Information Sciences 512 (2020) 278–294

Contents lists available at ScienceDirect

Information Sciences

journal homepage: www.elsevier.com/locate/ins

Objective reduction for visualising many-objective solution sets

Liangli Zhen[a,b], Miqing Li[b], Dezhong Peng[a,d,e], Xin Yao[b,c,*]

[a] Machine Intelligence Laboratory, College of Computer Science, Sichuan University, Chengdu 610065, China
[b] CERCIA, School of Computer Science, University of Birmingham, B15 2TT Birmingham, UK
[c] Shenzhen Key Laboratory of Computational Intelligence, University Key Laboratory of Evolving Intelligent Systems of Guangdong Province, Department of Computer Science and Engineering, Southern University of Science and Technology, Shenzhen 518055, China
[d] Chengdu Sobey Digital Technology Co., Ltd, Chengdu 610041, China
[e] Shenzhen Cyberspace Laboratory, Shenzhen 518055, China

ARTICLE INFO

Article history:
Received 16 October 2018
Revised 2 April 2019
Accepted 6 April 2019
Available online 8 April 2019

Keywords:
Objective reduction
Visualisation
Many-objective optimisation
Evolutionary algorithms

ABSTRACT

Visualising a solution set is of high importance in many-objective optimisation. It can help algorithm designers understand the performance of search algorithms and decision makers select their preferred solution(s). In this paper, an objective reduction-based visualisation method (ORV) is proposed to view many-objective solution sets. ORV attempts to map a solution set from a high-dimensional objective space into a low-dimensional space while preserving the distribution and the Pareto dominance relation between solutions in the set. Specifically, ORV sequentially decomposes objective vectors which can be linearly represented by their positively correlated objective vectors until the expected number of preserved objective vectors is reached. ORV formulates the objective reduction as a solvable convex problem. Extensive experiments on both synthetic and real-world problems have verified the effectiveness of the proposed method.

1. Introduction

Many-objective optimisation problems (MaOPs) which involve more than three (conflicting) objectives to be optimised exist in many industrial and engineering applications [20]. Over the last decade, many-objective optimisation has attracted increasing interest in the evolutionary computation community [16,20]. One inherent challenge in many-objective optimisation is that we cannot directly view solutions in such a four- or higher-dimensional objective space [28]. This brings difficulties for algorithm design, performance assessment, decision making, etc [21,41].

In many-objective optimisation, the goal of visualising solutions is to enable researchers and practitioners to understand the given problem (e.g., the shape of its Pareto front) and the characteristics of the solutions (e.g., their Pareto dominance relation) [28]. In other words, a good visualisation can reveal the underlying structure of the problem, helps researchers and practitioners make a proper decision and illustrates the relation between solutions and also between objectives.

There are many multi-dimensional data visualisation methods, some of which have been applied to show many-objective solution sets, such as parallel coordinates [15], radar chart [18,36], and heatmaps [29]. These methods directly plot the objective values of the solutions in a two-dimensional plane without any sophisticated transformations and can be easily

* Corresponding author at: CERCIA, School of Computer Science, University of Birmingham, B15 2TT Birmingham, UK.
E-mail address: x.yao@cs.bham.ac.uk (X. Yao).

https://doi.org/10.1016/j.ins.2019.04.014
0020-0255/

extended to cases with higher dimensionality and to more solutions. However, in order to represent the conflicts between objectives, they require the objectives of interest to be positioned adjacent to each other. Also, the contour information of a given approximate Pareto-front is unavailable.

In recent years, dimension reduction-based visualisation techniques, which transform the input vectors into a lower-dimensional space, have attracted much interest from the evolutionary computation community. These techniques can help data analysts obtain new observations and insights through viewing the mapped data in the reduced space. To preserve the distribution of solutions, many dimension reduction methods have been applied, such as principal component analysis (PCA) [17], Sammon mapping [30], and neuroscale [26]. They do not care much about preserving the dominance relation between solutions and the conflicts between the objectives.

There exist several dimension reduction-based visualisation methods that consider the dominance relation between solutions. For instance, Köppen and Yoshida [19] used two different strategies to map the dominated and non-dominated solutions, respectively. He and Yen [13] proposed to map high-dimensional solutions into a 2D polar coordinate plot so that a large number of solutions can be viewed in a plane and the dominance relation between them can be approximately persevered. Tušar and Filipič [38] used the prosection approach to visualise four-dimensional solution sets, and it can preserve the dominance relation between solutions despite working only for problems with four objectives.

Two desirable properties in the many-objective solution set visualisation are preserving (1) the distribution of solutions and (2) the dominance relation between solutions. The distribution of solutions can be implied by the diversity and density of solutions and the contour information about the Pareto front. The dominance relation between solutions is a fundamental criterion of convergence, and considering it can help algorithm designer and decision maker select the optimal solutions.

In this paper, we propose an objective reduction-based visualisation method (ORV) for showing many-objective solution set. In ORV, we sequentially decompose objective vectors which can be linearly represented by their positively correlated objective vectors until the expected number of preserved objective vectors is reached. The main contributions and novelty of this work include the following:

- An objective reduction-based method is proposed for viewing many-objective solution sets. The method is able to preserve the distribution of solutions and the dominance relation between solutions as far as possible.
- The strategy of sequentially decomposing objective vectors makes the whole objective reduction process visible, which helps researchers and participators understand the conflicts between objectives through viewing the objective vector decomposition in each iteration.
- It formulates the underlying objective vector representation as a convex optimisation problem, which can be solved efficiently. This formulation makes ORV capable of handling large-scale many-objective solution set.
- Using ORV, algorithm designers can observe the evolutionary behaviour of their algorithms; decision makers can read the distribution of solutions, which help them in both quality evaluation and preference articulation processes.

The rest of this paper is organised as follows. Section 2 reviews the terminology and related work. Section 3 is devoted to the description of the new objective reduction method for displaying many-objective solution sets. Section 4 provides experimental results to illustrate the effectiveness of the proposed method. Finally, Section 5 concludes the paper.

2. Related work

Without loss of generality, a multi-objective optimisation problem (MOP) can be formulated as a minimisation problem and defined as follows:

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

where $\Omega \subseteq \mathbb{R}^n$ is the decision space, $\mathbf{x} = (x_1, x_2, \ldots, x_n)^T$ is a candidate solution, and $F : \Omega \to \mathbb{R}^M$ consists of M (conflicting) objective functions. The multi-objective optimisation problem with more than three conflicting objectives, i.e., $M > 3$, is referred to as a many-objective optimisation problem (MaOP).

Let $\mathbf{a}$ and $\mathbf{b}$ be two feasible solutions for the above MOP, $\mathbf{a}$ is better than $\mathbf{b}$ if the following conditions hold:

$$\forall i \ f_i(\mathbf{a}) \leq f_i(\mathbf{b}) \ \text{ and } \ \exists j \ f_j(\mathbf{a}) < f_j(\mathbf{b}), \tag{2}$$

where $i, j \in \{1, 2, \ldots, M\}$. We can also say that $\mathbf{a}$ dominates $\mathbf{b}$, and denote it as $\mathbf{a} \prec \mathbf{b}$. A solution that is not dominated by any other solutions is Pareto optimal. The set of Pareto optimal solutions in the decision space is denoted as the Pareto set, and the corresponding outcome in this objective space is denoted as the Pareto front.

During the optimisation process or after obtaining the solution set, one of the important issues for algorithm designers and decision makers is the visualisation of the solutions in the objective space. In the following, we review several popular methods for visualisation of many-objective solution sets, which can be generally divided into directly plotted methods and dimension reduction-based methods.

2.1. Directly plotted visualisation methods

This class of methods show the original objective values of solutions without any sophisticated mapping and can be easily scaled to more objectives. The scatter plot matrix, parallel coordinate plot, radar chart, and heatmaps belong to this class. In the following, we will introduce them in detail.

2.1.1. Scatter plot matrix

Scatter plot matrix [37] considers two of objectives at one time, then draws solutions in the selected objective space. If this is done for all the combinations of M objectives, we will have $\binom{M}{2}$ subplots, which can be ordered as a scatter plot matrix. However, the number of subplots will increase dramatically as more objectives are involved in the optimisation problem, which makes it difficult to view solutions of a problem with more than five objectives.

2.1.2. Parallel coordinates

Parallel coordinates plot [15], also called value paths [28], is a general visualisation tool for high-dimensional data. To show a set of M-dimensional solutions, the parallel coordinates plot displays them on a two-dimensional plane with M vertically and equally spaced parallel axes. One solution is represented by a polygonal line with vertices on these parallel axes, and the position of the vertex on the ith axis corresponds to the value of the solution on the ith objective. Parallel coordinates plot is an effective visualisation method as it is easy to interpret and scales well to a large number of objectives (simply by adding more axes for new objectives). However, it makes interpretation and comparison of solutions difficult when the number of polygonal lines becomes large [46]. In addition, this way of presentation only shows $M-1$ relationships of a total of $\binom{M}{2}$ conflicting relationships existing among M objectives. A large portion of conflicts between objectives cannot be displayed.

2.1.3. Radar chart

Radar chart [18], also known as the spider-web chart, is originally provided for comparing multiple quantitative variables. It is easy to see which variables score high or low within a dataset, and making it ideal for displaying performance of the candidates. In a radar chart, each variable is provided an axis that starts from the origin. All axes are arranged radially, with equal relative angles between each other. For a data point, each variable value is plotted along its individual axis as a vertex, and these vertices are connected together to form a polygon. However, in some optimisation problems, there are hundreds of Pareto-optimal solutions, which involve hundreds of polygons in one radar chart and make them hard to read, confusing and cluttered. In addition, having more objectives creates more axes and makes the polygons more prone to overlap.

2.1.4. Heatmaps

Heatmaps are frequently used to view multivariate datasets and have recently been used to show multi-objective solution sets [29,40]. A heatmap represents objectives as columns, solutions by rows, and relative objective values in different colours. To show the values of different objectives with the same colour map, one should normalise values in each objective to a similar range, otherwise, the heatmap will be dominated by a small portion of colours. Also, the order of solutions is arbitrarily placed in the rows of a heatmap. Finally, perceiving the distribution of many-objective solutions is difficult in a heatmap.

2.2. Dimension reduction-based visualisation methods

Since it is difficult for human beings to perceive the solutions in a high-dimensional space, a natural way to understand a solution set is mapping the solution set into a low-dimensional space and viewing the solutions in this low-dimensional space. In the following, we briefly review several popular dimension reduction methods that have been used for visualisation of many-objective solution sets.

2.2.1. Principal component analysis

Principal component analysis (PCA) [17] is the most widely used linear dimension reduction method. It finds a new set of coordinates, known as principal components, so that projection of the data set onto these coordinates captures the maximum variance among all linear projections. These coordinates can be easily obtained since they can be the eigenvectors with the largest eigenvalues of the covariance matrix of the dataset. PCA has been used for visualisation of the multi-objective solution set of nurse scheduling problem in [27]. One of the problems of the PCA method is that it may misleadingly project two remote solutions to two adjacent points in the low-dimensional space. The dominance relation between solutions is also ignored in the visualisation of many-objective solution sets.

2.2.2. Sammon mapping and NeuroScale

Sammon mapping [30] aims to approximately preserve the latent structure of the dataset under the mapping from the high-dimensional space to the low-dimensional space by minimising the Sammon's stress function. Specifically, it enforces the distances between any two data points in the reduced low-dimensional space to be as close as possible to the corresponding distances in the original high-dimensional space and uses the steepest gradient descent procedure to search the

Table 1
Summary of the visualisation methods with regard to the desirable properties. Property 1 denotes the preservation of dominance relation, Property 2 denotes the preservation of distribution, Property 3 stands for it being scalable to the number of objectives, and Property 4 denotes it being able to handle large data sets. The symbol of "✓" indicates that the property holds, the "×" indicates that the property does not hold, and the "≈" indicates that the property does not necessarily hold despite the method being designed for it.

Method	Property 1	Property 2	Property 3	Property 4
Scatter plot matrix [37]	×	≈	×	≈
Parallel coordinates [15]	✓	≈	✓	×
Radar chart [18]	✓	×	✓	×
Heatmaps [29]	×	×	✓	×
PCA [17]	×	≈	✓	✓
Sammon mapping [30]	×	≈	✓	✓
Two-stage mapping [19]	≈	≈	✓	✓
Polar coordinate [13]	≈	≈	✓	✓
Prosection [38]	✓	≈	×	✓
The proposed method	≈	≈	✓	✓

low-dimensional mapped data points which achieve a minimum of the Sammon's stress function. It has been used to view solution sets of multi-objective optimisation problems in [39]. Neuroscale [26] follows the same goal as Sammon mapping to preserve the neighbourhood relationships in data by using neural network techniques and is based upon a radial basis function (RBF) architecture. However, it utilises an RBF neural network to predict the coordinates of the data points in the transformed low-dimensional space. It has been conducted to map the solutions on the Pareto front in the high-dimensional objective space into the two- or three-dimensional space for visualisation [8].

2.2.3. Two-stage mapping

Two-stage mapping [19] aims to preserve the Pareto dominance and the distribution relation between solutions as much as possible. It uses two different strategies to map the dominated and non-dominated solutions. In the first stage, all non-dominated solutions are mapped onto a quarter-circle in the first quadrant, where the radius reflects the average norm of these solutions. The two-stage mapping method finds a permutation for the ordering of the solutions by optimising a multi-objective problem so that both the relations of Pareto dominance and distribution among these solutions are preserved as much as possible. Then, it maps the solutions on the circle in the order given by the obtained permutation with new distances proportional to the original distances for all pairs of adjacent solutions. In the second stage, each dominated solution is mapped to the minimal vector of all non-dominated vectors that dominate it. While the two-stage mapping method splits the dominated and non-dominated solutions, it is rather complex as a multi-objective optimisation problem is involved in the mapping, and also this method fails to show the conflicts between objectives.

2.2.4. Polar coordinate method

This method [13] maps solutions from a high-dimensional space into a 2-D polar coordinate system. Each solution is assigned a radial coordinate value reflecting its convergence performance and an angular coordinate value revealing its diversity performance. The radial coordinate value is determined by the shape of the approximate front and the original objective values of the solution. A solution with a small radial coordinate implies it has a good convergence to the PF. Angular coordinates of all solutions show distribution of the solutions among each subregion and the crowdedness in each subregion of the high-dimensional objective space. The solutions with poor distribution and spread of the approximate PF would share the same angular coordinate values. This method can preserve the dominance relation. However, it fails to provide the distribution information of PF inside each subregion and relations between different subregions [12].

2.2.5. Prosection

The prosection method [38] tries to project the vectors in a section to visualise four-dimensional solution set. For a given solution set, it selects an origin and two objectives to construct a prosection plane. Then, it defines a section by choosing the angle and section width on the prosection plane. Finally, it projects all the solutions inside the previously defined section into a three-dimensional space by mapping these two-dimensional vectors into a line, and ignores other solutions outside the section. This method is able to correctly show the dominance relation between solutions but is not easily scalable to more than four objectives.

The main characteristics of the above visualisation methods are summarised in Table 1, and the readers can find the plot results of these methods in references [13,28] and [38]. As can be seen, the methods of direct plotting have difficulties to handle a large-size, high-dimensional solution set. The methods of PCA, Sammon mapping and neuroscale are unable to preserve the dominance relation between solutions. The two-stage mapping method splits the solutions into non-dominated and dominated sets, but the information loss is noticeable in terms of their distribution maintenance. The polar coordinate method can preserve the dominance relation between solutions, but it is not good at preserving the distribution of solutions.

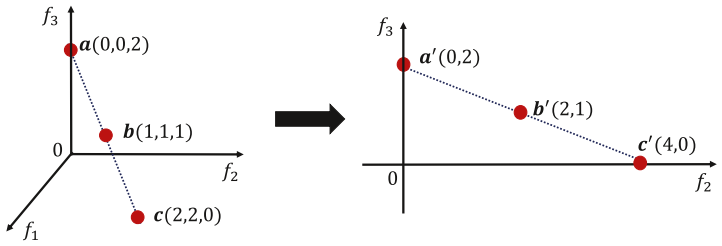


Fig. 1. A toy example of the basic idea of our method. Three non-dominated solutions in the three dimensional objective space are transformed into two dimensional objective space by representing the objective vector on f_1 with the objective vector on f_2.

While the prosection method tries to preserve both the dominance and distribution relations between solutions, it works only for a problem with less than five objectives.

There are also some other dimension reduction-based methods designed specially for many-objective optimisation (rather than for visualisation) [2,25,32]. For instance, by selecting the smallest set of objectives without changing the Pareto set of the original optimisation problem [9], Brockhoff and Zitzler [3] presented both exact and heuristic algorithms to reduce the number of objectives, while preserving as much as possible of the dominance structure. In [34], Singh et al. proposed to find out non-correlated objectives by analyzing the relation between the solutions from the search results of the corners of a Pareto front. In [33], Saxena et al. applied principal component analysis and maximum variance unfolding to select conflicting objectives according to the correlation information. In [42], Wang and Yao used nonlinear correlation information entropy [43] to measure the correlations among objectives, and proposed an online objective reduction-based method for evolutionary many-objective optimisation. In [44], Yuan et al. formulated the objective reduction as a tri-objective optimisation problem, where two of the objectives are based on the dominance structure and the third objective is based on the correlation between objectives of the considered solution set.

Note that the dimension reduction methods for optimisation and those for visualisation are of different purposes and focuses. The dimension reduction-based methods designed for many-objective optimisation focus mainly on preserving the conflicts between objectives; this may change the distribution of solutions and the dominance relation between solutions during the dimension reduction process. In contrast, the dimension reduction methods for visualisation of many-objective solution sets often attempt to preserve the dominance and the distribution relations between solutions, aiming to provide the user the original landscape of the solution sets.

3. Our proposed method

In this section, we present an objective reduction method to visualise a many-objective solution set. The basic idea of the proposed method is to decompose some objective vectors of a solution set into the remaining objective vectors in succession, thus forming a new set with lower dimensionality. In each iteration, we represent each objective vector by other objective vectors which are positively correlated with it. We then remove the objective vector that has the minimal representation error by decomposing it into other objective vectors. The reason for doing so is that we try to preserve the distribution and the dominance relations between solutions as much as possible, and the conflicts between objectives as well.

Fig. 1 gives a toy example of our proposed method. As can be seen from the figure, in the left plot there are three non-dominated solutions in the objective space $\mathbf{a}(0, 0, 2)$, $\mathbf{b}(1, 1, 1)$, and $\mathbf{c}(2, 2, 0)$. The first objective vector, which consists of the first objective values of these three solutions, is $\mathbf{y}_1 = (0, 1, 2)^T$. The second and third objective vectors are $\mathbf{y}_2 = (0, 1, 2)^T$ and $\mathbf{y}_3 = (2, 1, 0)^T$. It is clear that $\mathbf{y}_1$ is positively linearly dependent with $\mathbf{y}_2$, and $\mathbf{y}_3$ is negatively linearly dependent with $\mathbf{y}_2$. Following the basic idea of our proposed method, we represent $\mathbf{y}_1$ with $\mathbf{y}_2$ by transforming the information from the first objective to the second objective, and obtain the new objective vectors $\mathbf{y}'_2 = (1 + 1)\mathbf{y}_2 = (0, 2, 4)^T$ and $\mathbf{y}'_3 = \mathbf{y}_3 = (2, 1, 0)^T$ in the reduced objective space for these three solutions, as shown at the right plot of Fig. 1. The corresponding solutions in the reduced space are $\mathbf{a}'(0, 2)$, $\mathbf{b}'(2, 1)$, and $\mathbf{c}'(4, 0)$. They are still non-dominated with each other, and the distribution of them remains unchanged in terms of Manhattan distance. The conflict between the second objective and the third objective is preserved as well.

In the following, we introduce the main procedure of our proposed method and show some of its observations.

Given a set of N solutions (denoted as $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N$) for an optimisation problem with M objectives, we use their corresponding values on the ith objective to construct an objective vector $\mathbf{y}_i = (f_i(\mathbf{x}_1), \ldots, f_i(\mathbf{x}_N))^T \in \mathbf{R}^N$, where $i = 1, 2, \ldots, M$.

We sequentially decompose some objective vectors. In each iteration, we represent one objective vector by a few other objective vectors which are positively correlated with the target objective vector. Specifically, we first obtain the correlations between all the objective vectors, and find the positively correlated objective vectors for each objective vector. Here, we treat each objective as a variable, and the objective values of solutions on it as the corresponding scalar observations. Then, the Spearman's rank correlation coefficient of any two objective vectors, a measure of their rank correlation, is calculated. It equals the Pearson correlation between the rank values of the two variables. Technically, the two objective vectors $\mathbf{y}_i$ and $\mathbf{y}_j$ are converted to rank vectors $r_{\mathbf{y}_i}$ and $r_{\mathbf{y}_j}$, respectively, according to the ranks of the elements at the corresponding

objective vector. For example, for an given objective vector $\mathbf{y}_1 = (6, 9, 2)^T$, the corresponding rank vector is $r_{\mathbf{y}_1} = (2, 3, 1)^T$. The Spearman's rank correlation coefficient is defined as

$$\rho(\mathbf{y}_i, \mathbf{y}_j) = \frac{cov(r_{\mathbf{y}_i}, r_{\mathbf{y}_j})}{\sigma_{r_{\mathbf{y}_i}} \sigma_{r_{\mathbf{y}_j}}}, \tag{3}$$

where $\sigma_{r_{\mathbf{y}_i}}$ and $\sigma_{r_{\mathbf{y}_j}}$ are the standard deviation of $r_{\mathbf{y}_i}$ and $r_{\mathbf{y}_j}$, respectively, and $cov(r_{\mathbf{y}_i}, r_{\mathbf{y}_j})$ is the covariance of $r_{\mathbf{y}_i}$ and $r_{\mathbf{y}_j}$.

We can rewrite the correlations among all the objective vectors in a matrix form as:

$$\mathbf{R} = \begin{bmatrix} \rho(\mathbf{y}_1, \mathbf{y}_1) & \dots & \rho(\mathbf{y}_1, \mathbf{y}_M) \\ \dots & \dots & \dots \\ \rho(\mathbf{y}_M, \mathbf{y}_1) & \dots & \rho(\mathbf{y}_M, \mathbf{y}_M) \end{bmatrix}. \tag{4}$$

From the definition in (3), we have that the correlation coefficients between $\mathbf{y}_i$ and its positively correlated objective vectors are greater than zero, where $i = 1, \dots, M$. We thus can represent each objective vector with other objective vectors which have a positive correlation coefficients with the considered objective vector. Please note that to select positively correlated vectors for the target objective vector, some other criteria can also be adopted. In this paper we use the Spearman's rank correlation since it is a nonparametric measure of correlation and has a low computational complexity. Formally, for each objective vector $\mathbf{y}_i$, we select the objective vectors in the set $\mathcal{D}_i = \{\mathbf{y}_j | R_{ij} > 0, j \neq i\}$, where $j \in \{1, \dots, M\}$, to construct a dictionary matrix $\mathbf{D}_i$, and represent $\mathbf{y}_i$ by minimising the following optimisation problem:

$$\min_{\mathbf{c}_i} \ \|\mathbf{y}_i - \mathbf{D}_i \mathbf{c}_i\|_2 + \lambda \|\mathbf{c}_i\|_1, \quad \text{s.t.} \ \ \mathbf{c}_i \in \mathbb{R}^k_{\geq 0}, \tag{5}$$

where $\mathbf{c}_i$ is the representation coefficient vector corresponding to $\mathbf{y}_i$, the first term is the representation error, the second term is a regularisation term which enforces using the fewest number of columns in $\mathbf{D}_i$ to represent $\mathbf{y}_i$, λ is a tradeoff parameter (typically is set as 0.001), k is the number of columns in $\mathbf{D}_i$, and the constraint ensures that all the representation coefficients are non-negative. In view of this being a convex optimisation problem, we can solve it efficiently using existing convex optimisation methods [1]. Here, we use the solver from Matlab: CVX, a package for specifying and solving convex problems [10,11].

Next, we determine which objective vector to be removed according to the representation error. We consider the objective vector that has the minimal representation error for the minimal information loss. Technically, we remove the objective vector with the index

$$\alpha = \arg\min_i \ \frac{\|\mathbf{y}_i - \mathbf{D}_i \mathbf{c}_i^*\|_2}{\|\mathbf{y}_i\|_2}, \tag{6}$$

where $\mathbf{c}_i^*$ is the optimal representation coefficient vector of the ith objective vector via (5), and the effect of different scales of objective vectors is reduced by dividing by $\|\mathbf{y}_i\|_2$.

Assume that the positively correlated vectors in $\mathcal{D}_\alpha$ are $\mathbf{y}_{\sigma_1}, \dots, \mathbf{y}_{\sigma_k}$, we update them using

$$\mathbf{y}_{\sigma_j}^{(t)} = (1 + c_{ij}^*) \mathbf{y}_{\sigma_j}^{(t-1)}, j \in \{1, \dots, k\}, \tag{7}$$

where $t \in \{1, 2, \dots, M - m\}$ denotes the number of the current iteration, m is the expected dimensionality of the observed objective space, and $\mathbf{y}_{\sigma_j}^{(t)}$ denotes the value of $\mathbf{y}_{\sigma_j}$ at the tth iteration and $\mathbf{y}_{\sigma_j}^{(0)}$ is the input $\mathbf{y}_{\sigma_j}$.

In this manner, we can sequentially remove more objective vectors until the expected number of preserved objective vectors is reached.

Finally, we have the new objective function with m components as

$$\mathcal{F}(\mathbf{x}) = (f_{\tau_1}^{(M-m)}(\mathbf{x}), f_{\tau_2}^{(M-m)}(\mathbf{x}), \dots, f_{\tau_m}^{(M-m)}(\mathbf{x})), \tag{8}$$

where m is the number of preserved objective vectors, the index set of preserved objective vectors is $\{\tau_1, \tau_2, \dots, \tau_m\} \subset \{1, 2, \dots, M\}$, and $f_{\tau_1}^{(M-m)}(\mathbf{x})$, $f_{\tau_2}^{(M-m)}(\mathbf{x}), \dots, f_{\tau_m}^{(M-m)}(\mathbf{x})$ are objective values of $f_{\tau_1}(\mathbf{x}), f_{\tau_2}(\mathbf{x}), \dots, f_{\tau_m}(\mathbf{x})$ after $(M - m)$ iterations of objective reduction, respectively. The objective reduction process of our proposed method is summarised in Algorithm 1, from which we can see that the proposed method iteratively decomposes objective vectors to achieve objective reduction. In each iteration, it firstly constructs a dictionary matrix for each objective objective by selecting the positively correlated objective vectors to the target vector. Then, it represents each objective vector by the columns in the dictionary matrix and finds out the objective α which has the minimal reconstruction error at step 10 of Algorithm 1. At last, it updates the preserved objective vectors, which belong to the columns of the dictionary matrix $\mathbf{D}_\alpha$, and removes the objective α. The computational complexity of the proposed method is $O((M - m)(M + N)tNM^3)$, where N is the number of the solutions and t denotes the total number of iterations for the optimisation algorithm in solving step 7 in Algorithm 1.

The proposed method has the following two properties when the linear representation error is equal to zero in the objective reduction process.

Proposition 1. *For any $\mathbf{x}_i, \mathbf{x}_j \in \Omega$, it holds that $\mathbf{x}_i \prec \mathbf{x}_j$ in the reduced objective space if and only if $\mathbf{x}_i \prec \mathbf{x}_j$ in the original objective space.*

Proof. *Proof of Sufficiency*: For any $\mathbf{x}_i, \mathbf{x}_j \in \Omega$, if $\mathbf{x}_i \prec \mathbf{x}_j$ in the original objective space, and since the preserved objectives are included in the set consists of all original objectives, there exists two cases:

Algorithm 1 The objective reduction procedure of our proposed method for a many-objective solution set.

Input: The objective vectors of the solution sets $\mathbf{y}_i = (f_i(\mathbf{x}_1), \ldots, f_i(\mathbf{x}_N))^T \in \mathbb{R}^N, i = 1, 2, \ldots, M$, and the expected dimensionality of the observed objective space, m.

Initialise $\Upsilon^{(0)} = \{1, \ldots, M\}$, and $\mathbf{y}_i^{(0)} = \mathbf{y}_i$, **for all** $i \in \Upsilon^{(0)}$.
$R_{ij} \leftarrow \rho(\mathbf{y}_i, \mathbf{y}_j)$, **for all** $i, j \in \Upsilon^{t-1}$.
for each $t \in \{1, \ldots, M-m\}$ **do**
  for each $i \in \Upsilon^{(t-1)}$ **do**
    $\mathcal{S}_i \leftarrow \{j | R_{ij} > 0; j \neq i; j \in \Upsilon^{(t-1)}\}$.
    $\mathbf{D}_i \leftarrow \{\mathbf{y}_j^{(t-1)} | j \in \mathcal{S}_i\}$. /*Use the elements in the set to construct a matrix.*/
    $\mathbf{c}_i^* \leftarrow \arg\min_{\mathbf{c}_i \in \mathbb{R}^k_{\geq 0}} \|\mathbf{y}_i^{(t-1)} - \mathbf{D}_i\mathbf{c}_i\|_2 + \lambda\|\mathbf{c}_i\|_1$.
    $e_i \leftarrow \frac{\|\mathbf{y}_i - \mathbf{D}_i\mathbf{c}_i^*\|_2}{\|\mathbf{y}_i\|_2}$.
  end for
  $\alpha \leftarrow \arg\min_i \{e_i | i \in \Upsilon^{(t-1)}\}$.
  $\mathbf{y}_{\sigma_j}^{(t)} \leftarrow (1 + c_{\alpha\sigma_j}^*)\mathbf{y}_{\sigma_j}^{(t-1)}$, **for all** $\sigma_j \in \mathcal{S}_\alpha$.
  $\Upsilon^{(t)} \leftarrow \Upsilon^{(t-1)} - \{\alpha\}$.
end for
$\mathbf{Y} \leftarrow \{\mathbf{y}_l^{(M-m)} | l \in \Upsilon^{(M-m)}\}$. /*Use the elements in the set to construct a matrix.*/

Output: The representations of the solutions in the reduced m-dimensional objective space $\mathbf{Y}$, where each row represents the result for one solution.

(a) $\mathbf{x}_i \prec \mathbf{x}_j$ in the reduced objective space, which directly completes the proof.
(b) the objective values of $\mathbf{x}_i$ and $\mathbf{x}_j$ are equal on the preserved objectives, i.e.,

$$\begin{aligned} f_{\tau_1}^{(0)}(\mathbf{x}_i) &= f_{\tau_1}^{(0)}(\mathbf{x}_j), \\ f_{\tau_2}^{(0)}(\mathbf{x}_i) &= f_{\tau_2}^{(0)}(\mathbf{x}_j), \\ &\cdots \\ f_{\tau_m}^{(0)}(\mathbf{x}_i) &= f_{\tau_m}^{(0)}(\mathbf{x}_j), \end{aligned} \tag{9}$$

and $\exists \mu \in \{1, \ldots, M\} - \{\tau_1, \tau_2, \ldots, \tau_m\}$ satisfies

$$f_\mu^{(0)}(\mathbf{x}_i) < f_\mu^{(0)}(\mathbf{x}_j). \tag{10}$$

Considering the decomposition of the objective f_μ, we have

$$\begin{aligned} f_\mu^{(t-1)}(\mathbf{x}_i) &= c_{\varphi 1} f_{\sigma_1}^{(t-1)}(\mathbf{x}_i) + \cdots + c_{\varphi k} f_{\sigma_k}^{(t-1)}(\mathbf{x}_i), \\ f_\mu^{(t-1)}(\mathbf{x}_j) &= c_{\varphi 1} f_{\sigma_1}^{(t-1)}(\mathbf{x}_j) + \cdots + c_{\varphi k} f_{\sigma_k}^{(t-1)}(\mathbf{x}_j), \end{aligned} \tag{11}$$

where $t \in \{1, \ldots, M-m\}$, and $\{\sigma_1, \ldots, \sigma_k\} \subseteq \{\tau_1, \ldots, \tau_m\}$.

Combining (9) and (11), we have $f_\mu^{(M-m-1)}(\mathbf{x}_i) = f_\mu^{(M-m-1)}(\mathbf{x}_j)$, which contradicts with the result in (10). This means the case b) does not exist.

Proof of Necessity: For any $\mathbf{x}_i, \mathbf{x}_j \in \Omega$, since $\mathbf{x}_i \prec \mathbf{x}_j$ in the reduced objective space, we obtain that

$$\begin{aligned} f_{\tau_1}^{(M-m)}(\mathbf{x}_i) &\leq f_{\tau_1}^{(M-m)}(\mathbf{x}_j), \\ f_{\tau_2}^{(M-m)}(\mathbf{x}_i) &\leq f_{\tau_2}^{(M-m)}(\mathbf{x}_j), \\ &\cdots \\ f_{\tau_m}^{(M-m)}(\mathbf{x}_i) &\leq f_{\tau_m}^{(M-m)}(\mathbf{x}_j), \end{aligned} \tag{12}$$

and $\exists \tau_\eta \in \{\tau_1, \tau_2, \ldots, \tau_m\}$ satisfies

$$f_{\tau_\eta}^{(M-m)}(\mathbf{x}_i) < f_{\tau_\eta}^{(M-m)}(\mathbf{x}_j). \tag{13}$$

Consider any objective function f_φ in the objective space, if $\varphi \in \{\tau_1, \ldots, \tau_m\}$, then $f_\varphi^{(0)}(\mathbf{x}_i) \leq f_\varphi^{(0)}(\mathbf{x}_j)$, else if $\varphi \notin \{\tau_1, \tau_2, \ldots, \tau_m\}$,

We assume that f_φ is the last decomposed objective, and obtain that

$$\begin{aligned} f_\varphi^{(t-1)}(\mathbf{x}_i) &= c_{\varphi 1} f_{\sigma_1}^{(t-1)}(\mathbf{x}_i) + \cdots + c_{\varphi k} f_{\sigma_k}^{(t-1)}(\mathbf{x}_i), \\ f_\varphi^{(t-1)}(\mathbf{x}_j) &= c_{\varphi 1} f_{\sigma_1}^{(t-1)}(\mathbf{x}_j) + \cdots + c_{\varphi k} f_{\sigma_k}^{(t-1)}(\mathbf{x}_j), \end{aligned} \tag{14}$$

where $t = M - m$, and $\{\sigma_1, \sigma_2, \ldots, \sigma_k\} \subseteq \{\tau_1, \tau_2, \ldots, \tau_m\}$.

Moreover, based on the model in (5), we have

$$c_{\varphi 1}, c_{\varphi 2}, \ldots, c_{\varphi k} \in \mathbb{R}_{\geq 0} \tag{15}$$

Combining the results in (15), (12) and (14), we have $f_{\varphi}^{(M-m-1)}(\mathbf{x}_i) \leq f_{\varphi}^{(M-m-1)}(\mathbf{x}_j)$.

Following the same manner, we can obtain that objective values for $\mathbf{x}_j$ on the second last decomposed objective is not larger than that for $\mathbf{x}_j$. Likewise, we obtain the same conclusion for the first decomposed objective. Based on above results and (13), we have

$$\begin{aligned} f_1^{(0)}(\mathbf{x}_i) &\leq f_1^{(0)}(\mathbf{x}_j), \\ f_2^{(0)}(\mathbf{x}_i) &\leq f_2^{(0)}(\mathbf{x}_j), \\ &\cdots \\ f_M^{(0)}(\mathbf{x}_i) &\leq f_M^{(0)}(\mathbf{x}_j), \end{aligned} \tag{16}$$

and $\exists \varsigma \in \{1, 2, \ldots, M\}$ satisfies

$$f_{\varsigma}^{(0)}(\mathbf{x}_i) < f_{\varsigma}^{(0)}(\mathbf{x}_j), \tag{17}$$

i.e., $\mathbf{x}_i \prec \mathbf{x}_j$ in the original objective space.

This completes the proof. □

Proposition 2. *For any $\mathbf{x}_i, \mathbf{x}_j \in \Omega$, it holds that $\mathbf{x}_i$ and $\mathbf{x}_j$ are non-dominated to each other in the reduced objective space if and only if $\mathbf{x}_i$ and $\mathbf{x}_j$ are non-dominated to each other in the original objective space.*

Proof. *Proof of Sufficiency*: Assume that the sufficiency does not hold, i.e., $\exists \mathbf{x}_i, \mathbf{x}_j \in \Omega$, when $\mathbf{x}_i$ and $\mathbf{x}_j$ are non-dominated to each other in the original objective space, but $\mathbf{x}_i \prec \mathbf{x}_j$ or $\mathbf{x}_j \prec \mathbf{x}_i$ in the reduced objective space. It is clear that above conclusion contradicts with that in Proposition 1.

Proof of Necessity: Assume that the necessity does not hold, i.e., $\exists \mathbf{x}_i, \mathbf{x}_j \in \Omega$, when $\mathbf{x}_i$ and $\mathbf{x}_j$ are non-dominated to each other in the reduced objective space, but $\mathbf{x}_i \prec \mathbf{x}_j$ or $\mathbf{x}_j \prec \mathbf{x}_i$ in the original objective space. It is clear that above conclusion also contradicts with that in Proposition 1.

This completes the proof. □

Proposition 3. *For any $\mathbf{x}_i, \mathbf{x}_j \in \Omega$, the Manhattan distance between $\mathbf{x}_i$ and $\mathbf{x}_j$ in the reduced objective space equals the distance in the original objective space if the Spearman's rank coefficients between the removed objective vector and the vectors in its representation dictionary are equal to one.*

Proof. For any $\mathbf{x}_i, \mathbf{x}_j \in \Omega$, suppose the first decomposed objective is f_{φ}, and represented by the objectives $\sigma_1, \ldots, \sigma_k$ with corresponding coefficients as $c_{\varphi 1}, \ldots, c_{\varphi k}$, we have

$$\begin{aligned} f_{\varphi}^{(0)}(\mathbf{x}_i) &= c_{\varphi 1} f_{\sigma_1}^{(0)}(\mathbf{x}_i) + \cdots + c_{\varphi k} f_{\sigma_k}^{(0)}(\mathbf{x}_i), \\ f_{\varphi}^{(0)}(\mathbf{x}_j) &= c_{\varphi 1} f_{\sigma_1}^{(0)}(\mathbf{x}_j) + \cdots + c_{\varphi k} f_{\sigma_k}^{(0)}(\mathbf{x}_j), \end{aligned} \tag{18}$$

where $\Gamma = \{\sigma_1, \ldots, \sigma_k\} \subseteq \{1, \ldots, \varphi - 1, \varphi + 1, \ldots, M\}$.

If the Spearman's rank coefficients between $\mathbf{y}_{\varphi}$ and $\mathbf{y}_{\sigma_1}, \ldots, \mathbf{y}_{\sigma_k}$ are equal to one, we can obtain that

$$\left| f_{\varphi}^{(0)}(\mathbf{x}_i) - f_{\varphi}^{(0)}(\mathbf{x}_j) \right| = c_{\varphi 1} \left| f_{\sigma_1}^{(0)}(\mathbf{x}_i) - f_{\sigma_1}^{(0)}(\mathbf{x}_j) \right| + \cdots + c_{\varphi k} \left| f_{\sigma_k}^{(0)}(\mathbf{x}_i) - f_{\sigma_k}^{(0)}(\mathbf{x}_j) \right|. \tag{19}$$

Thus, the Manhattan distance between $\mathbf{x}_i$ and $\mathbf{x}_j$ in the original objective space

$$\begin{aligned} \|F(\mathbf{x}_i) - F(\mathbf{x}_j)\|_1 &= \sum_{\kappa \in \{1, \ldots, M\}} \left| f_{\kappa}^{(0)}(\mathbf{x}_i) - f_{\kappa}^{(0)}(\mathbf{x}_j) \right| \\ &= \sum_{\iota \in \{1, \ldots, M\} - \{\varphi\}} \left| f_{\iota}^{(0)}(\mathbf{x}_i) - f_{\iota}^{(0)}(\mathbf{x}_j) \right| + \left| f_{\varphi}^{(0)}(\mathbf{x}_i) - f_{\varphi}^{(0)}(\mathbf{x}_j) \right| \\ &= \sum_{\alpha \in \Psi} \left| f_{\alpha}^{(0)}(\mathbf{x}_i) - f_{\alpha}^{(0)}(\mathbf{x}_j) \right| \\ &\quad + \sum_{\beta \in \Gamma} (1 + c_{\varphi \beta}) \left| f_{\beta}^{(0)}(\mathbf{x}_i) - f_{\beta}^{(0)}(\mathbf{x}_j) \right|, \end{aligned} \tag{20}$$

where $\Psi = \{1, \ldots, M\} - \{\varphi\} - \Gamma$.

For the Manhattan distance between $\mathbf{x}_i$ and $\mathbf{x}_j$ in the new objective space

$$
\begin{aligned}
&\|F^{(1)}(\mathbf{x}_i) - F^{(1)}(\mathbf{x}_j)\|_1 = \sum_{\tau_\eta \in \Psi \cup \Gamma} \left| f^{(1)}_{\tau_\eta}(\mathbf{x}_i) - f^{(1)}_{\tau_\eta}(\mathbf{x}_j) \right| \\
&= \sum_{p \in \Psi} \left| f^{(1)}_p(\mathbf{x}_i) - f^{(1)}_p(\mathbf{x}_j) \right| + \sum_{q \in \Gamma} \left| f^{(1)}_q(\mathbf{x}_i) - f^{(1)}_q(\mathbf{x}_j) \right| \\
&= \sum_{p \in \Psi} \left| f^{(0)}_p(\mathbf{x}_i) - f^{(0)}_p(\mathbf{x}_j) \right| + \sum_{q \in \Gamma} \left| f^{(1)}_q(\mathbf{x}_i) - f^{(1)}_q(\mathbf{x}_j) \right| \\
&= \sum_{p \in \Psi} \left| f^{(0)}_p(\mathbf{x}_i) - f^{(0)}_p(\mathbf{x}_j) \right| \\
&\quad + \sum_{q \in \Gamma} \left| (1 + c_{\varphi q}) \left(f^{(0)}_q(\mathbf{x}_i) - f^{(0)}_q(\mathbf{x}_j) \right) \right| \\
&= \sum_{p \in \Psi} \left| f^{(0)}_p(\mathbf{x}_i) - f^{(0)}_p(\mathbf{x}_j) \right| \\
&\quad + \sum_{q \in \Gamma} (1 + c_{\varphi q}) \left| f^{(0)}_q(\mathbf{x}_i) - f^{(0)}_q(\mathbf{x}_j) \right|. \qquad (21)
\end{aligned}
$$

Combine the results in (20) and (21), we have

$$\|F^{(1)}(\mathbf{x}_i) - F^{(1)}(\mathbf{x}_j)\|_1 = \|F(\mathbf{x}_i) - F(\mathbf{x}_j)\|_1. \qquad (22)$$

Likewise, we have the similar results for the reminding representation of objective vectors, and finally obtain that

$$\|\mathcal{F}(\mathbf{x}_i) - \mathcal{F}(\mathbf{x}_j)\|_1 = \|F(\mathbf{x}_i) - F(\mathbf{x}_j)\|_1. \qquad (23)$$

This completes the proof. □

Propositions 1 and 2 provide the theory behind the proposed method with respect to the consistency of convergence (both Pareto dominance and Pareto non-dominance) before and after the objective reduction. Proposition 3 provides the theory with respect to the consistency of distribution (based on Manhattan distance). The Manhattan distance calculates the sum of absolute difference along with all the dimensions. Despite being less popular than the Euclidean distance, the Manhattan distance which treats each dimension equally can be more reliable than the Euclidean distance to reflect the difference between vectors in a high-dimensional space.

Note that the above properties hold theoretically only when the representation error is zero. However, when the representation error is small these properties can be reasonably preserved (as will be seen in the following experimental studies). A large representation error indicates that no objective can be decomposed, i.e., there is a lot of information loss if decomposing any one objective into others.

4. Experimental studies

4.1. Test problems

We evaluate the effectiveness of our method on both synthetic and real-world problems. The synthetic problems include the scalable DTLZ5(I, M) problem [6], a proposed test problem based on DTLZ7 [7], and the multi-line distance minimisation problem (ML-DMP) [22].

The DTLZ5(I, M) problem has the following three properties [6]: (1) the dimensionality of the Pareto front is I, where I is smaller or equal to the number of the problem's objectives M, (2) the first $M - I + 1$ objectives are linear dependent with each other, i.e., they are perfectly correlated, while the remaining objectives are in conflict with each other, and are also in conflict with the first $M - I + 1$ objectives, and (3) an essential objective set consists of one of the first $M - I + 1$ objectives and the last $I - 1$ objectives. Mathematically,

$$\mathcal{G} = \{f_\eta, f_{M-I+2}, \ldots, f_M\}, \qquad (24)$$

where $\eta \in \{1, \ldots, M - I + 1\}$.

The considered test problem DTLZ7V is modified from DTLZ7 [7]. The DTLZ7 problem has $2^{(M-1)}$ disconnected Pareto-optimal regions in the search space, and all of its objectives are completely conflicting. We only add a new objective as the one half sum of the first $(M - 1)$ objectives to existing DTLZ7. This problem has some positively correlated objectives, and is designed to demonstrate the results of the dimension reduction-based visualisation methods on the problem with disconnected Pareto front.

The definition of ML-DMP problem can be find in the reference [22], and this problem has the following two properties: (1) the Pareto optimal solutions lie in a regular polygon in the two-dimensional decision space, and (2) the Euclidean distance between any two Pareto optimal solutions in the decision space is equal to the Euclidean distance between their objective images multiplied by a constant.

In the DTLZ5(I, M) problem, some objectives are completely harmonious, while for the considered ML-DMP problem there is not any two objectives that are completely harmonious or conflicting.

The considered real-world problem is a many-objective optimisation problem, which chooses products from a software product line (SPL) feature model based on a set of user preferences (objectives). An SPL [4] is a set of software-intensive systems that share a common set of features satisfying specific needs of a particular market segment or mission and that are developed from a common set of core assets in a prescribed way. In our experiment, we test a real-world SPL model, Drupal [31], with seven objectives as adopted in [14]. Drupal represents the variability in the open source Web content management framework Drupal, and is the largest attributed feature model with non-synthetically generated attributes used for the problem of optimal SPL product selection. The seven objectives are as follows [14].

- Minimise the number of missing features.
- Minimise the number of lines of code.
- Minimise the cyclomatic complexity.
- Maximise the test assertions.
- Maximise the number of installations that contain the feature.
- Minimise the number of developers.
- Minimise the number of changes.

4.2. Metrics

To evaluate visualisation methods, we define several metrics to see how they perform in preserving the dominance relation and the distribution relation after the dimensionality reduction. Specifically, for the assessment of the preservation of the dominance relation between solutions, we define the metric of dominance ratio (DR) as:

$$DR = \frac{|\mathcal{D}|}{|\mathcal{W}|}, \tag{25}$$

where $\mathcal{W}$ denotes the set of all solution pairs and $\mathcal{D}$ denotes the set consisting of the pairs of solutions whose dominance relation are changed after the dimension reduction.

For evaluating the performance on preserving the distribution, we define two metrics to measure the difference of distances between the solutions in the original objective space and their corresponding distances in the reduced objective space. These two metrics are related to two widely-used distance measurements, i.e., the Manhattan distance and the Euclidean distance. Formally, the average distance difference metrics with the Manhattan distance and with the Euclidean distance are respectively defined as:

$$AD_1 = \frac{\sum_{i=1}^{N}\sum_{j>i}^{N}\left|\|F(\mathbf{x}_i) - F(\mathbf{x}_j)\|_1 - \|\mathcal{F}(\mathbf{x}_i) - \mathcal{F}(\mathbf{x}_j)\|_1\right|}{C_N^2}, \tag{26}$$

and

$$AD_2 = \frac{\sum_{i=1}^{N}\sum_{j>i}^{N}\left|\|F(\mathbf{x}_i) - F(\mathbf{x}_j)\|_2 - \|\mathcal{F}(\mathbf{x}_i) - \mathcal{F}(\mathbf{x}_j)\|_2\right|}{C_N^2}, \tag{27}$$

where N is the number of solutions. In (26) and (27), the absolute difference between the distance of every two solutions in the original objective space and their corresponding distance in the reduced objective space is summed up, then divided by the number of the solution pairs.

4.3. Verification

In this section, we examine the proposed method on the DTLZ5(I, M) problem with $I = 3$ and $M = 10$. In this problem, the first eight objectives are linearly dependent with each other and the last three objectives are conflicting. This allows us to verify the accuracy of the proposed method. In this experiment, we use a many-objective evolutionary algorithm, SPEA2+SDE [23], running for 200 generations with 200 individuals to find the Pareto-optimal solutions of DTLZ5(3, 10). SPEA2+SDE modifies the density estimator of SPEA2 [47] to make it suitable in many-objective optimisation. We apply ORV to visualise the obtained solutions in the three-dimensional space.

Fig. 2 shows the objective reduction process of the proposed method on DTLZ5(3, 10), from which we obtain that:

- The redundant objectives in the ten observed objective vectors are sequentially decomposed into three essential objective vectors $\mathbf{y}_3$, $\mathbf{y}_9$ and $\mathbf{y}_{10}$. It is consistent with the fact that DTLZ5(3, 10) has three essential objectives which include the last two objectives and one of the first eight objectives.
- The first eight objective vectors are sequentially decomposed to one of the first eight objective vectors. Specifically, the objective vector $\mathbf{y}_5$ is firstly decomposed to $0.3536\mathbf{y}_8$; then objective vectors $\mathbf{y}_4$, $\mathbf{y}_1$, $\mathbf{y}_8$, $\mathbf{y}_2$, $\mathbf{y}_6$ and $\mathbf{y}_7$ are decomposed to multiples of $\mathbf{y}_3$, sequentially; finally solutions, which are the rows of the matrix consisting of the objective vectors $\mathbf{y}_3$, $\mathbf{y}_9$ and $\mathbf{y}_{10}$, are obtained in the reduced objective space.

$$y_1^{(0)} + y_2^{(0)} + y_3^{(0)} + y_4^{(0)} + y_5^{(0)} + y_6^{(0)} + y_7^{(0)} + y_8^{(0)} + y_9^{(0)} + y_{10}^{(0)}$$
$$y_5^{(0)} = 0.3536 y_8^{(0)}; \quad (3.679 \times 10^{-6})$$
$$y_1^{(0)} + y_2^{(0)} + y_3^{(0)} + y_4^{(0)} + y_6^{(0)} + y_7^{(0)} + 1.3536 y_8^{(0)} + y_9^{(0)} + y_{10}^{(0)}$$
$$y_4^{(1)} = 0.1847 y_8^{(1)}; \quad (5.388 \times 10^{-6})$$
$$y_1^{(1)} + y_2^{(1)} + y_3^{(1)} + y_6^{(1)} + y_7^{(1)} + 1.1847 y_8^{(1)} + y_9^{(1)} + y_{10}^{(1)}$$
$$y_1^{(2)} = 0.07795 y_8^{(2)}; \quad (5.435 \times 10^{-6})$$
$$y_2^{(2)} + y_3^{(2)} + y_6^{(2)} + y_7^{(2)} + 1.07795 y_8^{(2)} + y_9^{(2)} + y_{10}^{(2)}$$
$$y_8^{(3)} = 2.445 y_7^{(3)}; \quad (6.539 \times 10^{-6})$$
$$y_2^{(3)} + y_3^{(3)} + y_6^{(3)} + 3.445 y_7^{(3)} + y_9^{(3)} + y_{10}^{(3)}$$
$$y_2^{(4)} = 0.05132 y_7^{(4)}; \quad (1.022 \times 10^{-5})$$
$$y_3^{(4)} + y_6^{(4)} + 1.05132 y_7^{(4)} + y_9^{(4)} + y_{10}^{(4)}$$
$$y_6^{(5)} = 0.1953 y_7^{(5)}; \quad (1.161 \times 10^{-5})$$
$$y_3^{(5)} + 1.1953 y_7^{(5)} + y_9^{(5)} + y_{10}^{(5)}$$
$$y_7^{(6)} = 17.31 y_3^{(6)}; \quad (1.223 \times 10^{-5})$$
$$18.31 y_3^{(6)} + y_9^{(6)} + y_{10}^{(6)}$$

Fig. 2. Decomposition process of our method for a set of solutions obtained by SPEA2+SDE on DTLZ5(3, 10).

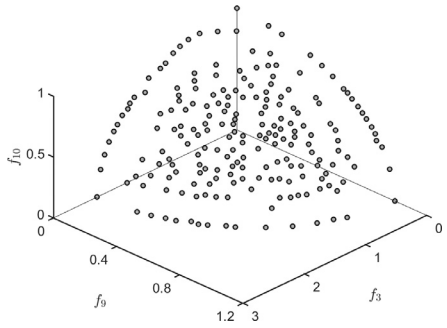


Fig. 3. The result of our method on the solution set obtained by SPEA2+SDE on DTLZ5(3, 10).

- The representation error of the objective vector representation at each iteration is very close to zero as shown in the right brackets in Fig. 2. This is because they are linearly dependent with each other.

Fig. 3 shows the solutions in the reduced objective space. We can see that the solutions in the reduced objectives lie on the surface of the one-eighth ellipsoid. The proposed method can discover the essential objectives automatically. Moreover, the proposed ORV does not discard the information on redundant objectives but transforms it into some other objectives which they are positively correlated with. Finally, the results $DR = 0$ and $AD_1 = 4.6423 \times 10^{-6}$ show that the dominance relation between the solutions is unchanged in the reduced objective space, and the Manhattan distances between them are nearly unchanged. However, the result $AD_2 = 0.3015$ shows that the Euclidean distances between the solutions changed to some extent.

4.4. Comparison with existing visualisation methods

In this section, we compare ORV with six existing methods consisting of two directly plotted methods (the radar chart [18] and the parallel coordinates [15]), and four dimension reduction-based methods (two-stage mapping [19], polar coordinate method [13], PCA [17] and the objective reduction method based on nonlinear correlation information entropy (NCIE) [42]) on DTLZ7V and ML-DMP.

For the DTLZ7V problem, we use the NSGA-II algorithm [5] running for 500 generations with a population size of 200 to obtain the solutions of the 4-objective DTLZ7V. In the DTLZ7V problem, the last objective is the one half sum of the first $M-1$ objective, i.e., $f_4(\mathbf{x}) = \frac{1}{2} f_1(\mathbf{x}) + \frac{1}{2} f_2(\mathbf{x})$.

Fig. 4 shows the decomposition process of ORV for the solution set obtained by NSGA-II. In the process, $\mathbf{y}_4$ is represented by its two nearby objective vectors $\mathbf{y}_1$ and $\mathbf{y}_2$ with the representation error of 4.664×10^{-11}. The representation error can

$$\boldsymbol{y}_1^{(0)} + \boldsymbol{y}_2^{(0)} + \boldsymbol{y}_3^{(0)} + \boldsymbol{y}_4^{(0)}$$

$$\downarrow \quad \boldsymbol{y}_4^{(0)} = 0.5\boldsymbol{y}_1^{(0)} + 0.5\boldsymbol{y}_2^{(0)}; \quad (4.664 \times 10^{-11})$$

$$\mathbf{1.5}\boldsymbol{y}_1^{(0)} + \mathbf{1.5}\boldsymbol{y}_2^{(0)} + \boldsymbol{y}_3^{(0)}$$

Fig. 4. Decomposition process of our method for a set of solutions obtained by NSGA-II on the 4-objective DTLZ7V.

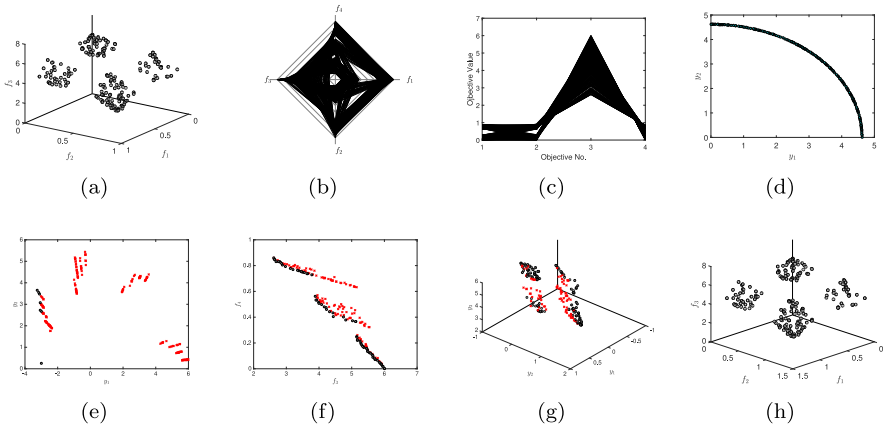


Fig. 5. Visualisation results of a non-dominanted solution set obtained by NSGA-II on the 4-objective DTLZ7V, which is based on DTLZ7 [7]. The red crosses and black circles denote the dominated and the non-dominated solutions in the reduced low-dimensional space, respectively. (a) The objective values of the solutions on the objectives f_1, f_2, f_3. (b) The solutions shown by the radar chart [18]. (c) The solutions shown by the parallel coordinates plot [15]. (d) The solutions plotted by the two-stage mapping method [19]. (e) The solutions plotted by the polar coordinate method [13] ($DR = 0.9222$, $AD_1 = 4.0721$, $AD_2 = 3.3485$). (f) The solutions after the objective reduction process of the NCIE-based method [42] ($DR = 0.7882$, $AD_1 = 0.7035$, $AD_2 = 0.1662$). (g) The solutions after the objective reduction process of PCA [17] ($DR = 0.6164$, $AD_1 = 0.1903$, $AD_2 = 4.256 \times 10^{-16}$). (h) The solutions after the objective reduction process of the proposed method ($DR = 0$, $AD_1 = 0.05590$, $AD_2 = 0.1265$). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

reflect the degree of conflicts between the represented objective vector f_4 and the objective vectors f_1 and f_2. It demonstrates that ORV can automatically detect the correlation between the objectives on DTLZ7V.

Fig. 5 shows the solution set obtained by NSGA-II on the 4-objective DTLZ7V. Fig. 5(a) is the result of the plot with the values on the first three objectives of the obtained solutions. From the results of the radar chart in Fig. 5(b) and the parallel coordinates plot in Fig. 5(c), we can observe the ranges of these four objectives. However, the shape of the Pareto front and the distances between different solutions are not easily understood. With more solutions involved, it becomes more difficult to distinguish one solution from others in these two plots.

In Fig. 5(d), all the solutions are mapped onto a quarter-circle, from which we cannot see the distribution pattern of the original set. In Fig. 5(e), the solutions are transformed into four groups on a 2-D plane, and the shape and distribution of the set are changed. From Fig. 5(f), we can see that the NCIE-based objective reduction method automatically selects f_3, and f_4 from the original four objectives and plots the solutions with their values on these two objectives. The results on DR, AD_1, and AD_2 are 0.7882, 0.7035, and 0.1662, respectively. The dominance and the distribution relations of the solutions are not well preserved. That is because the NCIE-based objective reduction method aims to preserve the conflict between the solutions rather than focusing on preserving the distribution and the dominance relations between the solutions. In Fig. 5(g), the distribution of the solutions is well preserved in terms of the Euclidean distance by using the method of PCA. This is also confirmed by the computed AD_2 which is equal to 4.256×10^{-16}. However, the Manhattan distances between the solutions are not well preserved with AD_1 is equal to 0.1903. Also, the dominance relation between the solutions is not well preserved, with DR equals 0.6164.

Comparing Fig. 5(a) and (h), we can easily find that all the solutions are distributed consistently in these two plots. The AD_1 and AD_2 values are equal to 0.05590 and 0.1265, respectively, which also illustrates that the proposed method preserves the distribution of solutions well. The result on DR equals zero, which means that the dominance relation between the solutions is unchanged in the reduced space.

For the ML-DMP problem, we use the decomposition-based multi-objective evolutionary algorithm with the penalty-based boundary intersection aggregation function (MOEA/D+PBI) [45] running for 500 generations with a population size of 126 to obtain the solutions of the 5-objective ML-DMP (Type I).

In the ML-DMP problem, the distribution of the solutions in the objective space is as same as that in the decision space. This provides us a straightforward way to evaluate the accuracy of a visualisation method by comparing the distribution of the solutions after the dimension reduction with the distribution of the solutions in the two-dimensional decision space.

$$\boldsymbol{y}_1^{(0)} + \boldsymbol{y}_2^{(0)} + \boldsymbol{y}_3^{(0)} + \boldsymbol{y}_4^{(0)} + \boldsymbol{y}_5^{(0)}$$

$$\boldsymbol{y}_1^{(0)} = 0.5209\boldsymbol{y}_2^{(0)} + 0.5209\boldsymbol{y}_5^{(0)}; \quad (0.2793)$$

$$1.5209\boldsymbol{y}_2^{(0)} + \boldsymbol{y}_3^{(0)} + \boldsymbol{y}_4^{(0)} + 1.5209\boldsymbol{y}_5^{(0)}$$

$$\boldsymbol{y}_3^{(1)} = 0.3426\boldsymbol{y}_2^{(1)} + 0.5208\boldsymbol{y}_4^{(1)}; \quad (0.2794)$$

$$1.3426\boldsymbol{y}_2^{(1)} + 1.5208\boldsymbol{y}_4^{(1)} + \boldsymbol{y}_5^{(1)}$$

Fig. 6. Decomposition process of our method for a set of solutions obtained by MOEA/D+PBI on the 5-objective ML-DMP.

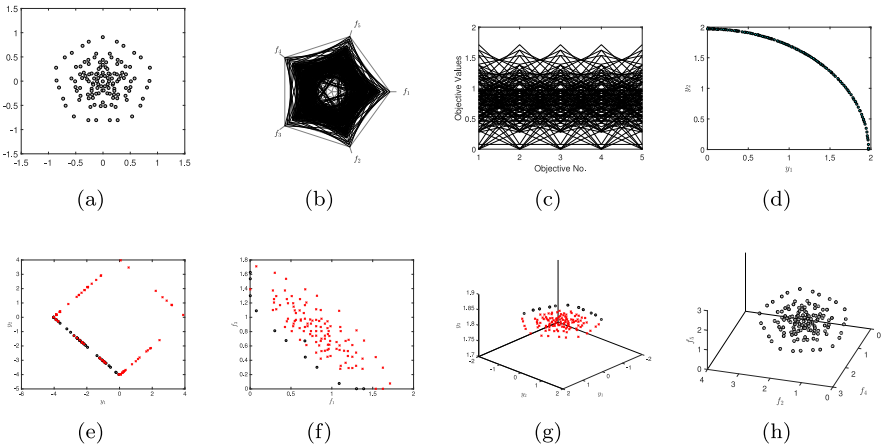


Fig. 7. Visualisation results of a non-dominated solution set obtained by MOEA/D-PBI on the 5-objective ML-DMP, where the solutions in the decision space are similar (in the sense of Euclidean geometry) to their images in the objective space [22]. The red crosses and black circles denote the dominated and the non-dominated solutions in the reduced low-dimensional space, respectively. (a) The original solutions in the decision space. (b) The solutions shown by the radar chart [18]. (c) The solutions shown by the parallel coordinates plot [15]. (d) The solutions plotted by the two-stage mapping method [19]. (e) The solutions plotted by the polar coordinate method [13] ($DR = 0.8559$, $AD_1 = 2.9542$, $AD_2 = 2.5335$). (f) The solutions after the objective reduction process of the NCIE-based method [42] ($DR = 0.9004$, $AD_1 = 1.2231$, $AD_2 = 0.4046$). (g) The solutions after the objective reduction process of PCA [17] ($DR = 0.9182$, $AD_1 = 0.7482$, $AD_2 = 5.6592 \times 10^{-14}$). (h) The solutions after the objective reduction process of the proposed method ($DR = 0$, $AD_1 = 0.5198$, $AD_2 = 0.3615$). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Fig. 6 shows the decomposition process of ORV for the solution set obtained by MOEA/D+PBI. In the process, $\mathbf{y}_3$ is represented by its two nearby objective vectors $\mathbf{y}_2$ and $\mathbf{y}_4$ with the representation error of 0.2793, and then the $\mathbf{y}_5$ is presented by its two nearby objective vectors $\mathbf{y}_1$ and $\mathbf{y}_4$, with the representation error of 0.2794. The representation errors are much larger than those for DTLZ5 (3, 10), because the decomposed objective has stronger conflicts with other objectives in this problem than those in the previous problems.

Fig. 7 shows the solution set obtained by MOEA/D-PBI on the 5-objective ML-DMP. From Fig. 7(a), we can see that the solutions are located inside a two-dimensional pentagon. From the results of the radar chart in Fig. 7(b) and the parallel coordinates plot in Fig. 7(c), we can observe the ranges of these five objectives. However, the shape of the Pareto front and the distances between different solutions are also not easily understood.

From Fig. 7(d), we can see the dominance relation between the solutions (here all the solutions are non-dominated to each other), but we cannot see the distribution between the solutions. In Fig. 7(e), similar to Fig. 7(d), the obtained solution set does not imply their original distribution in the high-dimensional space, despite the preservation of their dominance relation. From the result of the NCIE-based objective reduction method in Fig. 7(f), we can see that it automatically selects the first and the third objectives from the original five objectives and plots the solutions with their values on these two objectives. The results on DR, AD_1, and AD_2 are 0.9004, 1.2231, and 0.4046, respectively. It means that the dominance and the distribution relations of the solutions are not well preserved by automatically selecting two objectives with the NCIE-based method. In Fig. 7(g), the distribution of the solutions is well preserved in terms of the Euclidean distance by using the method of PCA. This is also confirmed by the computed AD_2 which is equal to 5.6592×10^{-14}. However, PCA cannot make the Manhattan distances between the solutions be well preserved as AD_1 is equal to 0.7482. Also, the dominance relation between the solutions is not well preserved, with DR equals 0.9182. Most of the non-dominated solutions in original space are been dominated by other solutions in the reduced low-dimensional space as denoted as cross in Fig. 7(g).

Comparing Fig. 7(a) and (h), we can easily find that all the solutions are distributed consistently in these two plots. The AD_1 and AD_2 values are equal to 0.5198 and 0.3615, respectively, which also illustrates that the proposed method preserves the distribution of solutions fairly well. The result on DR equals zero, which means that the dominance relation between

$$\boldsymbol{y}_1^{(0)} + \boldsymbol{y}_2^{(0)} + \boldsymbol{y}_3^{(0)} + \boldsymbol{y}_4^{(0)} + \boldsymbol{y}_5^{(0)} + \boldsymbol{y}_6^{(0)} + \boldsymbol{y}_7^{(0)}$$

$$\boldsymbol{y}_3^{(0)} = 0.002727\boldsymbol{y}_6^{(0)} + 0.01979\boldsymbol{y}_7^{(0)}; \quad (0.08378)$$

$$\boldsymbol{y}_1^{(0)} + \boldsymbol{y}_2^{(0)} + \boldsymbol{y}_4^{(0)} + \boldsymbol{y}_5^{(0)} + 1.002727\boldsymbol{y}_6^{(0)} + 1.01979\boldsymbol{y}_7^{(0)}$$

$$\boldsymbol{y}_5^{(1)} = 58430\boldsymbol{y}_1^{(1)} + 4013\boldsymbol{y}_4^{(1)}; \quad (0.09080)$$

$$58431\boldsymbol{y}_1^{(1)} + \boldsymbol{y}_2^{(1)} + 4014\boldsymbol{y}_4^{(1)} + \boldsymbol{y}_6^{(1)} + \boldsymbol{y}_7^{(1)}$$

$$\boldsymbol{y}_2^{(2)} = 51.25\boldsymbol{y}_6^{(2)} + 296.5\boldsymbol{y}_7^{(2)}; \quad (0.1383)$$

$$\boldsymbol{y}_1^{(2)} + \boldsymbol{y}_4^{(2)} + 52.25\boldsymbol{y}_6^{(2)} + 297.5\boldsymbol{y}_7^{(2)}$$

$$\boldsymbol{y}_7^{(3)} = 0.5722\boldsymbol{y}_6^{(3)}; \quad (0.2882)$$

$$\boldsymbol{y}_1^{(3)} + \boldsymbol{y}_4^{(3)} + 1.5722\boldsymbol{y}_6^{(3)}$$

Fig. 8. The decomposition process of the proposed method for the solution set obtained by SPEA2+SDE on a real-world problem.

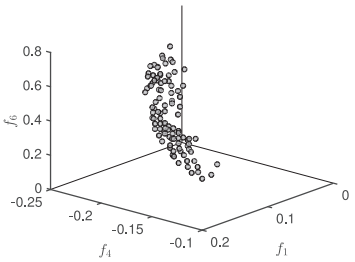


Fig. 9. The result of our method for the solution set obtained by SPEA2+SDE on a real-world problem.

the solutions is unchanged in the reduced space. These observations demonstrate that the proposed method can preserve the relation between the solutions, especially, the dominance relation between them.

4.5. On a real-world problem

In this section, we evaluate the performance of ORV on the 7-objective SPL product selection problem. We first use SPEA2+SDE[1] to obtain a set of 200 solutions. As the objective scale in this problem differs largely, we normalise the solution set by its range, and then use ORV to process this solution set, aiming to reveal some properties of the considered problem (e.g., the relationship between the objectives to be optimised).

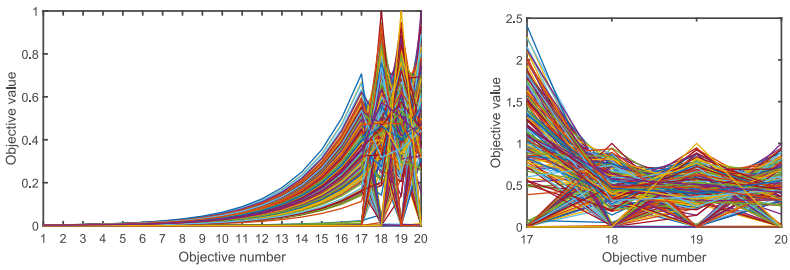
Fig. 8 shows the objective vector decomposition process, from which we have the following observations:

- In the first iteration, $\mathbf{y}_3$ is represented as a combination of $\mathbf{y}_6$ and $\mathbf{y}_7$. It means that the objective f_3 (i.e., minimising the cyclomatic complexity) correlates positively with the objectives f_6 (minimising the number of developers) and f_7 (minimising the number of changes). This sounds reasonable since simpler code might have fewer developers and also fewer changes (since it is less likely that faults will have been introduced).
- In the second iteration, $\mathbf{y}_5$ is represented as a combination of $\mathbf{y}_1$ and $\mathbf{y}_4$. It means that the objective f_5 (maximising the number of installations that contain the feature) is somehow correlated with the objectives f_1 (maximising the richness of features) and f_4 (maximising the test assertions).
- The representation error in the last iteration is much larger than that in the previous iterations, which means that it may not be accurate to represent $\mathbf{y}_6$ with $\mathbf{y}_7$.

The results of the solutions in the reduced objective space are shown in Fig. 9. From the figure, we can see that most of the solutions are non-dominated with respect to each other. Actually, the result on *DR* is 0.02828, i.e., there are around 3% pairs changed their dominance relation when reduced to the three-dimensional space.

The use of the proposed method is two-fold. On one hand, the decomposition process of ORV helps practitioners understand the relationships (i.e., conflict/harmony) among objectives, and also the produced representation error can tell practitioners the cost of reducing the objectives. On the other hand, the scatter plot of solutions after the objective reduction lets the decision makers be aware of the shape of solution set and also the distribution of solutions (e.g., the boundary solutions and knee solutions), which may facilitate the decision-making process.

[1] SPEA2+SDE had been found to be promising in the considered problem [14].



(a) (b)

Fig. 10. The solutions obtained by SPEA2+SDE on DTLZ5(4, 20) with the parallel coordinates plots. (a) The solutions in the original objective space. (b) The solutions after the objective reduction process of our method. The lines with the same colour in two plots are the same solution in two different spaces, respectively.

When executing the proposed method, it is desirable to provide the value of the expected dimensionality of the observed objective space, m. However, in practice, we usually do not know the number of essential objectives in advance. In this case, we suggest using the elbow method [35] to estimate the number of essential objectives. More specifically, one can run the objective decomposition of $M-1$ iterations and plot a line chart of the reconstruction error for each objective decomposition iteration, where M is the number of the problem's objectives. If the line chart looks like an arm, then the "elbow" on the arm corresponds to the suitable number of decomposition iterations $M-m$ and one can obtain the suitable value of the parameter m.

4.6. Extension to the case with more essential objectives

In the above experiments, since the number of the essential objectives is less than four, we can plot the solutions in the three-dimensional objective space. In this section, we apply our dimension reduction technique to show the solutions in the parallel coordinate system. Fig. 10(b) shows the result of ORV for the solution set obtained by SPEA2+SDE on DTLZ5(4, 20) in the parallel coordinate system along with the solutions in the original objective space shown in Fig. 10(a). The obtained values of the metrics DR, AD_1, and AD_2 are 0, 4.6423×10^{-6}, and 0.3015, respectively. From these evaluation results and Fig. 10, we have the following observations:

- The preserved objectives are f_{17}, f_{18}, f_{19}, and f_{20}. These objectives form an essential objective set for DTLZ5 (4, 20). This means that the proposed method successfully detects the essential objectives and preserves the conflicts among the objectives.
- The maximal value of the 17th objective in the reduced objective space is larger than that in the original objective space because the information on the first 16 objectives has been transformed to the 17th objective.
- The dominance relation and the Manhattan distances between the solutions are well preserved.
- Even if the proposed method can reduce the dimensionality of the observation objective space, it is still hard to see the distribution of the solutions in Fig. 10(b) since the overlapped lines in the parallel coordinates prevents solutions from being clearly presented [24].

5. Conclusion

Visualising a solution set of the many-objective optimisation problem is a challenging issue. Solutions lying in a high-dimensional space make them hard to be observed and perceived. This paper presented an objective reduction-based visualisation method, which tried to preserve the dominance and distribution relations between solutions during the objective reduction process. This allows users to visually know the behaviour of the solutions (e.g., their convergence and distribution shape) in certain cases.

Extensive experiments have been carried out to verify the effectiveness of the proposed method. From the results on three synthetic benchmarks, i.e., DTLZ5(I, M), DTLZ7V (Fig. 7), and ML-DMP (Fig. 5), we have found that the distribution and dominance relation between solutions are well preserved. The proposed method was also applied to discover the relation between objectives in a real-world software engineering problem, and some connections between objectives have been observed.

It is worth mentioning that the proposed method cannot be applied to the situation where there is no redundant objective at all, unless allowing the compromise of preserving the distribution and the dominance relation between solutions. For the situation where the number of objectives is not reducible but one still wants to proceed, the resultant solution set will be different from the original set in terms of solutions' distribution and dominance relations, despite that the proposed method tries to preserve them unchanged as much as possible.

Finally, we would like to note that the proposed method only considers the linear relation between objectives. However, there are many optimisation scenarios where the relation between objectives is nonlinear. For such problems, a potential

solution is to model the dominance relation preservation as a constraint to some existing nonlinear dimension reduction techniques (e.g., Sammon mapping). This would be one important topic of our subsequent work.

Acknowledgments

The authors would like to thank Prof. Robert M. Hierons and Dr. Sergio Segura for discussions on the SPL product selection problem. This work was supported by the National Natural Science Foundation of China under grants 61432012 and 61329302, the Engineering and Physical Sciences Research Council (EPSRC) of U.K. under grants EP/J017515/1 and EP/P005578/1, the Program for Guangdong Introducing Innovative and Entrepreneurial Teams (Grant no. 2017ZT07X386), Shenzhen Peacock Plan (Grant no. KQTD2016112514355- 531), the Science and Technology Innovation Committee Foundation of Shenzhen (Grant no. ZDSYS-201703031748284), the Program for University Key Laboratory of Guangdong Province (Grant no. 2017KSYS008), and the Sichuan Science and Technology Planning Projects (Grants nso. 2019YFH0075 and 2018-GZDZX0030).

References

[1] S. Boyd, L. Vandenberghe, Convex Optimization, Cambridge University Press, New York, NY, USA, 2004.
[2] D. Brockhoff, E. Zitzler, Are all objectives necessary? on dimensionality reduction in evolutionary multiobjective optimization, in: Parallel Problem Solving from Nature-PPSN IX, Springer, 2006, pp. 533–542.
[3] D. Brockhoff, E. Zitzler, Objective reduction in evolutionary multiobjective optimization: Theory and applications, Evol. Comput. 17 (2009) 135–166.
[4] P. Clements, L. Northrop, Software Product Lines: Practices and Patterns, Addison-Wesley Professional, 2001.
[5] K. Deb, A. Pratap, S. Agarwal, T. Meyarivan, A fast and elitist multiobjective genetic algorithm: NSGA-II, IEEE Trans. Evol. Comput. 6 (2002) 182–197.
[6] K. Deb, D. Saxena, Searching for Pareto-optimal solutions through dimensionality reduction for certain large-dimensional multi-objective optimization problems, in: Proceedings of the IEEE IEEE Congress on Evolutionary Computation, 2006, pp. 3352–3360.
[7] K. Deb, L. Thiele, M. Laumanns, E. Zitzler, in: Scalable Test Problems for Evolutionary Multiobjective Optimization, Springer London, 2005, pp. 105–145.
[8] R.M. Everson, J.E. Fieldsend, Multi-class ROC analysis from a multi-objective optimisation perspective, Pattern Recognit. Lett. 27 (2006) 918–927.
[9] T. Gal, T. Hanne, Consequences of dropping nonessential objectives for the application of mcdm methods, Eur. J. Oper. Res. 119 (1999) 373–378.
[10] M. Grant, S. Boyd, CVX: Matlab software for disciplined convex programming, version 2.1, 2014, http://cvxr.com/cvx.
[11] M.C. Grant, S.P. Boyd, in: Graph Implementations for Nonsmooth Convex Programs, Springer London, 2008, pp. 95–110.
[12] Z. He, G.G. Yen, An improved visualization approach in many-objective optimization, in: Proceedings of the 2016 IEEE Congress on Evolutionary Computation (CEC), 2016, pp. 1618–1625.
[13] Z. He, G.G. Yen, Visualization and performance metric in many-objective optimization, IEEE Trans. Evol. Comput. 20 (2016) 386–402.
[14] R.M. Hierons, M. Li, X. Liu, S. Segura, W. Zheng, SIP: optimal product selection from feature models using many-objective evolutionary optimization, ACM Trans. Softw. Eng. Methodol. 25 (2016) 17:1–17:39.
[15] A. Inselberg, B. Dimsdale, Parallel coordinates: a tool for visualizing multi-dimensional geometry, in: Proceedings of the 1st Conference on Visualization, 1990, pp. 361–378.
[16] H. Ishibuchi, N. Tsukamoto, Y. Nojima, Evolutionary many-objective optimization: a short review, in: Proceedings of the IEEE Congress on Evolutionary Computation, 2008, pp. 2419–2426.
[17] I. Jolliffe, Principal Component Analysis, Wiley Online Library, 2002.
[18] E. Kasanen, R. Östermark, M. Zeleny, Gestalt system of holistic graphics: new management support view of MCDM, Comput. Oper. Res. 18 (1991) 233–239.
[19] M. Koppen, K. Yoshida, Visualization of Pareto-sets in evolutionary multi-objective optimization, in: Proceedings of the International Conference on Hybrid Artificial Intelligence Systems, 2007, pp. 156–161.
[20] B. Li, J. Li, K. Tang, X. Yao, Many-objective evolutionary algorithms: A survey, ACM Comput. Surv. 48 (2015) 13:1–13:35.
[21] M. Li, Evolutionary many-objective optimisation: pushing the boundaries, Brunel University London, 2015 Ph.D. dissertation.
[22] M. Li, C. Grosan, S. Yang, X. Liu, X. Yao, Multiline distance minimization: a visualized many-objective test problem suite, IEEE Trans. Evolut. Comput. 22 (2018) 61–78.
[23] M. Li, S. Yang, X. Liu, Shift-based density estimation for Pareto-based algorithms in many-objective optimization, IEEE Trans. Evol. Comput. 18 (2014) 348–365.
[24] M. Li, L. Zhen, X. Yao, How to read many-objective solution sets in parallel coordinates, IEEE Comput. Intell. Mag. 12 (2017) 88–100.
[25] A. López Jaimes, C.A. Coello Coello, D. Chakraborty, Objective reduction using a feature selection technique, in: Proceedings of the 10th Annual Conference on Genetic and Evolutionary Computation, ACM, 2008, pp. 673–680.
[26] D. Lowe, M.E. Tipping, NeuroScale: novel topographic feature extraction using RBF networks, in: Proceedings of the 9th International Conference on Neural Information Processing Systems, 1996, pp. 543–549.
[27] Y. Masafumi, Y. Tomohiro, F. Takeshi, Study on effect of moga with interactive island model using visualization, in: Proceedings of the IEEE Congress on Evolutionary Computation, 2010, pp. 1–6.
[28] K. Miettinen, Survey of methods to visualize alternatives in multiple criteria decision making problems, OR Spectr. 36 (2014) 3–37.
[29] A. Pryke, S. Mostaghim, A. Nazemi, Heatmap visualization of population based multi objective algorithms, in: Proceedings of the First International Conference on Evolutionary Multi-Criterion Optimization, 2007, pp. 361–375.
[30] J.W. Sammon, A nonlinear mapping for data structure analysis, IEEE Trans. Comput. 100 (1969) 401–409.
[31] A.B. Sánchez, S. Segura, J.A. Parejo, A. Ruiz-Cortés, Variability testing in the wild: the Drupal case study, Softw. Syst. Model. 16 (2017) 173–194.
[32] D.K. Saxena, K. Deb, Non-linear dimensionality reduction procedures for certain large-dimensional multi-objective optimization problems: employing correntropy and a novel maximum variance unfolding, in: Proceedings of the International Conference on Evolutionary Multi-Criterion Optimization, Springer, 2007, pp. 772–787.
[33] D.K. Saxena, J.A. Duro, A. Tiwari, K. Deb, Q. Zhang, Objective reduction in many-objective optimization: Linear and nonlinear algorithms, IEEE Trans. Evol. Comput. 17 (2013) 77–99.
[34] H.K. Singh, A. Isaacs, T. Ray, A Pareto corner search evolutionary algorithm and dimensionality reduction in many-objective optimization problems, IEEE Trans. Evol. Comput. 15 (2011) 539–556.
[35] R.L. Thorndike, Who belongs in the family? Psychometrika 18 (1953) 267–276.
[36] H.L. Trinkaus, T. Hanne, KnowCube: a visual and interactive support for multicriteria decision making, Comput. Oper. Res. 32 (2005) 1289–1309.
[37] P.A. Tukey, J.W. Tukey, Preparation; prechosen sequences of views, in: Interpreting multivariate data, 1981, pp. 189–213.
[38] T. Tušar, B. Filipic, Visualization of Pareto front approximations in evolutionary multiobjective optimization: A critical review and the prosection method, IEEE Trans. Evol. Comput. 19 (2015) 225–245.
[39] J.J. Valdés, A.J. Barton, Visualizing high dimensional objective spaces for multi-objective optimization: a virtual reality approach, in: Proceedings of the IEEE Congress on Evolutionary Computation, 2007, pp. 4199–4206.

[40] D.J. Walker, R. Everson, J.E. Fieldsend, Visualizing mutually nondominating solution sets in many-objective optimization, IEEE Trans. Evol. Comput. 17 (2013) 165–184.
[41] H. Wang, M. Olhofer, Y. Jin, A mini-review on preference modeling and articulation in multi-objective optimization: current status and challenges, Complex Intell. Syst. 3 (2017) 233–245.
[42] H. Wang, X. Yao, Objective reduction based on nonlinear correlation information entropy, Soft Comput. 20 (2016) 2393–2407.
[43] Q. Wang, Y. Shen, J.Q. Zhang, A nonlinear correlation measure for multivariable data set, Physica D. 200 (2005) 287–295.
[44] Y. Yuan, Y. Ong, A. Gupta, H. Xu, Objective reduction in many-objective optimization: Evolutionary multiobjective approaches and comprehensive analysis, IEEE Trans. Evolut. Comput. 22 (2018) 189–210.
[45] Q. Zhang, H. Li, MOEA/D: a multiobjective evolutionary algorithm based on decomposition, IEEE Trans. Evol. Comput. 11 (2007) 712–731.
[46] L. Zhen, M. Li, R. Cheng, D. Peng, X. Yao, Adjusting parallel coordinates for investigating multi-objective search, in: Proceedings of the International Conference on Simulated Evolution and Learning, Shenzhen, China, 2017, pp. 224–235.
[47] E. Zitzler, M. Laumanns, L. Thiele, SPEA2: Improving the strength Pareto evolutionary algorithm for multiobjective optimization, in: Proceedings of the Evolutionary Methods for Design, Optimisation, and Control, 2002, pp. 95–100.