H. Ishibuchi, K. Doi, H. Masuda, and Y. Nojima, "Relation between weight vectors and solutions in MOEA/D," Proc. of 2015 IEEE Symposium on Computational Intelligence in Multicriteria Decision-Making (Accepted).

Relation between Weight Vectors and Solutions in MOEA/D

Hisao Ishibuchi, Ken Doi, Hiroyuki Masuda and Yusuke Nojima
Department of Computer Science and Intelligent Systems
Graduate School of Engineering, Osaka Prefecture University
Sakai, Osaka 599-8531, Japan
{hisaoi@, ken.doi@ci., hiroyuki.masuda@ci., nojima@}cs.osakafu-u.ac.jp

***Abstract*—An important implementation issue in MOEA/D (multiobjective evolutionary algorithm based on decomposition) is the specification of a scalarizing function. For its appropriate specification, it is necessary to understand the search behavior of MOEA/D for various settings of a scalarizing function. Especially, it is important to understand the relation between weight vectors and obtained solutions. The understanding of this relation is also very important for the incorporation of preference information into MOEA/D through weight vector specification. In this paper, we examine the mapping from weight vectors to solutions by monitoring which solution is obtained from each weight vector. MOEA/D with a number of different settings of a scalarizing function is applied to knapsack problems and DTLZ2 with 2-6 objectives. As a scalarizing function, we use the weighted sum, the weighted Tchebycheff and the PBI (penalty-based boundary intersection). We report some interesting observations obtained from computational experiments. Among them are the existence of many duplicated solutions, their positive and negative effects, and a dominant effect of the penalty parameter value in the PBI.**

I. Introduction

Evolutionary many-objective optimization [1] has been a hot topic in the field of EMO (evolutionary multiobjective optimization [2]). Since Pareto dominance-based algorithms such as NSGA-II [3] and SPEA2 [4] do not always work well on many-objective problems, various approaches have been proposed for the use of other fitness evaluation schemes. Most of them can be categorized into two classes. One is indicator-based algorithms where a performance indicator is used for fitness evaluation (e.g., IBEA [5], SMS-EMOA [6], HypE [7]). Hypervolume has been frequently used as an indicator due to its theoretical property called Pareto compliance [8]. However, its calculation for many-objective problems often needs heavy computation load. The other class is scalarizing function-based algorithms. MOEA/D [9] is a representative of this class. The popularity of this class has drastically increased after the proposal of MOEA/D in 2007 [9]. Scalarizing functions have been used for fitness evaluation in memetic EMO algorithms since mid-1990s [10], [11]. A cellular EMO algorithm similar to MOEA/D was also proposed in 2001 [12]. The main advantage of scalarizing function-based algorithms is their computational efficiency of fitness evaluation. The simplicity of algorithm structure is an additional advantage of MOEA/D.

A new trend in evolutionary many-objective optimization is the use of Pareto dominance-based and scalarizing function-based fitness evaluation schemes in a single algorithm such as NSGA-III [13] and I-DBEA (improved decomposition-based evolutionary algorithm [14]). It was reported in these studies that better results were obtained from NSGA-III and I-DBEA than MOEA/D. However, the use of Pareto dominance together with scalarizing functions may degrade the two advantages (i.e., efficient fitness evaluation and simple algorithm structure) while the search ability of MOEA/D may be improved.

In MOEA/D [9], the following three scalarizing functions were examined: the weighted sum, the weighted Tchebycheff with a reference point, and the PBI with a reference point and a penalty parameter value. As shown in the literature (e.g., see [15], [16]), the performance of MOEA/D depends strongly on the specification of a scalarizing function. Good specifications for some problems often lead to poor results for other problems. Moreover, good specifications for multiobjective problems are often inappropriate for many-objective problems. For example, it was reported in [15] that the PBI with a penalty parameter value $\theta = 5$ did not work well on a ten-objective knapsack problem while it was the best specification for a two-objective knapsack problem. The use of other scalarizing functions such as an inverted PBI function [17] was proposed in the literature. The decomposition of a many-objective problem into a number of simple multiobjective problems was also proposed [18].

As we have already mentioned, MOEA/D is an efficient EMO algorithm with a simple decomposition structure. It also has a high flexibility in its implementation (e.g., choice of a scalarizing function and weight vectors). However, it seems that its high flexibility has not been fully utilized or recognized in the literature. For example, a single fixed specification of a scalarizing function has often been used for all test problems in performance comparison. When a scalarizing function is not appropriately specified, high flexibility in the implementation of MOEA/D can be viewed as its disadvantage.

For fully utilizing the high flexibility of MOEA/D, we may need to know the behavior of MOEA/D under various settings. Especially, it is important to understand the relation between weight vectors and solutions. The understanding of this relation is also important for preference incorporation into MOEA/D through weight vector specification. In this paper, we visually examine this relation by monitoring which solution is obtained for each weight vector in each run of MOEA/D on knapsack problems [15], [19] and DTLZ2 [20] with 2-6 objectives. That is, the mapping from weight vectors to solutions is monitored

This work was supported by JSPS KAKENHI Grant Numbers 24300090 and 26540128.

in each run. As a scalarizing function, we use the weighted sum, the weighted Tchebycheff, and the PBI. Various specifications are examined for a reference point (Tchebycheff and PBI) and a penalty parameter value (PBI). Different specifications of the number of weight vectors (i.e., the population size) are also examined. We report some interesting observations such as positive and negative effects of many duplicated solutions in a population, which are observed in computational experiments by monitoring the mapping from weight vectors to solutions. Recently some ideas have been proposed for tackling negative effects of duplicated solutions on the diversity of a population in MOEA/D (e.g., use of a constraint condition [21] and new replacement mechanisms [22]-[24]). Most experimental results in this paper support the motivation of those ideas. However, some other results show positive effects of duplicated solutions.

This paper is organized as follows. In Section II, the three scalarizing functions are explained. Experimental results on knapsack problems and DTLZ2 are reported in Section III and Section IV, respectively. In Section V, this paper is concluded.

II. Scalarizing Functions in MOEA/D

In this section, we explain the three scalarizing functions in MOEA/D [9] for a multiobjective problem with k objectives: $f_1(\boldsymbol{x})$, $f_2(\boldsymbol{x})$, ..., $f_k(\boldsymbol{x})$. In MOEA/D, a multiobjective problem is decomposed into a number of single-objective problems. Each single-objective problem is defined by the same scalarizing function with a different weight vector. Thus the number of the single-objective problems is the same as the number of the weight vectors, which is the same as the population size.

Weight vectors $\boldsymbol{w} = (w_1, w_2, ..., w_k)$ are uniformly sampled using the following formulations in MOEA/D [9]:

$$w_1 + w_2 + \cdots + w_k = 1\,, \tag{1}$$

$$w_i \in \left\{0, \frac{1}{H}, \frac{2}{H}, ..., \frac{H}{H}\right\},\ i = 1, 2, ..., k\,, \tag{2}$$

where H is a positive integer. The number of weight vectors (i.e., population size) can be calculated from (1) and (2) as ${}_{H+k-1}C_{k-1}$ (see [9]). In computational experiments, we use two settings of H for each test problem as shown in Table I and Table II. The population size of 100-200 has been frequently used in the literature (e.g., see [13], [14]). Thus we use small populations in Table I. However, good results by MOEA/D are often obtained from large populations (e.g., see [16]). Thus we also examine large populations in Table II.

TABLE I. Setting for Small Population Size.

Number of objectives: k	2	3	4	5	6
Value of H	99	13	7	5	4
Population size: ${}_{H+k-1}C_{k-1}$	100	105	120	126	126

TABLE II. Setting for Large Population Size.

Number of objectives: k	2	3	4	5	6
Value of H	4999	98	29	16	11
Population size: ${}_{H+k-1}C_{k-1}$	5000	4950	4960	4845	4368

The weighted sum is written for a maximization problem as

$$\text{Maximize } f^{WS}(\boldsymbol{x}\,|\,\boldsymbol{w}) = w_1 \cdot f_1(\boldsymbol{x}) + \cdots + w_k \cdot f_k(\boldsymbol{x})\,. \tag{3}$$

This formulation is rewritten for a minimization problem as

$$\text{Minimize } f^{WS}(\boldsymbol{x}\,|\,\boldsymbol{w}) = w_1 \cdot f_1(\boldsymbol{x}) + \cdots + w_k \cdot f_k(\boldsymbol{x})\,. \tag{4}$$

Using a reference point $\boldsymbol{z}^* = (z_1^*, z_2^*, ..., z_k^*)$, the weighted Tchebycheff for a maximization problem is written as

$$\text{Minimize } f^{TE}(\boldsymbol{x}\,|\,\boldsymbol{w}, \boldsymbol{z}^*) = \max_{i=1,2,...,k}\{w_i \cdot |\, z_i^* - f_i(\boldsymbol{x})\,|\}\,. \tag{5}$$

In this paper, we specify the reference point $\boldsymbol{z}^*$ in (5) as

$$z_i^* = \alpha \cdot \max\{f_i(\boldsymbol{x})\,|\,\boldsymbol{x} \in \Omega(1) \cup \Omega(2) \cup \cdots \cup \Omega(t)\}\,,\quad i = 1, 2, ..., k, \tag{6}$$

where α is a real number and $\Omega(t)$ is the population at the t-th generation. The value of α is usually specified as $\alpha = 1.0$. However, $\alpha = 1.1$ was also used in [9]. In this paper, three values of α are examined ($\alpha = 1.0, 1.01, 1.1$) for maximization problems. The use of a value larger than 1.0 for α is to increase the diversity of solutions for maximization problems.

For minimization problems, (5) can be used after changing the definition of the reference point $\boldsymbol{z}^*$ in (6) as

$$z_i^* = \alpha \cdot \min\{f_i(\boldsymbol{x})\,|\,\boldsymbol{x} \in \Omega(1) \cup \Omega(2) \cup \cdots \cup \Omega(t)\}\,,\quad i = 1, 2, ..., k. \tag{7}$$

For the same reason as in maximization problems, we examine three specifications of α: $\alpha = 0.9, 0.99, 1.0$.

The PBI function for a maximization problem is written as

$$\text{Minimize } f^{PBI}(\boldsymbol{x}\,|\,\boldsymbol{w}, \boldsymbol{z}^*) = d_1 + \theta d_2\,, \tag{8}$$

where θ is a penalty parameter, and d_1 and d_2 are as follows:

$$d_1 = \left\| (\boldsymbol{z}^* - \boldsymbol{f}(\boldsymbol{x}))^T \boldsymbol{w} \right\| \Big/ \|\boldsymbol{w}\|\,, \tag{9}$$

$$d_2 = \left\| \boldsymbol{f}(\boldsymbol{x}) - \left(\boldsymbol{z}^* - d_1 \frac{\boldsymbol{w}}{\|\boldsymbol{w}\|} \right) \right\|. \tag{10}$$

The reference point $\boldsymbol{z}^*$ is specified in the same manner as in the weighted Tchebycheff. For a minimization problem, "–" in the parentheses in (10) should be replaces with "+". In [9], θ was specified as $\theta = 5$. In this paper, we examine four values of θ ($\theta = 0.1, 1, 5, 10$) for each value of α.

III. Experimental Results on Knapsack Problems

A. Setting of Computational Experiments

We applied MOEA/D (with no archive population) to 500-item knapsack problems with 2-6 objectives [15]. Fig. 1 shows randomly generated 100 solutions and the Pareto front of the two-objective 500-item knapsack problem. As shown in Fig. 1, EMO algorithms need a strong convergence property (since initial solutions are far from the Pareto front) as well as a strong diversification property (since the diversity of initial solutions is much smaller than that of the Pareto front). Computational experiments in this paper were performed under the same setting as in our former study [15] except for the neighborhood size (which was specified as 10 in [15]):

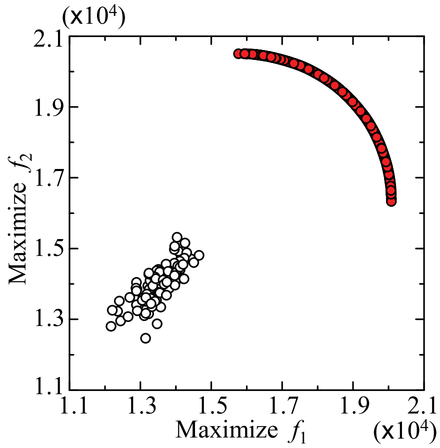


Fig. 1. Randomly generated 100 solutions (open circles) and the Pareto front (red circles) of the two-objective 500-item knapsack problem.

Coding: Binary string of length 500,
Population size: Table I and Table II,
Neighborhood size: 10% of the population size,
Termination condition: 400,000 solution evaluations,
Crossover probability: 0.8 (Uniform crossover),
Mutation probability: 2/500 (Bit-flip mutation),
Repair: Greedy repair based on the maximum profit ratio [19].

As a scalarizing function, we used the above-mentioned 16 settings: the weighed sum, the weighted Tchebycheff with α = 1, 1.01, 1.1, and the PBI with θ = 0.1, 1, 5, 10 for α = 1, 1.01, 1.1. In Tables III and IV, we show the average hypervolume value over 10 runs by each setting on each test problem. Small and large populations in Tables I and II were used in Tables III and IV, respectively. The hypervolume was calculated from the origin of the objective space. In each table, the best three results for each test problem are shown by red, the best result is underlined, and the worst result is shown by blue.

The best results on many-objective problems in Table III (small populations) were obtained by the weighted sum while the weighted Tchebycheff with α = 1.1 was the best in Table IV (large populations). This observation suggests that the choice of a scalarizing function depends on the population size. In other words, performance comparison results depend on the specification of the population size. From Table III and Table IV, we can also obtain the following observations.

Weighted sum: The weighted sum is a good choice for many-objective problems with 4-6 objectives in both tables. However, it is not a good choice for the two-objective problem.

Weighted Tchebycheff: The increase of the value of α from 1 to 1.01 and 1.1 improves the average hypervolume for all test problems in both tables. The weighted Tchebycheff with α = 1.1 is a good choice in both tables except for the six-objective problem in Table III with small populations.

PBI: The PBI with θ = 1 and α = 1 is always the worst in both table. This observation is consistent with the poor results of the weighted Tchebycheff with α = 1. The effect of the value of θ is not monotonic. Better results were obtained from θ = 0.1, 5, 10 than the case of θ = 1 for all test problems in both tables. The PBI with α = 1.1 and θ = 10 is a good choice for the problems with 2-3 objectives while α = 1.1 and θ = 0.1 are good specifications for the problems with 5-6 objectives.

TABLE III. AVERAGE HYPERVOLUME BY MOEA/D WITH SMALL POPULATION (KNAPSACK PROBLEMS).

Scalarizing Function	2-obj.	3-obj.	4-obj.	5-obj.	6-obj.
	x10^8	x10^{12}	x10^{17}	x10^{21}	x10^{25}
Weighted Sum	3.991	**7.637**	**1.439**	**2.664**	**4.800**
Tchebycheff: α = 1.00	3.844	7.141	1.269	2.285	4.015
Tchebycheff: α = 1.01	3.980	7.416	1.331	2.342	4.079
Tchebycheff: α = 1.10	**4.024**	**7.706**	**1.436**	**2.607**	4.521
PBI: α = 1.00, θ = 0.1	3.952	7.448	1.395	2.577	**4.640**
PBI: α = 1.00, θ = 1	**3.624**	**6.329**	**1.084**	**1.900**	**3.232**
PBI: α = 1.00, θ = 5	3.781	6.971	1.210	2.080	3.427
PBI: α = 1.00, θ = 10	3.815	7.123	1.230	2.114	3.474
PBI: α = 1.01, θ = 0.1	3.938	7.429	**1.396**	2.578	4.634
PBI: α = 1.01, θ = 1	3.704	6.478	1.105	1.910	3.255
PBI: α = 1.01, θ = 5	3.939	7.176	1.242	2.133	3.489
PBI: α = 1.01, θ = 10	3.967	7.276	1.257	2.167	3.546
PBI: α = 1.10, θ = 0.1	3.953	7.411	1.395	**2.579**	**4.635**
PBI: α = 1.10, θ = 1	3.834	6.635	1.136	1.937	3.263
PBI: α = 1.10, θ = 5	**4.032**	7.601	1.289	2.151	3.592
PBI: α = 1.10, θ = 10	**4.029**	**7.665**	1.298	2.190	3.660

TABLE IV. AVERAGE HYPERVOLUME BY MOEA/D WITH LARGE POPULATION (KNAPSACK PROBLEMS).

Scalarizing Function	2-obj.	3-obj.	4-obj.	5-obj.	6-obj.
	x10^8	x10^{12}	x10^{17}	x10^{21}	x10^{25}
Weighted Sum	3.993	7.570	**1.412**	**2.593**	**4.669**
Tchebycheff: α = 1.00	3.825	7.048	1.261	2.261	3.915
Tchebycheff: α = 1.01	3.980	7.421	1.350	2.427	4.213
Tchebycheff: α = 1.10	**4.027**	**7.769**	**1.483**	**2.733**	**4.846**
PBI: α = 1.00, θ = 0.1	3.947	7.329	1.359	2.462	4.380
PBI: α = 1.00, θ = 1	**3.631**	**6.360**	**1.097**	**1.904**	**3.213**
PBI: α = 1.00, θ = 5	3.765	6.852	1.224	2.127	3.554
PBI: α = 1.00, θ = 10	3.801	7.022	1.250	2.169	3.633
PBI: α = 1.01, θ = 0.1	3.943	7.341	1.362	2.464	4.395
PBI: α = 1.01, θ = 1	3.702	6.526	1.126	1.952	3.257
PBI: α = 1.01, θ = 5	3.945	7.259	1.280	2.200	3.672
PBI: α = 1.01, θ = 10	3.968	7.322	1.301	2.244	3.741
PBI: α = 1.10, θ = 0.1	3.947	7.340	1.355	**2.478**	**4.414**
PBI: α = 1.10, θ = 1	3.844	6.751	1.153	2.001	3.326
PBI: α = 1.10, θ = 5	**4.038**	**7.696**	1.348	2.272	3.796
PBI: α = 1.10, θ = 10	**4.041**	**7.733**	**1.367**	2.310	3.873

We further examine these observations in the next three subsections using the relation between weight vectors and obtained solutions for each scalarizing function. As additional information, the average percentage of the number of different non-dominated solutions over the population size in the final population is shown for each setting. The number of different solutions is counted in the objective space.

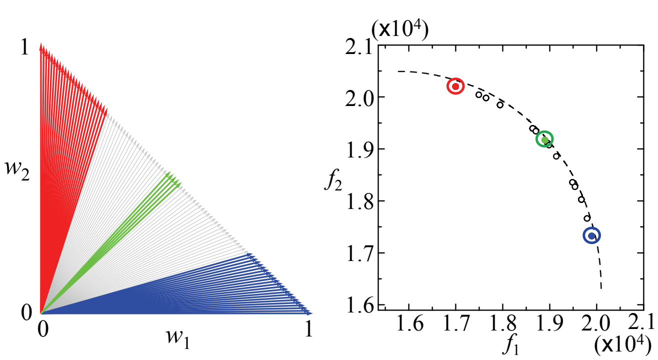
TABLE V. THE AVERAGE PERCENTAGE OF DIFFERENT NON-DOMINATED SOLUTIONS IN THE FINAL POPULATION WITH SMALL POPULATION (KNAPSACK PROBLEMS).

Scalarizing Function	2-obj.	3-obj.	4-obj.	5-obj.	6-obj.
Weighted Sum	17.5	48.2	63.7	73.1	75.1
Tchebycheff: $\alpha = 1.00$	75.5	93.0	96.5	97.5	97.5
Tchebycheff: $\alpha = 1.01$	81.6	92.8	90.1	94.2	97.9
Tchebycheff: $\alpha = 1.10$	47.6	61.9	61.5	40.1	59.4
PBI: $\alpha = 1.00$, $\theta = 0.1$	16.8	45.3	60.1	66.7	69.8
PBI: $\alpha = 1.00$, $\theta = 1$	7.6	26.4	21.8	12.4	12.3
PBI: $\alpha = 1.00$, $\theta = 5$	58.5	86.6	89.6	86.9	79.7
PBI: $\alpha = 1.00$, $\theta = 10$	62.5	84.9	91.3	91.4	88.1
PBI: $\alpha = 1.01$, $\theta = 0.1$	15.2	46.1	60.2	69.5	70.6
PBI: $\alpha = 1.01$, $\theta = 1$	30.9	36.3	24.3	16.0	14.0
PBI: $\alpha = 1.01$, $\theta = 5$	76.6	82.4	87.8	84.8	75.4
PBI: $\alpha = 1.01$, $\theta = 10$	74.6	82.4	87.9	88.7	83.1
PBI: $\alpha = 1.10$, $\theta = 0.1$	18.6	41.1	61.2	66.5	68.0
PBI: $\alpha = 1.10$, $\theta = 1$	30.8	34.8	20.2	14.4	7.9
PBI: $\alpha = 1.10$, $\theta = 5$	50.8	65.1	75.4	73.2	61.0
PBI: $\alpha = 1.10$, $\theta = 10$	50.6	61.3	79.3	78.1	69.3

B. Results by Weighted Sum

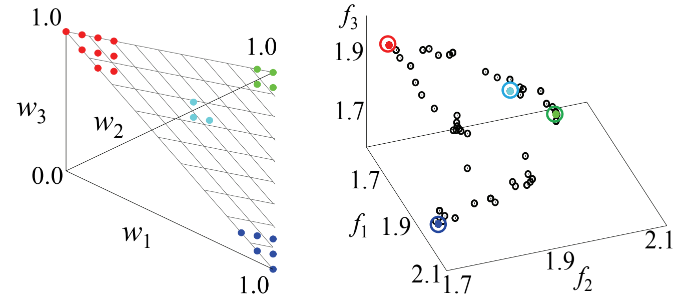
Fig. 2 shows the relation between the weight vectors and the solutions in the final population of a single run of MOEA/D with the weighted sum on the two-objective knapsack problem (population size 100). In Fig. 2 (b), the marked three solutions are the best solutions with respect to the first objective (blue), the second objective (red) and their sum (green). Fig. 2 (a) shows the weight vectors corresponding to each of those solutions (e.g., the red solution in Fig. 2 (b) is shared by many red weight vectors in Fig. 2 (a)). Since each solution is shared by multiple weight vectors, the total number of different solutions in Fig. 2 (b) is much smaller than the population size 100 (see Table V). This is the reason for poor performance of the weighted sum on the two-objective problem in Table III.

In the same manner, we show the results of a single run on the three-objective knapsack problem in Fig. 3. In Fig. 3 (b), the marked four solutions are the best solutions with respect to each individual objective and the sum of the three objectives.



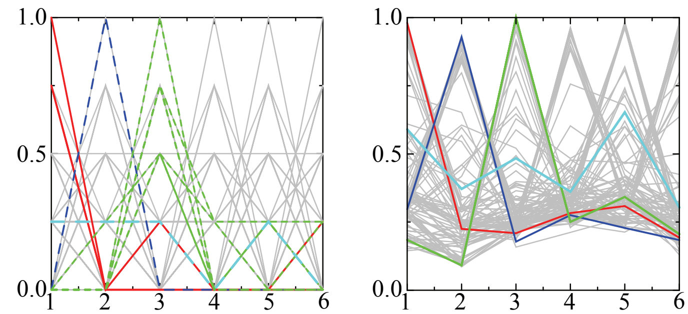
(a) Weight vectors. (b) Obtained solutions.

Fig. 2. Results by the weighted sum on the two-objective knapsack problem. In (a), weight vectors for each solution in (b) are shown by different colors.



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

Fig. 3. Results by the weighted sum on the three-objective knapsack problem.



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

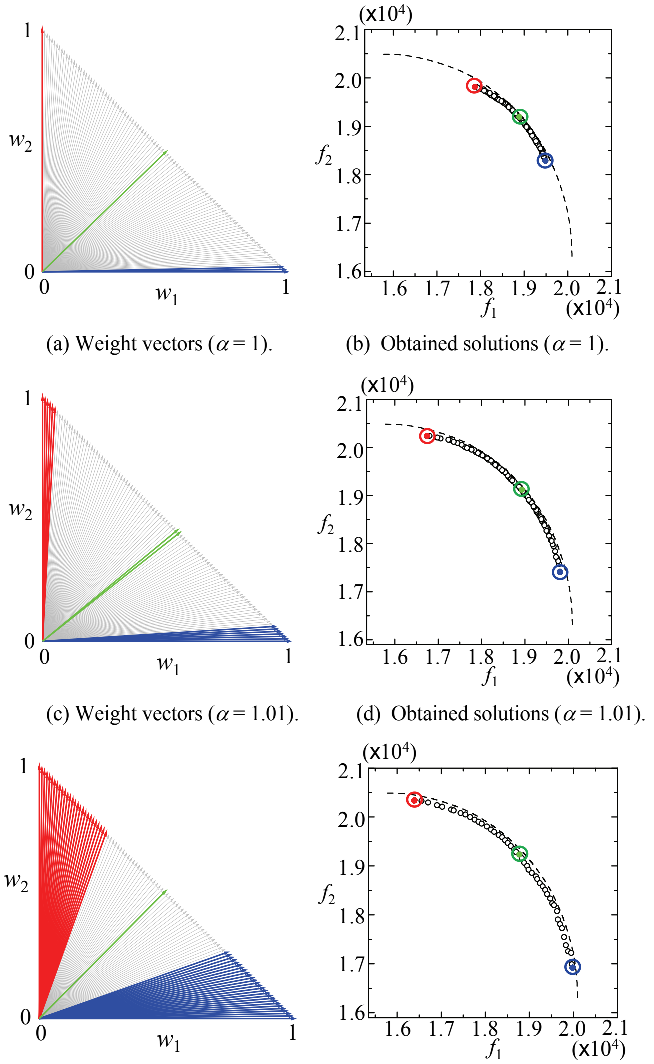
Fig. 4. Results by the weighted sum on the six-objective knapsack problem.

In Fig. 4, we show the results of a single run on the six-objective problem. Parallel coordinate plots are used for visualization. Each axis of the parallel coordinate plot in Fig. 4 (b) is normalized for each objective using all solutions obtained in our computational experiments in Table III and Table IV. In Fig. 4 (b), the highlighted four solutions are the best solutions with respect to each of the first three objectives and the sum of the six objectives. Whereas some solutions (e.g., red and green) are shared by multiple weight vectors, the number of those weight vectors is much smaller in Fig. 4 (a) than in Fig. 2 and Fig. 3 (see Table V). This is the reason for good results by the weighted sum on many-objective problems in Table III.

We can also see from Figs. 2-4 that the obtained solutions are directly related to the weight vectors in a straightforward manner. For example, by specifying the weight w_1 for the first objective $f_1(\boldsymbol{x})$ as $w_1 = 1.0$, the solutions with the best value of $f_1(\boldsymbol{x})$ were obtained in Figs. 2-4.

C. Results by Weighted Tchebycheff

For illustrating the effect of the location of the reference point on the obtained solutions by MOEA/D with the weighted Tchebycheff, we show results on the two-objective knapsack problem for the three settings of α: $\alpha = 1.0$, 1.01, 1.1 in Fig. 5. As shown in Fig. 5 (f), the use of $\alpha = 1.1$ clearly increased the diversity of solutions in comparison with the other cases in Fig. 5 (b) and Fig. 5 (d). We can also see that the two solutions with the best value of each objective were shared by many weight vectors in Fig. 5 (e). On the contrary, those solutions were shared by only a few weight vectors in the cases of $\alpha = 1.0$ in Fig. 5 (a). These observations suggest that many duplicated solutions in Fig. 5 (e) have a positive effect on the diversity of solutions in Fig. 5 (f).



(a) Weight vectors ($\alpha = 1$). (b) Obtained solutions ($\alpha = 1$).

(c) Weight vectors ($\alpha = 1.01$). (d) Obtained solutions ($\alpha = 1.01$).

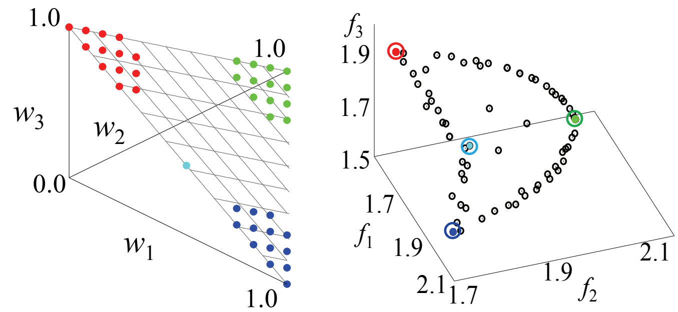
(e) Weight vectors ($\alpha = 1.1$). (f) Obtained solutions ($\alpha = 1.1$).

Fig. 5. Results by a single run of MOEA/D with the weighted Tchebycheff on the two-objective knapsack problem for each value of α.

In Fig. 6 and Fig. 7, we show the results on the three-objective and six-objective problems, respectively, for the case of $\alpha = 1.1$. As in Fig. 5, the colored solutions with the best values of each objective were shared by multiple weight vectors in Fig. 6 (a) and Fig. 7 (a). Those duplicated solutions have positive effects on the diversity of solutions. As a result, the highest average hypervolume was obtained from $\alpha = 1.1$ in Table III and Table IV among the three settings of α.

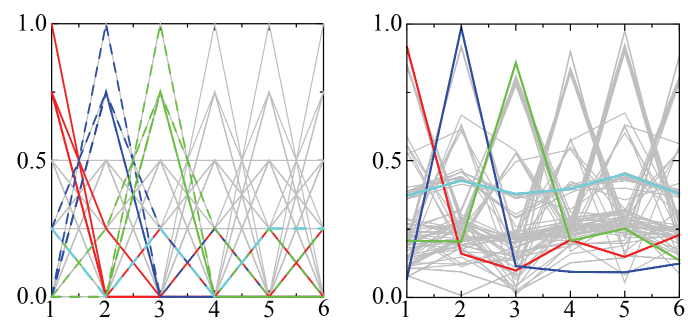
From the comparison between Fig. 3 (b) and Fig. 6 (b), we can see that a larger region of the Pareto front is covered by the obtained solutions in Fig. 6 (b) than Fig. 3 (b). This seems to be consistent with a large number of weight vectors in Fig. 6 (a) sharing the three extreme solutions in Fig. 6 (b). Actually the best result for the three objective problem was obtained by the weighted Tchebycheff with $\alpha = 1.1$ in Table III.

In Fig. 7 (b), solutions with good objective values close to 1 were not obtained except for $f_2(\boldsymbol{x})$ and $f_5(\boldsymbol{x})$. More solutions with bad objective values close to 0 were obtained. This is due to weak convergence ability of the weighted Tchebycheff for many-objective problems (see [15]). As a result, the weighted Tchebycheff is not a good choice in Table III for the six-objective problem. However, the weighted Tchebycheff with $\alpha = 1.1$ is the best choice in Table IV when the population size is large. In Fig. 8, we show the result of a single run with this setting in Table IV. A large diversity of solutions is observed in Fig. 8 (b). In Fig. 8 (a), many weight vectors share the extreme solution with the best value of each objective. The use of a large population clearly improved the result in Fig. 7 (b).



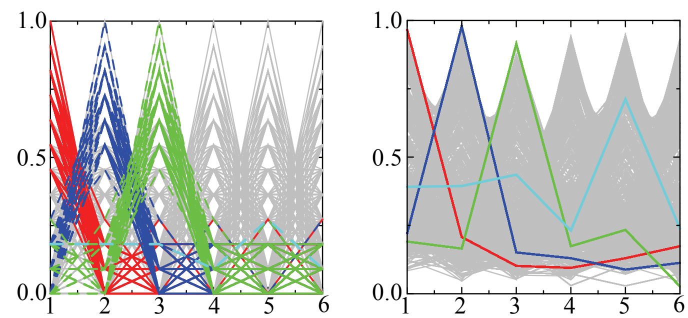
(a) Weight vectors. (b) Obtained solutions.

Fig. 6. Results of a single run by the weighted Tchebycheff with $\alpha = 1.1$ on the three-objective knapsack problem.



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

Fig. 7. Results of a single run by the weighted Tchebycheff with $\alpha = 1.1$ on the six-objective knapsack problem (the population size is 126).

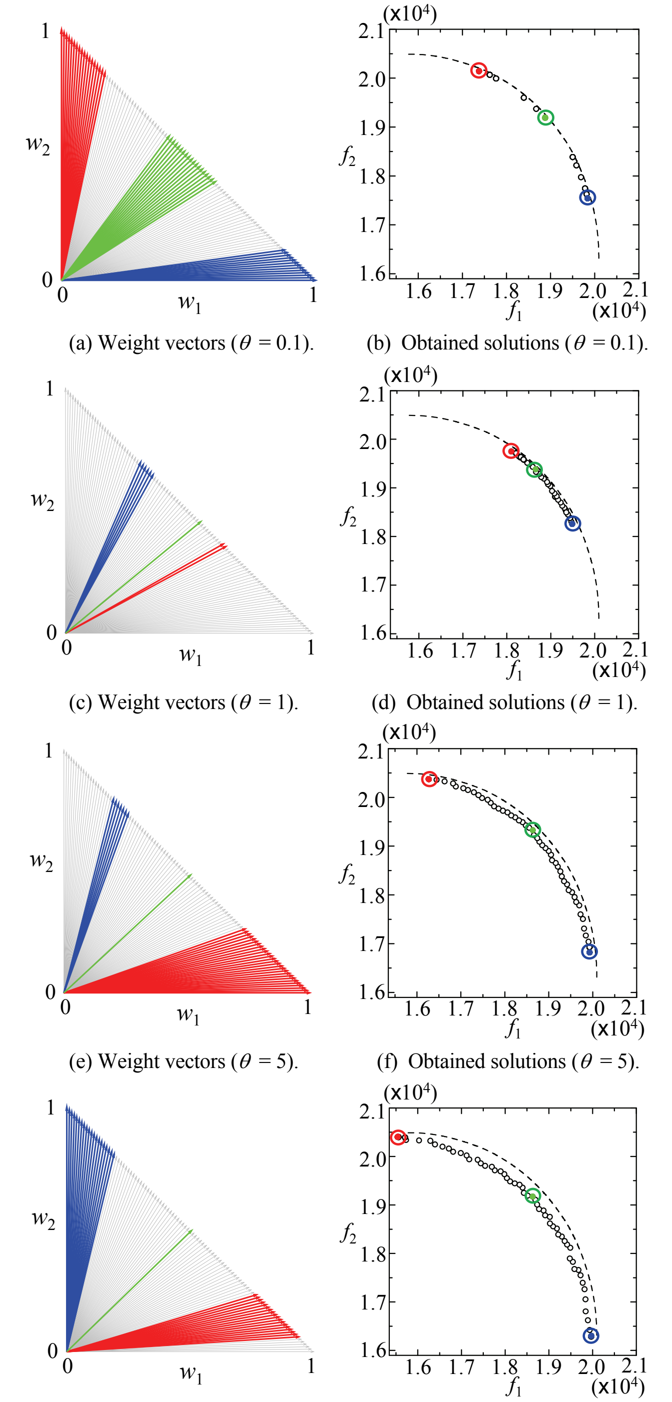


(a) Weight vectors. (b) Obtained solutions.

Fig. 8. Results of a single run by the weighted Tchebycheff with $\alpha = 1.1$ on the six-objective knapsack problem (the population size is 4368).

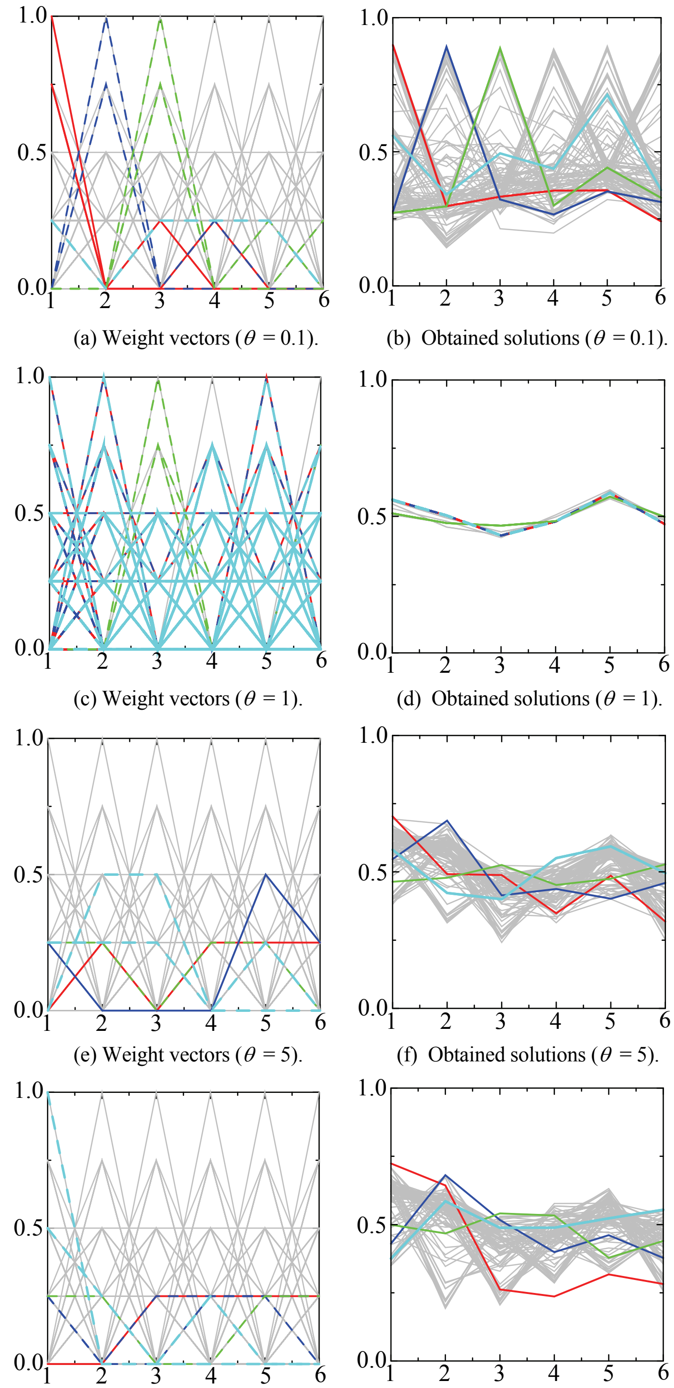
D. Results by PBI

Fig. 9 and Fig. 10 show the results of the PBI on the two-objective and six-objective knapsack problems, respectively. The value of α was specified as $\alpha = 1.1$. As shown in Fig. 9, totally different mappings from the weight vectors to the obtained solutions were realized depending on the value of θ.



(a) Weight vectors (θ = 0.1). (b) Obtained solutions (θ = 0.1).
(c) Weight vectors (θ = 1). (d) Obtained solutions (θ = 1).
(e) Weight vectors (θ = 5). (f) Obtained solutions (θ = 5).
(g) Weight vectors (θ = 10). (h) Obtained solutions (θ = 10).

Fig. 9. Results of a single run with the PBI on the two-objective knapsack problem. The value of α is specified as α = 1.1.



(a) Weight vectors (θ = 0.1). (b) Obtained solutions (θ = 0.1).
(c) Weight vectors (θ = 1). (d) Obtained solutions (θ = 1).
(e) Weight vectors (θ = 5). (f) Obtained solutions (θ = 5).
(g) Weight vectors (θ = 10). (h) Obtained solutions (θ = 10).

Fig. 10. Results of a single run with the PBI on the six-objective knapsack problem. The value of α is specified as α = 1.1.

When $\theta = 0.1$, the mapping is straightforward (i.e., the best solution with respect to the i-th objective was obtained when the corresponding weight w_i was 1.0). However, the mapping by $\theta = 10$ is totally different. The best objective value for each objective was obtained when the corresponding weight value was very small or zero. When $\theta = 1.0$, the diversity of solutions is small in Fig. 9 and Fig. 10. These observations suggest the difficulty of appropriately specifying the value of θ.

IV. Experimental Results on DTLZ2

Many-objective knapsack problems in Section III can be viewed as a representative of discrete many-objective maximization problems with convex Pareto fronts. In this section, we use DTLZ2 test problems with 2-6 objectives. DTLZ2 can be viewed as a representative of continuous many-objective minimization problems with concave Pareto fronts.

We use DTLZ2 since it has totally different characteristics from knapsack problems. In Fig. 11, we show randomly generated 100 solutions together with the Pareto front of the two-objective DTLZ2 problem. Fig. 11 may suggest that a good solution set will be easily obtained by simply pushing the initial population towards the Pareto front shown by the red curve. If compared with Fig. 1, Fig. 11 looks an easy problem.

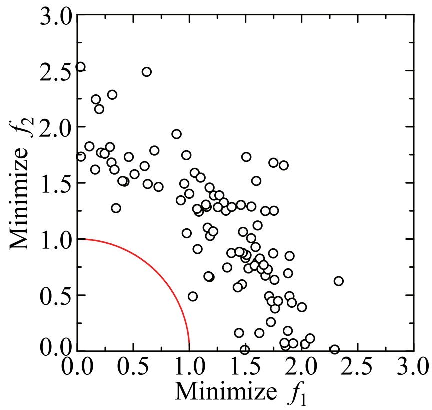


Fig. 11. Randomly generated 100 solutions (open circles) and the Pareto front (red curve) of the two-objective DTLZ2 problem.

MOEA/D with various scalarizing functions was applied to k-objective DTLZ2 problems under the following conditions:

Coding: Real number string of length $(9 + k)$,
Population size: Table I and Table II,
Neighborhood size: 10% of the population size,
Termination condition: 40,000 solution evaluations,
Crossover probability: 0.8 (SBX),
Mutation probability: $1/n$ where n is the string length (PM).

Due to the page limitation, experimental results only for small populations are shown in Table VI where the average hypervolume over 10 runs is reported for each test problem. The hypervolume was calculated for the reference point (1.1, 1.1, ..., 1.1). In Table VI, the best and worst results for each test problem are highlighted by red and blue, respectively. Table VII shows the average percentage of different non-dominated solutions in the final generation.

The worst results in Table VI were obtained from the weighted sum. This is because the weighted sum cannot handle concave Pareto fronts (see [25]). As shown in Table VII, many solutions were duplicated when the weighted sum was used. The best results in Table VI were obtained from the PBI with $\theta = 5$ independent of the value of α. Since initial solutions in Fig. 11 have larger diversity, no strong diversification property is needed. Thus the value of α has almost no effect in Table VI. On the contrary, the value of θ has a dominant effect on the performance of the PBI in Table VI. When θ is small, the PBI is similar to the weighted sum. Thus the PBI with a small value of θ did not work well on the DTLZ2 problems with concave Pareto fronts. The number of different non-dominated solutions was also very small as shown in Table VII when $\theta = 0.1$.

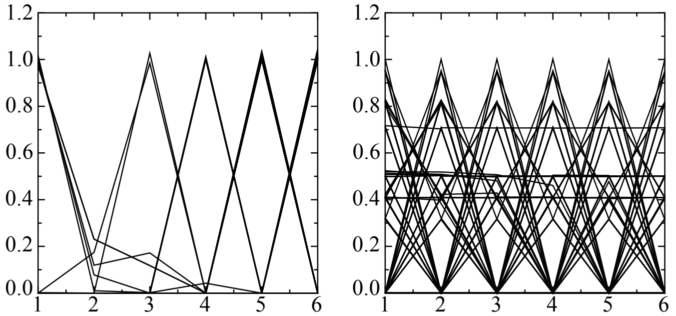
It is interesting to observe that almost the same results were obtained from the weighted sum and the weighted Tchebycheff for the six-objective DTLZ2 problem in Table VI and Table VII. Obtained solutions by a single run are shown in Fig. 12 for the weighted Tchebycheff ($\alpha = 1.0$) and the PBI ($\alpha = 1.0$ and $\theta = 5$). In Fig. 12 (b) with the PBI, all solutions are non-dominated (i.e., 100% non-dominated solutions in Table VII). The diversity of solutions in Fig. 12 (a) is much smaller than that in Fig. 12 (b). Moreover, some objective values in Fig. 12 (a) are larger than 1.0. That is, the weighted Tchebycheff has difficulties in both diversity and convergence. When large populations were used, the best results were obtained from the weighted Tchebycheff for DTLZ2 with 2-4 objectives. For 5-6 objectives, it was inferior to the PBI with $\alpha = 1.0$ and $\theta = 5$.

TABLE VI. AVERAGE HYPERVOLUME BY MOEA/D WITH SMALL POPULATION (DTLZ2).

Scalarizing Function	2-obj.	3-obj.	4-obj.	5-obj.	6-obj.
Weighted Sum	**0.210**	0.334	**0.470**	**0.625**	0.715
Tchebycheff: $\alpha = 0.90$	**0.420**	0.694	0.763	0.676	0.711
Tchebycheff: $\alpha = 0.99$	**0.420**	0.695	0.757	0.703	0.714
Tchebycheff: $\alpha = 1.00$	**0.420**	0.694	0.761	0.694	**0.706**
PBI: $\alpha = 0.90$, $\theta = 0.1$	**0.210**	**0.332**	0.472	0.653	0.773
PBI: $\alpha = 0.90$, $\theta = 1$	**0.420**	0.698	0.752	0.705	0.745
PBI: $\alpha = 0.90$, $\theta = 5$	**0.420**	**0.749**	**1.031**	**1.278**	1.510
PBI: $\alpha = 0.90$, $\theta = 10$	**0.420**	**0.749**	**1.031**	1.277	1.509
PBI: $\alpha = 0.99$, $\theta = 0.1$	**0.210**	**0.332**	0.472	0.650	0.760
PBI: $\alpha = 0.99$, $\theta = 1$	**0.420**	0.700	0.763	0.710	0.772
PBI: $\alpha = 0.99$, $\theta = 5$	**0.420**	**0.749**	**1.031**	**1.278**	**1.511**
PBI: $\alpha = 0.99$, $\theta = 10$	**0.420**	**0.749**	**1.031**	1.277	1.509
PBI: $\alpha = 1.00$, $\theta = 0.1$	**0.210**	**0.332**	0.472	0.650	0.771
PBI: $\alpha = 1.00$, $\theta = 1$	**0.420**	0.702	0.760	0.710	0.762
PBI: $\alpha = 1.00$, $\theta = 5$	**0.420**	**0.749**	**1.031**	**1.278**	**1.511**
PBI: $\alpha = 1.00$, $\theta = 10$	**0.420**	**0.749**	**1.031**	1.277	1.509

TABLE VII. THE AVERAGE PERCENTAGE OF DIFFERENT NON-DOMINATED SOLUTIONS IN THE FINAL POPULATION WITH SMALL POPULATION (DTLZ2).

Scalarizing Function	2-obj.	3-obj.	4-obj.	5-obj.	6-obj.
Weighted Sum	2.4	10.6	20.4	39.4	51.7
Tchebycheff: $\alpha = 0.90$	100.0	79.1	61.8	54.1	60.0
Tchebycheff: $\alpha = 0.99$	100.0	79.3	61.1	57.4	60.2
Tchebycheff: $\alpha = 1.00$	100.0	79.0	59.6	55.0	59.8
PBI: $\alpha = 0.90$, $\theta = 0.1$	2.0	7.4	17.7	24.5	27.9
PBI: $\alpha = 0.90$, $\theta = 1$	99.8	81.8	42.9	28.8	27.1
PBI: $\alpha = 0.90$, $\theta = 5$	100.0	100.0	100.0	100.0	100.0
PBI: $\alpha = 0.90$, $\theta = 10$	99.9	100.0	100.0	100.0	100.0
PBI: $\alpha = 0.99$, $\theta = 0.1$	2.0	7.4	16.9	25.3	27.9
PBI: $\alpha = 0.99$, $\theta = 1$	99.9	83.2	44.3	28.3	26.7
PBI: $\alpha = 0.99$, $\theta = 5$	99.9	100.0	100.0	100.0	100.0
PBI: $\alpha = 0.99$, $\theta = 10$	100.0	100.0	100.0	100.0	100.0
PBI: $\alpha = 1.00$, $\theta = 0.1$	2.0	7.4	17.2	24.0	27.7
PBI: $\alpha = 1.00$, $\theta = 1$	99.9	84.0	43.8	28.3	25.9
PBI: $\alpha = 1.00$, $\theta = 5$	99.9	100.0	100.0	100.0	100.0
PBI: $\alpha = 1.00$, $\theta = 10$	100.0	100.0	100.0	100.0	100.0



(a) Weighted Tchebycheff. (b) PBI.

Fig. 12. Obtained solutions by a single run with the weighted Tchebycheff (α = 1.0) and the PBI (α = 1.0 and θ = 5) on the six-objective DTLZ2 problem.

V. CONCLUDING REMARKS

We examined the mapping from weight vectors to solutions in MOEA/D. Due to the page limitation, we visually explained it only for knapsack problems with 2-6 objectives (discrete multiobjective maximization problems). Of course, we can perform the same examination for other test problems. One interesting observation was the existence of many duplicated solutions (i.e., many weight vectors sharing the same solution). Duplicated solutions seem to have a positive effect on the diversity of solutions of the knapsack problems with 2-3 objectives when each extreme solution with the best individual objective value is shared by many weight vectors. Another interesting observation was the very strong dependency of the mapping by the PBI function on the penalty parameter value θ. When θ was very small, the mapping was straightforward. An extreme solution with the best value of an individual objective was obtained by specifying the corresponding weight as its maximum value. However, such an extreme solution was obtained from small weight values when θ was large (e.g., $\theta = 5$ and $\theta = 10$). When $\theta = 1$, similar solutions were obtained from totally different weight vectors. As a result, the diversity of solutions was small. An unexpected observation was low percentages of non-dominated solutions in the final population when the weighted Tchebycheff was used for the DTLZ2 problems with 3-6 objectives (i.e., less than 80% of the population size). This issue should be further examined since all solutions in the final population were different non-dominated solutions (i.e., 100% non-dominated solutions) when we used the PBI with $\theta = 5$ and $\theta = 10$. Poor results of the weighted Tchebycheff together with small percentages of different non-dominated solutions clearly suggest the necessity of its improvement as explained in [21]-[24].

REFERENCES

[1] H. Ishibuchi, N. Tsukamoto, and Y. Nojima, "Evolutionary many-objective optimization: A short review," *Proc. of 2008 IEEE Congress on Evolutionary Computation*, 2424-2431, 2008.

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

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

[4] E. Zitzler, M. Laumanns, and L. Thiele, "SPEA2: Improving the strength Pareto evolutionary algorithm," *TIK-Report 103*, Department of Electrical Engineering, ETH, Zurich, 2001.

[5] E. Zitzler and S. Künzli, "Indicator-based selection in multiobjective search," *LNCS 3242: Parallel Problem Solving from Nature - PPSN VIII*, 832-842, 2004.

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

[7] J. Bader and E. Zitzler, "HypE: An algorithm for fast hypervolume-based many-objective optimization," *Evolutionary Computation* 19 (1) 45-76, 2011.

[8] E. Zitzler, D. Brockhoff, and L. Thiele, "The hypervolume indicator revisited: On the design of Pareto-compliant indicators via weighted integration," *LNCS 4403: Evolutionary Multi-Criterion Optimization - EMO 2007*, 862-876, 2007.

[9] Q. Zhang and H. Li, "MOEA/D: A multiobjective evolutionary algorithm based on decomposition," *IEEE Trans. on Evolutionary Computation* 11 (6) 712-731, 2007.

[10] H. Ishibuchi and T. Murata, "Multi-objective genetic local search algorithm," *Proc. of 1996 IEEE International Conference on Evolutionary Computation*, 119-124, 1996.

[11] H. Ishibuchi and T. Murata, "A multi-objective genetic local search algorithm and its application to flowshop scheduling," *IEEE Trans. on Systems, Man, and Cybernetics - Part C* 28 (3) 392-403, 1998.

[12] T. Murata, H. Ishibuchi, and M. Gen, "Specification of genetic search directions in cellular multi-objective genetic algorithm," *LNCS 1993*: *Evolutionary Multi-Criterion Optimization - EMO 2001*, 82-95, 2001.

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

[14] M. Asafuddoula, T. Ray, and R. Sarker, "A decomposition-based evolutionary algorithm for many objective optimization," *IEEE Trans. on Evolutionary Computation* 19 (3) 445-460, 2015.

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

[16] H. Ishibuchi, H. Masuda, and Y. Nojima, "Comparing solution sets of different size in evolutionary many-objective optimization," *Proc. of 2015 IEEE Congress on Evolutionary Computation*, 2859-2866, 2015.

[17] H. Sato, "Inverted PBI in MOEA/D and its impact on the search performance on multi and many-objective optimization," *Proc. of 2014 Genetic and Evolutionary Computation Conference,* 645-652, 2014.

[18] H. L. Liu, F. Gu, and Q. Zhang, "Decomposition of a multiobjective optimization problem into a number of simple multiobjective subproblems" *IEEE Trans. on Evolutionary Computation* 18 (3) 450-455, 2014.

[19] E. Zitzler and L. Thiele, "Multiobjective evolutionary algorithms: A comparative case study and the strength Pareto approach," *IEEE Trans. on Evolutionary Computation* 3 (4) 257-271, 1999.

[20] K. Deb, L. Thiele, M. Laumanns, and E. Zitzler, "Scalable multiobjective optimization test problems," *Proc. of 2002 IEEE Congress on Evolutionary Computation*, 825-830, 2002.

[21] L. Wang, Q. Zhang, A. Zhou, M. Gong, and L. Jiao, "Constrained subproblems in decomposition based multiobjective evolutionary algorithm," *IEEE Trans. on Evolutionary Computation* (Early Access).

[22] K. Li, Q. Zhang, S. Kwong, M. Li, and R. Wang, "Stable matching-based selection in evolutionary multiobjective optimization," *IEEE Trans. on Evolutionary Computation* 18 (6) 909-923, 2014.

[23] K. Li, S. Kwong, Q. Zhang, and K. Deb, "Interrelationship-based selection for decomposition multiobjective optimization," *IEEE Trans. on Cybernetics* 45 (10) 2076-2088, 2015.

[24] Y. Yuan, H. Xu, B. Wang, B. Zhang, and X. Yao, "Balancing convergence and diversity in decomposition-based many-objective optimizers," *IEEE Trans. on Evolutionary Computation* (Early Access).

[25] Y. Jin, M. Olhofer, and B. Sendhoff. "Dynamic weighted aggregation for evolutionary multiobjective optimization: Why does it work and how?" *Proc. of 2001 Genetic and Evolutionary Computation Conference*, 1042-1049, 2001.