

# A New Many-Objective Evolutionary Algorithm Based on Generalized Pareto Dominance

Shuwei Zhu<sup>1</sup>, Graduate Student Member, IEEE, Lihong Xu, Senior Member, IEEE, Erik D. Goodman<sup>2</sup>, and Zhichao Lu, Graduate Student Member, IEEE

**Abstract**—In the past several years, it has become apparent that the effectiveness of Pareto-dominance-based multiobjective evolutionary algorithms deteriorates progressively as the number of objectives in the problem, given by  $M$ , grows. This is mainly due to the poor discriminability of Pareto optimality in many-objective spaces (typically  $M \geq 4$ ). As a consequence, research efforts have been driven in the general direction of developing solution ranking methods that do not rely on Pareto dominance (e.g., decomposition-based techniques), which can provide sufficient selection pressure. However, it is still a nontrivial issue for many existing non-Pareto-dominance-based evolutionary algorithms to deal with unknown irregular Pareto front shapes. In this article, a new many-objective evolutionary algorithm based on the generalization of Pareto optimality (GPO) is proposed, which is simple, yet effective, in addressing many-objective optimization problems. The proposed algorithm used an “ $(M - 1) + 1$ ” framework of GPO dominance,  $(M - 1)$ -GPD for short, to rank solutions in the environmental selection step, in order to promote convergence and diversity simultaneously. To be specific, we apply  $M$  symmetrical cases of  $(M - 1)$ -GPD, where each enhances the selection pressure of  $M - 1$  objectives by expanding the dominance area of solutions, while remaining unchanged for the one objective left out of that process. Experiments demonstrate that the proposed algorithm is very competitive with the state-of-the-art methods to which it is compared, on a variety of scalable benchmark problems. Moreover, experiments on three real-world problems have verified that the proposed algorithm can outperform the others on each of these problems.

**Index Terms**—Evolutionary algorithms, generalized Pareto optimality, many-objective optimization, Pareto dominance.

Manuscript received May 17, 2020; revised September 21, 2020 and November 29, 2020; accepted January 3, 2021. This work was supported in part by the Natural Science Foundation of China under Grant 61973337; in part by the U.S. National Science Foundation’s BEACON Center under Cooperative Agreement DBI-0939454; and in part by the China Scholarship Council. This article was recommended by Associate Editor K.-C. Tan. (Corresponding author: Lihong Xu.)

Shuwei Zhu and Lihong Xu are with the Department of Electronics and Information Engineering, Tongji University, Shanghai 201804, China (e-mail: zswjiang@163.com; xulhk@163.com).

Erik D. Goodman is with the BEACON Center, Michigan State University, East Lansing, MI 48824 USA (e-mail: goodman@egr.msu.edu).

Zhichao Lu is with the Department of Computer Science and Engineering, Southern University of Science and Technology, Shenzhen 518055, China (e-mail: luzc@sustech.edu.cn).

This article has supplementary material provided by the authors and color versions of one or more figures available at <https://doi.org/10.1109/TCYB.2021.3051078>.

Digital Object Identifier 10.1109/TCYB.2021.3051078

## I. INTRODUCTION

MANY-OBJECTIVE optimization problems (MaOPs) refer to problems requiring optimization of more than three conflicting objectives and have recently gained wide interest in the evolutionary multiobjective optimization (EMO) community. As one of the best-known EMO methods based on the Pareto-dominance selection principle, NSGA-II [1] has gained huge success in addressing various kinds of multiobjective optimization problems (MOPs), including extensive application cases [2], [3]. The concept of Pareto dominance, an intuitive yet qualitative notion of compromise, has been commonly adopted to distinguish the quality of solutions for traditional 2-D or 3-D MOPs. However, the effectiveness of Pareto-based multiobjective evolutionary algorithms (MOEAs) degrades drastically when solving MaOPs. The main challenge of these approaches lies in the loss of Pareto-based selection pressure toward the true Pareto front (PF) as the objective number  $M$  grows [4], that is, solutions become incomparable due to dominance resistance, as well as the difficulty of balancing convergence against diversity [5].

To tackle the scalability problem of Pareto-based MOEAs, several techniques have been developed in recent years, which can be generally divided into two main groups. The first group modifies the Pareto-dominance criterion, also referred to as a relaxed Pareto-dominance relation, such as  $\epsilon$ -dominance and Cone  $\epsilon$ -dominance [6], CDAS [7], CN $\alpha$ -dominance [8], and generalization of Pareto optimality (GPO) [9], to better discriminate among solutions and select elite ones with enhanced selection pressure. The other group modifies the diversity maintenance operations, for example, shift-based density estimation (SDE) [10] and the use of associated reference points in NSGA-III [5] which was inspired by the earlier decomposition-based techniques. However, most of them, especially the former, have difficulty maintaining the delicate balance between convergence and diversity. Excessive selection pressure usually tends to cause diversity maintenance to deteriorate, for example, it may lead the population to converge into a subregion (or several small subregions) of the PF; while excessive diversity pressure may result in degraded convergence performance. Moreover, extra parameters are required that should be well tuned to control the dominance area for relaxed Pareto-dominance relations, which is not a trivial issue. Apart from the above, some non-Pareto-based dominance relations, for example,  $\theta$ -dominance [11], D-dominance [12],

and strengthened dominance relation (SDR) [13], have shown promising potential to address MaOPs.

Throughout the existing many-objective evolutionary algorithms (MaOEAs), a Pareto-dominance-based method (without using reference points) that is competitive in providing desirable performance across many different kinds of problems has not yet been reported. Our recent proposed GPO dominance criterion [9] can guarantee the identity of expanding the dominance area to improve selection pressure, but only at the cost of introducing some difficulty in diversity maintenance. In view of this, a strategy for distributed evolution was developed by evolving the PF in parallel with multiple processors, namely, asymmetric GPO (AGPO), such that each processor is responsible for a subregion of the entire PF [9], [14]. However, the difficulty of controlling the expanding angle  $\varphi$  has still not been well solved, and the distributed evolution with several relatively large subpopulations can lead to much extra computational burden. Hence, the AGPO technique is still not practical for addressing general MaOPs. Notice that GPO can provide nondiscriminatory expansion of the dominance envelope (i.e., the boundaries of the dominance area) for any solution regardless of the  $M$  value and PF shape [9], which is, to the best of our knowledge, not achieved by other modified Pareto-dominance methods, for example, [6]–[8]. With the aid of these generalizations, MOEAs could acquire the flexibility of improving their selection pressures within certain ranges in dealing with MaOPs. However, there still exist some restrictions (e.g., the selection of the expected dominating ratio level and the loss of diversity) that affect the applicability of GPO, which have not been addressed to this point.

In recent literature, research efforts on MaOPs have been mainly driven to develop solution ranking methods that do not rely on Pareto dominance, which can provide sufficient selection pressure. However, it is nontrivial for many existing non-Pareto-dominance-based MaOEAs to deal with unknown irregular PF shapes [15], [16]. Recently, Ishibuchi *et al.* [17] demonstrated that MaOPs are not always difficult for Pareto-dominance-based MOEAs. Moreover, they have illustrated that one reason for the poor performance of NSGA-II in the DTLZ-type test problems with triangular PFs is the existence of dominance-resistant solutions. Compared to decomposition-based methods (even some IGD-type-indicator-based methods), the performance of Pareto-dominance-based MOEAs is less related to the PF shapes. That is to say, Pareto dominance seems to have a more promising potential to address irregular many-objective realistic test problems and real-world problems, for example, as witnessed in addressing knapsack problems with highly correlated objectives [18].

To fully investigate the potential of GPO by addressing its aforementioned limitations in MaOPs, we propose a new version of GPO as the solution ranking method. Convergence capability can be guaranteed by using an “ $(M-1)+1$ ” framework of GPO dominance,  $(M-1)$ -GPD for short. Meanwhile, the population diversity is preserved by balancing  $M$  symmetrical processors of  $(M-1)$ -GPD in a single run each generation, together with adopting a simple, yet effective, distance-based subset selection strategy. In each case, all sets of  $M-1$  objectives enhance (via GPO) the selection pressure

to the same degree, and the one objective left out remains unchanged. To summarize, the main contributions of this article are as follows.

- 1) A new Pareto-dominance relation, that is,  $(M-1)$ -GPD, that is, reference-vector free, has been developed as the selection criterion, which can maintain a good balance of convergence and diversity. It can also be easily embedded into other existing MaOEAs, for example, NSGA-III.
- 2) The proposed  $(M-1)$ -GPD scheme is nearly parameterless and is used in a novel MaOEA, that is, multiple  $(M-1)$ -GPD-based optimization, called MultiGPO for short, which shows competitive performance compared with several state-of-the-art MaOEAs.
- 3) The superior performance of MultiGPO has also been confirmed on three real-world problems with irregular PFs. These results verify that pure Pareto-dominance-based MaOEA without predefined reference vectors can still work well on MaOPs, and deserves more attention.

The remainder of this article is organized as follows. Section II presents the background and our detailed motivation. The proposed MultiGPO method is described in Section III. Then, the experimental study on test problems and real-world applications are presented in Sections IV and V, respectively. Finally, Section VI concludes this article.

## II. BACKGROUND

### A. Preliminaries: Basic Definitions

For a minimization MOP, the formal definitions of Pareto dominance and its generalization are provided as follows.

**Definition 1 (Pareto Dominance):** A solution  $\mathbf{u}$  is said to dominate another solution  $\mathbf{v}$ , denoted as  $\mathbf{u} < \mathbf{v}$ , if and only if

$$\begin{cases} \forall i \in \{1, 2, \dots, M\} : f_i(\mathbf{u}) \leq f_i(\mathbf{v}) \\ \exists j \in \{1, 2, \dots, M\} : f_j(\mathbf{u}) < f_j(\mathbf{v}). \end{cases} \quad (1)$$

**Definition 2 (Generalized Pareto Dominance):** A solution  $\mathbf{u}$  is said to generally dominate another solution  $\mathbf{v}$  with respect to (w.r.t.) expanding angle  $\boldsymbol{\varphi} = [\varphi_1, \varphi_2, \dots, \varphi_M]$  (denoted by  $\mathbf{u} <^\varphi \mathbf{v}$ ), if and only if  $f(\mathbf{u})^\varphi$  is partially less than  $f(\mathbf{v})^\varphi$  [9], that is

$$\begin{cases} \forall i \in \{1, 2, \dots, M\} : f_i(\mathbf{u}) + \sum_{k \neq i} \delta_{ik} f_k(\mathbf{u}) \leq f_i(\mathbf{v}) \\ \quad + \sum_{k \neq i} \delta_{ik} f_k(\mathbf{v}) \\ \exists j \in \{1, 2, \dots, M\} : f_j(\mathbf{u}) + \sum_{k \neq j} \delta_{jk} f_k(\mathbf{u}) < f_j(\mathbf{v}) \\ \quad + \sum_{k \neq j} \delta_{jk} f_k(\mathbf{v}) \end{cases} \quad (2)$$

where  $\delta_i = \sqrt{(M-1)} \cdot \tan \varphi_i / (M-1)$ .

As the number of objectives  $M$  increases, most Pareto-optimal individuals are mutually nondominated, resulting in their incomparability. Given an appropriate  $\boldsymbol{\varphi}$ , the discriminability can usually be enhanced if we use  $<^\varphi$  instead.

### B. Existing Strategies for MaOEAs

Due to the deficiency associated with Pareto-domination schemes for solution ranking when solving MaOPs, some general directions of non-Pareto-dominance-based MaOEAs

are currently receiving more consideration within the research community. Among them, the two best-known representatives are decomposition based [12], [19]–[25] and indicator based [26]–[30]. In addition, some newly proposed methods are somewhat related to one of these groups, such as MaOEA/C [31], MDEA [32], PaRP/EA [33], etc.

1) *Decomposition-Based Methods*: During the past decade, the vast majority of work has been devoted to the development of decomposition-based algorithms for MaOPs. Generally, decomposition-based MaOEAs can be divided into two categories. The first is search-directions-based algorithms, which essentially decompose an MaOP into multiple single-objective subproblems (sometimes MOPs) and optimize them simultaneously, such as MOEA/DD [19], g-DBEA [20], MaOEA/D-2ADV [21], MOEA/D-SOM [34], MOEA/D-AM2M [22], PAEA [24], MOEA/AD [23], DrEA [12], DDEA [35], DECLA [25], MP-DEA [16], and so on. The other category is algorithms that select survival solutions with reference-point-based techniques, wherein the objective space is partitioned into a number of subregions, such as NSGA-III [5],  $\theta$ -DEA [11], and RVEA [36], as well as some newly proposed algorithms, for example, MaOEA-IT [37], CLIA [38], and ar-MOEA [39]. A brief summary of the very recent proposed notable MaOEAs, including most of the above methods and those discussed below, is relegated to the supplementary material, in Table S-I.

Recently, Ishibuchi *et al.* [15] have demonstrated that the performance of decomposition-based algorithms strongly depends on the PF shapes. This is because the predefined set of weight directions/reference points [5] plays a primary role in the performance of decomposition-based algorithms, in which reference points/directions are typically uniformly distributed and perform well on problems with regular PFs. However, methods with predefined reference information are likely to fail on problems with irregular PFs. To tackle this problem, some methods have been proposed that use reference point adaptation/learning [27], [38], for example, A-NSGA-III [40], MOEA/D-AWA [41], RVEA\* [36], and most of the above very newly proposed MaOEAs. These approaches seek to adjust the distribution of reference points to be consistent with the PF geometry. Unfortunately, for an unknown PF shape, the updating time and frequency, along with defining how to update the search directions, are both nontrivial tasks [16]. So adaptive updating of reference points/directions is still limited in its ability to cope with irregular PFs.

2) *Indicator-Based Methods*: These methods use performance indicators in seeking both convergence and diversity, for example, inverted generational distance (IGD), hypervolume (HV), R2, and  $I_{\epsilon+}$  indicators are used as the major selection criteria [26]–[28]. For example, MaOEA/IGD [28] employed the IGD indicator and an efficient dominance comparison to select solutions with favorable performance; and an enhanced IGD indicator (i.e., IGD-NS) was developed with a reference point adaptation method in AR-MOEA [27]. In R2HCA-EMOA [29], an R2 indicator variant was used to approximate the HV contribution. In MaOEA-IBP [30], a worst-elimination mechanism based on the  $I_{\epsilon+}$  indicator [42] and boundary protection strategy

was devised. To benefit from the complementary nature of multiple indicators, SRA [26] took advantage of indicators  $I_{\epsilon+}$  [10] and  $I_{SDE}$  [10] simultaneously; and the  $I_{SDE}^+$  [43] indicator was proposed using a combination of the sum of objectives (*sum* for short) and  $I_{SDE}$ .

In addition, the literature also witnessed the use of some uncommon indicators, such as the balanceable fitness estimation (BFE), an indicator proposed in NMPSO [44], and the polar-metric (*p*-metric) indicator in PMEA [45]. GFM-MOEA [46] introduced a generic front model (GFM) to approximate the PF; subsequently, a GFM-based indicator is defined. In both lbylEA [47] and 2REA [48], two indicators were designed as two separate selection criteria; they both defined the convergence indicator in terms of the *sum*, and the difference lies in the distribution indicators. It is worth noticing that the convergence indicator in terms of the sum of objectives (or the sum of distances to the ideal point) is also popular in other existing MaOEAs, for example, MDEA [32], MaOEA/C [31], NSGA-II/SDR [13], ASEA [49], and PaRP/EA [33], all of which can also be regarded as partial indicator-based methods.

However, the indicator-based MaOEAs do encounter their own difficulties. First but not foremost, it is a challenging task for the IGD-Like indicators [27]–[29], [45] to construct a set of uniformly distributed solutions. Second, exact HV calculation requires a growing computational burden as  $M$  increases.

3) *Miscellaneous*: Some other particular strategies have also shown desirable potential when tackling MaOPs. To treat this briefly, we list some of these related approaches in this part (also in Table S-I of the supplementary material), including the above partial indicator-based methods [13], [31], [32], [48], [49] that seek specifically to develop diversity maintenance strategies. The MaOEA/C algorithm [31] takes advantage of two clustering methods to divide the population into  $N$  clusters ( $N$  is the population size) for environmental selection. A tailored niching technique has been proposed in a new SDR [13] to distinguish solutions. In PaRP/EA [33], a global hyperplane was used to distinguish solutions with better convergence, then angular-distance-based selection (the *remove\_one\_by\_one* and *add\_one\_by\_one* procedures) was used for maintaining population diversity.

An alternative to extending existing MOEAs to improve their scalability w.r.t.  $M$  is to perform objective reduction, that is, finding the most relevant objectives and eliminating redundant ones [50]. This is a very straightforward way to simplify the original MaOPs. Recently, Yuan *et al.* [51] developed three EMO approaches to objective reduction through NSGA-II, namely, NSGA-II- $\delta$ , NSGA-II- $\eta$ , and NSGA-II- $\gamma$ . Although most real-world MaOPs contain some amount of redundancy among the objectives, objective reduction techniques have not yet been sufficiently explored to completely resolve the problem, which may partly be due to the lack of enough well-defined test problems with sufficient amounts of redundancy.

### C. Motivation

It is observable in Table S-I in the supplementary material that recent research efforts tackling MaOPs tend to be driven into either decomposition-based or indicator-based

methods—especially the former, since they have shown promising performance. As mentioned before, both of these two types of methods have their issues—especially the choice of reference information, regardless of whether predefined or self-learned. Recently, some very interesting reference-free MaOEAs have been proposed, for example, 1by1EA [47], MaOEA/C [31], MDEA [32], NSGA-II/SDR [13], PaRP/EA [33], and MaOES [52].

As mentioned in Section I, there have been several earlier attempts to improve Pareto-based MaOEAs. Nevertheless, although GPO shows promising potential, there exists almost no further research related to it (and even no much work to modify Pareto dominance) in the most recent literature. This is because the existing modified Pareto-dominance methods [6]–[9] cannot guarantee a general desirable performance in balancing convergence and diversity, as discussed in Section IV-B. Compared to these methods, some non-Pareto-based dominance relations, such as  $\theta$ -dominance [11], D-dominance [12], and SDR [13], have shown more promising potential to address MaOPs. However, they still confront some issues, for example,  $\theta$ -dominance and D-dominance, which belong to decomposition-based methods, depend highly on prior reference information; and SDR cannot work well on problems with low dimensionality (2 or 3 dimensions) or objective redundancy, and is not transitive among solutions [13]. As mentioned in Section I, although GPO has some restrictions, it can provide nondiscriminatory expansion of the dominance envelope for any solution, and has shown its flexibility to change the selection pressure, which inspires us to investigate the further use of GPO for MaOPs.

### III. PROPOSED METHOD

#### A. General Framework

The proposed MultiGPO algorithm starts from the initialization step by randomly generating a population, and subsequently, each individual is evaluated. Then, the evolutionary operators, including the mate selection and reproduction operations (typically crossover and mutation), are executed to generate offspring solutions. Here, we employ binary tournament selection to select promising parents for recombination, and the reproduction operation adopts simulated binary crossover (SBX) and polynomial mutation (PM) [5]. As for the environmental selection step, the  $(M - 1)$ -GPD-based ranking method (Algorithm 2) is used to choose the survival solutions. The evolution procedure is repeated until the termination criterion is met. The overall framework of MultiGPO is presented in Algorithm 1, and the source code is available online.<sup>1</sup>

#### B. Extreme Solutions

Extreme solutions, to which corner points may belong, play a key role in preserving population diversity. Note that extreme solutions have been given priority during environmental selection in some existing MaOEAs, for example, VaEA [53], MDEA [32], DDEA [35], 2REA [48], PaRP/EA [33], and so on.

#### Algorithm 1: General Framework of MultiGPO

---

**Input:** The expanding angle  $\varphi$ , and population size  $N$ ;  
**Output:** final population  $P$ .  
1  $P = \text{Population\_initialization}()$ ;  
2 Evaluate population  $P$ ;  
3 **while** *termination criterion is not fulfilled* **do**  
4      $P' \leftarrow \text{Mating\_Selection}(P)$ ;  
5      $Q \leftarrow \text{Reproduction}(P')$ ;  
6      $R \leftarrow P \cup Q$ ;  
7      $P \leftarrow (M - 1)\text{-GPD}(R, N, \varphi)$ ;  
8 **end**

---

In this study, each of the extreme solutions lies on a hyperplane where only one of the objectives  $f_k$  is minimum. The achievement scalarizing function (ASF) [5], [11], with  $M$  predefined weight vectors, is adopted to find  $m$  ( $m \leq M$ ) extreme solutions. The solution  $\mathbf{x}$  that minimizes the following ASF is then selected along each direction:

$$\text{ASF}(\mathbf{x}, \mathbf{w}_i) = \max_{k=1:M} \left\{ \frac{|f_k(\mathbf{x}) - z_k^*|}{w_k^i} \right\} \quad (3)$$

where  $\mathbf{w}_i$ ,  $i = 1, 2, \dots, M$ , is the axis direction of  $f_i$ , which satisfies that if  $i \neq j$  then  $w_{i,j} = 10^{-6}$ ; otherwise,  $w_{i,i} = 1$ . Also,  $\mathbf{z}^* = (z_1^*, \dots, z_M^*)$  is the ideal point. Note that sometimes one solution may obtain the smallest ASF value along several directions, such that only  $m$  ( $m < M$ ) solutions are selected.

#### C. Distance-Based Subset Selection

Distance-based subset selection can select uniformly distributed solutions from an archive for benchmarking MOPs/MaOPs [54], in which the procedure of iteratively selecting subsequent points that are most isolated from the current selected set is conducted until a prespecified set size  $K$  is achieved. It can be regarded as a kind of one-by-one selection method, namely: each time, identify the solution with the maximum value of the minimum pairwise distance between unselected and selected sets. For convenience, we term this strategy the max-min distance-based selection. It is noteworthy that some of the aforementioned reference-free MaOEAs, that is, 1by1EA [47], MDEA [32], 2REA [48], and PaRP/EA [33], have adopted similar max-min distance strategies during environmental selection in order to preserve population diversity.

Since solutions are distributed sparsely in a high-dimensional objective space [55], it is problematic to use the Euclidean distance (as it was used in [54]) for dissimilarity calculation of solutions in solving MaOPs. Instead, the usage of some other dissimilarity metrics is prevalent in the literature, such as the angular distance/cosine distance [13], [22], [31], [33], [36], [47] and the  $L_p$ -norm distance [32], [48], [55]. It is worth noting that the setting of  $p$  is not a nontrivial issue for the  $L_p$ -norm distance, which has been estimated based on the PF geometric shape in some previous research. For example,  $p < 1$  was suggested and used in [55]; while in the MDEA [32] and 2REA [48] algorithms, two different  $p$ -estimation strategies were designed for the  $L_p$ -norm distance, respectively. Besides the above two distance metrics, some other specific dissimilarity metrics have been proposed,

<sup>1</sup><https://github.com/ShuweiZhu/MultiGPO>

such as the hyperplane-assisted Euclidean distance [35], GFM-assisted angular distance [46], and the maximum extension distance (MED) [52], which have shown promising benefits.

For MaOPs, the direction of vectors is likely to be more important than their lengths in terms of diversity. The cosine distance is a sensible choice to measure dissimilarity between two solution vectors, which makes it simple, yet effective, in addressing MaOPs. Hence, we adopt it to compute the max-min distance in our framework. Given two  $d$ -dimensional vectors  $\mathbf{x}$  and  $\mathbf{y}$ , the cosine distance is

$$d_{\cos}(\mathbf{x}, \mathbf{y}) = 1 - \frac{\mathbf{x} \cdot \mathbf{y}}{\|\mathbf{x}\|_2 \times \|\mathbf{y}\|_2}. \quad (4)$$

#### D. $(M-1)$ -GPD for Environment Selection

The GPO criterion is mainly inspired by the CDAS strategy [7], in order to control the dominance area of solutions and regulate selection pressure. However, as clarified and illustrated in [9], the dominance envelope varies solution-by-solution with CDAS if  $M > 2$ ; that is, it cannot provide a nondiscriminatory optimality criterion for any solution (a detailed analysis of the expanding coefficient of CDAS can be found in [9]). Hence, a GPO criterion is desired that retains the common features of conventional Pareto optimality, to ensure the identity of dominance envelopes for all solutions. For a detailed demonstration and discussion of GPO, interested readers can refer to the original paper.

The original GPO can work in a dictionary-like “pseudo-2-D” principle for even a relatively higher  $M$ -D problem [9], which accordingly, usually leads the population to converge to a subset (or several subsets) of the PF. As mentioned in Section I, the AGPO based on the chopping principle [14] can compensate for this to some extent, and achieve a relatively complete PF for some problems. However, in AGPO, a subpopulation size of 300 is used for each processor, such that the entire population contains  $300 \times M$  individuals, leading to a heavy computational burden. Also, the AGPO technique is still very sensitive to the value of expanding angle  $\varphi$ , which should be dynamically tuned each generation during the whole optimization process. Besides, it is worth pointing out that the performance of AGPO was only studied on four test problems of 3-D and 5-D cases in the original paper.

1)  $(M-1)$ -GPD-Based Ranking Scheme: The “ $M-1$ ” + 1 framework of GPO dominance [or  $(M-1)$ -GPD] is the most important part of this study. Here, “ $M-1$ ” means the selection pressure of  $M-1$  objectives is enhanced by expanding the dominance area of solutions, while 1 denotes the one objective left out of that process remains unchanged. For clearer understanding, Fig. 1 shows a graphical explanation of  $(M-1)$ -GPD in the original  $f_1$ - $f_2$  space and the indirect objective spaces  $f_1$ - $\Omega_2$  (blue) and  $\Omega_1$ - $f_2$  (green).

Given an expanding angle  $\varphi$ , for solution P, the  $f_1$  objective remains unchanged while  $f_2$  expands by  $\varphi$  degrees, that is, “ $M-1$ ” is related to  $f_2$  and “1” is for  $f_1$ , such that  $\Lambda_P^{\varphi(1)}$  depicts the coverage of its dominance envelope. Correspondingly, the case of  $f_1$ -“ $M-1$ ” and  $f_2$ -“1” is presented for solution Q, and  $\Lambda_Q^{\varphi(2)}$  shows its dominating envelope. In the case of  $\Lambda_P^{\varphi(1)}$ , the contracted blue space  $f_1 - f_2$  (inside axes  $f_1$  and  $\Omega_2$ ) is

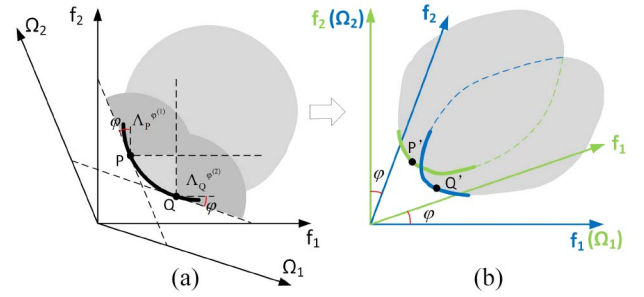


Fig. 1. Pictorial illustration of a 2-D objective space and the shrunk space after performing two  $(M-1)$ -GPD cases. (a) Original  $f_1$ - $f_2$  objective space. (b) Two indirect  $f_1$ - $\Omega_2$  (blue) and  $\Omega_1$ - $f_2$  (green) objective spaces (also the two contracted spaces inside) after generalization.

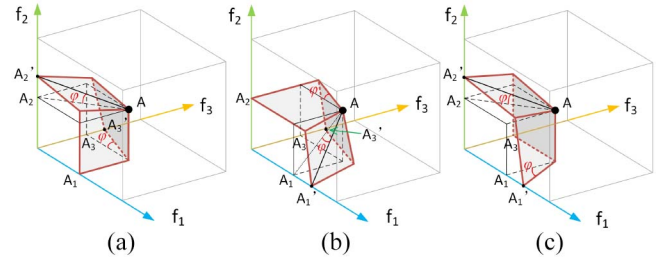


Fig. 2. Illustration of expanded dominance envelope for a 3-D example. (a)  $\Phi^1 = [0, \varphi, \varphi]$ . (b)  $\Phi^2 = [\varphi, 0, \varphi]$ . (c)  $\Phi^3 = [\varphi, \varphi, 0]$ .

obtained; while for dominating envelope  $\Lambda_Q^{\varphi(2)}$ , we have the contracted green space  $f_1$ - $f_2$  (inside axes  $\Omega_1$  and  $f_2$ ), as shown in Fig. 1(b). Moreover, the difference is clearly visible, marked in light gray, that is, the coverage of the true PF [bold PQ curve in Fig. 1(a)] in the original  $f_1 - f_2$  space and of the two partial generalized PFs [shown by the green arc of  $P'$  and the blue arc of  $Q'$ , respectively, in Fig. 1(b)]. In fact, a larger  $\varphi$  is likely to lower the coverage of the true PF. When  $\varphi$  reaches its maximum, the two transformed objective spaces will be close to the hyperlines, such that only two extreme solutions left in each of the extreme case, respectively.

In our proposed framework, we implement an  $(M-1)$ -GPD process, wherein the dominance area expansion is performed in  $M-1$  dimensions, and the left one remains nonexpansive. To take a 3-D problem as an example, the expanded dominance envelope of three symmetrical cases is shown in Fig. 2, where  $\Phi^1 = [0, \varphi, \varphi]$ ,  $\Phi^2 = [\varphi, 0, \varphi]$ , and  $\Phi^3 = [\varphi, \varphi, 0]$ , are used for the three processors with the same expanding angle  $\varphi$ . In Fig. 2(a), the dominance envelope of solution A (also point A) using the original Pareto dominance is depicted by the square  $AA_1A_2A_3$ , while its expanded dominance envelope using  $\Phi^1$ -based  $(M-1)$ -GPD is illustrated by the irregular polygon  $AA_1A_2'A_3'$ . Accordingly, in cases of  $\Phi^2$  and  $\Phi^3$ , the expanded dominance envelope of solution A is depicted by the irregular polygons  $AA_1'A_2A_3'$  in Fig. 2(b) and  $AA_1'A_2'A_3$  in Fig. 2(c), respectively. It is intuitive that among these three cases, each pair has an overlapping dominance area that has been expanded, which exhibits a complementary property.

With this subdivision, the selection pressure of each subtask is sufficient using a relatively small  $\varphi$ , which is fixed during the whole optimization process. For an  $M$ -objective optimization



problem, there are  $M$  different, parallel cases of  $(M-1)$ -GPD, and each one is responsible for converging to a relatively large subregion.

For convenience, we use  $\text{GPO}_{f_k}$  to mean that objective  $f_k$  remains unchanged, since the distinction among the processes is which single objective is not transformed [the “1” of “ $(M-1)+1$ ”]. To take  $\text{GPO}_{f_1}$  as an example, the transformed objectives of each solution are calculated by (5), where  $\varphi^* = \tan\varphi\sqrt{M-1}$  and only  $f'_1 = f_1$ . Thereafter, the fast nondominated sorting operation is performed on the new objectives  $[f'_1, f'_2, \dots, f'_M]$  of the population to achieve their ranking order, that is,  $(M-1)$ -GPD-based-Sort operation in line 10 of Algorithm 2. By conducting  $M$  different cases of  $\text{GPO}_{f_k}$ ,  $k \in [1, M]$ , each solution is labeled with  $M$  ranking orders

$$\begin{bmatrix} f'_1 \\ f'_2 \\ \vdots \\ f'_{M-1} \\ f'_M \end{bmatrix} = \begin{bmatrix} 1 & 0 & \cdots & 0 & 0 \\ \varphi^* & 1 & \cdots & \varphi^* & \varphi^* \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \varphi^* & \varphi^* & \cdots & 1 & \varphi^* \\ \varphi^* & \varphi^* & \cdots & \varphi^* & 1 \end{bmatrix} \cdot \begin{bmatrix} f_1 \\ f_2 \\ \vdots \\ f_{M-1} \\ f_M \end{bmatrix}. \quad (5)$$

2)  $(M-1)$ -GPD Environmental Selection: First, the fast nondominated sorting is adopted on the original objectives, for rough selection. The whole process of the  $(M-1)$ -GPD environmental selection is described in Algorithm 2, which roughly contains three steps: 1) identify the extreme solution set  $Q$  ( $|Q| \leq M$ ) from the combined population  $R$  and to append it to the surviving population  $P$  (lines 2 and 3); 2) simultaneously conduct the  $M$  symmetrical  $(M-1)$ -GPD-based-Sort operations to obtain an  $M \times 2N$  matrix  $\text{Front}^{(\text{GPD})}$  of sorted values, where each row denotes the sorting result of a specific case (lines 7–10 and 14); and 3) retain candidate solutions according to each of  $M$  rows of matrix  $\text{Front}^{(\text{GPD})}$  that are selected in a random order and resampled each generation (lines 15–30).

Note that solutions are selected almost evenly from each case of the  $(M-1)$ -GPD scheme, that is, the selected size of each case is around  $\lfloor (N - |Q|/M) \rfloor$  (given  $|Q|$  extreme solutions selected). Let  $t = \lfloor (N - |Q|/M) \rfloor$  for simplicity, the value of  $|Q| + M \times t$  is usually smaller than  $N$  in the case of either  $|Q| < M$  or  $N \nmid M$ . Thus, the remaining  $n_s$  (see line 5) solutions to  $P$  are selected based on  $n_s$  different  $(M-1)$ -GPD cases, such that  $t+1$  is assigned as the subset size of the corresponding random cases, that is, the randomly selected element of vector  $T_s$  (lines 4–6 and 11–13). By using the max–min distance-based selection, only solution set  $R_T$  in terms of smaller  $d_{\min}$  values (lines 24 and 25), which measured by the angular distance between the selected solutions and the remaining ones (lines 20–22), have the chance to be retained. Therefore, the solutions of  $R_T$  are selected one by one to further maintain diversity, that is, each time, one solution (among the subset of  $R_T$  that achieves the minimum value of  $PF^*$ , as shown in line 26) is selected that furthest isolates itself from the currently selected set (line 27).

3) *Illustrative Explanation*: To enhance understanding, we illustrate the  $(M-1)$ -GPD ranking method on a set of candidate solutions of a biobjective minimization problem, as shown in Fig. 3, where (a) depicts the case of  $\text{GPO}_{f_1}$ , while (b) is

## Algorithm 2: $(M-1)$ -GPD Environmental Selection

**Input:** Population size  $N$ , combined population  $R$  and the expanding angle  $\varphi$ ;  
**Output:** The survival solution set  $P$ .  
1 Initialize  $P = \emptyset$  and conduct fast non-dominated sorting on  $R$ ;  
2 Identify the extreme solution set  $Q$  in terms of the minimum ASF value for each  $\mathbf{w}_i$  axis direction,  $i = 1, 2, \dots, M$ ;  
3  $P \leftarrow P \cup Q$ , and  $R \leftarrow R \setminus Q$ ;  
4 Suppose a  $M$ -length vector  $T_s = [t, \dots, t]$ , where  $t = \lfloor \frac{N-|Q|}{M} \rfloor$ ;  
5  $n_s = N - |Q| - M \times t$ ;  
6 Generate a  $M$ -length zero vector  $\lambda = [0, \dots, 0]$ , then  $n_s$  randomly selected values in  $\lambda$  are set as 1;  
7 Initialize the expanding angle vector  $\hat{\Phi} = [\varphi, \dots, \varphi]$  of  $M$  same  $\varphi$ ;  
8 **for**  $k=1:M$  **do**  
9   Let  $\Phi^k = \hat{\Phi}$ , then  $\Phi^k(k) = 0$ ;  
10    $\text{Front}_k^{(\text{GPD})} = (M-1)\text{-GPD-based-Sort}(R, \Phi^k)$ ;  
11   **if**  $\lambda(k) = 1$  **then**  
12      $T_s(k) = t + 1$ ;  
13   **end**  
14 **end**  
15 Generate a value set  $\Gamma = \{1, 2, \dots, M\}$ ;  
16 **for**  $j=1:M$  **do**  
17   Let  $R_{\text{nd}}$  be the set of all non-dominated solutions in  $R$ ;  
18   Randomly select a value  $\tau$  from set  $\Gamma$ , then  $\Gamma \leftarrow \Gamma \setminus \tau$ ;  
19    $PF^* = \text{Front}_{\tau}^{(\text{GPD})}$ ;  
20   **for** each solution  $\mathbf{x} \in R_{\text{nd}}$  **do**  
21      $d_{\min}(\mathbf{x}) = \min_{\mathbf{y} \in P} \text{dist}(\mathbf{x}, \mathbf{y})$   
22   **end**  
23   **for**  $i = 1 : T_s(j)$  **do**  
24     Sort solutions of  $R_{\text{nd}}$  in descending order of  $d_{\min}$  values;  
25     Let  $R_T$  be the subset of the top ranked  $N_t$  individuals in  $R_{\text{nd}}$ ;  
26      $S = \{\mathbf{x} | \mathbf{x} \in R_T, PF^*(\mathbf{x}) = \min(PF^*(R_T))\}$ ;  
27      $s_i = \arg \max(d_{\min}(S))$ ;  $\% \text{ if } |S| = 1$ , then  $s_i = S$   
28      $P \leftarrow P \cup s_i$ ,  $R \leftarrow R \setminus s_i$ ;  
29   **end**  
30 **end**

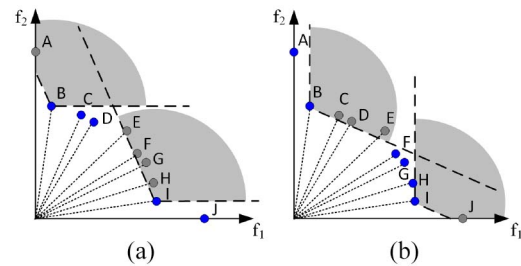


Fig. 3. Illustration of the  $(M-1)$ -GPD ranking method. (a)  $\text{GPO}_{f_1}$ : Objective  $f_1$  remains unchanged and  $f_2$  expands the dominance region. (b)  $\text{GPO}_{f_2}$ : Objective  $f_2$  remains unchanged and  $f_1$  expands the dominance region.

for  $\text{GPO}_{f_2}$ . The ten circles (labeled from A to J) denote the nondominated solutions of this problem, as analogous to those obtained on an MaOP, which are usually incomparable.

In Fig. 3(a), solutions B, C, D, I, and J highlighted in blue are ranked in the first order by  $\text{GPO}_{f_1}$ ; while Fig. 3(b) shows that solutions A, B, F, G, H, and I (also highlighted in blue) are ranked order 1 according to  $\text{GPO}_{f_2}$ . It is obvious that these two cases are mostly complementary to each other. That is to say, employing  $\text{GPO}_{f_1}$  or  $\text{GPO}_{f_2}$  can enhance selection pressure in terms of discriminability, and diversity can be preserved by adopting them simultaneously. Apart from the complementarity of the distinct symmetrical  $(M-1)$ -GPD schemes, the max–min distance-based selection also plays a role in maintaining diversity. For example, solution E is accorded a higher

rank ( $>1$ ) by either  $GPO_{f_1}$  or  $GPO_{f_2}$ ; however, its density in terms of the max–min distance is better than others. Hence, a priority will be provided to solution E for survival during environmental selection. Moreover, solutions C and D (or F and G) are closer to each other in terms of angular distance. With the max–min distance-based selection, so if C or F is selected, the nearest neighbor D or G can hardly be considered for survival to the next generation.

The  $(M-1)$ -GPD-based solution ranking method can be also easily embedded into other existing MaOEAs, for example, NSGA-III, to improve their environmental selection by generating more comparable solution ranks. For example, an improved version of NSGA-III, that is, NSGA-III\*, is discussed in Section IV-B.

#### E. Complexity and Property Analysis

In each generation, the computational complexity is dominated by the environmental selection procedure: suppose population size is  $N$  and objective dimension is  $M$ , each of the  $M$  different  $(M-1)$ -GPD cases takes  $O(MN^2)$  time (but can be executed in parallel), so the total complexity is  $O(M^2N^2)$ . But if the operations are executed in parallel, the wall clock time of MultiGPO remains  $O(MN^2)$  each generation.

In fact, MultiGPO is mainly simple from the framework perspective. In the literature, the main existing MaOEAs had several additional components and/or the adaptive reference vector tuning schemes, such as [32], [37], [38], [44], [56], [57], etc. In contrast, MultiGPO is designed in a simpler framework, as similar to the NSGA-II method for biobjective problems, whereas it performs effectively on MaOPs. A more detailed analysis is provided in Supplementary Section I-B.

According to the definition of GPO (Definition 2), the proposed  $(M-1)$ -GPD relationship is asymmetric, transitive, and irreflexive, and the theoretical proofs can be found in Supplementary Section I-B. Hence,  $(M-1)$ -GPD is compatible with the conventional Pareto dominance.

### IV. EXPERIMENTAL RESULTS

This section mainly presents the experimental comparison on a set of benchmark test problems. We first introduce the experimental design. Then, the effect of our proposed  $(M-1)$ -GPD is confirmed, and MultiGPO is compared with some state-of-the-art MaOEAs to show its advantages. Finally, we provide the sensitivity analysis regarding the expanding angle  $\varphi$  and investigate further the dissimilarity measure.

#### A. Experimental Design

1) *Algorithms in Comparison*: In this study, we compare the performance of MultiGPO with some state-of-the-art representative MaOEAs, including NSGA-III [5], MOEA/DD [19],  $\theta$ -DEA [11], lby1EA [47], and NSGA-II/SDR [13]. The parameter  $N_t$  (see line 25 of Algorithm 2) is set as  $\min(0.5|R|, |R_{nd}|)$  to balance selection pressure enhancement and the subset selection-based diversity preservation, as discussed in Supplementary Section I-B. In MOEA/DD, the neighborhood selection probability is  $\delta = 0.9$ , the neighborhood size  $T$  and the maximum number of solutions replaced

by each offspring  $n_r$  are  $T = \lceil 0.1N \rceil$  and  $n_r = \lceil 0.01N \rceil$ , respectively. According to [11], PBI is used with  $\theta = 5$  in  $\theta$ -DEA. In lby1EA, the parameters  $k$  and  $R$  are set to  $k = 0.1N$  and  $R = 1$ , respectively, following the guidelines in [47]. The time complexity of RVEA,  $\theta$ -DEA, MOEA/DD, and NSGA-II/SDR per generation is  $O(MN^2)$ , while it is  $\max\{O(N\log^{M-2}N), O(MN^2)\}$  for NSGA-III.

2) *Test Problems*: In our study, 31 scalable test problems from two test suites are used, that is, the WFG1–WFG9 test problems with differently scaled objectives [58] and the MaF1–MaF15, which have complicated PFs [59]. For each problem, we consider the number of objectives  $M$  varying from 5 to 20, that is,  $M \in \{5, 8, 10, 15, 20\}$ . For the nine WFG test problems, the number of decision variables is set to  $n = k + l$ , where  $k = M - 1$  and  $l = 10$ . For the MaF test problems, the setting of  $n$  is not very unified, for example,  $n = M + 9$  for MaF1–MaF6 and MaF10–MaF12;  $n = M + 19$  for MaF7;  $n = 2$  for MaF8 and MaF9,  $n = 5$  for MaF13; and  $n = 20 \times M$  for MaF14 and MaF15. These test MaOPs contain different problem properties, such as concavity, convexity, linearity, degeneracy, and disconnectedness. Due to page limitations, their main properties are summarized in Table S-II of the supplementary material.

3) *General Parameter Settings*: In all algorithms, SBX and PM are used to generate offsprings. The crossover probability  $p_c$  is set to 1 and mutation probability  $p_m$  is set to  $1/M$ ; the distribution index of them is set to 20. In principle, the population size  $N$  of MultiGPO can be arbitrarily assigned. However, NSGA-III, MOEA/DD, and  $\theta$ -DEA require  $N$  to be the same as the number of associated reference vectors, which cannot be arbitrarily set because the number of reference lines generated by Das and Dennis's method is a binomial coefficient [60]. To make a fair comparison, the population size of all the algorithms is set to the same in case of the same  $M$ , namely, 210, 240, 275, 240, and 210 for 5-, 8-, 10-, 15-, and 20-objective test problems. The maximum generation ( $G_{\max}$ ) is adopted as the termination criterion for all algorithms, which is set to 200 for the WFG problems and 500 for the MaF problems. Note that the termination criterion of WFG problems is the same as that in NSGA-II/SDR [13].

4) *Performance Metrics*: We adopted two popular performance metrics, namely, IGD and HV [5], together with the recently proposed pure distance (PD) measure [55] to assess the performance of each algorithm. The IGD and HV combine the convergence and diversity measurement together to obtain a single value, while the PD is used as an auxiliary criterion focusing only on the diversity assessment. A smaller IGD value indicates a better result, while a larger HV indicates a solution set of higher quality. To calculate IGD, a set of representative points (the size is around 10 000) that are uniformly distributed on the true PF of each test instance are sampled as the reference points. For HV, all the objective values are normalized by the ideal and nadir points (we here take the nadir point as 1.1 times the maximum value of each objective) before calculating the HV values. Then, the reference point is set to  $(1, \dots, 1)$  for the calculation of HV. Moreover, if  $M \geq 5$ , the Monte Carlo estimation method with 1 000 000 sample points is adopted to obtain

the approximate HV more efficiently. Note that all the experiments in this study were executed on the open-source PlatEMO platform [61], in which the reference points are generated mostly based on [60].

We independently conduct all the experiments for 30 runs and record the mean value and the standard deviation of each problem instance. Moreover, the Wilcoxon rank-sum test is used to analyze the results at a significance level of 0.05, where  $+$ ,  $-$ , and  $\approx$  indicate that the compared MaOEA performs significantly better than, worse than, or similar to MultiGPO, respectively, on that problem instance.

### B. Effect of $(M - 1)$ -GPD

1) *Comparing  $(M - 1)$ -GPD With Other Modified Pareto-Dominance Methods:* In the following, we verify the effectiveness of  $(M - 1)$ -GPD in balancing convergence and diversity by comparing it with some representative modified Pareto-dominance methods, for example, Cone  $\epsilon$ -dominance [6], CDAS [7],  $CN\alpha$ -dominance [8], and GPO [9] shown in Fig. 4. The detailed description of these methods is provided in supplementary Section I-E. It is evident that Cone  $\epsilon$ -dominance still cannot provide sufficient selection pressure if  $M$  increases. The CDAS method seems to be similar to GPO, however, the former cannot guarantee the identical expansion of the dominance area in all dimensions if  $M \geq 3$ , leading to difficulties of ranking solutions by the same degree. Compared to CDAS and GPO, the  $CN\alpha$ -dominance can provide a larger selection pressure by further expanding the dominated area; whereas none of them can well preserve population diversity in general. Therefore, in the proposed  $(M - 1)$ -GPD method, the diversity issue has been addressed in a better manner. We use  $M$  symmetrical cases to maintain a good balance between convergence and diversity, for example, on the right of Fig. 4(f), three  $\Phi^1$ ,  $\Phi^2$ , and  $\Phi^3$ -based  $(M - 1)$ -GPD schemes are adopted.

We also compare our MultiGPO with the MOEAs embedding the four modified Pareto-dominance methods (CDAS-, S-CDAS-,  $CN\alpha$ -, and GPO-dominance) and one non-Pareto strengthened dominance method (SDR [13]), as shown in Supplementary Figs. S-3–S-5 (three 3-objective and two 10-objective DTLZ problems). Note that the expanding angle  $\varphi$  of GPO and MultiGPO is set as the value of  $2M$ , rather than a relatively large value fixed for all  $M$ , as found in [13]. The parameters of other methods follow the settings as suggested in their original literature. From Figs. S-3 to S-5 in the supplementary material, MultiGPO can outperform the counterparts and obtain a good result for each case. On the contrary, most of the other modified Pareto-dominance methods perform poorly in maintaining a widely distributed population, although GPO can achieve better overall performance among them. It is worth pointing out that the recent proposed SDR criterion [13] performs very poorly on all the three 3-objective DTLZ problems.

In addition, we compare GPO and MultiGPO by setting different expanding angles (i.e.,  $\varphi = 10, 20, 30$ ) on three 3-objective DTLZ problems, respectively, as shown in Figs. S-6–S-8 of the supplementary material. It is evident that MultiGPO performs much better than GPO in all cases. The diversity of

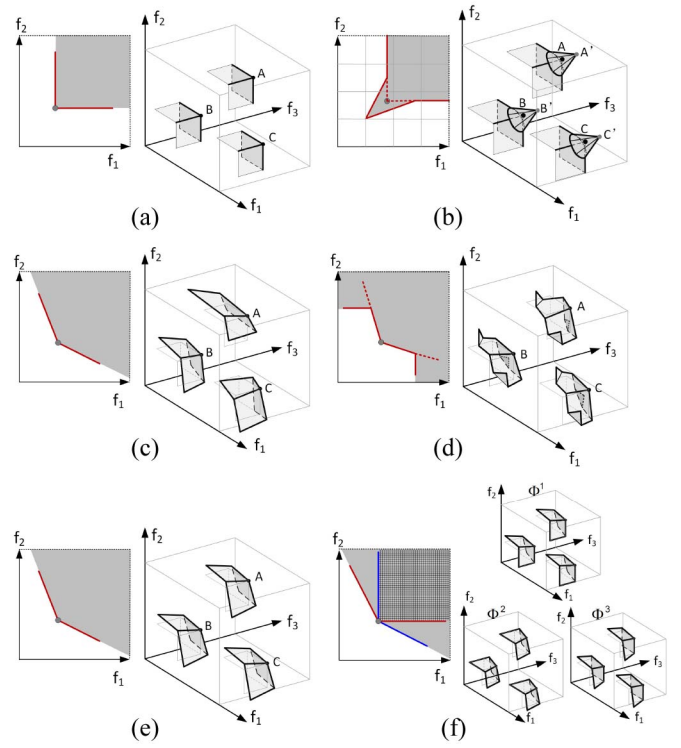


Fig. 4. Dominating areas obtained by six different relaxed Pareto-dominance relations in the biobjective space and the dominance envelope in three-objective spaces. (a) Pareto dominance. (b) Cone  $\epsilon$ -dominance. (c) CDAS (or S-CDAS). (d)  $CN\alpha$ -dominance. (e) GPO. (f)  $(M - 1)$ -GPD.

solutions obtained by MultiGPO is acceptable even by setting  $\varphi = 30$ , whereas GPO degrades dramatically when  $\varphi$  is getting larger.

2) *Embedding  $(M - 1)$ -GPD Into NSGA-III:* In NSGA-III, the convergence engine only depends on the traditional Pareto dominance, which is, however, insufficient to solve MaOPs with a large  $M$  given a limited budget. For example, NSGA-III usually needs a large number of generations (e.g.,  $G_{\max} > 1000$ ) to obtain a desirable result in cases of higher objectives. Hence, we replace the Pareto dominance of NSGA-III by our proposed  $(M - 1)$ -GPD schemes, in order to enhance selection pressure to determine solution ranks without losing diversity, which leads to the new algorithm NSGA-III\*. To be specific, NSGA-III\* adopts an improved environmental selection step, as presented in Algorithm S1 of the supplementary material. Instead of using WFG and MaF1 test suites mentioned in Section IV-A, we test the DTLZ problems (DTLZ1–DTLZ7) [62] in this section, which is more suitable for reference-weight-based MaOEAs due to their regular triangular properties [17]. The termination criterion is set to  $G_{\max} = 500$  here.

Table I reports the HV results of NSGA-III and NSGA-III\* ( $\varphi = 2 \times M$ ) on DTLZ1–DTLZ7 problems with 5, 10, and 15 objectives, where the best result for each instance is shown with bold font in gray shade. Moreover, the HV result of the five peer methods and NSGA-III\* on these problems is shown in Table S-III of the supplementary material. As observed from Table I, NSGA-III\* performs better or similar to NSGA-III in 26 out of 27 comparisons. Also, from Table S-III,



TABLE I  
HV COMPARISON OF NSGA-III AND NSGA-III\* ON DTLZ1–DTLZ7

Prob.( $M$ )	NSGA-III	NSGA-III*	Prob.( $M$ )	NSGA-III	NSGA-III*
DTLZ2(5)	<b>8.12e-1</b> $\approx$	<b>8.12e-1</b>	DTLZ5(5)	1.12e-1 –	<b>1.17e-1</b>
DTLZ2(10)	9.56e-1 $\approx$	<b>9.62e-1</b>	DTLZ5(10)	8.82e-2 –	<b>9.43e-2</b>
DTLZ2(15)	<b>9.80e-1</b> $\approx$	9.79e-1	DTLZ5(15)	8.34e-2 –	<b>9.10e-2</b>
DTLZ3(5)	7.94e-1 $\approx$	<b>7.99e-1</b>	DTLZ6(5)	1.03e-1 –	<b>1.10e-1</b>
DTLZ3(10)	4.94e-1 –	<b>9.44e-1</b>	DTLZ6(10)	6.36e-2 –	<b>9.30e-2</b>
DTLZ3(15)	2.04e-1 –	<b>7.59e-1</b>	DTLZ6(15)	1.81e-2 –	<b>9.12e-2</b>
DTLZ4(5)	8.12e-1 $\approx$	<b>8.12e-1</b>	DTLZ7(5)	2.56e-1 –	<b>2.60e-1</b>
DTLZ4(10)	9.69e-1 –	<b>9.70e-1</b>	DTLZ7(10)	1.70e-1 –	<b>1.90e-1</b>
DTLZ4(15)	9.85e-1 –	<b>9.91e-1</b>	DTLZ7(15)	<b>8.18e-2</b> +	6.85e-2

Note: The result of DTLZ1 is not shown as it is similar to that of DTLZ2—i.e., NSGA-III and NSGA-III\* perform significantly similar to each other for each  $M$ .

MOEA/DD,  $\theta$ -DEA, and NSGA-III\* obtain the overall better HV performance than the others, partly due to their usage of reference weights. However, NSGA-III is not competitive to the other three reference-weight-based MaOEAs (MOEA/DD,  $\theta$ -DEA, and NSGA-III\*), since the Pareto dominance is not sufficient to let the population converge with 500 generations. These results confirm the effect of the  $(M - 1)$ -GPD selection criterion. It can also be easily embedded into other existing MaOEAs, hence a good balance of convergence and diversity.

### C. Performance Comparison With Other MaOEAs

According to our prior experiments, the performance of MultiGPO in terms of the IGD and HV metrics is not very consistent, especially on WFG problems. Specifically, MultiGPO with a smaller  $\varphi$  tends to obtain better IGD results, while a higher  $\varphi$  is beneficial to obtain a better HV. Hence, we provide the results of both MultiGPO with  $\varphi = M$  and MultiGPO2 with  $\varphi = \eta \times M$ , where  $\eta = 3$  if  $M \leq 10$  and  $\eta = 2.5$  otherwise.

1) *Performance Comparisons on the WFG Test Suite:* The statistical results of all methods on the WFG1–WFG9 test problems in terms of IGD and HV are shown in Tables S-VI and S-VII, respectively, in the supplementary material. From Table S-VI in the supplementary material, MultiGPO and  $\theta$ -DEA achieve better overall IGD performances than the others. MultiGPO obtained the best or equaled the best results on 21 out of 45 comparisons,  $\theta$ -DEA did that on 17; while NSGA-III, MOEA/DD, 1by1EA, NSGA-II/SDR, and MultiGPO2 achieved the best performance on 8, 1, 0, 2, and 11, respectively. In total, MultiGPO performs better than NSGA-III, MOEA/DD,  $\theta$ -DEA, 1by1EA, and NSGA-II/SDR in 26, 35, 19, 43, and 32 out of 45 test cases, respectively. On the contrary, MOEA/DD and 1by1EA show the worst performance, which also happens in terms of HV, as stated later. MultiGPO2 performs a little worse than MultiGPO in terms of IGD, but is still competitive to other MaOEAs.

In contrast to the IGD result, MultiGPO2 tends to obtain better HV performance than MultiGPO, and the latter usually performs worse than other MaOEAs, except MOEA/DD and 1by1EA. From Table S-VII in the supplementary material, it is evident that MultiGPO is outperformed by NSGA-III,  $\theta$ -DEA, and NSGA-II/SDR in 24, 34, and 33 out of 45 comparisons, respectively, whereas MultiGPO2 outperforms these three methods in 32, 31, and 20 comparisons, respectively. This time, MultiGPO2 and NSGA-II/SDR can achieve better

overall HV performances than the others, while MOEA/DD and 1by1EA still show the worst performance. In total, MultiGPO2 obtained the best or equaled the best HV results on 26 out of 45 comparisons, and NSGA-II/SDR did that on 20. Table S-VIII in the supplementary material shows the PD results of all methods on the WFG1–WFG9 test problems with 5, 10, and 15 objectives. It is obvious that MultiGPO obtains the best or equivalent results in all cases, while MultiGPO2 performs a little worse but still mostly outperforms the others. This phenomenon indicates that a smaller  $\varphi$  is beneficial to better diversity performance. It is known that using the HV metric usually assigns more preference to the knees and borders of the approximate PF [45]. This can partly account for the above observation that MultiGPO2 shows much better HV results than MultiGPO, since MultiGPO assigns more equal preference across the search space, as the expanded dominance envelope of each solution is smaller.

2) *Performance Comparisons on the MaF Test Suite:* Tables S-IX and S-X (see the supplementary material) report the IGD and HV results of all methods on MaF1–MaF15, respectively. From these comparisons, both MultiGPO and MultiGPO2 can achieve overall better IGD and HV performances than other MaOEAs on most instances, which shows the competitive potential of the proposed  $(M - 1)$ -GPD selection criterion on problems with complicated PFs. To be specific, in terms of IGD, MultiGPO obtains better results than NSGA-III, MOEA/DD,  $\theta$ -DEA, 1by1EA, and NSGA-II/SDR in 55, 60, 53, 48, and 54 out of 75 cases, respectively. However, MultiGPO performs poorly on MaF3–MaF6 ( $M \geq 10$ ) problems, as do most of the other compared methods, which indicates that more search effort is required.

To aid the reader's intuition, parallel coordinate plots of the objective values for each algorithm on the ten-objective WFG and MaF problems are provided in Figs. S-10–S-12 of the supplementary material. Here, the resulting nondominated solutions of MultiGPO ( $\varphi = 3$ ,  $N = 200$ , and  $G_{\max} = 400$ ) on three-objective MaF1–MaF9 are shown in Fig. 5, while the plot comparison of all methods (with the same experimental settings) on the three-objective MaF test suite are provided in Figs. S-13 and S-14 of the supplementary material. As observed from Fig. 5, MultiGPO can in most cases attain widely distributed solutions, although it does not perform very well on MaF3 and MaF4.

From Figs. S-10 to S-12 in the supplementary material, MultiGPO shows effectiveness in solving most problems, except for WFG1–WFG3, MaF5, MaF10–11, and MaF14–15. Nevertheless, these eight problems cannot be well addressed by the other competitors, either. The convergence of NSGA-III (sometimes  $\theta$ -DEA as well) tends to be much worse than those of other MaOEAs on some MaF problems (e.g., MaF4, MaF6–9, and MaF14–15) due to the lack of sufficient selection pressure of the conventional Pareto-dominance criterion, despite using reference points that can help alleviate this to some degree. Furthermore, we can see from Figs. S-13 and S-14 in the supplementary material that NSGA-II/SDR performs very poorly on most three-objective MaF problems, although it does achieve second-best overall performance on problems with more objectives. This indicates that the

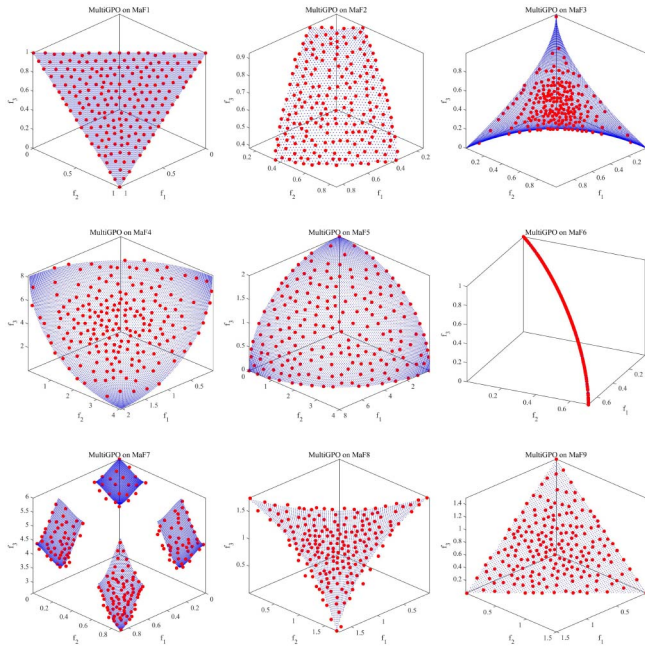


Fig. 5. Plot of results for MultiGPO on three-objective MaF1–MaF9.

SDR [13], considering convergence in terms of the sum of objectives and diversity based on the tailed niching technique, does not generalize well to fewer objectives.

3) *Overall Performance*: Table II summarizes the Wilcoxon signed-rank tests in terms of both IGD and HV between MultiGPO (or MultiGPO2) and the other methods at a significance level of 0.05. It is evident that MultiGPO and MultiGPO2 win in most cases out of 120 problem instances in terms of IGD, and MultiGPO is shown to be a little better than MultiGPO2. For the HV results, MultiGPO2 can clearly beat NSGA-III, MOEA/DD,  $\theta$ -DEA, and lby1EA, and most often beat NSGA-II/SDR, as well. However, in this case, MultiGPO performs much worse than MultiGPO2, and is even outperformed by  $\theta$ -DEA and NSGA-II/SDR on about 50% of the instances. This is mainly because  $\theta$ -DEA and NSGA-II/SDR beat MultiGPO on WFG instances over 2/3 of the time (see Table S-VI of the supplementary material) in terms of HV, while also showing performance competitive with MultiGPO. According to the above results, there may exist a weak trade-off between IGD and HV measures, which deserves further investigation.

#### D. Sensitivity Analysis of Expanding Angle $\varphi$

Here, we attempt to investigate further the influence of parameter  $\varphi$  on the final performance of  $(M - 1)$ -GPD. The MultiGPO is tested with  $\varphi = [0.5, 1, 1.5, 2, 2.5, 3, 3.5, 4] \times M$  for problems of  $M = 5$ ,  $M = 8$ , and  $M = 10$ ; while  $\varphi = [0.5, 1, 1.5, 2, 2.5, 3] \times M$  in cases of  $M = 15$  and  $M = 20$ .

Figs. 6 and 7 show the IGD and HV results of MultiGPO, respectively, on the WFG problems with different settings of  $\varphi$  for different numbers of objectives. To make a clearer comparison, the WFG5 test problem is not included, since it always has similar results to others. We can observe that mostly the

TABLE II  
SUMMARY OF THE WILCOXON SIGNED-RANK TESTS BETWEEN MULTIGPO AND THE OTHER ALGORITHMS AT SIGNIFICANCE LEVEL 0.05

Indicator	MultiGPO vs.			MultiGPO2 vs.		
	+	−	≈	+	−	≈
IGD	MultiGPO	/	/	25	47	48
	MultiGPO2	47	25	48	/	/
	NSGA-III	81	24	15	71	37
	MOEA/DD	95	19	6	92	21
	$\theta$ -DEA	72	31	17	68	39
	lby1EA	91	15	14	87	23
	NSGA-II/SDR	86	26	8	70	34
HV	MultiGPO	/	/	61	17	38
	MultiGPO2	17	61	38	/	/
	NSGA-III	47	44	25	74	23
	MOEA/DD	90	15	11	103	7
	$\theta$ -DEA	44	62	10	78	27
	lby1EA	74	30	12	82	18
	NSGA-II/SDR	43	60	13	65	36

Note: here +, − and ≈ indicate that MultiGPO (or MultiGPO2) performs significantly better, significantly worse and statistically similar to the counterpart in each row, respectively.

IGD value does not change much as the value of  $\varphi$  increases; however, a higher  $\varphi$  tends to help MultiGPO get a better result in terms of HV.

#### E. Further Investigation of MultiGPO

The distance-based subset selection, as mentioned in Section III, also plays an important role in preserving population diversity. Hence, we adopt some other dissimilarity metrics instead of the angular distance to investigate more about the  $(M - 1)$ -GPD scheme. Fig. 8 illustrates three different distances: (a) depicts the angular distance and (b) presents two transformed distances that also consider vector directions, by mapping the original points on a hyperplane (the blue line) or a roughly estimated PF (the red curve) along each associated direction.

Let  $F'(\mathbf{x}) = (f'_1(\mathbf{x}), f'_2(\mathbf{x}), \dots, f'_M(\mathbf{x}))$  denote the normalized objective vector of solution  $\mathbf{x}$ , as computed by max–min normalization with the ideal and nadir points. The hyperplane-assisted Euclidean distance between solutions  $\mathbf{x}$  and  $\mathbf{y}$  is [35]

$$d_\lambda(\mathbf{x}, \mathbf{y}) = \|\lambda(\mathbf{x}) - \lambda(\mathbf{y})\| \quad (6)$$

$$\text{where } \lambda(\mathbf{x}) = \frac{1}{\sum_{k=1}^M f'_k(\mathbf{x})} F'(\mathbf{x}). \quad (7)$$

Similarly, for the estimated PF-assisted Euclidean distance, the mapped points of (7) are replaced by

$$\lambda(\mathbf{x}) = \frac{1}{\left(\sum_{k=1}^M f'_k(\mathbf{x})^p\right)^{1/p}} F'(\mathbf{x}) \quad (8)$$

where  $p$  is estimated dynamically according to the shape of the approximated PF. The estimation of PF shapes is actually not a trivial issue; however, we do not pay more attention to it here. Instead, we adopt the simple, yet effective strategy of estimating PF shapes that is proposed in [33]. Interested readers can refer to the original paper for more details.

In addition, we also test other distance-based selection strategies instead of the max–min distance, for example, the MED [52], which is defined to guide the preservation of uniform distance and extension of individuals to approximate the

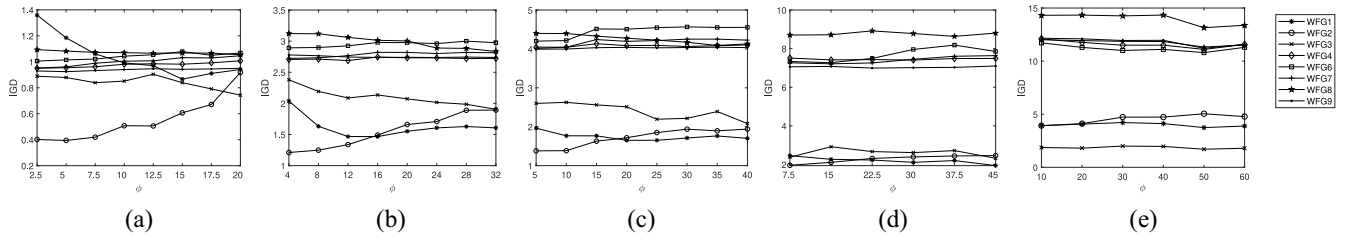


Fig. 6. IGD results of MultiGPO on WFG problems (except WFG5) with different settings of  $\phi$  for different numbers of objectives. (a) 5-objective. (b) 8-objective. (c) 10-objective. (d) 15-objective. (e) 20-objective.

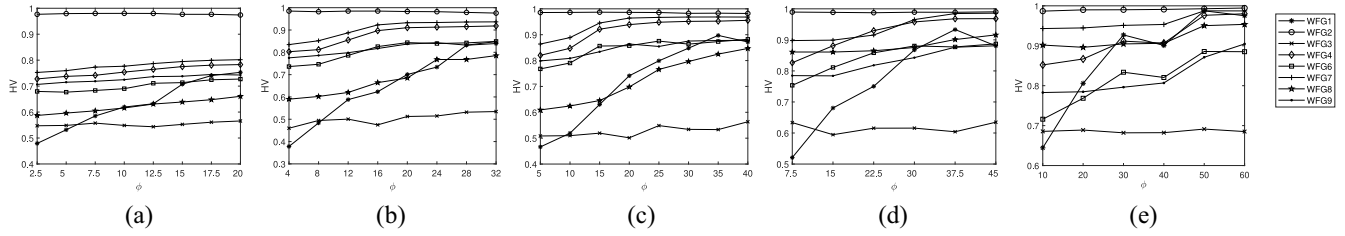


Fig. 7. HV results of MultiGPO on WFG problems (except WFG5) with different settings of  $\phi$  for different numbers of objectives. (a) 5-objective. (b) 8-objective. (c) 10-objective. (d) 15-objective. (e) 20-objective.

entire PF, that is

$$\text{MED}(\mathbf{x}_i) = \min_{j, j \neq i} d_{\cos}(\mathbf{x}_i, \mathbf{x}_j) \times \sum_{j=1}^N d_{\cos}(\mathbf{x}_i, \mathbf{x}_j). \quad (9)$$

However, the MED cannot be directly used to replace (4) of the proposed framework. A minor change is made here; that is, line 21 of Algorithm 2 is replaced by “ $d_{\min}(r) = \min_{y \in Q} \text{dist}(r, y)$ , and then  $d_{\min}(r) = d_{\min}(r) \times \sum_{y \in R^*} d(r, y)$ .” Note that here,  $R^*$  denotes the original combined population, as shown in line 6 of Algorithm 1.

By using hyperplane-assisted, PF-assisted distances, and MED, three variants of MultiGPO can be generated, denoted Variant1, Variant2, and Variant3, respectively. Their average IGD and HV results on 10-objective MaF1–MaF15 are shown in Tables S-IV and S-V of the supplementary material, respectively. We can see that there is mostly not much improvement with the three variants; this may be due to the fact that the  $(M - 1)$ -GPD scheme itself can provide desirable diversity maintenance. On the other hand, this phenomenon implicitly indicates that any of these distances that rely on vector directions can be effectively used in addressing MaOPs. Moreover, the variant based on the PF-assisted distance (Variant2) performs significantly better on MaF3 in terms of both IGD and HV; while Variant1 and Variant3 show better IGD values but worse HV values on MaF3.

Supplementary Fig. S-9 illustrates the results of MultiGPO and two variants (Variant1 and Variant2) on 3-objective MaF3. We can see that Variant1 improves the population distribution a little, while Variant2 gains a significant improvement to produce an evenly distributed PF. The PF of MaF3 is convex and multimodal, and the above observation suggests that the PF-assisted distance is more robust than the other distances in maintaining diversity on this problem. Despite the fact that the usage of the PF-assisted distance has not yielded a significant improvement across a majority of the test problems here, it deserves further efforts to explore its potential in some

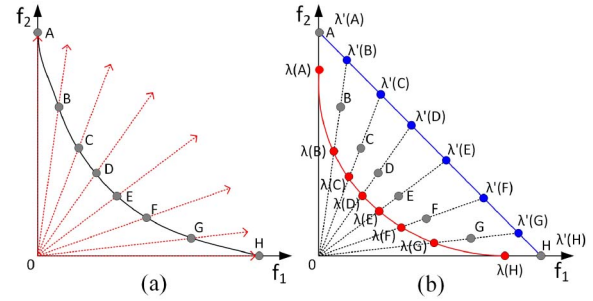


Fig. 8. Illustration of the distance between solutions. (a) Angular distance. (b) Hyperplane-assisted and the estimated PF-assisted Euclidean distances.

cases, especially when the PF shapes can be estimated more accurately than is proposed in [33].

It deserves pointing out that in some works  $G_{\max}$  is set as large as 2000 or even 3000 in addressing benchmark MaOPs, which is, however, impractical for many real-world applications. For example, the number of fitness evaluations is usually no larger than 10000 (or  $G_{\max} < 100$ ) when applying an EA/MOEA on data mining [2], [63], [64], deep learning [3], [65]–[67], or other tasks such as electroencephalogram (EEG) [68], [69], for which the objective function is expensive to evaluate. Therefore, we conduct additional experiments by setting  $G_{\max} = 200$  for the MaF test suite independent of  $M$ . The IGD and HV results are reported in Tables S-XI and S-XII of the supplementary material. We see that MultiGPO still obtains the overall best performance, which shows its promising potential when fewer function evaluations are available.

## V. APPLICATION TO REAL-WORLD PROBLEMS

In this part, three practical many-objective engineering problems with irregular PFs, that is, automotive frontal crashworthiness design [70], the water resource problem [40], and

TABLE III  
PERFORMANCE COMPARISON OF NSGA-III,  $\theta$ -DEA, 1by1EA, NSGA-II/SDR, AND MULTIGPO ON THREE REAL-WORLD PROBLEMS

Problem	Indicator	NSGA-III	$\theta$ -DEA	1by1EA	NSGA-II/SDR	MultiGPO
Crashworthiness	IGD <sup>†</sup>	2.7927e-1 (6.34e-2) $\approx$	2.5777e-1 (3.35e-2) $\approx$	1.2578e+0 (5.49e-1) $-$	1.5419e+0 (5.16e-1) $-$	<b>2.4503e-1 (7.89e-2)</b>
	HV <sup>‡</sup>	1.0299e+0 (3.51e-3) $-$	1.0236e+0 (6.83e-3) $-$	9.5573e-1 (1.27e-2) $-$	9.5914e-1 (1.15e-2) $-$	<b>1.0331e+0 (3.37e-3)</b>
	PD	2.9438e+5 (4.79e+4) $-$	3.0862e+5 (4.72e+4) $\approx$	1.6260e+5 (3.24e+4) $-$	1.1584e+5 (2.76e+4) $-$	<b>3.3520e+5 (5.12e+4)</b>
Order		2.67	2.33	4.33	4.67	1.00
Water Resource	IGD <sup>†</sup>	1.1379e+5 (7.03e+4) $-$	4.6759e+5 (1.32e+5) $-$	6.5556e+5 (5.56e+3) $-$	1.8400e+6 (1.49e+4) $-$	<b>3.3432e+4 (6.18e+3)</b>
	HV <sup>‡</sup>	1.3788e+0 (5.42e-3) $-$	1.3463e+0 (8.70e-3) $-$	3.0663e-1 (1.80e-2) $-$	6.7593e-2 (1.61e-2) $-$	<b>1.3984e+0 (3.73e-3)</b>
	PD	<b>1.693e+12 (1.12e+11) <math>\pm</math></b>	9.232e+11 (1.31e+11) $-$	1.142e+11 (1.28e+10) $-$	7.855e+10 (7.06e+9) $-$	1.334e+12 (5.39e+10)
Order		1.67	3	4	5	1.33
Car Side-Impact	IGD <sup>†</sup>	1.7817e+0 (1.61e-1) $-$	2.2910e+0 (3.55e-1) $-$	3.5631e+0 (5.61e-1) $-$	1.6752e+0 (1.06e-1) $-$	<b>1.2291e+0 (4.64e-2)</b>
	HV <sup>‡</sup>	1.2134e-1 (1.14e-2) $-$	1.4597e-1 (1.21e-2) $-$	7.2842e-2 (1.12e-2) $-$	1.2727e-1 (1.35e-2) $-$	<b>2.0258e-1 (6.53e-3)</b>
	PD	4.302e+10 (2.31e+9) $-$	2.733e+10 (2.17e+9) $-$	3.579e+10 (2.77e+9) $-$	4.088e+10 (1.78e+9) $-$	<b>6.060e+10 (1.27e+9)</b>
Order		3.00	3.67	4.67	2.67	1.00

<sup>†</sup> The reference Pareto front (denoted as  $P^*$ ) for calculating IGD is constructed by combining all the non-dominated solutions from lengthy runs.

<sup>‡</sup> The HV values are computed with reference point  $(1.1, \dots, 1.1)^T$  for normalized objective values of all the solutions.

automotive design for side impact [40], are used to further investigate the effectiveness of MultiGPO. The descriptions of these problems are provided in Supplementary Section I-C.

### A. Experimental Setups

MultiGPO is compared with four algorithms, that is, NSGA-III,  $\theta$ -DEA, 1by1EA, and NSGA-II/SDR, in terms of the three performance metrics IGD, HV, and PD. In MultiGPO, parameter  $\varphi$  is set the same as the number of objectives—here, that is,  $\varphi = 3$  for FCDP,  $\varphi = 6$  for WRP, and  $\varphi = 9$  for CSIP. For all methods, the population size is set to 156 for FCDP and 210 for the other two problems, and the maximum number of generations is 200. Since the true PF of the problem is unknown, a reference PF (denoted as  $P^*$ ) for use in the IGD measure is constructed by executing all of the above five algorithms for 800 generations, then combining all the nondominated solutions to compute an IGD measure for each problem. To calculate HV, the maximal and minimal objective values obtained from  $P^*$  are used to normalize the objective values of all the solutions; thereafter, the HV values are computed with reference point  $(1.1, \dots, 1.1)^T$ .

### B. Results and Discussion

Table III presents the means and standard deviations of IGD, HV, and PD values obtained by the above five algorithms over 30 runs on each of the three practical problems, and the best one of each case is highlighted with bold font in gray shade. It is obvious that MultiGPO shows a significant superiority over the others in terms of the three metrics, with MultiGPO obtaining the average ranks of 1, 1.33, and 1 on problems FCDP, WRP, and CSIP, respectively.

To promote a better understanding of the algorithms' performances, Figs. S-1 and S-2 shown in the supplementary material and Fig. 9 here provide visual comparisons. Note that the PF approximations presented are in each case from the run with the median HV values, and the objective space of each instance is normalized by using the maximal and minimal objective values obtained from  $P^*$ , in the cases of  $M > 3$ . Overall, MultiGPO maintains its population distributions uniformly, although it loses some small subregions of the PF in problem FCDP. As visible in Figs. S-1 and S-2 in the supplementary material, 1by1EA and NSGA-II/SDR attain very

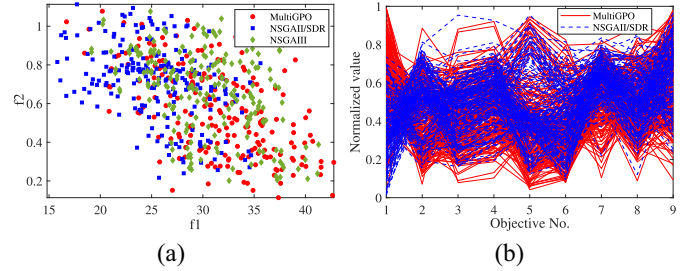


Fig. 9. Comparison of MultiGPO, NSGAII/SDR, and NSGA-III on the Car Side-Impact problem. (a) PF approximations obtained by the three algorithms are projected into a biobjective space of f1 and f2. (b) Complete normalized 9-objective space of MultiGPO and NSGAII/SDR.

poor distributions. In Fig. S-1(f) in the supplementary material, we also provide the result of MultiGPO with  $\varphi = 0$  (denoted as MultiGPO\*), in which the same original Pareto-dominance criterion is used for solution ranking as that used in NSGA-III. MultiGPO\* generates a better-distributed PF than MultiGPO and achieves the average performance metrics: IGD =  $1.6734e^{-1}$ , HV =  $1.036e^{+0}$ , and PD =  $4.449e^{+5}$ . This shows that the diversity preservation scheme based on the max-min distance works better than the reference-point-based strategy of NSGA-III for the FCDP problem.

Moreover, Fig. 9(a) depicts the solution sets of MultiGPO, NSGA-II/SDR, and NSGA-III (the top three methods based on the average ranking) that projected into a biobjective space considering objective f1 and f2; while in Fig. 9(b), the complete normalized 9-objective spaces of MultiGPO and NSGA-II/SDR are presented. In this problem, the ranges of f1 and f2 in  $P^*$  are [15.5851, 42.7680] and [0, 1.3679], respectively. The solutions of NSGA-II/SDR are located in the region of smaller values of f1, but it cannot obtain very low f2 values, as seen from Fig. 9(a). Also, the distribution of solutions obtained by NSGA-III is a little better than that by NSGA-II/SDR; however, both diversity and convergence properties are worse than those of MultiGPO. It is obvious from Fig. 9(b) that MultiGPO obtains lower values of objectives f2–f9, while NSGA-II/SDR gets lower f1 values only at the sacrifice of losing its population diversity. To test the sensitivity of  $\varphi$  ( $\varphi = 9$  for CSIP), we also perform the proposed algorithm with  $\varphi = 18$  (denoted by MultiGPO2) and  $\varphi = 27$  (denoted by MultiGPO3), respectively, on the CSIP problem. MultiGPO2



obtains the mean-metric values as  $\{IGD = 1.2784e^{+0}, HV = 2.0053e^{-1}, PD = 6.0846e^{+0}\}$  and MultiGPO3 reports the mean values as  $\{IGD = 1.2765e^{+0}, HV = 2.0175e^{-1}, PD = 5.9507e^{+0}\}$ , which are very close to the result of MultiGPO in Table III. Overall, the above results verify the effectiveness of the proposed MultiGPO in addressing practical many-objective applications.

## VI. CONCLUSION

In this article, a new, generalized-Pareto-based evolutionary algorithm is proposed for many-objective optimization. The proposed algorithm uses an “ $(M - 1) + 1$ ” framework of GPO dominance [i.e.,  $(M - 1)$ -GPD] for solution ranking, which can be simply implemented and efficiently performed to maintain a good balance of convergence and diversity. An EMO framework based on the proposed Pareto-based environmental selection, namely, MultiGPO, is shown to perform comparably to or better than the state of the art on a variety of scalable benchmark problems. We also applied the proposed algorithm on three real-world engineering problems with irregular PFs, and the results confirmed the superior performance of MultiGPO.

Since no reference vectors are employed, the performance of MultiGPO is less dependent on the PF shapes than methods using reference vectors, and is robust, especially in solving problems having irregular PFs. Nevertheless, the proposed  $(M - 1)$ -GPD scheme can still work collaboratively with reference-vector-based techniques, especially those using reference vector adaptation, to achieve further performance improvement on some problems. This deserves more attempt to fully explore the promising potential of  $(M - 1)$ -GPD.

## ACKNOWLEDGMENT

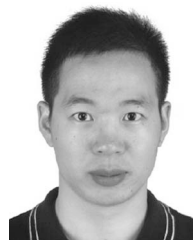
The authors would like to thank Prof. K. Deb for providing suggestions to this work.

## REFERENCES

- [1] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, “A fast and elitist multiobjective genetic algorithm: NSGA-II,” *IEEE Trans. Evol. Comput.*, vol. 6, no. 2, pp. 182–197, Apr. 2002.
- [2] S. Zhu and L. Xu, “Many-objective fuzzy centroids clustering algorithm for categorical data,” *Expert Syst. Appl.*, vol. 96, pp. 230–248, Apr. 2018.
- [3] Z. Lu, K. Deb, E. Goodman, W. Banzhaf, and V. N. Boddeti, “NSGANetV2: Evolutionary multi-objective surrogate-assisted neural architecture search,” in *Proc. Eur. Conf. Comput. Vis.*, 2020, pp. 35–51.
- [4] T. Santos and R. H. Takahashi, “On the performance degradation of dominance-based evolutionary algorithms in many-objective optimization,” *IEEE Trans. Evol. Comput.*, vol. 22, no. 1, pp. 19–31, Feb. 2018.
- [5] K. Deb and H. Jain, “An evolutionary many-objective optimization algorithm using reference-point-based nondominated sorting approach, part I: Solving problems with box constraints,” *IEEE Trans. Evol. Comput.*, vol. 18, no. 4, pp. 577–601, Aug. 2014.
- [6] L. S. Batista, F. Campelo, F. G. Guimarães, and J. A. Ramírez, “Pareto cone  $\epsilon$ -dominance: Improving convergence and diversity in multiobjective evolutionary algorithms,” in *Proc. Int. Conf. Evol. Multi Crit. Optimiz.*, 2011, pp. 76–90.
- [7] H. Sato, H. E. Aguirre, and K. Tanaka, “Controlling dominance area of solutions and its impact on the performance of MOEAs,” in *Proc. ACM Int. Conf. Evol. Multi Crit. Optimiz.*, 2007, pp. 5–20.
- [8] J. Liu, Y. Wang, X. Wang, S. Guo, and X. Sui, “A new dominance method based on expanding dominated area for many-objective optimization,” *Int. J. Pattern Recognit. Artif. Intell.*, vol. 33, no. 3, 2019, Art. no. 1959008.
- [9] C. Zhu, L. Xu, and E. D. Goodman, “Generalization of Pareto-optimality for many-objective evolutionary optimization,” *IEEE Trans. Evol. Comput.*, vol. 20, no. 2, pp. 299–315, Apr. 2016.
- [10] M. Li, S. Yang, and X. Liu, “Shift-based density estimation for Pareto-based algorithms in many-objective optimization,” *IEEE Trans. Evol. Comput.*, vol. 18, no. 3, pp. 348–365, Jun. 2014.
- [11] Y. Yuan, H. Xu, B. Wang, and X. Yao, “A new dominance relation-based evolutionary algorithm for many-objective optimization,” *IEEE Trans. Evol. Comput.*, vol. 20, no. 1, pp. 16–37, Feb. 2016.
- [12] L. Chen, H.-L. Liu, K. C. Tan, Y.-M. Cheung, and Y. Wang, “Evolutionary many-objective algorithm using decomposition-based dominance relationship,” *IEEE Trans. Cybern.*, vol. 49, no. 12, pp. 4129–4139, Dec. 2019.
- [13] Y. Tian, R. Cheng, X. Zhang, Y. Su, and Y. Jin, “A strengthened dominance relation considering convergence and diversity for evolutionary many-objective optimization,” *IEEE Trans. Evol. Comput.*, vol. 23, no. 2, pp. 331–345, Apr. 2019.
- [14] K. Deb, P. Zope, and A. Jain, “Distributed computing of Pareto-optimal solutions with evolutionary algorithms,” in *Proc. Int. Conf. Evol. Multi Crit. Optimiz.*, 2003, pp. 534–549.
- [15] H. Ishibuchi, Y. Setoguchi, H. Masuda, and Y. Nojima, “Performance of decomposition-based many-objective algorithms strongly depends on Pareto front shapes,” *IEEE Trans. Evol. Comput.*, vol. 21, no. 2, pp. 169–190, Apr. 2017.
- [16] M. Elarbi, S. Bechikh, C. A. Coello Coello, M. Makhoul, and L. B. Said, “Approximating complex Pareto fronts with pre-defined normal-boundary intersection directions,” *IEEE Trans. Evol. Comput.*, vol. 24, no. 5, pp. 809–823, Oct. 2020, doi: [10.1109/TEVC.2019.2958921](https://doi.org/10.1109/TEVC.2019.2958921).
- [17] H. Ishibuchi, T. Matsumoto, N. Masuyama, and Y. Nojima, “Many-objective problems are not always difficult for Pareto dominance-based evolutionary algorithms,” in *Proc. 24th Eur. Conf. Artif. Intell. (ECAI)*, 2020, pp. 291–298.
- [18] H. Ishibuchi, N. Akedo, and Y. Nojima, “Behavior of multiobjective evolutionary algorithms on many-objective knapsack problems,” *IEEE Trans. Evol. Comput.*, vol. 19, no. 2, pp. 264–283, Apr. 2015.
- [19] K. Li, K. Deb, Q. Zhang, and S. Kwong, “An evolutionary many-objective optimization algorithm based on dominance and decomposition,” *IEEE Trans. Evol. Comput.*, vol. 19, no. 5, pp. 694–716, Oct. 2015.
- [20] M. Asafuddoula, H. K. Singh, and T. Ray, “An enhanced decomposition-based evolutionary algorithm with adaptive reference vectors,” *IEEE Trans. Cybern.*, vol. 48, no. 8, pp. 2321–2334, Aug. 2018.
- [21] X. Cai, Z. Yang, Z. Fan, and Q. Zhang, “A decomposition-based many-objective evolutionary algorithm with two types of adjustments for direction vectors,” *IEEE Trans. Cybern.*, vol. 48, no. 8, pp. 2335–2348, Aug. 2018.
- [22] H.-L. Liu, L. Chen, Q. Zhang, and K. Deb, “Adaptively allocating search effort in challenging many-objective optimization problems,” *IEEE Trans. Evol. Comput.*, vol. 22, no. 3, pp. 433–448, Jun. 2018.
- [23] M. Wu, K. Li, S. Kwong, and Q. Zhang, “Evolutionary many-objective optimization based on adversarial decomposition,” *IEEE Trans. Cybern.*, vol. 50, no. 2, pp. 753–764, Feb. 2020.
- [24] Y. Zhou, Y. Xiang, Z. Chen, J. He, and J. Wang, “A scalar projection and angle-based evolutionary algorithm for many-objective optimization problems,” *IEEE Trans. Cybern.*, vol. 49, no. 6, pp. 2073–2084, Jun. 2019.
- [25] Y. H. Zhang *et al.*, “DECAL: Decomposition-based coevolutionary algorithm for many-objective optimization,” *IEEE Trans. Cybern.*, vol. 49, no. 1, pp. 27–41, Jan. 2019.
- [26] B. Li, K. Tang, J. Li, and X. Yao, “Stochastic ranking algorithm for many-objective optimization based on multiple indicators,” *IEEE Trans. Evol. Comput.*, vol. 20, no. 6, pp. 924–938, Dec. 2016.
- [27] Y. Tian, R. Cheng, X. Zhang, F. Cheng, and Y. Jin, “An indicator-based multiobjective evolutionary algorithm with reference point adaptation for better versatility,” *IEEE Trans. Evol. Comput.*, vol. 22, no. 4, pp. 609–622, Aug. 2018.
- [28] Y. Sun, G. G. Yen, and Z. Yi, “IGD indicator-based evolutionary algorithm for many-objective optimization problems,” *IEEE Trans. Evol. Comput.*, vol. 23, no. 2, pp. 173–187, Apr. 2019.
- [29] K. Shang and H. Ishibuchi, “A new hypervolume-based evolutionary algorithm for many-objective optimization,” *IEEE Trans. Evol. Comput.*, vol. 24, no. 5, pp. 839–852, Oct. 2020, doi: [10.1109/TEVC.2020.2964705](https://doi.org/10.1109/TEVC.2020.2964705).
- [30] Z. Liang, T. Luo, K. Hu, X. Ma, and Z. Zhu, “An indicator-based many-objective evolutionary algorithm with boundary protection,” *IEEE Trans. Cybern.*, early access, Jan. 13, 2020, doi: [10.1109/TCYB.2019.2960302](https://doi.org/10.1109/TCYB.2019.2960302).



- [31] Q. Lin *et al.*, "A clustering-based evolutionary algorithm for many-objective optimization problems," *IEEE Trans. Evol. Comput.*, vol. 23, no. 3, pp. 391–405, Jun. 2019.
- [32] H. Xu, W. Zeng, X. Zeng, and G. G. Yen, "An evolutionary algorithm based on Minkowski distance for many-objective optimization," *IEEE Trans. Cybern.*, vol. 49, no. 11, pp. 3968–3979, Nov. 2019.
- [33] Y. Xiang, Y. Zhou, X. Yang, and H. Huang, "A many-objective evolutionary algorithm with Pareto-adaptive reference points," *IEEE Trans. Evol. Comput.*, vol. 24, no. 1, pp. 99–113, Feb. 2020.
- [34] F. Gu and Y. M. Cheung, "Self-organizing map-based weight design for decomposition-based many-objective evolutionary algorithm," *IEEE Trans. Evol. Comput.*, vol. 22, no. 2, pp. 211–225, Apr. 2018.
- [35] X. He, Y. Zhou, Z. Chen, and Q. Zhang, "Evolutionary many-objective optimization based on dynamical decomposition," *IEEE Trans. Evol. Comput.*, vol. 23, no. 3, pp. 361–375, Jun. 2019.
- [36] R. Cheng, Y. Jin, M. Olhofer, and B. Sendhoff, "A reference vector guided evolutionary algorithm for many-objective optimization," *IEEE Trans. Evol. Comput.*, vol. 20, no. 5, pp. 773–791, Oct. 2016.
- [37] Y. Sun, B. Xue, M. Zhang, and G. G. Yen, "A new two-stage evolutionary algorithm for many-objective optimization," *IEEE Trans. Evol. Comput.*, vol. 23, no. 5, pp. 748–761, Oct. 2019.
- [38] H. Ge *et al.*, "A many-objective evolutionary algorithm with two interacting processes: Cascade clustering and reference point incremental learning," *IEEE Trans. Evol. Comput.*, vol. 23, no. 4, pp. 572–586, Aug. 2019.
- [39] J. Yi, J. Bai, H. He, J. Peng, and D. Tang, "Ar-MOEA: A novel preference-based dominance relation for evolutionary multiobjective optimization," *IEEE Trans. Evol. Comput.*, vol. 23, no. 5, pp. 788–802, Oct. 2019.
- [40] 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 Trans. Evol. Comput.*, vol. 18, no. 4, pp. 602–622, Aug. 2014.
- [41] Y. Qi, X. Ma, F. Liu, L. Jiao, J. Sun, and J. Wu, "MOEA/D with adaptive weight adjustment," *Evol. Comput.*, vol. 22, no. 2, pp. 231–264, 2014.
- [42] E. Zitzler, L. Thiele, M. Laumanns, C. Fonseca, and V. G. Da Fonseca, "Performance assessment of multiobjective optimizers: An analysis and review," *IEEE Trans. Evol. Comput.*, vol. 7, no. 2, pp. 117–132, Apr. 2003.
- [43] T. Pamulapati, R. Mallipeddi, and P. N. Suganthan, " $I_{SDE}^+$ —An indicator for multi and many-objective optimization," *IEEE Trans. Evol. Comput.*, vol. 23, no. 2, pp. 346–352, Apr. 2019.
- [44] Q. Lin *et al.*, "Particle swarm optimization with a balanceable fitness estimation for many-objective optimization problems," *IEEE Trans. Evol. Comput.*, vol. 22, no. 1, pp. 32–46, Feb. 2018.
- [45] H. Xu, W. Zeng, X. Zeng, and G. G. Yen, "A polar-metric-based evolutionary algorithm," *IEEE Trans. Cybern.*, early access, Feb. 4, 2020, doi: [10.1109/TCYB.2020.2965230](https://doi.org/10.1109/TCYB.2020.2965230).
- [46] Y. Tian, X. Zhang, R. Cheng, C. He, and Y. Jin, "Guiding evolutionary multiobjective optimization with generic front modeling," *IEEE Trans. Cybern.*, vol. 50, no. 3, pp. 1106–1119, Mar. 2020.
- [47] Y. Liu, D. Gong, J. Sun, and Y. Jin, "A many-objective evolutionary algorithm using a one-by-one selection strategy," *IEEE Trans. Cybern.*, vol. 47, no. 9, pp. 2689–2702, Sep. 2017.
- [48] Z. Liang, K. Hu, X. Ma, and Z. Zhu, "A many-objective evolutionary algorithm based on a two-round selection strategy," *IEEE Trans. Cybern.*, early access, Jun. 7, 2019, doi: [10.1109/TCYB.2019.2918087](https://doi.org/10.1109/TCYB.2019.2918087).
- [49] C. Liu, Q. Zhao, B. Yan, S. Elsayed, T. Ray, and R. Sarker, "Adaptive sorting-based evolutionary algorithm for many-objective optimization," *IEEE Trans. Evol. Comput.*, vol. 23, no. 2, pp. 247–257, Apr. 2019.
- [50] D. Brockhoff and E. Zitzler, "Objective reduction in evolutionary multiobjective optimization: Theory and applications," *Evol. Comput.*, vol. 17, no. 2, pp. 135–166, 2009.
- [51] Y. Yuan, Y.-S. Ong, A. Gupta, and H. Xu, "Objective reduction in many-objective optimization: Evolutionary multiobjective approaches and comprehensive analysis," *IEEE Trans. Evol. Comput.*, vol. 22, no. 2, pp. 189–210, Apr. 2018.
- [52] K. Zhang, Z. Xu, S. Xie, and G. G. Yen, "Evolution strategy-based many-objective evolutionary algorithm through vector equilibrium," *IEEE Trans. Cybern.*, early access, Jan. 10, 2020, doi: [10.1109/TCYB.2019.2960039](https://doi.org/10.1109/TCYB.2019.2960039).
- [53] Y. Xiang, Z. Yuren, M. Li, and Z. Chen, "A vector angle based evolutionary algorithm for unconstrained many-objective optimization," *IEEE Trans. Evol. Comput.*, vol. 21, no. 1, pp. 131–152, Feb. 2017.
- [54] H. K. Singh, K. S. Bhattacharjee, and T. Ray, "Distance-based subset selection for benchmarking in evolutionary multi/many-objective optimization," *IEEE Trans. Evol. Comput.*, vol. 23, no. 5, pp. 904–912, Oct. 2019.
- [55] H. Wang, Y. Jin, and X. Yao, "Diversity assessment in many-objective optimization," *IEEE Trans. Cybern.*, vol. 47, no. 6, pp. 1510–1522, Jun. 2017.
- [56] X.-F. Liu, Z.-H. Zhan, Y. Gao, J. Zhang, S. Kwong, and J. Zhang, "Coevolutionary particle swarm optimization with bottleneck objective learning strategy for many-objective optimization," *IEEE Trans. Evol. Comput.*, vol. 23, no. 4, pp. 587–602, Aug. 2019.
- [57] H. Chen, Y. Tian, W. Pedrycz, G. Wu, R. Wang, and L. Wang, "Hyperplane assisted evolutionary algorithm for many-objective optimization problems," *IEEE Trans. Cybern.*, vol. 50, no. 7, pp. 3367–3380, Jul. 2020.
- [58] S. Huband, P. Hingston, L. Barone, and L. While, "A review of multiobjective test problems and a scalable test problem toolkit," *IEEE Trans. Evol. Comput.*, vol. 10, no. 5, pp. 477–506, Oct. 2006.
- [59] R. Cheng *et al.*, "A benchmark test suite for evolutionary many-objective optimization," *Complex Intell. Syst.*, vol. 3, no. 1, pp. 67–81, 2017.
- [60] I. Das and J. E. Dennis, "Normal-boundary intersection: A new method for generating the Pareto surface in nonlinear multicriteria optimization problems," *SIAM J. Optim.*, vol. 8, no. 3, pp. 631–657, 1998.
- [61] Y. Tian, R. Cheng, X. Zhang, and Y. Jin, "PlatEMO: A MATLAB platform for evolutionary multi-objective optimization [educational forum]," *IEEE Comput. Intell. Mag.*, vol. 12, no. 4, pp. 73–87, Nov. 2017.
- [62] K. Deb, L. Thiele, M. Laumanns, and E. Zitzler, "Scalable test problems for evolutionary multiobjective optimization," in *Evolutionary Multiobjective Optimization*. Heidelberg, Germany: Springer, 2005, pp. 105–145.
- [63] Z. Zhou and S. Zhu, "Kernel-based multiobjective clustering algorithm with automatic attribute weighting," *Soft Comput.*, vol. 22, no. 11, pp. 3685–3709, 2018.
- [64] S. Zhu, L. Xu, and E. D. Goodman, "Evolutionary multi-objective automatic clustering enhanced with quality metrics and ensemble strategy," *Knowl. Based Syst.*, vol. 188, pp. 1–21, Jan. 2020.
- [65] W. Deng, H. Liu, J. Xu, H. Zhao, and Y. Song, "An improved quantum-inspired differential evolution algorithm for deep belief network," *IEEE Trans. Instrum. Meas.*, vol. 69, no. 10, pp. 7319–7327, Oct. 2020, doi: [10.1109/TIM.2020.2983233](https://doi.org/10.1109/TIM.2020.2983233).
- [66] Z. Lu *et al.*, "Multi-objective evolutionary design of deep convolutional neural networks for image classification," *IEEE Trans. Evol. Comput.*, early access, Sep. 21, 2020, doi: [10.1109/TEVC.2020.3024708](https://doi.org/10.1109/TEVC.2020.3024708).
- [67] Z. Lu, K. Deb, and V. N. Boddeti, "MUXConv: Information multiplexing in convolutional neural networks," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2020, pp. 12044–12053.
- [68] E. Q. Wu, G.-R. Zhou, L.-M. Zhu, C.-F. Wei, H. Ren, and R. S. Sheng, "Rotated sphere haar wavelet and deep contractive auto-encoder network with fuzzy Gaussian svm for pilot's pupil center detection," *IEEE Trans. Cybern.*, vol. 51, no. 1, pp. 332–345, Jan. 2021.
- [69] E. Q. Wu *et al.*, "Nonparametric Bayesian prior inducing deep network for automatic detection of cognitive status," *IEEE Trans. Cybern.*, early access, Mar. 20, 2020, doi: [10.1109/TCYB.2020.2977267](https://doi.org/10.1109/TCYB.2020.2977267).
- [70] X. Liao, Q. Li, X. Yang, W. Zhang, and W. Li, "Multiobjective optimization for crash safety design of vehicles using stepwise regression model," *Struct. Multidiscipl. Optimiz.*, vol. 35, no. 6, pp. 561–569, 2008.



**Shuwei Zhu** (Graduate Student Member, IEEE) received the M.S. degree in control science and engineering from Jiangnan University, Wuxi, China, in 2016. He is currently pursuing the Ph.D. degree in control science and engineering with Tongji University, Shanghai, China.

He has been awarded a scholarship from the China Scholarship Council (CSC) to pursue joint Ph.D. study in Michigan State University, East Lansing, MI, USA, from 2019 to 2020. His research interests focus on clustering algorithms, evolutionary multiobjective optimization, and machine learning.



**Lihong Xu** (Senior Member, IEEE) received the Ph.D. degree in control theory and control engineering from Southeast University, Nanjing, China, in 1991.

He is currently a Full Professor with the Department of Electronics and Information Engineering, Tongji University, Shanghai, China. He is currently conducting joint research work as a Visiting Professor and an Advisor with the Greenhouse Research Team, BEACON, Michigan State University, East Lansing, MI, USA. His

research fields include control theory, computational intelligence, and optimization theory.

Dr. Xu received a National Science and Technology Progress Award (Second Prize) of China in 2007. He is the President of IEEE CIS's Shanghai Chapter.



**Zhichao Lu** (Graduate Student Member, IEEE) received the B.Sc. and Ph.D. degrees in electrical and computer engineering from Michigan State University, East Lansing, MI, USA, in 2013 and 2020, respectively.

He is currently a Postdoctoral Research Fellow with the Department of Computer Science and Engineering, Southern University of Science and Technology, Shenzhen, China. His research interests are in the field of evolutionary machine learning, notably machine-learning-assisted evolutionary algo-

rithms, automated machine learning, and, in particular, evolutionary neural architecture search.

Dr. Lu received the Best Paper Award at GECCO 2019 under the evolutionary machine learning track.



**Erik D. Goodman** received the B.S. degree in mathematics and the M.S. degree in system science from Michigan State University (MSU), East Lansing, MI, USA, in 1966 and 1968, respectively, and the Ph.D. degree in computer and communication sciences from the University of Michigan at Ann Arbor, Ann Arbor, MI, USA, in 1972.

He was the Director of the BEACON Center for the Study of Evolution in Action, an NSF Science and Technology Center, MSU, from 2010 to 2019.

He has been a Professor with the Department of

Electrical and Computer Engineering, MSU, since 1984, where he also co-directs the Genetic Algorithms Research and Applications Group.

Dr. Goodman is a Senior Fellow of the International Society for Genetic and Evolutionary Computation. He was the Chair of the Executive Board of that society from 2001 to 2004. From 2005 to 2007, he was the Founding Chair of ACM SIG for Genetic and Evolutionary Computation.