On a Better Understanding of Unique Identifiers of Pareto Solutions for Multi-criterion Optimization, Visualization, and Decision-making

Anirudh Suresh, *Student Member, IEEE* and Kalyanmoy Deb, *Fellow, IEEE*

COIN Report Number 2024006

***Abstract*—A multi-criterion optimization task requires a decision-making activity during or after the Pareto solutions are found. A convenient and effective decision-making task can be achieved with well-represented Pareto solutions and by means of user-friendly and easily comprehensible visualization tools. To have a comprehensive idea of the spread of Pareto solutions obtained by evolutionary multi-criterion optimization (EMO) or other algorithms, each Pareto solution can be associated with a unique identifier – a vector of size of the dimension of the Pareto surface. While the Pareto solutions can lie on an arbitrary surface with non-domination properties but having its shape and size largely dependent on the problem being solved, these identifiers can have much simpler properties, such as lying on a unit simplex or a unit sphere, or directly relating to preference values. Once achieved, these unique identifiers can help EMO algorithms evaluate the extent of the spread of solutions during the optimization process and decision-makers to have a better understanding of trade-offs among solutions to make better decisions. Moreover, apparent gaps or other complexities of the obtained Pareto solutions can be more easily located in the identifier space. In this paper, we present and compare five identifiers – ideal point reference vectors (RV), nadir point RVs, projection RVs, pseudo-weights, and angle vectors – with respect to their advantages and disadvantages in decision-making and visualization purposes. We also demonstrate that, if desired by the decision maker, these alternative identifiers can be used during optimization to achieve a good distribution of solutions in this space.**

***Index Terms*—multi-objective optimization, multi-criterion decision-making, visualization, optimization**

I. Introduction

The aim of a typical multi- or many- objective evolutionary optimization algorithm (EM(a)O) is to find a good approximation of the respective Pareto-optimal (PO) front using a finite set of Pareto-optimal solutions that are well converged and uniformly distributed. This is then followed by a multi-criterion decision-making (MCDM) step, where the decision-maker (DM) has to visualize, analyze, and interpret the obtained solutions and choose a desired solution for further consideration. Decades of research has led to a number of algorithms [1]–[6] that achieve the first step reasonably well. Similarly, a number of successful MCDM methods as well as advanced visualization techniques [7], [8] have been developed to aid in the decision-making step. Methods which consider both optimization and decision-making steps together are increasingly getting attention.

From a decision-making point of view, DMs look for relative importance of objectives, in terms of their trade-off, direct relevance to preferred objectives, etc., which each Pareto solution may offer. Representation of the Pareto front (PF) in the objective space may not always be informative to DMs. On the other hand, many EM(a)O algorithms use a set of supplied reference vectors (RVs) to drive their search to locate a single optimal solution for each RV. In a sense, each RV is truly employed as an *identifier* for a unique and specific Pareto solution. An interesting thought discussed in the literature somewhat and is highlighted here amply is the possible visualization of PO solutions in the identifier space, rather than in the objective space. If a suitable identifier space is used, visualization might be more informative for decision-making purposes.

In the EM(a)O literature, it has been well demonstrated that the typical ideal point based RVs, although used commonly, are not adequate to handle different shapes of PO fronts (for example, an inverted simplex arising in inverted DTLZ1 problem) as many RVs do not contribute to any PO solution at the end, causing a waste of computations [9]. In such cases, there is a need to look for other identifiers, such as nadir point RVs [10], which may be more appropriate for these problems.

The above discussions indicate merits for exploring different identifier spaces for representing PO solutions for achieving both tasks – optimization and decision-making. In this paper, we discuss five identifier spaces and highlight their advantages and disadvantages for optimization, visualization and decision-making. We also provide simple mathematical relationships between them and also with the objective space, so that an easier transition from one identifier space to another in search of a better or a hybrid representation of PO solutions can be achieved.

The remainder of the study is arranged as follows: In Section II, we describe the identifiers used in this study along with the procedure to compute them. In Section IV, we experimentally demonstrate the advantages and disadvantages of the identifiers on a few test problems, and finally, in Section V, we summarize our findings and present potential future directions of research.

Anirudh Suresh is with Mechanical Engineering Department at Michigan State University, East Lansing, USA, e-mail: suresha2@msu.edu

Kalyanmoy Deb is with Electrical and Computer Engineering Department at Michigan State University, East Lansing, USA, e-mail: kdeb@egr.msu.edu

II. Unique identifiers for PO solutions

In this section, we describe the unique identifiers used in the study along with their existing applications in MOEAs. We consider ideal point RVs ($\mathbf{r}$), used in [2], [3], nadir point RVs ($\mathbf{R}$) used in [10], [11], projection RVs ($\mathbf{P}$) [12], [13], pseudo-weights ($\mathbf{w}$) introduced in [14] and angle based identifier (θ), used in [15]. First, we describe each of them and establish relationships between them.

A. Ideal point RVs

RVs originating from the ideal point and covering the entire positive orthant are commonly used in reference direction based MOEAs like NSGA-III [2], MOEA/D [3], RVEA [4], etc. For an M objective problem, for any $\mathbf{r}$, $\sum_{i=1}^{M} r_i = 1$, making each RV vector to lie on a *unit simplex* forcing $r_i \in [0,1]$. When a set of non-dominated (ND) or PO solutions ($\mathbf{f}$) are normalized so that normalized $\hat{f}_i \in [0,1]$, they can be directly compared with $\mathbf{r}$ vectors. Since $\mathbf{f}$ can occur with arbitrary values depending on the underlying problem, a unique identifier (an appropriate $\mathbf{r}$) can be used to represent an optimal objective vector:

$$r_i = \frac{\hat{f}_i}{\sum_{j=1}^{M} \hat{f}_j}, \quad \forall i. \tag{1}$$

Note that reverse mapping from $\mathbf{r}$ to a suitable $\hat{\mathbf{f}}$ is not possible, unless the Pareto surface is known.

Figure 1 illustrates the identification of a Pareto solution ($\mathbf{f}$) with ideal point RV ($\mathbf{r}$). Notice that since the unit simplex covers the entire positive orthant, any point in it can be represented by a suitable RV having $r_i \in [0,1]$ with $\sum_{i=1}^{M} r_i = 1$.

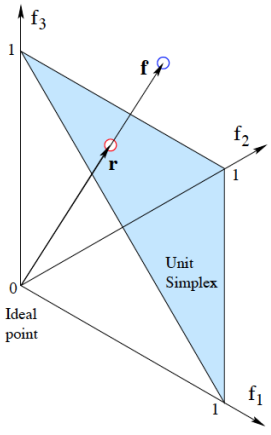


Fig. 1: Computation of $\mathbf{r}$ from $\mathbf{f}$.

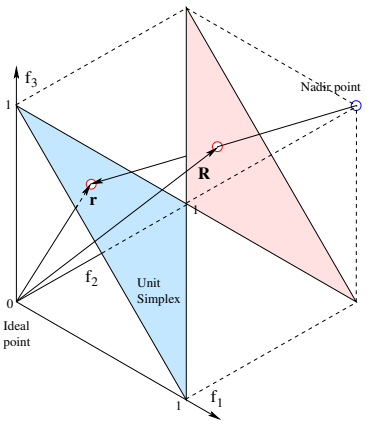


Fig. 2: Computation of $\mathbf{R}$ from $\mathbf{r}$.

Thus, for EM(a)O algorithms or for multi-criterion decision analysis (MCDA), $\mathbf{r}$ identifier vectors can be investigated for (i) evaluating the distributing ability of the algorithm, (ii) finding gaps and lean density of Pareto solutions on the PF, or (iii) directly relating decision-making preferences with $\mathbf{r}$ vectors. For a certain specific number of points, Das and Dennis method [16] provides a perfectly uniform set of points on the unit simplex. Recently proposed Riesz-energy methods are used to find an arbitrary number of RVs on the unit simplex for any dimensional space [17].

For the purpose of decision-making and visualization, recent machine learning based methods have been suggested to directly learn the mapping from RV ($\mathbf{r}$) to the respective optimal variable vectors from a set of EM(a)O obtained PO solutions [18], [19]. If a decision-maker notices a gap or less dense RVs on the unit simplex derived by transforming Pareto solutions to RVs, the trained machine learning system can be used to fill more solutions at the preferred regions of the unit simplex (hence on the PF).

Thus, we realize that having a way to identify Pareto solutions with a unique RV may be useful in both optimization and decision-making tasks.

B. Nadir point RVs

RVs originating from the nadir point have also been proposed as an alternative to ideal point RVs [10] where PBI from MOEA/D is reformulated to compute distance from the nadir point and is referred to as IPBI [10]. As discussed in Ishibuchi et al. [9], RVs from nadir point provide advantages in capturing a good distribution of solutions when the PF is inverted, like in Inverted DTLZ1, or for deeply convex PFs.

Like for ideal point based RVs, any identifier must be able to represent any non-negative objective vector. A little thought will reveal that to achieve this, a nadir point RV ($\mathbf{R}$) for an M objective problem must satisfy $\sum_{i=1}^{M} R_i = (M-1)$. Thus, $\mathbf{R}$ vectors also fall on a linear hyperplane on the M-dimensional objective space, except that it does not intercept the objective axis at one, rather it intersects at $R_i = (M-1)$ for all i. But, there is a limit to this hyperplane that is enough to cover all non-negative objective vectors. As shown in Figure 2, it is an inverted triangle for three-objective problems. For a normalized $\hat{\mathbf{f}}$ having nadir point at vector of ones, the respective $\mathbf{R}$ is given as follows:

$$R_i = 1 - \frac{1 - \hat{f}_i}{M - \sum_{j=1}^{M} \hat{f}_j}, \quad \forall i. \tag{2}$$

After arriving at $\mathbf{r}$ from $\mathbf{f}$ using Equation 1, a transition between $\mathbf{r}$ and $\mathbf{R}$ can be obtained as follows (derived from Figure 2):

$$R_i = \frac{M - 2 + r_i}{M - 1}, \quad \forall i, \tag{3}$$

in which $r_i \in [0,1]$, causing $R_i \in [(M-2)/(M-1), 1]$. Interestingly, R_i are always non-negative for every r_i from the unit simplex. Notice that for $M = 2$, both $\mathbf{r}$ and $\mathbf{R}$ are identical identifiers. Similarly, every $\mathbf{R}$ can also be represented by a specific $\mathbf{r}$:

$$r_i = (M-1)R_i - (M-2), \quad \forall i, \tag{4}$$

in which $R_i \in [0,1]$, causing $r_i \in [-(M-2), 1]$. Interestingly, certain $\mathbf{R}$ vectors will transform to ideal point based RVs taking negative r_i values. The geometry, shown in Figure 2, can be used to derive the above relationships.

Nadir point based RV may be of importance for certain problems, as described above, and a modification of EM(a)O algorithms and decision-making approaches using $\mathbf{R}$ may be beneficial in those cases. The above equations can allow an easy way to update the existing algorithms for optimization and decision-making to freely investigate the points in both identifier spaces.

C. Projection RVs

Another possible identifier is a vector on the extended unit simplex obtained by the orthogonal projection of objective vectors onto the unit simplex hyperplane. This is illustrated in Figure 3. A normalized $\hat{\mathbf{f}}$ point can be represented by a projection RV as follows:

$$P_i = \frac{1 + M\hat{f}_i - \sum_{j=1}^{M} \hat{f}_j}{M}, \quad \forall i. \tag{5}$$

The projection RVs ($\mathbf{P}$) can act as a middle ground compared to ideal and nadir point RVs. However, this also has the disadvantage at three or more dimensions that some regions in the unit hypercube can lead to negative P_i values.

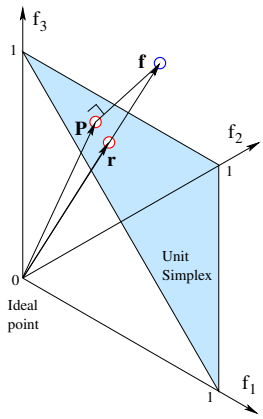


Fig. 3: Projection RV ($\mathbf{P}$) from $\mathbf{f}$.

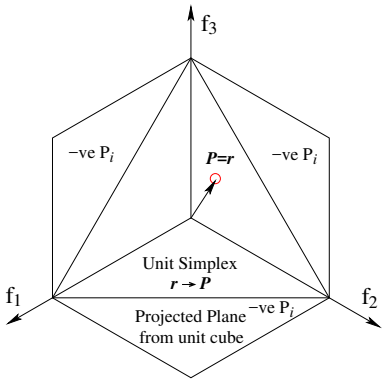


Fig. 4: $\mathbf{P}$-plane (hexagon) obtained from a unit cube in 3D space.

Some geometric considerations reveal that an $\mathbf{r}$ vector can be transformed into a $\mathbf{P}$, as follows:

$$P_i = r_i, \quad \forall i, \tag{6}$$

in which $r_i \in [0, 1]$, causing $P_i \in [0, 1]$ as well. Figure 4 shows that for three-dimensional problem, all non-negative points in unit hypercube $\hat{f}_i \in [0, 1]$ causes the respective $\mathbf{P}$ to lie on the hexagonal extension of the unit simplex, making some P_i's negative, whereas r_i for all points in the unit cube is always non-negative.

Projection RVs have been briefly studied in EMO literature [12], [13]. To achieve a better-distributed set of points on an orthogonally projected hyperplane of the PF hyper-surface, this identifier may be used. The above conversion relationships may be useful for an easier update of existing EM(a)O implementations or for identifying diversity and gaps on a projected plane on the objective space for decision-making purposes.

D. Pseudo-weight Vectors

Pseudo-weight vectors, introduced in Deb et al. [14] are computed as a normalized distance of Pareto points from the worst objective values, as given in Equation 7 for a normalized $\hat{\mathbf{f}}$ vector:

$$w_i = \frac{1 - \hat{f}_i}{M - \sum_{j=1}^{M} \hat{f}_j}, \quad \forall i. \tag{7}$$

Pseudo-weights are normalized to lie on the unit simplex plane ($\sum_i^M \mathbf{w}_i = 1$). The pseudo-weight w_i can be considered as an indicator of *importance* of i-th objective. This is because an objective value close to worst value of the objective function will be assigned a zero pseudo-weight, indicating that the respective solution produces least importance among all objectives. On the other hand, a solution having the minimum objective value for the i-th objective will make a large w_i value indicating that this specific solution indicates an excellent importance to the i-th objective. Similarly, a pseudo-weight value of (0.33, 0.33, 0.34) for a three-objective problem would mean equal priorities to all objectives and will correspond to a solution in the center of the three-objective PF.

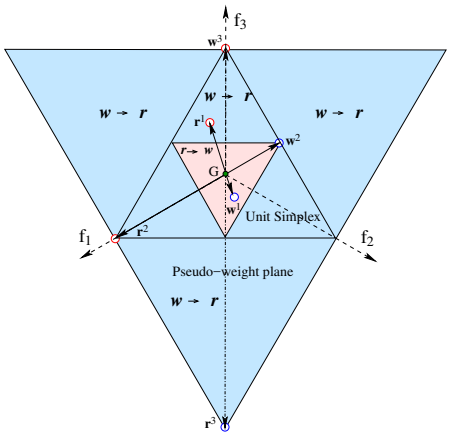


Fig. 5: $\mathbf{r}$ points from the entire unit simplex maps to smaller inverted triangle in the $\mathbf{w}$ space and $\mathbf{w}$ points from the entire unit simplex maps to larger inverted triangle in the $\mathbf{r}$ space spanning to the negative orthants.

Suresh and Deb compare pseudo-weights and ideal point RVs for the purpose of decision-making and visualization in another study [19]. The authors concluded that pseudo-weights, owing to the indicated relative importance of objectives, can be directly related to DM's objective priorities and may be suitable decision-making purposes. In some sense, ideal point based RV ($\mathbf{r}$) and pseudo-weights ($\mathbf{w}$) have contradictory meanings. Ideal point based RVs can be converted to a pseudo-weight vector, as follows:

$$w_i = \frac{1 - r_i}{M - 1}, \quad \forall i, \tag{8}$$

in which $r_i \in [0, 1]$ making $w_i \in [0, 1/(M-1)]$. Thus, pseudo-weights are always non-negative. Interestingly, both $\mathbf{r}$ points on the entire unit simplex map to the smaller inverted triangle as $\mathbf{w}$ points. The mapping of two $\mathbf{r}$ points ($\mathbf{r}^1$ and $\mathbf{r}^2$) to respective $\mathbf{w}$ points ($\mathbf{w}^1$ and $\mathbf{w}^2$) are shown in Figure 5.

The mapping of $\mathbf{w}$ to $\mathbf{r}$ is as follows:

$$r_i = 1 - (M-1)w_i, \quad \forall i, \tag{9}$$

in which $w_i \in [0,1]$ making $r_i \in [-(M-2), 1]$. For some $\mathbf{w}$ vectors ($w_i > 1/(M-1)$), certain r_i values can take negative value for $M > 2$, thereby making them fall outside the boundary of the unit simplex. Figure 5 shows that the mapping of $\mathbf{w}$ points on the entire unit simplex goes to the bigger inverted triangle on $\mathbf{r}$ space with some points extending to the negative orthants ($r_i \in [-1, 1]$). For example, the point $\mathbf{w}^3$ maps to $\mathbf{r}^3$ point on the ideal point based RV space.

An interesting observation about the mapping is that the pair $\mathbf{r}$ and $\mathbf{w}$ lies on the opposite side of the centroid point G on the simplex making G to fall exactly in a ratio $(M-1):1$ from $\mathbf{r}$ and $\mathbf{w}$ points on the extended unit simplex. For three-objective space, $\mathbf{w}$ is half the distance away from G as $\mathbf{r}$ is on the other side of G, and for bi-objective problems, they are equidistant from centroid on either side. Thus, except the point G, every other point from one indentifier space shifts to a different point on the other identifier space.

E. Angle based identifiers

Our final identifier comes from a representation of an objective vector with angle vectors, rather than Cartesian coordinates. A point in an M-dimensional space can be represented by polar coordinates [15], [20] – one distance value and remaining $(M-1)$ angles suitably measured from objective axes. If we represent points on a hyper-sphere (having a unit distance from the origin), then $(M-1)$ angles are enough. Figure 6a shows the angles (θ) for a point ($\mathbf{f}$) on the unit sphere in the 3D space. For any $\mathbf{f}$ vector or a normalized $\hat{\mathbf{f}}$ vector, the respective θ identifier is given as follows:

$$q_i = \frac{\hat{f}_i}{\sqrt{\sum_{j=1}^{M} \hat{f}_j^2}}, \quad \theta_i = \tan^{-1}\frac{q_{i+1}}{\sqrt{\sum_{j=1}^{i} q_j^2}}, \quad \forall i. \tag{10}$$

The respective $\mathbf{r}$ vector on the unit simplex that will represent the same $\mathbf{f}$ vector is also marked in the figure. Using geometry, the following conversion of $\mathbf{r}$ to an equivalent θ can be derived:

$$\theta_i = \tan^{-1}\frac{r_{i+1}}{\sqrt{\sum_{j=1}^{i} r_j^2}}, \quad \forall i, \tag{11}$$

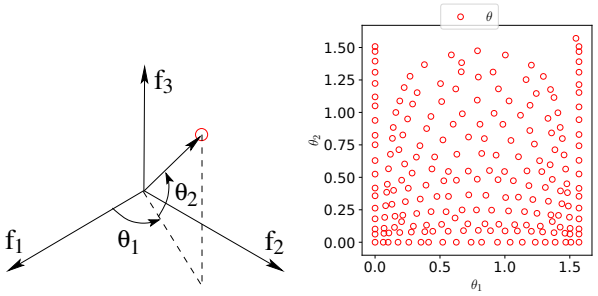
in which $r_i \in [0,1]$, causing $\theta_i \in [0, \pi/2]$. Similarly, the converse can be derived:

$$r_i = \begin{cases} \Pi_{j=1}^{M-1} \cos\theta_j, & \text{for } i = 1, \\ \sin\theta_{i-1} \Pi_{j=i}^{M-1} \cos\theta_j, & \text{for } i > 1, \end{cases} \tag{12}$$

in which $\theta_i \in [0, \pi/2]$, causing $r_i \in [0,1]$.

An example of visualizing tradeoffs with spherical coordinates is shown in Figure 6b. Here, PO solutions from three-objective DTLZ2 problem is transformed to spherical coordinates. Since the normalized PF is already on a spherical surface, this is ideal to be represented by θ vectors. As can be seen from Figure 6b, the distribution in the θ space is almost uniform. The apparent gaps near θ_2 close to $\pi/2$ happen due to the existence of fewer solutions near $f_3 \approx 1$ part of the PF.

The angles can be useful from a convenient decision-making point of view. According to our angle convention, a small θ_i means points have minimized f_{i+1} well (evident from Equation 11). Thus, if f_{i+1} is the most important objective of all, then identifying Pareto solutions with small θ_i would be a direct and easiest way to analyze the PF. Also, good values of f_1 occurs for large values of θ_1.



(a) Angle vectors for representing $\mathbf{f}$ (b) Angle vectors of PF of DTLZ2

Fig. 6: Angle based identifiers

III. Choice of Identifier Space based on Decision-making Preference

In the following, we highlight the relevance of choosing a specific identifier space during optimization to achieve a certain decision-making purpose.

1) **Ideal RV:** This identifier is desired to identify knee-like solutions with more dense points, or the DM is interested in the most compromise (middle) part in a convex-like PF, the DM is interested in obtaining a uniform distribution for a concave-like PF.
2) **Nadir RV:** This identifier is ideal to identify the middle part of a concave-like PF, or the DM is interested in a uniform distribution on a convex-like PF.
3) **Projection RV:** This identifier is ideal to achieve a uniform distribution on a linear-like PF, or the DM is interested in points on the entire PF with a well spread-out points, may be, to fit a surrogate Pareto surface.
4) **Pseudo-weight vectors:** This identifier allows the DM to obtain points corresponding to relative importance factors of objectives (for example, a weight vector (0.75, 0.25) indicating f_1 is three times more important than f_2).
5) **Angle vectors:** This identifier is cable of providing solutions close to a specific objective axis (say f_i), meaning solutions having better values of all objectives other than f_i.

IV. Experimental Study

As demonstrated in the previous section, every identifier space is bounded and is independent of the optimization problem being solved. Since the Pareto surface can come in any shape and form, except having non-dominated properties of any two points on the surface, visualizing and decision-making on the identifier space may enable generic frameworks

for the same. Moreover, every unique identifier for a PO solution needs to have the following properties:

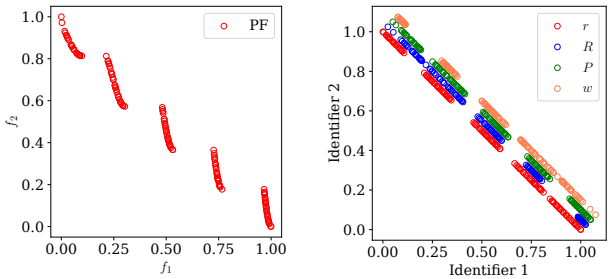
1) **Interpretable tradeoff**: The decision maker should be able to understand tradeoffs and the relationship between objectives from a relatively easier visualization of the distribution of points in the respective identifier space. This may provide greater flexibility to the DM to visualize and make decisions, which may not have been possible by directly investigating the objective vectors.
2) **Identify gaps and discontinuities**: Identifier spaces are intricately linked with objective space. Hence, any apparent gap or discontinuities on the obtained PF should also reflect the same way on the identifier space.
3) **Interpretable extremes**: The DM should understand which solutions in the identifier space correspond to the extreme solutions such that the overall tradeoffs between objectives can be understood. It is also important to be able to understand the bounds of the identifier spaces. Relationships derived in this paper should enable a clear understanding of the linking of objective and identifier spaces.

In this section, we evaluate and compare the identifiers based on the above properties. For this purpose, we analyze and visualize PO solutions of several test problems [21], [22] including inverted and disconnected PO fronts.

Typically, these test problems are solved by ideal point based RVs. Here, we show their representation on alternate identifiers and discuss their suitability for easier visualization and decision-making purposes. Ideal point RV based methods typically struggle in dealing with inverted problems [9], hence they are considered here. Lastly, disjointed PO fronts are also included to investigate if any specific identifier is more suitable to identify gaps and discontinuities in their PO fronts. For each problem, we collect up to 210 PO solutions (based on ideal point based reference direction survival, closest to RVs) from NSGA-III runs for our analysis. We independently compute the identifiers from $\mathbf{f}$ vectors.

A. Effect of Shape of PO Front

1) 2D example: Here, we discuss each identifier's properties from the context of a specific two-objective problem – ZDT3. All identifiers are arbitrarily shifted from each other in order to show them clearly in Figure 7b, while the normalized PF is shown in Figure 7a. In this plot, we clearly see the effect of different transformations of the objective vectors. In two-dimensions, both $\mathbf{r}$ and $\mathbf{R}$ fall on the same hyperplane (unit simplex line here). However, in higher dimensions, they will not lie on the same hyperplane. While the extreme solutions correspond to the same solutions, these solutions are shifted based on the perspective of the ideal or nadir point. Hence, the size of gaps between the islands are reversed for these identifiers. On the other hand, $\mathbf{P}$ have similar sizes gaps between the islands as they are projected perpendicularly to the unit simplex hyperplane. Contrary to RV based approaches, pseudo-weights provide a different perspective and have the islands reversed compared to ideal point RV. Here, the solutions at the maximum value of w_1 correspond to the best value of f_1 at the top left corner of Figure 7a.



(a) Normalized objective space (b) Identifier space

Fig. 7: PF and identifier space for ZDT3 problem. Identifiers lie on the equi-angled line from both axes.

2) Typical Multi-objective Test Problems: Typical test problems are problems where using ideal point RV leads to optimal distribution in the objective space. We can easily observe this to be true for DTLZ1 problem, as shown in Figure 8a. Here, the normalized PO solutions are present on the unit simplex and they map to a uniformly filled unit simplex hyperplane in the ideal point RV ($\mathbf{r}$) identifier. A similar observation can be made about the projection RV ($\mathbf{P}$) identifier as well. However, this is not the case with nadir point RV ($\mathbf{R}$) and pseudo-weight ($\mathbf{w}$). Nadir point RVs do not span the entire hyperplane, as discussed in subsection II-B and form a smaller triangle in the middle of the hyperplane. The same issue is also present with $\mathbf{w}$ identifier as they also do not span the whole unit simplex and instead form a smaller inverted triangle inside the unit simplex. The inverting aspect is further explained by selecting an inverted triangle of solutions for visualization, shown as a gray surface on Figure 8a. These points are marked in blue color in other identifiers (Figures 8b-8e). Contrary to other identifiers, $\mathbf{w}$ vectors have inverted the blue triangle of solutions.

Importantly, if any of these identifier spaces are considered for visualization and decision-making, the user would be expecting points on the entire bounded space marked by the triangular boundary (for 3D). Except for $\mathbf{r}$ and $\mathbf{P}$, other identifier spaces show gaps inside the triangular bounded space. This can cause an apparent concern to the decision-maker. Without a proper understanding of the properties of the different identifier spaces and their bounds, the DM might suspect that this could be an indication that all possible PO solutions are not found by the EMO algorithm, while clearly for DTLZ1 problem this is not the case. Any shape that spans the whole unit simplex with respect to ideal point RVs cannot cover the area with nadir point RVs and pseudo-weights. However, if EMO-created points show gaps inside the respective triangular regions for any of the four identifier spaces, it is expected that the Pareto front has gaps or that EMO was unable to find a well distributed set of points.

As discussed in Section II-E, uniformly distributed objective vectors will lead to some sparse regions and gaps in the angle vector space. This is seen in Figure 6b. This sparsity and gaps

are pathologies of the coordinate transformation and must be kept in mind by the decision maker.

3) Inverted PF problems: Inverted PF problems, like Inverted DTLZ1 show contrasting behavior in different identifier spaces. The PF and the corresponding identifier values for the Inverted DTLZ1 are shown in Figure 8f, and Figures 8g-8j and Figure 9b, respectively. Here, ideal point RVs only form a small inverted triangle on the unit simplex hyperplane. This property is reflected in the fact that ideal point RV based algorithms do not find sufficient solutions in the PF of the inverted problems. A similar effect is seen for projected RV space, as they also form an inverted triangle but are bigger and are spilling out of the first orthant with negative values. On the other hand, nadir point RVs cover the hyperplane completely. This can be vital from the perspective of decision making as an incomplete unit simplex hyper plane, as in the case of ideal point RVs ($\mathbf{r}$) can be misleading regarding the completeness of the PO front. On the other hand, the nadir point RVs show a much clearer picture regarding the fact that the PO front, like in the case of Inverted DTLZ1, covers the whole hyperplane.

A remarkable feature of the pseudo-weight identifier, as seen in Figure 8j, is that the pseudo-weight vectors of the Inverted DTLZ1 PF completely cover the unit simplex hyperplane. Since pseudo-weights are normalized by definition, the extreme solutions correspond to extremes of the unit simplex hyperplane. The mirroring effect of pseudo-weights conveniently fills the hyperplane here and can be of advantage during decision-making.

Angle based identifier spaces for Inverted DTLZ1 are shown in Figure 9b and the inverted nature of the PF is somewhat clear from the figure.

These properties of different identifiers affirm an interesting fact – a specific shape of a PF can be dealt better by a specific identifier. Thus, if some idea of the shape of a PF is known beforehand, EMO algorithms and ensuing decision-making can be performed using a specific identifier.

For inverted problems, nadir point RVs and pseudo-weights perform well in covering the identifier hyperplane and conveying the right information to the decision maker. Ideal point RVs, projection RVs and angle based identifiers only cover a small region in the identifier space and can be misleading to the decision maker regarding the completeness of the PF.

4) Disconnected PFs: An important purpose of visualization in the identifier space is to analyze apparent gaps and understand flaws, if any, with the EMO algorithm or in the PF. Disconnected PF along with the identifiers of DTLZ7 problem are shown in Figures 8k-8o and Figure 9a. The normalized PF (Figure 8k) has four parts with different sizes. All the identifiers form four yet differently-sized islands. For example, the island close to the maxima of f_1 and f_2 forms the smallest island in ideal point RVs and yet forms the largest island with nadir point RVs. Projection RVs forms relatively bigger and well-distrubuted islands that spill over to negative values. The nadir based RV plot skews the distribution providing more density of points near the top apex of the triangular PO front.

Pseudo-weights form differently sized islands conveying different trade-offs for the islands. This demonstrates that the rate at which the DM can navigate through the solutions for changing priorities is different for different islands. This could be important information to be conveyed to the DM. Similarly, angle vectors shown in Figure 9a also demonstrate different sizes and densities, showing different trade-offs as we navigate the $\theta_1 - \theta_2$ space.

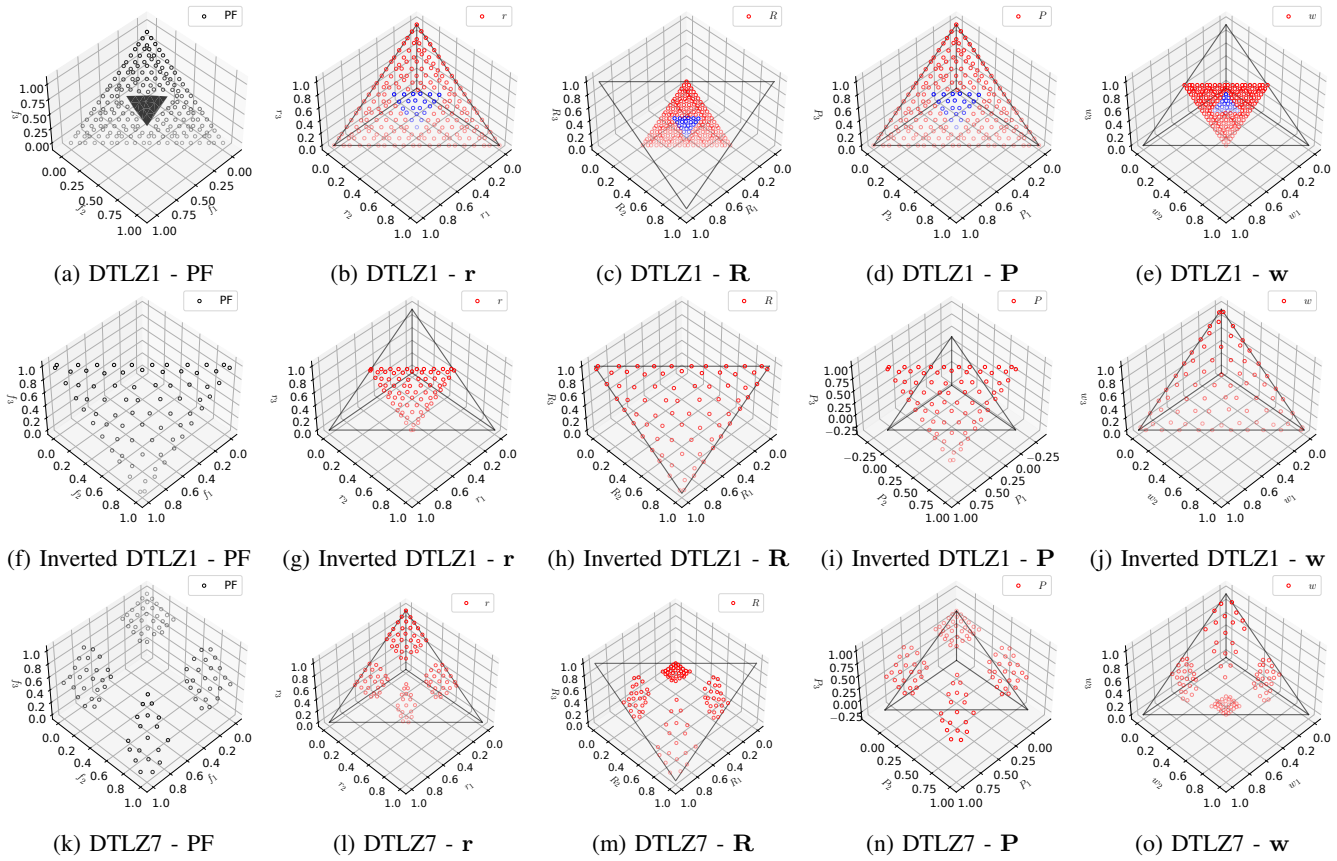
Gaps and islands in the objective space always correspond to gaps and islands in the identifier space. However, relationship between solutions in objective space and identifier space might not be the same. The identifiers can have different densities of solutions changing the interpretation of the trade-off between solutions. Desiring solutions in apparent gaps should be done carefully while taking into account the intricacies of each identifier space. Expecting solutions everywhere on the identifier hyperplane might not be feasible as many PO fronts cannot have solutions everywhere. For example, with angle-based identifier, as shown in Figure 9a on DTLZ7, one of the four disconnected islands seems to come from a wide variety of θ_1, but $\mathbf{r}$ or $\mathbf{P}$ identifiers find fairly uniform distributions of points within each island. For the Inverted DTLZ1 problem, the distribution of points in the θ space is more or less uniform, but the points appear in a bounded region indicating that the entire θ space is not covered by PO solutions.

B. Decision making at higher dimensions

Identifiers such as ideal point RVs, nadir point RVs and projection RVs retain the geometric properties of the PF and these identifiers do not reduce the dimension of the identifier space. This allows the DM to choose desirable solutions, identify gaps and interpret them in a spatial sense. If the DM knows the approximate location of the desired solution or can understand the geometry of the whole PO front, RV based identifiers can be directly used. If the DM does not know the geometric location of the desired solutions and can only navigate the Pareto solutions using priorities, pseudo-weights or angle based identifiers are better suited. However, visualizing solutions for decision-making is difficult at more than 3 objectives as the *geometry* of the PF cannot be taken into account anymore. Though visualization techniques like parallel coordinate plots (PCP), RadViz [23], RadViz3d [7] etc are available, these methods do not convey the spatial information regarding the PO fronts. However, identifiers such as pseudo-weights and angle-based identifiers ignore the geometry and capture the essential trade-off information necessary to choose a solution.

Pseudo-weights always sum up to one and the values convey information regarding the trade-off or priorities of the objectives. Even if the whole PF cannot be visualized due to the dimensionality, merely providing priority values can help the DM in finding a desirable solution or region of interest. PCP plots of pseudo-weights can aid in DM at any number of dimensions.

Angle based identifiers can provide similar advantages as these values convey specific trade-off information. Increasing

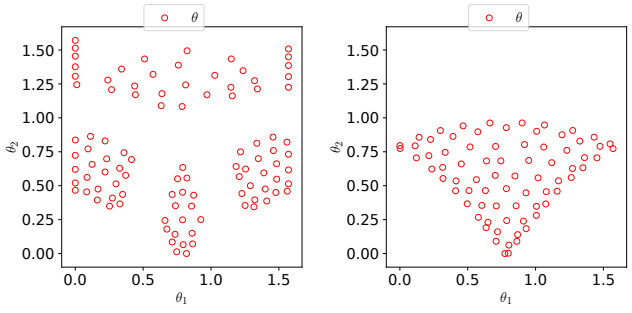


(a) DTLZ1 - PF (b) DTLZ1 - **r** (c) DTLZ1 - **R** (d) DTLZ1 - **P** (e) DTLZ1 - **w**

(f) Inverted DTLZ1 - PF (g) Inverted DTLZ1 - **r** (h) Inverted DTLZ1 - **R** (i) Inverted DTLZ1 - **P** (j) Inverted DTLZ1 - **w**

(k) DTLZ7 - PF (l) DTLZ7 - **r** (m) DTLZ7 - **R** (n) DTLZ7 - **P** (o) DTLZ7 - **w**

Fig. 8: PF and respective identifiers for typical, inverted and disconnected PO front problems.



(a) DTLZ7 (b) Inverted DTLZ1

Fig. 9: Angle identifiers of disconnected and inverted PFs.

or decreasing certain θ_i values can help navigate the Pareto solutions without understanding the shape of the PO front.

C. Suitability of an Identifier for Optimization and Decision-making

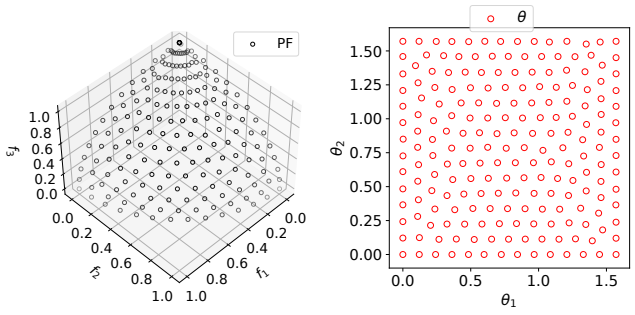
The varying properties of these identifiers take us to an important question: Can the PO solutions found by common EM(a)O algorithms (based on ideal point RVs or nadir point RVs) always produce the most suitable distributions for the DM? As demonstrated in this paper, different identifiers produce different distributions in their respective spaces. Moreover, some identifier spaces can be directly associated with decision-making priorities to the liking of the DM.

For example, Figure 8m has islands of vastly varying densities. Figure 9a has a sparse island of solutions near $\theta_2 = \pi/2$. These factors could be of concern for DMs if they want to analyze the PF in nadir point RVs or angle vector space respectively.

A simple yet effective solution to this issue is to use these identifiers within an EM(a)O algorithm, instead of using the standard ideal point or nadir point based RVs. This can lead to a different distribution of the objective space and yet may produce a better distribution of solutions in the space of identifiers.

Figure 10 shows the effect of preserving diversity in θ instead of the **r** using NSGA-III. As desired, a more uniformly distributed set of PO solutions can be obtained in the θ identifier space, shown in Figure 10b. The PF is abnormally dense near the minimum values of f_1 and f_2. While not uniform in the objective space, this PO set is ideal if the decision-making step is to be performed in the angle-based identifier space. For example, if DM wants best f_2 values, DM can choose a PO solution from small θ_1 region in the θ space. If DM wants small f_3 values, DM can choose a PO solution from small θ_2 region in the θ space, and so on, making a direct correlation of individual best objective solutions with an individual θ_i.

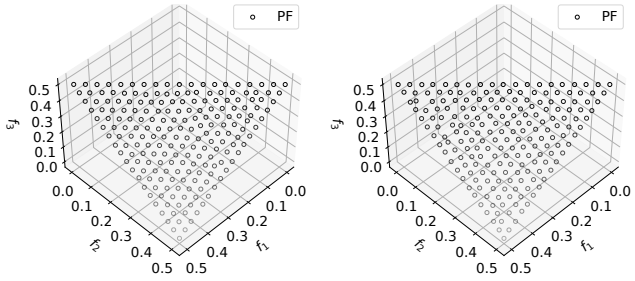
Similarly, the distinctive properties of nadir point RVs and pseudo-weights to capture inverted PO fronts can be exploited



(a) Pareto optimal front (b) Angle identifiers

Fig. 10: PF and respective angle identifiers obtained using angle-based (θ) NSGA-III on DTLZ2 problem.

by an EMO algorithm. Figure 11 shows the PF achieved by using nadir point RVs in NSGA-III instead of the ideal point based RVs (Figure 11a) and pseudo-weight vectors (Figure 11b) during NSGA-III's survival. These PO fronts are much denser than the one achieved using the baseline NSGA-III using $\mathbf{r}$-vectors, shown in Figure 8f.



(a) Using nadir point RVs (b) Using pseudo-weight vectors

Fig. 11: PO fronts obtained using $\mathbf{R}$ and $\mathbf{w}$-based NSGA-III for Inverted DTLZ1.

V. Conclusions

In this study, we have compared and discussed the advantages and disadvantages of several unique identifiers for Pareto-optimal solutions: ideal point RV, nadir point RV, projection RV, pseudo-weight vectors, and angle vectors from the perspective of optimization, visualization and decision-making. We have demonstrated that different identifiers provide varying properties and lead to different distributions of the same PO solutions. We have showed that nadir point vectors and pseudo-weights are capable of capturing inverted PO fronts better. These identifiers have different interpretations and can lead to distinct experiences during decision-making or visualization. For example, RV-based identifiers provide a geometric perspective, pseudo-weight vectors represent priorities of objectives, and angle vectors can describe direct objective-wise relationships. Hence, pseudo-weight vectors and angle vectors can be of use at higher dimensions when a geometric interpretation of the PF is not possible. An important aspect of using identifiers for visualization and decision making is identifying apparent gaps and discontinuities in PO fronts. Gaps in the PF will lead to gaps in the identifier space. However, gaps in the identifier space might not always mean gaps in the PF as the shape of the Pareto front can lead to gaps or incomplete regions in the identifier space. Thus, a proper understanding of the mapping of PO points and the respective identifier space, as derived through mathematical relationships in this paper, is important.

We have also demonstrated that in cases where the DM has a preference over the identifiers, EM(a)O algorithms, like NSGA-III can be modified to perform diversity preserving operations in a suitable identifier space instead of typically used ideal point RVs. This can lead to a satisfactory distribution of solutions in the identifier space for visualization and decision-making purposes. An interesting future direction could be to use multiple identifiers in an EM(a)O to find a satisfactory distribution of solutions for decision-makers with different preferences of identifiers. This also calls for development of further identifiers as they can pave the way for better decision making. While the current study has focused on simple visualization methods, identifiers need to be developed that can aid in advanced visualization techniques, like PaletteViz [8], for a more comprehensive decision-making task.

References

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

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

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

[4] R. Cheng, Y. Jing, M. Olhofer, and B. Sendhoff, "A reference vector guided evolutionary algorithm for many-objective optimization," *IEEE Trans. on Evolutionary Computation*, vol. 20, no. 5, pp. 773–791, 2016.

[5] K. Deb, C. L. do Val Lopes, F. V. C. Martins, and E. F. Wanner, "Identifying pareto fronts reliably using a multi-stage reference-vector-based framework," *IEEE Trans. on Evolutionary Computation*, in press.

[6] H. Seada and K. Deb, "U-NSGA-III: A unified evolutionary optimization pro- cedure for single, multiple, and many objectives – proof-of-principle results," in *Proceedings of Eighth Conference on Evolutionary Multi-Criterion Optimization (EMO-2015), Part II*. Heidelberg: Springer, 2015, pp. 34–49.

[7] A. Ibrahim, S. Rahnamayan, M. G. Martin, and K. Deb, "3D-RadVis: Visualization of Pareto front in many-objective optimization," in *Proceedings of the World Congress on Computational Intelligence (WCCI-2016)*, 2016, pp. 736–745.

[8] A. K. A. Talukder and K. Deb, "PaletteViz: A visualization method for functional understanding of high-dimensional Pareto-Optimal data-sets to aid multi-criteria decision making," *IEEE Computational Intelligence Magazine*, vol. 15, no. 2, pp. 36–48, 2020.

[9] H. Ishibuchi, Y. Setoguchi, H. Masuda, and Y. Nojima, "Performance of decomposition-based many-objective algorithms strongly depends on pareto front shapes," *IEEE Transactions on Evolutionary Computation*, vol. 21, no. 2, pp. 169–190, 2016.

[10] H. Sato, "Inverted pbi in moea/d and its impact on the search performance on multi and many-objective optimization," in *Proceedings of the 2014 annual conference on genetic and evolutionary computation*, 2014, pp. 645–652.

[11] Z. Wang, Q. Zhang, H. Li, H. Ishibuchi, and L. Jiao, "On the use of two reference points in decomposition based multiobjective evolutionary algorithms," *Swarm and Evol. Computation*, vol. 34, pp. 89–102, 2017.

[12] J. Yuan, H.-L. Liu, and F. Gu, "A cost value based evolutionary many-objective optimization algorithm with neighbor selection strategy," in *2018 IEEE Congress on Evolutionary Computation (CEC)*. IEEE, 2018, pp. 1–8.

[13] J. Yuan, H.-L. Liu, F. Gu, Q. Zhang, and Z. He, "Investigating the properties of indicators and an evolutionary many-objective algorithm using promising regions," *IEEE Transactions on Evolutionary Computation*, vol. 25, no. 1, pp. 75–86, 2020.
[14] K. Deb, *Multi-Objective Optimization Using Evolutionary Algorithms*. Chichester, UK: Wiley, 2001.
[15] R. Denysiuk, L. Costa, I. Espírito Santo, and J. C. Matos, "MOEA/PC: Multiobjective evolutionary algorithm based on polar coordinates," in *Evolutionary Multi-Criterion Optimization: 8th International Conference, EMO 2015, Guimarães, Portugal, March 29–April 1, 2015. Proceedings, Part I 8*. Springer, 2015, pp. 141–155.
[16] I. Das and J. E. Dennis, "Normal-boundary intersection: A new method for generating the pareto surface in nonlinear multicriteria optimization problems," *SIAM J. on Optimization*, vol. 8, no. 3, pp. 631–657, 1998.
[17] J. Blank, K. Deb, Y. Dhebar, S. Bandaru, and H. Seada, "Generating well-spaced points on a unit simplex for evolutionary many-objective optimization," *IEEE Transactions on Evolutionary Computation*, vol. 25, no. 1, pp. 48–60, 2021.
[18] I. Giagkiozis and P. J. Fleming, "Pareto front estimation for decision making," *Evolutionary computation*, vol. 22, no. 4, pp. 651–678, 2014.
[19] A. Suresh and K. Deb, "On the choice of unique identifiers for predicting pareto-optimal solutions using machine learning," in *2023 IEEE Symposium Series on Computational Intelligence (SSCI)*, 2023, pp. 1479–1484.
[20] D. Kuang and J. Zheng, "Strategies based on polar coordinates to keep diversity in multi-objective genetic algorithm," in *2005 IEEE Congress on Evolutionary Computation*, vol. 2. IEEE, 2005, pp. 1276–1281.
[21] 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.
[22] E. Zitzler, K. Deb, and L. Thiele, "Comparison of Multiobjective Evolutionary Algorithms: Empirical Results," *Evolutionary Computation*, vol. 8, no. 2, pp. 173–195, 06 2000.
[23] P. E. Hoffman, G. Grinstein, and D. Pinkney, "Dimensional anchors: A graphic primitive for multidimensional multivariate information visualizations," in *Proceedings of the 1999 Workshop on New Paradigms in Information Visualization and Manipulation in Conjunction with the Eighth ACM Internation Conference on Information and Knowledge Management*, 1999, pp. 9–16.