Honda Research Institute EU
Honda Research Institute Europe GmbH
https://www.honda-ri.de/

Test Problems for Large-Scale Multi- and Many-Objective Optimization

Ran Cheng, Yaochu Jin, Markus Olhofer, Bernhard Sendhoff

2016

Preprint:

This is an accepted article published in IEEE Transactions on Cybernetics. The final authenticated version is available online at: https://doi.org/10.1109/TCYB.2016.2600577

Test Problems for Large-Scale Multiobjective and Many-Objective Optimization

Ran Cheng, Yaochu Jin, *Fellow, IEEE*, Markus Olhofer and Bernhard Sendhoff, *Senior Member, IEEE*

***Abstract*—The interests in multi- and many-objective optimization have been rapidly increasing in the evolutionary computation community. However, most studies on multi- and many-objective optimization are limited to small-scale problems, despite the fact that many real-world multi- and many-objective optimization problems may involve a large number of decision variables. As has been evident in the history of evolutionary optimization, the development of evolutionary algorithms for solving a particular type of optimization problems has undergone a co-evolution with the development of test problems. To promote the research on large-scale multi- and many-objective optimization, we propose a set of generic test problems based on design principles widely used in the literature of multi- and many-objective optimization. In order for the test problems to be able to reflect challenges in real-world applications, we consider mixed separability between decision variables and non-uniform correlation between decision variables and objective functions. To assess the proposed test problems, six representative evolutionary multi- and many-objective evolutionary algorithms are tested on the proposed test problems. Our empirical results indicate that although the compared algorithms exhibit slightly different capabilities in dealing with the challenges in the test problems, none of them are able to efficiently solve these optimization problems, calling for the need for developing new evolutionary algorithms dedicated to large-scale multi- and many-objective optimization.**

***Index Terms*—Evolutionary algorithms, multi-objective optimization, many-objective optimization, large-scale optimization, test problems**

I. Introduction

MULTI- and many-objective optimization problems involve more than one conflicting objective to be optimized simultaneously, which can be formulated as follows:

$$\begin{aligned} \min_{\mathbf{x}} \quad & \mathbf{f}(\mathbf{x}) = (f_1(\mathbf{x}), f_2(\mathbf{x}), ..., f_M(\mathbf{x})) \\ \text{s.t.} \quad & \mathbf{x} \in X \end{aligned} \tag{1}$$

where $X \subseteq \mathbb{R}^n$ is the decision space with $\mathbf{x} = (x_1, x_2, ..., x_n) \in X$ being the decision vector. Due to the conflicting nature of the objectives, a set of optimal solutions representing the trade-off between different objectives, termed Pareto optimal solutions, can be achieved. The Pareto optimal solutions are known as the Pareto front (PF) in the objective space and the Pareto set (PS) in the decision space. Usually, multi-objective optimization problems (MOPs) refer to those with two or three objectives, while those with more than three objectives are known as many-objective optimization problems (MaOPs) [1]–[4].

Evolutionary algorithms (EAs), which are population-based, are well suited for solving MOPs and MaOPs as they can obtain a solution set in one single run [5]. In general, multi-objective evolutionary algorithms (MOEAs) can be categorized into three groups, including dominance based methods [6]–[8], decomposition based methods [9]–[11], and performance indicator based methods [12], [13]. In addition to the conventional MOEAs, a few MOEAs dedicated to solving MaOPs have recently been developed [14]–[19].

To evaluate the performance of various MOEAs, several test suites have been designed for empirical studies. Among many others, the ZDT test suite [20] and DTLZ test suite [21], [22] are the most widely used ones.

The ZDT test suite is one of the most popular test suites in the multi-objective optimization literature [20]. Based on the generic design principles proposed by Deb in [23], the test problems in the ZDT suite are constructed by introducing three basic functions, including a distribution function f_1, a distance function g and a shape function h, where function f_1 is designated to test the ability of an MOEA to maintain diversity along the PF, function g is meant for testing the ability of an MOEA to converge to the PF and function h for defining the shape of the PF. There are six test problems in the ZDT test suite, five of which (ZDT1 to ZDT4, ZDT6) are real-coded and one (ZDT5) is binary-coded. Different ZDT test problems have different characteristics. Specifically, ZDT3 has a disconnected PF, which is partly convex and partly concave; ZDT4 contains a large number of local PFs; ZDT6 has a non-uniform fitness landscape, which causes a biased distribution of the Pareto optimal solutions along the PF. However, despite its immense popularity, the ZDT test suite has a significant limitation, i.e., all test problems are bi-objective.

To address the limitation of the ZDT test suite (as well as many other bi-objective test problems), Deb *et al.* proposed another test suite, i.e., the DTLZ test suite [21], [22], in which the test problems are scalable to have any number of objectives. There are nine test problems in the DTLZ test suite, each of which is constructed with the same design principle, where the first $M-1$ decision variables define the PF, while the rest decision variables specify the convergence property. The DTLZ test suite has many unique characteristics. For instance, the fitness landscape of DTLZ1 and DTLZ3 contain a large number of local PFs; the distribution of the Pareto optimal solutions of DTLZ4 is highly non-uniform; the PFs of DTLZ5 and DTLZ6 are a degenerate curve; DTLZ7 has a disconnected

Ran Cheng and Yaochu Jin are with the Department of Computer Science, University of Surrey, Guildford, Surrey, GU2 7XH, United Kingdom (e-mail: ranchengcn@gmail.com; yaochu.jin@surrey.ac.uk). Y. Jin is also with the College of Information Sciences and Technology, Donghua University, Shanghai 201620, P. R. China. (*Corresponding author: Yaochu Jin*)

Markus Olhofer and Bernhard Sendhoff are with the Honda Research Institute Europe, 63073 Offenbach, Germany (e-mail: {markus.olhofer;bernhard.sendhoff}@honda-ri.de).

PF; and DTLZ8 and DTLZ9 are constrained problems. One significant contribution of the DTLZ test suite is the proposal of a generic design principle for constructing test problems that are scalable to have any number of objectives, as well as decision variables.

Some variants of the DTLZ test suite have also been developed. To assess the performance of MOEAs on highly scaled problems [16], Deb *et al.* proposed a method to scale the value of each objective function to a different range. In [24], some constrained DTLZ problems have been suggested to verify the constraint handling capability of MOEAs. However, in spite its immense popularity, the DTLZ test suite still has some shortcomings. For example, it doses not take into account some important characteristics commonly seen in the real-world problems, such as variable linkage and variable separability. In this work, separability refers to the correlation relationship between the decision variables in the entire decision space, while variable linkage is used to characterize the relationship between the decision variables of Pareto optimal solutions.

To remedy the shortcomings of the DTLZ test suite, Huband *et al.* proposed another test suite, i.e., the WFG test suite [25], [26]. The WFG test suite has incorporated a variety of important characteristics that widely exist in real-world problems. To construct a test problem, it only requires to specify a shape function, which determines the PF, and a transformation function, which describes the fitness landscape. In this test suite, WFG1, WFG7 and WFG9 have biased PFs; WFG5 and WFG9 have deceptive fitness landscapes; WFG2, WFG3, WFG6, WFG8 and WFG9 have non-separable fitness landscapes. Since these test problems are also scalable to have any number of objectives, the WFG test suite becomes another prevalent benchmark for many-objective optimization, in addition to the DTLZ test suite.

Apart from the above three general-purpose test suites, test problems have also been constructed to include some specific characteristics. In [27], Okabe *et al.* have proposed some design principles for constructing test problems with an arbitrarily complex PS, which were generalized in [28]. More recently, Saxena *et al.* have extended the test problems with complex PSs for many-objective optimization [29]. Other modified ZDT and DTLZ test problems by introducing linear or nonlinear variables into decision variables can be found in [30] [31]. While the test problems discussed above are static, dynamic multi-objective optimization test problems have also been proposed in [32], where the PFs and/or PSs change over time. Some variants of these test problems can also be found in [33], [34].

However, in spite of the various test problems above, none of them is designed for large-scale optimization by explicitly taking considering different characteristics of decision variables, even if they are theoretically scalable to have any number of decision variables. Despite the fact that large-scale optimization has attracted much attention in single-objective optimization [35]–[37], little work has been reported on large-scale multi- and many-objective optimization with few exceptions [38]–[40]. This can partly be attributed to the lack of large-scale multi- and many-objective optimization test problems, unlike the rich literature on large-scale single-objective optimization [41]–[44].

This work aims to propose a set of test problems for large-scale multi- and many-objective optimization. The main new contributions of this work can be summarized as follows:

1) Following a few basic design principles, new considerations have been introduced in constructing test problems for large-scale multi- and many-objective optimization, including non-uniform correlations between decision variables and objective functions, and mixed separability in the fitness landscape. To the best of our knowledge, this is the first time that such properties have been taken into account in multi- and many-objective test problems.
2) A generic method for defining correlations between the decision variables and the objective functions has been proposed. With a correlation matrix, any relationship between decision variables and objectives can be easily defined.
3) Based on the proposed design principles and considered characteristics, nine large-scale multi- and many-objective test problems (LSMOPs for short hereafter) have been constructed.

The remainder of this paper is organized as follows. Section II introduces a few fundamental concepts in large-scale optimization. Section III and Section IV present the basic design principles and problem characteristics, respectively. On the basis of these, nine LSMOPs are proposed in Section V. Section VI presents the empirical results of six selected MOEAs for multi- and many-objective optimization on the proposed test problems. Finally, Section VII draws the conclusion.

II. Background

In this section, we first go through a few basic concepts in large-scale optimization, such as separability and non-separability, followed by a brief review of some representative test suites for large-scale single-objective optimization.

A. Basic Concepts in Large-Scale Optimization

Large-scale optimization generally refers to the solution of optimization problems that involve hundreds or even thousands of decision variables [45]–[47]. As pointed by Weise *et al.* in [48], several factors make large-scale optimization problems extremely difficult. For example, with the increase of the number of decision variables, the volume of search space grows exponentially, and the complexity of the fitness landscape increases rapidly as well. Nevertheless, the major difficulty comes from the interactions between decision variables, often known as the *non-separability*. To be specific, an objective function $f(\mathbf{x})$ is said to be *separable* with respect to a decision variable x_k if the following condition is fulfilled [35]:

$$f(\mathbf{x}) < f(\mathbf{x}') \Rightarrow f(\mathbf{y}) < f(\mathbf{y}') \tag{2}$$

where $\mathbf{x} = [x_1, ..., x_k, ..., x_n]$, $\mathbf{x}' = [x'_1, ..., x'_k, ..., x'_n]$, $\mathbf{y} = [y_1, ..., x_k, ..., y_n]$ and $\mathbf{y}' = [y'_1, ..., x'_k, ..., y'_n]$ are four different decision vectors. Otherwise, $f(\mathbf{x})$ is considered to be *non-separable* with respect to x_k, which means that x_k has interactions with one or more other decision variables. Specially, if the decision variables only interact with some

(rather than all) of the other decision variables, the objective function $f(\mathbf{x})$ is known as a *partially separable* function.

Several important observations can be made with regard to separability, non-separability and partial separability [44], which are listed below.

A decision variable x_k can be optimized independently iff the objective function $f(\mathbf{x})$ is separable with respect to it:

$$\underset{\mathbf{x}}{\operatorname{argmin}} f(\mathbf{x}) = \left(\underset{x_k}{\operatorname{argmin}} f(\mathbf{x}), \underset{\forall x_j, j \neq k}{\operatorname{argmin}} f(\mathbf{x}) \right) \tag{3}$$

The decision variables can be optimized component-wise independently iff an objective function $f(\mathbf{x})$ is partially separable:

$$\underset{\mathbf{x}}{\operatorname{argmin}} f(\mathbf{x}) = \left(\underset{\mathbf{x_1}}{\operatorname{argmin}} f(\mathbf{x_1}, ...), ..., \underset{\mathbf{x_m}}{\operatorname{argmin}} f(..., \mathbf{x_m}) \right) \tag{4}$$

where $\mathbf{x_1}$, ..., $\mathbf{x_m}$ are disjoint sub-vectors of $\mathbf{x}$, and the interactions only exist among decision variables inside each sub-vector, but the decision variables in different sub-vectors do not interact with each other.

B. Benchmark Test Suites for Large-Scale Single-objective Optimization

The first benchmark suite for large-scale single-objective optimization was proposed by Tang *et al.* in the CEC'2008 special session and competition on large-scale global optimization [41]. This suite is now known as the CEC'2008 suite consisting of seven test problems. This is the first attempt to explicitly include the characteristics of separability and non-separability into test problems for large-scale optimization. However, a major weakness of this test suite is that the problems are designed to be either fully separable or fully non-separable, which are two extreme cases of most real-world problems.

The CEC'2008 test suite was extended by Tang *et al.* to construct the CEC'2010 test suite [42]. The major contribution of CEC'2010 test suite is the proposal of a modularity based design principle, which divides the decision variables into different subcomponents, and the separability of each subcomponent is independent. Consequently, the test problems can be fully separable, partially separable or fully non-separable by combining separable and non-separable subcomponents. The CEC'2010 test suite has motivated the development of techniques such as random grouping [35], cooperative co-evolution [36], and differential grouping [49], which are used to detect interactions between groups of decision variables and solve them in a *divide-and-conquer* manner.

The CEC'2013 test suite has been proposed on the basis of the CEC'2010 test suite to include a few new characteristics [43]. First, some subcomponents in the CEC'2013 test suite are overlapped, while the subcomponents of the decision variables in the CEC'2010 test suite are completely independent. Second, the size of all the subcomponents is fixed in the CEC'2010 test suite, while in the CEC'2013 test suite, the subcomponents are of different sizes, such that they have non-uniform contributions to the objective function.

III. Design Principles

In the design of the proposed test problems, we follow four basic principles as suggested in [21] and [26]:

1) The test problems can be generated with a uniform design formulation.
2) The test problems should be scalable to have an arbitrary number of objective functions.
3) The test problems should be scalable to have any number of decision variables.
4) The exact shapes and locations of the PFs are known.

The above four principles guarantee a good extensibility and generality of the test problems. To fulfil these four design principles, the following uniform design formulation is adopted:

$$\begin{cases} f_1(\mathbf{x}) = h_1(\mathbf{x}^f)(1 + g_1(\mathbf{x}^s)) \\ f_2(\mathbf{x}) = h_2(\mathbf{x}^f)(1 + g_2(\mathbf{x}^s)) \\ \quad ... \\ f_M(\mathbf{x}) = h_M(\mathbf{x}^f)(1 + g_M(\mathbf{x}^s)) \end{cases} \tag{5}$$

where f_1 to f_M are the objective functions, $\mathbf{x}^f = (x_1, ..., x_{m-1})$ is the first part of the decision vector and $\mathbf{x}^s = (x_m, ..., x_n)$ is the other part, functions $h_1, ..., h_M$ together define the shape of the PF, known as the *shape functions* hereafter, and functions g_i define the fitness landscape, known as the *landscape functions* hereafter.

If we denote $\mathbf{F}(\mathbf{x}) = [f_1(\mathbf{x}), ..., f_M(\mathbf{x})]$, $\mathbf{H}(\mathbf{x}^f) = [h_1(\mathbf{x}^f), ..., h_M(\mathbf{x}^f)]$, and $\mathbf{G}(\mathbf{x}^s) = \text{diag}(g(\mathbf{x}^s), ..., g(\mathbf{x}^s))$, (5) can be rewritten as:

$$\mathbf{F}(\mathbf{x}) = \mathbf{H}(\mathbf{x}^f)(\mathbf{I} + \mathbf{G}(\mathbf{x}^s)), \tag{6}$$

where $\mathbf{I}$ is an identity matrix, $\mathbf{H}(\mathbf{x}^f)$ is known as the *shape matrix* and $\mathbf{G}(\mathbf{x}^s)$ is known as the *landscape matrix*.

Given a problem as formulated by (6), the optimization target is to find the PS, denoted as $\mathbf{x}_* = (\mathbf{x}_*^f, \mathbf{x}_*^s)$, such that $\mathbf{G}(\mathbf{x}_*^s) = \mathbf{O}$ and $\mathbf{F}(\mathbf{x}_*) = \mathbf{H}(\mathbf{x}_*^f)$. Therefore, by using such a design formulation, $\mathbf{H}(\mathbf{x}^f)$ is able to test the ability of an algorithm to achieve diverse solutions along the PF, and $\mathbf{G}(\mathbf{x}^s)$ is able to test the ability of an algorithm to converge to the PF.

Since the motivation of this work is to introduce some important characteristics into large-scale fitness landscapes of the proposed test problems, we focus on the design of $\mathbf{G}(\mathbf{x}^s)$, which defines the fitness landscape. While for $\mathbf{H}(\mathbf{x}^f)$, which defines the shape of the PF, we simply refer to the existing definitions in [21]. In the following section, we will introduce the detailed characteristics of the proposed test problems with respect to the design of $\mathbf{G}(\mathbf{x}^s)$.

IV. Test Problem Characteristics

There are four main characteristics of the proposed test problems with regard to the design of the landscape matrix $\mathbf{G}(\mathbf{x}^s)$ as described in (6):

1) The decision variables are non-uniformly divided into a number of groups.
2) Different groups of decision variables are correlated with different objective functions.

3) The decision variables have mixed separability.
4) The decision variables have linkages on the PSs.

Characteristic (1) and characteristic (3) follow the design principles of large-scale single-objective optimization problems, as recommended in [44]; characteristic (4) follows the suggestions in [27] and [28] to take into consideration the variable linkages on the PSs, while characteristic (2) is to reflect the scenarios in real-world conceptual design [50]–[52].

A. Non-uniform Grouping of Decision Variables

The decision vector $\mathbf{x}^s$ is non-uniformly divided into M groups as follows:

$$\mathbf{x}^s = (\mathbf{x}_1^s, ..., \mathbf{x}_M^s), \tag{7}$$

where M is the number of objectives, and $\sum_{i=1}^{M} |\mathbf{x}_i^s| = |\mathbf{x}^s| = n_s$, with n_s being a predefined parameter that specifies the total number of decision variables in $\mathbf{x}^s$. Moreover, each group further consists of a number of n_k subcomponents:

$$\mathbf{x}_i^s = (\mathbf{x}_{i,1}^s, ..., \mathbf{x}_{i,n_k}^s), \tag{8}$$

where n_k is a predefined parameter that specifies the number of subcomponents in each group of decision variables.

To guarantee that the non-uniform group sizes are constant in different independent runs, we propose a chaos-based pseudo random number generator:

$$\begin{cases} |\mathbf{x}_i^s| = \lceil r_i \times n_s \rceil, \\ r_i = \dfrac{c_i}{\sum_{i=1}^{M} c_i} \end{cases} \tag{9}$$

with

$$c_i = \begin{cases} \alpha \times c_{i-1} \times (1 - c_{i-1}), i > 1 \\ \alpha \times c_0 \times (1 - c_0), i = 1 \end{cases}, \tag{10}$$

where α and c_0 are parameters for the logistic map[1] presented by (10). In this way, it is guaranteed that the pseudo random number list $r_1, ..., r_M$ is fixed and satisfies $\sum_{i=1}^{M} r_i = 1$, such that the random group sizes from $|\mathbf{x}_1^s|, ..., |\mathbf{x}_M^s|$ are fixed as well, as long as the same settings of α and c_0 are given. In this work, $\alpha = 3.8$ and $c_0 = 0.1$ are recommended. One thing to be noted is that since the ceiling function ($\lceil \cdot \rceil$) has been used in (9) to obtain integer numbers, the total number of decision variables in $\mathbf{x}^s$ after non-uniform grouping can be slightly larger than the predefined parameter n_s.

B. Different Correlations between Variable Groups and Objective Functions

Theoretically, a decision variable can be correlated with any (one or multiple) objective functions. For simplicity, we correlate a group of decision variables with a group of objective function(s), where the correlations can be described using a correlation matrix:

$$\mathbf{C} = \begin{pmatrix} c_{1,1} & c_{1,2} & \cdots & c_{1,M} \\ c_{2,1} & c_{2,2} & \cdots & c_{2,M} \\ \vdots & \vdots & \ddots & \vdots \\ c_{M,1} & c_{M,2} & \cdots & c_{M,M} \end{pmatrix} \tag{11}$$

where $c_{i,j} = 1$ means that objective function $f_i(\mathbf{x})$ is correlated with the decision variables in group $\mathbf{x}_j^s$, and $c_{i,j} = 0$ otherwise.

Correspondingly, if we denote $\mathbf{C} = (\mathbf{c}_1, ..., \mathbf{c}_M)^\top$, the landscape matrix $\mathbf{G}(\mathbf{x}^s)$ in (6) can be presented as:

$$\mathbf{G}(\mathbf{x}^s) = \text{diag}(\mathbf{c}_1 \bar{\mathbf{g}}_1(\mathbf{x}^s), ..., \mathbf{c}_M \bar{\mathbf{g}}_M(\mathbf{x}^s)) \tag{12}$$

with

$$\bar{\mathbf{g}}_i(\mathbf{x}^s) = (\bar{g}_i(\mathbf{x}_1^s), ..., \bar{g}_i(\mathbf{x}_M^s))^\top, \tag{13}$$

where $i = 1, ..., M$, $\mathbf{x}_i^s$ is the i-th group of decision variables, as defined in (7), and $\bar{g}_1$ to $\bar{g}_M$ are basic functions that define the fitness landscape for each objective function.

If we substitute the landscape matrix $\mathbf{G}(\mathbf{x}^s)$ in (6) with (11) to (13), the formulation of (5) can be rewritten as:

$$\begin{cases} f_1(\mathbf{x}) = h_1(\mathbf{x}^f)(1 + \sum_{j=1}^{M} c_{1,j} \times \bar{g}_1(\mathbf{x}_j^s)) \\ \quad ... \\ f_i(\mathbf{x}) = h_i(\mathbf{x}^f)(1 + \sum_{j=1}^{M} c_{i,j} \times \bar{g}_i(\mathbf{x}_j^s)) \\ \quad ... \\ f_M(\mathbf{x}) = h_M(\mathbf{x}^f)(1 + \sum_{j=1}^{M} c_{M,j} \times \bar{g}_M(\mathbf{x}_j^s)) \end{cases} \tag{14}$$

where $i = 1, ..., M$.

Thus, the correlations between the decision variables and the objective functions can be fully specified using the correlation matrix $\mathbf{C}$, where the correlations can be either separable or overlapped.

C. Mixed Separability

As pointed in [44], since real-world problems are rarely fully separable or non-separable, it is important to introduce mixed separability into test problems for large-scale optimization.

Specifically, mixed separability can be realized by applying different basic fitness landscape functions $\bar{g}_i$ in (14), which are some single-objective optimization functions of different types of separability.

D. Variable Linkages on PSs

The variable linkages on PSs is a kind of important characteristic to be taken into consideration in the design of multi-objective optimization test problems [27], [28], [54]. The main motivation of introducing variable linkages is to consider a more general distribution of the Pareto optimal solutions,

[1]Logistic map is an archetypal example to show how chaotic behavior can arise from very simple non-linear dynamical equations [53].

making it more challenging for MOEAs to achieve diverse solutions.

Generally, variable linkages can be defined using a *linkage function*:

$$\mathbf{x}^s \leftarrow L(\mathbf{x}^s), \tag{15}$$

where $\mathbf{x}^s$ consists of the decision variables that define the PS, and L is the linkage function that defines the linkages between the variables in $\mathbf{x}^s$.

V. The Proposed Test Problems

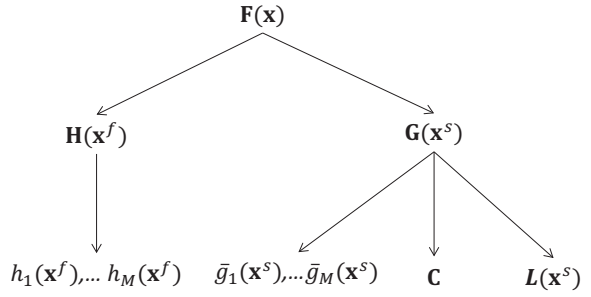


Fig. 1. The structure of the proposed test problems. A test problem $\mathbf{F}(\mathbf{x})$ consists of a shape matrix $\mathbf{H}(\mathbf{x}^f)$ and a landscape matrix $\mathbf{G}(\mathbf{x}^s)$, which are further composed of four components: a group of shape functions $h_i(\mathbf{x}^f)$, a group of landscape functions $\bar{g}_i(\mathbf{x}^s)$, a correlation matrix $\mathbf{C}$, and a linkage function $L(\mathbf{x}^s)$, with $i = 1, ..., M$.

As formulated in (6), generally, a test problem $\mathbf{F}(\mathbf{x})$ consists of two parts: the shape matrix $\mathbf{H}(\mathbf{x}^f)$ and the landscape matrix $\mathbf{G}(\mathbf{x}^s)$, where $\mathbf{x}^f$ and $\mathbf{x}^s$ are the first part and second part of the whole decision vector $\mathbf{x}$, respectively, and the Pareto optimal solution set of $\mathbf{F}(\mathbf{x})$ is defined as:

$$\begin{cases} \text{PS: } \mathbf{x}_* = (\mathbf{x}_*^f, \mathbf{x}_*^s) \\ \text{PF: } \mathbf{H}(\mathbf{x}_*^f) \end{cases}, \tag{16}$$

which satisfies $\mathbf{G}(\mathbf{x}_*^s) = \mathbf{O}$, such that $\mathbf{F}(\mathbf{x}_*) = \mathbf{H}(\mathbf{x}_*^f)$.

More specifically, for $i = 1, ..., M$, $\mathbf{H}(\mathbf{x}^f)$ consists of a group of shape functions $h_i(\mathbf{x}^f)$, as formulated in (5); $\mathbf{G}(\mathbf{x}^s)$ consists of a group of landscape functions $\bar{g}_i(\mathbf{x}^s)$ together with a correlation matrix $\mathbf{C}$, as formulated in (12). In addition, the variable linkages on the PSs are defined by a linkage function $L(\mathbf{x}^s)$, as described in Section IV-D. An illustration of the structure of the proposed test problems can be found in Fig. 1. Therefore, to define a test problem $\mathbf{F}(\mathbf{x})$, it is required to specify the following four basic components: $h_i(\mathbf{x}^f)$, $\bar{g}_i(\mathbf{x}^s)$, $\mathbf{C}$ and $L(\mathbf{x}^s)$.

This section first presents some instances of the four basic components that constitute the test problems. Afterwards, nine test problems are constructed based on the combinations of the instances of each component[2].

A. Instances of the Shape Functions $h_i(\mathbf{x}^f)$

In our design principles, the PF of a test problem is defined by a group of shape functions $h_i(\mathbf{x}^f)$. In general, the basic geometrical structure of a PF can be either linear, convex, concave, while the basic distribution property of a PF can be either continuous, disconnected or degenerate. As suggested in [26], a PF can be made complicated by mixing different basic geometrical structures and distribution properties.

Nevertheless, since this work focuses on the properties of the decision variables, here we only provide three basic instances of PF, including a linear PF (denoted as $\mathbf{H}_1$), a convex PF (denoted as $\mathbf{H}_2$) and a disconnected PF (denoted as $\mathbf{H}_3$) , which are taken from the DTLZ test suite suggested in [21]. To be specific, the linear PF is almost the same as that in DTLZ1, except that each objective is normalized to $[0, 1]$ instead of $[0, 0.5]$, while the convex PF and the disconnected PF are exactly the same as those of DTLZ2 and DTLZ7, respectively. The detailed definitions of the three instances can be found in Supplementary Materials I.

TABLE I
THE PROPERTIES OF THE BASIC SINGLE-OBJECTIVE OPTIMIZATION PROBLEMS.

Problem	Modality	Separability
η_1: Sphere function	Unimodal	Separable
η_2: Schwefel's problem 2.21	Unimodal	Non-separable
η_3: Rosenbrock's function	Multi-modal	Non-separable
η_4: Rastrigin's function	Multi-modal	Separable
η_5: Griewank's function	Multi-modal	Non-separable
η_6: Ackley's function	Multi-modal	Separable

B. Instances of the Basic Landscape Functions $\bar{g}_i(\mathbf{x}^s)$

In general, the basic landscape functions $\bar{g}_i(\mathbf{x}^s)$ can be constructed based on any (combinations of) single-objective functions. In the proposed test suite, we suggest six basic single-objective optimization problems for the construction of basic landscape functions $\bar{g}_i(\mathbf{x}^s)$, including the Sphere function, the Schwefel's problem 2.21, the Rosenbrock's function, the Rastrigin's function, the Griewank's function, and the Ackley's function, denoted as η_1 to η_6, which are not only widely adopted as test functions to assess the performance of large-scale optimization algorithms [35]–[37], [55], but also commonly used to construct large-scale optimization test problems [41]–[43]. The properties of these functions are summarized in Table I and the detailed definitions can be found in Supplementary Materials II.

It should be noted that although all the six basic single-objective functions have a fixed global optimum (except that the global optimum of η_3 is $\mathbf{x}^* = \mathbf{1}$, the global optimum of the other five functions is $\mathbf{x}^* = \mathbf{0}$), the Pareto optimal solutions have been shifted to different positions in the decision space by applying the variable linkages to be introduced in Section V-D.

To ensure that the objective function values are not too distant from the true PFs, the values of the single-objective functions are further normalized by dividing the length of the decision vectors as follows:

$$\bar{g}_i(\mathbf{x}_i^s) = \frac{1}{n_k} \sum_{j=1}^{n_k} \frac{\eta(\mathbf{x}_{i,j}^s)}{|\mathbf{x}_{i,j}^s|}, \tag{17}$$

[2]The Matlab code of the proposed test problems can be downloaded from: http://www.soft-computing.de/jin-pub_year.html

where η is one of the basic functions from η_1 to η_6 as listed in Table I, and $\mathbf{x}^s_{i,j}$ denotes the j-th subcomponent in the variable group $\mathbf{x}^s_i$.

In addition, since the test problems are proposed for both multi- and many-objective optimization, to make $\bar{g}_i(\mathbf{x}^i_s)$ scalable to any number of objectives, we divide $\bar{g}_i(\mathbf{x}^i_s)$ into two groups:

$$\begin{cases} \bar{g}^I(\mathbf{x}^s_i) = \{\bar{g}_{2k-1}(\mathbf{x}^s_i)\}, \\ \bar{g}^{II}(\mathbf{x}^s_i) = \{\bar{g}_{2k}(\mathbf{x}^s_i)\}. \end{cases} \tag{18}$$

where $k = 1, ..., \lceil M/2 \rceil$, $\bar{g}^I(\mathbf{x}^s_i)$ and $\bar{g}^{II}(\mathbf{x}^s_i)$ contain landscape functions with odd and even indices, respectively. In this way, given an arbitrary objective number M, the fitness landscapes can be defined by relating $\bar{g}^I(\mathbf{x}^s_i)$ and $\bar{g}^{II}(\mathbf{x}^s_i)$ to one of the basic functions among η_1 to η_6 according to (17).

C. Instances of the Correlation Matrix $\mathbf{C}$

Generally, the correlation matrix $\mathbf{C}$ can be arbitrarily defined to specify the correlation relationship between the variable groups and the objective functions. Here we present three typical instances to be used in the proposed test problems.

1) Separable Correlations:

$$\mathbf{C}_1 = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} \tag{19}$$

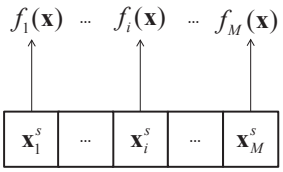


Fig. 2. The separable correlations between the variable groups and the objective functions.

The separable correlations between the variable groups and the objective functions can be described by an identity matrix $\mathbf{C}_1$, where the elements on the diagonal are set to 1 and all others are set to 0. In this way, each objective function is correlated with only one variable group. An illustration of the separable correlations is given in Fig. 2.

2) Overlapped Correlations:

$$\mathbf{C}_2 = \begin{pmatrix} 1 & 1 & 0 & \cdots & 0 \\ 0 & 1 & 1 & 0 & \vdots \\ \vdots & 0 & \ddots & \ddots & 0 \\ \vdots & 0 & 0 & 1 & 1 \\ 0 & \cdots & \cdots & 0 & 1 \end{pmatrix} \tag{20}$$

The overlapped correlations between the variable groups and the objective functions are presented by a matrix $\mathbf{C}_2$, where the neighboring elements $\mathbf{C}_2(i, i+1)$ are set to 1 in addition to the elements on the diagonal, with $i = 1, ..., M-1$, and the other elements are set to 0. In this way, each objective function has an overlapped correlation with two neighboring variable groups at the same time. Fig. 3 presents an illustrative example of the overlapped correlations.

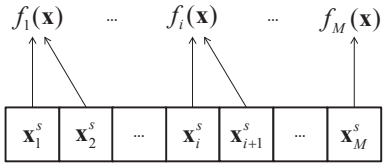


Fig. 3. The overlapped correlations between the variable groups and the objective functions.

3) Full Correlations:

$$\mathbf{C}_3 = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ 1 & 1 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & 1 \end{pmatrix} \tag{21}$$

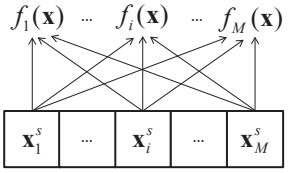


Fig. 4. The full correlations between the variable groups and the objective functions.

The full correlations between the variable groups and the objective functions are presented by a matrix $\mathbf{C}_3$, where all the elements are set to be 1. In this way, each objective function has a full correlation with all the variable groups at the same time. The full correlation relationship is visualized in Fig. 4.

D. Instances of the Linkage Function $L(\mathbf{x}^s)$

The general design principle of the linkage functions is similar to that in [30] and [31], where each decision variable in $\mathbf{x}^s$ (denoted as x^s_i, $i = 1, ..., n_s$) is linked with the first decision variable in $\mathbf{x}^f$, i.e., x^f_1. In this work, we consider both linear and nonlinear variable linkages, which will be elaborated in the following.

1) Linear Variable Linkage:

$$L_1(\mathbf{x}^s) = (1 + \frac{i}{|\mathbf{x}^s|}) \times (x^s_i - l_i) - x^f_1 \times (u_i - l_i), \tag{22}$$

2) Nonlinear Variable Linkage:

$$L_2(\mathbf{x}^s) = (1 + \cos(0.5\pi \frac{i}{|\mathbf{x}^s|})) \times (x^s_i - l_i) - x^f_1 \times (u_i - l_i), \tag{23}$$

where u_i and l_i is the upper and lower boundaries of decision variable x^s_i. In the proposed test problems, as will be presented later on in Section V-B, $u_i = 10$ and $l_i = 0$ are used. In addition to introducing variable linkages, another implicit effect of function L is to shift the optimal solutions to different positions, thereby increasing the difficulty of finding the global optimums of the landscape functions.

TABLE II
THE DEFINITIONS OF THE PROPOSED TEST PROBLEMS.

Problem	L	$\mathbf{H}$	g^I	g^{II}	$\mathbf{C}$
LSMOP1	L_1	$\mathbf{H}_1$	η_1	η_1	$\mathbf{C}_1$
LSMOP2	L_1	$\mathbf{H}_1$	η_5	η_2	$\mathbf{C}_1$
LSMOP3	L_1	$\mathbf{H}_1$	η_4	η_3	$\mathbf{C}_1$
LSMOP4	L_1	$\mathbf{H}_1$	η_6	η_5	$\mathbf{C}_1$
LSMOP5	L_2	$\mathbf{H}_2$	η_1	η_1	$\mathbf{C}_2$
LSMOP6	L_2	$\mathbf{H}_2$	η_3	η_2	$\mathbf{C}_2$
LSMOP7	L_2	$\mathbf{H}_2$	η_6	η_3	$\mathbf{C}_2$
LSMOP8	L_2	$\mathbf{H}_2$	η_5	η_1	$\mathbf{C}_2$
LSMOP9	L_2	$\mathbf{H}_3$	η_1	η_6	$\mathbf{C}_3$

E. Definitions of the Proposed Test Problems

The test problems are defined by specifying the shape functions h_i, the landscape functions $\bar{g}_i$, the correlation matrix $\mathbf{C}$ and the linkage function L, which are all taken from the subsets of the instances presented previously. As summarized in Table II, there are nine test problems in total. A detailed mathematical formulation of these test problems can be found in Supplementary Materials III and the properties are summarized as follows.

TABLE III
THE PROPERTY OF THE FITNESS LANDSCAPE OF EACH TEST PROBLEM.

Problem	Modality	Separability
LSMOP1	Unimodal	Fully Separable
LSMOP2	Mixed	Partially Separable
LSMOP3	Multi-modal	Mixed
LSMOP4	Mixed	Mixed
LSMOP5	Unimodal	Fully Separable
LSMOP6	Mixed	Partially Separable
LSMOP7	Multi-modal	Mixed
LSMOP8	Mixed	Mixed
LSMOP9	Mixed	Fully Separable

Regarding the PFs, LSMOP1 to LSMOP4 have a linear PF as defined by $\mathbf{H}_1$ in (1) of Supplementary Materials I, LSMOP5 to LSMOP8 have a convex PF as defined by $\mathbf{H}_2$ in (2) of Supplementary Materials I, and LSMOP9 has a disconnected PF as defined by $\mathbf{H}_3$ in (3) of Supplementary Materials I. As for the PSs, LSMOP1 to LSMOP4 have linear variable linkage as defined by L_1 in (22), while LSMOP5 to LSMOP9 have nonlinear variable linkage as defined by L_2 in (23).

As introduced in Section V-B, the fitness landscapes of a test problem can be defined by relating g^I and g^{II} to different basic single-objective functions (η_1 to η_6) according to (17) and (18). The combinations of the basic single-objective functions for the nine test problems are listed in Table II, and the properties of the landscapes thus defined are summarized in Table III. In summary, only LSMOP1 and LSMOP5 have a unimodal fitness landscape while the others have either multi-modal landscapes or mixed landscapes. As a result of the non-uniform grouping of the decision variables as introduced in Section IV-A, a landscape can only be fully separable, partially separable or the mixture of both. In summary, LSMOP1, LSMOP5 and LSMOP9 have a fully separable landscape; LSMOP2 and LSMOP6 have a partially separable landscape; while the other four have a landscape of mixed separability.

It should be noted that, the design of LSMOP1 to LSMOP9 also aims to maintain a proper balance between difficulty and diversity. For example, LSMOP3 has a simple linear variable linkage on the PS, but the fitness landscape is the most challenging one due to the adoption of η_4 and η_3, which are the most difficult separable and non-separable functions, respectively, while LSMOP9 has a complex disconnected PF, its fitness landscape is designed to be relatively simple by adopting η_1, which is a simple Sphere function.

TABLE IV
PROPERTIES OF THE SIX MOEAS USED IN EMPIRICAL STUDIES.

Algorithm	Reproduction Operator	Selection Operator
IM-MOEA	Inverse Models	Pareto Dominance
MOEA/D-DE	Differential Evolution	Tchebycheff Scalarization
NSGA-II	Genetic Operators	Pareto Dominance
IBEA	Genetic Operators	$I_{\epsilon+}$ Indicator
RVEA	Genetic Operators	Reference
NSGA-III	Genetic Operators	Pareto Dominance

VI. EMPIRICAL EVALUATIONS

This section presents some empirical evaluations of the proposed test functions using six representative MOEAs. The main motivation of these empirical studies is to assess the hardness of the proposed test functions for the popular existing MOEAs developed for solving small-scale (typically up to 50 decision variables) multi- and many-objective optimization problems. We first briefly describe the six MOEAs, then we present the performance indicators used to measure the quality of the results obtained by the six MOEAs, together with the experimental settings. Finally, we make some remarks on the performance of the compared MOEAs on the test problems.

A. MOEAs in Comparison

There are a variety of MOEAs in the literature, whereas most of them share similar frameworks. Since the most important components in an MOEA are the reproduction operator and selection operator, in our empirical studies, we have selected six representative MOEAs that differ with each other in these two aspects, where the properties are summarized in Table IV. To be specific, the six MOEAs adopted for empirical evaluations of the proposed test problems are inverse model based MOEA (IM-MOEA) [31], MOEA/D-DE [28], NSGA-II [6], indicator based evolutionary algorithm (IBEA) [12], reference vector guided evolutionary algorithm [14], and NSGA-III [16]. Among the six MOEAs, IM-MOEA, MOEA/D-DE and NSGA-II were proposed for solving MOPs having two or three objectives, whereas IBEA, RVEA and NSGA-III were shown to be promising for solving MaOPs having more than three objectives:

1) IM-MOEA is a recently proposed MOEA using Gaussian process based inverse modeling. Unlike traditional MOEAs, where the candidate solutions are directly generated in the decision space, the candidate solutions

in IM-MOEA are first sampled on the PF and then mapped back to the PS using a set of inverse models. IM-MOEA shows promising performance on a variety of problems with variable linkages on the PSs.

2) MOEA/D-DE is a variant of MOEA/D, which is a decomposition based approach and has been shown to perform well on various MOPs [56]–[58]. MOEA/D-DE was proposed for solving problems with complex PSs, where the simulated binary crossover (SBX) operator used in the original MOEA/D is replaced by the crossover operator in differential evolution [59].
3) NSGA-II is one of the most popular dominance based MOEAs [60]–[63]. In NSGA-II, the non-dominated sorting is adopted to achieve convergence of the population towards the PF, and crowding distance is used to manage the population diversity.
4) IBEA is a popular indicator based MOEA. In IBEA, the selection criterion is calculated based on a binary indicator ($I_{\epsilon+}$) which is considered to be compliant with the Pareto dominance. Specifically, a solution that has a higher contribution to the sum up of the indicator values for the whole population is considered to have better fitness, and vice versa.
5) RVEA is a recently proposed algorithm for solving MaOPs, where the search process is guided by a set of predefined reference vectors. RVEA first divides the objective space into a number of subspaces, and a scalarization method known as the angle penalized distance (APD) is designed to balance between convergence and diversity in many-objective optimization.
6) NSGA-III is an extended version of NSGA-II tailored for solving MaOPs. In NSGA-III, the secondary selection is performed with the help of a set of reference points, thereby addressing the issue of loss of selection pressure of the dominance based approaches.

In the empirical evaluations, IM-MOEA, MOEA/D-DE and NSGA-II are used to test the bi-/three-objective instances, while IBEA, NSGA-III and RVEA are used to test the many-objective instances. The recently proposed ENS approach is adopted in IM-MOEA, NSGA-II and NSGA-III to improve the efficiency of non-dominated sorting [64].

B. Performance Indicators

To measure the quality of the solution set obtained by an MOEA, the inverted generational distance (IGD) [65], [66] is used as the performance indicator.

1) Inverted Generational Distance (IGD): Let P^* denote a set of uniformly distributed solutions on the true PF, and P denote a set of approximate solutions obtained by an MOEA. The IGD value between P^* and P can be defined as:

$$IGD(P^*, P) = \frac{\sum_{v \in P^*} d(v, P)}{|P^*|}, \tag{24}$$

where $d(v, P)$ denotes the minimum Euclidean distance between a point v from P^* and each point in P. To use the IGD indicator, a set of uniformly distributed points should be generated to constitute P^*. In this study, a number of 500 and 10000 points are uniformly sampled on the true PFs as P^* for bi-/three-objective instances and six/eight-objective instances, respectively. A detailed description of the sampling approach for each true PF can be found in Supplementary Materials A.

TABLE V
SETTINGS OF POPULATION SIZES IN MOEA/D-DE, RVEA AND NSGA-III.

M	(H_1, H_2)	Population size
3	(99, 0)	100
3	(13, 0)	105
6	(4, 1)	132
10	(3, 2)	275

H_1 and H_2 are the simplex-lattice design factors for generating uniformly distributed reference (or weight) vectors on the outer boundaries and the inside layers, respectively.

C. Parameter Settings

1) Settings of the Test Problems: There are three parameters to be set in the proposed test problems, i.e., M, n_s and n_k, where M is the number of objectives, n_s is the number of decision variables that define the fitness landscape and n_k is the number of subcomponents in each variable group. We recommend that $n_s = 100 * M$ and $n_k = 5$ for general benchmark tests. Correspondingly, in this work, we have tested multi-objective instances with $M = 2, 3$ and $n_s = 200, 300$, many-objective instances with $M = 6, 10$ and $n_s = 600, 1000$, and the number of subcomponent n_k is constantly set to 5.

2) Experimental Settings: The maximum number of generations is used as the termination condition, which is set to 2000, 3000, 6000 and 10000 for the bi-objective 3-objective, 6-objective and 10-objective instances, respectively. To obtain statistical results, each algorithm is run for 20 times independently.

3) Settings of the Population Size: For MOEA/D-DE, NSGA-III and RVEA, the population size is determined by the simplex-lattice design factor H and the objective number M. As recommended in [16], a bi-layer vector generation method is adopted here to generate reference (or weight) vectors on both the outer boundaries and the inside layers of the PFs. The detailed settings of the H factors are summarized in Table V. For IM-MOEA, NSGA-II and IBEA, the population sizes are consistent with those in Table V with respect to the number of objectives.

4) Settings of the Genetic Operators: For the simulated binary crossover (SBX), the distribution index and the crossover probability are set to $\eta_c = 30$ and $p_c = 1$, and for polynomial mutation, the distribution index and the mutation probability are set to $\eta_m = 20$ and $p_m = 1/n$.

5) Settings of the MOEAs: For IM-MOEA, there are two parameters, i.e., the number of reference vectors K and the model group size L. In this study, we set $L = M$ and $K = 3$. For MOEA/D-DE, the neighborhood size T is set to 20, the selection probability δ is set to 0.9, the solution replacement parameter n_r is set to 2, and the parameters for the differential evolution operator is set as $CR = 1$ and $F = 0.5$. For IBEA, the parameter κ is set to 0.05. For RVEA, the adaptive

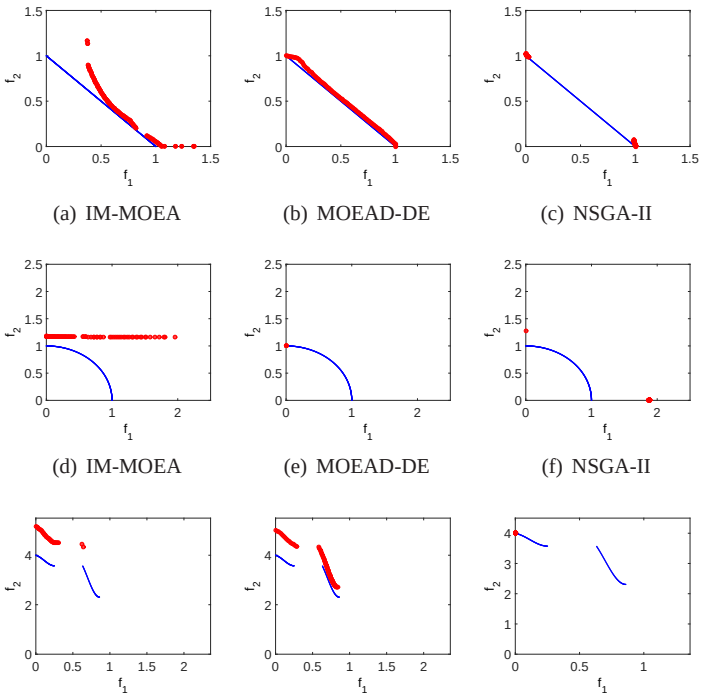
frequency $f_r = 0.1$ and the parameter for angle penalized distance $\alpha = 2.0$ are used.

TABLE VI
THE BEST, MEDIAN AND WORST IGD VALUES OBTAINED BY EACH COMPARED ALGORITHM ON BI-OBJECTIVE AND THREE-OBJECTIVE INSTANCES. THE BEST VALUES ARE HIGHLIGHTED.

Problem	M	IM-MOEA	MOEA/D-DE	NSGA-II
LSMOP1	2	1.10E-01	1.90E-02	3.05E-01
		1.32E-01	2.13E-02	3.13E-01
		1.82E-01	7.08E-02	3.23E-01
	3	4.60E-01	1.01E+00	8.07E-01
		8.03E-01	1.13E+00	9.27E-01
		9.75E-01	1.30E+00	9.98E-01
LSMOP2	2	6.27E-02	8.38E-02	8.39E-02
		6.39E-02	8.61E-02	8.90E-02
		6.80E-02	9.07E-02	9.42E-02
	3	8.27E-02	8.48E-02	9.61E-02
		8.54E-02	8.56E-02	9.93E-02
		9.61E-02	8.73E-02	1.00E-01
LSMOP3	2	1.45E+00	5.02E-01	1.35E+00
		1.72E+00	7.08E-01	1.42E+00
		2.49E+00	7.08E-01	1.75E+00
	3	3.26E+00	5.46E+00	3.71E+00
		4.40E+00	7.41E+00	4.33E+00
		7.99E+00	7.88E+00	4.91E+00
LSMOP4	2	7.12E-02	3.12E-02	1.28E-01
		7.14E-02	6.33E-02	1.28E-01
		7.62E-02	9.72E-02	1.33E-01
	3	2.06E-01	2.21E-01	2.54E-01
		2.12E-01	2.22E-01	2.56E-01
		2.20E-01	2.27E-01	2.71E-01
LSMOP5	2	2.17E-01	1.40E-02	3.41E-01
		2.77E-01	1.61E-02	3.42E-01
		4.66E-01	1.82E-02	3.44E-01
	3	6.91E-01	5.97E-01	1.48E+00
		9.85E-01	7.03E-01	1.64E+00
		1.40E+00	1.20E+00	1.82E+00
LSMOP6	2	5.44E-01	7.44E-01	7.18E-01
		6.17E-01	7.44E-01	7.74E-01
		7.76E-01	7.44E-01	8.72E-01
	3	2.86E+00	1.20E+00	1.80E+00
		1.06E+01	1.74E+00	2.45E+00
		1.43E+01	2.01E+00	2.62E+00
LSMOP7	2	2.88E+00	1.12E+00	1.71E+00
		4.35E+00	1.93E+00	2.20E+00
		5.67E+00	2.97E+00	2.41E+00
	3	1.21E+00	9.48E-01	1.49E+00
		1.25E+00	9.48E-01	1.50E+00
		1.36E+00	9.48E-01	1.54E+00
LSMOP8	2	1.11E-01	4.81E-02	3.46E-01
		1.91E-01	4.97E-02	3.47E-01
		2.19E-01	5.21E-02	3.55E-01
	3	3.70E-01	5.60E-01	3.15E-01
		4.02E-01	5.69E-01	3.28E-01
		4.26E-01	5.85E-01	3.74E-01
LSMOP9	2	6.85E-01	3.20E-01	8.11E-01
		9.95E-01	3.36E-01	8.11E-01
		1.28E+00	3.42E-01	8.11E-01
	3	1.39E+00	4.16E-01	1.63E+00
		2.40E+00	4.80E-01	2.53E+00
		2.43E+00	4.90E-01	2.60E+00

TABLE VII
THE BEST, MEDIAN AND WORST IGD VALUES OBTAINED BY THE COMPARED ALGORITHMS ON SIX-OBJECTIVE AND TEN-OBJECTIVE INSTANCES. THE BEST VALUES ARE HIGHLIGHTED.

Problem	M	IBEA	RVEA	NSGA-III
LSMOP1	6	6.41E-01	6.83E-01	2.27E+00
		9.13E-01	7.06E-01	3.47E+00
		9.34E-01	1.09E+00	4.18E+00
	10	7.11E-01	7.01E-01	5.05E+00
		9.15E-01	7.67E-01	5.62E+00
		1.06E+00	8.51E-01	6.33E+00
LSMOP2	6	1.66E-01	2.19E-01	2.22E-01
		1.76E-01	2.21E-01	2.23E-01
		1.92E-01	2.23E-01	2.24E-01
	10	2.24E-01	2.39E-01	2.51E-01
		2.36E-01	2.44E-01	2.52E-01
		2.45E-01	2.46E-01	2.53E-01
LSMOP3	6	7.42E+00	8.76E-01	1.14E+01
		1.63E+01	1.03E+00	1.83E+01
		2.05E+01	1.04E+00	2.27E+01
	10	1.06E+00	1.02E+00	7.73E-01
		2.27E+00	1.07E+00	4.27E+00
		3.34E+00	1.21E+00	1.16E+01
LSMOP4	6	1.81E-01	2.87E-01	2.79E-01
		1.82E-01	2.95E-01	2.82E-01
		1.90E-01	3.04E-01	2.84E-01
	10	2.32E-01	2.75E-01	2.93E-01
		2.36E-01	2.77E-01	2.94E-01
		2.45E-01	2.81E-01	2.96E-01
LSMOP5	6	4.33E-01	8.77E-01	5.37E+00
		1.00E+00	8.83E-01	6.39E+00
		1.22E+00	9.24E-01	7.79E+00
	10	7.54E-01	1.25E+00	4.23E+00
		7.55E-01	1.25E+00	4.70E+00
		1.26E+00	1.25E+00	1.56E+01
LSMOP6	6	1.54E+00	1.23E+00	1.42E+00
		1.78E+00	1.28E+00	1.43E+00
		1.89E+00	1.30E+00	2.10E+00
	10	1.68E+00	1.13E+00	1.95E+00
		1.75E+00	1.34E+00	2.17E+00
		2.01E+00	1.36E+00	3.67E+02
LSMOP7	6	1.97E+00	2.29E+00	6.10E+01
		2.21E+00	3.16E+00	6.83E+02
		2.24E+00	6.41E+00	2.88E+03
	10	1.44E+00	2.07E+00	2.23E+00
		1.58E+00	2.59E+00	4.93E+00
		2.19E+00	3.53E+00	4.69E+02
LSMOP8	6	6.69E-01	8.44E-01	2.16E+00
		6.88E-01	8.55E-01	3.06E+00
		8.07E-01	9.00E-01	3.45E+00
	10	7.48E-01	9.65E-01	9.59E-01
		7.55E-01	1.01E+00	9.81E-01
		8.04E-01	1.03E+00	4.52E+00
LSMOP9	6	7.87E+00	1.05E+01	7.84E+00
		8.74E+00	1.34E+01	8.63E+00
		9.14E+00	1.18E+02	9.13E+00
	10	1.19E+01	5.51E+01	3.74E+01
		1.31E+01	8.35E+01	3.81E+01
		1.35E+01	3.10E+02	4.36E+01

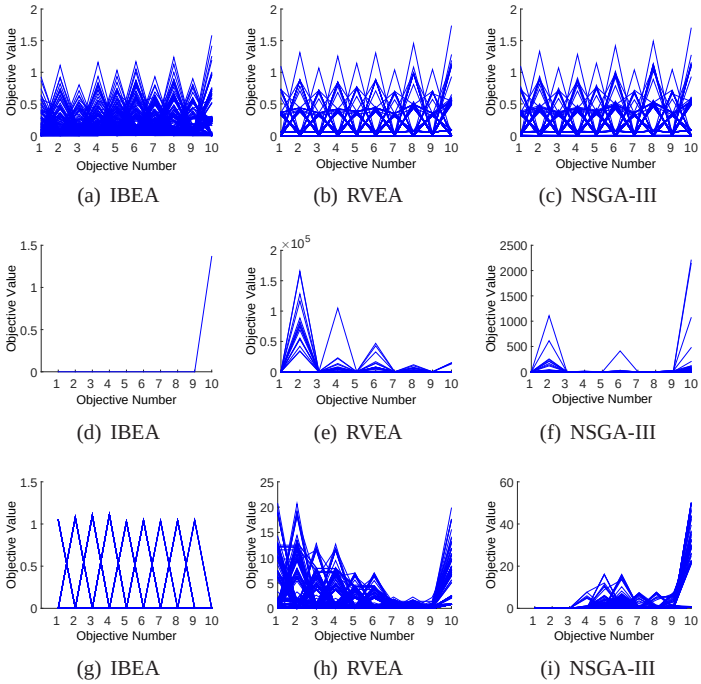


(a) IM-MOEA (b) MOEAD-DE (c) NSGA-II

(d) IM-MOEA (e) MOEAD-DE (f) NSGA-II

(g) IM-MOEA (h) MOEAD-DE (i) NSGA-II

Fig. 5. The non-dominated solutions obtained by each algorithm on 2-objective LSMOP1, LSMOP6, LSMOP9 in the run associated with the best IGD value.



(a) IBEA (b) RVEA (c) NSGA-III

(d) IBEA (e) RVEA (f) NSGA-III

(g) IBEA (h) RVEA (i) NSGA-III

Fig. 6. The parallel coordinates of non-dominated solutions obtained by each algorithm on 10-objective LSMOP2, LSMOP3, LSMOP8 in the run associated with the best IGD value.

D. Empirical Results

In the following, we present results obtained by each of the MOEA on the multi-objective and many-objective instances.

1) Results on Multi-Objective Instances: The test problems have 200 or 300 decision variables and two or three objectives, respectively. The statistical results of the IGD values obtained by IM-MOEA, MOEA/D-DE and NSGA-II are summarized in Table VI. In general, none of the three algorithms is able to efficiently solve all instances. To be more specific, IM-MOEA shows the best performance on LSMOP2 and LSMOP6, MOEA/D-DE shows the best performance on LSMOP5, LSMOP7 and LSMOP9, and NSGA-II shows the best performance on 3-objective LSMOP3 and LSMOP8. For further analysis, we have plotted the non-dominated solutions obtained by each algorithm in the best run on three typical problems (LSMOP1, LSMOP6 and LSMOP9) in Fig 5. In the following, we will present some discussions based on the observations made from Fig 5.

LSMOP1 is a relatively simple problem which has a unimodal and fully separable fitness landscape. Nevertheless, there is still one major difficulty in solving this problem, i.e., the non-uniform variable groups correlated with the two objectives. Our results show that, with the non-uniform group size generator presented in (9) and (10), a number of 60 and 145 decision variables are correlated with f_1 and f_2, respectively. This means that the decision variables have higher contribution to f_2 than f_1, thus causing an unbalanced convergence rate of the two objectives. This phenomenon can be observed from Fig. 5 (a), where the non-dominated solutions obtained by IM-MOEA have significantly better convergence on f_1 than on f_2. To confirm this observation, we have calculated the mean function values of the non-dominated solutions in Fig. 5 (a), and the results indicate that f_1 and f_2 have a mean value of 0.61 and 1.21, respectively.

LSMOP6 is a partially separable problem which has a complex fitness landscape. The difficulty of this problem is mainly caused by η_3 that is correlated with f_1. It can be clearly seen from Fig. 5 (d) to Fig 5 (e) that neither IM-MOEA nor MOEA/D-DE is able to obtain solutions with satisfactory convergence on f_1, especially for MOEA/D-DE, which has converged to a single point on the axis of f_2. By averaging the objective values of the non-dominated solutions in Fig. 5 (b), we can obtain the mean values of f_1 and f_2, which are around 1.0×10^5 (too large to be visible in the figure) and 0, respectively. Such a big difference between the values of the two objectives can cause difficulty for any existing MOEAs, because it leads to a biased convergence towards only one of them due to the highly unbalanced convergence pressures. This phenomenon can also be observed from the results obtained by IM-MOEA in Fig 5 (d). Interestingly, NSGA-II is the only algorithm that has obtained a solution of relatively proper convergence quality on f_2. We surmise that this can be attributed to the fact that the crowding distance based selection in NSGA-II has always preserved the extreme solutions, thus leading to convergence along the directions of the axis.

LSMOP9 has a disconnected PF, although the landscape is not complex. As shown in Fig. 5 (g) to Fig. 5 (i), the solutions obtained by IM-MOEA and MOEA/D-DE have similar convergence quality, while IM-MOEA has failed to approximate the second segment of the disconnected PF. Again, NSGA-

II has obtained an extreme solution that has achieved the best convergence on f_2, which, however, is the only solution thus obtained. This phenomenon implies that it is difficult for MOEAs to strike a balance between convergence and diversity on this problem.

Based on the above observations, it can be concluded that the proposed test problems are able to introduce various degrees of difficulties into different objective functions by applying different variable group sizes and correlation matrices.

2) Results on Many-objective Instances: This set of test problems are of 600- or 1000-dimensional and have six or ten objectives, respectively. The statistical results of the IGD values obtained by IBEA, RVEA and NSGA-III are summarized in Table VII. Again, none of the three algorithms is able to significantly outperform the others on all these instances. IBEA shows the best performance on LSMOP3 and LSMOP6. To further analyze the results, we have plotted the parallel coordinate of the best non-dominated solution set obtained by each algorithm on LSMOP2, LSMOP3 and LSMOP8 in Fig 6. In the following, we will present some discussions based on the observations from Fig 6.

LSMOP2 is a partially separable problem where the fitness landscape has mixed modality. As shown in Fig. 6 (a) to Fig. 6 (c), the solution sets obtained by the three algorithms show similar convergence quality, while the one obtained by IBEA shows the best distribution, thus leading to the smallest IGD value as shown in Table VII. Since the three algorithms adopt the same reproduction operators (SBX and polynomial mutation), we believe that the advantage of IBEA over the other two algorithms on this problem is due to its $I_{\epsilon+}$ indicator based selection operator.

LSMOP3 has the most difficult multi-modal fitness landscape among the nine test problems. As shown in Fig. 6 (d), interestingly, IBEA has only obtained one single solution on this problem although its solution set on LSMOP2 shows promising distribution, which implies that it is difficult to always achieve good performance using the same selection strategy on different test problems. However, although the solutions sets obtained by RVEA and NSGA-III show relatively better distribution, the convergence quality is quite poor.

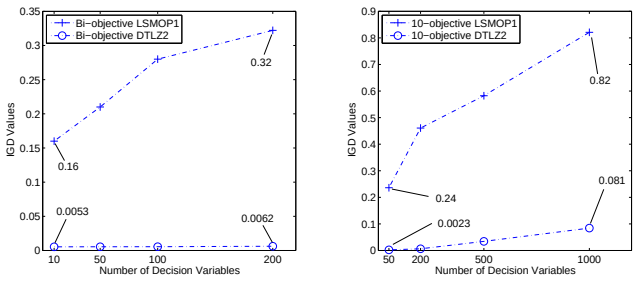
LSMOP8 is a problem with both mixed modality and separability. As shown in Fig. 6 (g), again, the solution set obtained by IBEA has achieved promising convergence while the distribution is quite sparse. By contrast, as shown in Fig. 6 (h) and Fig. 6 (i), the solution sets obtained by RVEA and NSGA-III show a denser distribution but worse convergence. Another interesting observation is that RVEA achieves better performance on f_1, f_2 and f_3 than NSGA-III.

To summarize, it turns out that none of the compared algorithms is able to obtain a solution set with both satisfying convergence and distribution performance on the many-objective instances, especially for problems with difficult fitness landscapes such as LSMOP3.

E. Scalability Analysis of Decision Variables

Not many existing test suites for multi- and many-objective optimization have considered the scalability to the number of objectives as they are expected to be only used in small-scale optimization. Therefore, if such test problems are directly applied to large-scale multi- and many-objective optimization, there will be a major issue: the fitness landscape is either too simple or too complex, thus leading to a too large or too small objective value. By contrast, the proposed test problems LSMOP1 to LSMOP9, which are tailored for large-scale multi- and many-objective optimization, have a proper scalability to the number of decision variables, ensuring that the objective values still fall into a reasonable range when the problems are applied to large-scale multi- and many-objective optimization.

To analyze the scalability of the decision variables in the proposed test problems, optimization of bi-objective and ten-objective is performed using NSGA-II and RVEA on two representative test problems, LSMOP1 and LSMOP3, which have the simplest and the most complex fitness landscapes, respectively. To make a comparison, we have also selected two representative test problems from the DTLZ test suite, DTLZ2 and DTLZ3, where DTLZ2 has a simple unimodal fitness landscape while DTLZ3 has a complex multimodal fitness landscape. The experimental settings are the same as those introduced in Section VI-C. For bi-objective instances, a number of 10, 50, 100 and 200 decision variables are adopted, while for 10-objective instances, a number of 50, 200, 500, 1000 decision variables are adopted, and the mean IGD values of the results obtained by NSGA-II and RVEA on each test instance are summarized in Fig. 7 and Fig. 8.

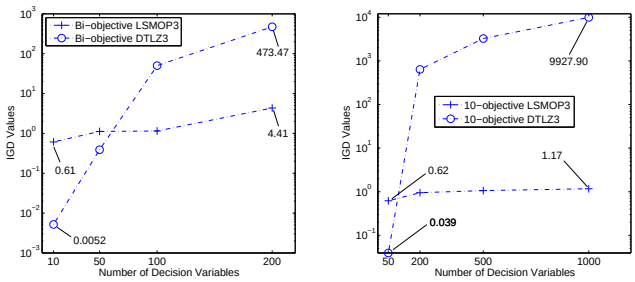


(a) NSGA-II (b) RVEA

Fig. 7. The IGD values of the solution sets obtained by NSGA-II and RVEA on bi-objective and 10-objective LSMOP1 and DTLZ2. For bi-objective instances, the number of decision variables increases from 10 to 200, while for 10-objective instances, the number of decision variables increases from 50 to 1000.

From Fig. 7 (a) and Fig. 7 (b), it can be observed that with the increase of decision variables, the IGD values obtained on both bi-objective and 10-objective LSMOP1 show continuous deterioration, from 0.16 to 0.32 and 0.24 to 0,82, respectively. By contrast, the IGD values obtained on DTLZ2 show very minor differences with the increase of decision variables, which implies that DTLZ2 is not suited for large-scale optimization as the fitness landscape is too simple to distinguish decision variables of different scales.

While the landscapes of LSMOP1 and DTLZ2 are simple unimodal, LSMOP3 and DTLZ3 have complex multimodal landscapes. Nevertheless, as evident in Fig. 8 (a) and Fig. 8 (b), with the increase of decision variables, a continuous deterioration of the IGD values is observed, from 0.61 to 4.41



(a) NSGA-II (b) RVEA

Fig. 8. The IGD values of the solution sets obtained by NSGA-II and RVEA on bi-objective and 10-objective LSMOP3 and DTLZ3. For bi-objective instances, the number of decision variables increases from 10 to 200, while for 10-objective instances, the number of decision variables increases from 50 to 1000.

and 0.62 to 1.17, respectively. By contrast, the IGD values obtained on DTLZ3 show drastic deterioration as the number of decision variables increases from 50 to 1000, which implies that DTLZ3 is not suited for large-scale optimization either, as the fitness landscape is too complex to constraint the objective values within a proper range.

Since LSMOP1 and LSMOP3 are two extreme cases where the fitness landscapes are the simplest one and most complex one, similar observations can also be obtained on the other seven test problems. Such empirical observations indicate that the proposed test problems have proper scalability to the number of decision variables such that the objective values can be constrained within a proper range.

VII. Conclusion

This paper proposes a new method for creating test problems for large-scale multi- and many-objective optimization by extending the design principles used in creating multi-objective optimization and large-scale optimization test problems. First, the decision variables are divided into a number of groups of different sizes. Second, the variable groups are correlated with different objective functions, and the correlations can be defined easily by means of a correlation matrix. Third, the decision variables have mixed separability.

Nine test problems are instantiated based on the proposed generic principle for designing large-scale multi- and many-objective test problems. Among them, LSMOP1 to LSMOP4 have a linear PF, LSMOP5 to LSMOP8 have a nonlinear PF, and LSMOP9 has a disconnected PF. In addition, the first four test problems (LSMOP1 to LSMOP4) are designed to have linear variable linkages, while the other five test problems (LSMOP5 to LSMOP9) have nonlinear variable linkages. Regarding the fitness landscapes, six widely used basic functions in the single-objective optimization literature have been adopted, three of which are separable and the other three are non-separable. Finally, three types of correlation relationships between the variable groups and the objective functions are proposed, including separable correlation, overlapped correlation and the full correlation.

To assess the proposed test problems, six representative MOEAs have been tested on the nine instantiations. We have first tested IM-MOEA, MOEA/D-DE and NSGA-II on bi- and three-objective instances, where the number of decision variables is 200 and 300, respectively. Furthermore, we have tested IBEA, RVEA and NSGA-III on six-objective and ten-objective instances, where the number of decision variables is 600 and 1000, respectively. Our experimental results indicate that the compared algorithms show various strength and weakness on different test problems, and none of them is able to efficiently solve all the proposed problems. This demonstrates that there is a demand for new MOEAs in order to solve such large-scale multi- or many-objective optimization problems with complex separability and correlations, where the new MOEAs should be able to address several basic issues. Firstly, the new MOEAs should be able to detect the non-uniform correlations between decision variables and objectives such that the objectives can be optimized in different orders. Secondly, the new MOEAs should be able to detect the mixed separability of decision variables such that the non-separable subcomponent of the decision vector can be optimized in a *divide-and-conquer* manner. Thirdly, the new MOEAs should be able to handle the variable linkages on the PSs. Finally, the new MOEAs should also be able to perform efficient and effective search on complex fitness landscapes. As recommended by in [67], the Borg framework can be promising in the design of such new MOEAs, where auto-adaptive techniques is adopted to address some of the above issues, e.g., different selection and reproduction operators can be applied to the search of different fitness landscapes.

It is worth noting that the nine test problems considered in the empirical study is a small set of instantiations created using the proposed design method. Thanks to the extensibility and generic nature of the proposed design principles, more test problems having different separability and correlation properties can be generated.

Acknowledgements

This work was supported in part by the Honda Research Institute Europe, the Joint Research Fund for Overseas Chinese, Hong Kong and Macao Scholars of the National Natural Science Foundation of China (Grant No. 61428302), and an EPSRC grant (No. EP/M017869/1).

References

[1] K. Ikeda, H. Kita, and S. Kobayashi, "Failure of Pareto-based MOEAs: does non-dominated really mean near to optimal?" in *Proceedings of the IEEE Congress on Evolutionary Computation*, 2001, pp. 957–962.

[2] D. Brockhoff, T. Friedrich, N. Hebbinghaus, C. Klein, F. Neumann, and E. Zitzler, "On the effects of adding objectives to plateau functions," *IEEE Transactions on Evolutionary Computation*, vol. 13, no. 3, pp. 591–603, 2009.

[3] 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 Transactions on Evolutionary Computation*, vol. 15, no. 4, pp. 444–455, 2011.

[4] H. Ishibuchi, N. Tsukamoto, and Y. Nojima, "Evolutionary many-objective optimization: A short review." in *Proceedings of the IEEE Congress on Evolutionary Computation*, 2008, pp. 2419–2426.

[5] A. Zhou, B.-Y. Qu, H. Li, S.-Z. Zhao, P. N. Suganthan, and Q. Zhang, "Multiobjective evolutionary algorithms: A survey of the state of the art," *Swarm and Evolutionary Computation*, vol. 1, no. 1, pp. 32–49, 2011.

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

[7] E. Zitzler, M. Laumanns, and L. Thiele, "SPEA2: Improving the strength Pareto evolutionary algorithm," in *Evolutionary Methods for Design, Optimisation and Control*, 2002, pp. 95–100.

[8] D. W. Corne, N. R. Jerram, J. D. Knowles, M. J. Oates *et al.*, "PESA-II: Region-based selection in evolutionary multiobjective optimization," in *Proceedings of the Genetic and Evolutionary Computation Conference*. ACM, 2001.

[9] H. Ishibuchi and T. Murata, "A multi-objective genetic local search algorithm and its application to flowshop scheduling," *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews*, vol. 28, no. 3, pp. 392–403, 1998.

[10] Y. Jin, T. Okabe, and B. Sendhoff, "Adapting weighted aggregation for multiobjective evolution strategies," in *Proceedings of the International Conference on Evolutionary multi-criterion optimization*. Springer, 2001, pp. 96–110.

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

[12] E. Zitzler and S. Künzli, "Indicator-based selection in multiobjective search," in *Proceedings of the Parallel Problem Solving from Nature*. Springer, 2004, pp. 832–842.

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

[14] R. Cheng, Y. Jin, M. Olhofer, and B. Sendhoff, "A reference vector guided evolutionary algorithm for many-objective optimization," *IEEE Transactions on Evolutionary Computation*, 2016, in press.

[15] H. Wang, L. Jiao, and X. Yao, "Two_arch2: An improved two-archive algorithm for many-objective optimization," *Evolutionary Computation, IEEE Transactions on*, vol. 19, no. 4, pp. 524–541, Aug 2015.

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

[17] X. Zhang, Y. Tian, and Y. Jin, "A knee point driven evolutionary algorithm for many-objective optimization," *IEEE Transactions on Evolutionary Computation*, accepted in 2014.

[18] K. Li, Q. Zhang, and S. Kwong, "An evolutionary many-objective optimization algorithm based on dominance and decomposition," *IEEE Transactions on Evolutionary Computation*, accepted in 2014.

[19] M. Li, S. Yang, and X. Liu, "Diversity comparison of pareto front approximations in many-objective optimization," *IEEE Transactions on Cybernetics*, vol. 44, no. 12, pp. 2568–2584, 2014.

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

[21] K. Deb, L. Thiele, M. Laumanns, and E. Zitzler, "Scalable test problems for evolutionary multi-objective optimization," in *2001 TIK-Technical Report*, no. 112, 2001.

[22] ——, *Evolutionary multiobjective optimization*. Springer, 2005, ch. Scalable test problems for evolutionary multiobjective optimization.

[23] K. Deb, "Multi-objective genetic algorithms: Problem difficulties and construction of test problems," *Evolutionary Computation*, vol. 7, no. 3, pp. 205–230, 1999.

[24] H. Jain and K. Deb, "An evolutionary many-objective optimization algorithm using reference-point based non-dominated sorting approach, part II: Handling constraints and extending to an adaptive approach," *IEEE Transactions on Evolutionary Computation*, vol. 18, no. 4, pp. 602–622, 2013.

[25] S. Huband, L. Barone, L. While, and P. Hingston, "A scalable multiobjective test problem toolkit," in *Proceedings of the International Conference on Evolutionary Multi-Criterion Optimization*, 2005, pp. 280–294.

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

[27] T. Okabe, Y. Jin, M. Olhofer, and B. Sendhoff, "On test functions for evolutionary multi-objective optimization," in *Parallel Problem Solving from Nature*. Springer, 2004, pp. 792–802.

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

[29] 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 *Proceedings of the Evolutionary Multi-Criterion Optimization*. Springer, 2011, pp. 197–211.

[30] Q. Zhang, A. Zhou, and Y. Jin, "RM-MEDA: A regularity model-based multiobjective estimation of distribution algorithm," *IEEE Transactions on Evolutionary Computation*, vol. 12, no. 1, pp. 41–63, 2008.

[31] R. Cheng, Y. Jin, K. Narukawa, and B. Sendhoff, "A multiobjective evoltuionary algorithm using Gaussian process based inverse modeling," *IEEE Transactions on Evolutionary Computation*, accepted in 2015.

[32] Y. Jin and B. Sendhoff, "Constructing dynamic optimization test problems using the multi-objective optimization concept," in *Applications of Evolutionary Computing*. Springer, 2004, pp. 525–536.

[33] A. Zhou, Y. Jin, Q. Zhang, B. Sendhoff, and E. Tsang, "Prediction-based population re-initialization for evolutionary dynamic multi-objective optimization," in *Evolutionary Multi-Criterion Optimization*. Springer, 2007, pp. 832–846.

[34] C.-K. Goh and K. C. Tan, "A competitive-cooperative coevolutionary paradigm for dynamic multiobjective optimization," *IEEE Transactions on Evolutionary Computation*, vol. 13, no. 1, pp. 103–127, 2009.

[35] Z. Yang, K. Tang, and X. Yao, "Large scale evolutionary optimization using cooperative coevolution," *Information Sciences*, vol. 178, no. 15, pp. 2985–2999, 2008.

[36] X. Li and X. Yao, "Cooperatively coevolving particle swarms for large scale optimization," *IEEE Transactions on Evolutionary Computation*, vol. 16, no. 2, pp. 210–224, 2012.

[37] R. Cheng and Y. Jin, "A competitive swarm optimizer for large scale optimization," *IEEE Transactions on Cybernetics*, vol. 45, no. 2, pp. 191–204, 2015.

[38] J. J. Durillo, A. J. Nebro, C. A. C. Coello, J. García-Nieto, F. Luna, and E. Alba, "A study of multiobjective metaheuristics when solving parameter scalable problems," *IEEE Transactions on Evolutionary Computation*, vol. 14, no. 4, pp. 618–635, 2010.

[39] L. M. Antonio and C. Coello Coello, "Use of cooperative coevolution for solving large scale multiobjective optimization problems," in *IEEE Congress on Evolutionary Computation*. IEEE, 2013, pp. 2758–2765.

[40] H. Masuda, Y. Nojima, and H. Ishibuchi, "Visual examination of the behavior of emo algorithms for many-objective optimization with many decision variables," in *Proceedings of the IEEE Congress on Evolutionary Computation*. IEEE, 2014, pp. 2633–2640.

[41] K. Tang, X. Yáo, P. N. Suganthan, C. MacNish, Y.-P. Chen, C.-M. Chen, and Z. Yang, "Benchmark functions for the CEC2008 special session and competition on large scale global optimization," *Nature Inspired Computation and Applications Laboratory, USTC, China*, pp. 153–177, 2007.

[42] K. Tang, X. Li, P. N. Suganthan, Z. Yang, and T. Weise, "Benchmark functions for the CEC2010 special session and competition on large-scale global optimization," Nature Inspired Computation and Applications Laboratory, USTC, China, Tech. Rep., 2009.

[43] X. Li, K. Tang, M. N. Omidvar, Z. Yang, and K. Qin, "Benchmark functions for the CEC2013 special session and competition on large-scale global optimization," 2013.

[44] M. N. Omidvar, X. Li, and K. Tang, "Designing benchmark problems for large-scale continuous optimization," *Information Sciences*, vol. 316, pp. 419–436, 2015.

[45] S. Sleesongsom, S. Bureerat, and K. Tai, "Aircraft morphing wing design by using partial topology optimization," *Structural and Multidisciplinary Optimization*, vol. 48, no. 6, pp. 1109–1128, 2013.

[46] H.-F. Teng, Y. Chen, W. Zeng, Y.-J. Shi, and Q.-H. Hu, "A dual-system variable-grain cooperative coevolutionary algorithm: satellite-module layout design," *Evolutionary Computation, IEEE Transactions on*, vol. 14, no. 3, pp. 438–455, 2010.

[47] B. Zhang, H.-F. Teng, and Y.-J. Shi, "Layout optimization of satellite module using soft computing techniques," *Applied Soft Computing*, vol. 8, no. 1, pp. 507–521, 2008.

[48] T. Weise, R. Chiong, and K. Tang, "Evolutionary optimization: Pitfalls and booby traps," *Journal of Computer Science and Technology*, vol. 27, no. 5, pp. 907–936, 2012.

[49] M. N. Omidvar, X. Li, Y. Mei, and X. Yao, "Cooperative co-evolution with differential grouping for large scale optimization," *IEEE Transactions on Evolutionary Computation*, vol. 18, no. 3, pp. 378–393, 2014.

[50] L. Graening and B. Sendhoff, "Shape mining: A holistic data mining approach for engineering design," *Advanced Engineering Informatics*, vol. 28, no. 2, pp. 166–185, 2014.

[51] N. E. Antoine and I. M. Kroo, "Framework for aircraft conceptual design and environmental performance studies," *AIAA journal*, vol. 43, no. 10, pp. 2100–2109, 2005.

[52] F. Duddeck, "Multidisciplinary optimization of car bodies," *Structural and Multidisciplinary Optimization*, vol. 35, no. 4, pp. 375–389, 2008.
[53] R. M. May *et al.*, "Simple mathematical models with very complicated dynamics," *Nature*, vol. 261, no. 5560, pp. 459–467, 1976.
[54] K. Deb, A. Sinha, and S. Kukkonen, "Multi-objective test problems, linkages, and evolutionary methodologies," in *Proceedings of the International Conference on Genetic and evolutionary computation*. ACM, 2006, pp. 1141–1148.
[55] R. Cheng and Y. Jin, "A social learning particle swarm optimization algorithm for scalable optimization," *Information Sciences*, vol. 291, no. 19, pp. 43–60, 2015.
[56] K. Li, S. Kwong, J. Cao, M. Li, J. Zheng, and R. Shen, "Achieving balance between proximity and diversity in multi-objective evolutionary algorithm," *Information Sciences*, vol. 182, no. 1, pp. 220–242, 2012.
[57] Z.-H. Zhan, J. Li, J. Cao, J. Zhang, H.-H. Chung, and Y.-H. Shi, "Multiple populations for multiple objectives: a coevolutionary technique for solving multiobjective optimization problems," *IEEE Transactions on Cybernetics*, vol. 43, no. 2, pp. 445–463, 2013.
[58] Y.-Y. Tan, Y.-C. Jiao, H. Li, and X.-K. Wang, "A modification to MOEA/D-DE for multiobjective optimization problems with complicated Pareto sets," *Information Sciences*, vol. 213, pp. 14–38, 2012.
[59] R. Storn and K. Price, "Differential evolution–a simple and efficient heuristic for global optimization over continuous spaces," *Journal of Global Optimization*, vol. 11, no. 4, pp. 341–359, 1997.
[60] M. Fazzolari, R. Alcala, Y. Nojima, H. Ishibuchi, and F. Herrera, "A review of the application of multiobjective evolutionary fuzzy systems: Current status and further directions," *IEEE Transactions on Fuzzy Systems*, vol. 21, no. 1, pp. 45–65, 2013.
[61] A. Ponsich, A. L. Jaimes, and C. A. C. Coello, "A survey on multiobjective evolutionary algorithms for the solution of the portfolio optimization problem and other finance and economics applications," *IEEE Transactions on Evolutionary Computation*, vol. 17, no. 3, pp. 321–344, 2013.
[62] E. K. Burke, M. Gendreau, M. Hyde, G. Kendall, G. Ochoa, E. Özcan, and R. Qu, "Hyper-heuristics: A survey of the state of the art," *Journal of the Operational Research Society*, vol. 64, no. 12, pp. 1695–1724, 2013.
[63] K. Deb, "Multi-objective optimization," in *Search Methodologies*. Springer, 2014, pp. 403–449.
[64] X. Zhang, Y. Tian, R. Cheng, and Y. Jin, "An efficient approach to non-dominated sorting for evolutionary multi-objective optimization," *IEEE Transactions on Evolutionary Computation*, 2014, DOI. 10.1109/TEVC.2014.2308305.
[65] Q. Zhang, A. Zhou, and Y. Jin, "RM-MEDA: a regularity model-based multiobjective estimation of distribution algorithm," *IEEE Transactions on Evolutionary Computation*, vol. 12, no. 1, pp. 41–63, 2008.
[66] M. Li and J. Zheng, "Spread assessment for evolutionary multiobjective optimization," in *Proceedings of the International Conference on Evolutionary Multi-Criterion Optimization*, 2009, pp. 216–230.
[67] D. Hadka and P. Reed, "Borg: An auto-adaptive many-objective evolutionary computing framework," *Evolutionary computation*, vol. 21, no. 2, pp. 231–259, 2013.



Ran Cheng received the B. E. degree from Northeastern University, Shenyang, China, in 2010. He is currently working towards the Ph.D. degree at the Department of Computer Science, University of Surrey, United Kingdom.

His current research interests include multi-/many-objective optimization, estimation of distribution algorithms, and swarm intelligence.



Yaochu Jin (M'98-SM'02-F'16) received the B.Sc., M.Sc., and Ph.D. degrees from Zhejiang University, Hangzhou, China, in 1988, 1991, and 1996 respectively, and the Dr.-Ing. degree from Ruhr University Bochum, Germany, in 2001.

He is a Professor in Computational Intelligence, Department of Computer Science, University of Surrey, Guildford, U.K., where he heads the Nature Inspired Computing and Engineering Group. He is also a Finland Distinguished Professor funded by the Finnish Agency for Innovation (Tekes) and a Changjiang Distinguished Visiting Professor appointed by the Ministry of Education, China. His science-driven research interests lie in the interdisciplinary areas that bridge the gap between computational intelligence, computational neuroscience, and computational systems biology. He is also particularly interested in nature-inspired, real-world driven problem-solving. He has (co)authored over 200 peer-reviewed journal and conference papers and been granted eight patents on evolutionary optimization. His current research is funded by EC FP7, UK EPSRC and industry. He has delivered 20 invited keynote speeches at international conferences.

He is Editor-in-Chief of the IEEE TRANSACTIONS ON COGNITIVE AND DEVELOPMENTAL SYSTEMS and Complex & Intelligent Systems. He is also an Associate Editor or Editorial Board Member of the IEEE TRANSACTIONS ON EVOLUTIONARY COMPUTATION, IEEE TRANSACTIONS ON CYBERNETICS, IEEE TRANSACTIONS ON NANOBIOSCIENCE, Evolutionary Computation, BioSystems, Soft Computing, and Natural Computing. Dr Jin was an IEEE Distinguished Lecturer (2013-2015) and Vice President for Technical Activities of the IEEE Computational Intelligence Society (2014-2015). He was the recipient of the Best Paper Award of the 2010 IEEE Symposium on Computational Intelligence in Bioinformatics and Computational Biology and the 2014 IEEE Computational Intelligence Magazine Outstanding Paper Award. He is a Fellow of IEEE.



Markus Olhofer obtained his Dipl.-Ing. in Electrical Engineering in 1997 and his PhD degree in 2000, both from Ruhr-University Bochum, Germany. He joined the Future Technology Research Division at Honda R&D Europe (Deutschland) in 1998 and has been Chief Scientist and head of the Complex Systems Optimisation and Analysis Group at the Honda Research Institute Europe since 2010. His research interests include the extension of soft computing methods to meet requirements in complex engineering problems, ranging from evolutionary design optimisation to engineering data mining.

Dr Olhofer is a Visiting Professor with the Department of Computer Science, University of Surrey, UK.



Bernhard Sendhoff (M'99-SM'05) obtained a PhD in Applied Physics in May 1998, from the Ruhr-Universit'at Bochum, Germany. From 1999 to 2002 he worked for Honda R&D Europe GmbH, and since 2003, he has been with the Honda Research Institute Europe GmbH. Since 2007 he is Honorary Professor of the School of Computer Science of the University of Birmingham, UK. Since 2008, he is Honorary Professor at the Technical University of Darmstadt, Germany. Since 2011 he is President of the Honda Research Institute Europe GmbH.

Bernhard Sendhoff is a Senior Member of the IEEE and of the ACM. His research focuses on methods from computational intelligence and their applications in development, production and services. He has authored and co-authored over 150 scientific papers and over 30 patents.