

H. Ishibuchi, N. Tsukamoto, and Y. Nojima, "Diversity Improvement by Non-Geometric Binary Crossover in Evolutionary Multiobjective Optimization," *IEEE Transactions on Evolutionary Computation*, vol. 14, no. 6, pp. 985 - 998, December 2010. (Published)

Diversity Improvement by Non-Geometric Binary Crossover in Evolutionary Multiobjective Optimization

Hisao Ishibuchi, *Senior Member, IEEE*, Noritaka Tsukamoto, *Student Member, IEEE*,
and Yusuke Nojima, *Member, IEEE*

Graduate School of Engineering, Osaka Prefecture University
1-1 Gakuen-cho, Naka-ku, Sakai, Osaka 599-8531, Japan

Corresponding author: Prof. Hisao Ishibuchi

Graduate School of Engineering, Osaka Prefecture University,
1-1 Gakuen-cho, Naka-ku, Sakai, Osaka 599-8531, Japan
Phone +81-72-254-9350, FAX +81-72-254-9915
E-mail: hisaoui@cs.osakafu-u.ac.jp

Diversity Improvement by Non-Geometric Binary Crossover in Evolutionary Multiobjective Optimization

Hisao Ishibuchi, *Senior Member, IEEE*, Noritaka Tsukamoto, *Student Member, IEEE*,
and Yusuke Nojima, *Member, IEEE*

Graduate School of Engineering, Osaka Prefecture University
1-1 Gakuen-cho, Sakai, Osaka 599-8531, Japan

Abstract - In the design of evolutionary multiobjective optimization (EMO) algorithms, it is important to strike a balance between diversity and convergence. Traditional mask-based crossover operators for binary strings (e.g., one-point, two-point and uniform) tend to decrease the spread of solutions along the Pareto front in EMO algorithms while they improve the convergence to part of the Pareto front. This is because such a crossover operator, which is called geometric crossover, always generates an offspring in the segment between its two parents under the Hamming distance in the genotype space. That is, the sum of the distances from the generated offspring to its two parents is always equal to the distance between the two parents. In this paper, we first propose a non-geometric binary crossover operator to generate an offspring outside the segment between its two parents. Next we show some properties of our crossover operator. Then we examine its effects on the behavior of EMO algorithms through computational experiments on knapsack problems with two, four and six objectives. Experimental results show that our crossover operator can increase the spread of solutions along the Pareto front in EMO algorithms without severely degrading their convergence property. As a result, our crossover operator improves some overall performance measures such as the hypervolume.

Index Terms - Evolutionary multiobjective optimization (EMO), geometric crossover, non-geometric crossover, diversity maintenance, multiobjective knapsack problems.

I. Introduction

Evolutionary multiobjective optimization (EMO) algorithms have been successfully used in various application areas [4], [31]. EMO algorithms are designed to find a set of well-distributed Pareto-optimal or near Pareto-optimal solutions with a wide range of objective values, which approximates the entire Pareto front of a multiobjective problem. It is important in the design of EMO algorithms to strike a balance between diversity and convergence [2], [17].

Usually no *a priori* information about the decision maker's preference is used when an EMO algorithm searches for Pareto-optimal solutions. A set of non-dominated solutions is presented to the decision maker as a result of the search by the EMO algorithm. The decision maker is supposed to

choose one of the presented non-dominated solutions. The choice of a final solution is performed based on the decision maker's preference. The EMO approach, which consists of the search for a number of non-dominated solutions and the choice of a final solution, is referred to as an ideal multiobjective optimization procedure in Deb [4]. It is implicitly assumed in the EMO approach that the choice of the most preferred one from the obtained non-dominated solutions is much easier for the decision maker than the elicitation of his/her preference before the search for non-dominated solutions (see [1] for various multicriteria decision making strategies).

It is essential for the success of the EMO approach to find a set of non-dominated solutions that well approximates the entire Pareto front. It is, however, not easy (usually very difficult) for EMO algorithms to find such a good non-dominated solution set of a large-scale combinatorial multiobjective problem as pointed out in some studies (e.g., Jaszkievicz [18]). This is the case even when multiobjective problems have only two objectives.

It was visually demonstrated [13] that crossover had a negative effect on the spread of solutions along the Pareto front (while it improved the convergence to part of the Pareto front) when EMO algorithms were applied to large-scale two-objective 0/1 knapsack problems. Several ideas of recombining similar parents were proposed to improve the performance of EMO algorithms by decreasing such a negative effect of crossover and increasing its positive effect [14], [15], [20], [28], [35]. It should be noted that crossover does not decrease the diversity of binary strings whereas it decreases the diversity of objective vectors. Emphasis is usually placed on diversity maintenance of objective vectors in the design of EMO algorithms rather than diversity maintenance of binary strings. This is different from the concept of diversity maintenance in single-objective optimization (see [19] for various differences between single-objective and multiobjective evolutionary algorithms).

Recently the concept of geometric crossover was proposed by Moraglio and Poli [21]-[23] to analyze crossover operators in terms of the distances among an offspring and its parents. Roughly speaking, a crossover operator is referred to as being geometric when the following relation always holds among an offspring C and its two parents $P1$ and $P2$ for a distance function d :

$$d(P1, C) + d(P2, C) = d(P1, P2), \quad (1)$$

where $d(A, B)$ denotes the distance between A and B in the genotype space (for details, see [21]-[23]).

Traditional mask-based crossover operators for binary strings (e.g., one-point, two-point and uniform) are geometric crossover [21], [22] because the relation in (1) always holds for the Hamming distance. On the other hand, many crossover operators for real number strings (i.e., real-coded genetic algorithms) such as simulated binary crossover (SBX [5], [7]), blend crossover (BLX- α [10]), extended line crossover [24], unimodal normal distribution crossover (UNDX [25]), and linear crossover [36] are non-geometric [23]. That is, such a crossover operator can generate an offspring C satisfying the following relation for its two parents $P1$ and $P2$ under the Euclidean distance d :

$$d(P1, C) + d(P2, C) > d(P1, P2). \quad (2)$$

Geometric crossover for real number strings has a bias toward the center of the feasible region in the continuous decision space [39]. Thus we usually need non-geometric crossover in real-coded genetic algorithms. Deb et al. [6] proposed a class of high-performance non-geometric crossover operators for real-coded genetic algorithms called parent-centric crossover (PCX), which was explained in contrast with mean-centric crossover. It was demonstrated through computational experiments in [6] that PCX outperformed some other crossover operators for real-coded genetic algorithms (e.g., UNDX).

On the other hand, geometric crossover is usually used together with bit-flip mutation in genetic algorithms with binary strings. Of course, there exist genetic algorithms with no mutation. Those genetic algorithms have some sort of safeguard mechanisms against premature convergence. For example, CHC [9], which does not use mutation to modify offspring generated by crossover, has a restart mechanism for partially randomizing the current population when convergence is detected.

In some studies, the use of non-geometric binary crossover was examined implicitly (i.e., without explicitly mentioning the non-geometric property of crossover operators). For example, it was reported in [29] that the transportation operator outperformed one-point, two-point and uniform crossover in computational experiments on 18 single-objective test problems. The transportation operator exchanges substrings between two parents. Its main characteristic is that a substring from one parent can be transported to a different position of the other parent. As a result, this operator can be viewed as a non-geometric binary crossover operator as shown later in Section III of this paper.

When geometric binary crossover operators are used in EMO algorithms for multiobjective optimization, they often decrease the spread of solutions along the Pareto front [13]. For diversity maintenance, we propose the probabilistic use of geometric and non-geometric crossover operators in EMO algorithms with binary strings. In this paper, first we propose a non-geometric binary crossover operator in Section II. The main characteristic of our crossover operator is to intentionally generate an offspring outside the segment between its two parents in the Hamming distance space. Its basic idea is similar to PCX [6]. Next we show properties of our crossover operator in Section III. Then we examine its effects on the behavior of EMO algorithms through computational experiments on two-objective problems in Section IV and many-objective problems in Section V. Experimental results are discussed in Section VI. Finally we conclude this paper in Section VII.

II. Non-Geometric Binary Crossover

Let \mathbf{x} and \mathbf{y} be two real number vectors (i.e., two real number strings). A simple line crossover operator for generating an offspring \mathbf{z} from the two parents \mathbf{x} and \mathbf{y} can be written as

$$\mathbf{z} = \alpha \mathbf{x} + (1 - \alpha) \mathbf{y}, \quad (3)$$

where α is a randomly specified real number. If α is always in the unit interval $[0, 1]$, this crossover is geometric under the Euclidean distance. On the other hand, this crossover is not geometric under the

Euclidean distance when α may assume a real number outside the unit interval $[0, 1]$. In the latter case, the crossover operator in (3) is referred to as extended line crossover [24], which can generate an offspring C satisfying the inequality relation in (2) under the Euclidean distance.

Non-geometric crossover operators under the Euclidean distance have been used in almost all real-coded genetic algorithms as in [5], [7], [10], [24], [25], [36]. On the other hand, traditional mask-based binary crossover operators are geometric under the Hamming distance. We illustrate uniform crossover in Fig. 1 where an offspring C is generated from two parents P1 and P2. In Fig. 1, the Hamming distance between the two parents (i.e., 10) is equal to the sum of the Hamming distances from the offspring C to its two parents P1 and P2 (i.e., 5 + 5). The same uniform crossover operator is illustrated in Fig. 2 in the Hamming distance space. The horizontal axis of Fig. 2 is the Hamming distance from Parent 1 (P1) to the offspring C while its vertical axis is the Hamming distance from Parent 2 (P2) to C. The offspring C in Fig. 1 is located at the point (5, 5) in Fig. 2. Each open circle in Fig. 2 shows a possible location of an offspring that can be generated from P1 and P2 in Fig. 1 by geometric crossover. As shown in Fig. 2, an offspring is always generated in the segment between its parents P1 and P2 in the Hamming distance space. The three arrows from C in Fig. 2 show possible moves by bit-flip mutation of a single bit of the offspring C in Fig. 1.

	* * * * *	* * * * *
Parent 1 (P1)	0000000000	1111111111
Offspring (C)	0000000000	0110010101
Parent 2 (P2)	0000000000	0000000000

Fig. 1. Illustration of uniform crossover.

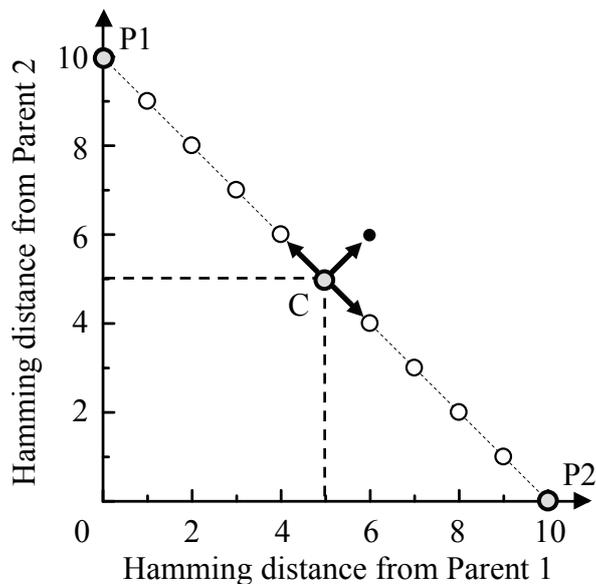


Fig. 2. Relation between the offspring C in Fig. 1 and its two parents P1 and P2 in the Hamming distance space.

As shown in Fig. 2, traditional mask-based binary crossover operators always generate an offspring in the segment between its two parents in the Hamming distance space. This corresponds to the situation where α is always in the unit interval $[0, 1]$ in line crossover in (3). In order to generate an offspring outside the segment between its two parents, we propose a non-geometric binary crossover operator which corresponds to the situation where α is outside the unit interval $[0, 1]$ in line crossover.

Let \mathbf{x} and \mathbf{y} be two binary strings of length n . We denote them as $\mathbf{x} = x_1 x_2 \dots x_n$ and $\mathbf{y} = y_1 y_2 \dots y_n$, respectively. Our non-geometric binary crossover operator generates an offspring $\mathbf{z} = z_1 z_2 \dots z_n$ satisfying the following relation for the Hamming distance d :

$$d(\mathbf{x}, \mathbf{z}) + d(\mathbf{y}, \mathbf{z}) > d(\mathbf{x}, \mathbf{y}). \quad (4)$$

Our idea is to generate an offspring from one parent in the opposite side of the other parent. First one parent is chosen as a primary parent (say \mathbf{x}). In this paper, we use the better parent (with respect to the fitness evaluation scheme of NSGA-II [8]) as the primary parent because better results were obtained from this better parent choice strategy than random parent choice in our preliminary experiments. The other parent is used as a secondary parent (say \mathbf{y}). Then an offspring \mathbf{z} is generated from the primary parent \mathbf{x} and the secondary parent \mathbf{y} in the following manner:

Non-Geometric Binary Crossover Operator

Step 1: Let $z_i := x_i$ for $i=1, 2, \dots, n$ (i.e., \mathbf{z} is a copy of the primary parent \mathbf{x}).

Step 2: Let $z_i := (1 - z_i)$ for $i=1, 2, \dots, n$ with a bit-flip probability P_{BF} only when $x_i = y_i$.

Our non-geometric binary crossover operator is illustrated in Fig. 3 where Parent 1 is used as the primary parent \mathbf{x} . First an intermediate offspring \mathbf{z} in Fig. 3 is generated as a copy of the primary parent \mathbf{x} in Step 1. Then the bit-flip operator is applied to z_i with a pre-specified bit-flip probability P_{BF} only when $x_i = y_i$ in Step 2. In Fig. 3, such a probabilistic bit-flip operator is applied to the first ten values of \mathbf{z} since the two parents \mathbf{x} and \mathbf{y} are the same at their first ten loci.

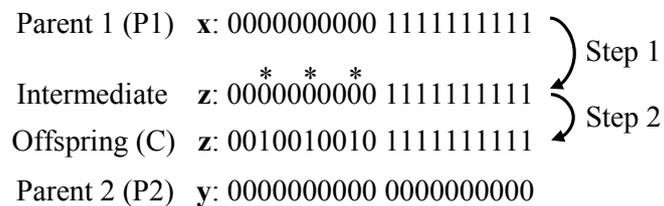


Fig. 3. Our non-geometric binary crossover operator.

One may think that our genetic operator is not crossover but mutation. In general, mutation maps a single string to a single offspring while crossover maps two (or more) parents to a single (or more) offspring. Since our genetic operator needs two parents, it can be viewed as crossover.

If we discuss our genetic operator as mutation, its main characteristic is the use of a different

mutation probability for each locus and each allele. Tan et al. [30], [32] assigned a different mutation probability to each locus for local exploration in EMO algorithms. For single-objective optimization, a number of adaptation mechanisms for mutation probabilities have been proposed in the literature. Those mechanisms usually use statistics about the current population such as the frequency of “1” at each locus and the average fitness over those strings with “1” at a specific locus [33], [37], [38]. Our genetic operator can be viewed as implicitly adjusting a mutation probability for each locus in a very simple manner without the use of population statistics.

In Fig. 3, the Hamming distance between Offspring C and Parent 2 is 13 while the distance between the two parents is 10. The generated offspring C in Fig. 3 is depicted in the Hamming distance space in Fig. 4. Its location shows its Hamming distances from Parent 1 and Parent 2. When Parent 1 is a primary parent, our crossover operator can generate an offspring at one of the open circles in the left-upper part of Fig. 4. On the other hand, the open circles in the right-lower part show possible offspring by our crossover operator when Parent 2 is a primary parent. The three arrows from C in Fig. 4 show possible moves by bit-flip mutation of a single bit of the offspring C as in Fig. 2.

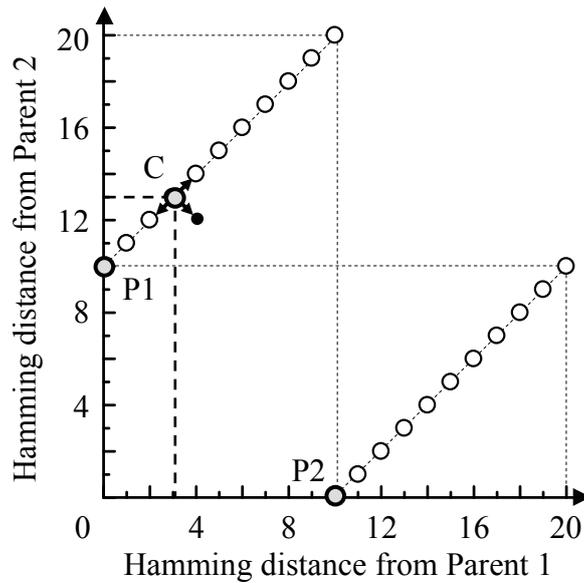


Fig. 4. Relation between the offspring C generated by our non-geometric binary crossover operator and its parents P1 and P2.

We propose the probabilistic use of geometric and non-geometric crossover operators. In this paper, we always use uniform crossover as a geometric crossover operator. Let P_C be the crossover probability. After a pair of parents is selected to generate an offspring, we first decide whether we use crossover (with the probability P_C) or not (with the probability $1-P_C$). Then we probabilistically choose one of the non-geometric and uniform crossover operators when crossover is to be used. Let us denote the probabilities of the non-geometric and uniform crossover operators by P and $(1-P)$,

respectively. In this case, the overall probabilities of the three choices for each pair of parents are calculated as follows: The non-geometric crossover operator with the probability $P \cdot P_C$, the uniform crossover operator with $(1-P) \cdot P_C$, and no crossover with $(1-P_C)$.

III. Properties of Our Crossover Operator

In this section, we explain the following properties of our non-geometric binary crossover operator:

- (i) Our crossover operator is not geometric for any distance function.
- (ii) The following relation always holds for the Hamming distance d among the primary parent P1, the secondary parent P2, and their offspring C generated by our crossover operator:

$$d(P1, P2) + d(P1, C) = d(P2, C). \quad (5)$$

- (iii) The primary parent P1 can be generated from P2 and C by the uniform crossover operator.
- (iv) The secondary parent P2 can be generated from P1 and C by our crossover operator.
- (v) Iterative use of our crossover operator, which starts from arbitrary two binary strings of length n , can generate any binary string of the same length.

Let \mathbf{x} and \mathbf{y} be the two parents of an offspring \mathbf{z} where \mathbf{x} , \mathbf{y} and \mathbf{z} are binary strings of length n : $\mathbf{x} = x_1 x_2 \dots x_n$, $\mathbf{y} = y_1 y_2 \dots y_n$ and $\mathbf{z} = z_1 z_2 \dots z_n$. A crossover operator is referred to as “geometric” with respect to a distance function d when the following relation always holds for arbitrary two strings \mathbf{x} and \mathbf{y} and their offspring \mathbf{z} [21]-[23]:

$$d(\mathbf{x}, \mathbf{z}) + d(\mathbf{y}, \mathbf{z}) = d(\mathbf{x}, \mathbf{y}). \quad (6)$$

On the other hand, a crossover operator is referred to as “non-geometric” if it is not geometric with respect to any distance function d (i.e., if there is no distance function d that always satisfies (6)).

Let us consider a special case where two parents are the same (i.e., $\mathbf{x} = \mathbf{y}$). In this case, (6) can be rewritten as follows for any offspring \mathbf{z} generated by a geometric crossover operator:

$$2d(\mathbf{x}, \mathbf{z}) = d(\mathbf{x}, \mathbf{y}) = 0. \quad (7)$$

This means that $d(\mathbf{x}, \mathbf{z}) = d(\mathbf{x}, \mathbf{y}) = 0$ (i.e., $\mathbf{x} = \mathbf{y} = \mathbf{z}$). That is, $\mathbf{x} = \mathbf{y} = \mathbf{z}$ always holds among two parents \mathbf{x} and \mathbf{y} and their offspring \mathbf{z} generated by a geometric crossover operator when $\mathbf{x} = \mathbf{y}$.

On the other hand, our non-geometric binary crossover operator can generate an offspring \mathbf{z} from its two parents \mathbf{x} and \mathbf{y} satisfying $\mathbf{x} = \mathbf{y} \neq \mathbf{z}$. In this case, the following relation always holds for any distance function d :

$$d(\mathbf{x}, \mathbf{y}) = 0 < d(\mathbf{x}, \mathbf{z}) + d(\mathbf{y}, \mathbf{z}). \quad (8)$$

This means that our crossover operator is not geometric for any distance function d .

Let us discuss geometric and non-geometric binary crossover operators by focusing on the values x_i , y_i and z_i at the i th locus of the two parents \mathbf{x} and \mathbf{y} and their offspring \mathbf{z} . In standard mask-based binary crossover (e.g., one-point, two-point, uniform), z_i is inherited from x_i or y_i . Thus $z_i = x_i$ or $z_i = y_i$ always holds. In this case, we can easily show that $d(x_i, z_i) + d(y_i, z_i) = d(x_i, y_i)$ holds for any distance functions. Thus we have $d(\mathbf{x}, \mathbf{z}) + d(\mathbf{y}, \mathbf{z}) = d(\mathbf{x}, \mathbf{y})$ for the Hamming distance. From these discussions, we can see that mask-based crossover (satisfying $z_i = x_i$ or $z_i = y_i$) is geometric. In Table 1, we show all the eight combinations of x_i, y_i and z_i . As shown in Table 1, six out of the eight combinations satisfy $z_i = x_i$ or $z_i = y_i$. If an binary crossover operator realizes only those combinations, it is geometric since $d(x_i, z_i) + d(y_i, z_i) = d(x_i, y_i)$ holds from $z_i = x_i$ or $z_i = y_i$. On the other hand, if an binary crossover operator can realize at least one of the other two combinations (i.e., $(x_i, y_i, z_i) = (0, 0, 1), (1, 1, 0)$), it is non-geometric since $d(x_i, z_i) + d(y_i, z_i) > d(x_i, y_i) = 0$ can be derived from $x_i = y_i \neq z_i$. Our crossover operator, which is non-geometric, can realize both of these two combinations.

The transportation operator [29] mentioned in Section I is also non-geometric. This operator can move x_j at the j th locus of \mathbf{x} to z_i at the i th locus of \mathbf{z} . Thus $x_i = y_i \neq z_i$ can be held among an offspring \mathbf{z} and its two parents \mathbf{x} and \mathbf{y} . This means that the transportation operator is non-geometric.

Table 1. All combinations of x_i, y_i and z_i .

x_i	y_i	z_i	$z_i = x_i$ or $z_i = y_i$
0	0	0	$x_i = y_i = z_i$
0	0	1	-
0	1	0	$z_i = x_i$
0	1	1	$z_i = y_i$
1	0	0	$z_i = y_i$
1	0	1	$z_i = x_i$
1	1	0	-
1	1	1	$x_i = y_i = z_i$

Our crossover operator satisfies the following relations for all i 's when \mathbf{x} is the primary parent:

Case 1: If $x_i \neq y_i$ then $z_i = x_i$.

Case 2: If $x_i = y_i$ then $x_i = y_i \neq z_i$ or $x_i = y_i = z_i$.

From these relations, we have the following relation:

$$\forall i; x_i = z_i \text{ or } x_i = y_i. \quad (9)$$

In this case, $d(x_i, y_i) + d(x_i, z_i) = d(y_i, z_i)$ holds. Thus we have $d(\mathbf{x}, \mathbf{y}) + d(\mathbf{x}, \mathbf{z}) = d(\mathbf{y}, \mathbf{z})$ for the Hamming distance. This relation is rewritten as follows (see Fig. 5 (a)):

$$d(\mathbf{P1}, \mathbf{P2}) + d(\mathbf{P1}, \mathbf{C}) = d(\mathbf{P2}, \mathbf{C}), \quad (10)$$

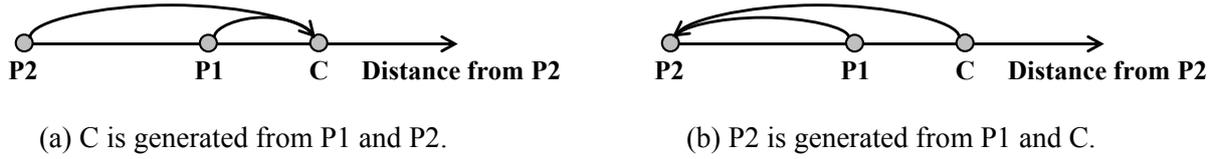


Fig. 5. Non-geometric crossover. In (a), P1 is the primary parent, P2 is the secondary parent, and C is their offspring. In (b), P1 is the primary parent, C is the secondary parent, and P2 is their offspring.

Eq. (9) shows that \mathbf{x} can be generated from \mathbf{y} and \mathbf{z} by the uniform crossover operator. That is, the primary parent P1 can be generated from the secondary parent P2 and the offspring C by the uniform crossover operator. In this case, (10) shows the geometric property among the two parents P2 and C and their offspring P1 (see Fig. 5 where P1 is in the segment between P2 and C).

Now let us explain that the secondary parent \mathbf{y} (P2) can be generated by our crossover operator from the primary parent \mathbf{x} (P1) and the offspring \mathbf{z} (C) as shown in Fig. 5 (b). Let $\mathbf{v} = v_1 v_2 \dots v_n$ be an offspring of \mathbf{x} and \mathbf{z} by our crossover operator with the primary parent \mathbf{x} . In the first step of our crossover operator, \mathbf{v} is a copy of \mathbf{x} . Thus $v_i = x_i$ for all i 's. In the second step, the bit-flip operator is probabilistically applied to v_i for each locus i if $x_i = z_i$. In order to make $\mathbf{v} = \mathbf{y}$ hold, we have to apply the bit-flip operator to all v_i such that $v_i \neq y_i$ (i.e., $x_i \neq y_i$). This is possible because $x_i = z_i$ always holds if $x_i \neq y_i$ (see Case 1 in the above). These discussions show that we can generate \mathbf{y} (i.e., P2) from \mathbf{x} (i.e., P1) and \mathbf{z} (i.e., C) by our crossover operator as shown in Fig. 5 (b).

An interesting property of our crossover operator is that its iterative use starting from arbitrary two parents of length n can generate any binary string of the same length. Let \mathbf{x} , \mathbf{y} and \mathbf{z} be the primary parent, the secondary parent, and their offspring, respectively. If the bit-flip operator in the second step is not applied to any z_i , \mathbf{z} is the same as \mathbf{x} (i.e., $x_i = z_i$ for all i 's). When \mathbf{x} and \mathbf{z} (such that $\mathbf{x} = \mathbf{z}$) are used as parents in our crossover operator in the next iteration, the bit-flip operator in the second step is probabilistically applied to all bits. Thus any binary string of the same length as \mathbf{x} and \mathbf{z} can be generated by our crossover operator from \mathbf{x} and \mathbf{z} . It should be noted that no geometric binary crossover operator has this property. This is because no geometric binary crossover operator can realize all the eight combinations in Table 1.

IV. Performance Evaluation on Two-Objective Problems

In this section, we examine the effect of our non-geometric binary crossover operator on the performance of NSGA-II [8] through computational experiments on a two-objective 500-item 0/1 knapsack problem in [43]. We used the following parameter specifications: The population size was 200, the crossover probability was 0.8 (i.e., $P_C = 0.8$), the mutation probability was 0.002 (i.e., $1/500$), and the termination condition was 2000 generations. When crossover was to be used (with the crossover probability 0.8), the non-geometric and uniform crossover operators were chosen with the

probabilities P and $(1-P)$, respectively. We examined six values of P : $P=0.0, 0.2, 0.4, 0.6, 0.8, 1.0$. In the second step of our non-geometric binary crossover operator, the bit-flip operator was applied with the probability P_{BF} . We examined eight values of P_{BF} : $P_{BF}=0.002, 0.004, 0.008, 0.012, 0.016, 0.020, 0.040, 0.080$. NSGA-II with each of the 6×8 combinations of P and P_{BF} was applied to the test problem 100 times. When infeasible solutions were generated, they were repaired into feasible ones by the greedy repair scheme based on the maximum profit/weight ratio as in [43]. This repair scheme was implemented in the Lamarckian manner [12].

The effect of our crossover operator was evaluated using four performance measures: the generational distance (GD), the $D1_R$ measure (i.e., inverse generational distance), the hypervolume measure, and the range measure. Let S and S^* be an obtained non-dominated solution set and the Pareto-optimal solution set, respectively. The convergence of the non-dominated solution set S to the Pareto-optimal solution set S^* has been often measured by the generational distance [34]:

$$GD(S) = \frac{1}{|S|} \sum_{\mathbf{x} \in S} \min \{ \|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{y})\| : \mathbf{y} \in S^* \}, \quad (11)$$

where $\|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{y})\|$ is the Euclidean distance between the two solutions \mathbf{x} and \mathbf{y} in the objective space, and $|S|$ is the number of solutions in S (i.e., $|S|$ is the cardinality of S).

Whereas the generational distance evaluates only the proximity of the non-dominated solution set S to the Pareto-optimal solution set S^* , the following $D1_R$ measure (i.e., inverse generational distance) can evaluate both the proximity and the diversity of the non-dominated solution set S :

$$D1_R(S) = \frac{1}{|S^*|} \sum_{\mathbf{y} \in S^*} \min \{ \|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{y})\| : \mathbf{x} \in S \}. \quad (12)$$

This measure was used in [3], [17]. The diversity of the non-dominated solution set S can be more directly measured by the sum of the range of objective values for each objective function:

$$\text{Range}(S) = \sum_{i=1}^K [\max_{\mathbf{x} \in S} \{f_i(\mathbf{x})\} - \min_{\mathbf{x} \in S} \{f_i(\mathbf{x})\}], \quad (13)$$

where $f_i(\mathbf{x})$ is the i th objective and K is the total number of objectives. This measure is similar to the maximum spread of Zitzler [40].

In order to measure both the diversity and the convergence, we can also use the hypervolume measure [42] that calculates the volume of the dominated region by the non-dominated solution set S in the objective space. As the reference point of the hypervolume calculation, we used the origin of the objective space. The boundary of the dominated region in the objective space is called the attainment surface [11]. From multiple attainment surfaces obtained by multiple runs of an EMO algorithm for a multiobjective problem, we can calculate the 50% attainment surface as a kind of their average result. For the calculation of the 50% attainment surface, see [4], [11].

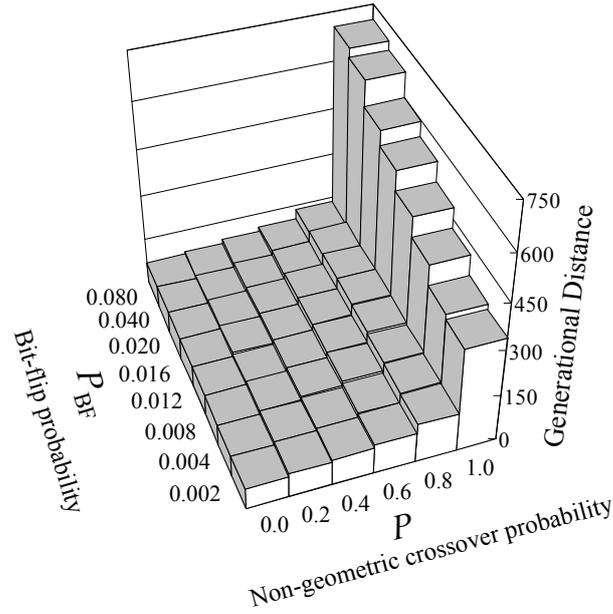


Fig. 6. Average values of the generational distance (the smaller is the better).

Average values of the generational distance are summarized in Fig. 6. The generational distance was degraded by increasing the non-geometric crossover probability P . This observation suggests that the convergence property of NSGA-II was degraded by the use of our crossover operator.

The above-mentioned observation is, however, somewhat misleading. Whereas the generational distance was clearly increased in Fig. 6, the convergence of solutions to the Pareto front was not so severely degraded. This is visually demonstrated in Fig. 7 by depicting the 50% attainment surfaces for the following two cases: $P = 0.0$ (i.e., only the uniform crossover operator) and $P = 0.8$ with $P_{BF} = 0.004$ (i.e., probabilistic use of the non-geometric (80%) and uniform (20%) crossover operators).

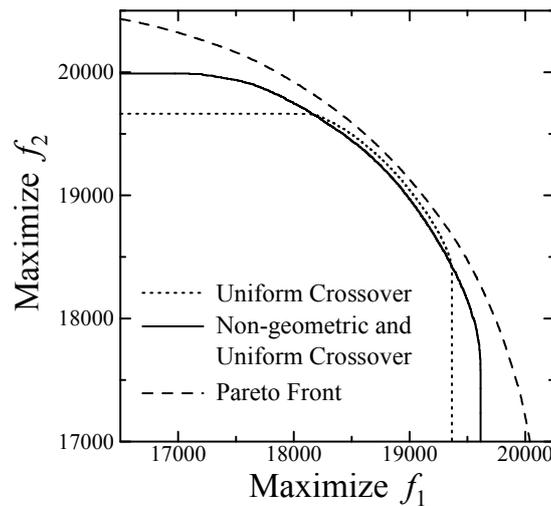


Fig. 7. 50% attainment surfaces obtained from the two cases: Only the uniform crossover operator ($P = 0.0$) and both the non-geometric and uniform crossover operators ($P = 0.8$ and $P_{BF} = 0.004$).

The effect of our non-geometric crossover operator on NSGA-II can be more clearly visualized by depicting all individuals at some generations. Fig. 8 and Fig. 9 are experimental results by NSGA-II with only the uniform crossover operator and NSGA-II with both of the non-geometric and uniform crossover operators, respectively (solutions A and B in Fig. 9 are used later). In Fig. 8 and Fig. 9, the same parameter specifications as in Fig. 7 were used. The comparison between Fig. 8 and Fig. 9 demonstrates a positive effect of our crossover operator on the spread of solutions along the Pareto front and its negative effect on the convergence toward the Pareto front.

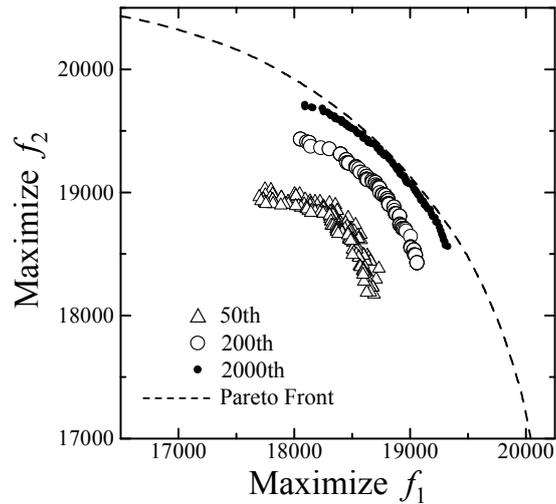


Fig. 8. All individuals at some generations of NSGA-II without non-geometric crossover ($P = 0$).

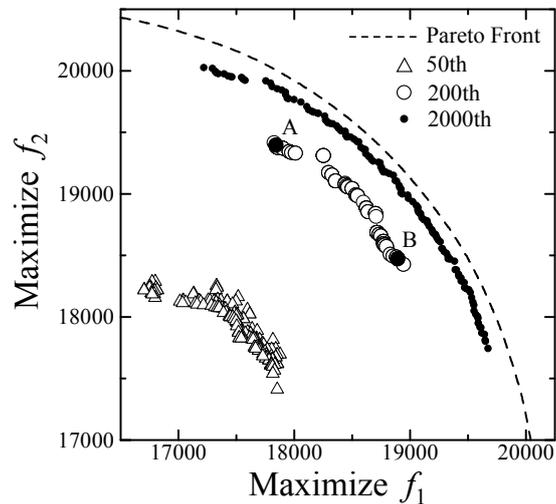


Fig. 9. All individuals at some generations of NSGA-II with non-geometric crossover ($P = 0.8$ and $P_{BF} = 0.004$).

Average values of the $D1_R$ measure are summarized in Fig. 10. The performance of NSGA-II was improved (i.e., the $D1_R$ measure was decreased) by the use of our crossover operator especially around

the right-bottom corner of Fig. 10 with $P = 1.0$ and $P_{BF} = 0.002$. We further examine two cases in detail: without non-geometric crossover ($P = 0.0$) and with non-geometric crossover ($P = 0.8$ and $P_{BF} = 0.004$). Experimental results of 100 runs for each case are shown in Fig. 11 where the horizontal axis is the discretized $D1_R$ measure with the width 40: $200 < D1_R \leq 240$, ..., $760 < D1_R \leq 800$. As shown in Fig. 11, the use of non-geometric crossover clearly improved (i.e., decreased) the $D1_R$ measure. This improvement is statistically significant. Since the results with non-geometric crossover were not normally distributed (according to the χ^2 test), we used the Wilcoxon rank-sum test with $\alpha = 0.05$. It was confirmed that the improvement in Fig. 11 is statistically significant ($p < 10^{-16}$)

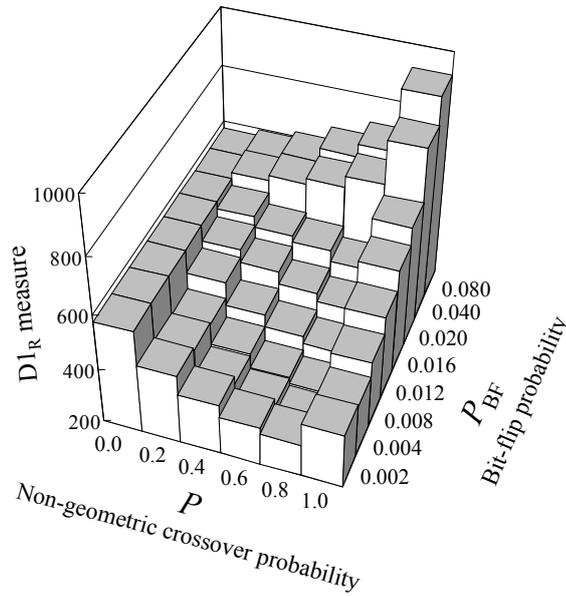


Fig. 10. Average values of $D1_R$ measure (the smaller is the better).

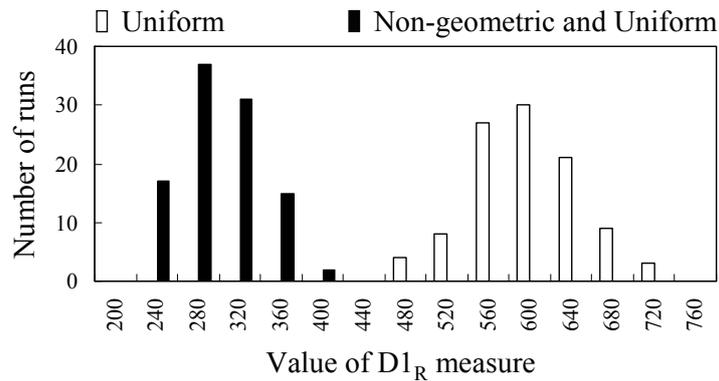


Fig. 11. Histograms of 100 values of the $D1_R$ measure obtained by 100 runs of NSGA-II with/without non-geometric crossover for the two-objective problem ($P = 0.8$ and $P_{BF} = 0.004$).

Performance improvement by the use of non-geometric crossover was also observed for the hypervolume and range measures as shown in Fig. 12 and Fig. 13 where the large values mean the better results. It should be noted that Fig. 13 was depicted from a different angle of vision for better visibility. As in Fig. 11 for the $D1_R$ measure, we also confirmed the statistical significance between the two cases (with/without non-geometric crossover) for the hypervolume and range measures. Actually we obtained separate distributions between the two cases for the hypervolume and range measures as in Fig. 11 (see Fig. 14 for the hypervolume measure).

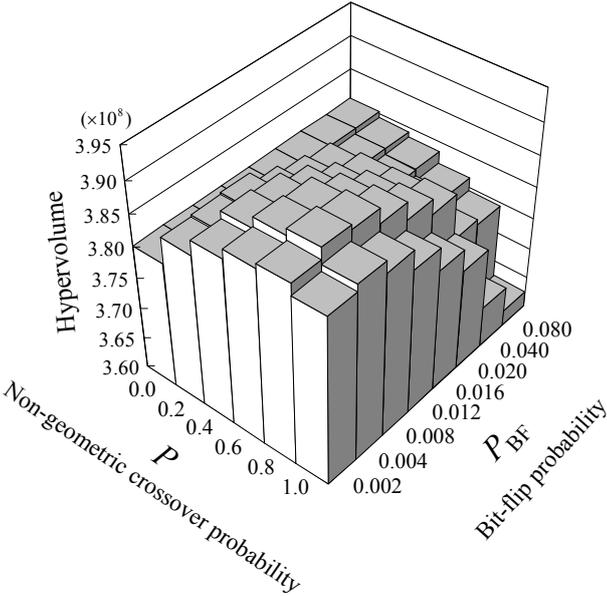


Fig. 12. Average values of the hypervolume measure.

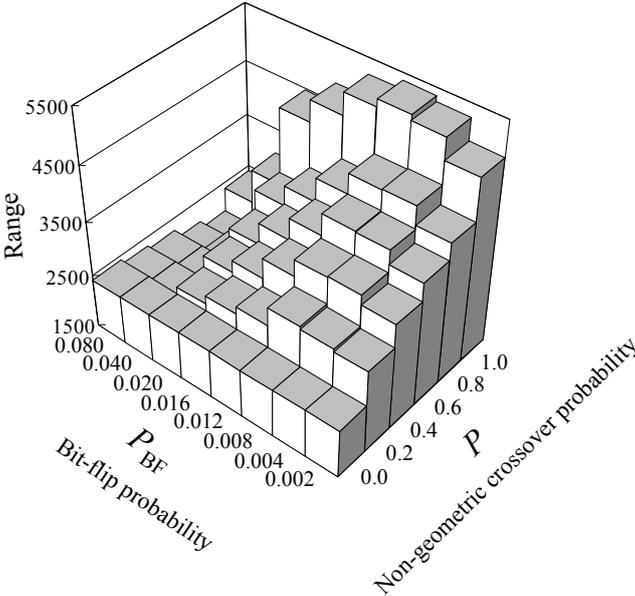


Fig. 13. Average values of the range measure.

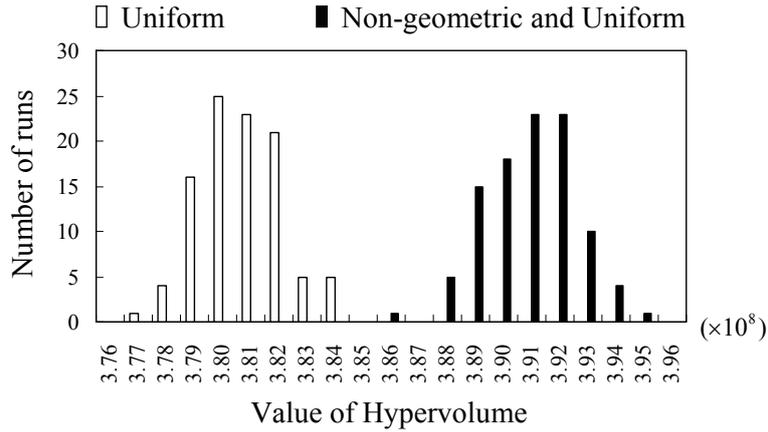


Fig. 14. Histograms of 100 values of the hypervolume obtained by 100 runs of NSGA-II with/without non-geometric crossover for the two-objective problem ($P = 0.8$ and $P_{BF} = 0.004$).

One may think that the performance of NSGA-II without non-geometric crossover can be improved by tuning the crossover and mutation probabilities. We performed computational experiments using various combinations of the crossover probability P_C and the mutation probability P_M : $P_C = 0.0, 0.2, 0.4, 0.6, 0.8, 1.0$ and $P_M = 0.002, 0.004, 0.008, 0.012, 0.016, 0.020$. Experimental results are summarized in Fig. 15 for the hypervolume measure.

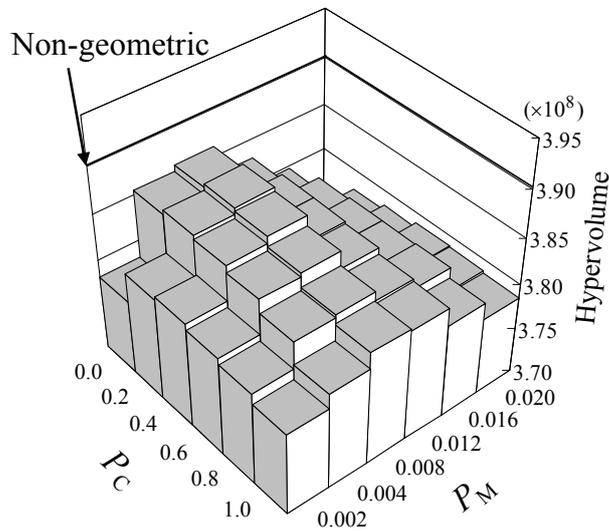


Fig. 15. Hypervolume by NSGA-II without non-geometric crossover.

In Fig. 15, the best average result 3.866×10^8 was obtained when $P_C = 0.2$ and $P_M = 0.008$. For comparison, the best average result 3.904×10^8 in Fig. 12 with non-geometric crossover ($P_C = 0.8, P_M = 0.002, P = 0.8$ and $P_{BF} = 0.004$) is also shown in Fig. 15. This result with non-geometric crossover is better than all results without non-geometric crossover in Fig. 15.

Let us examine the two cases (with/without non-geometric crossover) under the best parameter specifications: $P = 0.8$ and $P_{BF} = 0.004$ in Fig. 12 with non-geometric crossover, and $P_C = 0.2$ and $P_M = 0.008$ in Fig. 15 without non-geometric crossover. As in Fig. 11 and Fig. 14, Fig. 16 shows 100 values of the hypervolume measure for each case. In Fig. 16, the hypervolume measure was improved (i.e., increased) by the use of non-geometric crossover. This improvement is statistically significant. Since the results in Fig. 16 for each of the two cases are normally distributed with the same variance (according to the χ^2 test and the F test), we used the t test with $\alpha = 0.05$. It was confirmed that the improvement in Fig. 16 is statistically significant ($p < 10^{-16}$).

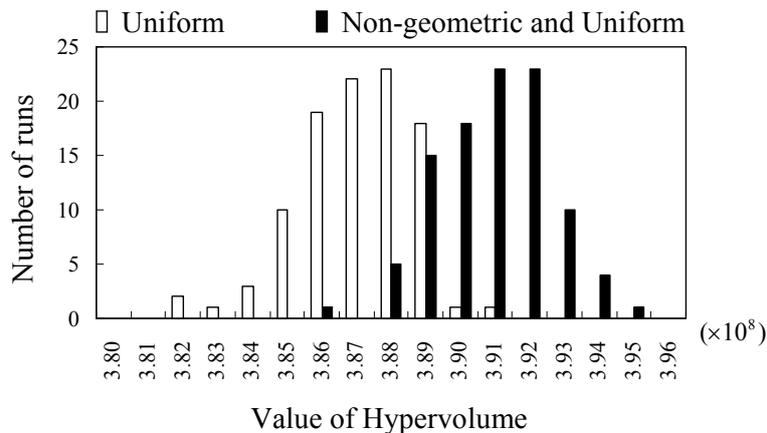


Fig. 16. Histograms of 100 values of the hypervolume measure obtained by 100 runs for each of the two cases (with/without non-geometric crossover) with the best parameter specifications.

One may also think that our non-geometric binary crossover operator involves a large computational overhead. This is not the case in our computational experiments. We did not observe any clear increase in computation time by the use of our crossover operator.

We further examine our non-geometric binary crossover operator by comparing it with the uniform crossover operator. We generated 200 offspring from the two parents A and B in Fig. 9 by each crossover operator ($P_{BF} = 0.004$ in our crossover operator). The generated offspring are shown in Fig. 17. The uniform crossover operator often generates offspring between the two parents in the objective space as shown in Fig. 17 (closed circles). On the other hand, our crossover operator usually generates offspring around their primary parent (open circles in Fig. 17 where each parent was used as the primary parent for 100 offspring). Fig. 17 clearly demonstrates the difference between the two crossover operators. Fig. 17 also shows a kind of locality [27] of our test problem: the two-objective 500-item 0/1 knapsack problem in [43]. That is, small changes in the search space from the primary parent by our crossover operator lead to small changes in the objective space.

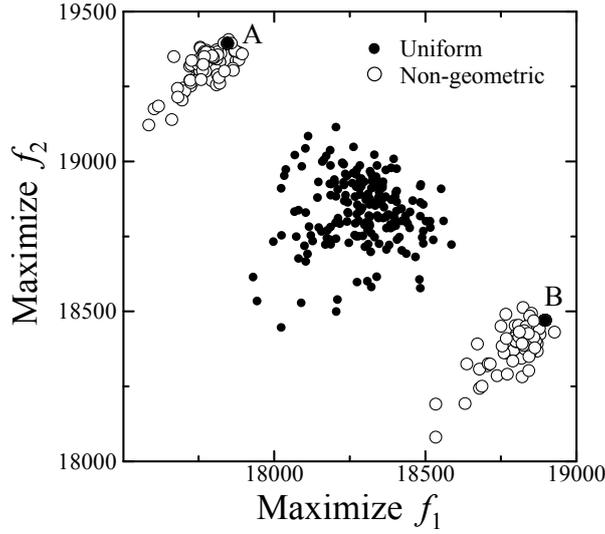


Fig. 17. Generated offspring by the uniform crossover operator (closed circles) and our non-geometric binary crossover operator (open circles) from the two parents A and B in Fig. 9.

In our former studies [14], [15], we proposed a similarity-based mating scheme for increasing the spread of solutions along the Pareto front in EMO algorithms, which is illustrated in Fig. 18. In our mating scheme, first α candidate solutions are selected in the left-hand side of Fig. 18 by iterating the binary tournament selection procedure α times. In tournament selection, each solution is evaluated by the non-dominated sorting and the crowding distance in the same manner as in NSGA-II [8]. Next the average objective vector over the selected α candidates is calculated. Then the most dissimilar candidate from the average objective vector is chosen as Parent A. The similarity between two solutions is measured by their Euclidean distance in the objective space. On the other hand, β candidates are selected in the right-hand side of Fig. 18 by iterating the binary tournament selection procedure β times. Then the most similar candidate to Parent A is selected as Parent B.

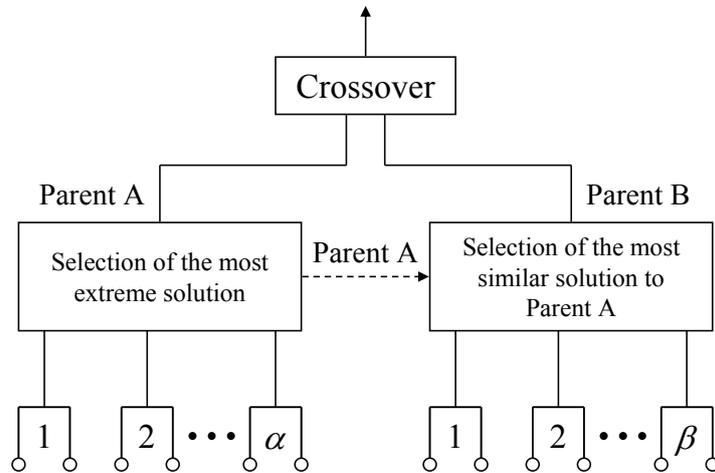


Fig. 18. Similarity-based mating scheme [14], [15].

We applied NSGA-II with the similarity-based mating scheme to our test problem under the same conditions as in the previous computational experiments in Fig. 8. We used only the uniform crossover operator (i.e., $P=0$). We examined the following two settings of the parameters α and β in the similarity-based mating scheme:

Case (1) They were handled as constant parameters and specified as $\alpha = 10$ and $\beta = 10$.

Case (2) They were handled as control parameters and changed after the 1000th generation from $\alpha = 10$ and $\beta = 10$ to $\alpha = 1$ and $\beta = 10$.

Experimental results of a single run with each of these two settings are shown in Fig. 19 where small closed circles and small open circles show individuals at the 2000th generation in Case (1) and Case (2), respectively. By biasing the selection pressure toward extreme solutions far from the average objective vector in each population (in the left part of Fig. 18), the similarity-based mating scheme increased the spread of solutions along the Pareto front as shown in Fig. 19. Its diversity improvement effect is stronger than the use of non-geometric crossover in Fig. 9. As a side-effect, the convergence of solutions was degraded especially in the center region of the Pareto front when we used the fixed parameter values $\alpha = 10$ and $\beta = 10$ (i.e., Case (1): Small closed circles in Fig. 19). Such a side-effect was remedied by handling α and β as control parameters (i.e., Case (2): Small open circles in Fig. 19). In this case, very good results were obtained in Fig. 19. The main difficulty of this approach is the necessity of a good control strategy for the two parameters α and β . On the other hand, the simplicity of implementation is the main advantage of our crossover operator. The performance of NSGA-II was improved by the use of our crossover operator in a wide range of its parameter values. For example, the hypervolume measure was improved in $0.2 \leq P \leq 1.0$ and $0.002 \leq P_{BF} \leq 0.008$ in Fig. 12.

An interesting question is what will happen if we use both the non-geometric binary crossover operator and the similarity-based mating scheme. We combined both of these diversity improvement schemes with NSGA-II. In such an NSGA-II algorithm, the similarity-based mating scheme was used for parent selection as in the computational experiments in Fig. 19. Then our crossover scheme (i.e., probabilistic use of the non-geometric and uniform crossover operators) was applied to each pair of parents in order to generate offspring in the same manner as in Fig. 9. Experimental results are shown in Fig. 20. The probabilistic use of non-geometric crossover increased the spread of solutions along the Pareto front and degraded the convergence toward the Pareto front in Fig. 20 (Compare Fig. 20 with Fig. 19). We specified the non-geometric crossover probability P as $P = 0.4$ in Fig. 20 whereas P was specified as $P = 0.8$ in the previous computational experiments. Since NSGA-II in Fig. 20 had the similarity-based mating scheme with a strong diversity improvement property, we used a smaller value of P in order to weaken the side-effect (i.e., convergence deterioration) of using the two diversity improvement schemes. As shown in Fig. 20, our non-geometric binary crossover operator can be used together with other diversity improvement mechanisms. Actually our crossover operator can be used in any single-objective and multiobjective evolutionary algorithms with binary strings. Such a wide

applicability is one merit of our crossover operator. Another merit is its simplicity of implementation.

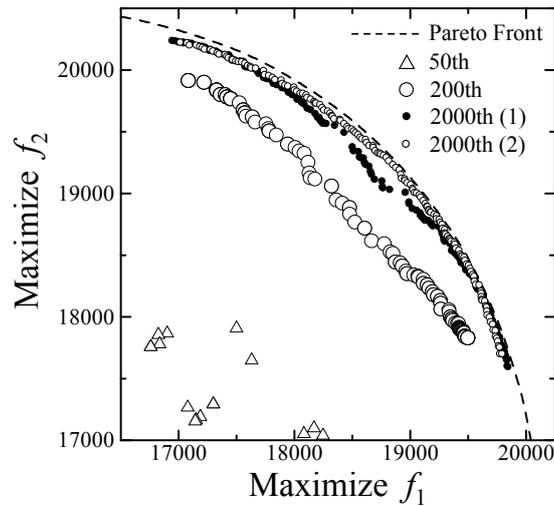


Fig. 19. All individuals in a single run of NSGA-II without non-geometric crossover. In the similarity-based mating scheme, (α, β) was specified as $(10, 10)$ throughout 2000 generations in Case (1). In Case (2), (α, β) was specified as $(10, 10)$ in the first 1000 generations and $(1, 10)$ in the last 1000 generations.

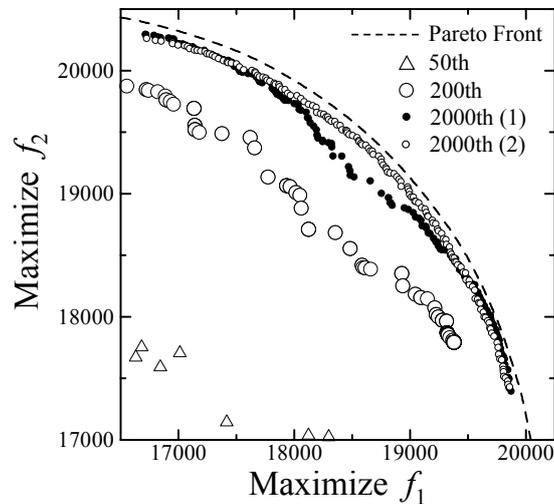


Fig. 20. All individuals in a single run of NSGA-II with non-geometric crossover ($P = 0.4$ and $P_{BF} = 0.004$). In the similarity-based mating scheme, (α, β) was specified in the same manner as in Fig. 19.

V. Performance Evaluation on Many-Objective Problems

It is well-known that the increase in the number of objectives severely degrades the convergence property of EMO algorithms [26], [44]. Convergence improvement mechanisms are often incorporated into EMO algorithms when they are applied to many-objective problems. Thus the use of our non-geometric binary crossover operator is not likely to improve the performance of EMO algorithms on

many-objective problems. In this section, we examine this issue through computational experiments on four-objective and six-objective problems.

We used a four-objective 500-item 0/1 knapsack problem in [43]. We also generated a six-objective 500-item 0/1 knapsack problem in the same manner as in [43]. We used NSGA-II with the same parameter specifications as in Section IV. The performance of NSGA-II was evaluated by the hypervolume measure. The generational distance (GD) and the $D1_R$ measure were not used because the true Pareto optimal solution set S^* or its good approximation set was not available for the test problems in this section.

Average results over 100 runs on the four-objective and six-objective problems are summarized in Fig. 21 and Fig. 22, respectively. Contrary to our expectations, the hypervolume measure was improved by the probabilistic use of non-geometric crossover (i.e., by specifying the non-geometric crossover probability P as $0 < P < 1$) when the bit-flip probability P_{BF} was not too large (e.g., $P_{BF} \leq 0.008$). The performance of NSGA-II, however, was severely degraded by non-geometric crossover when both P and P_{BF} were too large around the right-bottom corner of Fig. 21 and Fig. 22 with $P = 1.0$ and $P_{BF} = 0.080$. These observations are common among the experimental results on the three test problems with two objectives (in Fig. 12), four objectives (in Fig. 21) and six objectives (in Fig. 22). Careful comparison among these three figures shows that the range of values of the two parameters P and P_{BF} with the performance improvement becomes small when the number of objectives increases. This is because the negative effect of non-geometric crossover (i.e., convergence deterioration) becomes dominant over its positive effect (i.e., diversity improvement) when the number of objectives is large.

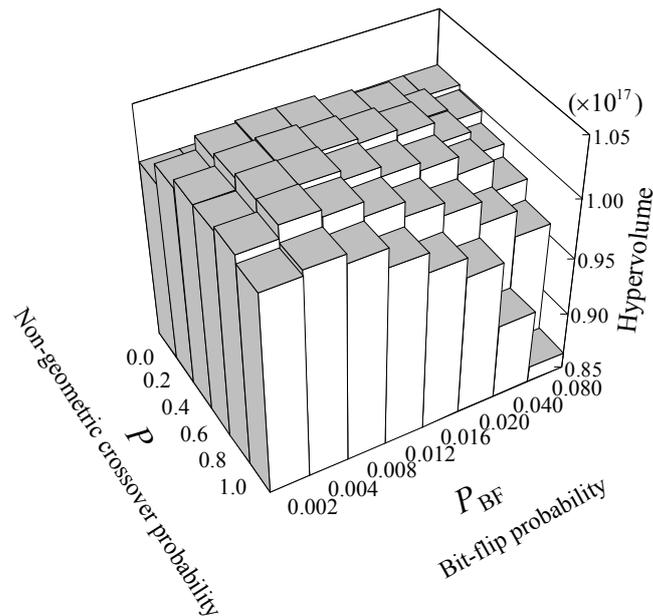


Fig. 21. Average value of the hypervolume measure on the four-objective problem.

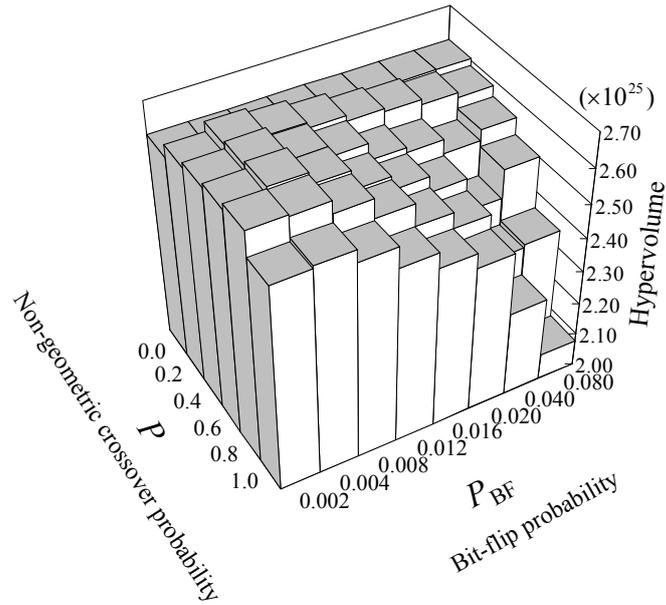


Fig. 22. Average value of the hypervolume measure on the six-objective problem.

In order to examine the statistical significance of the performance improvement by our crossover operator, we show the histogram of 100 values of the hypervolume measure obtained by 100 runs of NSGA-II with/without our crossover operator in Fig. 23 for the four-objective problem and Fig. 24 for the six-objective problem. In the case of NSGA-II with our crossover operator, we used the best parameter values of P and P_{BF} in Fig. 21 for the four-objective problem and Fig. 22 for the six-objective problem. In Fig. 23 and Fig. 24, we observe a clear increase in the values of the hypervolume measure by the use of non-geometric crossover.

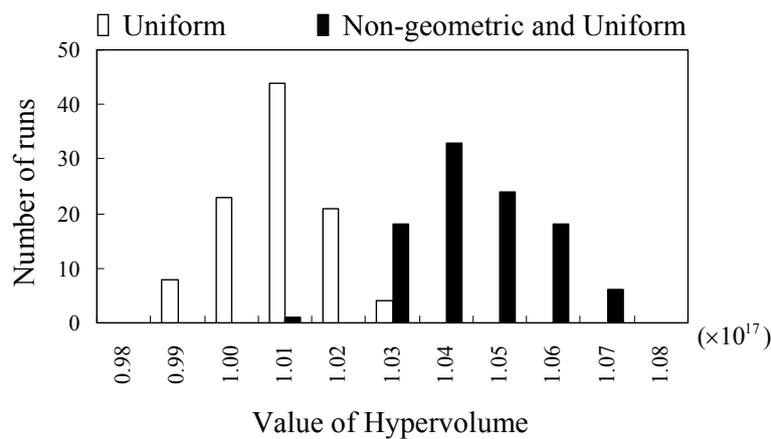


Fig. 23. Histograms of 100 values of the hypervolume measure obtained by 100 runs of NSGA-II with/without our crossover operator for the four-objective problem.

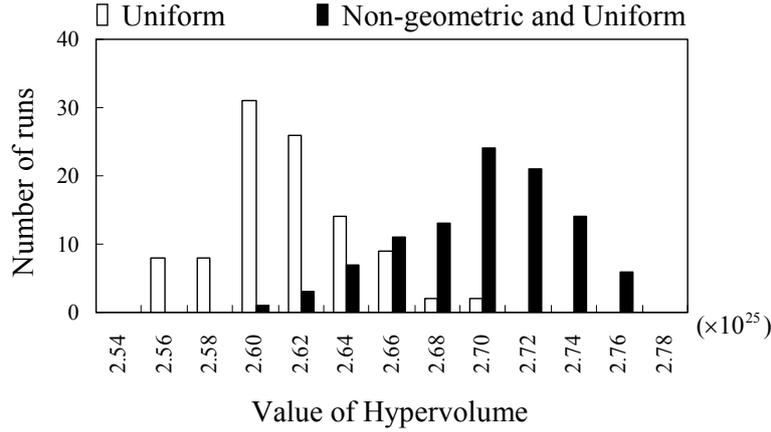


Fig. 24. Histograms of 100 values of the hypervolume measure obtained by 100 runs of NSGA-II with/without our crossover operator for the six-objective problem.

The performance improvement is statistically significant in Fig. 23 and Fig. 24. First, we tested the normality and the variance of the results using the χ^2 test and the F test. In Fig. 23, the two distributions are normally distributed with a common variance. So we used the t test ($\alpha = 0.05$) and confirmed that the performance improvement in Fig. 23 is statistically significant ($p < 10^{-16}$). On the other hand, the two distributions in Fig. 24 are normally distributed with different variances. So we used the Welch's t test ($\alpha = 0.05$) and confirmed that the performance improvement in Fig. 24 is statistically significant ($p < 10^{-16}$).

VI. Discussions

As we have visually demonstrated in Section IV, our crossover operator increases the spread of solutions along the Pareto front in EMO algorithms while it slows down the convergence toward the Pareto front. As a result, we observed the improvement of overall performance measures such as hypervolume and $D1_R$ by the probabilistic use of our crossover operator in a wide range of its parameter values. When the entire Pareto front is not covered by solutions, our crossover operator can be used to widen the population along the Pareto front in EMO algorithms. On the other hand, the use of our crossover operator is not recommended when the population has enough width to cover the entire Pareto front. The use of our crossover operator will improve the performance of EMO algorithms in the former case whereas it will simply slow down the convergence of solutions toward the Pareto front in the latter case. These discussions show that the usefulness of our crossover operator is problem-dependent.

A one-max zero-max problem is a typical example where our crossover operator works very well. This is a two-objective problem with binary strings. One objective is to maximize the number of 1's while the other is to maximize the number of 0's. This problem has a wide Pareto front ranging from

one extreme string of all 1's to the other extreme string of all 0's. Using the same setting as in the previous sections, we performed computational experiments on the 500-bit one-max zero-max problem. This problem has 501 Pareto-optimal objective vectors in the two-dimensional objective space: $(0, 500), (1, 499), \dots, (500, 0)$. When the two extreme points $(0, 500)$ and $(500, 0)$ are obtained, the value of the range measure is 1000. It should be noted that all binary strings are Pareto-optimal. Thus the generational distance is always zero for any solution sets. Experimental results are summarized in Fig. 25 for the range measure. Fig. 25 clearly shows that the use of our crossover operator (i.e., the use of a positive value for P) increased the spread of solutions along the Pareto front.

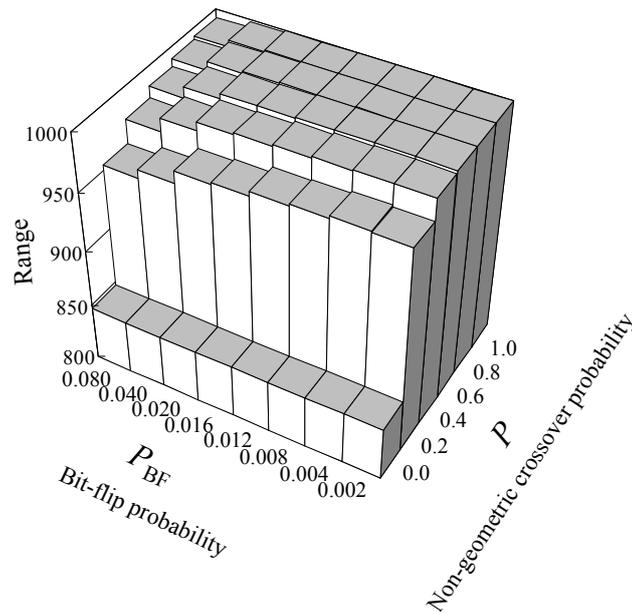


Fig. 25. Average values of the range measure for the 500-bit one-max zero-max problem.

We also performed computational experiments on ZDT test problems [41]. We used 30-variable versions of ZDT1, ZDT2, ZDT3 and a 10-variable version of ZDT4. Each variable was coded by a binary string of length 30. NSGA-II was applied to each test problem under the following setting: The population size was 100, the crossover probability was 0.8 ($P_C = 0.8$), the mutation probability was $1/n$ ($P_M = 1/n$) where n is the string length, and the termination condition was 250 generations. In computational experiments on ZDT1, ZDT2 and ZDT3, we did not observe the necessity to widen the population along the Pareto front. That is, NSGA-II with uniform crossover found solutions that covered the entire Pareto front of each test problem. So we do not need our crossover operator for ZDT1, ZDT2 and ZDT3. Actually the probabilistic use of our crossover operator showed only its negative effect: It simply slowed down the convergence of solutions toward the Pareto front.

We obtained totally different results from computational experiments on ZDT4. ZDT4 is a well-known two-objective problem with a large number of local Pareto fronts (e.g., see Deb [4]). As in

computational experiments on ZDT1, ZDT2 and ZDT3, we did not observe the necessity to widen the population along the Pareto front of ZDT4 in computational experiments by NSGA-II with uniform crossover. NSGA-II, however, almost always found solution sets along local Pareto fronts in the objective space. That is, NSGA-II got stuck with local Pareto fronts whereas obtained solutions had enough spread in the objective space. In this case, our crossover operator improved the performance of NSGA-II by increasing the diversity of binary strings (rather than the spread of objective vectors). Experimental results are summarized in Fig. 26 for the generational distance. The convergence toward the Pareto front was improved by the probabilistic use of our crossover operator in a wide range of its parameter values in Fig. 26.

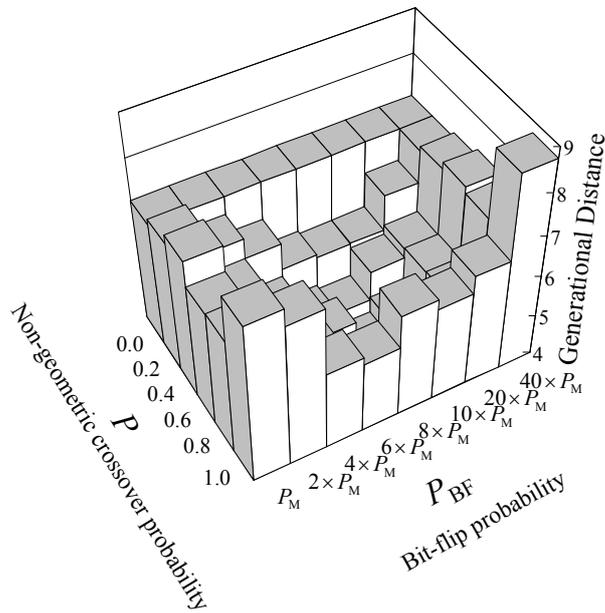


Fig. 26. Average values of the generational distance for ZDT4.

In order to visually illustrate the performance improvement of NSGA-II by our crossover operator for ZDT4, we show the 50% attainment surfaces in Fig. 27 for the two cases: NSGA-II with only the uniform crossover operator and NSGA-II with both crossover operators ($P = 0.6$ and $P_{BF} = 8 \times P_M$). We can see from Fig. 27 that the probabilistic use of our crossover operator improved the performance of NSGA-II by helping it to escape from local Pareto fronts (rather than by increasing the spread of solutions along local Pareto fronts).

In the previous sections, we demonstrated that the probabilistic use of our crossover operator improved the performance of NSGA-II for multiobjective 0/1 knapsack problems by increasing the spread of solutions along the Pareto front in the objective space. On the other hand, the performance improvement of NSGA-II with binary strings for ZDT4 in this section was achieved by increasing the diversity of binary strings to help NSGA-II to escape from local Pareto fronts.

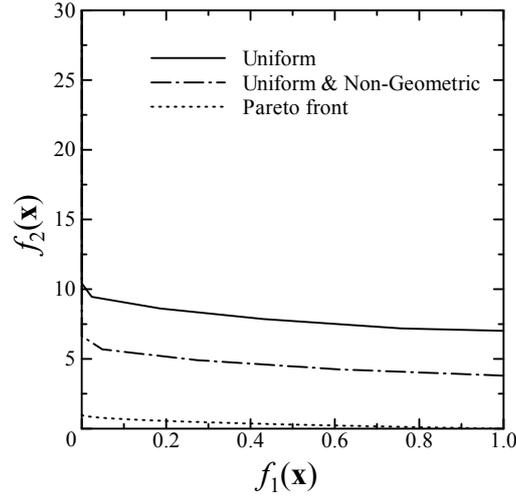


Fig. 27. 50% attainment surfaces obtained for ZDT4 from the two cases: NSGA-II with only the uniform crossover operator ($P = 0$) and NSGA-II with both crossover operators ($P = 0.6$ and $P_{BF} = 8 \times P_M$)).

In some other computational experiments, we also observed the performance improvement of NSGA-II by our crossover operator through the above-mentioned two positive effects (i.e., through the increase in the spread of solutions along the Pareto front and the escape from local Pareto fronts by increasing the diversity of binary strings). For example, experimental results of NSGA-II for multiobjective fuzzy rule selection [16] were improved with respect to both the spread of solutions along the Pareto front and the convergence toward the Pareto front. In Fig. 28, we show experimental results on a two-objective fuzzy rule selection problem for the glass identification data set where the number of fuzzy rules (horizontal axis) and the error rate (vertical axis) are minimized. In Fig. 28, 50% attainment surfaces over 100 runs are depicted for two cases: NSGA-II with only the uniform crossover ($P = 0$) and NSGA-II with both crossover operators ($P = 0.4$ and $P_{BF} = 1/n$ where n is the number of candidate fuzzy rules for rule selection).

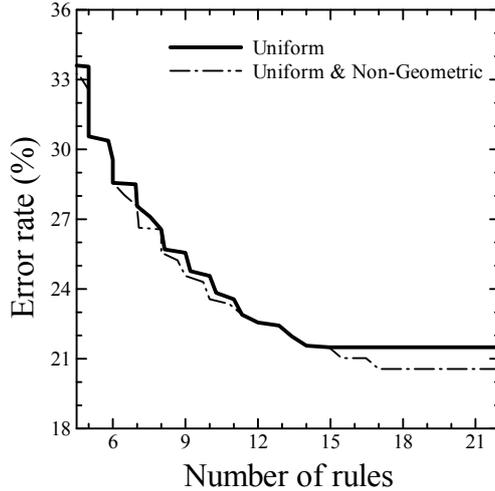


Fig. 28. 50% attainment surfaces obtained for a fuzzy rule selection problem for the glass identification data set from the two cases: NSGA-II with only the uniform crossover operator ($P = 0$) and NSGA-II with both crossover operators ($P = 0.4$ and $P_{BF} = 1/n$).

VII. Conclusion

In this paper, we proposed the probabilistic use of geometric and non-geometric binary crossover operators to increase the spread of solutions along the Pareto front in EMO algorithms. The use of not only geometric but also non-geometric crossover operators for binary strings was motivated by the fact that non-geometric crossover operators have been used for real-coded genetic algorithms. It was visually shown that almost all offspring generated by uniform crossover were between their two parents in the objective space when they were not similar to each other. The effect of our non-geometric binary crossover operator on the performance of EMO algorithms was clearly demonstrated through computational experiments on multiobjective 500-item 0/1 knapsack problems with two, four and six objectives using NSGA-II. Experimental results showed that the spread of solutions along the Pareto front was improved without severely degrading the convergence toward the Pareto front for a two-objective knapsack problem. It was also shown that our non-geometric binary crossover operator improved the performance of NSGA-II for four-objective and six-objective knapsack problems.

Diversity improvement can be also easily realized by increasing the mutation probability in NSGA-II. The use of a large mutation probability, however, severely degrades the convergence of solutions toward the Pareto front in NSGA-II [14]. As a result, the performance of NSGA-II was deteriorated when the mutation probability was too large (e.g., see Fig. 15). It was shown in this paper that the proposed non-geometric binary crossover operator has a larger positive effect on the performance of NSGA-II for the two-objective knapsack problem than the tuning of the crossover and mutation probabilities. Diversity maintenance effects were compared between our non-geometric binary crossover operator and the similarity-based mating scheme [14], [15]. Experimental results showed that better results with larger spread of solutions along the Pareto front were obtained by the

similarity-based mating scheme with an appropriate control strategy of its parameter values. It was also shown that these two schemes can be simultaneously used. This is because these two schemes work on different parts of EMO algorithms: The similarity-based mating scheme is used to select a pair of parents whereas our non-geometric crossover operator is used to generate an offspring from the selected parents.

An interesting observation in this paper is that the diversity improvement by our non-geometric binary crossover operator increased the hypervolume measure not only for the two-objective knapsack problem but also for the four-objective and six-objective knapsack problems. It is well-known that Pareto dominance-based EMO algorithms such as NSGA-II do not work well on many-objective optimization problems due to the lack of selection pressure toward the Pareto front. Our experimental results, however, suggest that not only selection pressure but also diversity maintenance should be improved in EMO algorithms when they are applied to many-objective problems.

We also examined the effect of our crossover operator on the performance of NSGA-II for other test problems. It was clearly demonstrated for a 500-bit one-max zero-max problem that our crossover operator can widen the population along the Pareto front. On the other hand, our crossover operator improved the performance of NSGA-II for ZDT4 by increasing the diversity of binary strings to help NSGA-II to escape from local Pareto fronts (rather than by increasing the spread of solutions along the Pareto front). These two positive effects (i.e., widening the population along the Pareto front and helping the escape from local Pareto fronts) were simultaneously involved in the performance improvement of NSGA-II for multiobjective fuzzy rule selections. We will report computational experiments in detail on these test problems in a future study.

Acknowledgement

This work was partially supported by Japan Society for the Promotion of Science (JSPS) through Grand-in-Aid for Scientific Research (B): KAKENHI (20300084).

References

- [1] P. P. Bonissone, R. Subbu, and J. Lizzi, "Multicriteria decision making (MCDM): A framework for research and applications," *IEEE Computational Intelligence Magazine*, vol. 4, no. 3, pp. 48-61, August 2009.
- [2] P. A. N. Bosman and D. Thierens, "The balance between proximity and diversity in multiobjective evolutionary algorithms," *IEEE Trans. on Evolutionary Computation*, vol. 7, no. 2, pp. 174-188, April 2003.
- [3] P. Czyzak and A. Jaskiewicz, "Pareto-simulated annealing – A metaheuristic technique for multi-objective combinatorial optimization," *Journal of Multi-Criteria Decision Analysis*, vol. 7, no. 1, pp. 34-47, January 1998.

- [4] K. Deb, *Multi-Objective Optimization Using Evolutionary Algorithms*, John Wiley & Sons, Chichester, 2001.
- [5] K. Deb and R. B. Agrawal, "Simulated binary crossover for continuous search space," *Complex Systems*, vol. 9, no. 2, pp. 115-148, April 1995.
- [6] K. Deb, A. Anand, and D. Joshi, "A computationally efficient evolutionary algorithm for real-parameter optimization," *Evolutionary Computation*, vol. 10, no. 4, pp. 371-395, Winter 2002.
- [7] K. Deb and A. Kumar, "Real-coded genetic algorithms with simulated binary crossover: Studies on multimodal and multiobjective problems," *Complex Systems*, vol. 9, no. 6, pp. 431-454, December 1995.
- [8] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: NSGA-II," *IEEE Trans. on Evolutionary Computation*, vol. 6, no. 2, pp. 182-197, April 2002.
- [9] E. L. Eshelman, "The CHC adaptive search algorithm: How to have safe search when engaging in nontraditional genetic recombination," *Foundation of Genetic Algorithms*, Morgan Kaufmann, San Mateo, pp. 265-283, 1991.
- [10] L. J. Eshelman and J. D. Schaffer, "Real-coded genetic algorithms and interval-schemata," *Foundation of Genetic Algorithms 2*, Morgan Kaufmann, San Mateo, pp. 187-202, 1993.
- [11] C. M. Fonseca and P. J. Fleming, "On the performance assessment and comparison of stochastic multiobjective optimizers," *Lecture Notes in Computer Science 1141: Parallel Problem Solving from Nature – PPSN IV*, Springer, Berlin, pp. 584-593, September 1996.
- [12] H. Ishibuchi, S. Kaige, and K. Narukawa, "Comparison between Lamarckian and Baldwinian repair on multiobjective 0/1 knapsack problems," *Lecture Notes in Computer Science 3410: Evolutionary Multi-Criterion Optimization – EMO 2005*, Springer, Berlin, pp. 370-385, March 2005.
- [13] H. Ishibuchi and K. Narukawa, "Recombination of similar parents in EMO algorithms," *Lecture Notes in Computer Science 3410: Evolutionary Multi-Criterion Optimization – EMO 2005*, Springer, Berlin, pp. 265-279, March 2005.
- [14] H. Ishibuchi, K. Narukawa, N. Tsukamoto, and Y. Nojima, "An empirical study on similarity-based mating for evolutionary multiobjective combinatorial optimization," *European Journal of Operational Research*, vol. 188, no. 1, pp. 57-75, July 2008.
- [15] H. Ishibuchi and Y. Shibata, "Mating scheme for controlling the diversity-convergence balance for multiobjective optimization," *Lecture Notes in Computer Science 3102: Genetic and Evolutionary Computation – GECCO 2004*, Springer, Berlin, pp. 1259-1271, June 2004.
- [16] H. Ishibuchi and T. Yamamoto, "Fuzzy rule selection by multi-objective genetic local search algorithms and rule evaluation measures in data mining," *Fuzzy Sets and Systems*, vol. 141, no. 1, pp. 59-88, January 2004.

- [17] H. Ishibuchi, T. Yoshida, and T. Murata, "Balance between genetic search and local search in memetic algorithms for multiobjective permutation flowshop scheduling," *IEEE Trans. on Evolutionary Computation*, vol. 7, no. 2, pp. 204-223, April 2003.
- [18] A. Jaszkievicz, "On the performance of multiple-objective genetic local search on the 0/1 knapsack problem – A comparative experiment," *IEEE Trans. on Evolutionary Computation*, vol. 6, no. 4, pp. 402-412, August 2002.
- [19] Y. Jin and B. Sendhoff, "A systems approach to evolutionary multiobjective structural optimization and beyond," *IEEE Computational Intelligence Magazine*, vol. 4, no. 3, pp. 62-76, August 2009.
- [20] M. Kim, T. Hiroyasu, M. Miki, and S. Watanabe, "SPEA2+: Improving the performance of the strength Pareto evolutionary algorithm 2," *Lecture Notes in Computer Science 3242: Parallel Problem Solving from Nature – PPSN VIII*, Springer, Berlin, pp. 742-751, September 2004.
- [21] A. Moraglio and R. Poli, "Topological interpretation of crossover," *Lecture Notes in Computer Science 3102: Genetic and Evolutionary Computation – GECCO 2004*, Springer, Berlin, pp. 1377-1388, September 2004.
- [22] A. Moraglio and R. Poli, "Product geometric crossover," *Lecture Notes in Computer Science 4193: Parallel Problem Solving from Nature – PPSN IX*, Springer, Berlin, pp. 1018-1027, September 2006.
- [23] A. Moraglio and R. Poli, "Inbreeding properties of geometric crossover and non-geometric recombinations," *Lecture Notes in Computer Science 4436: Foundation of Genetic Algorithms – FOGA 2007*, Springer, Berlin, pp. 1-14, August 2007.
- [24] H. Muhlenbein and D. Schlierkamp-Voosen, "Predictive models for the breeder genetic algorithm I: Continuous parameter optimization," *Evolutionary Computation*, vol. 1, no. 1, pp. 25-49, Spring 1993.
- [25] I. Ono and S. Kobayashi, "A real-coded genetic algorithm for function optimization using unimodal normal distribution crossover," *Proc. of 7th International Conference on Genetic Algorithms* (East Lansing, MI, July 19-23, 1997) pp. 246-253.
- [26] R. C. Purshouse and P. J. Fleming, "On the evolutionary optimization of many conflicting objectives," *IEEE Trans. on Evolutionary Computation*, vol. 11, no. 6, pp. 770-784, December 2007.
- [27] G. R. Raidl and J. Gottlieb, "Empirical analysis of locality, heritability and heuristic bias in evolutionary algorithms: A case study for the multidimensional knapsack problem," *Evolutionary Computation*, vol. 13, no. 4, pp. 441-475, December 2005.
- [28] H. Sato, H. Aguirre, and K. Tanaka, "Effects from local dominance and local recombination in enhanced MOEAs," *Proc. of 5th International Conference on Simulated Evolution and Learning* (Busan, Korea, October 26-29, 2004) CD-ROM Proceedings.

- [29] A. B. Simões and E. Costa, “Transposition versus crossover: An empirical study,” *Proc. of 1999 Genetic and Evolutionary Computation Conference* (Orlando, Florida, July 13-17, 1999) pp. 612-619.
- [30] K. C. Tan, C. K. Goh, Y. J. Yang and T. H. Lee, “Evolving better population distribution and exploration in evolutionary multi-objective optimization,” *European Journal of Operational Research*, vol. 171, no. 2, pp. 463-495, June 2006.
- [31] K. C. Tan, E. F. Khor, and T. H. Lee, *Multiobjective Evolutionary Algorithms and Applications*, Springer, Berlin, May 2005.
- [32] K. C. Tan, T. H. Lee and E. F. Khor, “Evolutionary algorithm with dynamic population size and local exploration for multiobjective optimization,” *IEEE Trans. on Evolutionary Computation*, vol. 5, no. 6, pp. 565-588, December 2001.
- [33] S. Uyar, S. Sariel, and G. Eryigit, “A gene based adaptive mutation strategy for genetic algorithms,” *Lecture Notes in Computer Science 3103: GECCO 2004*, Springer Berlin, pp. 271-281, June 2004.
- [34] D. A. Van Veldhuizen, “Multiobjective evolutionary algorithms: Classifications, analyses, and new innovations,” *Ph. D dissertation*, Air Force Institute of Technology, Wright-Patterson AFB, Ohio, May 1999.
- [35] S. Watanabe, T. Hiroyasu, and M. Miki, “Neighborhood cultivation genetic algorithm for multi-objective optimization problems,” *Proc. of 4th Asia-Pacific Conference on Simulated Evolution and Learning* (Singapore, November 18-22, 2002) pp. 198-202.
- [36] A. H. Wright, “Genetic algorithms for real parameter optimization,” *Foundation of Genetic Algorithm*, Morgan Kaufmann, San Mateo, pp. 205-218, 1991.
- [37] S. Yang, “Adaptive mutation using statistics mechanism for genetic algorithms,” In F. Coenen, A. Preece and A. Macintosh (eds.), *Research and Development in Intelligent Systems XX*, pp. 19-32, London, Springer-Verlag, March 2004.
- [38] S. Yang and S. Uyar, “Adaptive mutation with fitness and allele distribution correlation for genetic algorithms,” *Proc. of the 21st ACM Symposium on Applied Computing* (Dijon, France, April 23-27, 2006) pp. 940-944.
- [39] Y. Yoon, Y. H. Kim, A. Moraglio, and B. R. Moon, “Geometric crossover for real-coded representation,” *Proc. of 2007 Genetic and Evolutionary Computation Conference* (London, U.K., July 7-11, 2007) p. 1539.
- [40] E. Zitzler, “Evolutionary algorithms for multiobjective optimization: Methods and applications,” *Ph. D Dissertation*, Swiss Federal Institute of Technology, Zurich, November 1999.
- [41] E. Zitzler, K. Deb, and L. Thiele, “Comparison of multiobjective evolutionary algorithms: Empirical results,” *Evolutionary Computation*, vol. 8, no. 2, pp. 173-195, Summer 2000.

- [42] E. Zitzler and L. Thiele, “Multiobjective optimization using evolutionary algorithms – A comparative case study,” *Lecture Notes in Computer Science 1498: Parallel Problem Solving from Nature – PPSN V*, Springer, Berlin, pp. 292-301, September 1998.
- [43] E. Zitzler and L. Thiele, “Multiobjective evolutionary algorithms: A comparative case study and the strength Pareto approach,” *IEEE Trans. on Evolutionary Computation*, vol. 3, no. 4, pp. 257-271, November 1999.
- [44] X. Zou, Y. Chen, M. Liu, and L. Kang, “A new evolutionary algorithm for solving many-objective optimization problems,” *IEEE Trans. on Systems, Man, and Cybernetics: Part B - Cybernetics*, vol. 38, no. 5, pp. 1402-1412, October 2008.



Hisao Ishibuchi (M'93-SM'10) received the B.S. and M.S. Degrees in precision mechanics from Kyoto University, Kyoto, Japan, in 1985 and 1987, respectively, and the Ph.D. degree from Osaka Prefecture University, Osaka, Japan, in 1992.

Since 1987, he has been with Osaka Prefecture University where he is currently a Professor in Department of Computer Science and Intelligent Systems. His research interests include evolutionary multiobjective optimization, multiobjective memetic algorithms, evolutionary game, fuzzy rule-based classification, multiobjective genetic fuzzy systems, and fuzzy data mining.

Dr. Ishibuchi received the Best Paper Award from GECCO 2004 (GA Track), HIS-NCEI 2006 and FUZZ-IEEE 2009. He also received 2007 JSPS PRIZE from the Japan Society for the Promotion of Science. He is the IEEE CIS Vice-President for Technical Activities (2010-2011), an Associate Editor for a number of international journals such as *IEEE Trans. on Fuzzy Systems*, *IEEE Trans. on Evolutionary Computation*, and *IEEE Computational Intelligence Magazine*. He is the Program Chair of IEEE CEC 2010 and a Program Co-Chair of FUZZ-IEEE 2011.



Noritaka Tsukamoto received the M.S. Degrees in computer science and intelligent systems from Osaka Prefecture University, Osaka, Japan, in 2008. He is currently a doctor's course student in Department of Computer Science and Intelligent Systems, Osaka Prefecture University. His research interest is evolutionary multiobjective optimization.

Mr. Tsukamoto received the Best Paper Award from FUZZ-IEEE 2009.



Yusuke Nojima (M'00) received the B.S. and M.S. Degrees in mechanical engineering from Osaka Institute of Technology, Osaka, Japan, in 1999 and 2001, respectively, and the Ph.D. degree from Kobe University, Hyogo, Japan, in 2004.

Since 2004, he has been with Osaka Prefecture University where he is currently an Assistant Professor in Department of Computer Science and Intelligent Systems. His research interests include multiobjective genetic fuzzy systems, evolutionary multiobjective optimization, parallel distributed data mining, and ensemble classifier design.

Dr. Nojima received the Best Paper Award from HIS-NCEI 2006 and FUZZ-IEEE 2009.