

Visualisation of Pareto Front Approximation: A Short Survey and Empirical Comparisons*

Huiru Gao¹, Haifeng Nie¹ and Ke Li²

¹School of Computer Science and Engineering, University of Electronic Science and Technology of China

²Department of Computer Science, University of Exeter

*Email: k.li@exeter.ac.uk, {huir.gao, seapeak.nie}@gmail.com

Abstract: Visualisation is an effective way to facilitate the analysis and understanding of multivariate data. In the context of multi-objective optimisation, comparing to quantitative performance metrics, visualisation is, in principle, able to provide a decision maker better insights about Pareto front approximation sets (e.g. the distribution of solutions, the geometric characteristics of Pareto front approximation) thus to facilitate the decision-making (e.g. the exploration of trade-off relationship, the knee region or region of interest). In this paper, we overview some currently prevalent visualisation techniques according to the way how data is represented. To have a better understanding of the pros and cons of different visualisation techniques, we empirically compare six representative visualisation techniques for the exploratory analysis of different Pareto front approximation sets obtained by four state-of-the-art evolutionary multi-objective optimisation algorithms on the classic DTLZ benchmark test problems. From the empirical results, we find that visual comparisons also follow the *No-Free-Lunch* theorem where no single visualisation technique is able to provide a comprehensive understanding of the characteristics of a Pareto front approximation set. In other words, a specific type of visualisation technique is only good at exploring a particular aspect of the data.

Keywords: Visualisation, multi-objective optimisation, decision-making

1 Introduction

Multi-objective optimisation problems (MOPs), which consider simultaneously optimising more than one conflicting objective function, are ubiquitous across the breadth of science [1], engineer [2,3] and economics [4]. Specifically, this paper considers the MOP formally defined as:

$$\begin{aligned} & \text{minimize } \mathbf{F}(\mathbf{x}) = (f_1(\mathbf{x}), \dots, f_m(\mathbf{x}))^T \\ & \text{subject to } \mathbf{x} \in \Omega \end{aligned} \tag{1}$$

where $\mathbf{x} = (x_1, \dots, x_n)^T$ is a candidate solution, and $\Omega = [x_i^L, x_i^U]^n \subseteq \mathbb{R}^n$ defines the search (or decision variable) space. $\mathbf{F} : \Omega \rightarrow \mathbb{R}^m$ constitutes m conflicting objective functions, and \mathbb{R}^m is the objective space. Given two solutions $\mathbf{x}^1, \mathbf{x}^2 \in \Omega$, we said that \mathbf{x}^1 dominates \mathbf{x}^2 (denoted as $\mathbf{x}^1 \preceq \mathbf{x}^2$) in case $\mathbf{F}(\mathbf{x}^1)$ is not worse than $\mathbf{F}(\mathbf{x}^2)$ in any individual objective and it at least has one better objective. A solution \mathbf{x}^* is Pareto-optimal with respect to (1) in case $\nexists \mathbf{x} \in \Omega$ such that $\mathbf{x} \preceq \mathbf{x}^*$. The set of all Pareto-optimal solutions is called the Pareto set (PS). Accordingly, $PF = \{\mathbf{F}(\mathbf{x}) | \mathbf{x} \in PS\}$ is called the Pareto front (PF).

Evolutionary algorithms (EAs), which work with a population of solutions, are widely accepted as a major approach for multi-objective optimisation. In the past three decades and beyond, significant efforts have been devoted on the developments of evolutionary multi-objective optimisation (EMO) algorithms, such as fast non-dominated sorting genetic algorithm (NSGA-III) [5], indicator-based EA (IBEA) [6] and multi-objective EA based on decomposition (MOEA/D) [7], to find a set of well-converged and well-diversified efficient solutions that approximate the whole PF. Nevertheless, the

*This manuscript is submitted for possible publication. The reviewer can use this manuscript in peer review.

ultimate goal of MO is to help the DM find solution(s) that meet her/his preferences. Providing the DM with a large amount of raw data inarguably incurs severe cognitive obstacle for decision-making, especially when having many objectives (i.e. $m > 3$). Although there are some performance metrics that can evaluate the quality of a PF approximation set (e.g. inverted generational distance (IGD) [8] and Hypervolume (HV) [9]), it is still not intuitive to help the DM understand the trade-off relationship among objectives. Moreover, some recent studies [10] pointed out that the fairness of some currently prevalent performance metrics depends on their parameter settings. In contrast, high-dimensional data visualisation, which is an important and active research field in modern data analysis, is more effective and intuitive to facilitate the DM to understand the trade-off solutions and thus make a meaningful decision.

In the context of EMO, an ideal visualisation technique is expected to satisfy the following prerequisites [11].

- It is able to help the DM identify the Pareto dominance relation between different solutions. This will further facilitate the understanding of the trade-off relationship among objective functions.
- It is not only able to preserve the geometric characteristics (e.g. shape and location) of the approximated PF, but also can help the DM understand such information. This will finally lead to an intuitive comparison of the distribution of solutions obtained by different EMO algorithms.
- It is able to help the DM identify and exploit the region of interest (e.g. knee regions) thus facilitate the post hoc decision-making.
- It is well scalable to the number of dimensions and size of the approximation set without sacrificing its simplicity and computational complexity.
- It is resilient to some modifications in the approximation set, e.g. adding or removing some solutions.

Comparing to the development of new EMO algorithms, the research on visualisation is relatively lukewarm. Many currently prevalent visualisation methods in EMO are directly derived from the modern data analysis field. Generally speaking, they can be divided into three categories. Interested readers are encouraged to find a recent taxonomy from [12]. The first category aims to show all objective information (e.g. scatter plot, bubble chart [13], parallel coordinate plot [14] and heatmap [15]). From the cognitive perspective, human beings can only visually understand objects in a two- or three-dimensional space. Given this consideration, the second category mainly aims to implement the visualisation in a low-dimensional space by using some dimensionality reduction methods (e.g. principle component analysis [16], multidimensional scaling [17], isomap [18], self-organizing map [19]). Partially similar to the second category, the last one tries to map the original data into a new coordinate system (e.g. radial coordinate visualisation [20], projection plot [11], polar coordinate system [21]).

In this paper, we mainly aim to overview some currently prevalent visualisation techniques. In particular, we divide them into three different categories according to how data (i.e. population of solutions) is represented. To have a better understanding of the pros and cons of different visualisation techniques, we compare their visualisation results on various data sets. In particular, these data sets are collected by running four state-of-the-art EMO algorithms on the classic DTLZ benchmark test problems [22]. Note that these test problems have different characteristics and PF shapes.

The rest of this paper is organised as follows. Section 2 provides an overview of the currently prevalent visualisation techniques according to our taxonomy. Section 3 compares and analyses different visualisation techniques on various empirical data collected by running four state-of-the-art EMO algorithms on various test problems. Section 4 concludes this paper and provides some thoughts on future directions.

2 Overview of State-of-the-art Visualisation Techniques in EMO

In this section, we overview some currently prevalent visualisation techniques for visually comparing and understanding Pareto approximation sets in the EMO literature. In particular, these visualisation

techniques are classified into three categories according to how data is represented.

2.1 Visualisation of All Objective Information

The first category of visualisation techniques aims to reveal all individual objective information of the underlying approximation set. Specifically, scatter plot is the most commonly utilised data visualisation technique that provides a holistic exploration of the population distribution. However, the vanilla scatter plot is only useful in a two- or three-objective case. Although a matrix of scatter plots [23], which is an array of scatter plots displaying all possible pairwise combinations of coordinates, can be used to visualise m -dimensional data ($m > 3$), it yields at least $\frac{m(m-1)}{2}$ scatter plots with shared scales. In this case, the information will be too cluttered to provide the DM a holistic exploration of the characteristics of the approximation set, e.g. Pareto dominance relation, PF shape and knee region. Bubble chart is another variant of scatter plot which can visualise at most five-objective approximate set in a three-dimensional Cartesian coordinate system. In particular, the additional two objectives are represented as the sizes of those scatter points and their colours.

Instead of an affine projection of the approximation set into a two- or three-dimensional space as in the scatter plot, parallel coordinate plot (PCP) [14], the most frequently used visualisation technique in evolutionary many-objective optimisation, projects the data into a coordinate system with m parallel axes. In particular, each poly line in the PCP represents a solution that go across all parallel axes where the intersection at each axis is the corresponding objective value. However, due to the cluttering effect and the interference with crossing lines, it is nontrivial to interpret various characteristics of the approximation set. There have been many efforts devoted to the development of PCP variants [24]. For example, in order to enable the PCP to interpret the trade-off relation between objectives, Zhen et al. [25] proposed a parallel coordinates adjustment method. Specifically, conflicting objectives are gathered together where the correlation between any two objectives are measured by the Spearman's rank correlation coefficient. Although this method is simple and has shown promising results on some synthetic data, the first axis in the re-ordered PCP is chosen in a random manner. In [26], Lu et al. proposed to use singular value decomposition (SVD) to calculate the contribution of each objective. Afterwards, the objective associated with the highest eigenvalue is chosen as the first axis in the re-ordered PCP. In [27], Yuan et al. developed an interesting PCP variant that explores the synergy between scatter plot and PCP under the same framework. In particular, by using multidimensional scaling, any two or dimensions in the PCP can be conveniently converted to scatter points and vice versa. Furthermore, there is a brushing tool that allows the user to make selection on both points and line segments conveniently.

Similar to PCP, heatmap [15] is another frequently used technique for high-dimensional data visualisation. Instead of mapping a solution into parallel axes, heatmap translates the objective value into a colour map where a brighter colour usually represents a higher objective value and vice versa. Unfortunately, heatmap suffers from the same cluttering problem as PCP. In principle, the axes re-ordering techniques used in PCP can be directly transferred to improve the interpretability of heatmap. For example, Walker et al. [17] proposed to use spectral seriation to reorder objectives in heatmap. In addition, Radar chart [28] is another equivalent variant of PCP where the coordinate system is changed to a polar coordinate system and m radial axes represent m objectives.

2.2 Visualisation via Dimension Reduction

The second category of visualisation techniques aims to transform the high-dimensional data into a lower dimensional space to facilitate the human cognition. In modern data analytics, there are many dimension reduction techniques available to implement such transformation. For example, Walker et al. proposed to use multi-dimensional scaling (MDS) to map the original high-dimensional population into a two-dimensional space where pair-wise distance between any two solutions is expected to be preserved as much as possible. In fact, other prevalent dimension reduction techniques, e.g. principle component analysis [16], Isomap [18], locally linear embedding [29] and Laplacian eigenmaps [30] has been widely used to visualise and analyse high-dimensional data in the machine learning community [31]. However, their usefulness in the context of multi-objective optimisation is not validated yet. Note that since

it is inevitable to lose some information after the dimension reduction, the resulting plots might be distorted to recover the characteristics of the original PF.

Instead of using dimension reduction techniques from machine learning, Blasco et al. [32] proposed a new visualisation technique called level diagrams to visualise the approximation set in an objective-wise manner. More specifically, each diagram is a two-dimensional scatter plot where the horizontal coordinate represents the objective value at the corresponding objective while the vertical coordinate indicates the distance with respect to the ideal point. As claimed by the authors, the level diagrams are able to facilitate the investigation of some of the PF characteristics, i.e. discontinuities, closeness to ideal point and ranges of attainable values. However, since the plots are shown in an objective-wise manner, the level diagrams will end up with many subplots as the matrix of scatter plots.

In [33], Chiu et al. developed a simple dimension reduction method, called Hyper-Radial Visualisation, to visualise the population in a two-dimensional space by grouping the all objectives into two sets. Although this method is simple, most objective information is directly discarded and the trade-off relationship between objectives is ignored.

2.3 Visualisation via a Transformed Coordinate System

In principle, this last category share some similarity with the dimension reduction as introduced in Section 2.2. They also aim to visualise the original high-dimensional data in a lower-dimensional space. But they try to maintain the original information as much as possible. For example, Ibrahim et al. [20] developed a variant of the classic Radial coordinate visualisation (RadViz) [34], called 3D-RadViz, by adding an additional dimension. In particular, this additional dimension represents the perpendicular distance between a solution and a hyper-plane formed by the extreme points of the underlying population. By doing so, the 3D-RadViz is able to provide the information of the convergence of the population. As claimed by the authors, 3D-RadViz can also provide some information about the PF shape and the population distribution.

In [11], Tušar and Filipič proposed to use prosection method to visualise a four-dimensional approximation set in a three-dimensional space. Furthermore, they also provide some theoretical analysis about the property of using prosection for preserving most of the characteristics of the original PF. However, it is a bit shame that this technique is not easy to be scalable to a higher-dimensional case.

In [21], He and Yen proposed to use a polar coordinate system (PCS) to represent the approximation set. One of the major advantages of using PCS is the intrinsic dimension is two so that it is easy to interpret by the DMs. In order to map the population from a Cartesian coordinate system to a PCS, the DMs need to have some prior knowledge of the characteristics of the PF. In [21], they only considered three cases, i.e. linear, convex and concave. Accordingly, they developed an equation to calculate the new coordinates, i.e. radial value and angular value which represent the convergence and distribution respectively, in the PCS.

3 Experimental Results and Discussion

In order to have a better understanding of the pros and cons of those visualisation techniques introduced in Section 2, we choose six representative visualisation techniques (.e. PCP, heatmap, MDS, 3D-Radvis, prosection and polar coordinate system) and compare their visualisation results on PF approximate sets obtained by four state-of-the-art EMO algorithms (i.e. NSGA-III, MOEA/D, IBEA and HypE [35]). In particular, the classic DTLZ test problems [22] are chosen to form the benchmark suite. Note that these test problems have various characteristics (e.g. biased distribution and multi-modality) and different PF shapes (e.g. linear, concave, degenerate and disconnected) and have been widely used in the literature [36–45]. The number of objectives considered in our experiments is set to $m \in \{3, 5, 8, 10, 15\}$ while the settings for the population size and the number of generations for different test problems are in Table 1 and Table 2. Note that these settings follow the standard settings suggested in [46]. Furthermore, to make sure that all plots are in the same scale, we normalise the objective values of the data in each approximation set to the range $[0, 1]$. Beyond visual comparisons, we also calculate the inverted generational distance (IGD) and Hypervolume (HV) values of the

obtained approximation sets. In particular, we only plot the population with the best IGD value for comparison purpose.

Table 1: Settings of Population Size for Different Algorithms

m	NSGA-III, IBEA and HypE	MOEA/D
3	92	91
5	212	210
8	156	156
10	276	275
15	136	135

Table 2: Number of Generations for Different Test Problems

Test problem	$m = 3$	$m = 5$	$m = 8$	$m = 10$	$m = 15$
DTLZ1	400	600	750	1,000	1,500
DTLZ2, DTLZ5	250	350	500	750	1,000
DTLZ3, DTLZ7	1,000	1,000	1,000	1,500	2,000
DTLZ4, DTLZ6	600	1,000	1,250	2,000	3,000

3.1 Comparison of Six Visualisation Techniques

Due to the page limit, we only present some selected results while the complete plots can be found from the supplementary file of this paper¹. In the following paragraphs, we will separately discuss the observations for each of those six visualisation techniques used in this paper.

3.1.1 Observations on MDS

As shown in Fig. 1 to Fig. 4, we can clearly see that the performance of NSGA-III and MOEA/D are better than that of HypE and IBEA. Especially for HypE, it can only find some sparsely distributed points along the boundary or extreme of the PF. It is interesting to note that the solution distributions obtained for DTLZ1 and DTLZ4 are visually distinguishable in terms of geometry representations. In particular, the PF shapes of DTLZ1 and DTLZ4 are intrinsically different where the prior one is a convex hyper-plane while the latter one is a concave sphere. However, as shown in Fig. 5 and Fig. 6, MDS fails to interpret the geometrical characteristics for some more complex PFs, i.e. DTLZ5 (i.e. degenerate PF) and DTLZ7 (i.e. disconnected PF segments). This might be partially because all algorithms fail to find promising PF approximations on DTLZ5 to DTLZ7. On the other hand, due to the dimensionality reduction, it is inevitable to sacrifice some important information by using the MDS. In this case, the coordinates of those plotted points do not have any physical meaning. Therefore, we can not interpret neither convergence information of the population nor dominance relation by using MDS or other dimension reduction techniques.

3.1.2 Observations on Prosection

In this paper, we set the prosection plain as f_1 and f_2 , the prosection angle as 45° and the prosection width as 1. As shown in Fig. 7 to Fig. 10, prosection is able to preserve the geometrical characteristics as we can clearly distinguish the shape with respect to different test problems. Furthermore, since the prosection does not use any dimension reduction technique, the coordinates of different plots have their physical meaning. Therefore, prosection can facilitate the comparison in terms of convergence and dominance relationship. More specifically, as shown in Fig. 7 and Fig. 8, we can see that NSGA-III and

¹Supplementary file can be downloaded from this Google Drive share.

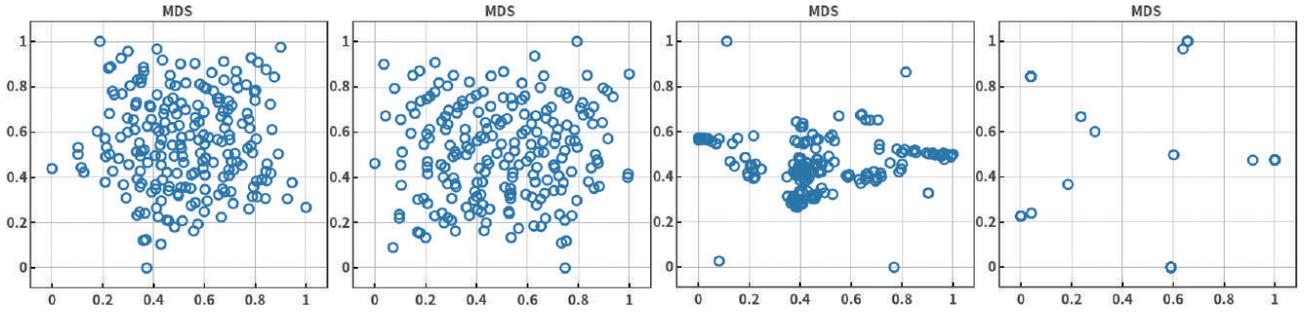


Figure 1: Solutions obtained by NSGA-III, MOEA/D, HypE and IBEA (from left to right) on 5-objective DTLZ1 by using MDS.

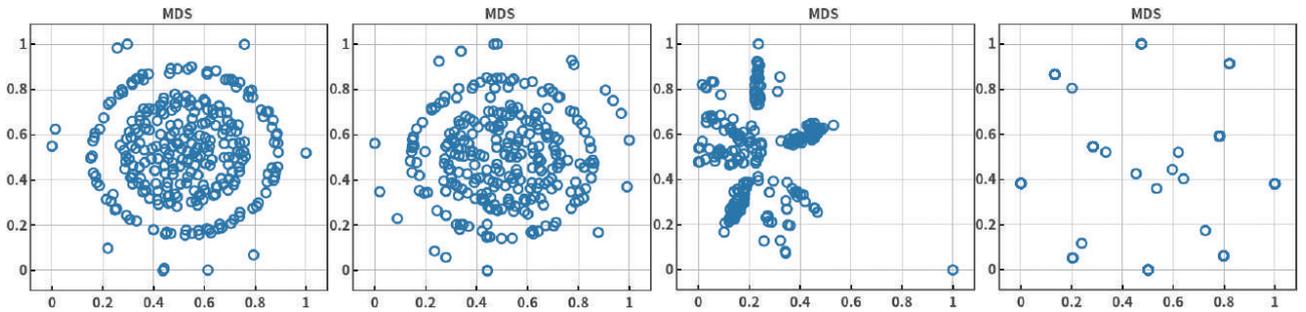


Figure 2: Solutions obtained by NSGA-III, MOEA/D, HypE and IBEA (from left to right) on 10-objective DTLZ1 by using MDS.

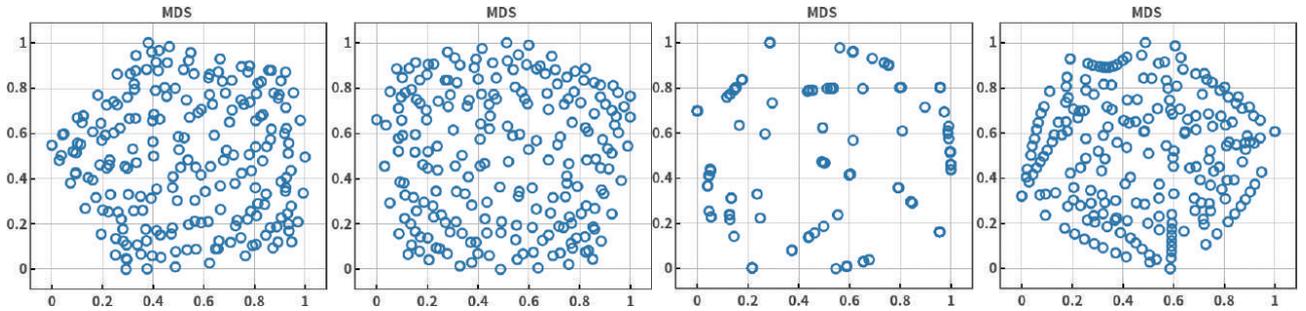


Figure 3: Solutions obtained by NSGA-III, MOEA/D, HypE and IBEA (from left to right) on 5-objective DTLZ4 by using MDS.

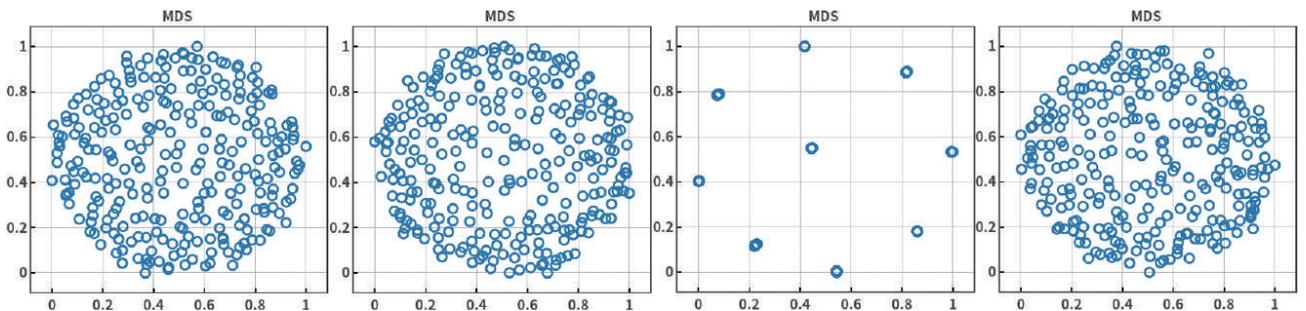


Figure 4: Solutions obtained by NSGA-III, MOEA/D, HypE and IBEA (from left to right) on 10-objective DTLZ4 by using MDS.

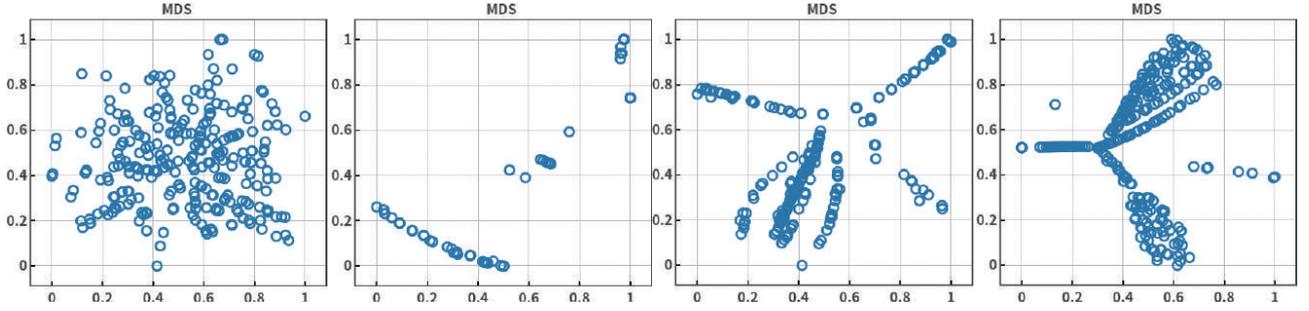


Figure 5: Solutions obtained by NSGA-III, MOEA/D, HypE and IBEA (from left to right) on 10-objective DTLZ5 by using MDS.

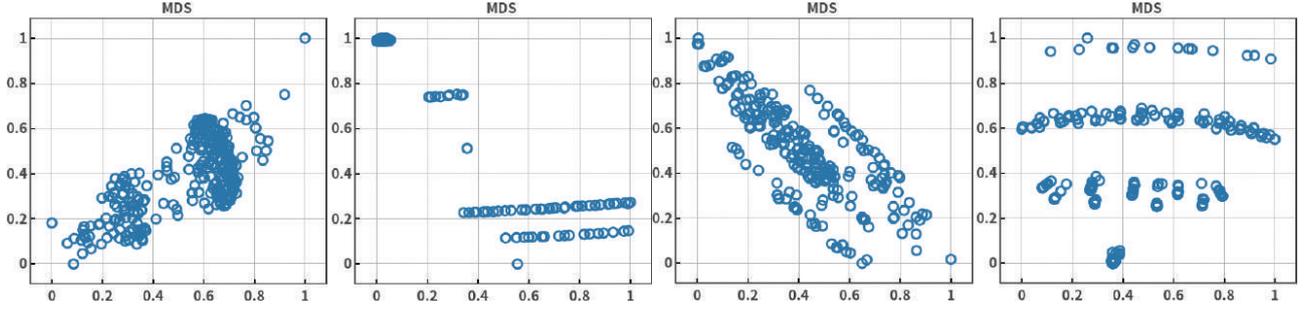


Figure 6: Solutions obtained by NSGA-III, MOEA/D, HypE and IBEA (from left to right) on 10-objective DTLZ7 by using MDS.

MOEA/D have shown better performance than HypE and IBEA. In particular, as observed in Fig. 1 and Fig. 2, IBEA can only find some sparsely distributed points on the boundary or extreme of the PF on DTLZ1. However, as shown in Fig. 9 and Fig. 10, the performance of NSGA-III and MOEA/D are not as promising as HypE and IBEA. This can be explained as the distribution of weight vectors used in NSGA-III and MOEA/D does not fit the PF shapes of DTLZ5 and DTLZ7 [47,48]. Note that such observations are not clear when using MDS for visualisation as discussed in Section 3.1.1. Although projection has shown very promising results for visually comparing different PF approximation sets on 4-dimensional space, it is not directly scalable to high-dimensional problems.

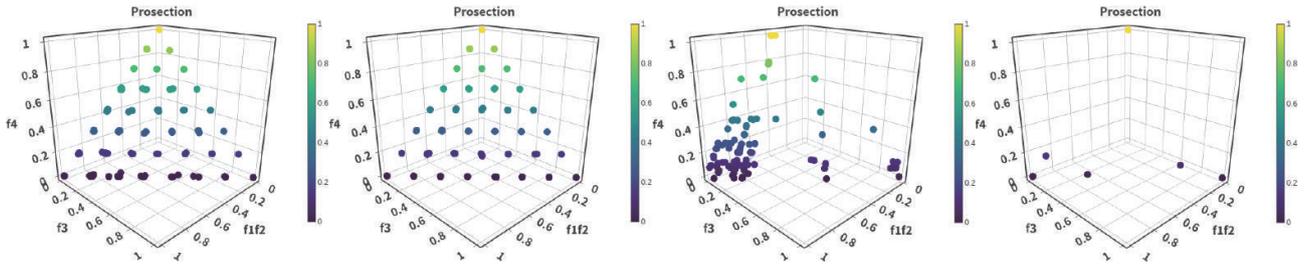


Figure 7: Solutions obtained by NSGA-III, MOEA/D, HypE and IBEA (from left to right) on 4-objective DTLZ1 by using MDS.

3.1.3 Observations on 3D-RadVis

3D-RadViz is a vibrant modification of the original RadViz. Because the additional dimension, which represents the perpendicular distance to a hyper-plane formed by the extreme point of each dimension, 3D-RadViz is able to interpret more information related to the geometry. As shown in Fig. 11 to Fig. 14, we can clearly tell the geometrical difference between the solutions with respect to DTLZ1 and DTLZ4. Note that we rotate the axis in order to provide a better view. However, similar to the observations on MDS, the plots on more complex PF shapes (i.e. DTLZ5 and DTLZ7) are not satisfactory where points are cluttered without an explicit regularity. In addition, due to the addi-

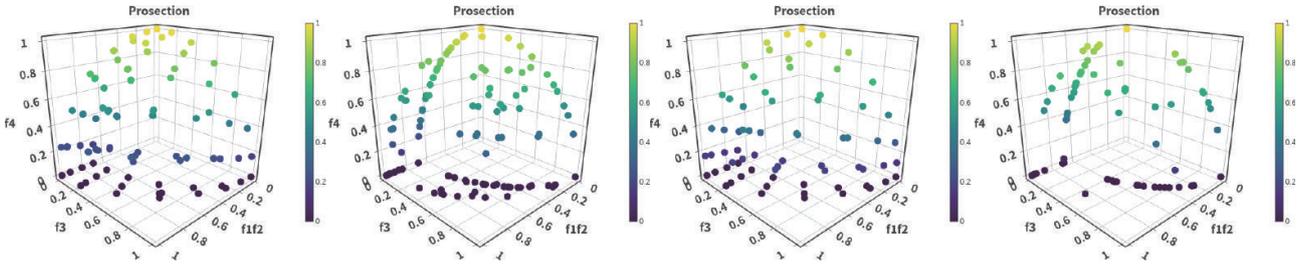


Figure 8: Solutions obtained by NSGA-III, MOEA/D, HypE and IBEA (from left to right) on 4-objective DTLZ4 by using MDS.

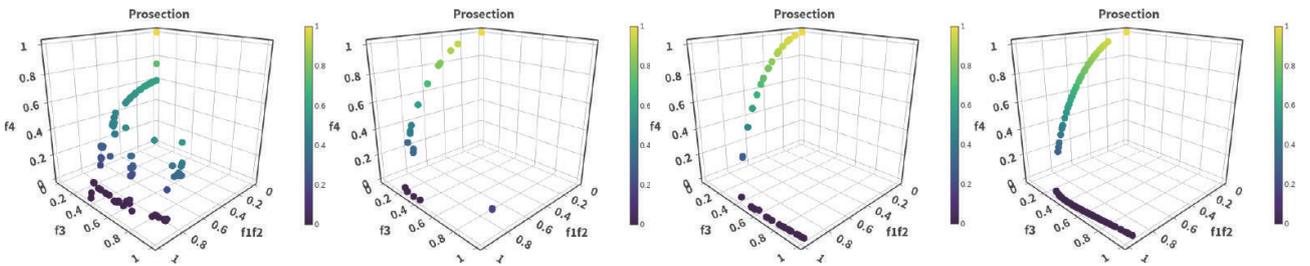


Figure 9: Solutions obtained by NSGA-III, MOEA/D, HypE and IBEA (from left to right) on 4-objective DTLZ5 by using MDS.

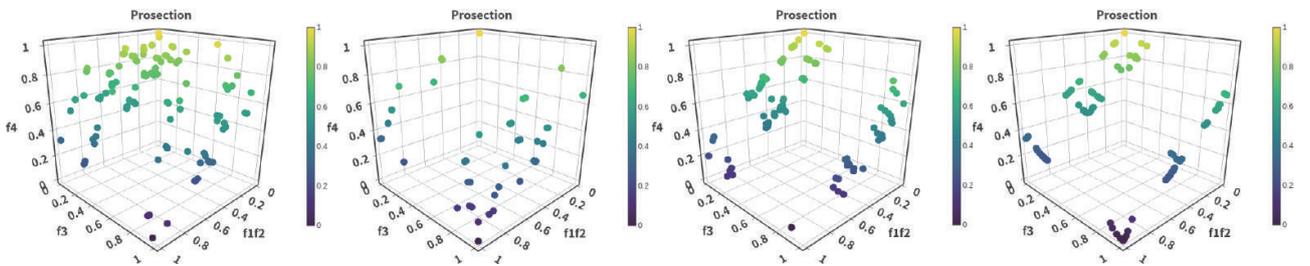


Figure 10: Solutions obtained by NSGA-III, MOEA/D, HypE and IBEA (from left to right) on 4-objective DTLZ7 by using MDS.

tional dimension, 3D-RadViz is able to help interpret the convergence of an approximation set but the dominance relationship between solutions is not interpretable.

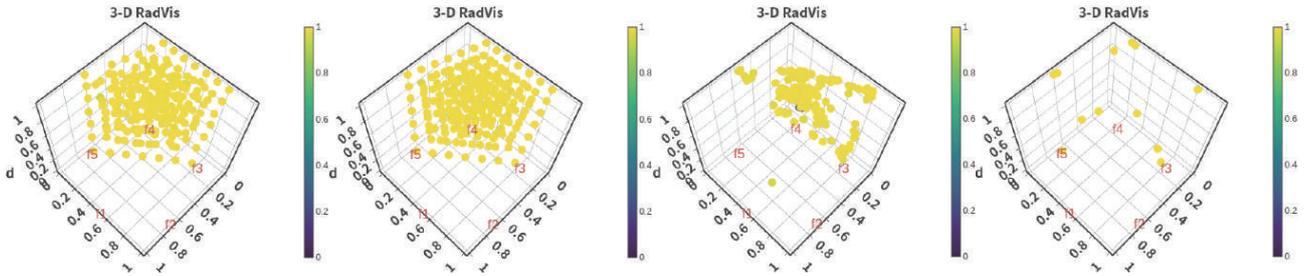


Figure 11: Solutions obtained by NSGA-III, MOEA/D, HypE and IBEA (from left to right) on 5-objective DTLZ1 by using 3D-RadViz.

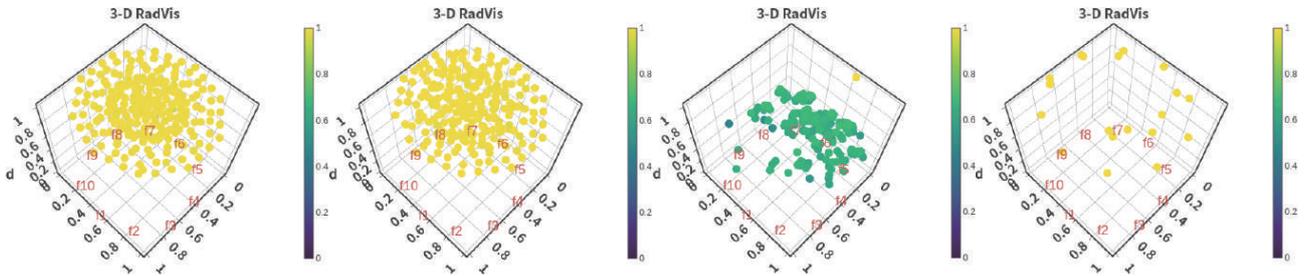


Figure 12: Solutions obtained by NSGA-III, MOEA/D, HypE and IBEA (from left to right) on 10-objective DTLZ1 by using 3D-RadViz.

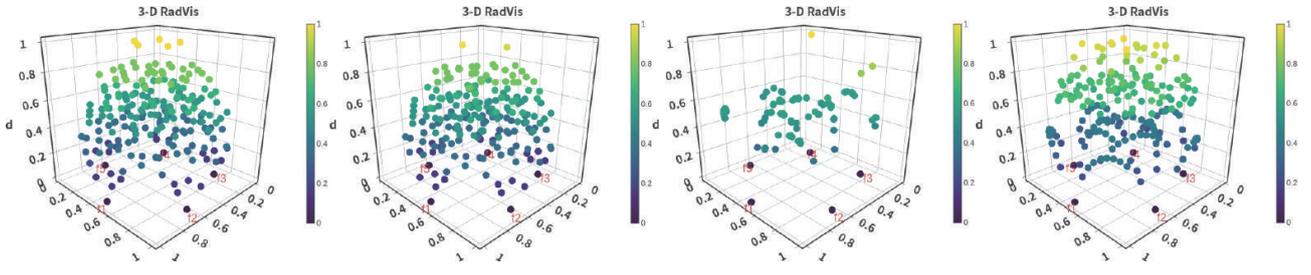


Figure 13: Solutions obtained by NSGA-III, MOEA/D, HypE and IBEA (from left to right) on 5-objective DTLZ4 by using 3D-RadViz.

3.1.4 Observations on PCS

Fig. 17 to Fig. 22 present the results of PCS on various test problems. One of the key characteristics of the PCS is the prior knowledge of the convexity of the target PF. For example, because the PF of DTLZ1 is a convex hyper-plane, the transformed plot in the PCS is a four-edge polygon which is irrelevant to the number of objectives. As shown in Fig. 17 and Fig. 18, solutions obtained by NSGA-III and MOEA/D show a good approximation to the target polygon while IBEA only presents some sparsely distributed points. As for DTLZ2, the PF of which is a concave hyper-sphere, the transformed plot in the PCS is an unit circle. As shown in Fig. 19 and Fig. 20, NSGA-III, MOEA/D and IBEA obtain a promising approximation to the target circle. In PCS, the closeness to the target plot, either a four-edge polygon or an unit circle, can indicate the convergence of the approximation set while the distribution along the target plot can represent the distribution of the approximation set. However, because the target plot is fixed according to the convexity of the PF itself, it is not easy to interpret the original objective axis in the PCS. Even worse, when the PF does not meet either convex or concave shape, the plotting results are even more difficult to interpret, e.g. Fig. 21 and

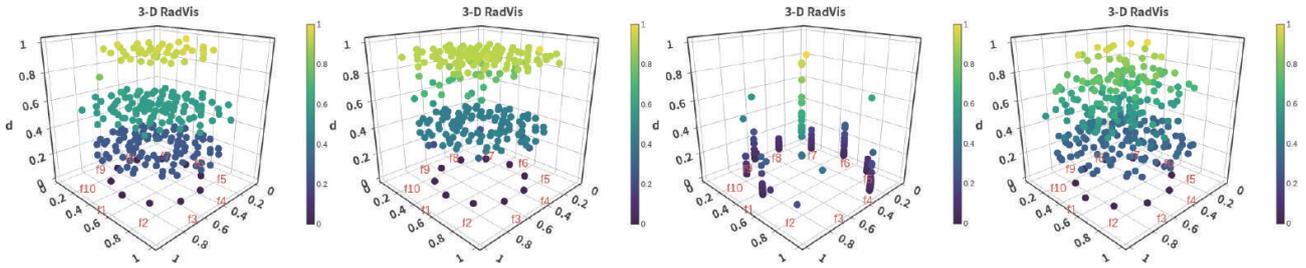


Figure 14: Solutions obtained by NSGA-III, MOEA/D, HypE and IBEA (from left to right) on 10-objective DTLZ4 by using 3D-RadViz.

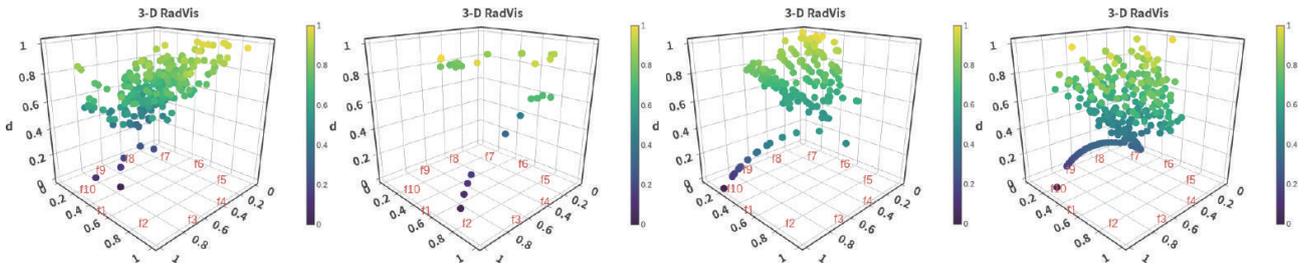


Figure 15: Solutions obtained by NSGA-III, MOEA/D, HypE and IBEA (from left to right) on 10-objective DTLZ5 by using 3D-RadViz.

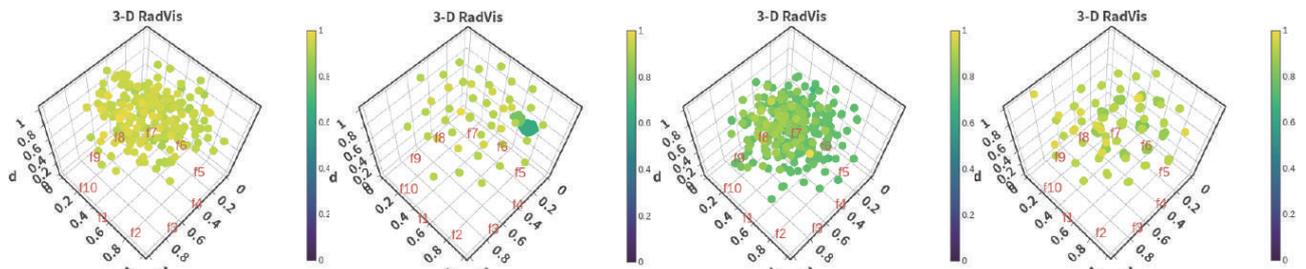


Figure 16: Solutions obtained by NSGA-III, MOEA/D, HypE and IBEA (from left to right) on 10-objective DTLZ7 by using 3D-RadViz.

Fig. 22. Unfortunately, because a real-world optimisation problem is always as a black-box system, it is impossible to have such prior knowledge beforehand.

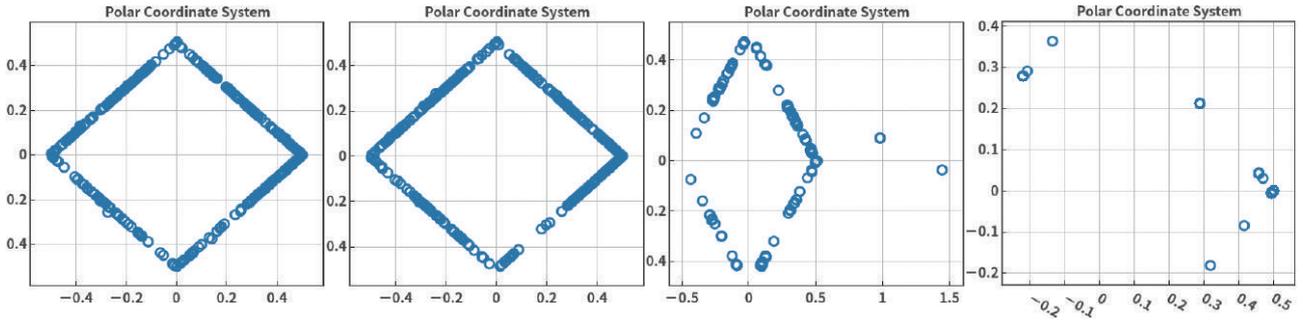


Figure 17: Solutions obtained by NSGA-III, MOEA/D, HypE and IBEA (from left to right) on 5-objective DTLZ1 by using Polar Coordinate System.

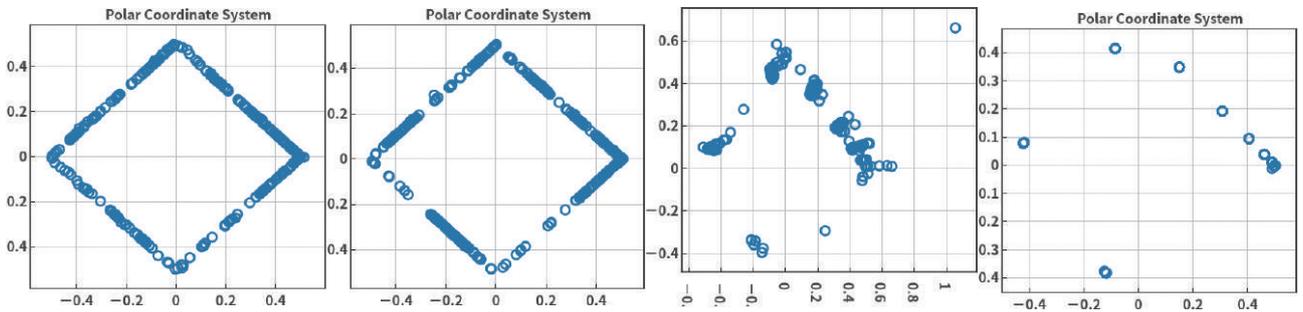


Figure 18: Solutions obtained by NSGA-III, MOEA/D, HypE and IBEA (from left to right) on 10-objective DTLZ1 by using Polar Coordinate System.

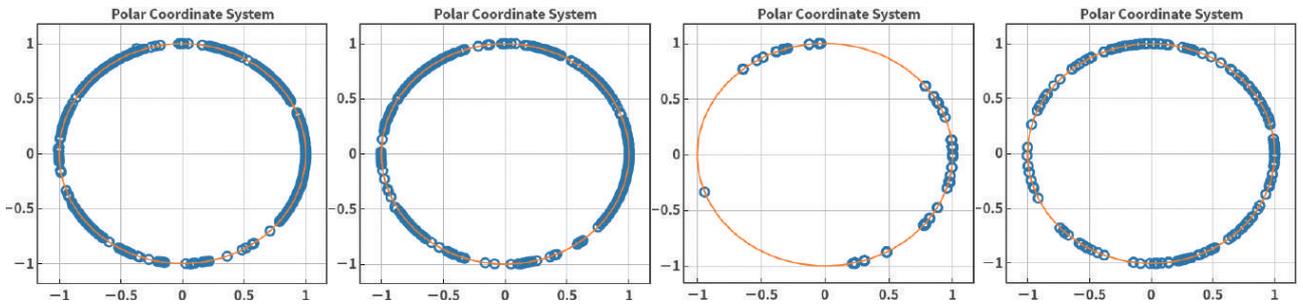


Figure 19: Solutions obtained by NSGA-III, MOEA/D, HypE and IBEA (from left to right) on 5-objective DTLZ4 by using Polar Coordinate System.

3.1.5 Observations on PCP and Heatmap

As discussed in Section 2.1, both PCP and heatmap aim to present all objective information in a two-dimensional plot. Because no information is reduced in PCP and heatmap, both of them have the ability to compare the convergence of the approximation set and infer the dominance relationship. But due to the cluttering effect of the original PCP and heatmap, the axis reordering is able to facilitate their interpretability. As shown in Fig. 23 to Fig. 26, we cannot find significant difference between the original PCP and Heatmap with respect to their reordering version. This can be explained as all objectives in DTLZ1 and DTLZ4 are conflicting to each other. In particular, we find that it is easier to investigate the distribution of objective values in heatmap than in PCP. For example, as shown in Fig. 24 and Fig. 26, there are many overlapping solutions obtained by MOEA/D than by NSGA-III. Furthermore, solutions obtained by IBEA on DTLZ1 and HypE on DTLZ4 mainly focus on the extreme region of the PF. In contrast, it is not quite intuitive to read such information from

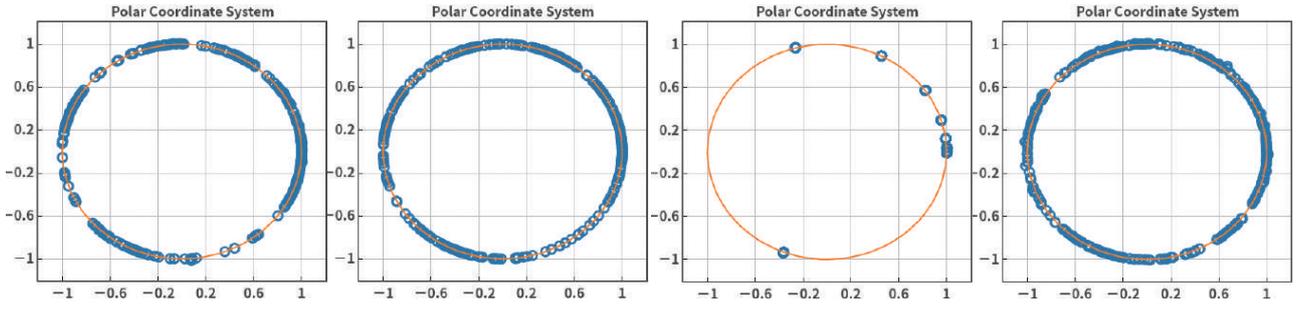


Figure 20: Solutions obtained by NSGA-III, MOEA/D, HypE and IBEA (from left to right) on 10-objective DTLZ4 by using Polar Coordinate System.

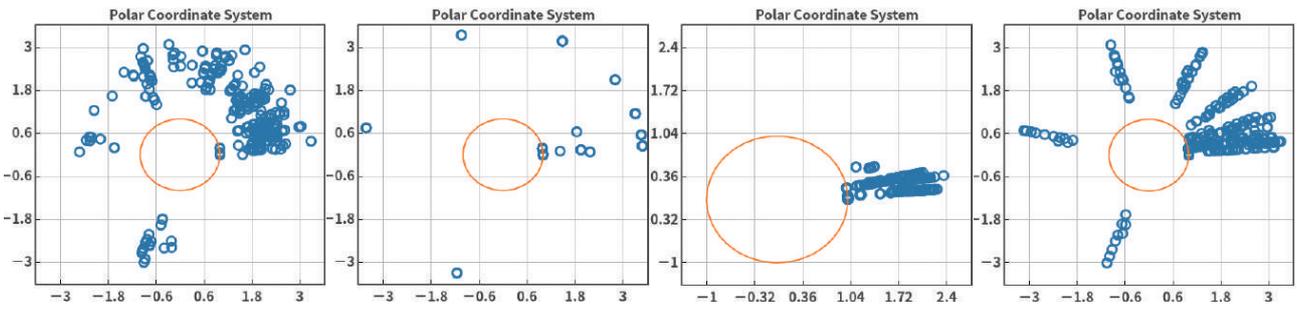


Figure 21: Solutions obtained by NSGA-III, MOEA/D, HypE and IBEA (from left to right) on 10-objective DTLZ5 by using Polar Coordinate System.

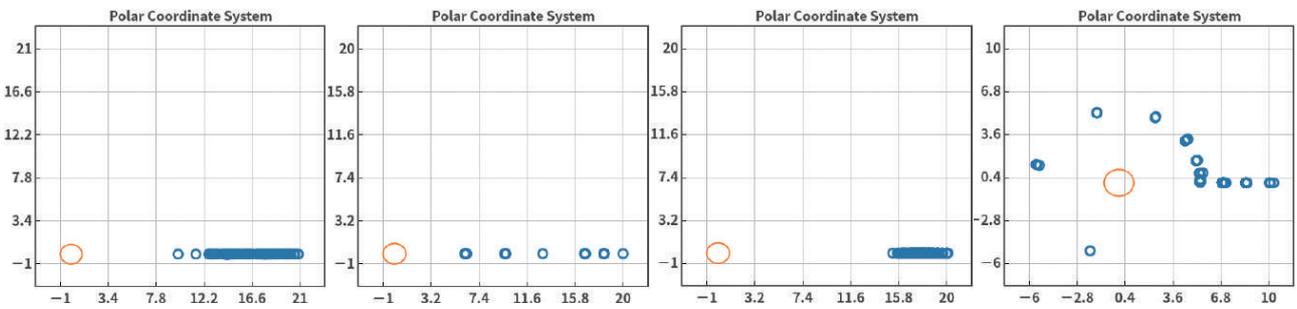


Figure 22: Solutions obtained by NSGA-III, MOEA/D, HypE and IBEA (from left to right) on 10-objective DTLZ7 by using Polar Coordinate System.

PCP as shown in Fig. 23 and Fig. 25. DTLZ5 is a degenerate test problem. As shown in Fig. 27 and Fig. 28, it is much easier to understand that f_1 and f_{10} are conflict with each other in DTLZ5 after reordering the axis in PCP and heatmap. Nevertheless, as shown in Fig. 29 and Fig. 30, both PCP and heatmap are not able to well interpret the disconnected PF segments in DTLZ7.

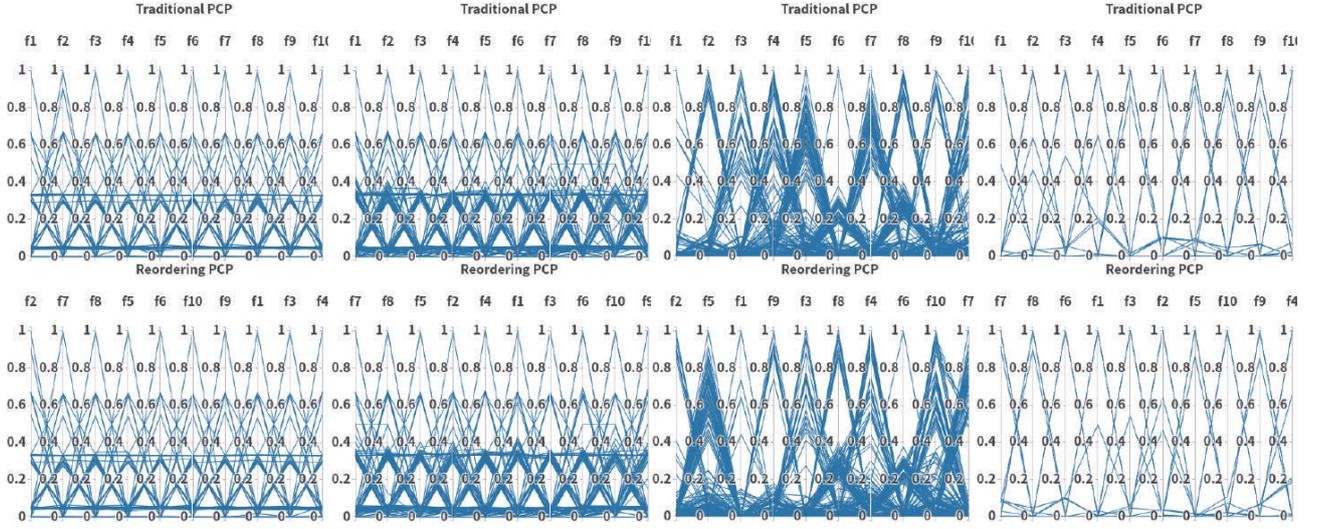


Figure 23: Solutions obtained by NSGA-III, MOEA/D, HypE and IBEA (from left to right) on 10-objective DTLZ1 by using PCP.

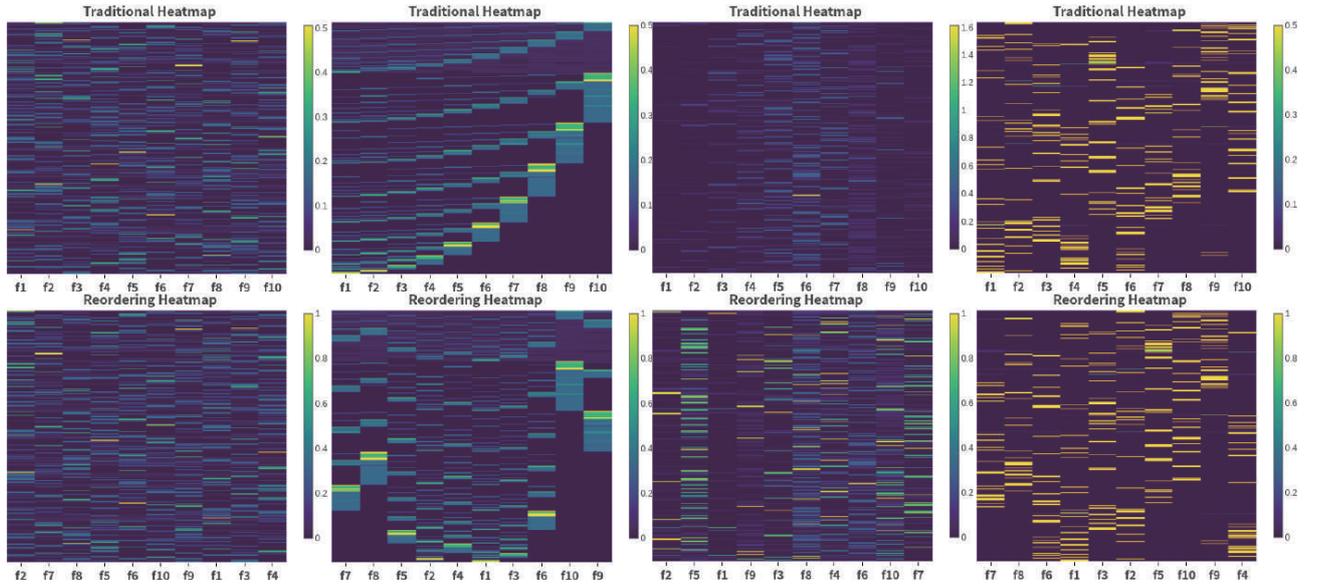


Figure 24: Solutions obtained by NSGA-III, MOEA/D, HypE and IBEA (from left to right) on 10-objective DTLZ1 by using Heatmap.

4 Conclusions and Future Directions

This paper overviews some currently prevalent visualisation techniques from the ways of how data is represented. Moreover, in order to have a holistic view of the pros and cons of different types of visualisation techniques, we conduct a series of experiments to visualise several PF approximation sets obtained by four state-of-the-art EMO algorithms on the DTLZ benchmark problems.

From the experimental results, it is interesting to find that the visualisation techniques also follow the *No-Free-Lunch* theorem [49] where no single visualisation technique is able to provide a comprehensive understanding of the characteristics of the approximation set. More specifically, we find that

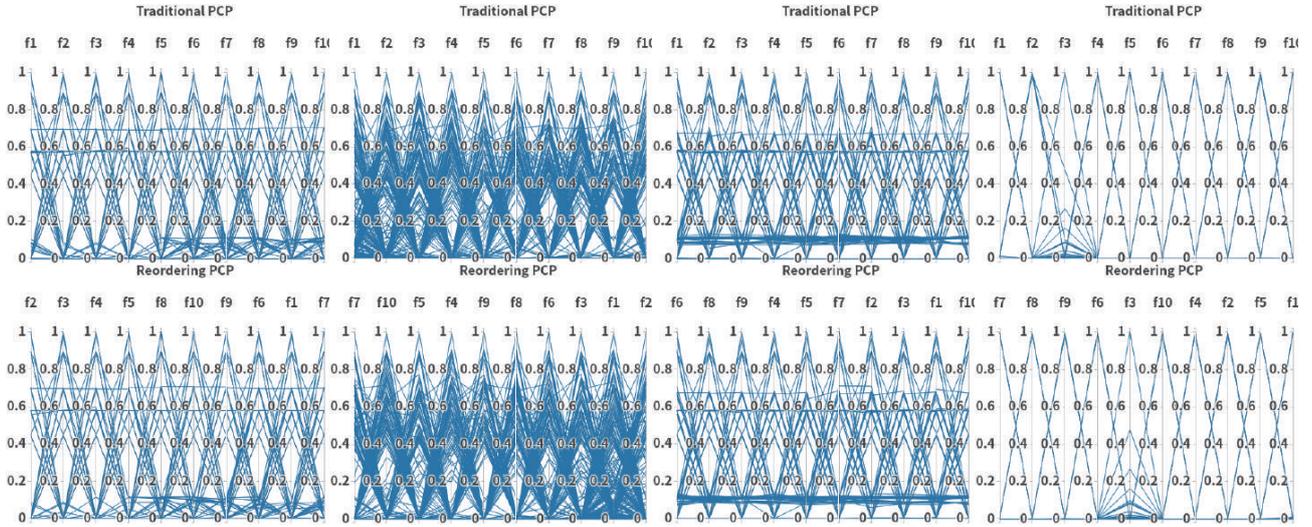


Figure 25: Solutions obtained by NSGA-III, MOEA/D, HypE and IBEA (from left to right) on 10-objective DTLZ4 by using PCP.

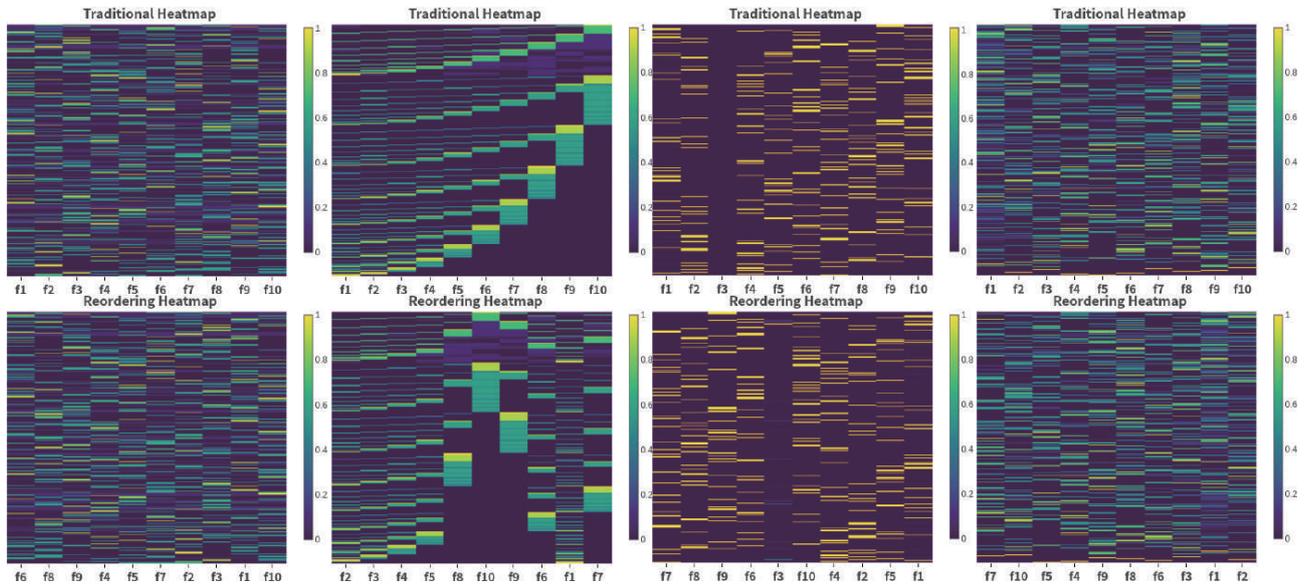


Figure 26: Solutions obtained by NSGA-III, MOEA/D, HypE and IBEA (from left to right) on 10-objective DTLZ4 by using Heatmap.

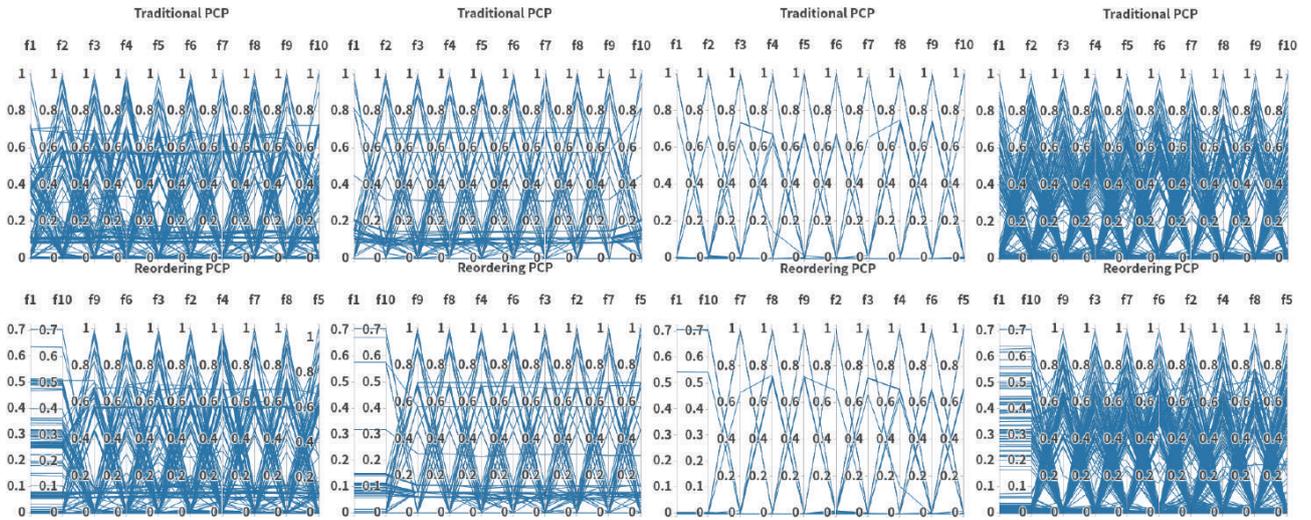


Figure 27: Solutions obtained by NSGA-III, MOEA/D, HypE and IBEA (from left to right) on 10-objective DTLZ5 by using PCP.

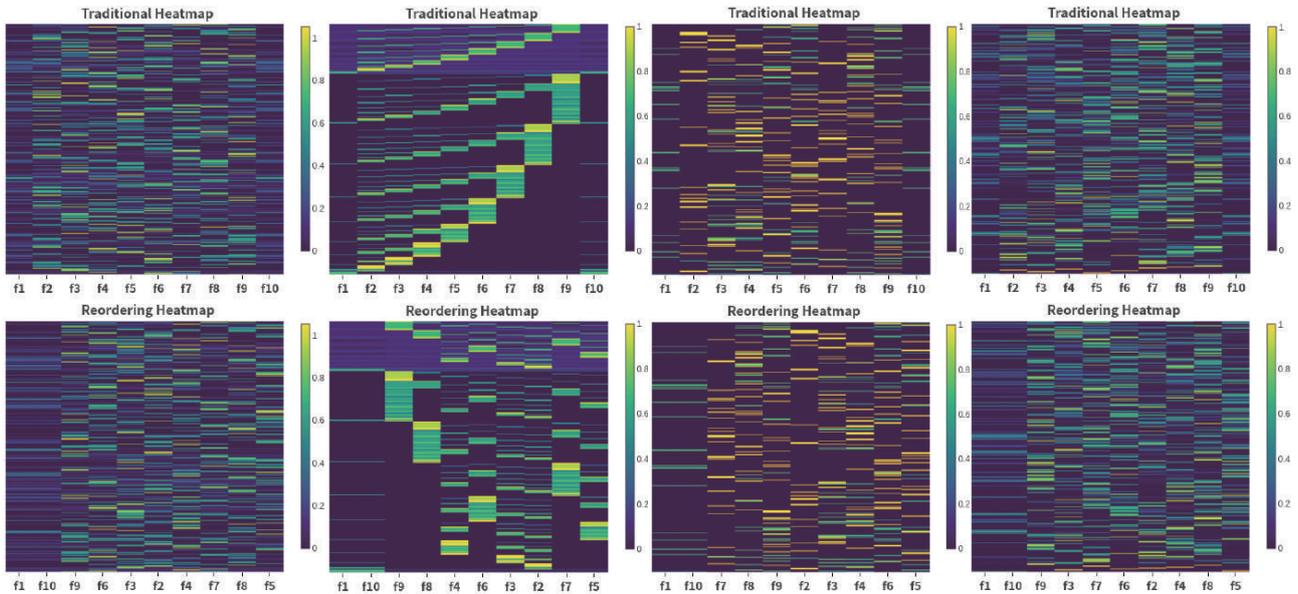


Figure 28: Solutions obtained by NSGA-III, MOEA/D, HypE and IBEA (from left to right) on 10-objective DTLZ5 by using Heatmap.

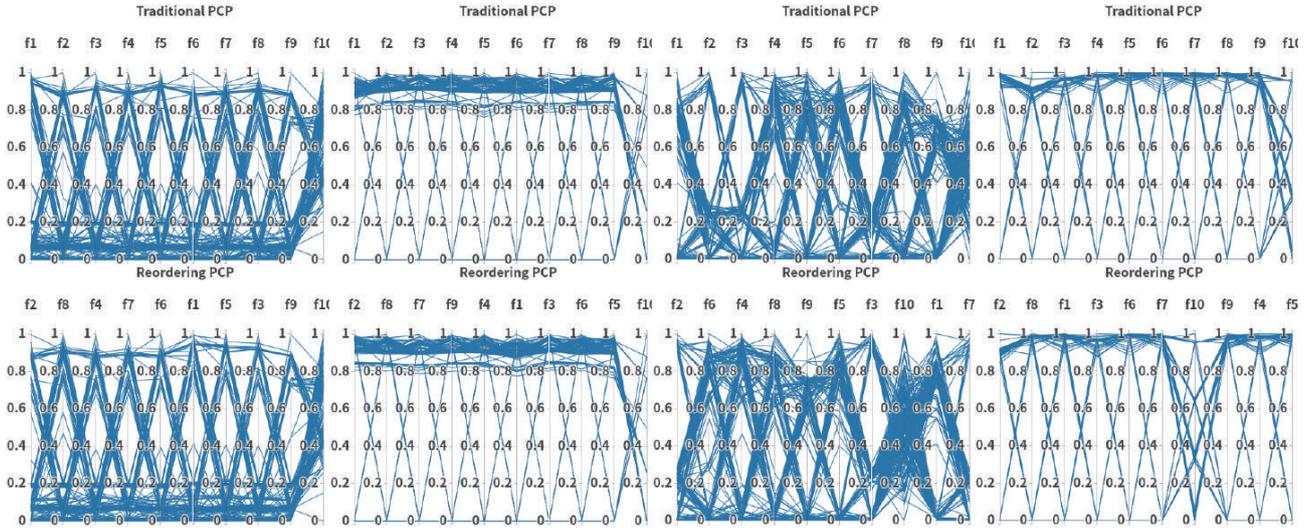


Figure 29: Solutions obtained by NSGA-III, MOEA/D, HypE and IBEA (from left to right) on 10-objective DTLZ7 by using PCP.

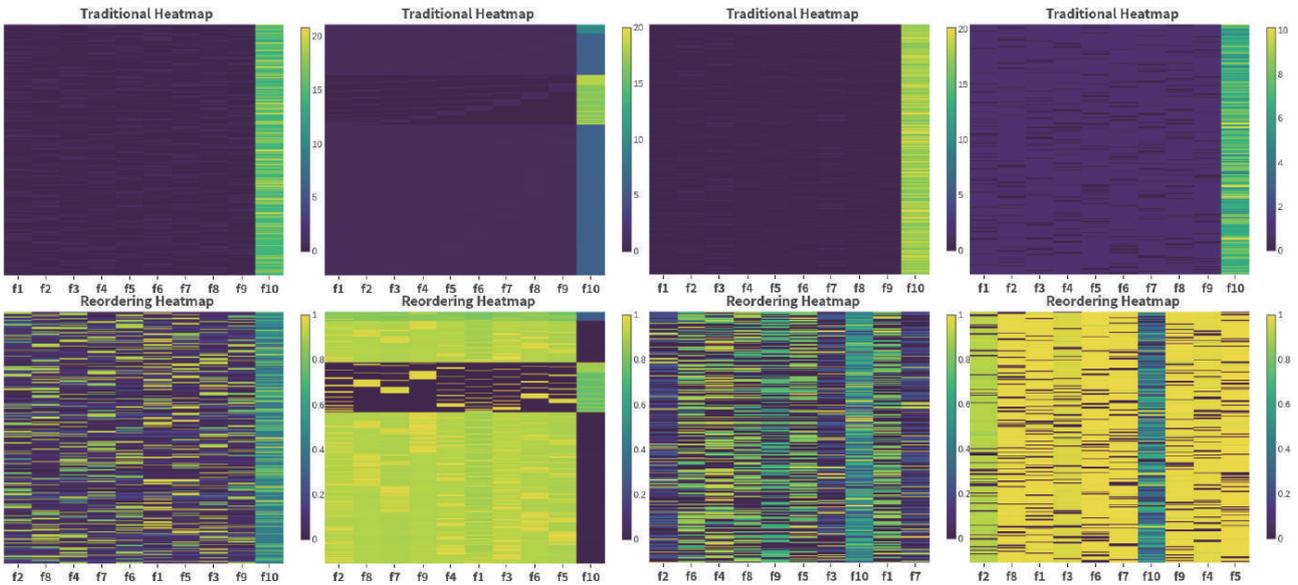


Figure 30: Solutions obtained by NSGA-III, MOEA/D, HypE and IBEA (from left to right) on 10-objective DTLZ7 by using Heatmap.

PCP and heatmap are relatively robust for interpreting the convergence and trade-off relationship among objectives. It is interesting to note that the projection has shown a promising result for understanding both the geometrical characteristics of the approximation set and its convergence. However, it is unfortunately not scalable to more than four objectives. In addition, we cannot interpret the trade-off relationship from the projection. Dimension reduction techniques like MDS has a potential to interpret the geometrical characteristics of the approximation set. However, due to the information loss, it cannot be used to investigate the convergence and trade-off relationship and it can even be misleading when the PF becomes complex. In principle, the PCS is an interesting idea but due to the requirement of prior knowledge of the convexity of the PF, it is not widely useful when dealing with real-world black-box optimisation problems.

Given these observations, as the next step, we plan to develop an interactive visualisation platform that is able to take advantages of various visualisation techniques under a unified framework. In particular, it is able to progressively interpret various characteristics of the approximation set in a hierarchical manner via an interaction process with DMs. It is also interesting to use the visualisation platform to compare and analysis the performance of more recently proposed EMO algorithms [50–52].

Acknowledgment

This work was supported by the Royal Society (Grant No. IEC/NSFC/170243).

References

- [1] C. A. C. Coello, “An updated survey of GA-based multiobjective optimization techniques,” *ACM Comput. Surv.*, vol. 32, no. 2, pp. 109–143, 2000.
- [2] A. A. Montaña, C. A. C. Coello, and E. Mezura-Montes, “Multiobjective evolutionary algorithms in aeronautical and aerospace engineering,” *IEEE Trans. Evol. Comput.*, vol. 16, no. 5, pp. 662–694, 2012.
- [3] K. Li, R. Chen, G. Fu, and X. Yao, “Two-archive evolutionary algorithm for constrained multi-objective optimization,” *IEEE Trans. Evolutionary Computation*, 2018, accepted for publication.
- [4] A. Ponsich, A. L. Jaimes, and C. A. C. Coello, “A survey on multiobjective evolutionary algorithms for the solution of the portfolio optimization problem and other finance and economics applications,” *IEEE Trans. Evol. Comput.*, vol. 17, no. 3, pp. 321–344, 2013.
- [5] K. Deb, S. Agrawal, A. Pratap, and T. Meyarivan, “A fast and elitist multiobjective genetic algorithm: NSGA-II,” *IEEE Trans. Evol. Comput.*, vol. 6, no. 2, pp. 182–197, 2002.
- [6] E. Zitzler and S. Künzli, “Indicator-based selection in multiobjective search,” in *PPSN’04: Proc. of 8th Parallel Problem Solving from Nature*, 2004, pp. 832–842.
- [7] Q. Zhang and H. Li, “MOEA/D: A multiobjective evolutionary algorithm based on decomposition,” *IEEE Trans. Evol. Comput.*, vol. 11, no. 6, pp. 712–731, 2007.
- [8] P. A. N. Bosman and D. Thierens, “The balance between proximity and diversity in multiobjective evolutionary algorithms,” *IEEE Trans. Evol. Comput.*, vol. 7, no. 2, pp. 174–188, 2003.
- [9] E. Zitzler and L. Thiele, “Multiobjective evolutionary algorithms: a comparative case study and the strength pareto approach,” *IEEE Trans. Evol. Comput.*, vol. 3, no. 4, pp. 257–271, 1999.
- [10] H. Ishibuchi, R. Imada, Y. Setoguchi, and Y. Nojima, “How to specify a reference point in Hypervolume calculation for fair performance comparison,” *Evolutionary Computation*, vol. 26, no. 3, 2018.

- [11] T. Tusar and B. Filipic, “Visualization of pareto front approximations in evolutionary multiobjective optimization: A critical review and the prosection method,” *IEEE Trans. Evol. Comput.*, vol. 19, no. 2, pp. 225–245, 2015.
- [12] B. Filipic and T. Tusar, “A taxonomy of methods for visualizing pareto front approximations,” in *GECCO’18: Proc. of 2018 Genet. Evol. Comput. Conf. (GECCO)*, 2018, pp. 649–656.
- [13] M. F. Ashby, “Multi-objective optimization in material design and selection,” *Acta materialia*, vol. 48, no. 1, pp. 359–369, 2000.
- [14] A. Inselberg, “Parallel coordinates: Visual multidimensional geometry and its applications,” in *KDIR’12: Proc. of the International Conference on Knowledge Discovery and Information Retrieval*, 2012.
- [15] A. Pryke, S. Mostaghim, and A. Nazemi, “Heatmap visualization of population based multi objective algorithms,” in *EMO’4: Proc. of 4th EMO Evolutionary Multi-Criterion Optimization*, 2006, pp. 361–375.
- [16] I. T. Jolliffe, “Principal component analysis,” in *IESS’11: Proc. of 2011 International Encyclopedia of Statistical Science*, 2011, pp. 1094–1096.
- [17] D. J. Walker, R. M. Everson, and J. E. Fieldsend, “Visualizing mutually nondominating solution sets in many-objective optimization,” *IEEE Trans. Evol. Comput.*, vol. 17, no. 2, pp. 165–184, 2013.
- [18] J. B. Tenenbaum, V. De Silva, and J. C. Langford, “A global geometric framework for nonlinear dimensionality reduction,” *science*, vol. 290, no. 5500, pp. 2319–2323, 2000.
- [19] T. Kohonen, “The self-organizing map,” *Proceedings of the IEEE*, vol. 78, no. 9, pp. 1464–1480, 1990.
- [20] A. Ibrahim, S. Rahnamayan, M. V. Martin, and K. Deb, “3d-radvis: Visualization of pareto front in many-objective optimization,” in *CEC’16: Proc. of 16th IEEE Congress on Evolutionary Computation*, 2016, pp. 736–745.
- [21] Z. He and G. G. Yen, “Visualization and performance metric in many-objective optimization,” *IEEE Trans. Evol. Comput.*, vol. 20, no. 3, pp. 386–402, 2016.
- [22] K. Deb, L. Thiele, M. Laumanns, and E. Zitzler, “Scalable test problems for evolutionary multiobjective optimization,” in *Evolutionary multiobjective optimization*. Springer, 2005, pp. 105–145.
- [23] D. F. Andrews, “Plots of high-dimensional data,” *Biometrics*, vol. 28, no. 1, pp. 125–136, 1972.
- [24] J. Heinrich and D. Weiskopf, “State of the art of parallel coordinates,” in *EG’13: Proc. of the 2013 Eurographics - State of the Art Reports*. The Eurographics Association, 2013, pp. 95–116.
- [25] L. Zhen, M. Li, R. Cheng, D. Peng, and X. Yao, “Adjusting parallel coordinates for investigating multi-objective search,” in *SEAL’17: Proc. of 11th Simulated Evolution and Learning*, 2017, pp. 224–235.
- [26] L. F. Lu, M. L. Huang, and J. Zhang, “Two axes re-ordering methods in parallel coordinates plots,” *J. Vis. Lang. Comput.*, vol. 33, pp. 3–12, 2016.
- [27] X. Yuan, P. Guo, H. Xiao, H. Zhou, and H. Qu, “Scattering points in parallel coordinates,” *IEEE Trans. Vis. Comput. Graph.*, vol. 15, no. 6, pp. 1001–1008, 2009.
- [28] B. Filipic and T. Tusar, “Visualization in multiobjective optimization,” in *GECCO’18: Proc. of 2018 Genet. Evol. Comput. Conf. (GECCO)*, 2018, pp. 858–879.

- [29] L. K. Saul and S. T. Roweis, “Think globally, fit locally: unsupervised learning of low dimensional manifolds,” *Journal of machine learning research*, vol. 4, no. Jun, pp. 119–155, 2003.
- [30] M. Belkin and P. Niyogi, “Laplacian eigenmaps and spectral techniques for embedding and clustering,” in *NIPS’01: Proc. of 14th Neural Information Processing Systems*, 2001, pp. 585–591.
- [31] J. Zhang, H. Huang, and J. Wang, “Manifold learning for visualizing and analyzing high-dimensional data,” *IEEE Intelligent Systems*, vol. 25, no. 4, pp. 54–61, 2010.
- [32] X. Blasco, J. M. Herrero, J. Sanchis, and M. Martínez, “A new graphical visualization of n -dimensional Pareto front for decision-making in multiobjective optimization,” *Inf. Sci.*, vol. 178, no. 20, pp. 3908–3924, 2008.
- [33] P.-W. Chiu and C. L. Bloebaum, “Hyper-radial visualization (HRV) method with range-based preferences for multi-objective decision making,” *Structural and Multidisciplinary Optimization*, vol. 40, no. 1-6, p. 97, 2010.
- [34] P. E. Hoffman, G. G. Grinstein, K. Marx, I. Grosse, and E. Stanley, “A survey of visualizations for high-dimensional data mining,” *Information visualization in data mining and knowledge discovery*, vol. 104, p. 4, 2002.
- [35] J. Bader and E. Zitzler, “HypE: An algorithm for fast hypervolume-based many-objective optimization,” *Evolutionary Computation*, vol. 19, no. 1, pp. 45–76, 2011.
- [36] K. Li, S. Kwong, R. Wang, J. Cao, and I. J. Rudas, “Multi-objective differential evolution with self-navigation,” in *SMC’12: Proc. of the 2012 IEEE International Conference on Systems, Man, and Cybernetics*, 2012, pp. 508–513.
- [37] K. Li, Á. Fialho, S. Kwong, and Q. Zhang, “Adaptive operator selection with bandits for a multiobjective evolutionary algorithm based on decomposition,” *IEEE Trans. Evolutionary Computation*, vol. 18, no. 1, pp. 114–130, 2014.
- [38] K. Li, Q. Zhang, S. Kwong, M. Li, and R. Wang, “Stable matching-based selection in evolutionary multiobjective optimization,” *IEEE Trans. Evolutionary Computation*, vol. 18, no. 6, pp. 909–923, 2014.
- [39] R. Chen, K. Li, and X. Yao, “Dynamic multiobjectives optimization with a changing number of objectives,” *IEEE Trans. Evolutionary Computation*, vol. 22, no. 1, pp. 157–171, 2018.
- [40] K. Li, S. Kwong, and K. Deb, “A dual-population paradigm for evolutionary multiobjective optimization,” *Inf. Sci.*, vol. 309, pp. 50–72, 2015.
- [41] K. Li, K. Deb, Q. Zhang, and Q. Zhang, “Efficient nondomination level update method for steady-state evolutionary multiobjective optimization,” *IEEE Trans. Cybernetics*, vol. 47, no. 9, pp. 2838–2849, 2017.
- [42] K. Li, S. Kwong, Q. Zhang, and K. Deb, “Interrelationship-based selection for decomposition multiobjective optimization,” *IEEE Trans. Cybernetics*, vol. 45, no. 10, pp. 2076–2088, 2015.
- [43] K. Li, K. Deb, and X. Yao, “R-metric: Evaluating the performance of preference-based evolutionary multiobjective optimization using reference points,” *IEEE Trans. Evolutionary Computation*, vol. 22, no. 6, pp. 821–835, 2018.
- [44] K. Li, R. Chen, G. Min, and X. Yao, “Integration of preferences in decomposition multiobjective optimization,” *IEEE Trans. Cybernetics*, vol. 48, no. 12, pp. 3359–3370, 2018.
- [45] M. Wu, K. Li, S. Kwong, and Q. Zhang, “Evolutionary many-objective optimization based on adversarial decomposition,” *IEEE Trans. Cybernetics*, 2018, accepted for publication.

- [46] K. Li, K. Deb, Q. Zhang, and S. Kwong, “An evolutionary many-objective optimization algorithm based on dominance and decomposition,” *IEEE Trans. Evolutionary Computation*, vol. 19, no. 5, pp. 694–716, 2015.
- [47] H. Ishibuchi, Y. Setoguchi, H. Masuda, and Y. Nojima, “Performance of decomposition-based many-objective algorithms strongly depends on pareto front shapes,” *IEEE Trans. Evolutionary Computation*, vol. 21, no. 2, pp. 169–190, 2017.
- [48] M. Wu, K. Li, S. Kwong, Q. Zhang, and J. Zhang, “Learning to decompose: A paradigm for decomposition-based multi-objective optimization,” *IEEE Trans. Evolutionary Computation*, 2018, accepted for publication.
- [49] D. H. Wolpert and W. G. Macready, “No free lunch theorems for optimization,” *IEEE Trans. Evolutionary Computation*, vol. 1, no. 1, pp. 67–82, 1997.
- [50] K. Li, Á. Fialho, and S. Kwong, “Multi-objective differential evolution with adaptive control of parameters and operators,” in *LION5: Proc. of the 5th International Conference on Learning and Intelligent Optimization*, 2011, pp. 473–487.
- [51] K. Li, S. Kwong, R. Wang, K. Tang, and K. Man, “Learning paradigm based on jumping genes: A general framework for enhancing exploration in evolutionary multiobjective optimization,” *Inf. Sci.*, vol. 226, pp. 1–22, 2013.
- [52] K. Li and S. Kwong, “A general framework for evolutionary multiobjective optimization via manifold learning,” *Neurocomputing*, vol. 146, pp. 65–74, 2014.