Proceedings Paper:
Wu, K.E. and Panoutsos, G. orcid.org/0000-0002-7395-8418 (2021) A visualisation method for Pareto Front approximations in many-objective optimisation. In: Proceedings of 2021 IEEE Congress on Evolutionary Computation (CEC). 2021 IEEE Congress on Evolutionary Computation (CEC), 28 Jun - 01 Jul 2021, Kraków, Poland (virtual conference). IEEE , pp. 1929-1937. ISBN 9781728183947

https://doi.org/10.1109/cec45853.2021.9504904

A Visualisation Method for Pareto Front Approximations in Many-objective Optimisation

Kai Eivind Wu
Department of ACSE
University of Sheffield
Sheffield, UK
kewu1@sheffield.ac.uk

George Panoutsos
Department of ACSE
University of Sheffield
Sheffield, UK
g.panoutsos@sheffield.ac.uk

***Abstract* — Visualisation of Pareto Front (PF) approximations of many-objective optimisation problems (MaOP) is critical in understanding and solving a MaOP. Research is ongoing on developing effective visualisation methods with desired properties, such as simultaneously revealing dominance relations, PF shape, and the diversity of approximations. State-of-the-art visualisation methods in the literature often retain some of the preferred properties, but there are still shortfalls to address others. A new visualisation method is proposed in this paper, which covers the majority of the desired properties for visualisation methods. The proposed method is based on displaying PF approximations via projections on a reference vector versus distances to the same reference vector. The reference vector is created using nominal Ideal and Nadir points of existing nondominated PF approximation sets. MaF benchmark problems are used to demonstrate the effectiveness; results show that the proposed method exhibits a more balanced performance than the state-of-the-art in capturing desired visualisation properties.**

Keywords — Many-objective Optimisation, Performance indicator, Diversity, Reference vectors, Benchmark testing

I. INTRODUCTION

Approximation sets and the approximated Pareto Front (PF) landscape of many-objective optimisation problems (MaOP) consist of vectors in an m dimensional objective space where m is the number of objective functions. It is nontrivial to visualise these quantities mainly because the number of axes in such a visualisation exceeds three, and a large amount of data – usually complex – is needed to be displayed simultaneously. However, visualisation of the approximations is crucial in optimisation research [1]. Effective visualisations of approximations may assist the decision-making process and support the work of interactively searching for optimised solutions. It can also be used to examine and improve the performance of optimisation algorithms as visual comparisons may convey important qualitative information, for example the dominance relationships of different approximations, the process of convergence towards the resulting PF as well as the diversity of the MaOP solutions. The graphical display of the approximated PF landscape is also helpful in identifying characteristics and challenges of the optimisation problem itself, such as distance of approximations to the constraint boundaries and local minima [1-2].

Visualisation of high dimensional data sets often involves compressing and mapping information into two- or three-dimensional spaces that can be displayed readily with traditional means of figures or charts, where properties of point data are easier to be analysed and understood. Although such mappings aim to maintain high dimensional properties as much as possible, some information is inevitably lost during the mapping process. Hence, an effective visualisation process consists of data compression and extraction of specific critical properties from the high-dimensional dataset, while omitting less essential features [1].

Effective visualisation methods should have the following desired properties, as outlined in [2]:

1) Dominance relation: shall preserve dominance relation between solutions by visualisation,
2) PF shape: should be able to display the PF shape.
3) Objective range: should reveal the ranges of objectives.
4) PF distribution: should exhibit the distribution of solutions.
5) Robustness: should maintain robustness when mapping data of high dimensions to 2D or 3D space
6) Handling large sets: it may present large approximation sets.
7) Handling multiple sets: it may simultaneously visualise multiple approximation sets.
8) Scalability: it should be scalable to any number of objectives
9) Simplicity: it should be simple to construct.

An additional desired property is introduced in this paper:

10) Uniqueness: the visualised image should be unique independent of the sequence of objective functions utilised in the display.

While they may be effective in their respective targeted desired properties, current visualisation methods possess only some of, and limitedly, the above-desired capabilities [2-5]. He and Yen [6] sort existing visualisation methodologies into five major groups. One representative method, out of each group, is chosen and tested for representing three commonly accepted PF patterns for visualisation approaches, i.e., simplex plane, sphere, and knee shape. Numerical results reveal that none of the five methods can satisfactorily display the expected basic patterns [6]. Specifically, the major weakness of the current state of the art methods is that they only retain some of the above-mentioned desired properties. Parallel Coordinates [7], for example, fulfils the demands for displaying Dominance Relation, Robustness, Scalability and Simplicity, but is not addressing the rest of the properties. Hence, research work is needed to expand the capability of methods that possess as many of the preferred properties simultaneously as possible.

This paper proposes a new visualisation method, which targets most of the desired properties for visualising high dimensional MaOP approximations. It is achieved via a

978-1-7281-8393-0/21/$31.00 ©2021 IEEE

Projection of solution vectors versus the solution vectors' **D**istance to a reference vector in the objective space - *ProD*. The reference vector is a vector that would link a nominal Ideal point and Nadir point. The actual Ideal point and Nadir point of a PF are nontrivial to be found; instead, nondominated approximations are used to find nominal Ideal and Nadir points.

The rest of the work is organised as follows. *Section II* includes a brief literature survey on relevant, up-to-date methods of visualisations in MaOP. *Section III* covers the proposed formulation and reasoning for ProD. *Section IV* includes testing and simulation results showing the effectiveness of ProD. The last two sections cover the discussion, *Section V*, and conclusions, *Section VI*.

II. Existing visualisation methods

It is nontrivial to develop a visualisation method for MaOP, which satisfies all the requirements listed in Section I simultaneously. Compromises are made in many cases. Various visualisation methods exist, for displaying dominance relations of approximations of MaOPs; a review on the topic can be found in [5] and [7].

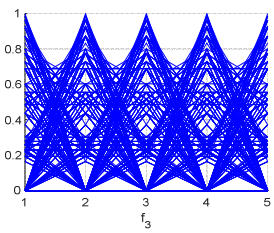
Filipič and Tušar [2] suggest a taxonomy of visualisation methodologies, consisting of two major categories: displaying a simple PF approximation set and showing repeated approximation sets.

Fig. 1 shows the results of a PF of a unit cube using different visualisation methods. Scatter plot [2]: A straightforward visualisation method frequently adopted where all vectors of approximations of non-dominated data points are projected to a 2D display by omitting higher dimensions of the vectors other than the two displayed. Similar plots are generated for all combinations of two objective functions, and as a result, a scatter plot matrix is formed. Although the method is simple, robust, and able to reveal the objective range and handle multiple sets, it is not scalable to high dimensional MaOPs [2]. Neither can it depict dominance relations of approximation sets. PF shape and its distribution are only shown to a limited degree [2].

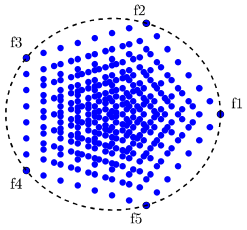
Parallel coordinates [7]: see Fig. 1 (a). High dimensional vectors are mapped onto a 2D figure using *m* equally spaced parallel axes in this method. The vectors are drawn as polylines through position on each axis corresponding to each component of the vector. Parallel coordinates are simple to construct and scalable to any dimensions of MaOP. No information is lost in the mapping process. The main weaknesses are three-fold: the method fails to display the shape of the PF approximation and possesses limited capability in handling large amounts of approximation data and simultaneous visualisation of several approximations [2]. Radar Chart [8] can be considered a further developed version of parallel coordinates with similar strengths and weaknesses, in which axes are placed radially. Radial Coordinate Visualisation (RadViz) [9]: see Fig. 1 (b). The *m*-dimensional vectors of approximations are mapped onto a two-dimensional plane by uniformly placing each m-dimensional vector's origin along a circle as various anchor points. Each approximation set is expressed by assuming that each anchor *i* is connected to a spring of force proportional to objective function f_i. An m-dimensional vector is identified inside the circle. The positions of the vectors are found and displayed on the two-dimensional plot. The method is simple and robust, may handle several sets of approximations simultaneously, and can be readily extended to any dimensions. However, RadViz fails in revealing the pattern of the PF front and the dominance relations between solution sets [2]. Multiple disparate vectors may share the same equilibrium position in the plot creating chaotic and unforeseeable patterns or data distribution. In contrast, the two vectors are neither neighbours nor belonging to any natural groups of vectors. RadViz has been further developed into a 3D version [10]. The above-stated shortcoming is also inherited in the 3D version of the method. PaletteViz: [3] High-dimensional and non-dominated objective vectors are mapped onto multiple two-dimensional Radviz plots in which vectors are sorted after their boundary to core location in their original high-dimensional space. As is the case for RadViz, the main weakness is that multiple disparate vectors may share the same equilibrium position, causing difficulties in interpreting data distribution.

2D Polar coordinates [11]: see Fig. 1 (c). The objective space is divided into subregions using reference vectors, which are evenly mapped to a 2D plane following the sequence of generated reference vectors. The best approximation from each subregion is chosen and plotted on the reference vector in a 2D map. The method can reveal a PF's basic patterns while it displays more complex PF shapes with reduced success [11]. Moreover, neighbour points on PF are disparately located in the display, and the outcome is dependent on the numbering of objectives, causing a non-unique result. Level diagrams [12]: see Fig. 1 (d). Euclidean distance of approximations to the ideal point is displayed as functional values of each objective. The main drawback here is that the method must utilise an equal number of separate plots like the number of objective functions. The number of figures for visualising MaOP approximations can be overwhelming. Pryke et al. [13] use a Heatmap visualisation where a tabulation of colour chart is formed. Objectives are taken in columns, and high dimensional vectors are taken in rows. Although the method is robust and scalable, it fails in showing PF shape, objective range, and distribution. Large or multiple data sets can hardly be displayed, as discussed in [2]. Yamamoto et al. [14] suggest Principal Component Analysis reduce the number of objectives necessary to be visualised. The challenge is that number of objectives after reduction may still exceed three, which makes visualisation difficult [14]. Lotov et al. [15] propose a visualisation method called "interactive decision maps" to visualise Pareto front approximations up to four or five objectives. With more cognitive effort, it is claimed that up to eight objectives can be revealed visually.

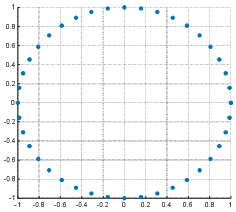
Other visualisation methods also exist. Freitas et al. [16] propose the Aggregation Trees approach where positively correlated objective functions are merged, and the total number of objective functions is thus reduced. Chiu and Bloebaum [17] introduce Hyper-radial visualisation (HRV) method. The mapping from high dimensions to 2D is done by expressing 2D vectors in Hyper-radial distance from the original vector to the ideal point. Koochaksaraei et al. [18] suggest a Chord diagram, where objective functions with their respective equal arc length on which scales are indicated and arcs are placed along with the circle's peripherals. Agrawal et al. [19] suggest a Hyperspace Diagonal Counting method to compress and group high dimensional vectors into 3D space mapping. T. Kohonen



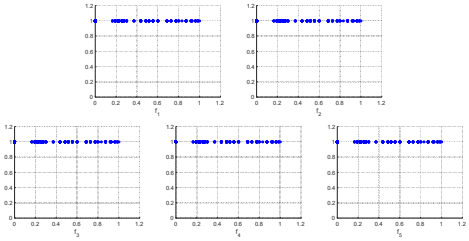
(a) Parallel coordinate



(b) RadViz.



(c) Polar coordinates



(d) Level diagram

Fig. 1. Visualisation of the true Pareto front of 5-D DTLZ2 using different visualisation methods. (a) Parallel coordinates. (b) RadViz. (c) Polar coordinates (f) level diagram.

[20] and S. Obayashi and D. Sasaki [21] utilise Self-Organised Maps to visualise solutions with many objectives, using trained neural networks to find nearby solutions. Yoshimi et al. [22] have further developed the methodology to present the result in a spherical form which improves the display of boundary points. Hence, the above methods target specific desired properties only (hence not cover all the desired properties described in Section I).

III. Proposed Visualisation Method – ProD

In this section, a new visualisation method, ProD, is proposed to identify PF patterns and monitor dominance relations as well as evaluate the diversity of approximations for MaOPs.

It is proposed that high dimensional vectors of PF approximations of a MaOP in objective function space are evaluated and displayed in their projections on and distances to a reference vector in the objective space. It is created by linking the nominal ideal point and Nadir point calculated based on existing nondominated approximations. The nominal ideal point is an auxiliary point with the least objective functions among current nondominated dataset(s) as coordinates. In contrast, the nominal Nadir point, consists of coordinates of the largest of objectives. Fig. 2 shows a schematic view in 2D space on how projections on, and distance to, the reference vector of a PF approximation vector are defined. The projection $r_{\parallel}$ is formulated as follows:

$$r_{\parallel} = [f_1\, f_2\, \dots\, f_m] * \frac{\overrightarrow{Reference\ vector}}{\left|\overrightarrow{Reference\ vector}\right|} \qquad (1)$$

in which $f_i, i \in (1, m)$ are objective functions, m is the number of objective functions, and the reference vector is defined as:

$$\overrightarrow{Reference\ vector} = \overrightarrow{Nadir\ point} - \overrightarrow{Ideal\ point} \qquad (2)$$

The distance to the reference vector is expressed as:

$$r_{\perp} = \sin(\theta)\sqrt{f_1^2 + f_2^2 + \dots + f_m^2} \qquad (3)$$

where angle θ between candidate vector and reference vector is calculated as:

$$\theta = cos^{-1}\left(\frac{[f_1\, f_2\, \dots\, f_m] * \overrightarrow{Reference\ vector}}{\sqrt{f_1^2 + f_2^2 + \dots + f_m^2} * \left|\overrightarrow{Reference\ vector}\right|}\right) \qquad (4)$$

From this point onwards, it is assumed that the origin is moved to the nominal Ideal Point. The results are displayed in a two-dimensional plot named ProD (Projection on a reference vector versus Distance to the reference vector).

As mentioned in the Introduction section, visualisation in high dimensional space compresses information and extracts specific fundamental properties from the dataset while omitting less critical details. The core idea of ProD is to compress all data of the same $r_{\parallel}$ and $r_{\perp}$ values into a single data point. In the 3D case, a data point in ProD contains or represents all raw data points on the PF surface that form a ring, which has an equal distance to the reference vector, the plane of which is normal to the reference vector; this is illustrated in Fig. 3. In high dimensions, such data points being compressed are located on a hyper ring.

Compared with definitions from the taxonomy of visualisation methodologies [2], ProD is classified into the category of visualisation of repeated approximation sets, which means it is a visualisation method for PF pattern recognition, convergence, and diversity monitoring.

It should be noted that the idea of expressing candidate solutions in the form of projections and distances to a reference vector in the objective space is frequently used in the formulation of MaOP algorithms, such as MOEA/D [23], however , to the best of our knowledge, this has not been used for compressing high dimensional data for visualisation purposes.

The normalisation of objective functions is performed prior to establishing ProD. Hence, each objective is assured to be in comparable range by employing coordinate values of nominal Nadir and Ideal points. Normalisation leads to a change of the final PF shape, and it sometimes creates more complicated ones. See more detailed discussions in Section IV. In practice, ProD can be used based on both normalised and non-normalised data, where the user can decide on the

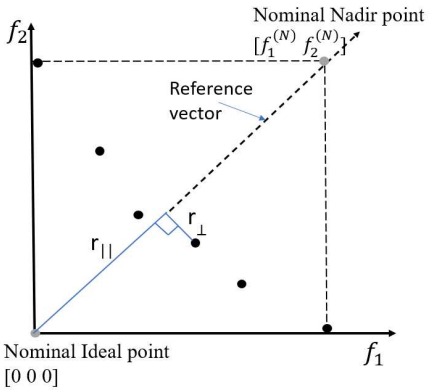


Fig. 2. A schematic view: Projection on ($r_{\|}$) and Distance to ($r_{\perp}$) reference vector of a data point in 2D space.

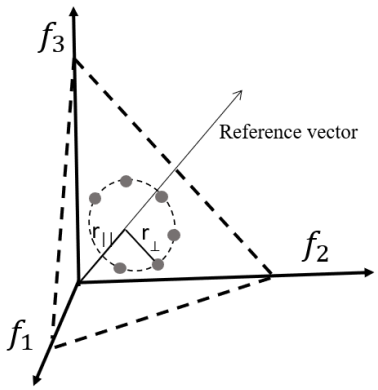


Fig. 3. A schematic view: Compression of data in ProD in 3D space.

approach to use (for example, simpler patters may reveal dominance relations better).

The primary benefit of the proposed visualisation approach is that it possesses in a balanced manner most of the desired properties of a visualisation method highlighted in Section I. ProD creates a unique image of PF approximations; the objective range is also visualisable. Moreover, it can handle large as well as multiple approximation sets. In the next section, ProD's ability to visualise PF shape and dominance relations between approximation sets is demonstrated.

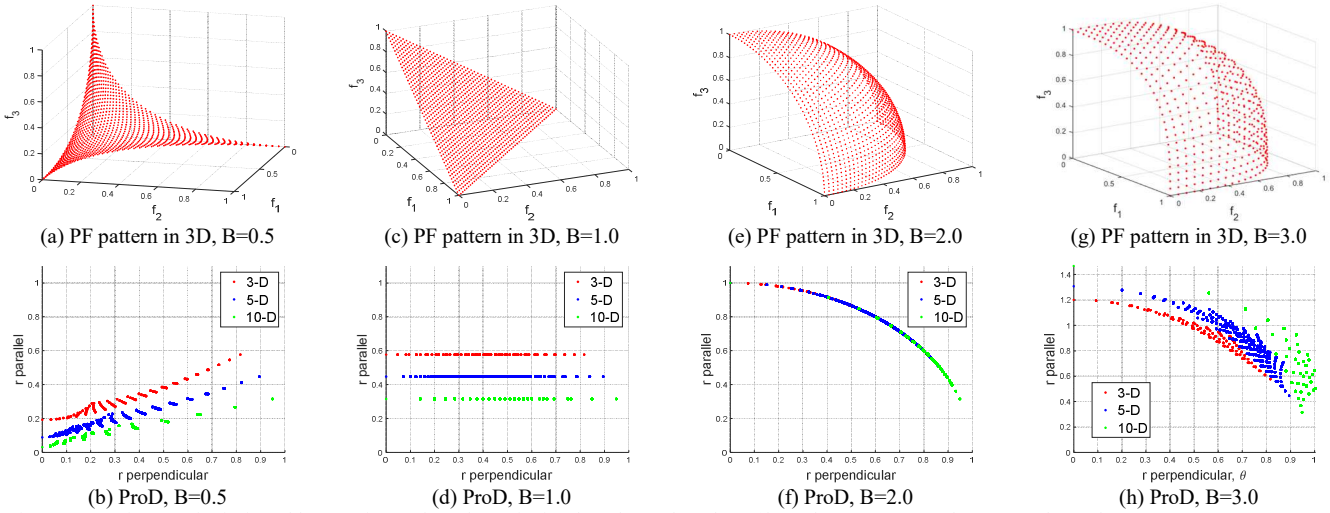
IV. Numerical Results

This section is divided into three main parts. First, three representative B-norm surfaces in 3D, 5D and 10D with corresponding B parameter values are visualised in ProD, imitating Pareto front of convex form with knee point, hyperplane, and concave shape; this will demonstrate the ability of ProD in revealing basic patterns of high dimensional PFs (B-norm functions are often named as F-norm functions in mathematics, see. Eq. 5. The functions are termed as B-norm functions in this paper to avoid confusion with the objective function vector F). Secondly, true PFs of thirteen scalable benchmarks of MaF 1-7 and 10-13 [24], used in the CEC'2017 competition on Evolutionary Many-Objective Optimisation [24], are selected and visualised by ProD. MaF Benchmark 8 and 9 are omitted here because their shapes are readily visualised with the two decision variables chosen [24]. MaF 14 and 15 are also omitted, mainly because their PF shapes are identical to those of MaF 1 and 4. Also, the purpose of their primary use is on studying large-scale problems [24], which is not the subject of this paper. Special attention is paid to the displaying effectiveness of ProD in showing PFs of irregular shape, partial coverage, and degenerate form. Finally, ProD is utilised to display approximations of NSGA III [1] between iterations, which captures the process of convergence, hence dominance relations. ProD is also used to evaluate the quality of PF approximations of MaF benchmark problems analysed using three well known MaOP algorithms: NSGA III, GrEA [25] and IBEA [26]. The convergence and diversity of the approximations are assessed. Due to space limitations, only the most characteristic MaF benchmark cases are shown in this paper.

A. PF form of B-norm types visualised in ProD showing its ability in revealing basic PF patterns

A prerequisite to a visualisation method's capability is to display three basic surfaces of MaOP patterns of any dimensions, i.e., knee, linear and sphere [1], [3], which in the current work are expressed as B-norm surfaces with different B parameter values. See Fig. 4(a)(c)(e) for illustration.

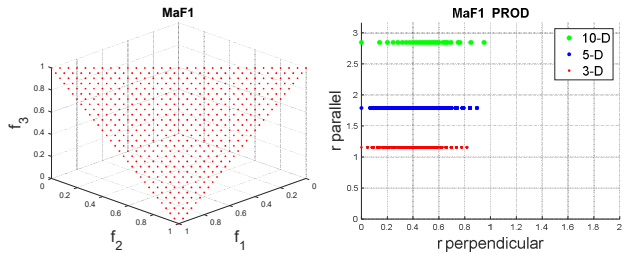
B-norm functions create a family of symmetric surfaces in high dimensional space, which are governed by Eq. 5,



(a) PF pattern in 3D, B=0.5 (c) PF pattern in 3D, B=1.0 (e) PF pattern in 3D, B=2.0 (g) PF pattern in 3D, B=3.0

(b) ProD, B=0.5 (d) ProD, B=1.0 (f) ProD, B=2.0 (h) ProD, B=3.0

Fig. 4. ProD images for imitated knee point surface (b), unit simplex plane (d), sphere (f), and stronger curved concave shape (h).



(a) PF of MaF1 in 3D (b) ProD for 3D, 5D and 10D

Fig. 5. PF displayed in ProD for Benchmark-MaF1.

$$\left(f_1^B + f_2^B + \cdots + f_m^B\right)^{1/B} = 1 \qquad (5)$$

in which B is a parameter determining the curvature.

In [27], Bn-PFt is used to generate more equally spaced reference points on B-norm surfaces in high dimensions is introduced, which is utilised in the creation of actual Pareto fronts in this paper.

Fig. 4(b), (d), and (f) display the capability and consistency of ProD to represent the three basic types of PF surfaces in three B parameter values corresponding to knee, plane, and sphere shapes in 3D 5D and 10D problems. The number of candidates covering the PF shown in Fig. 4 is listed in Table 1.

TABLE 1 NUMBER OF CANDIDATES GENERATED ON PARETO FRONT.

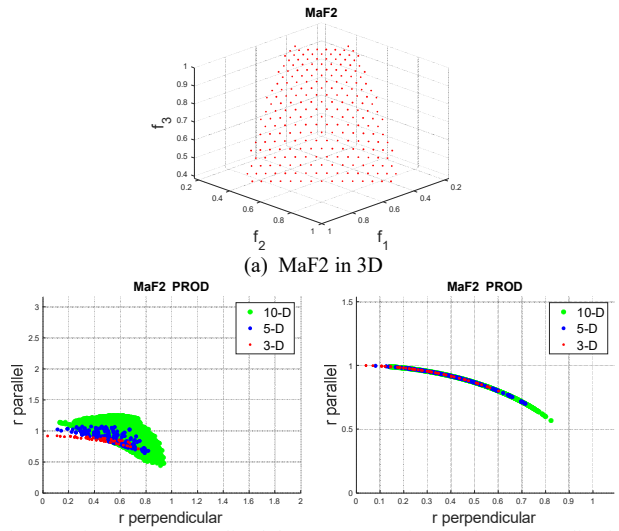
Number of objectives	3	5	10
Number of solutions used to generate authentic Pareto Front	496	10626	92376

In ProD, convex B-norm surfaces (B<1.0) in any dimensions bend upwards with some folded shape that resemble "clouds" on parts of the curves. See Fig. 4(a), (b). This is because the distribution of data is not fully symmetric about the reference vector. The folding degree reduces gradually, and the "cloud" disappears when B goes towards 1.0, which corresponds to the unit simplex plane. Unit simplex planes (B=1.0) in any dimensions appear as distinct horizontal line segments, where data is distributed symmetrically about the reference vector. See Fig. 4(c), (d). For the B-norm surface of concave sort (B>1.0), the curves bend downwards, and the form of curves is folded except for hyperspheres (B=2.0) which appear as distinct circular arcs bending downwards. See Figs. 4 (e), (f).

When B<2.0, higher dimensional curves locate lower in ProD, while for B>2.0, the opposite occur. See Fig. 4(g) (h). For B=2.0, curves of all dimensions fall together in distinct circular arcs. These observations are explained as follows: The mid-point of the B-norm surface of any B value has the property of $f_1 = f_2 = \cdots = f_m$ due to symmetry. When setting these equalities into Eq. 5, we have $f_1 = f_2 = \cdots = f_m = m^{-1/B}$. The distance of the midpoint to the Ideal point is:

$$r_m = \left(f_1^2 + f_2^2 + \cdots + f_m^2\right)^{1/2} = m^{\frac{1}{2}\left(1-\frac{2}{B}\right)} \qquad (6)$$

Eq. 6 indicates that midpoint coordinates on the B-norm surface decrease with the increasing number of objectives for B<2.0 and increases with the rising number of objectives for B>2.0 while falling together for B=2.0. See Fig. 4 for comparison.



(a) MaF2 in 3D

(b) ProD images in normalized data. (c) ProD images in unnormalized data.

Fig. 6. ProD for Benchmark-MaF2.

The results of PFs in 10D show limited number of data points for low $r_\perp$ values, which is due to the limited number of available intermediate reference points adopted in PF representation [1]. With 92376 reference points in total, only one intermediate point is generated when using the Das and Dennis method [28].

A comparable visualisation method to ProD is 2D Polar coordinates which displays PF shape of convex type as a rhombus shape with bent edges while of concave type as ellipses [6]. Since real-life PFs may have a mixture of several primary forms, e.g., PF of MaF7, Polar coordinates may create an indistinct pattern visually.

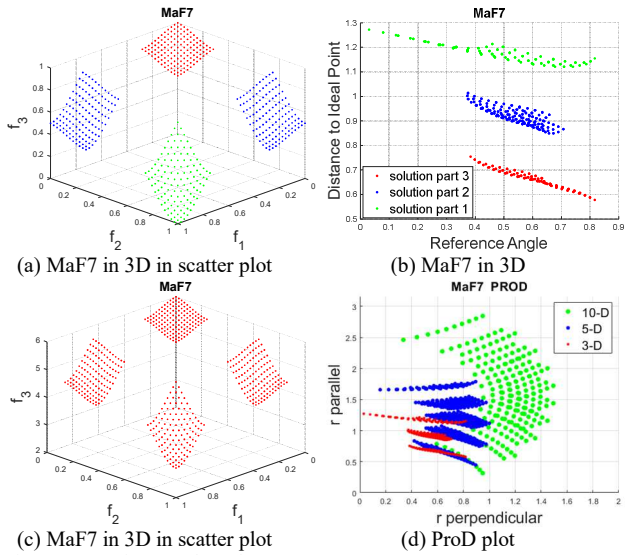
B. PF patterns of irregular shape, of partial coverage and in a degenerate form in objective space

True PFs of MaF benchmark problems, when presented in ProD, show that the visualisation method can convert PF of irregular shape, partial coverage, and degenerate form in high dimensional objective space into simple patterns in two-dimensional space.

Fig. 5 shows both the scatter plot for PF of 3D MaF1 (Fig. 5(a)) and the results of ProD of true PF of 3D (red), 5D (blue) and 10D (green) (see Fig. 5(b)). Three important points are discussed here in more detail. First, all three PFs consist of distinct horizontal line segments, indicating that MaF1 has a PF pattern of a simplex plane that is symmetrically distributed about the reference vector. Second, all three PF patterns are shown to cover only partially the objective space. Taking the 3D case as an example: the maximum $r_\perp$ value read in ProD is 0.82. If the PF covers the objective space completely, $r_\perp$ value would have been 1.633, which implies that the PF of MaF1 covers only partially the objective space. This maximal $r_\perp$ value is obtained by considering the equation of simplex plane of the general form:

$$f_1 + f_2 + f_3 = a_m \qquad (7)$$

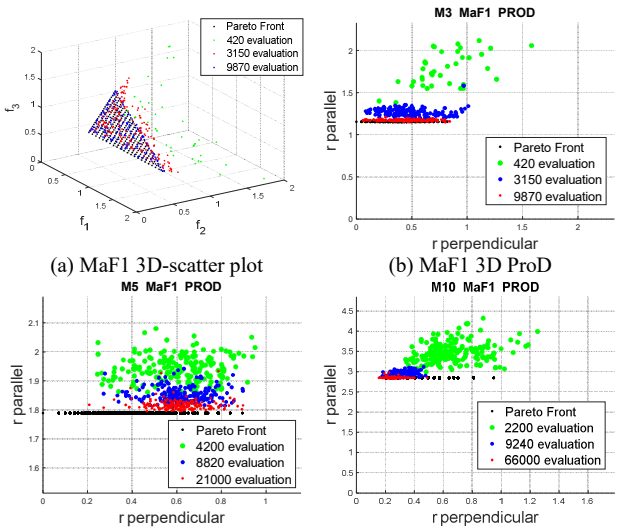
For $f_1 = f_2 = 1.0$ and $f_3 = 0$, $a_m = 2.0$, resulting $f_{1,\max} = 2.0$ (when $f_2 = f_3 = 0$). The angle θ between reference vector and f_1 axis is given by Eq. 8, where the unit



(a) MaF7 in 3D in scatter plot (b) MaF7 in 3D

(c) MaF7 in 3D in scatter plot (d) ProD plot

Fig. 7. ProD for Benchmark-MaF7



(a) MaF1 3D-scatter plot (b) MaF1 3D ProD

(c) MaF1 5D ProD (d) MaF1 10D ProD

Fig. 8. ProD showing the convergence process of Benchmark MaF1.

vector in f_1 direction is [1 0 0] and reference vector is [1 1 1]:

$$\theta = cos^{-1}\left(\frac{[1\,0\,0] * [1\,1\,1]}{\sqrt{1^2} * \sqrt{1^2+1^2+1^2}}\right) = 54.7° \qquad (8)$$

$$r_{\perp,max} = \sin(\theta) * f_{1,\max} = 1.663$$

Similar analysis can be done for any number of dimensions. The implication is that ProD may reveal the range of objectives and detect whether PF covers partially or fully the objective space. Third, curves locate higher up as the number of objective functions increases in contrast to what is the case with a unit simplex plane, indicating distances of solutions to ideal points increase in higher dimensions for this benchmark. Hence, a_m in Eq. 7 increases more rapidly than the impact of the increasing number of objectives (From Eqs. 6 and 7, we have $a_m/\sqrt{\mathrm{m}} > 1$).

PF of MaF2 turns out to be a partial sphere. See Fig. 6(a). When PF data are non-normalised, all three PF solutions in ProD are in pure circular arcs falling together, revealing that MaF2 are hyperspheres, see Fig. 6(c). With normalised data, PF patterns are more complicated, see Fig. 6(b). In this case, normalisation alters the PF shape.

PF of MaF7 is discontinuous and has four 'flakes' in 3D scatter plot (see Fig. 7(a)), but its image in ProD reveals that it consists of three groups of surfaces, see Fig 7 (b). The first piece locates closest to $f_1 - f_2$ plane, the third one with the highest f_3 values while the second one has two flakes situating in between, which are nearly the same surface. ProD of Fig. 7(b) also discloses that each data group appears as a thicker line indicating near symmetry about the reference vector, but not perfect. It is worth noting that ProDs of 5 and 10 objectives unveil five and more clusters of solutions, which means that PF patterns in higher objective space in this benchmark might have more clusters than observed in 3D space. See Fig. 7 (d).

C. *Dominance relations between candidate solutions and diversity of approximations visualised in ProD*

Benchmarks of MaF are analysed at various given amounts of iterations as well as the final stage of convergence using NSGA III, GrEA and IBEA. The results at various iterations and at the final convergence stage are displayed in ProD and compared in convergence and diversity. ProD reveals a necessary, but insufficient on its own, condition for a good diversity of approximations, i.e., the whole valid range in ProD must be covered by data. However, a good diversity is not guaranteed, therefore diversity metrics could be used to further investigate. It follows that, areas not fully covered by data have unsatisfactory diversity. 3D scatter plots are also shown to assist understanding of PF pattern in the actual benchmark problems.

1) Selected algorithms and related parameters

The adopted algorithms are acquired from PlatEMO [29]. NSGA III is a reference-point and nondominated sorting based genetic algorithm for MaOP. GrEA adopts grids drawn in high dimensional objective space to strengthen the selection pressure towards optimal direction while maintaining an extensive and uniform distribution among solutions. IBEA may adopt several binary performance indicators to select offspring as parents for the next round of iteration.

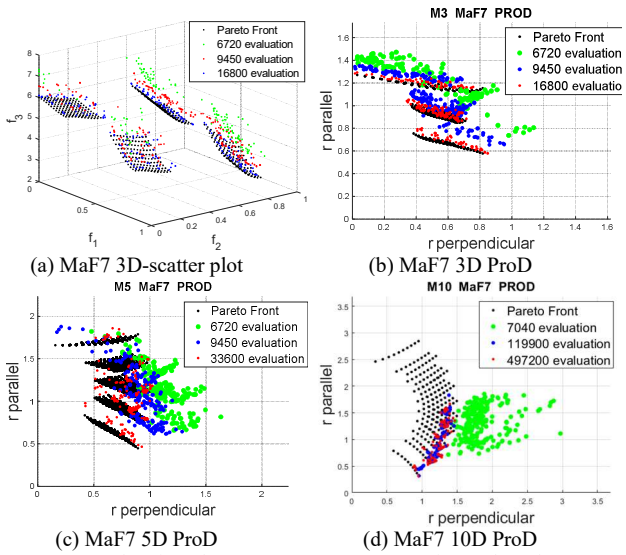
TABLE 2 - NUMBER OF SOLUTIONS GENERATED IN CHOSEN MAOP ALGORITHMS AND ACTUAL PARETO FRONT

Number of objectives	Number of solutions for algorithms	Number of solutions used to generate authentic Pareto Front
3	210	496
5	210	10626
10	220	92376

The number of candidate solutions adopted in each of the chosen algorithms, hence used for estimating the true PF, are listed in Table 2. The parameters applied in the three algorithms are based on default values acquired from PlatEMO version 2.7 [27].

2) Approximations between iterations reveal dominance relations between approximation sets

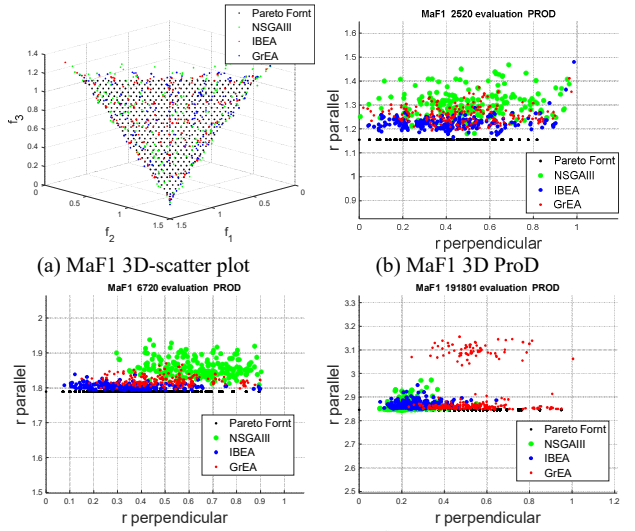
In this section, the ability of ProD to visualise the convergence process at a given time and over time is tested, which demonstrates ProD's ability in revealing dominance relations among PF approximation sets. NSGA III is used to illustrate that ProD can display the ongoing convergence of



(a) MaF7 3D-scatter plot (b) MaF7 3D ProD

(c) MaF7 5D ProD (d) MaF7 10D ProD

Fig. 9. ProD showing the convergence process of Benchmark MaF7 before normalisation on objectives.



(a) MaF1 3D-scatter plot (b) MaF1 3D ProD

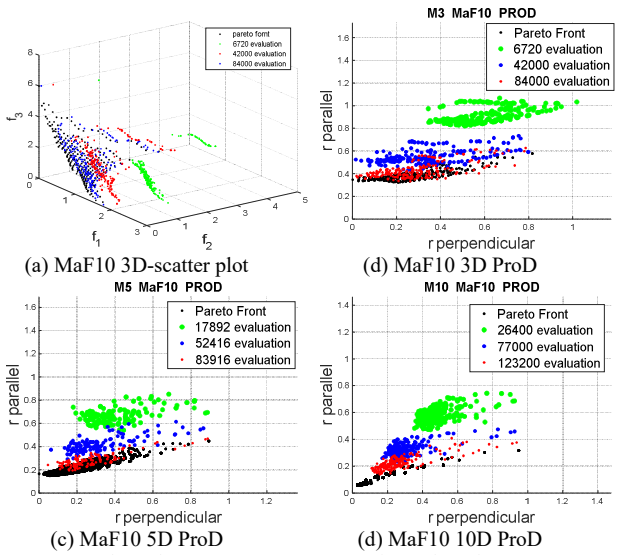
(c) MaF1 5D ProD (d) MaF1 10D ProD

Fig. 11. ProD show the convergence process of Benchmark MaF1 without normalisation on objectives.

an algorithm as the evaluation progresses. Other MaOP algorithms could be used for this purpose as well.

Fig. 8 displays PF approximations on MaF1 after various number of iterations. Results show the convergence process during the algorithm execution. In Fig. 8(c), a notable feature is the absence of data points for small $r_\perp$ values in 5D and 10D cases, a common phenomenon that occurred in most test cases, which is due to the limited amount of reference points created and used in NSGA III in these areas. In Fig.8(d), the absence of data points in the 10D case is observed both for small and large $r_\perp$ values, indicating that NSGA-III based on a small number of reference points can only find partial PF solutions [1]. The above observations imply that ProD reveals a necessary but insufficient condition for diversity, i.e., areas with no coverage with data have poor diversity. However, areas with good coverage do not necessarily mean having sufficiently good diversity.

Fig. 9 shows the convergent process of approximations of MaF7. A striking observation here is the limited diversity of solutions in cases of high dimensionality. Besides, unlike for 3D and 5D cases, NSGA III in the 10D case shows that



(a) MaF10 3D-scatter plot (d) MaF10 3D ProD

(c) MaF10 5D ProD (d) MaF10 10D ProD

Fig.10. ProD show the convergence process of Benchmark MaF10.

convergence stops after about 120,000 iterations. See Fig. 9(d).

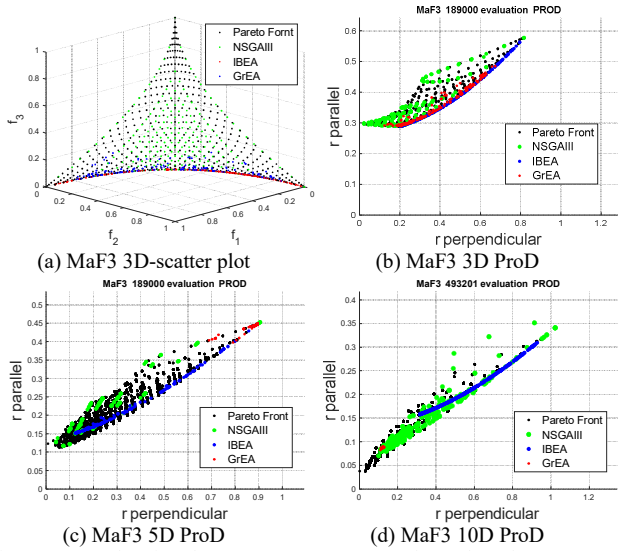
Fig. 10 shows the convergence of MaF10 towards the final goal, although the pattern of PF in ProD is irregular. Fig. 10(d) indicates, as in the earlier example, lack of data points in 10D approximations in low $r_\perp$ value range.

3) Comparison on approximations of various MaOP algorithms, showing the ability of ProD in revealing dominance relations and diversity properties

The approximations of three MaF Benchmarks are shown and discussed here, i.e., MaF1, MaF3, MaF7 and MaF10. Approximations from the three chosen algorithms after a various number of iterations are displayed and compared in ProD.

Fig. 11 (b) shows the diversity and convergence of 3D MaF1 after 2520 functional evaluations. Visually, the approximation set from IBEA has the best convergence for this 3D case after this number of assessments. GrEA ranks as the second-best, while NSGA III suffers a relatively slower speed of convergence. In terms of diversity, no superior algorithm can be identified among the three selected. Similar is the case for 5D, after 6720 function evaluations. However, the diversity of IBEA seems to be higher for a low and intermediate range of $r_\perp$ values. Simultaneously, NSGA III suffers in providing satisfactory diversity in the low range of $r_\perp$ values. For the 10D case, after 154000 evaluations, the situation is significantly changed. Most approximations of NSGA III are nearly converged to the PF, but solutions concentrate in a narrower region, indicating poor diversity. For IBEA, most solutions are converged, but there are some outliers of not fully converged candidates, while its diversity might be better than it is the case for NSGA III. For GrEA, a larger portion of candidates are not converged to PF but those already converged show better spread stretching far to areas away from the reference line or center region, indicating better coverage of the objective space.

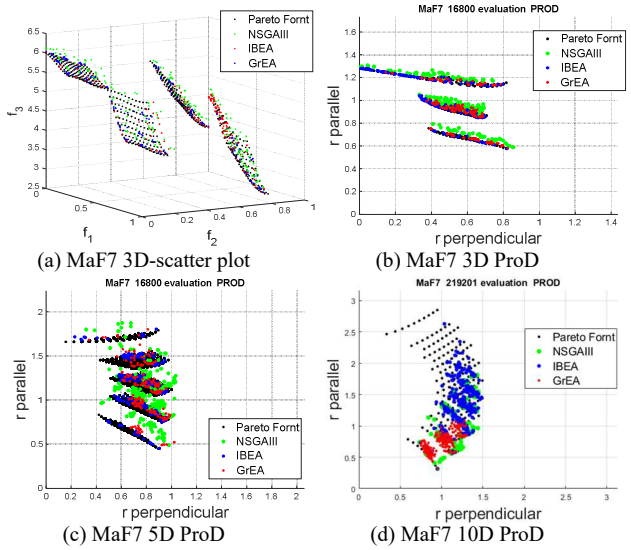
Fig. 12 depicts ProD for the convergence process of approximations of Benchmark MaF3 calculated using the three MaOP algorithms. In the 3D case, see Fig. 12 (a), the approximations of NSGA III have the best diversity. Most of the solutions of IBEA and GrEA locate near f_1 and f_2



(a) MaF3 3D-scatter plot (b) MaF3 3D ProD

(c) MaF3 5D ProD (d) MaF3 10D ProD

Fig.12. ProD showing the convergence process of Benchmark MaF3.



(a) MaF7 3D-scatter plot (b) MaF7 3D ProD

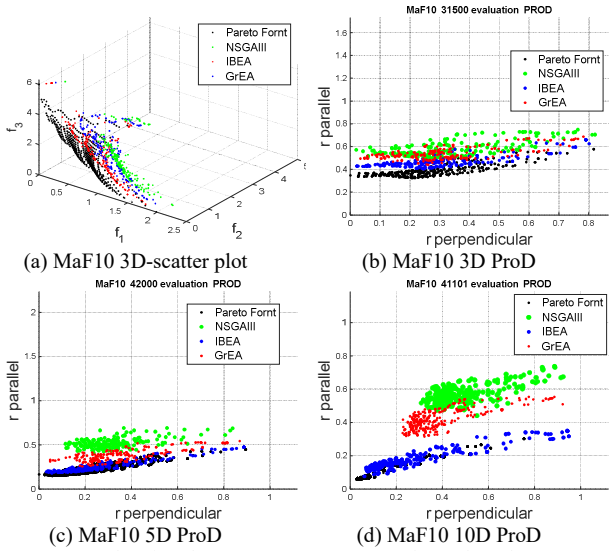
(c) MaF7 5D ProD (d) MaF7 10D ProD

Fig.13. ProD showing the convergence process of Benchmark MaF7.

plane, see Fig. 12(a), the diversity of them is thus inadequate. In ProD, solutions of NSGA III cover well the folded image of the true PF. The solutions of IBEA and GrEA are only located at the bottom of the folded PF image with a relatively good distribution. Still, the area of the lower range of $r_{\perp}$ values is not covered, which indicates poor diversity. The results for 5D and 10D on MaF3, yield similar observations.

Approximations from the 3D case of MaF7 after 16800 functional evaluations show almost equally good results, see Fig. 13(b). However, the convergence of NSGA III is somewhat weaker than it is for the other two algorithms. See Fig. 13(b). For the 5D case, after 16,800 evaluations, the presentation becomes less clear. However, one can still conclude that IBEA yields the best results in convergence and diversity while NSGA III generates the least converged approximations. See Fig. 13(c). For the 10D case, see Fig. 13(d), IBEA gives the best results in diversity compared to the other two algorithms. But none of the three algorithms generates satisfactory approximations in terms of diversity.

PF approximations of Benchmark problem MaF10 depict folded patterns. For the 3D case, after 31,500 functional evaluations, the three competing algorithms appear in layers, and none of them is fully converged. See



(a) MaF10 3D-scatter plot (b) MaF10 3D ProD

(c) MaF10 5D ProD (d) MaF10 10D ProD

Fig. 14. ProD showing the convergence process of Benchmark MaF10 without normalisation on objectives.

Fig. 14(b). Although they all generate similar good diversity results, IBEA yields the best results in convergence, followed by IBEA and then NSGA III. Similar conclusions can be made about the three competing algorithms in 5D after 42,000 functional evaluations and 10D after 176,000 iterations. See Fig. 14(c)(d).

V. Discussion

ProD can display convergence of an algorithm and overall dominance relations between approximation sets that are not too close in performance. When the sets' performance is comparable or similar, ProD has difficulty discriminating the good or poor ones, and a performance indicator, in this case, must be used in conjunction.

ProD may reveal PF shapes in high dimensional spaces in general. It can show PF patterns of symmetry/asymmetry about the reference vector, linearity, convexity/concavity, and their coverage in objective space.

It is recognised that ProD relies on knowing the ideal and nadir points, but these are unknown in real-world problems. However, ProD can be used to visualise the approximations on benchmark problems for the development of new MaOP algorithms, where PFs are known, and Ideal and Nadir points are well defined. For unknown PFs, the nominal Ideal and Nadir point may be estimated based on available data of approximations.

The normalisation of PF approximations may alter the shape of the PF. In practice, ProD can be used with both normalised and unnormalised data; users may select an appropriate approach that leads to a preferred – e.g., the most straightforward – PF pattern.

Multiple solutions in $r_{\parallel}$ may occur for given $r_{\perp}$ values in ProD, resulting in the folded shape of PF. These occur when PFs are asymmetric about the reference vector so that data with the same $r_{\perp}$ values may have different projections $r_{\parallel}$ on reference vector, resulting in several data points lying vertically on the same $r_{\perp}$. Moreover, most real-life MaOP would likely have PFs with highly irregular patterns resulting in the formation of "clouds" in ProD; this causes uncertainty when used to assess dominance relations

between approximations when the performance of competing approximations is similar; this can be mitigated with the use of suitable performance indicators. Future work may also include local diversity indicators to counteract the loss of such information due to the data compression nature of ProD.

VI. Conclusion and Future Work

A new visualisation method, ProD, is proposed to visualise high dimensional vectors of approximation sets of MaOP. All data are visualised in terms of projections on, and distance to, a reference vector; a vector linking the nominal Ideal point and nominal Nadir point based on approximations. Results show that ProD exhibits a more balanced performance compared to the state-of-the-art methods to capture the desired properties of a visualisation method. Satisfactory performance is observed in portraying convergence of PF approximations, thus in revealing dominance relationships.

ProD reveals a necessary, but insufficient condition on its own, for diversity, i.e., areas with no coverage with data in ProD have poor diversity. However, areas with good coverage do not necessarily mean good diversity. It was also shown in this paper that approximation sets of similar performance may be hard to distinguish via using ProD, which can be mitigated by the use of numerical performance indicators.

In future work, ProD should be compared with a wider range of visualisation methods and tested on more PF patterns of higher complexity e.g., of highly irregular distributions, as well as real world problems. Methods for estimating usable nominal Ideal and Nadir points can also be explored.

References

[1] K. Deb and H. Jain, "An evolutionary many-objective optimisation algorithm using reference-point-based nondominated sorting approach Part I: Solving problems with box constraints.", *IEEE Trans. Evol. Comput.*, vol. 18, no. 4, pp. 577_601, Apr. 2013.

[2] B. Filipič and T. Tušar, "Visualisation in Multiobjective Optimisation", Tutorial at GECCO '20.

[3] A. K. A. Talukder and K. Deb., "PaletteViz: A visualisation method for functional understanding of high-dimensional Pareto-optimal data-sets to aid multi-criteria decision making." *IEEE Computational Intelligence Magazine*, 15(2):36–48, 2020.

[4] S. Greco, K. Klamroth, J. D. Knowles, and G. Rudolph, "Understanding complexity in multiobjective optimisation", (Dagstuhl seminar 15031), Dagstuhl Reports, pages 96–163, 2015.

[5] T. Tušar and B. Filipič, "Visualisation of Pareto front approximations in evolutionary multiobjective optimisation: A critical review and the prosection method". *IEEE Transactions on Evolutionary Computation*, 19(2):225-245, 2015.

[6] Z. He, and G. G., Yen, "Comparison of visualisation approaches in many-objective optimisation. In Evolutionary Computation (CEC)", 2017 IEEE Congress on (pp. 357-363). IEEE.

[7] A. Inselberg, "Parallel Coordinates: Visual Multidimensional Geometry and its Applications", Springer, New York, NY, USA, 2009.

[8] F. Kudo and T. Yoshikawa, "Knowledge extraction in multi-objective optimisation problem based on visualisation of Pareto solutions", CEC 2012, 6 pages, 2012.

[9] D. Walker, R. Everson, J. Fieldsend, "Visualising mutually nondominating solution sets in many-objective optimisation", IEEE Trans. Evol. Comput. 17 (2013) 165–184.

[10] A. Ibrahim, S. Rahnamayan, M. V. Martin, K. Deb. "3D-RadVis: Visualisation of Pareto front in many-objective optimisation", CEC 2016, pages 736–745, 2016.

[11] Z. He and G. G. Yen, "Visualization and Performance Metric in Many-Objective Optimisation", IEEE TRANSACTIONS ON EVOLUTIONARY COMPUTATION, VOL. 20, NO. 3, JUNE 2016.

[12] X. Blasco, J. M. Herrero, J. Sanchis, and M. Martínez, "A new graphical visualisation of n-dimensional Pareto front for decision making in multiobjective optimisation", *Information Sciences*, 178(20):3908–3924, 2008.

[13] A. Pryke, S. Mostaghim, and A. Nazemi, "Heatmap visualisation of population based multiobjective algorithms", EMO 2007, pages 361–375, 2007.

[14] M. Yamamoto, T. Yoshikawa, and T. Furuhashi, "Study on effect of MOGA with interactive island model using visualisation", CEC 2010, 6 pages, 2010.

[15] Lotov, A., Bushenkov, V. A., Kamenev, G. K., "Interactive Decision Maps: Approximation and Visualisation of Pareto Frontier", Applied Optimization 89, Springer US, 2004. ISBN: 978-1-4613-4690-6, 978-1-4419-8851-5.

[16] A. R. R. de Freitas, P. J. Fleming, and F. G. Guimarães, "Aggregation Trees for visualisation and dimension reduction in many-objective optimisation", Information Sciences 298 (2015) 288–314

[17] P.-W. Chiu and C. Bloebaum, "Hyper-radial visualisation (HRV) method with range-based preferences for multi-objective decision making", Structural and Multidisciplinary Optimisation, 40(1–6):97–115, 2010.

[18] R. H. Koochaksaraei, R. Enayatifar, and F. G. Guimarães, "A new visualisation tool in many-objective optimisation problems", HAIS 2016, pages 213–224, 2016.

[19] G. Agrawal, C. Bloebaum, K. Lewis, "Intuitive design selection using visualised n-dimensional Pareto frontier, in: Intuitive Design Selection using Visualised n-Dimensional Pareto Frontier", American Institute of Aeronautics and Astronautics, 2005, http://dx.doi.org/10.2514/6.2005-1813.

[20] T. Kohonen, "Self-Organising Maps", Springer Series in Information Sciences, 2001.

[21] S. Obayashi and D. Sasaki, "Visualisation and data mining of Pareto solutions using self-organising map", EMO 2003, pages 796–809, 2003.

[22] M. Yoshimi, T. Kuhara, K. Nishimoto, M. Miki, T. Hiroyasu, "Visualisation of Pareto solutions by spherical self-organising map and its acceleration on a GPU", J. Softw. Eng. Appl. 5 (2012).

[23] Q. Zhang, H. Li, MOEA/D: a multiobjective evolutionary algorithm based on decomposition, IEEE Trans. Evol. Comput. 11 (2007) 712–731.

[24] R. Cheng, M. Li, Y. Tian, X. Zhang, S. Yang, Y, Jin, X. Yao, "A benchmark test suite for evolutionary many-objective optimisation", Complex Intell. Syst. (2017) 3:67–81 DOI 10.1007/s40747-017-0039-7

[25] Yang S. X., Li M., Q., Liu X. H., and Zheng J. H., "A Grid-Based Evolutionary Algorithm for Many-Objective Optimization," *IEEE Trans. Evol. Comput.*, vol. 17, no. 5, pp. 721–736, 2013, doi: 10.1109/TEVC.2012.2227145.

[26] E. Zitzler and S. Künzli, "Indicator-Based Selection in Multiobjective Search BT - Parallel Problem Solving from Nature - PPSN VIII," 2004, pp. 832–842.

[27] K. E. Wu and G. Panoutsos, "A New Method for Generating and Indexing Reference Points in Many Objective Optimisation," 2020 IEEE Congress on Evolutionary Computation (CEC), Glasgow, UK, 2020, pp. 1-8.

[28] J. E. Dennis, "NORMAL-BOUNDARY INTERSECTION : A NEW METHOD FOR GENERATING THE PARETO SURFACE IN NONLINEAR," *Soc. Ind. Appl. Mathematics*, vol. 8, no. 3, pp. 631–657, 1998, doi: 10.1137/S1052623496307510.

[29] Ye Tian, Ran Cheng, Xingyi Zhang, and Yaochu Jin, PlatEMO: A MATLAB platform for evolutionary multi-objective Optimisation [educational forum], IEEE Computational Intelligence Magazine, 2017, 12(4): 73-87".