Sampling Reference Points on the Pareto Fronts of Benchmark Multi-Objective Optimization Problems

Ye Tian[1], Xiaoshu Xiang[1], Xingyi Zhang[1], Ran Cheng[2], and Yaochu Jin[3]
[1]Institute of Bio-inspired Intelligence and Mining Knowledge, School of Computer Science and Technology, Anhui University, Hefei 230039, China *(Corresponding author: Xingyi Zhang)*
[2]School of Computer Science, University of Birmingham, Birmingham, B15 2TT, United Kingdom
[3]Department of Computer Science, University of Surrey, Guildford, Surrey, GU2 7XH, United Kingdom
Email: {field910921@gmail.com; xxs1394@163.com; xyzhanghust@gmail.com; ranchengcn@gmail.com; yaochu.jin@surrey.ac.uk}

***Abstract*—The effectiveness of evolutionary algorithms have been verified on multi-objective optimization, and a large number of multi-objective evolutionary algorithms have been proposed during the last two decades. To quantitatively compare the performance of different algorithms, a set of uniformly distributed reference points sampled on the Pareto fronts of benchmark problems are needed in the calculation of many performance metrics. However, not much work has been done to investigate the method for sampling reference points on Pareto fronts, even though it is not an easy task for many Pareto fronts with irregular shapes. More recently, an evolutionary multi-objective optimization platform was proposed by us, called PlatEMO, which can automatically generate reference points on each Pareto front and use them to calculate the performance metric values. In this paper, we report the reference point sampling methods used in PlatEMO for different types of Pareto fronts. Experimental results show that the reference points generated by the proposed sampling methods can evaluate the performance of algorithms more accurately than randomly sampled reference points.**

I. Introduction

Since the vector evaluated genetic algorithm (VEGA) was proposed in 1985 [1], a large number of evolutionary algorithms have been proposed and demonstrated high effectiveness in solving multi-objective optimization problems (MOPs), which are collectively known as multi-objective evolutionary algorithms (MOEAs) [2]. In order to compare the performance of existing MOEAs, some multi-objective test suites with various types of Pareto sets, Pareto fronts, and landscapes have been proposed. For example, ZDT [3] is one of the first test suites which contains six MOPs with two objectives; DTLZ [4] and WFG [5] are the two most popular test suites, which are scalable with respect to both decision variables and objective; LSMOP [6] is the first test suite designed for large-scale multi-objective optimization, and MaF [7] is a recently proposed test suite for many-objective optimization.

Since each solution of an MOP has multiple objective values, it is difficult to evaluate the performance of MOEAs by directly comparing the objective values of solutions. Therefore, some performance metrics have been proposed to quantitatively evaluate the results obtained by different MOEAs, such as inverted generational distance (IGD) [8] and hypervolume (HV) [9]. For some of the performance metrics, a set of uniformly distributed reference points sampled on the Pareto front (PF) is needed as the ground-truth, e.g.,

$$IGD(P,R) = \frac{\sum_{\mathbf{r}\in R} \min_{\mathbf{p}\in P} \|\mathbf{p}-\mathbf{r}\|}{|R|}, \tag{1}$$

where P is the objective values of a solution set, R is the reference point set, and $\|\cdot\|$ is the L_2-norm. In short, IGD calculates an average minimum distance from each point in R to those in P, where a smaller IGD value indicates a better convergence and diversity of P.

However, it is not an easy task to sample a set of uniformly distributed reference points on various PFs, and only little work has been done to investigate the reference point sampling method [10], [11]. To address this issue, this paper systematically presents the reference point sampling methods used in PlatEMO [12], which is an evolutionary multi-objective optimization platform recently proposed by us. PlatEMO can automatically generate a set of uniformly distributed reference points on the PF of each benchmark MOP, and use them to calculate the performance metric values of the solutions obtained by MOEAs.

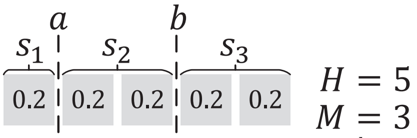
The rest of this paper is organized as follows. Section II reviews three methods for sampling reference points on unit simplex, which is the prerequisite for sampling reference points on most PFs. Section III introduces the reference point sampling methods for several types of PFs. Section IV presents the experimental results of five popular MOEAs on five MOPs, to verify the effectiveness of the proposed reference point sampling methods in performance metric calculation. Conclusions are drawn in Section V.

II. Sampling Reference Points on Unit Simplex

The uniformly distributed reference points on most PFs can be obtained by transforming the points uniformly sampled on unit simplex. Among many others [13], [14], three representative methods for sampling reference points on unit simplex are reviewed in the following.

A. Das and Dennis's Method

The Das and Dennis's method [15] is the most popular systematic approach for sampling uniformly distributed reference points on unit simplex, which is commonly employed



x_1	x_2	$a = x_1$	$b = x_2 - 0.2$	$s_1 = a$	$s_2 = b - a$	$s_3 = 1 - b$
0	0.2	0	0	0	0	1
0	0.4	0	0.2	0	0.2	0.8
0	0.6	0	0.4	0	0.4	0.6
0	0.8	0	0.6	0	0.6	0.4
0	1	0	0.8	0	0.8	0.2
0	1.2	0	1	0	1	0
0.2	0.4	0.2	0.2	0.2	0	0.8
0.2	0.6	0.2	0.4	0.2	0.2	0.6
0.2	0.8	0.2	0.6	0.2	0.4	0.4
0.2	1	0.2	0.8	0.2	0.6	0.2
0.2	1.2	0.2	1	0.2	0.8	0
0.4	0.6	0.4	0.4	0.4	0	0.6
0.4	0.8	0.4	0.6	0.4	0.2	0.4
0.4	1	0.4	0.8	0.4	0.4	0.2
0.4	1.2	0.4	1	0.4	0.6	0
0.6	0.8	0.6	0.6	0.6	0	0.4
0.6	1	0.6	0.8	0.6	0.2	0.2
0.6	1.2	0.6	1	0.6	0.4	0
0.8	1	0.8	0.8	0.8	0	0.2
0.8	1.2	0.8	1	0.8	0.2	0
1	1.2	1	1	1	0	0

Fig. 1. Example of generating reference points by Das and Dennis's method.

by decomposition based MOEAs [16]. A reference point on M-dimensional unit simplex generated by Das and Dennis's method can be defined by $\mathbf{s} = (s_1, s_2, \ldots, s_M)$, where

$$s_j \in \left\{ \frac{0}{H}, \frac{1}{H}, \ldots, \frac{H}{H} \right\}, \ \sum_{j=1}^{M} s_j = 1 \tag{2}$$

and H is the number of divisions in each objective.

Fig. 1 illustrates the procedure for generating all the points satisfying (2) with $M = 3$ and $H = 5$. Specifically, we can find all the combinations of a and b that satisfy $a, b \in \{0, 0.2, \ldots, 1\}$ and $a \leq b$, then let $s_1 = a - 0$, $s_2 = b - a$, and $s_3 = 1 - b$. Therefore, the above issue is equivalent to finding all the 2-multicombinations of $\{0, 0.2, \ldots, 1\}$, which can be further converted to finding all the 2-combinations of $\{0, 0.2, \ldots, 1.2\}$. To summarize, the procedure of Das and Dennis's method is:

1) Let X be all the $(M-1)$-combinations of $\{\frac{0}{H}, \frac{1}{H}, \ldots, \frac{H+M-2}{H}\}$;
2) For each $x_{ij} \in X$ (i.e., the j-th element of the i-th combination in X), $x_{ij} = x_{ij} - \frac{j-1}{H}$;
3) Let S be the reference point set, for each $s_{ij} \in S$ and $x_{ij} \in X$,

$$\begin{cases} s_{ij} = x_{ij} - 0, & j = 1 \\ s_{ij} = x_{ij} - x_{i(j-1)}, & 1 < j < M \\ s_{ij} = 1 - x_{i(j-1)}, & j = M \end{cases} . \tag{3}$$

Fig. 2 plots the reference points generated by Das and Dennis's method with $M = 3, H = 13$ and $M = 10, H = 3$. From Step 1), it is known that the number of reference points generated by Das and Dennis's method is C_{H+M-1}^{M-1}, hence the numbers of reference points shown in Fig. 2 are $C_{13+3-1}^{3-1} = 105$ and $C_{3+10-1}^{10-1} = 220$, respectively.

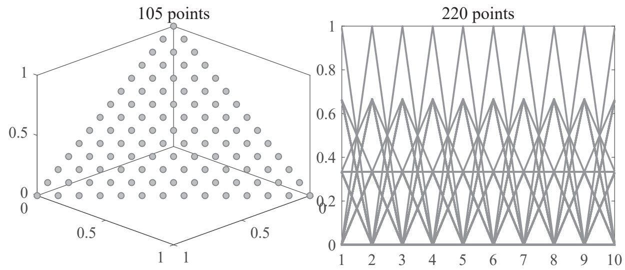


Fig. 2. The reference points sampled by Das and Dennis's method on 3- and 10-objective unit simplexes.

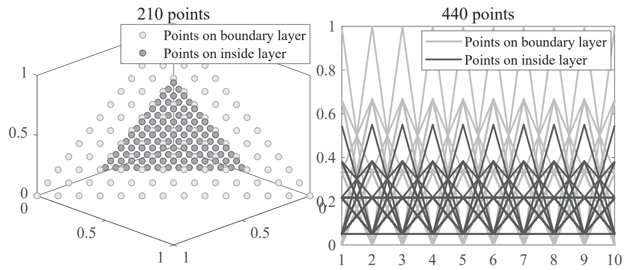


Fig. 3. The reference points sampled by Deb and Jain's method on 3- and 10-objective unit simplexes.

B. Deb and Jain's Method

As reported by Deb and Jain [17], no intermediate point will be generated by Das and Dennis's method as long as $H < M$. Therefore, at least $C_{10+10-1}^{10-1} = 92378$ points on 10-objective unit simplex are required to have at least one intermediate point. To avoid such a situation, they suggested to use two layers of reference points, which are generated as follows:

1) Generate S_1 by Das and Dennis's method as the point set on boundary layer;
2) Let S_2 be the point set on inside layer, for each $s'_{ij} \in S_2$ and $s_{ij} \in S_1$,

$$s'_{ij} = \frac{1}{2} s_{ij} + \frac{1}{2M}; \tag{4}$$

3) The reference point set $S = S_1 \cup S_2$.

Fig. 3 depicts the reference points generated by Deb and Jain's method with $M = 3, H = 13$ and $M = 10, H = 3$, where the numbers of reference points are 210 and 440, respectively. It can be seen from the figure that the reference points on boundary layer are the same to those generated by Das and Dennis's method, while the reference points on inside layer are shrunk to the center of the simplexes. So there can exist intermediate point even if $H < M$. In practice, the Das and Dennis's method is used when $M \leq 5$, and the Deb and Jain's method is used in other cases.

C. Mixture Uniform Design

A drawback of the above two sampling methods is that the number of reference points is restricted by the parameters M and H. By contrast, the mixture uniform design [18] provides a more flexible sampling method, where the number of reference points can be an arbitrary number. In short, it generates N points uniformly distributed in an $(M-1)$-dimensional

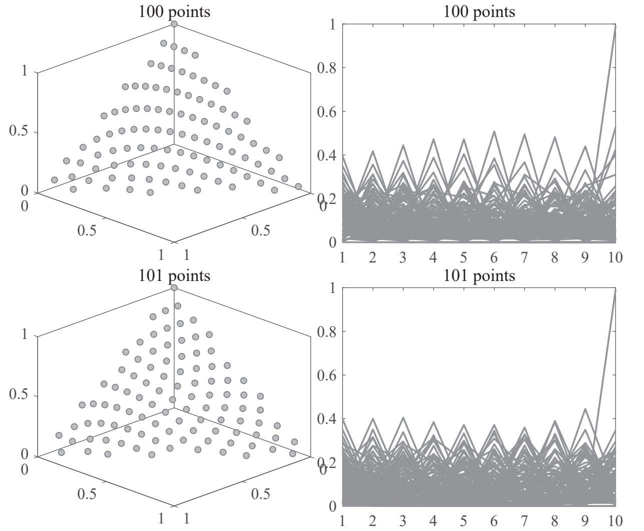


Fig. 4. The reference points sampled by mixture uniform design on 3- and 10-objective unit simplexes.

hypercube, then transforms them to the M-dimensional unit simplex. Specifically, the procedure of mixture uniform design consists of the following four steps:

1) Let W_1 be a row vector of all the positive integers that are coprime to and smaller than N; let $W_2 = (1, 2, \ldots, N)$;
2) Let $W = \mod (W_2^T W_1 - 1, N) + 1$;
3) Let X be a matrix consisting of $M-1$ columns from W, find the X having the largest CD_2 value;[1]
4) Let S be the reference point set, for each $s_{ij} \in S$ and $x_{ij} \in X$,

$$\begin{cases} s_{ij} = \frac{1}{N}[1-(x_{ij})^{\frac{1}{M-j}}]\prod_{k=1}^{j-1}(x_{ik})^{\frac{1}{M-k}}, & 1 \leq j < M \\ s_{ij} = \frac{1}{N}\prod_{k=1}^{j-1}(x_{ik})^{\frac{1}{M-k}}, & j = M \end{cases} . \tag{5}$$

In Step 3), the CD_2 (i.e., centered L_2-discrepancy) is used to measure the diversity of a point set X with N points filling a hypercube, which can be calculated by

$$\begin{aligned} CD_2(X) = \; & [(\tfrac{13}{12})^{M-1} - \tfrac{2^{2-M}}{N}\sum_{i=1}^{N}\prod_{j=1}^{M-1}(2+|x_{ij}-0.5| - \\ & |x_{ij}-0.5|^2) + \tfrac{1}{N^2}\sum_{i,k=1}^{N}\prod_{j=1}^{M-1}(1+0.5|x_{ij} - \\ & 0.5| + 0.5|x_{kj}-0.5| - 0.5|x_{ij}-x_{kj}|)]^{\frac{1}{2}}. \end{aligned} \tag{6}$$

Assume that the matrix W contains K columns, the calculation of CD_2 should be performed for C_K^{M-1} times, hence the computational complexity of mixture uniform design is obviously larger than those of Das and Dennis's method and Deb and Jain's method.

Fig. 4 presents 100 and 101 reference points generated by mixture uniform design with $M = 3$ and $M = 10$. It is obvious from the figure that the number of reference points can be an arbitrary number; however, the obtained reference points distribute less uniformly than those generated by Das and Dennis's method, and many extreme points are missed.

[1] Step 1) – Step 3) is called the good lattice point method, other methods such as the Latin hypercube sampling can also be adopted for generating X.

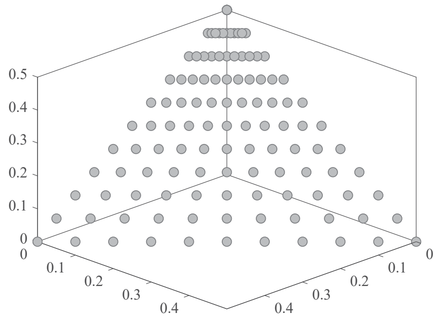


Fig. 5. Reference points for 3-objective DTLZ1, which are generated by uniformly sampling decision variables.

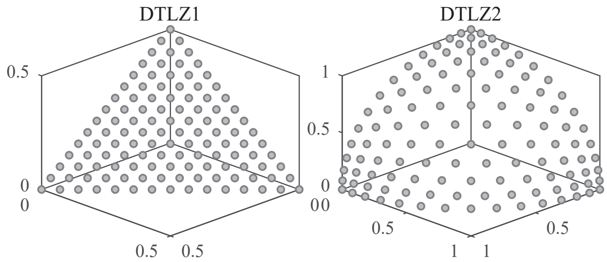


Fig. 6. Reference points sampled on the PFs of 3-objective DTLZ1 and DTLZ2.

III. Sampling Reference Points on Pareto Fronts

Intuitively, the reference points on the PF of an MOP can be obtained by uniformly sampling decision variables. For example, the PF of M-objective DTLZ1 [4] is defined by

$$\begin{cases} f_1 = 0.5x_1x_2\ldots x_{M-1} \\ f_2 = 0.5x_1x_2\ldots(1-x_{M-1}) \\ \ldots \\ f_M = 0.5(1-x_1) \end{cases} , \tag{7}$$

where $x_j \in [0,1]$ is the decision variable and f_j is the objective value. As shown in Fig. 5, if we uniformly generate x_j and calculate the reference points by (7), the obtained reference points will distribute nonuniformly. Therefore, it needs to sample the reference points in objective space directly. As suggested in [10], since the PF of DTLZ1 can be written as

$$f_1 + f_2 + \ldots + f_M = 0.5 \tag{8}$$

with $0 \leq f_j \leq 0.5$, the reference points for DTLZ1 can be obtained by halving the objective values of the points on unit simplex. Similarly, the PF of M-objective DTLZ2 [4] can be written as

$$f_1^2 + f_2^2 + \ldots + f_M^2 = 1 \tag{9}$$

with $0 \leq f_j \leq 1$, and the reference points for DTLZ2 can be obtained by mapping the points from unit simplex to unit hypersphere, i.e., calculating the intersection of the line connecting each point and the origin on unit hypersphere. Fig. 6 depicts 120 reference points sampled on the PFs of 3-objective DTLZ1 and DTLZ2. Note that here the Das and Dennis's method is adopted to generate the uniformly

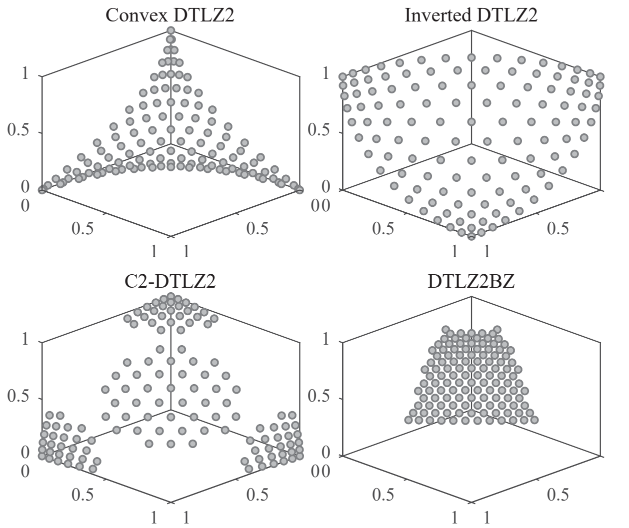


Fig. 7. Reference points sampled on the PFs of 3-objective convex DTLZ2, inverted DTLZ2, C2-DTLZ2, and DTLZ2BZ.

distributed reference points on unit simplex.

However, the above sampling method can only generate reference points on regular and smooth PFs. In this section, we introduce the reference point sampling methods for MOPs with more complex PFs.

A. Sampling Reference Points on Concave and Convex PFs (Variants of DTLZ2)

This subsection introduces the reference point sampling methods for four variants of DTLZ2, namely, convex DTLZ2 [17], inverted DTLZ2 [19], C2-DTLZ2 [19], and DTLZ2BZ [20]. The reference points sampled on the PFs of these four MOPs are plotted in Fig. 7.

The PF of M-objective convex DTLZ2 can be written as

$$f_1^{0.5} + \ldots + f_{M-1}^{0.5} + f_M = 1, \tag{10}$$

where $0 \le f_j \le 1$. The reference points for convex DTLZ2 can be obtained by mapping the points from unit simplex to the convex surface, i.e,

1) Uniformly sample a set of points S on unit simplex;
2) Let R be the reference point set, for each $r_{ij} \in R$ and $s_{ij} \in S$,

$$r_{ij} = \frac{s_{ij}}{t_i}, \tag{11}$$

where t_i can be obtained by solving

$$\left(\frac{s_{i1}}{t_i}\right)^{0.5} + \ldots + \left(\frac{s_{i(M-1)}}{t_i}\right)^{0.5} + \frac{s_{iM}}{t_i} = 1, \tag{12}$$

i.e.,

$$\begin{array}{l} t_i = \frac{\delta_i + 2s_{iM} + \sqrt{\delta_i^2 + 4\delta_i s_{iM}}}{2} \\ \delta_i = \left(\sum_{k=1}^{M-1} \sqrt{s_{ik}}\right)^2 \end{array}. \tag{13}$$

The PF of M-objective inverted DTLZ2 is

$$(1-f_1)^2 + (1-f_2)^2 \ldots + (1-f_M)^2 = 1, \tag{14}$$

where $0 \le f_j \le 1$. The reference points for inverted DTLZ2 can be obtained by inverting the reference points for DTLZ2, i.e.,

1) Generate a set of reference points R for DTLZ2;
2) For each $r_{ij} \in R$, $r_{ij} = 1 - r_{ij}$.

The PF of M-objective C2-DTLZ2 is the same to DTLZ2, except that a constraint is introduced

$$\begin{array}{l} \min\left\{ \min_{j=1}^{M} \left[(f_j-1)^2 + \sum_{k=1,k\neq j}^{M} f_k^2 - a^2\right],\right. \\ \quad \left. \left[\sum_{j=1}^{M}(f_j - 1/\sqrt{M})^2 - a^2\right]\right\} \le 0, \end{array} \tag{15}$$

where $a = 0.4$ for $M = 3$ and 0.5 otherwise. Obviously, the reference points for C2-DTLZ2 can be obtained by sampling reference points on the PF of DTLZ2 and eliminating those do not satisfy the constraint, i.e.,

1) Generate a set of reference points R for DTLZ2;
2) For each point $(r_{i1}, r_{i2}, \ldots, r_{iM})$ in R, eliminate it if (15) is not satisfied, where $f_1, f_2, \ldots, f_M$ are replaced by $r_{i1}, r_{i2}, \ldots, r_{iM}$ in (15).

The PF of M-objective DTLZ2BZ is

$$\begin{cases} f_1 = \cos(\theta_1)\ldots\cos(\theta_{M-2})\cos(\theta_{M-1}) \\ f_2 = \cos(\theta_1)\ldots\cos(\theta_{M-2})\sin(\theta_{M-1}) \\ \ldots \\ f_M = \sin(\theta_1) \end{cases}, \tag{16}$$

where $\theta_j = \frac{\pi}{2}(\frac{x_j}{2} + \frac{1}{4})$ and $0 \le x_j \le 1$. Similar to C2-DTLZ2, the PF of DTLZ2BZ is also a part of the PF of DTLZ2, but the constraint cannot be represented by $f_1, f_2, \ldots, f_M$ like (15). Considering that the value of each θ_j is always between $\frac{1}{8}\pi$ and $\frac{3}{8}\pi$, for each reference point sampled on the PF of DTLZ2, we can calculate the values of $\theta_1, \theta_2, \ldots, \theta_{M-1}$ according to (16), and eliminate the reference point if any $\theta_j < \frac{1}{8}\pi$ or $\theta_j > \frac{3}{8}\pi$. As a result, the procedure of sampling reference points for DTLZ2BZ is

1) Generate a set of reference points R for DTLZ2;
2) For each point $(r_{i1}, r_{i2}, \ldots, r_{iM})$ in R, successively solve the values of $\theta_{M-1}, \theta_{M-2}, \ldots, \theta_1$ by[2]

$$\begin{cases} \frac{\sin(\theta_{M-1})}{\cos(\theta_{M-1})} = \frac{r_{i2}}{r_{i1}} \\ \frac{\sin(\theta_{M-2})}{\cos(\theta_{M-2})} = \frac{r_{i3}}{r_{i2}}\sin(\theta_{M-1}) \\ \ldots \\ \frac{\sin(\theta_1)}{\cos(\theta_1)} = \frac{r_{iM}}{r_{i(M-1)}}\sin(\theta_2) \end{cases}, \tag{17}$$

and eliminate the point if any $\theta_j < \frac{1}{8}\pi$ or $\theta_j > \frac{3}{8}\pi$.

B. Sampling Reference Points on Degenerate PF (DTLZ5)

The PF of M-objective DTLZ5 [4] is

$$\begin{cases} f_j = (\frac{1}{\sqrt{2}})^{M-\max(j,2)}\cos(\frac{\pi}{2}x), & j < M \\ f_j = \sin(\frac{\pi}{2}x), & j = M \end{cases}, \tag{18}$$

where $0 \le x \le 1$. Since the PF of DTLZ5 is always a one-dimensional curve independent of the number of objectives, the reference points for DTLZ5 can be obtained by sampling points on one-dimensional curve and extending them to the other objectives, i.e.,

[2]For simplicity, $r_{ij} = 0$ can be replaced by a tiny value, e.g., 1e-6.

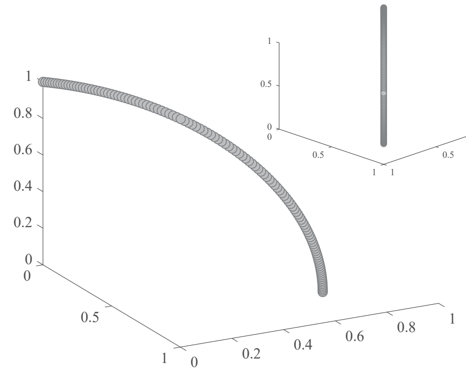


Fig. 8. 120 reference points sampled on the PF of 3-objective DTLZ5.

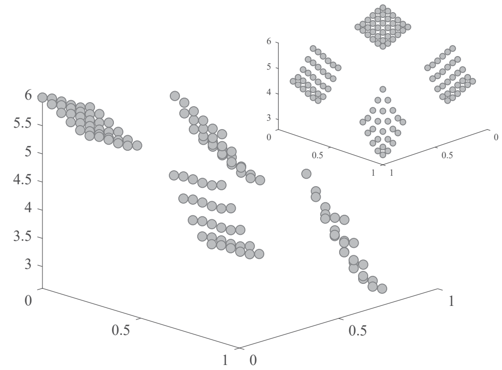


Fig. 9. 121 reference points sampled on the PF of 3-objective DTLZ7.

1) Let X be the set of a number of linearly equally spaced points between 0 and 1;
2) For each $x \in X$, calculate the objective values of the corresponding reference point by (18).

Fig. 8 shows 120 reference points sampled on the PF of 3-objective DTLZ5. It is worth noting that, as pointed out in [11], the PF of DTLZ5 with many objectives and decision variables is not completely a one-dimensional curve, which has an unknown non-degenerate part.

C. Sampling Reference Points on Disconnected PF (DTLZ7)

The PF of M-objective DTLZ7 [4] is:

$$f_M = 2M - \sum_{j=1}^{M-1} f_j\left[1 + \sin(3\pi f_j)\right], \tag{19}$$

where $0 \leq f_j \leq 1$ except for f_M. Note that DTLZ7 does not have a simplex-like PF; by contrast, the first $M-1$ objectives of the PF of DTLZ7 fill an $(M-1)$-dimensional unit hypercube, while the last objective is determined by (19). As a result, the procedure of sampling reference points for DTLZ7 is:

1) Replicate the vector $(\frac{0}{H}, \frac{1}{H}, \ldots, \frac{H}{H})^T$ to produce an $(M-1)$-dimensional point set X, i.e., the number of points in X is $(H+1)^{M-1}$;[3]
2) Let R be the reference point set, for each $r_{ij} \in R$ and $x_{ij} \in X$,
$$\begin{cases} r_{ij} = x_{ij}, & j < M \\ r_{ij} = 2M - \sum_{k=1}^{M-1} x_{ik}\left[1 + \sin(3\pi x_{ik})\right], & j = M \end{cases}; \tag{20}$$
3) Delete the dominated points from R.

Fig. 9 plots 121 reference points sampled on the PF of 3-objective DTLZ7. Since the PF of DTLZ7 is disconnected, there are many dominated points in R. For example, if R contains 1024 points with 3 objectives, only 289 of them will be non-dominated. In fact, according to the PF of 2-objective DTLZ7 shown in Fig. 10, it is obvious that only the points between the origin and a or b and c are non-dominated. According to (19), the first objective values of a, b and c are approximately 0.2514, 0.6316 and 0.8594, respectively,

[3]In order to obtain an arbitrary number of points in X, the good lattice point method or the Latin hypercube sampling can be adopted.

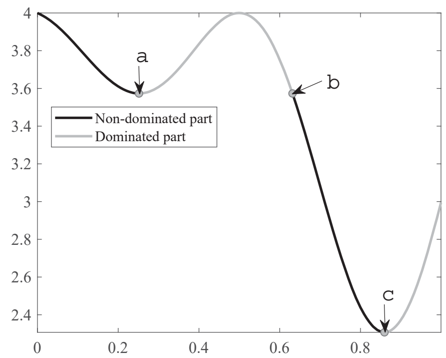


Fig. 10. The PF of 2-objective DTLZ7.

hence the ratio of non-dominated points in R for M-objective DTLZ7 is $(a_1 - b_1 + c_1)^{M-1} \approx 0.48^{M-1}$.

In order to increase the ratio of non-dominated points in R, we can map the elements in each column of X from $[0, 1]$ to $[0, a_1] \cup [b_1, c_1]$ before Step 2), such that all the points in R are non-dominated and Step 3) can be eliminated.

D. Sampling Reference Points on Highly Irregular PF (WFG2)

The PF of WFG2 [5] is defined by

$$\begin{cases} f_1 = 2(1-\cos(\frac{\pi}{2}x_1))(1-\cos(\frac{\pi}{2}x_2))\ldots(1-\cos(\frac{\pi}{2}x_{M-1})) \\ f_2 = 4(1-\cos(\frac{\pi}{2}x_1))(1-\cos(\frac{\pi}{2}x_2))\ldots(1-\sin(\frac{\pi}{2}x_{M-1})) \\ \ldots \\ f_{M-1} = 2(M-1)(1-\cos(\frac{\pi}{2}x_1))(1-\sin(\frac{\pi}{2}x_2)) \\ f_M = 2M(1-x_1\cos^2(5\pi x_1)) \end{cases}, \tag{21}$$

where $0 \leq x_j \leq 1$. Although the PF of WFG2 is simple-like, it cannot be defined as an equation represented by $f_1, f_2, \ldots, f_M$, hence the reference points cannot be mapped from unit simplex to the PF of WFG2 by the method used in convex DTLZ2. On the contrary, the mapping of reference points can be achieved by solving x_j, i.e.,

1) Uniformly sample a set of points S on unit simplex;
2) For each point $(s_{i1}, s_{i2}, \ldots, s_{iM})$ in S, successively

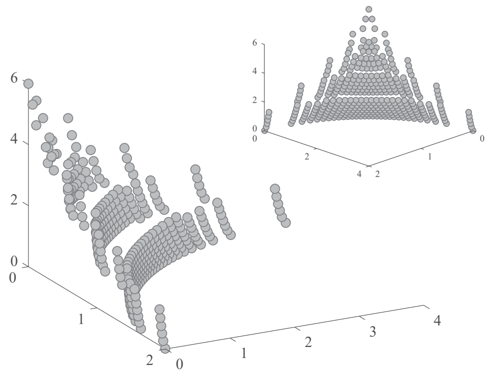


Fig. 11. 381 reference points sampled on the PF of 3-objective WFG2.

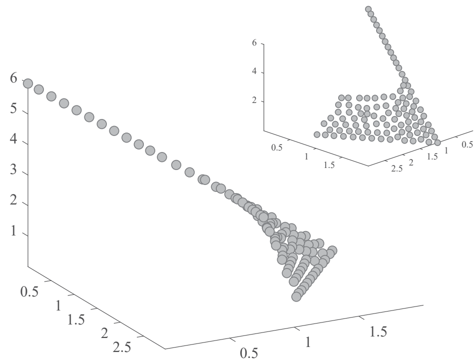


Fig. 13. 100 reference points sampled on the PF of 3-objective WFG3.

solve the values of $x_{M-1}, x_{M-2}, \ldots, x_1$ by[4]

$$\begin{cases} \frac{1-\sin(\frac{\pi}{2}x_{M-1})}{1-\cos(\frac{\pi}{2}x_{M-1})} = \frac{s_{i2}}{s_{i1}} \\ \frac{1-\sin(\frac{\pi}{2}x_{M-2})}{1-\cos(\frac{\pi}{2}x_{M-2})} = \frac{s_{i3}}{s_{i2}}(1-\sin(\frac{\pi}{2}x_{M-1})) \\ \ldots \\ \frac{1-x_1\cos^2(5\pi x_1)}{1-\cos(\frac{\pi}{2}x_1)} = \frac{s_{iM}}{s_{i(M-1)}}(1-\sin(\frac{\pi}{2}x_2)) \end{cases}, \quad (22)$$

and calculate the objective values of the corresponding reference point by (21).

3) Delete the dominated points from the reference point set.

Fig. 11 depicts 381 reference points sampled on the PF of 3-objective WFG2.

E. Sampling Reference Points for MOPs Whose PFs are Unknown

So far the PFs of some benchmark MOPs are still unknown (e.g., WFG3 [5] and most combinatorial MOPs [21]), hence it is impossible to obtain reference points on their PFs. In order to calculate the performance metric values of the solutions obtained on these MOPs, an alternative sampling method is to combine all the non-dominated solutions obtained in multiple runs, and truncate the solutions by the truncation method used in SPEA2 [22].

For example, Fig. 13 plots 100 reference points sampled for 3-objective WFG3. It is worth noting that although the reference points generated by this method may not be on the true PF, it has been evidenced that they are capable of distinguishing the performance of different MOEAs [11].

F. Discussions

The above reference point sampling methods are applicable to most existing multi-objective test suites. To summarize, Table I lists the reference point sampling methods for all the MOPs in three popular scalable test suites, i.e., DTLZ [4], WFG [5], and MaF [7].

It is noteworthy that the reference points obtained by the proposed sampling methods are not absolutely uniform. For example in Fig. 7, the reference points in the middle of convex/concave PFs are more/less crowded than those on the border. This is because the distribution of the reference points generated by Das and Dennis's method is slightly changed after being mapped from unit simplex to a convex or concave surface. To address this issue, some alternative methods [13], [23] can be adopted to replace the Das and Dennis's method. Nevertheless, these methods are less practical than Das and Dennis's method due to their low flexibility and efficiency.

[4]According to the definition of Pareto front, in the case that x_1 has multiple solutions, the one minimizing f_M is chosen.

TABLE I
THE REFERENCE POINT SAMPLING METHODS FOR ALL THE MOPS IN DTLZ, WFG, AND MAF.

Problem	PF shape	Sampling method
DTLZ1, C1-DTLZ1, MaF14	Linear	Same to DTLZ1
DTLZ2–DTLZ4, C3-DTLZ4, WFG4–WFG9, MaF5, MaF12, MaF13	Concave	Same to DTLZ2
DTLZ5, DTLZ6, MaF6	Degenerate	Same to DTLZ5
DTLZ7, MaF7	Disconnected	Same to DTLZ7
Convex DTLZ2, MaF3	Convex	Same to convex DTLZ2
Inverted DTLZ1, inverted DTLZ2, MaF1, MaF4, MaF15	Inverted	Same to inverted DTLZ2
C2-DTLZ2	Disconnected	Same to C2-DTLZ2
DTLZ2BZ, MaF2	Shrunk	Same to DTLZ2BZ
WFG1, WFG2, MaF10, MaF11	Highly irregular	Same to WFG2
WFG3	Unknown	Same to WFG3
MaF8, MaF9	Unknown	Uniformly sampling decision variables

IV. EMPIRICAL EVALUATIONS

To verify the effectiveness of the proposed reference point sampling methods in performance assessment, five selected MOEAs (i.e, NSGA-II [24], RVEA [16], KnEA [25], NSGA-III [17], and AR-MOEA [26]) are tested on four popular MOPs (i.e., C2-DTLZ2, DTLZ5, DTLZ7, and WFG2), where the results are analyzed by IGD based on the reference points generated by the proposed sampling methods.

For RVEA, the penalty parameter is set to 2 and the frequency of reference point adaption is set to 0.1. For KnEA,

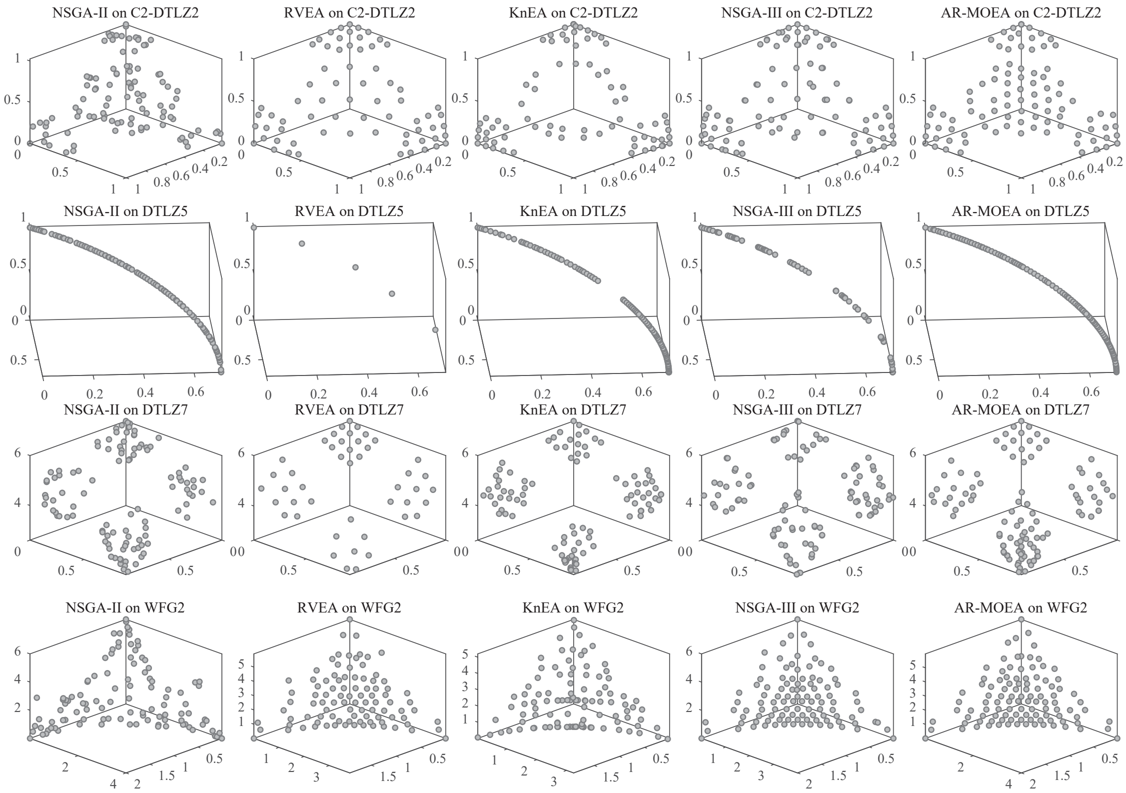


Fig. 12. The non-dominated solution sets with the median IGD among 30 runs obtained by five MOEAs on four 3-objective MOPs.

TABLE II
IGD VALUES BASED ON THE REFERENCE POINTS SAMPLED BY THE PROPOSED METHODS OF FIVE MOEAS ON FOUR 3-OBJECTIVE MOPS.

Problem	NSGA-II	RVEA	KnEA	NSGA-III	AR-MOEA
C2-DTLZ2	5.5011e-2	5.2301e-2	7.9084e-2	4.8558e-2	4.4312e-2
DTLZ5	5.5807e-3	6.9232e-2	1.7279e-2	1.8106e-2	4.2931e-3
DTLZ7	7.4402e-2	1.0457e-1	6.7288e-2	7.0937e-2	6.2788e-2
WFG2	2.1386e-1	1.4977e-1	1.8477e-1	1.6366e-1	1.4923e-1

the rate of knee points is set to 0.5. The population size of all the five MOEAs is set to 105, and the maximum number of generations is set to 500. The simulated binary crossover [27] and polynomial mutation [28] are employed as the genetic operators, where the probabilities of crossover and mutation are set to 1 and $1/D$ (D denotes the number of decision variables), respectively, and the distribution index is set to 20. All the tests are run for 30 times independently.

Table II lists the IGD values obtained by the five MOEAs, where roughly 10000 reference points are generated on each PF by the proposed sampling methods. It can be seen from the statistical results that RVEA, NSGA-III, and AR-MOEA have the best IGD values on WFG2, while AR-MOEA has significantly better IGD values than the others on C2-DTLZ2, DTLZ5, and DTLZ7. Fig. 12 plots the results with the median IGD among 30 runs, from which it is obvious that the IGD results based on the proposed reference point sampling methods are consistent with the visual conclusion.

By contrast, Table III presents the IGD values based on randomly sampled reference points, where some of the results are clearly counter-intuitive. For example, according to Table III, NSGA-III has better IGD value than KnEA on DTLZ5; however, the solutions obtained by NSGA-III distribute obviously less widely than those obtained by KnEA as shown in Fig. 12. Similar inconsistencies can also be found between the IGD values of NSGA-II and RVEA on C2-DTLZ2 and WFG2. Therefore, the randomly sampled reference points are unsuitable for the calculation of performance metrics.

TABLE III
IGD VALUES BASED ON RANDOMLY SAMPLED REFERENCE POINTS OF FIVE MOEAS ON FOUR 3-OBJECTIVE MOPS. THE RESULTS INCONSISTENT WITH FIG. 12 IN EACH ROW ARE FRAMED.

Problem	NSGA-II	RVEA	KnEA	NSGA-III	AR-MOEA
C2-DTLZ2	$\boxed{\text{5.0972e-2}}$	$\boxed{\text{5.1628e-2}}$	8.5015e-2	4.9519e-2	4.3762e-2
DTLZ5	5.4692e-3	6.8352e-2	$\boxed{\text{2.0950e-2}}$	$\boxed{\text{1.9120e-2}}$	4.2935e-3
DTLZ7	7.3551e-2	1.0276e-1	6.7255e-2	7.1555e-2	6.2552e-2
WFG2	$\boxed{\text{1.6601e-1}}$	$\boxed{\text{1.8192e-1}}$	2.6463e-1	1.9448e-1	1.9260e-1

To further compare the qualities of the reference points sampled by the proposed methods and the randomly sampled reference points, Table IV lists the HV [9] and Spacing [29] values of all the reference point sets, where a larger HV value indicates a better convergence and diversity of the point set, and a smaller Spacing value indicates a better evenness of the point set. It is clear from Table IV that the reference points sampled by the proposed methods are better than the randomly sampled reference points in terms of both HV and Spacing.

TABLE IV
HV AND SPACING VALUES OF THE REFERENCE POINTS SAMPLED BY THE PROPOSED METHODS AND THE RANDOMLY SAMPLED REFERENCE POINTS. THE BEST RESULTS IN EACH ROW ARE HIGHLIGHTED.

Problem	Proposed methods		Random sampling	
	HV	Spacing	HV	Spacing
C2-DTLZ2	5.4242e-1	5.8366e-3	5.4093e-1	5.8396e-3
DTLZ5	2.0285e-1	7.4365e-5	2.0210e-1	1.3007e-4
DTLZ7	2.9292e-1	3.3379e-3	2.9234e-1	1.2402e-2
WFG2	9.4721e-1	1.4195e-2	9.4610e-1	5.6458e-2

V. CONCLUSIONS

Sampling uniformly distributed reference points on PFs of benchmark MOPs is an important but rarely touched issue. In order to fill this gap, this paper first details three popular methods for sampling reference points on unit simplex, then introduces several methods for sampling reference points on different types of PFs. Experimental results demonstrate that the reference points generated by the proposed sampling methods can evaluate the performance of MOEAs accurately.

As mentioned before, the proposed reference point sampling methods can cover most of the existing multi-objective test suites. This is because existing benchmark MOPs contain only a few types of PFs, where most of these PFs are very similar to those presented in this paper. Therefore, in order to better evaluate the comprehensive performance of MOEAs, it is desirable to design new MOPs with more complex PFs to pose stiffer challenges to MOEAs in the future.

ACKNOWLEDGMENT

This work was supported in part by the National Natural Science Foundation of China (Grant No. 61672033, 61502004, and 61502001). The work of Y. Jin was supported in part by the U.K. EPSRC Under Grant EP/M017869/1.

REFERENCES

[1] J. D. Schaffer, "Multiple objective optimization with vector evaluated genetic algorithms," in *Proceedings of the 1st International Conference on Genetic Algorithms*, 1985, pp. 93–100.

[2] A. Zhou, B.-Y. Qu, H. Li, S.-Z. Zhao, P. N. Suganthan, and Q. Zhang, "Multiobjective evolutionary algorithms: A survey of the state of the art," *Swarm and Evolutionary Computation*, vol. 1, no. 1, pp. 32–49, 2011.

[3] E. Zitzler, K. Deb, and L. Thiele, "Comparison of multiobjective evolutionary algorithms: Empirical results," *Evolutionary Computation*, vol. 8, no. 2, pp. 173–195, 2000.

[4] K. Deb, L. Thiele, M. Laumanns, and E. Zitzler, "Scalable test problems for evolutionary multiobjective optimization," in *Evolutionary Multiobjective Optimization. Theoretical Advances and Applications*, 2005, pp. 105–145.

[5] L. B. S. Huband, P. Hingston and L. While, "A review of multiobjective test problems and a scalable test problem toolkit," *IEEE Transactions on Evolutionary Computation*, vol. 10, no. 5, pp. 477–506, 2006.

[6] R. Cheng, Y. Jin, M. Olhofer, and B. Sendhoff, "Test problems for large-scale multiobjective and many-objective optimization," *IEEE Transactions on Cybernetics*, vol. 47, no. 12, pp. 4108–4121, 2017.

[7] R. Cheng, M. Li, Y. Tian, X. Zhang, S. Yang, Y. Jin, and X. Yao, "A benchmark test suite for evolutionary many-objective optimization," *Complex & Intelligent Systems*, vol. 3, no. 1, pp. 67–81, 2017.

[8] A. Zhou, Y. Jin, Q. Zhang, B. Sendhoff, and E. Tsang, "Combining model-based and genetics-based offspring generation for multi-objective optimization using a convergence criterion," in *Proceedings of the 2006 IEEE Congress on Evolutionary Computation*, 2006, pp. 892–899.

[9] L. While, P. Hingston, L. Barone, and S. Huband, "A faster algorithm for calculating hypervolume," *IEEE Transactions on Evolutionary Computation*, vol. 10, no. 1, pp. 29–38, 2006.

[10] K. Li, K. Deb, Q. Zhang, and S. Kwong, "Combining dominance and decomposition in evolutionary many-objective optimization," *IEEE Transactions on Evolutionary Computation*, vol. 19, no. 5, pp. 694–716, 2015.

[11] H. Ishibuchi, H. Masuda, and Y. Nojima, "Pareto fronts of many-objective degenerate test problems," *IEEE Transactions on Evolutionary Computation*, vol. 20, no. 5, pp. 807–813, 2016.

[12] Y. Tian, R. Cheng, X. Zhang, and Y. Jin, "PlatEMO: A MATLAB platform for evolutionary multi-objective optimization," *IEEE Computational Intelligence Magazine*, vol. 12, no. 4, pp. 73–87, 2017.

[13] C. He, L. Pan, H. Xu, Y. Tian, and X. Zhang, "An improved reference point sampling method on Pareto optimal front," in *Proceedings of the 2016 IEEE Congress on Evolutionary Computation*, 2016, pp. 5230–5237.

[14] S. Jiang and S. Yang, "A strength Pareto evolutionary algorithm based on reference direction for multi-objective and many-objective optimization," *IEEE Transactions on Evolutionary Computation*, vol. 21, no. 3, pp. 329–346, 2017.

[15] I. Das and J. E. Dennis, "Normal-boundary intersection: A new method for generating the Pareto surface in nonlinear multicriteria optimization problems," *SIAM Journal on Optimization*, vol. 8, no. 3, pp. 631–657, 1998.

[16] R. Cheng, Y. Jin, M. Olhofer, and B. Sendhoff, "A reference vector guided evolutionary algorithm for many-objective optimization," *IEEE Transactions on Evolutionary Computation*, vol. 20, no. 5, pp. 773–791, 2016.

[17] K. Deb and H. Jain, "An evolutionary many-objective optimization algorithm using reference-point based non-dominated sorting approach, part I: Solving problems with box constraints," *IEEE Transactions on Evolutionary Computation*, vol. 18, no. 4, pp. 577–601, 2014.

[18] K. Fang and C. Ma, *Orthogonal and uniform experimental design*. Science and Technology Press, Beijing, 2001.

[19] H. Jain and K. Deb, "An evolutionary many-objective optimization algorithm using reference-point based nondominated sorting approach, part II: Handling constraints and extending to an adaptive approach," *IEEE Transactions on Evolutionary Computation*, vol. 18, no. 4, pp. 602–622, 2014.

[20] D. Brockhoff and E. Zitzler, "Objective reduction in evolutionary multiobjective optimization: Theory and applications," *Evolutionary Computation*, vol. 17, no. 2, pp. 135–166, 2009.

[21] X. Cai, Y. Li, Z. Fan, and Q. Zhang, "An external archive guided multi-objective evolutionary algorithm based on decomposition for combinatorial optimization," *IEEE Transactions on Evolutionary Computation*, vol. 19, no. 4, pp. 508–523, 2015.

[22] E. Zitzler, M. Laumanns, and L. Thiele, "SPEA2: Improving the strength Pareto evolutionary algorithm for multiobjective optimization," in *Proceedings of the Fifth Conference on Evolutionary Methods for Design, Optimization and Control with Applications to Industrial Problems*, 2001, pp. 95–100.

[23] I. Giagkiozis, R. C. Purshouse, and P. J. Fleming, "Generalized decomposition and cross entropy methods for many-objective optimization," *Information Sciences*, vol. 282, pp. 363–387, 2014.

[24] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A fast and elitist multi-objective genetic algorithm: NSGA-II," *IEEE Transactions on Evolutionary Computation*, vol. 6, no. 2, pp. 182–197, 2002.

[25] X. Zhang, Y. Tian, and Y. Jin, "A knee point driven evolutionary algorithm for many-objective optimization," *IEEE Transactions on Evolutionary Computation*, vol. 19, no. 6, pp. 761–776, 2015.

[26] Y. Tian, R. Cheng, X. Zhang, F. Cheng, and Y. Jin, "An indicator based multi-objective evolutionary algorithm with reference point adaptation for better versatility," *IEEE Transactions on Evolutionary Computation*, 2017, in press.

[27] K. Deb, *Multi-Objective Optimization Using Evolutionary Algorithms*. New York: Wiley, 2001.

[28] K. Deb and M. Goyal, "A combined genetic adaptive search (GeneAS) for engineering design," *Computer Science and Informatics*, vol. 26, no. 4, pp. 30–45, 1996.

[29] J. R. Schott, "Fault tolerant design using single and multicriteria genetic algorithm optimization," Master's thesis, Cambridge: Massachusetts Institute of Technology, 1995.