Adaptive Two-Level Matching-Based Selection for Decomposition Multi-Objective Optimization†

Mengyuan Wu[1], Ke Li[2], Sam Kwong[1], Yu Zhou[1], and Qingfu Zhang[1]

[1]Department of Computer Science, City University of Hong Kong
[2]CERCIA, School of Computer Science, University of Birmingham

The balance between convergence and diversity is a key issue of evolutionary multi-objective optimization. The recently proposed stable matching-based selection provides a new perspective to handle this balance under the framework of decomposition multi-objective optimization. In particular, the stable matching between subproblems and solutions, which achieves an equilibrium between their mutual preferences, implicitly strikes a balance between the convergence and diversity. Nevertheless, the original stable matching model has a high risk of matching a solution with a unfavorable subproblem which finally leads to an imbalanced selection result. In this paper, we propose an adaptive two-level stable matching-based selection for decomposition multi-objective optimization. Specifically, borrowing the idea of stable matching with incomplete lists, we match each solution with one of its favorite subproblems by restricting the length of its preference list during the first-level stable matching. During the second-level stable matching, the remaining subproblems are thereafter matched with their favorite solutions according to the classic stable matching model. In particular, we develop an adaptive mechanism to automatically set the length of preference list for each solution according to its local competitiveness. The performance of our proposed method is validated and compared with several state-of-the-art evolutionary multi-objective optimization algorithms on 62 benchmark problem instances. Empirical results fully demonstrate the competitive performance of our proposed method on problems with complicated Pareto sets and those with more than three objectives.

Keywords: Convergence and diversity, stable matching with incomplete lists, adaptive mechanism, decomposition, multiobjective optimization.

1 Introduction

The multi-objective optimization problem (MOP) considered in this paper is defined as follows [1]:

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

where $\mathbf{x} = (x_1, \cdots, x_n)^T$ is a n-dimensional decision vector and $\mathbf{F}(\mathbf{x})$ is a m-dimensional objective vector. $\Omega \subseteq \mathbb{R}^n$ is the feasible region of the decision space, while $\mathbf{F} : \Omega \to \mathbb{R}^m$ is the corresponding

†This work has been submitted to the IEEE for possible publication. Copyright may be transferred without notice, after which this version may no longer be accessible.

attainable set in the objective space $\mathbb{R}^m$. Given two solutions $\mathbf{x}^1, \mathbf{x}^2 \in \Omega$, $\mathbf{x}^1$ is said to dominate $\mathbf{x}^2$, denoted by $\mathbf{x}^1 \preceq \mathbf{x}^2$, if and only if $f_i(\mathbf{x}^1) \leq f_i(\mathbf{x}^2)$ for all $i \in \{1, \cdots, m\}$ and $\mathbf{F}(\mathbf{x}^1) \neq \mathbf{F}(\mathbf{x}^2)$. A solution $\mathbf{x}^* \in \Omega$ is said to be Pareto optimal if and only if no solution $\mathbf{x} \in \Omega$ dominates it. All Pareto optimal solutions constitute the Pareto-optimal set (PS) and the corresponding Pareto-optimal front (PF) is defined as $PF = \{\mathbf{F}(\mathbf{x}) | \mathbf{x} \in PS\}$.

Evolutionary multi-objective optimization (EMO) algorithms, which are capable of approximating the whole PS and PF in a single run, have been widely accepted as a major approach for solving MOPs. Convergence and diversity are two basic issues in design of multi-objective optimization algorithms: the convergence means the closeness to the PF while the diversity indicates the spread and uniformity along the PF. Selection, which selects the elites to survive to the next generation, plays a key role in balancing convergence and diversity. Based on different selection strategies, existing EMO algorithms can be roughly classified into three categories: 1) Pareto-based methods that use the Pareto dominance to emphasize the convergence and uses the density estimation technique to maintain the population diversity, e.g., [2–5]; 2) indicator-based methods that transform a MOP into a single-objective optimization problem by employing some performance indicator, e.g., hypervolume [6], to assign fitness values to solutions, e.g., [7–10]; 3) decomposition-based methods that decompose the original MOP into several subproblems, either scalar aggregation functions or simplified MOPs, and optimize them in a collaborative manner, e.g., [11–14].

This paper focuses on the multi-objective evolutionary algorithm based on decomposition (MOEA/D) [11], especially its selection mechanism. The original MOEA/D employs a steady-state selection scheme, where the parent population is updated immediately after the generation of an offspring. In particular, an offspring is able to replace its neighboring parents when it has a better aggregation function value for the corresponding subproblem. In terms of population diversity, [12] suggested to restrict the maximum number of replacements taken by an offspring. More recently, [13] developed a new perspective to understand the selection process of MOEA/D, which is modeled as one-to-one matching between subproblems and solutions. In particular, subproblems and solutions are treated as two sets of agents whose mutual preferences are defined representing convergence and diversity, respectively. The stable matching between subproblems and solutions achieves an equilibrium between their mutual preferences, and thus the balance between convergence and diversity. Nevertheless, as discussed in [15] and [16], partially due to the overrated convergence property, both the original MOEA/D and the stable matching based method fail to maintain the population diversity when solving some complicated problems. Bearing this consideration in mind, [16] modified the mutual preference definition and developed a straightforward but more effective selection mechanism based on the interrelationship between subproblems and solutions. Later on, [17] proposed an adaptive replacement strategy, which adjusts the replacement neighborhood size dynamically, to assign solutions to their most suitable subproblems. It is also interesting to note that some researchers have paved new avenues to take the advantages of the Pareto dominance- and decomposition-based selection mechanisms in a single paradigm [18–20].

To have a good balance between convergence and diversity, this paper suggests a two-level stable matching-based selection mechanism for MOEA/D. More specifically, borrowing the idea from the stable matching with incomplete lists [21], we restrict the number of subproblems, with which a solution is allowed to match, at the first-level stable matching. In this case, a solution can only be assigned to one of its favorite subproblems. At the second-level stable matching, the remaining subproblems, which have not found their matching pairs yet during the first-level matching procedure, are assigned with the suitable solutions according to the remaining preference information. From the preliminary results shown in [22], the two-level stable matching-based selection works well on some selected benchmark problems. Nevertheless, the setting of the length of the incomplete preference list has a significant impact on the performance and is problem dependent. By analyzing the underlying mechanism of the proposed two-level stable matching-based selection in deep, we further develop an effective and universal mechanism to set the length of the incomplete preference list for each solution on the fly. In particular, each solution is at first associated to its closest subproblem. Then, the length of the incomplete preference list for each solution is adaptively determined by the local competitiveness information extracted from its neighboring space. Comprehensive experiments on 62 benchmark problem instances with various characteristics fully demonstrate the effectiveness and competitiveness of our proposed algorithm for solving problems with complicated PSs and those with more than three objectives, comparing to several state-of-the-art EMO algorithms.

The rest of the paper is organized as follows. Section 2 provides some preliminaries of this paper. Thereafter, the proposed algorithm is described step by step in Section 3. Section 4 and Section 5 provide the experimental settings and the analysis of the empirical results. Finally, Section 6 concludes this paper and provides some future directions.

2 Preliminaries

In this section, we first introduce some background knowledge of MOEA/D and the stable matching-based selection. Then, our motivations are developed by analyzing their underlying mechanisms and drawbacks.

2.1 MOEA/D

As a representative of the decomposition-based method, MOEA/D has become an increasingly popular choice for *posterior* multi-objective optimization. Generally speaking, there are two basic components in MOEA/D: one is *decomposition* and the other is *collaboration*. The following paragraphs give some general descriptions of each component separately.

2.1.1 Decomposition

The basic idea of decomposition is transforming the original MOP into a single-objective scalar optimization subproblem. There are many established decomposition methods developed in the classic multi-objective optimization [23], among which the most popular ones are weighted sum, Tchebycheff (TCH) and boundary intersection approaches. Without loss of generality, this paper considers the inverted TCH approach [13], which is defined as follows:

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

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 $\sum_{i=1}^m w_i = 1$. In practice, w_i is set to be a very small number, say 10^{-6}, when $w_i = 0$. $\mathbf{z}^* = (z_1^*, \cdots, z_m^*)^T$ is an Utopian objective vector where $z_i^* = \min_{\mathbf{x}\in\Omega} f_i(x)$, $i \in \{1, \cdots, m\}$. Note that the search direction of the inverted TCH approach is $\mathbf{w}$, and the optimal solution of (2) is a Pareto-optimal solution of the MOP defined in (1) under some mild conditions. We can expect to obtain various Pareto-optimal solutions by using different weight vectors in (2). In MOEA/D, a set of uniformly distributed weight vectors are sampled from a unit simplex.

2.1.2 Collaboration

As discussed in [11], the neighboring subproblems, associated with the geometrically close weight vectors, share similar optima. In other words, the optimal solution of $g^{tch}(\cdot|\mathbf{w}^1, \mathbf{z}^*)$ is close to that of $g^{tch}(\cdot|\mathbf{w}^2, \mathbf{z}^*)$, given $\mathbf{w}^1$ and $\mathbf{w}^2$ are close to each other. In MOEA/D, each solution is associated with a subproblem. During the optimization process, the solutions cooperate with each other via a well defined neighborhood structure and they solve the subproblems in a collaborative manner. In practice, the collaboration is implemented as a restriction on the range of the mating and update procedures. More specifically, the mating parents are selected from neighboring subproblems and a newly generated offspring is only used to update its corresponding neighborhood. Furthermore, since different subproblems might have various difficulties, it is more reasonable to dynamically allocate the computational resources to different subproblems than treating all subproblems equally important. In [24], a dynamic resource allocation scheme is developed to allocate more computational resources to those promising ones according to their online performance.

2.2 Stable Matching-Based Selection

Stable marriage problem (SMP) was originally introduced in [25] and its related work won the 2012 Nobel Prize in Economics. In a nutshell, the SMP is about how to match two sets of agents, say men and women, which have mutual preferences over each other. A stable matching should not contain a man and a woman who are not matched together but prefer each other to their assigned spouses.

In MOEA/D, subproblems and solutions can be treated as two sets of agents which have mutual preferences over each other. In particular, a subproblem prefers a solution that optimizes its underlying single-objective scalar optimization problem as much as possible; while a solution hopes to have a well distribution in the objective space. The ultimate goal of selection is to select the best solution for each subproblem, and vice versa. In this case, we can treat the selection procedure as a matching procedure between subproblems and solutions. To the best of our knowledge, MOEA/D-STM [13] is the first one that models the selection procedure of MOEA/D as a SMP, and encouraging results have been reported therein. The framework of the stable matching-based selection proposed in [13] contains two basic components: one is *preference settings* and the other is *matching model*. The following paragraphs briefly describe these two components.

2.2.1 Preference Settings

The preference of a subproblem p on a solution $\mathbf{x}$ is defined as:

$$\Delta_P(p, \mathbf{x}) = g^{tch}(\mathbf{x}|\mathbf{w}, \mathbf{z}^*), \tag{3}$$

where $\mathbf{w}$ is the weight vector of p. Consequently, $\Delta_P(p, \mathbf{x})$ measures the convergence of $\mathbf{x}$ with respect to p. The preference of a solution $\mathbf{x}$ to a subproblem p is defined as:

$$\Delta_X(\mathbf{x}, p) = \|\overline{F}(\mathbf{x}) - \frac{\mathbf{w}^{\mathrm{T}} \cdot \overline{F}(\mathbf{x})}{\mathbf{w}^{\mathrm{T}} \cdot \mathbf{w}} \mathbf{w}\|, \tag{4}$$

where $\overline{F}(\mathbf{x})$ is the normalized objective vector $\mathbf{x}$ and $\| \cdot \|$ is the ℓ_2-norm. Since the weight vectors are usually uniformly distributed, it is desirable that the optimal solution of each subproblem has the shortest perpendicular distance to its corresponding weight vector. For the sake of simplicity, $\Delta_X(\mathbf{x}, p)$ can be used to measure the diversity of a solution [13].

2.2.2 Matching Model

Based on the above preference settings, [13] employed the classic deferred acceptance procedure (DAP) developed in [25] to find a stable matching between subproblems and solutions. The pseudo code of this stable matching-based selection procedure is given in Algorithm 1. The DAP presented in Algorithm 2 simulates a matching request sent from a subproblem to a solution. If the matching request is refused by the solution, the subproblem will remove that solution from its preference list. In particular, Ψ_P and Ψ_X are two sets containing subproblems' and solutions' preference lists, each element of which represents the preference list of a subproblem over all solutions, and vice versa. A preference list is built by sorting the preference values in ascending order. M indicates the set of

Algorithm 1: $\mathsf{STM}(P, S, \Psi_P, \Psi_X)$

Input:

- subproblem set P, solution set S
- sets of preference lists Ψ_P and Ψ_X

Output: stable matching set M

$P_u \leftarrow P$, $S_u \leftarrow S$, $M \leftarrow \emptyset$;
while $P_u \neq \emptyset$ **do**
    $p \leftarrow$ Randomly pick a subproblem from P_u;
    $\mathbf{x} \leftarrow$ First solution on p's preference list;
    $M \leftarrow \mathsf{DAP}(p, \mathbf{x}, P_u, S_u, M, \Psi_P, \Psi_X)$;
return M

Algorithm 2: $\mathsf{DAP}(p, \mathbf{x}, P_u, S_u, M, \Psi_P, \Psi_X)$

Input:

- current subproblem p and solution $\mathbf{x}$
- unmatched subproblem and solution sets P_u and S_u
- current stable matching set M
- sets of preference lists Ψ_P and Ψ_X

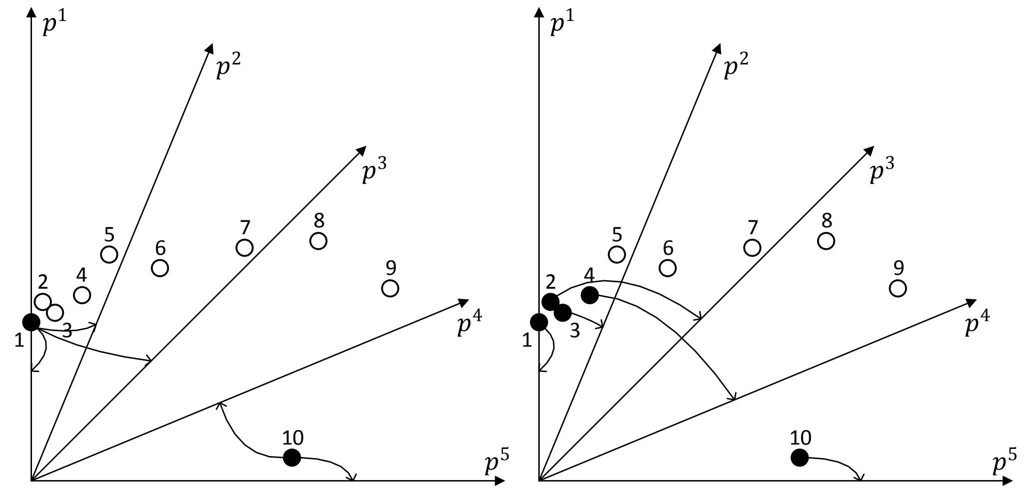
Output: stable matching set M

if $\mathbf{x} \in S_u$ **then**
    $P_u \leftarrow P_u \setminus p$, $S_u \leftarrow S_u \setminus \mathbf{x}$;
    $M \leftarrow M \cup (p, \mathbf{x})$; // match p and $\mathbf{x}$
else
    $p' \leftarrow$ Current matching mate of $\mathbf{x}$;
    if $\mathbf{x}$ *prefers* p *to* p' **then**
        Remove $\mathbf{x}$ from p''s preference list;
        $P_u \leftarrow P_u \cup p' \setminus p$;
        $M \leftarrow M \cup (p, \mathbf{x}) \setminus (p', \mathbf{x})$;
    else
        Remove $\mathbf{x}$ from p's preference list;
return M

all the constructed matching pairs. It is worth noting that the convergence and diversity have been aggregated into the preference settings. Therefore, the stable matching between subproblems and solutions is claimed to strike the balance between convergence and diversity.

2.3 Drawbacks of MOEA/D and MOEA/D-STM

In this subsection, we discuss some drawbacks of the selection mechanisms of MOEA/D and MOEA/D-STM.



(a) MOEA/D (b) MOEA/D-STM

Figure 1: Comparisons of MOEA/D and MOEA/D-STM.

2.3.1 MOEA/D

The update mechanism of the original MOEA/D is simple and efficient, yet greedy. In a nutshell, each subproblem simply selects its best solution according to the corresponding scalar optimization function value. As discussed in [24], since different parts of the PF might have various difficulties, it might be easier for some subproblems to find the optimal solutions. During some intermediate stages of the optimization process, the currently elite solutions of some relatively easier subproblems might also be good candidates for the others. In this case, these elite solutions can easily take over all subproblems. In addition, it is highly likely that the offspring solutions generated from these elite solutions crowd into the neighboring space of the corresponding subproblems. Therefore, this purely fitness driven selection mechanism can be severely harmful for the population diversity, and thus leading to the failure of MOEA/D on some challenging problems [15]. Let us consider an example shown in Fig. 1(a), where five out of ten solutions need to be selected for five subproblems. Since $\mathbf{x}^1$ is currently the best solution for $\{p^1, p^2, p^3\}$ and $\mathbf{x}^{10}$ is the current best candidate for $\{p^4, p^5\}$, these two elite solutions finally take over all five subproblems. Obviously, the population diversity of this selection result is not satisfied.

2.3.2 MOEA/D-STM

As discussed in [26], the DAP maximizes the satisfactions of the preferences of men and women in order to maintain the stable matching relationship. According to the preference settings for subproblems, solutions closer to the PF are always in the front of the subproblems' preference lists. In this case, the DAP might make some solutions match themselves with subprobolems lying in the rear of their preference lists. Even worse, as discussed in Section 2.3.1, these currently well converged solutions may crowd in a narrow space. This obviously goes against the population diversity. Let us consider the same example discussed in Section 2.3.1. The preference lists of subproblems and solutions are:

$$\Psi_P = \begin{array}{r} p^1: \\ p^2: \\ p^3: \\ p^4: \\ p^5: \end{array} \begin{array}{c} [\,1 \quad 2 \quad 3 \quad 4 \quad 5 \quad 6 \quad 7 \quad 8 \quad 10 \quad 9\,] \\ [\,1 \quad 3 \quad 2 \quad 4 \quad 5 \quad 6 \quad 7 \quad 8 \quad 10 \quad 9\,] \\ [\,1 \quad 3 \quad 2 \quad 4 \quad 6 \quad 5 \quad 7 \quad 10 \quad 8 \quad 9\,] \\ [\,10 \quad 1 \quad 3 \quad 2 \quad 4 \quad 9 \quad 6 \quad 5 \quad 7 \quad 8\,] \\ [\,10 \quad 1 \quad 3 \quad 2 \quad 4 \quad 9 \quad 6 \quad 5 \quad 7 \quad 8\,] \end{array} \tag{5}$$

$$\Psi_X = \begin{array}{r} \mathbf{x}^1: \\ \mathbf{x}^2: \\ \mathbf{x}^3: \\ \mathbf{x}^4: \\ \mathbf{x}^5: \\ \mathbf{x}^6: \\ \mathbf{x}^7: \\ \mathbf{x}^8: \\ \mathbf{x}^9: \\ \mathbf{x}^{10}: \end{array} \begin{array}{c} [\,1 \quad 2 \quad 3 \quad 4 \quad 5\,] \\ [\,1 \quad 2 \quad 3 \quad 4 \quad 5\,] \\ [\,1 \quad 2 \quad 3 \quad 4 \quad 5\,] \\ [\,2 \quad 1 \quad 3 \quad 4 \quad 5\,] \\ [\,2 \quad 1 \quad 3 \quad 4 \quad 5\,] \\ [\,2 \quad 3 \quad 1 \quad 4 \quad 5\,] \\ [\,3 \quad 2 \quad 4 \quad 1 \quad 5\,] \\ [\,3 \quad 4 \quad 2 \quad 5 \quad 1\,] \\ [\,4 \quad 3 \quad 5 \quad 2 \quad 1\,] \\ [\,5 \quad 4 \quad 3 \quad 2 \quad 1\,] \end{array}. \tag{6}$$

From (5), we can clearly see that $\mathbf{x}^1$ to $\mathbf{x}^4$ occupy the top positions of the preference lists of all subproblems. By using Algorithm 1, we have the selection/matching result shown in Fig. 1(b), where $\mathbf{x}^1$ to $\mathbf{x}^4$ crowd in a narrow area between p^1 and p^2. This is obviously harmful for the population diversity as well.

From the above discussions, we find that the original selection mechanism of MOEA/D is a convergence first and diversity second strategy, which might give excessive priority to the convergence requirement. On the other hand, although the stable matching-based selection mechanism intends to achieve an equilibrium between convergence and diversity, the stable matching between subproblems and solutions may fail to keep the population diversity. This is because no restriction is given to the subproblem with which a solution can match. In other words, a solution can match with a very unfavorable subproblem in order to achieve a stable matching. To relieve this side effect, the next section suggests a strategy to take advantages of partial information from the preference lists when finding the stable matching between subproblems and solutions.

3 Adaptive Two-Level Stable Matching-Based Selection

3.1 Two-Level Stable Matching-Based Selection

In the canonical SMP, each man/woman holds a complete and strictly ordered preference list over all agents from the other side. However, in practice, it may happen that a man/woman declares some unacceptable partners [26], and this results in a SMP with incomplete lists [26,27]. By these means, a man/woman is only allowed to match with an agent that appears on his/her incomplete preference list. Due to the restriction of the incomplete preference list, there is no guarantee that all agents can have a stable matching mate. A stable matching for the SMP with incomplete lists does not contain such a pair of man and woman who are not matched with each other, but acceptable to each other and even prefer each other to their current matching mate.

To overcome the drawbacks discussed in Section 2.3, here we propose a two-level stable matching-based selection mechanism for MOEA/D. In the first-level stable matching, the selection process is modeled as a SMP with incomplete lists. More specifically, let us assume that there are N subproblems and Q solutions, where $N < Q$. The pseudo code of the stable matching with incomplete lists is given in Algorithm 3 and works as follows:

- **Step 1)** For each solution $\mathbf{x}^i$, $i \in \{1, \cdots, Q\}$, we only keep the first $R[i]$, where $0 < R[i] \leq N$,

Algorithm 3: $\mathsf{STMIC}(P, S, \Psi_P, \Psi_X, R)$

Input:

- subproblem set P, solution set S
- sets of preference lists Ψ_P and Ψ_X
- lengths of solution's preference list R

Output: stable matching set M

1 $P_u \leftarrow P$, $S_u \leftarrow S$, $M \leftarrow \emptyset$;
2 **for** $i \leftarrow 1$ to Q **do**
3   Keep the first $R[i]$ subproblems on $\mathbf{x}^i$'s preference list and discard the remainders;
4   Remove $\mathbf{x}^i$ from the discarded subproblems' preference lists;
5 **while** $P_u \neq \emptyset$ **do**
6     $p \leftarrow$ Randomly select a subproblem from P_u;
7   **if** *p's preference list* $\neq \emptyset$ **then**
8         $\mathbf{x} \leftarrow$ First solution on p's preference list;
9         $M \leftarrow \mathsf{DAP}(p, \mathbf{x}, P_u, S_u, M, \Psi_P, \Psi_X)$;
10   **else**
11         $P_u \leftarrow P_u \setminus p$;
12 **return** M;

subproblems in the preference list of solution , while the remaining subproblems are discarded and not considered any longer (line 3). Correspondingly, $\mathbf{x}^i$ is removed from the preference lists of the discarded subproblems (line 4).

- **Step 2)** Based on the incomplete preference information, the DAP is employed to find a stable matching between subproblems and solutions (line 5 to line 11). During the matching process, the subproblems that fails to establish a stable matching pair with any solution on their preference list are removed from P_u (line 11).

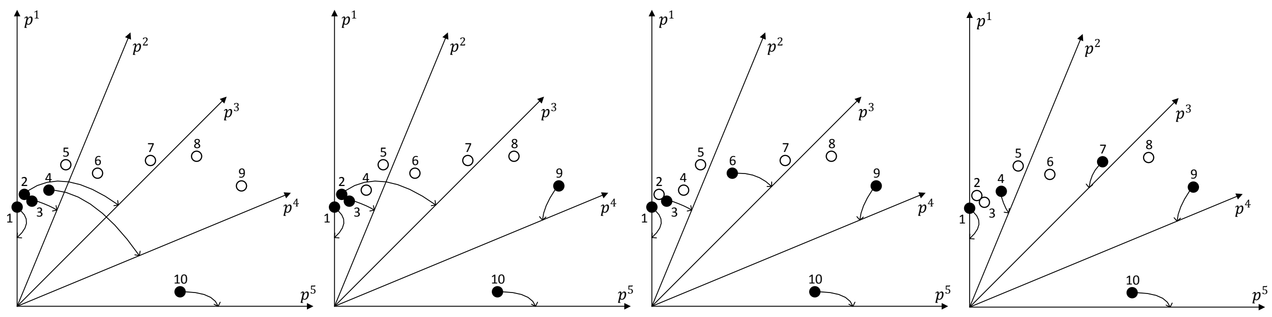
The matched solutions will be collected to form the selected solutions from the first-level stable matching. Thanks to the restriction from the incomplete preference information, each solution is only allowed to match with one of its $R[i]$ favorite subproblems that are close enough according to (4). By these means, we can expect that the population diversity is strengthened. This is because a solution is not prevented to match with an unfavorable subproblem which lies out of its incomplete preference list.

During the first-level stable matching, not all subproblems are assigned with a solution due to the incomplete preference lists. To remedy this issue, the second-level stable matching with complete preference lists is developed to find a stable solution for each unmatched subproblem. Algorithm 4 demonstrates the pseudo code of the two-level stable matching-based selection mechanism. After the first-level stable matching (line 1 to line 4 Algorithm 4), we collect the unmatched subproblems and solutions in line 5 and line 6. Then, we compute the preference lists of the unmatched subproblems and solutions in line 7. Afterwards, we employ Algorithm 1 to find the stable matching between them (line 8 of Algorithm 4). At the end, the matched solutions from the two levels of stable matching are gathered to form the final selection results (line 10).

Algorithm 4: SelectionSTM2L(P, S)

Input: subproblem set P and solution set S
Output: solution set S

```
/* First-level stable matching                                    */
```
Compute Ψ_P and Ψ_X for P and S;
$R \leftarrow$ Set the length of each solution's preference list;
$M \leftarrow \mathsf{STMIC}(P, S, \Psi_P, \Psi_X, R)$;
$(P_m, S_m) \leftarrow M$;
```
/* Second-level stable matching                                   */
```
$P_u \leftarrow P \setminus P_m$;
$S_u \leftarrow S \setminus S_m$;
Compute Ψ'_P and Ψ'_X for P_u and S_u;
$M' \leftarrow \mathsf{STM}(P_u, S_u, \Psi'_P, \Psi'_X)$;
$(P'_m, S'_m) \leftarrow M'$;
```
/* Gather matched solutions                                       */
```
$S \leftarrow S_m \cup S'_m$;
return S;



(a) STM2L ($r = 4$) (b) STM2L ($r = 3$) (c) STM2L ($r = 2$) (d) STM2L ($r = 1$)

Figure 2: Selection results of the two-level stable matching-based selection mechanism with different r settings.

3.2 Impacts of the Length of the Incomplete Preference list

The major difference between the selection mechanism developed in Section 3.1 and the one proposed in [13] is the use of incomplete preference lists in the first-level stable matching procedure. By restricting the number of favorite subproblems for each solution during the matching procedure, we can expect an improvement on the population diversity. A natural question is whether the length of the incomplete preference list affects the behavior of our proposed two-level stable matching-based selection? Let us consider the example discussed in Fig. 1 again. Here, we consider a static length of preference list for all solutions, i.e., $R[i] = r$, where $i \in \{1, \cdots, Q\}$. By using different settings of the length of the incomplete preference list, Fig. 2 shows the selection results of the two-level stable matching-based selection mechanism. From this figure, we find that the diversity of the selected solutions increase with the decrease of r; in the meanwhile, the improvement of the diversity is at the expense of the convergence. It is interesting to note that the two-level stable matching-based selection mechanism totally degenerates into the original stable matching-based selection mechanism shown in Fig. 1(b) when using $r = 4$. In a word, r controls the trade-off between convergence and diversity in the two-level stable matching-based selection. In the next subsection, we develop an adaptive mechanism to control the setting of R on the fly.

3.3 Adaptive Mechanism

To better understand the proposed adaptive mechanism, here we introduce the concept of *local competitiveness*. At first, all solutions are associated with their closest subproblems having the shortest perpendicular distance between the objective vector of the solution and the weight vector of the subproblem. Afterwards, for each subproblem having more than one associated solutions, we choose the one, which holds the best aggregation function value, as its representative solution. A solution is defined as a *locally competitive* solution in case it dominates at least one representative solution of its $\ell \geq 1$ nearest subproblems; otherwise, it is defined as a *locally noncompetitive* solution. In view of the population dynamics of the evolutionary process, we develop an adaptive mechanism to set the length of the incomplete preference list of a solution according to its local competitiveness. Generally speaking, the length of a solution's incomplete preference list is set as the maximum ℓ that can keep that solution locally noncompetitive.

Algorithm 5 provides the pseudo code of our proposed adaptive mechanism for setting the length of the incomplete preference list on the fly. The specific procedure is illustrated as follows:

- **Step 1)** Given Q solutions and N subproblems, each solution is associated with its closest subproblem as shown in line 1 and line 2. In particular, $\Phi[i]$ represents the index of the subproblem with which the solution $\mathbf{x}^i$ is associated, $i \in \{1, \cdots, Q\}$.

- **Step 2)** Collect the associated solutions of each subproblem p^j, $j \in \{1, \cdots, N\}$, to form a temporary set χ (line 4). If there exists any associated solution, determine the representative solution of p^j as the associated solution in χ that optimizes the objective of p^j (line 5 to line 8). $\varphi[j]$ represents the index of the representative solution.

- **Step 3)** For each solution $\mathbf{x}^i$, the length of the incomplete preference list is set as the maximum ℓ which keeps the solution locally noncompetitive among the neighboring representative solutions. More specifically, ℓ is gradually increased until $\mathbf{x}^i$ becomes locally competitive (line 10 to line 16).

On one hand, for a m-objective MOP, each solution locates within the space between m closest neighboring weight vectors in the objective space. In other words, this solution can be associated with any of these m subproblems in principle. On the other hand, it is usually desirable to keep the solution's preference list within a relatively small length as discussed in Section 3.2. To avoid a large number of unnecessary computations, here $R[i]$ is originally set to be m and gradually increases until it reaches a threshold ℓ_{max}. In particular, ℓ_{max} is set as the neighborhood size T used in MOEA/D.

Let us use the example shown in Fig. 1(b) to explain the underlying principle of our proposed adaptive mechanism. In this example, solutions $\mathbf{x}^2$ and $\mathbf{x}^3$ become locally competitive when $\ell > 1$; while solutions $\mathbf{x}^7$ and $\mathbf{x}^9$ are locally noncompetitive for all ℓ settings. It is worth noting that neither $\mathbf{x}^2$ nor $\mathbf{x}^3$ is the representative solution of any subproblem; in the meanwhile, they are crowded in a narrow area. Since these locally competitive solutions have a better rank in the preference lists than those locally noncompetitive ones, the original stable matching-based selection always give them a higher priority to form the stable matching pairs. However, this selection result is obviously harmful for the population diversity. In addition, we also notice that $\mathbf{x}^7$ and $\mathbf{x}^9$ are the current representative solutions of p^3 and p^4, thus they should contain some relevant information for optimizing these subproblems. In contrast, although $\mathbf{x}^2$ and $\mathbf{x}^3$ have a better aggregation function values, they are far away from p^3 and p^4 and should be less relevant to them. To resolve these issues, our proposed adaptive mechanism adaptively restricts the length of the preference list of each solution $\mathbf{x}$ by removing subproblems whose representative solution is dominated by $\mathbf{x}$. By these means, we can make sure that each solution does not consider a subproblem, which prefers this solutions to its own representative one, in its preference list. As a consequence, each subproblem is prevented from matching with a less relevant solution at the first-level stable matching. Note that this adaptive

Algorithm 5: AdaptiveSetR($\mathsf{P}, \mathsf{S}, \Psi_\mathsf{X}$)

Input:

- subproblem set P and solution set S
- set of solutions' preference lists Ψ_X

Output: lengths of solution's preference list R

1 **for** $i \leftarrow 1$ **to** Q **do**
2     $\Phi[i] \leftarrow \Psi_X[i][1]$;
3 **for** $j \leftarrow 1$ **to** N **do**
4     $\chi \leftarrow \{i|\Phi[i] = j,\ i \in 1, 2, ..., Q\}$;
5   **if** $\chi = \emptyset$ **then**
6         $\varphi[j] \leftarrow -1$;
7   **else**
8         $\varphi[j] \leftarrow \arg\min_{i \in \chi} g^{tch}(\mathbf{x}^i|\lambda^j, \mathbf{z}^*)$;
9 **for** $i \leftarrow 1$ **to** Q **do**
10     $R[i] \leftarrow m$;
11   **for** $\ell \leftarrow m + 1$ **to** ℓ_{max} **do**
12         $t \leftarrow \varphi[\Psi_X[i][\ell]]$;
13   **if** $t \neq -1$ **then**
14   **if** $\mathbf{x}^i \prec \mathbf{x}^t$ **then**
15   break;
16         $R[i] \leftarrow \ell$;
17 **return** R;

mechanism can be readily plugged into the two-level stable matching-based selection mechanism. This can be simply implemented by replacing line 2 of Algorithm 4 by Algorithm 5, and the final adaptive two-level stable matching-based selection mechanism is denoted as ASTM for short.

3.4 Time Complexity of ASTM

In this subsection, we analyze the time complexity of ASTM. During the first level stable matching, the calculation of $\Delta_P(p, \mathbf{x})$ and $\Delta_X(\mathbf{x}, p)$ cost $\mathcal{O}(NQlogQ)$ computations [13]. In Algorithm 5, the association operation between subproblems and solutions costs $\mathcal{O}(Q)$ calculations, i.e., line 1 and line 2. As for line 3 to line 8 of Algorithm 5, the identification of the representative solution for each subproblem requires $\mathcal{O}(mNQ)$ computations. Thereafter, the computation of R in line 9 to line 16 of Algorithm 5 costs $\mathcal{O}(mQ(\ell_{max} - m))$ computations in the worst case. Back to the Algorithm 4, the complexity of the stable matching with the incomplete lists in line 3 is $\mathcal{O}(NQ)$ [25]. Next, the complexity of line 4 to line 6 of Algorithm 4 is $\mathcal{O}(N + Q)$. During the second-level stable matching, i.e., line 7 to line 10, same complexity analysis can be done for the remaining subproblems and solutions. Overall, the total complexity is $\mathcal{O}(\max(NQlogQ, mQ(\ell_{max} - m)))$.

3.5 Incorporation of ASTM with MOEA/D

Similar to [13], we choose the MOEA/D-DRA [24] as the baseline framework and replace the STM selection mechanism of MOEA/D-STM by the ASTM selection mechanism developed in Section 3.3. The resulted algorithm is denoted as MOEA/D-ASTM, and its pseudo code is given in Algorithm 6.

Note that the normalization scheme proposed in [28] is adopted to handle MOPs with different scales of objectives. In the following paragraphs, some important components of MOEA/D-ASTM are further illustrated.

3.5.1 Initialization

Without any prior knowledge of the landscape, the initial population $S = \{\mathbf{x}^1, \cdots, \mathbf{x}^N\}$ is randomly sampled from Ω. Same as the original MOEA/D, we use the classic method suggested in [29] to generate a set of uniformly distributed weight vectors on a unit simplex, i.e., $W = \{\mathbf{w}^1, \cdots, \mathbf{w}^N\}$. In particular, each weight vector represents a subproblem. In addition, for each weight vector $\mathbf{w}^i$, $i \in \{1, \cdots, N\}$, we assign its T, $1 \leq T \leq N$, nearest weight vectors as its neighbors.

3.5.2 Reproduction

Reproduction operator is used to generate an offspring solution from its parent solutions. Any genetic operator or mathematical programming technique can serve this purpose. In MOEA/D-ASTM, we use the differential evolution (DE) operator [30] and polynomial mutation [31] as in [12]. More specifically, let $\mathbf{x}^{r_1}$, $\mathbf{x}^{r_2}$, and $\mathbf{x}^{r_3}$ be three parent solutions, an offspring solution $\overline{\mathbf{x}}^i = (\overline{x}^i_1, \cdots, \overline{x}^i_m)$ is generated as follows:

$$u^i_j = \begin{cases} x^{r_1}_j + F \times (x^{r_2}_j - x^{r_3}_j) & \text{if } rand < CR \text{ } or \text{ } j = j_{rand} \\ x^i_j & \text{otherwise} \end{cases} \tag{7}$$

where $j \in \{1, \cdots, m\}$, $rand \in [0, 1]$, CR and F are two control parameters and j_{rand} is a random integer uniformly chosen between 1 and m. Then, the polynomial mutation is applied on each $\mathbf{u}^i$ to generate $\overline{\mathbf{x}}^i$:

$$\overline{x}^i_j = \begin{cases} u^i_j + \sigma_j \times (b_j - a_j) & \text{if } rand < p_m \\ u^i_j & \text{otherwise} \end{cases} \tag{8}$$

with

$$\sigma_j = \begin{cases} (2 \times rand)^{\frac{1}{\eta+1}} - 1 & \text{if } rand < 0.5 \\ 1 - (2 - 2 \times rand)^{\frac{1}{\eta+1}} & \text{otherwise} \end{cases}, \tag{9}$$

where the distribution index η and the 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. For simplicity, the violated decision variable is set to its nearer boundary value.

3.5.3 Utility of Subproblem

The utility of subproblem p^i, denoted as π^i, where $i \in \{1, \cdots, N\}$, measures the improvement rate of p^i. It is formally defined as [24]:

$$\pi^i = \begin{cases} 1 & \text{if } \Delta^i > 0.001 \\ (0.95 + 0.05 \times \frac{\Delta^i}{0.001}) \times \pi^i & \text{otherwise} \end{cases} \tag{10}$$

where Δ^i represents the relative decrease of the scalar objective value of p^i, and it is evaluated as:

$$\Delta^i = \frac{g(\mathbf{x}^{i,old}|\mathbf{w}^i, \mathbf{z}^*) - g(\mathbf{x}^{i,new}|\mathbf{w}^i, \mathbf{z}^*)}{g(\mathbf{x}^{i,old}|\mathbf{w}^i, \mathbf{z}^*)} \tag{11}$$

where $\mathbf{x}^{i,old}$ is the solution of p^i in the previous generation.

Algorithm 6: MOEA/D–ASTM

```
Input: algorithm parameters
Output: final population S
Initialize the population S, a set of weight vectors W and its neighborhood structure;
Set neval ← 0, iteration ← 0;
while Stopping criterion is not satisfied do
    Select the current active subproblems to form I;
    for each i ∈ I do
        if uniform(0,1) < δ then
            E ← B(i);
        else
            E ← P;
        x^{r1} ← x^i, x^{r2} and x^{r3} ← two randomly selected solutions from E;
        Generate a candidate x̄ by using the method described in Section 3.5.2, S ← S ⋃ {x̄};
        Evaluate F(x̄), neval++;
    S ← SelectionSTM2L(P, S);
    iteration++;
    if mod(iteration, 30) = 0 then
        Update the utility of each subproblem;
return S;
```

4 Experimental Settings

This section presents the general setup of our empirical studies, including the benchmark problems, comparative algorithms, parameter settings and performance metrics.

4.1 Benchmark Problems

Three popular benchmark suites, i.e., UF [32], MOP [15] and WFG [33], 62 problem instances in total, are chosen as the benchmark set in our empirical studies. These problem instances have various characteristics, e.g., non-convexity, deceptive, multi-modality. According to the recommendations in the original references, the number of decision variables is set to as: $n = 30$ for the UF instances and $n = 10$ for the MOP instances. As the WFG instances are scalable to any number of objectives, here we consider $m \in \{2, 3, 5, 8, 10\}$ separately. In particular, when $m = 2$, $n = k + l$ [16], where the position-related variable $k = 2$ and the distance-related variable $l = 4$; while for $m > 2$, we use the recommended settings in [28] and [18], i.e., $k = 2 \times (m - 1)$ and $l = 20$.

4.2 Comparative Algorithms

Nine state-of-the-art EMO algorithms, i.e., MOEA/D-STM, MOEA/D-IR [16], gMOEA/D-AGR [17], MOEA/D-M2M [15], MOEA/D-DRA, HypE [9], NSGA-III [28], PICEA-g [34] and MOEA/DD [18], are considered in our empirical studies. In particular, the first seven algorithms are used for comparative studies on problems with complicated PSs; while the latter five are chosen to investigate the scalability on problems with more than three objectives. The following paragraphs briefly introduce the characteristics of different algorithms.

- *MOEA/D-IR*: its selection process is guided by the interrelationship between subproblems and solutions, which is built upon their mutual preferences. In particular, to further emphasize the population diversity, the niche count is added into the solution's preference setting.

- *gMOEA/D-AGR*: by specifying a particular replacement neighborhood as a set of closest solutions, it finds the appropriate solution for each subproblem. In particular, it employs a sigmoid function to dynamically control the size of the replacement neighborhood for all subproblems.
- *MOEA/D-M2M*: it decomposes the original MOP into a number of simplified MOPs and uses several independent NSGA-II procedures to solve these simplified MOPs in a collaborative manner.
- *MOEA/D-DRA*: it improves the original MOEA/D with a dynamically allocation of computation resources. For the problems with complicated PSs, we use TCH aggregation function and the reproduction operators are the same as introduced in Section 3.5.2. For problems with more than three objectives, we use the penalty-based boundary intersection (PBI) aggregation function in view of reported competitive performance [18, 28, 35]. In the meanwhile, were place the DE operator with the simulated binary crossover (SBX) [36] operator.
- *HypE*: it uses the individual hypervolume contribution to assign fitness to each solution. To speed up the hypervolume calculation, it uses the Monte Carlo simulation to estimate this contribution when $m > 5$.
- *NSGA-III*: it is an improved NSGA-II for solving problems with more than three objectives. Instead of the crowding distance used in NSGA-II, it uses a set of uniformly distributed reference points to adjust the population density. In particular, solutions associated with a less crowded reference point have a higher priority to survive to the next generation.
- *PICEA-g*: it co-evolves a family of decision maker's preference weight vectors, sampled in the objective space, together with a population of candidate solutions for solving problems with more than three objectives. It assigns fitness values to the preference weight vectors according to the number of satisfied solutions; while the fitness values of solutions are evaluated as the number of satisfied preference weight vectors.
- *MOEA/DD*: considering the complementary effects of Pareto- and decomposition-based techniques, it combines them in a single paradigm. In addition, to further emphasize the population diversity, it gives the dominated solutions lying on a less crowded area a second chance to survive to the next generation.

4.3 Parameter Settings

The settings of the population size for different benchmark problems are shown in Table 1. The stopping condition of each algorithm is the predefined number of function evaluations. In particular, it is set to 300,000 for the UF and MOP instances, and 25,000 for the bi-objective WFG instances. As for the many-objective WFG instances, i.e., $m \in \{3, 5, 8, 10\}$, the number of function evaluations is set as 36,400, 157,500, 234,000 and 550,000, respectively. The parameters of the comparative algorithms are set the same as their original references. The parameters of our proposed MOEA/D-ASTM are set as follows:

- *Reproduction operators*: As for problems with complicated properties, we use the reproduction operators introduced in Section 3.5.2 for offspring generation. In particular, as for the DE operator, we set $CR = 1.0$ and $F = 0.5$ for the UF and MOP instances; while $CR = F = 0.5$ for bi-objective WFG instances. We set the mutation probability as $p_m = \frac{1}{n}$ and the distribution index as $\mu_m = 20$ in the polynomial mutation. As for problems with more than three objectives, we use the SBX operator to replace the DE operator, where the crossover probability $p_c = 1$ and its distribution index $\mu_c = 30$.

Table 1: Settings of Population Size

Benchmark Problem	m	Population Size
UF1 to UF7	2	600
UF8 to UF10	3	1,000
MOP1 to MOP5	2	100
MOP6 to MOP7	3	300
WFG1 to WFG9	2	250
WFG1 to WFG9	3	91
WFG1 to WFG9	5	210
WFG1 to WFG9	8	156
WFG1 to WFG9	10	275

- *Neighborhood size*: $T = 20$.
- *Probability used to select in the neighborhood*: $\delta = 0.9$.

4.4 Performance metrics

To assess the performance of different algorithms, we choose the following two widely used performance metrics:

1. *Inverted Generational Distance* (IGD) [37]: Given P^* as a set of points uniformly sampled along the PF and P as the set of solutions obtained from an EMO algorithm. The IGD value of P is calculated as:
$$IGD(P, P^*) = \frac{\sum_{\mathbf{z}\in P^*} dist(\mathbf{z}, P)}{|P^*|}, \tag{12}$$
where $dist(\mathbf{z}, P)$ is the Euclidean distance of $\mathbf{z}$ to its nearest point in P.

2. *Hypervolume* (HV) [6]: Let $\mathbf{z}^r = (z_1^r, \cdots, z_m^r)^T$ be a point dominated by all the Pareto optimal objective vectors. The HV of P is defined as the volume of the objective space dominated by the solutions of P and bounded by $\mathbf{z}^r$:
$$HV(P) = \mathsf{VOL}(\bigcup_{\mathbf{z}\in P} [z_1, z_1^r] \times \cdots \times [z_m, z_m^r]), \tag{13}$$
where VOL indicates the Lebesgue measure.

Since the objective functions of WFG instances are in different scales, we normalize their PFs and the obtained solutions to $[0, 1]$ before calculating the performance metrics. In this case, we constantly set $\mathbf{z}^r = (1.2, \cdots, 1.2)^T$ in the HV calculation. Note that both IGD and HV can evaluate the convergence and diversity simultaneously. A smaller IGD value or a large HV value indicates a better approximation to the PF. Each algorithm is independently run 21 times. The mean and standard deviation of the IGD and HV values are presented in the corresponding tables. The best metric values are highlighted in bold face with a gray background. To have a statistically sound conclusion, we use the Wilcoxon's rank sum test at the 5% significant level to evaluate whether the proposed MOEA/D-ASTM is significantly better or worse than the others.

5 Empirical Studies

In this section, we first present the comparative results for problems with complicated properties. Afterwards, we show the results on problems with more than three objectives. At the end, we empirically investigate the effectiveness of the adaptive mechanism.

Table 2: IGD Results on MOP Test Instances.

Problem	IGD	MOEA/D-DRA	MOEA/D-STM	MOEA/D-IR	gMOEA/D-AGR	MOEA/D-M2M	NSGA-III	HypE	MOEA/D-ASTM
MOP1	Mean	3.330E-1	3.519E-1	4.773E-2	3.216E-2	**1.617E-2**	3.654E-1	8.019E-1	2.354E-2
	Std	6.508E-2	1.604E-2	3.444E-3	9.992E-3	**4.505E-4**	3.293E-3	1.092E-2	2.748E-3
	Rank	5 −	6 −	4 −	3 −	1 +	7 −	8 −	2
MOP2	Mean	2.800E-1	3.105E-1	2.863E-2	6.318E-2	**1.097E-2**	3.405E-1	5.980E-1	1.167E-2
	Std	6.871E-2	6.780E-2	1.864E-2	6.904E-2	**2.239E-3**	1.600E-2	2.155E-1	2.296E-2
	Rank	5 −	6 −	3 −	4 −	1 −	7 −	8 −	2
MOP3	Mean	5.015E-1	4.841E-1	3.991E-2	6.091E-2	**1.319E-2**	3.869E-1	6.202E-1	2.776E-2
	Std	2.964E-2	2.951E-2	1.368E-2	8.337E-2	**4.598E-3**	9.289E-17	1.712E-1	6.291E-2
	Rank	7 −	6 −	3 −	4 −	1 −	5 −	8 −	2
MOP4	Mean	3.087E-1	3.179E-1	4.069E-2	4.439E-2	**7.766E-3**	3.131E-1	7.099E-1	2.813E-2
	Std	2.754E-2	6.404E-3	4.040E-2	4.503E-2	**5.995E-4**	2.182E-2	1.077E-2	2.321E-2
	Rank	5 −	7 −	3 −	4 ≈	1 +	6 −	8 −	2
MOP5	Mean	3.174E-1	3.152E-1	5.587E-2	2.440E-2	2.147E-2	2.866E-1	1.024E+00	**2.024E-2**
	Std	7.897E-3	8.444E-3	1.969E-3	4.429E-3	8.993E-4	2.750E-2	2.255E-1	**1.478E-3**
	Rank	7 −	6 −	4 −	3 −	2 −	5 −	8 −	1
MOP6	Mean	3.061E-1	3.027E-1	1.163E-1	7.796E-2	8.570E-2	3.070E-1	5.745E-1	**6.385E-2**
	Std	2.062E-8	1.482E-2	9.058E-3	9.694E-3	3.703E-3	2.014E-4	1.588E-2	**3.535E-3**
	Rank	6 −	5 −	4 −	2 −	3 −	7 −	8 −	1
MOP7	Mean	3.512E-1	3.512E-1	1.765E-1	2.527E-1	1.177E-1	3.520E-1	6.374E-1	**1.091E-1**
	Std	2.002E-7	1.659E-7	1.183E-2	2.905E-2	8.498E-3	8.005E-4	9.289E-3	**1.326E-2**
	Rank	5 −	6 −	3 −	4 −	2 −	7 −	8 −	1
	Total Rank	40	42	24	24	11	44	56	11
	Final Rank	5	6	3	3	1	7	8	1

According to Wilcoxon's rank sum test, +, − and ≈ indicate that the corresponding EMO algorithm is significantly better than, worse than or similar to MOEA/D-ASTM.

Table 3: HV Results on MOP Test Instances.

Problem	HV	MOEA/D-DRA	MOEA/D-STM	MOEA/D-IR	gMOEA/D-AGR	MOEA/D-M2M	NSGA-III	HypE	MOEA/D-ASTM
MOP1	Mean	0.572	0.539	1.026	1.062	**1.080**	0.515	0.292	1.072
	Std	1.210E-1	3.318E-2	5.769E-3	1.169E-2	**8.839E-4**	8.855E-3	1.401E-2	3.503E-3
	Rank	5 −	6 −	4 −	3 −	1 +	7 −	8 −	2
MOP2	Mean	0.478	0.466	0.722	0.690	0.755	0.447	0.320	**0.756**
	Std	4.157E-2	4.375E-2	2.312E-2	8.592E-2	2.996E-3	1.086E-2	9.798E-2	**2.487E-2**
	Rank	5 −	6 −	3 −	4 −	2 −	7 −	8 −	1
MOP3	Mean	0.240	0.240	0.596	0.573	**0.636**	0.440	0.310	0.613
	Std	0.000E+00	0.000E+00	2.507E-2	1.132E-1	**4.616E-3**	2.220E-16	9.539E-2	9.177E-2
	Rank	7 −	8 −	3 −	4 −	1 −	5 −	6 −	2
MOP4	Mean	0.574	0.576	0.913	0.903	**0.945**	0.569	0.338	0.924
	Std	2.062E-2	1.352E-2	5.699E-2	5.807E-2	**1.723E-3**	1.183E-2	1.143E-2	3.291E-2
	Rank	6 −	5 −	3 −	4 ≈	1 +	7 −	8 −	2
MOP5	Mean	0.635	0.635	1.007	1.067	1.068	0.649	0.059	**1.075**
	Std	6.923E-9	4.542E-6	8.047E-3	8.612E-3	1.637E-3	2.874E-2	1.781E-1	**2.490E-3**
	Rank	7 −	6 −	4 −	3 −	2 −	5 −	8 −	1
MOP6	Mean	1.221	1.226	1.417	1.465	1.438	1.209	0.683	**1.483**
	Std	4.259E-7	2.279E-2	1.887E-2	1.676E-2	1.191E-2	2.041E-3	3.441E-2	**5.283E-3**
	Rank	6 −	5 −	4 −	2 −	3 −	7 −	8 −	1
MOP7	Mean	0.939	0.939	1.035	0.994	1.048	0.926	0.538	**1.064**
	Std	1.357E-6	1.309E-6	2.824E-2	4.818E-2	2.322E-2	1.447E-3	6.301E-3	**3.196E-2**
	Rank	5	6	3	4	2	7	8	1
	Total Rank	41	42	24	24	12	45	54	10
	Final Rank	6	7	3	3	2	5	8	1

According to Wilcoxon's rank sum test, +, − and ≈ indicate that the corresponding EMO algorithm is significantly better than, worse than or similar to MOEA/D-ASTM.

5.1 Performance Comparisons on MOP Instances

As discussed in [15], MOP benchmark suite, in which different parts of the PF have various difficulties, poses significant challenges for maintaining the population diversity. From the results shown in Table 2 and Table 3, we can see that MOEA/D-ASTM has shown the overall best performance. It obtains the significantly better results in 88 out of 98 comparisons according to the Wilcoxon's rank sum test. In particular, MOEA/D-ASTM has obtained the best IGD and HV values on MOP5 to MOP7 instances. The second best algorithm for MOP benchmark suite is MOEA/D-M2M. Especially for the IGD metric, MOEA/D-M2M has obtained the best IGD values on MOP1 to MOP4 instances. Following MOEA/D-ASTM and MOEA/D-M2M, MOEA/D-IR and gMOEA/D-AGR are able to obtain a set of non-dominated solutions having a moderate coverage accross the PF. As for MOEA/D-DRA, MOEA/D-STM, NSGA-III and HypE, they can only obtain some solutions lying on the boundaries.

Due to the page limit, we present the plots of the final solution sets with the best IGD values on all test instances in the supplemental file. From the plots, we can see that although the IGD and HV values obtained by MOEA/D-ASTM and MOEA/D-M2M are comparable, the solutions obtained by MOEA/D-ASTM have a more uniform distribution along the PF. This can be explained as the density estimation method, i.e., the crowding distance of NSGA-II, used in MOEA/D-M2M is too coarse to guarantee a uniform distribution. Nevertheless, the convergence achieved by MOEA/D-M2M is satisfied thus contributes to its promising IGD values. According to [17], gMOEA/D-AGR uses a sigmoid function to assign the same replacement neighborhood size to each subproblem. However, since different parts of the PF require various efforts, this same setting might not be appropriate. From these plots, we can see that the solutions obtained by gMOEA/D-AGR may miss some portions of the PF. This can be explained as the replacement neighborhood size grows too large for the corresponding subproblems. In order to emphasize the population diversity, for each subproblem, MOEA/D-IR selects the appropriate solution from a couple of related ones. However, its preference setting, which encourages the selection in a less crowded area, tends to result in a unstable selection result. In this case, some solutions far away from the PF can be selected occasionally. The reason behind the poor performance of NSGA-III, HypE and MOEA/D-DRA is their *convergence first and diversity second* selection strategy, which can easily trap the population in some narrow areas. As discussed in Section 2.3, the stable matching model used in MOEA/D-STM can easily match a solution with a unfavorable subproblem, thus results in a unbalanced selection.

5.2 Performance Comparisons on UF Instances

The comparison results on the IGD and HV metrics between MOEA/D-ASTM and the other EMO algorithms on UF benchmark suite are presented in Table 4 and Table 5. Different from the MOP benchmark suite, the major source of difficulty for the UF benchmark suite is not the diversity preservation but the complicated PS. Generally speaking, the overall performance of MOEA/D-ASTM ranks the second on the UF benchmark suite, just after its predecessor, i.e., MOEA/D-STM. More specifically, MOEA/D-ASTM performs the best on UF2 and UF3 instances and acts as the top three algorithm on UF1 and UF6 to UF9 instances. For UF1 and UF6, the performance of MOEA/D-ASTM does not show significant difference with MOEA/D-STM.

According to the performance of different algorithms on the UF problem instances, the analysis can be divided into three groups. For the UF4 and UF5 instances, MOEA/D-M2M, NSGA-III and HypE are able to provide a better performance than all other MOEA/D variants. All these three algorithms use the Pareto dominance as the major driving force in the environmental selection, which can improve the convergence to a great extent. For the UF1 to UF3 and UF6 to UF9, the MOEA/D variants outperform NSGA-III and HypE. In particular, MOEA/D-ASTM and MOEA/D-STM have shown very promising results on these six problem instances. The superior performance can be attributed to the well balance between convergence and diversity achieved by the stable matching relationship between subproblems and solutions. gMOEA/D-AGR has shown a medium performance for these two groups of problem instances. This might be because its adaptive mechanism can hardly make a satisfied prediction of the size of replacement neighborhood. UF10 is a difficult tri-objective problem, where none of these eight EMO algorithms are able to obtain a well approximation to the PF within the given number of function evaluations. Nevertheless, it is worth noting that the empirical studies in [13] demonstrate that the stable matching-based selection mechanism can offer a competitive result in case we double the maximum number of function evaluations.

Table 4: IGD Results on UF Test Instances.

Problem	IGD	MOEA/D-DRA	MOEA/D-STM	MOEA/D-IR	gMOEA/D-AGR	MOEA/D-M2M	NSGA-III	HypE	MOEA/D-ASTM
UF1	Mean	1.145E-3	**1.016E-3**	2.518E-3	1.823E-3	7.416E-3	9.620E-2	9.883E-2	1.033E-3
	Std	3.793E-4	**8.295E-5**	1.145E-4	1.011E-4	2.621E-3	1.266E-2	1.073E-2	5.942E-5
	Rank	3 ≈	1 ≈	5 −	4 −	6 −	7 −	8 −	2
UF2	Mean	3.059E-3	3.083E-3	5.422E-3	5.158E-3	3.950E-3	2.988E-2	2.090E-1	**2.976E-3**
	Std	2.041E-3	6.021E-4	1.097E-3	5.681E-4	4.642E-4	2.528E-3	6.061E-2	**1.129E-3**
	Rank	2 ≈	3 ≈	6 −	5 −	4 −	7 −	8 −	1
UF3	Mean	1.398E-2	6.031E-3	1.682E-2	7.863E-3	1.525E-2	2.141E-1	1.828E-1	5.126E-3
	Std	1.234E-2	5.794E-3	1.254E-2	7.149E-3	5.403E-3	4.700E-2	5.041E-2	4.190E-3
	Rank	4 −	2 ≈	6 −	3 ≈	5 −	8 −	7 −	1
UF4	Mean	5.346E-2	5.080E-2	5.602E-2	5.008E-2	**3.989E-2**	4.313E-2	4.884E-2	5.301E-2
	Std	3.019E-3	2.962E-3	2.976E-3	2.688E-3	**3.912E-4**	9.580E-4	6.817E-3	3.364E-3
	Rank	7 ≈	5 +	8 −	4 +	1 +	2 +	3 +	6
UF5	Mean	2.932E-1	2.397E-1	2.524E-1	2.315E-1	**1.816E-1**	2.113E-1	2.309E-1	2.521E-1
	Std	5.716E-2	1.938E-2	4.496E-2	4.452E-2	**2.620E-2**	2.146E-2	4.700E-2	3.114E-2
	Rank	8 −	5 ≈	7 ≈	4 ≈	1 +	2 +	3 ≈	6
UF6	Mean	1.710E-1	**7.509E-2**	1.078E-1	1.295E-1	8.358E-2	2.119E-1	2.332E-1	8.042E-2
	Std	1.650E-1	**4.245E-2**	4.784E-2	9.631E-2	4.746E-2	5.081E-2	7.273E-2	4.229E-2
	Rank	6 −	1 ≈	4 −	5 −	3 ≈	7 −	8 −	2
UF7	Mean	1.156E-3	**1.137E-3**	3.515E-3	2.115E-3	6.138E-3	6.526E-2	2.606E-1	1.372E-3
	Std	1.542E-4	**7.598E-5**	3.141E-4	3.209E-4	2.073E-3	8.587E-2	4.905E-2	1.899E-4
	Rank	2 +	1 +	5 −	4 −	6 −	7 −	8 −	3
UF8	Mean	3.121E-2	**3.052E-2**	6.457E-2	4.766E-2	9.800E-2	1.669E-1	3.107E-1	4.262E-2
	Std	4.219E-3	**5.057E-3**	8.971E-3	7.736E-3	1.122E-2	3.169E-3	3.335E-2	1.071E-2
	Rank	2 +	1 +	5 −	4 ≈	6 −	7 −	8 −	3
UF9	Mean	4.575E-2	**2.351E-2**	5.931E-2	5.529E-2	1.159E-1	1.735E-1	2.337E-1	4.236E-2
	Std	3.163E-2	**1.156E-3**	4.003E-2	4.366E-2	2.618E-2	2.817E-2	3.157E-2	3.967E-2
	Rank	3 −	1 +	5 −	4 −	6 −	7 −	8 −	2
UF10	Mean	5.343E-1	1.693E+00	7.398E-1	3.993E-1	5.567E-1	**2.279E-1**	2.618E-1	2.361E+00
	Std	3.886E-2	2.538E-1	1.262E-1	7.512E-2	5.802E-2	**5.410E-2**	7.152E-2	3.517E-1
	Rank	4 +	7 +	6 +	3 +	5 +	1 +	2 +	8
	Total Rank	41	27	57	40	43	55	63	34
	Final Rank	4	1	7	3	5	6	8	2

According to Wilcoxon's rank sum test, +, − and ≈ indicate that the corresponding EMO algorithm is significantly better than, worse than or similar to MOEA/D-ASTM.

5.3 Performance Comparisons on Bi-Objective WFG Instances

From the comparison results shown in Table 6 and Table 7, we find that MOEA/D-ASTM has shown the best overall performance on the bi-objective WFG instances, where it shows significantly better metric values in 103 out of 126 comparisons. In particular, it performs the best on WFG1, WFG6, WFG7 and WFG9 and obtains very promising results on WFG2 to WFG5 instances. It is interesting to note that MOEA/D-ASTM is significantly better than MOEA/D-STM on all WFG instances except for WFG8. One possible reason is the proper normalization method used in MOEA/D-ASTM. However, we also notice that the performance of NSGA-III fluctuates significantly on difficult problem instances, although it uses the same normalization method. MOEA/D-DRA and gMOEA/D-AGR have also shown competitive results to our proposed MOEA/D-ASTM.

5.4 Performance Comparisons on Many-Objective Problems

In order to further investigate the scalability of MOEA/D-ASTM, we choose the widely used WFG benchmark suite for empirical studies. As introduced in Section 4.3, we set the number of objectives to be 3, 5, 8 and 10, respectively. Due to the lack of appropriate sample points from the true PF, we only consider the HV metric in this study. Table 8 to Table 11 present the experimental results for 3-, 5-, 8- and 10-objective WFG test instances, respectively. For the 3-objective WFG instances, MOEA/D-ASTM performs significantly better in 36 out of 45 comparisons and loses in one comparison on WFG3. In particular, MOEA/D-ASTM has shown the best metric values on WFG1, WFG4 to WFG9 and the second best metric values on WFG2 and WFG3. NSGA-III, which obtains the second final rank, is outperformed by MOEA/D-ASTM on all test instances. MOEA/D-

Table 5: HV Results on UF Test Instances.

Problem	HV	MOEA/D-DRA	MOEA/D-STM	MOEA/D-IR	gMOEA/D-AGR	MOEA/D-M2M	NSGA-III	HypE	MOEA/D-ASTM
UF1	Mean	1.104	1.104	1.100	1.102	1.092	0.942	0.942	**1.104**
	Std	7.489E-4	4.088E-4	7.228E-4	3.883E-4	4.450E-3	3.068E-2	2.734E-2	**2.826E-4**
	Rank	3 −	2 ≈	5 −	4 −	6 −	8 −	7 −	1
UF2	Mean	1.099	1.100	1.093	1.097	1.099	1.054	0.892	**1.100**
	Std	3.766E-3	1.610E-3	4.580E-3	1.731E-3	1.961E-3	6.282E-3	4.373E-2	**2.412E-3**
	Rank	3 ≈	2 ≈	6 −	5 −	4 −	7 −	8 −	1
UF3	Mean	1.080	1.096	1.073	1.090	1.080	0.733	0.792	**1.098**
	Std	2.411E-2	9.063E-3	2.525E-2	1.718E-2	8.565E-3	5.360E-2	5.973E-2	**6.330E-3**
	Rank	5 −	2 ≈	6 −	3 ≈	4 −	8 −	7 −	1
UF4	Mean	0.672	0.678	0.667	0.681	**0.701**	0.698	0.685	0.674
	Std	5.628E-3	5.683E-3	5.823E-3	4.854E-3	**7.129E-4**	1.514E-3	1.403E-2	6.660E-3
	Rank	7 ≈	5 ≈	8 −	4 +	1 +	2 +	3 +	6
UF5	Mean	0.353	0.435	0.428	0.486	**0.573**	0.533	0.515	0.411
	Std	8.553E-2	4.441E-2	8.378E-2	8.540E-2	**5.384E-2**	4.230E-2	7.918E-2	5.590E-2
	Rank	8 −	5 ≈	6 ≈	4 +	1 +	2 +	3 +	7
UF6	Mean	0.577	0.647	0.597	0.645	**0.695**	0.622	0.587	0.644
	Std	1.266E-1	8.573E-2	9.178E-2	6.842E-2	**4.125E-2**	1.266E-2	7.213E-2	8.245E-2
	Rank	8 −	2 ≈	6 −	3 ≈	1 +	5 −	7 −	4
UF7	Mean	0.937	**0.937**	0.932	0.935	0.928	0.841	0.612	0.936
	Std	5.940E-4	**4.484E-4**	1.347E-3	1.801E-3	4.040E-3	9.773E-2	3.422E-2	7.995E-4
	Rank	2 ≈	1 +	5 −	4 −	6 −	7 −	8 −	3
UF8	Mean	**1.125**	1.125	1.051	1.086	0.932	0.780	0.783	1.094
	Std	**1.134E-2**	9.501E-3	2.365E-2	1.858E-2	3.199E-2	4.299E-3	3.824E-3	2.890E-2
	Rank	1 +	2 +	5 −	4 ≈	6 −	8 −	7 −	3
UF9	Mean	1.406	**1.463**	1.397	1.400	1.240	0.978	0.859	1.435
	Std	5.947E-2	**3.960E-3**	7.633E-2	8.590E-2	5.519E-2	5.214E-2	8.712E-2	8.144E-2
	Rank	3 −	1 ≈	5 −	4 −	6 −	7 −	8 −	2
UF10	Mean	0.173	0.000	0.059	0.321	0.167	**0.644**	0.602	0.000
	Std	3.752E-2	0.000E+00	5.012E-2	7.042E-2	2.181E-2	**1.132E-1**	1.321E-1	0.000E+00
	Rank	4 +	7 ≈	6 +	3 +	5 +	1 +	2 +	7
	Total Rank	44	29	58	38	40	55	60	35
	Final Rank	5	1	7	3	4	6	8	2

According to Wilcoxon's rank sum test, +, − and ≈ indicate that the corresponding EMO algorithm is significantly better than, worse than or similar to MOEA/D-ASTM.

DRA and HypE are the worst algorithms in this comparison, which might be due to the lack of normalization scheme. In contrast, PICEA-g and MOEA/DD, which does not apply normalization, performs a little better. For the WFG instances with 5 objectives, MOEA/D-ASTM keeps being the best algorithm on all problem instances except for WFG3 and achieves an overall best ranking. In terms of the Wilcoxon's rank sum test, MOEA/D-ASTM wins 39 out of 45 comparisons and shows similar results in 5 comparisons. Followed by NSGA-III and MOEA/DD, PICEA-g ranks just after the MOEA/D-ASTM. When the number of objectives increases to 8 and 10, MOEA/D-ASTM keeps its top position and it shows significantly better metric values in 72 out 90 comparisons. On the WFG1 and WFG4 to WFG9, MOEA/D-ASTM is always the best algorithm and significantly outperforms the others except for the 8-objective WFG6, where it is outperformed by NSGA-III.

5.5 Effectiveness of the Adaptive Mechanism

In order to investigate the effectiveness of the adaptive mechanism proposed in Section 3.3, we choose MOP1 as the problem instance and plot the trajectories of the r values of the selected solutions of 4 different subproblems, i.e., p_1, p_{34}, p_{67} and p_{100}, respectively. From trajectories shown in Fig. 3, we notice that the r value fluctuates significantly at the early stages of the evolution. Afterwards, it almost convergences to the threshold ℓ_{max}. This is because the local competitiveness varies dramatically when the population is away from but is also heading to the PF. With the progress of evolution, the selected solutions gradually become non-dominated from each other. As a consequence, the value of ℓ, which keeps the solution locally noncompetitive, grows and finally settles at ℓ_{max}. All in all, we can see that different solutions have different local competitiveness, thus it is meaningful to have a different r.

Table 6: IGD Results on Bi-Objective WFG Test Instances.

Problem	IGD	MOEA/D-DRA	MOEA/D-STM	MOEA/D-IR	gMOEA/D-AGR	MOEA/D-M2M	NSGA-III	HypE	MOEA/D-ASTM
WFG1	Mean	3.979E-1	4.091E-1	3.787E-1	3.879E-1	4.281E-1	4.739E-1	5.621E-1	**3.676E-1**
	Std	1.071E-2	6.783E-3	9.590E-3	1.227E-2	9.944E-3	7.885E-2	6.618E-2	**6.908E-3**
	Rank	4 −	5 −	2 −	3 −	6 −	7 −	8 −	1
WFG2	Mean	1.012E-2	1.025E-2	1.258E-2	1.017E-2	8.314E-3	**5.405E-3**	1.583E-2	6.554E-3
	Std	4.172E-4	2.275E-4	1.020E-3	1.525E-4	5.260E-4	**3.830E-4**	1.945E-2	4.590E-4
	Rank	4 −	6 −	7 −	5 −	3 −	1 +	8 −	2
WFG3	Mean	4.166E-2	**4.165E-2**	4.325E-2	4.167E-2	4.287E-2	4.176E-2	5.277E-2	4.166E-2
	Std	2.074E-5	**3.239E-5**	7.357E-4	2.010E-5	1.309E-4	7.884E-5	5.071E-3	8.429E-6
	Rank	3 ≈	1 ≈	7 −	4 ≈	6 −	5 −	8 −	2
WFG4	Mean	7.026E-3	1.078E-2	1.036E-2	7.007E-3	8.806E-3	**4.541E-3**	1.887E-2	4.757E-3
	Std	7.323E-4	1.019E-3	8.423E-4	4.930E-4	7.462E-4	**3.971E-5**	7.177E-3	1.189E-4
	Rank	4 −	7 −	6 −	3 −	5 −	1 +	8 −	2
WFG5	Mean	2.769E-2	2.773E-2	2.883E-2	2.767E-2	2.935E-2	**2.711E-2**	6.565E-2	2.735E-2
	Std	3.409E-5	3.998E-5	5.872E-4	1.367E-5	2.530E-4	**1.118E-3**	1.516E-2	1.325E-5
	Rank	4 −	5 −	6 −	3 −	7 −	1 −	8 −	2
WFG6	Mean	4.942E-3	4.946E-3	1.098E-2	4.960E-3	7.745E-3	1.164E-2	3.508E-2	**4.537E-3**
	Std	3.026E-5	4.153E-5	7.247E-4	2.451E-5	2.448E-4	7.786E-3	1.720E-2	**1.845E-5**
	Rank	2 −	3 −	6 −	4 −	5 −	7 −	8 −	1
WFG7	Mean	5.240E-3	5.239E-3	1.150E-2	5.267E-3	9.254E-3	4.771E-3	3.104E-2	**4.765E-3**
	Std	3.507E-5	2.977E-5	1.312E-3	2.427E-5	7.293E-4	1.513E-5	1.052E-2	**1.965E-5**
	Rank	4 −	3 −	7 −	5 −	6 −	2 ≈	8 −	1
WFG8	Mean	4.194E-2	4.388E-2	4.745E-2	3.074E-2	**1.249E-2**	6.431E-2	8.961E-2	4.289E-2
	Std	2.643E-2	2.519E-2	2.492E-2	2.638E-2	**9.623E-4**	1.928E-2	1.921E-2	2.698E-2
	Rank	3 ≈	5 ≈	6 ≈	2 +	1 ≈	7 −	8 −	4
WFG9	Mean	5.887E-3	6.611E-3	1.339E-2	5.625E-3	8.508E-3	5.344E-3	1.795E-2	**5.072E-3**
	Std	4.904E-4	4.503E-4	6.869E-4	3.111E-4	4.532E-4	3.408E-4	4.192E-3	**1.616E-4**
	Rank	4 −	5 −	7 −	3 −	6 −	2 −	8 −	1
	Total Rank	32	40	54	32	45	33	72	16
	Final Rank	2	5	7	2	6	4	8	1

According to Wilcoxon's rank sum test, +, − and ≈ indicate that the corresponding EMO algorithm is significantly better than, worse than or similar to MOEA/D-ASTM.

Table 7: HV Results on Bi-Objective WFG Test Instances.

Problem	HV	MOEA/D-DRA	MOEA/D-STM	MOEA/D-IR	gMOEA/D-AGR	MOEA/D-M2M	NSGA-III	HypE	MOEA/D-ASTM
WFG1	Mean	0.445	0.432	0.465	0.459	0.425	0.383	0.301	**0.486**
	Std	1.292E-2	7.023E-3	1.218E-2	1.550E-2	1.005E-2	7.566E-2	5.573E-2	**8.758E-3**
	Rank	4 −	5 −	2 −	3 −	6 −	7 −	8 −	1
WFG2	Mean	1.004	1.005	0.992	**1.006**	1.004	0.981	0.967	1.003
	Std	6.694E-3	8.531E-4	1.219E-2	**5.550E-4**	7.100E-4	2.080E-2	2.323E-2	9.182E-3
	Rank	4 ≈	2 ≈	6 −	1 +	3 ≈	7 ≈	8 −	5
WFG3	Mean	0.814	0.814	0.811	0.814	0.812	0.815	0.799	**0.815**
	Std	2.181E-4	1.814E-4	8.074E-4	2.011E-4	4.852E-4	2.948E-4	5.227E-3	**1.832E-4**
	Rank	3 −	5 −	7 −	4 −	6 −	2 −	8 −	1
WFG4	Mean	0.657	0.647	0.648	0.660	0.658	**0.666**	0.646	0.664
	Std	2.648E-3	2.901E-3	2.345E-3	2.880E-3	1.253E-3	**8.246E-4**	5.173E-3	1.109E-3
	Rank	5 −	7 −	6 −	3 −	4 −	1 +	8 −	2
WFG5	Mean	0.625	0.625	0.623	0.625	0.622	**0.630**	0.578	0.627
	Std	2.392E-4	4.255E-4	1.993E-3	1.041E-4	4.153E-4	**3.254E-3**	1.428E-2	1.096E-3
	Rank	5 −	4 −	6 −	3 −	7 −	1 +	8 −	2
WFG6	Mean	0.649	0.648	0.638	0.649	0.644	0.632	0.604	**0.649**
	Std	1.679E-4	2.973E-4	1.096E-3	2.069E-4	5.245E-4	1.467E-2	2.018E-2	**2.621E-4**
	Rank	2 −	4 −	6 −	3 −	5 −	7 −	8 −	1
WFG7	Mean	0.649	0.648	0.637	0.649	0.643	0.649	0.615	**0.650**
	Std	1.797E-4	2.562E-4	1.840E-3	1.351E-4	5.031E-4	5.094E-4	8.120E-3	**1.551E-4**
	Rank	4 −	5 −	7 −	3 −	6 −	2 −	8 −	1
WFG8	Mean	0.551	0.545	0.532	0.576	**0.626**	0.488	0.436	0.545
	Std	5.679E-2	5.333E-2	5.764E-2	5.387E-2	**4.984E-3**	4.463E-2	3.869E-2	6.089E-2
	Rank	3 ≈	5 ≈	6 ≈	2 +	1 +	7 −	8 −	4
WFG9	Mean	0.693	0.691	0.680	0.695	0.690	0.694	0.676	**0.696**
	Std	1.288E-3	1.212E-3	1.165E-3	1.183E-3	1.003E-3	1.500E-3	4.219E-3	**8.589E-4**
	Rank	4 −	5 −	7 −	2 −	6 −	3 −	8 −	1
	Total Rank	34	42	53	24	44	37	72	18
	Final Rank	3	5	7	2	6	4	8	1

According to Wilcoxon's rank sum test, +, − and ≈ indicate that the corresponding EMO algorithm is significantly better than, worse than or similar to MOEA/D-ASTM.

Table 8: HV Results on 3-Objective WFG Test Instances.

Problem	HV	MOEA/D-DRA	HypE	PICEA-g	NSGA-III	MOEA/DD	MOEA/D-ASTM
WFG1	Mean	0.849	0.331	0.772	0.539	0.805	**1.172**
	Std	5.971E-2	3.463E-2	1.013E-2	2.906E-2	5.968E-2	**5.437E-2**
	Rank	2 −	6 −	4 −	5 −	3 −	1
WFG2	Mean	1.293	1.363	**1.524**	1.484	1.483	1.492
	Std	7.078E-2	7.103E-2	**1.102E-1**	1.128E-1	1.176E-1	1.099E-1
	Rank	6 −	5 −	1 ≈	3 −	4 −	2
WFG3	Mean	1.041	1.026	**1.180**	1.147	1.087	1.153
	Std	4.840E-2	2.779E-2	**8.709E-3**	8.340E-3	1.311E-2	1.074E-2
	Rank	5 −	6 −	1 +	3 −	4 −	2
WFG4	Mean	1.050	1.013	1.101	1.103	1.100	**1.120**
	Std	7.363E-3	3.688E-2	6.978E-3	4.861E-3	3.877E-3	**5.930E-3**
	Rank	5 −	6 −	3 −	2 −	4 −	1
WFG5	Mean	1.021	0.969	1.033	1.050	1.039	**1.067**
	Std	5.433E-3	1.644E-2	8.101E-3	6.615E-3	4.654E-3	**4.661E-3**
	Rank	5 −	6 −	4 −	2 −	3 −	1
WFG6	Mean	1.007	0.953	1.062	1.060	1.054	**1.063**
	Std	2.143E-2	3.097E-2	9.001E-3	1.085E-2	9.550E-3	**1.067E-2**
	Rank	5 −	6 −	2 ≈	3 ≈	4 −	1
WFG7	Mean	1.011	0.953	1.116	1.117	1.109	**1.118**
	Std	3.725E-2	5.344E-2	3.238E-3	4.014E-3	1.530E-2	**3.488E-3**
	Rank	5 −	6 −	3 ≈	2 ≈	4 −	1
WFG8	Mean	0.949	0.842	0.972	0.989	0.986	**1.010**
	Std	1.252E-2	1.980E-2	9.594E-3	6.461E-3	7.113E-3	**3.448E-3**
	Rank	5 −	6 −	4 −	2 −	3 −	1
WFG9	Mean	0.923	0.861	0.981	1.003	1.002	**1.003**
	Std	2.295E-2	1.055E-1	2.927E-2	3.207E-2	3.399E-2	**4.385E-2**
	Rank	5 −	6 −	4 ≈	2 ≈	3 ≈	1
	Total Rank	43	53	26	24	32	11
	Final Rank	5	6	3	2	4	1

According to Wilcoxon's rank sum test, +, − and ≈ indicate that the corresponding EMO algorithm is significantly better than, worse than or similar to MOEA/D-ASTM.

Table 9: HV Results on 5-Objective WFG Test Instances.

Problem	HV	MOEA/D-DRA	HypE	PICEA-g	NSGA-III	MOEA/DD	MOEA/D-ASTM
WFG1	Mean	1.951	0.647	1.569	0.822	1.548	**2.272**
	Std	1.037E-1	1.999E-2	6.634E-2	3.185E-2	9.575E-2	**2.758E-2**
	Rank	2 −	6 −	3 −	5 −	4 −	1
WFG2	Mean	2.014	2.275	2.459	2.391	2.381	**2.480**
	Std	1.275E-1	1.423E-1	8.909E-2	1.417E-1	8.185E-2	**2.720E-3**
	Rank	6 −	5 −	2 ≈	3 −	4 −	1
WFG3	Mean	1.497	1.544	**1.781**	1.667	1.562	1.732
	Std	7.279E-2	2.823E-2	**1.062E-2**	1.522E-2	2.942E-2	1.516E-2
	Rank	6 −	5 −	1 +	3 −	4 −	2
WFG4	Mean	1.912	1.600	2.107	2.090	2.098	**2.161**
	Std	4.762E-2	7.751E-2	4.662E-2	7.334E-3	8.677E-3	**9.305E-3**
	Rank	5 −	6 −	2 −	4 −	3 −	1
WFG5	Mean	1.906	1.683	1.998	2.021	1.976	**2.063**
	Std	3.757E-2	8.138E-2	9.244E-3	6.035E-3	6.404E-3	**3.563E-3**
	Rank	5 −	6 −	3 −	2 −	4 −	1
WFG6	Mean	1.688	1.716	2.050	2.034	2.012	**2.052**
	Std	8.944E-2	6.873E-2	1.149E-2	9.940E-3	1.504E-2	**2.464E-2**
	Rank	6 −	5 −	2 ≈	3 −	4 −	1
WFG7	Mean	1.831	1.561	2.142	2.139	2.121	**2.169**
	Std	2.168E-2	1.171E-1	4.246E-3	5.516E-3	5.781E-3	**6.074E-3**
	Rank	5 −	6 −	2 −	3 −	4 −	1
WFG8	Mean	1.637	1.401	1.894	1.896	1.909	**1.914**
	Std	2.625E-1	7.628E-2	1.181E-2	4.913E-3	2.949E-2	**1.982E-2**
	Rank	5 −	6 −	4 −	3 −	2 −	1
WFG9	Mean	1.748	1.293	1.828	1.839	1.850	**1.897**
	Std	1.045E-1	1.758E-1	6.895E-3	3.775E-2	4.905E-2	**8.762E-2**
	Rank	5 −	6 −	4 ≈	3 ≈	2 ≈	1
	Total Rank	45	51	23	29	31	10
	Final Rank	5	6	2	3	4	1

According to Wilcoxon's rank sum test, +, − and ≈ indicate that the corresponding EMO algorithm is significantly better than, worse than or similar to MOEA/D-ASTM.

Table 10: HV Results on 8-Objective WFG Test Instances.

Problem	HV	MOEA/D-DRA	HypE	PICEA-g	NSGA-III	MOEA/DD	MOEA/D-ASTM
WFG1	Mean	3.364	1.076	3.318	1.213	3.075	**3.970**
	Std	2.230E-1	2.878E-2	1.718E-1	7.567E-2	1.655E-1	**6.899E-2**
	Rank	2 −	6 −	3 −	5 −	4 −	1
WFG2	Mean	3.416	3.818	4.013	**4.087**	3.998	3.963
	Std	2.305E-1	2.716E-1	3.373E-1	**3.025E-1**	1.871E-1	3.534E-1
	Rank	6 −	5 ≈	2 ≈	1 +	3 − ≈	4
WFG3	Mean	1.730	2.626	**2.981**	2.451	2.422	1.610
	Std	9.584E-2	7.088E-2	**7.869E-2**	1.559E-1	3.714E-2	1.025E-1
	Rank	5 +	2 +	1 +	3 +	4 +	6
WFG4	Mean	2.310	2.341	2.879	3.976	3.863	**4.075**
	Std	1.875E-1	2.006E-1	1.338E-2	1.435E-2	3.520E-2	**2.169E-2**
	Rank	6 −	5 −	4 −	2 −	3 −	1
WFG5	Mean	2.654	2.203	3.178	3.844	3.508	**3.845**
	Std	1.192E-1	2.976E-1	8.347E-2	8.222E-3	4.535E-2	**2.658E-2**
	Rank	5 −	6 −	4 −	2 ≈	3 −	1
WFG6	Mean	1.716	2.756	3.759	**3.883**	3.743	3.852
	Std	2.549E-1	1.532E-1	1.448E-1	**3.473E-2**	4.639E-2	4.409E-2
	Rank	6 −	5 −	3 −	1 +	4 −	2
WFG7	Mean	1.980	2.360	3.492	4.067	3.966	**4.079**
	Std	1.461E-1	1.485E-1	1.568E-1	9.422E-3	1.286E-2	**1.183E-2**
	Rank	6 −	5 −	4 −	2 −	3 −	1
WFG8	Mean	0.758	2.387	3.390	3.582	3.420	**3.662**
	Std	4.423E-1	2.087E-1	7.153E-2	3.265E-2	1.592E-1	**5.689E-2**
	Rank	6 −	5 −	4 −	2 −	3 −	1
WFG9	Mean	1.684	1.787	3.210	3.358	3.182	**3.524**
	Std	5.129E-1	3.649E-1	6.811E-2	9.816E-2	7.667E-2	**8.167E-2**
	Rank	6 −	5 −	3 −	2 −	4 −	1
	Total Rank	48	44	28	20	31	18
	Final Rank	6	5	3	2	4	1

According to Wilcoxon's rank sum test, +, − and ≈ indicate that the corresponding EMO algorithm is significantly better than, worse than or similar to MOEA/D-ASTM.

Table 11: HV Results on 10-Objective WFG Test Instances.

Problem	HV	MOEA/D-DRA	HypE	PICEA-g	NSGA-III	MOEA/DD	MOEA/D-ASTM
WFG1	Mean	4.341	1.574	5.660	1.867	5.287	**5.914**
	Std	3.530E-1	4.671E-2	2.302E-1	7.955E-2	1.273E-1	**1.486E-2**
	Rank	4 −	6 −	2 −	5 −	3 −	1
WFG2	Mean	5.193	5.621	6.088	**6.098**	5.918	5.663
	Std	3.276E-1	3.935E-1	2.814E-1	**2.201E-1**	2.801E-2	5.161E-1
	Rank	6 −	5 ≈	2 ≈	1 +	3 ≈	4 ≈
WFG3	Mean	1.577	3.941	**4.363**	4.046	3.434	2.298
	Std	2.198E-1	1.033E-1	**6.347E-2**	4.316E-1	8.074E-2	1.370E-1
	Rank	6 −	3 +	1 +	2 +	4 +	5
WFG4	Mean	3.150	3.706	4.425	5.919	5.543	**6.100**
	Std	2.214E-1	5.209E-1	3.142E-1	2.469E-2	5.042E-2	**5.694E-3**
	Rank	6 −	5 −	4 −	2 −	3 −	1
WFG5	Mean	3.559	3.841	4.753	5.692	5.060	**5.733**
	Std	3.018E-1	3.537E-1	1.674E-1	1.041E-2	9.347E-2	**6.724E-3**
	Rank	6 −	5 −	4 −	2 −	3 −	1
WFG6	Mean	2.418	4.204	5.375	5.735	5.349	**5.757**
	Std	3.890E-1	2.564E-1	2.283E-1	4.735E-2	9.130E-2	**5.643E-2**
	Rank	6 −	5 −	3 −	2 ≈	4 −	1
WFG7	Mean	2.813	4.193	5.317	6.037	5.779	**6.101**
	Std	3.709E-2	3.993E-1	1.891E-1	7.994E-3	1.378E-2	**3.216E-3**
	Rank	6 −	5 −	4 −	2 −	3 −	1
WFG8	Mean	0.878	4.031	5.054	5.477	5.056	**5.681**
	Std	2.695E-1	3.062E-1	1.884E-1	4.496E-2	3.086E-1	**5.365E-2**
	Rank	6 −	5 −	4 −	2 −	3 −	1
WFG9	Mean	2.283	2.828	4.760	4.961	4.562	**5.355**
	Std	6.462E-1	3.267E-1	6.239E-2	1.293E-1	1.054E-1	**1.124E-1**
	Rank	6 −	5 −	3 −	2 −	4 −	1
	Total Rank	52	44	27	20	30	16
	Final Rank	6	5	3	2	4	1

According to Wilcoxon's rank sum test, +, − and ≈ indicate that the corresponding EMO algorithm is significantly better than, worse than or similar to MOEA/D-ASTM.

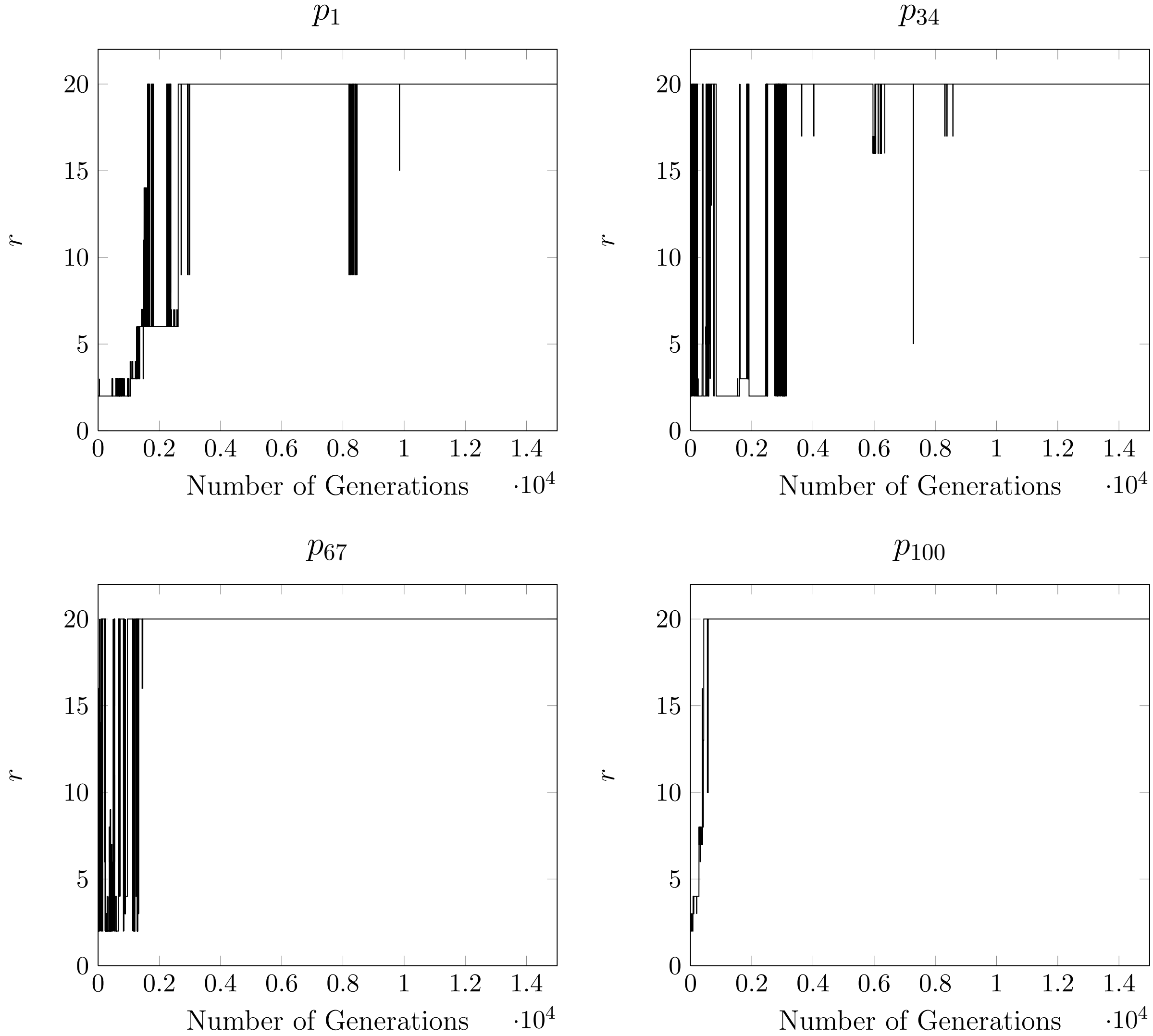


Figure 3: The illustration of the r values of the solutions selected by subproblems p_1, p_{34}, p_{67} and p_{100}.

In order to further investigate the effectiveness brought by our proposed adaptive mechanism, we develop a variant denoted as MOEA/D-ASTM-v in which all solutions share the same r setting. From the our offline parameter studies, we finally find that $r = 4$ and $r = 8$ are the best settings for MOP and UF benchmark suites respectively. Thereafter, in Table 12, we show the comparison results of MOEA/D-ASTM and MOEA/D-ASTM-v with the best r settings in terms of the IGD metric. From Table 12, we can see that MOEA/D-ASTM obtains better IGD values than MOEA/D-ASTM-v on 11 out of 17 problem instances. Note that the optimal r settings are obtained from a series of comprehensive try-and-error experiments, which are not as intelligent and flexible as our proposed adaptive mechanism. In addition, this experiment also demonstrate the importance of using different r values for different subproblems.

Table 12: Comparison Results on IGD Metric obtained by MOEA/D-ASTM and MOEA/D-ASTM-v.

IGD	MOEA/D-ASTM-v			MOEA/D-ASTM	
Problem	Mean	Std	Test	Mean	Std
MOP1	2.363E-2	2.159E-3	$\approx$	**2.354E-2**	**2.748E-3**
MOP2	4.389E-2	6.923E-2	$-$	**1.167E-2**	**2.296E-2**
MOP3	9.019E-2	1.133E-1	$-$	**2.776E-2**	**6.291E-2**
MOP4	3.596E-2	3.847E-2	$\approx$	**2.813E-2**	**2.321E-2**
MOP5	2.062E-2	2.000E-3	$\approx$	**2.024E-2**	**1.478E-3**
MOP6	**5.586E-2**	**2.705E-3**	$+$	6.385E-2	3.535E-3
MOP7	**7.617E-2**	**1.444E-3**	$+$	1.091E-1	1.326E-2
UF1	1.036E-3	6.294E-5	$\approx$	**1.033E-3**	**5.942E-5**
UF2	3.068E-3	1.389E-3	$\approx$	**2.976E-3**	**1.129E-3**
UF3	5.847E-3	5.365E-3	$\approx$	**5.126E-3**	**4.190E-3**
UF4	**5.245E-2**	**3.161E-3**	$\approx$	5.301E-2	3.364E-3
UF5	2.536E-1	2.141E-2	$\approx$	**2.521E-1**	**3.114E-2**
UF6	8.379E-2	4.294E-2	$\approx$	**8.042E-2**	**4.229E-2**
UF7	1.394E-3	3.194E-4	$\approx$	**1.372E-3**	**1.899E-4**
UF8	**4.069E-2**	**1.122E-2**	$\approx$	4.262E-2	1.071E-2
UF9	**2.692E-2**	**2.917E-3**	$\approx$	4.236E-2	3.967E-2
UF10	**2.140E+00**	**3.765E-1**	$\approx$	2.361E+00	3.517E-1

According to Wilcoxon's rank sum test, $+$, $-$ and $\approx$ indicate that the corresponding EMO algorithm is significantly better than, worse than or similar to MOEA/D-ASTM.

Table 13: Final Ranks of Mean IGD Metric Values on MOP, UF and Bi-Objective WFG Test Instances.

IGD	MOEA/D-DRA	MOEA/D-STM	MOEA/D-IR	gMOEA/D-AGR	MOEA/D-M2M	NSGA-III	HypE	MOEA/D-ASTM
Total Rank	113	109	135	96	99	132	191	61
Final Rank	5	4	7	2	3	6	8	1

Table 14: Final Ranks of Mean HV Metric Values on MOP, UF and Bi-Objective WFG Test Instances.

HV	MOEA/D-DRA	MOEA/D-STM	MOEA/D-IR	gMOEA/D-AGR	MOEA/D-M2M	NSGA-III	HypE	MOEA/D-ASTM
Total Rank	119	113	135	86	96	137	186	63
Final Rank	5	4	6	2	3	7	8	1

Table 15: Final Ranks of Mean HV Metric Values on 3-, 5-, 8- and 10-Objective WFG Test Instances.

HV	MOEA/D-DRA	HypE	PICEA-g	NSGA-III	MOEA/DD	MOEA/D-ASTM
Total Rank	188	192	104	93	124	55
Final Rank	5	6	3	2	4	1

6 Conclusions

The stable matching-based selection mechanism provides an avenue to address the balance between convergence and diversity from the perspective of achieving the equilibrium between the preferences of subproblems and solutions. However, considering of population diversity, it might not be appropriate to allow each solution to be matched with any subproblem in its preference list. This paper has developed a two-level stable matching-based selection in which the length of the preference list of each solution is restricted so that it is only allowed to match with one of its most favorite subproblems. In addition, the length of the preference list of each solution is problem dependent and is related to the difficulties of the corresponding local space. To address this issue, an adaptive mechanism is proposed to dynamically set the length of the preference list of each solution according to its local competitiveness. Comprehensive experiments are conducted on 62 benchmark problem instances which cover various characteristics, e.g., multi-modality, deceptive, complicated PSs and many objectives. From the experimental results discussed in Section 5 and the summary shown in Table 13 to Table 15, we can clearly observe the competitive performance obtained by our proposed MOEA/D-ASTM, comparing to a variety of state-of-the-art EMO algorithms.

Although our proposed MOEA/D-ASTM has shown very competitive performance in the empirical studies, we also notice that the stable matching relationship between subproblems and solutions may sacrifice the convergence ability of the algorithm to some extent. One possible reason might be the one-to-one matching restricts that each solution can only be selected by at most one subproblem. Future work could be focused on assigning higher priorities for elite solutions to produce offspring solutions or allowing each elite solution to be matched with more than one subproblem. It is also interesting to apply the proposed method to the real-world application scenarios.

References

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

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

[3] 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*, 2002, pp. 95–100.

[4] J. D. Knowles and D. Corne, "Approximating the nondominated front using the pareto archived evolution strategy," *Evolutionary Computation*, vol. 8, no. 2, pp. 149–172, 2000.

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

[6] E. Zitzler and L. Thiele, "Multiobjective evolutionary algorithms: a comparative case study and the strength pareto approach," *IEEE Trans. Evolutionary Computation*, vol. 3, no. 4, pp. 257–271, 1999.

[7] E. Zitzler and S. Künzli, "Indicator-based selection in multiobjective search," in *PPSN'04: Proc. of the 8th International Conference on Parallel Problem Solving from Nature*, 2004, pp. 832–842.

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

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

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

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

[12] H. Li and Q. Zhang, "Multiobjective optimization problems with complicated pareto sets, MOEA/D and NSGA-II," *IEEE Trans. Evolutionary Computation*, vol. 13, no. 2, pp. 284–302, 2009.

[13] K. Li, Q. Zhang, S. Kwong, M. Li, and R. Wang, "Stable matching-based selection in evolutionary multiobjective optimization," *IEEE Trans. Evolutionary Computation*, vol. 18, no. 6, pp. 909–923, 2014.

[14] K. Li, Á. Fialho, S. Kwong, and Q. Zhang, "Adaptive operator selection with bandits for a multiobjective evolutionary algorithm based on decomposition," *IEEE Trans. Evolutionary Computation*, vol. 18, no. 1, pp. 114–130, 2014.

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

[16] K. Li, S. Kwong, Q. Zhang, and K. Deb, "Interrelationship-based selection for decomposition multiobjective optimization," *IEEE Trans. Cybernetics*, vol. 45, no. 10, pp. 2076–2088, 2015.

[17] Z. Wang, Q. Zhang, A. Zhou, M. Gong, and L. Jiao, "Adaptive replacement strategies for MOEA/D," *IEEE Trans. Cybernetics*, vol. 46, no. 2, pp. 474–486, 2016.

[18] K. Li, K. Deb, Q. Zhang, and S. Kwong, "An evolutionary many-objective optimization algorithm based on dominance and decomposition," *IEEE Trans. Evolutionary Computation*, vol. 19, no. 5, pp. 694–716, 2015.

[19] X. Cai, Y. Li, Z. Fan, and Q. Zhang, "An external archive guided multiobjective evolutionary algorithm based on decomposition for combinatorial optimization," *IEEE Trans. Evolutionary Computation*, vol. 19, no. 4, pp. 508–523, 2015.

[20] K. Li, S. Kwong, and K. Deb, "A dual-population paradigm for evolutionary multiobjective optimization," *Inf. Sci.*, vol. 309, pp. 50–72, 2015.

[21] Q. Liu, G. J. Mailath, A. Postlewaite, and L. Samuelson, "Stable matching with incomplete information," *Econometrica*, vol. 82, no. 2, pp. 541–587, 2014.

[22] M. Wu, S. Kwong, Q. Zhang, K. Li, R. Wang, and B. Liu, "Two-level stable matching-based selection in MOEA/D," in *SMC'15: Proc. of the 2015 IEEE International Conference on Systems, Man, and Cybernetics*, 2015, pp. 1720–1725.

[23] K. Miettinen, *Nonlinear Multiobjective Optimization.* Kluwer Academic Publishers, 1999, vol. 12.

[24] Q. Zhang, W. Liu, and H. Li, "The performance of a new version of MOEA/D on CEC09 unconstrained MOP test instances," in *Proceedings of the IEEE Congress on Evolutionary Computation, CEC 2009, Trondheim, Norway, 18-21 May, 2009*, 2009, pp. 203–208.

[25] D. Gale and L. S. Shapley, "College admissions and the stability of marriage," *The American Mathematical Monthly*, vol. 120, no. 5, pp. 386–391, 1962.

[26] D. Gusfield and R. W. Irving, *The Stable marriage problem - structure and algorithms*, ser. Foundations of computing series. MIT Press, 1989.

[27] K. Iwama, D. Manlove, S. Miyazaki, and Y. Morita, "Stable marriage with incomplete lists and ties," in *Automata, Languages and Programming, 26th International Colloquium, ICALP'99, Prague, Czech Republic, July 11-15, 1999, Proceedings*, 1999, pp. 443–452.

[28] K. Deb and H. Jain, "An evolutionary many-objective optimization algorithm using reference-point-based nondominated sorting approach, part I: solving problems with box constraints," *IEEE Trans. Evolutionary Computation*, vol. 18, no. 4, pp. 577–601, 2014.

[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, pp. 631–657, 1998.

[30] S. Das and P. N. Suganthan, "Differential evolution: A survey of the state-of-the-art," *IEEE Trans. on Evolutionary Computation*, vol. 15, no. 1, pp. 4–31, 2011.

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

[32] 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, Colchester, UK and Nanyang technological University, Singapore, special session on performance assessment of multi-objective optimization algorithms, technical report*, vol. 264, 2008.

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

[34] R. Wang, R. C. Purshouse, and P. J. Fleming, "Preference-inspired coevolutionary algorithms for many-objective optimization," *IEEE Trans. Evolutionary Computation*, vol. 17, no. 4, pp. 474–494, 2013.

[35] H. Ishibuchi, N. Akedo, and Y. Nojima, "Behavior of multiobjective evolutionary algorithms on many-objective knapsack problems," *IEEE Trans. Evolutionary Computation*, vol. 19, no. 2, pp. 264–283, 2015.

[36] R. B. Agrawal, K. Deb, and R. B. Agrawal, "Simulated binary crossover for continuous search space," 1994.

[37] P. A. N. Bosman and D. Thierens, "The balance between proximity and diversity in multi-objective evolutionary algorithms," *IEEE Trans. Evolutionary Computation*, vol. 7, no. 2, pp. 174–188, 2003.