Inter-Relationship Based Selection for Decomposition Multiobjective Optimization

Ke Li[†], **Sam Kwong**[‡], **Qingfu Zhang**[‡], and **Kalyanmoy Deb**[†]

[†]Department of Electrical and Computer Engineering
Michigan State University, East Lansing, MI 48824, USA
http://www.egr.msu.edu/~kdeb/
[‡]Department of Computer Science
City University of Hong Kong, Kowloon, Hong Kong SAR

COIN Report Number 2014021

Abstract

Multiobjective evolutionary algorithm based on decomposition (MOEA/D), which bridges the traditional optimization techniques and population-based methods, has become an increasingly popular framework for evolutionary multiobjective optimization. It decomposes a multiobjective optimization problem (MOP) into a number of optimization subproblems. Each subproblem is handled by an agent in a collaborative manner. The selection of MOEA/D is a process of choosing solutions by agents. In particular, each agent has two requirements on its selected solution: one is the convergence towards the Pareto front, the other is the distinction with the other agents' choices. This paper suggests addressing these two requirements by defining mutual-preferences between subproblems and solutions. Afterwards, a simple yet effective method is proposed to build an inter-relationship between subproblems and solutions, based on their mutual-preferences. At each generation, this inter-relationship is used as a guideline to select the elite solutions to survive as the next parents. By considering the mutual-preferences between subproblems and solutions (i.e., the two requirements of each agent), the selection operator is able to balance the convergence and diversity of the search process. Comprehensive experiments are conducted on several MOP test instances with complicated Pareto sets. Empirical results demonstrate the effectiveness and competitiveness of our proposed algorithm.

1 Introduction

Multiobjective optimization problems (MOPs), which naturally arise in many disciplines, such as optimal design [1], economics [2] and electric power systems [3], involve more than one objective function to tackle simultaneously. Since evolutionary algorithm (EA) is able to approximate multiple non-dominated solutions, which portray the trade-offs among conflicting objectives, in a single run, it has been recognized as a major approach for multiobjective optimization [4]. Over the last two decades, much effort has been devoted to developing multiobjective evolutionary algorithms (MOEAs) (e.g., [5–9]).

There are two basic requirements in evolutionary multiobjective optimization:

- *Convergence*: the distance of solutions towards the efficient front (EF) should be as small as possible.
- *Diversity*: the spread of solutions along the EF should be as uniform as possible.

Most MOEAs use selection operators to address these two requirements. Depending on different selection mechanisms, the existing MOEAs can be classified into three categories.

- *Pareto-based method*: it uses Pareto dominance relation as the primary selection criterion to promote the convergence and the diversity is maintained by density metrics, such as crowding distance in NSGA-II [6] and clustering analysis in SPEA2 [10].

- *Indicator-based method*: it uses a performance indicator to guide the selection process (e.g., [11–13]). The most commonly used performance indicator is hypervolume [14], which can measure convergence and diversity simultaneously.

- *Decomposition-based method*: it decomposes a MOP into a number of single objective optimization subproblems by linear or nonlinear aggregation methods (e.g., [5] and [15]), or several *simple* multiobjective subproblems [16], and optimizes them in a collaborative manner. It is worth noting that a neighborhood concept among solutions was first proposed in [15] and then used in [5] as a key algorithmic component.

The selection of Pareto-based methods is a convergence first and diversity second strategy. As discussed in [16], such mechanism has difficulties in tackling some problems with particular requirements on diversity. Hypervolume indicator is a theoretically sound metric that measures both convergence and diversity of a solution set [17]. However, as a selection criterion, the balance between convergence and diversity of each step is not adjustable. Since the update of population only depends on the aggregation function of a subproblem, the decomposition-based method also can be regarded as a convergence first and diversity second strategy. However, due to the specification of different search directions a priori, the decomposition-based method provides a more flexible manner for balancing convergence and diversity.

MOEA/D [5], a decomposition-based method, has become an increasingly popular choice for posteriori multiobjective optimization [18]. It has a number of advantages [19], such as its scalability to problems with more than three objectives [20], applicability for combinatorial optimization problems [21], high compatibility with local search [22], and capability for tackling problems with complicated Pareto sets [23]. However, as discussed in some recent studies (e.g., [18,24] and [25]), evenly distributed weight vectors, used in the original MOEA/D, might not always lead to evenly distributed Pareto-optimal solutions (e.g., when the EF is disconnected, some weight vectors might not have solutions).

In MOEA/D, N subproblems are respectively handled by N collaborative agents. Each agent has two requirements on its selected solutions: one is the convergence towards the EF, the other is the distinction with respect to the other solutions in population. The selection of MOEA/D is a process of choosing solutions by agents. In each step, the achievement of these two requirements for each agent is therefore a balance between convergence and diversity of the search process. However, most, if not all MOEA/D implementations only explicitly consider the convergence requirement in selection, while the diversity issue is implicitly controlled by the wide distribution of weight vectors. Differently, our very recent work in [26] provides a first attempt to explicitly address those two requirements by considering the mutual-preferences between subproblems and solutions. Based on the preference articulations, a stable matching model [27] is suggested to guide the selection process. The encouraging results observed in [26] inspires our further explorations along this direction. This paper develops a simple yet effective method to establish an inter-relationship between subproblems and solutions. By simultaneously exploiting the mutual-preferences between subproblems and solutions, this inter-relationship can be used as a guideline to select the elite solutions to survive as the parents for the next generation. As a result, our proposed selection operator is able to allocate an appropriate solution to each agent (thus each subproblem), and balance the convergence and diversity of the search process.

In the remainder of this paper, we first provide some background knowledge in Section 2. Then, we provide a revisit of the selection process in MOEA/D in Section 3. Afterwards, the technical details of our proposed selection operator, based on the inter-relationship between subproblems and solutions, are described in Section 4. Next, its incorporation into MOEA/D is described in Section 5. The general experimental settings are described in Section 6, and the empirical results are presented and analyzed in Section 7. Finally, Section 8 concludes this paper and provides some future directions.

2 Preliminaries and Background

In this section, we first provide some basic definitions of multiobjective optimization. Afterwards, we introduce a classical decomposition method used in this paper.

2.1 Basic Definitions

A MOP can be stated as follows:

$$\begin{array}{ll} \text{minimize} & \mathbf{F}(\mathbf{x}) = (f_1(\mathbf{x}), \cdots, f_m(\mathbf{x}))^T \\ \text{subject to} & \mathbf{x} \in \Omega \end{array} \tag{1}$$

where $\Omega = \prod_{i=1}^{n}[a_i, b_i] \subseteq \mathbb{R}^n$ is the *decision* (*variable*) *space*, and a solution $\mathbf{x} = (x_1, \ldots, x_n)^T \in \Omega$ is a vector of decision variables. $\mathbf{F} : \Omega \to \mathbb{R}^m$ constitutes m real-valued objective functions, and $\mathbb{R}^m$ is called the *objective space*. The *attainable objective set* is defined as the set $\Theta = \{\mathbf{F}(\mathbf{x})|\mathbf{x} \in \Omega\}$. Due to the conflicting nature of MOP, only partial ordering can be specified among solutions. In other words, for two solutions $\mathbf{x}^1$, $\mathbf{x}^2 \in \Omega$, it can so happen that $\mathbf{F}(\mathbf{x}^1)$ and $\mathbf{F}(\mathbf{x}^2)$ are incomparable. Some definitions related to MOP are given as follows in the context of minimization problems.

Definition 1. *A solution $\mathbf{x}^1$ is said to Pareto dominate a solution $\mathbf{x}^2$, denoted as $\mathbf{x}^1 \preceq \mathbf{x}^2$, if and only if $f_i(\mathbf{x}^1) \leq f_i(\mathbf{x}^2)$ for every $i \in \{1, \cdots, m\}$ and $f_j(\mathbf{x}^1) < f_j(\mathbf{x}^2)$ for at least one index $j \in \{1, \cdots, m\}$.*

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

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

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

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

2.2 Decomposition Methods

In the literature of classical multiobjective optimization [28], there are several approaches for constructing aggregation functions to decompose the MOP, in question, into a single-objective optimization subproblem. Among them, the most popular ones are weighted sum, Tchebycheff (TCH) and boundary intersection approaches [29]. In this paper, without loss of generality, we only consider the TCH approach, which is mathematically defined as follows:

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

where $\mathbf{w} = (w_1, \cdots, w_m)^T$ is a user specified weight vector, $w_i \geq 0$ for all $i \in \{1, \cdots, m\}$ and $\sum_{i=1}^{m} w_i = 1$. As proved in [28], under some mild conditions, the optimal solution of TCH function with $\mathbf{w}$ is a Pareto-optimal solution of the MOP in question. By altering weight vectors, TCH approach is able to find different Pareto-optimal solutions.

Definition 6. *Let $\overline{\mathbf{w}} = (\frac{1}{w_1}/\sum_{i=1}^{m}\frac{1}{w_i}, \cdots, \frac{1}{w_m}/\sum_{i=1}^{m}\frac{1}{w_i})^T$ be the symmetry of $\mathbf{w}$, a line $\mathbf{L}$ that starts from the origin and passes $\overline{\mathbf{w}}$ is mathematically defined as:*

$$\frac{t_1}{1/w_1} = \cdots = \frac{t_m}{1/w_m} \tag{3}$$

where $w_i \neq 0$, $i \in \{1, \cdots, m\}$.

Based on Definition 6, we have the following theorem.

Theorem 1. *Let the optimal solution of (2), for a piece-wise continuous MOP, with weight vector $\mathbf{w} = \{w_1, \cdots, w_m\}$ is $\mathbf{x}$, then the intersection point of $\mathbf{L}$ with the PF, $\mathbf{T} = (t_1, \cdots, t_m)^T$, satisfies $t_i = f_i(\mathbf{x}) - z_i^*$, $i \in \{1, \cdots, m\}$.*

Proof. Assume that $\exists l \in \{1, \cdots, m\}$, $t_i \neq f_l(\mathbf{x}) - z_l^*$. According to (2), there should exist the following two nonempty sets:

$$\begin{aligned} \Theta &= \{j | w_j \times |f_j(\mathbf{x}) - z_j^*| < \max_{1 \leq i \leq m} \{w_i \times |f_i(\mathbf{x}) - z_i^*|\}\} \\ \Phi &= \{k | w_k \times |f_k(\mathbf{x}) - z_k^*| = \max_{1 \leq i \leq m} \{w_i \times |f_i(\mathbf{x}) - z_i^*|\}\} \end{aligned} \tag{4}$$

$$\Longrightarrow w_j \times |f_j(\mathbf{x}) - z_j^*| < w_k \times |f_k(\mathbf{x}) - z_k^*| \tag{5}$$

where $j \in \Theta$ and $k \in \Phi$.

Since the MOP, in question, is piece-wise continuous, suppose that there is another solution $\mathbf{x}'$, whose corresponding objective vector is $\mathbf{F}(\mathbf{x}') = (f_1(\mathbf{x}'), \cdots, f_m(\mathbf{x}')^T$ and $\mathbf{F}(\mathbf{x}')$ is in the ϵ-neighborhood of $\mathbf{F}(\mathbf{x})$. In addition, we assume that $f_j(\mathbf{x}') > f_j(\mathbf{x})$, $j \in \Theta$, and $f_k(\mathbf{x}') < f_k(\mathbf{x})$, $k \in \Phi$. According to equation (4), we have $w_j \times |f_j(\mathbf{x}) - z_j^*| < w_j \times |f_j(\mathbf{x}') - z_j^*| < w_k \times |f_k(\mathbf{x}') - z_k^*| < w_k \times |f_k(\mathbf{x}) - z_k^*|$.

$$\Longrightarrow g^{tch}(\mathbf{x}'|\mathbf{w}, \mathbf{z}^*) = \max_{1 \leq i \leq m} \{w_i \times |f_i(\mathbf{x}') - z_i^*|\} \tag{6}$$

$$= w_k \times |f_k(\mathbf{x}') - z_k^*| \tag{7}$$

$$< w_k \times |f_k(\mathbf{x}) - z_k^*| \tag{8}$$

$$= \max_{1 \leq i \leq m} \{w_i \times |f_i(\mathbf{x}) - z_i^*| \tag{9}$$

$$= g^{tch}(\mathbf{x}|\mathbf{w}, \mathbf{z}^*)$$

where $k \in \Phi$. This obviously violates the condition that $\mathbf{x}$ is the optimal solution of the TCH function with weight vector $\mathbf{w}$. □

Theorem 1 implies that the search direction of the TCH approach is the reciprocal of the corresponding weight vector as shown in Fig. 1. In order to make the search direction of TCH approach become the direction of weight vector, we make a simple modification on (2) as follows:

$$\begin{aligned} \text{minimize} \quad & g^{tch2}(\mathbf{x}|\mathbf{w}, \mathbf{z}^*) = \max_{1 \leq i \leq m} \{|f_i(\mathbf{x}) - z_i^*|/w_i\} \\ \text{subject to} \quad & \mathbf{x} \in \Omega \end{aligned} \tag{10}$$

where w_i is set to be a very small number, say 10^{-6}, in case $w_i = 0$.

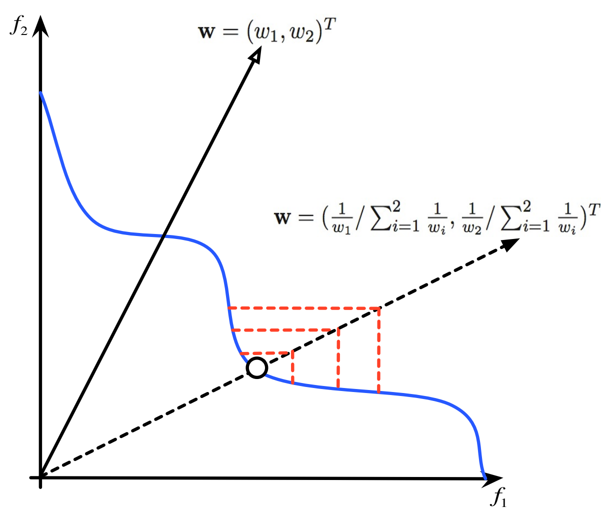

Figure 1: Illustration of the TCH approach

3 Revisits of the Selection Process in MOEA/D

In MOEA/D, N subproblems are respectively handled by N collaborative agents. Selection, in the context of MOEA/D, can thus be regarded as a process of choosing solutions by agents. In particular, each agent has the following two requirements on its selected solution.

- *Convergence*: the selected solution should have a as good as possible aggregation function value for the underlying subproblem.
- *Diversity*: the selected solution should be different from the other agents' choices as much as possible.

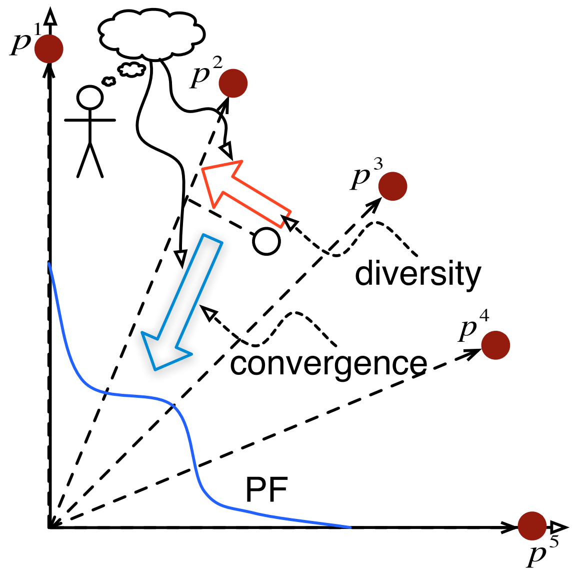

Figure 2: Illustration of two requirements of an agent

Fig. 2 gives an intuitive explanation on these two requirements. As discussed in Section 2.2, the optimal solution of a subproblem is a Pareto-optimal solution of the MOP in question. Therefore, the convergence requirement is easy to understand as the responsibility of an agent is to optimize its handled subproblem. As the optimal solution of a subproblem is usually on the direction line of the corresponding weight vector and weight vectors are designed to be evenly distributed [5], a unique and distinctive solution for each agent (thus for each subproblem) implies a promising population diversity and a well distribution along the PF. The more these two requirements are satisfied, the better is the quality of a solution. Depending on different specifications of these two requirements, the selection mechanisms of existing MOEA/D implementations can be understood from the following three ways.

- Most MOEA/D implementations update the population based on the aggregation function value of a solution. In this case, only the convergence requirement has been explicitly considered by an agent in selection, while the diversity issue is implicitly controlled by the wide distribution of weight vectors. As more than one solution might have similar aggregation function values for the same subproblem, this update mechanism might result in the loss of population diversity.
- In [5], the penalty-based boundary intersection (PBI) approach presents an avenue to aggregate these two requirements into a single criterion. In particular, the PBI approach is formulated as:

$$\begin{array}{ll} \text{minimize} & g^{pbi}(\mathbf{x}|\mathbf{w},\mathbf{z}^*) = d_1 + \theta d_2 \\ \text{subject to} & \mathbf{x} \in \Omega \end{array} \tag{11}$$

  where d_1 is the distance between $\mathbf{z}^*$ and the projection of $\mathbf{x}$ on the direction line of $\mathbf{w}$, and d_2 is the perpendicular distance between $\mathbf{x}$ and the direction line of $\mathbf{w}$. Intuitively, d_1 can be regarded as a measure of the convergence requirement and d_2 is a diversity measure. Although the PBI approach has explicitly integrated these two requirements, it still has the problem mentioned in the first issue.
- By defining the mutual-preferences between subproblems and solutions, [26] suggests a method to tackle these two requirements separately. In particular, the preference of a subproblem over a solution measures the convergence issue; while, similar to the effect of d_2 in PBI approach, the

preference of a solution over a subproblem measures the diversity issue. The selection process is thereafter coordinated by a stable matching model which finds a suitable matching between subproblems and solutions. Since the stable matching achieves an equilibrium between the mutual-preferences of subproblems and solutions, this selection mechanism strikes a balance between convergence and diversity of the search process.

From the encouraging results reported in [26], we find the effectiveness and advantages of treating agents' two requirements explicitly and separately. This paper presents a further attempt along this direction. In particular, according to the two requirements of each agent, an inter-relationship between subproblems and solutions is built upon the specifications of their mutual-preferences. Afterwards, based on this inter-relationship, a simple yet effective method is proposed to guide the selection process.

4 Selection Operator Based on Inter-Relationship

Given a set of subproblems $P = \{p^1, \cdots, p^N\}$ and a set of solutions $S = \{\mathbf{x}^1, \cdots, \mathbf{x}^M\}$ $(M > N)$, the selection process is to choose the appropriate solution for each subproblem.

4.1 Mutual-Preference Setting

As discussed in Section 3, an agent has two requirements (i.e., convergence and diversity) on its selected solution. From the perspectives of subproblems and solutions, these two requirements can be defined as their mutual-preferences. Different mutual-preference settings can lead to different behaviors of the selection process. Without loss of generality, we provide a simple way in setting mutual-preferences as follows:

1. A subproblem p prefers solution $\mathbf{x}$ that has a better aggregation function value. Therefore, the preference value of $\mathbf{x}$ with regard to p, denoted as $\Delta_P(p, \mathbf{x})$, is evaluated by the aggregation function of p:
$$\Delta_P(p, \mathbf{x}) = g(\mathbf{x}|\mathbf{w}^p, \mathbf{z}^{**}) \tag{12}$$
where $\mathbf{w}^p$ is the weight vector of p, and $g(*|*)$ is the aggregation function of p. Obviously, $\Delta_P(p, \mathbf{x})$ measures the achievement of convergence requirement of the agent, which handles p, on $\mathbf{x}$.

2. A solution $\mathbf{x}$ favors subproblem p on which $\mathbf{x}$ can have a as good as possible aggregation function value. Moreover, in MOEA/D, each weight vector also specifies a subregion in the objective space. Consider the population diversity, the subregion corresponding to p, in the objective space, should be as sparse as possible. In view of these two considerations, the preference value of p with regard to $\mathbf{x}$, denoted as $\Delta_X(\mathbf{x}, p)$, is calculated based on the following form:
$$\Delta_X(\mathbf{x}, p) = d^{\perp}(\mathbf{x}, p) + nc(p) \tag{13}$$
where $d^{\perp}(\mathbf{x}, p)$ is the perpendicular distance between $\mathbf{x}$ and the weight vector of p, it is calculated as follows:
$$d^{\perp}(\mathbf{x}, p) = \overline{\mathbf{F}}(\mathbf{x}) - \frac{\mathbf{w}^T\overline{\mathbf{F}}(\mathbf{x})}{\mathbf{w}^T\mathbf{w}}\mathbf{w} \tag{14}$$
where $\overline{F}(\mathbf{x})$ is the normalized objective vector of $\mathbf{x}$, and its k-th individual objective function is normalized as:
$$\overline{f}_k(\mathbf{x}) = \frac{f_k(\mathbf{x}) - z_k^*}{z_k^{nad} - z_k^*} \tag{15}$$
$\Delta_X(\mathbf{x}, p)$ measures the achievement of diversity requirement of the agent, which handles p, on $\mathbf{x}$. The first item of equation (13) plays the same effect as d_2 in PBI approach, while, inspired by [20], the second item $nc(p)$, the niche count of p, plays as a density estimator. In order to evaluate the niche count of the subregion corresponding to a subproblem, we first initialize it to be zero. Afterwards, based on the perpendicular distances between subproblems and solutions, the subproblem whose weight vector is closest to $\mathbf{x}$ is considered to be associated with $\mathbf{x}$. And the niche count of the subregion corresponding to the subproblem is incremented by one. A simple example to illustrate the evaluation of niche count is presented in Fig. 3, where the dotted line

connects a subproblem with its associated solutions and the dotted circle represents the niche around a subproblem. The niche count of a subproblem is the number of solutions associated with it, e.g., $nc(p^3) = 3$.

The pseudo-code of evaluating the mutual-preferences between subproblems and solutions is given in Algorithm 1.

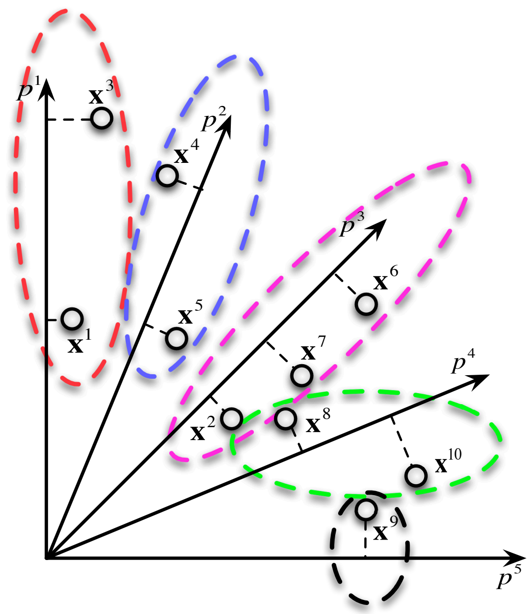

Figure 3: Illustration of the niche count, e.g., $nc(p^3) = 3$.

4.2 Inter-Relationship between Subproblems and Solutions

Based on the above mutual-preference settings, we build an inter-relationship between subproblems and solutions in the following manner.

1. *Selection of the related subproblems for each solution*: Sort each row of Δ_X in ascending order and keep the sorted indices in Ψ_X. For each solution $\mathbf{x}^i \in S$, $i \in \{1, \cdots, M\}$, the first K_d subproblems in the i-th row of Ψ_X are selected as the related subproblems of $\mathbf{x}^i$, where K_d is a user-specified parameter.

2. *Selection of the related solutions for each subproblem*: For each subproblem p^j, $j \in \{1, \cdots, N\}$, let Λ^j denote the set that contains all solutions whose related subproblems include p^j. Then, the set of related solutions of p^j, denoted as χ^j, is formulated as:

   (a) If $\Lambda^j = \emptyset$, χ^j is set to be an empty set, i.e., $\chi^j = \emptyset$ [1].

   (b) If $0 < |\Lambda^j| \leq \vartheta$, where $\vartheta > 0$ is a control parameter, χ^j is set to be Λ^j, i.e., $\chi^j = \Lambda^j$.

   (c) Otherwise, χ^j is set to include the ϑ closest solutions to p^j in Λ^j.

Considering the same example discussed in Fig. 3, Fig. 4 gives a simple illustration on the inter-relationship between subproblems and solutions. Let $K_d = \vartheta = 3$, according to the above rules, the related subproblems of $\mathbf{x}^7$ is $\{p^2, p^3, p^4\}$, and the related solutions of p^5 is $\{\mathbf{x}^8, \mathbf{x}^9, \mathbf{x}^{10}\}$. The pseudo-code of inter-relationship building is given in Algorithm 2.

[1] An example of empty χ is presented in the supplemental file of this paper.

Algorithm 1: COMPTPREF($S, P, \mathbf{z}^*, \mathbf{z}^{nad}$)

Input: solution set S, subproblem set P, the ideal and nadir objective vectors $\mathbf{z}^*$, $\mathbf{z}^{nad}$
Output: preference matrices Δ_X and Δ_P, distance ordering matrix $\Psi_{d^\perp}$

```
for i ← 1 to M do
    F̄(x^i) ← (F(x^i) − z*) / (z^nad − z*);
end
for i ← 1 to M do
    for j ← 1 to N do
        Δ_P(p^j, x^i) ← g(x^i | w^j, z*);
        d⊥(x^i, p^j) ← F̄(x^i) − (w^jT F̄(x^i)) / (w^jT w^j) w^j;
    end
end
Sort each row of d⊥ in ascending order and keep the sorted indices in Ψ_d⊥;
for i ← 1 to N do
    nc[i] ← 0;
end
Normalize d⊥ and nc to range [0, 1] respectively;
for i ← 1 to M do
    nc[Ψ_d⊥(i, 1)] ← nc[Ψ_d⊥(i, 1)] + 1;
end
for i ← 1 to M do
    for j ← 1 to N do
        Δ_X(x^i, p^j) ← d⊥(x^i, p^j) + nc(j);
    end
end
return Δ_X, Δ_P, Ψ_d⊥
```

4.3 Selection Operator

Based on the inter-relationship between subproblems and solutions, an agent selects its preferred solution as follows.

1. For a subproblem p^i, $i \in \{1, \cdots, N\}$, which has a nonempty χ^i, its hosted agent selects the best solution (in terms of the preference values in the i-th row of Δ_P) from χ^i to be included in $\overline{S}$.

2. For a subproblem p^j, $j \in \{1, \cdots, N\}$, which has an empty χ^j, its hosted agent selects the best solution (in terms of the preference values in the j-th row of Δ_P) from the set of unselected solutions to be included in $\overline{S}$. Particularly, this process is conducted in a random order for such subproblems.

It is worth noting that case 2) is not conducted until the completion of case 1). The selected solutions are survived as the parents for the next generation. Algorithm 3 presents the pseudo-code of this selection operator.

4.4 Computational Complexity Analysis

As the preliminary of inter-relationship building, we first evaluate the mutual-preferences between subproblems and solutions. In particular, the evaluations of Δ_p for all subproblems require $O(mMN)$ computations. Moreover, in the worst case, the calculations of distances between subproblems and solutions, and the evaluations of niche counts for all subproblems cost $O(mMN)$ and $O(MN)$ computations, respectively. Therefore, the evaluations of Δ_X for all solutions cost $O(mMN)$ computations. Afterwards, for each solution, the worst case complexity for selecting its related subproblems is $O(NlogN)$. For each subproblem, the selection of its related solutions requires $O(MK_d)$ comparisons. Therefore, the worst case complexity of the inter-relationship building is $\max\{O(MNlogN), O(MNK_d)\}$, whichever is larger. Based on the inter-relationship, the selection of solutions for the next generation requires

Algorithm 2: InterRelation($\Delta_X, \Psi_{d^\perp}$)

Input: preference matrix Δ_X, distance ordering matrix $\Psi_{d^\perp}$
Output: list of related solutions of subproblems χ

Sort each row of Δ_X in ascending order and keep the sorted indices in Ψ_X;
for $i \leftarrow 1$ **to** N **do**
    for $j \leftarrow 1$ **to** M **do**
        for $k \leftarrow 1$ **to** K_d **do**
            if $\Psi_X(j,k) = i$ **then**
                $\chi[i]$.add(j);
            end
        end
    end
end
for $i \leftarrow 1$ **to** N **do**
    if $\chi[i]$.*size*$> \vartheta$ **then**
        Keep the ϑ ones in $\chi[i]$ that have the highest ranks in $\Psi_{d^\perp}$;
    end
end
return χ

$O(N\vartheta)$ comparisons. In summary, the worst case complexity of the selection operator based on inter-relationship is $\max\{O(MNlogN), O(MNK_d)\}$.

4.5 Comparisons With MOEA/D-STM

As mentioned in Section 3, this paper is a further attempt along our previous work in [26]. This section discusses the similarities and differences between MOEA/D-IR and MOEA/D-STM. In particular, these two algorithms have the following two similarities:

1. Both of them are developed upon the high-level framework introduced in Section 3. The selection operators are developed from the perspectives of agents, and the convergence and diversity issues are treated separately and explicitly.

2. Both of them use the same metric to evaluate the preference of a solution with regard to a subproblem.

These two algorithms have the following two differences:

1. The preference of a subproblem p with regard to a solution $\mathbf{x}$ is different. In MOEA/D-STM, only the perpendicular distance between $\mathbf{x}$ and the weight vector of p is considered, while, in this paper, the niche count of p, used to estimate the local density, is considered as an additional term to the perpendicular distance. This modification further improves the population diversity.

2. The selection mechanisms of these two algorithms are essentially different.

   (a) MOEA/D-STM uses a stable matching model to find a suitable matching between subproblems and solutions. The stable matching between subproblems and solutions achieves an equilibrium between their mutual-preferences. Thus, this selection mechanism strikes a balance between convergence and diversity simultaneously.

   (b) In MOEA/D-IR, the selection of an appropriate solution for each subproblem is based on the inter-relationship between subproblems and solutions. In particular, the inter-relationship, built upon the preference values of solutions with regard to subproblems, concerns the diversity issue; while the selection process, depending on the preference values of subproblems to solutions, concerns the convergence issue. In principle, this selection mechanism is a diversity first and convergence second strategy.

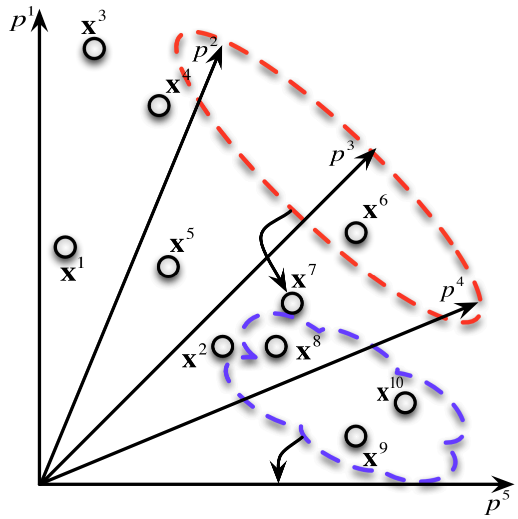

Figure 4: An example of the inter-relationship between subproblems and solutions.

5 Incorporation into MOEA/D

In this section, we present how to incorporate the proposed selection operator, based on the inter-relationship between subproblems and solutions, into the framework of MOEA/D. The pseudo-code of the resulted algorithm, denoted as MOEA/D-IR, is given in Algorithm 4. It is derived from MOEA/D-DRA [30], a MOEA/D variant with dynamic resource allocation scheme. MOEA/D-DRA was the winning algorithm in the CEC2009 MOEA competition [31]. It is worth noting that the difference between MOEA/D-IR and the original MOEA/D only lie in the selection process. Some important components of MOEA/D-IR are further illustrated in the following paragraphs.

5.1 Initialization

In case no prior knowledge about the search landscape at hand, the initial population $S_1 = \{\mathbf{x}^1, \cdots, \mathbf{x}^N\}$ can be randomly sampled from Ω via a uniform distribution. Since the exact ideal objective vector is usually unknown a priori, here we use its approximation, which is set as the minimum F-function value of each objective, i.e., $z_i^* = \min\{f_i(\mathbf{x})|\mathbf{x} \in S_1\}$, for all $i \in \{1, \cdots, m\}$, instead. Analogously, the nadir objective vector is approximately set as $z_i^{nad} = \max\{f_i(\mathbf{x})|\mathbf{x} \in S_1\}$, for all $i \in \{1, \cdots, m\}$.

We initialize a set of weight vectors $W = \{\mathbf{w}^1, \cdots, \mathbf{w}^N\}$ that are evenly spread in the objective space. These weight vectors also define the subproblems and their search directions. Here we set the number of weight vectors be equal to the population size. The chosen of weight vectors can either be predefined in a structured manner or supplied preferentially by the user. In this paper, we use the method proposed in [29] to generate the evenly spread weight vectors on a unit simplex. Each element of a weight vector $\mathbf{w}$ takes a value from $\{\frac{0}{H}, \frac{1}{H} \cdots, \frac{H}{H}\}$, where H is the number of divisions along each coordinate. In total, the number of weight vectors is $N = \binom{H+m-1}{m-1}$. After the generation of W, the Euclidean distance between any two weight vectors is computed. For each weight vector $\mathbf{w}^i$, $i \in \{1, \cdots, N\}$, let $B(i) = \{i_1, \cdots, i_T\}$ be the neighborhood set of $\mathbf{w}^i$, where $\mathbf{w}^{i_1}, \cdots, \mathbf{w}^{i_T}$ are the T $(1 \leq T \leq N)$ closest weight vectors of $\mathbf{w}^i$.

5.2 Reproduction

The reproduction process is to generate the offspring population $Q_t = \{\overline{\mathbf{x}}^1, \cdots, \overline{\mathbf{x}}^N\}$, where t is the generation counter. In general, any genetic operator can serve this purpose. In this paper, we use the differential evolution (DE) operator [32] and polynomial mutation [33] as done in [23]. To be specific,

Algorithm 3: SELECTION(S, Δ_P, χ)

Input: solution set S, preference matrix Δ_P, list of related solutions of subproblems χ
Output: solution set $\overline{S}$

$\overline{S} \leftarrow \emptyset$;
for $i \leftarrow 1$ **to** N **do**
  if $\chi[i].size! = 0$ **then**
    Choose one solution $\mathbf{x}$ in $\chi[i]$ with the best preference value in the i-th row of Δ_P;
    $\overline{S} \leftarrow \overline{S} \bigcup \{\mathbf{x}\}$;
  else
    φ.add(i);
  end
end
for $i \leftarrow 1$ **to** $\varphi.size$ **do**
  Choose one unselected solution $\mathbf{x}$ in S with the best preference value in the φ.get(i)-th row of Δ_P;
  $\overline{S} \leftarrow \overline{S} \bigcup \{\mathbf{x}\}$;
end
return $\overline{S}$

an offspring solution $\overline{\mathbf{x}}^i = \{\overline{x}_1^i, \cdots, \overline{x}_n^i\}$ is generated as follows:

$$u_j^i = \begin{cases} x_j^{r_1} + F \times (x_j^{r_2} - x_j^{r_3}) & \text{if } rand < CR \text{ } or \text{ } j = j_{rand} \\ x_j^i & \text{otherwise} \end{cases} \tag{16}$$

where CR and F are two control parameters of DE operator, $rand$ is a random real number uniformly sampled from $[0, 1]$, j_{rand} is a random integer uniformly chosen from 1 to n, $\mathbf{x}^{r_1}$, $\mathbf{x}^{r_2}$ and $\mathbf{x}^{r_3}$ are three solutions randomly chosen from E. Then, the polynomial mutation acts upon each $\mathbf{u}^i$ to obtain the $\overline{\mathbf{x}}^i$:

$$\overline{x}_j^i = \begin{cases} u_j^i + \sigma_j \times (b_j - a_j) & \text{if } rand < p_m \\ u_j^i & \text{otherwise} \end{cases} \tag{17}$$

with

$$\sigma_j = \begin{cases} (2 \times rand)^{\frac{1}{\eta+1}} - 1 & \text{if } rand < 0.5 \\ 1 - (2 - 2 \times rand)^{\frac{1}{\eta+1}} & \text{otherwise} \end{cases} \tag{18}$$

where $j \in \{1, \cdots, n\}$, the distribution index η and mutation rate p_m are two control parameters. a_j and b_j are the lower and upper bounds of the j-th decision variable.

6 Experimental Settings

This section devotes to the experimental design for the performance investigations of our proposed MOEA/D-IR. At first, we give the descriptions of benchmark problems and performance metrics. Then, we briefly introduce six MOEAs used for comparisons. At last, we illustrate the parameter settings of the empirical studies.

6.1 Test Instances

Twenty six unconstrained MOP test instances are employed here as the benchmark problems for empirical studies. To be specific, UF1 to UF10 are used as the benchmark in CEC2009 MOEA competition [31], and MOP1 to MOP7 are recently proposed in [16]. These test instances have distinct characteristics, and their PSs in the decision space are very complicated. We also consider WFG test suite [34], which has a wide range of problem characteristics, including non-separable, deceptive, degenerate problems, mixed PF shape and variable dependencies, for investigation. 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; for WFG1

Algorithm 4: MOEA/D-IR

Initialize the population $S \leftarrow \{\mathbf{x}^1, \cdots, \mathbf{x}^N\}$, a set of weight vectors $W \leftarrow \{\mathbf{w}^1, \cdots, \mathbf{w}^N\}$, the ideal and nadir objective vectors $\mathbf{z}^*$, $\mathbf{z}^{nad}$;
Set $neval \leftarrow 0$, $iteration \leftarrow 0$;
for $i \leftarrow 1$ **to** N **do**
    $B(i) \leftarrow \{i_1, \cdots, i_T\}$ where $\mathbf{w}^{i_1}, \cdots, \mathbf{w}^{i_T}$ are the T closest weight vectors to $\mathbf{w}^i$;
    $\pi^i \leftarrow 1$;
end
while *Stopping criterion is not satisfied* **do**
    Let all indices of the subproblems whose objectives are MOP individual objectives f_i form the initial I. By using 10-tournament selection based on π^i, select other $\lfloor N/5 \rfloor - m$ indices and add them to I.
    $Q \leftarrow \emptyset$;
    for *each* $i \in I$ **do**
        if $uniform(0, 1) < \delta$ **then**
            $E \leftarrow B(i)$;
        else
            $E \leftarrow S$;
        end
        Randomly select three solutions $\mathbf{x}^{r_1}$, $\mathbf{x}^{r_2}$, and $\mathbf{x}^{r_3}$ from E;
        Generate a candidate $\overline{\mathbf{x}}$ by using the method described in Section 5.2 and $Q \leftarrow Q \bigcup \{\overline{\mathbf{x}}\}$;
        Evaluate the F-function value of $\overline{\mathbf{x}}$;
        Update the current ideal objective vector $\mathbf{z}^*$;
        Update the current nadir objective vector $\mathbf{z}^{nad}$;
        $neval{+}{+}$;
    end
    $R \leftarrow S \cup Q$;
    $[\Delta_X, \Delta_P, \Psi_{d^\perp}] \leftarrow \textsc{ComptPref}(R, W, \mathbf{z}^*, \mathbf{z}^{nad})$;
    $\chi \leftarrow \textsc{InterRelation}(\Delta_X, \Psi_{d^\perp})$;
    $S \leftarrow \textsc{Selection}(R, \Delta_P, \chi)$;
    $iteration{+}{+}$;
    if $mod(iteration, 30) = 0$ **then**
        Update the utility of each subproblem;
    end
end
return S;

to WFG9, the number of objectives is set to 2, the numbers of position- and distance-related decision variables are set to 2 and 4, respectively. Interested readers are recommended to the corresponding references [31], [16] and [34] for more detailed information.

6.2 Performance Metrics

No unary performance metric can give a comprehensive assessment on the performance of an MOEA [35]. In our empirical studies, we consider the following two widely used performance metrics.

1. *Inverted Generational Distance* (IGD) metric [36]: Let P^* be a set of points uniformly sampled along the PF, and S be the set of solutions obtained by an MOEA. The IGD value of S is calculated as:
$$IGD(S, P^*) = \frac{\sum_{\mathbf{x} \in P^*} dist(\mathbf{x}, S)}{|P^*|} \tag{19}$$
where $dist(\mathbf{x}, S)$ is the Euclidean distance between the point $\mathbf{x}$ and its nearest neighbor in S, and $|P^*|$ is the cardinality of P^*. The PF of the underlying MOP is assumed to be known a priori when using the IGD metric. In our empirical studies, 1,000 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 [14]: Let $\mathbf{z}^r = (z_1^r, \ldots, z_m^r)^T$ be a 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 the solutions in S and bounded by $\mathbf{z}^r$.

$$HV(S) = \textsc{Vol}(\bigcup_{\mathbf{x} \in S} [f_1(\mathbf{x}), z_1^r] \times \ldots [f_m(\mathbf{x}), z_m^r]) \tag{20}$$

where $\textsc{Vol}(\cdot)$ indicates the Lebesgue measure. In our empirical studies, $\mathbf{z}^r = (2.0, 2.0)^T$ for bi-objective UF and MOP instances and $\mathbf{z}^r = (2.0, 2.0, 2.0)^T$ for three-objective ones, respectively. For WFG instances, $\mathbf{z}^r = (3.0, 5.0)^T$.

Both IGD and HV metrics can measure the convergence and diversity of S simultaneously. The lower is the IGD value (or the larger is the HV value), the better is the quality of S for approximating the entire 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 conducted to compare the significance of difference between two algorithms.

6.3 Six MOEAs Used for Comparisons

In order to validate our proposed algorithm, six MOEAs are considered here for comparative studies:

1. MOEA/D-DRA [30]: it was the winning algorithm in CEC2009 MOEA competition. Different from the previous MOEA/D variants, in which every subproblem receives the same amount of computational effort, it dynamically allocates the computational resources to different subproblems based on their utilities.

2. MOEA/D-FRRMAB [37]: it is a recently proposed MOEA/D variant that applies a multi-armed bandit model to adaptively select reproduction operators based on their feedbacks from the search process.

3. MOEA/D-M2M [16]: it is a recently proposed MOEA/D variant, which decomposes an 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 effort at each generation.

4. MOEA/D-STM [26]: it is a recently proposed MOEA/D variant, which employs a stable matching model to coordinate the selection process of MOEA/D. It is worth noting that both MOEA/D-STM and MOEA/D-IR are developed upon the high-level framework introduced in Section 3.

5. NSGA-II [6]: it is the most popular Pareto-based MOEA, which is characterized by the fast non-dominated sorting procedure for emphasizing the convergence and the crowding distance for maintaining the diversity. As in [23], we use the reproduction method described in Section 5.2 to generate new offspring solutions.

6. HypE [38]: it is a well known indicator-based MOEA, which uses the HV metric as the guideline of its selection process. In order to reduce the computational complexity in HV calculation, HypE employs Monte Carlo simulation to approximate the HV value.

6.4 General Parameter Settings

The parameters of MOEA/D-DRA, MOEA/D-FRRMAB, MOEA/D-M2M, MOEA/D-STM, NSGA-II and HypE are set according to their corresponding references [16,26,30,37,38] and [6]. All these MOEAs are implemented in JAVA [2], except MOEA/D-M2M in MATLAB and HypE in ANSI C [3]. The detailed parameter settings of our proposed MOEA/D-IR are summarized as follows.

[2]The source codes were developed upon the java MOEA framework jMetal, which can be downloaded from http://jmetal.sourceforge.net.

[3]The source code of HypE is downloaded from http://www.tik.ee.ethz.ch/sop/download/supplementary/hype/

1. *Settings for reproduction operators*: The mutation probability $p_m = 1/n$ and its distribution index $\mu_m = 20$ [4]. For UF and MOP instances, we set $CR = 1.0$ and $F = 0.5$ as recommended in [23], while for WFG instances, we set $CR = 0.5$ and $F = 0.5$.

2. *Population size*: $N = 600$ for UF1 to UF7 instances, $N = 1,000$ for UF8 to UF10 instances, $N = 100$ for MOP1 to MOP5 instances, and $N = 300$ for MOP6 and MOP7 instances, $N = 100$ for WFG1 to WFG9 instances.

3. *Number of runs and termination condition*: Each algorithm is independently launched 20 times on each test instance. The termination condition of an algorithm is the predefined number of function evaluations, which is set to be 300,000 for UF and MOP instances and 25,000 for WFG instances.

4. *Number of related subproblems chosen for a solution*: $K_d = 2$

5. *Number of related solutions chosen for a subproblem*: $\vartheta = 8$.

6. *Neighborhood size*: $T = 20$.

7. *Probability to select in the neighborhood*: $\delta = 0.9$.

7 Empirical Studies

In this section, the performance of our proposed MOEA/D-IR is comprehensively studied according to the experimental design described in Section 6.4. Generally speaking, the empirical studies can be divided into the following three parts.

1. Section 7.1 and Section 7.2, respectively, presents the performance comparisons of MOEA/D-IR and the six other MOEAs on UF, MOP instances [4].

2. Section 7.4 investigates the underlying rationality of our proposed selection operator by comparing with two other variants.

3. Section 7.5 empirically studies the effects of different parameter settings.

7.1 Performance Comparisons on UF instances

Table 1 and Table 2 show the performance comparisons of seven MOEAs on UF instances, regarding the IGD and HV metrics, respectively. Due to the page limit, all the simulation results described by figures (i.e., plots of the final solutions obtained by each MOEA in the run with the median IGD value) are presented in the supplemental file of this paper. Instead of plotting all obtained solutions, we only plot 100 promising ones for UF1 to UF7 and 153 solutions for UF8 to UF10, according to the method in [30]. Experimental results clearly demonstrate that MOEA/D-IR is promising in solving UF instances. It achieves the better mean metric values in 103 out of 120 comparisons. Wilcoxon's rank sum tests indicate that 99 of these better results achieved by MOEA/D-IR are with statistical significance.

To be specific, on UF1, UF2 and UF7, the distributions of solutions obtained by MOEA/D-IR, MOEA/D-DRA, MOEA/D-STM and MOEA/D-FRRMAB are not visually different. Although MOEA/D-M2M is able to approximate the entire EF, the distribution of solutions is not smooth. The other two MOEAs can only find some regions of the EF. For UF3, solutions obtained by MOEA/D-IR do not converge well to the EF. UF4 is a difficult instance with concave EF in the objective space. As shown in Figure 3 of the supplemental file, none of these MOEAs can find well-converged solutions. However, solutions obtained by MOEA/D-M2M are smoother than the other competitors. UF5 and UF6 are with discontinuous EFs in the objective space. All these MOEAs are unable to approximate the entire EF. However, as shown in Figure 3 of the supplemental file, solutions obtained by MOEA/D-IR and MOEA/D-STM are better than the other MOEAs in terms of convergence and diversity. For the three-objective instance UF8, only the solutions obtained by MOEA/D-IR and MOEA/D-STM can

[4]Due to the page limit, the empirical results on WFG instances are presented in the supplemental file of this paper.

Table 1: Performance comparisons of IGD values on UF test instances

Test Instance	MOEA/D-IR	MOEA/D-DRA	MOEA/D-M2M	MOEA/D-STM	MOEA/D-FRRMAB	HypE	NSGA-II
UF1	**9.932E-4(6.19E-5)**	1.516E-3(8.28E-4)†	6.827E-3(3.10E-3)†	1.064E-3(6.86E-5)	1.067E-3(1.71E-4)	1.005E-1(1.04E-2)†	3.478E-2(3.69E-3)†
UF2	3.203E-3(9.95E-4)	5.417E-3(2.98E-3)†	4.671E-3(3.14E-4)†	2.692E-3(1.17E-3)‡	**2.020E-3(5.23E-4)**‡	3.707E-2(8.88E-2)†	4.435E-2(2.54E-3)†
UF3	9.110E-3(4.56E-3)	8.547E-3(1.25E-2)‡	1.536E-2(5.61E-3)†	6.754E-3(7.79E-3)‡	**4.349E-3(4.99E-3)**‡	2.689E-1(4.45E-2)†	6.806E-2(1.26E-2)†
UF4	5.213E-2(3.38E-3)	5.495E-2(4.23E-3)†	**4.401E-2(6.70E-4)**‡	5.299E-2(3.79E-3)†	5.463E-2(3.17E-3)†	6.184E-2(5.52E-3)†	4.965E-2(2.69E-3)‡
UF5	2.625E-1(4.87E-2)	2.911E-1(7.12E-2)†	3.139E-1(2.98E-2)†	**2.471E-1(3.17E-2)**‡	2.969E-1(6.72E-2)†	2.362E-1(2.83E-2)†	1.547E+0(1.10E-1)†
UF6	**6.811E-2(3.48E-2)**	9.601E-2(4.30E-2)†	8.034E-2(4.13E-2)†	7.031E-2(2.72E-2)†	7.812E-2(5.89E-2)†	2.621E-1(3.58E-2)†	3.987E-1(2.03E-2)†
UF7	**1.089E-3(8.15E-5)**	1.172E-3(1.44E-4)†	6.186E-3(1.77E-3)†	1.114E-3(1.11E-4)†	1.123E-3(1.16E-4)	2.043E-1(1.16E-1)†	1.490E-2(9.39E-4)†
UF8	2.639E-2(3.43E-3)	3.577E-2(2.53E-3)†	8.805E-2(1.59E-2)†	**2.250E-2(1.46E-3)**‡	3.271E-2(5.26E-3)†	4.358E-1(2.90E-2)†	1.466E-1(8.73E-3)†
UF9	**2.046E-2(1.04E-3)**	1.052E-1(5.13E-2)†	9.244E-2(3.42E-2)†	2.120E-2(8.45E-4)†	4.320E-2(4.24E-2)†	1.458E-1(7.32E-2)†	9.912E-2(3.81E-2)†
UF10	**4.338E-1(6.55E-2)**	4.555E-1(4.75E-2)†	7.297E-1(7.61E-2)†	8.054E-1(1.76E-1)†	5.512E-1(7.65E-2)†	8.614E-1(1.77E-1)†	2.585E+0(1.84E-1)†

Wilcoxon's rank sum test at a 0.05 significance level is performed between MOEA/D-IR and each of the other MOEAs. † and ‡ denotes that the performance of the corresponding algorithm is significantly worse than or better than that of MOEA/D-IR, respectively. The best mean is highlighted in boldface with gray background.

Table 2: Performance comparisons of HV values on UF test instances

Test Instance	MOEA/D-IR	MOEA/D-DRA	MOEA/D-M2M	MOEA/D-STM	MOEA/D-FRRMAB	HypE	NSGA-II
UF1	**3.6633(1.45E-3)**	3.6531(7.15E-3)†	3.6472(9.06E-3)†	3.6631(4.74E-4)	3.6616(1.36E-3)†	3.3681(9.09E-2)†	3.5993(6.44E-3)†
UF2	3.6524(1.33E-2)	3.6465(1.31E-2)†	3.6461(1.07E-2)†	**3.6575(8.44E-3)**‡	3.6540(8.92E-3)‡	3.5641(4.74E-2)†	3.5852(4.02E-3)†
UF3	3.6473(9.65E-3)	3.6411(5.34E-2)†	3.6406(9.41E-3)†	3.6537(1.31E-2)‡	**3.6543(2.96E-2)**‡	2.7057(1.52E-1)†	3.5434(1.86E-2)†
UF4	3.1834(1.42E-2)	3.1709(1.42E-2)†	**3.2126(3.72E-3)**‡	3.1815(1.40E-2)†	3.1745(1.57E-2)†	3.1714(6.68E-3)†	3.2010(7.28E-3)‡
UF5	2.8180(1.56E-1)	2.6990(1.97E-1)†	2.6860(1.10E-1)†	**2.9426(8.94E-2)**‡	2.6724(2.12E-1)†	2.6436(1.03E-1)†	7.397E-1(1.81E-1)†
UF6	**3.2113(6.92E-2)**	3.1080(1.33E-1)†	3.1587(1.45E-1)†	3.2072(5.36E-2)†	3.1624(1.75E-1)†	2.6436(1.03E-1)†	2.4190(7.46E-2)†
UF7	**3.4973(1.24E-3)**	3.4902(6.20E-3)†	3.4862(3.86E-3)†	3.4968(5.97E-4)†	3.4962(1.03E-3)†	3.2110(3.49E-1)†	3.4688(1.86E-3)†
UF8	7.4170(6.89E-3)	7.3575(4.89E-3)†	7.1349(1.41E-1)†	**7.4241(2.91E-3)**‡	7.3622(1.53E-2)†	6.4437(8.40E-2)†	7.0026(4.30E-2)†
UF9	**7.7527(3.40E-3)**	7.3549(2.38E-1)†	7.3620(1.57E-1)†	7.7441(3.64E-3)†	7.6272(1.87E-1)†	7.0039(2.73E-1)†	7.4400(3.48E-2)†
UF10	**3.7044(4.68E-1)**	3.6674(2.59E-1)†	2.7269(4.26E-1)†	2.5199(6.15E-1)†	3.4797(3.04E-1)†	2.4083(5.91E-1)†	1.240E-3(6.69E-3)†

Wilcoxon's rank sum test at a 0.05 significance level is performed between MOEA/D-IR and each of the other MOEAs. † and ‡ denotes that the performance of the corresponding algorithm is significantly worse than or better than that of MOEA/D-IR, respectively. The best mean is highlighted in boldface with gray background.

approximate the entire EF. There are clear gaps on the non-dominated fronts obtained by the other MOEAs. UF9 is a three-objective instance with two disconnected EFs. It is clear that the solutions obtained by MOEA/D-IR and MOEA/D-STM are better than the other MOEAs in terms of convergence and diversity. UF10 is a multi-modal extension of UF8. All these MOEAs have troubles in converging to the true EF. It is worth noting that MOEA/D-IR and MOEA/D-STM show similar performance on many UF instance. This can be explained as both of them are developed upon the high-level framework introduced in Section 3. Their behaviors should share some similarities.

7.2 Performance Comparisons on MOP instances

MOP instances, modified from the ZDT [39] and DTLZ [40] test suites, are recently proposed benchmark problems. As reported in [16], the state-of-the-art MOEAs, such as MOEA/D and NSGA-II, have significant difficulties on tackling these instances. The major difficulties of these MOP instances are their twisted search landscapes in the decision space, which make the population be easily trapped in some specific regions. In this case, they pose new challenges to MOEAs for balancing convergence and diversity during the search process. In our experiments, all parameters are kept the same as Section 6.4 except $K_d = 4$ and $\vartheta = 30$. Table 3 and Table 4 compare the performances of all seven MOEAs on IGD and HV metrics, respectively. Similar to Section 7.1, we plot the solutions obtained in the run with the median IGD value of each MOEA for each MOP instance in the supplemental file. Comparing to the other algorithms, MOEA/D-IR and MOEA/D-M2M are the most competitive ones, as they show significantly better performance than the other competitors on all MOP instances. According to the Wilcoxon's rank sum test, 75 out of the 78 better results obtained by MOEA/D-IR are with statistical significance.

From the plots in Figure 7, Figure 8 and Figure 9 of the supplemental file, we find that only MOEA/D-IR and MOEA/D-M2M can approximate the entire EF, while the other MOEAs can only approximate a couple of regions along the EF. For MOP1, MOP3 and MOP5, the non-dominated fronts obtained by MOEA/D-IR are smoother than those found by MOEA/D-M2M. MOP2 and MOP3 have the same PF shape, but MOEA/D-M2M outperforms MOEA/D-IR on MOP2 instance. MOP4 is a modification of ZDT3 instance, whose EF contains three disconnected segments. Comparing to MOEA/D-M2M, solutions obtained by MOEA/D-IR cannot fully converge to the rightmost segment of the EF. It is also worth noting that both MOEA/D-IR and MOEA/D-M2M find some dominated solutions between the leftmost and middle segments of the EF. For MOP6, a three-objective instance developed from DTLZ1, solutions obtained by MOEA/D-IR have better convergence and spread over the EF than the other MOEAs. As for MOP7, a modification of DTLZ2, MOEA/D-M2M outperforms MOEA/D-IR. As shown in Figure 9 of the supplemental file, there are large gaps on the non-dominated fronts obtained by MOEA/D-IR and MOEA/D-M2M.

In contrast to the similar performance of MOEA/D-IR and MOEA/D-STM on UF instances, MOEA/D-IR performs significantly better than MOEA/D-STM on MOP instances. As discussed in Section 4.5, MOEA/D-STM tries to balance the convergence and diversity simultaneously, while the selection mechanism of MOEA/D-IR is a diversity first and convergence second strategy. In this case, MOEA/D-IR gives more emphasis on the diversity issue in selection. This is an essential requirement in algorithm design for solving these MOP instances, which pose significant challenges to the diversity p-reservation of a search process. As a consequence, the promising performance achieved by MOEA/D-IR benefit from this selection mechanism.

7.3 Performance Comparisons on WFG Instances

Table 5 and Table 6 present the IGD and HV metric values obtained by seven MOEAs on all WFG instances. From these empirical results, it is clear that the performance of MOEA/D-IR is better than the other algorithms. It obtains the better metric values in 89 out of 108 comparisons. According to the Wilcoxons rank sum tests, 80 of these better results are with statistical significance. Similar to the observations in Section 7.1, the differences of metric values obtained by MOEA/D-IR and MOEA/D-STM are not significant on many WFG instances, especially for WFG6, WFG7 and WFG9. In the supplemental file, Figure 10 presents the non-dominated fronts obtained by these seven MOEAs. From those nine subfigures, we notice that the differences between these seven algorithms are not significant in most cases. WFG1 is a difficult test instance with a strong bias away from the PF. From Figure 10, we find that only the non-dominated fronts obtained by MOEA/D-IR and MOEA/D-FRRMAB

Table 3: Performance comparisons of IGD values on MOP test instances

Test Instance	MOEA/D-IR	MOEA/D-DRA	MOEA/D-M2M	MOEA/D-STM	MOEA/D-FRRMAB	HypE	NSGA-II
MOP1	**1.724E-2(2.76E-3)**	3.453E-1(4.72E-2)†	1.750E-2(7.29E-4)	3.502E-1(2.15E-2)†	3.346E-1(6.64E-2)†	3.634E-1(6.18E-3)†	3.645E-1(4.19E-3)†
MOP2	6.027E-2(7.07E-2)	2.326E-1(5.45E-2)†	**1.565E-2(4.15E-3)**‡	2.755E-1(7.50E-2)†	2.724E-1(7.02E-2)†	3.051E-1(1.14E-2)†	3.543E-1(5.55E-7)†
MOP3	**1.201E-2(1.53E-2)**	1.295E-1(5.12E-2)†	1.860E-2(7.79E-3)†	1.218E-1(1.25E-2)†	1.257E-1(4.85E-2)†	1.676E-1(1.96E-2)†	1.119E-1(1.15E-2)†
MOP4	7.430E-2(8.33E-2)	2.885E-1(3.31E-2)†	**1.111E-2(2.57E-3)**‡	2.812E-1(2.87E-2)†	2.885E-1(3.24E-2)†	3.078E-1(1.79E-2)†	3.107E-1(2.04E-2)†
MOP5	**2.031E-2(1.65E-3)**	3.165E-1(5.47E-3)†	2.166E-2(1.17E-3)†	3.145E-1(1.38E-2)†	3.185E-1(3.31E-3)†	3.004E-1(2.82E-2)†	2.796E-1(3.26E-2)†
MOP6	**5.555E-2(2.82E-3)**	2.952E-1(1.88E-2)†	6.681E-2(1.12E-3)†	2.891E-1(5.32E-2)†	3.012E-1(9.16E-3)†	3.044E-1(6.22E-6)†	3.044E-1(5.75E-6)†
MOP7	1.171E-1(1.96E-2)	3.477E-1(1.85E-2)†	**8.786E-2(2.08E-3)**‡	3.401E-1(8.71E-2)†	3.406E-1(2.71E-2)†	3.559E-1(2.15E-3)†	3.509E-1(7.99E-6)†

Wilcoxon's rank sum test at a 0.05 significance level is performed between MOEA/D-IR and each of the other MOEAs. † and ‡ denotes that the performance of the corresponding algorithm is significantly worse than or better than that of MOEA/D-IR, respectively. The best mean is highlighted in boldface with gray background.

Table 4: Performance comparisons of HV values on MOP test instances

Test Instance	MOEA/D-IR	MOEA/D-DRA	MOEA/D-M2M	MOEA/D-STM	MOEA/D-FRRMAB	HypE	NSGA-II
MOP1	**3.6378(3.49E-3)**	3.1102(8.83E-2)†	3.6368(1.24E-3)	3.1009(4.20E-2)†	3.1284(1.23E-1)†	3.0794(1.58E-2)†	3.0772(1.11E-2)†
MOP2	3.1397(9.32E-2)	3.0664(4.72E-2)†	**3.3066(7.57E-3)**‡	3.0450(5.02E-2)†	3.0424(4.32E-2)†	3.0026(8.88E-3)†	3.0000(2.02E-6)†
MOP3	**3.1962(2.67E-2)**	3.0784(5.11E-2)†	3.1836(1.41E-2)†	3.0782(4.88E-2)†	3.0777(5.20E-2)†	3.0635(2.07E-2)†	3.1148(1.44E-2)†
MOP4	3.4191(1.19E-1)	3.1557(2.26E-2)†	**3.5033(4.30E-3)**‡	3.1501(1.95E-2)†	3.1493(2.00E-2)†	3.1431(6.04E-3)†	3.1413(7.67E-3)†
MOP5	**3.6411(1.19E-1)**	2.7598(1.14E-1)†	3.6266(3.01E-3)†	2.7513(1.64E-1)†	2.6972(1.22E-1)†	2.8390(1.23E-1)†	3.0753(1.42E-1)†
MOP6	**7.7545(2.08E-3)**	7.5096(2.83E-2)†	7.7346(3.21E-3)	7.5101(1.22E-4)†	7.5007(1.36E-2)†	7.4773(2.61E-5)†	7.4978(6.46E-5)†
MOP7	7.2533(7.19E-3)	7.2150(2.13E-2)†	**7.2771(4.09E-3)**	7.2101(1.77E-2)†	7.2190(2.64E-2)†	7.2011(3.12E-3)†	7.2130(3.80E-5)†

Wilcoxon's rank sum test at a 0.05 significance level is performed between MOEA/D-IR and each of the other MOEAs. † and ‡ denotes that the performance of the corresponding algorithm is significantly worse than or better than that of MOEA/D-IR, respectively. The best mean is highlighted in boldface with gray background.

well converge to the PF and have a satisfied distribution along the PF. Although the non-dominated front obtained by MOEA/D-STM approximate the entire PF, the convergence is not as good as those of MOEA/D-IR and MOEA/D-FRRMAB. In contrast, the other algorithms can hardly find a well converged non-dominated fronts. The PF of WFG2 consists of five disconnected segments. Almost all algorithms have no difficulty on this test instance. However, some solutions found by MOEA/D-IR do not converge to the PF, while HypE cannot find solutions on the rightmost segment. WFG4 to WFG9 have the same PF shape. On WFG4 to WFG7 and WFG9, the non-dominated fronts found by these seven algorithms are not visually different. However, we notice that HypE usually find the extreme points of the PF, but lose some intermediate parts along the PF. WFG8 is a difficult problem, all algorithms cannot find a well converged non-dominated front. However, solutions found by MOEA/D-IR have a better convergence and diversity.

7.4 Performance Comparisons with Other Variants

To further investigate the underlying rationality of our proposed selection operator, we extend it into other two variants.

1. *Variant*-I: Instead of finding the appropriate solution for each subproblem based on the inter-relationship between subproblems and solutions, this variant matches subproblems and solutions in a random manner. Specifically, for each subproblem, we randomly choose a solution from the hybrid population of parents and offspring as the next parent. It is worth noting that a solution is at most allowed to be chosen one time.

2. *Variant*-II: For each subproblem, this variant assigns the solution that owns the best aggregation function value to it. This is a purely greedy strategy, and different subproblems can be assigned with the same solution.

Similar to MOEA/D-IR, we instantiate two MOEA/D variants based on the above introduced s-election operator variants, respectively. Empirical studies are conducted on all UF, MOP and WFG instances. Table 7 presents the performance comparisons of IGD and HV metrics. From the experimental results, it is clear that *Variant*-I is the worst among all these algorithms. Since solutions are merely selected in a random manner, this variant makes the algorithm degenerate to a purely random search, which is obviously not effective for tackling problems with complicated properties. As for *Variant*-II, MOEA/D-IR outperforms it in 45 out of 52 comparisons, where all these better results are with statistical significance. As discussed in Section 3, an agent has two requirements on its selected solutions, i.e., convergence and diversity. From the perspectives of subproblems and solutions, the achievements of these two requirements are respectively treated as their mutual-preferences. The inferior performance of *Variant*-II should be attributed to its purely greedy strategy, in which only the convergence requirement has been explicitly considered, while the diversity issue has been ignored. This results in a severe loss of population diversity. In contrast, our proposed selection operator takes the mutual-preferences of subproblems and solutions into consideration. It strikes a balance between convergence and diversity.

7.5 Impacts of Parameter Settings

There are two major parameters in the proposed selection operator:

- K_d: This parameter determines how many subproblems are considered to be related with a solution. It controls the trade-off between exploration and exploitation. A large K_d results in an explorative behavior, while a small K_d leads to an exploitative behavior.

- ϑ: This parameter decides the niche size of a subproblem. It controls the selection pressure on the local diversity of the subregion specified by a subproblem. A large ϑ, which results in a large niche for a subproblem, tends to increase the local diversity of the corresponding subregion. In contrast, a small ϑ, which leads to a small niche, might decrease the local diversity of the corresponding subregion.

To study how these two parameters influence the behavior of our proposed selection operator, we have considered four values for K_d: 1, 2, 4, 10 and six values for ϑ: 1, 2, 4, 8, 20, 30. In total, there are 24 combinations of K_d and ϑ. In our experiments, all parameters are kept the same as Section 6.4,

Table 5: Performance comparisons of IGD values on WFG test instances

Test Instance	MOEA/D-IR	MOEA/D-DRA	MOEA/D-FRRMAB	MOEA/D-M2M	MOEA/D-STM	HypE	NSGA-II
WFG1	**1.166E-1(4.05E-2)**	2.117E-1(9.10E-2)†	1.250E-1(5.45E-2)†	1.610E-1(5.63E-2)†	1.277E-1(4.68E-2)†	3.389E-1(2.49E-1)†	4.146E-1(1.02E-1)†
WFG2	3.681E-2(4.07E-4)	3.560E-2(1.25E-3)‡	3.591E-2(2.75E-4)‡	3.891E-2(3.05E-4)†	3.562E-2(1.25E-2)‡	7.296E-2(1.02E-2)†	**1.179E-2(6.01E-4)**‡
WFG3	**1.097E-2(2.36E-5)**	1.297E-2(2.08E-5)†	1.298E-2(1.79E-5)†	1.498E-2(1.79E-5)†	1.296E-2(2.00E-5)†	1.286E-2(2.01E-4)†	1.575E-2(7.14E-4)†
WFG4	**1.382E-2(2.26E-3)**	1.652E-2(8.32E-4)†	1.771E-2(1.80E-3)†	2.803E-2(3.12E-3)†	1.600E-2(4,34E-4)†	1.836E-2(9.68E-4)†	1.442E-2(7.92E-4)†
WFG5	6.854E-2(5.95E-4)	**6.648E-2(3.14E-3)**‡	6.717E-2(1.52E-5)‡	6.801E-2(3.17E-4)	6.718E-2(4.15E-5)‡	7.156E-2(6.44E-4)†	6.748E-2(4.39E-4)‡
WFG6	**1.544E-2(2.08E-5)**	1.547E-2(2.88E-5)	1.546E-2(2.91E-5)	1.596E-2(5.77E-4)†	1.545E-2(2.84E-5)	4.229E-2(2.34E-2)†	1.625E-2(7.33E-4)†
WFG7	1.563E-2(1.95E-5)	**1.562E-2(1.65E-5)**	1.662E-2(8.73E-6)	1.687E-2(7.22E-4)†	1.563E-2(1.09E-5)	2.100E-2(8.98E-4)†	1.709E-2(8.79E-4)†
WFG8	**1.447E-1(5.57E-2)**	1.562E-1(1.65E-2)†	1.587E-1(4.83E-2)†	1.901E-1(3.58E-2)†	1.547E-1(5.94E-2)†	1.872E-1(3.73E-2)†	1.801E-1(1.55E-2)†
WFG9	1.515E-2(8.17E-4)	1.619E-2(8.82E-4)†	1.763E-1(5.88E-2)†	1.963E-2(6.17E-4)†	**1.512E-2(3.97E-4)**	1.996E-2(7.53E-4)†	1.942E-2(1.37E-3)†

Wilcoxon's rank sum test at a 0.05 significance level is performed between MOEA/D-IR and each of the other MOEAs. † and ‡ denotes that the performance of the corresponding algorithm is significantly worse than or better than that of MOEA/D-IR, respectively. The best mean is highlighted in boldface with gray background.

Table 6: Performance comparisons of HV values on WFG test instances

Test Instance	MOEA/D-IR	MOEA/D-DRA	MOEA/D-FRRMAB	MOEA/D-M2M	MOEA/D-STM	HypE	NSGA-II
WFG1	**11.6325(6.14E-1)**	10.3021(5.60E-1)†	11.0056(6.01E-1)†	10.8256(6.28E-1)†	10.7506(2.45E-1)†	9.0128(5.48E-1)†	9.7436(1.77E+0)†
WFG2	11.4030(1.88E-1)	11.4445(9.10E-4)‡	11.4405(5.31E-4)‡	11.3455(6.11E-4)†	11.4440(1.31E-3)‡	10.9607(4.13E-1)†	**11.4522(9.57E-4)**‡
WFG3	**10.9596(6.55E-4)**	10.9494(7.89E-4)†	10.9493(4.27E-4)†	10.9461(4.82E-4)†	10.9495(9.85E-4)†	10.9245(2.45E-3)†	10.9494(3.14E-3)†
WFG4	**8.6540(2.82E-2)**	8.6345(1.80E-2)†	8.6262(1.63E-2)†	8.5493(5.00E-4)†	8.6462(1.24E-2)†	8.6229(1.26E-2)†	8.6266(1.65E-3)†
WFG5	8.1348(2.38E-2)	8.1336(1.73E-2)†	8.1325(6.77E-4)†	8.1359(3.14E-2)‡	8.1379(1.74E-2)‡	8.1358(2.49E-2)‡	**8.1925(2.81E-2)**‡
WFG6	**8.6729(1.06E-3)**	8.6726(6.88E-4)	8.6723(2.57E-4)	8.6697(1.61E-4)†	8.6727(7.42E-4)	8.4882(1.68E-1)†	8.6627(1.70E-3)†
WFG7	**8.6742(4.31E-4)**	8.6740(4.24E-4)	8.6688(3.87E-4)	8.6683(2.71E-4)†	8.6741(2.74E-4)	8.6590(3.08E-3)†	8.6674(1.24E-3)†
WFG8	**7.2441(5.72E-1)**	7.2040(4.24E-1)†	7.1117(5.21E-1)†	6.9081(4.85E-1)†	7.2432(5.45E-1)†	7.0111(5.28E-1)†	6.9139(2.42E-1)†
WFG9	8.4353(3.11E-2)	8.4299(9.91E-3)†	7.1058(4.19E-1)†	8.4100(1.65E-2)†	**8.4530(7.48E-3)**‡	8.4101(2.39E-2)†	8.4161(1.74E-2)†

Wilcoxon's rank sum test at a 0.05 significance level is performed between MOEA/D-IR and each of the other MOEAs. † and ‡ denotes that the performance of the corresponding algorithm is significantly worse than or better than that of MOEA/D-IR, respectively. The best mean is highlighted in boldface with gray background.

Table 7: Performance comparisons of MOEA/D-IR and two variants

	IGD			HV		
Test Instance	MOEA/D-IR	*Variant*-I	*Variant*-II	MOEA/D-IR	*Variant*-I	*Variant*-II
UF1	**9.932E-4(6.19E-5)**	1.332E+0(1.19E-1)†	1.456E-3(7.61E-5)†	**3.6623(1.45E-3)**	5.850E-1(1.44E-1)†	3.6632(7.27E-4)†
UF2	3.203E-3(9.95E-4)	4.869E-1(5.34E-2)†	**1.885E-3(9.05E-4)**‡	3.6524(1.33E-2)	2.1394(1.29E-1)†	**3.6613(2.95E-3)**‡
UF3	**9.110E-3(4.56E-3)**	9.839E-1(1.00E-1)†	1.255E-2(7.13E-3)†	**3.6473(9.65E-3)**	9.717E-1(1.33E-1)†	3.6355(3.90E-2)†
UF4	**5.213E-2(3.38E-3)**	1.511E-1(4.18E-3)†	5.388E-2(3.90E-3)†	**3.1834(1.42E-2)**	2.9041(1.02E-2)†	3.1712(1.46E-2)†
UF5	**2.625E-1(4.87E-2)**	5.119E+0(3.32E-1)†	3.913E-1(1.34E-1)†	**2.8180(1.56E-1)**	0†	2.2029(3.64E-1)†
UF6	**6.911E-2(3.48E-2)**	5.447E+0(6.33E-1)†	5.106E-1(2.25E-1)†	**3.2103(6.92E-2)**	0†	2.5069(5.46E-1)†
UF7	**1.089E-3(8.15E-5)**	1.327E+0(1.35E-1)†	1.969E-2(9.15E-2)†	**3.4963(1.24E-3)**	4.210E-1(1.56E-1)†	3.4290(2.27E-1)†
UF8	**2.639E-2(3.43E-3)**	2.300E+0(3.35E-1)†	3.710E-2(2.30E-2)†	**7.4170(6.89E-3)**	1.576E-2(6.12E-2)†	7.4043(3.65E-2)
UF9	**2.146E-2(1.04E-3)**	2.371E+0(4.03E-1)†	1.222E-1(4.63E-2)†	**7.7527(3.40E-3)**	5.317E-2(1.07E-1)†	7.3129(2.20E-1)†
UF10	4.338E-1(6.55E-2)	1.207E+1(2.15E+0)†	**2.987E-1(1.13E-1)**‡	3.7044(4.68E-1)	0†	**4.8733(1.15E+0)**‡
MOP1	**1.724E-1(2.76E-3)**	4.013E-1(4.75E-2)†	3.662E-1(5.65E-3)†	**3.6378(3.49E-3)**	2.9567(2.25E-1)†	3.0689(1.32E-2)†
MOP2	**1.027E-2(7.07E-2)**	3.471E-1(9.13E-3)†	3.481E-1(2.75E-2)†	**3.3297(9.32E-2)**	3.0000(2.04E-6)†	3.0024(1.07E-2)†
MOP3	**1.201E-2(1.53E-2)**	2.502E-1(1.43E-2)†	1.406E-1(7.50E-2)†	**3.1962(2.67E-2)**	2.8656(4.48E-2)†	3.0670(5.85E-2)†
MOP4	**7.430E-2(8.33E-2)**	3.661E-1(8.24E-3)†	3.238E-1(5.08E-3)†	**3.4191(1.19E-1)**	3.0309(2.04E-2)†	3.1262(1.07E-2)†
MOP5	**2.031E-2(1.65E-3)**	3.149E-1(3.37E-2)†	3.023E-1(6.69E-2)†	**3.6411(1.19E-1)**	2.4436(1.65E-1)†	2.7649(2.42E-1)†
MOP6	**5.555E-2(2.82E-3)**	3.082E-1(3.11E-3)†	3.002E-1(1.61E-2)†	**7.7545(2.08E-3)**	7.4941(2.43E-3)†	7.5009(2.87E-2)†
MOP7	**1.171E-1(1.96E-2)**	3.541E-1(4.93E-3)†	3.511E-1(1.77E-7)†	**7.2533(7.19E-3)**	7.2118(4.98E-3)†	7.2131(4.02E-5)†
WFG1	**1.166E-1(4.05E-2)**	1.317E+0(5.13E-2)†	1.177E-1(4.05E-2)†	**11.6325(6.14E-1)**	4.1055(2.71E-1)†	11.3993(7.52E-1)†
WFG2	3.481E-2(4.07E-4)	5.885E-1(8.22E-2)†	**3.477E-2(2.29E-3)**	**11.4030(1.88E-1)**	7.5054(4.43E-1)†	11.2795(3.40E-1)†
WFG3	**1.097E-2(2.36E-5)**	5.091E-1(4.71E-2)†	1.296E-2(2.15E-5)†	**10.9596(6.55E-4)**	7.2916(2.82E-1)†	10.9497(6.02E-4)
WFG4	**1.382E-2(2.26E-3)**	2.921E-1(3.15E-2)†	1.589E-2(3.01E-4)†	**8.6540(2.82E-2)**	6.4075(2.42E-1)†	8.6492(9.73E-3)†
WFG5	6.854E-2(5.95E-4)	2.828E-1(3.86E-2)†	**6.719E-2(4.29E-5)**‡	8.1348(2.38E-2)	6.3418(2.77E-1)†	**8.1453(2.46E-2)**‡
WFG6	**1.544E-2(2.08E-5)**	5.317E-1(5.91E-2)†	1.646E-2(2.60E-5)†	**8.6729(1.06E-3)**	4.9292(3.93E-1)†	8.6706(1.01E-3)†
WFG7	**1.562E-2(1.95E-5)**	4.796E-1(1.08E-1)†	1.762E-2(1.16E-5)†	**8.6742(4.31E-4)**	5.3557(5.14E-1)†	8.6700(4.53E-4)†
WFG8	**1.447E-1(5.57E-2)**	6.455E-1(5.11E-2)†	1.770E-1(4.35E-2)†	**7.2441(5.72E-1)**	4.0060(2.31E-1)†	7.1215(4.53E-1)†
WFG9	**1.515E-2(8.17E-4)**	4.906E-1(8.24E-2)†	1.611E-2(5.33E-4)†	8.4353(3.11E-2)	5.3224(4.76E-1)†	**8.4527(1.02E-2)**‡

Wilcoxon's rank sum test at a 0.05 significance level is performed between MOEA/D-IR and each of the two variants. † and ‡ denotes that the performance of the corresponding algorithm is significantly worse than or better than that of MOEA/D-IR, respectively. The best mean is highlighted in boldface with gray background.

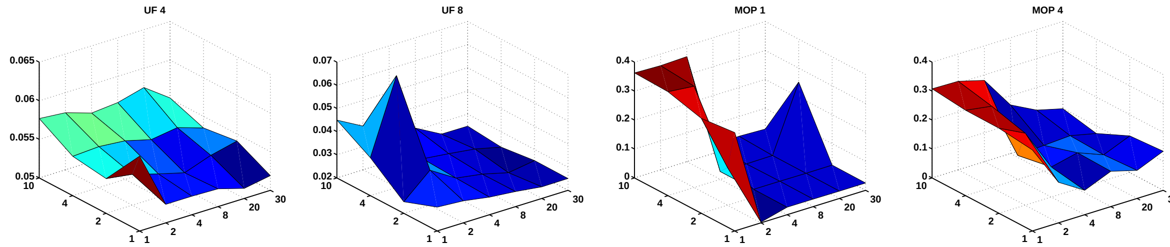

Figure 5: Parameter sensitivity studies of K_d and ϑ on UF4, UF8, MOP1 and MOP4 instances.

except the settings of K_d and ϑ. Twenty independent runs have been conducted for each combination of K_d and ϑ on each test instance introduced in Section 6.1. Considering the page limit, we only present the plots of the median IGD values found by 24 different combinations of K_d and ϑ on UF4, UF8, MOP1 and MOP4 instances in Fig. 5. The complete parameter sensitivity studies on all 26 MOP instances can be found in the supplemental file. From Fig. 5 and Figure 11 and Figure 12 of the supplemental file, we find that different parameter combinations lead to distinct performances of MOEA/D-IR. The performance of MOEA/D-IR, when both K_d and ϑ are small (e.g., $K_d = \vartheta = 1$), is usually inferior to the case that K_d is small (e.g., $K_d = 1$) but ϑ is large (e.g., $\vartheta = 30$). This is because the behavior of the selection operator is prone to be exploitative when K_d is small. On the other hand, a large ϑ setting, which is able to increase the local diversity and provides some explorative characteristics, exerts a complementary effect to the selection operator. These observations demonstrate the underlying mechanism of our proposed selection operator in trading off the convergence and diversity of the search process.

8 Conclusions

In MOEA/D, each subproblem is handled by an agent in a collaborative manner. The selection of MOEA/D can therefore be regarded as the process of choosing an appropriate solution by each agent. As an agent has two requirements, i.e., convergence and diversity, on its selected solution, it is judicious to treat these two requirements explicitly and simultaneously in designing selection mechanisms. This paper presents a simple yet effective attempt along this direction. It builds an inter-relationship between subproblems and solutions, according to their mutual-preferences. Based on this inter-relationship, each subproblem is able to be allocated with its desired solution, which is thus selected as the parent for the next generation. This selection operator trades off the mutual-preferences between subproblems and solutions, thus the convergence and diversity of the search process. Extensive experimental studies, conducting on several difficult problems with complicated PS shapes, demonstrate the effectiveness of our proposed MOEA/D-IR.

As for future directions, we make the following comments:

1. The effectiveness of recombination operators, such as crossover, usually relies on the selection of mating parents. However, most recombination operators choose mating parents in a random manner. This random mating scheme might leads to the inefficiency for offspring reproduction and premature convergence when tackling complicated problems. Exploitation versus exploration dilemma is a major issue in offspring reproduction. It is interesting to extend our idea in this paper to build the inter-relationship among solutions for mating selection, which can help to balance the exploration and exploitation of the search process.

2. Multiobjective optimization is usually used to assist decision makers (DMs) to find the solutions that fit their preferences. In this case, the agents might not be interested in optimizing all the subproblems that spread along the entire PF. Instead, they might be more interested in exploiting information around the subproblems that fit the characteristics of the DMs' preferences.

3. Many-objective optimization problem has become a major concern in evolutionary multiobjective optimization [20]. It is interesting to investigate the scalability of our proposed method for complicated problems with a large number of objectives.

The source codes and supplemental file of this paper can be obtained via request to the first author or downloaded from `http://www.cs.cityu.edu.hk/~51888309/`.

Acknowledgment

This work is supported by Hong Kong RGC GRF grant 9042038 (CityU 11205314), and the National Natural Science Foundation of China under Grant 6147324.

References

[1] A. Ropponen, R. Ritala, and E. N. Pistikopoulos, "Optimization issues of the broke management system in papermaking," *Computers & Chemical Engineering*, vol. 35, no. 11, pp. 2510–2520, 2011.

[2] M. G. C. Tapia and C. A. C. Coello, "Applications of multi-objective evolutionary algorithms in economics and finance: A survey," in *CEC'07: Proc. 2007 IEEE Congress on Evolutionary Computation*, Singapore, Sep. 2007, pp. 532–539.

[3] B. Amanulla, S. Chakrabarti, and S. Singh, "Reconfiguration of power distribution systems considering reliability and power loss," *IEEE Transactions on Power Delivery*, vol. 27, no. 2, pp. 918–926, Apr. 2012.

[4] K. Deb, *Multi-Objective Optimization Using Evolutionary Algorithms.* New York, NY, USA: John Wiley & Sons, Inc., 2001.

[5] Q. Zhang and H. Li, "MOEA/D: A Multiobjective Evolutionary Algorithm Based on Decomposition," *IEEE Transactions on Evolutionary Computation*, vol. 11, no. 6, pp. 712–731, Dec. 2007.

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

[7] K. Li, Á. Fialho, and S. Kwong, "Multi-objective differential evolution with adaptive control of parameters and operators," in *LION'11: Proc. of the 5th International Conference Learning and Intelligent OptimizatioN*, C. A. C. Coello, Ed. Springer LNCS, Jan. 2011, pp. 473–487.

[8] K. Li, S. Kwong, R. Wang, K.-S. Tang, and K.-F. Man, "Learning paradigm based on jumping genes: A general framework for enhancing exploration in evolutionary multiobjective optimization," *Information Sciences*, vol. 226, pp. 1–22, 2013.

[9] K. Li and S. Kwong, "A general framework for evolutionary multiobjective optimization via manifold learning," *Neurocomputing*, vol. 146, pp. 65–74, 2014.

[10] E. Zitzler, M. Laumanns, and L. Thiele, "SPEA2: Improving the Strength Pareto Evolutionary Algorithm for Multiobjective Optimization," in *Evolutionary Methods for Design, Optimisation and Control with Application to Industrial Problems (EUROGEN 2001)*, K. Giannakoglou *et al.*, Eds. International Center for Numerical Methods in Engineering (CIMNE), 2002, pp. 95–100.

[11] N. Beume, B. Naujoks, and M. Emmerich, "SMS-EMOA: Multiobjective selection based on dominated hypervolume," *European Journal of Operational Research*, vol. 181, no. 3, pp. 1653–1669, 2007.

[12] K. Li, J. Zheng, M. Li, C. Zhou, and H. Lv, "A novel algorithm for non-dominated hypervolume-based multiobjective optimization," in *SMC'09: Proc. of 2009 the IEEE International Conference on Systems, Man and Cybernetics.* San Antonio, TX, USA: IEEE, Oct. 2009, pp. 5220–5226.

[13] K. Li, S. Kwong, J. Cao, M. Li, J. Zheng, and R. Shen, "Achieving balance between proximity and diversity in multi-objective evolutionary algorithm," *Information Sciences*, vol. 182, no. 1, pp. 220–242, 2012.

[14] E. Zitzler and L. Thiele, "Multiobjective Evolutionary Algorithms: A Comparative Case Study and the Strength Pareto Approach," *IEEE Transactions on Evolutionary Computation*, vol. 3, no. 4, pp. 257–271, 1999.

[15] T. Murata, H. Ishibuchi, and M. Gen, "Specification of genetic search directions in cellular multi-objective genetic algorithms," in *EMO'01: Proc. of the 1st International Conference on Evolutionary Multi-Criterion Optimization*, 2001, pp. 82–95.

[16] H.-L. Liu, F. Gu, and Q. Zhang, "Decomposition of a multiobjective optimization problem into a number of simple multiobjective subproblems," *IEEE Transactions on Evolutionary Computation*, vol. 18, no. 3, pp. 450–455, Jun. 2014.

[17] K. Bringmann and T. Friedrich, "The maximum hypervolume set yields near-optimal approximation," in *GECCO'10: Proc. of the 12th annual conference on Genetic and evolutionary computation*, 2010, pp. 511–518.

[18] I. Giagkiozis, R. C. Purshouse, and P. J. Fleming, "Generalized decomposition," in *EMO'13: Proc. of the 7th International Conference on Evolutionary Multi-Criterion Optimization*, 2013, pp. 428–442.

[19] H. Ishibuchi, N. Akedo, and Y. Nojima, "Behavior of multi-objective evolutionary algorithms on many-objective knapsack problems," *IEEE Transactions on Evolutionary Computation*, 2014, accepted for publication.

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

[21] L. Ke, Q. Zhang, and R. Battiti, "MOEA/D-ACO: A multiobjective evolutionary algorithm using decomposition and ant colony," *IEEE Transactions on Cybernetics*, vol. 43, no. 6, 2013.

[22] S. Zapotecas Martínez and C. A. Coello Coello, "A direct local search mechanism for decomposition-based multi-objective evolutionary algorithms," in *CEC'12: Proc. of the 2012 IEEE Congress on Evolutionary Computation*, 2012, pp. 1–8.

[23] H. Li and Q. Zhang, "Multiobjective optimization problems with complicated Pareto sets, MOEA/D and NSGA-II," *IEEE Transactions on Evolutionary Computation*, vol. 12, no. 2, pp. 284–302, Apr. 2009.

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

[25] F. Gu and H.-L. Liu, "A novel weight design in multi-objective evolutionary algorithm," in *CIS'10: Proc. of the 2010 International Conference on Computational Intelligence and Security*, 2010, pp. 137–141.

[26] K. Li, Q. Zhang, S. Kwong, M. Li, and R. Wang, "Stable matching based selection in evolutionary multiobjective optimization," *IEEE Transactions on Evolutionary Computation*, 2013, accepted for publication.

[27] D. Gale and L. S. Shapley, "College admissions and the stability of marriage," *American Mathematical Monthly*, vol. 69, pp. 9–15, 1962.

[28] K. Miettinen, *Nonlinear Multiobjective Optimization.* Kluwer Academic Publisher, 1999, vol. 12.

[29] I. Das and J. E. Dennis, "Normal-Boundary Intersection: A New Method for Generating the Pareto Surface in Nonlinear Multicriteria Optimization Problems," *SIAM Journal on Optimization*, vol. 8, no. 3, pp. 631–657, 1998.

[30] Q. Zhang, W. Liu, and 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.

[31] Q. Zhang, A. Zhou, S. Zhao, P. N. Suganthan, W. Liu, and 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, 2008.

[32] S. Das and P. N. Suganthan, "Differential Evolution: A Survey of the State-of-the-Art," *IEEE Transactions on Evolutionary Computation*, vol. 15, no. 1, pp. 4–31, 2011.

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

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

[35] E. Zitzler, L. Thiele, M. Laumanns, C. Fonseca, and V. da Fonseca, "Performance assessment of multiobjective optimizers: an analysis and review," *IEEE Transactions on Evolutionary Computation*, vol. 7, no. 2, pp. 117–132, Apr. 2003.

[36] P. Bosman and D. Thierens, "The balance between proximity and diversity in multiobjective evolutionary algorithms," *IEEE Transactions on Evolutionary Computation*, vol. 7, no. 2, pp. 174–188, Apr. 2003.

[37] K. Li, Á. Fialho, S. Kwong, and Q. Zhang, "Adaptive operator selection with bandits for multiobjective evolutionary algorithm based decomposition," *IEEE Transactions on Evolutionary Computation*, no. 1, pp. 114–130, Feb. 2014.

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

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

[40] K. Deb, L. Thiele, M. Laumanns, and E. Zitzler, "Scalable test problems for evolutionary multiobjective optimization," in *Evolutionary Multiobjective Optimization*, ser. Advanced Information and Knowledge Processing, A. Abraham, L. Jain, and R. Goldberg, Eds. Springer London, 2005, pp. 105–145.