

---

# A Meta-Objective Approach for Many-Objective Evolutionary Optimization

**Dunwei Gong**

School of Information and Control Engineering, China University of Mining and Technology, Xuzhou 221116, China

**Yiping Liu** \*

yiping0liu@gmail.com

School of Information and Control Engineering, China University of Mining and Technology, Xuzhou 221116, China

Department of Computer Science and Intelligent Systems, Graduate School of Engineering, Osaka Prefecture University, Sakai 599-8531, Japan

**Gary G. Yen**

gyen@okstate.edu

School of Electrical and Computer Engineering, Oklahoma State University, Stillwater, OK 74078, USA

---

## Abstract

Pareto-based multi-objective evolutionary algorithms experience grand challenges in solving many-objective optimization problems due to their inability of maintaining both convergence and diversity in a high-dimensional objective space. Existing approaches usually modify the selection criteria to overcome this issue. Different from them, we propose a novel meta-objective (MeO) approach that transforms the many-objective optimization problems in which the new optimization problems become easier to solve by the Pareto-based algorithms. MeO converts a given many-objective optimization problem into a new one, which has the same Pareto optimal solutions and the number of objectives with the original one. Each meta-objective in the new problem consists of two components which measure the convergence and diversity performances of a solution, respectively. Since MeO only converts the problem formulation, it can be readily incorporated within any multi-objective evolutionary algorithms, including those non Pareto-based ones. Particularly, it can boost the Pareto-based algorithms' ability to solve many-objective optimization problems. Due to separately evaluating the convergence and diversity performances of a solution, the traditional density-based selection criteria, e.g., crowding distance, will no longer mistake a solution with poor convergence performance for a solution with low density value. By penalizing a solution in term of its convergence performance in the meta-objective space, the Pareto dominance becomes much more effective for a many-objective optimization problem. Comparative study validates the competitive performance of the proposed meta-objective approach in solving many-objective optimization problems.

## Keywords

Many-objective optimization, evolutionary multi-objective optimization, meta-objective, convergence, diversity.

## 1 Introduction

A great number of real-world applications involves multiple objectives to be optimized simultaneously. These Multi-objective Optimization Problems (MOPs) are commonly

---

\*Corresponding Author

seen in a variety of disciplines, such as industrial scheduling (Han et al., 2015, 2017; Li et al., 2018). The conflict of objectives implies that there is no single optimal solution to an MOP, rather a set of trade-off solutions, known as the Pareto optimal solutions.

Over the past two decades, a large number of Multi-Objective Evolutionary Algorithms (MOEAs) have been proposed. Compared to traditional mathematical programming techniques, MOEAs are particularly suited in searching for the Pareto optimal set in one single run. Generally, the task of MOEAs is to find solutions not only close to the Pareto optimal front (i.e., convergence measure), but also uniformly distributed (i.e., diversity measure).

Among various methods, Pareto-based MOEAs are the most popular ones, e.g., Non-dominated Sorting Genetic Algorithm II (NSGA-II) (Deb et al., 2002) and Strength Pareto Evolutionary Algorithm 2 (SPEA2) (Zitzler et al., 2001), to name a few. These algorithms first adopt the Pareto dominance principle to select non-dominated solutions which are always preferred for good convergence. Then a density-based selection criterion is employed to promote a good diversity among the solutions. Pareto-based MOEAs have been proven successful in solving a large number of MOPs.

However, for MOPs with more than three objectives which are often referred to as Many-objective Optimization Problems (He and Yen, 2016; Liu et al., 2017a, 2018b), the performance of Pareto-based MOEAs generally deteriorates appreciably. One main reason is that the Pareto dominance-based selection criterion becomes very ineffective, since the proportion of non-dominated solutions in the population rises considerably as the number of objectives increases, and then the density-based criterion alone plays a crucial role in environmental and mating selections. However, the density-based criterion often fails in a high-dimensional space, since a solution far away from the Pareto optimal front often has a lower density value.

Intuitively, there are two approaches to improving the performance of Pareto-based MOEAs in solving MaOPs: (1) redefining the dominance relationship (Laumanns et al., 2002; Jaimes et al., 2010; Sato et al., 2011; Yang et al., 2013; He et al., 2014); (2) improving the diversity-based secondary selection criterion or replacing it with another selection criterion (Deb and Jain, 2013; Li et al., 2014; Zhang et al., 2015). These approaches modify either the Pareto dominance or density-based selection criterion, and have shown great potential in many-objective optimization.

In this study, different from the exiting approaches of modifying the selection criteria in Pareto-based MOEAs, we propose a novel Meta-Objective approach (MeO for short) of modifying the many-objective optimization problems. By the proposed method, a many-objective optimization problem is transformed into a new optimization problem, which has the same Pareto optimal solution set with the original one, but is easier to be solved by the Pareto-based MOEAs. As mentioned above, the design goals of Pareto-based MOEAs are to utilize the Pareto dominance and density-based selection criteria to search for solutions with good convergence and diversity, respectively. However, the goals are often mistreated in many-objective optimization, due to the low efficiency of Pareto dominance and the misleading convergence information in density estimation. In our proposed meta-objective approach, these goals become apparent to reach when solving the new optimization problem. To be specific, the new optimization problem has the same number of objectives as the original optimization problem. Each meta-objective in the new optimization problem consists of two components which measure the convergence and diversity performances of a solution, respectively. The convergence component in each meta-objective is the same, and it represents the distance of a solution to the Pareto optimal front. On the other hand, the diversity

components in each meta-objective are different, and represent the positions of the solution according to different orientations. The diversity components lead to conflicts among the meta-objectives. Since all the meta-objective have the same convergence component, a solution closer to the Pareto optimal front, i.e., with better convergence performance, is more likely to dominate the other solution in the meta-objective space than that in the original objective space. For the same reason, the density-based selection criterion can focus on diversity components and will be no longer misguided by the convergence information. Therefore, the Pareto-based MOEAs are better equipped to solve the MaOPs by incorporating MeO.

The main contributions of this work can be summarized as follows:

First, a meta-objective method, termed MeO is proposed, which transforms a given MaOP into a new one. There are two components in each meta-objective in the new MaOP, which measure the convergence and diversity performances of a solution, respectively. Solving the new MaOP by an MOEA leads to a solution set with a good overall performance. This method provides a new idea of managing convergence and diversity in many-objective optimization, and it can be seamlessly integrated into any MOEA, including those of non Pareto-based ones. Moreover, it is computational cheap and does not require a set of pre-defined reference points or vectors.

Next, three alternatives to the convergence component are proposed based on the Pareto rank method, the  $L_p$  scalarizing method, and the double-rank method, respectively, since the true Pareto optimal front is actually unknown. The first two are feasible but ineffective. Therefore, the third one is recommended with the following advantages: 1) it is theoretically consistent with the Pareto dominance criterion; 2) it provides an adequate selection pressure in the high-dimensional objective space; and 3) it is insensitive to the settings of control parameters.

Last but most importantly, MeO provides a general extension of Pareto-based algorithms in many-objective optimization. By integrating MeO, the performances of Pareto-based MOEAs, e.g., NSGA-II and SPEA2, are significantly boosted in solving MaOPs. On one hand, the density-based selection criteria in these algorithms will no longer incorrectly estimate a solution's density value, since the convergence and diversity performances of a solution are separately evaluated in the meta-objective space. On the other hand, the Pareto dominance will become much more effective, since the meta-objective values of a solution are penalized in term of its convergence performance.

The remainder of this paper is organized as below. In Section 2, fundamental definitions in multi-objective optimization are given for the completeness of the presentation and the motivation of this work is also elaborated. The proposed MeO is then described in detail in Section 3. Section 4 demonstrates the effectiveness of MeO by incorporating it into two representative MOEAs. The comparison of MeO with the state-of-the-art MOEAs are given in Section 5. Section 6 concludes the paper and provides pertinent observations and future research directions.

## 2 Preliminaries

### 2.1 Basic Definitions in Multi-Objective Optimization

Without loss of generality, the MOP considered in this study is formulated as follows:

$$\begin{aligned} \min \mathbf{f}(\mathbf{x}) &= \min (f_1(\mathbf{x}), f_2(\mathbf{x}), \dots, f_M(\mathbf{x})) \\ \text{s.t. } \mathbf{x} &\in S \subset \mathbf{R}^n \end{aligned} \quad (1)$$

where  $\mathbf{x}$  represents an  $n$ -dimensional decision vector in space  $S$ .  $f_m(\mathbf{x})$ ,  $m = 1, \dots, M$ , is the  $m$ -th objective to be minimized, and  $M$  is the number of objectives. When  $M > 3$ ,

this problem is known as an MaOP.

In multi-objective optimization, the following concepts have been well defined and widely applied.

**Definition 1. Pareto Dominance:** For any two different solutions to the problem in (1),  $\mathbf{x}_1, \mathbf{x}_2 \in S$ , if  $\forall m = 1, \dots, M, f_m(\mathbf{x}_1) \leq f_m(\mathbf{x}_2)$ , and  $\exists i = 1, \dots, M, f_i(\mathbf{x}_1) < f_i(\mathbf{x}_2)$ , then  $\mathbf{x}_1$  dominates  $\mathbf{x}_2$ , denoted as  $\mathbf{x}_1 \prec \mathbf{x}_2$ .

**Definition 2. Pareto optimal set:** For a solution to the problem in (1),  $\mathbf{x}^* \in S$ , if there is no  $\mathbf{x} \in S$  satisfying  $\mathbf{x} \prec \mathbf{x}^*$ , then  $\mathbf{x}^*$  is the Pareto optimal solution. All such solutions form a set often called the Pareto optimal solution set (*PS*).

**Definition 3. Pareto optimal front:** The image of the Pareto optimal solution set in the objective space is known as the Pareto optimal front (*PF*).

**Definition 4. Ideal point:** The ideal point  $\mathbf{z}^* = (z_1^*, \dots, z_M^*)^T$ , where  $z_m^*$  is the infimum of  $f_m(\mathbf{x}), m = 1, \dots, M$ .

## 2.2 Many-Objective Evolutionary Optimization

As aforementioned in Section 1, many-objective optimization problems have more than three objectives. For the traditional Pareto-based MOEAs, their ability of solving MaOPs are usually improved by the following two approaches. The first is to re-defining the dominance relationship. A large body of researches along this line has been reported, e.g., dominance area control (Sato et al., 2011), Pareto partial dominance (Jaimes et al., 2010),  $\varepsilon$ -dominance (Laumanns et al., 2002), and fuzzy Pareto dominance (He et al., 2014). These modified dominance relationships have been shown to be able to improve the convergence of MOEAs for solving MaOPs. Recently, a Grid-based Evolutionary Algorithm (GrEA) (Yang et al., 2013) was developed based on the concept of  $\varepsilon$ -dominance and showed encouraging performance in many-objective optimization. The second is to improve the diversity-based secondary selection criterion or replacing it with another selection criterion. In [17], a diversity management operator is adopted to meet the requirement on diversity in selection. By shifting the position of a solution in sparse regions, a Shift-based Density Estimation (SDE) was proposed to discard solutions with poor convergence (Li et al., 2014). NSGA-III, an extension of NSGA-II, was developed in (Deb and Jain, 2013), which employs a reference-point-based selection criterion to promote diversity. Recently, a Knee point-driven Evolutionary Algorithm (KnEA) (Zhang et al., 2015) was proposed by preferring local knee solutions among non-dominated ones.

In addition to enhancing the Pareto-based MOEAs, the decomposition-based MOEAs have been found to be a promising alternative to solve MaOPs. A well-regarded representative of decomposition-based algorithms is MOEA/D (Zhang and Li, 2007). In MOEA/D, a set of well distributed reference vectors are first defined. Based on these reference vectors, a given MaOP is aggregated into a number of scalarizing functions. Individuals in a population are guided to search towards the Pareto optimal front in the directions specified by the reference vectors via minimizing the values of the scalarizing functions. MOEA/D has been demonstrated efficient in solving MaOPs. However, one known problem of MOEA/D is that uniformly distributed reference vectors do not necessarily lead to uniformly distributed solutions, particularly for problems with irregular (i.e., nonuniform) Pareto optimal fronts. Several methods to adaptively generating reference vectors have been proposed to address this issue (Qi et al., 2014; Liu et al., 2017b). In addition, Ensemble Fitness Ranking with a Ranking Restriction scheme (EFR-RR) (Yuan et al., 2016) provides another means to achieve uniformly distributed solutions in a high-dimensional objective space, in which only

solutions close to a reference vector in the objective space will be further considered.

Indicator-Based Evolutionary Algorithms (IBEA) (Zitzler and Künzli, 2004) are theoretically well-supported options to the Pareto-based MOEAs. It adopts a performance indicator to account for both convergence and diversity of a solution. Solutions with the best values of the indicator are preferred. Among others, hypervolume is a widely used indicator in IBEAs. Unfortunately, the computational complexity for calculating hypervolume increases exponentially as the number of objectives increases, which makes it computationally prohibitive for MaOPs. To address this issue, HypE (Bader and Zitzler, 2011) uses Monte Carlo simulations to estimate hypervolume, where the accuracy of the estimated hypervolume can be traded off against the available computational resources. Recently, employing the R2 indicator, an improved Many-Objective Metaheuristic Based on the R2 Indicator (MOMBI-II) (Hernández Gómez and Coello Coello, 2015) was proposed to further enhance the efficiency of IBEAs for solving MaOPs.

From the above, we can see that there have been various algorithms developed for solving MaOPs. On the other hand, if we can extract some characteristics from an MaOP and transform it into a new optimization problem, it would be easier for the existing algorithms to obtain the Pareto optimal solutions.

### 2.3 Meta-objective methods in Evolutionary Optimization

The meta-objective approach has been used in evolutionary optimization, where a meta-objective is referred to as a new formulation about the original objective(s) or the characteristic extracted from the optimization problem. In the meta-objective approach, a new optimization problem is formed based on the meta-objective(s), which has the same or similar optimal solutions with the original problem. This means we can solve the original problem by solving the new one, especially when we do not have effective means in solving the original one. For example, in constrained evolutionary optimization (Takahama et al., 2005; Wang and Cai, 2012), the solution's degree of constraint violations is regarded as a meta-objective. Optimizing this new objective can lead to feasible solutions to the original constrained optimization problem. For another example in multimodal evolutionary optimization (Wang et al., 2015), the meta-objectives are constructed based on the original objective and the solutions positions in the variable space. By solving the new multi-objective formed by these meta-objectives, multiple optimal solutions to the original multimodal optimization problem can be found.

Up to now, there exist very few meta methods for many-objective evolutionary optimization. Very recently, a Bi-Goal Evolutionary approach (BiGE) (Li et al., 2015) was proposed for solving MaOPs, in which two meta-objectives are constructed in terms of the convergence and diversity performances of a solution. The meta-objective about convergence is the sum of all objective functions, and the meta-objective about diversity is a share function based on niche technique. The experiments showed that a solution set with a well-balanced performance to the original MaOP can be achieved by optimizing the revised bi-objective optimization problem. However, the necessity of optimizing convergence and diversity as two conflicting objectives should be further investigated. In an ideal situation, only a single solution with both meta-objectives minimized can be obtained. This suggests that the original and new optimization problems do not have the same Pareto optimal set. Therefore, we may not achieve the whole Pareto optimal set of the original optimization problem by optimizing the new optimization problem.

In this study, we propose a novel meta-objective approach for many-objective op-

timization, where the meta-objectives are also constructed based on convergence and diversity. Different from (Li et al., 2015), convergence and diversity are not regarded as conflict objectives, but components in every meta-objective. Each meta-objective has the same convergence component, thus the Pareto optimal solutions to the original optimization problem have great advantages in the meta-objective space. In addition, the diversity components in each meta-objective are different, which leads to conflicts among the meta-objectives. In this way, the new optimization problem has a diverse Pareto optimal set, which is the same with that of the original optimization problem. We will prove this in Subsection 3.1.

## 2.4 Motivation

As mentioned in the last subsection, a meta-objective can be constructed based on the characteristics extracted from the optimization problem. For a multi-objective optimization problem, its characteristics distinguished from a single-objective optimization problem is that not only the convergence but also the diversity performances of solutions should be simultaneously considered by an optimizer. Therefore, constructing meta-objectives in terms of these two performances is considered a natural idea. In this study, we propose a meta-objective approach to convert the given MOP formulated by (1) into a new MOP. When constructing this new MOP, we should make clear that these meta-objectives conflict with each other. If not, the new MOP will be inherently a single-objective optimization problem, and it would be difficult to achieve diverse solutions by solving the new MOP.

Recently, there has been a growing interest in balancing convergence and diversity in multi- and many-objective optimization (Adra and Fleming, 2011; Li et al., 2012, 2015; Yuan et al., 2016; Cheng et al., 2015). Some researchers argue that convergence and diversity are two requirements that conflict with each other. Particularly, these requirements are formulated as a bi-objective optimization problem in (Li et al., 2015). In fact, the rationale for balancing the convergence and diversity performances of solutions is that the selection criteria are involved with both performances. For instance, the PBI (Penalty-Boundary Intersection) function in MOEA/D is composed of  $d_1$  and  $d_2$ , which measure the distances of a solution to the origin (i.e., convergence measure) and to the reference vector (i.e., diversity measure), respectively. Then a solution with a smaller value of  $d_1$  but a large value of  $d_2$  might be preferred. Selecting this solution is not helpful to maintaining diversity in a high-dimensional objective space. Therefore, MOEA/D may not perform well in some MaOPs, since it lacks a strategy of balancing  $d_1$  and  $d_2$ . To overcome this shortcoming in MOEA/D, EFR-RR (Yuan et al., 2016) only considers solutions with the smaller values of  $d_2$ . As mentioned in the last subsection, the meta-objective about convergence in BiGE (Li et al., 2015) is the sum of all the objectives. Since the contour curves of this meta-objective is usually different from the geometric shape of the Pareto optimal front, solely optimizing it results in solutions in specific regions, which means that it is inherently involved with diversity information. This is why the meta-objective about convergence can be traded off with the other meta-objective about diversity. What conflict with each other are the diversity information in both meta-objectives, but not convergence and diversity.

From the above discussions, we notice that whether or not convergence and diversity conflict with each other depends on how to define them. If solutions located exactly on the Pareto optimal front and uniformly distributed are regarded as having good convergence and diversity performances, then these two performances do not conflict with each other. For example, in Figure 1, when searching for an ideal solu-

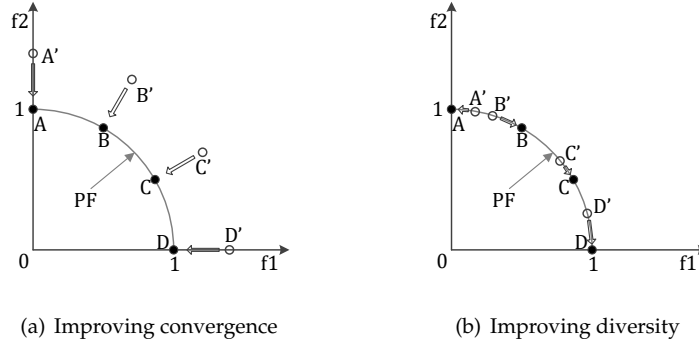


Figure 1: An illustration of searching for the ideal solution set by improving the convergence or diversity performances of a solution set. In this example, solid dots  $A$ - $D$  are the evenly distributed Pareto optimal solutions, and open dots  $A'$ - $D'$  are the solutions to be improved.

tion set, improving the convergence or diversity performance of a solution set does not imply that the other performance should be degraded. In Figure 1.(a), the solutions remain uniformly distributed when approximating to the true Pareto optimal front. Similarly, in Figure 1.(b), diversifying the archived solution set does not result in degeneration in convergence performance. Therefore, convergence and diversity are not necessarily balanced, but simultaneously optimized.

In this study, we suggest that what really conflict with each other are the positions of solutions in the objective space. In (Wang et al., 2015), a multimodal optimization problem is converted into a multi-objective optimization problem based on the positions of solutions in the variable space. The tradeoff between these meta-objectives leads to multiple optimal solutions for the multimodal optimization problem. Hence in this study, we will build the meta-objectives based on the position of a solution in the objective space. Figure 2.(a) provides an illustrative example, where  $v_1(1, 0)$  and  $v_2(0, 1)$  are vectors on the coordinate axes,  $\theta_1$  ( $\theta_2$ ) is the angle between  $v_1$  ( $v_2$ ) and  $f(x)$ , and  $d(x)$  is the distance between  $f(x)$  and the Pareto optimal front in the objective space. It is evident that when  $\theta_1$  becomes large,  $\theta_2$  will decrease, and vice versa. This means that  $\theta_1$  and  $\theta_2$  conflict with each other. Then, a new bi-objective optimization problem can be formulated as follows:

$$\begin{aligned} \min f'(x) &= \min(f'_1(x), f'_2(x)) \\ &= \min(\theta_1, \theta_2) \end{aligned} \quad (2)$$

Balancing these two objectives will lead to solutions well scattered between  $v_1$  and  $v_2$ .

Besides, as discussed above, the convergence performance of a solution, i.e.,  $d(x)$ , should be simultaneously optimized. Integrating (2) with  $d(x)$ , we get:

$$\begin{aligned} \min f'(x) &= \min(f'_1(x), f'_2(x)) \\ &= \min(\theta_1 + \eta \cdot d(x), \\ &\quad \theta_2 + \eta \cdot d(x)) \end{aligned} \quad (3)$$

where  $\eta$  is a scaling factor. Optimizing  $f'_1(x)$  implies to seek solutions not only clos-

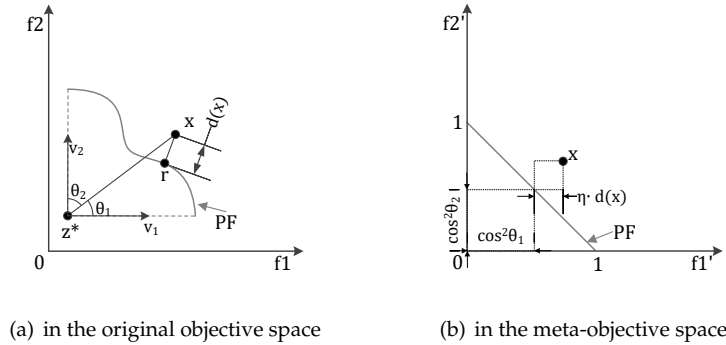


Figure 2: A solution in the original objective and meta-objective spaces, respectively.

est to the  $f_1$ -axis but also the Pareto optimal front, and optimizing  $f'_2(x)$  has a similar meaning. Since both objectives contain  $d(x)$ , optimizing (3) will lead to solutions with  $d(x)$  close to 0. Consequently, a solution set that is almost located on the Pareto optimal front and well distributed can be achieved.

There may exist two issues in the above method. The first is that  $d(x)$  cannot be calculated since the true Pareto optimal front is actually unknown. To address this issue, in Subsection 3.2, we will give three alternative measures based on the Pareto rank, the  $L_p$  scalarizing function, and the double-rank method for estimating the convergence performance of a solution. The first two are feasible but ineffective. Thus we propose the third method by taking the advantages and avoiding the disadvantages of the first two. The other issue is that the new optimization problem has the same number of objectives as the given one, which means that the new problem is also an MaOP. However, this issue is actually trivial. In Subsection 3.5, we will explain the reason why the new optimization problem can be easily solved by a traditional Pareto-based MOEA, even though the number of objectives could be large. Additionally, our experimental results will demonstrate the effectiveness of the proposed method in Sections 4 and 5.

### 3 The Proposed Method

#### 3.1 The Meta-Objective Method (MeO) for MOPs

MeO converts the MOP formulated by (1) into the following optimization problem:

$$\begin{aligned} \min f'(x) &= \min (f'_1(x), f'_2(x), \dots, f'_M(x)) \\ \text{s.t. } x &\in S \subset \mathbf{R}^n \end{aligned} \quad (4)$$

where

$$f'_m(x) = f_m^D(x) + \eta \cdot f^C(x) \quad (5)$$

$f_m^D(x)$ ,  $m = 1, \dots, M$  is the diversity component that measures the position of a solution,  $x$ , with respect to the  $f_m$ -axis in the objective space,  $f^C(x)$  is the convergence component that reflects the distance between a solution and the Pareto optimal front.

In the following, we present the methods of calculating the two components. For the first component,

$$f_m^D(x) = \cos^2\theta(v_m, f(x) - z^*) \quad (6)$$



where  $v_m = (0, \dots, 1_m, \dots, 0)$ , and  $\theta(v_m, f(x) - z^*)$  is the angle between  $v_m$  and  $f(x) - z^*$ . It is clear that  $f_m^D(x) \in [0, 1]$ .  $f_m^D(x) = 0$  implies that  $f(x) - z^*$  is perpendicular to the  $f_m$ -axis, while  $f_m^D(x) = 1$  indicates that  $f(x) - z^*$  and the  $f_m$ -axis are colinear. Note that  $\cos^2\theta$  is adopted in the diversity component to make Pareto optimal solutions located on a hyperplane, which will be explained in Theorem 2.

With respect to the second component,

$$f^C(x) = d(r, f(x)) \quad (7)$$

where  $r$  is a point located on the Pareto optimal front and closest to  $f(x)$ .  $d(r, f(x)) \geq 0$  defines the distance between  $r$  and  $f(x)$ .

For an intuitive understanding, Figure 2.(b) shows how to map a solution in Figure 2.(a) from the original objective space to the meta-objective space for a bi-objective problem, where  $d(x)$  is short for  $d(r, f(x))$  and  $\theta_m, m = 1, 2$  for  $\theta(v_m, f(x) - z^*)$ .

According to the definitions in Subsection 2.1, we have the following two theorems.

**Theorem 1.** Any Pareto optimal solution to the original optimization problem defined as in (1) is a Pareto optimal solution to the new optimization problem as (4).

*Proof of Theorem 1.* Since the convergence component,  $f^C(x^*)$ , of a Pareto optimal solution to the original optimization problem must be zero, the theorem will be proven if we can show that when  $f^C(x^*) = 0$ ,  $x^*$  is definitely a Pareto optimal solution to the new optimization problem. If this assertion would not hold, then there exists another solution,  $x'$ , that dominates  $x^*$ . According to **Definition 1**,  $\sum_{m=1}^M f'_m(x') < \sum_{m=1}^M f'_m(x^*)$ , i.e.,  $\eta M f^C(x') + \sum_{m=1}^M f_m^D(x') < \eta M f^C(x^*) + \sum_{m=1}^M f_m^D(x^*)$ . Since  $\sum_{m=1}^M f_m^D(x) = \sum_{m=1}^M \cos^2\theta(v_m, f(x) - z^*) = \sum_{m=1}^M \frac{|f_m(x) - z_m^*|^2}{|f(x) - z^*|^2} = 1$ , then  $f^C(x') < f^C(x^*) = 0$ . This contradicts the fact that  $f^C(x') \geq 0$ . Consequently, we can infer that there are no other solutions that dominate  $x^*$ . This means that  $x^*$  is a Pareto optimal solution to the new optimization problem. We have thus proved the theorem.  $\square$

**Theorem 2.** The Pareto optimal front of the new optimization problem is located on a hyperplane formed by  $\sum_{m=1}^M f'_m = 1$ .

*Proof of Theorem 2.* According to **Theorem 1**, for any Pareto optimal solution,  $x^*$ ,  $f^C(x^*) = 0$ , and we get  $\sum_{m=1}^M f_m^D(x^*) = 1$ . This means that  $x^*$  is on a hyperplane formed by  $\sum_{m=1}^M f'_m = 1$  in the meta-objective space.  $\square$

**Theorem 1** implies that all Pareto optimal solutions to the original optimization problem can be achieved by successfully solving the new optimization problem. However, there exists a critical issue when using this method. Since the true Pareto optimal front is unavailable *a priori*, it is impractical to obtain  $d(r, f(x))$  and  $z^*$ . In multi-objective evolutionary optimization,  $z_m^*$  can be well replaced with the known minimum of  $f_m(x)$  during the evolution. Therefore, how to estimate  $d(r, f(x))$  as the convergence component is the major issue. In view of this, we present three alternatives to the convergence component, which will be described in detail in the following subsection.

### 3.2 Alternatives to the convergence component

In this subsection, we give three alternatives to compute the convergence component, which are based on the Pareto rank method (Deb et al., 2002), the  $L_p$  scalarizing method (Wang et al., 2016b), and the double-rank method, respectively.

#### 3.2.1 The Pareto rank method

It is often believed that the Pareto rank is ineffective in distinguishing superior solutions when the number of objectives is large. However, when cooperating with a good diversity based selection criterion, it is also capable of selecting solutions with good convergence. One of the most successful examples is NSGA-III (Deb and Jain, 2013). Therefore, in this study, the Pareto rank is regarded as one of alternatives to the convergence component, which is represented as follows:

$$\hat{f}^C(\mathbf{x}) = \text{rank}(\mathbf{x}) - 1 \quad (8)$$

where  $\hat{f}^C$  represents the alternative of  $f^C$ ,  $\text{rank}(\mathbf{x})$  is the Pareto rank of  $\mathbf{x}$  in the current solution set, with its value being in the range of it is  $[1, N]$  ( $N$  is the size of the solution set). Please refer to (Deb et al., 2002) for the details of calculating  $\text{rank}(\mathbf{x})$ .  $\eta$  defined in (5) will be set to 1 when using this method. Under this circumstance, a solution dominating others in the original optimization problem will be definitely superior to these solutions in the new optimization problem, since each diversity component is in the range of  $[0, 1]$ . Solutions in the first Pareto rank has zero values of  $\hat{f}^C$  and are assumed as the Pareto optimal solutions. However, this method has a low efficiency, which results in a slow convergence rate for some optimization problems.

#### 3.2.2 The $L_p$ scalarizing method

The  $L_p$  scalarizing method is commonly used in the decomposition- and preference-based MOEAs (Wang et al., 2016b; Zhang and Li, 2007; Goulart and Campelo, 2016). In this study, however, no weight vectors are predefined for the objectives, since all the objectives are considered equally important. The alternative to the convergence component based on this method is formulated as follows:

$$\hat{f}^C(\mathbf{x}) = \frac{f^S(\mathbf{x}) - f_{\min}^S}{f_{\max}^S - f_{\min}^S}, \quad (9)$$

where

$$f^S(\mathbf{x}) = \left( \sum_{m=1}^M (f_m(\mathbf{x}) - z_m^*)^p \right)^{\frac{1}{p}} \quad (10)$$

$f_{\min}^S$  and  $f_{\max}^S$  represent the minimum and maximum values of  $f^S(\mathbf{x})$  during the evolution, respectively. The contour curves of the  $L_p$  scalarizing function varies according to the value of  $p$ . Figure 3 shows three examples. When  $p = 1$ ,  $f^S(\mathbf{x})$  refers to the sum of all the objective values, and when  $p = 2$  (or  $p = \infty$ ),  $f^S(\mathbf{x})$  means the Euclidean (or Chebyshev) distance between the solution,  $\mathbf{x}$ , and the ideal point. Intuitively, this convergence component will work well for a problem whose Pareto optimal front fits its contour curves. The solution with the smallest  $L_p$  scalarizing function value is termed as the optimal solution.

The main advantage of the  $L_p$  scalarizing method is that it provides a larger selection pressure than that of the Pareto rank method. However, this method has the following three deficiencies. The first is that the process of converting objectives is not

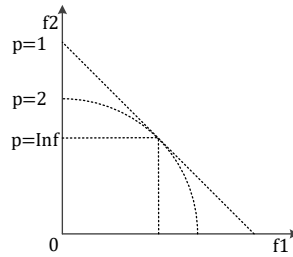


Figure 3: The contour curves of the  $L_p$  scalarizing method with different values of  $p$ .

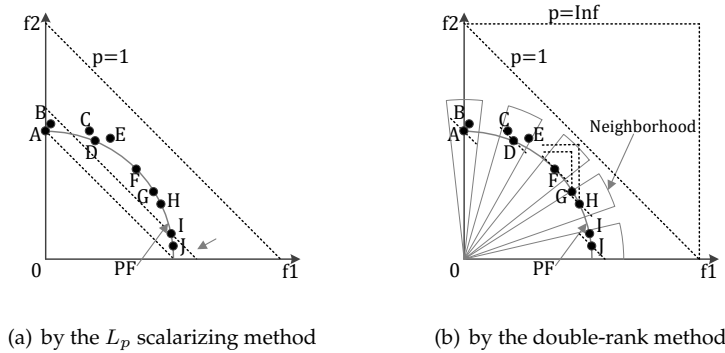


Figure 4: The evaluation of the convergence performances of solutions, where  $p = 1$ .

strictly consistent with the Pareto dominance criterion. In other words, a dominated solution to the original optimization problem may become non-dominated under the new optimization problem, and be preferentially selected. The second is that this method is very sensitive to the setting of parameter  $\eta$  as defined in (5). Obtaining a proper setting of  $\eta$  is non-trivial, if not impossible, especially when the problem has a complex Pareto optimal front. The last but not least is that although Wang et al. (2016b) presented a strategy of adaptively selecting an appropriate value of  $p$  for MOEA/D, one single scalarizing function cannot properly measure the solutions on the whole Pareto front in this study. The bias with respect to its contour curves may result in solutions located in specific regions. An illustrative example is provided in Figure 4.(a). Solutions  $A$ - $J$  are candidates for a bi-objective optimization problem with a concave Pareto optimal front. Since the geometric shape of the Pareto optimal front is actually unknown, the  $L_p$  scalarizing method with  $p = 1$  is simply chosen. Then solutions  $A$ ,  $B$ ,  $I$  and  $J$  will be preferred, since they have smaller values of the  $L_p$  scalarizing function values than the rest. Consequently, solutions in the final solution set may be overly crowded in edge regions. However, the double-rank method can overcome this issue (as shown in Figure 4.(b)), which will be presented in the next subsection.

### 3.2.3 The double-rank method

Taking the advantages of the above two designs and addressing their disadvantages, we propose the third alternative to the convergence component based on the double-rank method, which is formulated as follows:

$$\hat{f}^C(\mathbf{x}) = 2 \cdot (\text{rank}(\mathbf{x}) - 1) + \frac{1}{n_{\max}} \cdot (\text{rank}_1(\mathbf{x}) - 1) \quad (11)$$

where  $\text{rank}_1(\mathbf{x})$  is the local rank of  $\mathbf{x}$  in its neighborhood in terms of the  $L_p$  scalarizing function value. Solutions whose angles to  $\mathbf{x}$  smaller than a threshold value,  $\theta_1$ , are termed as the neighbours of  $\mathbf{x}$ .  $\theta_1$  is set to  $C \cdot \pi \cdot N^{-\frac{1}{M-1}}$ , where  $N$  is the population size,  $M$  is the objective number, and  $C$  is constant parameter. Increasing  $N$  and reducing of  $M$  decrease the neighborhood size. Based on our preliminary investigations,  $C = 0.22$  would work well for most problems. For a problem whose Pareto front does not cover the whole objective space, the value of  $C$  can be set a bit smaller to achieve a better performance. The number of neighbours of  $\mathbf{x}$  is denoted as  $n_1(\mathbf{x})$ , then  $n_{\max}$  is the maximum of  $n_1(\mathbf{x})$  among the current population.

When employing this method,  $\eta$  can simply be set to 1. Only non-dominated solutions with the smallest  $L_p$  scalarizing function value among their neighbors have zero value of  $\hat{f}^C$  and are termed as the optimal solutions. This method has the following four advantages: (1) the process of converting objective is strictly consistent with the Pareto dominance criterion due to  $\text{rank}(\mathbf{x})$ ; (2) it offers a large selection pressure in the high-dimensional objective space due to  $\text{rank}_1(\mathbf{x})$ ; (3) it is insensitive to the settings of parameters. Although  $\theta_1$  is an adjustable parameter, it does not need to be tuned in most situations; and (4) the bias with respect to the contour curves of the  $L_p$  scalarizing function can be neglected. Please see Figure 4.(b) for an example. Compared to the method in Figure 4.(a), solutions in the non-edge regions will also be preferred, since some solutions (e.g.,  $D$ ,  $F$  and  $H$ ) have the smallest values of the  $L_p$  scalarizing function in their neighborhoods. Additionally, it can be seen from Figure 4.(b) that if  $p$  is set to  $\infty$ ,  $H$  will no longer be assumed as the best solution, since its neighbor,  $G$ , has a smaller value of the  $L_p$  scalarizing function. However, if the population is infinite and the neighborhood size is infinitesimal, this discrepancy resulted from different settings of  $p$  will become negligibly small.

The proposed double-rank method inherently utilizes the idea of niching (Liu et al., 2018a,c) and shares some common ideas with constrained decomposition based methods, e.g., localized weighted sum (Wang et al., 2018) and localized PBI methods (Wang et al., 2016a), where they all evaluate solutions according to their scalarizing function values in their local areas. The main difference between them is that the double-rank method does not require any reference vector. Another difference is that the double-rank method allows solutions in the same rank to compete on equal terms, which further promotes diversity in the solution set.

### 3.3 Incorporating MeO into MOEAs

MeO can be integrated within any existing MOEAs. Since the generational MOEAs are most commonly seen, we give a flow chart of incorporating MeO into them in Figure 5. We denote an MOEA combined with MeO as MeO+MOEA in this study, e.g., MeO+NSGA-II. Specifically, the versions with the Pareto rank method, the  $L_p$  scalarizing method, and the double-rank method are denoted as MeO<sup>I</sup>+MOEA, MeO<sup>II</sup>+MOEA and MeO<sup>III</sup>+MOEA, respectively. The unique difference of MeO+MOEA from the original MOEA is that individuals are evaluated in terms of their meta-objectives, which are

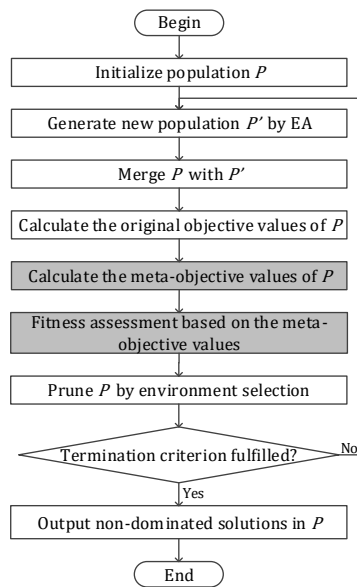


Figure 5: The flow chart of MeO+MOEA.

shown in the shaded parts of the flow chart. Note that in this study, when employing a normalization procedure to tackling a scaled problem, the original objective values will be normalized according to the maximum and minimum original objective values in the current population before calculating the meta-objective values.

In addition, the computational complexity analysis of MeO+MOEA is given in Section I in the Supplementary Material.

### 3.4 Why Does MeO work for MaOPs?

After the above description of MeO, one may be interested in the reason why MeO+MOEA could be effective for MaOPs when the MOEA is a traditional Pareto-based one and the new problem also has a large number of objectives. We now explain the reasons under the following two situations: (1) selecting solutions with the same value of the convergence component; and (2) selecting solutions with different values of the convergence component.

#### 3.4.1 Situation 1

This usually happens when the Pareto rank method is employed as an alternative to the convergence component. In the original Pareto-based MOEAs, when solutions are in the same Pareto rank, the diversity-based selection criteria, e.g., the crowding distance (Deb et al., 2002) or the  $k$ -nearest neighbour (Zitzler et al., 2001), may prefer the solutions far from the Pareto optimal front since they are located in sparse regions. In other words, the poor convergence performance of a solution is mistaken for its good diversity performance. This is one of reasons why the Pareto-based MOEAs fail in solving MaOPs. In NSGA-III, this issue is solved by a new diversity-based selection criterion. The solutions are mapped into a hyperplane and evaluated by their distances to the predefined reference points. In the proposed method in this study, if different

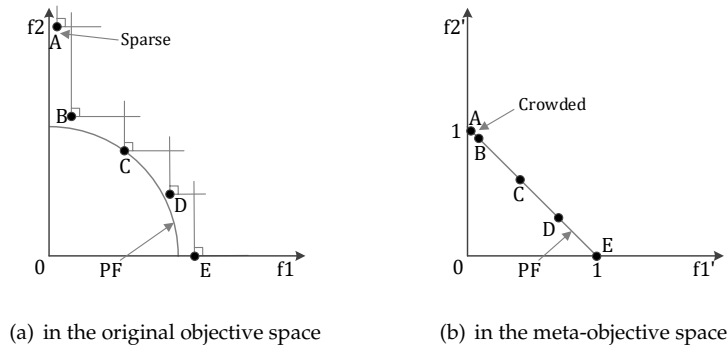


Figure 6: Density estimation of solutions.

solutions have the same value of the convergence component, they will be located on a hyperplane, however, in the meta-objective space. In this situation, only their diversity performances will be traded off. Therefore, the aforementioned diversity-based selection criteria (Deb et al., 2002; Zitzler et al., 2001) can also be effective. Figure 6 provides an illustrative example, where solutions  $A-E$  are non-dominated solutions. Solution  $A$  will be estimated to have a low density value in the original objective space. In contrast, when employing the Pareto rank method as an alternative to the convergence component, all these solutions will be mapped into a hyperplane in the meta-objective space, and solution  $A$  will be estimated to have a high density value. Compared to NSGA-III, the main advantage of the proposed method is that any density-based selection criterion can be adopted and no reference points need to be predefined.

### 3.4.2 Situation 2

This often happens when the  $L_p$  scalarizing method or the double-rank method is employed as an alternative to the convergence component. In this situation, solutions with large values of the convergence component will be penalized in the meta-objective space, whilst those with small values of the convergence component will have more chances to become the dominators. This effect is somewhat similar to that of  $\epsilon$ -dominance which has been proven capable for solving MaOPs (Hadka and Reed, 2012; Yang et al., 2013). When comparing two solutions which do not dominate each other in the objective space, if the difference of the convergence components of the two solutions is larger than a positive constant, i.e.,  $\epsilon$ , the one with a smaller value will dominate the other in the meta-objective space. Figure 7 provides an illustrative example. In Figure 7.(a), solutions  $A$  and  $B$  do not dominate each other in the original objective space, and  $f^C(A)$  and  $f^C(B)$  are their distances to the Pareto optimal front, respectively. If we only consider the diversity components of solutions  $A$  and  $B$ , we can obtain points  $A'$  and  $B'$  in Figure 7.(b). Assuming that  $\eta = 1$  in this case, after adding the convergence components, i.e.,  $f^C(A)$  and  $f^C(B)$ , they are mapped to  $A$  and  $B$  in the meta-objective space. In this mapping process, we notice that if the difference of  $f^C(A)$  and  $f^C(B)$  is larger than the vertical or horizontal distance (these two distances are the same) between  $A'$  and  $B'$ ,  $B$  will be located in the domination area of  $A$  in the meta-objective space. This means  $\epsilon$  is actually the vertical or horizontal distance between  $A'$  and  $B'$ . From Figure 7, we can draw the following observation. For any two solutions  $A$  and

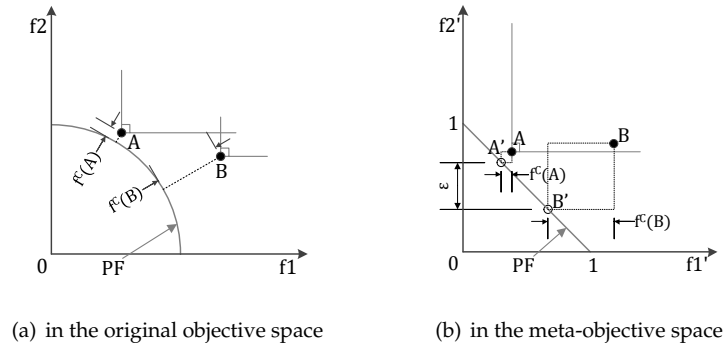


Figure 7: Pareto dominance relationship of solutions.

$B$ , if  $f^C(A) < f^C(B)$  and  $\eta \cdot |f^C(A) - f^C(B)| > \epsilon$ , solution  $A$  will dominate  $B$  in the meta-objective space. Consequently, the Pareto dominance becomes much more effective in MaOPs after incorporating MeO. It is worth noting that  $A'$  and  $B'$  are in the range of  $[0, 1]$ , and then  $\epsilon$  is in the range of  $(0, 1]$ . This is the reason why  $\eta$  can be set to 1 in the Pareto rank method and the double-rank method. Under this circumstances, if solution  $A$  dominates solution  $B$  in the original objective space,  $\eta \cdot |f^C(A) - f^C(B)|$  will be definitely larger than  $\epsilon$ . As a result, solution  $A$  will necessarily dominate  $B$  in the meta-objective space.

In addition, SDE (Li et al., 2014) is another algorithm that penalizes solutions with poor convergence performances, where the penalization is considered in the density-based selection criterion. In SDE, solutions with poor convergence performances are shifted (mapped) to new locations that close to other solutions, thus they will have larger density values. SDE has been integrated with NSGA-II and SPEA2, respectively. However, it only works well with the latter and is computationally very expensive. On the other hand, both NSGA-II and SPEA2 perform very well after integrating with MeO.

#### 4 Investigations on the Effect of MeO on Two Pareto-based MOEA

In this section, we validate the performance of MeO by integrating it into two Pareto-based MOEAs, i.e., NSGA-II and SPEA2, which results in two new algorithms, denoted as MeO+NSGA-II and MeO+SPEA2, respectively. There are two reasons for choosing these two MOEAs. The first is that both are classic MOEAs and have been applied in a wide range of optimization problems with success. The second is that both are generally believed incapable of solving MaOPs. The experiments in this section are divided into the following three parts. Firstly, we compare MeO<sup>I</sup>+NSGA-II and MeO<sup>I</sup>+SPEA2 with their corresponding original versions in high-dimension DTLZ benchmark problems. Following that, we further compare three different alternatives to the convergence component in the design of MeO+NSGA-II.

##### 4.1 Test Problems, Performance Indicators, and Parameter Settings

We choose two widely used test suites, i.e., DTLZ (Deb et al., 2005) and WFG (Huband et al., 2006), for empirical comparisons in this study. We consider these scalable problems with 4, 6, 8, and 10 objectives in this paper. The number of variables is set to

$M + 4$  for DTLZ1,  $M + 19$  for DTLZ7, and  $M + 9$  for the other DTLZ problems. The distance- and position-related parameters are set to 24 and  $M - 1$ , respectively, for the WFG problems. Please refer to (Deb et al., 2005) and (Huband et al., 2006) for detailed descriptions of the DTLZ and WFG suites.

In order to compare different algorithms on these test problems, we adopt the Inverted Generation Distance Plus (IGD<sup>+</sup>) (Ishibuchi et al., 2015) as a performance metric which gives a comprehensive quantification of both the convergence and diversity of a solution set. It is weakly Pareto compliant and thus more accurate on evaluation than its original version, IGD. The smaller the value of IGD<sup>+</sup> of an algorithm, the better the performance of the algorithm. IGD<sup>+</sup> requires a reference set, whose points are typically uniformly distributed on the true Pareto optimal front of a test problem. In this study, over 5,000 reference points are evenly sampled from the true Pareto optimal front of each test problem. Note that solutions and the reference points will be normalized based on the true Pareto optimal front, when calculating IGD<sup>+</sup>.

The following parameter settings are adopted by all the compared algorithms. Simulated binary crossover and polynomial mutation are applied as the crossover and mutation operators, respectively, where both distribution indexes are set to 20. The crossover and mutation probabilities are 1.0 and  $1/n$ , respectively, where  $n$  is the number of decision variables. The termination criterion is that the population has evolved for the predefined maximum number of generations. For DTLZ1, DTLZ3, DTLZ6, and WFG1, the maximum number of generations is set to 1,000, while 300 is used for the rest test functions. The setting of termination criterion is the same as those in the recent studies (Li et al., 2015; Yang et al., 2013). The population size,  $N$ , is set to 120, 132, 156, and 275, when  $M$  is 4, 6, 8, and 10, respectively. All these parameter settings are the same as those in Section 5 for a fair comparison. All the experimental results in this paper are obtained by executing 50 independent runs of each algorithm on each optimization problem. The Wilcoxon's rank sum test is employed to determine whether one algorithm shows a statistically significant difference with the other, and the null hypothesis is rejected at a significant level of 0.05.

## 4.2 Comparisons among NSGA-II, SPEA2, MeO<sup>I</sup>+NSGA-II and MeO<sup>I</sup>+SPEA2

In this subsection, we apply NSGA-II, SPEA2 and their new versions (i.e., MeO<sup>I</sup>+NSGA-II and MeO<sup>I</sup>+SPEA2 with the Pareto rank method as the convergence measure) to solve the DTLZ problems. Table 1 shows the mean values of IGD<sup>+</sup> obtained by all competing algorithms on the DTLZ problems. The results in boldface indicate that MeO<sup>I</sup>+NSGA-II (MeO<sup>I</sup>+SPEA2) performs significantly better its original version. In Table 1, the performance score (Bader and Zitzler, 2011) of each algorithm on each test problem is given in gray scale after the mean IGD<sup>+</sup> value. A darker tone corresponds to a larger performance score. For a test problem, the performance score of an algorithm is the number of the competing algorithms which perform significantly worse than it according to IGD<sup>+</sup>. Moreover, we give the average performance score (APS) of each algorithms on all the test problems at the bottom of Table 1. Note that in this paper, notation DTLZ $i$ - $M$  refers to DTLZ $i$  (i.e.,  $i \in 1, 2, 3, 4, 5, 6, 7$ ) with  $M$  objectives (i.e.,  $M \in 4, 6, 8, 10$ ), and WFG $i$ - $M$  has a similar definition.

As can be seen from Table 1, the performances of both NSGA-II and SPEA2 are significantly boosted on most problems when integrating with MeO. These results show the effectiveness of MeO when solving MaOPs. However, MeO<sup>I</sup>+SPEA2 does not show significantly better performance than SPEA2 on some 4- and 6-objective test problems, e.g., DTLZ1-4 and -6, DTLZ5-4, DTLZ6-4, and DTLZ7-4 and -6. In addition, neither



Table 1: Comparisons among NSGA-II, SPEA2, MeO+NSGA-II and MeO+SPEA2 based on the mean value of IGD<sup>+</sup>.

IGD+	SPEA2	MeO <sup>I</sup> + SPEA2	NSGA-II	MeO <sup>I</sup> + NSGA-II	MeO <sup>II</sup> + NSGA-II (p=1)	MeO <sup>II</sup> + NSGA-II (p=2)	MeO <sup>III</sup> + NSGA-II (p=1)	MeO <sup>III</sup> + NSGA-II (p=2)	MeO <sup>*</sup> + NSGA-II
DTLZ1-4	5.958E-2 6	7.490E-2 2	7.226E-2 3	<b>6.637E-2</b> 5	1.423E-1 1	1.826E-1 0	5.910E-2 7	5.919E-2 7	6.777E-2 4
DTLZ1-6	1.952E-1 3	3.672E-1 2	8.706E-1 0	<b>5.026E-1</b> 1	3.452E-1 2	5.553E-1 0	1.009E-1 7	1.010E-1 7	1.450E-1 4
DTLZ1-8	2.246E+2 0	<b>1.692E-1</b> 6	5.436E+0 1	<b>1.955E-1</b> 4	2.095E-1 3	3.918E-1 2	1.256E-1 7	1.253E-1 7	1.704E-1 5
DTLZ1-10	3.090E+2 0	<b>2.903E+0</b> 2	1.432E+1 1	<b>2.233E+0</b> 3	7.277E-1 5	1.216E+0 4	2.054E-1 6	2.029E-1 7	2.032E-1 7
DTLZ2-4	6.478E-2 4	<b>5.619E-2</b> 8	8.236E-2 0	<b>7.397E-2</b> 1	7.251E-2 1	5.896E-2 7	6.111E-2 5	6.132E-2 5	6.851E-2 3
DTLZ2-6	6.039E-1 1	<b>1.173E-1</b> 8	1.506E+0 0	<b>1.420E-1</b> 2	1.421E-1 2	1.243E-1 5	1.188E-1 6	1.195E-1 6	1.271E-1 4
DTLZ2-8	2.349E+0 0	<b>1.546E-1</b> 6	1.971E+0 1	<b>1.849E-1</b> 3	1.913E-1 2	1.709E-1 4	1.561E-1 6	1.564E-1 6	1.632E-1 5
DTLZ2-10	2.399E+0 0	<b>3.623E-1</b> 2	1.456E+0 1	<b>3.396E-1</b> 3	2.246E-1 4	2.025E-1 5	1.852E-1 7	1.845E-1 7	1.957E-1 6
DTLZ3-4	5.729E-2 7	5.706E-2 7	7.709E-2 2	<b>7.464E-2</b> 3	4.260E-1 1	1.952E+0 0	5.953E-2 5	6.010E-2 5	6.842E-2 4
DTLZ3-6	3.636E+1 0	<b>7.364E+0</b> 5	1.834E+1 1	<b>9.840E+0</b> 3	8.484E+0 3	1.597E+1 1	2.662E-1 7	2.533E-1 7	2.533E-1 6
DTLZ3-8	1.188E+3 0	<b>5.770E+0</b> 3	9.989E+2 1	<b>1.291E+1</b> 2	1.548E+0 5	4.498E+0 3	1.554E-1 7	1.568E-1 7	1.671E-1 6
DTLZ3-10	1.435E+3 0	<b>2.101E+2</b> 2	1.379E+3 0	<b>2.074E+2</b> 2	1.316E+1 5	2.268E+1 4	2.937E-1 7	2.931E-1 7	4.972E-1 6
DTLZ4-4	1.070E-1 0	<b>7.807E-2</b> 2	8.353E-2 2	<b>7.247E-2</b> 5	7.396E-2 5	7.303E-2 6	8.354E-2 1	7.669E-2 2	6.659E-2 3
DTLZ4-6	7.824E-1 1	<b>1.404E-1</b> 4	1.063E+0 0	<b>1.577E-1</b> 2	1.503E-1 3	1.197E-1 7	1.227E-1 6	1.191E-1 7	1.356E-1 5
DTLZ4-8	2.155E+0 0	<b>1.948E-1</b> 3	1.949E+0 1	<b>2.148E-1</b> 2	1.938E-1 3	1.475E-1 8	1.573E-1 6	1.572E-1 6	1.632E-1 5
DTLZ4-10	2.371E+0 0	<b>5.300E-1</b> 2	2.053E+0 1	<b>4.448E-1</b> 3	2.256E-1 4	1.678E-1 8	1.866E-1 5	1.868E-1 5	1.837E-1 7
DTLZ5-4	8.682E-2 2	1.255E-1 1	3.395E-2 3	<b>1.314E-2</b> 7	4.785E-1 0	1.813E-2 4	1.755E-2 4	1.803E-2 4	2.849E-3 3
DTLZ5-6	9.392E-1 0	<b>5.865E-1</b> 2	1.591E-1 3	<b>6.112E-2</b> 4	7.154E-1 1	2.031E-2 7	3.446E-2 5	3.434E-2 5	2.914E-3 3
DTLZ5-8	1.874E+0 1	2.751E+0 0	2.080E-1 3	1.832E-1 3	7.953E-1 2	2.346E-2 7	3.814E-2 5	3.741E-2 5	2.758E-3 3
DTLZ5-10	1.770E+0 1	2.693E+0 0	2.345E-1 4	4.910E-1 3	8.700E-1 2	2.527E-2 7	3.415E-2 6	3.768E-2 5	1.712E-3 3
DTLZ6-4	9.253E-2 2	1.503E-1 1	2.650E-1 0	<b>1.985E-2</b> 5	7.559E-2 2	6.478E-2 2	1.846E-2 2	1.795E-2 5	2.713E-3 3
DTLZ6-6	1.050E+1 0	<b>7.611E+0</b> 1	5.298E+0 2	<b>4.107E+0</b> 3	1.341E-1 4	7.621E-2 5	3.611E-2 6	3.535E-2 6	2.690E-3 3
DTLZ6-8	1.135E+1 0	<b>1.031E+1</b> 1	6.379E+0 3	9.415E+0 2	1.366E-1 4	8.764E-2 5	4.165E-2 6	4.278E-2 6	2.587E-3 3
DTLZ6-10	1.304E+1 0	<b>1.018E+1</b> 1	6.491E+0 3	9.107E+0 2	3.115E-1 4	1.721E-1 5	5.176E-2 6	5.347E-2 6	1.592E-3 3
DTLZ7-4	7.795E-2 6	8.770E-2 4	1.043E-1 3	<b>8.647E-2</b> 4	3.148E-1 0	1.484E-1 1	5.500E-2 7	5.595E-2 7	1.324E-1 2
DTLZ7-6	1.570E-1 6	2.009E-1 2	4.037E-1 0	<b>1.832E-1</b> 3	3.863E-1 1	2.025E-1 2	1.284E-1 7	1.284E-1 7	1.617E-1 5
DTLZ7-8	5.960E-1 1	<b>2.439E-1</b> 4	9.757E-1 0	<b>2.662E-1</b> 3	4.051E-1 2	2.148E-1 6	1.908E-1 6	1.908E-1 6	2.203E-1 5
DTLZ7-10	9.598E-1 1	<b>2.841E-1</b> 4	1.367E+0 0	<b>2.920E-1</b> 3	3.968E-1 2	2.245E-1 7	2.315E-1 5	2.312E-1 5	1.890E-1 8
APS	1.500	3.036	1.393	3.071	2.607	4.357	5.821	5.893	6.107

MeO<sup>I</sup>+NSGA-II nor MeO<sup>I</sup>+SPEA2 performs better than its original version on some DTLZ5 and DTLZ6 test problems. The most possible reason is that the Pareto rank method does not provide a great selection pressure. The Pareto fronts of DTLZ5 and DTLZ6 only have two dimensions. Therefore, when solving DTLZ5 and DTLZ6 with any number of objectives and some other test problems with a relatively smaller number of objectives (e.g. 4 or 6), MeO<sup>I</sup>+NSGA-II and MeO<sup>I</sup>+SPEA2 do not show evident strength than their original versions.

Since MeO<sup>I</sup>+SPEA2 not only performs worse than MeO<sup>I</sup>+NSGA-II according to APS but also is extraordinarily time-consuming (see Table 1 in the Supplementary Material), we further investigate the performance of MeO+NSGA-II with different convergence components in the next subsection.

### 4.3 Comparisons among MeO<sup>I</sup>+NSGA-II, MeO<sup>II</sup>+NSGA-II, MeO<sup>III</sup>+NSGA-II and MeO<sup>\*</sup>+NSGA-II

In this subsection, we investigate the performance of MeO+NSGA-II with three different alternatives to the convergence component. We employ two settings of  $p$  (1 and 2) in MeO<sup>II</sup>+NSGA-II and MeO<sup>III</sup>+NSGA-II to observe their different behaviors. We also include the results obtained by the method that calculates the convergence component precisely (denoted as MeO<sup>\*</sup>+NSGA-II), since the true Pareto fronts of DTLZ problems are known. Note that  $\eta$  is set to 1 in all the designs.

Table 1 lists the mean values of IGD<sup>+</sup> obtained by different algorithms. According to Table 1, we can make the following observations: (1) MeO<sup>III</sup>+NSGA-II can achieve satisfactory results on most problems. There is nearly no significant difference between

the results obtained by MeO<sup>III</sup>+NSGA-II with  $p = 1$  and  $p = 2$ , which implies that the performance of MeO<sup>III</sup> is not sensitive to the parameter  $p$ . (2) MeO<sup>II</sup>+NSGA-II with  $p = 1$  performs significantly better on DTLZ1 than that with  $p = 2$ . This attributes to the fact that the geometrical shape of the Pareto optimal front of DTLZ1 perfectly fit the contour curves of the  $L_p$  scalarizing function with  $p = 1$ . Generally, MeO<sup>II</sup>+NSGA-II with  $p = 2$  show better performance than that with  $p = 1$  according to APS. (3) MeO<sup>II</sup>+NSGA-II does not work as well as MeO<sup>III</sup>+NSGA-II on most problems. The reason is that  $\eta$  is not tuned for each problem. However, MeO<sup>II</sup>+NSGA-II with  $p = 2$  performs very well on DTLZ4 and DTLZ5. (4) MeO\*+NSGA-II achieve the best APS on all the problems. However, it is not as good as MeO<sup>III</sup>+NSGA-II on some test problems. One reason is that it behaves like MeO<sup>II</sup>+NSGA-II to some degree and thus is relatively sensitive to the setting of parameter  $\eta$ . Another reason is that the population could be far away from the true Pareto front at the beginning of evolution. Precisely calculating the convergence component based on the true Pareto front may result in the population converging into a small region, which degenerates the diversity performance.

In addition, we show the Pareto fronts achieved by NSGA-II and MeO<sup>III</sup>+NSGA-II on some 3-objective DTLZ test problems in Figure 1 in the Supplementary Material for a visual understanding of MeO<sup>III</sup>+NSGA-II's performance.

## 5 Comparisons with the State-of-the-Art Algorithms

In this section, we compare MeO<sup>III</sup>+NSGA-II with  $p = 2$  with seven state-of-the-art MaOEAs, i.e., BiGE (Li et al., 2015), SDE (Li et al., 2014), NSGA-III (Deb and Jain, 2013), EFR-RR (Yuan et al., 2016), MOMBI-II (Hernández Gómez and Coello Coello, 2015), GrEA (Yang et al., 2013), and KnEA (Zhang et al., 2015) on DTLZ and WFG suites with 4-, 6-, 8-, and 10-objective.

### 5.1 Competing Algorithms

In this subsection, we briefly introduce the seven chosen algorithms. These algorithms cover all the main categories mentioned in Section 1 for many-objective optimization. All the competing algorithms as well as MeO are implemented by PlatEMO (Tian et al., 2017). The source code of MeO is available on <https://github.com/yiping01iu>.

(1) BiGE is a known meta-objective method designed for many-objective optimization. It simultaneously considers the convergence and diversity performances of a solution. Different from MeO proposed in this paper, BiGE transforms an MaOP into a bi-objective optimization problem. A solution set with a good performance can be achieved by optimizing a bi-objective optimization problem.

(2) SDE is a recently proposed method that modifies the density estimation strategy in traditional Pareto-based MOEAs. By shifting the positions of individuals in sparse regions, individuals with poor convergence are penalized to have a high density value, and then discarded during the evolution. In this method, the version that integrates SDE into SPEA2 (denoted as SPEA2+SDE) is employed for comparisons due to its better performance compared to other variants.

(3) NSGA-III is advanced from a popular Pareto dominance-based MOEA, NSGA-II, for handling MaOPs. NSGA-III uses a reference-point-based selection criterion instead of a density-based counterpart (i.e., the crowding distance) in NSGA-II. When employing the Pareto rank method in MeO, both NSGA-III and MeO+MOEA map non-dominated solutions into a hyperplane, whereas no reference points need to be defined in the latter.

(4) EFR-RR is based on the framework of NSGA-II. However, it is inherently a

Table 2: The setting of parameters,  $p_1$  and  $p_2$ , for controlling the number of reference points (population size) along the boundary of the Pareto optimal front and inside it, respectively.

$M$	$p_1$	$p_2$	$N$
4	7	0	120
6	4	1	132
8	3	2	156
10	3	2	275

decomposition-based method like MOEA/D. EFR-RR also concentrates in balancing convergence and diversity for MaOPs. It first considers the diversity performance of a solution. Only when the perpendicular distance from the solution to the reference vector is small, its scalarizing function value will be calculated. This idea has also been applied to MOEA/D, termed MOEA/D-DU. Since EFR-RR performs slightly better than MOEA/D-DU according to the original study, we choose the former for comparisons.

(5) MOMBI-II is an indicator-based algorithm that uses the R2 indicator to guide the search. The R2 indicator is an attractive alternative to hypervolume, due to its low computational cost and weak-Pareto compatibility. MOMBI-II takes two key aspects into account, i.e., using the scalarizing function and statistical information about the population's convergence to select individuals.

(6) GrEA exploits the potential of the grid-based approach to deal with MaOPs. The grid dominance and the grid difference are introduced to determine the relationship between individuals in a grid environment. Three grid-based criteria are incorporated into calculating the fitness of an individual to distinguish individuals in both mating and environmental selection processes. Moreover, a fitness adjustment strategy is developed to avoid partial overcrowding as well as to guide the search towards various directions in the archive.

(7) KnEA is established under the framework of NSGA-II, and it gives a priority to the knee points among non-dominated solutions. To identify local knee points, a niche method based on hyperboxes is employed, in which the niche size can be adaptively tuned in term of information from the population at the previous generations.

## 5.2 Extra Parameter Settings

In this subsection, we give the settings of extra parameters used in the above algorithms. For the common parameter settings, their settings are the same as those in Subsection 4.1.

To avoid the situation in which all reference points are located on the boundary of the Pareto optimal front for problems with a large number of objectives, the strategy of two-layered reference points (vectors) is used for NSGA-III, EFR-RR, and MOMBI-II. As a result, the population size of NSGA-III and EFR-RR cannot be arbitrarily specified. We set the population size (i.e.,  $N$ ) of all algorithms to be the same value for a fair comparison. The settings of parameters for controlling the number of reference points are listed in Table 2.

As suggested in the original study of EFR-RR, the Chebychev approach is chosen as the scalarizing method, the neighborhood size is set to  $0.1N$ , and  $K$  is set to 2. Since the number of objectives and the population size are different from those of the original KnEA and GrEA, we adjust the threshold,  $T$ , in KnEA and the grid division,  $div$ , in GrEA according to the guidelines provided in the original studies to achieve the best

performances of these algorithms. The settings of  $T$  and  $div$  for the DTLZ and WFG problems are listed in Table 3 in the Supplementary Material.

### 5.3 Results and Discussions

Table 3 shows the mean values of  $IGD^+$  and the performance scores obtained by different algorithms on the DTLZ and WFG problems. A darker tone corresponds to a larger performance score. The average performance score (APS) of each algorithms on all the test problems at the bottom of Table 3.

From Table 3 we can draw the following conclusions.

MeO<sup>III</sup>+NSGA-II shows a competitive performance on most problems. It achieves the best APS among the eight competitors. It significantly outperforms the other algorithms on more test problems. However, it performs relatively poor on some test problems, e.g., WFG2 and WFG3, since both the setting of  $p$  and  $\theta_1$  are not perfect for these problems. This suggests that if developing a method to better estimate the convergence component, the performance of MeO+MOEA can be further enhanced. One promising way is to adaptively tune  $p$  and  $\theta_1$  for a given optimization problem. However, this is left for future study.

BiGE shows appealing results on some test problems, such as DTLZ5, WFG2, and WFG3, although it does not achieve any best result among the algorithms considered. BiGE works poorly on DTLZ1, DTLZ3, and DTLZ7. The reason is that the normalization procedure results in an improper niche size for these problems.

SPEA2+SDE shows a competitive performance on DTLZ1, DTLZ5, DTLZ6, DTLZ7, and WFG1. However, SPEA2+SDE is a very time-consuming design, whose runtime is no less than the original SPEA2. Please see Table 1 in the Supplementary for an elaborated discussion. Although SDE can also be combine with NSGA-II which is computationally efficient, the performance of NSGA-II+SDE is much worse. Please refer to (Li et al., 2014) for more details.

NSGA-III works quite well on most test problems. However, its performance scales poorly with the increasing number of objectives on DTLZ5, DTLZ6, WFG1, and WFG3. This attributes to the fact that the reference vectors employed in NSGA-III are uniformly distributed in the whole objective space, whereas the Pareto optimal solutions (or non-dominated solutions during the evolution) of these problems are not.

EFR-RR can effectively deal with most problems. Although it has the same issue as NSGA-III, it generally performs better than for NSGA-III on DTLZ5, DTLZ6, WFG1, and WFG3. The performance of both EFR-RR and NSGA-III can be further improved by integrating a strategy for adaptively generating the reference vectors (Qi et al., 2014).

MOMBI-II performs poorly on some scaled test problems, e.g., DTLZ7, and WFG4 to 9, since it does not employ a normalization operator. In addition, it adopts the R2 indicator based on the reference vectors, which results in a low time consumption but share the similar disadvantages as those in the decomposition-based algorithms.

GrEA achieves encouraging results on most problems. However, it is very sensitive to the choice of parameter  $div$ . In our experiments, we test a number of settings of  $div$  for each instance to make sure that GrEA can obtain a satisfactory performance. If a method of adaptively adjusting  $div$  can be further developed, the performance of GrEA will be considerably enhanced on various problems.

KnEA has an advantage of approximating the true Pareto optimal front by giving preferences to the knee points among non-dominated solutions. It performs very well on DTLZ7, WFG1, and WFG2. Similar to GrEA, KnEA needs a proper setting of the parameter,  $T$ . A smaller value of  $T$  can lead to more solutions with good convergence.

Table 3: Comparisons between MeO<sup>III</sup>+NSGA-II and seven state-of-the-art algorithms regarding the mean value of IGD<sup>+</sup>.

IGD <sup>+</sup>	MeO <sup>III</sup> + NSGA-II (p=2)	BiGE	SPEA2+SDE	NSGA-III	EFR-RR	MOMBI-II	GrEA	KnEA
DTLZ1-4	5.919E-2	1.241E-1	5.481E-2	5.599E-2	5.597E-2	5.615E-2	1.799E-1	1.983E-1
DTLZ1-6	1.010E-1	2.519E-1	9.584E-2	1.073E-1	1.070E-1	1.242E-1	2.793E-1	3.870E-1
DTLZ1-8	1.253E-1	3.503E-1	1.200E-1	1.298E-1	1.240E-1	2.639E-1	3.445E-1	4.091E-1
DTLZ1-10	2.029E-1	5.007E-1	1.495E-1	1.883E-1	1.746E-1	3.123E-1	5.406E-1	7.168E+0
DTLZ2-4	6.132E-2	7.808E-2	6.986E-2	5.707E-2	5.734E-2	5.716E-2	6.053E-2	6.602E-2
DTLZ2-6	1.195E-1	1.425E-1	1.248E-1	1.231E-1	1.176E-1	1.169E-1	1.172E-1	1.276E-1
DTLZ2-8	1.564E-1	1.743E-1	1.571E-1	1.488E-1	1.455E-1	1.415E-1	1.521E-1	1.668E-1
DTLZ2-10	1.845E-1	2.057E-1	1.933E-1	2.080E-1	2.044E-1	1.997E-1	1.857E-1	1.986E-1
DTLZ3-4	6.010E-2	1.217E-1	6.891E-2	5.665E-2	5.679E-2	5.696E-2	3.200E-1	3.298E-1
DTLZ3-6	2.533E-1	8.190E-1	2.685E-1	2.498E-1	2.522E-1	2.544E-1	8.492E-1	7.854E-1
DTLZ3-8	1.568E-1	3.881E+0	1.630E-1	1.864E-1	1.464E-1	1.837E-1	5.178E-1	2.215E+1
DTLZ3-10	2.931E-1	1.181E+1	4.271E-1	9.038E-1	4.472E-1	6.017E-1	1.240E+0	1.382E+0
DTLZ4-4	7.669E-2	7.427E-2	1.050E-1	7.361E-2	7.050E-2	8.491E-2	8.195E-2	6.334E-2
DTLZ4-6	1.191E-1	1.402E-1	1.432E-1	1.324E-1	1.167E-1	1.270E-1	1.179E-1	1.255E-1
DTLZ4-8	1.572E-1	1.737E-1	1.599E-1	1.608E-1	1.499E-1	1.606E-1	1.531E-1	1.646E-1
DTLZ4-10	1.868E-1	2.834E-1	1.902E-1	2.012E-1	2.069E-1	2.066E-1	1.857E-1	2.038E-1
DTLZ5-4	1.803E-2	3.235E-2	1.954E-2	2.453E-2	5.183E-2	3.642E-2	4.728E-2	7.567E-2
DTLZ5-6	3.434E-2	7.816E-2	5.186E-2	1.959E-1	3.090E-1	1.132E-1	1.119E-1	1.282E-1
DTLZ5-8	3.741E-2	1.115E-1	6.644E-2	2.590E-1	1.955E-1	1.967E-1	1.373E-1	1.578E-1
DTLZ5-10	3.768E-2	1.945E-1	7.754E-2	3.342E-1	2.541E-1	4.385E-1	1.611E-1	1.857E-1
DTLZ6-4	1.795E-2	4.099E-1	2.128E-2	8.869E-2	5.309E-2	3.638E-2	1.082E-1	7.745E-2
DTLZ6-6	3.535E-2	4.371E-1	6.897E-2	2.721E-1	2.565E-1	1.418E-1	1.746E-1	1.969E-1
DTLZ6-8	4.278E-2	4.384E-1	1.132E-1	5.439E-1	2.170E-1	3.569E-1	3.105E-1	3.053E-1
DTLZ6-10	5.347E-2	3.765E-1	2.1049E-1	9.070E-1	2.255E-1	3.557E-1	2.917E-1	4.987E-1
DTLZ7-4	5.595E-2	1.011E-1	6.519E-2	6.432E-2	8.432E-2	1.176E-1	4.994E-2	5.722E-2
DTLZ7-6	1.284E-1	2.148E-1	1.146E-1	1.543E-1	1.520E-1	2.132E-1	1.202E-1	1.119E-1
DTLZ7-8	1.908E-1	3.571E-1	1.637E-1	1.949E-1	2.267E-1	2.4248E-1	1.985E-1	1.349E-1
DTLZ7-10	2.312E-1	3.821E-1	2.122E-1	2.337E-1	2.418E-1	4.625E-1	3.326E-1	2.275E-1
WFG1-4	3.078E-2	3.504E-2	3.213E-2	5.298E-2	3.230E-2	4.333E-2	3.644E-2	3.186E-2
WFG1-6	3.698E-2	4.236E-2	3.909E-2	7.106E-2	4.299E-2	4.830E-2	4.194E-2	3.608E-2
WFG1-8	3.699E-2	4.348E-2	3.766E-2	5.962E-2	3.174E-2	4.277E-2	4.675E-2	3.131E-2
WFG1-10	3.119E-2	3.969E-2	3.276E-2	5.363E-2	2.687E-2	3.509E-2	3.360E-2	2.834E-2
WFG2-4	4.631E-2	2.236E-2	3.357E-2	2.153E-2	3.214E-2	4.855E-2	2.919E-2	2.009E-2
WFG2-6	6.020E-2	3.406E-2	6.079E-2	2.626E-2	3.467E-2	4.455E-2	3.871E-2	2.051E-2
WFG2-8	5.008E-2	2.593E-2	4.520E-2	3.260E-2	2.452E-2	2.916E-2	3.100E-2	1.833E-2
WFG2-10	4.681E-2	1.870E-2	5.706E-2	4.154E-2	3.539E-2	2.1996E-2	2.242E-2	1.646E-2
WFG3-4	1.280E-1	5.218E-2	1.430E-1	1.201E-1	8.831E-2	1.046E-1	5.115E-2	1.194E-1
WFG3-6	6.446E-1	3.954E-1	8.569E-1	9.447E-1	9.054E-1	4.742E-1	1.379E-1	6.688E-1
WFG3-8	3.175E+0	2.1067E+0	5.4282E+0	2.530E+0	3.1521E+0	4.1021E+0	5.6381E-1	3.810E+0
WFG3-10	1.122E+1	4.103E+0	1.892E+1	1.032E+1	6.809E+0	1.321E+0	6.670E+0	1.722E+1
WFG4-4	6.649E-2	7.712E-2	7.930E-2	6.220E-2	7.039E-2	1.042E-1	6.601E-2	7.068E-2
WFG4-6	1.411E-1	1.591E-1	1.610E-1	1.396E-1	1.559E-1	2.477E-1	1.391E-1	1.515E-1
WFG4-8	2.054E-1	2.161E-1	2.275E-1	1.894E-1	2.074E-1	2.403E-1	1.994E-1	2.121E-1
WFG4-10	2.309E-1	2.509E-1	2.619E-1	2.494E-1	2.544E-1	3.123E-1	2.441E-1	2.496E-1
WFG5-4	8.663E-2	5.1012E-1	1.004E-1	1.7894E-2	7.9019E-2	3.127E-1	8.398E-2	8.895E-2
WFG5-6	1.619E-1	1.829E-1	1.817E-1	1.550E-1	1.784E-1	3.2507E-1	1.570E-1	1.697E-1
WFG5-8	2.262E-1	2.394E-1	2.443E-1	1.996E-1	2.239E-1	2.490E-1	2.171E-1	2.303E-1
WFG5-10	2.456E-1	2.705E-1	2.774E-1	2.601E-1	2.704E-1	2.832E-1	2.466E-1	2.630E-1
WFG6-4	8.899E-2	1.008E-1	1.036E-1	8.069E-2	9.167E-2	1.393E-1	8.685E-2	9.472E-2
WFG6-6	1.638E-1	1.811E-1	1.904E-1	1.560E-1	1.803E-1	2.933E-1	1.592E-1	1.742E-1
WFG6-8	2.254E-1	2.345E-1	2.492E-1	1.989E-1	2.231E-1	2.452E-1	2.189E-1	2.380E-1
WFG6-10	2.454E-1	2.633E-1	2.806E-1	2.690E-1	2.720E-1	2.773E-1	2.454E-1	2.620E-1
WFG7-4	6.426E-2	7.848E-2	7.914E-2	6.152E-2	7.024E-2	3.1051E-1	6.655E-2	6.811E-2
WFG7-6	1.357E-1	1.561E-1	1.605E-1	1.328E-1	1.537E-1	2.413E-1	1.354E-1	1.478E-1
WFG7-8	1.937E-1	2.094E-1	2.208E-1	1.755E-1	1.988E-1	2.159E-1	1.893E-1	2.071E-1
WFG7-10	2.181E-1	2.395E-1	2.564E-1	2.348E-1	2.469E-1	2.552E-1	2.223E-1	2.396E-1
WFG8-4	1.230E-1	2.180E-1	3.182E-1	9.621E-2	7.1097E-1	6.1668E-1	1.149E-1	5.1272E-1
WFG8-6	2.162E-1	2.147E-1	2.168E-1	1.858E-1	2.110E-1	3.696E-1	2.038E-1	2.210E-1
WFG8-8	2.836E-1	3.206E-1	2.918E-1	2.824E-1	2.551E-1	3.154E-1	2.677E-1	2.907E-1
WFG8-10	2.886E-1	3.001E-1	3.179E-1	3.457E-1	2.989E-1	3.780E-1	3.891E-1	2.972E-1
WFG9-4	7.317E-2	8.874E-2	8.436E-2	7.366E-2	7.883E-2	3.1166E-1	7.179E-2	7.253E-2
WFG9-6	1.601E-1	1.785E-1	1.764E-1	1.531E-1	1.746E-1	3.2990E-1	1.518E-1	1.623E-1
WFG9-8	2.326E-1	2.502E-1	2.510E-1	2.110E-1	2.262E-1	2.594E-1	2.212E-1	2.311E-1
WFG9-10	2.591E-1	2.792E-1	2.880E-1	2.682E-1	2.747E-1	3.3021E-1	2.566E-1	2.710E-1
APS	4.641	2.203	2.797	3.953	3.781	1.938	4.141	3.078

To sum up, MeO<sup>III</sup>+NSGA-II is very competitive compared with the state-of-the-art MOEAs. Unlike KnEA and GrEA, MeO<sup>III</sup>+NSGA-II is relatively insensitive to its design parameter,  $\eta$ . Additionally, different from NSGA-III, EFR-RR, and MOMBI-II, it does not require any reference vectors. Finally, it is computationally much cheaper than that of SPEA2+SDE.

In addition, we show the results of hypervolume obtained by these algorithms on DTLZ and WFG test problems in Table 2 in the Supplementary Material. Readers can refer to them if interested.

## 6 Conclusion

In this paper, we have proposed a novel meta-objective method for many-objective evolutionary optimization, which converts a given MOP into a new MOP. Each meta-objective in the new MOP contains a unique diversity component that represents the position of a solution according to a coordinate axis in the original objective space, which results in the conflict among the meta-objectives. Furthermore, all the meta-objectives contain the same convergence component that measures the distance between the solution and the Pareto optimal front. A well converged and distributed solution set can be obtained by optimizing the new MOP, even if the number of objectives is large. This method provides a new way to manage the requirements of convergence and diversity in many-objective optimization.

Since the Pareto optimal front is actually unknown, we have proposed three alternatives to the convergence component based on the Pareto rank method, the  $L_p$  scalarizing method, and the double-rank method, respectively. The first is effective on solving MaOPs but occasionally has a low efficiency for some optimization problems. The second provides a needed selection pressure but is very sensitive to the settings of parameters. The third is recommended by taking the advantages of the above two methods while avoiding their disadvantages at the same time.

In this study, we combine MeO with two classic Pareto-based MOEAs, i.e., NSGA-II and SPEA2, which results in two new algorithms, denoted as MeO+NSGA-II and MeO+SPEA2. The performances of the two new algorithms are significantly better than the original counterparts when solving MaOPs. Additionally, MeO+NSGA-II performs slightly better than MeO+SPEA2 and is computationally much cheaper.

To demonstrate the effectiveness of MeO, we test MeO+NSGA-II with the double-rank method on the DTLZ and WFG problems in comparison with seven state-of-the-art designs, namely, BiGE, SDE, NSGA-III, EFR-RR, MOMBI-II, GrEA, and KnEA. The experimental results demonstrate that MeO+NSGA-II is very competitive among the chosen algorithms.

To further promote the performance of MeO+MOEA, better estimating the convergence component is a must. Therefore, our future work includes the development of a strategy of adaptively tuning the parameters,  $\theta_1$  and  $p$ , in each neighborhood for the double-rank method. In addition, combining MeO with other MOEAs, e.g., indicator- and decomposition-based algorithms, and investigating the behavior of the new corresponding algorithms in many-objective optimization are of great interest.

## Acknowledgments

This work was jointly supported by National Natural Science Foundation of China with grant No. 61803192, 61876075, 61773384, 61763026, 61673404, 61573361, and 61503220, National Basic Research Program of China (973 Program) with grant No. 2014CB046306-2, National Key R&D Program of China with grant No.

2018YFB1003802-01, and China Scholarship Council with grant No. 201606420005.

## References

- Adra, S. F. and Fleming, P. J. (2011). Diversity management in evolutionary many-objective optimization. *IEEE Transactions on Evolutionary Computation*, 15(2):183–195.
- Bader, J. and Zitzler, E. (2011). HypE: An algorithm for fast hypervolume-based many-objective optimization. *Evolutionary Computation*, 19(1):45–76.
- Cheng, J., Yen, G. G., and Zhang, G. (2015). A many-objective evolutionary algorithm with enhanced mating and environmental selections. *IEEE Transactions on Evolutionary Computation*, 19(4):592–605.
- Deb, K. and Jain, H. (2013). 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*, 18(4):577–601.
- Deb, K., Pratap, A., Agarwal, S., and Meyarivan, T. (2002). A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Transactions on Evolutionary Computation*, 6(2):182–197.
- Deb, K., Thiele, L., Laumanns, M., and Zitzler, E. (2005). *Scalable test problems for evolutionary multiobjective optimization*. Springer.
- Goulart, F. and Campelo, F. (2016). Preference-guided evolutionary algorithms for many-objective optimization. *Information Sciences*, 329:236–255.
- Hadka, D. and Reed, P. (2012). Diagnostic assessment of search controls and failure modes in many-objective evolutionary optimization. *Evolutionary Computation*, 20(3):423–452.
- Han, Y., Gong, D., Jin, Y., and Pan, Q. (2017). Evolutionary multi-objective blocking lot-streaming flow shop scheduling with machine breakdowns. *IEEE Transactions on Cybernetics*, Early Access.
- Han, Y., Gong, D., and Sun, X. (2015). A discrete artificial bee colony algorithm incorporating differential evolution for the flow-shop scheduling problem with blocking. *Engineering Optimization*, 47(7):927–946.
- He, Z. and Yen, G. G. (2016). Visualization and performance metric in many-objective optimization. *IEEE Transactions on Evolutionary Computation*, 20(3):386–402.
- He, Z., Yen, G. G., and Zhang, J. (2014). Fuzzy-based Pareto optimality for many-objective evolutionary algorithms. *IEEE Transactions on Evolutionary Computation*, 18(2):269–285.
- Hernández Gómez, R. and Coello Coello, C. A. (2015). Improved metaheuristic based on the r2 indicator for many-objective optimization. In *Proceedings of the 2015 on Genetic and Evolutionary Computation Conference*, pages 679–686, Madrid, Spain. ACM.
- Huband, S., Hingston, P., Barone, L., and While, L. (2006). A review of multiobjective test problems and a scalable test problem toolkit. *IEEE Transactions on Evolutionary Computation*, 10(5):477–506.

- Ishibuchi, H., Masuda, H., Tanigaki, Y., and Nojima, Y. (2015). Modified distance calculation in generational distance and inverted generational distance. In *International Conference on Evolutionary Multi-Criterion Optimization*, pages 110–125. Springer.
- Jaimes, A. L., Aguirre, H., Tanaka, K., and Coello, C. A. C. (2010). Objective space partitioning using conflict information for many-objective optimization. In *Parallel Problem Solving from Nature, PPSN XI*, pages 657–666. Springer, Krakw, Poland.
- Laumanns, M., Thiele, L., Deb, K., and Zitzler, E. (2002). Combining convergence and diversity in evolutionary multiobjective optimization. *Evolutionary Computation*, 10(3):263–282.
- Li, J., Sang, H., Han, Y., Wang, C., and Gao, K. (2018). Efficient multi-objective optimization algorithm for hybrid flow shop scheduling problems with setup energy consumptions. *Journal of Cleaner Production*, 181:584–598.
- Li, K., Kwong, S., Cao, J., Li, M., Zheng, J., and Shen, R. (2012). Achieving balance between proximity and diversity in multi-objective evolutionary algorithm. *Information Sciences*, 182(1):220–242.
- Li, M., Yang, S., and Liu, X. (2014). Shift-based density estimation for Pareto-based algorithms in many-objective optimization. *IEEE Transactions on Evolutionary Computation*, 18(3):348–365.
- Li, M., Yang, S., and Liu, X. (2015). Bi-goal evolution for many-objective optimization problems. *Artificial Intelligence*, 228:45–65.
- Liu, Y., Gong, D., Sun, J., and Jin, Y. (2017a). A many-objective evolutionary algorithm using a one-by-one selection strategy. *IEEE Transactions on Cybernetics*, 47(9):2689–2702.
- Liu, Y., Gong, D., Sun, X., and Zhang, Y. (2017b). Many-objective evolutionary optimization based on reference points. *Applied Soft Computing*, 50:344–355.
- Liu, Y., Ishibuchi, H., Nojima, Y., Masuyama, N., and Shang, K. (2018a). A double-niched evolutionary algorithm and its behavior on polygon-based problems. In *International Conference on Parallel Problem Solving from Nature*, pages 262–273. Springer.
- Liu, Y., Ishibuchi, H., Nojima, Y., Masuyama, N., and Shang, K. (2018b). Improving 1by1ea to handle various shapes of pareto fronts. In *International Conference on Parallel Problem Solving from Nature*, pages 311–322. Springer.
- Liu, Y., Yen, G. G., and Gong, D. (2018c). A multi-modal multi-objective evolutionary algorithm using two-archive and recombination strategies. *IEEE Transactions on Evolutionary Computation*, Early Access.
- Qi, Y., Ma, X., Liu, F., Jiao, L., Sun, J., and Wu, J. (2014). MOEA/D with adaptive weight adjustment. *Evolutionary Computation*, 22(2):231–264.
- Sato, H., Aguirre, H., and Tanaka, K. (2011). Improved s-cdas using crossover controlling the number of crossed genes for many-objective optimization. In *Proceedings of the 13th annual conference on Genetic and evolutionary computation*, pages 753–760, Deblin, Ireland. ACM.



- Takahama, T., Sakai, S., and Iwane, N. (2005). Constrained optimization by the  $\varepsilon$  constrained hybrid algorithm of particle swarm optimization and genetic algorithm. *AI 2005: Advances in Artificial Intelligence*, pages 389–400.
- Tian, Y., Cheng, R., Zhang, X., and Jin, Y. (2017). PlatEMO: A MATLAB platform for evolutionary multi-objective optimization. *IEEE Computational Intelligence Magazine*, 12(4):73–87.
- Wang, R., Ishibuchi, H., Zhang, Y., Zheng, X., and Zhang, T. (2016a). On the effect of localized pbi method in moea/d for multi-objective optimization. In *Computational Intelligence (SSCI), 2016 IEEE Symposium Series on*, pages 1–8. IEEE.
- Wang, R., Zhang, Q., and Zhang, T. (Early access, 2016b). Decomposition based algorithms using pareto adaptive scalarizing methods.
- Wang, R., Zhou, Z., Ishibuchi, H., Liao, T., and Zhang, T. (2018). Localized weighted sum method for many-objective optimization. *IEEE Transactions on Evolutionary Computation*, 22(1):3–18.
- Wang, Y. and Cai, Z. (2012). Combining multiobjective optimization with differential evolution to solve constrained optimization problems. *IEEE Transactions on Evolutionary Computation*, 16(1):117–134.
- Wang, Y., Li, H., Yen, G. G., and Song, W. (2015). Mommop: Multiobjective optimization for locating multiple optimal solutions of multimodal optimization problems. *IEEE Transactions on Cybernetics*, 45(4):830–843.
- Yang, S., Li, M., Liu, X., and Zheng, J. (2013). A grid-based evolutionary algorithm for many-objective optimization. *IEEE Transactions on Evolutionary Computation*, 17(5):721–736.
- Yuan, Y., Xu, H., Wang, B., Zhang, B., and Yao, X. (2016). Balancing convergence and diversity in decomposition-based many-objective optimizers. *IEEE Transactions on Evolutionary Computation*, 20(6):180–198.
- Zhang, Q. and Li, H. (2007). MOEA/D: A multiobjective evolutionary algorithm based on decomposition. *IEEE Transactions on Evolutionary Computation*, 11(6):712–731.
- Zhang, X., Tian, Y., and Jin, Y. (2015). A knee point driven evolutionary algorithm for many-objective optimization. *IEEE Transactions on Evolutionary Computation*, 19(6):761–776.
- Zitzler, E. and Künzli, S. (2004). Indicator-based selection in multiobjective search. In *Parallel Problem Solving from Nature-PPSN VIII*, pages 832–842, Birmingham, UK. Springer.
- Zitzler, E., Laumanns, M., and Thiele, L. (2001). SPEA2: improving the strength Pareto evolutionary algorithm. Technical report, Eidgenössische Technische Hochschule Zürich (ETH), Institut für Technische Informatik und Kommunikationsnetze (TIK).