A Multi-objective Evolutionary Algorithm for Finding Knee Regions Using Two Localized Dominance Relationships

Guo Yu, *Student Member, IEEE,* Yaochu Jin, *Fellow, IEEE,* and Markus Olhofer

***Abstract*—In preference based optimization, knee points are considered the naturally preferred trade-off solutions, especially when the decision-maker has little *a priori* knowledge about the problem to be solved. However, identifying all convex knee regions of a Pareto front remains extremely challenging, in particular in a high-dimensional objective space. This paper presents a new evolutionary multi-objective algorithm for locating knee regions using two localized dominance relationships. In the environmental selection, the α-dominance is applied to each subpopulation partitioned by a set of predefined reference vectors, thereby guiding the search towards different potential knee regions while removing possible dominance resistant solutions. A knee-oriented dominance measure making use of the extreme points is then proposed to detect knee solutions in convex knee regions and discard solutions in concave knee regions. Our experimental results demonstrate that the proposed algorithm outperforms the state-of-the-art knee identification algorithms on a majority of multi-objective optimization test problems having up to eight objectives and a hybrid electric vehicle controller design problem with seven objectives.**

***Index Terms*—Multi-objective evolutionary optimization, knees, knee-oriented dominance, α-dominance, preference**

I. Introduction

MANY real-world optimization problems have multiple conflicting objectives, to which a set of Pareto optimal solutions will be found [1]. Without loss of generality, a multi-objective optimization problem (MOP) can be formulated as the following m-objective minimization problem:

$$\text{minimize } F(x) = (f_1(x), \cdots, f_m(x))^T, \tag{1}$$

where $x = (x_1, \cdots, x_n) \in \Omega$ is the decision vector. $x_i^L \leq x_i \leq x_i^U, i = 1, \cdots, n$, where x_i^L and x_i^U are the lower and upper bounds, respectively, of the i-th decision variable. $\Omega \subseteq \mathbb{R}^n$ is the decision space, and n is the number of decision variables. $F : \Omega \rightarrow \mathbb{R}^m$ consists of m objectives. When m is larger than three, the MOP is also known as a many-objective optimization problem (MaOP).

Recent decades have witnessed a great success in developing multi-objective evolutionary algorithms (MOEAs) for solving MOPs [2]. Over the past a few years, research on MOEAs has focused on solving MaOPs [3], [4], mainly due to the deteriorated performance of the MOEAs designated for solving bi- or three-objective optimization problems. Although considerable progress has been made in finding a set of diverse and well converged trade-off solutions in dealing with MaOPs, an implicit hypothesis made is that the obtained set of solutions, which is typically small (e.g., up to a few hundreds), is able to represent the entire Pareto optimal front (PoF) of an MaOP. Unfortunately, this hypothesis usually does not hold, in particular when the number of objectives is large [5].

In practice, the DM may be interested in only a few subregions of the PoF instead of the whole PoF, e.g., the central part of PoF, the boundary regions, the extreme regions. If user preferences are available, we can use them to guide the search towards the regions of interest (ROIs) [6], [7], thereby making it easier for the DM to select a small number of solutions for final implementation [8]. For the above reasons, preference based evolutionary optimization algorithms have attracted much research interest in the past decades [9], [10].

When user preferences are not available, knee points are considered as the preferred solutions, since they need a large compromise in at least one objective to gain a small improvement in other objectives [11], [12]. Besides, knee solutions are often prioritized in many MOEAs since they usually contribute to a large hypervolume [13]. Many algorithms have been proposed by taking advantage of knee solutions to more efficiently solve MaOPs [13], [14] or dynamic optimization problems [15]. Knee solution based MOEAs have already found successful applications in solving real-world problems, such as self-adaptive software [16], sparse reconstruction [17], and driving strategy for electric vehicles [18].

In consequence, several *a posterioi* methods have been proposed to characterize the knee points among a set of non-dominated solutions. Das et al. [19], [20] suggested to identify the knee points with the "maximum bulge" on the Pareto front using the normal boundary intersection [21]. Branke et al. [22] took advantage of the expected marginal utility (EMU) to locate the knee regions. The niching method [23] defines possible knee points in convex and concave regions based on the density of the solution distribution. Other methods have also been reported for identifying the knees of two-objective problems, such as the reflex/bend angle based approaches [22], [24], and its variant, the (α, β)-approach [12].

Notably, there is an assumption in all *a posteriori* approaches that a large set of well distributed and well converged solutions is available. However, it is computationally expensive to achieve such a solution set, especially for MaOPs. There-

Guo Yu and Yaochu Jin are with the Department of Computer Science, University of Surrey, Guildford, Surrey GU2 7XH, UK. e-mail: gysearch@163.com, yaochu.jin@surrey.ac.uk. (*Corresponding author: Yaochu Jin*)

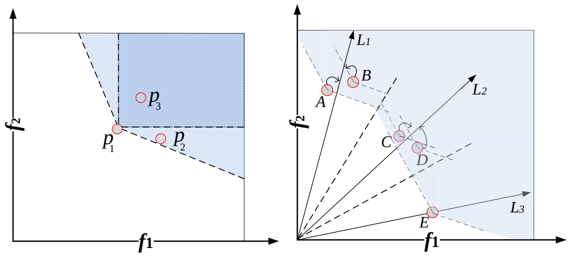
Markus Olhofer is with Honda Research Institute Europe GmbH, Carl-Leigien-Strasse 30, D-63073 offenbach/Main, Germany. e-mail: markus.olhofer@honda-ri.de.

fore, *a priori* approaches to the search of knee regions are popular. For example, in [25], [26], methods for characterizing knee points with the max-min utility value is incorporated in the environmental selection. However, the boundary points usually have a larger utility value than other solutions and will most likely be kept in the environmental selection, which may mislead the search process. Zhang et al. developed a selection method for solving MaOPs by prioritizing knee points [13] identified based on the extreme points [19]. However, the extreme solutions may become dominance resistant solutions (DRSs) [27] that will seriously slow down the convergence of the population. In [27], the DRSs are defined as the solutions which are extremely inferior in at least one objective and there exist very few solutionse that are able to dominate them. The work in [28] recursively uses the EMU [22] to locate the most promising knee candidates during the environmental selection. However, this approach also favors the boundary points and DRSs in the search, which degrades the convergence performance. The angle-based pruning strategy [29] was adopted to detect the knee regions [30] in the environmental selection, although the issue of DRSs remains. Most recently, we introduced an α-dominance to eliminate the DRSs for the search of knee solutions [31]. However, uninterested solutions like the boundary points and solutions from the concave knee regions will still be selected. Too much convergence pressure of the modified dominance may degrade the diversity of the solutions by eliminating the knee candidates in a potential knee regions, so that less knee regions will be finally reserved.

A common issue as found in the above discussions is that some particular solutions, such as some extreme and boundary solutions, are detrimental to the effective search of knee solutions. To address this issue, this work firstly introduces a set of reference vectors to partition the objective space into a number of subregions. Then, the α-dominance [27] is applied separately in each subregion to find the potential knee regions and to remove dominance resistant solutions, thereby guiding the search towards multiple potential knee regions. Afterwards, a knee-oriented dominance is proposed to identify the knee solutions in each potential knee regions and eliminate the boundary solutions as well as solutions in the concave knee regions. However, boundary and extreme solutions may be of interest to the decision-maker (DM). In practice, we can store the boundary and extreme solutions in a separate archive. After the environmental selection, the reference vectors will be updated according to the number of associated solutions. With the help of the localized α-dominance together with the knee-oriented dominance, the proposed algorithm is able to efficiently locate knee regions and knee solutions.

The main contributions of this paper are summarized as follows:

1) A localized α-dominance based sorting is designed to identify potential knee regions and get rid of the DRSs.
2) A knee-oriented dominance measure making use of the extreme solutions is proposed to accurately locate the knee solutions and eliminate the boundary solutions, and solutions in the concave knee regions.
3) A framework for detecting knee regions and knee solutions is developed by embedding the localized α-dominance and knee-oriented dominance in the environmental selection. The proposed framework is compared with a few state-of-the-art methods on a set of widely used test problems to show its superior performance in both convergence and accuracy in detecting knee regions and knee solutions.



(a) (b)

Fig. 1. (a) An illustrative example of Pareto dominance and α-dominance, where p_1 dominates p_3 in terms of both Pareto dominance and α-dominance, while p_1 and p_2 are non-dominated with each other in terms of Pareto dominance. (b) An example of localized α-dominance, where the objective space is divided into three subspaces by a set of reference vectors (L_1, L_2, L_3). As a result, solutions A and B are associated with L_1, C and D with L_2, and E with L_3. According to the conventional α-dominance, A and E are non-dominated and the rest are dominated. According to the localized α-dominance, however, A, C, D, and E are non-dominated with each other, but B is dominated by A.

The rest of the paper is organized as follows. Section II introduces the related definitions and dominance relationships, based on which a new knee-oriented dominance relationship is proposed. A new environmental selection strategy is then suggested in Section III, in which the population is first sorted based on the localized α-dominance and further locally sorted according to the knee-oriented dominance before the environmental selection. Section IV presents the sensitivity analysis and experimental results, together with a discussion of the comparative results. Section VI concludes the paper.

II. A KNEE-ORIENTED DOMINANCE RELATIONSHIP

In this section, the definitions of the related dominance relationships and knee points are introduced, before we present the new knee-oriented dominance relationship proposed in this work. All discussions are based on minimization problems as defined in 1.

A. Related definitions

Definition 1 (*Pareto dominance*) Given two solutions $x_1, x_2 \in \Omega$, x_1 is said to Pareto dominate x_2, denoted by $x_1 \prec x_2$, if and only if the following equation is satisfied:

$$\begin{aligned}&\forall i \in \{1,2,\ldots,m\}, f_i(x_1)\\&\quad \leq f_i(x_2) \wedge \exists j \in \{1,2,\ldots,m\} : f_j(x_1) < f_j(x_2).\end{aligned} \tag{2}$$

The PoF is composed of all the Pareto optimal solutions in the objective space and the collection of Pareto optimal solutions in the decision space is denoted by the Pareto optimal set (PoS). For example in Fig. 1(a), $p_1 \prec p_3$, but p_1 and p_2 are non-dominated.

Definition 2 (*α-dominance* [27]) A solution x is said to α-dominate solution y, denoted by $x \prec_\alpha y$, if the following condition holds:

$$\begin{aligned} &\forall i \in \{1, 2, \dots, m\}, g_i(x, y) \le 0 \quad \wedge \\ &\exists j \in \{1, 2, \dots, m\}, g_j(x, y) < 0, \end{aligned} \tag{3}$$

where $g_i(x, y) = f_i(x) - f_i(y) + \sum_{j \ne i}^{m} \alpha_{ij}(f_j(x) - f_j(y))$, and α_{ij} is the predefined bound of the trade-off rates.

We can see from the above definition, α-dominance makes the Pareto dominance relationship stronger, when $\alpha > 0$. In Fig. 1(a) for instance, $p_1 \prec_\alpha p_2$ and $p_1 \prec_\alpha p_3$, although p_1 and p_2 are non-dominated according to the Pareto dominance.

Definition 3 (Localized *α-dominance* [31]) A solution x is said to localized α-dominate solution y, if the following condition holds:

$$\begin{aligned} &I(x) = I(y) \quad \wedge \\ &\forall i \in \{1, 2, \dots, m\}, g_i(x, y) \le 0 \quad \wedge \\ &\exists j \in \{1, 2, \dots, m\}, g_j(x, y) < 0, \end{aligned} \tag{4}$$

where $g_i(x, y) = f_i(x) - f_i(y) + \sum_{j \ne i}^{m} \alpha_{ij}(f_j(x) - f_j(y))$, and α_{ij} is the predefined bound of trade-off rates. For example in [31], α_{ij} is recommended to set as 0.75 for knee identification. I is the index of a reference vector, where $I(x) = I(y)$ means that x and y are associated with the same reference vector.

Fig. 1(b) illustrates how the localized α-dominance can change the dominance relationship. If the conventional α-dominance relationship is applied to sort the five solutions in the plot, then A and E are in the first frontier, and the rest are all in the second frontier. By contrast, A, C, D and E are non-dominated according to the localized α-dominance and will be in the first frontier, while B will be in the second frontier.

Definition 4 (*Knee points* [19]) A knee (k) is defined to be the one having the maximum distance from the convex hull of individual minima (CHIM) to the hyperplane S constructed by the extreme points.

$$k = \arg\max_p (d(p, S)) \tag{5}$$

where p is a solution on the PoF. $d(F(p), S)$ denotes the distance from solution p to the hyperplane $S : f_1 + \dots + f_m = 1$ in a normalized coordinator system.

In the above definition, an extreme point x in the i-th objective can be described as follows for a given population P:

$\forall y \in P, \exists i \in \{1, \dots, m\}, x = arg \max_x f_i(y) \wedge \forall j \in \{1, \dots, i-1, i+1, \dots, m\}, f_j(x) = \min f_j(y)$.

In Fig. 2, point B is the knee point on the PoF, which has the largest distance to the hyperplane constructed by the extreme points (A and C). Additional definitions of knee points can be found in [12], [32].

B. Motivation

Most existing *a priori* knee search methods introduce a secondary criterion, such as the max-min utility [26], expected marginal utility [28], angle-based pruning [30], and distance to the hyperplane [13] into the environmental selection to guide the search towards the potential knee regions. It has also been found that a selection method favoring knee points can enhance the convergence because the knee candidates are shown to be able to contribute to the hypervolume more than other solutions [13].

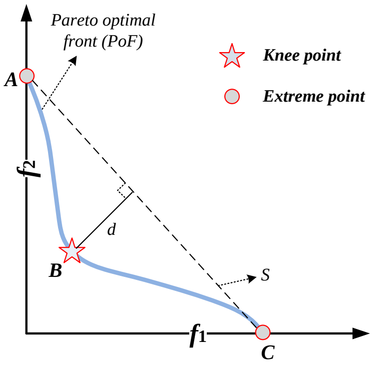


Fig. 2. An illustrative example of the knee point (B) of a PoF. Solutions A and C are the extreme points.

It should be noted that most existing *a priori* knee search methods also favor the boundary points or extreme points during the search. These solutions, however, may easily become the DRSs in the environmental selection, seriously degrading the convergence of the population and misleading the search process. Besides, even though a modified dominance measure is introduced to deal with the DRSs in the optimization [31], some undesired solutions such as the boundary solutions or the solutions in the concave knee regions cannot be eliminated in the selection, which slows down the convergence. On the contrary, some potential knee regions may get lost during the search if a modified dominance relationship results in an overly large selection pressure.

Therefore, this work aims to design a selection mechanism that is able to get rid of the DRSs, boundary solutions, and solutions in the concave knee regions, while properly guiding the population towards all knee regions of the PoF, and finally detecting as many knee points as possible.

C. Proposed knee-oriented dominance relationship

In knee solution detection, it is essential to locate potential knee regions before the knee solutions can be identified. In this section, we introduce a new dominance relationship, called *knee-oriented dominance*, that favors solutions in potential knee regions in environmental selection.

Given two solutions A and B from CHIM, A is said to knee-oriented-dominate solution B if the following conditions are satisfied.

$$\begin{cases} \mu(A, B) < 0, \\ \text{subject to:} \\ \quad \mu(A, B) = \langle \overrightarrow{N_{id}A}, \overrightarrow{AB} \rangle - \tau \cdot (\max\limits_{i=1,\dots,m}\{\delta_i(A)\} + \min\limits_{i=1,\dots,m}\{\delta_i(A)\}), \\ \quad \delta_i(A) = arc\tan(\frac{\sqrt{\sum_{j=1, j \ne i}^{m}(f_j(A) - f_j(N_{id}))^2}}{|f_i(A) - \max f_i(E) - \epsilon|}), \end{cases} \tag{6}$$

where $\mu(A, B) < 0$ means solution A knee-oriented-dominates B, $\delta_i(A)$ is an acute angle determined by the i-

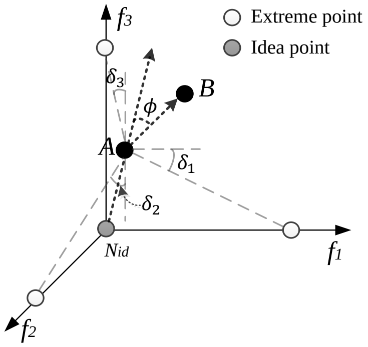


Fig. 3. An illustrative example of the knee-oriented dominance relationship, where ϕ is the acute angle between $\overrightarrow{N_{id}A}$ and $\overrightarrow{AB}$, denoted by $\langle \overrightarrow{N_{id}A}, \overrightarrow{AB} \rangle$. Here, $\delta_3 = \min_{i=1,\ldots,m} \{\delta_i(A)\}$ and $\delta_2 = \max_{i=1,\ldots,m} \{\delta_i(A)\}$.

Algorithm 1 : Overall framework of LBD-MOEA

```
Input: Population Size: $n$, termination condition: $\mathcal{T}$, extreme point set: $E$, Number of reference vectors: $N$
Output: Population: $P = \{x_1, x_2, \cdots, x_n\}$
$P = Initialization\,(n)$
$Evaluation\,(P)$
$W = Reference - Vector - Generator(N)$
$UpdateExe(E, P)$ //*Initialize the extreme points.*//
while $\neg\mathcal{T}$ do
    $Q = MatingSelection\,(P)$
    $Q = Crossover\,(Q)$
    $Q = Mutation\,(Q)$
    $Evaluation\,(Q)$
    $R = P \cup Q$
    $(R_I, R_C) = Association(W, R)$
    $UpdateExe(E, R)$ //* update extreme points.*//
    $P = BiEnvironmentalSelection\,(R, n, E, R_I)$
    $UpdateRef(W, R_C)$ //*Update reference vectors.*//
end while
$Output(P)$
```

th objective value of solution A, the ideal point N_{id}. The ideal point is defined by $f_j(N_{id}) = \min f_j(E) - \epsilon$, where $E = \{E^i | i = 1, 2, ..., m\}$, is the set of extreme points, and ϵ is a small positive constant [1] to ensure the denominator is not equal to zero. In the above equation, $\tau \in [1/2, 1]$[2]is a parameter controlling the size of the knee region to be achieved. Fig. 3 provides an example of the knee-oriented dominance relationship between two solutions, where $\delta_3 = \min_{i=1,\ldots,m} \{\delta_i(A)\}$ and $\delta_2 = \max_{i=1,\ldots,m} \{\delta_i(A)\}$. A knee-oriented-dominates B, provided that ϕ is smaller than $\tau \cdot (\delta_2 + \delta_3)$.

In this work, we use $\max_{i=1,\ldots,m}\{\delta_i(A)\} + \min_{i=1,\ldots,m}\{\delta_i(A)\}$ to indicate the size of the area solution A dominates with the help of the extreme points. The corresponding proof is presented in Section I of the Supplementary material. $\mu(A, B) < 0$ in Eq. 6 means that solution A knee-oriented-dominates B, if the sum of these two angles is larger than the acute angle $\langle \overrightarrow{N_{id}A}, \overrightarrow{AB} \rangle$ (when $\tau = 1$). The reason to choose these two angles is that different knee regions may have different curvatures and different solutions in the same knee region can have different values of $\max_{i=1,\ldots,m}\{\delta_i(A)\}$ and $\min_{i=1,\ldots,m}\{\delta_i(A)\}$. In this work, we adopt the maximum and minimum of the m angles to roughly characterize how big the region should a solution knee-oriented-dominate. Fig. 4 (a) shows three solutions A, B and C and their dominated region. We can see that solutions A and B are non-knee-oriented-dominated from each other, while C is knee-oriented-dominated by both A and B. We can also see from Fig. 4 (a) that the farther a solution from the hyperplane is, the more likely it is a knee point, and the less likely such a solution will be dominated by other solutions. Here, we do not use the average of all $\delta_i, i = 1, \ldots, m$, simply because the average angle value may be less capable of capturing the differences of the solutions in different knee regions. As shown in Fig. 4 (b), the dominated area of a solution (shaded area) becomes larger when the solution moves closer to the hyperplane S.

The definition of knee-oriented-dominance in Eq. (6) and the discussions above assume that there is one knee region only. One potential issue with such global knee-oriented-dominance comparison is that solutions in a knee region can be knee-oriented-dominated by solutions in another knee region that have a larger degree of curvature, leading to the loss of knee solutions in the search process. For example, Fig. 5 shows three knee solutions, A, B and C. According to the definition of knee-oriented-dominance, solution B is knee-oriented-dominated by solution A, and actually, all solutions in the knee region in which solution B is located are knee-oriented-dominated by A. As a result, all solutions in the knee region of solution B will get lost during the search, which is not desirable. This issue can be resolved if the knee-oriented-dominance is applied for comparing solutions in a local region only. To this end, a set of reference vectors is adopted in this work in the environmental selection to partition the overall objective space into a number of subspaces and the knee-oriented dominance comparisons are restricted to each local subspace, thereby enabling the search towards multiple knee regions. Section III will detail how to group solutions before the knee-oriented non-dominated sorting is performed.

III. An MOEA Driven by Two Localized Dominance Relationships

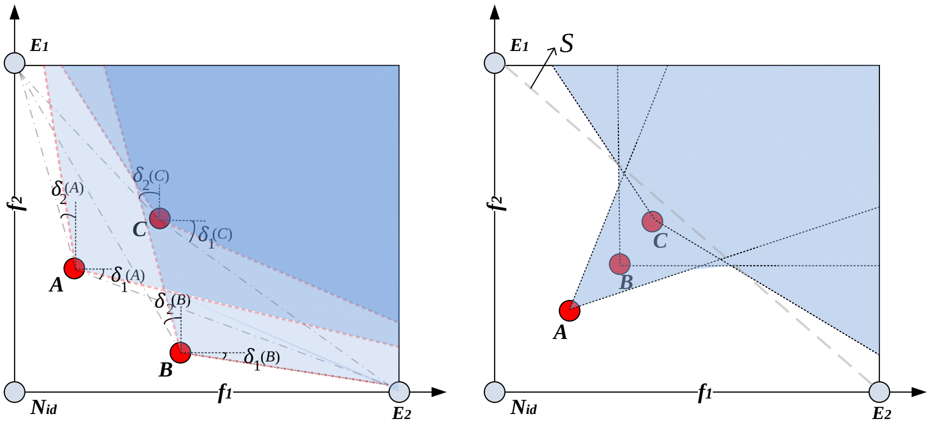
In this section, we firstly present the overall framework of the proposed localized bi-dominance driven MOEA, called LBD-MOEA, followed by a description of the details of its main components. Finally, an analysis of the computational complexity of the algorithm is given.

A. Overall framework

The overall framework of LBD-MOEA is presented in Algorithm 1. Firstly, the population P is initialized and

[1] ϵ = E-05.

[2] Section VI in the Supplementary materials provides a sensitive analysis and a self-adjusting strategy on the parameter τ in the knee-oriented dominance.



(a)

(b)

Fig. 4. (a) Illustration of three solutions and their dominated regions denoted by the shaded area. (b) It is shown that more closer a solution to the hyperplane is, the wider its dominated will become.

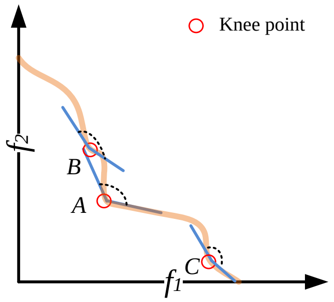


Fig. 5. An illustrative example showing the importance of patitioning the objective space into a number of subspaces in order to keep solutions in multiple knee regions. If the knee-oriented-dominance is used to compare solutions in the whole objective space, solution A will dominate all solutions in the knee region in which B is located.

evaluated, followed by the generation of a set of reference vectors in Line 3 and the initialization of the extreme points E_p in Line 4. A number of genetic operations, including mating selection, crossover, and mutation are then performed from Line 6 to Line 9 to generate an offspring population Q. After that, Q and P are merged into a combined population R. Then all individuals in R are associated with their closest reference vectors in Line 11. After that, the bi-dominance driven environmental selection described in Line 13 is applied on R to select the solutions to be passed to the next generation P. Finally, the reference vectors are updated in Line 14. The above steps (Lines 6 to 14) are repeated until the termination condition is satisfied.

The main components of LBD-MOEA include reference vector generation, update of the extreme solutions, objective partition, bi-dominance driven environmental selection, and the update of reference vectors. In the following, we present the details of each component.

B. Reference vector generation

The method for reference vector generation in NSGA-III [33] is adopted in this work, which is based on the normal-boundary intersection [21]. The number of the reference vectors (N) is determined as follows.

Given two predefined positive integers (H_1 and H_2) and the number of objectives (m),

$$N = \binom{H_1 + m - 1}{m - 1} + \binom{H_2 + m - 1}{m - 1}, \tag{7}$$

where H_1 and H_2 are introduced to equally divide the boundary layer and inner layer into H_1 and H_2 parts, respectively.

Suppose that a point $x = (x_1, \cdots, x_m)$ is generated satisfying the following condition:

$$\sum_{i=1}^{m} x_i = H, \quad x_i \in \mathbb{N}, \tag{8}$$

then the corresponding reference vector $v_i = (v_1, \cdots, v_m)$ can be calculated as follows:

$$v_{i,j} = \frac{x_i}{H}, \quad j = 1, 2, \ldots, m \tag{9}$$

where H is H_1 when reference vectors for the boundary layer are to be generated, and H is H_2 when reference vectors for the inner layer are generated.

C. Update of extreme points

The extreme points are important in knee-oriented dominance comparisons because they are used for calculating the angles in Eq. 6, which need to be constantly updated as the evolution proceeds.

In this work, the method for detection of extreme points introduced in [34] is adopted to update the extreme points. Given a solution set P, the extreme point set $E = \{E^1, \cdots, E^m\}$ is updated as follows:

$$E^i = \arg\min_{p \in P} \sqrt{\sum_{j=1, j\neq i}^{m} (f_j(p) - z_j^*)^2}, \tag{10}$$

where $z_i^* = \min_{p \in P} f_i(p)$, and $i = 1, \cdots, m$. m is the number of objectives.

In the Section V of the Supplementary material, this study also investigates the effect of the extreme points to the performance of the proposed algorithm. The investigation shows that the extreme points from the whole population play a much larger role in improving the performance of LBD-MOEA than the extreme points from the first front obtained by the localized α-dominance sorting.

D. Association

The association operator is to partition the objective space into a number of subregions, where each solution is associated with its closest reference vector. This work adopts the association method presented in [33], which is defined as follows:

$$R_I(x) = arg \min_{i=1,\cdots,N} \|F(x) - (N_{id} - d \cdot v_i)\| \tag{11}$$

$$R_C(v_i) = count(R_I(P) == i)$$

where

$$d = \frac{\|(N_{id} - F(x))^T v_i\|}{\|v_i\|}$$

where v_i is the i-th reference vector in the reference set and $i = 1, \ldots, N$. N_{id} is the ideal point. R_I records the indices of the reference vectors of solution x, and R_C is the number of solutions associated with each reference vector.

E. Bi-dominance driven environmental selection

The proposed bi-dominance driven environmental selection is detailed in Algorithm 2, which consists of two major steps. One is to sort the population using the localized α-dominance relationship (Line 3), and the other is to re-sort the solutions in the critical frontier resulting from the first step using the localized knee-oriented-dominance (Line 9). Refer to the next paragraph for a definition of the critical front.

In the first step in Algorithm 2, the population is divided into a number of sub-populations using a set of reference vectors. Each sub-population is sorted separately using the α-dominance so that each individual is assigned a front number. The sub-populations are then combined and divided into a number of fronts according to their front number (Line 3). Then, the solutions are selected front by front according to their front number in an ascending order (refer to Lines 5 – 6). The selection continues until it starts to select solutions from the critical front denoted by $L_\imath$. The critical front is defined as the last front from which only pat of its solutions will be selected, i.e., $|L_\imath| \wedge |P| > n$, where n is the population size, $L_\imath$ is the number of solutions in the critical front, and P is the number of solutions that have been selected so far (Line 9 of Algorithm 2).

The second step of Algorithm 2 is the knee-oriented dominance based selection (Line 9 of Algorithm 2). In this step, the algorithm is going to select $n - |P|$ solutions from the critical front, which becomes more important for identifying multiple knee regions when most solutions are on the critical front after the localized α-dominance based sorting. While the first step is mainly to drive the population towards the Pareto front, the second step is meant to select a set of solutions from each knee region close to the knee point, and discard boundary solutions and the solutions in concave regions. The details of knee-oriented dominance based selection are presented in the next subsection.

Algorithm 2 : *BiEnvironmentalSelection*

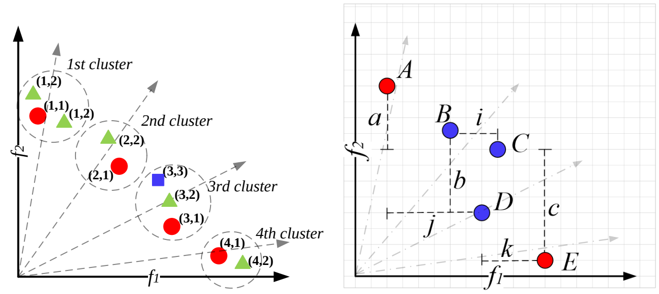
```
Input: Population: $R = \{x_1, x_2, \cdots, x_r\}$, output population
    size: $n \leq r$, extreme point set: $E$, indices of the reference
    vectors: $R_I$
Output: Population: $P = \{x_1, x_2, \cdots, x_n\}$
 1: $P = \emptyset$
 2: //* Do localized $\alpha$-dominance sorting.*//
 3: $\alpha\_nondominatedSorting(R, R_I) = \{L_1, L_2, \cdots\}$
 4: for each $L_\imath \in \{L_1, L_2, \cdots\}$ do
 5:     if $|P| + |L_\imath| \leq n$ then
 6:         $P = P \cup L_\imath$
 7:     else
 8: //*Do localized knee-oriented dominance selection on critical layer.*//
 9:         $P = P \cup KDSelection(L_\imath, n - |P|, E, R_I)$
10:     end if
11: end for
```

Algorithm 3 : *KDSelection*

```
Input: Population: $L = \{x_1, x_2, \cdots, x_l\}$, output population
    size: $\ell \leq l$, extreme point set: $E$, the set of indices of the
    solutions: $R_I$
Output: Population: $P = \{x_1, x_2, \cdots, x_\ell\}$
 1: $P = \emptyset$
 2: $\mathscr{U} = \{\mathscr{L}_1, \mathscr{L}_2, \cdots\} \wedge \mathscr{L}_1 = \emptyset, \mathscr{L}_2 = \emptyset, \cdots$. //* A set of
    empty lists in $\mathscr{U}$.*//
 3: $CR : \{CR_1, \cdots, CR_k\} = Grouping(R_I)$ //*Do grouping
    on the solutions from the critical layer by using the indices of their
    associated reference vectors.*//
 4: for each $CR_i \in CR$ do
 5:     $KDdominanceSorting(CR_i) = \{S_1^i, S_2^i, \cdots\}$ //*Do
    localized knee-oriented sorting on each sub-population.*//
 6:     $\mathscr{U} = \{S_1^i \cup \mathscr{L}_1, S_2^i \cup \mathscr{L}_2, \cdots\}$
 7: end for
 8: for each $\mathscr{L}_i \in \mathscr{U} \wedge |P| < \ell$ do
 9:     if $|P| + |\mathscr{L}_i| \leq \ell$ then
10:         $P = P \cup \mathscr{L}_i$
11:     else
12:         $P = P \cup CrowdingDistance(\mathscr{L}_i, \ell - |P|)$
13:     end if
14: end for
```

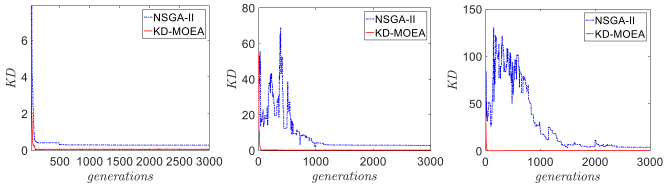
F. Localized knee-oriented-dominance based selection

The localized knee-oriented dominance based selection consists of four steps. First, solutions on the critical front are again divided into the sub-populations according to the reference vectors each individual is associated with in the localized α-dominance based sorting. Second, the solutions in each sub-population are re-sorted according to the knee-oriented-dominance relationship and a front number is assigned to each



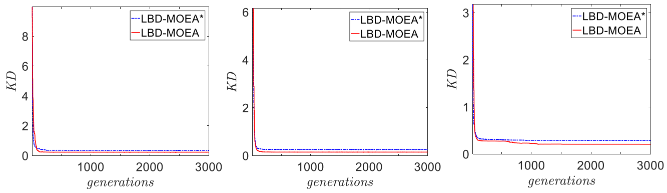
(a) (b)

Fig. 6. (a) An illustration of the localized knee-oriented-dominance based selection on the critical front. Ten solutions are first grouped into four sub-populations, sorted separately using the knee-oriented-dominance, and assigned a sub-front number. The sorted solutions are combined again and sorted into three layers based on their sub-front number, where these layers consist of the solutions highlighted in different shapes and colors. Each solution is assigned with two numbers. The first number indicates the sub-population the solution is grouped into and the second number is its sub-front number. (b) A illustration of the crowding distance, where A, B, C, D, E are the solutions, and a, b, c, i, j, k represent the distances computed according to the research [35]. Notably, A and E are the extreme points of the whole population.



(a) three-objective (b) five-objective (c) eight-objective

Fig. 7. The KD values of the solutions obtained by KD-MOEA and NSGA-II over the generations on PMOP2 with three, five, and eight objectives.



(a) three-objective (b) five-objective (c) eight-objective

Fig. 8. The KD values of the solutions obtained by LBD-MOEA and its variant (LBD-MOEA*) over the generations on PMOP2 with three, five, and eight objectives.

solution. As previously discussed, the localized α-dominance based sorting is able to prevent a knee region having a large curvature from dominating other knee regions. Third, the sorted solutions in different sub-populations are combined again, which are then grouped into a number of sub-frontiers according to their knee-oriented-dominance front number. Then, the crowding distance is calculated for individuals on each sub-front. By now, all solutions on the critical front are sorted into sub-fronts according to their front number in an ascending order and solutions on the same sub-front are sorted according to the crowding distance in a descending order. Finally, the knee oriented selection can be completed based on the rank of the sub-frontiers at first and then based on the crowding distance, similar to the selection in the non-dominated sorting based genetic algorithm (NSGA-II) [35]. Algorithm 3 lists the pseudo code of the knee-oriented-dominance based selection, where the knee-oriented sorting is a non-dominated sorting that uses the proposed knee-oriented dominance instead of the Pareto dominance as used in NSGA-II.

Fig. 6 (a) gives an example of selecting solutions from the critical front using the localized knee-oriented-dominance sorting. In this example, ten solutions on the critical front are grouped into four sub-populations. Then, solutions in each sub-populations are sorted into sub-fronts according to the knee-oriented-dominance. For instance, three solutions in the first sub-population are sorted into two sub-fronts, where solution (1,1) means a solution in sub-population 1 has been assigned a front number 1 (the first sub-front), while two solutions, both labelled (1,2), are assigned a front number of 2. Similarly, the two solutions in sub-population 2 are sorted into two sub-fronts, the three solutions in the third sub-population are sorted into three sub-fronts, and the two solutions in sub-population 4 are sorted into two sub-fronts. Afterwards, the ten solutions are combined again and sorted into three sub-fronts. That is, four solutions labelled (1,1), (2,1), (3,1), and (4,1) are on the first sub-front, five solutions labelled (1,2), (2,2), (3,2), and (4,2) are on the second sub-front, and one solution labelled (3,3) is on the third sub-front. Then, the crowding distance will be calculated for solutions on the same sub-front. For example, if seven out of ten solutions need to be selected for the next generation in Fig. 6 (a), the solutions highlighted in red circles in the first sub-front will be selected at first, then three out of five solutions from the second sub-front (consisting of the solutions highlighted in green triangles) will be chosen according to their crowding distances. Hence, solution (1,2) (the one that is closer to the f2 axis in the 1st cluster), solution (3,2) in the 2nd cluster, and solution (4,2) in the 4th cluster will be selected since their crowding distances are the three largest among all solutions on the same front. Another example is shown in Fig. 6 (b) to illustrate the crowding distance when there are Pareto dominated solutions, where solution D Pareto dominates C, because the bi-dominance relationship is compliant with the Pareto dominance in a CHIM or cluster but not compliant with the Pareto dominance when the solutions from different CHIMs or clusters are selected for comparison, investigated in Section II of the Supplementary material. Different from the crowding distance defined in [35], solutions A and E are the extreme points of the whole population in this study, and assigned with an infinitely large crowding distance, while the crowding distance of B, C, D is $a + j$, $b + k$, and $i + c$, respectively. Hence, solutions A, B, D, E having a larger crowding distance than C are selected, if four out of five solutions need to be selected. In summary, the crowding distance is to choose the solutions from different clusters regarding to their closeness to each other.

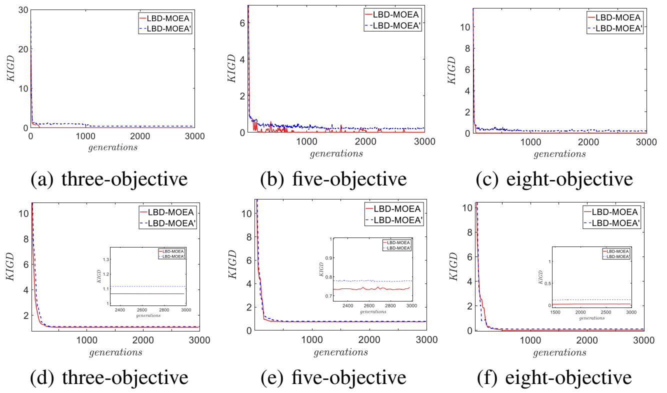
A pilot study is conducted to verify that the proposed knee-oriented selection is able to help drive the population towards the Pareto front in the early search stage and then guide the population to knee regions in the later search stage. To this end, we replace the crowding distance in NSGA-II with the proposed knee-oriented sorting, called KD-MOEA and compare the convergence performance of KD-MOEA with NSGA-II in terms of the knee-driven dissimilarity (KD) [32] a knee-oriented benchmark problem (PMOP2) [32] with three, five and eight objectives, respectively. The experimental results are plotted Fig. 7, respectively. Recall that KD describes the obtained solution set whether contains at least one solution close to each true knee point of the Pareto front, and the smaller the KD value, the better the performance.

Fig. 7 shows that KD-MOEA and NSGA-II perform very differently on the three-, five- and eight-objective PMOP2 test instances, where the knee regions with slightly different degrees of convexity are distributed on an asymmetrical PoF, illustrating three typical different cases in search for knee regions. Both KD-MOEA and NSGA-II converge quickly in terms of KD on the three-objective PMOP2, as shown in Fig. 7 (a), since the Pareto dominance works well for driving the population to the Pareto front for three-objective problems. However, since the selection strategy in NSGA-II is not meant for finding knee regions, its performance in terms of KD is poor because the solutions that are not in a knee region will also be kept in the final population. By contrast, the knee-oriented dominance in KD-MOEA favoring knee solutions will discard solutions not in a knee region, resulting in a much better KD value than that of NSGA-II. Furthermore, the difference in Fig. 7 (b) becomes more apparent as the number of objectives increases, in which case the number of the knee regions also significantly increases. Notably, in Fig. 7 (c), the KD values of NSGA-II increase in the initial stage of the search. This is because the solutions in the high-dimensional objective space are easily Pareto non-dominated, especially in the early search stage. However, the KD value of KD-MOEA decreases quickly and is close to zero in the final stage of the search, indicating that the knee-oriented sorting is able to guide the population towards the PoF and find the knee candidates in the knee regions of the PoF.

Another pilot study is to verify that the proposed knee-oriented dominance can improve the search of the potential knee regions after the α-dominance sorting, so that we compare the LBD-MOEA embedded with both localized α-dominance and knee-oriented dominance, and LBD-MOEA* without the knee-oriented dominance. From Fig. 8, we can see that both algorithms converge fast to the PoF according to the KD values, which indicates that both algorithms are able to drive the subpopulation towards the potential knee regions. By contrast, LBD-MOEA has better KD performance than LBD-MOEA* on PMOP2 with 3, 5, and 8 objectives, mainly because the knee-oriented dominance can help the search concentrate on the potential knee regions. Consequently, LBD-MOEA will find better knee candidates in the knee regions than the LBD-MOEA* in the final stage of the search of knee candidates.

G. Reference vector update

The reference vectors are updated at each generation to make sure that the partition of the sub-population roughly reflects the distribution of the knee regions, which is unknown in the beginning. Refer to Line 14 of Algorithm 1. During the optimization, some reference vectors may have no solution associated with, which may indicate that these regions are not of interest in terms of search for knee regions. Consequently, it is essential to update of the reference vectors.



(a) three-objective (b) five-objective (c) eight-objective

(d) three-objective (e) five-objective (f) eight-objective

Fig. 9. The KIGD values of the solutions obtained by LBD-MOEA and its variant (LBD-MOEA′) over the generations on PMOP2 and PMOP14 with three, five, and eight objectives, where plots (a) – (c) are the results on PMOP2 with unimodal functions, and (d) – (f) on PMOP14 with degenerate PoFs.

In this work, reference vectors with no solution or only one solution associated with will be updated (Line 14 of Algorithm 1). Reference vectors having no solution associated with will be replaced with a random vector, e.g., $v = (r_1/\sum_{i=1}^{m} r_i, \cdots, r_m/\sum_{i=1}^{m} r_i), r_i = rand(0,1)$ and $i = 1, \ldots, m$. In addition, reference vectors associated with one solution will also be updated in the following way. Given a solution p whose objectives are $(f_1(p), \cdots, f_m(p))$, the reference vector is updated by $v = (f_1(p)/\sum_{i=1}^{m} f_i(p), \cdots, f_m(p)/\sum_{i=1}^{m} f_i(p))$. The previous strategy is mainly for the diversity of the population towards the knee regions. Instead of the waste of the search in the subregions associated with no solutions, random sampling the reference vectors may assist the search towards undetected knee regions or may drive the search close to the detected knee regions when the reference vectors are close to the knee regions. The latter strategy aims to adjust the distribution of the reference vectors associated with only one solution, which may assist in the search of the subspaces. With the help of this strategy, the associated solution is easily remained because it is located along the reference vector. Consequently, new solutions may be generated in next generations and the search of potential knee region in this subspace will be conducted. An example is shown in Fig. 9 to investigate the diversity performance of LBD-MOEA and its variant (LBD-MOEA′ with fixed reference vectors) on PMOP2 and PMOP14 with three, five and eight objectives, respectively. The plots show that the KIGD values of LBD-MOEA is smaller than that of its variant. KIGD [32] evaluates the diversity of the solutions covering the knee regions, and the smaller the KIGD value, the better the performance. The results indicate that the strategy to update the reference vectors can improve the diversity of LBD-MOEA in search for more potential knee regions. As a result, the algorithm is able to explore potential knee regions and the ability to identify knee points is enhanced.

H. Computational complexity

The computational complexity of LBD-MOEA comes mainly from the knee-oriented environmental selection, which consists of the localized α-dominance sorting and knee-oriented sorting. The localized α-dominance sorting follows the same procedure of the Pareto non-dominated sorting [35] whose complexity is $O(n^2 \times m)$. But the calculation of $g_i(x, y)$ in α-dominance introduces an additional complexity of $O(m)$ on each objective. Therefore, the complexity of the α-dominance based non-dominated sorting is $O(n^2 \times m^2)$, where n and m are the population size and the number of objectives, respectively. In the knee-oriented selection, the angle between two solutions needs to be calculated, which requires a computational complexity of $O(m)$. Note, however, that the knee-oriented sorting is only applied on the critical front. The worst case occurs when all solutions are on the critical front. Thus, the complexity of knee-oriented sorting is $O(n^2 \times m)$. The complexity of the crowding distance is $O(m \times n \log n)$.

Overall, the expected computational complexity of LBD-MOEA is $O(n^2 \times m^2)$.

IV. Experimental Results and Discussion

A. Experimental setting

To examine the performance of LBD-MOEA, six knee identification algorithms are compared, including KD-MOEA, a variant of NSGA-II replacing the crowding distance with the proposed knee-oriented selection, TKR [26], EMU^r [28], KnEA [13], K-ASA [29], and α-MOEA-KI [31]. All parameters are set following the recommended settings in the original papers. Specifically, TKR uses mobile reference points and a utility function to search the knee candidates, where the utility is based on the ratio between the improvement and deterioration when the objectives of two solutions are exchanged, where a cleaning parameter is set to 0.001. EMU^r recursively uses the expected marginal utility to detect knee regions and the internal solutions will be kept for comparison, where the number of weight vectors is set the same as the population size. KnEA is based on the distance from the solution to the hyperplane constructed by the extreme points [19] and the knee identification will continue on the final set, where the rate of knee points is set to 0.5 as the same as the default setting. K-ASA adopts the angle-based pruning strategy for the search of knee regions, where threshold of the angle size is set to 0.95 for two-objective, and 0.90 for three- and many-objective problems. α-MOEA-KI uses a localized α-dominance for the search of knee regions, where α is suggested to set to 0.75. The (H_1, H_2) in this work is set to (1, 5), (1, 3), (1, 2), and (1, 3) for reference vector generation for the problems with two, three, five, and eight objectives, respectively. We refer the readers to Section III in the Supplementary materials for a sensitive analysis on the number of reference vectors and Section VI in the Supplementary materials for a sensitive analysis on the parameter τ in the knee-oriented dominance. The corresponding population size is set to 100, 105, 126 and 156. In the experiments, the binary tournament selection is applied as the mating selection. The distribution indices in both the simulated binary crossover and polynomial mutation are set to 20. The crossover probability and mutation probability are set to 1.0 and $1/n$, respectively, where n is the number of decision variables.

Two sets of knee-oriented benchmarks are introduced. The first set includes DO2DK [22], CKP [23], DEB2DK [22], and DEB3DK [22]. The second set is the PMOP test suite recently proposed in [32]. The former is mainly designed for the knee identification in two- and three-objective problems. The latter is for the identification of knees in high-dimension objective spaces. All parameter settings are presented in Table S1 in the Supplementary material, where m and n are the number of objectives and decision variables, respectively. (A, B, s, p) and (K, l) are the parameters of the basic knee functions in different sets of benchmarks to control the shape and number of the knee regions. Each algorithm is executed for 30 independent runs on each test instance. The termination is set to 1000 generations for the first set of benchmark problems, except for DO2DK with 5000. For the second set (PMOP test suite), the maximum number of generations is set to 3000 for PMOP1-PMOP3, PMOP6-PMOP9, and PMOP13, 5000 for PMOP10-PMOP12 and PMOP14, and 10000 for PMOP4 and PMOP5, respectively. In the comparative experiments, the Wilcoxon rank sum test (a significance level is 0.05) is adopted to analyze the results, where "+", "$-$", and "$\approx$" indicate that the result is significantly better, significantly worse and statistically comparable to the solutions obtained by LBD-MOEA, respectively.

All the compared algorithms are run on the PlatEMO [36] in Matlab 2018b using the CPU with an Intel(R) Core(TM) i5-8250U CPU @ 1.8GHz and 8.00 GB RAM. The operation system is the 64 Microsoft Windows 10 on a 64-bit processor.

B. Performance indicators

For quantificationally analyzing LBD-MOEA, three knee-oriented indicators from [32] are adopted for performance evaluation, including the knee-driven generational distance (KGD), knee-driven inverted generational distance (KIGD), and the knee-driven dissimilarity (KD).

Given a reference point set ($\mathbb{Q}$) of the convex knee regions and true knee point set ($\mathbb{K}$), the KGD, KIGD, and KD values of an achieved solution set ($\mathbb{P}$) are calculated as follows:

- Knee-driven generational distance (KGD):

$$\text{KGD} = \frac{1}{|\mathbb{P}|}\sum_{i=1}^{|\mathbb{P}|} d(\nu_i, \mathbb{Q}) \tag{12}$$

  where $d(\nu_i, \mathbb{Q})$ means the shortest Euclidean distance from a solution ν_i in $\mathbb{P}$ to the reference set $\mathbb{Q}$.

- Knee-driven inverted generational distance (KIGD):

$$\text{KIGD} = \frac{1}{|\mathbb{Q}|}\sum_{i=1}^{|\mathbb{Q}|} d(\nu_i, \mathbb{P}) \tag{13}$$

  where $d(\nu_i, \mathbb{P})$ means the shortest Euclidean distance from the reference point ν_i in $\mathbb{Q}$ to the obtained solution set $\mathbb{P}$.

- Knee-driven dissimilarity (KD):

$$\mathrm{KD} = \frac{1}{|\mathbb{K}|}\sum_{i=1}^{|\mathbb{K}|} d(\nu_i, \mathbb{P}) \tag{14}$$

  where $d(\nu_i, \mathbb{P})$ means the shortest Euclidean distance from a the true knee point ν_i in $\mathbb{K}$ to the obtained solution set $\mathbb{P}$.

The KGD evaluates the proximity of the obtained solutions to the reference points in the knee regions of the Pareto front. The KIGD measures the diversity of the obtained solutions covering the knee regions. The KD describes the obtained solution set whether contains at least one solution close to each true knee point. The smaller the values of the indicators are, the better the performance of the algorithm is.

C. Experimental results and analysis

This section aims to compare the performance of LBD-MOEA in comparison with six knee identification methods in terms of three knee-driven indicators, KGD, KIGD, and KD. The experiments are conducted on two sets of problems listed in Table S1 in the Supplementary material. Tables S2-S4 in the Supplementary material present the comparative results (mean and variance values) obtained by the seven algorithms on 50 test instances with two, three, five, and eight objectives.

1) Comparison with KGD indicator: The KGD values of the seven algorithms are presented in Table S2 in the Supplementary material. The results indicate that LBD-MOEA performs the best according to the best values and rank values in comparison with the other six algorithms. Specifically, LBD-MOEA ranks the first with 20 best records, followed by K-ASA and α-MOEA-KI with 12 and 10 best records, respectively. According to the rank sum test, LBD-MOEA achieves better convergence performance on 42, 44, 47 and 40 out of 50 instances than TKR, KnEA, EMU^r and KD-MOEA, respectively. It may be because TKR, KnEA and EMU^r favor the extreme solutions or boundary solutions, which may easily become DRSs. As a consequence, these solutions may deteriorate the convergence performance. Specifically, TKR uses the ratio between the improvement and deterioration when the objectives of two solutions are exchanged, and since the solutions from the extreme regions have a larger ratio, they are more likely to be selected during the environmental selection. KnEA uses the extreme solutions to construct the hyperplane and locates the knee candidates that have the maximum distance to the hyperplane. In EMU^r, the expected marginal utility value of the boundary points are larger than that of some solutions in the knee regions. Additionally, KD-MOEA may preserve the solutions from the concave regions when the parameter to control the dominated area of a solution is small, in which case the preserved solutions may mislead the search process and further degrade the convergence performance. In contrast, LBD-MOEA adopts the localized α-dominance based non-dominated sorting during the environmental selection, which is able to get rid of the DRSs. LBD-MOEA outperforms α-MOEA-KI on most problems too, which also adopts the localized α-dominance. This may be attributed to the proposed knee-oriented selection used by LBD-MOEA, which can drive the population towards the knee regions and eliminate boundary solutions and solutions in the concave regions. K-ASA performs worse than LBD-MOEA on 34 instances. K-ASA adopts a angle-based pruning strategy in the environmental selection. However, the angle between an extreme point (or a boundary point) and its adjacent solution can be very small, and consequently, the boundary points will be kept according to the pruning strategy. Because of this selection strategy, the boundary points may become the DRSs and the convergence of K-ASA may be degraded by the DRSs. The performance of LBD-MOEA is worse than that of other algorithms on some problems such as PMOP5 and PMOP6. PMOP5 has many knee regions close to each other, which will make LBD-MOEA perform much local search, slowing down the convergence speed. On the contrary, PMOP6 only has one global knee region. As a result, many reference vectors do not have any solutions associated with them and they are frequently adjusted, degrading the search performance of LBD-MOEA. Overall, the experimental results demonstrate that LBD-MOEA can effectively guide the evolutionary search to find the knee regions on the majority of the test functions investigated in this study compared with six state-of-the-art algorithms.

2) Comparison with KIGD indicator: Table S3 in the Supplementary material presents the comparative results in terms of the KIGD indicator. The results show that LBD-MOEA outperforms others on most instances. According to the best records, LBD-MOEA achieves the best with 21 best records, while α-MOEA-KI ranks the second with 9 best results. According to the rank values, LBD-MOEA has better diversity performance over 36, 40, 40, 42, 32, and 39 out of 50 instances compared with TKR, KnEA, EMU^r, K-ASA, α-MOEA-KI, and KD-MOEA, respectively. LBD-MOEA shows better performance on most PMOP test problems including PMOP1-PMOP4, PMOP7-PMOP13. Most of them are multimodal and have more knee regions than other test functions such as PMOP6, making it more challenging for a search algorithm to find the knee regions of these test problems. Thus, the better performance of LBD-MOEA on the PMOP test problems can be attributed to the fact that it is able to search for multiple potential knee regions and converge to the knee regions, while the selection mechanisms in the compared algorithms tend to favor the solutions in the non knee regions. For example, in the environmental selection of TKR and EMU^r, the solutions from the boundary regions have a larger chance to be selected than the solutions from the knee regions with a relatively smaller curvature. As a result, the algorithm prefers to search the boundary regions and makes little effort on exploring potential knee regions. Similarly, KnEA favors solutions in the extreme regions, while K-ASA, α-MOEA-KI and KD-MOEA prioritize solutions in the concave regions. As a result, their KIGD performances are worse than that of LBD-MOEA on these problems. However, the performance of LBD-MOEA is less competitive than that of the compared methods on the first set of test problems, probably because the selection pressure of LBD-MOEA focuses too much on the knee regions and the obtained solutions will crowded around the true knee points. Consequently, its diversity performance is relatively poor. By contrast, KD-MOEA shows better performance than LBD-MOEA. It may be because these problems are relatively easy to be converged in comparison with the PMOP test suite, where the distance functions of DEB2DK and DEB3DK problems are unimodal but most PMOPs multimodal. As a result, the localized α-dominance improving the convergence rate cannot be reflected on these problems but more on the PMOPs. Notably, LBD-MOEA also shows worse performance on PMOP5 and PMOP6, probably due to the frequent adjustment of reference vectors to search for multiple knee regions. From the above comparative experiments, we demonstrate that LBD-MOEA is able to find good knee candidates in the knee regions on most test problems, especially those having multiple knees regions.

3) Comparison with KD indicator: A further observation on LBD-MOEA is made by comparing the KD values of the solution sets obtained by the knee identification methods. The results are given in Table S4 in the Supplementary material. LBD outperforms others with 20 best records, followed by α-MOEA-KI, KD-MOEA, and KnEA with nine, six, and six best records, respectively. According to the rank values, the results show that LBD-MOEA is competitive against the compared algorithms on most instances, indicating that LBD-MOEA is able to achieve good knee points. Specifically, LBD-MOEA outperforms TKR, KnEA, EMU^r, K-ASA, α-MOEA-KI, and KD-MOEA on 38, 42, 39, 41, 32, and 38 out of 50 instances, respectively. It is mainly due to the fact that the localized dominated sorting and knee-oriented selection can guide the search towards multiple potential knee regions during the optimization and the solutions closer to the center of the knee regions are favored over their neighbors in the environmental selection. Besides, LBD-MOEA is relatively insensitive to the DRSs and boundary solutions because these solutions are eliminated during the environmental selection. Recall that both α-dominance and knee-oriented dominance relationships are able to enlarge the dominated area of the solutions in the knee regions. Consequently, once a solution having a larger curvature and is closer to the boundaries is obtained, the solution will dominate the solutions on the boundaries according to the two localized dominance relationships. On the contrary, the boundary regions, extreme regions or concave regions are easily retained in other identification methods, which may become DRSs and degrade the KD performance. To sum up, LBD-MOEA shows a stronger capability of achieving knee solutions than the compared algorithms on most test instances studied in this work.

4) Visualization of the results and analysis: In the following, we visually compare some solution sets obtained in Section IV-C to take a closer look at the performance of the compared algorithms.

Figs. S1-S4 in the Supplementary material plot the knee candidate solutions obtained by seven algorithms on the DO2DK, DEB2DK, CKP, and DEB3DK problems. These results show that LBD-MOEA, and KD-MOEA outperform other methods in acquiring good knee candidates to the knees or knee regions on these problems. α-MOEA-KI shows similar results on DO2DK, DEB2DK, and CKP problems but worse performance than LBD-MOEA on DEB3DK problems. TKR shows good performance on DO2DK, DEB2DK, and CKP problems. The following is the EMU^r which is easily impacted by the global and boundary solutions, so that it shows good performance on DO2DK and DEB2DK problems. KnEA easily finds the global knees but the local knee regions are easily ignored. It is main because LBD-MOEA introduces the techniques (modified dominance relationships) to deal with the DRSs and boundary points, which can balance the optimization and locating the knee regions. Consequently, LBD-MOEA acquires better results than other methods on these problems. Notably, Fig. S4 (h) shows that LBD-MOEA has found seven out of nine knee regions. LBD-MOEA cannot distinguish very close knee regions, mainly because the solutions in the closely located neighboring knee regions will be partitioned in the same sub-population and as a result, only the knee region with a large curvature will be kept during the knee-oriented environmental selection. In dealing with DEB3DK problems, KD-MOEA and α-MOEA-KI also find seven knee regions, but they also provide non-interested solutions. Both K-ASA and TKR find five knee regions but TKR cannot eliminate the boundary solutions or solutions in concave regions. KnEA and EMU^r are easily influenced by the boundary solutions and extreme solutions, and consequently they perform worse than the other compared algorithms.

Figs. S5-S8 in the Supplementary material present the results obtained by seven algorithms on PMOP2, PMOP10, PMOP11, and PMOP13 with eight objectives. PMOP2 is relatively easy to be optimized. But the results indicate that LBD-MOEA and α-MOEA-KI have better performance than others to get candidate solutions close to the true knee points. However, KnEA and EMU^r favor the boundary solutions so that their convergence performance is worse than others. By contrast, the solutions obtained by K-ASA are diverse but not properly located in the knee regions. KD-MOEA also shows worse convergence performance. Similar performance can be also shown on PMOP10 and PMOP11, which are hard to be converged because different convergence speeds are designed on different objectives. Consequently, the modified dominance driven MOEAs like LBD-MOEA and α-MOEA-KI show better performance than others. PMOP13 is very challenging, because this problem is degenerated and only one global knee region is on the PoF. Hence, most algorithms cannot deal with the problem. LBD-MOEA shows better results as it has good balance between the optimization and the search of the knee regions. The following is α-MOEA-KI. The rest algorithms cannot find good knee solutions mainly because they are more sensitive to the DRSs during the optimization and the convergence speed will be slowed down.

All in all, Tables S2-S4 in the Supplementary material summarize the experimental results obtained by seven compared algorithms on 50 instances in terms of the KGD, KIGD, and KD indicators, and the results indicate that LBD-MOEA is competitive in search for knee regions and location of knee solutions. Its performance is also verified by the results presented in Figs. S1 to S3 in the Supplementary material.

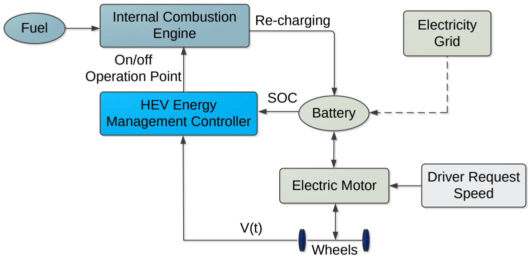


Fig. 10. An illustration of the architecture of the HEV. The propulsions to drive the car are from the combination of internal combustion engine (ICE) and electric motor (EM), which are powered by the fuel and battery, respectively. The battery can be charged by electricity grid from any changing station or recharged during the braking. The HEV energy management controller adjusts the ICE in terms of the current speed ($v(t)$) and state-of-charge (SOC). The torque generated by the electric motor is determined in terms of the request speed from the driver.

TABLE I
THE INDICATOR RESULTS OBTAINED BY SEVEN ALGORITHMS ON HEV PROBLEM. THE BEST RESULTS ARE HIGHLIGHTED IN GREY.

Algorithm	KGD	KIGD	KD
TKR	1.86E-01	9.73E-01	1.06E+00
KnEA	3.88E-01	1.77E+00	1.79E+00
EMU^r	1.25E-01	7.91E-01	7.68E-01
K-ASA	5.18E-02	7.30E-01	6.80E-01
α-MOEA-KI	7.06E-02	7.97E-01	8.86E-01
KD-MOEA	7.97E-02	9.51E-01	8.93E-01
LBD-MOEA	7.57E-02	7.06E-01	5.33E-01

V. Experiment on Hybrid Electric Vehicle Controller Design

In this section, we compare seven algorithms on a hybrid electric vehicle (HEV) controller design model [37] to investigate the effectiveness of LBD-MOEA. It should be pointed out that no ground true about the knee points and knee regions of the HEV controller design problem is known.

The general architecture of HEV is shown in Fig. 10. The HEV energy management controller aims to minimize seven objectives by switching the power sources between an internal combustion engine (ICE) and an electric motor (EM), where the objectives include the minimization of the fuel consumption (FC), battery stress (BS), operation charges (OC), emission, noise, urban operations (UO), and battery state-of-change (SOC). Interested readers are referred to [37] for more details of the seven-objective HEV controller design problem.

All algorithms are run on the seven-objective HEV controller design problem with a maximum number of 78000 fitness evaluations. In the seven compared algorithms, (H_1, H_2) is set to (3,2) for the reference vector generator, and the population size is set to 156. Other parameters are set the same as in Section IV-A. The knee solutions obtained by seven compared algorithms are shown in Fig. 11 (a) – (g). Since the number and locations of the knee solutions of the HEV controller design problems is unknown, we run three popular MOEAs, namely MOEA/D [38], NSGA-III [33], and RVEA [39] on the problems with 100,000 fitness evaluations for each to collect as set of 828 Pareto non-dominated solutions. This large set of solutions will be used as the “ground truth" of the PoF of the HEV problem for evaluating the performance of the solutions obtained by the compared algorithms. Then, three *a posteriori* knee identification algorithms, TKR, KnEA, and EMU^r are adopted to identify the knee points from the 828 non-dominated solutions, which are shown in Fig. 11 (h). The three sets of knee solutions are merged and redundant knee solutions are removed, resulting in 52 knee points, referring to Fig. 11 (i). The neighboring solutions of these knee points, as shown in Fig. 11 (j), will then be used as the reference set for calculating the knee-driven performance indicators, KGD, KIGD, and KD for evaluating the quality of the solution sets obtained by the seven compared algorithms.

The evaluation results are presented in Table I, where the best results are highlighted. From these results, we can see that LBD-MOEA outperforms all compared algorithms in terms of KIGD and KD, although it is slightly than K-ASA in terms of KGD. It indicates that LBD-MOEA has better performance in locating more knee regions and is able to find solutions close to the knee solutions of the approximated PoF. By contrast, K-ASA may find more solutions close the reference solutions in the knee regions but relatively far away from the knees.

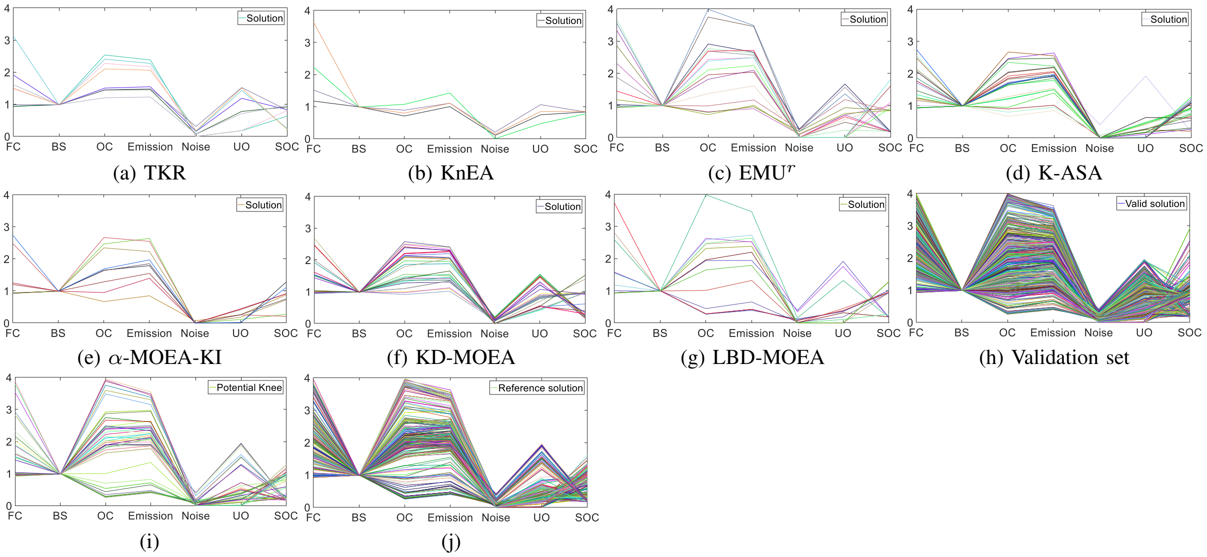
Overall, LBD-MOEA has shown competitive performance in acquiring knee solutions on the seven-objective HEV controller design problem.

VI. Conclusion

In preference-driven evolutionary optimization, the lack of *a* priori knowledge makes it difficult for the decision-makers to explicitly express their preferences. In these cases, the knee points are considered as the naturally preferred solutions. Several online algorithms have been proposed to search for knee regions by embedding different knee-oriented measures into the environmental selection, although most existing methods do not perform well in striking a good balance between converging to the knee solutions and searching for multiple knee regions.

To address the issue, this paper proposed a localized knee-oriented environmental selection for online detecting knee solutions and knee regions. A localized α-dominance sort and a localized knee-oriented-dominance sort proposed in this work are embedded in the environmental selection. The localized α-dominance based selection can alleviate impact of the dominance resistant solutions and guide the search towards different knee regions, whereas the localized knee-oriented-dominance based selection can locate the knee solutions in a potential knee regions and keep the knee solutions that may be missed by the α-dominance based selection. Our empirical results demonstrated that the proposed environmental selection combining the localized α-dominance and the localized knee-oriented dominance is able to maintain a good balance between approximating multiple knee regions and locating the knee solutions. The results also verified that the proposed method outperforms its competitors on most problems studied in this work having up to eight objectives and on a hybrid electric vehicle controller design problem with seven objectives.

The experiments also show that the proposed algorithm cannot distinguish multiple knee regions that are close to



(a) TKR (b) KnEA (c) EMU^r (d) K-ASA

(e) α-MOEA-KI (f) KD-MOEA (g) LBD-MOEA (h) Validation set

(i) (j)

Fig. 11. Plots (a) - (g) are the knee candidate solutions obtained by seven algorithms on hybrid electric vehicle controller design problem. Plot (h) presents the approximated PoF of HEV controller design problem. (i) plots the potential knees of the approximated PoF. (j) presents the representative solutions in the corresponding knee regions.

each other because solutions in these knee regions will be very likely associated with the same sub-population, making the localized α-dominance and the localized knee-oriented-dominance ineffective. Thus, our future work will be dedicated to developing new methods for distinguishing multiple knee regions in a close neighborhood and develop more efficient reference adaption method to deal with the issue. Another line of research is to improve the search ability of the algorithm to find knee solutions in higher dimensional objective spaces by automatically setting the number of reference vectors during the optimization.

Acknowledgment

This work was supported in part by the Honda Research Institute Europe GmbH.

References

[1] K. Miettinen, *Nonlinear multiobjective optimization*. Springer Science & Business Media, 2012, vol. 12.

[2] D. A. Van Veldhuizen and G. B. Lamont, "Multiobjective evolutionary algorithms: Analyzing the state-of-the-art," *Evolutionary Computation*, vol. 8, no. 2, pp. 125–147, 2000.

[3] T. Wagner, N. Beume, and B. Naujoks, "Pareto-, aggregation-, and indicator-based methods in many-objective optimization," in *Proceeddings of the International Conference on Evolutionary Multi-criterion Optimization*. Springer, 2007, pp. 742–756.

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

[5] G. Yu, Y. Jin, and M. Olhofer, "References or preferences–rethinking many-objective evolutionary optimization," in *2019 IEEE Congress on Evolutionary Computation (CEC)*. IEEE, 2019, pp. 2410–2417.

[6] S. F. Adra, I. Griffin, and P. J. Fleming, "A comparative study of progressive preference articulation techniques for multiobjective optimisation," in *International Conference on Evolutionary Multi-Criterion Optimization*. Springer, 2007, pp. 908–921.

[7] G. Yu, J. Zheng, R. Shen, and M. Li, "Decomposing the user-preference in multiobjective optimization," *Soft Computing*, vol. 20, no. 10, pp. 4005–4021, 2016.

[8] J. Zheng, G. Yu, Q. Zhu, X. Li, and J. Zou, "On decomposition methods in interactive user-preference based optimization," *Applied Soft Computing*, vol. 52, pp. 952–973, 2017.

[9] R. C. Purshouse, K. Deb, M. M. Mansor, S. Mostaghim, and R. Wang, "A review of hybrid evolutionary multiple criteria decision making methods," in *Proceedings of the 2014 Congress on Evolutionary Computation*. IEEE, 2014, pp. 1147–1154.

[10] H. Wang, M. Olhofer, and Y. Jin, "A mini-review on preference modeling and articulation in multi-objective optimization: current status and challenges," *Complex & Intelligent Systems*, vol. 3, no. 4, pp. 233–245, 2017.

[11] J. Dyer, "Multiple criteria decision analysis: state of the art surveys," *International Series in Operations Research and Management Science*, vol. 78, no. 4, pp. 265–292, 2005.

[12] K. Deb and S. Gupta, "Understanding knee points in bicriteria problems and their implications as preferred solution principles," *Engineering Optimization*, vol. 43, no. 11, pp. 1175–1204, 2011.

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

[14] S. M. Alzahrani and N. Wattanapongsakorn, "Comparative study of knee-based algorithms for many-objective optimization problems," *ECTI Transactions on Computer and Information Technology (ECTI-CIT)*, vol. 12, no. 1, pp. 7–16, 2018.

[15] J. Zou, Q. Li, S. Yang, H. Bai, and J. Zheng, "A prediction strategy based on center points and knee points for evolutionary dynamic multi-objective optimization," *Applied Soft Computing*, vol. 61, pp. 806–818, 2017.

[16] T. Chen, K. Li, R. Bahsoon, and X. Yao, "FEMOSAA: Feature-guided and knee-driven multi-objective optimization for self-adaptive software," *ACM Transactions on Software Engineering and Methodology (TOSEM)*, vol. 27, no. 5, pp. 1–50, 2018.

[17] C. Yue, J. Liang, B. Qu, H. Song, G. Li, and Y. Han, "A knee point driven particle swarm optimization algorithm for sparse reconstruction," in *Asia-Pacific Conference on Simulated Evolution and Learning*. Springer, 2017, pp. 911–919.

[18] A. K. Nandi, D. Chakraborty, and W. Vaz, "Design of a comfortable optimal driving strategy for electric vehicles using multi-objective optimization," *Journal of Power Sources*, vol. 283, pp. 1–18, 2015.

[19] I. Das, "On characterizing the knee of the Pareto curve based on normal-boundary intersection," *Structural Optimization*, vol. 18, no. 2-3, pp. 107–115, 1999.

[20] O. Schütze, M. Laumanns, and C. A. C. Coello, "Approximating the knee of an MOP with stochastic search algorithms," in *Proceedings of the International Conference on Parallel Problem Solving from Nature*. Springer, 2008, pp. 795–804.

[21] I. Das and J. E. Dennis, "Normal-boundary intersection: A new method

for generating the Pareto surface in nonlinear multicriteria optimization problems," *SIAM Journal on Optimization*, vol. 8, no. 3, pp. 631–657, 1998.

[22] J. Branke, K. Deb, H. Dierolf, and M. Osswald, "Finding knees in multi-objective optimization," in *Proceedings of the International Conference on Parallel Problem Solving from Nature*. Springer, 2004, pp. 722–731.

[23] G. Yu, Y. Jin, and M. Olhofer, "A method for *a posteriori* identification of knee points based on solution density," in *Proceedings of the 2018 IEEE Congress on Evolutionary Computation*. IEEE, 2018, pp. 1–8.

[24] M. Braun, P. Shukla, and H. Schmeck, "Angle-based preference models in multi-objective optimization," in *Proceedings of the International Conference on Evolutionary Multi-Criterion Optimization*. Springer, 2017, pp. 88–102.

[25] L. Rachmawati and D. Srinivasan, "Multiobjective evolutionary algorithm with controllable focus on the knees of the Pareto front," *IEEE Transactions on Evolutionary Computation*, vol. 13, no. 4, pp. 810–824, 2009.

[26] B. Slim, S. L. Ben, and K. Ghédira, "Searching for knee regions of the Pareto front using mobile reference points," *Soft Computing*, vol. 15, no. 9, pp. 1807–1823, 2011.

[27] K. Ikeda, H. Kita, and S. Kobayashi, "Failure of Pareto-based MOEAs: Does non-dominated really mean near to optimal?" in *Proceedings of the 2001 Congress on Evolutionary Computation*, vol. 2. IEEE, 2001, pp. 957–962.

[28] K. S. Bhattacharjee, H. K. Singh, M. Ryan, and T. Ray, "Bridging the gap: Many-objective optimization and informed decision-making," *IEEE Transactions on Evolutionary Computation*, vol. 21, no. 5, pp. 813–820, 2017.

[29] S. Sudeng and N. Wattanapongsakorn, "Post pareto-optimal pruning algorithm for multiple objective optimization using specific extended angle dominance," *Engineering Applications of Artificial Intelligence*, vol. 38, pp. 221–236, 2015.

[30] ——, "A knee-based multi-objective evolutionary algorithm: an extension to network system optimization design problem," *Cluster Computing*, vol. 19, no. 1, pp. 411–425, 2016.

[31] G. Yu, Y. Jin, and M. Olhofer, "An a priori knee identification multi-objective evolutionary algorithm based on α-dominance," in *Proceedings of the Genetic and Evolutionary Computation Conference Companion*. ACM, 2019, pp. 241–242.

[32] ——, "Benchmark problems and performance indicators for search of knee points in multi-objective optimization," *IEEE Transactions on Cybernetics*, pp. 1–14, 2019.

[33] 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 Transactions on Evolutionary Computation*, vol. 18, no. 4, pp. 577–601, 2014.

[34] Y. Liu, D. Gong, J. Sun, and Y. Jin, "A many-objective evolutionary algorithm using a one-by-one selection strategy," *IEEE Transactions on Cybernetics*, vol. 47, no. 9, pp. 2689–2702, 2017.

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

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

[37] R. Cheng, T. Rodemann, M. Fischer, M. Olhofer, and Y. Jin, "Evolutionary many-objective optimization of hybrid electric vehicle control: From general optimization to preference articulation," *IEEE Transactions on Emerging Topics in Computational Intelligence*, vol. 1, no. 2, pp. 97–111, 2017.

[38] Q. Zhang and H. Li, "MOEA/D: A multiobjective evolutionary algorithm based on decomposition," *IEEE Transactions on Evolutionary Computation*, vol. 11, no. 6, pp. 712–731, 2007.

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

Guo Yu (S'18-S'20) received the B.S. degree in information and computing science and the M.Eng. degree in computer technology from Xiangtan University, Xiangtan, China, in 2012 and 2015, respectively. He received the Ph.D. degree in computer science from University of Surrey, Guildford, U.K., in 2020.

His current research interests include preference learning, evolutionary optimization, and machine learning. He is a regular reviewer of the IEEE-TRANSACTIONS ON EVOLUTIONARY COMPUTATION, Artificatial Intelligence, IEEE Computational Intelligence Magazine, IEEE Access, Complex & Intelligent Systems, Applied Soft Computing, and Soft Computing.

Yaochu Jin (M'98-SM'02-F'16) received the B.Sc., M.Sc., and Ph.D. degrees in automatic control from Zhejiang University, Hangzhou, China, in 1988, 1991, and 1996, respectively, and the Dr.-Ing. degree from Ruhr-University Bochum, Bochum, Germany, in 2001.

He is a Distinguished Chair, Professor in Computational Intelligence, Department of Computer Science, University of Surrey, Guildford, U.K., where he heads the Nature Inspired Computing and Engineering Group. He was a Finland Distinguished Professor funded by the Finnish Funding Agency for Innovation (Tekes) and a Changjiang Distinguished Visiting Professor appointed by the Ministry of Education, China. His main research interests include data-driven surrogate-assisted evolutionary optimization, secure machine learning, multi-objective evolutionary deep learning, swarm robotics, and evolutionary developmental systems.

Dr Jin is presently the Editor-in-Chief of the IEEE TRANSACTIONS ON COGNITIVE AND DEVELOPMENTAL SYSTEMS and the Editor-in-Chief of Complex & Intelligent Systems. He was an IEEE Distinguished Lecturer, and the Vice President for Technical Activities of the IEEE Computational Intelligence Society. He was the General Co-Chair of the 2016 IEEE Symposium Series on Computational Intelligence and the 2020 IEEE Congress on Evolutionary Computation. He is the recipient of the 2018 IEEE Transactions on Evolutionary Computation Outstanding Paper Award, the 2015, 2017, and 2020 IEEE Computational Intelligence Magazine Outstanding Paper Award, and the Best Paper Award of the CIBCB 2010. He is recognized as a Highly Cited Researcher 2019 by the Web of Science Group. He is a Fellow of IEEE.

Markus Olhofer received the Dipl.-Ing. degree in electrical engineering and the Ph.D. degree in electrical engineering from Ruhr-University Bochum, Bochum, Germany, in 1997 and 2000, respectively.

He joined the Future Technology Research Division, Honda R&D Europe (Germany) GmbH, Offenbach, Germany, in 1998 and has been the Chief Scientist and the Head of the Complex Systems Optimization and Analysis Group, Honda Research Institute Europe, Offenbach, since 2010. He is a Visiting Professor with the Department of Computer Science, University of Surrey, Guildford, U.K. His current research interests include the extension of soft computing methods to meet requirements in complex engineering problems, ranging from evolutionary design optimization to engineering data mining.