

---

# How to Specify a Reference Point in Hypervolume Calculation for Fair Performance Comparison

**Hisao Ishibuchi**

hisao@sustc.edu.cn

Shenzhen Key Laboratory of Computational Intelligence, Department of Computer Science and Engineering, Southern University of Science and Technology, Shenzhen, 518005, China

**Ryo Imada**

ryo.imada@ci.cs.osakafu-u.ac.jp

Department of Computer Science and Intelligent Systems, Osaka Prefecture University, Sakai, 5998531, Japan

**Yu Setoguchi**

yu.setoguchi@ci.cs.osakafu-u.ac.jp

Department of Computer Science and Intelligent Systems, Osaka Prefecture University, Sakai, 5998531, Japan

**Yusuke Nojima**

nojima@cs.osakafu-u.ac.jp

Department of Computer Science and Intelligent Systems, Osaka Prefecture University, Sakai, 5998531, Japan

---

doi:10.1162/EVCO\_a\_00226

## Abstract

The hypervolume indicator has frequently been used for comparing evolutionary multi-objective optimization (EMO) algorithms. A reference point is needed for hypervolume calculation. However, its specification has not been discussed in detail from a viewpoint of fair performance comparison. A slightly worse point than the nadir point is usually used for hypervolume calculation in the EMO community. In this paper, we propose a reference point specification method for fair performance comparison of EMO algorithms. First, we discuss the relation between the reference point specification and the optimal distribution of solutions for hypervolume maximization. It is demonstrated that the optimal distribution of solutions strongly depends on the location of the reference point when a multi-objective problem has an inverted triangular Pareto front. Next, we propose a reference point specification method based on theoretical discussions on the optimal distribution of solutions. The basic idea is to specify the reference point so that a set of well-distributed solutions over the entire linear Pareto front has a large hypervolume and all solutions in such a solution set have similar hypervolume contributions. Then, we examine whether the proposed method can appropriately specify the reference point through computational experiments on various test problems. Finally, we examine the usefulness of the proposed method in a hypervolume-based EMO algorithm. Our discussions and experimental results clearly show that a slightly worse point than the nadir point is not always appropriate for performance comparison of EMO algorithms.

## Keywords

Evolutionary multi-objective optimization, hypervolume, reference point, performance comparison.

## 1 Introduction

In the field of evolutionary multi-objective optimization (EMO), a number of new algorithms continue to be proposed every year. The proposed algorithms are compared with existing ones using performance indicators. The hypervolume indicator (Zitzler and Thiele, 1998), which is also abbreviated as "hypervolume" in this article, has frequently been used for performance comparison in the EMO community while a wide variety of indicators are available in the literature (Zitzler et al., 2003). This is because no other Pareto compliant indicator is known (Zitzler et al., 2007). The hypervolume indicator has also frequently been used in indicator-based EMO algorithms such as LAHC (Knowles et al., 2003), SMS-EMOA (Emmerich et al., 2005; Beume et al., 2007), and HypE (Bader and Zitzler, 2011). In these algorithms, multi-objective optimization is handled as single-objective optimization of a solution set for hypervolume maximization.

An important research issue is the analysis of the optimal distribution of solutions for hypervolume maximization. Theoretical studies (Emmerich et al., 2007; Auger et al., 2012; Brockhoff, 2010; Friedrich et al., 2015) show that the hypervolume indicator is maximized by a set of uniformly distributed solutions on a linear Pareto front of a two-objective problem. It is also shown that such a solution set is not optimal when the Pareto front is nonlinear. This is because the optimal distribution of solutions depends on the slope of the Pareto front in the two-dimensional objective space. Except for the case of single-dimensional degenerate Pareto fronts (Shukla et al., 2014), no optimal distribution of solutions has been theoretically derived for multi-objective problems with three or more objectives. General theoretical analysis of the optimal distribution of solutions for hypervolume maximization is difficult for multi-objective problems with three or more objectives.

Selection of a prespecified number of solutions from a given solution set has been actively studied for hypervolume maximization under the name of hypervolume subset selection. This is a single-objective combinatorial optimization problem. Efficient subset selection methods have been proposed in the literature (Basseur et al., 2016; Bringmann et al., 2014; Guerreiro et al., 2015, 2016; Kuhn et al., 2016). Hypervolume subset selection can be also viewed as a preprocessing procedure for selecting a small number of solutions to be presented to the decision maker from a large number of solutions obtained by EMO algorithms (Ishibuchi et al., 2014). Recently, such a preprocessing procedure has been used for performance comparison of EMO algorithms with an unbounded archive population (Ishibuchi et al., 2016; Tanabe et al., 2017).

A reference point is needed to calculate the hypervolume indicator of a given solution set. For two-objective problems, the effect of the reference point specification on the optimal distribution of solutions for hypervolume maximization has been theoretically studied (Auger et al., 2012; Brockhoff, 2010; Friedrich et al., 2015). For three-objective problems with linear Pareto fronts, the optimal distribution of solutions has been empirically studied through a number of computation experiments (Ishibuchi et al., 2017a). It is also shown through computational experiments that performance comparison results among different EMO algorithms depend on the location of the reference point (Ishibuchi et al., 2015; Ishibuchi, Setoguchi et al., 2017). However, to the best of our

---

This article is an extended version of our conference paper (Ishibuchi et al., 2017b), which includes many new experiments, explanations, and discussions on the reference point specification for multi-objective problems with nonlinear Pareto fronts and also for many-objective problems. The visualization method of hypervolume contributions is also improved from the conference paper (Ishibuchi et al., 2017b).

knowledge, the reference point specification has not been studied in detail from a viewpoint of fair performance comparison of EMO algorithms, especially for many-objective problems, except for our former conference paper (Ishibuchi et al., 2017b) where it was discussed for two multi-objective test problems with linear Pareto fronts (i.e., DTLZ1 and inverted DTLZ1). This article is an extended version of our former conference paper. We discuss the reference point specification not only for multi-objective test problems with linear Pareto fronts but also for those with non-linear Pareto fronts (i.e., DTLZ2 and Minus-DTLZ2), a multi-objective knapsack problem and a multi-objective car-side impact problem. A visualization method of hypervolume contributions is improved from the conference paper. Moreover, detailed discussions on the reference point specification in SMS-EMOA (Emmerich et al., 2005; Beume et al., 2007) are added.

A general guideline for the reference point specification is to use a slightly worse point than the nadir point so that the reference point is dominated by all Pareto optimal solutions (i.e., so that all Pareto optimal solutions in a solution set have positive hypervolume contributions). However, there exists no specific numerical guideline. As a result, different specifications were used in different studies as pointed out by Chugh et al. (2018). For example, the reference point was specified for DTLZ1 (Deb et al., 2002) with the ideal point  $(0.0, \dots, 0.0)$  and the nadir point  $(0.5, \dots, 0.5)$  as follows: 1% larger than the nadir point, i.e.,  $(0.505, \dots, 0.505)$  in Seada and Deb (2016), 10% larger in Yuan, Xu, Wang, and Yao (2016) and Yuan, Xu, Wang, Zhang et al. (2016) and Maltese et al. (2018), 40% larger in Wagner et al. (2007), and 100% larger in Li et al. (2015). In Zhang et al. (2018), the nadir point was used as the reference point. These observations raise the following two questions: "How sensitive are the performance comparison results to the reference point specification?" and "How should we specify the reference point for fair performance comparison?"

The goal of this article is to answer these two questions. First, we discuss the optimal distribution of solutions for hypervolume maximization. We clearly demonstrate that the sensitivity of the optimal distribution of solutions to the location of the reference point is problem-dependent. The effect of the reference point is minor when a multi-objective problem has a triangular Pareto front such as DTLZ1-4 (Deb et al., 2002) and WFG4-9 (Huband et al., 2006). However, the reference point specification has a large effect on the optimal distribution of solutions when a multi-objective problem has an inverted triangular Pareto front such as Minus-DTLZ1-4 and Minus-WFG4-9 (Ishibuchi, Setoguchi et al., 2017). We discuss why the reference point specification has a different effect depending on the shape of the Pareto front. Then, we propose a reference point specification method for fair performance comparison of EMO algorithms for multi-objective and many-objective problems. Our idea is based on the following two assumptions with respect to the performance evaluation of solution sets by the hypervolume indicator.

1. A uniformly distributed solution set over an entire linear Pareto front should be highly evaluated by the hypervolume indicator. That is, such a solution set should have a large hypervolume.
2. All solutions from such a solution set should have similar hypervolume contributions. That is, no solution should have a dominant or too small effect on hypervolume calculation.

This article is organized as follows. In Section 2, we examine the optimal distribution of solutions for hypervolume maximization using SMS-EMOA (Emmerich et al., 2005; Beume et al., 2007) for DTLZ1-4 and Minus-DTLZ1-4. It is clearly demonstrated

that the optimal distribution of solutions strongly depends on the location of the reference point when a multi-objective problem has an inverted triangular Pareto front. In Section 3, we propose a reference point specification method for fair performance comparison based on the above-mentioned two assumptions and the theoretical studies on the optimal distribution of solutions for a two-objective problem with a linear Pareto front. In Section 4, we examine whether the proposed method can appropriately specify a reference point for fair performance comparison of EMO algorithms through computational experiments on various test problems. In Section 5, we examine the usefulness of the proposed method in SMS-EMOA through computational experiments. Finally, we conclude this article in Section 6.

## 2 Optimal Distribution of Solutions for Hypervolume Maximization: Empirical Discussions

In the EMO community, we implicitly assume that a set of well-distributed solutions over the entire Pareto front of a multi-objective problem is a good solution set. We also implicitly assume that such a solution set is highly evaluated by the hypervolume indicator and the IGD (inverted generational distance indicator). Thus these two indicators have frequently been used for performance comparison in the EMO community. However, we do not know the optimal distribution of solutions with respect to each indicator except for some special cases such as two-objective problems. That is, we are comparing different EMO algorithms using these indicators without knowing which solution set of a fixed size (what type of solution sets of a fixed size) is most highly evaluated by each indicator. If a set of solutions on a particular region of the Pareto front is evaluated as being better than a set of well-distributed solutions over the entire Pareto front, performance comparison results may be inconsistent with our intuition (i.e., they may be unreliable). In this sense, it is important to know the optimal distribution of solutions with respect to each indicator. Actually, a set of well-distributed solutions over the entire Pareto front is not always the best solution set. For example, it is reported in a recent study (Ishibuchi et al., 2018) that the best solution set for IGD minimization has much smaller diversity than a fully expanded solution set over the entire Pareto front when a large number of uniformly distributed reference points are used for IGD calculation for many-objective problems.

In this section, we examine the optimal distribution of solutions for hypervolume maximization using SMS-EMOA (Emmerich et al., 2005; Beume et al., 2007) for DTLZ1-4 (Deb et al., 2002) and Minus-DTLZ1-4 (Ishibuchi, Setoguchi et al., 2017). The objective space of each test problem is normalized for uniform visualization so that the ideal and nadir points are  $(0, \dots, 0)$  and  $(1, \dots, 1)$ , respectively. Since DTLZ2-4 have the same Pareto front, they have the same optimal distribution of solutions. For the same reason, Minus-DTLZ2-4 have the same optimal distribution of solutions.

In order to focus on the search for the optimal distribution of solutions on the Pareto front, the number of distance variables in each test problem is specified as zero in our computational experiments in this section. Under this setting, all feasible solutions of each test problem are Pareto optimal. SMS-EMOA with the SBX crossover (crossover probability: 1.0, distribution index: 20) and the polynomial mutation (mutation probability:  $1/n$  where  $n$  is the number of decision variables, distribution index: 20) is applied to each test problem using the following parameter specifications: The population size is 50, and the termination condition is 100,000 generations. Five settings of the reference point  $r = (r, \dots, r)$  are examined in SMS-EMOA:  $r = 1.0$  (the same as the nadir point),  $r = 1.1$  (10% larger than the nadir point),  $r = 1.5$  (50% larger),  $r = 2.0$  (100% larger),

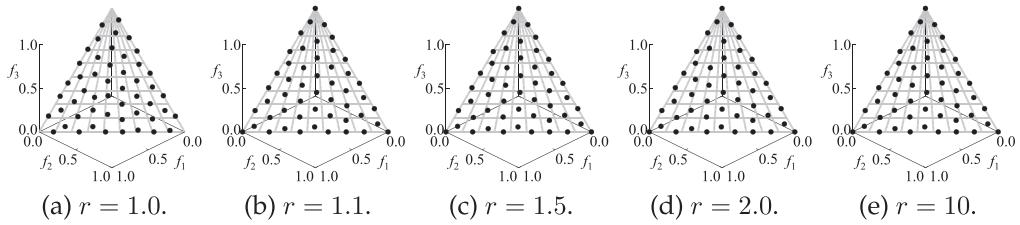


Figure 1: Obtained solution sets for the three-objective normalized DTLZ1.

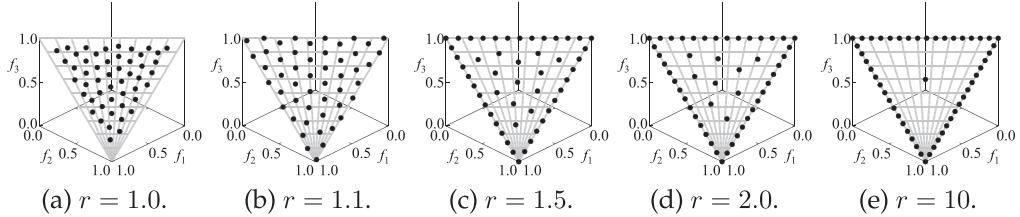


Figure 2: Obtained solution sets for the three-objective normalized Minus-DTLZ1.

and  $r = 10$  (ten times larger). During the execution of SMS-EMOA, the reference point is fixed at the pre-specified point (i.e., one of the five settings). For each setting of the reference point, SMS-EMOA is executed five times. In this manner, we obtained 25 solution sets. Each solution set is evaluated using the hypervolume indicator with each setting of the reference point. Then, the best solution set with respect to the hypervolume indicator is selected among the 25 solution sets as the final result for each setting of the reference point.

## 2.1 Three-Objective DTLZ1 and Minus-DTLZ1 Test Problems

The obtained solution set for each setting of the reference point is shown in Figure 1 for the three-objective normalized DTLZ1. When the nadir point is used as the reference point in Figure 1a, the three extreme solutions of the Pareto front are not included in the best solution set. In Figures 1b–e with  $r = 1.1, 1.5, 2.0, 10$ , the same solution set is selected independent of the reference point specification. The solution set in Figure 1a with  $r = 1.0$  is similar to the solution set in Figures 1b–e. These observations suggest that a set of well-distributed solutions over the entire Pareto front is highly evaluated by the hypervolume indicator, independent of the reference point specification. As a result, it is likely that similar results are obtained from hypervolume-based comparison, independent of the reference point specification for the three-objective DTLZ1.

In Figure 2, we show the obtained solution sets for the three-objective normalized Minus-DTLZ1. When the nadir point is used as the reference point in Figure 2a, all solutions are inside the Pareto front. On the contrary, when  $r = 10$  (i.e., when the reference point is far away from the nadir point) in Figure 2e, almost all solutions are on the sides of the Pareto front. Only a single inside solution is included in the best solution set in Figure 2e. By increasing the distance of the reference point from the nadir point, the optimal distribution of solutions is biased toward the sides of the Pareto front. Figure 2 clearly shows that the optimal distribution of solutions strongly depends on the location of the reference point. This means that hypervolume-based comparison results for



Figure 3: Hypervolume calculation from the nadir point for the three-objective normalized DTLZ1.



Figure 4: Hypervolume calculation from the nadir point for the three-objective normalized Minus-DTLZ1.

the three-objective Minus-DTLZ1 strongly depend on the reference point specification. Different specifications of the reference point may lead to different comparison results of EMO algorithms by the hypervolume indicator.

The question is why our simulation results in Figures 1 and 2 are totally different from each other. First, we explain the reason for the difference between Figures 1a and 2a where the nadir point is used as the reference point (i.e.,  $r = 1.0$ ). In this case, the hypervolume calculation is performed within the triangular pyramid stretched from the nadir point  $(1, 1, 1)$  as shown in Figures 3 and 4. In Figure 3, the triangular Pareto front of DTLZ1 is inscribed in the green inverted triangle, which is the base of the triangular pyramid for the hypervolume calculation. Only the three extreme solutions of the Pareto front are on the sides of the green inverted triangle (i.e., only these extreme solutions have zero hypervolume contribution). Thus, Pareto optimal solutions on the sides of the Pareto front have positive hypervolume contributions. As a result, some solutions on the sides of the Pareto front are included in the best solution set in Figure 1a. However, in Figure 4, the green inverted triangle perfectly overlaps the inverted triangular Pareto front of Minus-DTLZ1. This means that any solution on the sides of the Pareto front cannot have a positive hypervolume contribution. As a result, all solutions in the best solution set in Figure 2a are inside the Pareto front.

Next, let us discuss the difference between Figures 1e and 2e when the reference point is far away from the nadir point. This difference can be explained by projecting two types of Pareto fronts onto a two-dimensional subspace as shown in Figure 5. In Figure 5a, the triangular Pareto front of DTLZ1 is projected to the red triangle in the  $f_2-f_3$  space. Only the extreme solution  $(1, 0, 0)$  projected to  $(0, 0)$  is nondominated in the  $f_2-f_3$  space. This means that only this solution has a positive hypervolume contribution in the  $f_2-f_3$  space. By moving the reference point away from the Pareto front along the  $f_1$

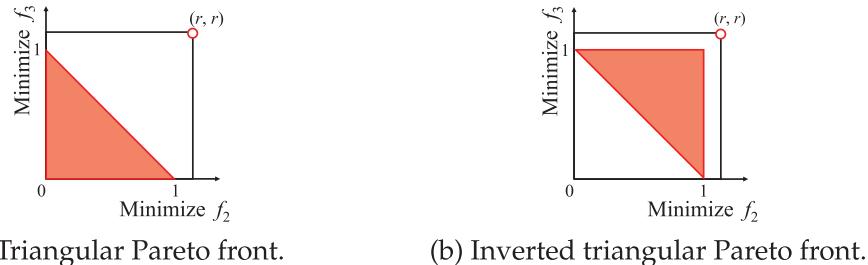
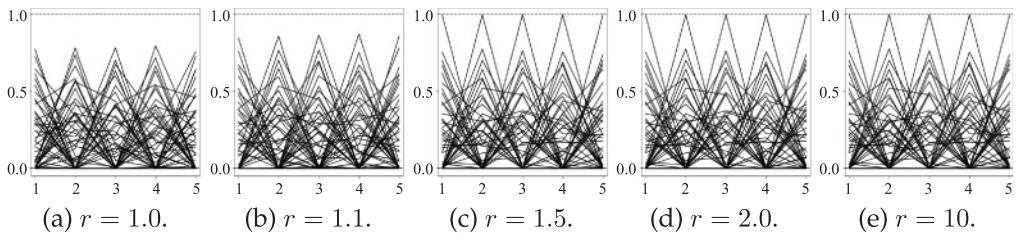
Figure 5: Projection of the Pareto front to the  $f_2$ - $f_3$  space.

Figure 6: Obtained solution sets for the five-objective normalized DTLZ1 problem.

axis, the hypervolume contribution of only this solution increases in the original three-dimensional objective space. From the same discussion in the  $f_1$ - $f_2$  and  $f_1$ - $f_3$  spaces, we can see that the hypervolume contributions of the three extreme solutions increase by moving the reference point far away from the nadir point. The location of the reference point has no effect on the hypervolume contributions of all the other solutions when the three extreme solutions are included in the solution set. Thus the same best solution set is obtained in Figures 1b–e.

In Figure 5b, the inverted triangular Pareto front of Minus-DTLZ1 is projected to the red triangle in the  $f_2$ - $f_3$  space. All solutions on the line between  $(1, 0)$  and  $(0, 1)$  are nondominated in the  $f_2$ - $f_3$  space. This means that all solutions on this line have positive hypervolume contributions in the  $f_2$ - $f_3$  space. By moving the reference point away from the Pareto front along the  $f_1$  axis, their hypervolume contributions monotonically increase in the original three-dimensional objective space. From the same discussion in the  $f_1$ - $f_2$  and  $f_1$ - $f_3$  spaces, we can see that the hypervolume contributions of all solutions on the sides of the inverted triangular Pareto front monotonically increase with the increase in the distance of the reference point from the nadir point. As a result, more solutions on the sides of the Pareto front are included in the best solution set in Figure 2. When  $r$  is very large, almost all solutions in the best solution set are on the sides of the Pareto front as in Figure 2e. In Section 3, we will explain these discussions about the relation between the location of the reference point and the optimal solution set through visualization of the hypervolume contribution of each solution.

## 2.2 Five-Objective DTLZ1 and Minus-DTLZ1 Test Problems

Experimental results on the five-objective normalized DTLZ1 and Minus-DTLZ1 are shown in Figures 6 and 7, respectively. In Figure 6, the same solution set is obtained for the three settings of the reference point ( $r = 1.5, 2.0, 10$ ) for the five-objective

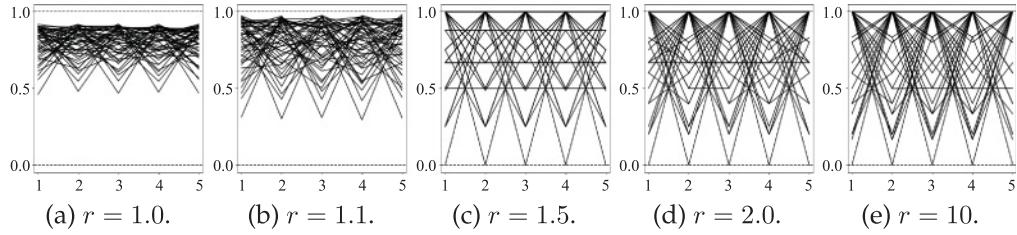


Figure 7: Obtained solution sets for the five-objective Minus-DTLZ1 problem.

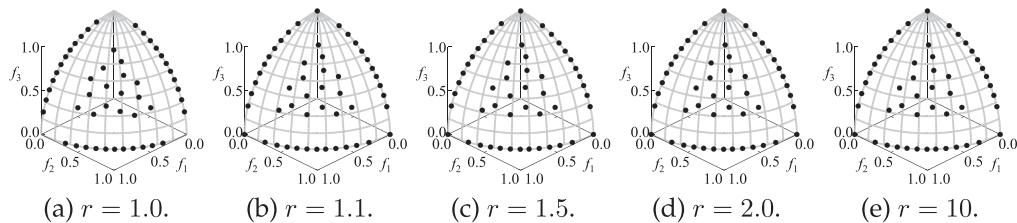


Figure 8: Obtained solution sets for the three-objective DTLZ2 problem.

normalized DTLZ1. This is because the location of the reference point has an effect only on the hypervolume contributions of the five extreme solutions of the five-objective normalized DTLZ1. When  $r = 1.0$  (i.e., when the nadir point is used as the reference point), the best solution set does not cover the entire Pareto front. That is, the diversity of the best solution set in Figure 6a is much smaller than that in Figures 6c–e. The best solution set for  $r = 1.1$  does not cover the entire Pareto front, either. That is,  $r = 1.1$  is too small for the five-objective normalized DTLZ1 whereas it is a good setting for the three-objective normalized DTLZ1. This observation suggests that the appropriate specification of the reference point depends on the number of objectives. This issue will be further discussed in Section 3.

In Figure 7, a different solution set is obtained for each setting of the reference point for the five-objective Minus-DTLZ1. When  $r = 1.0$ , the diversity of the obtained solution set is very small in Figure 7a. Almost all solutions are close to the center of the Pareto front ( $0.8, 0.8, 0.8, 0.8, 0.8$ ). When  $r = 1.1$ , the diversity is still small in Figure 7b whereas it is larger than that in Figure 7a. However, when  $r = 10$ , no solutions are close to the center of the Pareto front in Figure 7e. Our experimental results in Figures 6 and 7 suggest that the reference point specification is much more important for Minus-DTLZ1 in Figure 7 than DTLZ1 in Figure 6. This is consistent with our observation in Figures 1 and 2 for the three-objective normalized DTLZ1 and Minus-DTLZ1.

### 2.3 Three-Objective DTLZ2 and Minus-DTLZ2 Test Problems

Whereas DTLZ1 and Minus-DTLZ1 have linear Pareto fronts, the Pareto fronts of DTLZ2 and Minus-DTLZ2 are non-linear. Experimental results on the three-objective DTLZ2 and normalized Minus-DTLZ2 are shown in Figures 8 and 9, respectively. As in Figures 1 and 6 for DTLZ1, the location of the reference point does not have a large effect on the optimal distribution of solutions in Figure 8 for DTLZ2. This is because DTLZ2 has a triangular Pareto front. In Figure 9 for Minus-DTLZ2, the location of the

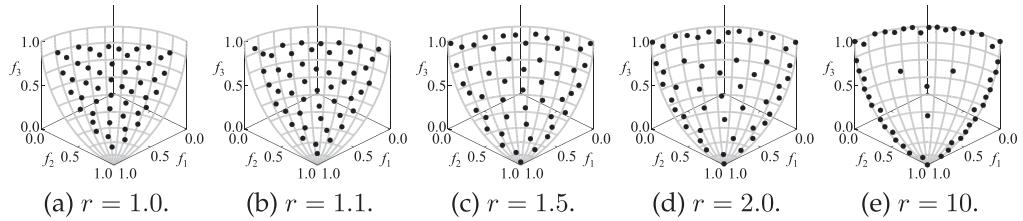


Figure 9: Obtained solution sets for the three-objective Minus-DTLZ2 problem.

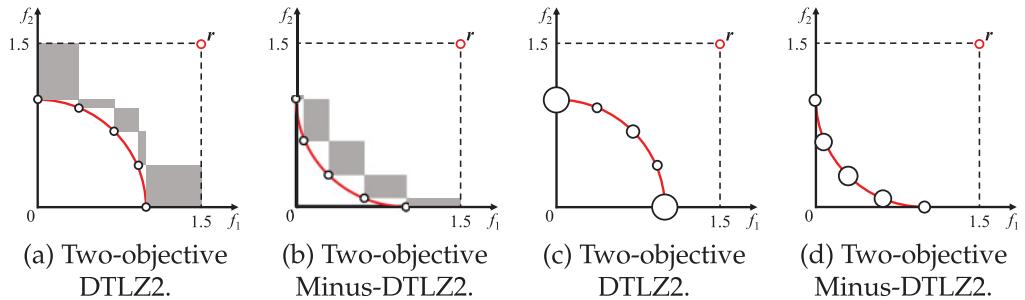


Figure 10: Hypervolume contribution of each of the five uniformly distributed solutions.

reference point has a large effect on the optimal distribution of solutions. This is because Minus-DTLZ2 has an inverted Triangular Pareto front. However, the optimal distributions of solutions in Figure 9 for Minus-DTLZ2 are clearly different from Figure 2 for Minus-DTLZ1. For example, in Figure 2e, almost all solutions are on the sides of the inverted triangular Pareto front. However, most solutions are not on the sides of the inverted triangular Pareto front in Figure 9e whereas they are close to the sides. This is related to the nonlinear shape of the Pareto front of Minus-DTLZ2.

As explained in Auger et al. (2012) for a nonlinear Pareto front of a two-objective problem, the two extreme solutions of the Pareto front are not always included in the optimal distribution of solutions even when the reference point is far away from the Pareto front. In Figure 10, we show the hypervolume contribution of each of the five uniformly distributed solutions on the entire Pareto fronts of the two-objective DTLZ2 and normalized Minus-DTLZ2. The hypervolume contribution of each solution is shown by the corresponding shaded rectangle in Figures 10a and b, and by the size of the corresponding open circle in Figures 10c and d. The size of each open circle in Figures 10c and d is proportional to the hypervolume contribution of the corresponding solution. This hypervolume visualization method is used for three-objective problems later in this article. When the reference point is far away from the Pareto front, the two extreme solutions of DTLZ2 have large hypervolume contributions as shown in Figures 10a and c. Thus they are included in the optimal distribution of solutions of DTLZ2. However, the two extreme solutions of Minus-DTLZ2 have small hypervolume contributions as shown in Figures 10b and d. Thus they are not included in the optimal distribution of solutions of Minus-DTLZ2. Similar explanations hold for the three-objective normalized Minus-DTLZ2 in Figure 9. Thus most solutions in the obtained best solution set

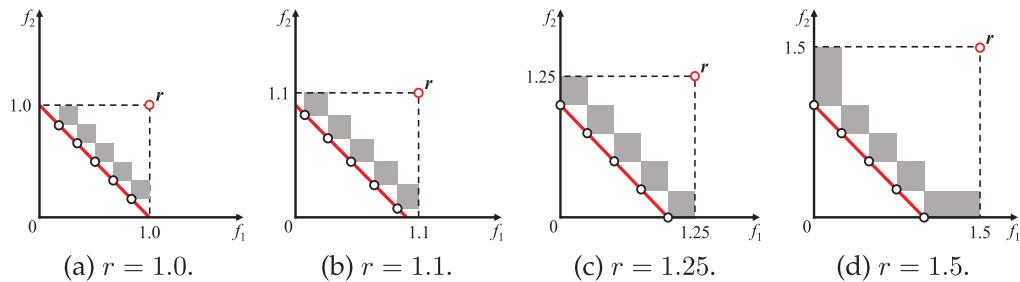


Figure 11: Optimal distributions of solutions for hypervolume maximization for a linear Pareto front.

are inside the Pareto front (whereas they are close to the sides of the Pareto front) even when the reference point is far away from the Pareto front.

In this section, we demonstrated that the optimal distribution of solutions for hypervolume maximization depends on the location of the reference point, the curvature property of the Pareto front (linear, convex, or concave) and the Pareto front shape (triangular or inverted triangular). Among the six combinations of the curvature property and the Pareto front shape, four combinations are examined using DTLZ1 (linear triangular), Minus-DTLZ1 (linear inverted triangular), DTLZ2 (concave triangular), and Minus-DTLZ2 (concave inverted triangular).

### 3 Reference Point Specification Method

In this section, we propose a reference point specification method in hypervolume calculation for fair performance comparison of EMO algorithms. First, we derive a simple specification method for two-objective problems from theoretical analysis of the optimal distribution of solutions on the linear Pareto front of a normalized two-objective problem. Next, the specification method is generalized for multi-objective problems. Then, its validity is examined through computational experiments.

#### 3.1 Basic Idea for Two-Objective Problems: Theoretical Analysis

Let us consider a two-objective minimization problem with a linear Pareto front, which is a line between  $(0, 1)$  and  $(1, 0)$  in the two-dimensional objective space. For this problem, the hypervolume indicator is maximized by a set of equidistant points on the Pareto front (Emmerich et al., 2007; Auger et al., 2012). The optimal distribution of five solutions is illustrated in Figure 11 for different settings of the reference point  $r = (r, r)$ :  $r = 1.0, 1.1, 1.25, 1.5$ . Each shaded region in Figure 11 shows the hypervolume contribution of the corresponding solution. As illustrated in Figure 11a with  $r = 1.0$ , the two extreme solutions of the Pareto front are not included in the optimal solution set when the nadir point is used as the reference point. When  $r = 1.1$  in Figure 11b, they are not included in the optimal solution set either. The two extreme solutions are included in the optimal solution set only when the reference point satisfies the inequality condition  $r \geq 1 + 1/(\mu - 1)$ , where  $\mu$  is the number of solutions. This condition has already been proven in Theorem 3 of Brockhoff (2010). In Figure 11 with  $\mu = 5$ , this condition means  $r \geq 1.25$ . Thus the two extreme solutions are included in Figures 11c and d.

In Figure 12, we show the hypervolume contribution of each of the five uniformly distributed solutions. When the nadir point is used as the reference point in Figure 12a, the two extreme solutions have no hypervolume contribution. This means that they

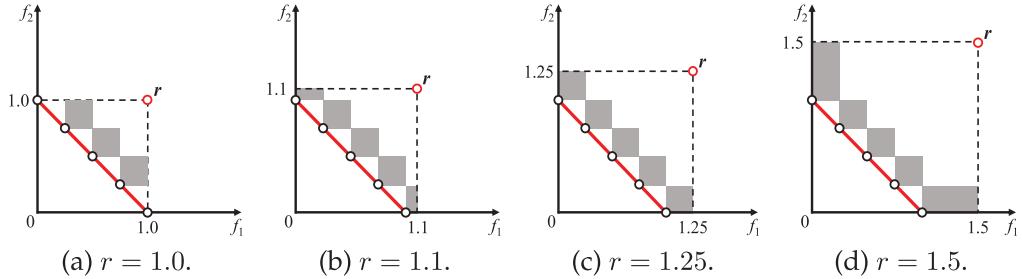


Figure 12: Hypervolume contribution of each of the uniformly distributed solutions on a linear Pareto front.

have no effect on hypervolume-based performance comparison results. This is not desirable for fair performance comparison. When  $r = 1.1$  in Figure 12b, the two extreme solutions have much smaller hypervolume contributions than the other three solutions. However, when  $r = 1.5$  in Figure 12d, they have much larger hypervolume contributions than the other three solutions. All of these settings are not desirable from a viewpoint of fair performance comparison since some solutions in the solution set have much smaller or much larger effects on hypervolume calculation (i.e., on hypervolume-based performance comparison results). Only when  $r = 1.25$  (i.e.,  $r = 1 + 1/(\mu - 1)$ ), all solutions in Figure 12 have the same hypervolume contribution. Based on these discussions, we propose to specify the reference point  $r = (r, r)$  as  $r = 1 + 1/(\mu - 1)$  for two-objective problems.

The optimal distribution of solutions has been actively studied theoretically for the case of two objectives. For example, an interesting relation between the optimal distribution and the hypervolume contribution was proven by Bader (2009): Each solution in the optimal distribution for hypervolume maximization has the same hypervolume contribution when the number of solutions increases to infinity (Theorem 3.26). It was shown by Auger et al. (2012) that the density of solutions in the optimal distribution depends only on the slope of the Pareto front. Generalization of the hypervolume indicator was also proposed. For example, cone-based hypervolume indicators were proposed in Emmerich et al. (2013) where the optimal distribution of solutions depends on an angle parameter of cones. When the angle parameter is small (i.e., cones are acute), solutions are more evenly distributed on convex and concave Pareto fronts than the case of the standard hypervolume indicator. Reference point free weighted hypervolume indicators were proposed in Emmerich et al. (2014) where it was shown that the standard hypervolume indicator belongs to a more general class of Pareto compliant measures. Whereas those generalized versions were proposed, we focus on the standard hypervolume indicator since it has frequently been used for performance comparison of EMO algorithms in the literature.

### 3.2 Generalization to Multi-Objective Problems: Proposed Method

In Figure 13, we show 15 uniformly distributed solutions on the entire Pareto front of the three-objective normalized DTLZ1. In this article, uniformly distributed solutions are generated by the simplex lattice-based method in the same manner as MOEA/D for generating uniformly distributed weight vectors (whereas other methods may generate other distributions, e.g., simplex centroid). In Figure 13, each side of the triangular Pareto front is divided into four intervals as in Figures 11c and 12c with five solutions on

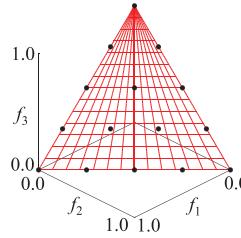


Figure 13: Fifteen uniformly distributed solutions for the three-objective normalized DTLZ1 problem.

the linear Pareto front. Our idea is to apply the reference point specification method in the previous subsection to the sides of the triangular Pareto front. Since each side is divided into four intervals in Figure 13 as in Figures 11c and 12c, we propose to use  $(1.25, 1.25, 1.25)$  as the reference point in Figure 13. In general, let us assume that each side of the triangular Pareto front of the normalized  $m$ -objective DTLZ1 problem is divided into  $H$  intervals. Then, we propose to specify the reference point  $\mathbf{r} = (r, r, \dots, r)$  as  $r = 1 + 1/H$ . In the case of two-objective problems,  $H$  is calculated from the number of solutions  $\mu$  as  $H = \mu - 1$ . Thus the proposed specification method by  $r = 1 + 1/H$  can be viewed as a generalization of our basic idea (i.e.,  $r = 1 + 1/(\mu - 1)$ ) for two-objective problems in the previous subsection.

In MOEA/D (Zhang and Li, 2007), the population size  $\mu$  is defined for an  $m$ -objective problem using a user-defined integer parameter  $H$  as  $\mu = C_{m-1}^{H+m-1}$  where  $C_m^n$  is the number of  $m$ -combinations from a set of  $n$  elements, that is,  $C_m^n = n!/m!(n-m)!$ . The integer parameter  $H$  in MOEA/D corresponds to the number of intervals of each side of the triangular Pareto front (e.g.,  $H = 4$  in Figure 13). When the population size  $\mu$  is specified by this formulation, we can directly use our formulation (i.e.,  $r = 1 + 1/H$ ) for specifying the reference point  $\mathbf{r} = (r, r, \dots, r)$ . However, in general, the population size  $\mu$  is not always specified as  $\mu = C_{m-1}^{H+m-1}$ . For handling such a general case, first we specify the value of  $H$  using the following formulation:

$$C_{m-1}^{H+m-1} \leq \mu < C_{m-1}^{H+m}, \quad (1)$$

where  $\mu$  is the population size and  $m$  is the number of objectives. Then, the reference point  $\mathbf{r} = (r, r, \dots, r)$  is specified in the normalized objective space with the ideal point  $(0, 0, \dots, 0)$  and the nadir point  $(1, 1, \dots, 1)$  as follows:

$$r = 1 + 1/H. \quad (2)$$

Our proposal is to use (1) and (2) for specifying the reference point  $\mathbf{r} = (r, r, \dots, r)$  in the normalized objective space from the population size  $\mu$  and the number of objectives  $m$ .

### 3.3 Numerical Examples: Hypervolume Contribution of Each Solution on Linear Pareto Fronts

In this subsection, we examine whether each of the uniformly distributed solutions over the entire Pareto front has a similar hypervolume contribution when the reference point  $\mathbf{r} = (r, r, \dots, r)$  is specified by the proposed method. In Figure 14, we generate 66 solutions for the three-objective normalized DTLZ1 using the weight vector generation method in MOEA/D with  $H = 10$ . The suggested reference point by the proposed

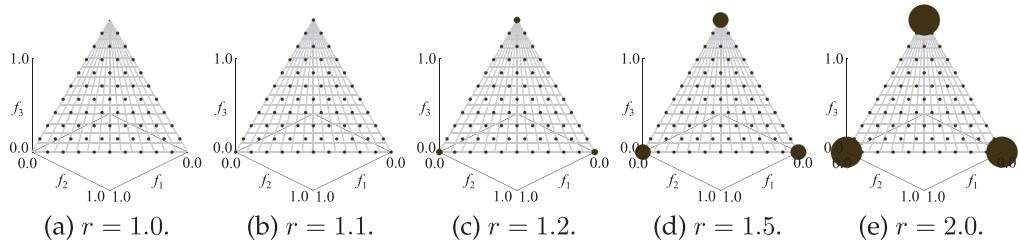


Figure 14: Hypervolume contribution of each solution for the three-objective normalized DTLZ1 ( $H = 10$ ).

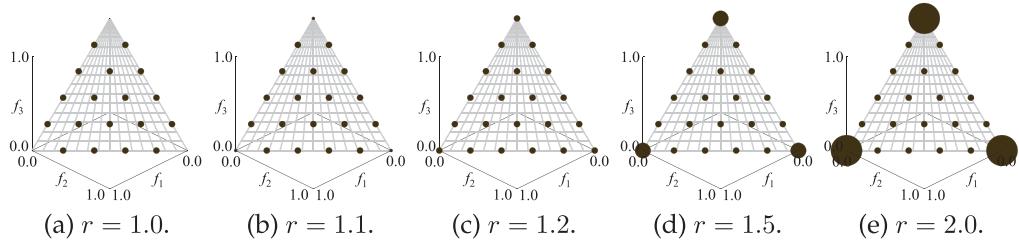


Figure 15: Hypervolume contribution of each solution for the three-objective normalized DTLZ1 ( $H = 5$ ).

method is  $r = 1.1$ . The hypervolume contribution of each solution is shown by the size of the corresponding circle. That is, the area of each circle is proportional to the hypervolume contribution of the corresponding solution. In this subsection, the size of each circle is normalized so that the largest circle in each figure has the same maximum radius. So, we can compare the size of each circle within each figure (e.g., between a circle in Figure 14a and another one in Figure 14e). However, we cannot compare the size of each circle between different figures (e.g., between Figures 14 and 15).

As we have already explained, the hypervolume contributions of the three extreme solutions are zero when the nadir point is used as the reference point (i.e.,  $r = 1.0$ ) in Figure 14a. When  $r$  is large (e.g.,  $r = 2.0$ ), those three extreme solutions have much larger hypervolume contributions than the other solutions. Only when the reference point is specified by the proposed method as  $r = 1.1$ , all the 66 solutions have the same hypervolume contribution.

For examining the effect of the number of solutions on the appropriate reference point specification, we also generate 21 solutions using the weight vector generation method in MOEA/D with  $H = 5$ . The suggested reference point by the proposed method is  $r = 1.2$ . The hypervolume contribution of each solution is shown in Figure 15 in the same manner as in Figure 14 (i.e., the area of each circle is proportional to the hypervolume contribution of the corresponding solution). As we have already explained, the size of each circle is normalized in each figure. So, the three extreme solutions in Figures 14e and 15e have the largest circles of the same size. However, their hypervolume contributions in Figure 15 are larger than those in Figure 14.

When  $r = 1.1$  in Figure 15b, the three extreme solutions have much smaller hypervolume contributions than the other solutions. This observation suggests that the

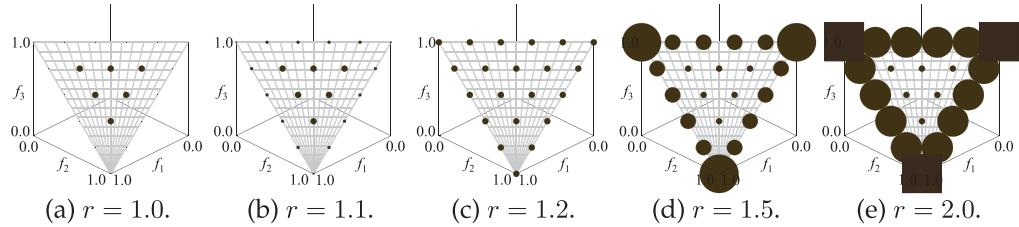


Figure 16: Hypervolume contribution of each solution for the three-objective normalized Minus-DTLZ1 ( $H = 5$ ).

reference point is too close to the Pareto front. When  $r = 1.5$  and  $r = 2.0$ , the three extreme solutions have much larger hypervolume contributions than the other solutions. This observation suggests that those specifications are too far from the Pareto front. Only when the reference point is specified by the proposed method as  $r = 1.2$ , all the 21 solutions have the same hypervolume contribution.

We also perform the same computational experiment for the three-objective normalized Minus-DTLZ1 for the case of  $H = 5$  (i.e., 21 solutions). Experimental results are shown in Figure 16. Only in this figure, the three extreme solutions which are shown by squares (instead of circles) in Figure 16e are excluded from the normalization of the size of each circle. This is because they have too large hypervolume contributions to be visualized in the same figure.

Minus-DTLZ1 has an inverted triangular Pareto front. When the nadir point is used as the reference point in Figure 16a, all solutions on the sides of the Pareto front have no hypervolume contribution. When  $r = 1.1$ , those solutions have much smaller hypervolume contribution than the other solutions inside the Pareto front. Only when we specify the reference point as  $r = 1.2$  using the proposed method, all the 21 solutions have the same hypervolume contribution. By increasing the value of  $r$  (i.e., by moving the reference point away from the Pareto front), the hypervolume contributions of the three extreme points increase as in the case of the triangular Pareto front in Figure 15. However, at the same time, the hypervolume contributions of the other solutions on the sides of the inverted triangular Pareto front also increase in Figure 16. This is the reason why almost all solutions in the best solution sets are on the sides of the inverted triangular Pareto front when the reference point is far away from the Pareto front (e.g., see Figure 9e). It should be noted that the hypervolume contributions of those solutions do not change by moving the location of the reference point in the case of the triangular Pareto front as shown in Figure 15. That is, when the Pareto front is triangular, the location of the reference point has an effect only on the hypervolume contributions of the extreme solutions.

To examine the effect of the number of objectives on the appropriate reference point specification, we generate 15 solutions for the five-objective normalized DTLZ1 using the weight vector generation method in MOEA/D with  $H = 2$ . The suggested reference point by the proposed method is  $r = 1.5$ . The hypervolume contribution of each solution is shown by the line thickness in Figure 17. When  $r = 1.0$ , the five extreme solutions with  $f_i = 1$  for one of the five objectives have no hypervolume contribution. Thus they are not depicted in Figure 17a. When  $r = 1.1$  and  $r = 1.2$ , their hypervolume contributions are much smaller than the other solutions. When the suggested reference point  $r = 1.5$  is used in Figure 17d, all solutions have the same hypervolume

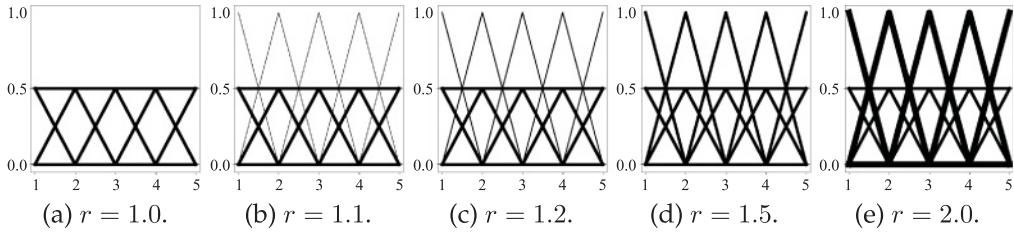


Figure 17: Hypervolume contribution of each solution for the five-objective normalized DTLZ1 ( $H = 2$ ). The hypervolume contribution is shown by the line thickness.

Table 1: Hypervolume contribution of each solution of the five-objective normalized DTLZ1 ( $\times 10^{-5}$ ).

Solution Type (Number of solutions)	$r = 1.0$	$r = 1.05$	$r = 1.1$	$r = 1.2$	$r = 1.5$
One non-zero element (5)	0.000	0.500	1.000	2.000	5.000
Two non-zero elements (90)	1.000	1.000	1.000	1.000	1.000
Three non-zero elements (360)	1.000	1.000	1.000	1.000	1.000
Four non-zero elements (420)	1.000	1.000	1.000	1.000	1.000
Five non-zero elements (126)	1.000	1.000	1.000	1.000	1.000

contribution. By increasing the value of  $r$  from  $r = 1.5$ , the five extreme solutions have larger hypervolume contributions than the other solutions as shown in Figure 17e.

We further examine the hypervolume contributions of a large number of solutions of the five-objective normalized DTLZ1. First, we generate its 1001 solutions uniformly on the triangular Pareto front using the weight vector generation mechanism of MOEA/D with  $H = 10$ . The suggested reference point by the proposed method is 1.1. The 1001 solutions (i.e., 1001 points in the five-dimensional objective space) satisfy the following relations:  $f_1 + f_2 + f_3 + f_4 + f_5 = 1$  and  $0 \leq f_i \leq 1$  for  $i = 1, 2, \dots, 5$ . Among them, 126 solutions are inside the Pareto front. That is, they have no zero element (i.e., they have five non-zero elements). All the others are boundary solutions with at least one zero element (i.e., with less than five non-zero elements). For example, the five extreme solutions have only a single non-zero element:  $(1, 0, 0, 0, 0)$ ,  $(0, 1, 0, 0, 0)$ ,  $(0, 0, 1, 0, 0)$ ,  $(0, 0, 0, 1, 0)$ , and  $(0, 0, 0, 0, 1)$ . Ninety solutions have only two non-zero elements such as  $(0.9, 0.1, 0, 0, 0)$  and  $(0, 0, 0.4, 0, 0.6)$ . The hypervolume contribution of each solution is summarized in Table 1. When the reference point is specified by the proposed method as  $r = 1.1$ , all solutions have exactly the same hypervolume contribution. The hypervolume contributions of only the five extreme solutions depend on the location of the reference point in Table 1. This is because the shape of the Pareto front is triangular as we have already explained.

In the same manner as in Table 1, we also examine the hypervolume contribution of 1001 uniformly distributed solutions on the inverted triangular Pareto front of the five-objective normalized Minus-DTLZ1. The suggested reference point by the proposed method is 1.1. The 1001 solutions satisfy the following relations:  $f_1 + f_2 + f_3 + f_4 + f_5 = 4$  and  $0 \leq f_i \leq 1$  for  $i = 1, 2, \dots, 5$ . Among the 1001 solutions, 126 are inside

Table 2: Hypervolume contribution of each solution of the five-objective normalized Minus-DTLZ1 ( $\times 10^{-5}$ ).

Solution Type (Number of solutions)	$r = 1.0$	$r = 1.05$	$r = 1.1$	$r = 1.2$	$r = 1.5$
One non-one element (5)	0.000	0.0625	1.000	16.00	625.0
Two non-one elements (90)	0.000	0.125	1.000	8.000	125.0
Three non-one elements (360)	0.000	0.250	1.000	4.000	25.00
Four non-one elements (420)	0.000	0.500	1.000	2.000	5.000
Five non-one elements (126)	1.000	1.000	1.000	1.000	1.000

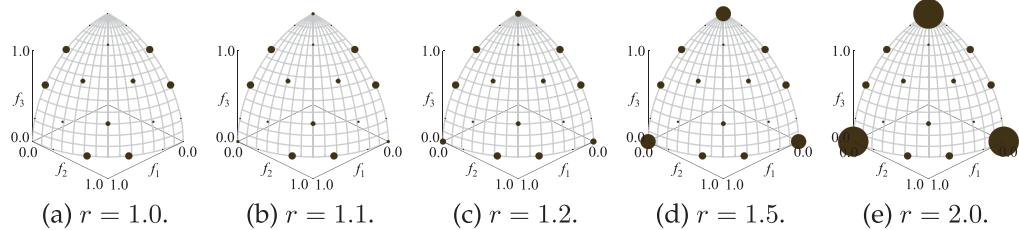


Figure 18: Hypervolume contribution of each solution for the three-objective DTLZ2 ( $H = 5$ ).

the Pareto front. Those inside solutions have five non-one elements. All the others are boundary solutions with less than five non-one elements such as the five extreme solutions with only a single non-one element:  $(0, 1, 1, 1, 1)$ ,  $(1, 0, 1, 1, 1)$ ,  $(1, 1, 0, 1, 1)$ ,  $(1, 1, 1, 0, 1)$ , and  $(1, 1, 1, 1, 0)$ . The hypervolume contribution of each solution is summarized in Table 2. When the reference point is specified by the proposed method as  $r = 1.1$ , all solutions have exactly the same hypervolume contribution. When the nadir point is used as the reference point (i.e.,  $r = 1.0$ ), only the 126 inside solutions have positive hypervolume contributions. Their hypervolume contributions are the same, independent of the location of the reference point in Table 2. The hypervolume contribution of each of the other boundary solutions increases as the value of  $r$  increases. It is interesting to observe that the increase of the hypervolume contribution is more rapid when solutions have fewer non-one elements. For example, the hypervolume contributions of the five extreme solutions with only a single non-one element drastically increase when the value of  $r$  increases from  $r = 1.1$  to  $r = 1.5$  in Table 2.

### 3.4 Numerical Examples: Hypervolume Contribution of Each Solution on Nonlinear Pareto Fronts

Experimental results in the previous subsection show that each solution in a uniformly distributed solution set on a linear Pareto front has the same hypervolume contribution when the reference point  $r = (r, r, \dots, r)$  is specified by the proposed method as  $r = 1 + 1/H$ . For comparison, we also examine the hypervolume contribution of each solution on nonlinear Pareto fronts. In Figure 18, we generate 21 solutions for the three-objective DTLZ2 using the weight vector generation method in MOEA/D with  $H = 5$ . More specifically, we calculate the best solution for each weight vector with respect to

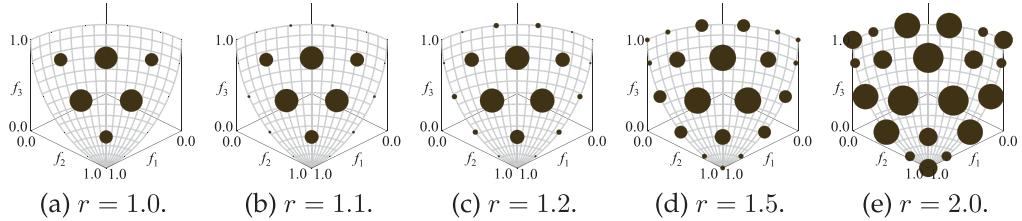


Figure 19: Hypervolume contribution of each solution for the three-objective normalized Minus-DTLZ2 ( $H = 5$ ).

the PBI function with  $\theta = \infty$  in MOEA/D. Then, we calculate the hypervolume contribution of each solution. When the nadir point is used as the reference point (i.e.,  $r = 1.0$ ) in Figure 18a, the three extreme solutions have no hypervolume contribution. When  $r = 2.0$  in Figure 18e, their hypervolume contributions are much larger than the other solutions. Due to the nonlinearity of the Pareto front, each solution has a different hypervolume contribution in Figure 18. The proposed specification  $r = 1.2$  in Figure 18c is more appropriate than  $r = 1.0$  and  $r = 2.0$  in the sense that the difference of the hypervolume contributions among the 66 solutions is small. Due to the concave shape of the Pareto front, solutions on the sides of the Pareto front have larger hypervolume contributions than inside solutions in Figure 18 as explained for the two-objective DTLZ2 in Figure 10.

We also generate 21 solutions for the three-objective normalized Minus-DTLZ2 using the inverted PBI function with  $\theta = \infty$  in MOEA/D (Sato, 2014). Then, we calculate the hypervolume contribution of each solution. When the nadir point is used as the reference point (i.e.,  $r = 1.0$ ) in Figure 19a, all solutions on the sides of the Pareto fronts have no hypervolume contribution. By increasing the distance of the reference point from the nadir point, their hypervolume contributions increase. However, due to the convex shape of the Pareto front, solutions on the sides of the Pareto front have smaller hypervolume contributions than inside solutions in Figure 19 as explained for the two-objective Minus-DTLZ2 in Figure 10.

### 3.5 Numerical Examples: Hypervolume-Based Comparison of Solution Sets

In this subsection, we examine whether a uniformly distributed solution set has a larger hypervolume than biased solution sets when the reference point is specified by the proposed method. In our computational experiments, first we generate a set of 91 uniformly distributed solutions for the three-objective normalized DTLZ1 by specifying the value of  $H$  as  $H = 12$  in the weight vector specification mechanism in MOEA/D as shown in Figure 20d. For comparison, we also generate other solution sets with 91 solutions. For generating a biased solution set around the center of the Pareto front, first we generate 136 uniformly distributed solutions for  $H = 15$ . This solution set consists of 91 inside solutions and the other 45 solutions on the sides of the Pareto front. We use a set of the 91 inside solutions for comparison as shown in Figure 20c. We also generate a further biased solution set. First, we generate 190 solutions by specifying  $H$  as  $H = 18$ . Next, we remove the 54 solutions on the sides of the Pareto front. Then, we remove the 45 solutions on the side of the remaining solution set. We use the other 91 inside solutions as a solution set for comparison as shown in Figure 20b. In the same manner, we also

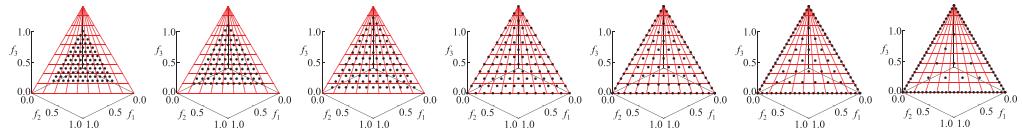
(a)  $H = 21$ . (b)  $H = 18$ . (c)  $H = 15$ . (d)  $H = 12$ . (e)  $H = 10$ . (f)  $H = 8$ . (g)  $H = 6$ .

Figure 20: Generated solution sets with 91 solutions of the three-objective normalized DTLZ1 for hypervolume-based performance comparison. The value of  $H$  shows the resolution of inside solutions. A large value of  $H$  means a dense distribution of solutions around the center of the Pareto front.

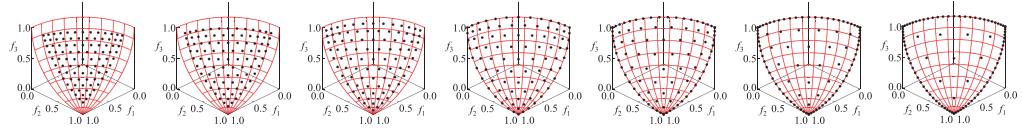
(a)  $H = 21$ . (b)  $H = 18$ . (c)  $H = 15$ . (d)  $H = 12$ . (e)  $H = 10$ . (f)  $H = 8$ . (g)  $H = 6$ .

Figure 21: Generated solution sets with 91 solutions of the three-objective normalized Minus-DTLZ2.

generate a solution set with 91 solutions selected from 253 uniformly distributed solutions for  $H = 21$  in Figure 20a.

We also generate biased solution sets toward the sides of the Pareto front in the following manner. First, we generate a set of 66 uniformly distributed solutions by specifying the value of  $H$  as  $H = 10$ . This solution set includes 36 inside solutions. In addition to those 36 inside solutions, 55 solutions are uniformly distributed over the sides of the Pareto front to generate a solution set with 91 solutions as shown in Figure 20e. In the same manner, we generate a solution set from  $H = 8$  in Figure 20f where the 21 inside solutions are generated from  $H = 8$  and the other 70 solutions are uniformly distributed over the sides of the Pareto front. We also generate a solution set with 10 inside solutions from  $H = 6$  and 81 uniformly distributed solutions on the sides of the Pareto front in Figure 20g.

In the same manner, we also generate seven solution sets with 91 solutions of the three-objective normalized Minus-DTLZ1, DTLZ2, and Minus-DTLZ2. The generated solution sets of Minus-DTLZ2 are shown in Figure 21. The generated seven solution sets for each test problem are compared by the hypervolume indicator for each of the following specifications of the reference point:  $r = 1.00, 1.02, 1.04, 1.06, 1.08, 1.10, 1.20, 1.40, 1.60, 1.80, 2.00$ . Table 3 shows the solution set with the largest hypervolume for each specification of the reference point. For DTLZ1, the well-distributed solution set in Figure 20d is evaluated as the best solution set, independent of the specification of the reference point. For DTLZ2, such a solution set is also evaluated as the best solution set, independent of the specification of the reference point. However, for Minus-DTLZ1 and Minus-DTLZ2, the hypervolume-based performance comparison results strongly depend on the specification of the reference point. The suggested reference point by the proposed method for a three-objective problem with 91 solutions (i.e.,  $m = 3$  and  $\mu = 91$ ) is  $r = 1 + 1/12 = 1.0833$ . In Table 3, the well-distributed solution set over the entire Pareto front of Minus-DTLZ1, which is the inverted distribution of Figure 20d, is

Table 3: Best solution set by the hypervolume-based comparison with each specification of the reference point.

Test Problem	Reference Point Specification ( $r$ )										
	1.01	1.02	1.04	1.06	1.08	1.1	1.2	1.4	1.6	1.8	2
DTLZ1	(d)	(d)	(d)	(d)	(d)	(d)	(d)	(d)	(d)	(d)	(d)
Minus-DTLZ1	(c)	(c)	(c)	(d)	(d)	(d)	(e)	(e)	(f)	(f)	(f)
DTLZ2	(d)	(d)	(d)	(d)	(d)	(d)	(d)	(d)	(d)	(d)	(d)
Minus-DTLZ2	(b)	(b)	(b)	(b)	(b)	(b)	(c)	(c)	(d)	(d)	(d)

evaluated as the best solution set only when  $r = 1.06, 1.08, 1.10$ . This observation supports the validity of the proposed reference point specification method. However, for Minus-DTLZ2, the solution set in Figure 21b is evaluated as the best solution set when  $r = 1.06, 1.08, 1.10$ . This observation, which is consistent with the above-mentioned experimental results on Minus-DTLZ2 with respect to the best solution set and the hypervolume contribution of each solution, suggests the necessity of special care in the hypervolume-based performance comparison for multi-objective problems with convex Pareto fronts.

## 4 Computational Experiments on Various Test Problems

We have already explained that an appropriate reference point can be obtained by the proposed method for DTLZ1, DTLZ2, and Minus-DTLZ1. We have also demonstrated that the location of the reference point has a large effect on the optimal distribution of solutions and performance comparison results through computational experiments on Minus-DTLZ1 and Minus-DTLZ2 with inverted triangular Pareto fronts. However, the effect of the location of the reference point is small when test problems have triangular Pareto fronts (e.g., DTLZ1-4 and WFG4-9). In this section, we examine the effect of the reference point specification using other frequently-used three-objective test problems: knapsack problem, car-side impact problem, and WFG3. The number of solutions is specified as 50. For  $m = 3$  (three objectives) and  $\mu = 50$  (50 solutions), the suggested reference point by the proposed method is calculated as  $r = 1.125$  since  $H$  is determined as  $H = 8$  from Eq. (1). The three test problems in this section are normalized as three-objective minimization problems with the ideal point  $(0, 0, 0)$  and the nadir point  $(1, 1, 1)$ .

### 4.1 Three-Objective Knapsack Problem

Multi-objective knapsack problems have frequently been used as combinatorial multi-objective test problems in the EMO community (Zitzler and Thiele, 1999; Zhang and Li, 2007; Ishibuchi et al., 2015). As a test problem, we use the three-objective 500-item knapsack problem in Ishibuchi et al. (2015). First, we search for an approximate Pareto front by MOEA/D with Tchebycheff function (Zhang and Li, 2007) using a large computation load (population size: 10,011, termination condition: 1,000 generations, execution: ten times). All solutions obtained from ten runs of MOEA/D are merged into a single solution set. Next, we select only nondominated solutions from the merged solution set. In this manner, we obtain an approximate Pareto front with 31,509 non-dominated solutions in Figure 22a where the three-dimensional objective space is normalized as a

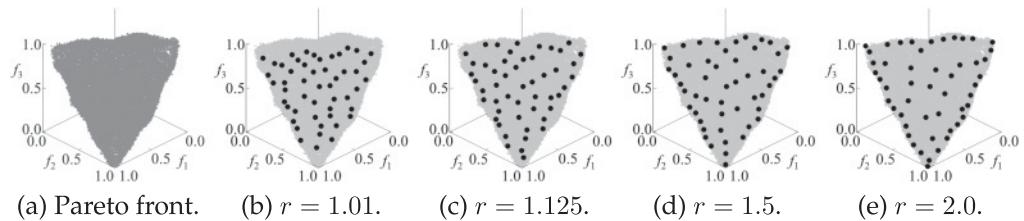


Figure 22: Pareto front in (a) and selected 50 solutions for each specification of the reference point for hypervolume maximization in (b)–(e) for the three-objective 500-item knapsack problem.

minimization problem with the ideal point  $(0, 0, 0)$  and the nadir point  $(1, 1, 1)$ . Then, a greedy algorithm (Guerreiro et al., 2016) is used to select a set of 50 solutions for hypervolume maximization. In the greedy algorithm, the first solution is selected by examining the hypervolume indicator of each solution. That is, the best solution with the largest hypervolume is selected as the first solution and included in the solution set. Next, the second solution is included in the solution set to maximize the hypervolume indicator after examining all combinations of the first solution and one of the other solutions. Then, the third solution is included in the solution set to maximize the hypervolume indicator after examining all combinations of the first two solutions and one of the other solutions. In this manner, 50 solutions are selected for each of the following settings of the reference point:  $r = 1.01, 1.125, 1.5, 20$ . The obtained solution sets are shown in Figures 22b–e.

When the reference point is slightly worse than the nadir point (i.e.,  $r = 1.01$ ) in Figure 22a, only a few solutions are close to the sides of the Pareto front. When the reference point is far away from the Pareto front (i.e.,  $r = 2.0$ ) in Figure 22e, many solutions are close to the sides of the Pareto front. When the reference point is specified by the proposed method (i.e.,  $r = 1.125$ ) in Figure 22c, solutions are well distributed over the entire Pareto front. Our experimental results in Figure 22 show that an appropriate reference point is specified by the proposed method for the three-objective 500-item knapsack problem. They also show that the optimal solution set strongly depends on the reference point specification. That is, hypervolume-based performance comparison results are sensitive to the reference point specification. As shown in Figure 22, the Pareto front of the three-objective knapsack problem is inverted triangular. As a result, the experimental results in Figure 22 are similar to those on the three-objective Minus-DTLZ1 problem in Figure 2.

#### 4.2 Three-Objective Car-Side Impact Problem

The car-side impact problem has frequently been used as a multi-objective problem with constraint conditions in the EMO community (Jain and Deb, 2014; Asafuddoula et al., 2015). This problem has three objectives, ten constraint conditions, and eleven decision variables. First, we search for an approximate Pareto front by NSGA-II (Deb et al., 2002) using a large computation load (population size: 5,000, termination condition: 10,000 generations, execution: 20 times). In NSGA-II, a set of 5,000 feasible solutions are generated as an initial population by randomly generating a large number of solutions and selecting only feasible solutions. During the execution of NSGA-II, infeasible solutions are handled as being dominated by feasible ones. All solutions obtained from 20 runs of

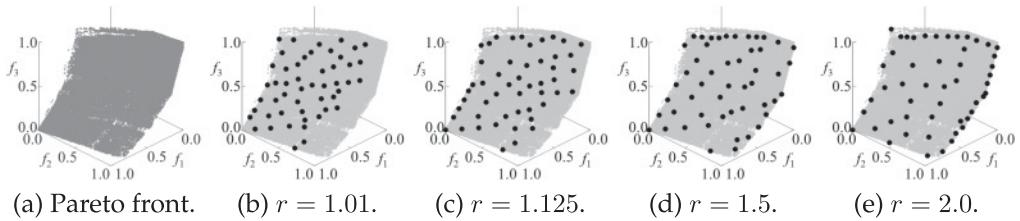


Figure 23: Pareto front in (a) and selected 50 solutions for each specification of the reference point for hypervolume maximization in (b)–(e) for the car-side impact problem.

NSGA-II are merged into a single solution set, from which nondominated solutions are selected. In this manner, we obtain an approximate Pareto front with 54,874 nondominated solutions in Figure 23a in the three-dimensional objective space. Then, a greedy algorithm is used to select a set of 50 solutions for hypervolume maximization.

In Subsection 4.1, we used MOEA/D to search for an approximate Pareto front of the three-objective knapsack problem. This is because much better results were reported for multi-objective knapsack problems with 2–4 objectives from MOEA/D than NSGA-II in the literature (e.g., Zhang and Li, 2007; Ishibuchi et al., 2015). However, in this subsection, we used NSGA-II to search for an approximate Pareto front of the car-side impact problem. This is because each objective of the car-side impact problem has a different scale. Since MOEA/D has no scaling mechanism of the objective space, usually it does not work well on such a multi-objective problem (e.g., WFG test problems of Huband et al., 2006).

The selected 50 solutions are shown in Figures 23b–e. When a reference point is slightly worse than the nadir point (i.e.,  $r = 1.01$ ) in Figure 23b, no solutions are close to the right-hand side boundary of the Pareto front. By specifying the reference point by the proposed method as  $r = 1.125$ , well-distributed solutions are obtained in Figure 23c except for the right-bottom corner of the Pareto front. By increasing the value of  $r$  (i.e., by moving the reference point far away from the Pareto front), more solutions are obtained around the top and the right sides of the Pareto front as shown in Figures 23d and e. Our experimental results on the car-side impact problem show that an appropriate reference point is specified by the proposed method. They also show that the optimal solution set strongly depends on the reference point specification. That is, hypervolume-based performance comparison results are sensitive to the reference point specification. The shape of the Pareto front of the car-side impact problem is a mixture of triangle and inverted triangle. In Figure 23, the bottom-left half of the Pareto front is triangular. The solution at the bottom-right corner is the single optimal solution in the  $f_2 - f_3$  space as in Figure 5a. So, many solutions are not around the bottom and left sides of the Pareto front even when  $r = 2.0$  in Figure 23e. On the contrary, the top-right half of the Pareto front is inverted triangular. So, many solutions are around the top and right sides of the Pareto front in Figure 23e.

### 4.3 Three-Objective WFG3 Problem

WFG3 was originally proposed as a degenerate test problem (Huband et al., 2006). However, it was shown in Ishibuchi et al. (2016) that WFG3 has a partially degenerate Pareto front as shown in Figure 24a. In the same manner as in Section 2, SMS-EMOA is applied to the three-objective WFG3. The number of distance variables is specified as two

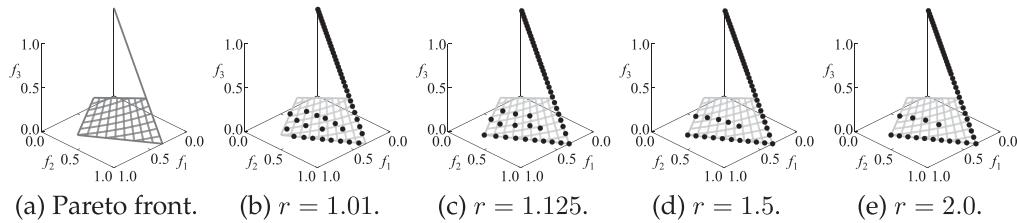


Figure 24: Pareto front in (a) and obtained 50 solutions for each specification of the reference point for hypervolume maximization in (b)–(e) for the three-objective WFG3.

due to the restriction of the WFG problem formulation whereas it was specified as zero for DTLZ1 and DTLZ2 in Section 2. Experimental results are shown in Figures 24b–e. When  $r = 1.01$  in Figure 24b, the two extreme solutions at the bottom-left and bottom-right corners are not selected. When  $r = 2.0$  in Figure 24e, only four solutions are inside the Pareto front. If compared with these results, the solution set from  $r = 1.125$  in Figure 24c looks better. As in Figures 22 and 23, Figure 24 shows that an appropriate reference point is specified by the proposed method. Figure 24 also shows that the best solution set depends on the reference point specification. That is, hypervolume-based performance comparison results depend on the reference point specification.

## 5 Effect on the Search Ability of SMS-EMOA

In this section, we examine whether the proposed specification method is useful in SMS-EMOA or not. In Sections 2–4, we discussed the reference point specification method for fair performance comparison where the true or approximate Pareto front is assumed to be known. However, in the use of the proposed method in SMS-EMOA, the ideal and nadir points are estimated from the current population. Their estimations are not accurate, especially in early generations. In our computational experiments in this section, the normalization is performed using non-dominated solutions among solutions in the current population at each generation. As mentioned in Emmerich et al. (2005), SMS-EMOA can work with an infinite reference point. This is a good specification for two-objective problems and multi-objective problems with triangular Pareto fronts. However, this is not appropriate for multi-objective problems with inverted triangular Pareto fronts. In this section, we examine the usefulness of the proposed reference point specification method in SMS-EMOA for multi-objective problems with triangular and inverted triangular Pareto fronts.

We apply SMS-EMOA to the three-objective normalized DTLZ1 and Minus-DTLZ1 under the following parameter specifications:

Number of distance variables: 5,

Population size ( $\mu$ ): 15 and 50,

Crossover: SBX with the index 20 (probability: 1.0),

Mutation: Polynomial mutation with the index 20 (probability:  $1/n$  where  $n$  is the number of variables),

Termination conditions: 100,000 generations,

Number of runs: 5.

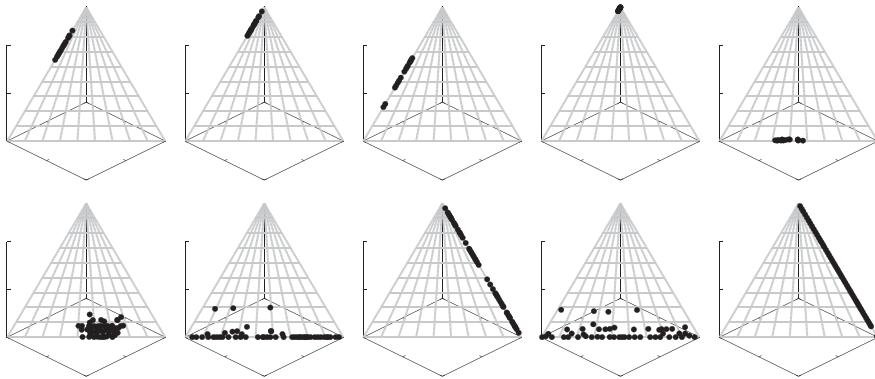


Figure 25: Obtained solution sets for the three-objective normalized DTLZ1 with  $r = 1.01$  (top: population size 15, bottom: population size 50).

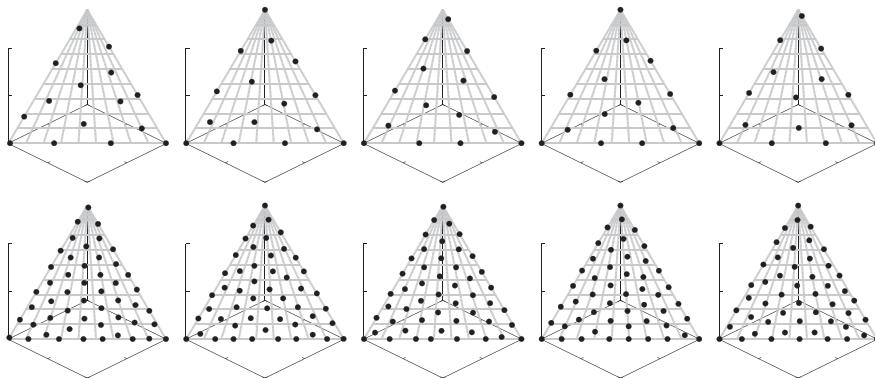


Figure 26: Obtained solution sets for the three-objective normalized DTLZ1 with the suggested reference point (top: population size 15 and  $r = 1.25$ , bottom: population size 50 and  $r = 1.125$ ).

In our computational experiments, we examine the following three-specification of the reference point: a slightly worse reference point than the nadir point (i.e.,  $r = 1.01$ ), the suggested reference point (i.e.,  $r = 1.25$  for  $\mu = 15$  and  $r = 1.125$  for  $\mu = 50$ ), and a much worse reference point than the nadir point (i.e.,  $r = 2.0$ ).

Experimental results on DTLZ1 are shown in Figures 25–27 where the final solution sets obtained from the five runs are shown for each specification of the population size. When  $r = 1.01$  in Figure 25, good results are not obtained. The final population always converges to a small region of the Pareto front. When the reference point is specified by the proposed method in Figure 26, well-distributed solutions are always obtained. Good results are also obtained in Figure 27 where the reference point is far away from the Pareto front. As we have already explained, the reference point can be far away from the triangular Pareto front. However, such a reference point specification has a clear negative effect when the Pareto front is inverted triangular.

Experimental results on Minus-DTLZ1 are shown in Figures 28–30. When the reference point is far away from the Pareto front in Figure 30, many solutions are on the

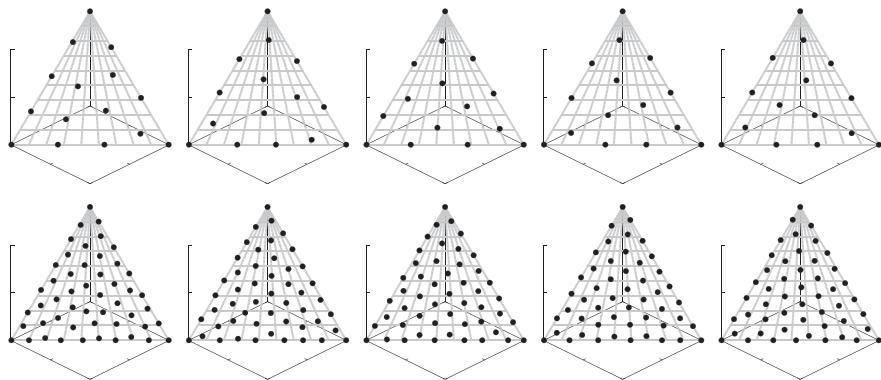


Figure 27: Obtained solution sets for the three-objective normalized DTLZ1 with  $r = 2.0$  (top: population size 15, bottom: population size 50).

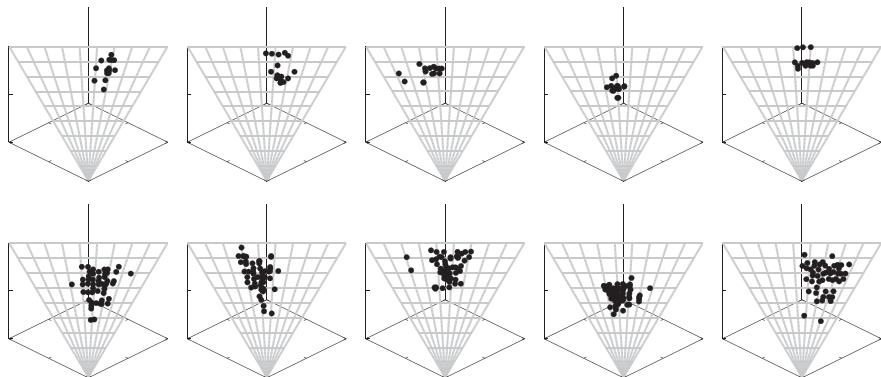


Figure 28: Obtained solution sets for the three-objective normalized Minus-DTLZ1 with  $r = 1.01$  (top: population size 15, bottom: population size 50).

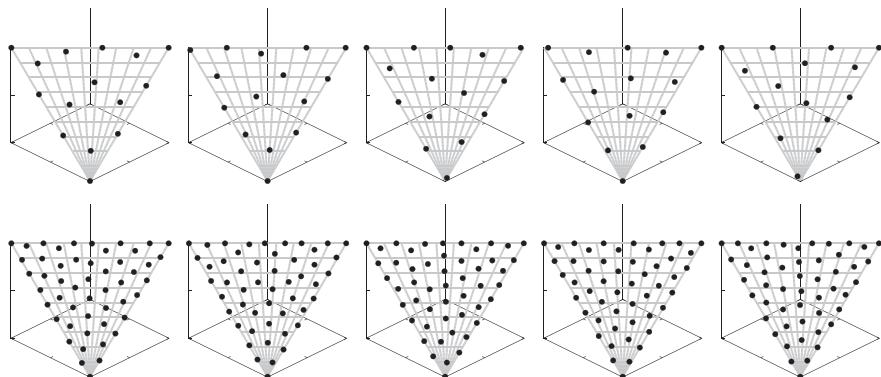


Figure 29: Obtained solution sets for the three-objective normalized Minus-DTLZ1 with the suggested reference point (top: population size 15 and  $r = 1.25$ , bottom: population size 50 and  $r = 1.125$ ).

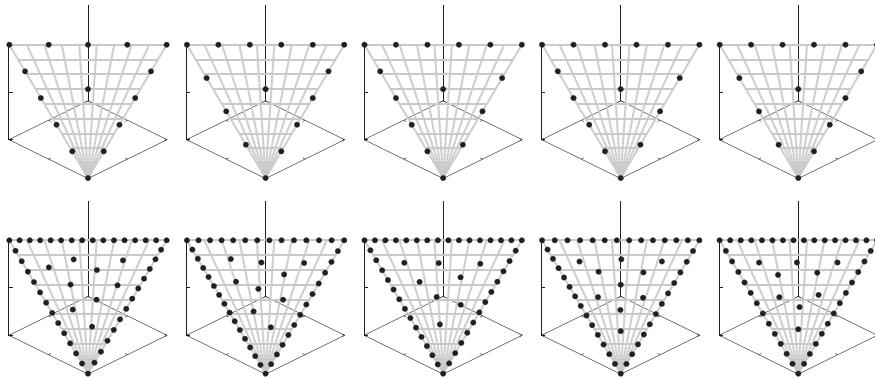


Figure 30: Obtained solution sets for the three-objective normalized Minus-DTLZ1 with  $r = 2.0$  (top: population size 15, bottom: population size 50).

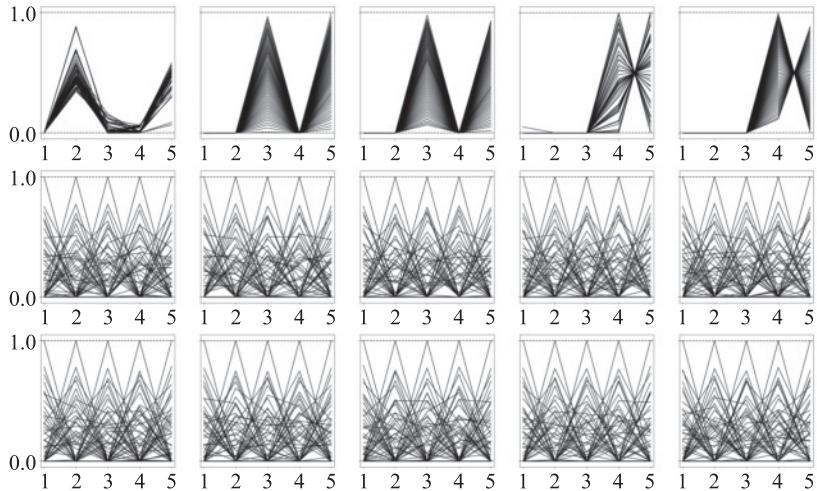


Figure 31: Experimental results on the five-objective normalized DTLZ1 with the population size 50 (top:  $r = 1.01$ , middle:  $r = 1.333$ , bottom:  $r = 2.0$ ).

sides of the inverted triangular Pareto front. When the reference point is slightly worse than the nadir point (i.e.,  $r = 1.01$ ), the final population always converges to a small region of the Pareto front in Figure 28. When the reference point is specified by the proposed method in Figure 29, well-distributed solutions are obtained. Our experimental results in Figures 28–30 show that the performance of SMS-EMOA on multi-objective problems with inverted triangular Pareto fronts strongly depends on the reference point specification. Only when the reference point is appropriately specified, well-distributed solutions over the entire Pareto front are obtained.

We also apply SMS-EMOA to the five-objective normalized DTLZ1 and Minus-DTLZ1. The population size is specified as 50. The suggested reference point for this setting (i.e.,  $m = 5$  and  $\mu = 50$ ) is  $r = 1.333$ . Experimental results are shown in Figures 31 and 32. When the reference point is slightly worse than the nadir point (i.e.,

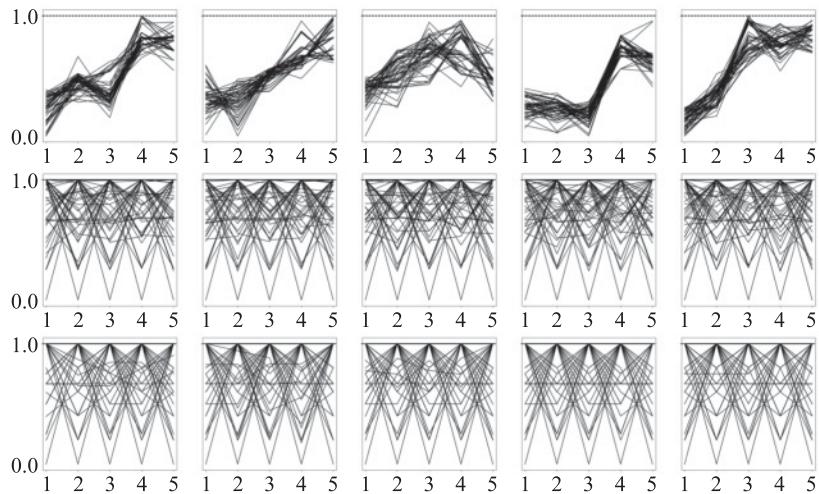


Figure 32: Experimental results on the five-objective normalized Minus-DTLZ1 with the population size 50 (top:  $r = 1.01$ , middle:  $r = 1.333$ , bottom:  $r = 2.0$ ).

$r = 1.01$ ), good results are not obtained for these test problems as shown in the top five results in each figure. When the reference point is specified by the proposed method (i.e.,  $r = 1.333$ ), good results are obtained as shown in the middle five results in each figure. When the reference point is far away from the Pareto front (i.e.,  $r = 2.0$ ), good solution sets are obtained for DTLZ1 in the bottom five results of Figure 31. This is because DTLZ1 has a triangular Pareto front. However, in the application of SMS-EMOA to Minus-DTLZ1, most solutions are around the sides of the Pareto front (i.e., at least one objective value is 0 or 1). Many solutions are not obtained inside the Pareto front (whereas this is not very clear in the bottom five results in Figure 32).

## 6 Concluding Remarks

In this article, we discussed the reference point specification in hypervolume calculation for fair performance comparison of EMO algorithms. First, we demonstrated that the optimal distribution of solutions for hypervolume maximization strongly depends on the location of the reference point except for multi-objective problems with triangular Pareto fronts such as DTLZ1-4 and WFG4-9. This means that hypervolume-based performance comparison results strongly depend on the location of the reference point. It was shown through computational experiments that a slightly worse reference point than the nadir point is not always appropriate especially for the case of many-objective problems and/or small population size. It was also shown that a much worse reference point, which is far away from the nadir point, is not appropriate except for multi-objective problems with triangular Pareto fronts. Based on the theoretical analysis of the optimal distribution of solutions for hypervolume maximization, we proposed a reference point specification method. The proposed method calculates the location of the reference point based on the number of solutions (i.e., population size) and the number of objectives. In the proposed method, the distance between the reference point and the nadir point increases with the increase in the number of objectives and the decrease in the number of solutions (i.e., it increases with the increase in the distance

between adjacent solutions). Our experimental results showed that an appropriate reference point can be specified by the proposed method for multi-objective and many-objective problems with linear triangular and linear inverted triangular Pareto fronts. It was also shown that the reference point specified by the proposed method is more appropriate for other test problems than a slightly or much worse reference point.

Since the proposed method was derived from the theoretical analysis of the optimal distribution of solutions for two-objective problems with linear Pareto fronts, the specified reference point is not necessarily appropriate for multi-objective problems with nonlinear Pareto fronts. Modification of the proposed method for nonlinear Pareto fronts is an interesting future research topic. Since the proposed method was derived for fair performance comparison where the true Pareto front is assumed to be known, the specified reference point is not necessarily appropriate in hypervolume-based EMO algorithms where the ideal and nadir points are estimated from the current population. Modification of the proposed method for hypervolume-based EMO algorithms is also an interesting future research topic. Especially, online adaptation of the reference point seems to be a promising future research direction.

In this article, we discussed the reference point specification in hypervolume calculation using the optimal distribution of solutions for hypervolume maximization. This approach (i.e., use of the optimal distribution of solutions) applies to the analysis of other performance indicators (Ishibuchi et al., 2018). Comparison of performance indicators using the optimal distribution of solutions for each indicator is an important future research direction for the further understanding and fair performance comparison of EMO algorithms.

In this article, we implicitly assumed that the size of each solution set to be compared was the same. That is, the number of non-dominated solutions obtained by each EMO algorithm was the same. However, this is not always the case. As demonstrated in Ishibuchi et al. (2016), an appropriate population size for each EMO algorithm is different. As a result, performance comparison results of EMO algorithms depend on the setting of the population size when they are compared under the same population size. Moreover, even when they are compared under the same population size, the number of obtained non-dominated solutions by each EMO algorithm may be different since all the obtained solutions are not always non-dominated. Thus, fair performance comparison of EMO algorithms is difficult. In order to address this difficulty, two scenarios were proposed in Ishibuchi et al. (2016) and used in Tanabe et al. (2017). One scenario is the use of an unbounded archive where all nondominated solutions among examined ones are stored. Performance comparison is based on all nondominated solutions in the obtained archive by each run of an EMO algorithm. The other scenario is the selection of a prespecified number of nondominated solutions from the obtained archive. By using several settings of the number of nondominated solutions, performance comparison results may be unbiased if compared with the case of the common population size.

## Acknowledgments

This work was supported by the Science and Technology Innovation Committee Foundation of Shenzhen (Grant No. ZDSYS201703031748284).

## References

- Asafuddoula, M., Ray, T., and Sarker, R. (2015). A decomposition-based evolutionary algorithm for many objective optimization. *IEEE Transactions on Evolutionary Computation*, 19(3):445–460.

- Auger, A., Bader, J., Brockhoff, D., and Zitzler, E. (2012). Hypervolume-based multiobjective optimization: Theoretical foundations and practical implications. *Theoretical Computer Science*, 425:75–103.
- Bader, J. (2009). Hypervolume-based search for multiobjective optimization: Theory and methods. PhD thesis, ETH Zurich.
- Bader, J., and Zitzler, E. (2011). HypE: An algorithm for fast hypervolume-based many-objective optimization. *Evolutionary Computation*, 19(1):45–76.
- Basseur, M., Derbel, B., Goffon, A., and Liefoghe, A. (2016). Experiments on greedy and local search heuristics for  $d$ -dimensional hypervolume subset selection. In *Proceedings of the 2016 Genetic and Evolutionary Computation Conference (GECCO)*, pp. 541–548.
- Beume, N., Naujoks, B., and Emmerich, M. (2007). SMS-EMOA: Multiobjective selection based on dominated hypervolume. *European Journal of Operational Research*, 181(3):1653–1669.
- Bringmann, K., Friedrich, T., and Klitzke, P. (2014). Two-dimensional subset selection for hypervolume and epsilon-indicator. In *Proceedings of the 2014 Genetic and Evolutionary Computation Conference (GECCO)*, pp. 589–596.
- Brockhoff, D. (2010). Optimal  $\mu$ -distributions for the hypervolume indicator for problems with linear bi-objective fronts: Exact and exhaustive results. In *8th International Conference on Simulated Evolution and Learning*, pp. 24–34.
- Chugh, T., Jin, Y., Miettinen, K., Hakkanen, J., and Sindhy, K. (2018). A surrogate-assisted reference vector guided evolutionary algorithm for computationally expensive many-objective optimization. *IEEE Transactions on Evolutionary Computation*, 22(1):129–142.
- Deb, K., Thiele, L., Laumanns, M., and Zitzler, E. (2002). Scalable multi-objective optimization test problems. In *2002 IEEE Congress on Evolutionary Computation*, pp. 825–830.
- Emmerich, M., Beume, N., and Naujoks, B. (2005). An EMO algorithm using the hypervolume measure as selection criterion. In *2005 International Conference on Evolutionary Multi-Criterion Optimization*, pp. 62–76.
- Emmerich, M., Deutz, A., and Beume, N. (2007). Gradient-based/evolutionary relay hybrid for computing Pareto front approximations maximizing the S-metric. In *4th International Workshop on Hybrid Metaheuristics*, pp. 140–156. Lecture Notes in Computer Science, Vol. 4771.
- Emmerich, M., Deutz, A., Kruisselbrink, J., and Shukla, P. K. (2013). Cone-based hypervolume indicators: Construction, properties, and efficient computation. In *2013 International Conference on Evolutionary Multi-Criterion Optimization*, pp. 111–127.
- Emmerich, M., Deutz, A., and Yevseyeva, I. (2014). On reference point free weighted hypervolume indicators based on desirability functions and their probabilistic interpretation. *Procedia Technology*, 16:532–541.
- Friedrich, T., Neumann, F., and Thyssen, C. (2015). Multiplicative approximations, optimal hypervolume distributions, and the choice of the reference point. *Evolutionary Computation*, 23(1):131–159.
- Guerreiro, P., Fonseca, C. M., and Paquete, L. (2015). Greedy hypervolume subset selection in the three-objective case. In *Proceedings of the 2015 Genetic and Evolutionary Computation Conference (GECCO)*, pp. 671–678.
- Guerreiro, P., Fonseca, C. M., and Paquete, L. (2016). Greedy hypervolume subset selection in low dimensions. *Evolutionary Computation*, 24(3):521–544.
- Huband, S., Hingston, P., Barone, L., and While, L. (2006). A review of multiobjective test problems and a scalable test problem toolkit. *IEEE Transactions on Evolutionary Computation*, 10(5):477–506.

- Ishibuchi, H., Akedo, N., and Nojima, Y. (2015). Behavior of multi-objective evolutionary algorithms on many-objective knapsack problems. *IEEE Transactions on Evolutionary Computation*, 19(2):264–283.
- Ishibuchi, H., Imada, R., Setoguchi, Y., and Nojima, Y. (2017a). Hypervolume subset selection for triangular and inverted triangular Pareto fronts of three-objective problems. In *14th ACM/SIGEVO Conference on Foundations of Genetic Algorithms*, pp. 95–110.
- Ishibuchi, H., Imada, R., Setoguchi, Y., and Nojima, Y. (2017b). Reference point specification in hypervolume calculation for fair comparison and efficient search. In *Proceedings of the 2017 Genetic and Evolutionary Computation Conference (GECCO)*, pp. 585–592.
- Ishibuchi, H., Imada, R., Setoguchi, Y., and Nojima, Y. (2018). Reference point specification in inverted generational distance for triangular linear Pareto front. *IEEE Transactions on Evolutionary Computation* (Early Access Paper: Available from IEEE Xplore).
- Ishibuchi, H., Masuda, H., and Nojima, Y. (2014). Selecting a small number of non-dominated solutions to be presented to the decision maker. In *2014 IEEE International Conference on Systems, Man, and Cybernetics*, pp. 3850–3855.
- Ishibuchi, H., Setoguchi, Y., Masuda, H., and Nojima, Y. (2016). How to compare many-objective algorithms under different settings of population and archive sizes. In *2016 IEEE Congress on Evolutionary Computation*, pp. 1149–1156.
- Ishibuchi, H., Setoguchi, Y., Masuda, H., and Nojima, Y. (2017). Performance of decomposition-based many-objective algorithms strongly depends on Pareto front shapes. *IEEE Transactions on Evolutionary Computation*, 21(2):169–190.
- Jain, H., and Deb, K. (2014). An evolutionary many-objective optimization algorithm using reference-point based non-dominated sorting approach, Part II: Handling constraints and extending to an adaptive approach. *IEEE Transactions on Evolutionary Computation*, 18(4):602–622.
- Knowles, J. D., Corne, D. W., and Fleischer, M. (2003). Bounded archiving using the Lebesgue measure. In *Proceedings of the 2003 IEEE Congress on Evolutionary Computation*, pp. 2490–2497.
- Kuhn, T., Fonseca, C. M., Paquete, L., Ruzika, S., Duarte, M. M., and Figueira, J. R. (2016). Hypervolume subset selection in two dimensions: Formulations and algorithms. *Evolutionary Computation*, 24(3):411–425.
- Li, K., Deb, K., Zhang, Q., and Kwong, S. (2015). An evolutionary many-objective optimization algorithm based on dominance and decomposition. *IEEE Transactions on Evolutionary Computation*, 19(5):694–716.
- Maltese, J., Ombuki-Berman, B. M., and Engelbrecht, A. P. (2018). A scalability study of many-objective optimization algorithms. *IEEE Transactions on Evolutionary Computation* (Early Access Paper: Available from IEEE Xplore).
- Sato, H. (2014). Inverted PBI in MOEA/D and its impact on the search performance on multi and many-objective optimization. In *2014 Genetic and Evolutionary Computation Conference (GECCO)*, pp. 645–652.
- Seada, H., and Deb, K. (2016). A unified evolutionary optimization procedure for single, multiple, and many objectives. *IEEE Transactions on Evolutionary Computation*, 20(3):358–369.
- Shukla, P. K., Doll, N., and Schmeck, H. (2014). A theoretical analysis of volume based Pareto front approximations. In *Proceedings of the 2014 Genetic and Evolutionary Computation Conference (GECCO)*, pp. 1415–1422.
- Tanabe, R., Ishibuchi, H., and Oyama, A. (2017). Benchmarking multi- and many-objective evolutionary algorithms under two optimization scenarios. *IEEE Access*, 5:19597–19619.

- Wagner, T., Beume, N., and Naujoks, B. (2007). Pareto-, aggregation-, and indicator-based methods in many-objective optimization. In *4th International Conference on Evolutionary Multi-Criterion Optimization*, pp. 742–756.
- Yuan, Y., Xu, H., Wang, B., and Yao, X. (2016). A new dominance relation based evolutionary algorithm for many-objective optimization. *IEEE Transactions on Evolutionary Computation*, 20(1):16–37.
- Yuan, Y., Xu, H., Wang, B., Zhang, B., and Yao, X. (2016). Balancing convergence and diversity in decomposition-based many-objective optimizers. *IEEE Transactions on Evolutionary Computation*, 20(2):180–198.
- Zhang, Q., and Li, H. (2007). MOEA/D: A multiobjective evolutionary algorithm based on decomposition. *IEEE Transactions on Evolutionary Computation*, 11(6):712–731.
- Zhang, X., Tian, Y., Cheng, R., and Jin, Y. (2018). A decision variable clustering based evolutionary algorithm for large-scale many-objective optimization. *IEEE Transactions on Evolutionary Computation*, 22(1):97–112.
- Zitzler, E., Brockhoff, D., and Thiele, L. (2007). The hypervolume indicator revisited: On the design of Pareto-compliant indicators via weighted integration. In *Proceedings of the 4th International Conference on Evolutionary Multi-Criterion Optimization*, pp. 862–876.
- Zitzler, E., and Thiele, L. (1998). Multiobjective optimization using evolutionary algorithms—A comparative case study. In *5th International Conference on Parallel Problem Solving from Nature*, pp. 292–301.
- Zitzler, E., and Thiele, L. (1999). Multiobjective evolutionary algorithms: A comparative case study and the strength Pareto approach. *IEEE Transactions on Evolutionary Computation*, 3(4):257–271.
- Zitzler, E., Thiele, L., Laumanns, M., Fonseca, C. M., and da Fonseca, V. G. (2003). Performance assessment of multiobjective optimizers: An analysis and review. *IEEE Transactions on Evolutionary Computation*, 7(2):117–132.

Copyright of Evolutionary Computation is the property of MIT Press and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.