

1. NSGA-II, MOEA/D and HypE showed totally different search behavior on WFG3 in Fig. 4, Fig. 5 and Fig. 8.
2. Inconsistent results were obtained from the hypervolume and the IGD with reference points on the entire Pareto front.
3. Consistent results were obtained from the hypervolume and the IGD with reference points only from the degenerate part.

REFERENCES

[1] H. Ishibuchi, N. Tsukamoto, and Y. Nojima, "Evolutionary many-objective optimization: A short review," in *Proc. 2008 IEEE Congr. Evol. Comput.*, pp. 2424-2431, Hong Kong, June 1-6, 2008.

[2] C. von Lücken, B. Barán, and C. Brizuela, "A survey on multi-objective evolutionary algorithms for many-objective problems," *Computational Optimization and Applications*, vol. 58, no. 3, pp. 707-756, July 2014.

[3] K. Deb, L. Thiele, M. Laumanns, and E. Zitzler, "Scalable multi-objective optimization test problems," in *Proc. 2002 IEEE Congr. Evol. Comput.*, pp. 825-830, Honolulu, May 12-17, 2002.

[4] K. Deb, L. Thiele, M. Laumanns, and E. Zitzler, "Scalable test problems for evolutionary multiobjective optimization," in A. Abraham, L. Jain, and R. Goldberg (Eds.), *Evolutionary Multiobjective Optimization*, pp. 105-145, Springer-Verlag, London, 2005.

[5] S. Huband, L. Barone, L. While, and P. Hingston, "A scalable multi-objective test problem toolkit," *Lecture Notes in Computer Science 3410: EMO 2005*, pp. 280-295, Springer, Berlin, March 2005.

[6] S. Huband, P. Hingston, L. Barone, and L. While, "A review of multiobjective test problems and a scalable test problem toolkit," *IEEE Trans. Evol. Comput.*, vol. 10, no. 5, pp. 477-506, 2006.

[7] K. Deb and D. K. Saxena, "On finding Pareto-optimal solutions through dimensionality reduction for certain large-dimensional multi-objective optimization problems," *KanGAL Report*, No. 2005011, Indian Institute of Technology Kanpur, 2005.

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

[9] D. K. Saxena, Q. Zhang, J. A. Duro, and A. Tiwari, "Framework for many-objective test problems with both simple and complicated Pareto-set shapes," *Lecture Notes in Computer Science 6576: EMO 2011*, pp. 197-211, Springer, Berlin, April 2011.

[10] H. Li, Q. Zhang, and J. Deng, "Multiobjective test problems with complicated Pareto fronts: Difficulties in degeneracy," in *Proc. 2014 IEEE Congr. Evol. Comput.*, pp. 2156-2163, Beijing, July 6-11, 2014.

[11] D. A. Van Veldhuizen, "Multiobjective evolutionary algorithms: Classifications, analyses, and new innovations," *Ph. D dissertation*, Air Force Institute of Technology, Wright-Patterson AFB, Ohio, May 1999.

[12] C. A. C. Coello and M. R. Sierra, "A study of the parallelization of a coevolutionary multi-objective evolutionary algorithm," *Lecture Notes in Computer Science 2972: MICAI 2004*, pp. 688-697, April 2004.

[13] E. Zitzler and L. Thiele, "Multiobjective optimization using evolutionary algorithms – A comparative case study," *Lecture Notes in Computer Science 1498: Parallel Problem Solving from Nature – PPSN V*, pp. 292-301, Springer, Berlin, September 1998.

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

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

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

[17] O. Schütze, X. Esquivel, A. Lara, and C. A. C. Coello, "Using the averaged Hausdorff distance as a performance measure in evolutionary multiobjective optimization," *IEEE Trans. Evol. Comput.*, vol. 16, no. 4, pp. 504-522, August 2012.

[18] S. Jiang, Y. Ong, J. Zhang, and L. Feng, "Consistencies and contradictions of performance metrics in multiobjective optimization," *IEEE Trans. on Cybernetics*, vol. 44, no. 12, pp. 2391-2404, Dec. 2014.