Objective Extraction for Many-Objective Optimization Problems: Algorithm and Test Problems

Yiu-ming Cheung, *Senior Member, IEEE,* Fangqing Gu, and Hai-Lin Liu

***Abstract*—For many-objective optimization problems (MaOPs), in which the number of objectives is greater than three, the performance of most existing evolutionary multi-objective optimization algorithms generally deteriorates over the number of objectives. As some MaOPs may have redundant or correlated objectives, it is desirable to reduce the number of the objectives in such circumstances. However, the Pareto solution of the reduced MaOP obtained by most of the existing objective reduction methods, based on objective selection, may not be the Pareto solution of the original MaOP. In this paper, we propose an objective extraction method (OEM) for MaOPs. It formulates the reduced objective as a linear combination of the original objectives to maximize the conflict between the reduced objectives. Subsequently, the Pareto solution of the reduced MaOP obtained by the proposed algorithm is that of the original MaOP, and the proposed algorithm can thus preserve the dominance structure as much as possible. Moreover, we propose a novel framework that features both simple and complicated Pareto set shapes for many-objective test problems with an arbitrary number of essential objectives. Within this framework, we can control the importance of essential objectives. As there is no direct performance metric for the objective reduction algorithms on the benchmarks, we present a new metric that features simplicity and usability for the objective reduction algorithms. We compare the proposed OEM with three objective reduction methods, i.e., REDGA, L-PCA, and NL-MVU-PCA, on the proposed test problems and benchmark DTLZ5 with different numbers of objectives and essential objectives. Our numerical studies show the effectiveness and robustness of the proposed approach.**

***Index Terms*—Evolutionary algorithm, many-objective optimization, objective reduction, test problem.**

Manuscript received February 1, 2015; revised June 26, 2015, September 23, 2015, and December 31, 2015; accepted December 31, 2015. Date of publication January 19, 2016; date of current version September 30, 2016. This work was supported in part by the Faculty Research Grant of Hong Kong Baptist University under Project FRG2/14-15/075 and Project FRG1/14-15/041; in part by the National Science Foundation of China under Grant 61272366; in part by the Natural Science Foundation of Guangdong Province (2014A030313507), and in part by the Projects of Science and Technology of Guangzhou (2014J4100209, 201508010008). *(Corresponding author: Yiu-ming Cheung.)*

Y.-m. Cheung is with the Department of Computer Science, Hong Kong Baptist University (HKBU), Hong Kong 999077, the Institute of Research and Continuing Education, HKBU, Hong Kong, and with the United International College, Beijing Normal University–HKBU, Zhuhai 519000, China (e-mail: ymc@comp.hkbu.edu.hk).

F. Gu is with the Department of Computer Science, Hong Kong Baptist University, Hong Kong (e-mail: fqgu@comp.hkbu.edu.hk).

H.-L. Liu is with the Guangdong University of Technology, Guangzhou 510520, China (e-mail: hlliu@gdut.edu.cn).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TEVC.2016.2519758

I. Introduction

EVOLUTIONARY multiobjective optimization (EMO) algorithms have been successfully applied to multiobjective optimization problems [1], [2]. However, the performances of most well-known EMO algorithms, such as the nondominated sorting genetic algorithm II (NSGA-II) [3] and the improved strength Pareto evolutionary algorithm [4], seriously deteriorate over the number of objectives in solving many-objective optimization problems (MaOPs) [5], in which the number of objectives is greater than three. This deterioration is due to the poor scalability of most existing EMO algorithms, difficulty visualizing the Pareto front (PF) [6]–[8], and high computational cost. Currently, the decomposition-based multiobjective evolutionary algorithm (MOEA/D) [9], [10], indicator-based HypE [11], and the hypervolume metric selection EMO algorithm [12] have shown good performance in their application domains. Nevertheless, MOEA/D needs to design the weight vectors, which is a nontrivial task, and indicator-based algorithms suffer from high computational costs. Undoubtedly, MaOPs are more challenging than two- or three-objective problems [13].

Recently, a number of efforts have been made to deal with MaOPs, which can be roughly divided into three categories: 1) approaches that deal with MaOPs without *a priori* knowledge; 2) preference-based algorithms with *a priori* knowledge of the users' preference; and 3) objective reduction methods with *a priori* knowledge of the current nondominated solutions. In the first category, examples include modify Pareto dominance algorithms [14]–[16], indicator-based methods [17], [18], and substitute distance assignments [19]–[22]. Modify Pareto dominance algorithms [23]–[27] aim to increase the selection pressure toward the PF by decreasing the number of nondominated solutions in the population. The indicator-based methods [11] order the solutions using different fitness evaluation mechanisms. The substitute distance assignments [28]–[30] can be used instead of crowding distance to create a selection pressure and improve the convergence rate toward the PF. These algorithms cannot overcome the intrinsic difficulty of MaOPs, i.e., high computational cost and visualization difficulties. An adaptive divide-and-conquer methodology is also proposed for MaOPs, in which some objectives are independent of others [31]. The second category, i.e., preference-based algorithms [32], [33], aims for only a part of the PF and improves convergence by sacrificing diversity. From a practical perspective, it is difficult to provide the appropriate users' preference if there is a lack of knowledge of the problems, which therefore limits the applications. In contrast, objective reduction methods [34] aim to find either a minimum objective set with a threshold of an indicator measuring the PF preserving (e.g., δ-Minimum Objective Subset (δ-MOSS) [35], the off-line linear and nonlinear objective reduction algorithm based on PCA, namely, L-PCA and NL-MVU-PCA, respectively [7]) or an objective set with a prespecified size via minimizing an indicator (e.g., the minimum objective subset of size k with minimum error (k-EMOSS) [35]). In short, it transforms the MaOP into a problem with fewer objectives, and preserves the PF of the original problem, i.e., the dominance structure of the solutions, as much as possible. In particular, in some applications [36], [37] in which some of the objectives in the original MaOP would be correlated or redundant, i.e., its PF may not be spread over a wide range of the objective space, such MaOP would be reduced to a problem with fewer essential objectives by an objective reduction scheme without any loss in PF. Moreover, the objective reduction scheme can be utilized to analyze which objectives are conflicting and which are harmonious. Hereinafter, our studies will concentrate on the work within the last category.

As summarized in [7] and [34], a number of objective reduction schemes for MaOPs have been presented in recent years. The early seminal work discussing the issue of redundancy in objectives was presented by Gal and Leberling [38]. More extensive studies in this domain have been conducted in the last decade. For example, Saxena *et al.* [7] proposed a correlation-based reduction method [39] in which a set of nondominated solutions for dimensionality analysis is obtained by running NSGA-II for a large number of generations. Subsequently, the correlation matrix $\mathbf{R}$ is computed by using the objective values of the nondominated solutions. The eigenvalues and corresponding eigenvectors are then analyzed to reduce the objectives. Also, Brockhoff and Zitzler [6], [35] and Jaimes *et al.* [40] proposed a dominance structure-based reduction method that investigates how adding and omitting an objective affects the problem characteristics. Further, formal definitions of conflict and redundancy among objective sets have been discussed to find a subset of objectives such that either the entire or most of the dominance structure is preserved. Recently, Jaimes *et al.* [41], [42] developed an objective reduction scheme based on feature selection, and denoted the NSGA-II equipped with the reduction method as REDGA. In their approach, the objective set is first divided into homogeneous neighborhoods based on the correlation matrix of the objective values of a set of nondominated solutions obtained by an EMO algorithm. Then, a distance matrix based on the correlation matrix is defined to measure the conflict between the objectives. The more conflict there is between two objectives, the more distance between them in the objective "conflict" space. Thereafter, the most compact neighborhood is chosen, in which all objectives except the center one are dropped because they are the least conflicting. The above mentioned methods all try to find an objective subset from the original objectives. The numerical results have shown that these algorithms based on objective selection can successfully identify the redundant objectives. However, the Pareto solution of the reduced MaOP obtained by the objective selection-based reduction methods may not be the Pareto solution of the original problems. Given that the population size is very limited for MaOPs, the dominant solutions for the original problem in the population may greatly degrade the algorithms' performance.

Considering the limitations of the selection-based reduction methods, we propose an objective extraction method (OEM) for MaOPs. In general, a more negative correlation between two objectives indicates more conflict among them [41], [43]. That is, the correlation can be used to measure the degree of conflict among the objectives. This ideal was used in [41] and [42]. In [44], we preliminarily proposed an online objective reduction method for MaOPs based on the correlations between the reduced objectives. It works well for MaOPs with two essential objectives. This paper further proposes a novel OEM for MaOPs. It formulates the reduced objective as a linear combination of the original objectives to maximize the conflict between the reduced objectives, i.e., minimize the correlation between each pair of reduced objectives. It is a continuous and piecewise differentiable constrained optimization problem and can be solved by the subgradient projection method [45], [46]. Moreover, it can work for MaOPs with multiple essential objectives. Compared with the objective selection-based reduction methods, the proposed algorithm has the following two advantages.

1) The Pareto solution of the reduced MaOP obtained by the proposed algorithm must be that of the original MaOP.
2) It can reduce an MaOP in which the number of objectives is smaller than the number of essential objectives, in the case that the dimension of the PF is less than the number of essential objectives minus 1.

Under certain smoothness assumptions, it can be induced from the Karush–Kuhn–Tucker condition that the minimum number of objectives required (intrinsic number of conflicting objectives) is m for an $(m-1)$-dimensional PF [47]. In the example illustrated in Fig. 2 of Section II-B, the dimensionality of the PF is 1, i.e., the intrinsic number of the conflicting objectives should be 2. For this case, the proposed algorithm can ideally obtain the whole PF by solving a reduced two-objective optimization problem, whereas the objective selection-based reduction methods cannot.

In addition, performance metrics play an important role in understanding the strengths and weaknesses of the existing EMO algorithms. To the best of our knowledge, there is no direct performance metric for the objective reduction algorithms. Their performance can only be indirectly evaluated by metrics of the solutions obtained by an EMO algorithm equipped with the objective reduction method, such as IGD-metric [9] and H-metric [17]. In fact, many factors beyond the objective reduction scheme influence the quality of the solutions. It is not suitable to evaluate the performance of off-line objective reduction methods because the solutions are given. This paper presents a direct performance metric featuring the simplicity and usability of the objective reduction algorithms. However, it is not easily computed on the benchmarks, i.e., DTLZ and WFG [48], and the test problems proposed in [49]. Thus, we propose a framework that features both simple and complicated PS for many-objective test problems with an arbitrary number of essential objectives. Furthermore, one can control the importance of essential objectives and easily show the Pareto solutions visually in the decision space. The experimental results show the effectiveness and robustness of the proposed method in comparison with its existing counterparts. In summary, the major contributions of this paper are as follows.

1) Develop an OEM for MaOPs, which formulates a reduced objective as a combination of the original objectives to minimize the correlation between each pair of reduced objectives.
2) Present a framework for many-objective test problems with an arbitrary number of the essential objectives.
3) Propose a usable and intuitive performance metric for evaluating the performance of the objective reduction algorithms.
4) Empirically compare the proposed OEM with three well-known objective reduction algorithms, i.e., REDGA, L-PCA, and NL-MVU-PCA.

The remainder of this paper is organized as follows. Section II briefly introduces the basic concepts and the objective reduction problem for MaOPs. Section III gives a detailed description of the proposed objective reduction method and its characteristics, integrating the proposed objective reduction method into NSGA-II. We propose a framework for many-objective test problems in Section IV. In Section V, we briefly introduce the existing objective reduction counterparts, and the test suite constructed by the proposed framework in Section IV for the subsequent comparative studies. In Section VI, we compare the proposed OEM with three existing counterparts on ten test problems and benchmark DTLZ5 with the different numbers of objectives and essential objectives. Moreover, we compare the algorithm with the size of reduced objectives less than the number of essential objectives on two test problems. The experimental results show the effectiveness and robustness of the proposed method. Finally, we draw the conclusion in Section VII.

II. Background

A. Basic Concepts and Notations

Without loss of generality, a multiobjective optimization problem can be formulated as follows:

$$\begin{aligned} \min \quad & f(\mathbf{x}) = (f_1(\mathbf{x}), f_2(\mathbf{x}), \ldots, f_M(\mathbf{x}))^T \\ \text{s.t.} \quad & \mathbf{x} \in \Omega \end{aligned} \tag{1}$$

where $\Omega \subset \mathbb{R}^n$ is the decision space and n is the dimension of the decision variable $\mathbf{x} = (x_1, x_2, \ldots, x_n)^T$. $f : \Omega \to \mathbb{R}^M$ consists of M real-value objective functions, and T is the transpose of a matrix. When $M \geq 4$, (1) is considered as an MaOP. The optimality of a multiobjective optimization problem is defined by the concept of dominance [50]. That is, a solution $\mathbf{x}_1$ is said to dominate a solution $\mathbf{x}_2$ with respect to objective set $f' \subseteq f$ [denoted as $f'(\mathbf{x}_1) \prec f'(\mathbf{x}_2)$] if for any $f_i \in f'$, we have $f_i(\mathbf{x}_1) \leq f_i(\mathbf{x}_2)$, and there exists $f_i \in f'$ such that $f_i(\mathbf{x}_1) < f_i(\mathbf{x}_2)$. $\mathbf{x}^*$ is called a Pareto optimal solution if there is no solution $\mathbf{x} \in \Omega$ such that $f(\mathbf{x}) \prec f(\mathbf{x}^*)$. The set of all Pareto optimal solutions in Ω is called PS and denoted as $E(f, \Omega)$. Moreover, the set of Pareto optimal solutions in the objective space is called PF and denoted as $\Delta(f, \Omega)$. For a given solution set $\mathcal{X}$, a solution $\mathbf{x}^* \in \mathcal{X}$ is called the nondominated solution if it is not dominated by any member of the set $\mathcal{X}$. The set of all the nondominated solutions in $\mathcal{X}$ with respect to f is called nondominated set and is denoted as $E(f, \mathcal{X})$.

The essential objective set is defined as the smallest subset f' of objectives with $f' \subseteq f$ that can generate the same PS, i.e., $E(f, \Omega) = E(f', \Omega)$. A redundant objective set (denoted as $f \backslash f'$) refers to the set of objectives that can be eliminated without affecting the PS. The dimension of the problem refers to the number of essential objectives. Under the regularity condition, for the $(m-1)$-dimensional PF, the number of intrinsic conflicting objectives is m. The intrinsic dimension of the problem refers to the number of intrinsic conflicting objectives. Generally, the intrinsic dimension of the problem is less than or equal to its dimension.

Let $\mathcal{P} = \{\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_m\}$ with $\mathbf{p}_i \in \mathbb{R}^M$ for $i = 1, 2, \ldots, m$ be a given point set. The set

$$C(\mathcal{P}) = \left\{ \sum_{i=1}^{m} \theta_i \mathbf{p}_i \middle| \theta_i \geq 0, i = 1, \ldots, m, \sum_{i=1}^{m} \theta_i = 1 \right\} \tag{2}$$

is called the convex hull spanned by $\{\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_m\}$. Obviously, it is a convex bound set. We denote the volume of the convex hull $C(\mathcal{P})$ as $V(\mathcal{P})$. Accordingly, for each point $\mathbf{p}_i$, if $C(\mathcal{P} \backslash \mathbf{p}_i) = C(\mathcal{P})$, it is called an interior point of $C(\mathcal{P})$. Otherwise, it is called the vertex of $C(\mathcal{P})$, where $\mathcal{P} \backslash \mathbf{p}_i = \{\mathbf{p}_1, \ldots, \mathbf{p}_{i-1}, \mathbf{p}_{i+1}, \ldots, \mathbf{p}_m\}$. If $\mathbf{p}_i$ is the interior point of $C(\mathcal{P})$, $\mathbf{p}_i$ can be regarded as a combination of $\mathcal{P} \backslash \mathbf{p}_i$, i.e., $\mathbf{p}_i = \sum_{j=1, j \neq i}^{m} \theta_j \mathbf{p}_j$.

In the case in which all objectives are differentiable, the following theorem [51] states a necessary condition for Pareto optimality for unconstrained multiobjective optimization problems.

Theorem 1: Let $\mathbf{x}^*$ be a Pareto solution of (1). There exists a vector $\theta = (\theta_1, \ldots, \theta_M) \in \mathbb{R}^M$ with $\mathbf{0} \prec \theta$ and $\|\theta\|_1 = 1$ such that

$$\sum_{i=1}^{M} \theta_i \nabla f_i\big(\mathbf{x}^*\big) = 0 \tag{3}$$

where $\nabla f_i(\mathbf{x}^*)$ is the gradient of $f_i(\mathbf{x})$ at point $\mathbf{x}^*$.

B. Objective Reduction for Many-Objective Optimization Problems

As the size m of the essential objectives may be close to the number M of the original objectives, to achieve a more substantial reduction of the objective sets, there are two perspectives, in general, for objective reduction. First, given a bound of the indicator, e.g., the error δ [35] and correlation threshold T_{cor} [7], which measures the change level of the PF, the objective reduction is to find a minimum objective set by minimizing the size of the reduced objectives. Second, the objective reduction aims to find an objective set, given a prespecified size k, in which the change in the PF is as small as possible. This can be achieved by minimizing the error δ in k-EMOSS [35], or the correlation between the reduced objectives in the proposed algorithm. For example, it can be seen that either perspective mentioned above allows the PF of the original objectives to be changed to a certain degree. A problem is said to be reducible if the change of its PF through an objective reduction is not too much and bounded. In this sense, the requirement of objective reduction without PF loss is usually too strict to be used from a practical perspective.

From the first perspective, the nondominated solution set obtained by an algorithm might not be able to provide a good PF-representation, as described in [7]. It makes the number of conflict objectives with respect to the nondominated solution greater than the dimension of the original problem, as shown in [7]. Under such circumstances, it is difficult to assign an appropriate value to the threshold of the indicator. Thus, we focus on the second perspective in this paper.

Typically, the objective reduction methods with k-sized objective set can be formulated as follows. Let $\mathcal{X} = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N\}$ be N nondominated solutions with respect to f obtained by an EMO algorithm. Then, a prespecified k-sized $(k \leq M)$ objective subset can be defined as

$$f'(\mathbf{x}) = \left(f_{i_1}(\mathbf{x}), \ldots, f_{i_k}(\mathbf{x})\right)^T = \mathcal{I}^T (f_1(\mathbf{x}), \ldots, f_M(\mathbf{x}))^T \quad (4)$$

where $i_1, i_2, \ldots, i_k \in \{1, 2, \ldots, M\}$ and $\mathcal{I}_{M\times k} = [\mathbf{e}_{i_1}, \mathbf{e}_{i_2}, \ldots, \mathbf{e}_{i_k}]$ is an index matrix with

$$\mathbf{e}_{i_j} = \big(\underbrace{0, \ldots, 0}_{i_j-1}, 1, \underbrace{0, \ldots, 0}_{M-i_j}\big)^T.$$

We have $f_{i_j} = \mathbf{e}_{i_j}^T (f_1, f_2, \ldots, f_M)^T$. The objective reduction, or more accurately objective selection, can be regarded as finding the index matrix $\mathcal{I}$ according to a given rule, e.g., preserving the dominance structure with respect to $\mathcal{X}$ as much as possible in k-EMOSS [35], and maximizing the distance between the reduced objectives in REDGA [42], etc.

There are at least two limitations of the objective reduction methods based on objective selection. First, as the size of the reduced MaOP $f'(\mathbf{x})$ is smaller than the dimension of the problem, it cannot give a good approximation of that for the original MaOP $f(\mathbf{x})$. For example, the PF of a synthetic test problem is composed of three line segments that join in the center of the plane $f_1 + f_2 + f_3 = 1$ with the points $(1, 0, 0)^T$, $(0, 1, 0)^T$, and $(0, 0, 1)^T$, as illustrated in Fig. 2(a). Obviously, the intrinsic dimension and the dimension of this problem are 2 and 3, respectively. Three hundred points are randomly generated on the PF as the nondominated solutions. Generally, this problem cannot be reduced using the objective selection-based reduction methods. In this problem, the distribution of the projections of the original objective on subspaces $\{f_1, f_2\}$, $\{f_1, f_3\}$, and $\{f_2, f_3\}$ is the same. We forcibly reduce the problem to two objectives, denoted as $\{f_1, f_2\}$, and plot the projection of the nondominated solutions on subspace $\{f_1, f_2\}$, i.e., the objective values of the reduced MaOP $f'(\mathbf{x}) = (f_1, f_2)^T$, as in Fig. 2(b). In this figure, we can see that only one nondominated solution of the original MaOP is also that of the reduced MaOP. The PF of the reduced MaOP is not a good approximation of that of the original MaOP. Second, a Pareto solution of the reduced problem $f'(\mathbf{x})$ may not be a Pareto solution of $f(\mathbf{x})$, i.e., $E(f', \Omega) \nsubseteq E(f, \Omega)$, as shown in Fig. 1(a) and illustrated in Fig. 2. Suppose that a dominated solution $(0, 0, 1.1)^T$, denoted as '★', is on the f_3-axis. Its projection on subspace $\{f_1, f_2\}$ is $(0, 0)^T$, i.e., the origin in Fig. 2(b). Obviously, it is the nondominated solution of the reduced problem but not that of the original problem. In other words, for the objective reduction methods based on objective selection, the PS of the reduced problem is not a subset of that of the original problem in general. However, we expect that the PS of the reduced MaOP is a subset of the PS of the original MaOP, and the size of the PS of the reduced problem is as great as possible. As such, we propose a novel OEM in the next section.

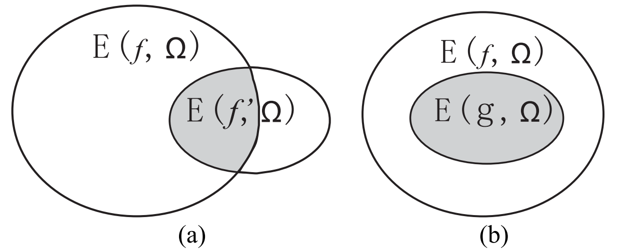

Fig. 1. Relation between the PS of the original problem and that of the reduced problem. (a) Reduced MaOP obtained by the objective reduction methods based on objective selection. (b) Reduced MaOP obtained by the proposed OEM.

III. Proposed Objective Extraction Method Based on Correlation

A. Novel Correlation-Based Objective Extraction Model

In general, a negative correlation between each pair of objectives means that one objective increases while the other decreases, and vice versa [43]. Thus, this paper infers that the more negative the correlation between two objectives, the more conflict between them. We can utilize the correlation between two objectives to measure the degree of conflict between them. Subsequently, we propose an objective extraction method based on correlation to deal with MaOPs. It is expected that the conflict between the reduced objectives should be as high as possible, i.e., the correlation between the reduced objectives with respect to the nondominated solutions $\mathcal{X}$ should be minimized. Then, the model of the objective extraction method proposed in this paper can be formulated as follows:

$$\begin{aligned} \min \quad & L(\mathbf{W}) = \sum_{i=1}^{k} \left[\max_{j, j\neq i} \rho\left(g_i, g_j\right) + \lambda \sum_{j=1, j\neq i}^{k} \mathbf{w}_i^T \mathbf{w}_j \right] \\ \text{s.t.} \quad & \mathbf{w}_i \in \mathbb{R}_+^M, \quad i = 1, \ldots, k \\ & \|\mathbf{w}_i\|_1 = 1, \quad i = 1, \ldots, k \end{aligned} \quad (5)$$

where $\mathbf{W} = [\mathbf{w}_1, \ldots, \mathbf{w}_k]$ is the weighted matrix and $\mathbf{w}_i = (w_{i1}, w_{i2}, \ldots, w_{iM})^T$ for $i = 1, \ldots, k$. The ith reduced objective $g_i(\mathbf{x}) = \mathbf{w}_i^T (f_1, \ldots, f_M)^T$ is formulated as a linear combination of the original objectives. We call "f_j within g_i" if w_{ij} is not close to zero. According to the definition of a correlation, the correlation between objectives g_i and g_j can

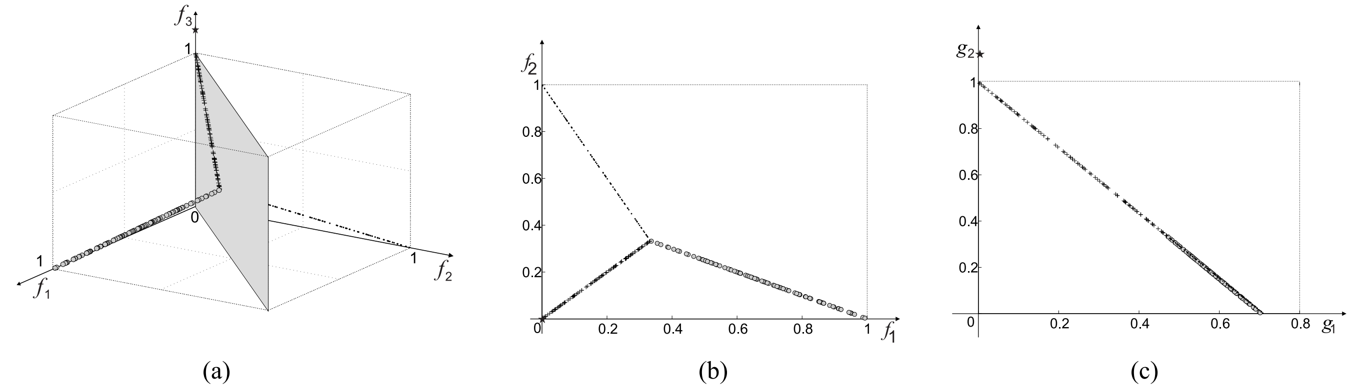

Fig. 2. Proposed objective reduction method on a synthetic test problem. (a) Nondominated set and the projection subspace spanned by w_1 and w_2. (b) Projection of the nondominated set on the subspace of f_1 versus f_2. (c) Projection of the nondominated set on the subspace spanned by w_1 and w_2.

be calculated as

$$\rho\left(g_i, g_j\right) = \frac{\mathbf{w}_i^T \mathbf{R} \mathbf{w}_j}{\sqrt{\mathbf{w}_i^T \mathbf{R} \mathbf{w}_i}\sqrt{\mathbf{w}_j^T \mathbf{R} \mathbf{w}_j}} \tag{6}$$

where $\mathbf{R}$ is the correlation matrix of $f(\mathbf{x})$ with respect to $\mathcal{X}$. The second term of the model in (5) is a penalty term that makes the solution of this model sparse and avoids overfitting the current nondominated solutions. That is, most elements of $\mathbf{w}_i$, $i = 1, \ldots, k$, approach zero. λ is the penalty coefficient, which is used to control the sparsity level. The greater the value of λ, the sparser the solution. It also ensures that the weight vectors $\mathbf{w}_i$, $i = 1, \ldots, k$, are all approximately orthogonal to each other. That is, each objective f_j is within a reduced objective g_i at most. For simplicity, λ is set at 1 in this paper. Evidently, a reduced MaOP with a size of k, denoted as $g(\mathbf{x})$, is achieved

$$\begin{aligned} &g(\mathbf{x}) = \mathbf{W}^T (f_1(\mathbf{x}), \ldots, f_M(\mathbf{x}))^T \\ &\text{s.t. } \mathbf{x} \in \Omega. \end{aligned} \tag{7}$$

Putting (6) into (5), the model of (5) can be rewritten as

$$\begin{aligned} \min \quad & L(\mathbf{W}) = \sum_{i=1}^{k} \left[\max_{j, j \neq i} \frac{\mathbf{w}_i^T \mathbf{R} \mathbf{w}_j}{\|\mathbf{w}_i\|_R \|\mathbf{w}_j\|_R} + \lambda \sum_{j=1, j \neq i}^{k} \mathbf{w}_i^T \mathbf{w}_j \right] \\ \text{s.t.} \quad & \mathbf{w}_i \in \mathbb{R}_+^M, \quad i = 1, \ldots, k \\ & \|\mathbf{w}_i\|_1 = 1, \quad i = 1, \ldots, k \end{aligned} \tag{8}$$

where $\|\mathbf{w}_i\|_{\mathbf{R}} = \sqrt{\mathbf{w}_i^T \mathbf{R} \mathbf{w}_i}$.

To solve (8), we optimize $\mathbf{w}_i$ alternately for $i = 1, \ldots, k$. That is, suppose that $\mathbf{w}_j s$ with $j = 1, \ldots, i-1, i+1, \ldots, k$ are fixed, the ith optimization problem with respect to $\mathbf{w}_i$, which is deduced from (8), is given as follows:

$$\begin{aligned} \min \quad & l_i(\mathbf{w}_i) = \max_{j, j \neq i} \frac{\mathbf{w}_i^T \mathbf{R} \mathbf{w}_j}{\|\mathbf{w}_i\|_{\mathbf{R}} \left\|\mathbf{w}_j\right\|_{\mathbf{R}}} + \lambda \sum_{j=1, j \neq i}^{k} \mathbf{w}_i^T \mathbf{w}_j \\ \text{s.t.} \quad & \mathbf{w}_i \in \mathbb{R}_+^M \\ & \|\mathbf{w}_i\|_1 = 1. \end{aligned} \tag{9}$$

As (9) is a continuous and piecewise differentiable constrained optimization problem, we can utilize the subgradient projection method [46] to solve it. The subgradient projection method is based on projecting the search direction into the subspace, tangential to the active constraints. The subgradient of the objective of (9) at point $\mathbf{w}_i^{t-1}$ obtained in the $(t-1)$th iteration is given as

$$\nabla l_i = \frac{\mathbf{R} \mathbf{w}_\tau}{\left\|\mathbf{w}_i^{t-1}\right\|_{\mathbf{R}} \|\mathbf{w}_\tau\|_{\mathbf{R}}} - \frac{\left(\mathbf{w}_i^{t-1}\right)^T \mathbf{R} \mathbf{w}_\tau \mathbf{R} \mathbf{w}_i^{t-1}}{\left\|\mathbf{w}_i^{t-1}\right\|_{\mathbf{R}}^3 \|\mathbf{w}_\tau\|_{\mathbf{R}}} + \lambda \sum_{j=1, j \neq i}^{k} \mathbf{w}_j \tag{10}$$

where $\tau = \arg\max_{j, j \neq i} (((\mathbf{w}_i^{t-1})^T \mathbf{R} \mathbf{w}_j) / (\|\mathbf{w}_i^{t-1}\|_{\mathbf{R}} \|\mathbf{w}_j\|_{\mathbf{R}}))$. To ensure that the constraints $\mathbf{w}_i \in \mathbb{R}_+^M$ and $\|\mathbf{w}_i\|_1 = 1$, the new point can be given as

$$\mathbf{w}_i^t = P\left(\mathbf{w}_i^{t-1} - \alpha \nabla l_i\right) \tag{11}$$

where $P(\mathbf{x}) = (\mathbf{x}_\varepsilon / \|\mathbf{x}_\varepsilon\|_1)$, $\mathbf{x}_\varepsilon$ is the componentwise max of ε and $\mathbf{x}$, ε is a small positive constant and let $\varepsilon = 0.0001$ in this paper. $\|\mathbf{x}_\varepsilon\|_1$ is the 1-norm of vector $\mathbf{x}_\varepsilon$, and α is the step length, which can be solved by the following problem:

$$\min_{0 < \alpha < \alpha_{\max}} l_i\left(P\left(\mathbf{w}_i^{t-1} - \alpha \nabla l_i\right)\right) \tag{12}$$

where $\alpha_{\max}$ is the maximum value of α to ensure that $w_i \in \mathbb{R}_+^M$. There exists at least one component greater than zero. This means that if $\mathbf{0} \prec \nabla l_i$, we have

$$\alpha_{\max} = \max_j \left\{ \frac{w_{ij}}{\nabla l_{ij}} \right\} \tag{13}$$

where w_{ij} and ∇l_{ij} are the jth element of $\mathbf{w}_i$ and ∇l_i, respectively. Otherwise, we set $\alpha_{\max} = 100$. Equation (12) is a one-variable optimization problem. We use the linear search to solve the best step length α. The pseudocode for computing the weighted matrix $\mathbf{W}$ is given in Algorithm 1.

B. Characteristics of the Proposed OEM

The characteristics of the proposed OEM are as follows.

1) The Pareto set of the reduced problem $g(\mathbf{x})$ obtained by the proposed OEM must be a subset of the Pareto set of the original MaOP, as shown in Fig. 1(b), which can be shown by the following proposition.

Proposition 1: Let positive matrix $\mathbf{W} = [\mathbf{w}_1, \ldots, \mathbf{w}_k]_{M \times k}$ and $g(\mathbf{x}) = \mathbf{W}^T f(\mathbf{x})$, then for any $\mathbf{x}_1, \mathbf{x}_2 \in \Omega$, if $f(\mathbf{x}_1) \prec f(\mathbf{x}_2)$, we have $g(\mathbf{x}_1) \prec g(\mathbf{x}_2)$, i.e., $E(g, \Omega) \subseteq E(f, \Omega)$.

Algorithm 1: Compute the Weighted Matrix **W**

input :
- A stopping criterion;
- The correlation matrix **R**;
- The initial weighted matrix **W**.

output: The weighted matrix **W**.

while *the stopping criteria is not met* **do**
  for $i \leftarrow 1$ **to** k **do**
    Compute the subgradient with respect to w_i by Eq. (10); perform the linear search on α to obtain the optimal step length α by solving problem:
$$\alpha_{best} \leftarrow arg \min_{0<\alpha<\alpha_{max}} l_i(P(\mathbf{w}_i - \alpha \nabla l_i));$$
    $\mathbf{w}_i \leftarrow P(\mathbf{w}_i - \alpha_{best} \nabla l_i)$.
  end
end

Proof: We only need to prove that for any $\mathbf{x}^* \notin E(f, \Omega)$, it must not be in $E(g, \Omega)$, i.e., $\mathbf{x}^* \notin E(g, \Omega)$. As $\mathbf{x}^* \notin E(f, \Omega)$, by the definition of $E(f, \Omega)$, there exists a solution $\mathbf{x} \in \Omega$ such that $f(\mathbf{x}) \prec f(\mathbf{x}^*)$, i.e., for any $f_i \in f$, we have $f_i(\mathbf{x}) \leq f_i(\mathbf{x}^*)$. Meanwhile, there exists $f_i \in f$ such that $f_i(\mathbf{x}) < f_i(\mathbf{x}^*)$. Then, we have
$$g_i(\mathbf{x}) = \sum_{j=1}^{M} w_{ij} f_j(\mathbf{x}) < \sum_{j=1}^{M} w_{ij} f_j(\mathbf{x}^*) = g_i(\mathbf{x}^*)$$
where $i = 1, \ldots, k$. This means that $g(\mathbf{x}) \prec g(\mathbf{x}^*)$, i.e., $\mathbf{x}^* \notin E(g, \Omega)$. ■

This proposition claims that any nondominated solution of the problem $g(\mathbf{x})$ is also a nondominated solution of the original problem $f(\mathbf{x})$ but not vice versa. Fig. 2(c) plots the projection of the nondominated solutions, i.e., the objective values of $g(\mathbf{x})$, on the subspace spanned by $\mathbf{w}_1 = (0.5051, 0.4948, 0.0001)^T$ and $\mathbf{w}_2 = (0.0001, 0.0001, 0.9998)^T$, which are the optimal solution of the problem in (5). Then, the projection of the dominated solution $(0, 0, 1.1)^T$ in the original objective space, also denoted as '★', is $(0.00011, 1.09978)^T$ in this subspace. Obviously, it is also a dominated solution of the reduced problem. This means that the Pareto solution obtained by the reduced problem $g(\mathbf{x})$ is also the Pareto solution of the original problem $f(\mathbf{x})$.

2) The proposed objective reduction method can preserve the dominance structure as much as possible in terms of the correlation. Subsequently, we have the following proposition.

Proposition 2: Let $\mathcal{X} = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N\}$ be N nondominated solutions with respect to the original problem f, and $g_i = \mathbf{w}_i^T f(\mathbf{x})$ with $\mathbf{w}_i$ obtained by solving the model of (5), $i = 1, \ldots, k$. Then, for any $i_1, i_2, \ldots, i_k \in \{1, 2, \ldots, M\}$, we have
$$\sum_{i=1}^{k} \max_{j, j \neq i} \rho(g_i, g_j) \leq \sum_{i=1}^{k} \max_{j, j \neq i} \rho(f_{i_i}, f_{i_j}).$$

Proof: It is easy to verify that
$$L(\mathcal{I}) = \sum_{i=1}^{k} \left[\max_{j, j \neq i} \rho(f_{i_i}, f_{i_j}) + \lambda \sum_{j=1, j \neq i}^{k} \mathbf{e}_{i_i}^T \mathbf{e}_{i_j} \right].$$

Algorithm 2: Pseudocode for the EMO Algorithm With the Proposed OEM

input :
- Maximum number of generations: T_{max};
- Number of reductions during the search: O;
- Size of the population: N;
- Number of reduced objectives: k; and
- Genetic operators and their associated parameters.

output: All of the nondominated solutions in final population $P_{O,pre}$.

◇ Initialize a random population P_0 in the decision space.
◇ Run the EMO algorithm s generations for original objectives until all of the solutions in the population are nondominated solutions.
◇ Compute the weighted matrix **W** on current nondominated solutions.
◇ $T_{pre} \leftarrow \lceil (T_{max} - s)/O \rceil$.
for $o \leftarrow 1$ **to** O **do**
  for $t \leftarrow 1$ **to** T_{pre} **do**
    ◇ Run the EMO algorithm on the new optimization problem $g(x) = \mathbf{W}^T f(\mathbf{x})$ and obtain the population $P_{o,t}$;
  end
  ◇ Compute the correlation matrix **R** of $f(\mathbf{x})$ w.r.t. $P_{o,t}$;
  ◇ Update the weighted matrix **W** as described in Algorithm 1;
end

Noting that $\mathbf{e}_{i_i}^T \mathbf{e}_{i_j} = 0$ as $i \neq j$, we have
$$L(\mathcal{I}) = \sum_{i=1}^{k} \max_{j, j \neq i} \rho(f_{i_i}, f_{i_j}).$$

Moreover, based on the model in (5), we have
$$\sum_{i=1}^{k} \max_{j, j \neq i} \rho(g_i, g_j) \leq \sum_{i=1}^{k} \left[\max_{j, j \neq i} \rho(g_i, g_j) + \lambda \sum_{j=1, j \neq i}^{k} \mathbf{w}_i^T \mathbf{w}_j \right] = L(\mathbf{W}).$$

As **W** is the optimal solution of (5), we have $L(\mathbf{W}) \leq L(\mathcal{I})$. This implies that
$$\sum_{i=1}^{k} \max_{j, j \neq i} \rho(g_i, g_j) \leq \sum_{i=1}^{k} \max_{j, j \neq i} \rho(f_{k_i}, f_{k_j}).$$
■

From this proposition, we can know that, as $\mathcal{I}$ is the optimal solution of (5), the proposed OEM is equivalent to the objective selection. This means that the objective selection-based reduction method using a correlation analysis of the objectives [41] is a special case of the proposed OEM. Moreover, as the intrinsic dimension of the problem is smaller than the dimension of the problem, the proposed algorithm can reduce the MaOP with fewer objectives compared with objective selection-based reduction methods. From the projection of the nondominated solution on the subspace spanned by $\mathbf{w}_1$ and

$\mathbf{w}_2$ [see Fig. 2(c)], we can see that the dominance structures of most solutions are preserved. We can find most of the nondominated solutions in this subspace. Thus, it is conjectured that the proposed OEM can preserve the dominance structure with fewer objectives.

C. Integration of the Proposed OEM Into EMO Algorithm

This section integrates the OEM into an EMO algorithm to deal with MaOPs. We obtain a set of nondominated solutions for an M-objective problem and initialize the weighted matrix $\mathbf{W}$. Then, we compute the weighted matrix $\mathbf{W}$ on the current nondominated solutions by Algorithm 1, and run the EMO algorithm corresponding to the new problem $g(\mathbf{x})$ as described in (7). Algorithm 2 gives the details of the EMO algorithm with the proposed OEM.

IV. Proposed Framework for Many-Objective Test Problems

A. Measuring the Performance of Given Objective Set

We introduce a direct σ-metric for measuring the performance of a given objective set f'. It is defined as

$$\sigma = \frac{\left|E\left(f', \Omega\right) \cap E(f, \Omega)\right|}{|E(f, \Omega)|} \tag{14}$$

where $|A|$ is the size/cardinality of set A. Obviously, $\sigma \in [0, 1]$. When $\sigma = 1$, it implies that the PS with respect to f' is the same as that of the original problem f. The condition that PS must not change is too strict from a practical perspective. Such a measure is helpful in that it allows the researcher to gradually tune the acceptable σ in the PS. The greater the value of σ, the better f' is. Thus, it can be used to evaluate the performance of objective reduction methods.

In this way, we define a metric below for measuring the contribution/importance of objective $f_i \in f'$ to the objective set f_i

$$\sigma_{f_i2f'} = \frac{\left|E\left(f', \Omega\right) - E\left(f' \setminus f_i, \Omega\right)\right|}{|E(f', \Omega)|} \tag{15}$$

where $E(f', \Omega) - E(f' \setminus f_i, \Omega)$ indicates that set $E(f', \Omega)$ subtracts set $E(f' \setminus f_i, \Omega)$. From the definition, we have that $\sigma_{f_i2f'} \in [0, 1]$. As $\sigma_{f_i2f'} = 0$, f_i is a redundant objective with respect to objective set f'. Moreover, the greater the value of $\sigma_{f_i2f'}$, the greater the contribution of f_i to the objective set f'.

For a given solution set $\mathcal{X}$, (14) and (15) can also be utilized to measure the performance of the objective set and the contribution of an objective to an objective set on this solution set. Then, they can be rewritten as

$$\sigma = \frac{\left|E\left(f', \mathcal{X}\right) \cap E(f, \mathcal{X})\right|}{|E(f, \mathcal{X})|} \tag{16}$$

and

$$\sigma_{f_i2f'} = \frac{\left|E\left(f', \mathcal{X}\right) - E\left(f' \setminus f_i, \mathcal{X}\right)\right|}{|E(f', \mathcal{X})|}. \tag{17}$$

Such a metric is utilized to construct the many-objective test problems, as it is convenient to control the importance of the essential objectives and thereby evaluate the performances of the objective reduction methods.

B. Framework for Many-Objective Test Problems

Some many-objective test problems are proposed in [48]. For these benchmarks, i.e., DTLZ and WFG [48], the essential objective f_i is totally indispensable to the essential objective set f', i.e., $\sigma_{f_i2f'} \to 1$. This, however, is not common from a practical perspective. Moreover, it is not easy to compute the value of σ-metric on these benchmarks.

Thus, we propose a framework for MaOPs, in which the decision space is

$$\Omega = \prod_{i=1}^{n}[a_i, b_i] \tag{18}$$

where $-\infty < a_i < b_i < \infty$ for $i = 1, \ldots, n$.

Then, the objectives of the problem $f(\mathbf{x}) = (f_1(\mathbf{x}), f_2(\mathbf{x}), \ldots, f_M(\mathbf{x}))^T$ are given as

$$f_i(\mathbf{x}) = \alpha_i(\mathbf{x}_I) + \beta_i(\mathbf{x}_{II} - \gamma(\mathbf{x}_I)) \quad \forall\ i = 1, \ldots, M \tag{19}$$

where

1) $\mathbf{x} = (x_1, x_2, \ldots, x_n) \in \Omega$, $\mathbf{x}_I = (x_1, \ldots, x_l)$ and $\mathbf{x}_{II} = (x_{l+1}, \ldots, x_n)$ are two subvectors of $\mathbf{x}$, and l is the dimension of $\mathbf{x}_I$;
2) $\mathcal{P} = \{\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_M\}$ is a point set with $\mathbf{p}_i \in \prod_{i=1}^{l}[a_i, b_i]$, and $i = 1, 2, \ldots, M$;
3) $\alpha(\mathbf{x}_I) = (\alpha_1(\mathbf{x}_I), \ldots, \alpha_M(\mathbf{x}_I))^T$ with $\alpha_i(\mathbf{x}_I) = \|\mathbf{x}_I - \mathbf{p}_i\|_2^2$ is defined as the Euclidean distance from $\mathbf{x}_I$ to $\mathbf{p}_i$, and $i = 1, 2, \ldots, M$;
4) β_i is a function from $\mathbb{R}^{n-l}$ to $\mathbb{R}^+$ for $i = 1, \ldots, M$;
5) γ is a function from $\prod_{i=1}^{l}[a_i, b_i]$ to $\prod_{i=l+1}^{n}[a_i, b_i]$.

Subsequently, regarding the PS $E(f, \Omega)$ and PF $\Delta(f, \Omega)$ of the test problem, we have the following theorem.

Theorem 2: Supposing $\beta_i(z) = 0$ if and only if $z = 0$ for $i = 1, \ldots, M$, we have:

1) $E(f, \Omega) = \{\mathbf{x} | \mathbf{x}_I \in C(\mathcal{P}), \mathbf{x}_{II} = \gamma(\mathbf{x}_I)\}$;
2) $\Delta(f, \Omega) = \{\alpha(\mathbf{x}_I) | \mathbf{x}_I \in C(\mathcal{P})\}$, where $C(\mathcal{P})$ is the convex hull spanned by $\mathcal{P}$.

Proof of Theorem 2:

1) Let $\mathbf{x} = (\mathbf{x}_I, \mathbf{x}_{II})^T \in E(f, \Omega)$. According to Schütze *et al.* [51], $E(\alpha, \prod_{i=1}^{l}[a_i, b_i]) = C(\mathcal{P})$. Thus, we only show that $\mathbf{x}_{II} = \gamma(\mathbf{x}_I)$. As $\beta_i(z) = 0$ if and only if $z = 0$, suppose that $\mathbf{x}_{II} \neq \gamma(\mathbf{x}_I)$, we then have $\beta_i(\mathbf{x}_{II} - \gamma(\mathbf{x}_I)) > 0$ for all $i = 1, \ldots, M$. That is, there exists $\mathbf{x}' = (\mathbf{x}_I, \gamma(\mathbf{x}_I))$ such that $f_i(\mathbf{x}') = \alpha_i(\mathbf{x}_I) < \alpha_i(\mathbf{x}_I) + \beta_i(\mathbf{x}_{II} - \gamma(\mathbf{x}_I)) = f_i(\mathbf{x})$ for $i = 1, \ldots, M$. This means that $f(\mathbf{x})$ is dominated by $f(\mathbf{x}')$. That is, $\mathbf{x}$ is not Pareto optimal of $f(\mathbf{x})$. This conflicts with $\mathbf{x} \in E(f, \Omega)$. Thus, we have that $\mathbf{x}_{II} = \gamma(\mathbf{x}_I)$.
2) As $\beta_i(\mathbf{x}) = 0$ for any $\mathbf{x} \in E(f, \Omega)$, we have $f(\mathbf{x}) = \alpha(\mathbf{x})$. Thus, $\Delta(f, \Omega) = \Delta(\alpha(\mathbf{x}), \prod_{i=1}^{l}[a_i, b_i])$. Consequently, we have that $\Delta(f, \Omega) = \{\alpha(\mathbf{x}_I) | \mathbf{x}_I \in C(\mathcal{P})\}$. ■

By this theorem, the PS and PF of the reduced problem $f' = (f_{i_1}, f_{i_2}, \ldots, f_{i_k})^T \subseteq f$ obtained by the objective selection-based reduction methods can be given as:

1) $E(f', \Omega) = \{\mathbf{x} | \mathbf{x}_I \in C(\mathcal{P}'), \mathbf{x}_{II} = \gamma(\mathbf{x}_I)\}$;
2) $\Delta(f', \Omega) = \{\alpha'(\mathbf{x}_I) | \mathbf{x}_I \in C(\mathcal{P}')\}$, where $\alpha'(\mathbf{x}_I) = (\alpha_{i_1}(\mathbf{x}_I), \alpha_{i_2}(\mathbf{x}_I), \ldots, \alpha_{i_k}(\mathbf{x}_I))^T$ and $\mathcal{P}' = \{\mathbf{p}_{i_1}, \mathbf{p}_{i_2}, \ldots, \mathbf{p}_{i_k}\}$ is a point set that is used to construct the reduced problem f'.

With regard to the PS and the PF of the reduced MaOP obtained by the proposed OEM on such test problem, we have the following theorem.

Theorem 3: Let non-negative matrix $\mathbf{W} = [\mathbf{w}_1, \ldots, \mathbf{w}_k]_{M\times k}$ with $\|\mathbf{w}_i\|_1 = 1$ for $i = 1, \ldots, k$, $f(\mathbf{x})$ as defined in (19). If $g(\mathbf{x}) = \mathbf{W}^T f(\mathbf{x})$, we have:

1) $E(g, \Omega) = \left\{\mathbf{x}|\mathbf{x}_I \in C(\tilde{\mathcal{P}}), \mathbf{x}_{II} = \gamma(\mathbf{x}_I)\right\}$;
2) $\Delta(g, \Omega) = \{\tilde{\alpha}(\mathbf{x}_I)|\mathbf{x}_I \in C(\tilde{\mathcal{P}})\}$, where $\tilde{\alpha}(\mathbf{x}_I) = (\tilde{\alpha}_1(\mathbf{x}_I), \tilde{\alpha}_2(\mathbf{x}_I), \ldots, \tilde{\alpha}_k(\mathbf{x}_I))^T$ with $\tilde{\alpha}_i(\mathbf{x}_I) = \sum_{j=1}^{M} w_{ij}\alpha_j(\mathbf{x}_I)$ and $\tilde{\mathcal{P}} = \left\{\tilde{\mathbf{p}}_1, \ldots, \tilde{\mathbf{p}}_k\right\}$ with $\tilde{\mathbf{p}}_i = \sum_{j=1}^{M} w_{ij}\mathbf{p}_j$, $i = 1, \ldots, k$.

Proof of Theorem 3:

1) Suppose $\mathbf{x} = (\mathbf{x}_I, \mathbf{x}_{II}) \in E(g, \Omega)$ is a Pareto solution of $g(\mathbf{x})$. We first prove that $\mathbf{x}_I \in C(\tilde{\mathcal{P}})$. As $g(\mathbf{x})$ is a differentiable function of $\mathbf{x}_I$, by Theorem 1, we have that there exist $\lambda_1, \ldots, \lambda_k \geq 0$ with $\sum_{i=1}^{k} \lambda_i = 1$ such that $\sum_{i=1}^{k} \lambda_i \bigtriangledown_{\mathbf{x}_I} g_i(\mathbf{x}) = \mathbf{0}$. The gradient of $g_i(\mathbf{x})$ on x_I is
$$\bigtriangledown_{\mathbf{x}_I} g_i(\mathbf{x}) = 2\sum_{j=1}^{M} w_{ij}\left(\mathbf{x}_I - \mathbf{p}_j\right) = 2\left(\mathbf{x}_I - \tilde{\mathbf{p}}_i\right).$$
Then
$$\sum_{i=1}^{k} \lambda_i \bigtriangledown_{\mathbf{x}_I} g_i(\mathbf{x}) = 2\sum_{i=1}^{k} \lambda_i\left(\mathbf{x}_I - \tilde{\mathbf{p}}_i\right) = 2\left(\mathbf{x}_I - \sum_{i=1}^{k} \lambda_i\tilde{\mathbf{p}}_i\right) = \mathbf{0}.$$
This means that $\mathbf{x}_I = \sum_{i=1}^{k} \lambda_i\tilde{\mathbf{p}}_i$, i.e., $\mathbf{x}_I \in C(\tilde{\mathcal{P}})$. Moreover, $g(\mathbf{x})$ is a convex function of $\mathbf{x}_I$. By Theorem 1, the Pareto set on $\mathbf{x}_I$ is equal to $C(\tilde{\mathcal{P}})$. As proven in Theorem 2, we can also show that $\mathbf{x}_{II} = \gamma(\mathbf{x}_I)$.
2) As proven in Theorem 2, for any $\mathbf{x} \in E(g, \Omega)$, we have that $\beta_i(\mathbf{x}) = 0$. This means that, for any $\mathbf{x} \in E(g, \Omega)$, $g(\mathbf{x}) = \tilde{\alpha}(\mathbf{x})$. Thus, $\Delta(g, \Omega) = \Delta(\tilde{\alpha}(\mathbf{x}), \prod_{i=1}^{l}[a_i, b_i])$. Consequently, we have that $\Delta(g, \Omega) = \{\tilde{\alpha}(\mathbf{x}_I)|\mathbf{x}_I \in C(\tilde{\mathcal{P}})\}$. ■

Theorem 4: For the MaOP $f(\mathbf{x})$ defined in (19), $f_i(\mathbf{x})$ is an essential objective, if and only if $\mathbf{p}_i$ is a vertex of $C(\mathcal{P})$.

Proof of Theorem 4:

1) *Necessary Condition:* Let $f_i(\mathbf{x})$ be an essential objective, then $E(f\backslash f_i, \Omega) \subset E(f, \Omega)$. By Theorem 2, we have $E(f\backslash f_i, \Omega) = C(\mathcal{P}\backslash\mathbf{p}_i)$ and $E(f, \Omega) = C(\mathcal{P})$. As a result, we have $C(\mathcal{P}\backslash\mathbf{p}_i) \subset C(\mathcal{P})$. Thus, $\mathbf{p}_i$ is a vertex of $C(\mathcal{P})$.
2) *Sufficient Condition:* Let $\mathbf{p}_i$ be a vertex of $C(\mathcal{P})$. Recalling the definition of a vertex, we have $C(\mathcal{P}\backslash\mathbf{p}_i) \subset C(\mathcal{P})$. On the other hand, by Theorem 2, we have $E(f\backslash f_i, \Omega) = C(\mathcal{P}\backslash\mathbf{p}_i)$ and $E(f, \Omega) = C(\mathcal{P})$. We obtain $E(f\backslash f_i, \Omega) \subset E(f, \Omega)$. This means that f_i is an essential objective. ■

Obviously, $f_i(\mathbf{x})$ is a redundant objective if and only if $\mathbf{p}_i$ is an interior point of $C(\mathcal{P})$.

[1]It can also be applied to the reduced problem obtained by the proposed OEM because the PS of the reduced problem is also a convex hull spanned $\tilde{\mathcal{P}}$ by Theorem 3.

In the proposed test problem, the size $|E(f, \Omega)|$ of the PS of f is equivalent to the volume of $C(\mathcal{P})$. Then, the value of the σ-metric defined in (14) can be calculated as
$$\sigma = \frac{V\left(\mathcal{P}'\right)}{V(\mathcal{P})}{}^{1} \tag{20}$$
and the importance $f_i \in f'$ to f' defined in (15) is equivalent to
$$\sigma_{f_i 2f'} = \frac{V\left(\mathcal{P}'\right) - V\left(\mathcal{P}' \setminus \mathbf{p}_i\right)}{V(\mathcal{P}')} \tag{21}$$
where $V(\mathcal{P})$ is the volume of the convex hull of $C(\mathcal{P})$.

It is noted as follows.

1) The PS of the problem is determined by $\alpha(\mathbf{x}_I)$ and γ. Precisely, the distribution of the solutions obtained by an algorithm can be represented by the distribution of the solutions in the l dimensional space and controlled by $\alpha(\mathbf{x}_I)$. As $l = 2$, by Theorem 2, the shape of the PS is a convex polygon in 2-D space. It is therefore easily visually represented.
2) Plotting the convex hull of $C(\tilde{\mathcal{P}})$ clearly shows that which is the identified essential objective and the importance of each reduced objective. The performance of the reduced objectives can be revealed by comparing $C(\tilde{\mathcal{P}})$ with $C(\mathcal{P})$.
3) For the reduced problems obtained by either objective selection or objective extraction, we can easily obtain the PS and PF of the reduced problem using Theorems 2 and 3, and compute the value of the σ-metric by using (20). This is convenient for understanding the performance of the algorithms equipped with the objective reduction method.
4) By Theorem 4, the dimension and the intrinsic dimension of the problem are the number of vertices of $C(\mathcal{P})$, and the number of redundant objectives is the number of the interior points of $C(\mathcal{P})$. Thus, it is easy to control the number of essential objectives.
5) By (21), we can see that $\sigma_{f_i 2f'}$ can be any value in [0, 1]. Thus, one is able to control the importance of f_i for $i = 1, \ldots, M$.

C. Modular Approach to Test Instances

This section proposes a modular approach for composing test instances with different numbers of essential objectives and variations in the importance of the essential objectives. It involves the following steps.

1) Decide the number of objectives (M), the dimension ($m \leq M$) of the problem, the dimension l of $\mathbf{x}_I$ and the dimension ($n \geq l$) of the decision variable.
2) For simplicity, we suppose that all functions β_i are the same and let $\beta_i = \beta$ for $i = 1, \ldots, M$. In this paper, $\mathbf{x}_{II} - \gamma(\mathbf{x}_I)$ is represented by $\mathbf{y}_{l:n}$, then $\beta(y_{l:n})$ in (19) is defined as follows:
$$\beta(y_{l:n}) = \sum_{j=l}^{n} y_j^2. \tag{22}$$
To visualize the PS with an arbitrary number of the essential objectives, we only consider the case of $l = 2$

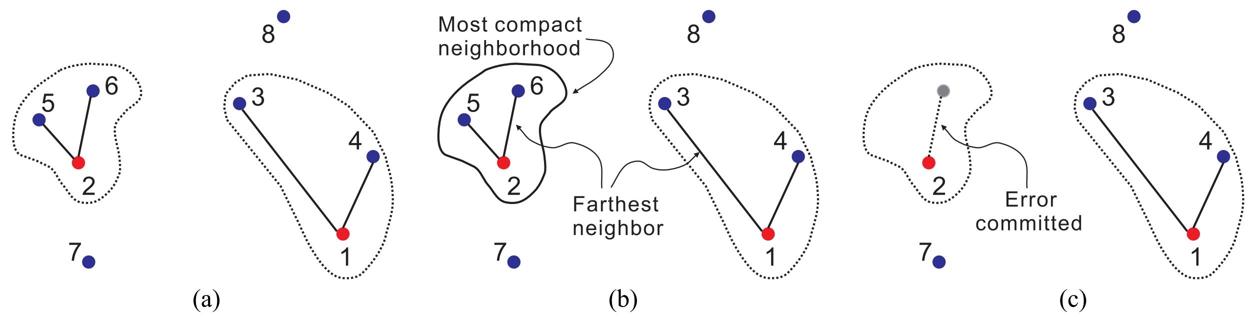

Fig. 3. Basic strategy of the objective reduction method-based feature selection (taken from [42]). (a) Divide the objective set into neighborhoods around each objective. (b) Select the most compact neighborhood. (c) Retain the center and remove the neighbors.

in this paper. Then, as in [49], we define

$$y_j = x_j - 2x_2 \sin\left(2\pi x_1 + \frac{j\pi}{n}\right)$$

where $j = 3, \ldots, n$, and $x_j \in [-2, 2]$.

3) Generate m vertices $\mathbf{p}_1, \ldots, \mathbf{p}_m$ to compose the essential objectives. As $l = 2$, we can generate the points on the circle to compose the set of vertices

$$\mathbf{p}_i = (\sin\phi_i, \cos\phi_i) \tag{23}$$

where $i = 1, \ldots, m$, and $\phi_i \in [0, 2\pi]$. If $\phi_i, i = 1, \ldots, m$ is uniformly distributed in $[0, 2\pi]$, these essential objectives have the same degree of importance. We can control the objectives' importance using the distance between the points. That is, the farther the point is from the other points, the more important the objective. Obviously, the points on the circle must be the vertices of the convex hull spanned by these points. We can use the interior points to compose the redundant objectives. Specifically, $\mathbf{p}_i, i = m+1, \ldots, M$ can be given as follows:

$$\mathbf{p}_i = \sum_{j=1}^{m} \theta_{ij}\mathbf{p}_j, i = m+1, \ldots, M \tag{24}$$

where $\theta_{ij} \geq 0$ is a randomly generated constant and $\sum_{j=1}^{m} \theta_{ij} = 1$. We also let the domain of $x_i \in [-2, 2]$ for $i = 1, 2$.

V. Compared Algorithms and Test Suite

A. Algorithms Utilized for Comparison

We compared the proposed algorithm with REDGA [42], L-PCA, and NL-MVU-PCA [7].

1) Online Objective Reduction Based on Feature Selection: Jaimes *et al.* [42] proposed an online objective reduction algorithm based on feature selection for MaOPs, namely REDGA. They have shown that this objective reduction algorithm has a lower time complexity and can improve the convergence of an EMO algorithm in MaOPs. It finds an objective subset by using a correlation matrix to estimate the conflict between each pair of objectives. It first defines the distance between two objectives (features) f_i and f_j as $1 - \rho(f_i, f_j)$, where $\rho(f_i, f_j)$ is the correlation between f_i and f_j on a given nondominated set $\mathcal{X}$. Thus, a value of 2 indicates that f_i and f_j are completely conflicting, while a value of 0 indicates that the objectives have no conflict. Then, the objective reduction algorithm based on the feature selection is divided into three steps.

1) It divides the objective set into homogeneous neighborhoods of size q around each objective. As $q = 2$, Fig. 3(a) shows a hypothetical situation with two neighborhoods and eight objectives.
2) Select the neighborhood with the minimum distance to its qth nearest neighbor. In the example, the neighborhood on the left is the most compact [see Fig. 3(b)].
3) Retain the center of the neighborhood and discard its q neighbors. As shown in Fig. 3(c), it retains the objective 2 and discards neighboring objectives 5 and 6.

Then, [42] incorporates the objective reduction method into NSGA-II and presents an online objective reduction to deal with MaOPs.

2) Linear Objective Reduction Based on Correlation Matrix: Saxena *et al.* [7] proposed a framework for the linear objective reduction algorithm, namely L-PCA. The basic idea is to find the essential objectives by eigenvalue and reduced correlation matrix analyses. In the following, we overview the L-PCA algorithm [7]. Given a nondominated set $\mathcal{X}$, the objective reduction method works as follows.

1) Compute the correlation matrix $\mathbf{R}$ of the objective values with respect to $\mathcal{X}$.
2) Compute the eigenvalues and eigenvectors of $\mathbf{R}$ with the jth principal component denoted as $\mathbf{V}_j$. The ith component of $\mathbf{V}_j$ reflects the contribution of the ith objective f_i toward $\mathbf{V}_j$.
3) Perform the eigenvalue analysis to identify the set of the important objectives. For each significant principal component $\mathbf{V}_j$, if all components have the same sign, i.e., all positive or all negative, the pair of objectives with the top two elements in terms of magnitude is picked; otherwise, the objective with the highest contribution to $\mathbf{V}_j$ by magnitude and all the other objectives with opposite sign contributions to $\mathbf{V}_j$ are selected.
4) Perform the reduced correlation matrix analysis to identify the subset of identically correlated objectives within the important objectives. If such a subset exists, the most significant objective in each subset is retained while the rest is discarded. It can further reduce the set of important objectives. The eigenvalue and reduced correlation matrix analyses are repeatedly performed until no objective can be reduced. Consequently, the reduced objective set is regarded as the essential objective set.

3) Nonlinear Objective Reduction Based on Kernel Matrix: Saxena *et al.* [7] also proposed a nonlinear objective algorithm based on a kernel matrix, namely NL-MVU-PCA. It is similar to the L-PCA, but it uses the kernel matrix $\mathbf{K}$ instead of the correlation matrix $\mathbf{R}$. The kernel matrix $\mathbf{K}$ can be obtained by solving the following problem:

$$\begin{aligned} \max \quad & \text{trace}(\mathbf{K}) = \frac{1}{2M}\sum_{ij}\left(K_{ii} + K_{jj} - 2K_{ij}\right) \\ \text{s.t.} \quad & K_{ii} + K_{jj} - 2K_{ij} = R_{ii} + R_{jj} - 2R_{ij} \ \forall \ \eta_{ij} = 1, \\ & \sum_{ij} K_{ij} = 0 \end{aligned} \tag{25}$$

where $\mathbf{K}$ is positive-semidefinite, and K_{ij} and R_{ij} are the (i, j)th element of the kernel matrix $\mathbf{K}$ and correlation matrix $\mathbf{R}$, respectively. $\eta = [\eta_{ij}]_{M\times M}$ is the neighborhood matrix, in which each element $\eta_{ij} \in \{0, 1\}$ and $\eta_{ij} = 1$ only when objective function f_i is the neighbor of f_j.

In their original versions of REDGA, L-PCA, and NL-MVU-PCA, all finds the minimum objective set as given the number q of the neighbors in REDGA and the correlation threshold T_{cor} in L-PCA and NL-MVU-PCA. The number of the reduced objectives may not equal the prespecified size k. We adjust the value of parameter q in REDGA and T_{cor} in L-PCA and NL-MVU-PCA so that the number of reduced objectives equals k.

B. Test Suite Constructed by the Proposed Framework

Ten test problems constructed by the above-stated framework are used to evaluate the performance of the proposed algorithm in this paper. The parameters for these test problems are presented in Table I, where V is the area of the convex polygon spanned by $\{\mathbf{p}_1, \ldots, \mathbf{p}_M\}$ and ϕ is used in (23) to generate the points $\{\mathbf{p}_1, \ldots, \mathbf{p}_M\}$. The dimension of the decision variable is set at 5, i.e., $n = 5$. To observe the effects of the number of objectives and essential objectives, these test problems have the different number m of objectives and the number M of essential objectives given in Table I. For the odd-numbered test problems, i.e., MAOP1, MAOP3, MAOP5, MAOP7, and MAOP9, the points used to compose the essential objective set are uniformly distributed on the unit circle. That is, all essential objectives have the same degree of importance. For the other test problems, the essential objectives have different degrees of importance.

TABLE I
PARAMETERS OF MAOP1–MAOP10

Instance	M	m	V	ϕ
MAOP1	5	3	1.2990	$[0, \frac{2}{3}\pi, \frac{4}{3}\pi]$
MAOP2	5	3	0.9330	$[0, \frac{5}{6}\pi, \frac{7}{6}\pi]$
MAOP3	10	4	2	$[0, \frac{1}{2}\pi, \pi, \frac{3}{2}\pi]$
MAOP4	10	4	1	$[0, \frac{5}{6}\pi, \pi, \frac{7}{6}\pi]$
MAOP5	10	6	2.5980	$[0, \frac{1}{3}\pi, \frac{2}{3}\pi, \pi, \frac{4}{3}\pi, \frac{5}{3}\pi]$
MAOP6	10	6	2.3660	$[0, \frac{1}{2}\pi, \frac{5}{6}\pi, \pi, \frac{7}{6}\pi, \frac{3}{2}\pi]$
MAOP7	15	5	2.3776	$[0, \frac{2}{5}\pi, \frac{4}{5}\pi, \frac{6}{5}\pi, \frac{8}{5}\pi]$
MAOP8	15	5	1.7990	$[0, \frac{2}{3}\pi, \frac{5}{6}\pi, \frac{7}{6}\pi, \frac{4}{3}\pi]$
MAOP9	15	7	2.7364	$[0, \frac{2}{7}\pi, \frac{4}{7}\pi, \frac{6}{7}\pi, \frac{8}{7}\pi, \frac{10}{7}\pi, \frac{12}{7}\pi]$
MAOP10	15	7	2.4550	$[0, \frac{1}{2}\pi, \frac{3}{4}\pi, \frac{9}{10}\pi, \frac{11}{10}\pi, \frac{5}{4}\pi, \frac{3}{2}\pi]$

C. Redundant Problems

This paper also considers the benchmark DTLZ5(m, M) [34], which is formulated as follows:

$$\begin{cases} f_1 = (1+g)\Pi_{i=1}^{M-1}\cos(\theta_i) \\ f_{l=2:M-1} = (1+g)\Pi_{i=1}^{M-l}\cos(\theta_i)\sin(\theta_{M-l+1}) \\ f_M = (1+g)\sin(\theta_1) \end{cases}$$

where $g = \sum_{i=M}^{M+k-1}(x_i - 0.5)^2$ and

$$\theta_i = \begin{cases} \dfrac{\pi}{2}x_i & i = 1, \ldots, m-1 \\ \dfrac{\pi}{4(1+g)}(1+2gx_i) & i = m, \ldots, M-1. \end{cases}$$

Subject to $\sum_{j=0}^{m-2} f_{M-j}^2 + 2^{p_i} f_i^2 \geq 1$ for $i = 1, \ldots, M-m+1$, where

$$p_i = \begin{cases} M - m & i = 1 \\ (M-m+2) - i & i = 2 : M-m+1 \end{cases}$$

and $0 \leq x_i \leq 1$ for $i = 1, \ldots, n$, $k = n - M + 1$ is the number of variables, which is used to design function g. As suggested in [48], $k = 10$ is used in the experiments. The PF occurs for the minimum of g function, i.e., at $x_i = 0.5$ for $i = M, \ldots, n$ and satisfies $\sum_{i=1}^{M} f_i^2 = 1$. The first $M - m + 1$ objectives are perfectly correlated on the PF, which makes it easier to identify the redundant objectives. The essential objective set is $\{f_i, f_{M-m+2}, \ldots f_M\}$, where $i \in \{1, \ldots, M - m+1\}$. In this problem, the dimension of the PF is $m-1$. This means that the problem has the same intrinsic dimension and dimension.

D. Problems With Different Intrinsic Dimension and Dimension

The above-mentioned test problems have the same intrinsic dimension and dimension. We construct two test instances with different intrinsic dimension and dimension using the same formulation as (19). A detailed description of the constructed test instances is as follows.

1) *MF1:* One test instance with the intrinsic dimension and dimension being 2 and 3, respectively:

$$\begin{cases} f_1 = \alpha(x_1x_2, x_1(1-x_2), 1-x_1) + \sum_{j=3}^{n}\left(x_j - 0.5\right)^2 \\ f_2 = \alpha(x_1(1-x_2), 1-x_1, x_1x_2) + \sum_{j=3}^{n}\left(x_j - 0.5\right)^2 \\ f_3 = \alpha(1-x_1, x_1x_2, x_1(1-x_2)) + \sum_{j=3}^{n}\left(x_j - 0.5\right)^2 \end{cases}$$

where

$$\alpha(a, b, c) = \begin{cases} a, & \text{if } a = \max(a, b, c) \\ 0.5 * (a + \min(b, c)), & \text{otherwise} \end{cases}$$

and, $\mathbf{x} \in [0, 1]^n$. The PF of this problem occurs at $x_j = 0.5$ for $j = 3, \ldots, n$. The PF is shown in Fig. 2(a).

2) *MF2:* The test instance MF2 is the same as MF1, with the exception of α that is defined as follows:

$$\alpha(a, b, c) = \begin{cases} 0, & \text{if } a = \min(a, b, c) \\ \dfrac{a(a+b+c)}{a + \max(b, c)}, & \text{otherwise.} \end{cases}$$

TABLE II
VALUES OF σ OBTAINED BY OEM, REDGA, L-PCA, AND NL-MVU-PCA ON THE SOLUTIONS SAMPLED ON THE TRUE PF

Instance	OEM	REDGA	L-PCA	NL-MVU-PCA
MAOP1	**0.9577**	0.4751	0.1644	0.3398
MAOP2	0.8496	0.5558	0.5558	**1.0000**
MAOP3	0.9594	0.3048	0.1683	**1.0000**
MAOP4	**0.8289**	0.6739	0.5000	0.5000
MAOP5	0.8815	0.8333	0.4941	**1.0000**
MAOP6	0.8589	0.8382	0.9717	**1.0000**
MAOP7	**0.8951**	0.4552	0.0713	0.7572
MAOP8	0.7656	0.4580	0.6576	**0.8983**
MAOP9	0.8595	0.2381	0.5680	**1.0000**
MAOP10	**0.8820**	0.7947	0.7947	0.5927

Its PF is a triangle whose vertexes are $(1, 0, 0)^T$, $(0, 1, 0)^T$, and $(0, 0, 1)^T$. Its intrinsic dimension and dimension are also 2 and 3, respectively.

VI. SIMULATION RESULTS

To study the effectiveness and robustness of the proposed algorithm, we compared it with REDGA, L-PCA, and NL-MVU-PCA. We conducted the following four experiments: 1) objective reduction methods on the solutions sampled on the true PF for MaOP suite test problems in Section VI-A; 2) NSGA-II equipped with the objective reduction methods for MaOP suite test instances in Section VI-B; 3) NSGA-II equipped with the objective reduction methods for DTLZ5 with a different number of objectives and essential objectives in Section VI-C; and 4) NSGA-II equipped with the objective reduction methods for MF1 and MF2, which have a dimension that differs from the intrinsic dimension, in Section VI-D. For the sake of convenience, we denote the NSGA-II equipped with OEM, REDGA, L-PCA, and NL-MVU-PCA as OEM, REDGA, L-PCA, and NL-MVU-PCA, respectively.

A. Experiment for the True PF of MaOP

We first applied the objective reduction algorithms to the solutions sampled on the true PF, symbolizing an unnoised signal, for the MaOP suite test problems proposed in this paper to evaluate the effectiveness of the proposed algorithm. It can quantitatively evaluate the performance of the objective reduction methods in the best of circumstances.

1) Experimental Setting: For each test problem, 1000 solutions were generated as follows. Let $\beta = 0$ in (19), and then 1000 points were uniformly generated in the convex polygon spanned by $\{\mathbf{p}_1, \ldots, \mathbf{p}_M\}$. The number of reduced objective k was set at the number of essential objectives, which is given in Table I.

2) Experimental Results: Table II lists the values of the σ-metric obtained by OEM, REDGA, L-PCA, and NL-MVU-PCA on the solutions sampled on the true PF. The results in bold are the best of those obtained using these algorithms on the test instances. It can be seen that the proposed OEM was much better than REDGA and L-PCA in terms of the σ-metric, except in the case of MAOP6. Moreover, OEM was slightly worse than NL-MVU-PCA on MAOP2, MAOP3, MAOP5, MAOP6, MAOP8, and MAOP9, but much better on MAOP1, MAOP4, MAOP7, and MAOP10. To sum up, the performance of the proposed OEM algorithm was comparable to that of NL-MVU-PCA in terms of the σ-metric. Nevertheless, the stability of the proposed algorithm outperformed NL-MVU-PCA because the OEM obtained a better result for all test instances tried thus far.

Fig. 4 shows the results obtained by OEM, REDGA, L-PCA, and NL-MVU-PCA on the solutions that were uniformly distributed on the true PF. In this figure, the convex polygon with the legend is "ideally" spanned by $\mathbf{p}_1, \ldots, \mathbf{p}_M$. The convex polygon with the legend of algorithm names is spanned by $\tilde{\mathbf{p}}_1, \ldots, \tilde{\mathbf{p}}_m$ obtained by the corresponding algorithms. The larger the area of the convex polygon, the better the performance of the objective reduction algorithm. The proposed algorithm identified the essential objectives for most test problems. Moreover, the value of its σ-metric in Table II was smaller than that obtained by NL-MVU-PCA on MAOP2, MAOP3, MAOP5, MAOP6, MAOP8, and MAOP9, but the shape of the convex polygon obtained by the OEM was very similar to the ideal convex polygon on these test problems.

B. Experiment for MaOP Suite Test Problems

We performed the NSGA-II equipped with our proposed OEM, REDGA, L-PCA, and NL-MVU-PCA, respectively, on MaOP suite test instances to study the robustness of the proposed OEM.

1) Experimental Setting: The σ-metric value defined in (20) and IGD-metric [9] were used to evaluate the performance of the algorithm. The IGD-metric is defined as follows. Suppose Q^* is a set of points that are uniformly distributed along the PF in objective space, and Q is an approximation of the PF. The distance between Q^* and Q can be defined as

$$\mathrm{IGD}\big(Q^*, Q\big) = \frac{\sum_{v \in Q^*} d(v, Q)}{|Q^*|}$$

where $d(v, Q)$ is the minimum Euclidean distance from the point v to Q and $|Q^*|$ is the cardinality of Q^*. The IGD-metric not only gives a good evaluation of the solutions' accuracy, but also their uniformity. Obviously, the smaller the value of IGD, the better the performance of the algorithm. As it is difficult to generate a set of uniform solutions in the objective space for MaOPs, we utilize the objective values of the 1000 points generated in Section VI-A to form Q^*. The distribution of these points may not be uniform in the objective space.

All algorithms ran 20 times independently for each test problem. The parameters of the algorithms were given as follows.

1) The crossover and mutation operators with the same control parameters in [10] were utilized in the four algorithms to generate new solutions.
2) The maximum number of generations is set at 300 for all test problems, i.e., $T_{\max} = 300$.

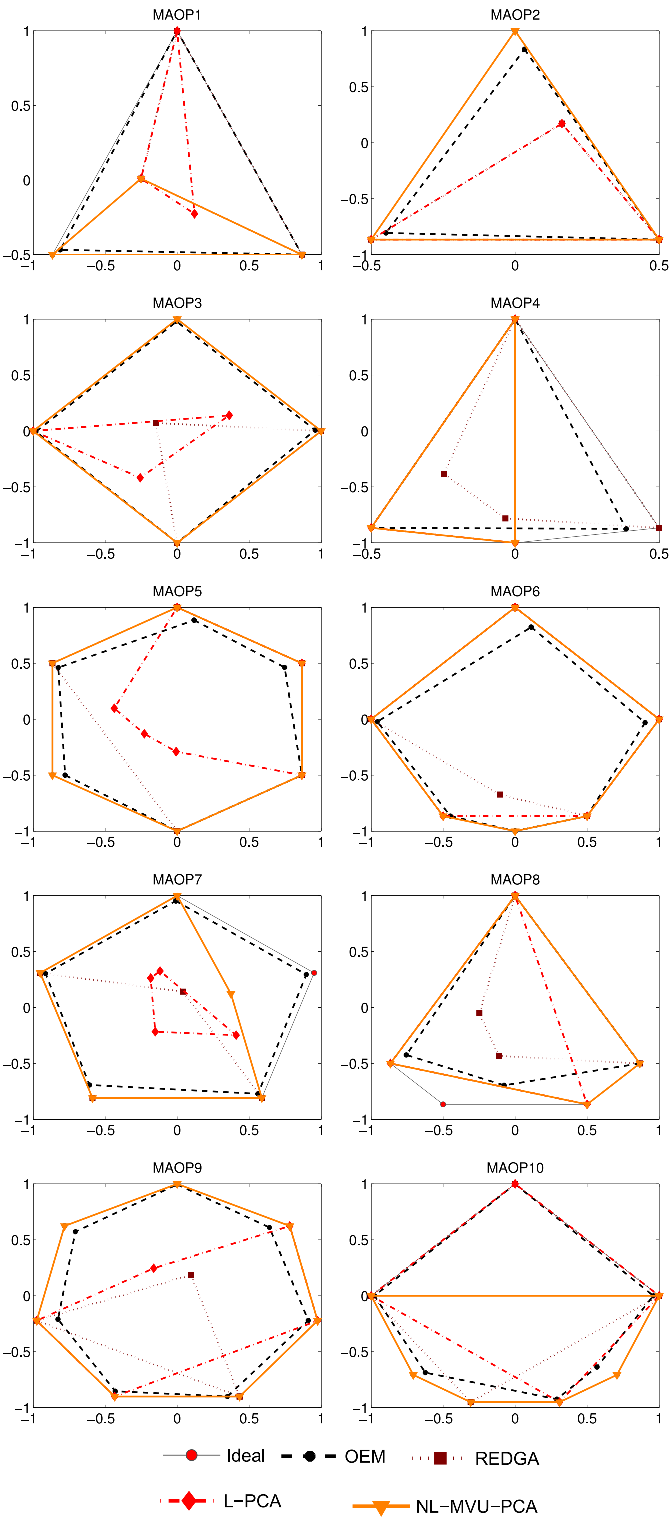

Fig. 4. Convex hull $C(\mathcal{P})$ of the original MaOP and $C(\tilde{\mathcal{P}})$ of the reduced MaOP found by OEM, REDGA, L-PCA, and NL-MVU-PCA on the true PF for the test problems MAOP1–MAOP10.

3) The number of reductions during the search is set at 10, i.e., $O = 10$.
4) The size of the population is set at 300 in all the algorithms for each test instance.
5) The size k of the reduced objectives is set at the number of essential objectives given in Table I.

TABLE III
MEAN VALUE OF σ ON THE FINAL SOLUTIONS OBTAINED BY OEM, REDGA, L-PCA, AND NL-MVU-PCA IN 20 INDEPENDENT RUNS FOR TEST PROBLEMS MAOP1–MAOP10

Instance	OEM	REDGA	L-PCA	NL-MVU-PCA
MAOP1	**0.7597**	0.1625	0.6490	0.5048
MAOP2	**0.5998**	0.5280	0.5808	0.5658
MAOP3	**0.6465**	0.0850	0.6093	0.0944
MAOP4	0.4546	**0.6841**	0.2461	0.4310
MAOP5	0.7036	**0.8333**	0.6808	0.7250
MAOP6	0.9153	0.8382	0.7509	**1**
MAOP7	0.6198	0.4259	0.2623	**0.7236**
MAOP8	**0.7931**	0.6901	0.4826	0.7766
MAOP9	**0.8014**	0.5658	0.4620	0.7785
MAOP10	**0.8169**	0.7725	0.6469	0.7578

2) Experimental Results: Table III lists the means of σ-metric values obtained by OEM, REDGA, L-PCA, and NL-MVU-PCA in 20 independent runs for test problems MAOP1–MAOP10. The results in bold are the best of those obtained using these algorithms in each test instance. This table shows that, in terms of the mean of the σ-metric, the results obtained by the proposed OEM were much better than those obtained by its counterparts, i.e., REDGA, L-PCA, and NL-MVU-PCA on MAOP1, MAOP3, MAOP9, and MAOP10. For the other test problems, the proposed algorithm also provided promising results. Furthermore, the stability of the proposed OEM was better than that of its counterparts.

Table IV lists the minimum (best), mean, and maximum (worst) values of the IGD-metric for the final solutions obtained by OEM, REDGA, L-PCA, and NL-MVU-PCA for MAOP1–MAOP10. The results in bold are the best of those obtained using these algorithms in each test instance. The results obtained by the proposed algorithm were better than those produced by its counterparts for most test problems.

Fig. 5 plots the final population in the subspace of x_1 versus x_2 and the convex polygon $C(\tilde{\mathcal{P}})$ with the median σ-metric value found by these algorithms on MAOP1–MAOP5 in the 20 independent runs. Fig. 6 plots the results on MAOP6–MAOP10. These figures reveal not only the reduced objectives obtained by the objective reduction methods but also the performance of the reduced objectives.

We demonstrated the performance of the proposed algorithm over the different values of k for MAOP6 with 20 independent runs. We chose MAOP6 as an example because it had the modest number of objectives and essential objectives, and the latter exhibited the varied importance. Table V shows the result. It can be seen that the proposed algorithm still worked well, as the value of k was no less than the number of essential objectives.

C. Experiment for Benchmark DTLZ5

In this section, we performed the algorithms on benchmark DTLZ5 with different numbers of objectives and essential objectives.

TABLE IV
MINIMUM (BEST), MEAN, AND MAXIMUM (WORST) OF IGD-METRIC VALUES FOR THE FINAL SOLUTIONS OBTAINED BY OEM, REDGA, L-PCA, AND NL-MVU-PCA IN 20 INDEPENDENT RUNS FOR TEST PROBLEMS MAOP1–MAOP10

IGD-metric	OEM			REDGA			L-PCA			NL-MVU-PCA		
Instance	Best	Mean	Worst	Best	Mean	Worst	Best	Mean	Worst	Best	Mean	Worst
MAOP1	**0.083168**	**0.187171**	**0.311488**	0.231081	0.412639	0.492908	0.127172	0.261471	0.491163	0.089532	0.280609	0.349649
MAOP2	0.113412	**0.133159**	**0.152493**	0.201748	0.223610	0.363476	0.107055	0.260699	0.750450	**0.080893**	0.220488	0.324578
MAOP3	**0.264968**	**0.330757**	**0.395403**	0.519531	0.751670	0.799833	0.265915	0.403957	0.564981	0.580062	0.590937	0.600168
MAOP4	0.122350	0.198706	0.252007	**0.117642**	**0.145814**	**0.201424**	0.156996	0.229674	0.269458	0.153059	0.200983	0.284660
MAOP5	**0.172940**	**0.201738**	**0.232778**	0.180315	0.210858	0.282748	0.197061	0.291951	0.397879	0.217554	0.307131	0.386753
MAOP6	0.155742	0.176611	0.207027	0.173579	0.187072	0.209295	0.161982	0.230852	0.399356	**0.149516**	**0.162526**	**0.173595**
MAOP7	**0.273243**	**0.342538**	**0.408537**	0.301354	0.556549	0.963198	0.529126	0.669814	0.993346	0.343980	0.378731	0.412512
MAOP8	0.178550	**0.255851**	**0.358388**	0.194515	0.261946	0.404274	0.206709	0.428795	0.987707	**0.154454**	0.257489	0.589273
MAOP9	**0.211137**	**0.227047**	**0.253707**	0.313233	0.437544	0.608080	0.364772	0.510618	0.652650	0.277228	0.297326	0.321113
MAOP10	**0.188516**	**0.205524**	**0.251822**	0.205840	0.259910	0.364430	0.228860	0.295434	0.596041	0.200985	0.274888	0.468556

TABLE V
BEST, MEAN, AND WORST σ-METRIC AND IGD-METRIC VALUES OBTAINED BY OEM WITH DIFFERENT k IN 20 INDEPENDENT RUNS FOR TEST PROBLEM MAOP6

k	σ-metric			IGD-metric		
	Best	Mean	Worst	Best	Mean	Worst
3	0.4250	0.3800	0.3646	0.3355	0.4235	0.4571
4	0.7363	0.6577	0.5678	0.2462	0.2660	0.2792
5	0.9446	0.8151	0.6685	0.1648	0.1865	0.2077
6	0.9717	0.9153	0.8507	0.1557	0.1766	0.2070
7	1.0000	0.9363	0.7925	0.1543	0.1700	0.2046
8	1.0000	0.9664	0.9405	0.1476	0.1587	0.1702

1) Experiment Setting: The IGD-metric was also used to evaluate the performances of the compared algorithms. In this experiment, as the exact PF was known, we used a set of evenly spread Pareto solutions on the PF that were generated by the methods presented in [52] to constitute Q^*. The cardinality of Q^* for these problems with the number of essential objectives $m = \{2, 3, 5, 7\}$ were set as $|Q^*| = \{1000, 1891, 10626, 18564\}$, respectively. The size k of the reduced objectives was set to be the number of essential objective given in Table VI. The other parameter settings used were as in Section VI-B.

2) Experiment Results: Table VI lists the minimum (best), mean, and maximum (worst) values of the IGD-metric for the final solutions obtained by OEM, REDGA, L-PCA, and NL-MVU-PCA for DTLZ5 with the different numbers of objectives and essential objectives. The results in bold are the best of those obtained using these algorithms in each test instance. This table shows the following.

1) All compared algorithms can obtain a perfect result on the test problems with fewer essential objectives, i.e., DTLZ5(2, 5), DTLZ5(2, 20), and DTLZ5(2, 50).
2) In terms of IGD-metric values, the results obtained by the OEM were excellent compared with those of its counterparts.

Further, we utilized DTLZ5(3, 5) as an example to analyze the weighted matrix $\mathbf{W}$ obtained by the proposed algorithm. A snapshot of the correlation matrix of the final solutions is

$$\mathbf{R} = \begin{bmatrix} 1.000 & 0.583 & 0.755 & -0.338 & -0.417 \\ 0.583 & 1.000 & 0.792 & -0.325 & -0.388 \\ 0.755 & 0.792 & 1.000 & -0.342 & -0.391 \\ -0.338 & -0.325 & -0.342 & 1.000 & -0.514 \\ -0.417 & -0.388 & -0.391 & -0.514 & 1.000 \end{bmatrix}.$$

Then, we optimized the problem (8) and obtained the weighted matrix

$$\mathbf{W}^T = \begin{bmatrix} 0.4557 & 0.3720 & 0.1723 & 0_+ & 0_+ \\ 0_+ & 0_+ & 0_+ & 1_- & 0_+ \\ 0_+ & 0_+ & 0_+ & 0_+ & 1_- \end{bmatrix}$$

where 0_+ is a small positive number and 1_- is a number smaller than 1 and very close to 1. From $\mathbf{W}^T$, it can be seen that the objectives $\{f_1, f_2, f_3\}$ are within the first reduced objectives. That is, the original objectives within the same reduced objective, i.e., $\{f_1, f_2, f_3\}$, are harmonious and the objectives within different objectives, i.e., $\{f_k, f_4, f_5\}, k \in \{1, 2, 3\}$ are conflicting.

D. Experiment for MF1 and MF2

In this section, we conducted the algorithms on MF1 and MF2 to demonstrate that the nondominated solutions of the reduced problem obtained by the proposed algorithm were also those of the original, and thus preserving the dominance structure as much as possible.

1) Experiment Setting: The IGD-metric was also used to evaluate the algorithms' performance. There were 1000 points uniformly distributed on PF to form Q^*. The success rate (SR) was the percentage of nondominated solutions for both the original problem and the reduced problem in the final solutions obtained by an algorithm. It was used to measure whether the final solutions obtained by the algorithm on the reduced algorithm were the nondominated solutions of the original problems. The maximum number of generations was set at 600 and the population size was set at 150 for the two test problems. The size m of reduced objectives was set at 2. It

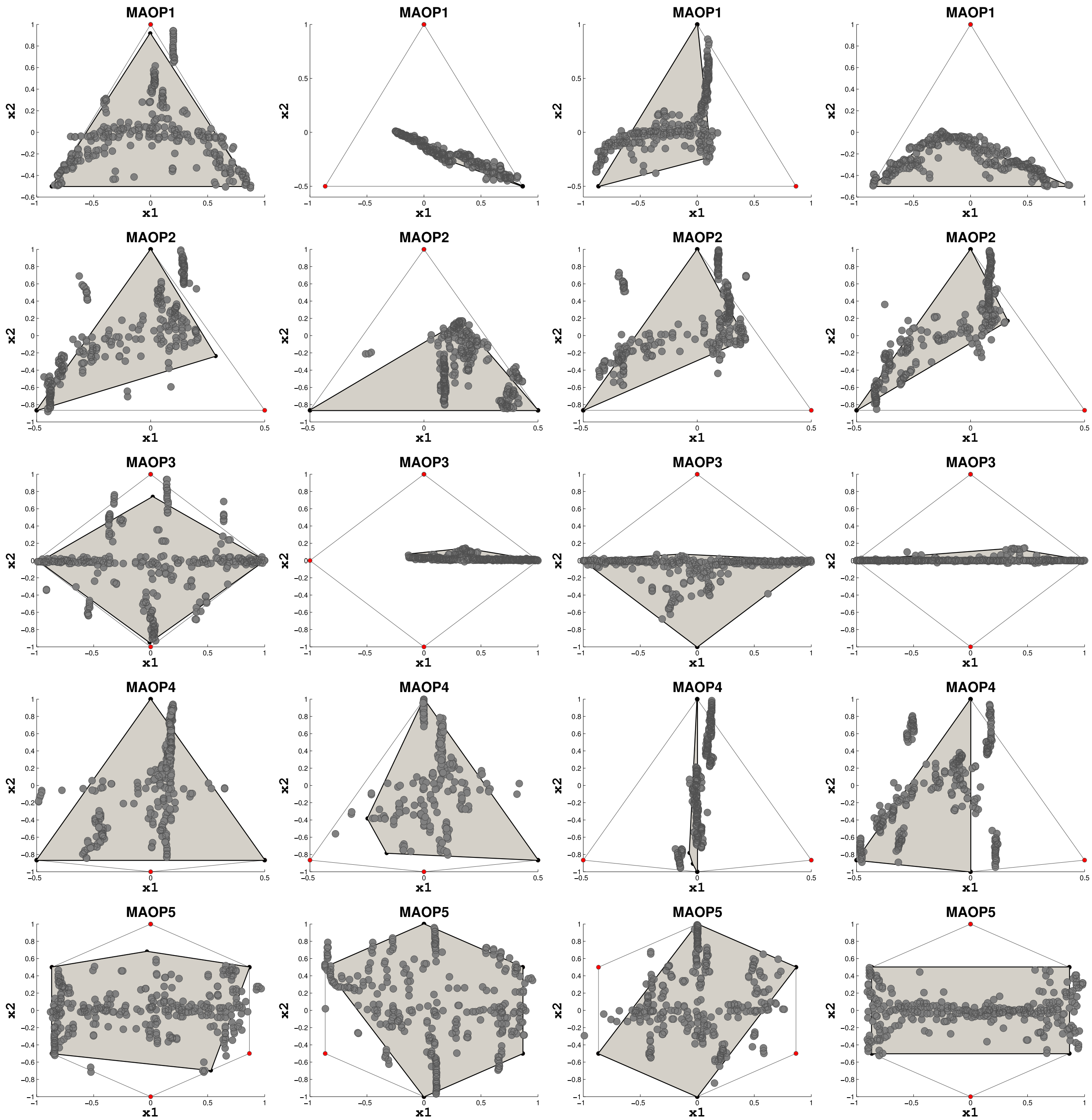

Fig. 5. Final population in the subspace of x_1 versus x_2 and $C(\tilde{\mathcal{P}})$ with the median σ-metric value found by OEM (the leftmost column), REDGA (the second leftmost column), L-PCA (the third column), and NL-MVU-PCA (the rightmost column) on MAOP1–MAOP5.

was smaller than the number of essential objectives. The other parameters were used as in Section VI-A.

2) Experiment Results: Table VII lists the minimum IGD-metric value (IGD-metric-B), mean IGD-metric value (IGD-metric-M), maximum IGD-metric value (IGD-metric-W), maximum SR value (SR-B), mean SR value (SR-M), and the minimum SR value (SR-W) obtained by OEM, REDGA, L-PCA, and NL-MVU-PCA on MF1 and MF2. From this table, we can see that the proposed algorithm, in terms of the IGD-metric, was significantly superior to the compared algorithms on these two problems. Moreover, the value of SR obtained by the proposed algorithm was 100% for all test problems. This shows that the nondominated solutions for the reduced problems obtained by the OEM must be those of the original problems. In contrast, we cannot draw such conclusion on REDGA, L-PCA, and NL-MVU-PCA.

Fig. 7 plots the PF and the final solutions with the median IGD-metric value obtained by OEM, REDGA, L-PCA, and NL-MVU-PCA on MF1 and MF2. The results obtained by the proposed algorithm were better than those of its counterparts. Ideally, the proposed algorithm would obtain the whole PF, but it may not from a practical perspective because the distribution

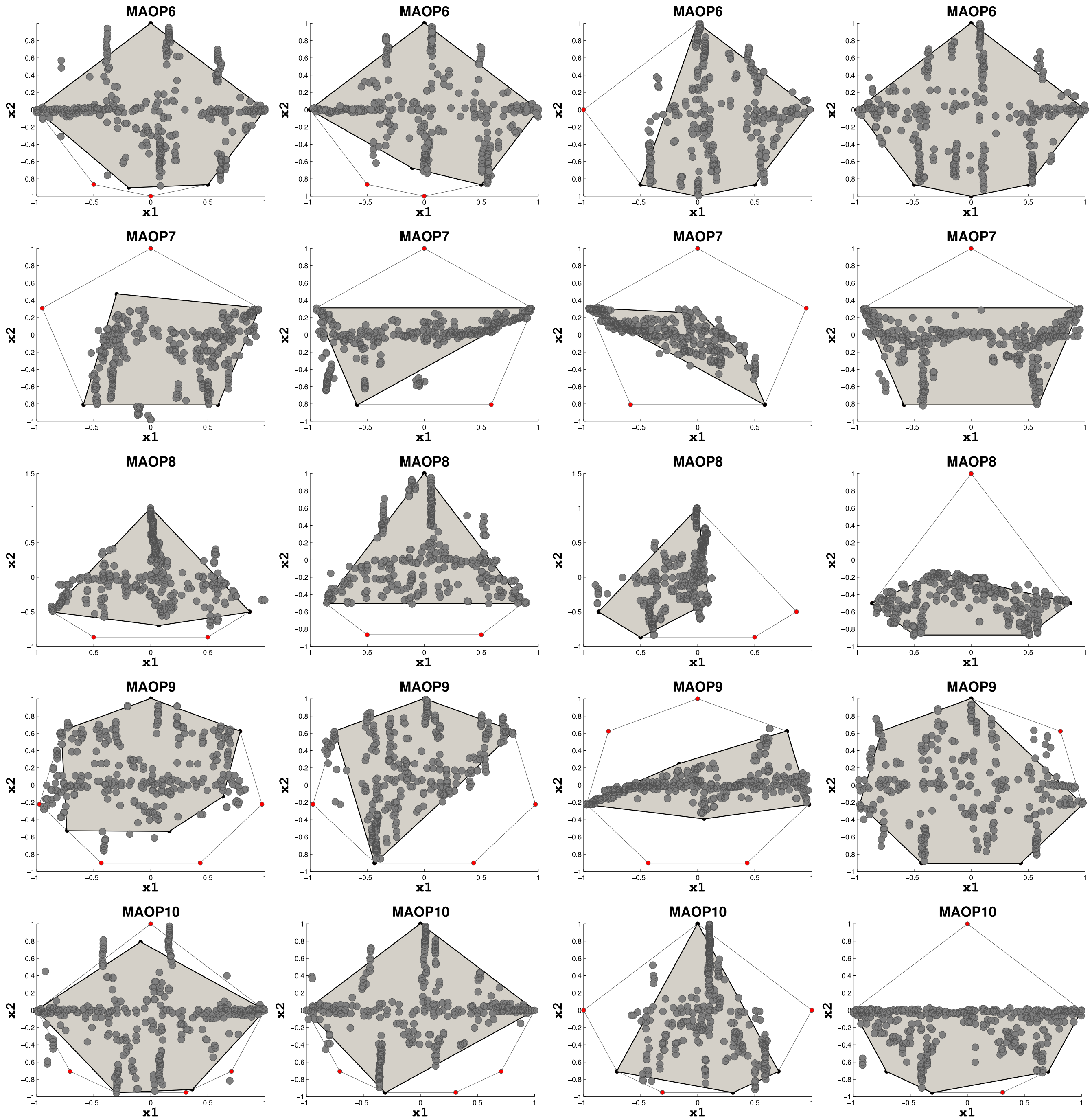

Fig. 6. Final population in the subspace of x_1 versus x_2 and $C(\tilde{\mathcal{P}})$ with the median σ-metric value found by OEM (the leftmost column), REDGA (the second leftmost column), L-PCA (the third column), and NL-MVU-PCA (the rightmost column) on MAOP6–MAOP10.

of the solution is not uniform on the PF, such that the perfect projection plane cannot be achieved.

E. Summary and Discussion

In summary, the simulation experiment results show that:

1) the proposed algorithm is of good stability, and the mean of the performance metrics is better than that of the compared algorithms for most test problems;
2) the proposed algorithm can preserve the dominance structure as much as possible, based on the results given in Section VI-D;
3) the nondominated solutions to the reduced problem obtained by the OEM are also those of the original problem.

In this paper, we assume that the value of k is prespecified. From a practical perspective, the value of k is needed when the proposed method is applied to real-world problems. Undoubtedly, an optimal value of k, i.e., the dimension of the true PF, is crucial to achieving optimal performance from the proposed algorithm. Unfortunately, as far as we know, the best way to identify an optimal value of k remains unknown, and this falls beyond the scope of this paper. We therefore

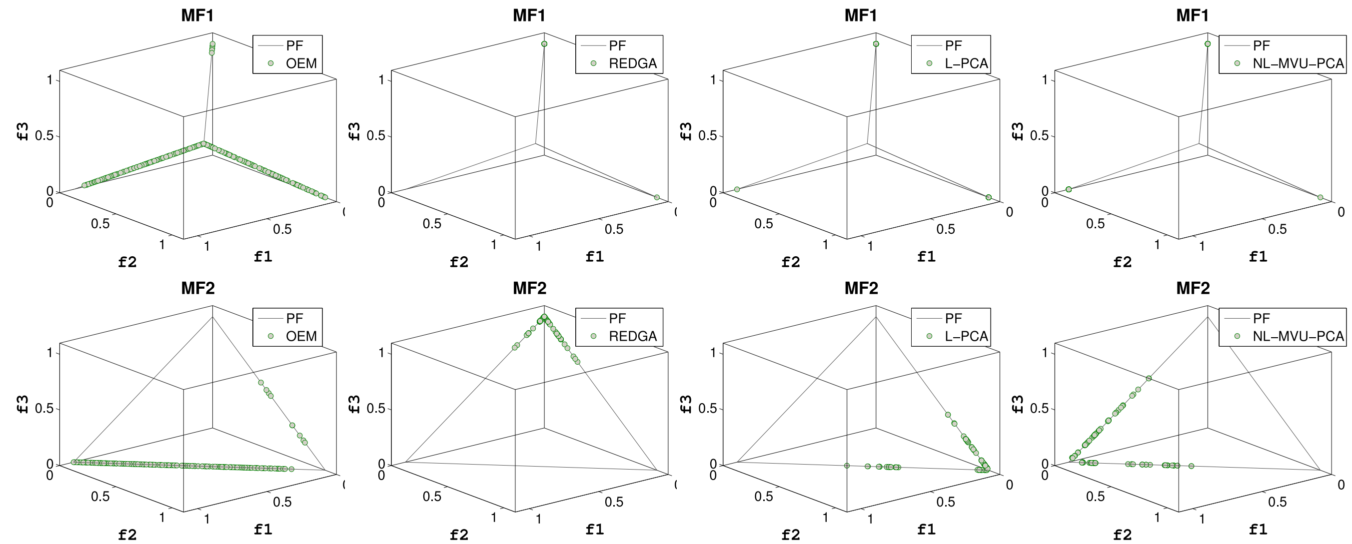

Fig. 7. PF and the final solutions with the median IGD-metric values obtained by OEM (the leftmost column), REDGA (the second leftmost column), L-PCA (the third column), and NL-MVU-PCA (the rightmost column) on test problems MF1 and MF2.

TABLE VI
MINIMUM (BEST), MEAN, AND MAXIMUM (WORST) OF IGD-METRIC VALUES FOR THE FINAL SOLUTIONS OBTAINED BY OEM, REDGA, L-PCA, AND NL-MVU-PCA IN 20 INDEPENDENT RUNS FOR DTLZ5 WITH DIFFERENT NUMBERS OF OBJECTIVES AND ESSENTIAL OBJECTIVES

IGD-metric		OEM			REDGA			L-PCA			NL-MVU-PCA		
M	m	Best	Mean	Worst	Best	Mean	Worst	Best	Mean	Worst	Best	Mean	Worst
5	2	0.001869	0.001918	0.001967	**0.001825**	0.001938	0.001988	0.001885	0.001916	**0.001962**	0.001858	**0.001913**	0.001963
5	3	0.039668	**0.040287**	**0.040975**	0.040901	0.122521	0.363443	0.039503	0.040638	0.041706	**0.038903**	0.040424	0.041622
10	5	**0.213865**	0.293870	0.466914	0.258099	**0.286460**	**0.340694**	0.542846	0.760212	1.005242	0.418291	0.688089	0.925311
10	7	0.405034	**0.630103**	**0.890672**	0.678103	0.836657	1.022040	0.890072	1.053582	1.201576	**0.397884**	0.637925	1.039015
20	2	0.001864	0.001892	0.001921	**0.001827**	**0.001869**	**0.001902**	0.001945	0.001991	0.002099	0.001914	0.001953	0.001980
20	3	0.040170	0.190676	0.343895	**0.039092**	**0.040322**	**0.040877**	0.039356	0.040806	0.042562	0.040618	0.171731	0.564702
20	5	**0.179945**	0.396376	**0.564267**	0.220436	0.453091	1.128258	0.253042	**0.385620**	0.730629	0.425335	0.507261	0.595729
20	7	**0.434135**	**0.681352**	1.028603	0.585878	0.810545	**1.027336**	0.488619	0.781078	1.205715	0.720448	0.915160	1.031057
50	2	0.001904	**0.001918**	**0.001940**	0.001918	0.002029	0.002232	**0.001891**	0.001954	0.002026	0.001907	0.002006	0.002131

TABLE VII
RESULTS OBTAINED BY OEM, REDGA, L-PCA, AND NL-MVU-PCA ON TEST PROBLEMS MF1 AND MF2

Instances	Algorithms	IGD-metric-B	IGD-metric-M	IGD-metric-W	SR-B	SR-M	SR-W
MF1	OEM	**0.059630**	**0.119737**	**0.214898**	**100**	**100**	**100**
	REDGA	0.463901	0.653663	0.863851	**100**	91	0.6
	L-PCA	0.407832	0.408170	0.408249	**100**	96	45
	NL-MVU-PCA	0.408206	0.408235	0.408249	**100**	92	6
MF2	OEM	**0.157401**	**0.219009**	**0.274432**	**100**	**100**	**100**
	REDGA	0.349411	0.607573	0.901870	**100**	83	72
	L-PCA	0.248694	0.491650	0.844487	**100**	82	72
	NL-MVU-PCA	0.213606	0.392342	0.662250	96	80	68

leave it to be studied in future work. Nevertheless, one feasible way to determine an appropriate value of k is to make multiple runs of the proposed algorithm over the different values of k, and use them to determine its value. For example, Table V shows that an appropriate value of k should be around 6 because the performance of the proposed algorithm drops significantly when $k < 6$, but it does not change much when $k \geq 6$.

Furthermore, we use correlation to measure the degree of conflict between the reduced objectives. The correlation is not totally equivalent to the conflict between the reduced objectives in some cases, especially for problems with multiple essential objectives. Thus, the results obtained by the proposed OEM for the test problems with multiple essential objectives are not very good. In the future, we will try to present a more accurate measure of the degree of conflict. In addition, a linear transformation of the space has been presented in this paper. It would be natural to study a nonlinear transformation of the space to reduce the objectives.

VII. Conclusion

In this paper, we have proposed a novel OEM for MaOPs. It formulates a reduced objective as a linear combination of the original objectives to maximize the conflict between the reduced objectives, i.e., minimize the correlation between the reduced objectives. One of its features is that the PS of the reduced MaOP obtained by the proposed algorithm is a subset of the PS of the original MaOP, and the proposed algorithm can preserve the dominance structure as much as possible. Moreover, we have introduced a framework, featuring both simple and complicated Pareto set shapes, for many-objective test problems with an arbitrary number of the essential objectives. More importantly, it can directly evaluate the performance of objective reduction algorithms through a usable and intuitive performance metric. Subsequently, we have compared the proposed objective reduction method with three objective reduction methods, i.e., REDGA, L-PCA, and NL-MVU-PCA, on the proposed test problems and a benchmark (i.e., DTLZ5) with different numbers of objectives and essential objectives. The experimental results confirm the effectiveness and robustness of the proposed approach.

References

[1] S. Nguyen, M. Zhang, M. Johnston, and K. C. Tan, "Automatic design of scheduling policies for dynamic multi-objective job shop scheduling via cooperative coevolution genetic programming," *IEEE Trans. Evol. Comput.*, vol. 18, no. 2, pp. 193–208, Apr. 2014.

[2] H.-L. Liu, F. Gu, Y.-M. Cheung, S. Xie, and J. Zhang, "On solving WCDMA network planning using iterative power control scheme and evolutionary multiobjective algorithm," *IEEE Comput. Intell. Mag.*, vol. 9, no. 1, pp. 44–52, Feb. 2014.

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

[4] Z. Eckart, L. Marco, and T. Lothar, "SPEA2: Improving the strength Pareto evolutionary algorithm for multiobjective optimization," in *Proc. Evol. Methods Design Optim. Control Appl. Ind. Probl.*, Athens, Greece, 2001, pp. 95–100.

[5] O. Schutze, A. Lara, and C. A. C. Coello, "On the influence of the number of objectives on the hardness of a multiobjective optimization problem," *IEEE Trans. Evol. Comput.*, vol. 15, no. 4, pp. 444–455, Aug. 2011.

[6] D. Brockhoff and E. Zitzler, "Are all objectives necessary? On dimensionality reduction in evolutionary multiobjective optimization," in *Proc. Parallel Probl. Solving Nat. (PPSN IX)*, Reykjavik, Iceland, 2006, pp. 533–542.

[7] D. K. Saxena, J. A. Duro, A. Tiwari, K. Deb, and Q. Zhang, "Objective reduction in many-objective optimization: Linear and nonlinear algorithms," *IEEE Trans. Evol. Comput.*, vol. 17, no. 1, pp. 77–99, Feb. 2013.

[8] D. J. Walker, R. M. Everson, and J. E. Fieldsend, "Visualizing mutually nondominating solution sets in many-objective optimization," *IEEE Trans. Evol. Comput.*, vol. 17, no. 2, pp. 165–184, Apr. 2013.

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

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

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

[12] B. Nicola, B. Naujoks, and M. Emmerich, "SMS-EMOA: Multiobjective selection based on dominated hypervolume," *Eur. J. Oper. Res.*, vol. 181, no. 3, pp. 1653–1669, 2007.

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

[14] M. Farina and P. Amato, "A fuzzy definition of 'optimality' for many-criteria optimization problems," *IEEE Trans. Syst., Man, Cybern. A, Syst., Humans*, vol. 34, no. 3, pp. 315–326, May 2004.

[15] H. Ishibuchi, N. Tsukamoto, and Y. Nojima, "Evolutionary many-objective optimization," in *Proc. 3rd Int. Workshop Genet. Evol. Syst.*, Bommerholz, Germany, 2008, pp. 47–52.

[16] C. Dai, Y. Wang, and L. Hu, "An improved α-dominance strategy for many-objective optimization problems," *Soft Comput.*, vol. 20, no. 3, pp. 1105–1111, 2016.

[17] E. Zitzler and S. Künzli, "Indicator-based selection in multiobjective search," in *Proc. Parallel Probl. Solving Nat. (PPSN VIII)*, Birmingham, U.K., 2004, pp. 832–842.

[18] H. Ishibuchi, H. Masuda, Y. Tanigaki, and Y. Nojima, "Difficulties in specifying reference points to calculate the inverted generational distance for many-objective optimization problems," in *Proc. IEEE Symp. Comput. Intell. Multi-Criteria Decis.-Making*, Orlando, FL, USA, 2014, pp. 170–177.

[19] K. Deb, J. Sundar, N. U. B. Rao, and S. Chaudhuri, "Reference point based multi-objective optimization using evolutionary algorithms," *Int. J. Comput. Intell. Res.*, vol. 2, no. 3, pp. 273–286, 2006.

[20] H. K. Singh, A. Isaacs, T. Ray, and W. Smith, "A study on the performance of substitute distance based approaches for evolutionary many objective optimization," in *Proc. Simulat. Evol. Learn.*, Melbourne, VIC, Australia, 2008, pp. 401–410.

[21] P. C. Roy, M. M. Islam, K. Murase, and X. Yao, "Evolutionary path control strategy for solving many-objective optimization problem," *IEEE Trans. Cybern.*, vol. 45, no. 4, pp. 702–715, Apr. 2015.

[22] Z. He and G. G. Yen, "Diversity improvement in decomposition-based multi-objective evolutionary algorithm for many-objective optimization problems," in *Proc. IEEE Int. Conf. Syst. Man Cybern.*, San Diego, CA, USA, 2014, pp. 2409–2414.

[23] X. Zou, Y. Chen, M. Liu, and L. Kang, "A new evolutionary algorithm for solving many-objective optimization problems," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 38, no. 5, pp. 1402–1412, Oct. 2008.

[24] H. Wang and X. Yao, "Corner sort for Pareto-based many-objective optimization," *IEEE Trans. Cybern.*, vol. 44, no. 1, pp. 92–102, Jan. 2014.

[25] L. S. Batista, F. Campelo, F. G. Guimaraes, and J. A. Ramirez, "A comparison of dominance criteria in many-objective optimization problems," in *Proc. IEEE Congr. Evol. Comput.*, New Orleans, LA, USA, 2011, pp. 2359–2366.

[26] D. Hadka and P. Reed, "Borg: An auto-adaptive many-objective evolutionary computing framework," *Evol. Comput.*, vol. 21, no. 2, pp. 231–259, 2013.

[27] S. Yang, M. Li, X. Liu, and J. Zheng, "A grid-based evolutionary algorithm for many-objective optimization," *IEEE Trans. Evol. Comput.*, vol. 17, no. 5, pp. 721–736, Oct. 2013.

[28] M. Asafuddoula, T. Ray, and R. Sarker, "A decomposition-based evolutionary algorithm for many objective optimization," *IEEE Trans. Evol. Comput.*, vol. 19, no. 3, pp. 445–460, Jun. 2015.

[29] M. Köppen and K. Yoshida, "Substitute distance assignments in NSGA-II for handling many-objective optimization problems," in *Proc. Evol. Multi-Criterion Optim.*, Matsushima, Japan, 2007, pp. 727–741.

[30] Y. Yuan, H. Xu, and B. Wang, "An improved NSGA-III procedure for evolutionary many-objective optimization," in *Proc. Conf. Genet. Evol. Comput.*, Vancouver, BC, Canada, 2014, pp. 661–668.

[31] R. C. Purshouse and P. J. Fleming, "An adaptive divide-and-conquer methodology for evolutionary multi-criterion optimisation," in *Proc. Evol. Multi-Criterion Optim.*, Faro, Portugal, 2003, vol. 2632, pp. 133–147.

[32] U. K. Wickramasinghe and X. Li, "A distance metric for evolutionary many-objective optimization algorithms using user-preferences," in *Proc. Adv. Artif. Intell.*, Melbourne, VIC, Australia, 2009, pp. 443–453.

[33] A. L. Jaimes and C. A. C. Coello, "Study of preference relations in many-objective optimization," in *Proc. 11th Annu. Conf. Genet. Evol. Comput.*, Montreal, QC, Canada, 2009, pp. 611–618.

[34] H. K. Singh, A. Isaacs, and T. Ray, "A Pareto corner search evolutionary algorithm and dimensionality reduction in many-objective optimization problems," *IEEE Trans. Evol. Comput.*, vol. 15, no. 4, pp. 539–556, Aug. 2011.

[35] D. Brockhoff and E. Zitzler, "Objective reduction in evolutionary multiobjective optimization: Theory and applications," *Evol. Comput.*, vol. 17, no. 2, pp. 135–166, 2009.

[36] A. Mukhopadhyay, U. Maulik, S. Bandyopadhyay, and C. A. C. Coello, "A survey of multiobjective evolutionary algorithms for data mining: Part I," *IEEE Trans. Evol. Comput.*, vol. 18, no. 1, pp. 4–19, Feb. 2014.

[37] M. Asafuddoula, H. K. Singh, and T. Ray, "Six-sigma robust design optimization using a many-objective decomposition-based evolutionary algorithm," *IEEE Trans. Evol. Comput.*, vol. 19, no. 4, pp. 490–507, Aug. 2015.

[38] T. Gal and H. Leberling, "Redundant objective functions in linear vector maximum problems and their determination," *Eur. J. Oper. Res.*, vol. 1, no. 3, pp. 176–184, 1977.

[39] S. Bandyopadhyay and A. Mukherjee, "An algorithm for many-objective optimization with reduced objective computations: A study in differential evolution," *IEEE Trans. Evol. Comput.*, vol. 19, no. 3, pp. 400–413, Jun. 2015.

[40] A. L. Jaimes, C. A. C. Coello, H. Aguirre, and K. Tanaka, "Objective space partitioning using conflict information for solving many-objective problems," *Inf. Sci.*, vol. 268, pp. 305–327, Jun. 2014.

[41] A. L. Jaimes, C. A. C. Coello, and D. Chakraborty, "Objective reduction using a feature selection technique," in *Proc. 10th Annu. Conf. Genet. Evol. Comput.*, Atlanta, GA, USA, 2008, pp. 673–680.

[42] A. L. Jaimes, C. A. C. Coello, and J. E. U. Barrientos, "Online objective reduction to deal with many-objective problems," in *Proc. Evol. Multi-Criterion Optim.*, Nantes, France, 2009, pp. 423–437.

[43] D. Freedman, R. Pisani, and R. Purves, *Statistics*. New York, NY, USA: Norton, 1998.

[44] Y.-M. Cheung and F. Gu, "Online objective reduction for many-objective optimization problems," in *Proc. IEEE Congr. Evol. Comput.*, Beijing, China, 2014, pp. 1165–1171.

[45] Y.-M. Cheung and F. Gu, "On solving complex optimization problems with objective decomposition," in *Proc. IEEE Int. Conf. Syst. Man Cybern.*, Manchester, U.K., 2013, pp. 2264–2269.

[46] P. H. Calamai and J. J. Moré, "Projected gradient methods for linearly constrained problems," *Math. Program.*, vol. 39, no. 1, pp. 93–116, 1987.

[47] K. Miettinen, *Nonlinear Multiobjective Optimization*. New York, NY, USA: Springer, 1999.

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

[49] D. K. Saxena, Q. Zhang, J. A. Duro, and A. Tiwari, "Framework for many-objective test problems with both simple and complicated Pareto-set shapes," in *Proc. Evol. Multi-Criterion Optim.*, Ouro Preto, Brazil, 2011, pp. 197–211.

[50] K. Deb, *Multi-Objective Optimization Using Evolutionary Algorithms*. New York, NY, USA: Wiley, 2001.

[51] O. Schütze, A. Lara, and C. A. C. Coello, "On the influence of the number of objectives on the hardness of a multiobjective optimization problem," *IEEE Trans. Evol. Comput.*, vol. 15, no. 4, pp. 444–455, Aug. 2011.

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

Yiu-ming Cheung (SM'06) received the Ph.D. degree from the Department of Computer Science and Engineering, Chinese University of Hong Kong, Hong Kong, in 2000.

He is currently a Full Professor with the Department of Computer Science, Hong Kong Baptist University, Hong Kong. His research interests include machine learning, information security, image and video processing, and pattern recognition.

Prof. Cheung is the Founding Chair of the Computational Intelligence Chapter of the IEEE Hong Kong Section and the Vice Chair of Technical Committee on Intelligent Informatics of the IEEE Computer Society. He is also a Senior Member of the Association for Computing Machinery.

Fangqing Gu received the B.S. degree from Changchun University, Jilin, China, in 2007, and the M.S. degree from the Guangdong University of Technology, Guangzhou, China, in 2011. He is currently pursuing the Ph.D. degree with the Department of Computer Science, Hong Kong Baptist University, Hong Kong.

His research interests include data mining, machine learning, and evolutionary computation.

Hai-Lin Liu received the B.S. degree in mathematics from Henan Normal University, Xinxiang, China, the M.S. degree in applied mathematics from Xidian University, Xi' an, China, the Ph.D. degree in control theory and engineering from the South China University of Technology, Guangzhou, China, and the Post-Doctoral degree from the Institute of Electronic and Information, South China University of Technology.

He is currently a Professor with the School of Applied Mathematics, Guangdong University of Technology, Guangzhou. His research interests include evolutionary computation and optimization and blind source separation.