# Optimal Distributions of Solutions for Maximizing the Minimum Crowding Distance for Two-Objective and Three-Objective Linear Pareto Fronts: Search Behavior Analysis of NSGA-II

Hisao Ishibuchi†
Southern University of Science and Technology
Shenzhen, China
hisao@sustech.edu.cn

Yang Nan
Southern University of Science and Technology
Shenzhen, China
nany@mail.sustech.edu.cn

Lie Meng Pang
Southern University of Science and Technology
Shenzhen, China
panglm@sustech.edu.cn

## ABSTRACT

In this paper, we discuss the optimal distributions of solutions for maximizing the minimum crowding distance on linear Pareto fronts of two- and three-objective problems. Our discussions are twofold. One is theoretical discussions to analytically derive the optimal distributions. The other is empirical discussions, which are based on a newly implemented indicator-based algorithm to maximize the minimum crowding distance. First, we show the upper bound on the minimum crowding distance, which is derived from the upper bound on the total crowding distance. Next, we theoretically derive the optimal distributions for maximizing the minimum crowding distance on two-objective linear Pareto fronts. Our analysis shows that two solutions always overlap in the optimal distributions when the population size is an even number. For an odd number population size, the optimal distribution cannot be uniquely specified except for some anchor solutions (i.e., many different distributions are optimal). The theoretically obtained optimal distributions are compared with experimental results by our indicator-based algorithm. Then, we show an example of optimal distributions for a three-objective linear Pareto front where the population size is a multiple of three. Our indicator-based algorithm is also compared with the standard NSGA-II algorithm and its steady-state variant.

## CCS CONCEPTS

**Mathematics of computing → Evolutionary algorithms.**

## KEYWORDS

Evolutionary multi-objective optimization, NSGA-II, crowding distance, optimal distribution of solutions, steady-state algorithms.

†Corresponding Author

*FOGA '25, August 27–29, 2025, Leiden, Netherlands*
 ACM ISBN 979-8-4007-1859-5/25/08
https://doi.org/10.1145/3729878.3746702

**ACM Reference format:**

Hisao Ishibuchi, Yang Nan and Lie Meng Pang. 2025. Optimal Distributions of Solutions for Maximizing the Minimum Crowding Distance for Two-Objective and Three-Objective Linear Pareto Fronts: Search Behavior Analysis of NSGA-II. In *Proceedings of Conference on Foundations of Genetic Algorithms (FOGA '25) Aug 27-29, 2025, Leiden, The Netherlands.* ACM, New York, NY, USA, 12 pages. https://doi.org/10.1145/3729878.3746702

## 1 Introduction

An $m$-objective minimization problem can be written as follows:

$$\text{Minimize } \boldsymbol{f}(\boldsymbol{x}) = (f_1(\boldsymbol{x}), \ldots, f_m(\boldsymbol{x})) \text{ subject to } \boldsymbol{x} \in X, \quad (1)$$

where $f_i(\boldsymbol{x})$ is the $i$th objective to be minimized ($i$ = 1, ..., $m$), $\boldsymbol{x}$ is an $n$-dimensional decision vector, and $X$ is the feasible region. In general, objectives are conflicting with each other. Thus, there is no solution which minimizes all objectives. In multi-objective optimization, the following Pareto dominance relation is used to define Pareto optimal solutions: solution $\boldsymbol{x}$ is referred to as being dominated by solution $\boldsymbol{y}$ if and only if $f_i(\boldsymbol{y}) \le f_i(\boldsymbol{x})$ for all $i$ and $f_i(\boldsymbol{y}) \ne f_i(\boldsymbol{x})$. Solution $\boldsymbol{x}^*$ is a Pareto optimal solution if there is no solution which dominates $\boldsymbol{x}^*$. The set of all Pareto optimal solutions $\boldsymbol{x}^*$ is the Pareto set, and the set of all Pareto optimal vectors $\boldsymbol{f}(\boldsymbol{x}^*)$ is the Pareto front. Various evolutionary multi-objective optimization (EMO) algorithms have been proposed to search for a solution set to approximate the entire Pareto front.

In the last two decades, NSGA-II [1] has always been the most frequently-used EMO algorithm. It has a $(\mu+\mu)$ generation update scheme where $\mu$ offspring solutions are created from $\mu$ solutions in the current population. The next population with $\mu$ solutions is selected from the merged population with $(\mu+\mu)$ solutions using two mechanisms: non-dominated sorting and crowding distance. Except for very early generations, the number of non-dominated solutions in the merged population is usually larger than the populations size $\mu$. In this case, NSGA-II removes solutions with the smallest crowding distance from the non-dominated solutions to choose $\mu$ solutions. Thus, the distribution of finally selected solutions depends on crowding distance. However, the effect of crowding distance on the distribution of obtained solutions has not

been carefully analyzed in many studies. To examine this effect, we discuss the optimal distributions of solutions for maximizing the minimum crowding distance for linear Pareto fronts of two- and three-objective problems.

With respect to the search behavior of NSGA-II, its severe performance deterioration by increasing the number of objectives has been clearly reported in the literature (e.g., [1], [2]). The reason for the performance deterioration was explained as follows. When we have many objectives (e.g., ten objectives), almost all solutions are non-dominated in the merged population. As a result, NSGA-II cannot generate strong selection pressure towards the Pareto front. Various mechanisms for enhancing the search ability of NSGA-II have been proposed for many-objective problems (see survey papers on evolutionary many-objective optimization [4]-[7]).

Crowding distance-related undesirable search behavior was also pointed out in the EMO community. In Figure 1, crowding distance calculation is explained for a two-objective problem. For simplicity, normalization is omitted. For solution C in Figure 1, the distance between its two adjacent solutions B and D is calculated for each objective, and the sum of the distance for all objectives (i.e., $c_1 + c_2$) is the crowding distance of C. For two extreme solutions A and F in Figure 1 with only one adjacent solution, an infinitely large value (i.e., ∞ in Figure 1) is assigned as their crowding distance. In NSGA-II, solutions with large crowding distance have high fitness. The basic idea of crowding distance-based solution selection is to delete solutions in crowded regions in the objective space. The special handling of the extreme solutions is useful as a diversification mechanism.

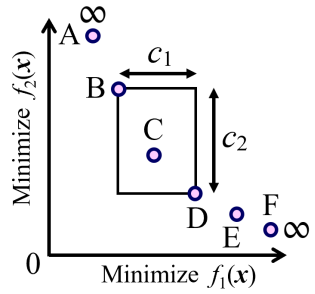



**Figure 1: Crowding distance of solution C.**

One difficulty in crowding distance-based solution selection in NSGA-II is related to its $(\mu+\mu)$ scheme. When all solutions in the merged population are non-dominated, $\mu$ solutions with the smallest crowding distance are removed with no crowding distance recalculation as pointed out in some studies (e.g., Kukkonen and Deb [8]). This issue is illustrated in Figure 2 where five solutions with the smallest crowding distance are removed from the merged population with ten solutions. In the right-hand side figure in Figure 2, there is a large gap around the center of the merged population after the five gray solutions are removed by the crowding distance-based generation update. The reason for this undesirable search behavior of NSGA-II is clear: the five gray solutions are removed without crowding distance recalculation. A modified mechanism was examined in [8] where the solution removal is performed one by one with crowding distance recalculation. This modification is illustrated in Figure 3 for the same example as in Figure 2. In Figure 3, well-distributed solutions are obtained as the next population. Almost the same effect as the crowding distance recalculation can be achieved by implementing NSGA-II as a steady-state algorithm (i.e., using the $(\mu+1)$ generation update scheme instead of the $(\mu+\mu)$ scheme). Steady-state variants of NSGA-II were examined in the literature [9]-[11], and better results were reported.

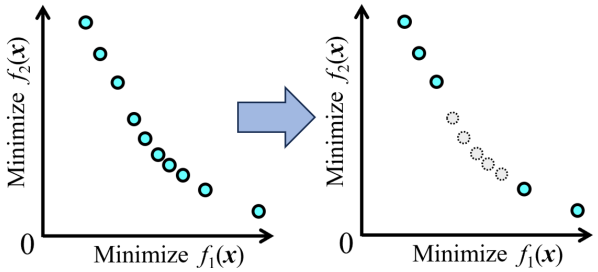



**Figure 2: Generation update in NSGA-II where five gray solutions with the smallest crowding distance are removed from the merged population with ten solutions.**

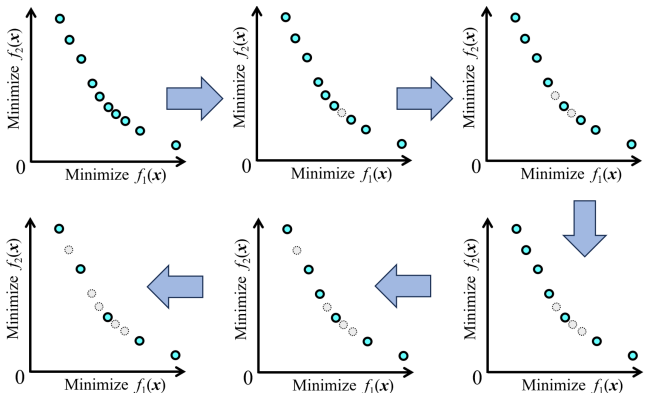


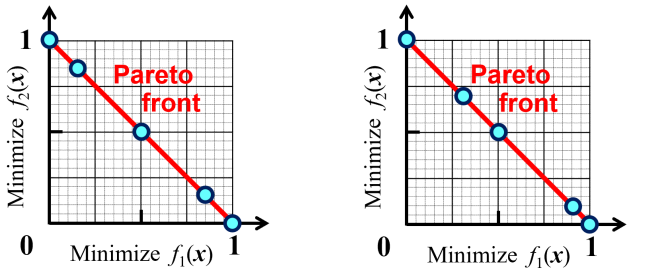

**Figure 3: Generation update in a modified NSGA-II algorithm in [8] where crowding distance of each solution is recalculated after one solution is removed (i.e., five gray solutions are removed from the merged population one by one).**

Another issue of crowding distance is related to the distance calculation between adjacent solutions on each axis of the objective space. As pointed out in some studies in the literature (e.g., [8], [12]), except for the case of two objectives, adjacent solutions on each axis of the objective space can be far away in the objective space with more than two objectives. That is, crowding distance cannot evaluate actual crowdedness of each solution in the objective space in the case of more than two objectives. As a result, uniformly distributed solutions cannot be obtained by NSGA-II except for the case of two objectives even for simple test problems (e.g., see [11]) for which uniformly distributed solutions can be easily obtained by other EMO algorithms such as MOEA/D [13] and NSGA-III [14].

Whereas various difficulties have been pointed out for NSGA-II, it was also demonstrated that NSGA-II outperforms other EMO algorithms (including some state-of-the-art algorithms) in their applications to non-standard multi-objective test problems [15], real-world multi-objective problems [16] and constrained real-world multi-objective problems [17], [18]. High performance of NSGA-II is due to its ability to handle realistic features of multi-objective problems such as objectives with totally different scales, correlated objectives, and irregular and degenerate Pareto fronts (e.g., [19]-[21]). Recently, NSGA-II has also been actively studied theoretically [22]-[27]. That is, whereas NSGA-II was proposed more than 20 years ago and various difficulties have been pointed out, it is still frequently used and continues to receive much attention in theoretical studies.

One approach for examining the search behavior of an EMO algorithm is to find the optimal distribution of solutions for that algorithm. For example, we can analyze the search behavior of an indicator-based algorithm (e.g., SMS-EMOA [28], [29]) using the optimal distribution of solutions for the indicator in the algorithm [30]. For a decomposition-based algorithm (e.g., MOEA/D [14], NSGA-III [15]), the optimal distribution of solutions is the set of the best solution for each subproblem. Recently, the optimal distribution of solutions was examined for maximizing the minimum crowding distance on a linear Pareto front of two-objective problems by Ishibuchi and Pang [11]. In their analysis, symmetric distributions were assumed for deriving the optimal distribution on the linear Pareto front of two-objective problems as shown in Figure 4 (a).

This paper is different from [11] in the following three points. (i) In this paper, we derive the optimal distributions under a more general setting including asymmetric distributions whereas only symmetric distributions were used in [11]. This difference is illustrated in Figure 4. (ii) We discuss the optimal distributions for two- and three-objective problems whereas only two-objective problems were used in [11]. (iii) We implement an indicator-based EMO algorithm for maximizing the minimum crowding distance whereas no indicator-based algorithm was used in [11].




**(a) Symmetric distribution.** **(b) Asymmetric distribution.**

**Figure 4: Symmetric distribution of solutions in [11] in (a) and general asymmetric distribution of solutions in this paper in (b).**

This paper is organized as follows. In Section 2, we derive the upper bound on the total crowding distance (i.e., the sum of the crowding distance of each solution). Based on this upper bound, we derive the upper bound on the minimum crowding distance. In Section 3, we derive the optimal distribution of solutions on the linear Pareto front for the case of two objectives under various settings of the population size. We also implement an indicator-based EMO algorithm where the minimum crowding distance is used as the indicator to be maximized. Using this algorithm, the theoretically derived optimal distributions are compared with the numerically obtained results. In Section 4, based on the results for two-objective problems in Section 3, we construct an optimal distribution of solutions for the case of three objectives. We also visually compare three algorithms: the original NSGA-II algorithm, its steady-state variant, and the newly implemented indicator-based algorithm. Finally, in Section 5, we conclude this paper.

## 2 Upper Bounds on Total and Minimum Crowding Distance

### 2.1 Assumptions in This Paper

In this paper, we discuss the optimal distribution of solutions on the Pareto front. Thus, we assume that all solutions are on the Pareto front. We also assume that the objective space has already been normalized so that the ideal and nadir points are (0, ..., 0) and (1, ..., 1), respectively. Since our discussions are for the maximization of the minimum crowding distance, we assume that there exists a solution with the minimum value 0 and the maximum value 1 for each objective as in Figure 4. In this section, we do not need any other assumptions about the shape of the Pareto front. This is because all discussions in this section are on each axis of the objective space.

However, in Section 3, we assume two-objective minimization problems with the following linear Pareto front in $[0, 1] \times [0, 1]$:

$$f_1 + f_2 = 1,\ 0 \le f_1 \le 1,\ 0 \le f_2 \le 1. \quad (2)$$

This Pareto front is the same as the red line in Figure 4.

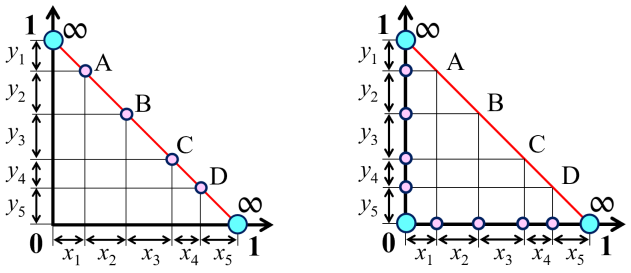
In Section 4, we assume three-objective minimization problems with the following linear Pareto front in $[0, 1] \times [0, 1] \times [0, 1]$:

$$f_1 + f_2 + f_3 = 1,\ 0 \le f_1 \le 1,\ 0 \le f_2 \le 1,\ 0 \le f_3 \le 1. \quad (3)$$

The normalized versions of the two- and three-objective DTLZ1 test problems [31] have the same Pareto fronts as in (2) and (3), respectively. The original $m$-objective DTLZ1 test problems have their Pareto fronts in $[0, 0.5] \times \ldots \times [0, 0.5]$. That is, their ideal and nadir points are (0, ..., 0) and (0.5, ..., 0.5), respectively.

### 2.2 Total Crowding Distance

Since we assume that all solutions are on the Pareto front and the objective space has already been normalized, all objective values of each solution are in the unit interval [0, 1]. Thus, the crowding distance calculation is also performed in the unit interval [0, 1] on each axis of the objective space as shown in Figure 5.




**(a) In the objective space.** **(b) On each axis.**

**Figure 5: Crowding distance calculation of each solution.**

In Figure 5, the crowding distance of each solution (except for the two blue solutions with an infinitely large crowding distance) is calculated based on its two adjacent solutions on each axis. For example, the crowding distance of solution A is calculated as $x_1 + x_2 + y_1 + y_2$. In Figure 5, the total crowding distance excluding an infinitely large crowding distance is calculated as follows (using $x_1 + x_2 + x_3 + x_4 + x_5 = 1$ and $y_1 + y_2 + y_3 + y_4 + y_5 = 1$):

$$x_1 + 2x_2 + 2x_3 + 2x_4 + x_5 + y_1 + 2y_2 + 2y_3 + 2y_4 + y_5 = 2 - x_1 - x_5 + 2 - y_1 - y_5. \quad (4)$$

When we have $N$ solutions with finite crowding distance, this equation is generalized as follows:

$$x_1 + 2x_2 + \ldots + 2x_N + x_{N+1} + y_1 + 2y_2 + \ldots + 2y_N + y_{N+1} = 2 - x_1 - x_{N+1} + 2 - y_1 - y_{N+1}. \quad (5)$$

Thus, the upper bound on the total crowding distance is 4 for two-objective problems independent of the number of solutions. This upper bound is obtained by decreasing $x_1$, $x_5$, $y_1$, and $y_5$ to zero in Figure 5. That is, the total crowding distance is maximized when two solutions overlap at each edge of the Pareto front in the case of two objectives (i.e., at (0, 1) and (1, 0)).

In general, the upper bound is $2m$ for $m$-objective problems since the upper bound for each objective is 2 (e.g., the upper bound of $(2 - x_1 - x_{N+1})$ in (5) is 2). This upper bound is obtained when two solutions overlap at $f_i = 0$ and $f_i = 1$ on each axis of the objective space (e.g., $x_1 = 0$ and $x_{N+1} = 0$ in (5)). Except for the case of two objectives, the overlapping of two solutions at $f_i = 0$ and $f_i = 1$ on each axis does not mean the overlapping in the objective space. In two-objective problems, two solutions overlap at each edge of the Pareto front (i.e., blue points in Figure 5 (a)).

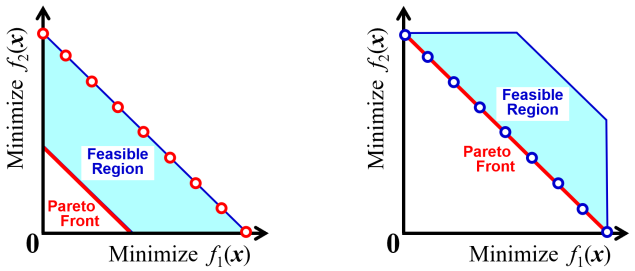
## 2.3 Minimum Crowding Distance

Let us denote the number of solutions with finite crowding distance by $N$ (e.g., $N = 4$ in Figure 5 (a)). Since the upper bound of the total crowding distance over $N$ solutions is $2m$ for $m$-objective problems, it is clear that the upper bound of the minimum crowding distance is $2m/N$ for $m$-objective problems. One issue is the number of solutions with finite crowding distance in the population $P$ of size $\mu$. On each axis, two extreme solutions have an infinitely large crowding distance. The number of those solutions in $m$-objective problems is $2m$ or less. This is because one solution can be an extreme solution for multiple axes such as Figure 5 (a). This issue will be discussed later using experimental results in this paper.

From the above discussions, the upper bound of the minimum crowding distance for the population $P$ of an $m$-objective problem is $2m/N \leq 2m/(\mu - 2m)$.

It should be noted that all discussions in this section hold independently of the shape of the Pareto front. For example, for the case of two objectives, even when the Pareto front is nonlinear, the upper bounds of the total crowding distance and the minimum crowding distance are obtained when two solutions overlap at (1, 0) and (0, 1) in the normalized objective space.

We discuss the optimal distributions of solutions by assuming that all solutions are on the Pareto front. This assumption is used since the final goal of all EMO algorithms including NSGA-II is to find a pre-specified number of solutions on the Pareto front. In NSGA-II, non-dominated sorting is used as the primary solution selection criterion to push the population towards the Pareto front. This assumption is also used to make sure that all solutions are in the first front since the crowding distance calculation is performed in each front. If we remove this assumption, we can obtain a larger value of the minimum crowding distance. In Figure 6 (a), the nine points far away from the Pareto front have a larger minimum crowding distance than any set of nine solutions on the Pareto front. Thus, the maximization of the minimum crowding distance is conflict with the convergence towards the Pareto front. However, in Figure 6 (b), solutions are pushed towards the Pareto front by maximizing the minimum crowding distance. As shown in Figure 6, the effect of crowding distance on the convergence ability of NSGA-II depends on the shape of the feasible region in the objective space. This means that the performance of NSGA-II also strongly depends on its shape.




**(a) Negative effects.** **(b) Positive effects.**

**Figure 6: Different effects of crowding distance on the convergence of solutions towards the Pareto front.**

# 3 Optimal Distributions for Two-Objective Problem

In this section, first we derive the optimal distributions of solutions on the linear Pareto front in Eq. (2) of two-objective problems for maximizing the minimum crowding distance. Next, we implement an indicator-based EMO algorithm to maximize the minimum crowding distance. Then, experimental results are compared with the analytically derived optimal distributions.

### 3.1 Theoretical Analysis

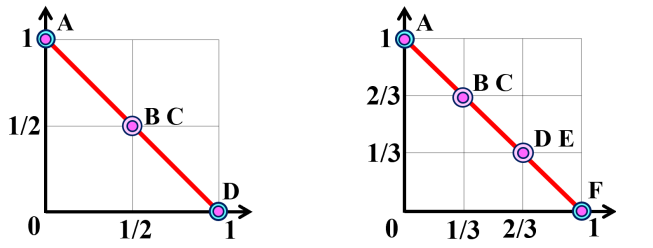
Let us consider the optimal distribution of four solutions A, B, C and D in Figure 5 (a). Since the number of solutions with finite crowding distance is 4, the upper bound of the minimum crowding distance is $2m/N = 1$. We denote the crowding distance of each solution by Dis(.) such as Dis(A). Since the Pareto front is linear, the following relation holds: $x_i = y_i$ for $i = 1, 2, ..., 5$. Thus, the crowding distance of each solution is Dis(A) = $2x_1 + 2x_2$, Dis(B) = $2x_2 + 2x_3$, Dis(C) = $2x_3 + 2x_4$, and Dis(D) = $2x_4 + 2x_5$.

Our optimization problem for Figure 5 (a) can be written as the following LP problem using an additional decision variable $k$:

**LP Problem ($N = 4$)**: Maximize $k$,

$$\text{subject to } x_1 + x_2 + x_3 + x_4 + x_5 = 1,$$
$$2x_1 + 2x_2 \geq k, \quad 2x_2 + 2x_3 \geq k,$$
$$2x_3 + 2x_4 \geq k, \quad 2x_4 + 2x_5 \geq k.$$

This LP problem has the unique solution: $k^* = 1$, $x_1^* = 0$, $x_2^* = 1/2$, $x_3^* = 0$, $x_4^* = 1/2$, $x_5^* = 0$. In this case, the minimum crowding distance is equal to its upper bound 1. The optimal distribution of solutions is A = (0, 1), B = (1/2, 1/2), C = (1/2, 1/2), and D = (1, 0) as in Figure 7 (a). That is, two solutions overlap at each of the following three locations: (0, 1), (1/2, 1/2), (1, 0). In this case, the crowding distance of each solution (A, B, C and D) is 1.




**(a) $N = 4$ ($\mu = 6$).** **(b) $N = 6$ ($\mu = 8$).**

**Figure 7: Optimal distributions of solutions when the number of solutions with finite crowding distance is 6 and 8.**

For a nonlinear Pareto front, we can obtain a similar result. From the discussions in Section 2, we have A = (0, 1) and D = (1, 0) independently of the Pareto front shape. In this case, the crowding distance of each solution is Dis(A) = MD(A, B), Dis(B) = MD(A, C), Dis(C) = MD(B, D) and Dis(D) = MD(C, D) where MD(A, B) is the Manhattan distance between A and B. Thus, the minimum crowding distance is maximized when MD(A, B) = MD(A, C) = MD(D, B) = MD(D, C) = 1. That is, B and C overlap at the point on the Pareto front with the same Manhattan distance from A = (0, 1) and D = (1, 0) in a similar manner as in Figure 7 (a). We leave more detailed rigorous discussions for the case of nonlinear Pareto fronts as a future research topic.

When the number of solutions with finite crowding distance is an even number (i.e., when $N$ is an even number), our linear programming problem has a similar unique solution to the case of $N = 4$. For example, when $N = 6$, we have seven decision variables $x_i$ ($i = 1, 2, ..., 7$) and an additional decision variable $k$. Then, the corresponding linear programming problem has the following unique optimal solution: $k^* = 2/3$, $x_1^* = 0$, $x_2^* = 1/3$, $x_3^* = 0$, $x_4^* = 1/3$, $x_5^* = 0$, $x_6^* = 1/3$, $x_7^* = 0$. The minimum crowding distance 2/3 is the same as its upper bound $2m/N = 4/6 = 2/3$. The optimal distribution is shown in Figure 7 (b) where two solutions overlap at each of the following four points (0, 1), (1/3, 2/3), (2/3, 1/3) and (1, 0) on the linear Pareto front. In this case, the crowding distance of each of the six solutions A, B, C, D, E and F is 2/3.

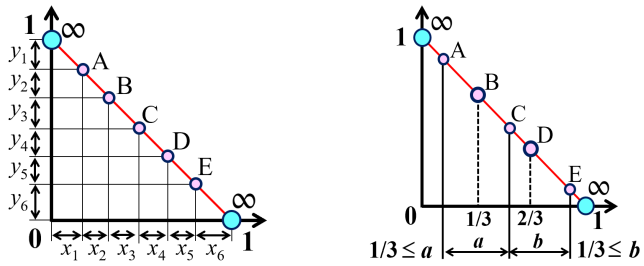
When $N$ is an odd number, our linear programming problem has many optimal solutions. Let us consider the case of $N = 5$. In this case, we have six decision variables $x_i$ ($i = 1, 2, ..., 6$) as in Figure 8 (a) and an additional decision variable $k$. Then, the corresponding LP problem is formulated as follows:

**LP Problem ($N = 5$)**: Maximize $k$,

$$\text{subject to } x_1 + x_2 + x_3 + x_4 + x_5 + x_6 = 1,$$
$$2x_1 + 2x_2 \geq k, \quad 2x_2 + 2x_3 \geq k,$$
$$2x_3 + 2x_4 \geq k, \quad 2x_4 + 2x_5 \geq k,$$
$$2x_5 + 2x_6 \geq k.$$

From the 1st, 3rd and 5th inequality conditions, we have the following relation.

$$2x_1 + 2x_2 + 2x_3 + 2x_4 + 2x_5 + 2x_6 \geq 3k. \quad (6)$$




**(a) Six decision variables.** **(b) $N = 5$ ($\mu = 7$).**

**Figure 8: Six decision variables in the LP problem in (a) and an example of the optimal distribution of solutions for $N = 5$ in (b).**

Using the equality condition of the LP Problem ($N = 5$), we have the following inequality relation from Eq. (6): $2 \geq 3k$. Since our problem is to maximize $k$, $k^* = 2/3$ is the optimal value. From the 1st, 3rd and 5th inequality conditions and the equality condition, we have the following relation for the optimal solutions: $x_1^* + x_2^* = 1/3$, $x_3^* + x_4^* = 1/3$, and $x_5^* + x_6^* = 1/3$. We can see from Figure 8 (a) that these relations fix the locations of solutions B and D. The locations of the other solutions A, C and E are not uniquely specified. From the 2nd and 4th inequality relations, the locations of these solutions should satisfy the following relations: $x_2^* + x_3^* \geq 1/3$ and $x_4^* + x_5^* \geq 1/3$. An example of the optimal distribution of solutions is shown in Figure 8 (b). Whereas the locations of solution B and D are fixed at (1/3, 2/3) and (2/3, 1/3), respectively, solutions A, C and E can be at any other locations satisfying the two conditions: $x_2^* + x_3^* \geq 1/3$ and $x_4^* + x_5^* \geq 1/3$. In Figure 8 (b), the crowding distance

of solutions A, C and E is 2/3, the crowding distance of solution B is $2a$, and the crowding distance of solution D is $2b$. The minimum crowding distance is 2/3, which is smaller than the upper bound of the crowding distance $2m/N = 4/5$. This is because the total crowding distance cannot be equally distributed to each solution. The total crowding distance is maximized by $x_1 = 0$ and $x_6 = 0$ in Figure 8 (a). In this case, if solution C is located at the midpoint (1/2, 1/2) of the Pareto front, the crowding distance of solutions B and D is 1, which is larger than the crowding distance 2/3 of solutions A, C and E.

When $N$ is an odd number, the corresponding LP problem has a similar structure as in the case of $N = 5$. For example, when $N = 7$, we have the following relation for the optimal solution of the corresponding LP problem: $x_1^* + x_2^* = x_3^* + x_4^* = x_5^* + x_6^* = x_7^* + x_8^* = 1/4$, $x_2^* + x_3^* \geq 1/4$, $x_4^* + x_5^* \geq 1/4$, and $x_6^* + x_7^* \geq 1/4$. This means that the locations of three solutions (say, solutions B, D and F) are fixed at (1/4, 3/4), (1/2, 1/2), and (3/4, 1/4). The other solutions (say, solutions A, C, E, and G) can be at any locations satisfying $x_2^* + x_3^* \geq 1/4$, $x_4^* + x_5^* \geq 1/4$, and $x_6^* + x_7^* \geq 1/4$. The minimum crowding distance is 1/2, which is smaller than the upper bound of the minimum crowding distance $2m/N = 4/7$.

## 3.2 Computational Experiments

To search for the optimal distribution of solutions for maximizing the minimum crowding distance, we implement an indicator-based EMO algorithm using the minimum crowding distance as an indicator to be maximized. We use a steady-state variant of NSGA-II as the base algorithm where only one solution is created in each generation. Except for the number of offspring in each generation, the steady-state NSGA-II is the same as the original NSGA-II. The next population with $\mu$ solutions is selected from the merged population with $(\mu + 1)$ solutions. In our indicator-based algorithm, only one solution is created in each generation. When at least one solution is a dominated solution in the merged population, we use exactly the same generation update mechanism as in the steady-state NSGA-II. Only when all solutions are non-dominated, we use the indicator to identify a single solution to be removed from the merged population. More specifically, we remove a single solution to maximize the minimum crowding distance of the remaining solutions. In this manner, we search for the optimal distribution of solutions for maximizing the minimum crowding distance.

In our computational experiments, we apply the implemented indicator-based algorithm to the two-objective DTLZ1 problem [31] with a linear Pareto front. Since the ideal and nadir points of DTLZ1 are (0, 0) and (0.5, 0.5), we show our experimental results in the normalized objective space with the ideal point (0, 0) and the nadir point (1, 1). In order to focus on the optimization of the solution distribution on the Pareto front, we specify the number of distance variables in DTLZ1 as zero. Under this setting, all feasible solutions are Pareto optimal solutions. This special setting about the number of distance variables has been used in some studies on the optimal distributions of solutions for other indicators (e.g., [32]). In the next section, we also use DTLZ1 with both distance and position variables to compare three algorithms: NSGA-II, its steady-state variant, and our indicator-based algorithm.

Our indicator-based algorithm is applied to the two-objective DTLZ1 problem with no distance variables under the following setting in the PlatEMO platform [33]:

Population size $\mu$: $\mu = 3, 4, 5, \ldots, 100$.
Termination condition: $50{,}000\mu$ solution evaluations.
Crossover: SBX crossover with distribution index 20.
Crossover probability: 1.
Mutation: Polynomial mutation with distribution index 20.
Mutation probability: $1/n$ ($n$: the number of decision variables).
Number of independent runs: 10.

Experimental results of ten runs for $\mu = 8, 9, 10, 11$ are shown in Figures 9-12. We can see that the obtained solutions in Figure 9 ($\mu = 8$) are the same as Figure 7 (a) for $N = 4$ ($\mu = 6$) with the minimum crowding distance 1 ("Min CD = 1.0000" in Figure 9), which is the same as its upper bound. This means that the optimal distributions of solutions are obtained.

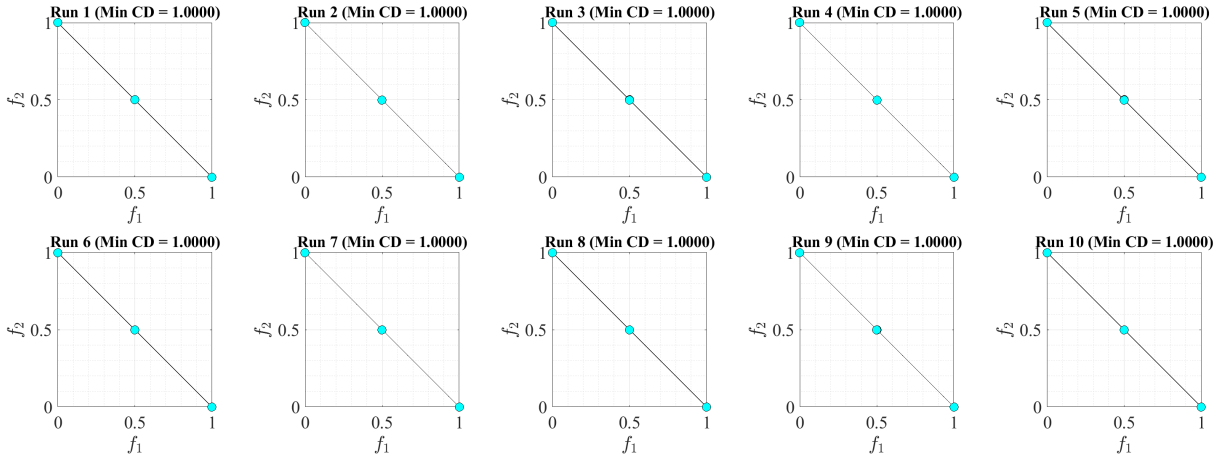



**Figure 9: Results by our indicator-based algorithm ($\mu = 8$), which are consistent with theoretical analysis for $\mu = 6$ in Figure 7 (a).**

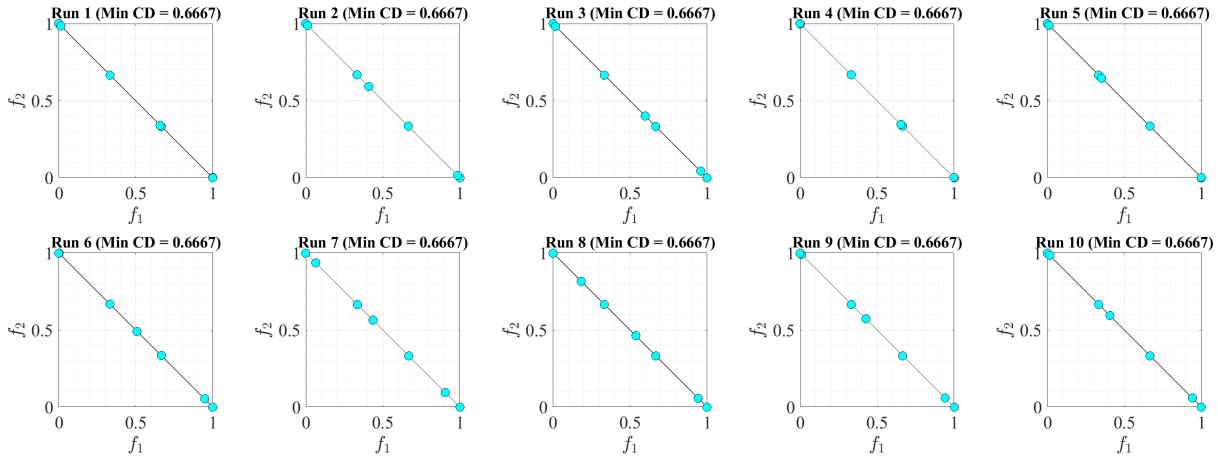



**Figure 10: Results by our indicator-based algorithm ($\mu = 9$), which are consistent with theoretical analysis for $\mu = 7$ in Figure 8 (b).**

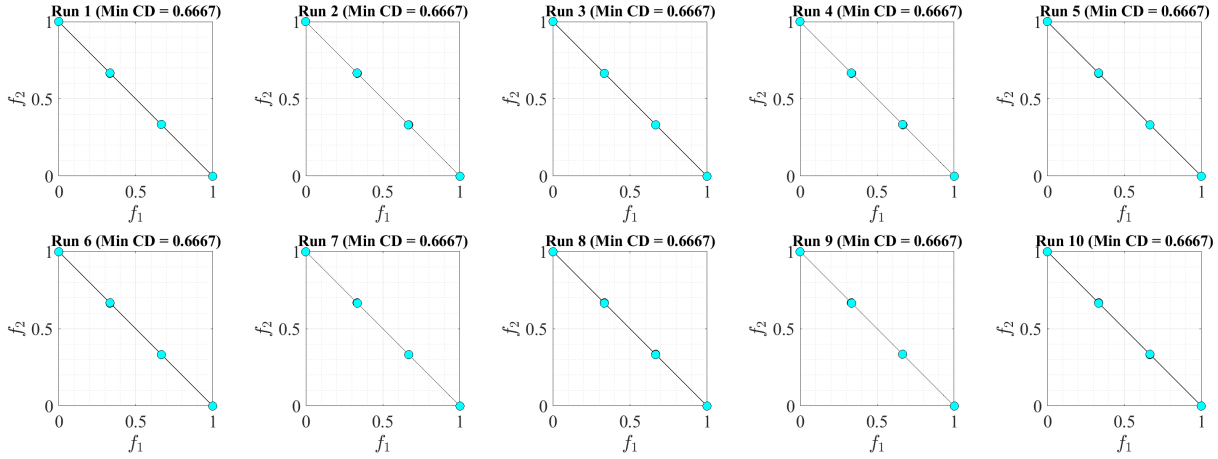



**Figure 11: Results by our indicator-based algorithm ($\mu = 10$), which are consistent with theoretical analysis for $\mu = 8$ in Figure 7 (b).**

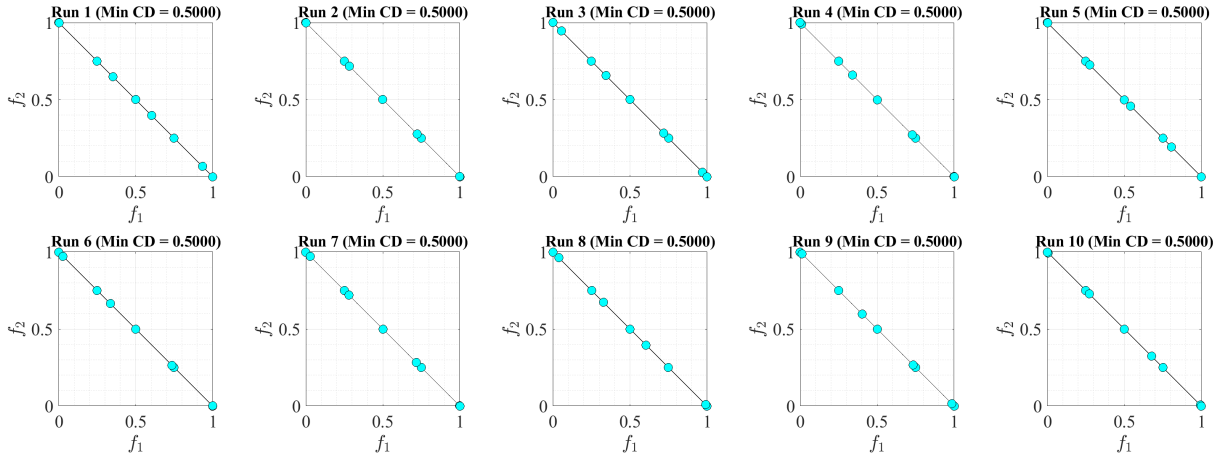



**Figure 12: Results by our indicator-based algorithm ($\mu = 11$), which are consistent with theoretical analysis for $\mu = 9$.**

We can also see that the obtained solutions in Figure 11 for $\mu = 10$ are the same as Figure 7 (b) for $N = 6$. In Figure 10 for $\mu = 9$, we can observe the two anchor solutions at (1/3, 2/3) and (2/3, 1/3) as in Figure 8 (b) for $N = 5$. In Figure 12 for $\mu = 11$, we can observe the three anchor solutions for the case of $N = 7$. These observations show that the number of solutions with finite crowding distance is $(\mu - 4)$ instead of $(\mu - 2)$. This means that there exist two extreme solutions with an infinitely large crowding distance at each edge of the linear Pareto front (i.e., at (0, 1) and (1, 0)).

Let us explain this observation for the case of $\mu = 8$ with eight solutions A, B, C, D, E, F, G, H in an ascending order of the first objective values. If all solutions are different, these eight solutions are ordered as H, G, F, E, D, C, B, A based on the second objective values. In this case, only two solutions A and H have an infinitely large crowding distance. Let us assume that two solutions A and B overlap at (0, 1) and two solutions G and H overlap at (1, 0). In this case, the solutions are ordered on each axis as shown in Figure 13. If solution A is handled as having a smaller $f_1$ value than solution B, it is likely that A is also handled as having a smaller $f_2$ value than B. Then, A is an extreme solution on the $f_1$ axis (since it has the smallest $f_1$ value), and B is an extreme solution on the $f_2$ axis (since it has the largest $f_2$ value). Thus, both A and B have an infinitely large crowding distance.

For comparison, we show the final populations of NSGA-II and its steady-state variant in Figures 14 and 15, respectively, for the case of population size 11. Uniform distributions are obtained by the steady-state NSGA-II in Figure 15. The obtained distributions by NSGA-II are not so good in Figure 14. We can also observe that only nine solutions are obtained in all results in Figure 15 and most results in Figure 14.

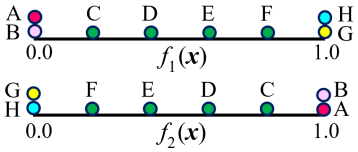



**Figure 13: Eight solutions on each axis.**

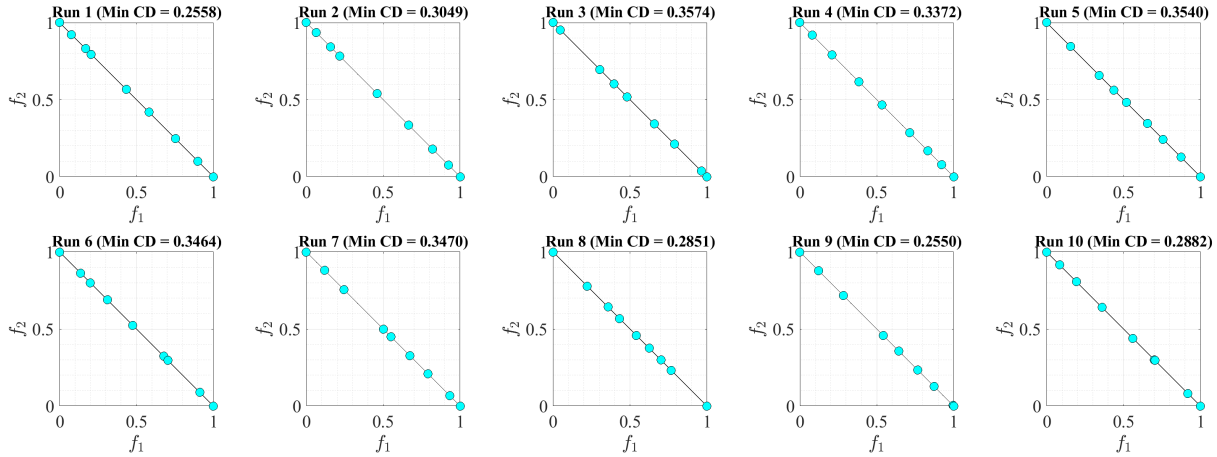



**Figure 14: Experimental results by the original NSGA-II (population size 11).**

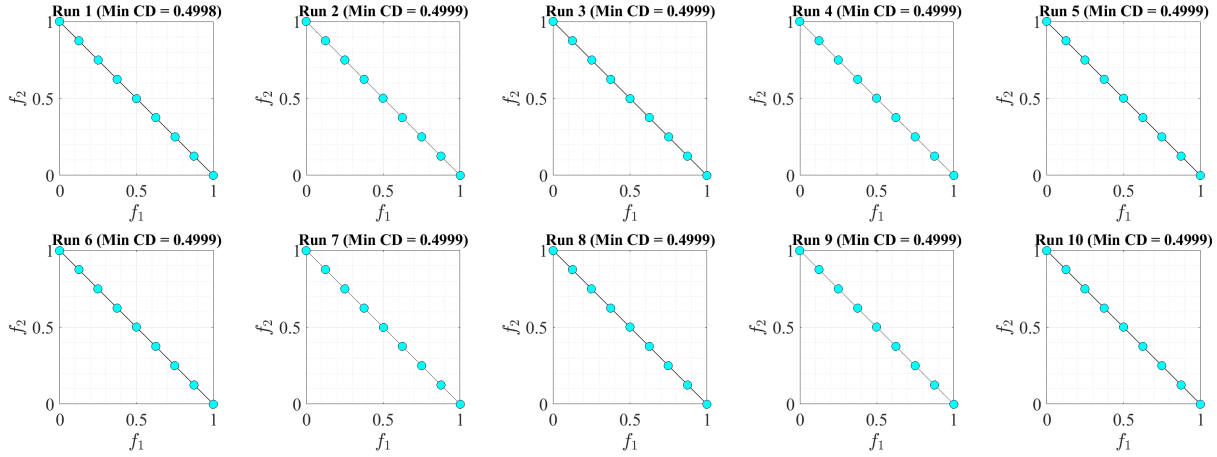



**Figure 15: Experimental results by the steady-state NSGA-II (population size 11).**

In those results, two solutions are overlapping at each edge of the linear Pareto front as in Figures 9-12 by the indicator-based algorithm. When the population size is large (i.e., 100, 200) for two-objective problems, the overlapping two extreme solutions at each edge of the Pareto front are not a serious issue. However, if the population size is small (e.g., 10 for expensive problems), it may be a good idea to use a special mechanism for removing duplicated solutions or avoiding duplicated solution creation. When the population size is very small, such a minor modification can improve the performance of NSGA-II.

Our experimental results show that the best results (i.e., well-distributed solutions) are obtained by the steady-state variant of NSGA-II as explained in some other studies (e.g., [9]-[11]).

## 4 Optimal Distributions for Three-Objective Problems

In this section, first we show an example of the optimal distributions of solutions on the linear Pareto front in Eq. (3) of three-objective problems. Next, we show experimental results on the three-objective DTLZ1 problem by the three algorithms.

### 4.1 Theoretical Analysis

For any performance indicator (e.g., hypervolume), it is not easy to discuss the optimal distribution of solutions for three-objective problems. In this section, we focus on the distribution of solutions along the boundary of the linear Pareto front in Eq. (3). More specifically, we explain that the distribution of solutions in Figure 16 is an example of the optimal distribution of 18 solutions. In Figure 17, those solutions are shown on the $f_1$ axis.

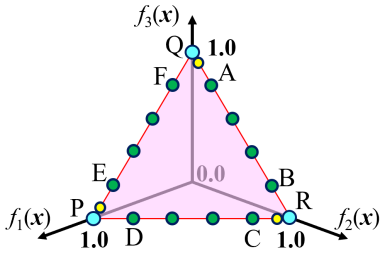



**Figure 16: An example of the optimal distribution of solutions on the linear Pareto front of a three-objective problem.**

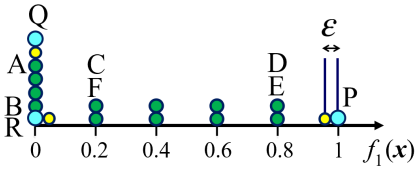



**Figure 17: The projected solutions on the $f_1$ axis.**

In Figure 16, we assume that only three corner solutions (i.e., solutions P, Q and R) have an infinitely large crowding distance. Thus, the number of solutions with finite crowding distance is 15 (i.e., $N = 15$). The upper bound of the minimum crowding distance is $2m/N = 2/5 = 0.4$. This value is used to explain that the distribution of 18 solutions in Figure 16 is optimal.

In Figure 17, we can see that two overlapping solutions are uniformly located at 0.2, 0.4, 0.6 and 0.8. This is the same as the optimal distribution of 12 solutions on the linear Pareto front of two-objective problems for the case of $N = 10$ in Section 3.1. One small yellow solution around 1.0 is very close to 1.0 (i.e., $\varepsilon$ is very close to zero). Another small yellow solution around 0.0 is very close to 0.0. Thus, the crowding distance of those two small yellow solutions on the $f_1$ axis is very close to 0.2. The other eight solutions at 0.2, 0.4, 0.6 and 0.8 have the crowding distance of 0.2 (or very close to 0.2). If $\varepsilon$ is zero, the crowding distance of all of those ten solutions on the $f_1$ axis is 0.2. It should be noted that this value "0.2" is the calculation result on the $f_1$ axis (which is not the final crowding distance calculation result). All solutions at 0.0 in Figure 17 (except for one of the two extreme solutions Q and R) has zero crowding distance on the $f_1$ axis. Those solutions are located on the line QR in the $f_2$-$f_3$ plane in Figure 16. The crowding distance is also calculated on the $f_2$ axis and the $f_3$ axis. Each of the 15 solutions in Figure 16 (except for the three corner solutions) have crowding distance 0.2 on two out of the three axes. Thus, all of those solutions have crowding distance 0.4. This is the same as the upper bound of the minimum crowding distance. Thus, the 18 solutions in Figure 16 are optimal when each of the three small yellow solutions approaches to the nearest corner solution (i.e., $\varepsilon$ becomes zero).

As we have already explained, $2m$ extreme solutions can have an infinitely large crowding distance (e.g., four solutions in the case of two objectives). In this case, we can add one additional solution at each corner in Figure 16. Then we have an example of the optimal distribution of 21 solutions.

### 4.2 Computational Experiments

In the same manner as in Section 3.2, we apply the three algorithms to the three-objective DTLZ1. Whereas we examined two settings with and without distance variables, we cannot observe any clear differences in the final populations. Thus, we only show the final population of each run of each algorithm in Figures 18-20 for the case of no distance variables. For the three-objective DTLZ1, it is difficult to say which algorithm is the best from Figures 18-20.

Since the population size is 66 in these figures, the upper bound of the minimum crowding distance is $2m/N = 6/(66 - 6) = 0.1$ since $2m = 6$ solutions can have an infinitely large crowding distance. The obtained minimum crowding distance by the indicator-based algorithm is from 0.0740 to 0.0887 in its 10 runs in Figure 20. Whereas the optimal distributions of solutions are always obtained by our indicator-based algorithm in Figures 9-12 for the two-objective DTLZ1 in Section 3, the optimal distributions are not obtained for the three-objective DTLZ1 in Figure 20. It seems from the experimental results in Figure 20 that the maximization of the minimum crowding distance is difficult for the three-objective DTLZ1. Having said that, the minimum crowding distance in [0.0740, 0.0887] in Figure 20 is not so bad if compared with its upper bound 0.1. However, the distribution of

solutions is not so good in Figure 20. This observation supports the well-known difficulty in using crowding distance for three or more objectives. Using the standard three-objective DTLZ1 with the distance variables, we also examine the anytime performance of each algorithm. Figures 21-23 show experimental results for the case of population size 91. We calculate the average values over 10 runs for the minimum crowding distance, the HV indicator, and the IGD indicator.

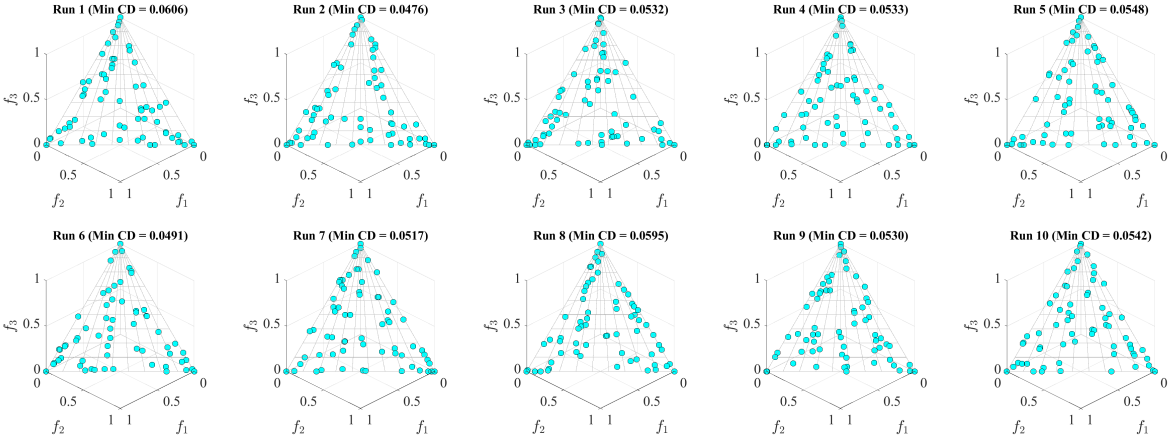



**Figure 18: Experimental results by the standard NSGA-II (population size $\mu = 66$).**

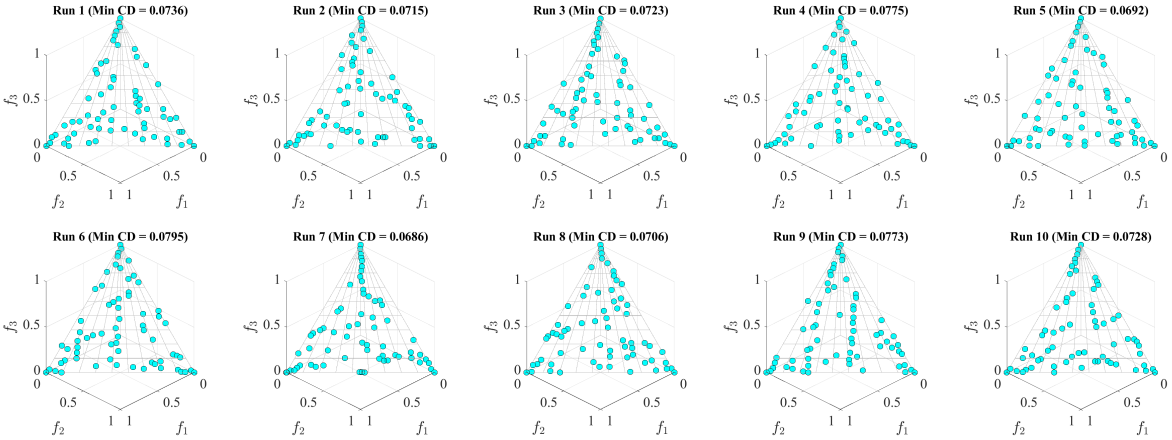



**Figure 19: Experimental results by the steady-state NSGA-II (population size $\mu = 66$).**

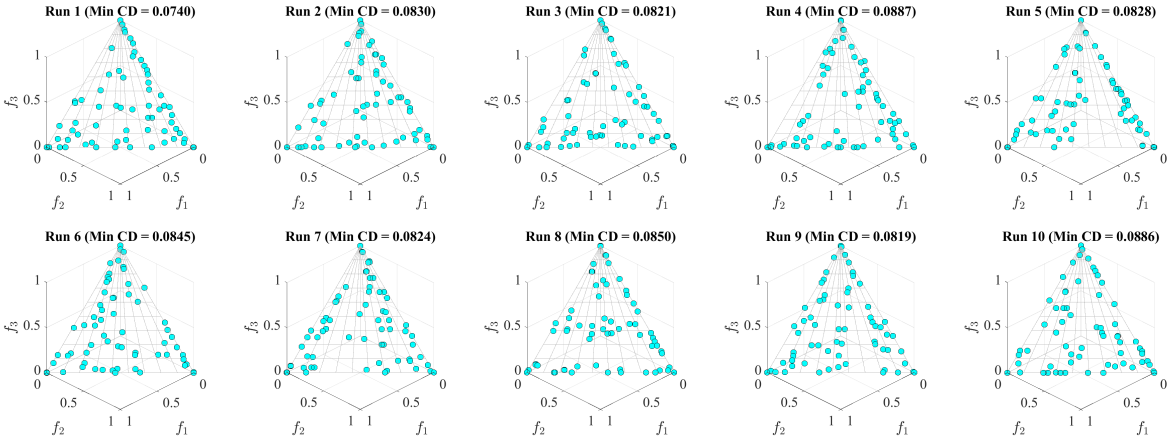



**Figure 20: Experimental results by the indicator-based algorithm (population size $\mu = 66$).**

When we use the minimum crowding distance to examine the anytime performance in Figure 21, slightly better results are obtained from our indicator-based algorithm than the steady-state variant of NSGA-II, and much better results are obtained than the original NSGA-II. As we have already explained using Figure 20, the maximization of the minimum crowding distance is difficult for three-objective problems. So, we may need much more generations in Figure 21 for obtaining the near optimal results by our indicator-based algorithm. When we use HV (in Figure 22) and IGD (in Figure 23), almost the same results are obtained from the three algorithms. This observation is consistent with our visual comparison results in Figures 18-20.

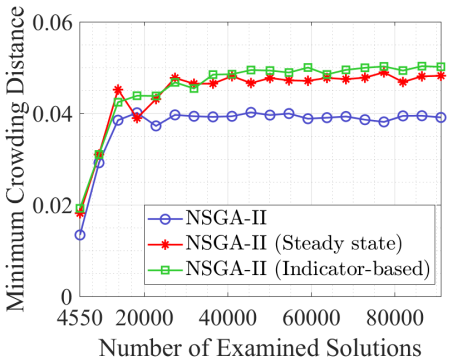



**Figure 21: Average minimum crowding distance values.**

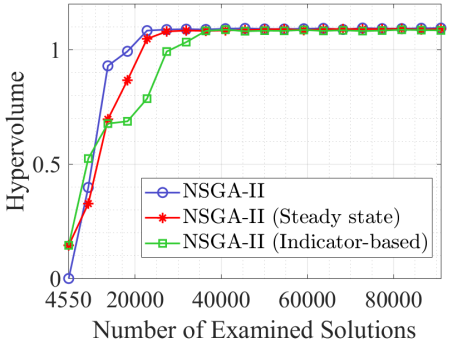



**Figure 22: Average HV values.**

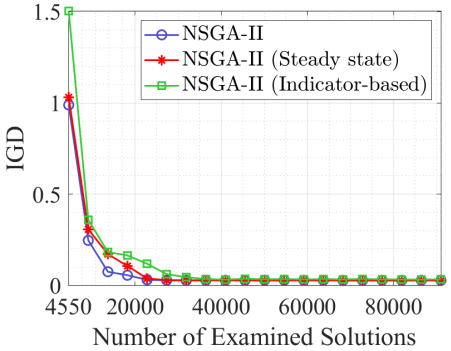



**Figure 23: Average IGD values.**

## 5 Conclusions and Discussions

In this paper, we theoretically and empirically examined the optimal distributions of solutions for maximizing the minimum crowding distance. Our theoretical considerations for two-objective problems showed that uniformly distributed solutions do not maximize the minimum crowding distance. For three-objective problems, we explained that solutions only on the boundary of the linear Pareto front can maximize the minimum crowding distance. However, this does not mean that solutions in the optimal distribution for maximizing the minimum crowding distance are always on the boundaries of the Pareto front. Our computational experiments were performed using three EMO algorithms: NSGA-II, its steady-state variant, and an indicator-based algorithm for maximizing the minimum crowding distance. For the two-objective DTLZ1, the best results in terms of the uniformity of solutions were obtained by the steady-state NSGA-II. All the three algorithms found solution sets with overlapping solutions at each edge of the two-objective linear Pareto front. However, solution distributions obtained by each algorithm are clearly different. This observation shows that NSGA-II does not maximize the minimum crowding distance. For three-objective DTLZ1, similar results were obtained by all the three algorithms. No good distributions were obtained by the examined algorithms.

One may think that our assumption (i.e., all solutions are on a linear Pareto front) and our experimental setting (i.e., DTLZ1 with no distance variables) are too special to explain the search behavior of NSGA-II. To address this issue, we apply the three algorithms in the same manner as in Section 3 to the two-objective minimization problem with the following two objectives and the two-dimensional decision vector $\boldsymbol{x}$ in $[0, 10] \times [0, 10]$:

$$f_1(\boldsymbol{x}) = \sqrt{(x_1 - 2)^2 + (x_2 - 9)^2}\,, \quad f_2(\boldsymbol{x}) = \sqrt{(x_1 - 8)^2 + (x_2 - 1)^2}\,. \qquad (7)$$

The final population of a single run after 100,000 generations is shown in Figure 24 for each algorithm with the population size 10. While some solutions are not on the Pareto front, we obtain similar observations from Figure 24 to those from the results in Section 3. For example, in Figure 24, two solutions are overlapping at five locations in the final population of the indicator-based algorithm in the right figure. In the final populations of NSGA-II (in the left figure) and its steady-state variant (in the center figure), two solutions are overlapping at each edge of the Pareto front (i.e., (0, 10) and (10, 0)). The best solution set is obtained from the steady-state variant of NSGA-II.

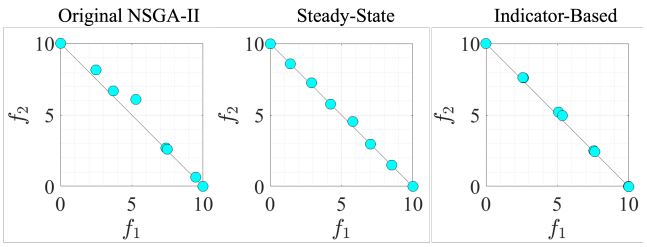



**Figure 24: Final population of each algorithm on the two-objective minimization problem in (7).**

The final population of NSGA-II in Figure 24 clearly shows its difficulty with respect to both the convergence and the uniformity. Whereas one solution is not close to the Pareto front, all solutions are non-dominated in the final population of NSGA-II in Figure 24. Moreover, the obtained solutions by NSGA-II are not uniformly distributed as explained in this paper about the use of crowding distance. Our theoretical and experimental discussions clearly show that NSGA-II does not maximize the minimum crowding distance whereas it removes the solutions with the smallest crowding distance. One future research direction is to extend our discussions to the case of many objectives and also the case of nonlinear Pareto fronts. Another interesting research direction is to examine why NSGA-II works well on real-world application problems whereas it has clear difficulties with respect to both the convergence and the uniformity of solutions. These future studies may lead to the design of more practically-useful EMO algorithms based on the NSGA-II framework.

## ACKNOWLEDGMENTS

This work was supported by the National Key R&D Program of China (Grant No. 2023YFE0106300), National Natural Science Foundation of China (Grant No. 62376115), Guangdong Provincial Key Laboratory (Grant No.2020B121201001).

## REFERENCES

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

[2] T. Wagner, N. Beume, and B. Naujoks. 2007. Pareto-, aggregation-, and indicator-based methods in many-objective optimization, In *Proceedings of the 4th International Conference on Evolutionary Multi-Criterion Optimization* (EMO 2007), Springer, Matsushima, Japan, 742-756.

[3] S. Mostaghim and H. Schmeck. 2008. Distance based ranking in many-objective particle swarm optimization. In *Proceeding of the 10th International Conference on Parallel Problem Solving from Nature* (PPSN 2008), Springer, Dortmund, Germany, 753-762.

[4] H. Ishibuchi, N. Tsukamoto, and Y. Nojima. 2008. Evolutionary many-objective optimization: A short review. In *Proceedings of 2008 IEEE Congress on Evolutionary Computation* (IEEE CEC 2008), IEEE, Hong Kong, 2424-243.

[5] C. von Lücken, B. Barán, and C. Brizuela. 2014. A survey on multi-objective evolutionary algorithms for many-objective problems. *Computational Optimization and Applications* 58 (2014) 707–756.

[6] B. Li, J. Li, K. Tang, and X. Yao. 2015. Many-objective evolutionary algorithms: A survey. *ACM Computing Surveys* 48, 1 (2015) Article No. 13.

[7] H. Sato and H. Ishibuchi. 2023. Evolutionary many-objective optimization: Difficulties, approaches, and discussions. *IEEJ Trans. on Electrical and Electronic Engineering* 18, 7 (2023) 1048-1058.

[8] S. Kukkonen and K. Deb. 2006. Improved pruning of non-dominated solutions based on crowding distance for bi-objective optimization problems. In *Proceedings of 2006 IEEE Congress on Evolutionary Computation* (IEEE CEC 2006), IEEE, Vancouver, Canada, 1179-1186.

[9] J. J. Durillo, A. J. Nebro, F. Luna, and E. Alba. 2009. On the effect of the steady-state selection scheme in multi-objective genetic algorithms. In *Proceedings of 2009 International Conference on Evolutionary Multi-Criterion Optimization* (EMO 2009), Springer, Nantes, France, 183-197.

[10] W. Zheng and B. Doerr. 2025. Approximation guarantees for the Non-dominated Sorting Genetic Algorithm II (NSGA-II). *IEEE Trans. on Evolutionary Computation* (Early Access).

[11] H. Ishibuchi and L. M. Pang. 2025. Optimal Distribution of Solutions for Crowding Distance on Linear Pareto Fronts of Two-Objective Optimization Problems. *arXiv preprint arXiv:2504.17222* (2025).

[12] W. Zheng and B. Doerr. 2024. Runtime analysis for the NSGA-II: Proving, quantifying, and explaining the inefficiency for many objectives. *IEEE Trans. on Evolutionary Computation* 28, 5 (2024) 1442-1454.

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

[14] K. Deb and H. Jain. 2014. 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 (2014) 577-601.

[15] H. Ishibuchi and L. M. Pang. 2025. Visual explanations of some problematic search behaviors of frequently-used EMO algorithms. In *Proceedings of 13th International Conference on Evolutionary Multi-Criterion Optimization* (EMO 2025) Vol. II, Springer Canberra, Australia, 3-16.

[16] H. Ishibuchi, Y. Nan, and L. M. Pang. 2023. Performance evaluation of multi-objective evolutionary algorithms using artificial and real-world problems. In *Proceedings of 12th International Conference on Evolutionary Multi-Criterion Optimization* (EMO 2023), Springer, Leiden, Netherland, 333-347.

[17] C. Picard and J. Schiffmann. 2021. Realistic constrained multiobjective optimization benchmark problems from design. *IEEE Trans. on Evolutionary Computation* 25, 2 (2021) 234-246.

[18] Y. Nan, H. Ishibuchi, T. Shu, and K. Shang. 2024. Analysis of real-world constrained multi-objective problems and performance comparison of multi-objective algorithms. In *Proceeding of 2024 Genetic and Evolutionary Computation Conference* (GECCO 2024), ACM, Melbourne, Australia, 576-584.

[19] R. Tanabe and H. Ishibuchi. 2020. An easy-to-use real-world multi-objective optimization problem suite. *Applied Soft Computing* 89 (2020) Article 106078.

[20] C. He, Y. Tian, H. Wang, and Y. Jin. 2020. A repository of real-world datasets for data-driven evolutionary multiobjective optimization. *Complex & Intelligent Systems* 6 (2020) 189-197.

[21] A. Kumar, G. Wu, M. Z. Ali, Q. Luo, R. Mallipeddi, P. N. Suganthan, and S. Das. 2021. A benchmark-suite of real-world constrained multi-objective optimization problems and some baseline results. *Swarm and Evolutionary Computation* 67 (2021) Article 100961.

[22] W. Zheng, Y. Liu, and B. Doerr. 2022. A first mathematical runtime analysis of the non-dominated sorting genetic algorithm II (NSGA-II). In *Proceedings of the 36th AAAI Conference on Artificial Intelligence* (AAAI 2022), AAAI, Virtual, 10408-10416.

[23] C. Bian and C. Qian. 2022. Better running time of the non-dominated sorting genetic algorithm II (NSGA-II) by using stochastic tournament selection. In *Proceedings of 17th International Conference Parallel Problem Solving from Nature* (PPSN 2022) Part II, Springer, Dortmund, Germany, 428-442.

[24] B. Doerr and Z. Qu, Runtime Analysis for the NSGA-II: Provable Speed-ups from Crossover. In *Proceedings of the 37th AAAI Conference on Artificial Intelligence* (AAAI 2023), AAAI, Washington, DC, USA, 12399-12407.

[25] B. Doerr and Z. Qu. 2023. From understanding the population dynamics of the NSGA-II to the first proven lower bounds. In *Proceedings of the 37th AAAI Conference on Artificial Intelligence* (AAAI 2023), AAAI, Washington, DC, USA, 12408-12416.

[26] B. Doerr and Z. Qu. 2023. A first runtime analysis of the NSGA-II on a multimodal problem. *IEEE Trans. on Evolutionary Computation* 27, 5 (2023) 1288-1297.

[27] B. Doerr, T. Ivan, and M. S. Krejca. 2025. Speeding up the NSGA-II with a simple tie-breaking rule. In *Proceedings of the 39th AAAI Conference on Artificial Intelligence* (AAAI 2025), AAAI, Philadelphia, USA, 26964-26972.

[28] M. Emmerich, N. Beume, and B. Naujoks. 2005. An EMO algorithm using the hypervolume measure as selection criterion. In *Proceedings of the Third International Conference on Evolutionary Multi-Criterion Optimization* (EMO 2005), Springer, Guanajuato, Mexico, 62-76.

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

[30] R. Tanabe and H. Ishibuchi. 2020. An analysis of quality indicators using approximated optimal distributions in a 3-D objective space. *IEEE Trans. on Evolutionary Computation* 24, 5 (2020) 853-867. [

[31] K. Deb, L. Thiele, M. Laumanns, and E. Zitzler. 2002. Scalable multi-objective optimization test problems. In *Proceedings of 2002 Congress on Evolutionary Computation* (CEC 2002), IEEE, Honolulu, USA, 825-830.

[32] H. Ishibuchi, R. Imada, N. Masuyama, and Y. Nojima. 2019. Comparison of hypervolume, IGD and IGD+ from the viewpoint of optimal distributions of solutions. In *Proceeding of the 10th International Conference on Evolutionary Multi-Criterion Optimization* (EMO 2019), East Lansing, USA, 332-345.

[33] Y. Tian, R. Cheng, X. Zhang, and Y. Jin. 2017. PlatEMO: A MATLAB platform for evolutionary multi-objective optimization. *IEEE Computational Intelligence Magazine* 12, 4 (2017), 73-87.