

A dual-population paradigm for evolutionary multiobjective optimization



Ke Li ^{a,b}, Sam Kwong ^{a,*}, Kalyanmoy Deb ^b

^a Department of Computer Science, City University of Hong Kong, Tat Chee Avenue, Kowloon, Hong Kong

^b Department of Electrical and Computer Engineering, Michigan State University, East Lansing, MI 48824, USA

ARTICLE INFO

Article history:

Received 19 June 2014

Received in revised form 29 December 2014

Accepted 1 March 2015

Available online 7 March 2015

Keywords:

Dual-population paradigm

Pareto dominance

Decomposition

Evolutionary computation

Multiobjective optimization

ABSTRACT

Convergence and diversity are two basic issues in evolutionary multiobjective optimization (EMO). However, it is far from trivial to address them simultaneously, especially when tackling problems with complicated Pareto-optimal sets. This paper presents a dual-population paradigm (DPP) that uses two separate and co-evolving populations to deal with convergence and diversity simultaneously. These two populations are respectively maintained by Pareto- and decomposition-based techniques, which arguably have complementary effects in selection. In particular, the so called Pareto-based archive is assumed to maintain a population with competitive selection pressure towards the Pareto-optimal front, while the so called decomposition-based archive is assumed to preserve a population with satisfied diversity in the objective space. In addition, we develop a restricted mating selection mechanism to coordinate the interaction between these two populations. DPP paves an avenue to integrate Pareto- and decomposition-based techniques in a single paradigm. A series of comprehensive experiments is conducted on seventeen benchmark problems with distinct characteristics and complicated Pareto-optimal sets. Empirical results fully demonstrate the effectiveness and competitiveness of the proposed algorithm.

© 2015 Elsevier Inc. All rights reserved.

1. Introduction

Multiobjective optimization problems (MOPs) involve more than one objective function to be optimized simultaneously. They typically arise in various fields of science (e.g., [29,1,31]) and engineering (e.g., [33,27,32]) where optimal decisions need to be taken in the presence of trade-offs between two or more conflicting objectives. Maximizing the expected value of portfolio returns and minimizing the potential risk in portfolio management is an example of MOP involving two objectives. Due to the population-based property, evolutionary algorithms (EAs) have been widely recognized as a major approach for multiobjective optimization. Over the last two decades, much effort has been dedicated to developing evolutionary multiobjective optimization (EMO) algorithms, such as non-dominated sorting genetic algorithm II (NSGA-II) [12], improved strength Pareto EA (SPEA2) [42] and multiobjective EA based on decomposition (MOEA/D) [36]. Interested readers can refer to [40] for a more comprehensive survey on recent developments of EMO algorithms.

Generally speaking, there are two distinct goals, common but often conflicting, in multiobjective optimization [30]:

* Corresponding author.

E-mail address: cssamk@cityu.edu.hk (S. Kwong).

1. *Convergence*: find a set of solutions that approximates the Pareto-optimal set (PS) as close as possible.
2. *Diversity*: find a set of solutions that represents the entire range of the Pareto-optimal front (PF) as diverse as possible.

Achieving a balance between convergence and diversity is important, but usually far from trivial, in multiobjective optimization.

In traditional Pareto-based EMO algorithms (e.g., NSGA-II and SPEA2), they adopt a convergence first and diversity second selection strategy [23] to fulfill the above two goals. In particular, the convergence issue can be easily achieved by various dominance-preserving techniques, such as the non-dominated sorting in NSGA-II and the Pareto strength value in SPEA2. As for maintaining the population diversity, many diversity-preserving strategies have been proposed to emphasize the solutions in a less crowded niche. However, these diversity-preserving strategies lead to another trade-off between the achievable diversity and the computational complexity. Some algorithms use a quick-but-dirty method (e.g., the crowding distance estimation in NSGA-II) to find a reasonably good distribution quickly; whereas some other ones use a more time consuming strategy (e.g., the clustering analysis in SPEA2) to obtain a better distribution. To achieve both progress towards the PS and coverage over the entire range of PF, some modifications have been introduced to the Pareto dominance relation, such as the ε -dominance [18], ε -MOEA [11], argued by their authors, achieves a good compromise in terms of convergence towards the PS, well distribution along the PF and small computational complexity. However, its capabilities for solving difficult problems with complicated PSs are still questionable.

Another way to simultaneously consider the convergence and diversity is to apply a set-based indicator, which can evaluate the quality of a PF approximation, as the selection criterion of EMO algorithms. Such indicator can be regarded as a function that assigns a scalar value to each possible PF approximation reflecting its quality and fulfilling certain monotonicity conditions. Hypervolume indicator [43], which provides a combined information about convergence and diversity of the obtained approximation set, is widely used in developing EMO algorithms (e.g., indicator-based EA (IBEA) [41] and S -metric selection EMO algorithm (SMS-EMOA) [4]). However, one of its major criticisms is the large computational complexity, which increases exponentially with the number of objectives. Furthermore, the choice of reference points has a large influence on its calculation result [2]. Unfortunately, in practice, this is far from trivial and is problem dependent.

One other avenue to integrate convergence and diversity into a single criterion is the aggregation-based method (e.g., multiple single objective Pareto sampling (MSOPS) [16] and MOEA/D [36]). As the state-of-the-art, MOEA/D [36] is a representative of this sort. It decomposes a MOP into a number of single objective optimization subproblems through aggregation functions and optimizes them in a collaborative manner. MOEA/D paves a new avenue to balance the convergence and diversity in an efficient manner. Particularly, the convergence of a solution is ensured by the optimum of a subproblem, while the population diversity is guaranteed by the wide distribution of subproblems. During the past few years, MOEA/D, as a major framework to design EMO algorithms, has spawned a large amount of research works, e.g., introducing adaptive mechanism in reproduction [20], hybridizing with local search [34] and incorporating stable matching in selection [22]. Nevertheless, as discussed in [23], MOEA/D struggles to maintain the population diversity when tackling problems with complicated search landscapes. As a consequence, the population might be trapped in some limited regions and fail to achieve a good coverage of the PF.

Based on the above discussions, developing an EMO algorithm that finds a set of well-converged and well-distributed solutions in an efficient manner, especially when tackling problems with complicated PSs [19], is challenging and important. To address these considerations, this paper presents a novel technique, termed dual-population paradigm (DPP), to handle the convergence and diversity in two separate and co-evolving populations. In particular, DPP maintains two populations: one population, termed Pareto-based archive, maintains a repository of solutions with a competitive selection pressure towards the PF; the other one, called decomposition-based archive, preserves an archive of solutions with a satisfied distribution in the objective space. As the name suggests, the Pareto- and decomposition-based archives use the Pareto- and decomposition-based techniques, respectively, to update their populations. In addition, the objective space is divided into several subregions by a set of evenly distributed unit vectors. Each solution in a population is associated with a subregion. Then, a restricted mating selection mechanism is developed to coordinate the interaction between these two co-evolving populations. To validate the effectiveness of DPP, we present four DPP instantiations, which apply the existing Pareto- and decomposition-based techniques in a plug-in manner, for empirical studies. Furthermore, we not only compare each DPP instantiation with its baseline algorithms, we also compare a representative DPP instantiation with four state-of-the-art EMO algorithms. Comprehensive experiments on seventeen benchmark problems with distinct characteristics and complicated PSs fully demonstrate the superiority of the proposed algorithm.

It is worth noting that using two or multiple populations is not a brand new technique in the EMO community. For example, ε -MOEA [11] uses two populations and different archiving mechanisms in evolution. Differently, the idea of DPP is to deal with convergence and diversity by two separate and co-evolving populations. Moreover, by respectively using Pareto- and decomposition-based techniques to update these two populations, DPP paves an avenue to integrate the Pareto- and decomposition-based techniques in EMO. To the best of the authors' knowledge, very few reports [25,28] have addressed this topic in the literature. In [25], a hybrid algorithm is proposed to combine MOEA/D and NSGA-II for solving multiobjective capacitated arc routing problems. In [28], the Pareto-based method is used to build the leaders' archive while a decomposition-based method is used for fitness assignment.

The rest of this paper is organized as follows. Some preliminary knowledge of multiobjective optimization is provided in Section 2. Afterwards, technical details of DPP are illustrated step by step in Section 3, and instantiations of DPP are

presented in Section 4. Then, empirical results, performance comparisons and discussions are given in Section 5. Finally, conclusions and future works are given in Section 6.

2. Preliminary concepts

This section first provides some basic definitions of multiobjective optimization. Afterwards, we briefly introduce the concept of decomposing a MOP.

2.1. Multiobjective optimization

A continuous MOP considered in this paper can be mathematically defined as follows:

$$\begin{aligned} & \text{minimize } \mathbf{F}(\mathbf{x}) = (f_1(\mathbf{x}), \dots, f_m(\mathbf{x}))^T \\ & \text{subject to } \mathbf{x} \in \Omega \end{aligned} \quad (1)$$

where $\Omega = \prod_{i=1}^n [a_i, b_i] \subset \mathbb{R}^n$ is the *decision (variable) space*, $\mathbf{x} = (x_1, \dots, x_n)^T \in \Omega$ is a candidate solution, $\mathbf{F}: \Omega \rightarrow \mathbb{R}_+^m$ ¹ constitutes of m real-valued objective functions, and \mathbb{R}_+^m is called the *objective space*. The *attainable objective set* is defined as $\Theta = \{\mathbf{F}(\mathbf{x}) | \mathbf{x} \in \Omega\}$.

Definition 1. \mathbf{x}^1 is said to *Pareto dominate* \mathbf{x}^2 , denoted as $\mathbf{x}^1 \preceq \mathbf{x}^2$, if and only if

1. $\forall i \in \{1, \dots, m\} : f_i(\mathbf{x}^1) \leq f_i(\mathbf{x}^2)$.
2. $\exists j \in \{1, \dots, m\} : f_j(\mathbf{x}^1) < f_j(\mathbf{x}^2)$.

Definition 2. A decision vector $\mathbf{x}^* \in \Omega$ is said to be *Pareto-optimal* if there is no other decision vector $\mathbf{x} \in \Omega$ such that $\mathbf{x} \preceq \mathbf{x}^*$.

Definition 3. The set of all Pareto-optimal solutions is called the Pareto-optimal set (PS). The set of all the Pareto-optimal objective vectors, $PF = \{\mathbf{F}(\mathbf{x}) \in \mathbb{R}_+^m | \mathbf{x} \in PS\}$, is called the Pareto-optimal front (PF).

Definition 4. The ideal objective vector is $\mathbf{z}^* = (z_1^*, \dots, z_m^*)^T$, where $z_i^* = \min_{\mathbf{x} \in \Omega} f_i(\mathbf{x}), i \in \{1, \dots, m\}$.

Definition 5. The nadir objective vector is $\mathbf{z}^{nad} = (z_1^{nad}, \dots, z_m^{nad})^T$, where $z_i^{nad} = \max_{\mathbf{x} \in PS} f_i(\mathbf{x}), i \in \{1, \dots, m\}$.

2.2. Decomposition of a MOP

It is well-known that a Pareto-optimal solution of a MOP, under some mild conditions, is an optimal solution of a scalar optimization problem whose objective function is an aggregation of all individual objectives. Therefore, the task of approximating the PF can be decomposed into a number of scalar optimization subproblems. In the literature of mathematical programming [26], there are several approaches for constructing aggregation functions. Among them, weighted Tchebycheff (TCH) approach, applicable to both convex and non-convex PF shapes, is widely used in many decomposition-based EMO algorithms (e.g., [36,34]). In particular, the TCH approach is mathematically defined as the following single objective optimization problem:

$$\begin{aligned} & \text{minimize } g^{tch1}(\mathbf{x} | \mathbf{w}, \mathbf{z}^*) = \max_{1 \leq i \leq m} \{w_i |f_i(\mathbf{x}) - z_i^*|\} \\ & \text{subject to } \mathbf{x} \in \Omega \end{aligned} \quad (2)$$

where $\mathbf{w} = (w^1, \dots, w^m)^T$ is a user specified weight vector, $w_i \geq 0$ for all $i \in \{1, \dots, m\}$ and $\sum_{i=1}^m w_i = 1$. However, as discussed in [9], the following modification on (2) can result in a more uniform distribution of solutions in the objective space.²

$$\begin{aligned} & \text{minimize } g^{tch2}(\mathbf{x} | \mathbf{w}, \mathbf{z}^*) = \max_{1 \leq i \leq m} \{|f_i(\mathbf{x}) - z_i^*| / w_i\} \\ & \text{subject to } \mathbf{x} \in \Omega \end{aligned} \quad (3)$$

¹ Here we only consider the non-negative case, i.e., $f_i \geq 0$ for all $i \in \{1, \dots, m\}$; otherwise, we replace the negative objective function f_i by $f_i + M$, where $M > 0$ is an appropriate constant.

² Fig. 1 presents a comparison of MOEA/D [19] with two different TCH approaches on DTLZ1 test instance [13]. It is clear that the TCH approach introduced in (3) leads to a better distribution. More detailed discussion on this issue can be found in [10,15].

where w_i is set to be a very small number, say 10^{-6} , in case $w_i = 0$. Therefore, without loss of generality, this paper only considers (3) as an example to serve the decomposition purpose.

2.3. General descriptions of two popular EMO algorithms

Since this work is developed upon the Pareto- and decomposition-based techniques, for better understanding, we give a general description on two popular Pareto- and decomposition-based EMO algorithms in this section.

- NSGA-II** [12]: This is one of the most popular EMO algorithms in the past decade. At each generation of NSGA-II, it at first uses non-dominated sorting to divide the hybrid population of parents and offspring into several non-domination levels (i.e., F_1, \dots, F_l). Starting from F_1 , one non-domination level is selected at a time to construct a new population, until its size equals or exceeds the limit. The exceeded solutions in the last acceptable non-domination level will be eliminated according to some density indicator (i.e., crowding distance). In this case, the selection strategy of NSGA-II implies its convergence first and diversity second behavior.
- MOEA/D** [36]: This is a seminal decomposition-based method, where a MOP is transformed into multiple single objective optimization subproblems and a population-based algorithm is employed to solve them simultaneously. To exploit information of neighboring subproblems, the neighborhood of each subproblem is defined in the initialization phase. At each generation, MOEA/D maintains a population of solutions, each of which corresponds to a subproblem. One major feature of MOEA/D is mating restriction, where the mating parents are selected from some neighboring subproblems. The other major feature is local replacement, where an offspring $\bar{\mathbf{x}}$ is compared with parent solutions from its neighboring subproblems. In particular, $\bar{\mathbf{x}}$ replaces a parent solution \mathbf{x} in case $g^{tch2}(\bar{\mathbf{x}}|\mathbf{w}, \mathbf{z}^*) < g^{tch2}(\mathbf{x}|\mathbf{w}, \mathbf{z}^*)$.

3. Dual-population paradigm

The general architecture of our proposed dual-population paradigm (DPP) is given in Fig. 2. It maintains two co-evolving populations: one population, called the Pareto-based archive $A^P = \{\mathbf{x}^1, \dots, \hat{\mathbf{x}}^M\}$, uses a Pareto-based technique to maintain the population; the other one, termed the decomposition-based archive $A^d = \{\hat{\mathbf{x}}^1, \dots, \hat{\mathbf{x}}^N\}$, employs a decomposition-based technique to maintain the population. For simplicity, A^P and A^d are set the same size, i.e., $M = N$, in this paper. Given these two populations, each time, they interact with each other via our proposed restricted mating selection (RMS) mechanism. Generally speaking, RMS chooses some solutions, from each of these two populations, respectively, as mating parents for offspring generation. Thereby, the offspring which inherits the information from both populations is used to update each of A^P and A^d , respectively, based on the corresponding archiving mechanism. DPP is a simple and general framework that integrates Pareto- and decomposition-based techniques in a single paradigm. Through the interaction between A^P and A^d , DPP takes advantages of both populations to facilitate the search process. We argue that any technique, used in the existing Pareto- and decomposition-based EMO algorithms, can be applied to A^P and A^d , respectively, in a plug-in manner. For common users, they need to consider two issues: one is the choice of Pareto- and decomposition-based techniques for A^P and A^d ; the other is the design of the interaction mechanism between these two populations. In the following paragraphs, we will elaborate them one by one.

3.1. Pareto and decomposition-based archives

Fig. 3 presents an intuitive comparison of the selection results of NSGA-II and MOEA/D. There are four solutions in a two-dimensional objective space, i.e., A (5, 8), B (5, 4), C (10, 3) and D (11, 1). Assume that the maximum archive size is three, thus one solution should be eliminated after the selection procedure. For NSGA-II, A will be eliminated from the population, as it is dominated by B. For MOEA/D, C will be eliminated from the population, as it cannot beat any of the other three solutions on subproblems $\mathbf{w}^1, \mathbf{w}^2$ and \mathbf{w}^3 , in terms of aggregation functions. From this comparison, we observe that NSGA-II and MOEA/D have some complementary effects in selection. The selection result of NSGA-II might preserve a better convergence property, while the diversity is not satisfied enough (as A has been eliminated and the selected solutions all crowd in the lower half of the objective space). By contrast, although the selection result of MOEA/D might have a better diversity property, its convergence is not satisfied enough (as the dominated solution A has been selected). In addition, a recent empirical study in [35] has also demonstrated that NSGA-II and MOEA/D are suitable for different problems. Based on these discussions, we suggest using two co-evolving populations, which have complementary effects in selection, to balance the convergence and diversity of the search process. More specifically, one population uses a Pareto-based technique to strengthen the selection pressure for the convergence towards the PF; the other one uses a decomposition-based technique to facilitate the preservation of population diversity. It is worth noting that any existing technique can be applied to these two populations in a plug-in manner (some simple instantiations will be presented in Section 4).

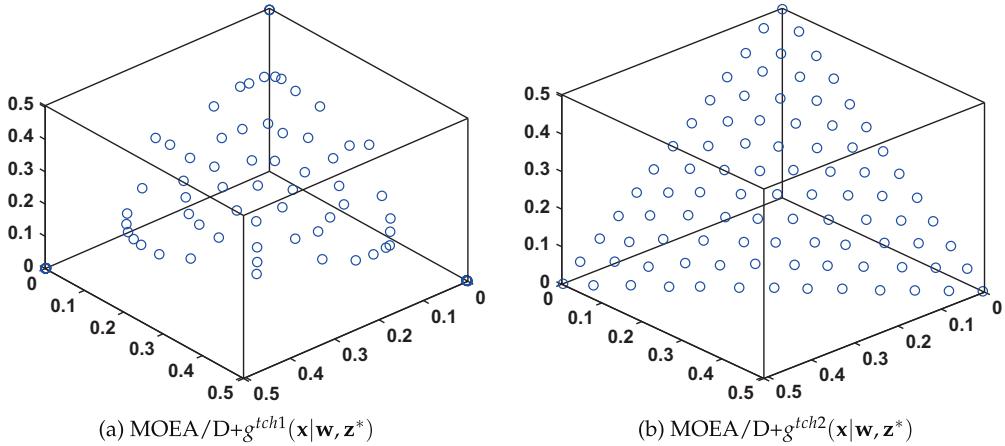


Fig. 1. Comparison results of MOEA/D [19] (the population size is 91 and number of function evaluations is 30,000) with two different TCH approaches on DTLZ1 test instance [13]. It is clear that solutions obtained by MOEA/D with the original TCH approach have a biased distribution along the PF. In contrast, MOEA/D with the modified TCH approach obtains a set of more uniformly distributed solutions.

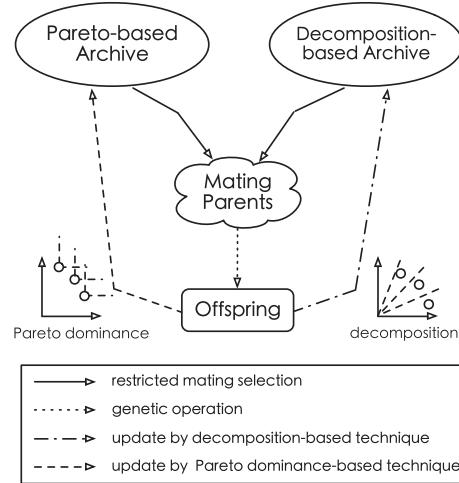


Fig. 2. System architecture of the dual-population paradigm.

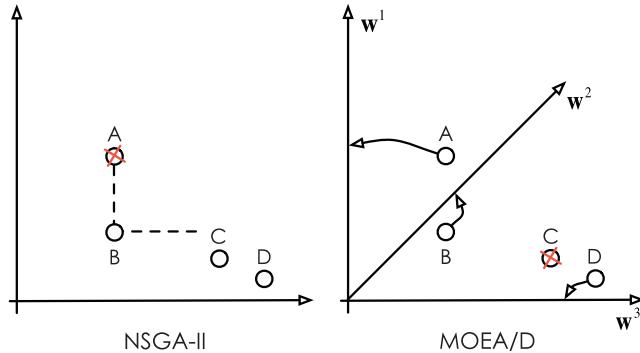


Fig. 3. Comparison of selection results of NSGA-II and MOEA/D.

3.2. Interaction between two co-evolving populations

The interaction between A^p and A^d is a vital step of DPP. In this paper, we define this interaction as the selection of mating parents, from each of these two populations, respectively, for offspring generation, based on our proposed RMS mechanism. As a consequence, the offspring inherits the advantages from both populations, and the search process is expected to strike a

balance between convergence and diversity. As reported in [7], differential evolution (DE) has shown powerful search abilities for continuous optimization. Due to its vector-based property, DE facilitates the design of RMS mechanism for exploiting guidance information towards the optima. Therefore, this paper uses DE and polynomial mutation [8] for offspring generation. However, any other genetic operator with an appropriate adaptation can also serve this purpose. Specifically, an offspring solution $\bar{\mathbf{x}} = \{\bar{x}_1, \dots, \bar{x}_n\}$ is generated as follows:

$$u_j = \begin{cases} x_j^i + F(x_j^{r_1} - x_j^{r_2}) & \text{if } rand < CR \text{ or } j = j_{rand} \\ x_j & \text{otherwise} \end{cases} \quad (4)$$

where $j \in \{1, \dots, n\}$, $rand \in [0, 1]$, CR and F are two control parameters and j_{rand} is a random integer uniformly chosen from 1 to n . Then, a polynomial mutation is applied on \mathbf{u} to obtain $\bar{\mathbf{x}}$:

$$\bar{x}_j = \begin{cases} u_j + \sigma_j(b_j - a_j) & \text{if } rand < p_m \\ u_j & \text{otherwise} \end{cases} \quad (5)$$

with

$$\sigma_j = \begin{cases} (2rand)^{\frac{1}{\eta+1}} - 1 & \text{if } rand < 0.5 \\ 1 - (2 - 2rand)^{\frac{1}{\eta+1}} & \text{otherwise} \end{cases} \quad (6)$$

where the distribution index η and the mutation rate p_m are two user-defined parameters. For simplicity, the violated variable is set to its nearest boundary.

In the traditional DE operator [7], \mathbf{x}^{r_1} and \mathbf{x}^{r_2} in Eq. (4) are usually randomly sampled from the entire population. This random mating selection mechanism has a good exploration ability. However, since no guidance information towards the optima has been considered, it might cause a degeneration problem. Take Fig. 4 as an example, if we set $\mathbf{x}^i \leftarrow \mathbf{x}^1$, $\mathbf{x}^{r_1} \leftarrow \mathbf{x}^2$ and $\mathbf{x}^{r_2} \leftarrow \mathbf{x}^3$, based on Eq. (4), the offspring might be the green star \mathbf{u}^1 . In this case, although \mathbf{x}^1 and \mathbf{x}^3 are close to the PS (denoted as the blue dotted curve), \mathbf{u}^1 is guided away from it. On the other hand, if we set the mating parents as $\mathbf{x}^i \leftarrow \mathbf{x}^4$, $\mathbf{x}^{r_1} \leftarrow \mathbf{x}^6$ and $\mathbf{x}^{r_2} \leftarrow \mathbf{x}^5$, according to Eq. (4), we might obtain the offspring solution \mathbf{u}^2 , which is closer to the PS.

In principle, our proposed RMS mechanism is to exploit the guidance information towards the optima, so that the offspring is expected to approach the PS as much as possible. To this end, we at first specify N unit vectors $\mathbf{v}^1, \dots, \mathbf{v}^N$ in \mathbb{R}_+^m , which divide \mathbb{R}_+^m into N subregions $\Lambda^1, \dots, \Lambda^N$. In particular, a subregion $\Lambda^i, i \in \{1, \dots, N\}$ is defined as:

$$\Lambda^i = \{\mathbf{F}(\mathbf{x}) \in \mathbb{R}_+^m | \langle \mathbf{F}(\mathbf{x}), \mathbf{v}^i \rangle \leq \langle \mathbf{F}(\mathbf{x}), \mathbf{v}^j \rangle, \forall j \in \{1, \dots, N\}\} \quad (7)$$

where $\langle \mathbf{F}(\mathbf{x}), \mathbf{v} \rangle$ is the Euclidian distance between $\mathbf{F}(\mathbf{x})$ and \mathbf{v} . After the setup of subregions, each solution in A^p and A^d is associated with a subregion based on its position in \mathbb{R}_+^m . Specifically, for a solution \mathbf{x} , the index of its associated subregion Λ^k is determined as:

$$k = \operatorname{argmin}_{i \in \{1, \dots, N\}} \langle \bar{\mathbf{F}}(\mathbf{x}), \mathbf{v}^i \rangle \quad (8)$$

where $\bar{\mathbf{F}}(\mathbf{x})$ is the normalized objective vector of \mathbf{x} , and its k th objective function is calculated as:

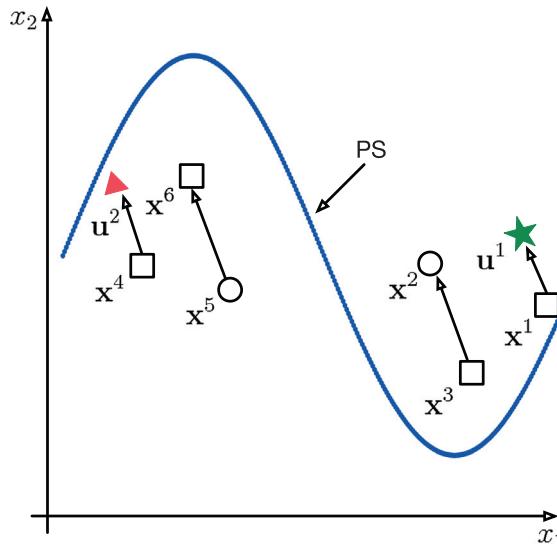


Fig. 4. A simple illustration of the RMS mechanism.

$$\bar{f}_k(\mathbf{x}) = \frac{f_k(\mathbf{x}) - z_k^*(\mathbf{x})}{z_k^{nad}(\mathbf{x}) - z_k^*(\mathbf{x})} \quad (9)$$

where $k \in \{1, \dots, m\}$. It is worth noting that a recently proposed algorithm MOEA/D-M2M [23] also employs a similar method to divide \mathbb{R}_+^m . However, its idea is in essence different from our method. In particular, the specification of a subregion in this paper is to identify the position of a solution; while, in MOEA/D-M2M, different weight vectors are used to specify different subpopulations for solving several simplified MOPs.

Under some mild conditions, solutions in a neighboring area should have similar structure when solving continuous MOPs [38]. To take advantages of neighborhood information, similar to [36], we specify the neighborhood of each subregion based on the Euclidean distance between unit vectors. Thereafter, we restrict the mating parents to be selected from several neighboring subregions with a large probability δ . On the other hand, as discussed in [19], in order to enhance the exploration ability, we also allow the mating parents to be selected from the whole population with a low probability $1 - \delta$. Furthermore, we restrict \mathbf{x}^{r_1} to be selected from A^P , which is assumed to provide solutions with good convergence properties; while \mathbf{x}^{r_2} is restricted to be chosen from A^d , which is supposed to maintain a population with a promising diversity. In this way, we have a large chance to generate offspring solutions closer to the optima. The pseudo-code of the RMS mechanism is given in [Algorithm 1](#). Since the archiving mechanism of A^P does not depend on the uniformly spread unit vectors, there is no guarantee that each subregion could be associated with a solution in A^P . If A^P does not contain any solution in the selected subregion (line 6 of [Algorithm 1](#)), we just choose one from the corresponding subregion in A^d . Let us consider a simple example in [Fig. 5](#). \mathbf{v}^1 to \mathbf{v}^6 are six uniformly distributed unit vectors, which specify six subregions Λ^1 to Λ^6 . Let $\mathbf{x}^i \leftarrow \mathbf{x}^1$, where \mathbf{x}^1 is associated with Λ^2 . Assume that \mathbf{x}^2 , associated with Λ^4 , is selected from A^P , and set $\mathbf{x}^{r_1} \leftarrow \mathbf{x}^2$; \mathbf{x}^3 , associated with Λ^3 , is selected from A^d , and set $\mathbf{x}^{r_2} \leftarrow \mathbf{x}^3$. After the reproduction operation, we might obtain the offspring $\bar{\mathbf{x}}$ (denoted as the red triangle).

Algorithm 1. RE_{MATSEL}(A^P, A^d, i, B_i)

Input: Pareto-based archive A^P , decomposition-based archive A^d , subregion index i , neighborhood index set B_i
Output: mating parents $Q = \{\mathbf{x}^{r_1}, \mathbf{x}^{r_2}\}$

```

1 if rand <  $\delta$  then
2   | Randomly select two indices  $j$  and  $k$  from  $B_i$ ;
3 else
4   | Randomly select two indices  $j$  and  $k$  from  $\{1, 2, \dots, N\}$ ;
5 end
6 if  $\nexists \dot{\mathbf{x}} \in A^P$  belongs to  $\Lambda^j$  then
7   | Randomly select a solution  $\ddot{\mathbf{x}} \in A^d$  that belongs to  $\Lambda^j$ , and set  $\mathbf{x}^{r_1} \leftarrow \ddot{\mathbf{x}}$ ;
8 else
9   | Randomly select a solution  $\dot{\mathbf{x}} \in A^P$  that belongs to  $\Lambda^j$ , and set  $\mathbf{x}^{r_1} \leftarrow \dot{\mathbf{x}}$ ;
10 end
11 Randomly select a solution  $\ddot{\mathbf{x}} \in A^d$  that belongs to  $\Lambda^k$ , and set  $\mathbf{x}^{r_2} \leftarrow \ddot{\mathbf{x}}$ ;
12 return  $Q \leftarrow \{\mathbf{x}^{r_1}, \mathbf{x}^{r_2}\}$ 
```

4. Instantiations of DPP

The instantiation of DPP only needs to consider two major issues: one is the choice of Pareto- and decomposition-based techniques for A^P and A^d ; and the other is the design of the interaction mechanism between A^P and A^d . Since the latter issue has been elaborated in [Section 3.2](#), this section only discusses the prior one. Note that the instantiations introduced here employs a steady-state evolution scheme to update the population. In other words, after the generation of an offspring solution, it is used to update each of A^P and A^d , respectively.

As mentioned in [Section 3.1](#), DPP is a simple and general framework, to which any existing EMO technique can be applied in a plug-in manner, with appropriate modifications. For the Pareto-based archive A^P , we use the techniques from two classic Pareto-based EMO algorithms:

1. NSGA-II [12]: Since DPP uses a steady-state evolution scheme to update the population, the original archiving mechanism of NSGA-II is modified as a three-stage procedure: a) if the current offspring solution $\bar{\mathbf{x}}$ is dominated by any member in A^P , it is rejected by A^P ; b) if $\bar{\mathbf{x}}$ dominates some members in A^P , the one in the worst non-domination level will be eliminated (tie is broken by crowding distance); c) if $\bar{\mathbf{x}}$ is non-dominated with all members in A^P , the one with the smallest crowding distance is eliminated.

2. ε -MOEA [11]: It is a steady-state EMO algorithm that uses the ε -dominance relation [18] to strengthen the selection pressure towards the PF. Without any modification, the archiving mechanism of ε -MOEA for the external archive is directly applied to A^P . More detailed information can be found from the original reference [11].

For the decomposition-based archive A^d , we employ the methods of two classic MOEA/D variants:

1. MOEA/D-DE [19]: It is a variant of MOEA/D [36], whose reproduction operator has been replaced by DE operator. Moreover, in order to increase the population diversity, only a limited number of solutions can be replaced by the high quality offspring.
2. MOEA/D-DRA [37]: It is the winning algorithm in CEC2009 MOEA competition. The major difference between MOEA/D-DRA and MOEA/D-DE is its dynamical resource allocation scheme, which dynamically allocates computational resources to different subproblems according to their utilities.

Note that the update procedure of A^d is a slightly different from the original MOEA/D. In particular, each offspring can only be compared with the parent solutions in its associated subregion. This guarantees that each subregion in A^d owns one solution.

Without loss of generality, we present the pseudo-code of one DPP instantiation (denoted as ND/DPP), which applies the modified archiving mechanisms of NSGA-II and MOEA/D-DE for A^P and A^d , in [Algorithm 2](#)³. Other instantiations, i.e., ND/DPP-DRA which combines NSGA-II and MOEA/D-DRA, ED/DPP which combines ε -MOEA and MOEA/D-DE, and ED/DPP-DRA which combines ε -MOEA and MOEA/D-DRA, can be done in a similar way.

Algorithm 2. ND/DPP

Output: final population P

```

1  $[A^P, A^d, W, V, B, \mathbf{z}^*, \mathbf{z}^{nad}] \leftarrow \text{INITIALIZATION}();$ 
2  $neval \leftarrow 0$            //number of function evaluations
3 while stopping criterion is not satisfied do
4   for  $i \leftarrow 1$  to  $N$  do
5      $Q \leftarrow \text{REMATSEL}(A^P, A^d, i, B_i);$ 
6     Generate an offspring solution  $\bar{\mathbf{x}}$  based on the mating parents in  $Q$ , according to Eqs. (4) and (5);
7     Identify the associated subregion of  $\bar{\mathbf{x}}$  in  $A^P$  and  $A^d$ , respectively, according to Eq. (8);
8     Update  $A^P$  and  $A^d$  based on the archiving mechanisms of NSGA-II and MOEA/D-DE;
9     Update ideal and nadir objective vectors  $\mathbf{z}^*, \mathbf{z}^{nad}$ ;
10     $neval++;$ 
11  end
12 end
13 return  $P \leftarrow A^P \cup A^d$ 
```

Algorithm 3. INITIALIZATION()

Output: Pareto-based archive A^P , decomposition-based archive A^d , weight vector set W , unit vector set V , neighborhood index sets B , ideal objective vector \mathbf{z}^* , nadir objective vector \mathbf{z}^{nad}

```

1 Initialize the weight vector set  $W \leftarrow \{\mathbf{w}^1, \dots, \mathbf{w}^N\}$  according to the method introduced in [6], and set the unit vector set  $V \leftarrow \{\mathbf{v}^1, \dots, \mathbf{v}^N\}$  as  $W$            //weight vectors and unit vectors are set the same
2 for  $i \leftarrow 1$  to  $N$  do
3    $B_i \leftarrow \{i_1, \dots, i_T\}$  where  $\mathbf{w}^{i_1}, \dots, \mathbf{w}^{i_T}$  are the  $T$  closest weight vectors to  $\mathbf{w}^i$            //this neighborhood index set is directly applicable for subregions
4 end
5 Randomly sample a population of solutions from  $\Omega$  and assign them to  $A^P$  and  $A^d$ ;
6 Randomly assign each solution in  $A^P$  and  $A^d$  to a subregion;
7 Initialize the ideal and nadir objective vectors  $\mathbf{z}^*, \mathbf{z}^{nad}$ ;
8 end
9 return  $A^P, A^d, W, V, B, \mathbf{z}^*, \mathbf{z}^{nad}$ 
```

³ In initialization procedure, depicted in [Algorithm 3](#), the unit vectors for specifying subregions are set the same as the weight vectors for specifying subproblems.

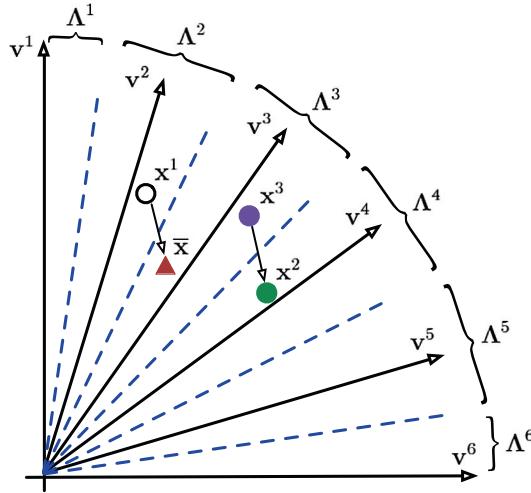


Fig. 5. A simple illustration of offspring generation via RMS mechanism.

The time complexity of a DPP instantiation depends on the designs of A^p and A^d . Here, without loss of generality, we only take ND/DPP as an example for illustration. Specifically, the most time consuming part of initialization procedure (line 1 in [Algorithm 2](#)) is the identification of neighborhood index sets B (line 2 to line 4 in [Algorithm 3](#)) which needs $O(N^2)$ operations. In the main for-loop, line 5 to line 6 in [Algorithm 2](#) just pick two parent solutions for offspring generation. In line 7, the identification of associated subregion costs $O(N)$ comparisons. For updating A^p , the non-dominated sorting, which divides the population into several non-dominated levels, costs $O(N \log N)$ comparisons. In addition, the modified archiving mechanism of NSGA-II described above costs $O(N)$ comparisons. For updating A^d , the only operation is the comparison of aggregation function values between the offspring \bar{x} and the parent in its corresponding subregion. Therefore, the time complexity of line 8 is $O(N \log N)$. In summary, the time complexity of ND/DPP is $O(N^2 \log N)$ since it has N passes for one generation (i.e., generating N offspring solutions). Since A^p and A^d are with the same size, N solutions are stored in each archive per generation. Therefore, the memory complexity of ND/DPP is $\Theta(N)$.

5. Performance verification of DPP

5.1. Experimental setup

Seventeen unconstrained MOP test instances (UF1 to UF10, MOP1 to MOP7) are employed here as the benchmark problems. Specifically, UF1 to UF10 are used as the benchmark problems for CEC2009 MOEA competition [39], and MOP1 to MOP7 are recently suggested in [23]. These test instances are with distinct characteristics and complicated PSs in the decision space. According to the settings in [39,23], the number of decision variables of UF1 to UF10 is set to 30; for MOP1 to MOP7, the number of decision variables is set to 10. More detailed information can be found in their corresponding references [39,23].

As discussed in [44], no unary performance metric can give a comprehensive assessment on the performance of an EMO algorithm. In our empirical studies, we consider the following two widely used performance metrics.

1. *Inverted Generational Distance* (IGD) metric [5]: Let P^* be a set of points uniformly sampled along the PF, and P be the set of solutions obtained by an EMO algorithm. The IGD value of P is calculated as:

$$IGD(P, P^*) = \frac{\sum_{\mathbf{x} \in P^*} dist(\mathbf{x}, P)}{|P^*|} \quad (10)$$

where $dist(\mathbf{x}, P)$ is the Euclidean distance between the solution \mathbf{x} and its nearest neighbor in P , and $|P^*|$ is the cardinality of P^* . The PF of the underlying MOP is assumed to be known a priori when calculating the IGD metric. In our empirical studies, 1000 uniformly distributed points are sampled along the PF for the bi-objective test instances, and 10,000 for three-objective ones, respectively.

2. *Hypervolume* (HV) metric [43]: Let $\mathbf{z}^r = (z_1^r, \dots, z_m^r)^T$ be an anti-optimal reference point in the objective space that is dominated by all Pareto-optimal objective vectors. HV metric measures the size of the objective space dominated by solutions in P and bounded by \mathbf{z}^r .

Table 1

Parameter settings of the EMO algorithms.

MOEAs	Parameter settings
NSGA-II	$p_c = 0.9, p_m = 1/nreal, \mu_c = 20, \mu_m = 20$
ε -MOEA	$p_c = 0.9, p_m = 1/nreal, \mu_c = 20, \mu_m = 20, \varepsilon = 1/600$ for UF1 to UF7, $\varepsilon = 1/60$ for UF8 to UF10 $\varepsilon = 1/13$ for MOP1 to MOP5, $\varepsilon = 1/23$ for MOP6 to MOP7
MOEA/D-DE	$p_m = 1/nreal, \mu_m = 20, CR = 1.0, F = 0.5, \delta = 0.9, T = 20, n_r = 2$
MOEA/D-DRA	$p_m = 1/nreal, \mu_m = 20, CR = 1.0, F = 0.5, \delta = 0.9, T = 0.1 * N, n_r = 0.01 * N$

$$HV(P) = \text{Vol} \left(\bigcup_{\mathbf{x} \in P} [f_1(\mathbf{x}), z_1^r] \times \dots \times [f_m(\mathbf{x}), z_m^r] \right) \quad (11)$$

where $\text{Vol}(\cdot)$ is the Lebesgue measure. In our empirical studies, $\mathbf{z}^r = (2.0, 2.0)^T$ for bi-objective test instances and $\mathbf{z}^r = (2.0, 2.0, 2.0)^T$ for three-objective ones.

Both IGD and HV metrics can measure the convergence and diversity of P simultaneously. The lower is the IGD value (the larger is the HV value), the better is the quality of P for approximating the true PF. Comparison results are presented in the corresponding data tables, where the best mean metric values are highlighted in bold face with gray background. In order to have statistically sound conclusions, Wilcoxon's rank sum test at a 5% significance level is adopted to compare the significance of difference between the performances of two algorithms. It is worth noting that the output of a DPP instantiation is a combination of A^P and A^d , thus its size is larger than the regular EMO algorithms. For peer comparison, a data preprocessing is carried out on P before evaluating performance metric values. First of all, M evenly spread weight vectors are generated according to the method introduced in [6]. Afterwards, for each weight vector, the solution that achieves the best aggregation function⁴ value is selected for the final performance assessment. The pseudo-code of this data preprocessing is given in [Algorithm 4](#). Specifically, $M = 600$ for UF1 to UF7, $M = 990$ for UF8 to UF10, $M = 100$ for MOP1 to MOP5 and $M = 300$ for MOP6 to MOP7.

Algorithm 4. DATAPROCESS(S)

Input: solution set S

Output: final population P

- 1 Initialize the weight vector set $W = \{\mathbf{w}^1, \dots, \mathbf{w}^M\}$;
- 2 Find the ideal objective vector \mathbf{z}^* of S ;
- 3 $P \leftarrow \emptyset$;
- 4 **for** $i \leftarrow 1$ **to** M **do**
- 5 $k \leftarrow \text{argmin}_{\mathbf{x} \in S} \{g^{tch2}(\mathbf{x}|\mathbf{w}^i, \mathbf{z}^*)\}$;
- 6 $P \leftarrow P \cup \{\mathbf{x}^k\}$;
- 7 $S \leftarrow S \setminus \{\mathbf{x}^k\}$;
- 8 **end**
- 9 **return** P

The parameters of NSGA-II, ε -MOEA, MOEA/D-DE, MOEA/D-DRA are set according to their corresponding references [12,11,19,37], as summarized in [Table 1](#). According to the settings in [39,23], the population size is set as $N = 600$ for UF1 to UF7, $N = 1000$ for UF8 to UF10, $N = 100$ for MOP1 to MOP5 and $N = 300$ for MOP6 to MOP7. Each algorithm is independently launched 20 times on each test instance. The termination criterion of an algorithm is the predefined number of function evaluations, which is constantly set to 300,000.

5.2. DPP instantiations vs baseline algorithms

This section presents the performance comparisons of four DPP instantiations with their corresponding baseline algorithms.

5.2.1. ND/DPP vs NSGA-II and MOEA/D-DE

[Table 2](#) shows the performance comparisons of ND/DPP, NSGA-II and MOEA/D-DE on UF and MOP instances, regarding the IGD and HV metrics, respectively. From the experimental results, we find that ND/DPP shows a clearly better performance than both NSGA-II and MOEA/D-DE. It achieves better metric values in 62 out of 68 comparisons, where 61 of those

⁴ Here we use the TCH approach introduced in (3).

Table 2

Performance comparisons between ND/DPP and NSGA-II, MOEA/D-DE on IGD and HV metrics.

Test instances	IGD			HV		
	NSGA-II	MOEA/D-DE	ND/DPP	NSGA-II	MOEA/D-DE	ND/DPP
UF1	6.953E–2(9.10E–3) [†]	9.854E–4(1.13E–4) [†]	9.173E–4(3.28E–5)	3.4957(5.44E–2) [†]	3.6617(2.01E–3)	3.6618(1.28E–3)
UF2	1.843E–2(4.92E–3) [†]	5.792E–3(2.03E–3) [†]	2.492E–3(4.42E–4)	3.6139(2.62E–2) [†]	3.6518(6.51E–3) [†]	3.6589(1.55E–3)
UF3	9.479E–2(2.53E–2) [†]	1.491E–2(1.42E–2) [†]	6.667E–3(3.58E–3)	3.1720(9.29E–2) [†]	3.6256(6.69E–2) [†]	3.6547(8.34E–3)
UF4	4.097E–2(5.86E–4)[‡]	5.628E–2(2.59E–3) [†]	5.134E–2(8.19E–4)	3.220T(4.09E–3)[‡]	3.1667(1.58E–2) [†]	3.1927(3.13E–3)
UF5	2.186E–1(3.72E–2) [†]	3.107E–1(4.97E–2) [†]	1.405E–1(4.78E–2)	2.7682(1.21E–1) [†]	2.6482(1.49E–1) [†]	3.1607(1.07E–1)
UF6	1.737E–1(6.16E–2) [†]	9.826E–2(9.15E–2) [†]	6.218E–2(1.31E–3)	2.8759(1.86E–1) [†]	3.1047(2.53E–1) [†]	3.2160(8.18E–3)
UF7	5.099E–2(7.31E–2) [†]	1.573E–3(3.56E–4) [†]	1.041E–3(7.88E–5)	3.3500(2.32E–1) [†]	3.4904(6.05E–3) [†]	3.4943(9368E–4)
UF8	1.553E–1(3.43E–2) [†]	4.772E–2(9.82E–3) [†]	2.830E–2(3.25E–3)	6.6380(2.21E–1) [†]	7.3650(2.27E–2) [†]	7.4071(1.23E–2)
UF9	1.075E–1(4.78E–2) [†]	5.412E–2(3.98E–2) [†]	2.475E–2(1.25E–3)	7.3048(2.68E–1) [†]	7.5659(1.77E–1) [†]	7.7237(1.06E–2)
UF10	4.523E–1(2.47E–2)[‡]	5.258E–1(6.78E–2) [‡]	9.261E–1(1.25E–1)	3.9050(2.21E–1)[‡]	3.4731(2.77E–1) [‡]	1.5329(3.62E–1)
MOP1	3.645E–1(4.19E–3) [†]	3.575E–1(9.83E–3) [†]	2.192E–2(1.56E–3)	3.0772(1.11E–2) [†]	3.0885(2.18E–2) [†]	3.6346(1.91E–3)
MOP2	3.543E–1(5.55E–2) [†]	2.891E–1(6.77E–2) [†]	5.958E–3(3.35E–4)	3.0000(2.02E–2) [†]	3.0313(3.83E–2) [†]	3.3226(4.42E–4)
MOP3	1.119E–1(1.15E–2) [†]	1.180E–1(5.10E–2) [†]	2.106E–2(8.00E–5)	3.1148(1.44E–2) [†]	3.0855(5.46E–2) [†]	3.1779(1.59E–4)
MOP4	3.107E–1(2.04E–2) [†]	2.872E–1(3.32E–2) [†]	1.626E–2(1.29E–3)	3.1413(7.67E–3) [†]	3.1517(2.07E–2) [†]	3.4998(1.01E–3)
MOP5	2.796E–1(3.26E–2) [†]	3.163E–1(6.94E–3) [†]	1.749E–2(6.90E–4)	3.0753(1.42E–1) [†]	2.7448(1.44E–1) [†]	3.6393(9.17E–4)
MOP6	3.044E–1(5.75E–6) [†]	2.975E–1(1.81E–2) [†]	5.441E–2(1.84E–3)	7.4978(6.46E–5) [†]	7.5062(2.74E–2) [†]	7.7610(2.20E–3)
MOP7	3.509E–1(7.99E–6) [†]	3.389E–1(2.93E–2) [†]	8.397E–2(4.94E–3)	7.2130(3.80E–5) [†]	7.2214(3.00E–2) [†]	7.3671(5.18E–3)

Wilcoxon's rank sum test at a 0.05 significance level is performed between ND/DPP and each of NSGA-II and MOEA/D-DE. [†] and [‡] denote the performance of the corresponding algorithm is significantly worse than and better than that of the proposed ND/DPP, respectively. And the best mean metric value is highlighted in boldface with gray background.

Table 3

Performance comparisons between ND/DPP-DRA and NSGA-II, MOEA/D-DRA on IGD and HV metrics.

Test instances	IGD			HV		
	NSGA-II	MOEA/D-DRA	ND/DPP-DRA	NSGA-II	MOEA/D-DRA	ND/DPP-DRA
UF1	6.953E–2(9.10E–3) [†]	9.787E–4(6.22E–5) [†]	9.397E–4(4.84E–5)	3.4957(5.44E–2) [†]	3.6624(9.64E–4) [†]	3.6628(6.99E–4)
UF2	1.843E–2(4.92E–3) [†]	2.541E–3(1.18E–3) [†]	1.353E–3(4.11E–4)	3.6139(2.62E–2) [†]	3.6533(1.81E–2) [†]	3.6588(1.42E–3)
UF3	9.479E–2(2.53E–2) [†]	9.188E–3(1.16E–2) [†]	4.929E–3(3.82E–3)	3.1720(9.29E–2) [†]	3.6409(5.57E–2) [†]	3.6582(5.94E–3)
UF4	4.097E–2(5.86E–4)[‡]	5.471E–2(3.68E–3) [†]	5.070E–2(1.31E–3)	3.220T(4.09E–3)[‡]	3.1691(1.52E–2) [†]	3.1828(4.47E–3)
UF5	2.186E–1(3.72E–2) [†]	2.976E–1(5.37E–2) [†]	1.949E–1(5.42E–2)	2.7682(1.21E–1) [†]	2.6912(1.52E–1) [†]	3.0216(1.29E–1)
UF6	1.737E–1(6.16E–2) [†]	1.592E–1(2.23E–1) [†]	6.646E–2(3.53E–2)	2.8759(1.86E–1) [†]	3.0432(3.66E–1) [†]	3.1673(7.49E–2)
UF7	5.099E–2(7.31E–2) [†]	1.111E–3(1.96E–4) [†]	9.089E–4(5.32E–5)	3.3500(2.32E–1) [†]	3.4947(4.14E–3)	3.4959(1.33E–3)
UF8	1.553E–1(3.43E–2) [†]	3.101E–2(5.26E–3) [†]	2.481E–2(1.84E–3)	6.6380(2.2E–1) [†]	7.4063(9.12E–3) [†]	7.4152(9.33E–3)
UF9	1.075E–1(4.78E–2) [†]	4.478E–2(3.96E–2) [†]	2.174E–2(1.02E–3)	7.3048(2.68E–1) [†]	7.6388(1.75E–1) [†]	7.7457(7.59E–3)
UF10	4.523E–1(2.47E–2)[‡]	4.805E–1(4.85E–2) [‡]	5.157E–1(7.13E–2)	3.9050(2.21E–1)[‡]	3.5769(2.46E–1) [‡]	3.8575(3.19E–1)
MOP1	3.645E–1(4.19E–3) [†]	3.453E–1(4.72E–2) [†]	2.313E–2(9.29E–4)	3.0772(1.11E–2) [†]	3.1102(8.83E–2) [†]	3.6323(1.22E–3)
MOP2	3.543E–1(5.55E–2) [†]	2.326E–1(5.45E–2) [†]	5.425E–3(1.15E–4)	3.0000(2.02E–2) [†]	3.0664(4.72E–2) [†]	3.3239(2.03E–4)
MOP3	1.119E–1(1.15E–2) [†]	1.295E–1(5.12E–2) [†]	4.773E–3(4.40E–5)	3.1148(1.44E–2) [†]	3.0784(5.11E–2) [†]	3.2076(1.29E–4)
MOP4	3.107E–1(2.04E–2) [†]	2.885E–1(3.31E–2) [†]	1.623E–2(6.79E–4)	3.1413(7.67E–3) [†]	3.1557(2.26E–2) [†]	3.5002(7.86E–4)
MOP5	2.796E–1(3.26E–2) [†]	3.165E–1(5.47E–3) [†]	1.785E–2(8.47E–4)	3.0753(1.42E–1) [†]	2.7598(1.14E–1) [†]	3.6366(1.49E–3)
MOP6	3.044E–1(5.75E–6) [†]	2.952E–1(1.88E–2) [†]	5.099E–2(1.49E–3)	7.4978(6.46E–5) [†]	7.5096(2.83E–2) [†]	7.7664(1.73E–3)
MOP7	3.509E–1(7.99E–6) [†]	3.477E–1(1.85E–2) [†]	7.834E–2(4.50E–3)	7.2130(3.80E–5) [†]	7.2150(2.13E–2) [†]	7.3686(3.75E–3)

Wilcoxon's rank sum test at a 0.05 significance level is performed between ND/DPP-DRA and each of NSGA-II and MOEA/D-DRA. [†] and [‡] denote the performance of the corresponding algorithm is significantly worse than and better than that of the proposed ND/DPP-DRA, respectively. And the best mean metric value is highlighted in boldface with gray background.

better results achieved by ND/DPP are with statistical significance. It is worth noting that although the performance of NSGA-II is not satisfied on most test instances, it obtains the best metric values on UF4 and UF10 instances. By contrast, ND/DPP shows extremely poor a performance on UF10 instance, which is featured by many local optima. MOP instances, recently proposed in [23], not only have nonlinear PSs, but also require the algorithm to have a good ability to maintain the population diversity. As reported in [23], the performance of both NSGA-II and MOEA/D-DE are extremely poor on these instances. However, by integrating the Pareto- and decomposition-based techniques in DPP, ND/DPP shows significantly better performance than both of its baseline algorithms on all MOP instances. These observations demonstrate the effectiveness of DPP for balancing convergence and diversity.

5.2.2. ND/DPP-DRA vs NSGA-II and MOEA/D-DRA

Table 3 provides the performance comparisons of ND/DPP-DRA, NSGA-II and MOEA/D-DRA on UF and MOP instances, regarding the IGD and HV metrics, respectively. Similar to the observations in Section 5.2.1, ND/DPP-DRA achieves better

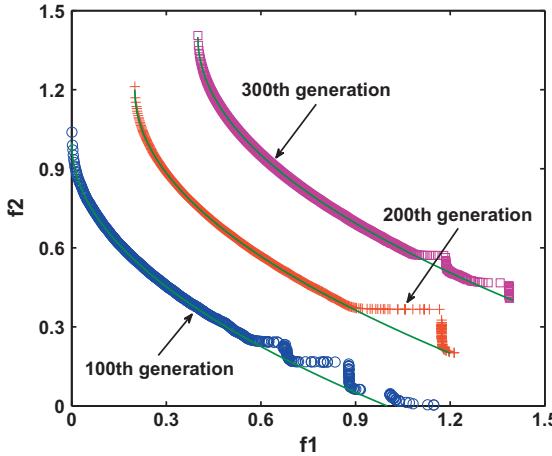


Fig. 6. Solutions obtained by MOEA/D-DE [19] in different stages of evolution ($N = 600$).

Table 4

Performance comparisons between ED/DPP and ε -MOEA, MOEA/D-DE on IGD and HV metrics.

Test instances	IGD			HV		
	ε -MOEA	MOEA/D-DE	ED/DPP	ε -MOEA	MOEA/D-DE	ED/DPP
UF1	8.090E-2(2.22E-2) [†]	9.854E-4(1.13E-4) [†]	9.059E-4(3.30E-5)	3.4274(1.03E-1) [†]	3.6617(2.01E-3) [†]	3.6624(3.38E-4)
UF2	3.617E-2(9.79E-3) [†]	5.792E-3(2.03E-3) [†]	1.492E-3(8.76E-4)	3.5524(5.17E-2) [†]	3.6518(6.51E-3) [†]	3.6588(1.25E-3)
UF3	1.331E-1(3.15E-2) [†]	1.491E-2(1.42E-2) [†]	3.113E-3(4.75E-3)	3.0059(1.23E-1) [†]	3.6256(6.69E-2) [†]	3.6614(8.11E-3)
UF4	4.108E-2(9.35E-4)[‡]	5.628E-2(2.59E-3) [†]	4.464E-2(1.25E-3)	3.2166(4.28E-3)[‡]	3.1667(1.58E-2) [†]	3.2071(3.41E-3)
UF5	2.129E-1(2.92E-2) [†]	3.107E-1(4.97E-2) [†]	1.936E-1(7.08E-2)	2.7010(1.43E-1) [†]	2.6482(1.49E-1) [†]	3.0835(1.54E-1)
UF6	2.041E-1(7.39E-2) [†]	9.826E-2(9.15E-2) [†]	3.989E-2(1.09E-2)	2.7675(2.19E-1) [†]	3.1047(2.53E-1) [†]	3.2521(3.77E-2)
UF7	1.199E-1(1.26E-1) [†]	1.573E-3(3.56E-4) [†]	1.009E-3(3.36E-4)	3.1619(3.93E-1) [†]	3.4904(6.05E-3) [†]	3.4941(1.26E-3)
UF8	3.084E-1(6.74E-2) [†]	4.772E-2(9.82E-3) [†]	2.304E-2(5.34E-3)	6.5239(2.23E-1) [†]	7.3650(2.27E-2) [†]	7.4148(1.54E-2)
UF9	1.281E-1(5.22E-2) [†]	5.412E-2(3.98E-2) [†]	1.910E-2(1.63E-3)	7.0648(3.24E-1) [†]	7.5659(1.77E-1) [†]	7.7491(7.68E-3)
UF10	6.672E-1(7.07E-2) [†]	5.258E-1(6.78E-2) [†]	4.142E-1(2.16E-1)	3.4617(9.17E-1) [†]	3.4731(2.77E-1) [†]	4.1711(7.44E-1)
MOP1	3.793E-1(5.29E-3) [†]	3.575E-1(9.83E-3) [†]	1.711E-2(2.88E-3)	3.0374(1.36E-2) [†]	3.0885(2.18E-2) [†]	3.6409(3.65E-3)
MOP2	3.539E-1(2.19E-3) [†]	2.891E-1(6.77E-2) [†]	5.444E-3(7.27E-4)	3.0002(9.48E-4) [†]	3.0313(3.83E-2) [†]	3.3234(1.25E-3)
MOP3	1.591E-1(2.28E-2) [†]	1.180E-1(5.10E-2) [†]	6.547E-3(3.66E-3)	3.0674(2.12E-2) [†]	3.0855(5.46E-2) [†]	3.2040(6.55E-3)
MOP4	3.318E-1(1.27E-2) [†]	2.872E-1(3.32E-2) [†]	1.396E-2(4.36E-3)	3.1001(1.01E-2) [†]	3.1517(2.07E-2) [†]	3.5016(4.07E-3)
MOP5	2.877E-1(3.11E-2) [†]	3.163E-1(6.94E-3) [†]	1.515E-2(1.44E-3)	2.9981(1.58E-1) [†]	2.7448(1.44E-1) [†]	3.6430(1.99E-3)
MOP6	3.046E-1(4.72E-5) [†]	2.975E-1(1.81E-2) [†]	4.657E-2(2.79E-3)	7.4933(6.33E-4) [†]	7.5062(2.74E-2) [†]	7.7714(3.27E-3)
MOP7	3.517E-1(5.29E-4) [†]	3.389E-1(2.93E-2) [†]	7.697E-2(4.13E-3)	7.2081(1.46E-3) [†]	7.2214(3.00E-2) [†]	7.3711(5.32E-3)

Wilcoxon's rank sum test at a 0.05 significance level is performed between ED/DPP and each of ε -MOEA and MOEA/D-DE. [†] and [‡] denote the performance of the corresponding algorithm is significantly worse than and better than that of the proposed ED/DPP, respectively. And the best mean metric value is highlighted in boldface with gray background.

metric values than its baseline NSGA-II and MOEA/D-DRA in 63 out of 68 performance comparisons, where 62 of the better results are with statistical significance. Comparing to MOEA/D-DE, the only difference of MOEA/D-DRA is its dynamical resource allocation scheme, which allocates different amounts of computational efforts to different subproblems, based on their utilities. As discussed in [37], different regions of the PF might have distinct approximation difficulties. In other words, some regions might be easier to converge than the others. To validate this consideration, Fig. 6 plots the solutions obtained by MOEA/D-DE at the 100th, 200th and 300th generation, respectively, on UF2 instance. It is clear that the northwestern part of the PF is easier to approximate than the southeastern part (as the northwestern part does not change any more after the 200th generation). Therefore, it makes sense to allocate more computational effort to those difficult regions. Comparing the results in Tables 3 and 2, we also observe that, by using the dynamical resource allocation scheme, MOEA/D-DRA and ND/DPP-DRA perform better than their corresponding standard versions, i.e., MOEA/D-DE and ND/DPP.

5.2.3. ED/DPP vs ε -MOEA and MOEA/D-DE

Table 4 presents the performance comparisons of ED/DPP, ε -MOEA and MOEA/D-DE on UF and MOP instances, regarding the IGD and HV metrics, respectively. In contrast with the slight difference between ND/DPP and its baseline algorithms, ED/DPP shows significant improvements over ε -MOEA and MOEA/D-DE. It achieves better metric values in 66 out of 68 comparisons, where all those better results are with statistical significance. In fact, the only difference between ND/DPP and ED/DPP is the specification of the Pareto-based archive, where the latter one uses the archiving mechanism of ε -MOEA. As discussed

in [18], the ε -dominance, modified by Pareto dominance, strengthens the selection pressure towards the PF. In this case, the Pareto-based archive of ED/DPP is thus able to provide solutions with a better convergence property. Therefore, better guidance information towards the optima can be exploited by the RMS mechanism. However, it is interesting to note that the performance of ε -MOEA is not as good as NSGA-II. Thereby, the good performance of ED/DPP should be attributed to the interaction and collaboration between the Pareto- and decomposition-based archives.

5.2.4. ED/DPP-DRA vs ε -MOEA and MOEA/D-DRA

Table 5 shows the performance comparisons of ED/DPP, ε -MOEA and MOEA/D-DRA on UF and MOP instances, regarding the IGD and HV metrics, respectively. Comparing to the baseline algorithms, ED/DPP-DRA also shows very competitive performance on all test instances. It obtains better metric values in 66 out of 68 comparisons, where all better results are with statistical significance. Furthermore, similar to the observations in Section 5.2.2, ED/DPP-DRA shows better performance than ED/DPP in 27 out of 34 comparisons. These observations further confirm the importance of dynamically allocating the computation resources.

In summary, empirical studies in this section fully demonstrate the effectiveness of DPP. It is so simple and general that any existing Pareto- and decomposition-based techniques can be applied in a plug-in manner. Depending on the designs of Pareto- and decomposition-based archives, DPP instantiations show different magnitudes of improvements over their corresponding baseline algorithms.

5.3. Effects of the usage of dual populations

In order to understand the underlying mechanisms of DPP, this section considers answering the following three questions:

- Can we have similar or better performance when the two populations in DPP use the same archiving mechanism?
- Can we have similar or better performance by using only a single population in DPP?
- Can we have similar or better performance when using an indicator-based technique in DPP?

In order to answer the above three questions, ED/DPP-DRA is chosen as the representative DPP instantiation for further investigations, in view of its good performance reported in the Section 5.2. Empirical results are reported and discussed in the following paragraphs.

5.3.1. DPP with the same archiving mechanism

In order to answer the first question, we have developed two DPP variants: one (denoted as Dual-NSGA-II) uses the modified archiving mechanism of NSGA-II, in both A^P and A^D , while the other one (denoted as Dual-MOEA/D) employs the modified archiving mechanism of MOEA/D in both populations. The performance comparisons of ED/DPP-DRA, Dual-NSGA-II and Dual-MOEA/D, regarding the IGD and HV metrics, are reported in **Table 6**. From these results, we find that ED/DPP-DRA

Table 5
Performance comparisons between ED/DPP-DRA and ε -MOEA, MOEA/D-DRA on IGD and HV metrics.

Test instances	IGD			HV		
	ε -MOEA	MOEA/D-DRA	ED/DPP-DRA	ε -MOEA	MOEA/D-DRA	ED/DPP-DRA
UF1	8.090E-2(2.26E-2) [†]	9.787E-4(6.22E-5) [†]	8.789E-4(5.75E-5)	3.4274(1.03E-1) [†]	3.6624(9.64E-4) [†]	3.6643(6.89E-4)
UF2	3.617E-2(9.79E-3) [†]	2.541E-3(1.18E-3) [†]	1.313E-3(3.67E-4)	3.5524(5.17E-2) [†]	3.6533(1.81E-2)	3.6566(9.77E-4)
UF3	1.331E-1(3.15E-2) [†]	9.188E-3(1.16E-2) [†]	1.378E-3(4.23E-4)	3.0059(1.23E-1) [†]	3.6409(5.57E-2) [†]	3.6639(6.93E-4)
UF4	4.108E-2(9.35E-4)[‡]	5.471E-2(3.68E-3) [†]	4.738E-2(8.75E-4)	3.2166(4.28E-3)[‡]	3.1691(1.52E-2) [†]	3.1907(3.15E-3)
UF5	2.129E-1(2.92E-2) [†]	2.976E-1(5.37E-2) [†]	1.154E-1(6.72E-2)	2.7010(1.43E-1) [†]	2.6912(1.52E-1) [†]	3.2509(1.47E-1)
UF6	2.041E-1(7.39E-2) [†]	1.592E-1(2.23E-1) [†]	3.027E-2(1.19E-2)	2.7675(2.19E-1) [†]	3.0432(3.66E-1) [†]	3.3165(4.94E-2)
UF7	1.199E-1(1.26E-1) [†]	1.111E-1(1.96E-4) [†]	9.809E-4(5.35E-5)	3.1619(3.93E-1) [†]	3.4947(4.14E-3)	3.4950(1.38E-3)
UF8	3.084E-1(6.74E-2) [†]	3.101E-2(5.26E-3) [†]	2.012E-2(1.77E-3)	6.5239(2.23E-1) [†]	7.4063(9.12E-3) [†]	7.4220(5.56E-3)
UF9	1.281E-1(5.22E-2) [†]	4.478E-2(3.96E-2) [†]	2.042E-2(1.68E-3)	7.0648(3.24E-1) [†]	7.6388(1.75E-1) [†]	7.7509(6.85E-3)
UF10	6.672E-1(7.07E-2) [†]	4.805E-1(4.85E-2) [†]	4.276E-1(2.59E-2)	3.4617(9.17E-1) [†]	3.5769(2.46E-1) [†]	4.1699(2.14E-1)
MOP1	3.793E-1(5.29E-3) [†]	3.453E-1(4.72E-2) [†]	1.528E-2(1.54E-3)	3.0374(1.36E-2) [†]	3.1102(8.83E-2) [†]	3.6429(2.14E-3)
MOP2	3.539E-1(2.19E-3) [†]	2.326E-1(5.45E-2) [†]	5.151E-3(5.45E-4)	3.0002(9.48E-4) [†]	3.0664(4.72E-2) [†]	3.3244(9.46E-4)
MOP3	1.591E-1(2.28E-2) [†]	1.295E-1(5.12E-2) [†]	5.882E-3(1.85E-3)	3.0674(2.12E-2) [†]	3.0784(5.11E-2) [†]	3.2055(2.85E-3)
MOP4	3.318E-1(1.27E-2) [†]	2.885E-1(3.31E-2) [†]	1.642E-2(3.91E-3)	3.1001(1.01E-2) [†]	3.1557(2.26E-2) [†]	3.4998(3.58E-3)
MOP5	2.877E-1(3.11E-2) [†]	3.165E-1(5.47E-3) [†]	1.511E-2(1.82E-3)	2.9981(1.58E-1) [†]	2.7598(1.14E-1) [†]	3.6427(2.79E-3)
MOP6	3.046E-1(4.72E-5) [†]	2.952E-1(1.88E-2) [†]	4.509E-2(2.51E-3)	7.4933(6.33E-4) [†]	7.5096(2.83E-2) [†]	7.7740(3.21E-3)
MOP7	3.517E-1(5.29E-4) [†]	3.477E-1(1.85E-2) [†]	7.276E-2(3.45E-3)	7.2081(1.46E-3) [†]	7.2150(2.13E-2) [†]	7.3696(4.55E-3)

Wilcoxon's rank sum test at a 0.05 significance level is performed between ED/DPP-DRA and each of ε -MOEA and MOEA/D-DRA. [†] and [‡] denote the performance of the corresponding algorithm is significantly worse than and better than that of the proposed ED/DPP-DRA, respectively. And the best mean metric value is highlighted in boldface with gray background.

Table 6

Performance comparisons between ED/DPP-DRA and Dual-NSGA-II, Dual-MOEA/D on IGD and HV metrics.

Test instances	IGD			HV		
	Dual-NSGA-II	Dual-MOEA/D	ED/DPP-DRA	Dual-NSGA-II	Dual-MOEA/D	ED/DPP-DRA
UF1	1.445E−1(1.74E−2) [†]	1.150E−2(1.05E−2) [†]	8.798E−4(5.75E−5)	3.2992(3.49E−2) [†]	3.6459(1.62E−2) [†]	3.6643(6.89E−4)
UF2	1.518E−1(8.06E−3) [†]	9.659E−3(1.96E−3) [†]	1.313E−3(3.67E−4)	3.2444(2.84E−2) [†]	3.6506(2.73E−3) [†]	3.6566(9.77E−4)
UF3	3.429E−1(2.66E−3) [†]	1.472E−2(5.98E−3) [†]	1.378E−3(4.23E−4)	3.0855(6.41E−3) [†]	3.6364(1.15E−2) [†]	3.6639(6.93E−4)
UF4	7.092E−2(9.75E−3) [†]	5.112E−2(1.40E−3) [†]	4.738E−2(8.75E−4)	3.1340(2.71E−2) [†]	3.1754(6.12E−3) [†]	3.1907(3.15E−3)
UF5	2.206E+0(8.74E−2) [†]	5.519E−1(6.05E−2) [†]	1.154E−1(6.72E−2)	0 [†]	2.1715(1.58E−1) [†]	3.2509(1.47E−1)
UF6	1.057E+0(5.68E−2) [†]	1.070E−1(6.95E−2) [†]	3.027E−2(1.19E−2)	0.9097(1.01E−1) [†]	3.1463(1.67E−1) [†]	3.3165(4.94E−2)
UF7	1.950E−1(1.17E−2) [†]	6.799E−3(1.35E−3) [†]	9.809E−4(5.35E−5)	3.0433(3.05E−2) [†]	3.4867(2.33E−3) [†]	3.4950(1.38E−3)
UF8	6.759E−1(6.65E−2) [†]	8.055E−2(1.04E−2) [†]	2.012E−2(1.77E−3)	3.2840(2.27E−1) [†]	7.2576(6.48E−2) [†]	7.4220(5.56E−3)
UF9	7.703E−1(5.91E−2) [†]	5.125E−2(2.36E−3) [†]	2.042E−2(1.68E−3)	3.4471(2.83E−1) [†]	7.6354(1.40E−2) [†]	7.7509(6.85E−3)
UF10	5.513E+0(4.61E−1) [†]	2.071E+0(1.06E−1) [†]	4.276E−1(2.59E−2)	0 [†]	0.0300(3.10E−2) [†]	4.1699(2.14E−1)
MOP1	3.733E−1(2.80E−3) [†]	3.260E−1(5.96E−2) [†]	1.528E−2(1.54E−3)	3.0513(6.64E−3) [†]	3.1376(9.19E−2) [†]	3.6429(2.14E−3)
MOP2	2.975E−1(4.36E−3) [†]	2.482E−1(3.10E−2) [†]	5.151E−3(5.45E−4)	3.0057(6.52E−3) [†]	3.0383(1.11E−2) [†]	3.3244(9.46E−4)
MOP3	1.322E−1(3.95E−3) [†]	3.045E−2(1.23E−2) [†]	5.882E−3(1.85E−3)	3.0071(1.10E−2) [†]	3.1725(1.45E−2) [†]	3.2055(2.85E−3)
MOP4	2.902E−1(7.12E−2) [†]	2.669E−1(1.95E−2) [†]	1.642E−2(3.91E−3)	3.1041(1.13E−2) [†]	3.1586(7.18E−3) [†]	3.4998(3.58E−3)
MOP5	2.762E−1(1.28E−2) [†]	9.868E−2(2.22E−1) [†]	1.511E−2(1.82E−3)	2.8801(1.36E−1) [†]	3.4037(3.88E−1) [†]	3.6427(2.79E−3)
MOP6	2.709E−1(2.65E−2) [†]	2.962E−1(1.92E−2) [†]	4.509E−2(2.51E−3)	7.5206(4.01E−2) [†]	7.4988(2.46E−2) [†]	7.7740(3.21E−3)
MOP7	3.246E−1(1.81E−2) [†]	3.516E−1(7.12E−4) [†]	7.276E−2(3.45E−3)	7.1962(1.59E−2) [†]	7.2055(3.51E−3) [†]	7.3696(4.55E−3)

Wilcoxon's rank sum test at a 0.05 significance level is performed between ED/DPP-DRA and each of Dual-NSGA-II and Dual-MOEA/D. [†] and [‡] denote the performance of the corresponding algorithm is significantly worse than and better than that of the proposed ED/DPP-DRA, respectively. And the best mean metric value is highlighted in boldface with gray background.

shows significantly better performance than both Dual-NSGA-II and Dual-MOEA/D in all comparisons. In addition, it is worth noting that Dual-NSGA-II and Dual-MOEA/D even perform worse than their corresponding baseline algorithms in most cases. From this experiment, we conclude that DPP does not merely benefit from using two populations. Instead, as discussed in Section 3.1, the archiving mechanisms of Pareto- and decomposition-based archives have complementary effects, which provide more useful information to the search process.

5.3.2. DPP with only one population

In fact, there is not much difference between DPP and its baseline algorithm, e.g., NSGA-II, in case it only uses one population. However, besides the use of dual populations, one other important characteristic of DPP is the use of RMS mechanism for selecting mating parents. In case only one population is maintained, the original RMS mechanism proposed in Section 3.2 cannot be directly applied here. To mimic the effect of RMS mechanism, we make a simple modification on the mating selection of the DE operator. Specifically, as for the two mating parents \mathbf{x}^{r_1} and \mathbf{x}^{r_2} in Eq. (4), they exchange the positions with each other in case $\mathbf{x}^{r_2} \preceq \mathbf{x}^{r_1}$. To answer the second question, we have designed two variants, denoted as NSGA-II-RMS and MOEA/D-RMS, by using NSGA-II and MOEA/D-DRA as the baseline algorithms. The performance comparisons of ED/DPP-DRA, NSGA-II-RMS and MOEA/D-RMS are presented in Table 7. It is clear that ED/DPP-DRA performs better than the two variants, as it wins in 66 out of 68 comparisons. Similar to the RMS mechanism, the modified mating selection mechanism tries to exploit some guidance information towards the convergence direction. From the empirical results, we observe that the performance of two baseline algorithms, NSGA-II and MOEA/D-DRA, have been improved to a certain extent by using the modified mating selection mechanism. However, it is also worth noting that NSGA-II-RMS performs significantly worse than the original NSGA-II on UF5 and UF10 instances. UF5 is featured by 21 discrete points in the objective space and the characteristic of UF10 is many local optima which might obstruct the search towards the global PF. The inferior performance of NSGA-II-RMS on these two instances indicates that the deterministic mating scheme provided by the modified mating selection mechanism might trap the search when tackling this kind of problems.

5.3.3. DPP with indicator-based techniques

As described in Section 1, the indicator-based technique is another avenue to achieve a balance between convergence and diversity. HV indicator, which simultaneously measures the convergence and diversity of a PF approximation, is widely used in designing indicator-based EMO algorithms. Inspired by the method developed in [21], here we use the individual HV contribution as the criterion to select solutions for an indicator-based archive. Fig. 7 gives an illustration of the individual HV contribution. The larger is the individual HV contribution, the better is the quality of a solution. In a word, each time, the indicator-based archive eliminates the solution that has the smallest individual HV contribution.

We have developed other two DPP variants, denoted as ID/DPP and NI/DPP, which, respectively, replaces the Pareto- and decomposition-based archiving mechanism by the above suggested indicator-based archiving mechanism. The performance comparisons of ID/DPP, NI/DPP and ED/DPP-DRA, regarding IGD and HV metrics, are presented in Table 8. From these results, we find that ED/DPP-DRA shows better metric values than the two variants on all comparisons, where all better results are with statistical significance. This observation demonstrates that using the indicator-based technique in DPP is not a good

Table 7

Performance Comparisons between ED/DPP-DRA and NSGA-II-RMS, MOEA/D-RMS on IGD and HV metrics.

Test instances	IGD			HV		
	NSGA-II-RMS	MOEA/D-RMS	ED/DPP-DRA	NSGA-II-RMS	MOEA/D-RMS	ED/DPP-DRA
UF1	3.575E–2(2.85E–3) [†]	9.544E–4(7.98E–5) [†]	8.789E–4(5.75E–5)	3.5969(5.26E–3) [†]	3.6625(1.11E–3) [†]	3.6643(6.89E–4)
UF2	4.446E–2(3.02E–3) [†]	2.841E–3(2.20E–3) [†]	1.313E–3(3.67E–4)	3.5850(6.02E–3) [†]	3.6560(1.05E–2)	3.6566(9.77E–4)
UF3	6.427E–2(1.41E–2) [†]	1.132E–2(1.57E–2) [†]	1.378E–3(4.23E–4)	3.5482(2.19E–2) [†]	3.6294(7.75E–2) [†]	3.6639(6.93E–4)
UF4	4.894E–2(2.29E–3) [†]	5.668E–2(4.21E–3) [†]	4.738E–2(8.75E–4)	3.2030(6.98E–3)[‡]	3.1605(1.51E–2) [†]	3.1907(3.15E–3)
UF5	1.130E+0(1.09E–1) [†]	2.847E–1(4.59E–2) [†]	1.154E–1(6.72E–2)	0.7796(1.89E–1) [†]	2.7219(1.54E–1) [†]	3.2509(1.47E–1)
UF6	3.968E–1(2.24E–2) [†]	1.064E–1(1.00E–1) [†]	3.027E–2(1.19E–2)	2.4020(7.65E–2) [†]	3.0908(2.41E–1) [†]	3.3165(4.94E–2)
UF7	1.538E–2(9.87E–4) [†]	1.056E–3(1.56E–4) [†]	9.809E–4(5.35E–5)	3.4684(2.09E–3) [†]	3.4960(2.01E–3)[‡]	3.4950(1.38E–3)
UF8	1.440E–1(6.28E–3) [†]	3.414E–2(1.08E–2) [†]	2.012E–2(1.77E–3)	7.0111(3.08E–2) [†]	7.3981(2.47E–2) [†]	7.4220(5.56E–3)
UF9	1.283E–1(1.20E–2) [†]	5.068E–2(4.37E–2) [†]	2.042E–2(1.68E–3)	7.4328(3.44E–2) [†]	7.6101(1.89E–1) [†]	7.7509(6.85E–3)
UF10	2.572E+0(1.73E–1) [†]	4.777E–1(6.10E–2) [†]	4.276E–1(2.59E–2)	0 [†]	3.6638(2.99E–1) [†]	4.1699(2.14E–1)
MOP1	3.433E–1(2.05E–2) [†]	3.567E–1(3.39E–2) [†]	1.528E–2(1.54E–3)	3.1169(4.02E–2) [†]	3.0884(6.37E–2) [†]	3.6429(2.14E–3)
MOP2	3.010E–1(6.35E–2) [†]	2.790E–1(7.41E–2) [†]	5.151E–3(5.45E–4)	3.0236(3.26E–2) [†]	3.0415(5.04E–2) [†]	3.3244(9.46E–4)
MOP3	5.865E–2(6.33E–2) [†]	1.194E–1(4.13E–2) [†]	5.882E–3(1.85E–3)	3.1483(7.40E–2) [†]	3.0835(4.42E–2) [†]	3.2055(2.85E–3)
MOP4	3.055E–1(1.99E–2) [†]	2.948E–1(2.91E–2) [†]	1.642E–2(3.91E–3)	3.1551(1.85E–2) [†]	3.1493(1.53E–2) [†]	3.4998(3.58E–3)
MOP5	2.736E–1(5.21E–2) [†]	3.176E–1(5.39E–3) [†]	1.511E–2(1.82E–3)	2.9717(2.20E–1) [†]	2.7103(1.09E–1) [†]	3.6427(2.79E–3)
MOP6	2.989E–1(1.54E–2) [†]	2.952E–1(1.88E–2) [†]	4.509E–2(2.51E–3)	7.5033(1.67E–2) [†]	7.5097(2.83E–2) [†]	7.7740(3.21E–3)
MOP7	3.509E–1(1.07E–5) [†]	3.511E–1(3.03E–7) [†]	7.276E–2(3.45E–3)	7.2129(5.73E–5) [†]	7.2112(1.03E–6) [†]	7.3696(4.55E–3)

Wilcoxon's rank sum test at a 0.05 significance level is performed between ED/DPP-DRA and each of NSGA-II-RMS and MOEA/D-RMS. [†] and [‡] denote the performance of the corresponding algorithm is significantly worse than and better than that of the proposed ED/DPP-DRA, respectively. And the best mean metric value is highlighted in boldface with gray background.

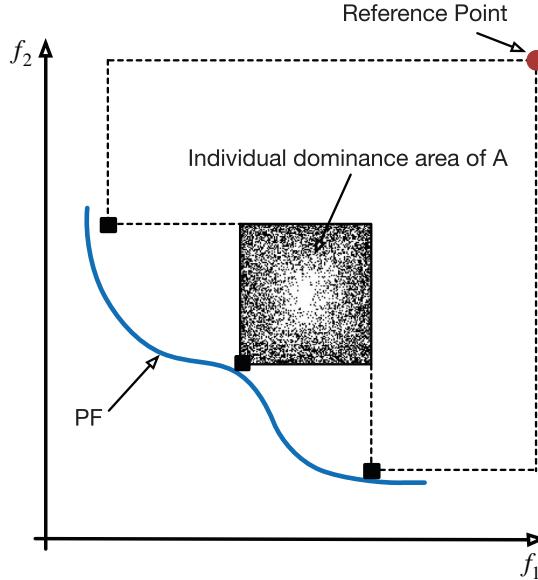


Fig. 7. An illustration of individual HV contribution in 2-D space.

choice. One possible explanation is that no complementary effects exist between the indicator- and either of Pareto- and decomposition-based techniques. Therefore, such DPP design does not provide any useful information to the search process.

In summary, from the empirical studies in this section, we conclude that the encouraging results obtained by our proposed DPP is not merely a simple hybridization of dual populations. Instead, the success of DPP can be attributed to two aspects:

- The archiving mechanisms of two populations have complementary effects: one is good at strengthening convergence towards the optima, the other is good at maintaining population diversity.
- These two populations co-evolve towards the optima by using an appropriate interaction mechanism (e.g., RMS mechanism in this paper).

Table 8

Performance comparisons between ED/DPP-DRA and ID/DPP, NI/DPP on IGD and HV metrics.

Test instances	IGD			HV		
	ED/DPP-DRA	ID/DPP	NI/DPP	ED/DPP-DRA	ID/DPP	NI/DPP
UF1	8.789E-4(5.75E-5)	1.427E-2(1.54E-3) [†]	1.643E-2(6.28E-4) [†]	3.6643(6.89E-4)	3.6290(4.62E-3) [†]	3.6308(2.31E-3) [†]
UF2	1.313E-3(3.67E-4)	1.852E-2(3.75E-3) [†]	3.274E-2(1.43E-3) [†]	3.6566(9.77E-4)	3.6270(7.63E-3) [†]	3.6132(3.98E-3) [†]
UF3	1.378E-3(4.23E-4)	2.931E-2(4.64E-2) [†]	2.529E-2(1.89E-3) [†]	3.6639(6.93E-4)	3.6154(7.83E-2) [†]	3.6205(3.38E-3) [†]
UF4	4.738E-2(8.75E-4)	6.559E-2(9.78E-4) [†]	6.439E-2(1.33E-3) [†]	3.1907(3.15E-3)	3.1649(4.05E-3) [†]	3.1651(5.56E-3) [†]
UF5	1.154E-1(6.72E-2)	3.957E-1(1.28E-1) [†]	2.095E-1(4.68E-2) [†]	3.2509(1.47E-1)	2.3595(2.91E-1) [†]	2.7302(1.54E-1) [†]
UF6	3.027E-2(1.19E-2)	1.157E-1(1.42E-2) [†]	1.067E-1(3.34E-2) [†]	3.3165(4.94E-2)	2.9054(4.95E-2) [†]	3.0002(5.58E-2) [†]
UF7	9.809E-4(5.35E-5)	9.682E-3(1.63E-3) [†]	1.654E-2(8.03E-4) [†]	3.4905(1.38E-3)	3.4527(4.10E-3) [†]	3.4435(2.44E-3) [†]
UF8	2.012E-2(1.77E-3)	5.167E-2(5.26E-3) [†]	1.440E-2(1.31E-2) [†]	7.4220(5.56E-3)	7.0982(2.76E-2) [†]	7.2616(3.91E-2) [†]
UF9	2.042E-2(1.68E-3)	3.238E-2(4.39E-3) [†]	3.404E-2(5.80E-3) [†]	7.7509(6.85E-3)	7.6892(1.78E-2) [†]	7.0194(2.31E-2) [†]
UF10	4.276E-1(2.59E-2)	2.655E+0(1.62E-1) [†]	1.941E+0(1.89E-1) [†]	4.1699(2.14E-1)	2.435E-4(1.33E-3) [†]	0.0441(7.13E-2) [†]
MOP1	1.528E-2(1.54E-3)	3.314E-1(5.87E-3) [†]	3.481E-1(3.40E-3) [†]	3.6429(2.14E-3)	3.1173(8.56E-3) [†]	3.0896(4.60E-3) [†]
MOP2	5.151E-3(5.45E-4)	1.895E-1(1.60E-2) [†]	2.804E-1(1.03E-2) [†]	3.3244(9.46E-4)	3.1104(2.50E-3) [†]	3.0896(1.48E-3) [†]
MOP3	5.882E-3(1.85E-3)	1.374E-1(2.31E-2) [†]	1.194E-1(1.76E-2) [†]	3.2055(2.85E-3)	3.0039(4.52E-4) [†]	3.0764(4.07E-4) [†]
MOP4	1.642E-2(3.91E-3)	1.598E-1(3.95E-2) [†]	2.022E-2(6.42E-2) [†]	3.4998(3.58E-3)	3.1060(4.39E-3) [†]	3.0975(3.84E-3) [†]
MOP5	1.511E-2(1.82E-3)	3.005E-1(1.77E-3) [†]	2.895E-1(1.44E-3) [†]	3.6427(2.79E-3)	2.8181(2.79E-3) [†]	3.0023(2.00E-3) [†]
MOP6	4.509E-2(2.51E-3)	2.875E-1(1.18E-2) [†]	3.005E-2(2.86E-3) [†]	7.7740(3.21E-3)	7.5067(1.17E-2) [†]	7.5386(3.61E-3) [†]
MOP7	7.276E-2(3.45E-3)	3.131E-1(5.16E-3) [†]	3.284E-1(3.97E-3) [†]	7.3696(4.55E-3)	7.2612(9.13E-4) [†]	7.2346(2.90E-3) [†]

Wilcoxon's rank sum test at a 0.05 significance level is performed between ED/DPP-DRA and each of ID/DPP and EI/DPP. [†] and [‡] denote the performance of the corresponding algorithm is significantly worse than and better than that of the proposed ED/DPP-DRA, respectively. And the best mean metric value is highlighted in boldface with gray background.

Table 9

IGD results of ED/DPP-DRA and the other four variants.

Test instances	Variant-I	Variant-II	Variant-III	Variant-IV	ED/DPP-DRA
UF1	9.703E-4(7.84E-5) [†]	7.089E-3(8.83E-3) [†]	1.527E-2(1.54E-3) [†]	1.043E-3(9.58E-4) [†]	8.789E-4(5.75E-5)
UF2	1.555E-3(5.72E-4) [†]	5.583E-2(4.18E-2) [†]	1.852E-2(3.75E-3) [†]	1.618E-3(4.59E-4) [†]	1.313E-3(3.67E-4)
UF3	1.180E-3(2.04E-3) [‡]	4.999E-2(4.34E-2) [†]	2.931E-2(4.64E-2) [†]	1.060E-3(1.91E-3)[‡]	1.378E-3(4.23E-4)
UF4	5.059E-2(1.62E-3) [†]	5.270E-2(1.47E-3) [†]	4.559E-2(9.78E-4)[‡]	5.116E-2(2.70E-3) [†]	4.738E-2(8.75E-4)
UF5	2.311E-1(4.29E-2) [†]	2.189E-1(4.10E-2) [†]	6.957E-1(1.28E-1) [†]	2.397E-1(2.79E-3) [†]	1.154E-1(6.72E-2)
UF6	5.397E-2(2.88E-2) [†]	6.508E-2(1.25E-2) [†]	4.157E-1(1.42E-2) [†]	5.222E-2(3.69E-2) [†]	3.027E-2(1.19E-2)
UF7	1.776E-3(2.75E-3) [†]	1.919E-2(1.62E-2) [†]	6.682E-3(1.63E-3) [†]	1.429E-4(2.12E-5) [†]	9.809E-4(5.35E-5)
UF8	2.674E-2(4.91E-3) [†]	2.467E-1(1.34E-1) [†]	1.167E-1(5.26E-3) [†]	1.941E-2(1.47E-3)[‡]	2.012E-2(1.77E-3)
UF9	2.833E-2(1.31E-3) [†]	9.244E-2(3.42E-2) [†]	9.238E-2(4.39E-3) [†]	2.350E-2(2.48E-3) [†]	2.042E-2(1.68E-3)
UF10	5.244E-1(8.79E-2) [†]	1.758E+0(1.01E+0) [†]	2.655E+0(1.62E-1) [†]	5.253E-1(6.90E-2) [†]	4.276E-1(2.59E-2)
MOP1	1.756E-2(1.78E-3) [†]	5.093E-2(2.59E-2) [†]	3.314E-2(5.87E-3) [†]	1.696E-2(1.66E-3) [†]	1.528E-2(1.54E-3)
MOP2	6.060E-3(7.74E-4) [†]	4.034E-2(5.35E-2) [†]	1.335E-2(1.60E-3) [†]	5.791E-3(6.94E-4) [†]	5.151E-3(5.45E-4)
MOP3	1.822E-2(1.89E-3) [†]	2.136E-2(1.91E-2) [†]	6.374E-3(2.31E-4) [†]	4.981E-3(9.62E-5)[‡]	5.882E-3(1.85E-3)
MOP4	1.003E-2(4.48E-3) [‡]	5.211E-2(8.01E-2) [†]	7.298E-3(3.95E-3)[‡]	1.634E-2(2.64E-3) [†]	1.642E-2(3.91E-3)
MOP5	1.733E-2(1.34E-3) [†]	2.340E-2(2.81E-3) [†]	3.005E-2(1.77E-3) [†]	1.645E-2(1.67E-3) [†]	1.511E-2(1.82E-3)
MOP6	6.126E-2(1.42E-3) [†]	2.248E-1(7.77E-2) [†]	1.075E-1(1.18E-2) [†]	5.336E-2(1.58E-3) [†]	4.509E-2(2.51E-3)
MOP7	7.906E-2(4.05E-3) [†]	1.153E-1(1.15E-2) [†]	1.731E-1(5.16E-3) [†]	7.619E-2(2.53E-3) [†]	7.276E-2(3.45E-3)

Wilcoxon's rank sum test at a 0.05 significance level is performed between ED/DPP-DRA and each of the other variants. [†] and [‡] denote the performance of the corresponding algorithm is significantly worse than and better than that of the proposed ED/DPP-DRA, respectively. And the best mean metric value is highlighted in boldface with gray background.

5.4. Effects of the RMS mechanism

RMS mechanism, which plays the key role in the interaction between Pareto- and decomposition-based archives, is the vital step in DPP. By selecting the appropriate mating parents for offspring generation, RMS mechanism exploits useful information from the population to propel the search process. In order to further understand its rationality, this section extends the RMS mechanism into four other mating variants. The principles of these four mating variants are briefly described as follows:

1. **Variant-I:** Instead of carefully selecting the mating parents from each of the Pareto- and decomposition-based archives, this variant selects two mating parents, \mathbf{x}^{r_1} and \mathbf{x}^{r_2} , from the union of Pareto- and decomposition-based archives, in a random manner.
2. **Variant-II:** This variant selects \mathbf{x}^{r_1} from the decomposition-based archive, and it selects \mathbf{x}^{r_2} from the Pareto-based archive.
3. **Variant-III:** This variant has no restriction on the neighborhood structure. In other words, the mating parents are chosen from several randomly selected subregions. This corresponds to the case that $\delta = 0$ in Algorithm 1.

Table 10

HV results of ED/DPP-DRA and the other four variants.

Test instances	Variant-I	Variant-II	Variant-III	Variant-IV	ED/DPP-DRA
UF1	3.6608(2.17E–3) [†]	3.6430(2.45E–2) [‡]	3.6290(4.62E–3) [†]	3.6597(2.20E–3) [†]	3.6643(6.89E–4)
UF2	3.6514(6.79E–3) [†]	3.5022(1.31E–1) [†]	3.6270(7.63E–3) [†]	3.6502(7.00E–4) [†]	3.6566(9.77E–4)
UF3	3.6644(4.65E–3) [‡]	3.5770(7.71E–2) [†]	3.6154(7.83E–2) [†]	3.6649(2.93E–3)[‡]	3.6639(6.93E–4)
UF4	3.1751(5.80E–3) [†]	3.1772(1.09E–2) [†]	3.2049(4.05E–3)[‡]	3.1783(1.70E–2) [†]	3.1907(3.15E–3)
UF5	2.9870(1.19E–1) [†]	2.9897(9.35E–2) [†]	1.7595(2.90E–1) [†]	2.9821(2.78E–2) [†]	3.2509(1.47E–1)
UF6	3.2266(6.24E–2) [†]	3.2211(4.89E–2) [†]	2.4054(4.95E–2) [†]	3.2307(8.08E–2) [†]	3.3165(4.94E–2)
UF7	3.4838(3.67E–2) [†]	3.4499(3.71E–2) [†]	3.4827(4.10E–3) [†]	3.4895(2.50E–3) [†]	3.4950(1.38E–3)
UF8	7.4042(1.43E–2) [†]	5.5178(1.46E+0) [†]	7.0982(2.76E–2) [†]	7.4286(5.49E–3)	7.4220(5.56E–3)
UF9	7.7320(8.87E–3) [†]	7.3620(1.57E–1) [†]	7.5092(1.78E–2) [†]	7.7401(1.37E–2) [†]	7.7509(6.85E–3)
UF10	4.0055(6.05E–1) [†]	1.1748(1.27E+0) [†]	2.435E–4(1.33E–3) [†]	4.0059(6.41E–1) [†]	4.1699(2.14E–1)
MOP1	3.6296(2.33E–3) [†]	3.5937(3.68E–2) [†]	3.6173(8.56E–3) [†]	3.6318(2.22E–3) [†]	3.6429(2.14E–3)
MOP2	3.3226(1.57E–3) [†]	3.2756(6.40E–2) [†]	3.3104(2.50E–3) [†]	3.3242(1.24E–3)	3.3244(9.46E–4)
MOP3	3.1834(3.75E–3) [†]	3.1779(3.35E–2) [†]	3.2039(4.52E–4) [†]	3.2081(2.73E–4)	3.2055(2.85E–3)
MOP4	3.5047(3.79E–3) [‡]	3.4525(1.13E–1) [†]	3.5060(4.39E–3)[‡]	3.4996(2.00E–3)	3.4998(3.58E–3)
MOP5	3.6306(2.23E–3) [†]	3.6301(4.92E–3) [†]	3.6181(2.79E–3) [†]	3.6320(2.37E–3) [†]	3.6427(2.79E–3)
MOP6	7.7548(2.02E–3) [†]	7.5880(8.23E–2) [†]	7.7067(1.17E–2) [†]	7.7612(2.02E–3) [†]	7.7740(3.21E–3)
MOP7	7.3585(2.72E–3) [†]	7.3289(7.41E–3) [†]	7.3212(9.13E–4) [†]	7.3603(2.83E–3) [†]	7.3696(4.55E–3)

Wilcoxon's rank sum test at a 0.05 significance level is performed between ED/DPP-DRA and each of the other variants. [†] and [‡] denote the performance of the corresponding algorithm is significantly worse than and better than that of the proposed ED/DPP-DRA, respectively. And the best mean metric value is highlighted in boldface with gray background.

4. **Variant-IV:** In contrast to *Variant-III*, the mating selection of this variant is entirely restricted to some specific neighboring subregions. This corresponds to the case that $\delta = 1$ in [Algorithm 1](#).

We design four EMO algorithms, each of which, respectively, applies one of the above mating variant to replace the RMS mechanism in ED/DPP-DRA. The performance comparisons, regarding IGD and HV metrics, are presented in [Tables 9 and 10](#). It is clear that the original ED/DPP-DRA is the best candidate, which obtains better metric values, with statistical significance, in 120 out of 136 comparisons. To be specific, the selection mechanism of *Variant-I* is the same as the traditional DE [7], where the mating parents are randomly sampled from the entire population. It is outperformed by ED/DPP-DRA in almost all test instances except UF3 and MOP4. The inferior performance of *Variant-I* validates our conjecture in Section 3.2 that the random mating selection mechanism does not well utilize the optimal information of the current population, and the search process will be stuck when tackling complicated problems. *Variant-II* is the worst among these four variants. This observation validates that it is more reasonable to choose \mathbf{x}^{t_1} from the Pareto-based archive and to choose \mathbf{x}^{t_2} from the decomposition-based archive. *Variant-III* and *Variant-IV* are two extreme cases of the RMS mechanism. In particular, the inferior performance of *Variant-III* can be explained by its purely exploration-oriented scheme, which might slow down the convergence rate of the search. However, it is also worth noting that *Variant-III* obtains the best metric values on UF4 and MOP4. In contrast with *Variant-III*, *Variant-IV* is a purely exploitation-oriented scheme, which restricts the search within some local areas. This variant performs best among these four variants, where it obtains the best metric values on UF3, UF8 and MOP3. However, this purely exploitation-oriented scheme has the risk to trap the population in some particular areas, thus slows down the search process. In summary, our proposed RMS mechanism, which fully utilizes the guidance information towards the optima and well balance the exploitation and exploration, is reasonable and effective.

5.5. Comparisons with other state-of-the-art EMO algorithms

After the investigations of all important aspects of DPP, this section compares the performance of ED/DPP-DRA with the following four EMO algorithms.

1. **MOEA/D-FRRMAB** [20]: This is a recently proposed MOEA/D variant that applies multi-armed bandit model to adaptively select reproduction operators based on their feedbacks from the search process.
2. **MOEA/D-M2M** [23]: This is a recently proposed MOEA/D variant, which decomposes a MOP into a set of simple multiobjective subproblems. Different from the other MOEA/D variants, each subproblem in MOEA/D-M2M has its own population and receives the corresponding computational resource at each generation.
3. **D²MOPSO** [28]: This is a recently proposed multiobjective particle swarm optimizer that combines the Pareto- and decomposition-based techniques in a single paradigm. Specifically, Pareto-based method plays a major role in building the leaders' archive, whereas the decomposition-based technique simplifies the fitness assignment by transforming a MOP into a set of aggregated subproblems.
4. **HypE** [3]: This is a well known indicator-based EMO algorithm, which uses the HV indicator to guide the selection process. In order to reduce the computational complexity of its HV calculation, HypE employs Monte Carlo simulation to approximate the HV value.

Table 11

Performance comparisons between ED/DPP-DRA and the other four EMO algorithms on IGD metric.

IGD	ED/DPP-DRA	MOEA/D-FRRMAB	MOEA/D-M2M	D^2 MOPSO	HypE
UF1	8.789E-4(5.75E-5)	1.067E-3(1.71E-4) [†]	6.827E-3(3.10E-3) [†]	8.642E-2(9.37E-4) [†]	1.005E-1(1.04E-2) [†]
UF2	1.313E-3(3.67E-4)	2.020E-3(5.23E-4) [†]	4.671E-3(3.14E-4) [†]	3.414E-2(5.46E-4) [†]	3.707E-2(8.88E-2) [†]
UF3	1.378E-3(4.23E-4)	4.349E-3(4.99E-3) [†]	1.536E-2(5.61E-3) [†]	2.792E-1(7.14E-2) [†]	2.689E-1(4.45E-2) [†]
UF4	4.738E-2(8.75E-4)	5.463E-2(3.17E-3) [†]	4.401E-2(6.70E-4)[‡]	7.609E-2(5.55E-3) [†]	5.584E-2(5.52E-3) [†]
UF5	1.154E-1(6.72E-2)	2.969E-1(6.72E-2) [†]	2.139E-1(2.98E-2) [†]	3.485E-1(1.21E-1) [†]	2.362E-1(2.83E-2) [†]
UF6	3.027E-2(1.19E-2)	7.812E-2(5.89E-2) [†]	8.034E-2(4.13E-2) [†]	1.077E-1(5.39E-2) [†]	2.621E-1(3.58E-2) [†]
UF7	9.809E-4(5.35E-5)	1.172E-3(1.44E-4) [†]	6.186E-3(1.77E-3) [†]	3.925E-2(8.56E-4) [†]	1.043E-1(1.16E-1) [†]
UF8	2.012E-2(1.77E-3)	3.271E-2(5.26E-3) [†]	8.805E-2(1.59E-2) [†]	6.251E-2(8.64E-2) [†]	4.358E-1(2.90E-2) [†]
UF9	2.042E-2(1.68E-3)	4.320E-2(4.24E-2) [†]	9.244E-2(3.42E-2) [†]	2.942E-2(1.76E-2) [†]	1.458E-1(7.32E-2) [†]
UF10	4.276E-1(2.59E-2)	5.512E-1(7.65E-2) [†]	7.297E-1(7.61E-2) [†]	4.501E-1(1.09E-1) [†]	8.614E-1(1.77E-1) [†]
MOP1	1.528E-2(1.54E-3)	3.346E-1(6.64E-2) [†]	1.750E-2(7.29E-4) [†]	3.642E-1(4.55E-3) [†]	3.634E-1(6.18E-3) [†]
MOP2	5.151E-3(5.45E-4)	2.724E-1(7.02E-2) [†]	1.565E-2(4.15E-3) [†]	3.513E-1(1.31E-2) [†]	3.0506E-1(1.14E-2) [†]
MOP3	5.882E-3(1.85E-3)	1.257E-1(4.85E-2) [†]	1.860E-2(7.79E-3) [†]	1.288E-1(1.35E-2) [†]	1.676E-1(1.96E-2) [†]
MOP4	1.642E-2(3.91E-3)	2.885E-1(3.24E-2) [†]	1.111E-2(2.57E-3)[‡]	2.684E-1(8.27E-2) [†]	3.078E-1(1.79E-2) [†]
MOP5	1.511E-2(1.82E-3)	3.185E-1(3.31E-3) [†]	2.166E-2(1.17E-3) [†]	3.202E-2(9.85E-2) [†]	2.804E-1(2.82E-2) [†]
MOP6	4.509E-2(2.51E-3)	3.012E-1(9.16E-3) [†]	6.681E-2(1.12E-3) [†]	3.044E-1(9.37E-2) [†]	3.044E-1(6.22E-6) [†]
MOP7	7.276E-2(3.45E-3)	3.406E-1(2.71E-2) [†]	8.786E-2(2.08E-3) [†]	3.202E-2(9.85E-2) [†]	3.559E-1(2.15E-3) [†]

Wilcoxon's rank sum test at a 0.05 significance level is performed between ED/DPP-DRA and each of the other EMO algorithms. [†] and [‡] denote the performance of the corresponding algorithm is significantly worse than and better than that of the proposed ED/DPP-DRA, respectively. And the best mean metric value is highlighted in boldface with gray background.

Table 12

Performance comparisons between ED/DPP-DRA and the other four EMO algorithms on HV metric.

IGD	ED/DPP-DRA	MOEA/D-FRRMAB	MOEA/D-M2M	D^2 MOPSO	HypE
UF1	3.6643(6.89E-4)	3.6616(1.36E-3)	3.6472(9.06E-3) [†]	3.5133(7.21E-3) [†]	3.3681(9.09E-2) [†]
UF2	3.6566(9.77E-4)	3.6540(8.92E-3)	3.6461(1.07E-2) [†]	3.5774(1.49E-2) [†]	3.5641(4.74E-2) [†]
UF3	3.6639(6.93E-4)	3.6543(2.96E-2) [†]	3.6406(9.41E-3) [†]	2.6472(2.35E-1) [†]	2.7057(1.52E-1) [†]
UF4	3.1907(3.15E-3)	3.1745(1.57E-2) [†]	3.2126(3.72E-3)[‡]	3.1697(1.27E-2) [†]	3.1714(6.68E-3) [†]
UF5	3.2509(1.47E-1)	2.6724(2.12E-1) [†]	2.9160(1.10E-1) [†]	2.5503(3.76E-1) [†]	2.6436(1.03E-1) [†]
UF6	3.3165(4.94E-2)	3.1624(1.75E-1) [†]	3.1587(1.45E-1) [†]	3.1070(1.34E-1) [†]	2.6436(1.03E-1) [†]
UF7	3.4950(1.38E-3)	3.4902(6.20E-3)	3.4862(3.86E-3) [†]	3.4248(1.27E-3) [†]	3.2110(3.49E-1) [†]
UF8	7.4220(5.56E-3)	7.3622(1.53E-2) [†]	7.1349(1.41E-1) [†]	7.2971(2.99E-2) [†]	6.4437(8.40E-2) [†]
UF9	7.7509(6.85E-3)	7.6272(1.87E-1) [†]	7.3620(1.57E-1) [†]	7.6754(4.89E-2) [†]	7.0039(2.73E-1) [†]
UF10	4.1699(2.14E-1)	3.4797(3.04E-1) [†]	2.7269(4.26E-1) [†]	3.6701(3.22E-1) [†]	2.4083(5.91E-1) [†]
MOP1	3.6429(2.14E-3)	3.1284(1.23E-1) [†]	3.6368(1.24E-3) [†]	3.0758(1.12E-2) [†]	3.0794(1.58E-2) [†]
MOP2	3.3244(9.46E-4)	3.0424(4.32E-2) [†]	3.3066(7.57E-3) [†]	3.0022(9.87E-3) [†]	3.0026(8.88E-3) [†]
MOP3	3.2055(2.85E-3)	3.0777(5.20E-2) [†]	3.1866(1.41E-2) [†]	3.0482(1.54E-2) [†]	3.0635(2.07E-2) [†]
MOP4	3.4998(3.58E-3)	3.1493(2.00E-2) [†]	3.5033(4.30E-3)[‡]	0.3154(9.70E-1) [†]	3.1431(6.04E-3) [†]
MOP5	3.6427(2.79E-3)	2.6972(1.22E-1) [†]	3.6266(3.01E-3) [†]	0.2561(7.88E-1) [†]	3.0390(1.23E-1) [†]
MOP6	7.7740(3.21E-3)	7.5007(1.36E-2) [†]	7.7346(3.21E-3) [†]	7.4980(1.19E-4) [†]	7.4773(2.61E-5) [†]
MOP7	7.3696(4.55E-3)	7.2190(2.64E-2) [†]	7.2771(4.09E-3) [†]	7.2205(5.12E-4) [†]	7.2011(3.12E-3) [†]

Wilcoxon's rank sum test at a 0.05 significance level is performed between ED/DPP-DRA and each of the other EMO algorithms. [†] and [‡] denote the performance of the corresponding algorithm is significantly worse than and better than that of the proposed ED/DPP-DRA, respectively. And the best mean metric value is highlighted in boldface with gray background.

Parameters of these four EMO algorithms are set the same as recommended in their corresponding references. The performance comparisons of ED/DPP-DRA and these four EMO algorithms, regarding IGD and HV metrics, are presented in Tables 11 and 12, respectively. From these results, it is clear that ED/DPP-DRA is the best optimizer, which obtains better metric values in 132 out of 136 comparisons (129 of these better results are with statistical significance).

Figs. 8–11 show the plots of the final solutions obtained by each algorithm, in the run with the median IGD value, on UF1 to UF10. As for UF1 to UF3, which have the same PF shapes in the objective space, ED/DPP-DRA, MOEA/D-FRRMAB and MOEA/D-M2M can find the approximation set that covers the entire PF. However, solutions obtained by ED/DPP-DRA have a smoother distribution. D^2 MOPSO and HypE can only find some parts of the PF. For UF4, whose PF is concave, the solutions obtained by MOEA/D-M2M have a better distribution than those obtained by ED/DPP-DRA and MOEA/D-FRRMAB. HypE finds several disconnected segments to approximate the PF, while D^2 MOPSO only finds some discrete points. The PFs of UF5 and UF6 are several discrete points. From the plots shown in Fig. 8, we find that ED/DPP-DRA can approximate more points on the PF than the other EMO algorithms. For UF7, the performance of ED/DPP-DRA and MOEA/D-FRRMAB are quite similar. Although the solutions obtained by MOEA/D-M2M approximate the entire PF, their distributions are rather rugged. The PF of UF8 is the first octant of a unit sphere. From Fig. 9, we find that only ED/DPP-DRA approximate the entire PF, while

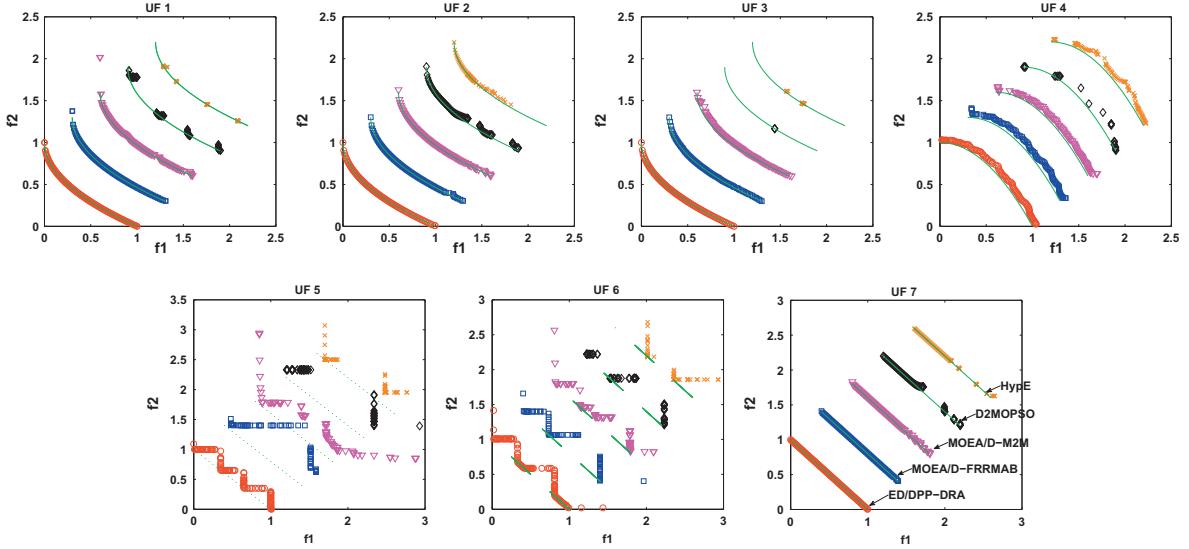


Fig. 8. Plots of final solutions with the median IGD value found by each EMO algorithm on UF1 to UF7.

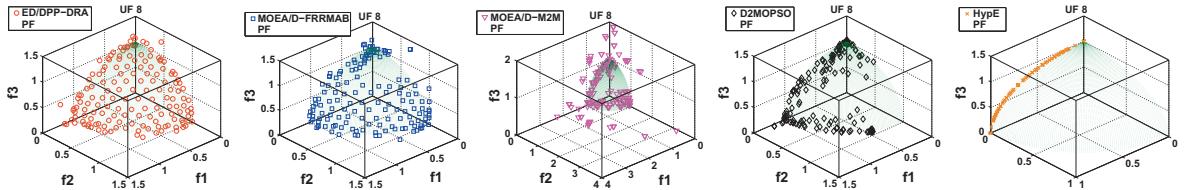


Fig. 9. Plots of final solutions with the median IGD value found by each EMO algorithm on UF8.

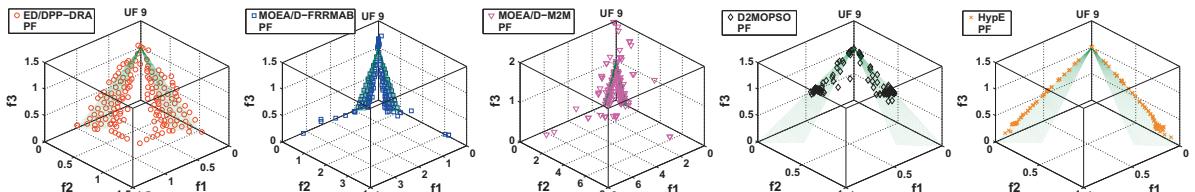


Fig. 10. Plots of final solutions with the median IGD value found by each EMO algorithm on UF9.

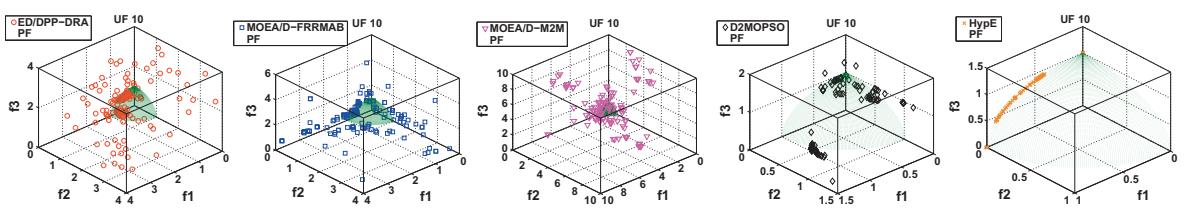


Fig. 11. Plots of final solutions with the median IGD value found by each EMO algorithm on UF10.

the solutions obtained by the other EMO algorithms have clear gaps on the non-dominated fronts. The PF of UF9 is a combination of two disconnected segments. Solutions obtained by ED/DPP-DRA have a better convergence and spread. UF10 has the same PF shape as UF8, but it has many local optima in the search space. Only the solutions obtained by HypE fully converge to the PF, but their spread is terrible.

Figs. 12–14 present the plots of the final solutions obtained by each EMO algorithm, in the run with the median IGD value, on MOP1 to MOP7. From these figures, we find that only ED/DPP-DRA and MOEA/D-M2M can approximate the entire PF,

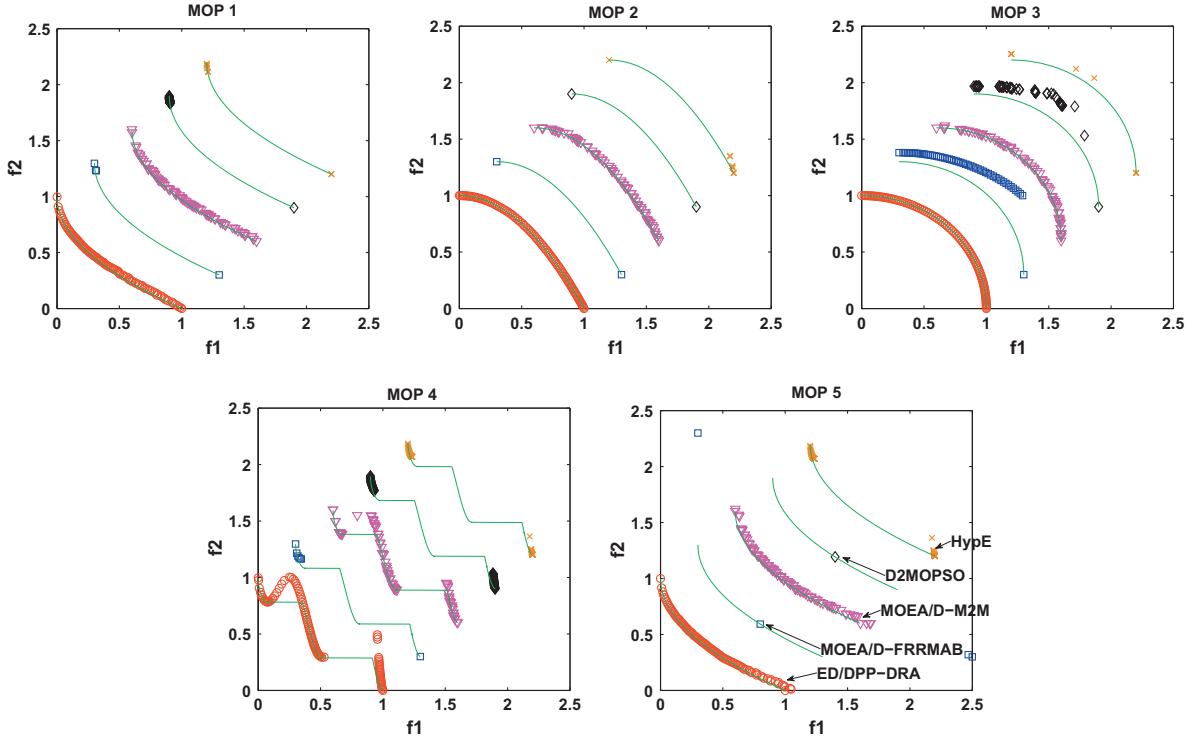


Fig. 12. Plots of final solutions with the median IGD value found by each EMO algorithm on MOP1 to MOP5

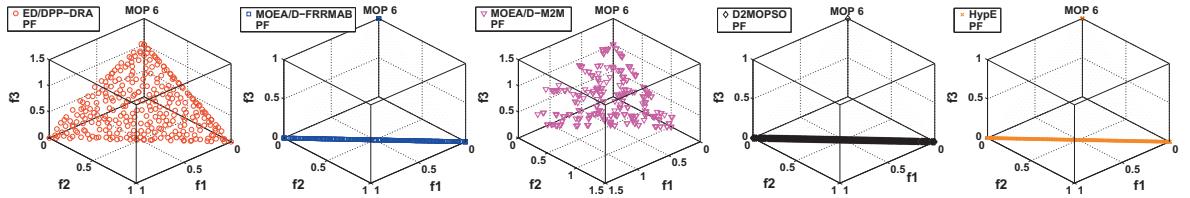


Fig. 13. Plots of final solutions with the median IGD value found by each EMO algorithm on MOP6.

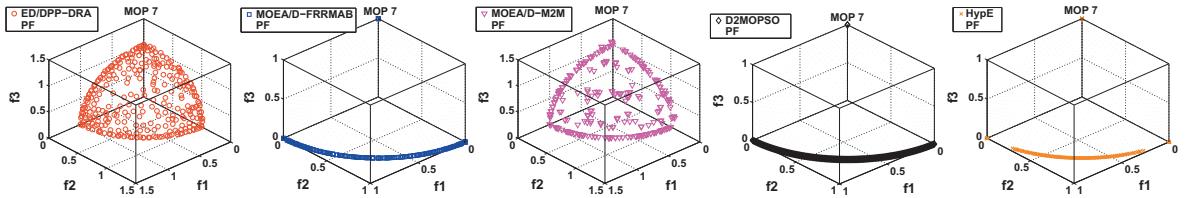


Fig. 14. Plots of final solutions with the median IGD value found by each EMO algorithm on MOP7.

while the other EMO algorithms can only approximate limited regions along the PF. For MOP1 to MOP3 and MOP5, the non-dominated fronts obtained by ED/DPP-DRA are smoother than those found by MOEA/D-M2M. MOP4 is a modification of ZDT3 instance, whose PF contains three disconnected segments. Comparing to MOEA/D-M2M, solutions obtained by ED/DPP-DRA cannot fully converge to the rightmost segment of the PF. It is also worth noting that both ED/DPP-DRA and MOEA/D-M2M find some dominated solutions between the leftmost and middle segments of the PF. For MOP6, a three-objective instance developed from DTLZ1, solutions obtained by ED/DPP-DRA have a better convergence and spread over the PF than the other EMO algorithms. As for MOP7, a modification of DTLZ2, MOEA/D-M2M outperforms ED/DPP-DRA. However, as shown in Fig. 14, there are large gaps on the non-dominated fronts obtained by ED/DPP-DRA and MOEA/D-M2M.

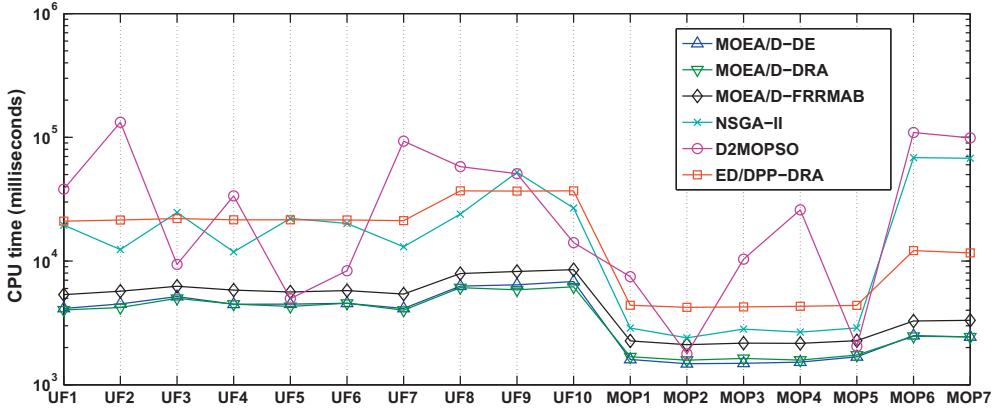


Fig. 15. CPU time comparisons on different test instances.

5.6. CPU time comparison

In this section, we compare the CPU time cost by ED/DPP-DRA and some other EMO algorithms on all seventeen test instances (i.e., UF1 to UF10 and MOP1 to MOP7). Since ED/DPP-DRA is implemented by JAVA, we choose NSGA-II, MOEA/D-DE, MOEA/D-DRA, MOEA/D-FRRMAB and D^2 MOPSO,⁵ which are also implemented in JAVA, for comparative studies. The experimental settings are the same as introduced in Sections 5.1 and 5.5. Fig. 15 shows the results of CPU time comparisons. From this figure, we can find that MOEA/D-DE and MOEA/D-DRA cost the least CPU time, compared to the other algorithms. This is easy to understand as the time complexity of MOEA/D is only $O(NT)$ [36], where T ($T \ll N$) is the neighborhood size. As MOEA/D-FRRMAB employs an adaptive operator selection module for offspring generation, it costs a little more CPU time than its baseline MOEA/D algorithms. Since ED/DPP-DRA maintains two co-evolving populations by different archiving mechanisms, it should cost more CPU time than its baseline MOEA/D algorithm. Moreover, we also notice that the CPU time costs by ED/DPP-DRA, NSGA-II and D^2 MOPSO is similar.

6. Conclusion

This paper presents a dual-population paradigm to achieve a balance between convergence and diversity of the search process. In DPP, one population, termed the Pareto-based archive, focuses on maintaining a repository of solutions that provides a strong selection pressure towards the PF; while the other one, called the decomposition-based archive, concentrates on preserving an archive of well-distributed solutions. As the name suggests, these two populations are updated by Pareto- and decomposition-based techniques, respectively. In addition, a restricted mating selection mechanism, which exploits the guidance information of the search process, is proposed to coordinate the interaction between these two populations. DPP provides an avenue to integrate the Pareto- and decomposition-based techniques in a single paradigm. It is a simple and general framework that any existing Pareto- and decomposition-based techniques can be applied in a plug-in manner. The performance of DPP is investigated on seventeen benchmark problems with various characteristics and complicated PSs. Empirical results demonstrate the effectiveness and competitiveness of the proposed algorithm.

DPP is a simple and general framework that many issues can be further explored in future. At first, although, in this paper, the Pareto- and decomposition-based archives use some existing techniques in a plug-in manner, other problem specific techniques can also be developed and applied in these two populations. Moreover, it is also interesting to investigate different subregion settings (e.g., with user-preference information) for two populations. In addition, the interaction mechanism between two populations is not limited to a mating restriction. For example, solution migration between two populations, which has been widely used in parallel genetic algorithms [24], can be considered. One other major issue in future is to further investigate the proposed algorithm in solving MOPs with many objectives (i.e., more than three objectives) [17], and those problems in dynamic, noisy and uncertain environments [14].

Acknowledgement

This work was supported by the Hong Kong RGC GRF grant 9042038 (CityU 11205314).

⁵ The source code of D^2 MOPSO, implemented in JAVA, is obtained from its authors. The source codes of NSGA-II, MOEA/D-DE, MOEA/D-DRA, MOEA/D-FRRMAB are implemented in jMetal, an integrated JAVA framework, which can be downloaded from <http://jmetal.sourceforge.net>. MOEA/D-M2M, obtained from its authors, is implemented in MATLAB. HypE, downloaded from <http://www.tik.ee.ethz.ch/sop/download/supplementary/hype/>, is implemented in ANSI C.

References

- [1] M. Antonelli, P. Ducange, F. Marcelloni, A fast and efficient multi-objective evolutionary learning scheme for fuzzy rule-based classifiers, *Inform. Sci.* 283 (2014) 36–54.
- [2] A. Auger, J. Bader, D. Brockhoff, E. Zitzler, Hypervolume-based multiobjective optimization: theoretical foundations and practical implications, *Theoret. Comput. Sci.* 425 (2012) 75–103.
- [3] J. Bader, E. Zitzler, HypE: an algorithm for fast hypervolume-based many-objective optimization, *Evol. Comput.* 19 (1) (2011) 45–76.
- [4] N. Beume, B. Naujoks, M. Emmerich, SMS-EMOA: multiobjective selection based on dominated hypervolume, *Eur. J. Oper. Res.* 181 (3) (2007) 1653–1669.
- [5] P. Bosman, D. Thierens, The balance between proximity and diversity in multiobjective evolutionary algorithms, *IEEE Trans. Evol. Comput.* 7 (2) (2003) 174–188.
- [6] I. Das, J.E. Dennis, Normal-boundary intersection: a new method for generating the Pareto surface in nonlinear multicriteria optimization problems, *SIAM J. Optimiz.* 8 (1998) 631–657.
- [7] S. Das, P.N. Suganthan, Differential evolution: a survey of the state-of-the-art, *IEEE Trans. Evol. Comput.* 15 (1) (2011) 4–31.
- [8] K. Deb, M. Goyal, A combined genetic adaptive search (GeneAS) for engineering design, *Comput. Sci. Inform.* 26 (1996) 30–45.
- [9] K. Deb, H. Jain, An evolutionary many-objective optimization algorithm using reference-point based non-dominated sorting approach, Part I: Solving problems with box constraints, *IEEE Trans. Evol. Comput.* 18 (4) (2014) 577–601.
- [10] K. Deb, H. Jain, An Improved NSGA-II Procedure for Many-objective Optimization, Part I: Solving Problems with Box Constraints. Tech. Rep. 2012009, KanGAL, 2012.
- [11] K. Deb, M. Mohan, S. Mishra, Evaluating the ϵ -domination based multi-objective evolutionary algorithm for a quick computation of Pareto-optimal solutions, *Evol. Comput.* 13 (4) (2005) 501–525.
- [12] K. Deb, A. Pratap, S. Agarwal, T. Meyarivan, A fast and elitist multiobjective genetic algorithm: NSGA-II, *IEEE Trans. Evol. Comput.* 6 (2) (2002) 182–197.
- [13] K. Deb, L. Thiele, M. Laumanns, E. Zitzler, Scalable test problems for evolutionary multiobjective optimization, in: A. Abraham, L. Jain, R. Goldberg (Eds.), *Evolutionary Multiobjective Optimization. Advanced Information and Knowledge Processing*, Springer, London, 2005, pp. 105–145.
- [14] J.E. Fieldsend, R.M. Everson, The rolling tide evolutionary algorithm: a multi-objective optimiser for noisy optimization problems, *IEEE Trans. Evol. Comput.* 19 (1) (2015) 103–117.
- [15] I. Giagkiozis, R. Purshouse, P. Fleming, Generalized decomposition and cross entropy methods for many-objective optimization, *Inform. Sci.* 282 (2014) 363–387.
- [16] E.J. Hughes, Evolutionary many-objective optimisation: many once or one many? in: CEC'05: Proc. of the IEEE Congress on Evolutionary Computation, 2005, pp. 222–227.
- [17] H. Ishibuchi, Y. Sakane, N. Tsukamoto, Y. Nojima, Evolutionary many-objective optimization by NSGA-II and MOEA/D with large populations, in: SMC'09: Proc. of the 2009 IEEE International Conference on Systems, Man and Cybernetics, 2009, pp. 1758–1763.
- [18] M. Laumanns, L. Thiele, K. Deb, E. Zitzler, Combining convergence and diversity in evolutionary multiobjective optimization, *Evol. Comput.* 10 (3) (2002) 263–282.
- [19] H. Li, Q. Zhang, Multiobjective optimization problems with complicated Pareto sets, MOEA/D and NSGA-II, *IEEE Trans. Evol. Comput.* 13 (2) (2009) 284–302.
- [20] K. Li, Á. Fialho, S. Kwong, Q. Zhang, Adaptive operator selection with bandits for a multiobjective evolutionary algorithm based on decomposition, *IEEE Trans. Evol. Comput.* 18 (1) (2014) 114–130.
- [21] K. Li, S. Kwong, J. Cao, M. Li, J. Zheng, R. Shen, Achieving balance between proximity and diversity in multi-objective evolutionary algorithm, *Inform. Sci.* 182 (1) (2012) 220–242.
- [22] K. Li, Q. Zhang, S. Kwong, M. Li, R. Wang, Stable matching based selection in evolutionary multiobjective optimization, *IEEE Trans. Evol. Comput.* 18 (6) (2014) 909–923.
- [23] H. Liu, F. Gu, Q. Zhang, Decomposition of a multiobjective optimization problem into a number of simple multiobjective subproblems, *IEEE Trans. Evol. Comput.* 18 (4) (2014) 450–455.
- [24] F. Luna, G.R. Zavala, A.J. Nebro, J.J. Durillo, C.A.C. Coello, Distributed multi-objective metaheuristics for real-world structural optimization problems, *Comput. J.* 58 (3) (2015).
- [25] Y. Mei, K. Tang, X. Yao, Decomposition-based memetic algorithm for multiobjective capacitated arc routing problem, *IEEE Trans. Evol. Comput.* 15 (2) (2011) 151–165.
- [26] K. Miettinen, *Nonlinear Multiobjective Optimization*, vol. 12, Kluwer Academic Publishers, 1999.
- [27] A.A. Montaño, C.A.C. Coello, E. Mezura-Montes, Multiobjective evolutionary algorithms in aeronautical and aerospace engineering, *IEEE Trans. Evol. Comput.* 16 (5) (2012) 662–694.
- [28] N.A. Moubayed, A. Petrovski, J. McCall, D² MOPSO: MOPSO based on decomposition and dominance with archiving using crowding distance in objective and solution spaces, *Evol. Comput.* 22 (1) (2014) 47–77.
- [29] A. Ponsich, A. Jaimes, C. Coello, A survey on multiobjective evolutionary algorithms for the solution of the portfolio optimization problem and other finance and economics applications, *IEEE Trans. Evol. Comput.* 17 (3) (2013) 321–344.
- [30] R.C. Purshouse, P.J. Fleming, On the evolutionary optimization of many conflicting objectives, *IEEE Trans. Evol. Comput.* 11 (6) (2007) 770–784.
- [31] S.N. Qasem, S.M. Shamsuddin, S.Z.M. Hashim, M. Darus, E. Al-Shammary, Memetic multiobjective particle swarm optimization-based radial basis function network for classification problems, *Inform. Sci.* 239 (1) (2013) 165–190.
- [32] H. Rajabalipour Cheshmehgaz, M. Desa, A. Bibowono, A flexible three-level logistic network design considering cost and time criteria with a multi-objective evolutionary algorithm, *J. Intell. Manuf.* (2011) 1–17.
- [33] S. Rostami, D. O'Reilly, A. Shenfield, N. Bowring, A novel preference articulation operator for the evolutionary multi-objective optimisation of classifiers in concealed weapons detection, *Inform. Sci.* 295 (2014) 494–520.
- [34] K. Sindhy, K. Miettinen, K. Deb, A hybrid framework for evolutionary multi-objective optimization, *IEEE Trans. Evol. Comput.* 17 (4) (2013) 495–511.
- [35] G.G. Yen, Z. He, Performance metric ensemble for multiobjective evolutionary algorithms, *IEEE Trans. Evol. Comput.* 18 (1) (2014) 131–144.
- [36] Q. Zhang, H. Li, MOEA/D: a multiobjective evolutionary algorithm based on decomposition, *IEEE Trans. Evol. Comput.* 11 (6) (2007) 712–731.
- [37] Q. Zhang, W. Liu, H. Li, The performance of a new version of MOEA/D on CEC09 unconstrained MOP test instances, in: CEC'09: Proc. of the 2009 IEEE Congress on Evolutionary Computation, 2009, pp. 203–208.
- [38] Q. Zhang, A. Zhou, Y. Jin, RM-MEDA: a regularity model-based multiobjective estimation of distribution algorithm, *IEEE Trans. Evol. Comput.* 12 (1) (2008) 41–63.
- [39] Q. Zhang, A. Zhou, S. Zhao, P.N. Suganthan, W. Liu, S. Tiwari, Multiobjective Optimization Test Instances for the CEC 2009 Special Session and Competition, University of Essex and Nanyang Technological University, Tech. Rep. CES-487, 2008b.
- [40] A. Zhou, B. Qu, H. Li, S. Zhao, P.N. Suganthan, Q. Zhang, Multiobjective evolutionary algorithms: a survey of the state of the art, *Swarm Evol. Comput.* 1 (1) (2011) 32–49.
- [41] E. Zitzler, S. Künnli, Indicator-based selection in multiobjective search, in: PPSN VIII: Proc. of the 8th International Conference on Parallel Problem Solving from Nature, Springer, 2004, pp. 832–842.
- [42] E. Zitzler, M. Laumanns, L. Thiele, SPEA2: improving the strength Pareto evolutionary algorithm for multiobjective optimization, in: K. Giannakoglou, et al. (Eds.), *Evolutionary Methods for Design, Optimisation and Control with Application to Industrial Problems (EUROGEN 2001)*, International Center for Numerical Methods in Engineering (CIMNE), 2002, pp. 95–100.

- [43] E. Zitzler, L. Thiele, Multiobjective evolutionary algorithms: a comparative case study and the strength Pareto approach, *IEEE Trans. Evol. Comput.* 3 (4) (1999) 257–271.
- [44] E. Zitzler, L. Thiele, M. Laumanns, C.M. Fonseca, V.G. da Fonseca, Performance assessment of multiobjective optimizers: an analysis and review, *IEEE Trans. Evol. Comput.* 7 (2) (2003) 117–132.