Coordinated Adaptation of Reference Vectors and Scalarizing Functions in Evolutionary Many-objective Optimization

Qiqi Liu, Yaochu Jin, Martin Heiderich, Tobias Rodemann

2023

Preprint:

This is an accepted article published in IEEE Transactions on Systems, Man and Cybernetics: Systems. The final authenticated version is available online at: https://doi.org/10.1109/TSMC.2022.3187370

Coordinated Adaptation of Reference Vectors and Scalarizing Functions in Evolutionary Many-objective Optimization

Qiqi Liu, Yaochu Jin, *Fellow, IEEE,* Martin Heiderich, and Tobias Rodemann

***Abstract*—It is highly desirable to adapt the reference vectors to unknown Pareto fronts in decomposition based evolutionary many-objective optimization. While adapting the reference vectors enhances the diversity of the achieved solutions, it often decelerates the convergence performance. To address this dilemma, we propose to adapt the reference vectors and the scalarizing functions in a coordinated way. On the one hand, the adaptation of the reference vectors is based on a local angle threshold, making the adaptation better tuned to the distribution of the solutions. On the other hand, the weights of the scalarizing functions are adjusted according to the local angle thresholds and the reference vectors' age, which is calculated by counting the number of generations in which one reference vector has at least one solution assigned to it. Such coordinated adaptation enables the algorithm to achieve a better balance between diversity and convergence, regardless of the shape of the Pareto fronts. Experimental studies on MaF, DTLZ and DPF test suites demonstrate the effectiveness of the proposed algorithm in solving problems with both regular and irregular Pareto fronts.**

***Index Terms*—Evolutionary many-objective optimization, reference vector, scalarizing function, irregular Pareto fronts**

I. Introduction

In real-world applications, there are a great deal of multi-objective optimization problems (MOPs) that comprise two or three objectives, or many-objective optimization problems (MaOPs) with more than three objectives. Over the past decades, multi-objective evolutionary algorithms (MOEAs) have been shown to be promising in solving both MOPs [1], [2] and MaOPs [3]. Generally speaking, MOEAs can be grouped into four main categories, namely decomposition based [4]–[7], performance indicator based [8]–[10], modified dominance relationship based [11], [12], and preference based algorithms [13], [14]. Among these MOEAs, decomposition based algorithms with evenly distributed reference vectors work particularly well in solving MaOPs with regular Pareto fronts (PFs) that cover the whole objective space. Nonetheless, they become less efficient in handling MaOPs with irregular PFs, i.e., those with discontinuous, inverted, or degenerate PFs. One reason is that in solving such problems, some reference vectors in decomposition based algorithms may become inactive, in other words, with no solutions being associated, making the search process inefficient. For example in NSGA-III [5] or RVEA [6], each solution is associated to the reference vector that is closest to it according to the Euclidean distance or angle. Consequently, some reference vectors may have multiple solutions associated to them, while others may have no solution, resulting in inactive reference vectors.

In recent years, a large body of research has been dedicated to designing decomposition based MOEAs for solving irregular MaOPs and a detailed survey of MOEAs for handling irregular problems can be found in [15], [16]. Since no a priori information is given about the shape of the true Pareto fronts (PFs), a major difficulty arises when there is a mismatch between the distribution of the reference vectors and the shape of the PF. To address this challenge, either solutions in the current population as in MOEA/D-AM2M [17] or solutions in the archive as in MOEA/D-AWA [18], AdaW [19] and AR-MOEA [20], can be utilized to adapt the distribution of reference vectors after a given number of generations. Note that the reference vectors should not be adjusted too frequently and the solutions that survive for a consecutive number of generations can reflect the distribution of Pareto optimal solutions to some extent. This class of methods can ensure that the convergence speed is not seriously slowed down since the reference vectors are kept unchanged during a number of generations. Note, however, that adjusting reference vectors after a number of fixed generations may lead to the loss of promising solutions.

Another intuitive idea is to replace inactive reference vectors with newly generated ones during the search process [6], [21], [22]. As pointed out in [21], however, frequent adaptation or even maladaptation of reference vectors will seriously slow down the convergence. Thus, only one inactive reference vector is adjusted in each generation in [21]. In A-NSGA-III [22] and [22], multiple reference points are added around the active reference points if the number of active reference points is not enough. However, the above ways of adding new reference points may become ineffective if the objective space is very large, as it is unpractical to expect that newly generated solutions will always locate near the newly added reference points. As a result, newly added reference points may remain

Qiqi Liu is with the Department of Computer Science, University of Surrey, Guildford, Surrey GU2 7XH, UK. e-mail: qiqi.liu@surrey.ac.uk.

Yaochu Jin is with the Faculty of Technology, Bielefeld University, 33619 Bielefeld, Germany. He is also with the Department of Computer Science, University of Surrey, Guildforld, GU2 7XH, UK. Email: yaochu.jin@uni-bielefeld.de. (*Corresponding author: Yaochu Jin*)

Martin Heiderich is with Department of Advanced Vehicle Technology Research, Honda R&D Europe (Deutschland) GmbH, Carl-Legien-Strasse 30, D-63073 Offenbach/Main, Germany. e-mail: martin_heiderich@de.hrdeu.com.

Tobias Rodemann is with Department of Optimization and Creativity, Honda Research Institute Europe, Carl-Legien-Strasse 30, D-63073 Offenbach/Main, Germany. e-mail: tobias.rodemann@honda-ri.de.

inactive even if they locate in a promising region.

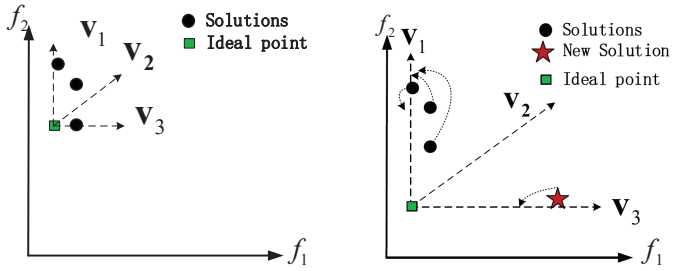
Furthermore, the best distribution of the reference vectors may be learned using machine learning methods. For instance, in MOEA/D-SOM [23], DEA-GNG [24], and RVEA-iGNG [25], the topology of approximated PFs is learned by training a self-organizing mapping (SOM) [26] and a growing neural gas network (GNG) [27], respectively. This class of methods does not require to determine which reference vectors should be adjusted and when, since the distribution of reference vectors are learned by training the SOM or GNG using the solutions in the population.

Apart from adjusting the distribution of reference vectors based on the framework of decomposition based algorithms, clustering based algorithms have also been widely used to deal with problems with various PF shapes. This can be attributed to the relative insensitivity of the clustering algorithms to the shapes of PFs, regardless whether the MaOP is regular or irregular. For instance, in EMyO/C [28] and CA-MOEA [29], hierarchical agglomerative clustering is adopted to cluster the non-dominated solutions according to the Euclidean distance between normalized solutions. Despite the great success achieved by the clustering based approaches such as [30], [31], it is well recognized, however, that it is nontrivial to specify the number of clusters beforehand, nor is it easy to determine which solution should be selected from each cluster to balance the proximity and diversity. It is also hard to determine the frequency of performing clustering since overly frequent clustering may slow down the convergence.

II. Motivation and Contributions

No matter whether we use reference vector based or clustering based methods to solve MaOPs, a key issue is the normalization of solutions. Before clustering the solutions or associating solutions to reference vectors, the solutions in the population need to be normalized because of different scales of objectives. In [32], the effect of the ideal point (the minimum value of each objective) for normalization on the performance of MOEA/D is analyzed. Since the ideal point and the scale of the objectives keep changing during the optimization, it becomes more challenging for the MOEAs to properly adapt the reference vectors. Fig. 1 plots an example of solutions and reference vectors in generation t and $t+1$. In Fig. 1 (b), the new solution denoted by a star will change the relative position of all solutions in the population. Consequently, no solutions can be associated with reference vector $\mathbf{v}_2$, and all solutions denoted by a solid circle will be associated with $\mathbf{v}_1$, resulting in the loss of some promising solutions. Therefore, even if the reference vectors are not adjusted, the changes in the relative position of the solutions within a population because of normalization during the search process may lead to the loss of promising solutions.

Preservation of solutions in promising regions can be even more challenging for solving irregular problems, since reference vectors may become inactive also due to the irregular shapes of true PFs. For irregular PFs, only part of the objective space



(a) generation t (b) generation $t+1$

Fig. 1: Loss of promising solutions caused by normalization. (a) Three solutions denoted by circles are associated with $\mathbf{v}_1$ to $\mathbf{v}_3$, respectively. (b) After a new solution (denoted by a star) emerges, no solutions can be associated with reference vector $\mathbf{v}_2$, and all solutions denoted by a solid circle will be associated with $\mathbf{v}_1$, resulting in the loss of some promising solutions.

is covered; thus, some of the predefined reference vectors covering the whole objective space may become inactive. That is to say, one needs to take both the normalization mechanism and the shape of the PFs into consideration in reference vector adaptation to preserve promising solutions or promising regions emerging in the search process. By promising solutions, we mean the solutions having better convergence and/or diversity properties. In this study, we aim to preserve these promising solutions to assure that they do not easily get lost, especially when the objective space is very large. If the promising solutions cannot be kept as soon as they emerge, promising search regions may get lost forever when handling irregular problems.

Faced with these challenges, some researchers propose to adjust the distribution of reference vectors after a number of fixed generations for dealing with irregular problems so that the reference vectors do not change too frequently. However, keeping the reference vectors unchanged even for a few generations may lose promising solutions that are newly generated within these generations, resulting in failures to generate reference vectors that can cover some unexplored PFs. On the other hand, the convergence speed will be decelerated if the reference vectors are adjusted whenever the number of active reference vectors is insufficient. Thus, when and where to add new reference vectors is extremely important to ensure that promising regions can be preserved without impairing the convergence performance. In summary, it is argued that the adaptation of the reference vectors is non-trivial, which should simultaneously consider the frequency of performing adaptation and where reference vectors should be adjusted. Thus, a mechanism that is able to adapt reference vectors to explore more promising regions without impairing the convergence performance is highly desirable.

Motivated by the distance threshold of nodes in the growing neural gas network [37], we propose to introduce a local angle threshold for each reference vector to determine when and where to add new reference vectors to tackle the above issues. The angle threshold of a reference vector is defined as the range in angle that the reference vector can cover. A new reference

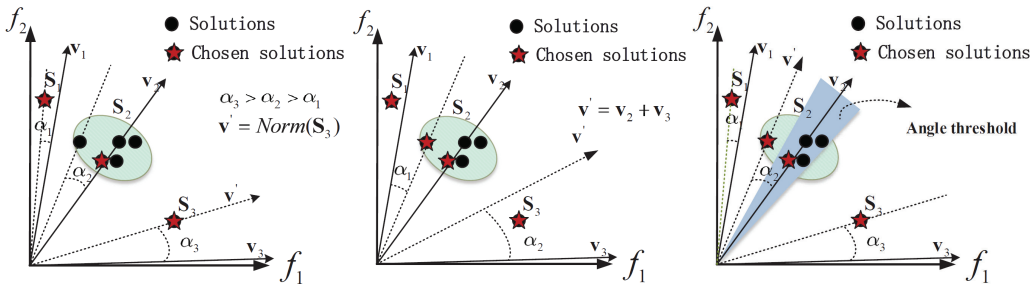


Fig. 2: Comparisons of different ways of adding a new reference vector: a) a solution $\mathbf{s}_3$ which is of the largest similarity in terms of angle or Euclidean distance is chosen as the new reference vector $\mathbf{v}'$ [21], [33], [34]; b) the new reference vector $\mathbf{v}'$ is generated by interpolating the existing active reference vectors $\mathbf{v}_2$ and $\mathbf{v}_3$ since these two have the largest angle between them [5], [35], [36]; c) the new reference vector $\mathbf{v}'$ is generated by the proposed angle threshold.

vector can be added only when there is a solution outside the range, i.e., when the acute angle between the solution vector and the reference vector is larger than the angle threshold. To better understand the advantages of using an angle threshold for adding new reference vectors, we compare it with other two commonly used methods. In Fig. 2, solutions in the population are denoted by solid circles and those that will be selected to the next generation are denoted by solid stars. Suppose three solution sets S_1, S_2 and S_3 are associated with three reference vectors $\mathbf{v}_1$, $\mathbf{v}_2$ and $\mathbf{v}_3$, respectively. Figs. 2 (a) and (b) illustrate two existing ways of adding a new reference vector, and Fig. 2 (c) shows the idea proposed in this work. In Fig. 2 (a), the solution that has the largest cosine distance to the existing active reference vectors $\mathbf{v}_1$ to $\mathbf{v}_3$, that is, solution S_3, will be used for creating a new reference vector $\mathbf{v}'$. It can be found, however, that the addition of $\mathbf{v}'$ will not have any influence on the assignment of the solutions because still only three solutions denoted by stars can be associated with a reference vector, even if the new reference vector $\mathbf{v}'$ is added. A similar situation can happen in Fig. 2 (b), where a new vector, $\mathbf{v}'$ is added by using the weighted sum of two active reference vectors $\mathbf{v}_2$ and $\mathbf{v}_3$, which have the maximum angle distance among all active reference vectors. By contrast, by using an angle threshold as proposed in this work, the solution outside the angle threshold of reference vector $\mathbf{v}_2$ will be chosen as a new reference vector $\mathbf{v}'$, as shown in Fig. 2 (c). The angle threshold of a reference vector proposed in this work utilizes information such as the number of solutions associated with this reference vector, the number of solutions associated with its two neighbouring reference vectors, the angle between this reference vector and its two neighbouring reference vectors, and the angle between this reference vector and its associated solutions. This means the angle threshold of each reference vector depends on the local distribution of both the reference vectors and solutions in its neighborhood. In Fig. 2 (c), as a result, one more solution will be associated with the newly generated vector, which will survive in the environmental selection. The main idea here is that the addition of new reference vectors should take into account the number of solutions associated with each reference vector and its neighboring reference vectors. For example, five solutions are associated with $\mathbf{v}_2$, meaning that the region near $\mathbf{v}_2$ is very promising. We should allocate more search resources to the region near $\mathbf{v}_2$.

For decomposition based algorithms, the design of scalarizing functions should take both the convergence and diversity into consideration, since eventually, the scalarizing functions will be used to rank the solutions before selection based on the reference vector with which the solutions are associated. Even though many scalarizing functions are proposed over the past decade, little work has been reported that coordinates the adaptation of the reference vectors and the adaptation of the scalarizing function. Note that most scalarizing functions are designed based on a set of predefined fixed reference vectors. If the reference vectors are frequently adjusted, a solution in the population may change its associated reference vector very frequently, resulting in reduced selection pressure along a particular search direction and slowing down the convergence. For example, the selection pressure for convergence along a reference vector should be emphasized if the reference vector has not been adjusted and has been active in many generations. By contrast, diversity should be emphasized for a newly generated reference vector. Therefore, in this work, to preserve the promising solutions as soon as they emerge in dealing with irregular problems, we propose a new adaptive scalarizing function tailored for adaptive reference vectors so that the adaptation of the weights in the scalarizing function can take into account the age and angle threshold of each reference vector, thereby better balancing the trade-off between convergence and diversity.

The contributions of this work are summarized as follows.

- An angle threshold specific to each reference vector is designed to determine when and where new reference vectors should be inserted so that more regions can be explored and promising regions can be exploited whenever they emerge. To further enhance the exploratory capability of the proposed algorithm, inactive or overly crowded reference vectors will be deleted in a predefined frequency.
- A new adaptive scalarizing function dedicated to adaptive reference vectors is proposed. Both the angle threshold and the age of reference vectors are considered in the scalarizing function. This way, a coordinated adaptation of the reference vectors and the scalarizing function can

be achieved and a better balance between convergence and diversity can be realized.

The rest of this paper is organized as follows. Section III describes the proposed algorithm in greater detail and Section IV presents the experimental results by comparing the proposed algorithm with eight state-of-the-art algorithms. Finally, Section V concludes this paper and discusses promising future research directions.

III. Proposed Algorithm

The reference vectors in the initialization are generated using the canonical simplex-lattice design method [38]. As discussed in Section II, a new reference vector is typically generated based on the solution whose angle with the reference vector it is associated is particularly large [21], or by interpolating two existing reference vectors that have a large angle between them [35]. The potential problem of these methods is that they may fail to preserve newly generated promising solutions mainly due to two reasons. First, a newly generated reference vector may be far from newly generated promising solutions due to the high-dimensional objective space. Second, the scalarizing function is not able to give a higher priority to the newly generated reference vectors in environmental selection. To address the above problems, we hypothesize that the adaptation of the reference vectors should be coordinated with that of the scalarizing functions. To this end, an MOEA with coordinated adaptation of reference vectors, termed as CARV-MOEA is proposed. CARV-MOEA consists of four main components: reproduction, reference vector adaptation, scalarizing function adaptation, and environmental selection, as shown in Algorithm 1.

Algorithm 1: General Framework of CARV-MOEA.

Input : Population P, Maximum generations t_{max}, The frequency of adapting reference vectors f_r
Output : The final Population P

Initialize: reference vectors $V = \{\mathbf{v_1}, \mathbf{v_2}, \cdots, \mathbf{v_N}\}$, $o = (1, 1, \cdots, 1)$, f_r = 0.05;
while $t < t_{max}$ **do**
  /* Reproduce the offspring O */
  $O \leftarrow$ Reproduction(P);
  $P \leftarrow P \cup O$;
  /* Adapt reference vectors V according to proposed angle threshold */
  $V, \Phi \leftarrow$ Reference Vector Adaptation(V, P, f_r);
  /* Environmental Selection */
  $P \leftarrow$ Scalarizing Function Adaptation and Environmental Selection(P, V,Φ);
  $t \leftarrow t + 1$;
end

1) Reproduction: Offspring O are generated using the simulated binary crossover (SBX) [39] and polynomial mutation [40] operator.
2) Reference vector adaptation: This step consists of angle threshold based reference vector generation, and deletion of inactive or overly crowded reference vectors.
3) Scalarizing function adaptation and environmental selection: The scalarizing function is adjusted based on the angle threshold and age of the reference vectors and then environmental selection is performed.

Next, the details of reference vector adaptation, scalarizing function adaptation and environmental selection in CARV-MOEA will be presented.

A. Reference Vector Adaptation

Since the adaptation of reference vectors heavily depends on the normalization of solutions, we propose to keep the scale of different objectives unchanged for a number of fixed generations to avoid too frequent changes of the relative position between the normalized solutions and the reference vectors. For this purpose, we normalize the solutions in the population as follows:

$$f_j^{'} = (f_j - f_j^*)/o_j, \tag{1}$$

where f_j is the j-th objective value, f_j^* is the minimum value of the j-th objective among all non-dominated solutions, and $\mathbf{o}$ is obtained by the scale of the objectives of the non-dominated solutions obtained during every f_r generations ($o_j = f_j^{max} - f_j^*$, where f_j^{max} and f_j^* are set as the maximum and minimum value of all non-dominated solutions.) Note that $\mathbf{o}$ is changed in every $f_r \cdot t_{max}$ generations only, where f_r is the frequency of adjusting the $\mathbf{o}$ and t_{max} is the maximum number of generations. Note that once $\mathbf{o}$ is calculated, it remains unchanged for the next $f_r \cdot t_{max}$ generations before it is recalculated. The reason is that the relative position of the normalized solutions will change frequently if $\mathbf{o}$ is calculated in each generation, making the adaptation of reference vectors more difficult. Thus, in this study, we propose to keep $\mathbf{o}$ unchanged in every $f_r \cdot t_{max}$ generations to strike a better balance between convergence and diversity.

1) Angle threshold based reference vector generation: One main challenge in reference vector adaptation is to determine when reference vectors should be adjusted and where new reference vectors should be inserted. Inappropriate adjustment of the reference vectors may slow down the convergence, reduce the population diversity or even mislead the search process. Thus, in this work, we propose an angle threshold that is specific to each reference vector to decide when to add a new reference vector so that newly generated promising solutions are most likely to survive during the search process.

In this study, a solution is associated with the reference vector with which it has the maximum cosine similarity. For the proposed angle threshold based reference vector generation method, the main idea here is that a new reference vector should be generated only if the angle between a solution and its associated reference vector is larger than a threshold that is dependent on the density of the solutions around the reference vector, i.e., the number solutions associated with the

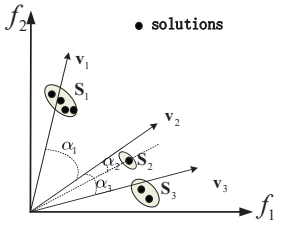


Fig. 3: An example illustrating the definition of the angle threshold. Calculation of the angle threshold of $\mathbf{v}_2$ requires the information of its neighbouring reference vectors $\mathbf{v}_1$ and $\mathbf{v}_3$, including the angle between reference vectors and the number of solutions associated with $\mathbf{v}_1$ to $\mathbf{v}_3$.

reference vector and the angles between these solutions and the reference vector, and the angles between the vector with its M neighbouring reference vectors, where M is the number of objectives. Specially, the angle threshold Φ_j of reference vector $\mathbf{v}_j$ is defined as follows:

$$\Phi_j = \frac{\sum_{\mathbf{s}\in S_j} \alpha(\mathbf{s},\mathbf{v}_j) + \sum_{i\in\{j_1,j_2,\ldots,j_M\}} |S_i|\,\alpha(\mathbf{v}_i,\mathbf{v}_j)\cdot(\frac{N_{av}}{N})^k}{|S_j| + \sum_{i\in\{j_1,j_2,\ldots,j_M\}} |S_i|} \tag{2}$$

where S_j are the solutions associated with $\mathbf{v}_j$, $\alpha(\mathbf{s},\mathbf{v}_j)$ is the angle between a solution $\mathbf{s}$ and $\mathbf{v}_j$ ($\mathbf{s}\in S_j$), $j_1, j_2, \ldots, j_M$ are the index of M nearest neighbouring reference vectors of reference vector $\mathbf{v}_j$, $|S_i|$ ($i\in\{j_1,j_2,\ldots,j_M\}$) is the number of solutions associated with each of the neighbouring reference vectors. Note that if several neighboring reference vectors are on the same side of reference vector $\mathbf{v}_i$, then only one of the neighboring reference vectors will be randomly selected. $\alpha(\mathbf{v}_i,\mathbf{v}_j)$ ($i\in\{j_1,j_2,\ldots,j_M\}$) is the angle between reference vector $\mathbf{v}_j$ and each of its neighbouring reference vectors $\mathbf{v}_{j_1}$ to $\mathbf{v}_{j_M}$. N_{av} is the number of active reference vectors associated with all non-dominated solutions and N is the predefined population size. k is a user-defined parameter and k is set to 2 in this study. $(\frac{N_{av}}{N})^k$ is used to penalize the angle between the active reference vectors $\alpha(\mathbf{v}_i,\mathbf{v}_j)$ in case the number of active reference vectors N_{av} is far less than the population size N.

Fig. 3 gives an example to illustrate the definition of the angle threshold for a 2-objective problem. In this example, the angle threshold of reference vector $\mathbf{v}_2$ is equal to $\frac{1*\alpha_2+4*\alpha_1+2*\alpha_3}{1+4+2}$. Among them, 1, 4, 2 are the number of solutions associated with $\mathbf{v}_2$, and its two neighbouring reference vectors $\mathbf{v}_1$ and $\mathbf{v}_3$, respectively. α_2 is the angle between the solution $\mathbf{s}_2$ and $\mathbf{v}_2$. α_1 and α_3 are the angle between $\mathbf{v}_1$ and $\mathbf{v}_2$, and $\mathbf{v}_3$ and $\mathbf{v}_2$, respectively.

Note that the angle threshold of $\mathbf{v}_2$ is weighted by both the number of solutions associated with $\mathbf{v}_2$, the angles between $\mathbf{v}_2$ and its M neighbors, and the number of solutions associated with its neighbouring reference vectors. As plotted in Fig. 2 (c), suppose seven solutions denoted by solid circles and star are associated with $\mathbf{v}_1$, $\mathbf{v}_2$ and $\mathbf{v}_3$, respectively, among which five solutions inside the oval are associated with $\mathbf{v}_2$. Since most of solutions are associated with $\mathbf{v}_2$, the angle between these five solutions and $\mathbf{v}_2$ is more important than the angle between $\mathbf{v}_2$ and $\mathbf{v}_1$ and the angle between $\mathbf{v}_2$ and $\mathbf{v}_3$. As a result, the angle threshold of $\mathbf{v}_2$ will be smaller. In Fig. 2 (c), the angle threshold is indicated by the shaded area and we can find that the solution denoted by the star is outside the angle threshold of $\mathbf{v}_2$. Consequently, the solution denoted by the star will be used to generate a new reference vector. By defining this angle threshold, we are able to determine when and where a new reference vector should be added according to the location of a solution and the distribution of the solutions as well as the vectors in its neighborhood, thereby improving the effectiveness of the newly added reference vectors.

To more clearly show rationale behind the proposed angle threshold, we have rigorously proven that the proposed angle threshold can ensure that only solutions with a large contribution to the diversity will have a chance to be selected for generating a new reference vector. We have derived the condition based on the angle threshold to generate new reference vectors in Equation 1 to 3 in Section I in the Supplementary material.

The pseudo code of the angle threshold based reference vector adaptation is given in Algorithm 2. Firstly, all solutions in the current population are sorted by the non-dominated sorting, and then N solutions (N is the population size) are randomly chosen for calculating the angle threshold of each reference vector, as given in Lines 1 to 6 in Algorithm 2. A solution vector will be added as a new reference vector if the angle between this solution and its associated reference vector $\mathbf{v}_t^a$ is larger than the angle threshold Φ_t, as described in Lines 8 to 18 in Algorithm 2.

2) Deletion of inactive and crowded reference vectors: As mentioned above, new reference vectors can be added if they satisfy the angle threshold defined for each reference vector. Meanwhile, we also delete inactive reference vectors and those in overly crowded regions according to the angle between the reference vectors, when the total number of reference vectors is larger than the population size. In every $f_r \cdot t_{max}$ generations, redundant reference vectors are deleted after new reference vectors are generated. At first, all inactive reference vectors are deleted. If the number of active reference vectors is still larger than the population size, reference vectors will be deleted one by one according to the angle between the reference vectors. To this end, we calculate the angles between all reference vectors, and the one with the minimum angle will be deleted. Then we recalculate the angle between the remaining reference vectors and repeat this process until the number of reference vectors reaches the population size. Note that reference vector deletion is carried out in every $f_r \cdot t_{max}$ generations only, when $\mathbf{o}$ in Equation (1) for solution normalization is also updated. After deleting the inactive or overly crowded reference vectors or adding new reference vectors according to the angle threshold, we will recalculate the angle threshold for each reference vector, as described in Lines 26 to 28 in Algorithm 2.

B. Adaptation of Scalarizing Function and Environmental Selection

Various scalarizing functions have been proposed for decomposition based MOEAs, such as the weighted sum, Tchebycheff,

Algorithm 2: Reference vector adaptation.

Input : Population P, reference vectors V, generation number t, maximum generations t_{max}
Output : The reference vectors V, angle threshold Φ, the scale of reference vectors $\mathbf{o}$

/* F_1^P denotes the non-dominated solutions in P */
$F_1^P \leftarrow$ Pareto non-dominated sorting(P);
if $\left|F_1^P\right| > N$ **then**
  $P_a \leftarrow$ Randomly select N solutions from F_1^P;
else
  $P_a \leftarrow$ Randomly select N solutions from P ;
end
$V^a \leftarrow$ Find the reference vectors that are associated with solutions in P_a;
/* Insert new reference vectors based on the proposed angle threshold */
for *t=1:size(V^a)* **do**
  $S \leftarrow$ Find solutions associated with $\mathbf{v}_t^a$;
  $\mathbf{v}_{t_1}, \mathbf{v}_{t_2}, \cdots, \mathbf{v}_{t_M} \leftarrow$ Find M nearest neighbouring reference vectors of V_t^a;
  $\Phi_t \leftarrow$ Calculate the angle threshold of $\mathbf{v}_t^a$;
  for *j=1:size(S)* **do**
    if $angle(\boldsymbol{s}_j, \boldsymbol{v}_t^a) > \Phi_t$ **then**
      $\mathbf{v}_{new} \leftarrow$ Normalize the solution $\mathbf{s}_j$;
    end
    $V \leftarrow V \cup \mathbf{v}_{new}$;
  end
end
/* Delete inactive reference vectors generated over every $f_r \cdot t_{max}$ generations */
if *mod(t,$f_r \cdot t_{max}$)==0* **then**
  $V \leftarrow$ Identify active reference vectors that the age is bigger than zero over a period;
  if $size(V) > N$ **then**
    $V \leftarrow$ Delete crowed reference vectors by angle(V);
  end
  $\mathbf{o} \leftarrow$ Recalculate the scale.
end
/* Recalculate the angle threshold. */
if *$\boldsymbol{v}_{new}$ is not empty or mod(t,$f_r \cdot t_{max}$)==0* **then**
  $\Phi \leftarrow$ Repeat step 8 to step 18 to recalculate the angle threshold of each active reference vector ;
end

penalty-based boundary intersection (PBI) in MOEA/D [4], and their variants [41], [42]. In [43], the weighted sum and weighted Tchebycheff scalarizing functions are adaptively adopted for each individual considering that the convexity of PFs differs over the whole objective space. In MOEA/D-PaS [44], p value in L_p scalarizing function is adapted for each weight vector according to the PF geometry. In the angle penalized distance (APD) in RVEA [6], a scalarizing function with a generation-varying parameter is introduced to adapt the priority over convergence and diversity. An adaptive hybridation of different scalarizing functions is presented in [45]. However, most of the above scalarizing functions are designed for a fixed set of reference vectors, and they may fail to perform properly if the reference vectors are frequently adjusted. This is because if a solution is frequently associated with different reference vectors, it will then be compared with different solutions, reducing the selection pressure towards a particular search direction. To ensure the promising solutions to be selected without severely slowing down the convergence speed, we propose an adaptive scalarizing function, called adaptive angle penalized distance (A-APD) that is able to take the adaptation of the reference vectors into account.

As its name suggests, A-APD is a variant of the angle penalized distance [6], which is defined as follows:

$$d_{i,j} = (1 + (1 - \frac{\rho_j^i}{\Delta}) \cdot \frac{\alpha_{i,j}}{\Phi_j} \cdot \frac{\gamma_j}{\gamma_{\max}}) \cdot \left\|\hat{\mathbf{f}}^i\right\| \tag{3}$$

and

$$\gamma_j = \begin{cases} T \cdot \Phi_j & \text{if reference vector } j \text{ is new} \\ & \text{or has only one solution} \\ \min_{i' \subset 1,2,\ldots N} \alpha_{\mathbf{v}_{i'}, \mathbf{v}_j} & \text{if otherwise} \end{cases} \tag{4}$$

where i is a solution index, representing one of the solutions associated to the j-th reference vector, $\alpha_{i,j}$ is the angle between solutions i and reference vector j, and Φ_j is the angle threshold of reference vector j. γ_j is the smallest angle between reference vector j and all pairs of reference vectors, $\gamma_{\max}$ is the maximum of all γ_j. $\left\|\hat{\mathbf{f}}^i\right\|$ represents the Euclidean distance from the solutions in the population to the ideal point $\mathbf{f}_*$, that is $\hat{\mathbf{f}}^i = \mathbf{f}^i - \mathbf{f}_*$. In the above definition, ρ_j^i is the age of the reference vector j associated with solution i, which counts the number of generations in which the reference vector is active. $\Delta = f_r \cdot t_{max}$, where f_r is the frequency of deleting inactive reference vectors and t_{max} is the maximum number of generations. The age of the reference vector j, i.e., ρ_j^i, starts from zero and increases by one if there are solutions associated with it in every generation. If there is no solutions to be associated with the reference vector, then the age is not changed. The larger ρ_j^i is, the larger the age of the reference vector j is. We assume that the convergence should be more emphasized for an reference vector with a larger age. Thus, $1 - \frac{\rho_j^i}{\Delta}$ is used as a penalization term.

Note from Equation (4) that there are two scenarios for determining γ_j. For newly generated reference vectors or for those having only one solution associated to them, at least two solutions will be identified for environmental selection, which is achieved by enlarging the range of solution association by T times, where T is a parameter depending on the location of the assigned solutions. In this case, $\gamma_j = T \cdot \Phi_j$, which is meant to make sure that very poorly converged solutions will be unlikely to survive. For other active reference vectors, γ_j is the smallest angle between reference vector j and all other reference vectors.

From the definition of A-APD, we can make the following observations. First, an age is introduced for reference vectors to

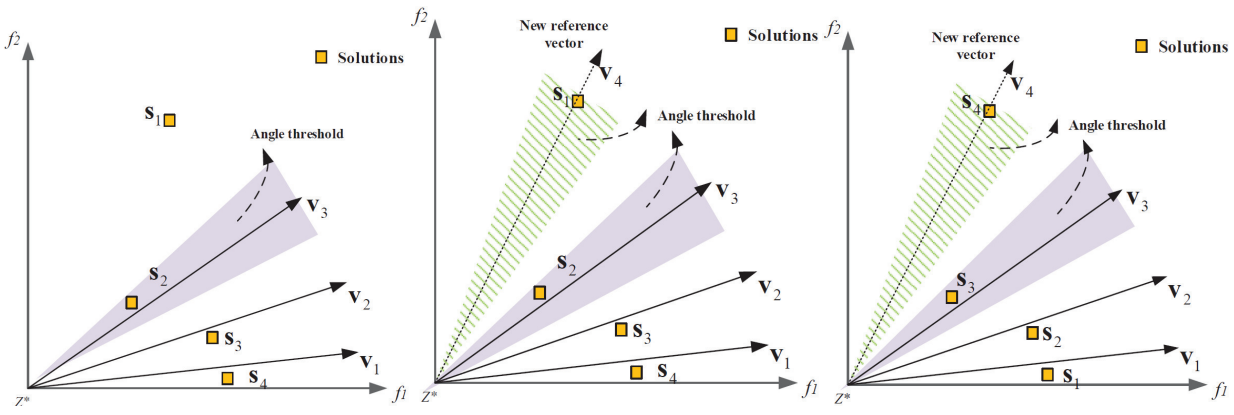


Fig. 4: An illustrative example of how A-APD works : a) a solution $\mathbf{s}_4$ is outside the angle threshold of $\mathbf{v}_3$. b) The solution vector of $\mathbf{s}_4$ is used to generate a new reference vector $\mathbf{v}_4$. $\mathbf{s}_4$ is a very poorly converged solution. $\mathbf{s}_4$ is not going to be selected based on the A-APD. c) $\mathbf{s}_4$ is a well converged solution. $\mathbf{s}_4$ will be selected based on the design of A-APD.

increase the selection pressure along those stable (consistently active) reference vectors to accelerate convergence. In addition, the angle threshold Φ_j of reference vector j is used for normalizing the angle between a solution and the reference vector $\alpha_{i,j}$. This indicates that a priority will be given to diversity, if $\alpha_{i,j}$ is larger than the threshold Φ_j, meaning that solution i is outside the angle threshold. Finally, an adaptation of the selection pressure is included, making sure that very poorly converged solutions associated to a newly generated reference vectors or those having a single solution associated with will not be selected to enhance convergence.

We give an illustrative example here to show why the proposed adaptive scalarizing function based on the age and the angle threshold of the reference vectors is effective. As shown in Fig. 4 (a), there are three reference vectors, $\mathbf{v}_1$ to $\mathbf{v}_3$, and four solutions $\mathbf{s}_1$ to $\mathbf{s}_4$, of which $\mathbf{s}_3$ and $\mathbf{s}_4$ are associated with $\mathbf{v}_3$ according to their cosine similarity to the vectors. Based on the definition of the angle threshold, since $\mathbf{s}_4$ is outside the angle threshold of $\mathbf{v}_3$, thus, solution vector $\mathbf{s}_4$ is used to generate a new reference vector $\mathbf{v}_4$. Note that if the angle between one solution and its associated reference vector is larger than the calculated angle threshold, then we say the solution is outside the angle threshold of the reference vector. In Fig. 4 (b), the angle threshold of $\mathbf{v}_4$ is highlighted with the shade region. Based on Equation (3), the A-APD value of $\mathbf{s}_4$ is $\left\|\hat{\mathbf{f}}^{\mathbf{s}_4}\right\|$ without angle penalization, since the angle between $\mathbf{s}_4$ and $\mathbf{v}_4$ (that is $\alpha_{4,4}$) is zero. In Equation (4), if a reference vector is new or has one solution associated with it only, at least two solutions will be identified for environmental selection. Thus, in Fig. 4 (b) and (c), $\mathbf{s}_3$ will be compared with $\mathbf{s}_4$ in terms of the A-APD value. In Fig. 4 (b), the A-APD value of $\mathbf{s}_3$ is penalized with the relatively large angle value since the angle between $\mathbf{s}_3$ and $\mathbf{v}_4$ ($\alpha_{3,4}$) is larger than the angle threshold of $\mathbf{v}_4$ (Φ_4). Suppose $\mathbf{s}_4$ is a poorly converged solution in Fig. 4 (b), and $\mathbf{s}_4$ will need to be compared with $\mathbf{s}_3$ based on Equation (4). Compared with $\mathbf{s}_3$, the very poorly converged solution $\mathbf{s}_4$ will have little chance to survive in environmental selection since the A-APD value of $\mathbf{s}_4$ is still much larger than that of $\mathbf{s}_3$, because $\left\|\hat{\mathbf{f}}^{\mathbf{s}_4}\right\|$ is much larger than $\left\|\hat{\mathbf{f}}^{\mathbf{s}_3}\right\|$. In this case, we can ensure that very poorly converged solutions will have no chance to be selected. Moreover, it is also important to make sure that well converged solutions associated with new reference vectors will be selected by comparing them with their neighboring solutions. For example, in Fig. 4 (c), $\mathbf{s}_4$ and $\mathbf{s}_3$ are both well converged. The Euclidean distance of $\mathbf{s}_3$ to the ideal point is similar to that of $\mathbf{s}_4$, however, since the A-APD value of $\mathbf{s}_3$ is penalized with the angle value, the A-APD value of $\mathbf{s}_3$ will be larger than that of $\mathbf{s}_4$, making the well converged $\mathbf{s}_4$ have a larger chance to be selected.

The pseudo code for environmental selection based on A-APD is presented in Algorithm 3. Line 6 to Line 20 present the A-APD selection criterion. The solutions from the first layer to the critical layer will involve in the selection process. If the front F_k satisfies the condition $|F_1 \bigcup F_2 \bigcup \cdots F_{k-1}| < N$ and $|F_1 \bigcup F_2 \bigcup \cdots F_k| \geqslant N$ (N is the population size), we define F_k as the critical layer. Only one solution will be selected for each active reference vector. Note that the number of active reference vectors may be larger than the population size N. Thus, a number of solutions will be deleted according to the angle after the A-APD selection process, which is shown in Line 21 to Line 22.

IV. Experimental Results

This section presents the empirical results to demonstrate the competitiveness of CARV-MOEA by comparing it with eight state-of-art algorithms on both regular and irregular MaOPs, namely VaEA [46], DEA-GNG [24], RVEA$*$ [6], AdaW [19], MOEA/D-AM2M [17], MOEA/D-SOM [23], MaOEA/SRV [47] and CA-MOEA [29]. Among these algorithms, DEA-GNG, RVEA$*$ are based on two sets of reference vectors: one is fixed and the other is adaptive. VaEA does not use reference vectors and is designed for solving both regular and irregular problems. AdaW and MOEA/D-AM2M are both based on a set of adaptive reference vectors for solving irregular problems, and an archive or the current population is preserved for determining the distribution of adaptive reference vectors after a number

Algorithm 3: Adaptive scalarizing function and environmental selection.

Input : Population P, Reference Vectors V, angle threshold Φ
Output : Population P_o

```
/* F_1^P, F_2^P, ..., F_c^P denotes the solutions from the
   first layer to the critical layer                    */
F_1^P ⋃ F_2^P, ..., ⋃ F_c^P ← Non-dominated sorting(P);
Associate the solutions in F_1^P ⋃ F_2^P, ..., ⋃ F_c^P with
 the reference vectors;
Increase the age ρ of each active reference vector by 1;
α(S, V) ← Calculate the angle between solutions and
 reference vectors(F_1^P, F_2^P, ..., F_c^P, V);
γ ← Calculate the angle between reference vectors(V);
for j=1:size(V) do
    S ← Identify the solutions associated with v_j;
    /* enlarging the range of solution
       association by T times                           */
    if size(S)== 1 or v_j is new added then
        T = 1;
        S_neighbor ← empty;
        while S_neighbor is empty do
            T ← T + 1;
            γ_j ← T * Φ_j;
            /* If the angle between one solution
               and its associated reference
               vector is smaller than the
               calculated angle threshold, then
               we say the solution is within the
               angle threshold of the reference
               vector                                   */
            S_neighbor ← find solutions within the
             angle γ_j;
        end
    end
    S ← S ∪ S_neighbor;
    /* A-APD is defined in Equation (3)                 */
    s_b ← Pick up a solution with the best A-APD
     values among S;
    P_o ← P ∪ s_b;
end
P_o ← Delete duplicated solutions (P_o) ;
P_o ← Delete crowded solutions (P_o) ;
```

of fixed generations. MaOEA/SRV [47] adopts a set of self-guided reference vectors in each generation to guide the search, which has been shown to be very competitive in handling irregular problems. The distribution of the reference vectors in MOEA/D-SOM and DEA-GNG are learned by training a self-organizing mapping network or a growing neural gas network, which does not need to determine when and where to adjust the reference vectors, showing competitive performance on handling irregular problems.

Three suites of test problems are considered for testing the compared algorithms, i.e., MaF benchmarks (MaF1-15) [48], DTLZ benchmarks (DTLZ1-4) [49], IDTLZ2 [36] and DPF benchmarks (DPF1-5) [50]. We also test the performance of the proposed CARV-MOEA on a real world application, hybrid electric vehicles (HEVs). HEVs has 7 objectives and 11 decision variables. All algorithms under comparison except for MaOEA/SRV are implemented on the PlatEMO platform [51] with MATLAB 2018a on Intel Core i7-8700 (3.20GHz). MaOEA/SRV is implemented in Java. The code of CARV-MOEA is uploaded to https://github.com/qiqi6770304/CARV_MOEA.git.

A. Parameter Settings

The proposed CARV-MOEA and all compared algorithms adopt SBX [39] as the crossover operator with a crossover probability $P_c = 1.0$ and a distribution index of crossover $n_c = 20$ except for MaOEA/SRV, in which $n_c = 30$, and the polynomial mutation [40] as the mutation operator with a mutation probability $P_m = 1/D$ (D is the dimension of the decision space), and a distribution index of mutation $n_m = 20$.

The population size for 5-, 10- and 15-objective problems are set to 126, 230 and 240 for CARV-MOEA and the compared algorithms except for MOEA/D-SOM. For MOEA/D-SOM, the population size is set to 144, 256 and 256, respectively. The maximum number of fitness evaluations for problems with 5, 10 and 15 objectives is set to $M * 1e4$, respectively, for all test functions, where M is the number of objectives. The frequency of adjusting the scale of different objectives f_r is set to 0.05. Parameter k in the angle threshold is set to 2. The sensitivity analysis of f_r and k is given in Section III of the Supplementary material. The population size set for HEVs is 112 and the maximum number of fitness evaluations for HEVs is set to 30000.

B. Performance Indicators

Two widely used performance indicators, hypervolume (HV) [52] and inverted generational distance plus (IGD$^+$) [53] are adopted for comparison. IGD$^+$ and HV can both be used to account for convergence and diversity. For calculating the IGD$^+$ indicator, a set of uniformly distributed reference points sampled from the true PFs are needed. We use the canonical simplex-lattice design method [38] to generate 10,000 reference points for calculating IGD$^+$ indicator, as in PlatEMO [51]. In this study, IGD$^+$ is used for evaluating the performance of algorithms as the first performance indicator and HV is served as the secondary performance indicator. The comparative results in terms of the IGD$^+$ values on regular problems is presented in Table SI in the Supplementary material. The comparative results in terms of the HV values on regular and irregular problems are presented in Table SII and SIII in the Supplementary material for reference. Note that all solutions are normalized by $1.1\,\mathbf{z}_{nadir}$ and the reference point for calculating the HV value is set to $(1, 1, ..., 1)$ for problems except for MaF1, IDTLZ2 and MaF4. As suggested in [54], since MaF1 has the same PF as IDTLZ1 [36], the reference point for MaF1 is set to $1+\frac{1}{H}$, where H is the number of

TABLE I: IGD$^+$ VALUES OF THE SOLUTION SETS OBTAINED BY NINE ALGORITHMS ON IRREGULAR TEST INSTANCES. THE BEST RESULT ON EACH TEST INSTANCE IS HIGHLIGHTED IN DARK GREY.

Shape of PFs	Problem	M	D	CARV-MOEA	DEA-GNG	CA-MOEA	VaEA	RVEA*	AdaW	MOEA/D-AM2M	MOEA/D-SOM	MaOEA/SRV
inverted	MaF1	5	14	8.5062e-2	8.7768e-2 −	1.0557e-1 −	9.1329e-2 −	1.2807e-1 −	8.4376e-2 +	1.2816e-1 −	2.0409e-1 −	8.4215e-2 +
		10	19	1.5926e-1	1.7110e-1 −	2.0869e-1 −	1.6583e-1 −	2.7853e-1 −	1.6429e-1 −	3.2512e-1 −	3.0841e-1 −	1.5971e-1 ≈
		15	24	1.9221e-1	2.1414e-1 −	2.5101e-1 −	2.0101e-1 −	3.4280e-1 −	1.9738e-1 −	3.4953e-1 −	2.8524e-1 −	1.9025e-1 +
	IDTLZ2	5	14	9.9964e-2	1.1196e-1 ≈	1.3522e-1 −	1.2065e-1 −	1.1694e-1 −	1.0779e-1 −	1.3139e-1 −	2.7307e-1 −	1.1359e-1 −
		10	19	2.2221e-1	2.4809e-1 −	4.1981e-1 −	2.7255e-1 −	2.6222e-1 −	2.5469e-1 −	3.5543e-1 −	4.1452e-1 −	2.2861e-1 −
		15	24	2.9882e-1	3.8220e-1 −	6.6241e-1 −	3.4004e-1 −	3.6854e-1 −	3.2396e-1 −	4.4364e-1 −	4.4901e-1 −	3.2053e-1 −
	MaF4	5	14	1.3400e+0	1.4703e+0 −	2.9718e+0 −	1.7868e+0 −	7.8799e-1 +	1.3875e+0 −	4.2391e+1 −	3.9728e+0 −	1.4474e+0 −
		10	19	1.5183e+1	2.9429e+1 −	2.1477e+2 −	1.7414e+1 −	9.3343e+0 +	1.7473e+1 −	9.9954e+1 −	2.5477e+1 −	2.9219e+1 −
		15	24	3.1978e+2	8.6420e+2 −	6.8301e+3 −	4.4098e+2 −	1.7063e+2 +	4.7777e+2 −	5.2425e+2 −	5.8270e+2 −	8.9977e+2 −
discontinuous	MaF7	5	24	1.2912e-1	1.3501e-1 ≈	2.4744e-1 −	1.9111e-1 −	1.4596e-1 −	1.5432e-1 −	3.2754e-1 −	1.2681e+0 −	1.8976e-1 −
		10	29	5.5002e-1	5.9114e-1 ≈	2.1217e+0 −	8.6838e-1 −	7.7203e-1 −	8.2637e-1 −	8.4388e-1 −	3.8417e+0 −	1.4787e+0 −
		15	34	1.0681e+0	4.8306e+0 −	1.3393e+1 −	1.4978e+0 −	1.2990e+0 −	6.4073e+0 −	1.5533e+0 −	1.0344e+1 −	6.4577e+0 −
	MaF11	5	14	1.2750e-1	1.7445e-1 −	3.5120e-1 −	1.8693e-1 −	1.5685e-1 −	2.5300e-1 −	1.9861e-1 −	1.9042e+0 −	1.2359e-1 +
		10	19	1.9680e-1	5.4820e-1 −	7.5030e-1 −	3.0715e-1 −	1.6876e-1 +	7.2778e-1 −	3.5574e-1 −	1.8190e+0 −	2.8462e-1 −
		15	24	2.0047e-1	1.3095e+0 −	1.0538e+0 −	3.9765e-1 −	1.9946e-1 +	1.0078e+0 −	3.8493e-1 −	4.3465e+0 −	4.1667e-1 −
Pareto box	MaF8	5	2	6.5497e-2	7.7107e-2 −	7.0513e-2 −	6.6564e-2 −	2.8134e-1 −	1.0300e-1 −	2.6965e-1 −	6.5657e-1 −	6.9383e-2 −
		10	2	7.5299e-2	8.8618e-2 −	7.6296e-2 ≈	7.4843e-2 ≈	6.4748e-1 −	7.8518e-2 ≈	4.5476e-1 −	1.0935e+0 −	7.0633e-2 +
		15	2	9.1473e-2	1.0130e-1 −	9.3871e-2 −	9.5303e-2 −	8.4391e-1 −	8.8318e-2 +	5.1935e-1 −	9.8047e-1 −	8.6112e-2 +
	MaF9	5	2	1.3563e-1	3.0405e-1 −	1.8325e-1 −	3.4557e-1 −	2.1043e-1 −	1.8894e-1 ≈	1.8891e-1 −	1.7299e+0 −	5.9357e-1 −
		10	2	6.8620e-1	5.0061e-1 +	2.9959e+1 −	1.3957e-1 +	6.4961e-1 ≈	1.1629e+1 −	2.8566e-1 +	2.0458e+0 −	2.0707e-1 +
		15	2	1.2180e-1	2.6622e-1 −	2.4032e-1 −	1.7641e-1 −	7.8249e-1 −	8.3170e-1 −	6.9765e-1 −	4.7149e+0 −	3.5086e-1 −
degenerate	MaF6	5	14	1.9893e-3	1.4368e-3 +	1.6644e-3 +	1.8853e-3 +	1.2418e-2 −	1.4839e-3 +	3.5676e-3 −	2.6906e-2 −	3.2647e-3 −
		10	19	1.9172e-1	1.5110e+0 −	4.0019e+0 −	7.5322e-1 −	2.2362e-2 ≈	8.5844e-4 +	1.5001e-3 +	7.9989e-3 ≈	6.5856e-1 −
		15	24	2.8296e-1	2.1227e+1 −	6.7430e+0 −	4.0222e-1 −	5.9677e-2 +	8.4888e-4 +	1.1987e-3 +	3.0643e-2 +	4.1965e-1 −
	MaF13	5	5	1.2590e-1	1.4939e-1 −	1.0187e-1 +	1.2561e-1 ≈	4.0283e-1 −	6.3177e-2 +	9.2191e-2 +	1.6446e-1 −	1.0093e-1 +
		10	5	1.4283e-1	1.4608e-1 ≈	7.8926e-2 +	1.0021e-1 +	2.6393e-1 −	5.1029e-2 +	1.1130e-1 +	1.2017e-1 +	9.0302e-2 +
		15	5	1.3732e-1	1.4222e-1 ≈	7.6691e-2 +	1.0617e-1 +	3.1331e-1 −	4.7576e-2 +	1.1943e-1 +	1.3057e-1 ≈	8.7151e-2 +
	DPF1	5	12	2.6863e-2	4.8123e-2 −	2.0711e-2 ≈	4.5097e-2 ≈	2.5934e-2 ≈	1.9522e-2 +	7.5581e+0 −	1.6413e-1 −	4.3628e-2 ≈
		10	14	4.3796e-2	1.7213e-1 −	9.9788e-2 −	8.9416e-2 −	9.0616e-2 −	5.1037e-2 −	3.8249e+0 −	1.4723e-1 −	5.4345e-2 −
		15	17	7.9489e-2	2.0899e-1 −	2.9100e+1 −	1.4729e-1 −	1.2972e-1 −	9.8974e-2 −	3.8024e+0 −	2.0912e-1 −	1.2226e-1 −
	DPF2	5	22	4.5667e-2	6.0193e-2 −	5.4132e-2 −	4.1478e-2 +	1.6726e-1 −	3.7026e-2 +	8.6942e-2 −	1.3052e+0 −	3.0513e-2 +
		10	24	2.2141e-1	1.6505e-1 +	5.0288e-1 −	2.8906e-1 −	5.3678e-1 −	2.0005e-1 +	5.2271e-1 −	8.2411e+0 −	1.2118e-1 ≈
		15	27	3.8754e-1	3.9458e-1 −	5.4038e+0 −	5.4799e-1 −	1.3864e+0 −	4.0990e-1 −	1.5570e+0 −	5.8077e+1 −	3.1550e-1 +
	DPF3	5	9	2.4585e-2	2.6253e-2 −	3.1119e-2 −	2.8793e-2 −	4.4765e-2 −	2.5271e-2 −	5.6442e-2 −	1.6040e-1 −	2.6063e-2 −
		10	14	8.4801e-2	8.7481e-2 −	1.2199e-1 −	1.1846e-1 −	9.1265e-2 −	1.3525e-1 −	2.2531e-1 −	2.4081e-1 −	9.4307e-2 −
		15	19	1.6456e-1	1.8826e-1 −	2.8172e-1 −	2.4408e-1 −	2.2301e-1 −	2.5517e-1 −	3.8346e-1 −	4.6172e-1 −	1.9077e-1 −
	DPF4	5	12	9.1992e-2	2.8401e-2 ≈	6.6655e-2 ≈	4.8006e-2 ≈	2.1213e-2 +	5.8740e-2 ≈	7.3763e+1 −	1.9863e+4 −	4.0785e-2 +
		10	14	9.4762e-2	1.2771e-1 −	1.0951e-1 −	1.0424e-1 −	7.2398e-2 +	1.0072e-1 ≈	6.5817e+0 −	4.9710e+4 −	5.2717e+0 −
		15	17	1.9585e-1	9.2144e-1 −	3.7816e-1 −	2.5531e-1 −	1.7070e-1 +	2.3127e-1 −	2.4868e+0 −	1.4720e+5 −	2.3217e+1 −
	DPF5	5	14	7.3827e-2	1.2499e-1 −	1.1629e-1 −	9.2459e-2 −	7.6878e-2 −	7.7846e-2 −	1.0754e-1 −	2.5928e-1 −	5.0894e-2 +
		10	19	1.7287e-1	2.8527e-1 −	1.4093e+0 −	2.2101e-1 −	2.1497e-1 −	1.9949e-1 −	2.3085e-1 −	2.4730e-1 −	1.2809e-1 +
		15	24	2.3279e-1	3.6736e-1 −	1.2436e+0 −	3.0172e-1 −	2.5663e-1 −	3.4208e-1 −	3.5844e-1 −	3.9504e-1 −	1.6721e-1 +
other	MaF2	5	14	5.7329e-2	5.4894e-2 +	7.4300e-2 −	6.4094e-2 −	5.5877e-2 +	5.6337e-2 +	8.1271e-2 −	1.0899e-1 −	6.1057e-2 −
		10	19	1.1915e-1	1.4187e-1 −	1.1182e-1 +	1.2220e-1 −	1.4569e-1 −	1.1076e-1 +	1.3882e-1 −	2.0424e-1 −	1.2676e-1 −
		15	24	1.2712e-1	1.6039e-1 −	1.2624e-1 ≈	1.2457e-1 +	1.4665e-1 −	1.1977e-1 +	1.5354e-1 −	2.0216e-1 −	1.4288e-1 −
$+/-/\approx$					4/35/6	5/36/4	6/35/4	10/32/3	14/27/4	6/39/0	2/41/2	15/27/3

[1] Here $+/-/\approx$ indicates whether the compared algorithms are statistically significantly better, worse, or equivalent to CARV-MOEA.

intervals for generating the uniform reference vectors. The reference point for inverted IDTLZ2 and MaF4 with nonlinear PFs is set to (2,2,...,2) according to [54].

C. Comparison on Irregular Problems

Table I presents the IGD$^+$ values of the solutions obtained by CARV-MOEA and eight compared algorithms on solving irregular test instances, consisting of inverted problems (MaF1, IDTLZ2 and MaF4), discontinuous problems (MaF7 and MaF11), Pareto box problems (MaF8 and MaF9), and degenerate problems (MaF6, MaF13, and DPF1 to DPF5).The best mean value in terms of IGD$^+$ indicator are highlighted in dark grey in Table I, respectively. From these results, we can see that CARV-MOEA achieves the best performance on 15 out of 45 test instances. Overall, CARV-MOEA, MaOEA/SRV, AdaW show competitive results in dealing with irregular problems. On all irregular test instances, CARV-MOEA outperforms MaOEA/SRV on 27 test instances and performs worse than MaOEA/SRV on 15 test instances. Among eight compared algorithms, AdaW adjusts the distribution of reference vectors using the solutions in an archive after a number of predefined generations. As shown in Table I, AdaW performs the best on degenerate MaF6 and MaF13, but is outperformed by CARV-MOEA on most test instances. Recall that DEA-GNG, and RVEA$*$ use a set of fixed reference vectors along with a set of adaptive reference vectors, which aims to solve both regular problems as well as the irregular problems. It can be found that RVEA$*$ performs the best on inverted MaF4 and degenerate DPF4. DEA-GNG is outperformed by CARV-MOEA on most test instances. CA-MOEA, MOEA/D-SOM and MOEA/D-AM2M show competitive performance on degenerate problems such as MaF6 and MaF13.

We plot the solution sets obtained by CARV-MOEA and MaOEA/SRV on 10-objective inverted MaF4, discontinuous MaF7, degenerate MaF6, DPF1 and DPF4, respectively, in Fig. 5. The PF of DPF4 is both degenerate and discontinuous. From these results, we can observe that the approximated PFs obtained by CARV-MOEA are very close to the true PFs. The true PFs of adopted test problems are also given

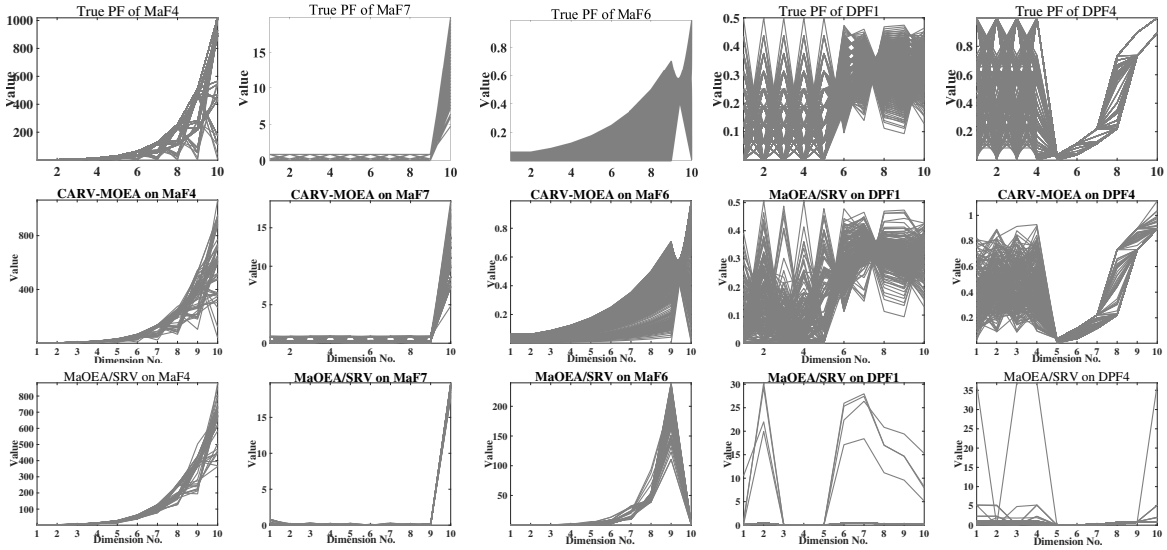


Fig. 5: The non-dominated solution set obtained by CARV-MOEA and MaOEA/SRV on solving MaF4 (inverted), MaF7 (discontinuous), MaF6, DPF1, DPF4 (degenerate) with 10 objectives. The set with the median IGD^+ value out of 30 independent runs is plotted.

Fig. 5. By contrast, MaOEA/SRV fails to converge to the true PFs on MaF6 and DPF4. Although both CARV-MOEA and MaOEA/SRV converge to the true PFs on 10-objective MaF4 and MaF7, CARV-MOEA seems to cover larger regions of true PFs than MaOEA/SRV. The Parto front of MaF7 contains a large number of discontinuous segments and the reason for CARV-MOEA's ability to handle MaF7 well could be attributed to the fact that the solution vectors of promising solutions with better diversity will be used to generate new reference vectors based on the angle threshold in CARV-MOEA, ensuring that promising regions can be preserved when they emerge.

D. Comparison on Regular Problems

Table SI in the Supplementary material presents the IGD^+ values of the solution sets obtained by CARV-MOEA and eight compared MOEAs on the regular test instances. As shown in Table SI, CARV-MOEA wins on 14 out of 30 test instances, while DEA-GNG, CA-MOEA, VaEA, RVEA$*$, Adaw, MOEA/D-AM2M, MOEA/D-SOM, and MaOEA/SRV wins on zero, zero, zero, seven, two, two, zero, and five test instances, respectively, out of 30 test instances in terms of the IGD^+ value. CARV-MOEA and MaOEA/SRV achieve the best performance on the regular problems with concave PFs. On the concave problems, MaOEA/SRV achieves the best in solving MaF12 and DTLZ2, while CARV-MOEA achieves the best on MaF5, DTLZ3 and DTLZ4. DTLZ3 contains a large number of local Pareto-optimal fronts and DTLZ4 is of biased density in the search space, which shows CARV-MOEA is competitive in convergence. However, the performance of CARV-MOEA on the problem with mixed PFs, MaF10 is worse than AdaW. The performance of CARV-MOEA on linear DTLZ1 and convex MaF3 is the best in terms of IGD^+.

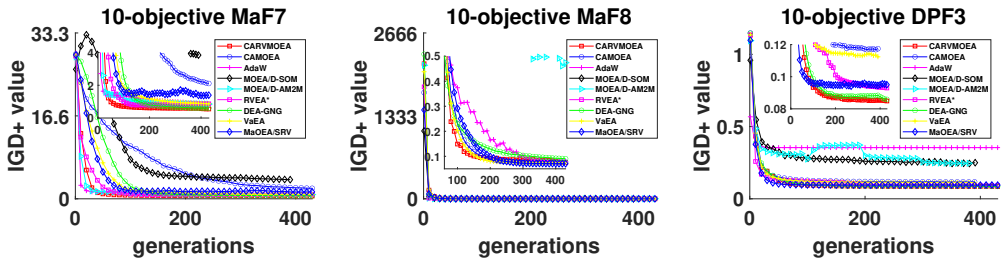
As mentioned above, RVEA$*$ has one set of fixed reference vectors, making it also capable of handling regular problems. Overall, CARV-MOEA achieves the best performance in terms of the IGD^+ value on all regular problems studied in this work, followed by MaOEA/SRV and RVEA$*$.

The Wilcoxon signed-rank test [55] is adopted to assess the performance between CARV-MOEA and the eight compared algorithms. Fig. S4 and Fig. S5 in the Supplementary material plot the performance score in terms of IGD^+ and HV indicator value averaged over all regular and irregular test instances. It can be seen that CARV-MOEA ranks the first, followed by MaOEA/SRV and AdaW, RVEA$*$ in terms of both IGD^+ and HV values. In summary, we can conclude that CARV-MOEA achieves competitive results on solving both irregular and regular problems and performs particularly well on discontinuous and degenerate problems.

E. Analysis of Convergence

To study the convergence speed of CARV-MOEA, we plot the IGD^+ values of solutions over the generations obtained by CARV-MOEA and eight compared algorithms on solving discontinuous MaF7, Pareto box MaF8 and degenerate DPF3 in Fig. 6. It can be seen that CARV-MOEA converges a little slower than RVEA$*$ and MOEA/D-AM2M on MaF7 over the first 50 generations, but then it quickly converges much faster than the compared algorithms after the first 50 generations. On solving MaF8, we can find that the performance of CARV-MOEA converges much better than the compared algorithms, then finally CARV-MOEA, MaOEA/SRV,VaEA and CA-MOEA converge to almost the same point. On solving degenerate DPF3, CARV-MOEA converges slightly more slowly than AdaW, but much faster than other seven compared algorithms over the first 100 generations, then CARV-MOEA surpasses eight compared algorithms after 100 generations.

In addition, we also give the IGD^+ indicator values obtained by CARV-MOEA and eight compared algorithms on solving regular problems, such as MaF5, DTLZ3 and DTLZ4. The plots are given in Fig. S3 in the Supplementary material.



(a) MaF7 (b) MaF8 (c) DPF3

Fig. 6: Mean IGD$^+$ indicator value of solutions obtained by CARV-MOEA and eight compared algorithms on handling discontinuous MaF7, Pareto box MaF8 and degenerate DPF3 over 30 independent runs.

F. The Effectiveness of A-APD and the Angle Threshold

In this subsection, we perform a set of experiments to show the importance of the coordinated adaptation of the reference vectors and scalarizing function. To this end, we compare CARV-MOEA with three of its variants. In the first variant, the proposed A-APD is replaced with the original APD, termed as CARV-MOEA-APD. The second variant uses a simplified version of A-APD, where γ_j is determined in the same way as in the original APD, which is denoted as CARV-MOEA-noC. In the third variant, we replace the proposed reference adaptation method with the one in MOEA/D-2ADV [35], which adds new reference vectors between sparsely distributed reference vectors, which is denoted as CARV-MOEA-IRV. The average performance scores in terms of IGD$^+$ over all test instances of CARV-MOEA, CARV-MOEA-APD, CARV-MOEA-IRV and CARV-MOEA-noC are 1.9333, 2.7067, 3.08 and 2.28, respectively. All comparative results obtained by CARV-MOEA and three of its variants in terms of IGD$^+$ and HV values are listed in Table SIV and Table SV in the Supplementary material, respectively.

As shown in the Table SIV, we can find CARV-MOEA-APD performs significantly worse than, comparable to, and better than CARV-MOEA on 52, 10, and 13, respectively, out of 75 test instances in terms of IGD$^+$. These results confirm that the proposed A-APD performs better than APD on MaOPs when the reference vectors are adapted during the search process.

In addition, we observe from Table SIV that CARV-MOEA-IRV performs significantly worse than, comparable to and better than CARV-MOEA on 50, 12, and 13 among 75 test instances, respectively. Note that, the interpolated reference vectors in CARV-MOEA-IRV will be deleted only after a number of fixed generations, which is the same as in CARV-MOEA. As a result, the algorithm will have a better chance to exploit the search space along the interpolated reference vectors during the search space, which, however, may limit the exploration capability of the algorithm.

Finally, we note that CARV-MOEA-noC performs significantly worse than CARV-MOEA on those problems with a huge number of local PFs, such as DTLZ1, DTLZ3 and MaF3. The results demonstrate that the adaptation of γ in A-APD is helpful to ensure that convergence is not significantly slowed down. In Fig. S6, we plot the non-dominated solution set obtained by CARV-MOEA and CARV-MOEA-noC on solving 10-objective DTLZ3 and MaF1. It can be seen that CARV-MOEA-noC performs much worse than CARV-MOEA on DTLZ3. Both CARV-MOEA and CARV-MOEA-noC achieve competitive results on MaF1. Thus, we can conclude that the adaptation of γ in A-APD can help avoid selecting the poorly converged solutions. At the same time, it can prevent the well converged solutions associated with new reference vectors from getting lost.

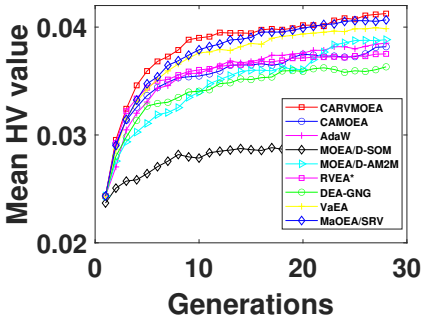


Fig. 7: Mean HV value of the solutions obtained by CARV-MOEA and eight compared algorithms on the HEV controller design problem over the generations averaged over 24 independent runs.

G. Hybrid Electric Vehicle Controller Design

The performance of the proposed CARV-MOEA is further tested on a real-world application, i.e., the design of a controller of hybrid electric vehicles (HEVs). The design of HEV controller contains seven objectives and 11 decision variables. A description of objectives and the design variables are given in Section V in the Supplementary material and a more detailed description of the HEV controller optimization problem can be found in [14].

We compare CARV-MOEA on the HEV controller design problem with eight compared algorithms. In Fig. 7, we give the HV indicator values of the solutions obtained by CARV-MOEA and the compared algorithms over the generations averaged over 24 independent runs. The maximum objective values of all non-dominated solutions obtained by these nine algorithms are used as the nadir point. From Fig. 7, it can be seen that CARV-MOEA performs the best in terms of the HV indicator values over all generations, which further demonstrates the

competitiveness of CARV-MOEA in solving many-objective problems.

V. Conclusion

In this study, we propose a new criterion for adapting the reference vectors based on the angle threshold specific to each reference vector. Additionally, we design an adaptive scalarizing function, termed A-APD in coordination with the adaptation of the reference vectors. This way, the algorithm is able to preserve promising solutions and explore new regions in the search space without severely slowing down the convergence. Experimental results show that the coordinated adaptation of the reference vectors and scalarizing functions is effective in enhancing the search performance on MaOPs with both regular and irregular PFs, compared with eight state-of-the-art algorithms.

However, the proposed CARV-MOEA is ineffective in preserving the extreme solutions when the PF is convex but has a very large curvature. Thus, our future work will focus on detecting the shape of the PFs during the search process and then adapt the scalarizing function to the shape of Pareto fronts. In addition, diversity measures will be introduced to enhance the evenness of the distribution of the reference vectors in reference vector adaption to further enhance the diversity of the obtained solutions. Finally, efforts will be made to develop more dedicated solution generation methods [56] by taking the distribution of the Pareto optimal solutions into account to enhance the search efficiency.

VI. Acknowledgements

This work was supported in part by the Honda Research Institute Europe GmbH, Offenbach am Main, Germany. Yaochu Jin is funded by an Alexander von Humboldt Professorship for Artificial Intelligence endowed by the Federal Ministry of Education and Research, Germany.

References

[1] C. A. C. Coello, S. G. Brambila, J. F. Gamboa, M. G. C. Tapia, and R. H. Gómez, "Evolutionary multiobjective optimization: open research areas and some challenges lying ahead," *Complex & Intelligent Systems*, vol. 6, no. 2, pp. 221–236, 2020.

[2] C. A. C. Coello, "A comprehensive survey of evolutionary-based multiobjective optimization techniques," *Knowledge and Information Systems*, vol. 1, no. 3, pp. 269–308, 1999.

[3] B. Li, J. Li, K. Tang, and X. Yao, "Many-objective evolutionary algorithms: A survey," *ACM Computing Surveys (CSUR)*, vol. 48, no. 1, pp. 1–35, 2015.

[4] 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.

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

[6] 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.

[7] H. Chen, R. Cheng, W. Pedrycz, and Y. Jin, "Solving many-objective optimization problems via multistage evolutionary search," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 51, no. 6, pp. 3552–3564, 2021.

[8] Y. Sun, G. G. Yen, and Z. Yi, "IGD indicator-based evolutionary algorithm for many-objective optimization problems," *IEEE Transactions on Evolutionary Computation*, vol. 23, no. 2, pp. 173–187, 2018.

[9] J. Bader and E. Zitzler, "HypE: An algorithm for fast hypervolume-based many-objective optimization," *Evolutionary Computation*, vol. 19, no. 1, pp. 45–76, 2011.

[10] Y. Zhou, M. Zhu, J. Wang, Z. Zhang, Y. Xiang, and J. Zhang, "Tri-goal evolution framework for constrained many-objective optimization," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 50, no. 8, pp. 3086–3099, 2020.

[11] Y. Yuan, H. Xu, B. Wang, and X. Yao, "A new dominance relation-based evolutionary algorithm for many-objective optimization," *IEEE Transactions on Evolutionary Computation*, vol. 20, no. 1, pp. 16–37, 2016.

[12] 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 Transactions on Evolutionary Computation*, vol. 23, no. 2, pp. 331–345, 2018.

[13] G. Yu, Y. Jin, and M. Olhofer, "A multi-objective evolutionary algorithm for finding knee regions using two localized dominance relationships," *IEEE Transactions on Evolutionary Computation*, vol. 25, no. 1, pp. 145–158, 2020.

[14] 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.

[15] X. Ma, Y. Yu, X. Li, Y. Qi, and Z. Zhu, "A survey of weight vector adjustment methods for decomposition based multi-objective evolutionary algorithms," *IEEE Transactions on Evolutionary Computation*, vol. 24, no. 4, pp. 634–649, 2020.

[16] Y. Hua, Q. Liu, K. Hao, and Y. Jin, "A survey of evolutionary algorithms for multi-objective optimization problems with irregular Pareto fronts," *IEEE/CAA Journal of Automatica Sinica*, vol. 8, no. 2, pp. 303–318, 2021.

[17] H.-L. Liu, L. Chen, Q. Zhang, and K. Deb, "Adaptively allocating search effort in challenging many-objective optimization problems," *IEEE Transactions on Evolutionary Computation*, vol. 22, no. 3, pp. 433–448, 2017.

[18] Y. Qi, X. Ma, F. Liu, L. Jiao, J. Sun, and J. Wu, "MOEA/D with adaptive weight adjustment," *Evolutionary Computation*, vol. 22, no. 2, pp. 231–264, 2014.

[19] M. Li and X. Yao, "What weights work for you? Adapting weights for any Pareto front shape in decomposition-based evolutionary multi-objective optimisation," *Evolutionary Computation*, vol. 28, no. 2, pp. 227–253, 2020.

[20] Y. Tian, R. Cheng, X. Zhang, F. Cheng, and Y. Jin, "An indicator based multi-objective evolutionary algorithm with reference point adaptation for better versatility," *IEEE Transactions on Evolutionary Computation*, vol. 22, no. 4, pp. 609–622, 2017.

[21] Q. Liu, Y. Jin, M. Heiderich, and T. Rodemann, "Adaptation of reference vectors for evolutionary many-objective optimization of problems with irregular Pareto fronts," in *2019 IEEE Congress on Evolutionary Computation (CEC)*. IEEE, 2019, pp. 1726–1733.

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

[23] F. Gu and Y.-M. Cheung, "Self-organizing map-based weight design for decomposition-based many-objective evolutionary algorithm," *IEEE Transactions on Evolutionary Computation*, vol. 22, no. 2, pp. 211–225, 2017.

[24] Y. Liu, H. Ishibuchi, N. Masuyama, and Y. Nojima, "Adapting reference vectors and scalarizing functions by growing neural gas to handle irregular Pareto fronts," *IEEE Transactions on Evolutionary Computation*, vol. 24, no. 3, pp. 439–453, 2019.

[25] Q. Liu, Y. Jin, M. Heiderich, T. Rodemann, and G. Yu, "An adaptive reference vector guided evolutionary algorithm using growing neural gas for many-objective optimization of irregular problems," *IEEE Transactions on Cybernetics*, 2020, doi:10.1109/TCYB.2020.3020630 .

[26] T. Kohonen, "The self-organizing map," *Proceedings of the IEEE*, vol. 78, no. 9, pp. 1464–1480, 1990.

[27] B. Fritzke, "A growing neural gas network learns topologies," in *Advances in Neural Information Processing Systems*, 1995, pp. 625–632.

[28] R. Denysiuk, L. Costa, and I. E. Santo, "Clustering-based selection for evolutionary many-objective optimization," in *International Conference on Parallel Problem Solving from Nature*. Springer, 2014, pp. 538–547.

[29] Y. Hua, Y. Jin, and K. Hao, "A clustering-based adaptive evolutionary algorithm for multiobjective optimization with irregular Pareto fronts," *IEEE Transactions on Cybernetics*, vol. 49, no. 7, pp. 2758–2770, 2018.

[30] Q. Lin, S. Liu, K.-C. Wong, M. Gong, C. A. C. Coello, J. Chen, and J. Zhang, "A clustering-based evolutionary algorithm for many-objective optimization problems," *IEEE Transactions on Evolutionary Computation*, vol. 23, no. 3, pp. 391–405, 2018.

[31] H. Ge, M. Zhao, L. Sun, Z. Wang, G. Tan, Q. Zhang, and C. P. Chen, "A many-objective evolutionary algorithm with two interacting processes: Cascade clustering and reference point incremental learning," *IEEE Transactions on Evolutionary Computation*, vol. 23, no. 4, pp. 572–586, 2018.

[32] R. Wang, J. Xiong, H. Ishibuchi, G. Wu, and T. Zhang, "On the effect of reference point in MOEA/D for multi-objective optimization," *Applied Soft Computing*, vol. 58, pp. 25–34, 2017.

[33] H. Li and D. Landa-Silva, "An adaptive evolutionary multi-objective approach based on simulated annealing," *Evolutionary Computation*, vol. 19, no. 4, pp. 561–595, 2011.

[34] H. Zhao, C. Zhang, B. Zhang, P. Duan, and Y. Yang, "Decomposition-based sub-problem optimal solution updating direction-guided evolutionary many-objective algorithm," *Information Sciences*, vol. 448, pp. 91–111, 2018.

[35] X. Cai, Z. Mei, and Z. Fan, "A decomposition-based many-objective evolutionary algorithm with two types of adjustments for direction vectors," *IEEE Transactions on Cybernetics*, vol. 48, no. 8, pp. 2335–2348, 2017.

[36] H. Jain and K. Deb, "An improved adaptive approach for elitist nondominated sorting genetic algorithm for many-objective optimization," in *International Conference on Evolutionary Multi-Criterion Optimization*. Springer, 2013, pp. 307–321.

[37] M.-R. Bouguelia, Y. Belaïd, and A. Belaïd, "An adaptive incremental clustering method based on the growing neural gas algorithm," in *2nd International Conference on Pattern Recognition Applications and Methods-ICPRAM 2013*. SciTePress, 2013, pp. 42–49.

[38] J. A. Cornell, *Experiments with mixtures: designs, models, and the analysis of mixture data*. John Wiley & Sons, 2011, vol. 403.

[39] R. B. Agrawal, K. Deb, and R. Agrawal, "Simulated binary crossover for continuous search space," *Complex Systems*, vol. 9, no. 2, pp. 115–148, 1995.

[40] K. Deb and M. Goyal, "A combined genetic adaptive search (GeneAS) for engineering design," *Computer Science and Informatics*, vol. 26, pp. 30–45, 1996.

[41] Y.-H. Zhang, Y.-J. Gong, T.-L. Gu, H.-Q. Yuan, W. Zhang, S. Kwong, and J. Zhang, "DECAL: Decomposition-based coevolutionary algorithm for many-objective optimization," *IEEE Transactions on Cybernetics*, vol. 49, no. 1, pp. 27–41, 2017.

[42] S. Jiang, S. Yang, Y. Wang, and X. Liu, "Scalarizing functions in decomposition-based multiobjective evolutionary algorithms," *IEEE Transactions on Evolutionary Computation*, vol. 22, no. 2, pp. 296–313, 2018.

[43] H. Ishibuchi, Y. Sakane, N. Tsukamoto, and Y. Nojima, "Adaptation of scalarizing functions in MOEA/D: An adaptive scalarizing function-based multiobjective evolutionary algorithm," in *International Conference on Evolutionary Multi-Criterion Optimization*. Springer, 2009, pp. 438–452.

[44] R. Wang, Q. Zhang, and T. Zhang, "Decomposition-based algorithms using Pareto adaptive scalarizing methods," *IEEE Transactions on Evolutionary Computation*, vol. 20, no. 6, pp. 821–837, 2016.

[45] M. Pescador-Rojas and C. A. C. Coello, "Collaborative and adaptive strategies of different scalarizing functions in MOEA/D," in *2018 IEEE Congress on Evolutionary Computation (CEC)*. IEEE, 2018, pp. 1–8.

[46] Y. Xiang, Y. Zhou, M. Li, and Z. Chen, "A vector angle-based evolutionary algorithm for unconstrained many-objective optimization," *IEEE Transactions on Evolutionary Computation*, vol. 21, no. 1, pp. 131–152, 2017.

[47] S. Liu, Q. Lin, K.-C. Wong, C. A. C. Coello, J. Li, Z. Ming, and J. Zhang, "A self-guided reference vector strategy for many-objective optimization," *IEEE Transactions on Cybernetics*, vol. 52, no. 2, pp. 1164–1178, 2022.

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

[49] K. Deb, "Multi-objective genetic algorithms: Problem difficulties and construction of test problems," in *Evolutionary Computation*. MIT Press, 1999, vol. 7, no. 3, pp. 205–230.

[50] L. Zhen, M. Li, R. Cheng, D. Peng, and X. Yao, "Multiobjective test problems with degenerate Pareto fronts," *arXiv preprint arXiv:1806.02706*, 2018.

[51] 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.

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

[53] H. Ishibuchi, H. Masuda, Y. Tanigaki, and Y. Nojima, "Modified distance calculation in generational distance and inverted generational distance," in *International Conference on Evolutionary Multi-criterion Optimization*. Springer, 2015, pp. 110–125.

[54] H. Ishibuchi, R. Imada, Y. Setoguchi, and Y. Nojima, "How to specify a reference point in hypervolume calculation for fair performance comparison," *Evolutionary Computation*, vol. 26, no. 3, pp. 411–440, 2018.

[55] F. Wilcoxon, "Individual comparisons by ranking methods," in *Biometrics Bulletin*, 1945, vol. 1, no. 6, pp. 80–83.

[56] C. He, R. Cheng, and D. Yazdani, "Adaptive offspring generation for evolutionary large-scale multiobjective optimization," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 52, no. 2, pp. 786–798, 2022.

Qiqi Liu received the B.Eng. degree in industrial engineering from Xi'an University of Science and Technology and the M.E. degree in information and communication engineering from Shenzhen University in 2013 and 2016, respectively. She is pursuing the Ph.D. degree in computer science at University of Surrey. Her current research interests include evolutionary many-objective optimization, surrogate assisted evolutionary optimization and preference learning. She is a regular reviewer of Complex & Intelligent Systems, Soft Computing, and Swarm and Evolutionary Computation.

Yaochu Jin (Fellow, IEEE) 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 an Alexander von Humboldt Professor for Artificial Intelligence endowed by the German Federal Ministry of Education and Research, Chair of Nature Inspired Computing and Engineering, Faculty of Technology, Bielefeld University, Germany. He is also a Distinguished Chair, Professor in Computational Intelligence, Department of Computer Science, University of Surrey, Guildford, U.K. He was a "Finland Distinguished Professor" of University of Jyväskylä, Finland, "Changjiang Distinguished Visiting Professor", Northeastern University, China, and "Distinguished Visiting Scholar", University of Technology Sydney, Australia. His main research interests include evolutionary optimization, evolutionary learning, trustworthy machine learning, and evolutionary developmental systems.

Prof Jin is currently the Editor-in-Chief of Complex & Intelligent Systems. He is the recipient of the 2018 and 2021 IEEE Transactions on Evolutionary Computation Outstanding Paper Award, and the 2015, 2017, and 2020 IEEE Computational Intelligence Magazine Outstanding Paper Award. He is a Member of the Academia Europaea.

Martin Heiderich Biography not available.

Tobias Rodemann studied physics and neuroinformatics at the Ruhr Universität Bochum, Germany, and received his Masters (Diploma) in Neuroinformatics from the Universität Bochum in 1998 and a Ph.D. degree in Computational Neuroscience from the Technische Universität Bielefeld, Germany in 2003. In 1998 he joined the Future Technology Research Division of Honda R&D Europe in Offenbach, Germany, where he worked on various research topics like biologically-inspired vision systems, computational neuroscience, and auditory processing. Since 2011 he is working at the Honda Research Institute Europe as a principal scientist on system modeling and the application of many-objective optimization methods for energy management.