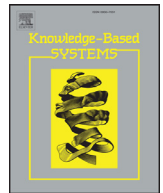




Contents lists available at ScienceDirect

Knowledge-Based Systems

journal homepage: www.elsevier.com/locate/knosysReference line-based Estimation of Distribution Algorithm for many-objective optimization^{*}Yanan Sun^a, Gary G. Yen^{b,*}, Zhang Yi^a^a College of Computer Science, Sichuan University, Chengdu, Sichuan 610065, China^b School of Electrical and Computer Engineering, Oklahoma State University, Stillwater, OK 74078, USA

ARTICLE INFO

Article history:

Received 27 January 2017

Revised 13 June 2017

Accepted 14 June 2017

Available online xxx

Keywords:

Many-objective evolutionary algorithm

Evolutionary computation

Reference line

Estimation of Distribution Algorithm

ABSTRACT

Multi-Objective Evolutionary Algorithms (MOEAs) are preferred in solving multi-objective optimization problems due to their considerable performance giving decision-maker a set of not only convergent but diversified promising solutions. However, the scalability of MOEAs deteriorates in addressing many-objective optimization problems which involve more than three conflicting objectives. The principal reason is largely due to the deficiency of the existing genetic operators which cannot generate promising offspring from parents chosen by the Pareto-dominance rule in these MOEAs. Estimation of Distribution Algorithms (EDAs) generate offspring with a probabilistic model built from the statistics extracting upon existing solutions to expectedly alleviate the weakness arisen in genetic operators. In this paper, a reference line-based EDA is proposed for effectively solving many-objective optimization problems. Specifically, the estimation model is built based on the reference lines in the decision space to sample solutions with *favorable proximity*. Then solutions with *considerable diversity* in Pareto-optimal front are selected. These two phases collectively promote the needed convergence and diversity for the proposed algorithm. To evaluate the performance, extensive experiments are performed against four state-of-the-art many-objective evolutionary algorithms and two EDAs over DTLZ and WFG test suites with 5-, 8-, 10-, and 15-objective. Experimental results quantified by the selected performance metrics indicate that the proposed algorithm shows significant competitiveness in tackling many-objective optimization problems.

© 2017 Elsevier B.V. All rights reserved.

1. Introduction

A reference line-based Estimation of Distribution Algorithm (EDA) Many-Objective Evolutionary Algorithm (MaOEA) is proposed in this paper for efficiently tackling Many-objective Optimization Problems (MaOPs). Specifically, MaOEAs aim at solving m conflicting objectives concurrently and m is greater than three [1,2]. Generally, a given MaOP is mathematically formulated by Eq. (1):

$$f = (f_1(x), \dots, f_m(x))^T \quad \text{subject to } x \in \Omega \quad (1)$$

where $\Omega \in \prod_{i=1}^n [x_i^l, x_i^u]$ is the feasible space and $-\infty < x_i^l \leq x_i^u < \infty$ for $i = 1, \dots, n$. The mapping f from Ω to \mathbb{R}^m defines m conflicting

objective functions $f_i(x)$ and $i = 1, \dots, m$. Without loss of generality, it is assumed that $f_1(x), \dots, f_m(x)$ are to be minimized, as maximization problems can be transformed into minimization problems by duality principle. Because MaOPs widely exist in many real-world applications, such as calibration problems of automotive engine with 10-objective [3], management in land exploitation with 14-objective [4], to name a few, there is a strong incentive for efficiently and effectively solving MaOPs.

In contrast to a single-objective optimization problem in which one single global optimal solution is desired, a set of trade-off solutions, namely Pareto-optimal solutions, is preferred in a MaOP to give decision-makers more alternatives for their preferences due to the conflicting nature among the objectives. Moreover, the Pareto-optimal solutions in the decision space constitute the Pareto-optimal Set (PS) while their images correspond to the Pareto-optimal Front (PF) in the objective space [5]. MaOEA has been recognized as one of the promising paradigms to address MaOPs due to its population-based nature and meta-heuristics search ability to be capable of obtaining a set of solutions approximating the PF in a single run. Generally speaking, MOEAs aim at pursuing two distinct, yet complement, goals throughout the evolution process:

^{*} This work is supported in part by the scholarship from China Scholarship Council, under Grant 201506240048; in part by the Miaozi Project in Science and Technology Innovation Program of Sichuan Province under Grant 16-YCG061, China; and in part by the National Natural Science Foundation of China under Grants 61432012 and U1435213.

^{*} Corresponding author.

E-mail addresses: yanan_sun@outlook.com (Y. Sun), gyen@okstate.edu (G.G. Yen), zhangyi@scu.edu.cn (Z. Yi).

1) improving the convergence and 2) preserving the diversity among solutions.

1.1. Introductions to algorithms in solving MaOPs

During the past several decades, numerous Multi-Objective Evolutionary Algorithms (MOEAs), such as elitist Non-dominated Sorting Genetic Algorithm (NSGA-II) [6], advanced version of Strength Pareto Evolutionary Algorithm (SPEA2) [7], have been developed for dealing with Multi-objective Optimization Problems (MOPs) [8,9] in which two or three objectives are to be optimized simultaneously [1,2]. However, these MOEAs severely degrade their performance in handling MaOPs. The major reason is the loss of selection pressure, i.e., a large proportion of individuals is all non-dominated, which leads to the Pareto-based primary selection approach losing its selecting effectiveness [10]. Furthermore, most of these solutions are only the bests at very few objectives but fairly worse in others, which is known as the Dominance Resistant (DR) solutions [11–13]. As a result, the density-based secondary diversity improving mechanism is activated, which is recognized as the Active Diversity Promotion (ADP) [12], to determine which solutions survive into the next generation. Due to the DR and ADP phenomena, the selected solutions are not moving uniformly towards the PF as the evolution continues and could even be stagnated at far away [14].

To this end, various evolutionary algorithms for solving MaOPs, namely MaOEAs,¹ have been specifically designed for addressing MaOPs. In summary, these MaOEAs manage the adverse impact of DR and ADP from different schemes: 1) exploiting novel approaches to strengthen the diversity of selected solutions, and 2) enhancing the comparison mechanism of the traditional dominance relationship. For example, the reference point-based Non-dominated Sorting Genetic Algorithm for many-objective optimization (NSGA-III) [15] improves the diversity by checking which one has the nearest perpendicular distance to the reference lines, and these reference lines are manually set or evenly sampled in the unit hyper-plane of the objective space. The Multi-Objective Evolutionary Algorithm based on Decomposition (MOEA/D) [16] decomposes one considered MaOP into a set of scalar optimization sub-problems, and then optimizes them simultaneously by a group of predefined well distributed weight vectors aggregating all objectives. Especially, the diversity of MOEA/D is maintained by these weight vectors which direct the population towards different areas of the PF. Furthermore, the dominance comparison occurs only in the neighboring solutions, which reduces the adverse impact of the DR to some extent. The Grid-based Evolutionary Algorithm for many-objective optimization (GrEA) [17] employs the grid-based fitness comparison technique to relax the dominance relationship of solutions. Other grid-based quantitative measurements are also incorporated into the fitness value to strengthen the diversity in mating and environmental selections. In the Hypervolume-based many-objective Evolutionary algorithm (HypE) [18], the selection is preferred given the fitness assigned by the corresponding hypervolume contribution which is estimated by the Monte Carlo simulation. This fitness assignment concurrently avoids the deficiency of traditional dominance comparison mechanism and improves the diversity. In addition, the same selection principle is also employed in its mating selection.

1.2. Launching of EDAs

In evolutionary algorithms, individuals evolving towards the regions of the promising solutions are generated from their elite

parents by genetic operators, such as crossover and mutation. Recently, existing genetic operators have been extensively investigated over a set of scalable benchmark test problems, which concludes that existing genetic operators are crucial to the performance of MOEAs, and cannot guarantee promising solutions to be generated especially in MaOPs [19]. Furthermore, most genetic operators require a set of parameters to be empirically assigned in advance, such as the probabilities of crossover and mutation, the distribution indexes in widely used simulated binary crossover (SBX) [20] and polynomial mutation [21] operators. As a consequence, these deficiencies of the existing genetic operators collectively motive the development of the Estimation of Distribution Algorithm (EDA) [22–25] which is a kind of computing paradigm to solve MOPs [26–29] by generating new promising offspring with the probabilistic models built based on the solutions the algorithms have visited.

Typically, EDAs-based MOEAs are classified into two distinct categories based on their employment of estimation models. The first category is often known as the mixture probability model-based EDAs. For instance, the multi-objective mixture-based iterated density estimation evolutionary algorithm [26] utilized the mixed probability distributions to sample well-distributed solutions and the multi-objective Parzen-based EDA [30] learnt from the Gaussian and Cauchy kernels to build its models. In [31], the multi-objective hierarchical Bayesian optimization algorithm was designed by the mixture Bayesian network-based probabilistic model for discrete multi-objective optimization problems. In addition, the multi-objective extended compact genetic algorithm [32] employed a marginal product model as the mixture of probability model. Furthermore, a Regularity-based Model EDA (RM-MEDA) was proposed lately in [28] in which the model is built based on the mixture normal distribution over the regularity. The other category covers the Bayesian network-based EDAs. Example include the multi-objective Bayesian optimization algorithm [33] utilizing the Bayesian Optimization Algorithm (BOA) to build a Bayesian network as its model for generating offspring. In addition, a related work was investigated in [34] to predict the model by strengthening Pareto ranking approach [35] and BOA. Furthermore, Laumanns and Ocenasek in [36] proposed a Bayesian multi-objective optimization algorithm whose model was built over the solutions selected by the ϵ -Pareto ranking method [37]. Moreover, an improved non-dominated sorting approach was employed by decision tree-based multi-objective EDA [38] to select a subset of solutions serving for a regression decision tree to learn the model. Recently, the Multi-dimensional Bayesian Network EDA (MBN-EDA) was proposed in [29] specifically for addressing MaOPs.

1.3. Motivation of the proposed algorithm

It is believed that EDAs are capable of solving MaOPs without suffering the weaknesses of MOEAs employing genetic operators. Although, MBN-EDAs have been tested to be with the ability in addressing MaOPs, the development of Many-objective Optimization EDAs (MaOEDA) is still in its infancy. Especially, probability models based on regularity have been extensively investigated in the discipline of statistical learning [39,40], and regularity-based models are easier to build, and fairly effective. In addition, regularity-based EDAs can visually help decision-makers to easily select their desirable solutions. Therefore, a variety of regularity-based MOEAs have recently been successfully proposed for dealing with MOPs [28,41,42]. Moreover, because uniformly reference points are capable of directing the search towards the exceptional solution locations, instead of a heuristic exploration in a large space of MaOPs, the reference points related paradigms have been successfully employed in MaOEAs, such as MOEA/D and NSGA-

¹ In this context, MaOEAs include the algorithms which are originally designed for MOPs, and now are extended for MaOPs.

III. In addition, the approximation of the PS is also necessary to ease the multi-criteria decision-making in some real-world applications [43–47], while most existing MaOEAs solely concern on the approximation of the PF. Consequently, it is reasonable to take advantage of the regularity into the design of EDAs combined with the reference point techniques to collectively address MaOPs.

As the first attempt to seize as well as employ the regularity of the PS combined with reference points and based on our recent research findings in addressing MaOPs [48,49–53], we have proposed the reference line-based EDA for solving continuous MaOPs, in short named MaOEDA/RL, to not only obtain a group of solutions uniformly distributed in the PS but also in the PF. Compared to the existing MaOEAs, the DR solutions were not discriminated in the proposed algorithm but encouraged for building an approximate utopian PS (i.e., extreme points lying at the axis of decision space were detected for constructing a hyper-plane at which uniformly distributed reference points stand). Next, offspring were generated with a heuristic step size by sampling along the reference lines which are constructed by connecting the origin to all the reference points. In addition, dimension reduction in the decision space was performed to accelerate the exploration of the sampling process. In further, a set of uniformly distributed solutions lying at the PS is selected with the help of the non-dominated comparison technique. Finally, the reference lines-based method is utilized once more in the objective space to select a set of well-distributed solutions approximating the PF. In summary, the convergence of the proposed algorithm is guaranteed by sampling new solutions along the reference lines, which steers the search towards the PS. To be specific, sampling solutions in the boundary reference lines which are generated by connecting origin to the reference points lying at the axes improve the spread of the solutions in the objective space. Combined with the sampling process in the internal reference lines (which are not the boundary reference lines) and selection with the help of reference lines generated in the objective space, the diversity of the proposed algorithm is improved.

1.4. Major contributions

In summary, the contributions of the proposed algorithm compared to the existing MaOEAs as well as the EDAs are outlined as follows:

1. The DR phenomenon is harmful to evolutionary progress in solving MaOPs. Existing MaOEAs try to employ new dominance comparison or diversity mechanism to alleviate the adverse effects. However, the proposed algorithm recognizes this phenomenon and takes full advantage of these DR solutions to assist in solving MaOPs.
2. Convergence and diversity are conflicting goals since both of them are processed simultaneously in the existing MaOEAs. The proposed algorithm considers them in two different phases, and does not pursue the trade-off between the convergence and diversity, but maximize the benefit over each of them. Finally, a group of solutions approximating the PS as well as the PF is obtained.
3. Compared to the evolutionary-based algorithms, no genetic operator is employed in the proposed EDA design. An explicit sampling model is proposed to generate promising offspring which direct the search towards the region with promising convergence. Moreover, as one type of EDAs that takes the consideration of regularity, the proposed algorithm can be effectively used in solving MaOPs, while RM-MEDA is only designed for MOPs. In addition, the dimension reduction technique is utilized by the proposed algorithm in the decision variable space to speed up the exploration of sampling process.

The rest of this paper is organized as follows. In Section 2, state-of-the-art MaOEAs and EDAs are reviewed which naturally lead to the motivation of the proposed algorithm. Section 3 is devoted to elaborate the details of the proposed algorithm. In addition, chosen benchmark test problems, experimental settings, and considered performance metrics are illustrated in Section 4 for quantitative and qualitative comparisons of the proposed algorithm against some peer competitors. Furthermore, statistical experimental results and discussions are also presented. Finally, Section 5 concludes this paper and provides directions for future works.

2. Related works and motivation

In this section, state-of-the-art designs related to the proposed algorithm are thoroughly surveyed. For convenience, these algorithms are classified into two different categories. One is the traditional MaOEAs, represented by NSGA-III, MOEA/D, HypE, and GrEA. The other one is EDAs, composed of RM-MEDA and MBN-EDA. Next, the motivation of the proposed algorithm compared to these MaOEAs, and the superiority of the proposed algorithm against chosen EDAs are presented. Especially, the significant distinctions to RM-MEDA which is the most similar EDA to the proposed algorithm are highlighted.

Furthermore, the problem formulated by Eq. (1) is utilized for conveniently discussing these state-of-the-art algorithms. In addition, it is assumed that there are P solutions and Q offspring which have been generated in the current generation t , while only N slots are available. Moreover, $P + Q > N$.

2.1. Traditional many-objective evolutionary algorithms

2.1.1. NSGA-III

The diversity maintaining scheme of NSGA-III is the major difference compared to its predecessor NSGA-II designed specifically to address MaOPs. To be specific, it takes effect when $P + Q > N$ in the environmental selection which is composed of three distinct phases. First, these $P + Q$ individuals are sorted into the different fronts F_1, \dots, F_r based on their dominance relationship. Let R_t and S_t denote all the solutions in $\bigcup_{j=1}^{l-1} F_j$ and $\bigcup_{j=1}^l F_j$, respectively, where $|R_t| < N < |S_t|$ and $|\cdot|$ is a countable operator. The second phase is the normalization in which all the objectives of the individuals in S_t and the reference points are normalized. Next, the perpendicular distances of each solution to the reference lines generated by connecting the origin to the reference points are calculated, then the reference point to whom the solution has the smallest perpendicular distance is marked. Finally, $N - |R_t|$ individuals are selected one by one from F_l based on the niche preservation operation.

2.1.2. MOEA/D

Like most peer algorithms, MOEA/D starts to work with a group of randomly initialized solutions based on which the evolution takes effect. In addition, an empty external population is created, and a set of well distributed weight vectors w_1, \dots, w_k in the objective space is given, where $w_i \in \mathbb{R}^n$ ($i = 1, \dots, k$). Initially, the mutual Euclidean distances of weight vectors are calculated, then top T vectors are selected as the neighbors of w_i by ascending the distances of w_j to w_i ($j \in \{1, \dots, k\}$ and $j \neq i$), and the indexes of the neighbors of w_i are denoted as $B(i) = \{i_1, \dots, i_T\}$. Hopefully, each weight vector w_i has a corresponding solution x_i . Next, two elements j and l are randomly selected from $B(i)$, and the offspring z^i is generated by the solution x^j and x^l . Moreover, the problem-specific repair or heuristic improvement is applied on z^i to generate $z^{i'}$. Finally, $z^{i'}$ is included into the external population if no solutions in the external population dominate $F(z^{i'})$, and the solutions which are dominated by $F(z^{i'})$ are removed from the external population. When the stopping criteria are satisfied, MOEA/D stops

and gives decision-makers the final solutions that exist in the external population.

2.1.3. HypE

HypE is a hypervolume-based evolutionary algorithm, and all the offspring are selected based on their fitness which are calculated from their respective contributions in hypervolume. To be specific, the hypervolume contribution of the solution x is calculated as the overall hypervolume differences between F_i' and F_i , and F_i is denoted by $F_i' \cup x = F_i$. Consequently, the entire hypervolume of F_i are divided into $|F_i|$ partitions which denote the hypervolume contributions of each solution in F_i . In order to effectively calculate the hypervolume, especially in a high dimensional objective space, the Monte Carlo Simulation is utilized to estimate the contribution. To be specific, one reference point is given, and a set of uniformly distributed vectors are sampled in the objective space first. Then, the hit vector is identified by filtering whether it is dominated by the solutions in F_i and dominates the reference point. Finally, the number of hits are counted and multiplied by the hit ratio with the volume of the sampling box. Moreover, one is randomly selected when multiple solutions have the same hypervolume fitness. Noted here that the contribution of hypervolume is also employed in the mating selection phase.

2.1.4. GrEA

GrEA exploits the potential of the grid-based method to improve the selection pressure towards the PF, and collect the wide spread and evenly distributed solutions. In contrast to the other MOEAs based on the grid technique, GrEA proposes an effective approach to construct the grid coordinate system, and employs the modified Pareto dominance criteria as well as the density estimator to tackle MaOPs. To be specific, an adaptive grid system is constructed to adjust the size and the location of a grid so as to be capable of covering new generated solutions. Moreover, the location of the grid at which the solution lies is utilized to replace its raw fitness for comparing the dominance relationship, which apparently relaxes the Pareto dominance relationship thus improves the selection pressure. Furthermore, the enlarged region is considered for estimating the density by introducing the neighbors. In addition, three other criteria are also introduced to collectively assign the fitness which is employed further for the mating and the environmental selections.

2.2. Estimation of distribution algorithms

2.2.1. RM-MEDA

RM-MEDA is one kind of regularity-based EDAs which employs the probabilistic model built from the regularity of the PS to generate new offspring. To this end, one probabilistic model needs to be built first with the statistical regularity information extracted from the solutions visited by the population. To be specific, the local PCA [54] approach is utilized to partition the selected solutions into a number of subparts, then each subpart is modeled. Specifically, for each part S^j , its eigenvalues are descended and denoted by $\lambda_1^j, \dots, \lambda_m^j$, their corresponding eigenvectors are denoted by U_1^j, \dots, U_m^j , and the mean of all the solutions in S^j is denoted by \bar{x}^j . Consequently, S^j is modeled with the manifold consumption that the dimension of the PS in a m -dimensional continuous problem is $(m - 1)$. Next, offspring are sampled from S^j with the proportion being equal to how much the volume of S^j accounts for the volumes of all the models. Once all offspring are sampled, non-dominated solutions with maximum diversity are selected from the union of the current population and offspring, then new models are rebuilt and offspring are generated until the stop criteria are satisfied.

2.2.2. MBN-EDA

MBN-EDA is a recently proposed EDA and also the first EDA addressing MaOPs. It employs a Multi-dimensional Bayesian Network (MBN) to build the jointly probabilistic model of the variables and the objectives. To be specific, the variables and the objectives are denoted by the features and the classes of one MBN, respectively, and the relationship from the classes to the features are denoted by bridges. Then, features, classes, and bridges with their interactions are formulated as three sub-graphs which encode the dependencies of variables, objectives, variables and objectives, respectively. In the process of training this MBN, a subset of solutions are sampled as the training data. Next, one search algorithm is employed to find all the possible structures that match the training data as much as possible. Finally, Bayesian information criterion is employed to rank which structure is the best based on the log-likelihood criterion. When the structure is confirmed, the dependencies are learnt and the original problem can be decomposed into several sub-problems without any correlations. Furthermore, the probabilistic logic sampling approach [55] is used to generate new solutions from this model. To address the deficiency of the widely used non-dominated sorting algorithm in solving MaOPs, four special ranking methods are incorporated (more details can be found in [29]).

In summary, existing MaOEAs and the MBN-EDA are designed by explicit or implicit favor to strengthen the selection pressure towards the PF and keep the diversity as much as possible. To be specific, because of the loss of selection pressure and the DR phenomenon becoming increasingly severe as the number of the objectives increases, new fitness assignment strategies are proposed or effective Pareto-optimal comparison techniques are employed, such as GrEA and HypE. In order to explicitly strengthen the diversity of the solutions, NSGA-III proposes the evenly distributed reference points-based mechanism in the objective space to ensure the individuals which are closer to these well distributed reference lines to be selected. Furthermore, MOEA/D decomposes one considered MaOP into a group of sub-problems by a set of evenly distributed weights to improve the diversity, then the objectives are accumulated by the weights to be solved to improve the selection pressure in each sub region. In addition, the recombination in MOEA/D is performed with the parents which are selected from the neighbor regions to avoid the unpromising offspring, which is adverse to the convergence. Although MBN-EDA is based on EDA, and it proposes special ranking methods for solving MaOPs with the purpose of strengthening selection pressure and maintaining diversity, it is not easy to train a MBN with good performance to sample promising offspring. For example, in the optimization problems with many-to-one mappings, a set of balance training data is difficult to obtain, which affects the full dependencies between the features and the classes of the MBN to be captured.

The proposed algorithm works on the basis of the semi-finished solutions which are produced by solving the scalar problems transformed by accumulating the objectives with a set of evenly distributed weight vectors. Although most part of these semi-finished solutions are DR, they introduce the information to construct a utopian PS. Then uniformly distributed reference points are simulated and translated into this utopian PS. Furthermore, the semi-finished solutions, as a part of Pareto-optimal solutions, provide sufficient statistics for the dimension reduction technique. Finally, the reference lines generated by connecting the origin to the boundary reference points enlarge the spread of the solutions, and the reference lines generated by connecting the origin to the interior reference points maintain the diversity. However, we do not know the exact regularity of the PF in real-world applications, and the diversity in the PS does not necessarily map the same images in the PF. To this end, the reference lined-based selection in the objective space is also presented in the proposed algorithm to

solve this issue. Although RM-MEDA is the most similar algorithm to the proposed algorithm based on the regularity-based model, the proposed algorithm presents significant differences which are presented as follows:

1. The model built in RM-MEDA employs local PCA technique that involves the matrix factorization operation which is often time-consuming especially for MaOPs which have a large search space. However, the model in the proposed algorithm is built by the reference lines, which is simple and effective.
2. Because of the DR solutions in MaOPs, local PCA takes effects in RM-MEDA just like the results of sampling new solutions in the boundary reference lines of the proposed algorithm, which is effective for MOPs whose PS are a front and in which the search space is small. The proposed algorithm sampling solutions with the help of the interior reference lines is more effective for the problems whose PS are the closed hypercubes in a large search space (i.e., MaOPs).
3. The models in RM-MEDA are built with the manifold assumption which requires sufficient points to represent its structure, and this is not suitable for the MaOPs because they have a large search space and cannot have enough points to meet the manifold assumption. However, the proposed algorithm does not demand such an assumption.

3. Proposed algorithm

In this section, the details of the proposed Reference Line-based Estimation of Distribution Algorithm for Many-Objective optimization problems (in short named MaOEDA/RL) are documented. Particularly, the framework of the proposed algorithm is presented first in Section 3.1. This is followed by the details of the proposed MaOEDA/RL in Sections 3.2–3.5 with the context of Eq. (1). Next, in Section 3.6, the parameter settings in the proposed algorithm are discussed, and the reasonable values are recommended. Finally, the computational complexity of the proposed MaOEDA/RL is analyzed, and the proposed algorithm is further discussed in Sections 3.7 and 3.8, respectively.

3.1. Framework of the proposed algorithm

The proposed algorithm employs the potential of the reference lines which are evenly distributed in a subspace of the decision space to build models, and then new promising solutions are sampled based on these built models. To be specific, a set of individuals are randomly initialized at first. Then, the dimension reduction technique is applied to expectedly reduce the volume of the search space in the proposed algorithm (Section 3.2). Next, the models are built upon the generated reference points (Section 3.3), and a set of uniformly distributed solutions in the decision space are sampled (Section 3.4). Finally, the representative solutions which are evenly distributed in the objective space are selected (Section 3.5). In summary, the framework of MaOEDA/RL is illustrated in Algorithm 1.

3.2. Dimension reduction

The dimension reduction of the proposed algorithm is performed upon the decision space, which expectedly reduces the volume of the feasible space for speeding up the search. Obviously, the best space with reduced dimension is the subspace of the PS, because the solutions arbitrarily sampled from the PS are all with promising convergence, and the remaining work is to pursue the diversity [6,28]. For this purpose, a set of training data for performing the dimension reduction is required. Noted here that, the training data only require the solutions with convergence while

Algorithm 1 Framework of the proposed algorithm.

Input: The required number of solutions in the decision space n_p .
The required number of solutions in the objective space n_o .
The step size for sampling solution α .
Output: n_p representative solutions in the decision space.
 n_o representative solutions in the objective space.
1: Randomly initializing the population;
2: Dimension reduction;
3: Generating reference points in the utopian Pareto set;
4: Building models and sampling solutions P_{gene} ;
5: $Q \leftarrow$ Final Selection;
6: **return** P_{gene} , Q

the diversity is not necessary. Consequently, these training data are called the semi-finished solutions (see Definition 1), because the full-finished solutions in MOEAs and MaOEAs are with promising both convergence and diversity properties, while only the convergence of these training data is required.

Definition 1. The semi-finished solutions refer to the solutions with promising convergence but inferior diversity.

Generally speaking, a set of solutions with only convergence is easy to obtain by the conventional weighted aggregation method [56]. For convenience of the development, let P_{semi} denote the semi-finished solutions. Specifically, Algorithm 2 presents the

Algorithm 2 Obtaining the semi-finished solutions.

Input: The initialized population P_{init} .
Output: The semi-finished solutions P_{semi} .
1: $W = \{w^1, \dots, w^k\} \leftarrow$ Initialize k well distributed weight vectors in the objective space;
2: $P_{semi} \leftarrow \emptyset$;
3: **for** $i \leftarrow 1$ **to** k **do**
4: $f(w^i) \leftarrow w_1^i f_1 + \dots + w_m^i f_m$;
5: $P_{w^i} \leftarrow$ Solve $f(w^i)$;
6: $P_{semi} \leftarrow P_{semi} \cup P_{w^i}$;
7: **end for**
8: **return** P_{semi}

details of obtaining P_{semi} . Firstly, a set of well-distributed weight vectors in the objective space are sampled (line 1). Noted here that the Das and Dennis's systematic method [57] is employed to generate these weight vectors. Secondly, these weight vectors individually translate the m -dimensional problem to be optimized into a single-objective problem by weighing each objective with the corresponding elements in the weight vector (line 4). Thirdly, these single-objective optimization problems are solved and their solutions are returned as the P_{semi} (lines 5 and 8). Actually, any algorithm can be utilized to solve the translated single-objective optimization problems, while a genetic algorithm implemented by ourselves is employed in the proposed algorithm for the purposed of simplicity.

When the semi-finished solutions P_{semi} are obtained, the dimension reduction can be performed. It is assumed that the population size of P_{semi} is N . The matrix X is used to denote the representation of P_{semi} , and each row of X denotes one solution in P_{semi} , i.e., $X \in \mathbb{R}^{N \times n}$. Algorithm 3 illustrates the process of the dimension reduction on P_{semi} . Specifically, lines 3–7 are intended to find the dimensions that can be temporarily removed based on the comparison whether the values of their respective standard derivations are less than or equal to a pre-defined threshold β , then the indexes of these dimensions are saved in A . In addition, the mean values on these reduced dimensions are calculated then placed in

Algorithm 3 Dimension reduction.**Input:** Threshold β .**Output:** The removed dimension indexes A .

The mean values of the removed dimension B .
 The same-finished solutions X with reduced dimensions.

```

1:  $[std_1, \dots, std_n] \leftarrow$  Compute the standard deviation of each column in  $X$ .
2:  $A \leftarrow \emptyset, B \leftarrow \emptyset$ ;
3: for  $i \leftarrow 1$  to  $n$  do
4:   if  $std_i \leq \beta$  then
5:      $A \leftarrow A \cup i$ ;
6:   end if
7: end for
8: for each  $i$  in  $A$  do
9:    $v \leftarrow$  Compute the mean value of  $i$ -th column in  $X$ ;
10:   $B \leftarrow B \cup v$ ;
11: end for
12:  $X \leftarrow$  All the columns whose indexes are in  $A$  are removed from  $X$ ;
13: return  $A, B, X$ 
```

B , details of which are described in lines 8–11. Finally, X is updated in line 12 by removing these columns whose indexes are placed in A . It is noticed here that consecutive operations of MaOEDA/RL are mainly based on the updated X . A and B are used in further development. To conveniently discuss further, let $p = m - |A|$, i.e., the dimension of the updated X is p .

3.3. Generating reference points

This procedure is divided into two steps. The first step is to generate N' reference points $rp_{norm} \in \mathbb{R}^{N' \times p}$ in a normalized hyperplane. The other is to map these reference points to the utopian PS denoted as H_d , which is constructed by the extreme points and motivated by the works in [6,28]. To be specific, in the first step, we employ the Das and Dennis's method [57] to generate uniformly distributed reference points in the p -dimensional normalized hyperplane. We noted here that this approach is also employed in NSGA-III [15]. In the second step, the minimum and the maximum values in the j -th dimension of X are calculated first by Eqs. (2) and (3), respectively,

$$z_j = \min X_i^j, \quad i = 1, \dots, N \quad (2)$$

$$z_j^* = \max X_i^j, \quad i = 1, \dots, N \quad (3)$$

then the j -th extreme point ep_j is set to be $[0, \dots, z_j^*, \dots, 0]$. Furthermore, H_d which is constructed by the extreme points ep_i ($i = 1, \dots, p$) will be mapped to the one that has unit magnitude intercepts on each axis, which involves computing the solutions of a group of linear equations. Specifically, we denote the scaled intercepts on each axis as $C = [a_1, \dots, a_p]^T$, then C is found by solving the Eq. (4)

$$EP \times C = I, \quad (4)$$

where $EP = [ep_1, \dots, ep_p]^T$ and $I = [1, \dots, 1]^T$. Note here that, all the elements in EP are linear independent and a deterministic C is obtained. Finally, rp_{norm} are mapped to rp_{dec} by Eq. (5)

$$(rp_{dec})_j^i = (rp_{norm})_j^i \times (a_i - z_i) + z_i, \quad (5)$$

where $(rp_{dec})_j^i$ and $(rp_{norm})_j^i$ denote the value of the j -th element of i -th reference point in rp_{dec} and rp_{norm} , respectively. In summary, the complete steps of this procedure are summarized in Algorithm 4.

Algorithm 4 Generating reference points in H_d .**Input:** The reference points rp_{norm} uniformly generated in the normalized p -dimensional hyperplane.**Output:** A set of uniformly distributed reference points rp_{des} in H_d .

```

1:  $[z_1, \dots, z_p] \leftarrow$  Compute the minimum value of each dimension in  $X$ ;
2:  $[z_1^*, \dots, z_p^*] \leftarrow$  Compute the maximum value of each dimension in  $X$ ;
3: Translate each element in  $X$ ;
4:  $[a_1, \dots, a_p] \leftarrow$  Calculate the scaled intercept of each dimension;
5:  $rp_{des} \leftarrow$  Map  $rp_{norm}$  into the hyperplane  $H_d$ ;
6: return  $rp_{des}$ 
```

3.4. Building model and sampling solutions

In the process of building the model, the reference line \vec{l}_i is constructed first by connecting the origin to the reference point $(rp_{des})^i \in rp_{des}$. Then the model is built by Eqs. (6) and (7) ($i = 1, \dots, p$), and new solutions are sampled.

Algorithm 5 Building model and sampling solutions.**Input:** The step size α .The stop size k .Uniformly distributed reference points rp_{des} .**Output:** Non-dominated solutions P_{gene} .

```

1: Building models;
2:  $P'_{gene} \leftarrow \emptyset$ ;
3:  $P_{gene} \leftarrow P_{semi}$ ;
4: while termination conditions are not satisfied do
5:    $P' \leftarrow$  Sample new solution;
6:    $P'_{gene} \leftarrow P'_{gene} \cup P'$ ;
7:    $F \leftarrow$  Evaluate the fitness of  $P'$ ;
8:   if  $P'$  is non-dominated to solutions in  $P_{gene}$  then
9:      $P_{gene} \leftarrow P_{gene} \cup P'$ ;
10:  end if
11: end while
12: return  $P_{gene}$ 
```

$$\vec{l}_i + k_1 \cdot \alpha \cdot \vec{l}_i / |\vec{l}_i| \quad (6)$$

$$\vec{l}_i - k_2 \cdot \alpha \cdot \vec{l}_i / |\vec{l}_i| \quad (7)$$

It is noted here that sampling solutions in each reference line is composed of two distinct aspects: 1) forward sampling and 2) backward sampling which are illustrated by Eqs. (6) and (7), respectively. Moreover, k_1 and k_2 denote how many solutions that have been sampled when the sampling process terminates, and their values are assigned from $\{0, \dots, \max(k_1)\}$ as well as $\{1, \dots, \max(k_2)\}$, respectively. In addition, termination conditions are that 1) any element of the new generated solution is beyond its scope or 2) generated solutions are all dominated by solutions in P_{gene} when $k_1 \leq k$ and $k_2 \leq k$ (k is the stop size) or 3) the generated solution is dominated by the solutions from P_{semi} . Because rp_{des} are uniformly distributed in the H_d , and α is a constant, the solutions sampled in P_{gene} are approximately uniformly distributed. Moreover, the solutions sampled from the boundary reference lines improve the spread of the PS (See Fig. 1c and d). In addition, the fitness of new sampled solution is evaluated with Algorithm 6. As a consequence, a group of evenly distributed solutions with widely spread approximating the PS are generated with non-dominated comparison when Algorithm 5 terminates.

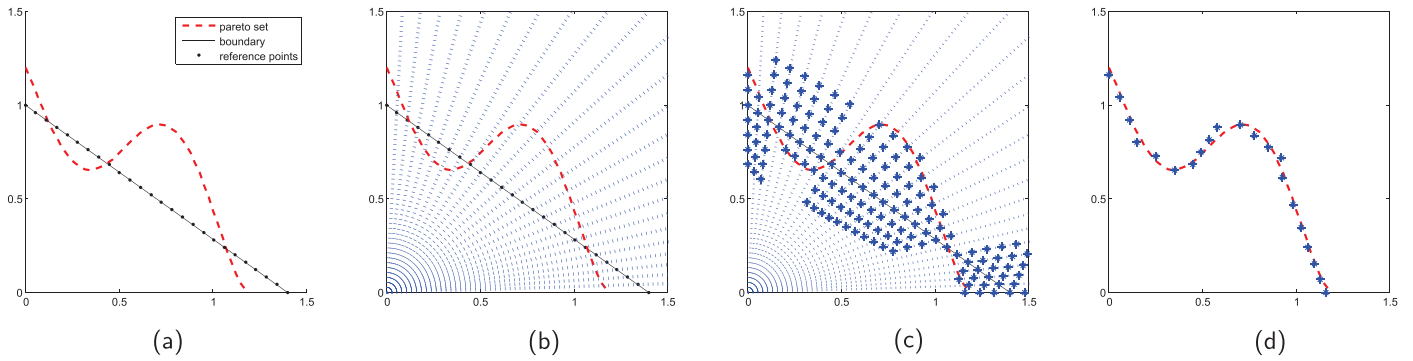


Fig. 1. A schematic diagram of Algorithm 5 in a bi-dimensional decision space. More specifically, in (a), the dashed line denotes the Pareto-optimal set, the solid line denotes the utopian boundary constructed by connecting two extreme points laid at the axes, and the circles are the transformed reference points. The dotted lines in (b) denote the reference lines in the space constrained by the range of the decision variables. Note that, the lines from origin to extreme points are also the reference lines, particularly, the boundary reference lines. The cross points in (c) denote the generated solutions (i.e., P_{gene}) by models 6 and 7 with $k = 3$ in this example. Finally, The cross points in (d) denote the uniformly distributed solutions, with wide spread approximating the Pareto-optimal set, selected by non-dominated comparison.

Algorithm 6 Evaluating fitness on P' .

Input: The solutions P' .

The removed dimension indexes A .

The mean values of the removed dimension B .

Output: Fitness F .

- 1: **for each** element j in A **do**
 - 2: $(P')^j \leftarrow$ insert B_j into the j -th interspace of P' ;
 - 3: **end for**
 - 4: $F \leftarrow (f((P')^1), \dots, f((P')^m))$;
 - 5: **return** F
-

3.5. Final selection

This is the final step of the proposed algorithm in which n_0 solutions with maximum diversity in the PF are selected from the Pareto-optimal solutions P_{gene} . To be specific, this part is composed of two steps in which the first step is the normalization of the PF, while the second one is the environmental selection. Moreover, the normalization of the PF is operated on the objective values F of P_{semi} , and the details are presented as follows.

1. Minimum values in the j -th dimension of F is calculated first by Eq. (8) where F_j^i denotes the j -th element of the i -th solution in F , and $|F|$ denotes the cardinality of F .

$$z_j = \min_i F_j^i, \quad i = 1, \dots, |F| \quad (8)$$

2. Extreme points $ep_i (i = 1, \dots, |F|)$ in each dimension are obtained by Eq. (9) where $w_k = 10^{-6}$ for $k \neq i$ while $w_k = 1$ for $k = i$.

$$ep_i = \min_{l=1}^{|F|} \max_{k=1}^m \frac{\bar{F}_k^l}{w_k} \quad (9)$$

3. Intercepts are calculated by Eq. (4).
4. Each element F_j^i of F is translated by Eq. (10).

$$F_j^i = \frac{F_j^i - z_j}{a_j - z_j} \quad (10)$$

Noted here that, an alternative method proposed in [58] is employed when EP has duplicate points which lead to multiple solutions to Eq. (4).

In addition, the final selection is presented in Algorithm 7. To be specific, n_0 solutions are selected from lines 8–12 by controlling the magnitude of the parameters. Furthermore, lines 12–15 take effect when extreme bias exists or the feature of PF is disconnected in the solving problem.

Algorithm 7 Final selection.

Input: The translated objectives F .

Output: The selected solutions Q .

- 1: Let $N = n_0$ and generate N uniformly distributed reference points in a m -dimensional unit hyperplane;
 - 2: Connected the origin to each reference point to formulate reference lines l_1, \dots, l_N ;
 - 3: Let $c_1 = \dots = c_N \leftarrow \emptyset$;
 - 4: **for** $i \leftarrow 1$ to $|F|$ **do**
 - 5: Calculate the perpendicular distance between the i -th element F^i of F to all the reference lines;
 - 6: Find the reference line l_k that has the nearest distance to F^i , and let $c_k = c_k \cup F^i$;
 - 7: **end for**
 - 8: **for** $i \leftarrow 1$ to N **do**
 - 9: **if** $|c_i| > 0$ **then**
 - 10: Find the element F^j from $|c_i|$ that F^j has the smallest perpendicular to l_i , and put the corresponding solution into Q ;
 - 11: **end if**
 - 12: **end for**
 - 13: **while** $|Q| < N$ **do**
 - 14: Randomly select one solution from $P_{semi} \setminus Q$ and put it into Q ;
 - 15: **end while**
 - 16: **return** Q
-

3.6. Parameter settings

In this section, all the parameters used in the proposed algorithm are discussed, and the reasonable values are recommended. In order to reasonably assign the values of these parameters, it is assumed that N solutions is obtained when the algorithm terminates (i.e., $n_0 = N$). To our experiences, it is sufficient in most cases that P_{init} , P_{semi} , and N' are set to be N . In addition, the mating pool size is set to be $N/2$, the probabilities of SBX and polynomial mutation are set to be 0.9 and $1/N$, respectively, and the maximum number of the generations is set to be 300 based on the conventions in literatures. Furthermore, the size of P_{gene} is obviously larger than or equal to N based on the settings above. Next, we mainly discuss how to assign the values of β , k_1 , k_2 , α , and k which have great effects on the performance of the proposed algorithm. To this end, we first sort the standard deviations std_1, \dots, std_n in an ascending order, and calculate the absolute differences between each two neighbor elements. Based on the mag-

nitute of the absolute difference, it is easy to find the value of β by checking the first largest value of these absolute differences. In addition, k is first set to $+\infty$ and updated by $k_1 + \gamma$ or $k_2 + \gamma$. γ is set to an integer with a small value (such as $\{1, 2, 3, 4, 5\}$) when the non-dominated solution is found for the first time in the continuous problems. Once the reference points are generated in H_d , the average distance of two neighboring reference points actually are confirmed as well. For this purpose, α is appropriately set to be the same value as this distance since a set of uniformly distributed solution in the PS are desired. In summary, the maximum value of $k_1 + k_2$ is approximately equal to N' .

3.7. Computational complexity

To analyze the computational complexity of the proposed algorithm, the symbols are unified at first, and it is assumed that N solutions are selected for an M -objective optimization problem. Furthermore, the parameters are adopted from the suggestions presented in Section 3.6. To be specific, the semi-finished solutions are obtained by solving N single objective problems with a genetic algorithm in which the size of mating pool is set to be $N/2$, and the maximum generation is set to be a constant integer $C = 300$. Consequently, it totally requires $C \times (N/2) \times (n + n)$ computations, and has a $O(Nn)$ time complexity. Furthermore, line 2, lines 3–7, lines 8–11, and line 12 execute Nn , n , Np , and p times in Algorithm 3, and the overall time complexity is $O(Nn)$. In Algorithm 4, the first and second steps require $O(N')$ and $O(N + N + N + N)$ computations, respectively, therefore, the complexity is $O(N)$. In addition, constructing reference lines in Algorithm 5 takes $O(N')$ computations. Because the exact values of k_1 and k_2 in the building model are subject to a particular problem. The worst complexity of the forward and the backward sampling is $O(N'^2)$ while the best is $O(N')$ based on the details presented in Section 3.6. Noted here that, the non-dominated comparisons consume the major computational complexity of the termination conditions, which is $O(N^4)$ under the worst case scenario and $O(N^2)$ under the best case scenario. Consequently, Algorithm 4 requires the complexity $O(N^4)$ at the worst while $O(N^2)$ at the best. Moreover, line 5 in Algorithm 1 requires $O(N^2M)$ computations. Because n is often less than N , in summary, the total computational complexity of the proposed algorithm is $O(N^4)$ under the worst case scenario and $O(N^2M)$ under the best case scenario. Especially, the worst complexity can be interpreted from two aspects:

1. EDAs generate new solutions based on the estimation models. In order to improve the accuracy of the estimation, much information is required, this lead to a higher computational complexity than other types of algorithms. In addition, the number of dependencies utilized by EDAs between variables is larger than that of other algorithms. As a consequence, the computational complexity of EDAs is worse than others, which has also been reported in literature [59].
2. The worst complexity of the proposed algorithm only caused by the non-dominated comparisons when the PS covers the entire decision space with reduced dimension. However, it is more frequently that the PS is a subspace of the decision space even with the reduced dimension, which will not give rise to such computational complexity in solving most problems. i.e., the worst complexity does not always emerge in practice.

3.8. Discussions

In the proposed algorithm, the selection of solutions exists in two phases: selecting the solutions in the decision space and the objective space, respectively. Both phases are used for maintaining diversity of solutions. To be specific, the reasons of selecting in the

objective space are the number of the available slots and the different regularities between the PS and the PF (i.e., the solutions in the decision space do not necessarily have the same images in their corresponding objective space. See Fig. 2), which is reported commonly in real-world problems and widely used benchmark test problems [60]. Furthermore, the phase of sampling the solutions in the decision space gives decision-markers' more choices for their preferences, and finds the approximated boundary of the PS (this can be seen in Fig. 1a and d). In Fig. 1a, the approximated boundary is constructed by the two circle points laying at the axes. With the help of sampling and selection of solutions, they move to the place where the Pareto-optimal solutions are interested. In addition, the semi-finished solutions are a part of the Pareto-optimal solutions to the problem, and have the sufficient statistics of the final solutions (i.e., it is reasonable that the result of dimension reduction performing on the semi-finished solutions is adapted for the final solutions of the problem, which is the fundamental of the success in the proposed algorithm). Moreover, the dimension reduction technique reduces the computational complexity of sampling the solutions at a large extent, and speeds up the exploration of the sampling process. Although the normal-boundary intersection (NBI) approach [57] is similar to the manner of the proposed sampling solutions method, the proposed algorithm has the significant differences to the NBI method. First, the NBI cannot solve problems whose PFs are concave, however, the proposed algorithm does not suffer from the restrictions of these aspects. Second, the quality of the solutions generated in NBI is highly dependent on the utopian PS whose exact form is difficult to obtain. The proposed algorithm initially employs an approximated utopian PS and update it in each generation based on the sampling solutions in the boundary reference lines. Third, the proposed algorithm is capable of solving MaOPs, while the NBI is solely for addressing MOPs.

Noted here that, the selection in the objective space is based on the solutions selected from the decision space, in order to obtain a sufficient population in the objective space, a larger value of $|P_{\text{gene}}|$ than n_o should be ensured (this can be implemented by scaling other parameters in the proposed algorithm and the details are presented in Section 3.6). Moreover, it is hopeful that each final solution is selected according to the corresponding reference point, and the number of reference points is calculated by Eq. (11)

$$H = \binom{m + p - 1}{p} \quad (11)$$

where p denotes how many divisions in each boundary edge. For this purpose, an estimate of the number of the reference points should be first given based on the amount of the available slots according to Eq. (11), then n_p is derived in further.

Although, the proposed algorithm is one kind of EDA, it is different from the existing EDAs in which explicit probabilistic models are built to direct the sampling process of the new solutions. However, in the proposed algorithm, the approximated boundary of the PS is built based on P_{semi} , which is an implicit probabilistic inference. Then the sampling model denoted by Eqs. (6) and (7) combined with the termination conditions narrows the exploration area to improve the efficacy of searching for the PS, which is also in the context of the implicit probabilistic model. Furthermore, the forward sampling and the backward sampling are the heuristic search, because the proposed algorithm does not know in advance where the Pareto-optimal solutions stand. In summary, the proposed algorithm belongs to the field of EDAs.

4. Experiments

To evaluate the performance of the proposed algorithm against peer competitors in solving MaOPs, a series of experiments are

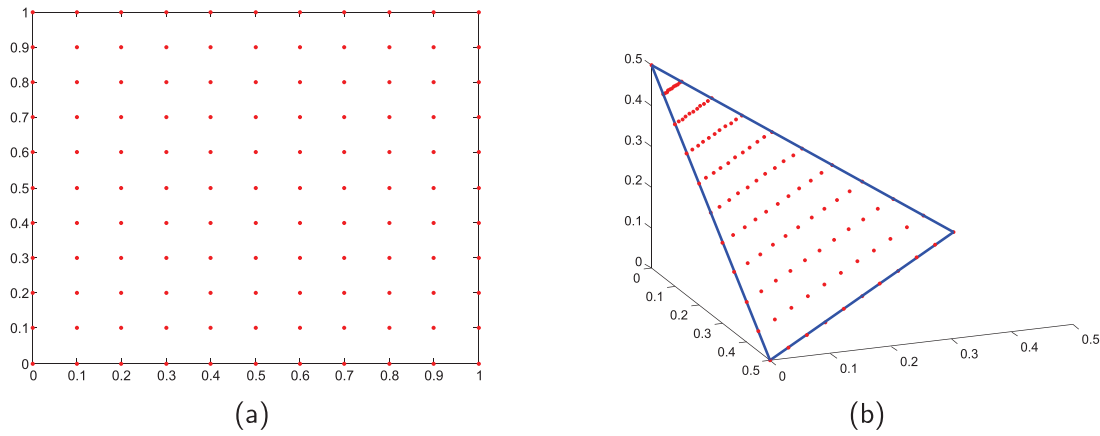


Fig. 2. An example to illustrate the different regularity between the Pareto set and the Pareto front. Specifically, a group of evenly distributed solutions (red circles) in the Pareto set are illustrated in (a). (b) plots their corresponding objectives (red circles) which do not uniformly distributed. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

performed on two widely used scalable benchmark test suits. The experimental results indicated by the considered performance metrics are then analyzed. To be specific, this section is organized as follows. First, the details of the benchmark test suits employed in these experiments are presented. This is followed by the description of the considered performance metrics. Then the parameter settings of the proposed algorithm as well as compared competitors are given. Finally, the qualitative experiments displayed through the parallel coordinate system and the quantitative experiments measured by the considered performance metrics are illustrated. Furthermore, their experimental results are analyzed.

4.1. Benchmark test problems

Based on the bottom-up approach, Deb et al. in [61] presented the DTLZ test suite. To be specific, each objective function of one given problem in this test suite takes the form as following:

$$f_i(x) = h(x_1, \dots, x_{m-1}) \cdot g(x_m, \dots, x_{k+m-1}), \quad (12)$$

where $n = k + m - 1$ is the number of the decision variables, k is set to be 5 for DTLZ1, and 10 for DTLZ2-DTLZ4, respectively. Furthermore, $h(\cdot)$ is mapped by the first $m - 1$ decision variables to measure the distribution of the solutions in the PF, and $g(\cdot)$ which is mapped by the last $k + m - 1$ decision variables determines the distance to the PF. In addition, various complicated non-linear forms of the mapping function $h(\cdot)$ raise the challenges for the algorithms in solving MaOPs. Noted here that, various versions of DTLZ have been reported in literatures. In the experiments herein, the version introduced in the original paper [61] is employed.

Working fish group (WFG) [60] is another widely used scalable benchmark test suite. Particularly, a variety of properties not manifested in DTLZ are included in WFG. To be specific, WFG is composed of nine problems, and each objective function f_i of one given problem in WFG is formulated by Eq. (13)

$$f_i(z) = D \cdot z_m + S_i \cdot h_i(z_1, \dots, z_{m-1}), \quad (13)$$

where D and S_i are used for scaling the problem. Combined with $h_i(\cdot)$, the feature of the PF in this problem, such as concave, convex, to name a few, are determined. In addition, z is mapped by applying a series of transformation functions with different challenges (e.g., multi-modality, non-separability, etc) to the decision variables $x \in \mathbb{R}^n$. Furthermore, the first $m - 1$ parameters of z (i.e., z_1, \dots, z_{m-1}) are obtained from the position-related parameters (i.e., the first k elements in x), and the last element z_m is obtained

Table 1

The features of the Pareto-optimal fronts in the test problems.

Test problem	Features of Pareto optimal front
DTLZ1	Linear, multi-modal
DTLZ2	Concave
DTLZ3	Concave, multi-modal
DTLZ4	Concave, biased
WFG1	Mixed, biased
WFG2	Convex, disconnected, multi-modal, non-separable
WFG3	Linear, degenerate, non-separable
WFG4	Concave, multi-modal
WFG5	Concave, deceptive
WFG6	Concave, non-separable
WFG7	Concave, biased
WFG8	Concave, biased, non-separable
WFG9	Concave, biased, multi-modal, deceptive, non-separable

by the distance-related parameters (the last l elements in x). Obviously, $n = k + l$.

In summary, the characteristics of the test problems employed in our experiments are summarized in Table 1.

4.2. Performance metrics

In our experiments, we consider two popular metrics, which are Inverted Generational Distance (IGD) [62] and Hypervolume (HV) [63], to measure the performance of evolutionary algorithms in solving MaOPs. IGD and HV evaluate the convergence and the diversity of the resulted PFs. It is assumed that F^* is a set of uniformly distributed solutions sampled from the PF, and P denotes the approximated solutions generated by a given algorithms, then the IGD value is estimated by Eq. (14)

$$\text{IGD} = \frac{\sum_{f^* \in F^*} \text{dist}(f^*, P)}{|F^*|}, \quad (14)$$

where $|F^*|$ is the cardinality of F^* , and $\text{dist}(f^*, P)$ denotes the shortest Euclidean distance from f^* to the elements in P . Furthermore, the smaller the IGD value, the better is the quality of P . It is noted here that, IGD is only available for DTLZ1-DTLZ4 test problems whose PF are analytical. Furthermore, HV is calculated by Eq. (15)

$$\text{HV} = \text{VOL} \left(\bigcup_{x \in P} [f_1(x), z_1] \times \dots [f_m(x), z_m] \right), \quad (15)$$

where $\text{VOL}(\cdot)$ is the Lebesgue measurement, and z_1, \dots, z_m are the elements of a reference point which is dominated by all the

Table 2

The setting of divisions and the corresponding population size with different number of the objectives.

# of Objectives (m)	Division (p)	Population size (N)
5	5	126
8	3, 2	156
10	3, 2	275
15	2, 1	135

Pareto-optimal solutions. In contrast to IGD, a larger value of HV indicates the higher quality of the obtained P . Because the selection of the reference point is a key issue in HV, we follow the suggestion in [64] to select the reference point. In addition, a Monte Carlo sampling is utilized to approximate the HV [18] due to the high computation in calculating the exact HV, and the number of the sampling points is set to be 100,000 for guaranteeing the accuracy. Furthermore, the HV values obtained in our experiments are all normalized, and the solutions which do not dominate the reference point are rejected from the calculation of HV according to the recommendation in [14,65].

4.3. Parameter settings

Because the proposed algorithm is based on the EDA for solving MaOPs, the compared algorithms include RM-MEDA, MBN-EDA, NSGA-III, HypE, GrEA, and MOEA/D. In this section, general parameter settings for all the chosen competitors as well as the proposed algorithm are given as follows, special parameter settings are provided when they are required.

- **Objectives of test instance:** The number of the objectives considered in these experiments is set to be 5, 8, 10, and 15 due to the focus of this paper on solving MaOPs.
- **Number of runs and stopping criterion:** Each test instance is performed with 20 independent runs, and a maximum generation with 500 is set as the stopping condition if required.
- **Statistical significance:** Because of the heuristic nature, statistics of experimental results is needed. We utilize the Mann-Whitney-Wilcoxon rank-sum test [66] with a 5% significance level to quantify the results.
- **Population size:** The population sizes of the proposed algorithm, NSGA-III, and MOEA/D are restricted by the parameter p in Eq. (11). For a fair comparison, the population size is consistent for all considered algorithms, and the details are shown in Table 2. In addition, the population size of GrEA is a multiple of 4 as suggested by the developers. As a result, its population size is set to be 128 for 8-objective, 256 for 10-objective, and 136 for 15-objective, which are the closest numbers to the unified population size. Because the reference points for IGD are generated only in the boundary when $p < m$, and the number of the reference points is unaffordable when $p \geq m$ with 8, 10, and 15 objectives. To this end, the two-layer approach suggested in [15] is employed.
- **Crossover and mutation:** Simulated binary crossover (SBX) [20] and polynomial mutation are employed as the recombination operator if required. Furthermore, their settings are shown in Table 3 where n denotes the number of decision variables. In addition, the distribution index of crossover in NSGA-III is set to be 30 according to [15,67,68].

4.4. Experimental results

In this subsection, quantitative and qualitative experiments are performed against the peer competitors and the experimental results are analyzed. Because the proposed algorithm is a MaOEA

Table 3

The settings for crossover and mutation.

Parameter	Value
Probability of SBX	1.0
Probability of polynomial mutation	$1/n$
Distribution index of crossover	20
Distribution index of mutation	20

based on the EDA, state-of-the-art related to the EDAs (i.e., RM-MEDA and MBN-EDA) and traditional MaOEA (i.e., NSGA-III, HypE, GrEA, and MOEA/D) are selected to be the peer competitors. Specifically, the experimental results against EDAs are illustrated in Section 4.3 while those against the traditional MaOEA are detailed in Section 4.3, respectively.

4.4.1. Comparisons to EDAs

To measure the performance of the proposed algorithm focusing on the estimation of distribution, two representative EDAs, RM-MEDA² which is a promising EDA based on regularity, and MBN-EDA which is proposed recently for addressing MaOPs, are selected. In addition, IGD and HV metrics are utilized to perform quantitative comparisons. Furthermore, the parallel coordinate system is employed for the qualitative analysis.

Because RM-MEDA is originally designed for solving MOPs, in order to perform a fair comparison in solving MaOPs, we try the number of the clusters of local PCA in RM-MEDA varying in {10, 20, 30, 50}, and the number of the generations varying in {500, 800, 1000, 2000, 3000} with the maximum iterations 50 of local PCA. Furthermore, the best performance based on the smallest mean value of IGD is selected for comparisons. Moreover, because MBN-EDA is specifically designed for solving MaOPs, the configurations suggested in its seminal work [29] are used.

Specifically, IGD results of the proposed algorithm compared against the considered EDAs (i.e., RM-MEDA and MBN-EDA) upon 5-, 8-, 10-, and 15-objective DTLZ1, DTLZ2, DTLZ3, and DTLZ4 test problems are shown in Table 4, while HV results of that with the reference points $\{1, \dots, 1\}$ for DTLZ1 and $\{2, \dots, 2\}$ for DTLZ2, DTLZ3, and DTLZ4 are shown in Table 5. In these two tables, the best mean values are highlighted in bold face, and the symbols “+”, “=”, and “−” indicate the result of the proposed algorithm is statistically better than, equal to, and worse than that of the corresponding compared algorithm, respectively. It is clearly observed from Table 4 that the proposed algorithm outperforms peer competitors upon DTLZ1, DTLZ2, DTLZ3, and DTLZ4 test problems with 10- and 15-objective. In addition, the proposed algorithm wins the best IGD results upon 5- and 8-objective DTLZ3 test problem. Although, MBN-EDA obtains the best results of DTLZ1, DTLZ2 and DTLZ4 with 8-, 5-, and 8-objective, respectively, the proposed algorithm wins the best results of these three test problems with 5-, 8-, and 5-objective, respectively. Furthermore, Table 5 clearly shows the proposed algorithm outperforms RM-MEDA and MBN-EDA upon DTLZ1, DTLZ2, DTLZ3, and DTLZ4 with all considered objective numbers. Noted in Table 5 that the symbol “=” indicates that the statistical results are not better or worse than that of MaOEA/RL, which is caused by the solutions obtained by the corresponding algorithms are all dominated by the selected reference points for estimating HV. In summary, the proposed algorithm shows the best performance indicated by both the IGD and HV metrics upon 5-, 8, 10-, and 15-objective DTLZ1, DTLZ2, DTLZ3, and DTLZ4 test problems against its peer competitors based on EDAs, which illustrates that the promising conver-

² The code is available at: <http://dces.essex.ac.uk/staff/zhang/IntrotoResearch/RegEDA.htm>.

Table 4

IGD results of MaOEDA/RL, RM-MEDA, and MBN-EDA over DTLZ1, DTLZ2, DTLZ3, and DTLZ4 test problems with 5-, 8-, 10-, and 15-objective. Each compared algorithm is independently performed 20 runs, and the best median IGD results are highlighted in bold face. The symbols “+,” “=,” and “−” denote whether the IGD results of the proposed MaOEDA/RL are statistically better than, equal to, and worse than that of the corresponding peer competitors with a significant level 5%, respectively.

	<i>m</i>	MaOEDA/RL	RM-MEDA	MBN-EDA
DTLZ1	5	1.872E−02	2.982E−02(+)	2.128E−02(=)
	8	1.219E−02	3.214E−01(+)	5.223E−03(−)
	10	2.146E−03	1.434E+00(+)	4.672E−01(+)
	15	2.516E−03	1.426E+00(+)	3.956E−01(+)
DTLZ2	5	8.147E−03	9.058E−03(+)	1.270E−03(−)
	8	9.575E−03	5.469E−03(+)	2.785E−02(+)
	10	1.758E−03	7.563E−03(+)	4.052E−03(+)
	15	3.124E−03	8.467E−03(+)	5.332E−03(+)
DTLZ3	5	1.419E−02	4.218E−01(+)	3.157E−01(+)
	8	4.522E−02	4.915E−02(+)	6.527E−02(+)
	10	1.759E−03	5.649E−02(+)	2.087E−02(+)
	15	3.123E−03	6.325E−02(+)	2.118E−02(+)
DTLZ4	5	3.570E−02	8.491E−02(+)	9.340E−02(+)
	8	6.787E−03	7.577E−02(+)	6.431E−03(−)
	10	2.193E−03	2.522E−03(+)	3.132E−03(+)
	15	3.382E−03	4.638E−03(+)	3.786E−03(+)

Table 5

HV results of MaOEDA/RL, RM-MEDA, and MBN-EDA over DTLZ1, DTLZ2, DTLZ3, and DTLZ4 test problems with 5-, 8-, 10-, and 15-objective. Each compared algorithm is independently performed 20 runs, and the best median HV results are highlighted in bold face. The symbols “+,” “=,” and “−” denote whether the HV results of the proposed MaOEDA/RL are statistically better than, equal to, and worse than that of the corresponding peer competitors with a significant level 5%, respectively.

	<i>m</i>	MaOEDA/RL	RM-MEDA	MBN-EDA
DTLZ1	5	0.9821	0.9323(+)	0.9810(=)
	8	0.9910	0.0000(=)	0.8321(+)
	10	0.9940	0.0000(=)	0.0000(=)
	15	0.9877	0.0000(=)	0.0000(=)
DTLZ2	5	0.9999	0.9982(+)	0.9341(+)
	8	0.9999	0.4555(+)	0.9838(+)
	10	0.9998	0.3289(+)	0.8717(+)
	15	0.9993	0.4716(+)	0.9101(+)
DTLZ3	5	0.9993	0.7438(+)	0.9876(+)
	8	0.9790	0.0000(=)	0.0000(=)
	10	0.9999	0.0000(=)	0.0000(=)
	15	0.9991	0.0000(=)	0.0000(=)
DTLZ4	5	0.9319	0.8893(+)	0.9217(+)
	8	0.9811	0.9554(+)	0.9710(+)
	10	0.9997	0.9993(+)	0.9977(+)
	15	0.9999	0.9992(+)	0.9992(+)

gence and diversity are simultaneously achieved by the proposed algorithm.

Actually, it is no surprise that a superior performance against RM-MEDA and MBN-EDA, which are based on the EDAs, is reached by the proposed algorithm. Firstly, although the parameters of RM-MEDA have been extensively tuned upon these tested MaOPs, there are fundamental differences between MaOPs and MOPs, and RM-MEDA is originally designed for only solving MOPs. To be specific, RM-MEDA employs the local PCA [54] to estimate the PF, and then uniformly samples new solutions from the estimated PF for maintaining good diversity. Because a large proportion of DR solutions exists in the current population in solving MaOPs (this phenomenon has been discussed in Section 1), the local PCA could not properly estimate the whole PF, thus the diversity cannot be guaranteed. Fig. 3 illustrates the deficiencies of the local PCA in

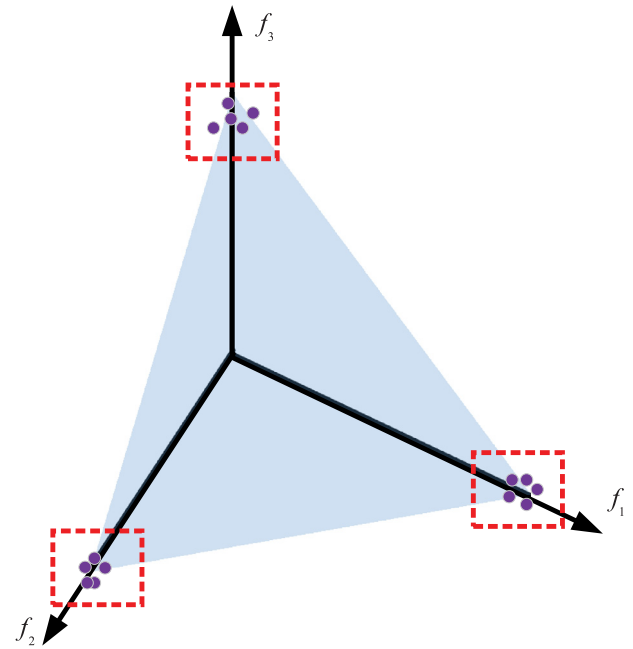


Fig. 3. An example of a tri-objective optimization problem to illustrate the deficiencies of the local PCA. Specifically, the solid circles denote the dominance resistant solutions, and the dotted rectangles denote the estimated models by the local PCA employing the clustering technique.

solving MaOPs. In Fig. 3, the shaded region denotes the whole PF of a tri-objective optimization problem, the solid circles represent the DR solutions, and the dotted rectangles refer to the estimated models based on these DR solutions by the local PCA. Because the local PCA expectedly span the whole PF by clustering solutions, the estimated PFs (denoted by the dotted rectangles) would cover only the corners of the whole PF under this scenario. Consequently, the sampled solutions would only surround these corners, but not the whole PF, thus the diversity is lost. Secondly, MBN-EDA employs the solutions sampled in advance from the problems to be solved for building models, and then solutions with promising convergence and diversity are generated from these built models. While it is difficult to sample solutions with good diversity both in the decision space and the objective space for MBN-EDA building models, which is caused by the different images from the decision space to the objective space of the problems to be solved, such as the DLTZ4, and most WFG instances. Therefore, these built models are not significant [29], which leads to the sampled solutions to be not necessarily promising. Thirdly, the proposed algorithm employs the DR solutions to construct the utopian PF, and then manually generates uniformly distributed reference lines for sampling solutions with the proposed forward sampling and backward sampling approaches. These strategies collectively mitigate the adverse impacts of RM-MEDA and MBN-EDA, thus promote the promising performance to be achieved by the proposed algorithm.

In addition, obtained solutions of DTLZ4 with 10- and 15-objective are plotted in Fig. 4 by the parallel coordinate system to visually compare the diversity of the compared EDAs. For convenience, all the values have been normalized. Based on the priori of DTLZ4, it is clearly shown that the diversity of RM-MEDA is better than that of MBN-EDA for DTLZ4 with 10-objective, which is contrast to that in DTLZ4 with 15-objective. However, the diversity of MaOEDA/RL is the best in DTLZ4 with both 10- and 15-objective (seen from Fig. 4a and b, respectively).

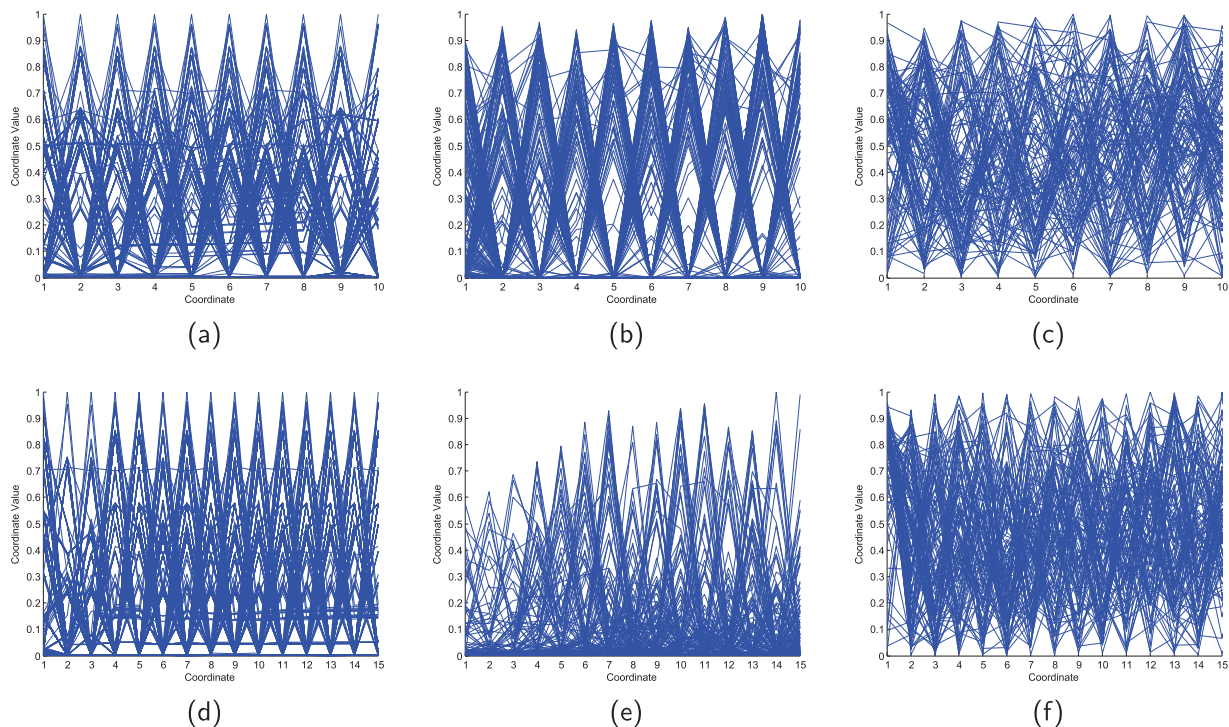


Fig. 4. Solutions of MaOEDA/RL, RM-MEDA, and MBN-EDA in DTLZ4 with 10- and 15-objective are illustrated by the parallel coordinate system. Especially, a, b, and c denote the solutions of MaOEDA/RL, RM-MEDA, and MBN-EDA over DTLZ4 test instance with 10-objective. d, e, and f are the solutions of MaOEDA/RL, RM-MEDA, and MBN-EDA over DTLZ4 test instance with 15-objective.

4.4.2. Comparisons to MaOEAs

Four state-of-the-art MaOEAs (NSGA-III,³ HypE,⁴ GrEA,⁵ and MOEA/D⁶) are selected for comparing the performance of the proposed algorithm on addressing all the problems in WFG test suite. Especially, HV is considered to measure their performance. In addition, extra parameter settings are presented as follows.

- *Archive size*: It is set to be the same as N if required.
- *Neighborhood size in MOEA/D*: The neighborhood size is set to be 20.
- *Number of divisions in GrEA*: According to the suggestion given by its developers in [17], the numbers of divisions varying in {7, 8, 9, 10, 11} are tested individually on each test instance, then the best value is selected based on the best mean result of HV, which are listed in Table 6.
- *Reference points for HV*: The reference points for calculating the HV of WFG1-WFG9 are set to $\{3.0, \dots, 2m + 1\}$ based on the suggestion from [64].

The results measured by HV upon WFG1-WFG9 test problems with 5-, 8-, 10- and 15-objective are shown in Table 7. In this table, the best mean values are highlighted in bold face, and the symbols “+,” “=,” and “-” indicate the result of the proposed algorithm is statistically better than, equal to, and worse than that of the associated compared algorithm, respectively, and the last row summarizes how many times the proposed algorithm better than, equal to, and worse than that of the corresponding peer

Table 6

Grid division settings in GrEA.

Problem	# of Objective	Grid division
WFG1	5, 8, 10, 15	9, 7, 11, 11
WFG2	5, 8, 10, 15	11, 11, 10, 11
WFG3	5, 8, 10, 15	9, 11, 7, 7
WFG4	5, 8, 10, 15	10, 7, 9, 11
WFG5	5, 8, 10, 15	11, 11, 11, 9
WFG6	5, 8, 10, 15	10, 9, 11, 9
WFG7	5, 8, 10, 15	7, 7, 11, 9
WFG8	5, 8, 10, 15	11, 9, 10, 11
WFG9	5, 8, 10, 15	11, 11, 11, 9

competitor. Specifically, it is clearly observed that the best performances upon all considered 10-objective WFG test problems are achieved by the proposed algorithm except for WFG2 whose winner is MOEA/D. Furthermore, the proposed algorithm is better than others upon 8-objective test problems of WFG1, WFG2, WFG4, WFG6, and WFG7, while inferior to NSGA-III upon WFG3 and WFG8, GrEA upon WFG5, and MOEA/D upon WFG9. In addition, the proposed algorithm wins the best HV scores on 15-objective WFG1, WFG3, WFG5, WFG8, and WFG9, and the winner of other test problems is NSGA-III. For the 5-objective test problems, the proposed algorithm outperforms peer competitors upon WFG2, WFG3, WFG7, and WFG8, while underperforms MOEA/D upon WFG1, WFG5, and WFG9, and GrEA upon WFG4. Although, NSGA-III achieves the best mean value of WFG8 with 5-objective, the proposed algorithm shows the equal statistical performance. In summary, the proposed algorithm wins 112 times out of 144 experiments, and significantly outperforms peer MaOEA competitors on the WFG test suit measured by HV.

The experimental results from Table 7 show a superior performance has been achieved by the proposed algorithm through

³ The source code is available at: <http://web.ntnu.edu.tw/~tcchiang/publications/nsga3cpp/nsga3cpp.htm>.

⁴ The source code is available at: <http://www.tik.ee.ethz.ch/sop/download/supplementary/hype/>.

⁵ The source code is available at: <http://www.brunel.ac.uk/~cspgmm1/Codes/GrEA.rar>.

⁶ The source code is available at: <http://dces.essex.ac.uk/staff/zhang/webofmoead.htm>.

Table 7

HV results of MaOEDA/RL, NSGA-III, GrEA, HypE, and MOEA/D on WFG1-WFG9 test problems with 5-, 8-, 10-, and 15-objective. Each compared algorithm is independently performed 20 runs, and the best median HV results are highlighted in bold face. The symbols “+,” “=,” and “−” denote whether the HV results of the proposed MaOEDA/RL are statistically better than, equal to, and worse than that of the corresponding peer competitors with a significant level 5%, respectively.

	<i>m</i>	MaOEDA/RL	NSGA-III	GrEA	HypE	MOEA/D
WFG1	5	8.1472E−01	9.3211E−01(−)	6.4322E−01(+)	2.7855E−01(+)	9.5750E−01(−)
	8	9.6496E−01	9.1340E−01(+)	9.0759E−01(+)	5.4694E−01(+)	9.0582E−01(+)
	10	9.6953E−01	8.4096E−01(+)	3.1026E−01(+)	4.0088E−02(+)	9.4458E−01(+)
	15	9.4533E−01	8.3636E−01(+)	4.7829E−01(+)	8.4675E−03(+)	8.9462E−01(+)
WFG2	5	9.5022E−01	8.2346E−01(+)	9.7132E−02(+)	6.9483E−01(+)	8.4072E−01(+)
	8	7.6552E−01	7.9520E−01(=)	6.4631E−01(+)	7.0936E−01(+)	5.6782E−01(+)
	10	7.8260E−01	9.7558E−01(−)	9.6238E−01(−)	7.4963E−01(+)	9.9454E−01(−)
	15	7.1982E−01	7.9038E−01(−)	9.5234E−01(−)	5.9507E−01(+)	9.6913E−01(−)
WFG3	5	9.9613E−01	9.6190E−01(=)	9.1334E−01(+)	8.0007E−01(+)	8.6869E−01(+)
	8	7.7917E−01	7.8528E−01(−)	6.5408E−01(+)	6.0198E−01(+)	7.4815E−01(+)
	10	6.4951E−01	6.2528E−01(=)	6.0187E−01(+)	1.2390E−01(+)	6.2325E−01(+)
	15	6.4327E−01	5.6635E−01(+)	5.7974E−01(+)	5.8788E−02(+)	6.2365E−01(=)
WFG4	5	9.0005E−01	9.4205E−01(−)	9.5613E−01(−)	7.8025E−01(+)	9.4479E−01(−)
	8	9.6455E−01	9.0272E−01(+)	8.2119E−01(+)	8.1126E−02(+)	7.4469E−01(+)
	10	8.6752E−01	7.9006E−01(+)	6.4474E−01(+)	4.3599E−02(+)	5.9538E−01(+)
	15	7.1577E−01	7.8331E−01(+)	5.4342E−01(+)	2.7506E−02(+)	3.5799E−01(+)
WFG5	5	9.2939E−01	8.1763E−01(+)	8.1158E−01(+)	5.3283E−01(+)	9.3900E−01(−)
	8	7.7571E−01	7.9483E−01(−)	8.7594E−01(−)	6.2248E−01(+)	6.4432E−01(+)
	10	8.8391E−01	8.2598E−01(+)	8.3914E−01(+)	2.5366E−01(+)	5.3683E−01(+)
	15	8.0213E−01	8.0120E−01(=)	7.1181E−01(+)	1.5036E−02(+)	3.7874E−01(+)
WFG6	5	9.7838E−01	9.7975E−01(=)	9.0488E−01(+)	6.0284E−01(+)	8.4431E−01(+)
	8	9.2885E−01	8.0101E−01(+)	8.5516E−02(+)	4.3887E−01(+)	7.1122E−01(+)
	10	8.5995E−01	8.5037E−01(+)	8.5933E−01(=)	4.3294E−02(+)	6.3284E−01(+)
	15	8.3141E−01	8.5269E−01(−)	7.7053E−01(+)	2.5226E−02(+)	4.5396E−01(+)
WFG7	5	9.8798E−01	9.6309E−01(+)	8.8517E−01(+)	6.7973E−01(+)	9.8712E−01(=)
	8	9.0472E−01	9.0372E−01(=)	8.5944E−01(+)	5.7672E−01(+)	8.9092E−01(+)
	10	8.8499E−01	8.4639E−01(+)	8.1858E−01(+)	4.5930E−02(+)	5.8109E−01(+)
	15	8.3576E−01	8.6573E−01(−)	7.9172E−01(+)	3.0985E−02(+)	3.1785E−01(+)
WFG8	5	8.8651E−01	8.1755E−01(+)	7.1269E−01(+)	4.2431E−02(+)	8.1815E−01(+)
	8	8.0033E−01	8.3470E−01(−)	8.3138E−01(−)	6.0471E−02(+)	8.0336E−01(=)
	10	8.3222E−01	7.2493E−01(+)	7.7672E−01(+)	3.3228E−02(+)	4.2684E−01(+)
	15	8.2378E−01	7.1342E−01(+)	7.4198E−01(+)	2.4417E−02(+)	2.7479E−01(+)
WFG9	5	9.5163E−01	9.8305E−01(−)	9.2033E−01(+)	7.3786E−01(+)	9.8406E−01(−)
	8	8.8187E−01	8.5552E−01(+)	7.0110E−01(+)	3.2601E−02(+)	8.9464E−01(−)
	10	8.2617E−01	7.1271E−01(+)	7.8000E−01(+)	2.7937E−02(+)	4.8844E−01(+)
	15	7.4124E−01	6.3940E−01(+)	7.0578E−01(+)	9.9595E−03(+)	3.0149E−01(+)
+ / = / −			20/6/10	30/1/5	36/0/0	26/3/7

the comparisons against the peer MaOEA competitors. Specifically, the strengths of the proposed algorithm in this experiment can be explained as follows. Firstly, these peer MaOEA competitors employ the genetic operators for performing exploration search and exploitation search to expectedly guarantee the promising performance. However, because of the deficiencies of these employed genetic operators in solving MaOPs, which has also been extensively investigated in [19], these MaOEAs are not capable to achieve the best performance. Secondly, although the proposed algorithm is based on EDAs and also possesses the heuristic characteristics, the proposed algorithm obtains a promising performance with a higher probability than the traditional MaOEAs, which can be viewed as the advantages of EDAs. For example in the proposed algorithm, once the utopian PF has been constructed, we would know the approximate position of PF, which has guaranteed the global search. Combined with the proposed forward sampling and backward sampling approaches, which can be viewed as the local search, promising solutions are found with higher accuracy.

5. Conclusion

In this paper, a reference line-based Estimation of Distribution Algorithm for efficiently solving many-objective optimization problems has been proposed. In the proposed algorithm, the utopian Pareto Set is constructed from the dominance resistant solutions, which are commonly regarded as harmful to the traditional many-

objective evolutionary algorithms, for generating uniformly distributed reference points in the decision space. Then, new solutions are sampled by the proposed sampling strategies from the reference lines, which lie at a subspace of the decision space to find the PS, and the subspace is estimated by dimension reduction technique. Next, Pareto-optimal solutions with promising diversity in the objective space are selected based on these new sampled solutions. Specifically, sampling solutions from the reference lines in the subspace promotes the convergence, while selecting solutions further from the sampled solutions maintains the diversity. These two strategies collectively guarantee the promising convergence and diversity to be achieved by the proposed algorithm. For quantifying the performance of the proposed algorithm, a series of well-designed experiments are performed against peer competitors, which includes two state-of-the-art EDA competitors and four many-objective evolutionary algorithm competitors, upon a couple of widely used scalable benchmark test suits with 5-, 8-, 10-, and 15-objective. Experimental results measured by the chosen performance metrics indicate that the proposed algorithm shows superiority in tackling many-objective optimization problems. In our future study, more attentions will be investigated in how to employ dependencies between the decision variables and the objective variables to speed up the estimation in the proposed algorithm.

References

- [1] M. Farina, P. Amato, On the optimal solution definition for many-criteria optimization problems, in: *Proceedings of the NAFIPS-FLINT International Conference*, 2002, pp. 233–238.
- [2] B. Li, J. Li, K. Tang, X. Yao, An improved two archive algorithm for many-objective optimization, in: *Proceedings of 2014 IEEE Congress on Evolutionary Computation*, CEC, IEEE, 2014, pp. 2869–2876.
- [3] R.J. Lygoe, M. Cary, P.J. Fleming, A real-world application of a many-objective optimisation complexity reduction process, in: *Evolutionary Multi-Criterion Optimization*, Springer, 2013, pp. 641–655.
- [4] O. Chikumbo, E. Goodman, K. Deb, Approximating a multi-dimensional pareto front for a land use management problem: a modified moea with an epigenetic silencing metaphor, in: *Proceedings of 2012 IEEE Congress on Evolutionary Computation*, CEC, IEEE, 2012, pp. 1–9.
- [5] K. Miettinen, *Nonlinear Multiobjective Optimization*, 12, Springer Science & Business Media, 2012.
- [6] K. Deb, A. Pratap, S. Agarwal, T. Meyarivan, A fast and elitist multiobjective genetic algorithm: nsga-ii, *IEEE Trans. Evol. Comput.* 6 (2) (2002) 182–197.
- [7] E. Zitzler, M. Laumanns, L. Thiele, E. Zitzler, L. Thiele, L. Thiele, *Spea2: Improving the Strength Pareto Evolutionary Algorithm*, 2001.
- [8] K. Deb, *Multi-objective Optimization Using Evolutionary Algorithms*, 16, John Wiley & Sons, 2001.
- [9] C.A.C. Coello, D.A. Van Veldhuizen, G.B. Lamont, *Evolutionary Algorithms for Solving Multi-objective Problems*, 242, Springer, 2002.
- [10] D.W. Corne, J.D. Knowles, Techniques for highly multiobjective optimisation: some nondominated points are better than others, in: *Proceedings of the 9th Annual Conference on Genetic and Evolutionary Computation*, ACM, 2007, pp. 773–780.
- [11] C.M. Fonseca, P.J. Fleming, Multiobjective optimization and multiple constraint handling with evolutionary algorithms. I. A unified formulation, *IEEE Trans. Syst. Man Cybern. Part A: Syst. Hum.* 28 (1) (1998) 26–37.
- [12] R.C. Purshouse, P.J. Fleming, On the evolutionary optimization of many conflicting objectives, *IEEE Trans. Evol. Comput.* 11 (6) (2007) 770–784.
- [13] J. Knowles, D. Corne, Quantifying the effects of objective space dimension in evolutionary multiobjective optimization, in: *Evolutionary Multi-criterion Optimization*, Springer, 2007, pp. 757–771.
- [14] T. Wagner, N. Beume, B. Naujoks, Pareto-, aggregation-, and indicator-based methods in many-objective optimization, in: *Evolutionary Multi-criterion Optimization*, Springer, 2007, pp. 742–756.
- [15] K. Deb, H. Jain, An evolutionary many-objective optimization algorithm using reference-point-based nondominated sorting approach, part I: solving problems with box constraints, *IEEE Trans. Evol. Comput.* 18 (4) (2014) 577–601.
- [16] Q. Zhang, H. Li, Moea/d: a multiobjective evolutionary algorithm based on decomposition, *IEEE Trans. Evol. Comput.* 11 (6) (2007) 712–731.
- [17] S. Yang, M. Li, X. Liu, J. Zheng, A grid-based evolutionary algorithm for many-objective optimization, *IEEE Trans. Evol. Comput.* 17 (5) (2013) 721–736.
- [18] J. Bader, E. Zitzler, Hype: an algorithm for fast hypervolume-based many-objective optimization, *Evol. Comput.* 19 (1) (2011) 45–76.
- [19] K. Deb, A. Sinha, S. Kukkonen, Multi-objective test problems, linkages, and evolutionary methodologies, in: *Proceedings of the 8th Annual Conference on Genetic and Evolutionary Computation*, ACM, 2006, pp. 1141–1148.
- [20] K. Deb, R.B. Agrawal, Simulated binary crossover for continuous search space, *Complex Syst.* 9 (2) (1995) 115–148.
- [21] K. Deb, M. Goyal, A combined genetic adaptive search (geneas) for engineering design, *Comput. Sci. Inform.* 26 (1996) 30–45.
- [22] P. Larranaga, J.A. Lozano, *Estimation of Distribution Algorithms: A New Tool for Evolutionary Computation*, 2, Springer Science & Business Media, 2002.
- [23] J.A. Lozano, Towards a New Evolutionary Computation: Advances on Estimation of Distribution Algorithms, 192, Springer Science & Business Media, 2006.
- [24] M. Pelikan, K. Sastry, E. Cantú-Paz, Scalable Optimization via Probabilistic Modeling: From Algorithms to Applications, 33, Springer, 2007.
- [25] P. Larrañaga, H. Karshenas, C. Bielza, R. Santana, A review on probabilistic graphical models in evolutionary computation, *J. Heuristics* 18 (5) (2012) 795–819.
- [26] P.A. Bosman, D. Thierens, Multi-objective optimization with diversity preserving mixture-based iterated density estimation evolutionary algorithms, *Int. J. Approx. Reason.* 31 (3) (2002) 259–289.
- [27] M. Pelikan, K. Sastry, D.E. Goldberg, Multiobjective estimation of distribution algorithms, in: *Scalable Optimization via Probabilistic Modeling*, Springer, 2006, pp. 223–248.
- [28] Q. Zhang, A. Zhou, Y. Jin, Rm-meda: a regularity model-based multiobjective estimation of distribution algorithm, *IEEE Trans. Evol. Comput.* 12 (1) (2008) 41–63.
- [29] H. Karshenas, R. Santana, C. Bielza, P. Larranaga, Multiobjective estimation of distribution algorithm based on joint modeling of objectives and variables, *IEEE Trans. Evol. Comput.* 18 (4) (2014) 519–542.
- [30] M. Costa, E. Minisci, Moped: a multi-objective parzen-based estimation of distribution algorithm for continuous problems, in: *Evolutionary Multi-Criterion Optimization*, Springer, 2003, pp. 282–294.
- [31] M. Pelikan, K. Sastry, D.E. Goldberg, Multiobjective hboa, clustering, and scalability, in: *Proceedings of the 7th Annual Conference on Genetic and Evolutionary Computation*, ACM, 2005, pp. 663–670.
- [32] K. Sastry, D.E. Goldberg, M. Pelikan, Limits of scalability of multiobjective estimation of distribution algorithms, in: *Proceedings of the 2005 IEEE Congress on Evolutionary Computation*, 2005, 3, IEEE, 2005, pp. 2217–2224.
- [33] N. Khan, D.E. Goldberg, M. Pelikan, Multi-objective bayesian optimization algorithm, *Urbana* 51 (2002) 61801.
- [34] J. Schwarz, J. Ocenasek, Multiobjective Bayesian optimization algorithm for combinatorial problems: theory and practice, *Neural Netw. World* 11 (5) (2001) 423–442.
- [35] E. Zitzler, K. Deb, L. Thiele, Comparison of multiobjective evolutionary algorithms: empirical results, *Evol. Comput.* 8 (2) (2000) 173–195.
- [36] M. Laumanns, J. Ocenasek, Bayesian optimization algorithms for multi-objective optimization, in: *Parallel Problem Solving from Nature PPSN VII*, Springer, 2002, pp. 298–307.
- [37] M. Laumanns, L. Thiele, K. Deb, E. Zitzler, Combining convergence and diversity in evolutionary multiobjective optimization, *Evol. Comput.* 10 (3) (2002) 263–282.
- [38] X. Zhong, W. Li, A decision-tree-based multi-objective estimation of distribution algorithm, in: *Proceedings of 2007 International Conference on Computational Intelligence and Security*, IEEE, 2007, pp. 114–118.
- [39] V. Cherkassky, F.M. Mulier, *Learning from Data: Concepts, Theory, and Methods*, John Wiley & Sons, 2007.
- [40] T. Hastie, R. Tibshirani, J. Friedman, J. Franklin, The elements of statistical learning: data mining, inference and prediction, *Math. Intell.* 27 (2) (2005) 83–85.
- [41] A. Zhou, Q. Zhang, Y. Jin, E. Tsang, T. Okabe, A model-based evolutionary algorithm for bi-objective optimization, in: *Proceedings of the 2005 IEEE Congress on Evolutionary Computation*, 3, IEEE, 2005, pp. 2568–2575.
- [42] A. Zhou, Y. Jin, Q. Zhang, B. Sendhoff, E. Tsang, Combining model-based and genetics-based offspring generation for multi-objective optimization using a convergence criterion, in: *Proceedings of IEEE Congress on Evolutionary Computation*, CEC 2006, IEEE, 2006, pp. 892–899.
- [43] H. Benson, S. Sayin, Optimization over the efficient set: four special cases, *J. Optim. Theory Appl.* 80 (1) (1994) 3–18.
- [44] P. Thach, H. Konno, D. Yokota, Dual approach to minimization on the set of pareto-optimal solutions, *J. Optim. Theory Appl.* 88 (3) (1996) 689–707.
- [45] R. Horst, N.V. Thoai, Utility function programs and optimization over the efficient set in multiple-objective decision making, *J. Optim. Theory Appl.* 92 (3) (1997) 605–631.
- [46] G. Eichfelder, *Adaptive Scalarization Methods in Multiobjective Optimization*, Springer, 2008.
- [47] A. Zhou, Q. Zhang, Y. Jin, Approximating the set of pareto-optimal solutions in both the decision and objective spaces by an estimation of distribution algorithm, *IEEE Trans. Evol. Comput.* 13 (5) (2009) 1167–1189.
- [48] Z. He, G.G. Yen, Visualization and performance metric in many-objective optimization, *IEEE Trans. Evol. Comput.* 20 (3) (2016) 386–402.
- [49] J. Cheng, G.G. Yen, G. Zhang, A many-objective evolutionary algorithm with enhanced mating and environmental selections, *IEEE Trans. Evol. Comput.* 19 (4) (2015) 592–605.
- [50] G.G. Yen, Z. He, Performance metric ensemble for multiobjective evolutionary algorithms, *IEEE Trans. Evol. Comput.* 18 (1) (2014) 131–144.
- [51] Z. He, G.G. Yen, J. Zhang, Fuzzy-based pareto optimality for many-objective evolutionary algorithms, *IEEE Trans. Evol. Comput.* 18 (2) (2014) 269–285.
- [52] Z. He, G.G. Yen, Many-objective evolutionary algorithm: objective space reduction and diversity improvement, *IEEE Trans. Evol. Comput.* 20 (1) (2016) 145–160.
- [53] Z. He, G.G. Yen, Many-objective evolutionary algorithms based on coordinated selection strategy, *IEEE Trans. Evol. Comput.* (2017) 1. (Submitted for publication)
- [54] N. Kambhatla, T.K. Leen, Dimension reduction by local principal component analysis, *Neural Comput.* 9 (7) (1997) 1493–1516.
- [55] J. Lemmer, L. Kanal, Propagating uncertainty in Bayesian networks by probabilistic logic sampling, *Uncertain. Artif. Intell.* 2 5 (2014) 149.
- [56] P. Fleming, Computer aided control systems using a multiobjective optimization approach, in: *Proceedings of IEEE Control Conference*, 1985, pp. 174–179.
- [57] I. Das, J.E. Dennis, Normal-boundary intersection: a new method for generating the pareto surface in nonlinear multicriteria optimization problems, *SIAM J. Optim.* 8 (3) (1998) 631–657.
- [58] Y. Yuan, H. Xu, B. Wang, An experimental investigation of variation operators in reference-point based many-objective optimization, in: *Proceedings of the 2015 on Genetic and Evolutionary Computation Conference*, ACM, 2015, pp. 775–782.
- [59] E. Bengoetxea, *Inexact Graph Matching Using Estimation of Distribution Algorithms*, 2, Ecole Nationale Supérieure des Télécommunications, Paris, 2002, p. 4.
- [60] S. Huband, P. Hingston, L. Barone, L. While, A review of multiobjective test problems and a scalable test problem toolkit, *IEEE Trans. Evol. Comput.* 10 (5) (2006) 477–506.
- [61] K. Deb, L. Thiele, M. Laumanns, E. Zitzler, *Scalable Test Problems for Evolutionary Multiobjective Optimization*, Springer, 2005.
- [62] P.A. Bosman, D. Thierens, The balance between proximity and diversity in multiobjective evolutionary algorithms, *IEEE Trans. Evol. Comput.* 7 (2) (2003) 174–188.
- [63] E. Zitzler, L. Thiele, Multiobjective evolutionary algorithms: a comparative case study and the strength pareto approach, *IEEE Trans. Evol. Comput.* 3 (4) (1999) 257–271.
- [64] K. Li, K. Deb, Q. Zhang, S. Kwong, An evolutionary many-objective optimization algorithm based on dominance and decomposition, *IEEE Trans. Evol. Comput.* 19 (5) (2015) 694–716.

- [65] X. Zou, Y. Chen, M. Liu, L. Kang, A new evolutionary algorithm for solving many-objective optimization problems, *IEEE Trans. Syst. Man Cybern. Part B: Cybern.* 38 (5) (2008) 1402–1412.
- [66] R.G. Steel, D. JH Dickey, et al., *Principles and Procedures of Statistics a Biometrical Approach*, WCB/McGraw-Hill, 1997.
- [67] M. Asafuddoula, T. Ray, R. Sarker, A decomposition-based evolutionary algorithm for many objective optimization, *IEEE Trans Evol. Comput.* 19 (3) (2015) 445–460.
- [68] Y. Yuan, H. Xu, B. Wang, Evolutionary many-objective optimization using ensemble fitness ranking, in: *Proceedings of the 2014 Conference on Genetic and Evolutionary Computation*, ACM, 2014, pp. 669–676.