Use of Piecewise Linear and Nonlinear Scalarizing Functions in MOEA/D

Hisao Ishibuchi(✉), Ken Doi, and Yusuke Nojima

Department of Computer Science and Intelligent Systems,
Graduate School of Engineering, Osaka Prefecture University,
1-1 Gakuen-cho, Naka-ku, Sakai, Osaka 599-8531, Japan
{hisaoi,nojima}@cs.osakafu-u.ac.jp,
ken.doi@ci.cs.osakafu-u.ac.jp

Abstract. A number of weight vector-based algorithms have been proposed for many-objective optimization using the framework of MOEA/D (multi-objective evolutionary algorithm based on decomposition). Those algorithms are characterized by the use of uniformly distributed normalized weight vectors, which are also referred to as reference vectors, reference lines and search directions. Their common idea is to minimize the distance to the ideal point (i.e., convergence) and the distance to the reference line (i.e., uniformity). Each algorithm has its own mechanism for striking a convergence-uniformity balance. In the original MOEA/D with the PBI (penalty-based boundary intersection) function, this balance is handled by a penalty parameter. In this paper, we first discuss why an appropriate specification of the penalty parameter is difficult. Next we suggest a desired shape of contour lines of a scalarizing function in MOEA/D. Then we propose two ideas for modifying the PBI function. The proposed ideas generate piecewise linear and nonlinear contour lines. Finally we examine the effectiveness of the proposed ideas on the performance of MOEA/D for many-objective test problems.

Keywords: Evolutionary multi-objective optimization (EMO) · Many-objective optimization · Decomposition-based evolutionary algorithm · MOEA/D

1 Introduction

In the EMO (evolutionary multi-objective optimization) community, many-objective optimization has been a hot topic in the last decade [9, 10]. The difficulty of many-objective optimization for EMO algorithms is explained as follows [9]: When a Pareto dominance-based EMO algorithm such as NSGA-II [4] and SPEA [14] is applied to a multi-objective problem with many objectives, all solutions in a population become non-dominated with each other in a very early stage of evolution. As a result, no strong selection pressure towards the Pareto front can be generated by its Pareto dominance-based fitness evaluation mechanism.

Recently a number of weight vector-based algorithms were proposed for many-objective problems in the framework of MOEA/D [13] such as I-DBEA [1], RVEA [2], NSGA-III [3] and MOEA/DD [11]. Those algorithms are characterized by

J. Handl et al. (Eds.): PPSN XIV 2016, LNCS 9921, pp. 503–513, 2016.
DOI: 10.1007/978-3-319-45823-6_47

the use of uniformly distributed normalized weight vectors. They also have similar fitness evaluation mechanisms. In Fig. 1, we show the angle a between the solution $\boldsymbol{f}(\boldsymbol{x})$ and the nearest reference line l, the distance d_1 from $\boldsymbol{f}(\boldsymbol{x})$ to the ideal point $\boldsymbol{z}^*$ along l, and the distance d_2 from $\boldsymbol{f}(\boldsymbol{x})$ to l. Each solution is usually assigned to the nearest reference line l using the angle a or the distance d_2. Then the fitness of the assigned solution is evaluated by the closeness to the nearest reference line (i.e., a or d_2) and the closeness to the ideal point (i.e., d_1).

An important issue is how to strike a balance between the convergence (i.e., minimization of d_1) and the uniformity (i.e., minimization of d_2 or a). In MOEA/D [13], this balance was handled by the penalty parameter θ for the distance d_2 in the following PBI (penalty-based boundary intersection) function:

$$\text{Minimize}\, f^{PBI}(\mathbf{x}|\mathbf{w},\mathbf{z}^*) = d_1 + \theta d_2, \tag{1}$$

where the penalty parameter θ is a non-negative real number. This parameter is used to handle the balance between the convergence d_1 and the uniformity d_2.

In this paper, we first discuss the difficulty of the penalty parameter specification in MOEA/D in Sect. 2. We also discuss a desired shape of the contour lines of a scalarizing function in MOEA/D. Next we propose two ideas for modifying the PBI function in Sect. 3. One is a piecewise linear function, and the other is a non-linear function. Then the performance of MOEA/D with each function is examined in Sect. 4. Finally we conclude this paper in Sect. 5.

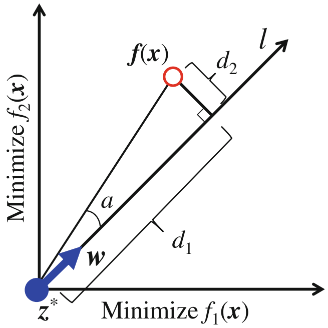
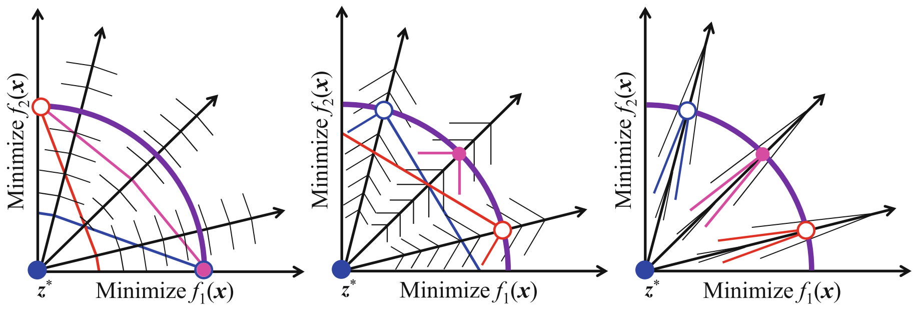

Minimize $f_2(\boldsymbol{x})$
$\boldsymbol{f}(\boldsymbol{x})$
d_2
l
a
$\boldsymbol{w}$
d_1
$\boldsymbol{z}^*$
Minimize $f_1(\boldsymbol{x})$

Fig. 1. The weight vector $\boldsymbol{w}$, the reference line l, and the solution $\boldsymbol{f}(\boldsymbol{x})$.

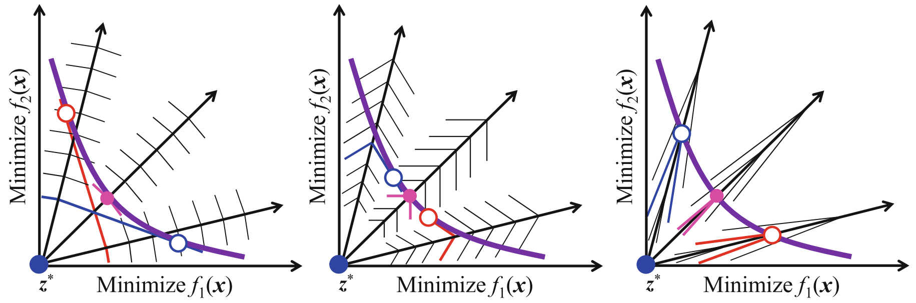
2 Parameter Specification in the PBI Function

In Fig. 2, we show the relation between the contour lines of the PBI function and the optimal solution for a concave Pareto front. When θ is not small, the optimal solution is on the intersection of the reference line and the Pareto front as shown in Fig. 2(b) and (c). However, when θ is small, the optimal solution is far from the reference line. For example, the red circle on the f_2 axis in Fig. 2(a) is the optimal solution for the reference line with the direction (0.8, 0.2). Moreover, when θ is small, it is difficult to find a solution on the concave region of the Pareto front as shown in Fig. 2(a).


Minimize $f_2(\boldsymbol{x})$
Minimize $f_1(\boldsymbol{x})$
z^*

(a) $\theta = 0.1$. (b) $\theta = 1$. (c) $\theta = 10$.

Fig. 2. Relation between the contour lines of the PBI function for three directions (($\boldsymbol{w}$ = (0.2, 0.8), (0.5, 0.5), (0.8, 0.2)) and the optimal solution for the case of a concave Pareto front. (Color figure online)


Minimize $f_2(\boldsymbol{x})$
Minimize $f_1(\boldsymbol{x})$
z^*

(a) $\theta = 0.1$. (b) $\theta = 1$. (c) $\theta = 10$.

Fig. 3. Relation between the contour lines of the PBI function for three directions (($\boldsymbol{w}$ = (0.2, 0.8), (0.5, 0.5), (0.8, 0.2)) and the optimal solution for the case of a convex Pareto front.

Well-distributed solutions are obtained from a small value of θ in Fig. 3(a) and a large value of θ in Fig. 3(c). However, in Fig. 3(b) with $\theta = 1$, the three solutions are close to each other around the center of the Pareto front (i.e., well-distributed solutions are not obtained). Thus an intermediate value of θ is not a good choice.

From these discussions, one may think that a large value of θ is a good choice. The use of a large value of θ is also consistent with the emphasis of the uniformity in the above-mentioned weight vector-based algorithms. However, a large value of θ degrades the convergence property of the PBI function in the same manner as the performance deterioration of Pareto dominance-based EMO algorithms for many-objective problems [7]. In Fig. 4, we show the region of solutions which are evaluated as being better than the red circle by the PBI function. When θ is small in Fig. 4(a), the solution has a large improved region. However, when θ is large in Fig. 4 (c), the improved region is very small. So, it is not likely that a better solution is easily found by crossover and mutation. The increase in the number of objectives exponentially decreases the ratio of this improved region in the neighborhood of the solution.

This exponential decrease explains poor performance of the PBI function with a large value of θ for many-objective knapsack problems [7]. Similar discussions were given about the specification of p in the weighted L_p scalarizing function in [12].

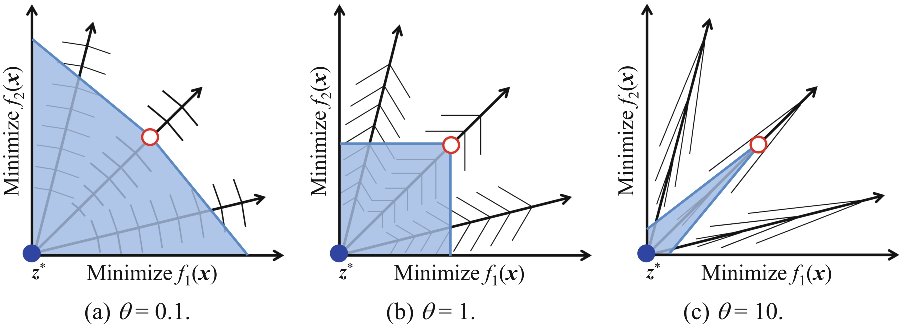

Minimize $f_2(\boldsymbol{x})$
Minimize $f_1(\boldsymbol{x})$
z^*
(a) $\theta = 0.1$. (b) $\theta = 1$. (c) $\theta = 10$.

Fig. 4. Improved region for a solution (red circle) with respect to the PBI function. (Color figure online)

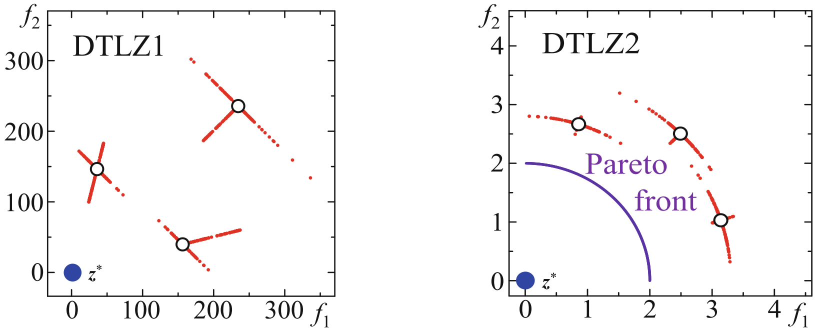


Fig. 5. 100 solutions generated by polynomial mutation of a randomly selected variable [8].

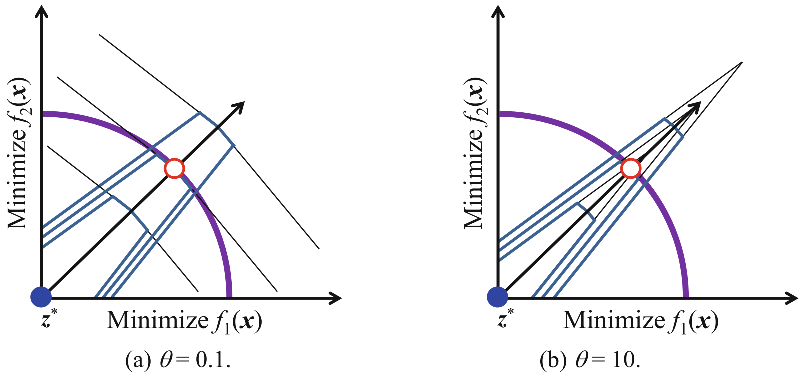
These discussions, however, are not consistent with experimental results in our former study [8] where good results were obtained from a large value of θ for many-objective DTLZ1 and DTLZ2 problems. This inconsistency can be explained by the special features of DTLZ 1-4 [5]. In DTLZ 1-4, the decision variable vector $\boldsymbol{x}$ is separable into the distance variable vector $\boldsymbol{x}_M$ and the position variable vector $\boldsymbol{x}_{\text{pos}}$. The objective vector $\boldsymbol{f}(\boldsymbol{x})$ is written as $\boldsymbol{f}(\boldsymbol{x}) = (1 + g(\boldsymbol{x}_M))\boldsymbol{h}(\boldsymbol{x}_{\text{pos}})$. Pareto optimal solutions are obtained by minimizing the scalar function $g(\boldsymbol{x}_M)$ to $g(\boldsymbol{x}_M) = 0$. Thus, the convergence improvement can be viewed as separate single-objective optimization.

In Fig. 5, we show 100 solutions generated by the polynomial mutation with the distribution index 20 to a randomly selected single variable from each of three solutions (open circles). When a distance variable in $\boldsymbol{x}_M$ is mutated, only the distance from the ideal point z^* is decreased or increased without changing any value of $\boldsymbol{h}(\boldsymbol{x}_{\text{pos}})$. Thus improved solutions are obtained on the line between the ideal solution z^* and the current solution in Fig. 5. If the current solution is on the reference line, those solutions are evaluated as being better than the current solution by the PBI function independent of the value of θ. When the mutation is applied to a position variable in $\boldsymbol{x}_{\text{pos}}$, the location of the solution is changed without changing the value of $g(\boldsymbol{x}_M)$ as shown in Fig. 5. Thus the uniformity can be improved separately from the convergence. Thanks to these special features, good experimental results were reported when a large value of θ was used for many-objective DTLZ 1-4 test problems. WFG 4-9 test problems [6] also have similar special features.

Discussions on the specification of θ in this section are summarized as follows.

(a) Small values of θ: The PBI function has high convergence ability even for many-objective problems. Its main difficulty is the handling of concave Pareto fronts.
(b) Values between (a) and (c): The diversity of solutions can be very small for values around $\theta = 1$ when the shape of the Pareto front is convex.
(c) Large values of θ: Uniformly distributed solutions are likely to be obtained. However, the convergence is degraded by the increase in the number of objectives.

These discussions may suggest two directions for improving the PBI function. One is to improve the uniformity for the PBI function with a small value of θ. This direction is illustrated in Fig. 6(a). The other is to improve the convergence for the PBI function with a large value of θ as illustrated in Fig. 6(b). The contour lines after the modification are similar between Fig. 6(a) and (b). That is, the convergence is emphasized only when a solution is close to the reference line. The uniformity is emphasized when a solution is far from the reference line.


Minimize $f_2(\boldsymbol{x})$
z^*
Minimize $f_1(\boldsymbol{x})$
Minimize $f_2(\boldsymbol{x})$
z^*
Minimize $f_1(\boldsymbol{x})$

(a) $\theta = 0.1$. (b) $\theta = 10$.

Fig. 6. Modifications of the contour lines of the PBI function.

3 Modifications of the PBI Functions

The PBI function after the modifications in Fig. 6 can be formulated using two penalty parameters θ_1 and θ_2 as the following two-level PBI function:

$$\text{Minimize}\, f_{2-Level}^{PBI}(\mathbf{x}|\mathbf{w},\mathbf{z}^{*}) = \begin{cases} d_1 + \theta_1 d_2, & \text{if}\, d_2 \le d^{*}, \\ d_1 + \theta_1 d^{*} + \theta_2 (d_2 - d^{*}), & \text{if}\, d_2 > d^{*}, \end{cases} \quad (2)$$

where $\theta_1 < \theta_2$ and d^* is a parameter to switch the penalty value between θ_1 and θ_2. If d_2 is smaller than d^*, a small penalty value θ_1 is used. If d_2 is larger than d^*, a large penalty value θ_2 is used for the amount of the violation: d_2 - d^*. In this paper, we specify the two penalty parameters θ_1 and θ_2 as θ_1 = 0.1 and θ_2 = 10.

The value of d^* is specified by solutions in the current population as follows:

$$d^{*} = \alpha \frac{1}{H}\frac{1}{m}\sum_{i=1}^{m}\left(f_i^{Max}(\mathbf{x}) - f_i^{Min}(\mathbf{x})\right), \quad (3)$$

where α is a parameter, H is an integer parameter used for generating uniformly distributed weight vectors in MOEA/D, m is the number of objectives, and $f_i^{Max}(\boldsymbol{x})$ and $f_i^{Min}(\boldsymbol{x})$ are the maximum and minimum values of the ith objective in the current population, respectively. In (3), the average width of the domain of each objective is divided by H to obtain a rough estimation for the distance between adjacent solutions. The parameter α is used to examine the validity of the formulation (3) through computational experiments with various values of α.

Our idea in (2) is to use a small penalty value only when a solution is close to the reference line. This idea can be also implemented as the following quadratic function.

$$\text{Minimize} f_{Quadratic}^{PBI}(\mathbf{x}|\mathbf{w},\mathbf{z}^{*}) = d_1 + \theta d_2 \frac{d_2}{d^{*}}, \quad (4)$$

where d^* is the same parameter as in (2), which is calculated by (3). The effect of the penalty parameter θ is decreased by the factor (d_2/d^*) when d_2 is small (i.e., $d_2 < d^*$) and increased by (d_2/d^*) when d_2 is large (i.e., $d_2 > d^*$). When $d_2 = d^*$, this formulation is the same as the PBI function in (1). The value of θ is specified as $\theta = 1$ in (4).

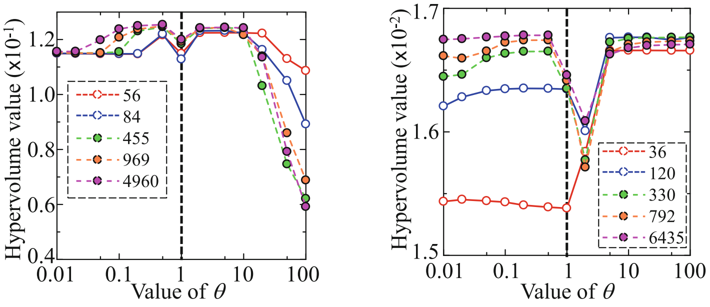
4 Computational Experiments

4.1 Experimental Results of the PBI Function

We applied MOEA/D with the PBI function to DTLZ 1-2 with four and eight objectives. Various values of θ between 0.01 and 100 were examined. The total number of examined solutions was used as the termination condition: $m \times 10{,}000$ solutions for m-objective problems. We examined various settings of the population size. The neighborhood size in MOEA/D was specified as 10 % of the population size. The number of decision variables (n) was 5 + m − 1 (DTLZ1) and 10 + m − 1 (DTLZ2). We used the SBX crossover with the distribution index 15 and the crossover probability 0.8, and the polynomial mutation with the distribution index 20 and the mutation probability $1/n$. The average hypervolume was calculated over 50 runs for the reference point (0.6, …, 0.6) of DTLZ1 and (1.1, …, 1.1) of DTLZ2.

In the same manner, we applied MOEA/D to 500-item 0/1 knapsack problems with four and eight objectives [7] except for the following settings: 400,000 solution evaluations, uniform crossover with the probability 0.8, bit-flip mutation with the probability 2/500, the reference point (0, …, 0) for the hypervolume calculation, and the reference point z^* for the PBI function as $z_i^* = 1.1 \times \max\{f_i(\boldsymbol{x})\}$ for $i = 1$, 2, …, m where $\max\{f_i(\boldsymbol{x})\}$ is the maximum value of $f_i(\boldsymbol{x})$ among all the examined solutions [7].

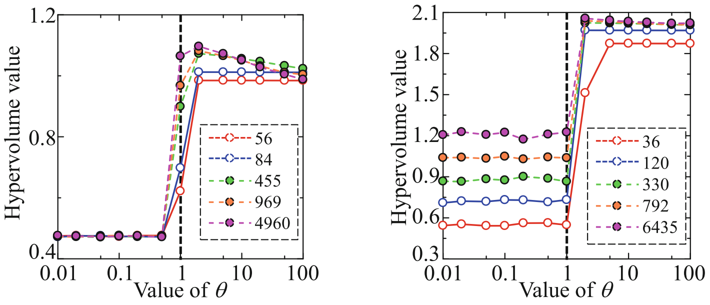
The average hypervolume value over 50 runs is shown in Figs. 7, 8 and 9. Each circle shows the average result from the corresponding setting of the population size (e.g., 56) and the value of θ (e.g., 0.01). The range of appropriate values of θ in each figure are as follows: $5 \le \theta \le 20$ in Fig. 7, $2 \le \theta \le 100$ in Fig. 8, and $0.01 \le \theta \le 0.1$ in Fig. 9. The PBI function with a small values of θ cannot handle the concave Pareto front of DTLZ2 in Fig. 8. Large values for θ deteriorate the convergence performance of the PBI function for many-objective knapsack problems in Fig. 9(b). Clear performance deterioration is also observed around $\theta = 1$ in Figs. 7 and 9. Figures 7, 8 and 9 show the difficulty and the importance of an appropriate parameter specification of θ.



(a) Four-Objective DTLZ1.

(b) Eight-Objective DTLZ1.

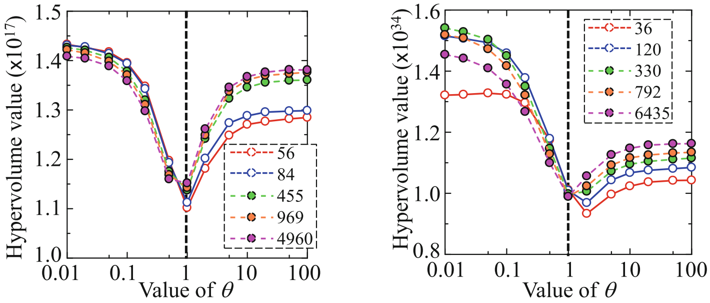
Fig. 7. Results of the PBI function on DTLZ1 (Linear Pareto front).



(a) Four-Objective DTLZ2.

(b) Eight-Objective DTLZ2.

Fig. 8. Results of the PBI function on DTLZ2 (Concave Pareto front).

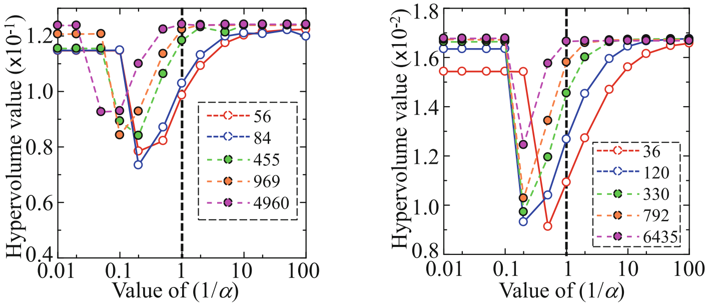


(a) Four-Objective knapsack problem. (b) Eight-Objective knapsack problem.

Fig. 9. Results of the PBI function on the knapsack problems (Convex Pareto front).

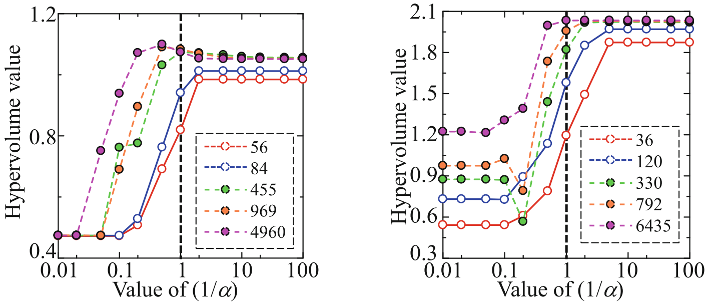
4.2 Experimental Results of the Two-Level PBI Function

Experimental results of the two-level PBI function are shown in Figs. 10, 11 and 12 where the horizontal axis is $1/\alpha$. At the leftmost (rightmost) point of each figure with a small (large) value of $1/\alpha$, $\theta_1 = 0.1$ ($\theta_2 = 10$) is mainly used. Thus the obtained results at the leftmost (rightmost) point of each figure are almost the same as those by $\theta = 0.1$ ($\theta = 10$) in Subsect. 4.1. Only for the knapsack problems, we use larger values of α (see the horizontal axis of each figure in Figs. 10, 11 and 12).



(a) Four-Objective DTLZ1. (b) Eight-Objective DTLZ1.

Fig. 10. Results of the two-level PBI function on DTLZ1.

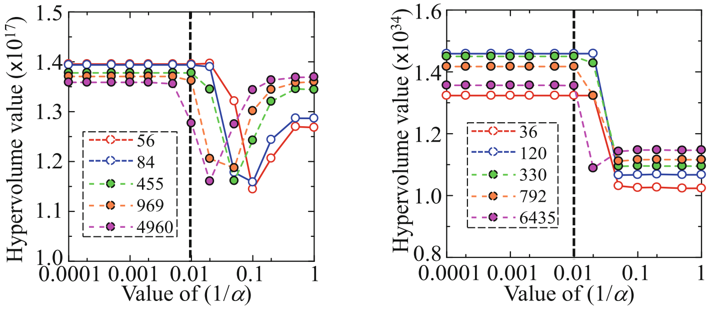


(a) Four-Objective DTLZ2. (b) Eight-Objective DTLZ2.

Fig. 11. Results of the two-level PBI function on DTLZ2.

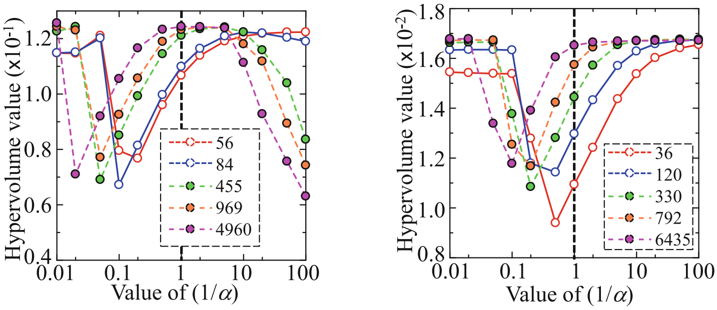
4.3 Experimental Results of the Quadratic PBI Function

Experimental results of the quadratic PBI function are shown in Figs. 13, 14 and 15. The obtained results at the leftmost (rightmost) point of each figure are similar to those by $\theta = 0.01$ ($\theta = 100$) in Subsect. 4.1. This is because the penalty value is very small (very large) on average at the leftmost (rightmost) point in Figs. 13, 14 and 15. We can



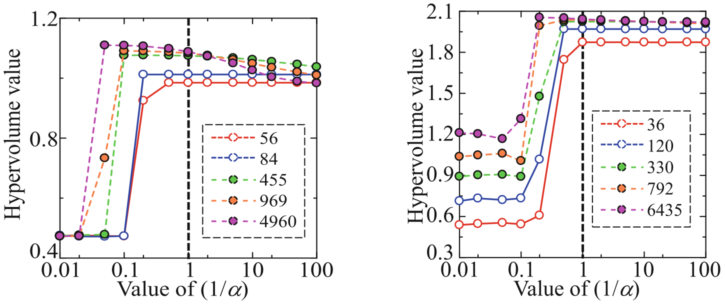
(a) Four-Objective knapsack problem. (b) Eight-Objective knapsack problem.

Fig. 12. Results of the two-level PBI function on the knapsack problems.



(a) Four-Objective DTLZ1. (b) Eight-Objective DTLZ1.

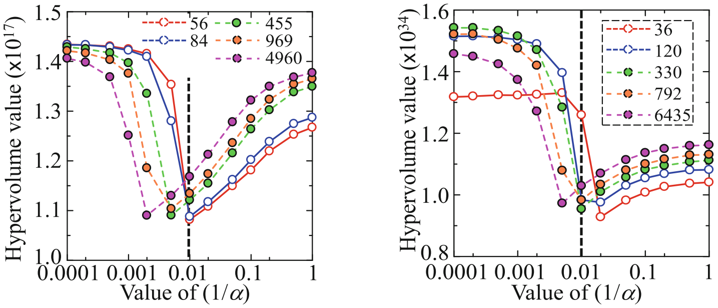
Fig. 13. Experimental results of the quadratic PBI function on DTLZ1.



(a) Four-Objective DTLZ2. (b) Eight-Objective DTLZ2.

Fig. 14. Experimental results of the quadratic PBI function on DTLZ2.

also observe some similarity among the obtained results on each test problem in the three subsections such as the V-shape results in Figs. 9(a), 12(a) and 15(a).



(a) Four-Objective knapsack problem.

(b) Eight-Objective knapsack problem.

Fig. 15. Experimental results of the quadratic PBI function on the knapsack problems.

5 Conclusions

We first explained why the specification of the penalty value θ is difficult in the PBI function of MOEA/D. Then we proposed an idea of modifying the shape of the contour lines of the PBI function to strike a convergence-uniformity balance. This idea was implemented as two-level and quadratic PBI functions. By the proposed idea, we obtained interpolative results between small and large penalty value cases in Figs. 11 and 14 for the DTLZ2 problems. In Figs. 11(a) and 14(a), improvement was observed by the proposed idea from the interpolative results when $0.1 < 1/\alpha < 1$. However, for the DTLZ1 and knapsack problems, clear performance deterioration was observed from the interpolative results, which was similar to the performance deterioration by θ around 1.0 in the original PBI function.

References

1. Asafuddoula, M., Ray, T., Sarker, R.: A decomposition-based evolutionary algorithm for many objective optimization. IEEE Trans. Evol. Comput. **19**, 445–460 (2015)
2. Cheng, R., Jin, Y., Olhofer, M., Sendhoff, B.: A reference vector guided evolutionary algorithm for many-objective optimization. IEEE Trans. Evol. Comput. (in press). doi:10.1109/TEVC.2016.2519378
3. Deb, K., Jain, H.: An evolutionary many-objective optimization algorithm using reference-point-based non-dominated sorting approach, part I: solving problems with box constraints. IEEE Trans. Evol. Comput. **18**, 577–601 (2014)
4. Deb, K., Pratap, A., Agarwal, S., Meyarivan, T.: A fast and elitist multiobjective genetic algorithm: NSGA-II. IEEE Trans. Evol. Comput. **6**, 182–197 (2002)
5. Deb, K., Thiele, L., Laumanns, M., Zitzler, E.: Scalable multi-objective optimization test problems. In: Proceedings of IEEE CEC 2002, pp. 825–830

6. Huband, S., Hingston, P., Barone, L., While, L.: A review of multiobjective test problems and a scalable test problem toolkit. IEEE Trans. Evol. Comput. **10**, 477–506 (2006)
7. Ishibuchi, H., Akedo, N., Nojima, Y.: Behavior of multi-objective evolutionary algorithms on many-objective knapsack problems. IEEE Trans. Evol. Comput. **19**, 264–283 (2015)
8. Ishibuchi, H., Doi, K., Nojima, Y.: Characteristics of many-objective test problems and penalty parameter specification in MOEA/D. In: Proceedings of IEEE CEC 2016, pp. 1115–1122 (2016)
9. Ishibuchi, H., Tsukamoto, N., Nojima, Y.: Evolutionary many-objective optimization: a short review. In: Proceedings of IEEE CEC 2008, pp. 2424–2431 (2008)
10. Li, B., Li, J., Tang, K., Yao, X.: Many-objective evolutionary algorithms: a survey. ACM Comput. Surv. **48**(1), 13 (2015)
11. Li, K., Deb, K., Zhang, Q., Kwong, S.: An evolutionary many-objective optimization algorithm based on dominance and decomposition. IEEE Trans. Evol. Comput. **19**, 694–716 (2015)
12. Wang, R., Zhang, Q., Zhang, T.: Decomposition based algorithms using pareto adaptive scalarizing methods. IEEE Trans. Evol. Comput. (in press). 10.1109/TEVC.2016.2521175
13. Zhang, Q., Li, H.: MOEA/D: a multiobjective evolutionary algorithm based on decomposition. IEEE Trans. Evol. Comput. **11**, 712–731 (2007)
14. Zitzler, E., Thiele, L.: Multiobjective evolutionary algorithms: a comparative case study and the strength pareto approach. IEEE Trans. Evol. Comput. **3**, 257–271 (1999)