Towards identification of solutions of interest for multi-objective problems considering both objective and variable space information

Tapabrata Ray, Hemant Singh, Kamrul Rahi, Tobias Rodemann, Markus Olhofer

2022

Preprint:

Contents lists available at ScienceDirect

Applied Soft Computing

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

Highlights

Towards identification of solutions of interest for multi-objective problems considering both objective and variable space information

Applied Soft Computing xxx (xxxx) xxx

Tapabrata Ray, Hemant Kumar Singh*, Kamrul Hasan Rahi, Tobias Rodemann, Markus Olhofer

- Development of an approach to identify solutions of interest (SOI) to aid decision making in presence of multiple conflicting objectives.
- The proposed approach incorporates both variable and objective space information in identification of SOI.
- The approach is capable of for both offline and online identification of SOIs.
- The utility of the approach is demonstrated on a range of benchmark and practical design problems containing up to five objectives.

Graphical abstract and Research highlights will be displayed in online search result lists, the online contents list and the online article, but will not appear in the article PDF file or print unless it is mentioned in the journal specific style requirement. They are displayed in the proof pdf for review purpose only.

Please cite this article as: T. Ray, H.K. Singh, K.H. Rahi et al., Towards identification of solutions of interest for multi-objective problems considering both objective and variable space information, Applied Soft Computing (2022) 108505, https://doi.org/10.1016/j.asoc.2022.108505.

ASOC: 108505

ARTICLE IN PRESS

Applied Soft Computing xxx (xxxx) xxx

Contents lists available at ScienceDirect

Applied Soft Computing

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

Towards identification of solutions of interest for multi-objective problems considering both objective and variable space information

Tapabrata Ray [a], Hemant Kumar Singh [a,*], Kamrul Hasan Rahi [a], Tobias Rodemann [b], Markus Olhofer [b]

[a] School of Engineering and Information Technology, The University of New South Wales, Canberra ACT 2600 Australia
[b] Honda Research Institute Europe, Carl-Legien-Strasse 30, 63073, Offenbach/Main Germany

ARTICLE INFO

Article history:
Received 26 October 2020
Received in revised form 16 December 2021
Accepted 20 January 2022
Available online xxxx

Keywords:
Solutions of interest
Multi/many-objective optimization
Robust solutions

ABSTRACT

In multi/many-objective optimization, a decision maker (DM) may often be interested in examining only a small set of solutions instead of the entire Pareto optimal front (PF). Such solutions are referred to as *solutions of interest* (SOI) in some recent studies. A number of methods have been proposed to identify SOIs in an offline or online setting using measures based on reflex angle, bend angle, expected marginal utility, etc. However, these measures only account for the desirable trade-offs in the *objective* space. On the other hand, the variable space information is often critical in practical scenarios as it relates directly to the implemented design. For example, a DM may additionally require that the obtained solutions are robust, i.e., insensitive to variable perturbations, or look significantly different in the variable space, thereby offering multiple equivalent designs to achieve similar performance. These require formulation of new measures and search strategies that simultaneously consider both objective and variable spaces while identifying SOIs. In this paper, we develop an approach that can identify a given number of SOIs for DM's consideration for three different scenarios: (a) purely based on objective space, (b) simultaneous consideration of objectives and robustness, and (c) simultaneous considerations of objectives and equivalent designs. Towards this end, we first define the relevant quantitative measures and illustrate their use for offline selection for a few 2–3 objective test problems. Thereafter, we design an online algorithm that can identify the SOIs and bias the search towards the SOIs based on the scenarios listed above. Lastly, we also present results on two practical examples: a 2-objective welded beam and a 5-objective wind-turbine design problem.

1. Introduction

Optimization involving multiple conflicting criteria is commonly encountered in a number of real-world problems in diverse domains such as engineering design 1[1], finance [2] and operations research [3]. Formally, such problems are defined as multi-objective optimization problems (MOPs). Given the research challenges involved in solving MOPs as well as their wide applicability, MOPs are of significant interest to researchers and practitioners. Consequently, a rich literature exists on the topic, including a number of review papers [4–6].

Theoretically, the optimum of an MOP consists of a set of trade-off solutions in the objective space, referred to as the Pareto optimal front (PF). The corresponding solutions in the variable space are referred to as the Pareto-optimal set (PS). For an $M-$objective problem, the PF can be a manifold of up to $M-1$ dimensions, comprising of an unlimited number of solutions. When solving MOPs using metaheuristics such as evolutionary algorithms (EAs), the goal is to achieve a good approximation of the PF using a finite number (N) of solutions. The solutions should exhibit good *convergence* (proximity to the PF) and good *diversity* (uniform spread across the entire PF) [7]. Though N is often a large number (few hundreds or thousands), for real-world applications, only one or at most a few solutions are required [8]. Therefore, from a practical perspective, it is of interest to identify only a few solutions, which have certain trade-off advantages over the rest of the solutions in the PF, even though they may be equivalent based on Pareto-dominance [9]. The concept of "knee" solutions was introduced in some of the seminal works (e.g. [10]) to characterize solutions with preferable trade-off properties over others in the PF. The knee solutions were defined as those that involve a large sacrifice in at least one of the objectives for a small gain in the other objective. Geometrically, this occurs in the convex regions of the PF that exhibit noticeable change in slope

* Corresponding author.
E-mail addresses: t.ray@adfa.edu.au (T. Ray), h.singh@adfa.edu.au (H.K. Singh), k.rahi@student.adfa.edu.au (K.H. Rahi), tobias.rodemann@honda-ri.de (T. Rodemann), markus.olhofer@honda-ri.de (M. Olhofer).

https://doi.org/10.1016/j.asoc.2022.108505
1568-4946/

of the curve/surface in the objective space, resembling a knee. In some of the more recent works [9,11,12], the term "solutions of interest" (SOI) has been used to denote a small set of preferred solutions in a generic sense, which may not resemble knee points as per the conventional definition. As discussed in [9,12], the term SOI can be used to refer to any solution preferred by the DM. Evidently, such preferences can vary depending on the scenario, and cannot be captured by a universal rule. To cite some examples, in [13], a scenario is considered where anticipated changes could make a given design infeasible, and hence alternative designs are sought which could provide similar performance in the (multi-)objective space. In [9], the diversity of the solutions in the objective space, along with their expected marginal utility was considered to define a preferred set of solutions. In [14], the DM prefers a solution that has the best net normalized gain. That is, relative to the alternative solutions, the improvement in the objectives of a selected solution outweigh the compromises made in the other objectives. In [15], the preferred solution was the one with the highest "convex bulge", likely to be (for certain geometries) in the middle of the PF providing an equal compromise between the objectives. The knee points can thus be considered a category of SOIs [12], where it is required that a small improvement in any objective must come at a large sacrifice in the other(s). For a visualization of how these various types of SOIs could look like on a PF, the readers are referred to the examples in [9,13–16] among others. A number of illustrations will also be presented for the SOIs based on the metrics considered in this study from Section 3 onward.

Furthermore, in many cases, it may not be possible for the DM to provide explicit preferences in absence of problem knowledge. A number of preference modeling and articulation approaches have been therefore proposed in the literature, and a recent review can be found in [17]. The range of available methods could cater to different levels of DM involvement to identify a set of SOIs. These methods can be broadly classified as *a priori*, *a posteriori* and *interactive* [14]. In *a priori* approaches, the preferences are articulated in advance, for example through scalarization using weighted sum of the objectives. In *a posteriori* (also referred to as 'offline selection' approaches), multi-objective optimization is conducted first and then a set of SOIs is selected from the PF approximation [9]. In *interactive* methods (also known as 'online selection' approaches), the optimization is driven with a bias towards potential SOIs [8]. The interactive methods may involve human decision-makers (DM) in loop, or be driven by certain quantitative measures calculated based on current solutions [10,18]. The advantage of the latter category of methods is that it allows the identification of SOIs without inputs from the DM. Therefore a number of such methods have been proposed, such as those based on convex bulge [15], reflex/bend angles [19,20], trade-off information [19,21], expected marginal utility (EMU) [9,10], density [22], curvature [16], Manhattan distance [23–25], etc.

Although a range of above-mentioned measures are available to aid a decision maker by quantifying the trade-off information, the selection of preferred solutions is derived from the objective space alone. However, in real-world problems, it could also be of significant interest to a DM to additionally consider the variable space since it captures the features of the candidate designs [14]. This is relevant, for example, when a DM wants multiple alternative designs that look very different to each other but achieve similar performance in the objective space [8,13]. Another scenario could be where the design performance needs to be *robust* to changes in the variables. That is, small variations in the implemented design due to manufacturing tolerance or other uncertainties should not significantly deteriorate its performance.

In this study, we aim to contribute towards the above research gap by including variable space information in the multi-objective search and selection of SOIs. This extends our previous study [14] wherein we developed a mechanism based on $L1$ norm and angle of influence (Φ) to identify SOIs based on the objective space only (referred to as Scenario-I subsequently). In this work, we introduce additional measures ($T1, T2$) based on combined variable and objective space analysis. The measures are used to identify solutions incorporating the above-mentioned DM considerations around robustness and equivalent designs; referred to as Scenarios-II and III subsequently. Moreover, the measures are integrated in the ranking process to achieve high density of solutions around the prescribed number of SOIs during the online search.

Following this introduction, we discuss some of the prominent works in identification of SOIs in Section 2, followed by the proposed approach in Section 3. The numerical experiments are detailed in Section 4, followed by concluding remarks in Section 5.

2. Related work

As mentioned previously, a number of quantitative measures have been introduced in the literature to characterize knees or SOIs in general. The methods can be broadly classified into two categories: one is based on geometric characteristics of the PF, while the other is based on trade-off information. Although the eventual goal of SOI identification is to achieve superior trade-off solutions, the former category of methods do not consider trade-offs explicitly in their formulation. Instead, they analyze the geometry of the PF to implicitly target such points. The latter category uses the trade-offs in objective space more explicitly to bias the search towards the SOIs. Some representative works from each category are discussed below.

2.1. Geometry based SOI identification

Among the methods that identify SOIs based on geometry of the PF, the method of 'convex bulge' is among the earliest and most commonly used [15,18,26–28]. This method constructs a hyperplane through the extremities of the PF and measures the perpendicular distances of each point in the set to the hyperplane. The hyperplane divides the objective space in two non-overlapping regions. One of the regions contains the ideal point, while the other contains the nadir point. Then, among the solutions that lie in the region containing the ideal point, the one with the maximum distance from the hyperplane is considered as the knee solution. The basic premise in this convex bulge method is that a DM is likely to prefer a balanced trade-off in the central portion of the PF rather than towards the extremities [15]. The method is scalable for problems with more than two objectives, but the reliability of finding the correct PF extremities reduces due to geometric complexities in the shape of the PF in higher dimensions [29]. Another common way to characterize SOIs is to use the *angle* between the points in the PF. In [10], a reflex angle was defined as the external angle formed by a given point with its left and right neighbors on the non-dominated front for a bi-objective problem. The solutions were ranked based on descending order of reflex angle. The high-ranked solutions were considered more likely to be the knee(s). The key issue with the measure is that the reflex angle is highly sensitive to the distribution of the solutions in the obtained non-dominated front. To overcome this, a bend angle was suggested in [19], where the angles with the extremities of the PF were used rather than the neighbors. While these measures were restricted to bi-objective problems, extensions based on angle dominance have also been proposed to prune a non-dominated set [30] and drive evolutionary search [20,31]. Another proposal to identify SOIs

geometrically is by estimating local curvatures among the non-dominated solutions in the objective space [16]. In [16], the objective functions $f_i; i = 1 \ldots M$ were locally fitted to a family of curves $\sum_i^M f_i^\alpha = 1$, where M denotes the number of objectives). Relatively lower values of α denoted proximity to the knee regions. However, fitting the solutions in this manner implicitly assumes that the PF is a continuous manifold. Consequently, the approach is may not be suitable for problems with disconnected PFs.

2.2. Trade-off based SOI identification

Some of the methods use trade-off information in the objective space explicitly to determine the proximity of a solution to the knee region. In [19], a trade-off approach was presented for bi-objective optimization, where the aim was to find solutions where a unit gain in f_1 results in at least α sacrifice in f_2. Similarly, a unit gain in f_2 should result in at least β sacrifice in f_1; where (α, β) are user defined parameters. The scalability of the approach for > 2 objectives is challenging, as it would involve specifying several threshold values. A scalable approach was presented in [10], where a measure known as *expected marginal utility* (EMU) was formulated. A linear utility function of the form $U(\mathbf{x}, \lambda) = \sum \lambda_i f_i$; where $\sum \lambda_i = 1$ was calculated for each point. The EMU is calculated as the sacrifice in utility along a particular direction if the individual with the best utility along that direction becomes unavailable. A recursive application of EMU (EMU^r) was further proposed in [9,11] to achieve complete ordering of the solutions and deliver a given number of SOIs, especially when dealing with higher number of objectives. In [23], a Minimum Manhattan Distance (MMD) from the ideal point in the normalized objective space was used to identify the knee point, and was further integrated in an online method EvoKnee^r in [25]. MMD can also be classified under geometry based methods, but for pairwise comparison it also directly represents the cumulative trade-offs made in all objectives and hence included here in the trade-off category. Along similar lines, net normalized gain along with angle of influence (NNGA) was presented in [14]. In NNGA, the reference for calculation of $L1$ norm and the angles was set as the nadir point instead of the ideal point. This is because for desirable trade-offs, the knees lie in convex regions as discussed previously. The angles with nadir point as the origin map the convex PFs more uniformly compared to ideal points [32,33], hence the choice of nadir point in NNGA. The trade-off utility approach in [21] also compares the solutions based on 'knee-dominance', which considers the overall gain and deterioration between the normalized objective values, akin to the methods above.

There also exist some other approaches that may not directly fit into the above two categories. For example, solution density on a projected hyperplane was used [22] as a primary measure for identifying knee regions. The concave regions are also considered to be of interest in the work [22]. In [9,11], diversity was considered as a secondary criteria (in conjunction with EMU^r) to identify the specified number of SOIs. Hypervolume contribution was used as a criteria to characterize knee solutions in [34].

Apart from supporting decision making, knee solutions have also been used to drive the search more efficiently, even where the final aim was to achieve a better approximation of the full PF rather than achieving a few SOIs [24,31,34–37]. Recently, new problems have also been introduced for benchmarking the algorithms which aim to find solutions/regions of interest [38].

In the methods discussed above, the identification of SOIs is solely based on objective space considerations. However, often, the decision-maker may benefit from the associated variable space analysis and its incorporation into the measures can help in locating more useful SOIs [14]. Such considerations could involve, for example, having a choice of several alternative designs that are very different in the variable space, but achieve similar performance (with desirable trade-offs) in the objective space. Another consideration may be robustness, i.e., the performance of the implemented design should not deteriorate significantly under unavoidable uncertainties. In the preliminary study [14], these cases were analyzed in an offline setting. However, the quantitative measures and the online search for such solutions were not incorporated within the multi-objective search itself. In this work, we build upon these concepts to develop a method that will bias the search towards the SOIs considering both the objective and variable spaces. To achieve this, we complement the previously proposed objective space measures $(L1, \Phi)$ [14] with newly proposed online measures $(T1, T2)$ that consider both objective and variable space information. Subsequently, we construct search strategies to obtain the SOIs for three different scenarios, and illustrate the performance of the method on a range of benchmark problems and practical examples.

3. Methodology

In this section, we briefly discuss the targeted scenarios and label them as Scenarios I–III for ease of reference. Thereafter, we discuss the quantitative measures for ranking the solutions to induce selection pressure towards the SOIs for the above scenarios. Lastly, we discuss how these measures are incorporated in the online search, so that the algorithm intensifies the search around the given number of SOIs rather than the entire PF.

3.1. Target scenarios in this study

Inline with several works in the existing literature [9,11], we assume that a small, fixed number N_s of SOIs is required to be presented to a DM instead of the entire PF approximation. When viewed in terms of online search, this would further imply that it is desirable to have higher intensity of solutions around the SOIs, rather than uniform distribution over the entire PF. The main motivation behind doing so is to aid informed decision-making, as past studies have shown that large number of solutions in the presence of multiple criteria are not easy to process cognitively [9]. While searching for the given number of SOIs, we consider the following three scenarios in this study:

- Scenario-I: The DM requires a small number of SOIs purely based on the desirable trade-off properties in the objective space. This is inline with the existing methods that only consider the objective space.
- Scenario-II: The DM requires a small number of SOIs considering the performance in objective space as well as the robustness of the recommended designs. The rationale behind such preferences is that under minor (often inevitable) variations in implementing a design, the performance will not change significantly. In practice, such situations are encountered frequently in many disciplines, particularly engineering [39,40].
- Scenario-III: The DM requires a small number of SOIs that consider the performance in the objective space as well as the variable space diversity. The latter implies the presence of multiple equivalent designs for similar objective performance. Such solutions are of relevance to DMs involved in, e.g., design of re-configurable product families [41,42]. It is desirable to have options for replacing a given design with another one to achieve similar performance in case the former becomes infeasible/impractical..

It is relevant to mention that the Scenarios-II and III have some commonalities with the field of multi-objective robust optimization [40,43] and multi-objective multimodal optimization [44],

respectively. In multi-objective robust optimization [43], PF approximation(s) are sought that satisfy some criteria of robustness, normally quantified using additional constraints/objectives based on *expected* response values in presence of uncertainties. In multi-objective multi-modal optimization [44], the aim is to find all designs in the variable space that correspond to each solution on the obtained PF approximation; thus requiring convergence and diversity in both objective and variable spaces. The main aspect in which this work differs from the majority of the above-mentioned literature is that the complete PF is not sought in this study. Instead, the aim is only to focus the search around a selected number of SOIs on the PF, where the definition of SOI implicitly considers the robustness or multimodality. Also worth noting is that there do exist other studies that incorporate preferences for certain solutions instead of searching for the entire PF. For example, in [13], multiple alternative designs were sought in the search space to deliver a similar performance in the multi-objective space as that of a baseline design preferred by a DM. Thus, the current study can be thought of as a preference-based search, where the preferences specifically cover the above-mentioned scenarios. Moreover, it is assumed that the scenarios are known, but no explicit preference articulations or existing designs have been provided a priori by the DM. The aim is thus to identify a limited number of solutions for the above scenarios in an offline or online manner for an informed decision-making.

3.2. Quantitative measures for SOI selection/ranking

For calculating the measures discussed in this section, the first step involves normalization of the given solutions in the objective and variable spaces. This is to avoid bias towards any particular objective(s)/variable(s) while calculating geometric measures such as angles and Euclidean distances. For the variable space, normalization is straightforward. Since the variable ranges are known, the variables can simply be linearly mapped to [0,1] using lower and upper variable bounds $(\mathbf{x}_l, \mathbf{x}_u)$ as $\frac{\mathbf{x}-\mathbf{x}_l}{\mathbf{x}_u-\mathbf{x}_l}$. However, for the objectives, since the limits of the PF are typically unknown to start with, the normalization needs to be done based on the estimates of the ideal point (Z^I) and nadir point (Z^N) from the existing non-dominated set of solutions. In doing so, it is important to remove the *dominance resistant solutions* (DRS) from the set. DRS solutions are those that are poor in some of the objectives (thus far from the PF), but remain in the non-dominated set due to superior value(s) in at least one of the remaining objectives [45]. The identification of DRS in itself is a challenging and open problem in the field [46]. In this work, we employ a relatively simple strategy to remove potential DRS. Firstly, all solutions are normalized in the objective space using the ranges of the current non-dominated set of solutions. To do so, the coordinates of the ideal and nadir points (Z^I, Z^N) are identified using the minimum and maximum values, respectively, of each objective. The normalized objective values are then calculated as $\frac{f_i-Z_i^I}{Z_i^N-Z_i^I}; i = 1 \ldots M$, where M is the number of objectives. In the normalized space, the coordinates of (Z^I) and (Z^N) map to $(0, 0, \ldots, 0)$ and $(1, 1, \ldots, 1)$, respectively. Then, we generate a set of uniformly distributed reference vectors using systematic sampling [47] and attach the current solutions to these reference vectors using a minimum angle assignment. Then, for each axial reference vector (those for which all but one weight values are 0), we identify all the solutions attached to it. Among these solutions, we pick the solution with the lowest Euclidean distance to the ideal point $(0, 0 \ldots 0)$. A solution corresponding to each axial vector is thus identified through the above process. The maximum value corresponding to each objective in this set of solutions yields the coordinates of the estimated nadir point. It is important to take note that the reference vectors are used only for DRS removal, unlike previous proposals (e.g. [9]) where they were used for trade-off quantification. The generation of weight vectors requires the spacing parameter (h). Any other form of DRS removal can also be used instead of the above. In addition, to avoid the degenerate cases specially when the estimated Z^I and Z^N is same due to having only one solution in the non-dominated set, we adopt the strategy proposed in [48]. Accordingly, solutions from consecutive fronts are added to the non-dominated set until the difference between Z_i^N and Z_i^I in each objective i is larger than a predefined threshold (10^{-4}).

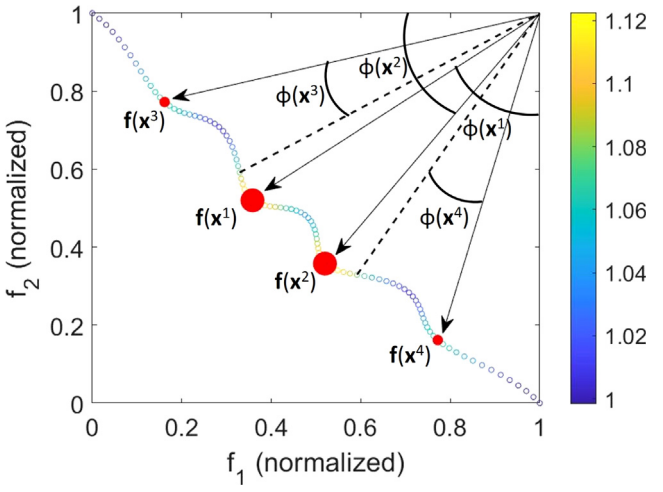


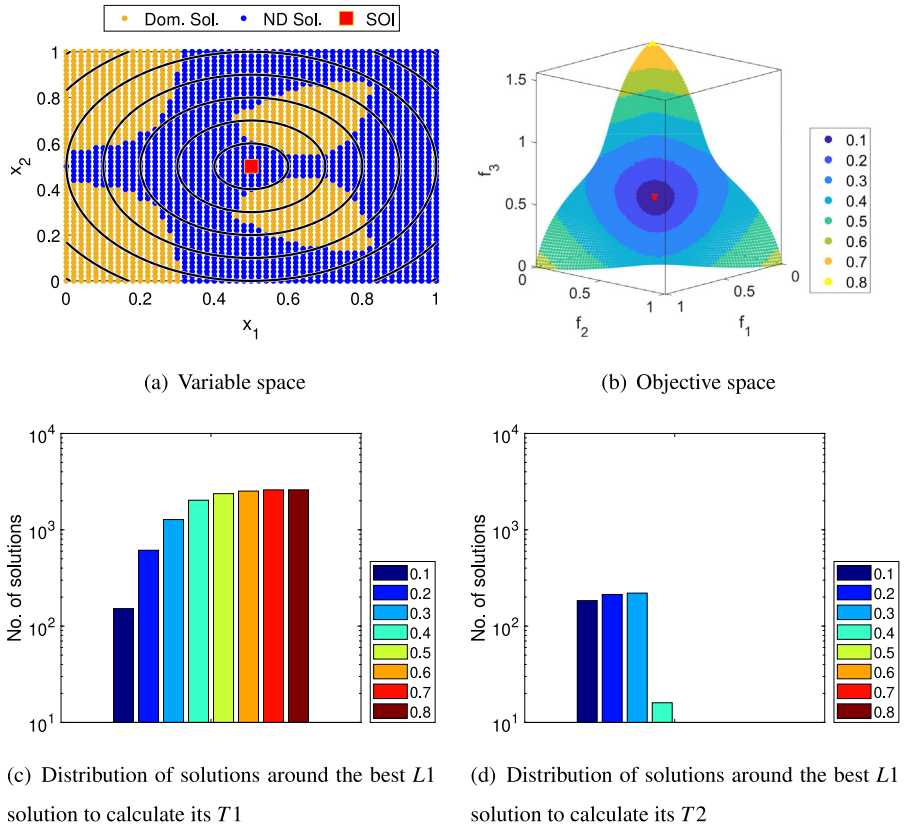
Fig. 1. Illustration of $L1$ measure and angle of influence Φ. The $L1$ value for each point in the given non-dominated set are shown using colorbar. Top four solutions with the highest angle of influence are presented with red solid circles whose sizes are proportionate to their Φ measure.

3.2.1. L1 Norm and angle of influence (Φ)

These two measures operate on the information of the objective space alone, and have been inherited from the previous work [14]. The details are provided here for completeness, in particular since the measures are also used in the new scenarios considered later in this study.

The first step is the normalization of the objective values of the given non-dominated set of solutions. This proceeds in the same manner as discussed in the previous section, but with the DRS removed. In this normalized objective space, we calculate $L1$ norm between a given solution $\mathbf{x}$ and the nadir vector as $L1 = \sum_i^M 1 - f_i(\mathbf{x})$. Here, $f_i(\mathbf{x})$ denotes the ith objective value, and $\mathbf{f}(\mathbf{x})$ denotes the normalized objective vector corresponding to $\mathbf{x}$. This quantity can be interpreted as the normalized net gain of the solution's performance over the nadir point, since the gain in ith objective is simply $1 - f_i(\mathbf{x})$. An illustration of the $L1$ values for a bi-objective DEB2DK4 problem [10] using a set of 100 solutions on the PF is shown in Fig. 1. The solutions marked $\mathbf{x}^1$ and $\mathbf{x}^2$ have the highest normalized net gain.

It is worth noting that if the Minimum Manhattan distance (MMD) [25] from ideal point is used to select the point instead of the proposed $L1$ measure, then the same solution $\mathbf{x}^1$ or $\mathbf{x}^2$ will be obtained. If only one SOI is required as the final output, then either of these solutions can be thought of as the most preferred solution given that it exhibits the best trade-off (i.e., maximum net normalized gain). However, if the number of required SOIs (N_s) is more than one, then an additional measure is required to generate a complete preference order. We do this using an *angle of influence* Φ. The calculation of the $\Phi(\mathbf{x})$ measure for a solution $\mathbf{x}$ is outlined below.

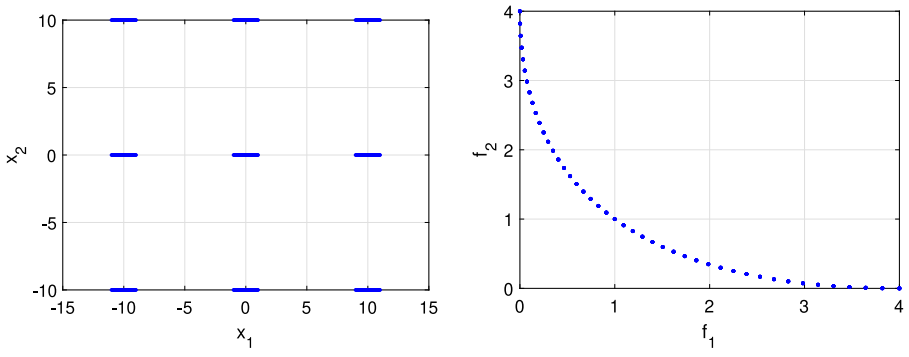


(a) Variable space (b) Objective space

(c) Distribution of solutions around the best $L1$ solution to calculate its $T1$

(d) Distribution of solutions around the best $L1$ solution to calculate its $T2$

Fig. 2. Illustration of proximity in objective and variable spaces used for calculation of $T1, T2$ measures around a given solution, for DEB3DK1 problem.



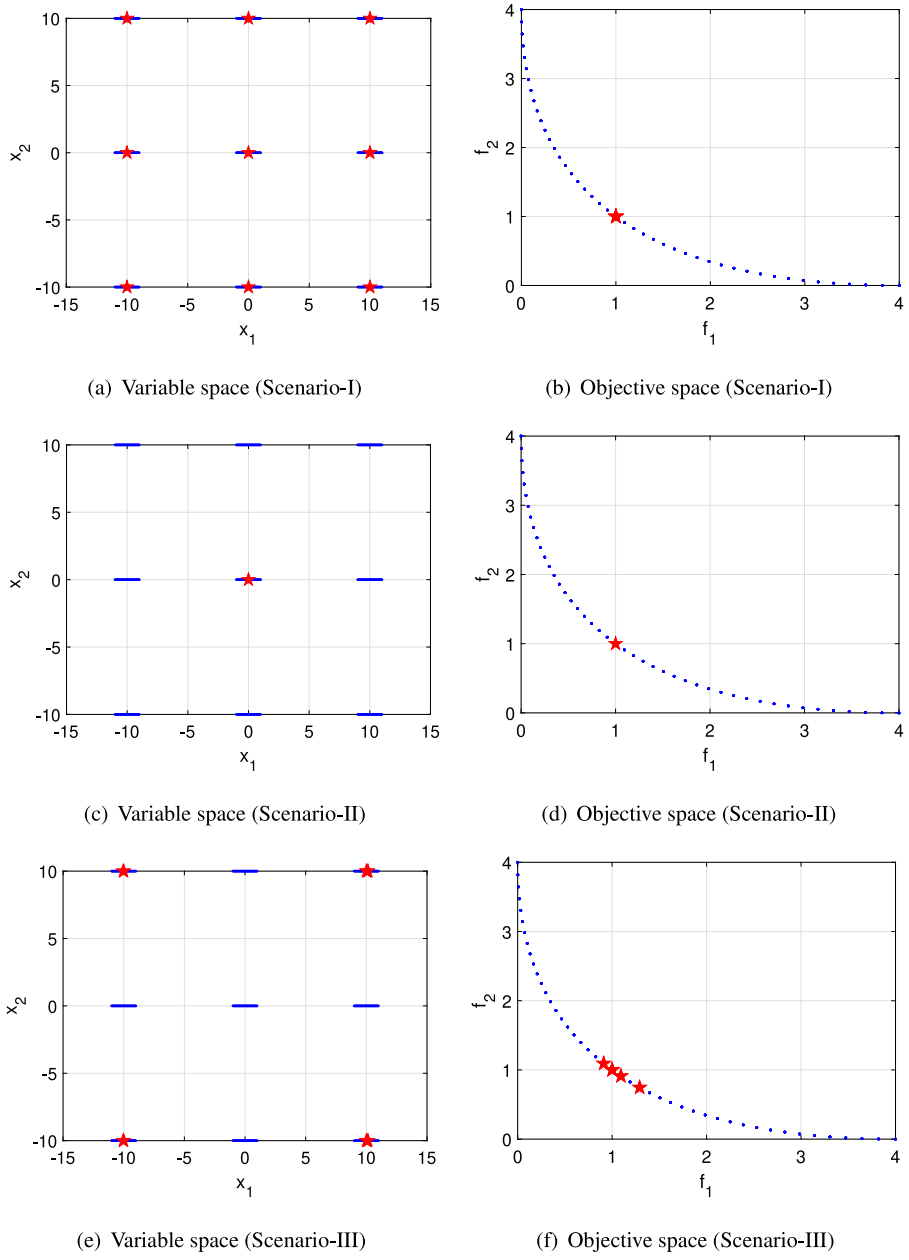
(a) Variable space (b) Objective space

Fig. 3. PS and PF illustration for SYM-PART problem in (a) and (b).

Using the nadir point $(1, \ldots, 1)$ as the origin, we construct a reference vector to the point under consideration. The reference vector corresponding to solution $\mathbf{x}$ is denoted as $\vec{R}_{\mathbf{x}}$. The same is done for all the solutions in the given non-dominated set. Then, the angle of influence of $\mathbf{x}$, i.e., $\Phi(\mathbf{x})$, is defined as the smallest angle between $\vec{R}_{\mathbf{x}}$ and $\vec{R}_{\mathbf{x}^j}$, such that the solution $\mathbf{x}^j$ has a higher $L1$ than solution $\mathbf{x}$ (if such a point exists). Mathematically, it can be computed as shown in Eq. (1) (first condition), where Q represents the non-dominated front in the normalized objective space, and $\mathbf{x}^j$ is the nearest solution with a higher $L1$ than $\mathbf{x}$. If there does not exist $\mathbf{x}^j$ with a better $L1$ than $\mathbf{x}$, then the angle of influence is simply the largest angle that any $\mathbf{x}^j \in Q$ makes with $\mathbf{x}$, as shown in Eq. (1) (second condition).

$$\Phi(\mathbf{x}) = \begin{cases} \min\limits_{\mathbf{x}^j \in Q; L1(\mathbf{x}^j) > L1(\mathbf{x})} \left(\cos^{-1} \frac{\vec{R}_{\mathbf{x}} . \vec{R}_{\mathbf{x}^j}}{\|\vec{R}_{\mathbf{x}}\| \|\vec{R}_{\mathbf{x}^j}\|} \right) & \text{if } \{\mathbf{x}^j \in Q; \\ & L1(\mathbf{x}^j) > L1(\mathbf{x})\} \neq \emptyset \\ \max\limits_{\mathbf{x}^j \in Q} \left(\cos^{-1} \frac{\vec{R}_{\mathbf{x}} . \vec{R}_{\mathbf{x}^j}}{\|\vec{R}_{\mathbf{x}}\| \|\vec{R}_{\mathbf{x}^j}\|} \right) & \text{otherwise} \end{cases} \tag{1}$$

In Fig. 1, $\mathbf{x}^1$ and $\mathbf{x}^2$ have the equal highest $L1$ among all solutions. The corresponding $\Phi(\mathbf{x}^1)$ and $\Phi(\mathbf{x}^2)$ will be the largest as well, and equal to the maximum angle formed by the respective solutions with any other solution in the set. Now consider another solution $\mathbf{x}^3$ shown in Fig. 1. For this solution, $\Phi(\mathbf{x}^3)$ is depicted in Fig. 1.



(a) Variable space (Scenario-I) (b) Objective space (Scenario-I)

(c) Variable space (Scenario-II) (d) Objective space (Scenario-II)

(e) Variable space (Scenario-III) (f) Objective space (Scenario-III)

Fig. 4. Offline illustration of SYM-PART problem considering Scenario-I, II and III respectively. Blue dots show the archive of solutions. For Scenario-I, solutions with the highest Φ is shown with a red star. For Scenario-II and III, the solutions non-dominated with respect to normalized $L1$ and $T1$ (Scenario-II) and normalized $L1$ and $T2$ (Scenario-III) are shown with red stars.

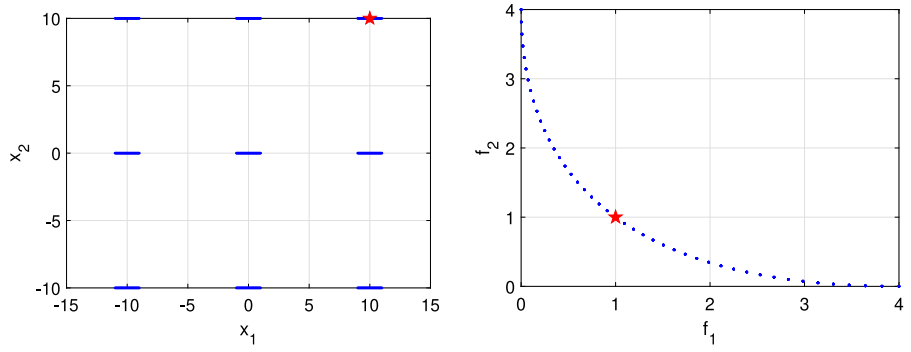
Note that based on the above definitions, $L1(\mathbf{x}^1) > L1(\mathbf{x}^3)$ implies that $L1_{1-3} = (1 - f_1(\mathbf{x}^1) + 1 - f_2(\mathbf{x}^1)) - (1 - f_1(\mathbf{x}^3) + 1 - f_2(\mathbf{x}^3)) > 0$, i.e., $L1_{1-3} = f_1(\mathbf{x}^3) - f_1(\mathbf{x}^1) + f_2(\mathbf{x}^3) - f_2(\mathbf{x}^1) > 0$. The quantity $L1_{1-3}$ can be interpreted as the net gain that the solution $\mathbf{x}^1$ has over solution $\mathbf{x}^3$.

Thus, to summarize the interpretation of the above process, the selection of the solution $\mathbf{x}^1$ or $\mathbf{x}^2$ (with highest $L1$) asserts that no other solution in the set will have a net positive gain in the objectives over $\mathbf{x}^1$ or $\mathbf{x}^2$. Similarly, for another solution $\mathbf{x}^3$, there will exist no solution within its angle of influence $\Phi(\mathbf{x}^3)$ that will have net positive gain over $\mathbf{x}^3$. Thus, if more than one SOI is required, the remainder of the solutions can be selected as per descending values of the angle of influence. Top four SOIs selected based on the above scheme are presented in Fig. 1. The size of the solid marker indicates the order of SOIs; with the largest markers corresponding to $\mathbf{x}^1$ and $\mathbf{x}^2$ followed by solutions $\mathbf{x}^3$ and $\mathbf{x}^4$.

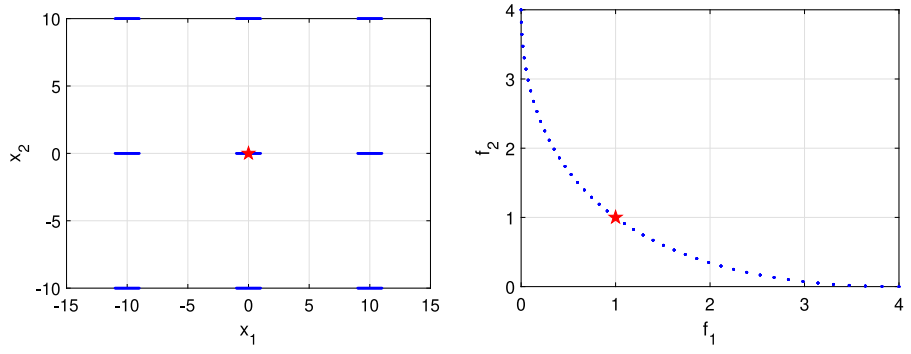
3.2.2. $T1$ And $T2$ measures

In this section, we propose two new measures, $T1$ and $T2$, which are designed to handle Scenarios-II and III respectively. The main idea behind these measures is to quantify the spatial distribution of the solutions in the objective and variable spaces. Subsequently, the measures are used in the ranking of the solutions to induce higher preference for the solutions that align with the given decision-making scenario.

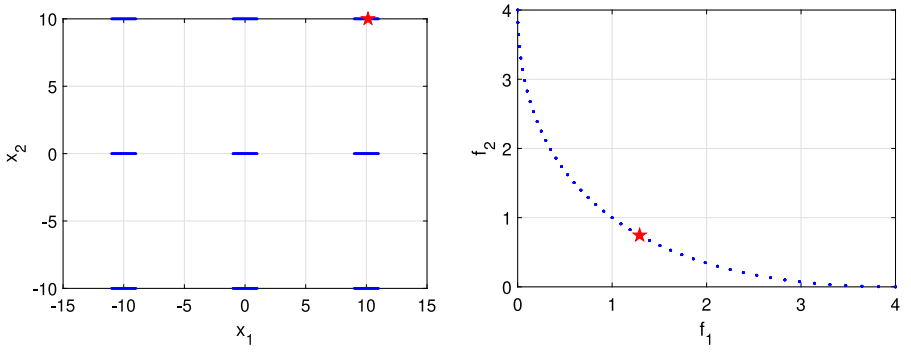
The process of calculating the two measures is outlined in Algo. 1. The steps involved can be more conveniently understood using an example. For this illustration, we use the 2-variable DEB3DK1 problem, so that the objective and variable spaces can



(a) Variable space (Scenario-I) (b) Objective space (Scenario-I)



(c) Variable space (Scenario-II) (d) Objective space (Scenario-II)



(e) Variable space (Scenario-III) (f) Objective space (Scenario-III)

Fig. 5. Offline selection of $N_s = 1$ solution for Scenario-I, II and III for SYM-PART problem.

be both visualized in a scatter plot. Consider a series of increasing Euclidean distance thresholds $E = \{\epsilon_j; j = 1, 2, \ldots, N_\epsilon\}$. Here, we use ϵ_j as 0.1 to 0.8 in uniform increments of 0.1. The main idea is to divide the *normalized* objective and variable spaces using these thresholds as radii, and count the number of solutions that exist in the sub-spaces to estimate the distribution of the solutions. For example, let us say that we are interested in finding the $T1(\mathbf{x})$ measure for solution $\mathbf{x}$, marked in red in Fig. 2(a). In the variable space, the different thresholds are shown using circles of increasing radii. Likewise, the normalized objective space is presented in Fig. 2(b) with different colors indicating the different regions that fall within the given thresholds.

For each threshold ϵ_j, we identify three interim sets of solutions:

- The set of solutions that are within ϵ_j distance from $\mathbf{f}(\mathbf{x})$ in the normalized objective space. We denote this set as $S_{j,f}$.
- The set of solutions that are within ϵ_j distance from $\mathbf{x}$ in the normalized variable space. We denote this set as $S_{j,x\leq}$.
- The set of solutions that are beyond ϵ_j distance from $\mathbf{x}$ in the normalized variable space. We denote this set as $S_{j,x>}$.

From the above sets, the solutions that are suitable for Scenario-II can be thought of as those that belong to both $S_{j,f}$ and $S_{j,x\leq}$; since they are close to each other in both objective and variable spaces relative to threshold ϵ_j. This set is denoted as $S_{j,T1}$. The number of elements in the set $(N_{j,T1})$ is shown in Fig. 2(c). It is evident that the higher these numbers are, more options one has in the vicinity of the given solution in the variable space to find solutions with similar performance in the objective space. Thus, a sum of these numbers, i.e., $N_{T1}(\mathbf{x}^i) = \sum_{j=1}^{N_\epsilon} N_{j,T1}$, can give an overall indication of the robustness of the solution $\mathbf{x}^i$ under consideration. To capture this in a normalized form across the set of solutions A, the final value of $T1$ can be calculated as shown in

Eq. (2). From this expression, $T1$ varies between 0 and 1, with a higher value indicating higher preference of a solution with regards to Scenario-II.

$$T1(\mathbf{x}^i) = \frac{N_{T1}(\mathbf{x}^i) - \min\limits_{\mathbf{x}\in A}(N_{T1}(\mathbf{x}))}{\max\limits_{\mathbf{x}\in A}(N_{T1}(\mathbf{x})) - \min\limits_{\mathbf{x}\in A}(N_{T1}(\mathbf{x}))} \tag{2}$$

On the other hand, the solutions suitable for Scenario-III are those that belong to both $S_{j,f}$ and $S_{j,x>}$, since they are close in the objective space but far in the variable space, relative to threshold ϵ_j. This set is denoted as $S_{j,T2}$. The number of elements in the set ($N_{j,T2}$) are shown in Fig. 2(d). In this case, higher numbers would imply that a DM has more diverse options (equivalent designs) that are far apart in the variable space that have similar performance in the objective space. If we denote the sum of these numbers as $N_{T2}(\mathbf{x}^i) = \sum_{j=1}^{N_\epsilon} N_{j,T2}$ for a solution $\mathbf{x}^i$, then $T2$ can be calculated in a similar way as $T1$, as shown in Eq. (3). A higher value of $T2$ indicates a higher preference for solutions sought in Scenario-III.

$$T2(\mathbf{x}^i) = \frac{N_{T2}(\mathbf{x}^i) - \min\limits_{\mathbf{x}\in A}(N_{T2}(\mathbf{x}))}{\max\limits_{\mathbf{x}\in A}(N_{T2}(\mathbf{x})) - \min\limits_{\mathbf{x}\in A}(N_{T2}(\mathbf{x}))} \tag{3}$$

Algorithm 1 Calculation of $T1$ and $T2$ metrics for a given (normalized) solution set A

Input: Set of solutions (archive) A, Threshold values $E = \{\epsilon_i; i = 1, 2, \ldots N_\epsilon\}$
Output: $T1, T2$ measures for each solution in the archive A

for $i = 1$ to $|A|$ **do**
  $\mathbf{x}^i = A[i]$; ▷ i^{th} solution in the archive
  for $j = 1$ to N_ϵ **do**
    $S_{j,f} = \{\mathbf{x} | \{\mathbf{x} \in A; \mathbf{x} \neq \mathbf{x}^i \& ||\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{x}^i)|| \leq \epsilon_j\}$ ▷ Solutions within ϵ_j in objective space
    $S_{j,x\leq} = \{\mathbf{x} | \mathbf{x} \in A; \mathbf{x} \neq \mathbf{x}^i \& ||\mathbf{x} - \mathbf{x}^i|| \leq \epsilon_j\}$ ▷ Solutions within ϵ_j in variable space
    $S_{j,x>} = \{\mathbf{x} | \mathbf{x} \in A \& ||\mathbf{x} - \mathbf{x}^i|| > \epsilon_j\}$ ▷ Solutions further than ϵ_j in variable space
    $S_{j,T1} = S_{j,f} \cap S_{j,x\leq}$; $N_{j,T1} = |S_{j,T1}|$;
    $S_{j,T1} = S_{j,f} \cap S_{j,x>}$; $N_{j,T2} = |S_{j,T2}|$
  end for
  $N_{T1}(\mathbf{x}^i) = \sum_{j=1}^{N_\epsilon} N_{j,T1}$; $N_{T2}(\mathbf{x}^i) = \sum_{j=1}^{N_\epsilon} N_{j,T2}$
end for
for $i = 1$ to $|A|$ **do**
  $\mathbf{x}^i = A[i]$;
  $T1(\mathbf{x}^i) = \frac{N_{T1}(\mathbf{x}^i) - \min(N_{T1}(\mathbf{x}))}{\max(N_{T1}(\mathbf{x})) - \min(N_{T1}(\mathbf{x}))}$; $T2(\mathbf{x}^i) = \frac{N_{T2}(\mathbf{x}^i) - \min(N_{T2}(\mathbf{x}))}{\max(N_{T2}(\mathbf{x})) - \min(N_{T2}(\mathbf{x}))}$
end for

3.3. Online identification of SOIs

Having illustrated various quantitative measures offline to rank a given set of solutions, we now construct an algorithmic framework to identify the SOIs and intensify search within an *online* setting. That is, we attempt to bias the search by incorporating these measures so that it intensifies the density of the solutions around the SOIs instead of trying to cover the entire PF uniformly.

The overall framework is presented in Algo. 2. The inputs to the algorithm include the problem details, the scenario specified by the DM and the algorithmic parameters such as population size, function evaluation budget and evolution/variation parameters.

The algorithm begins with the generation of an initial set of N_p solutions using Latin Hypercube Sampling (LHS) design. The objectives and constraint(s) are evaluated for each of these solutions using the given problem formulation. An archive (A) of all evaluated solutions so far is maintained and updated whenever a new solution is evaluated. The current population then uses the variation operators (crossover, mutation) to generate N_p offspring solutions. The parents are selected from the current population using binary tournament, and recombined using simulated binary crossover (SBX) and polynomial mutation (PM) [49]. The unique offspring solutions with respect to the existing archive A are then fully evaluated. Thereafter, the evaluated child solutions are added to the archive A. The solutions in A then undergo a ranking process as outlined in Algo. 3.

Algorithm 2 Algorithmic framework for online identification of SOIs

Input: Problem formulation, DM's preferred Scenario (I, II or III), number of SOIs N_s, algorithmic parameters: population size N_p, Maximum evaluation count G, Crossover, mutation probabilities (P_c, P_m), Crossover, mutation index (I_c, I_m)
Output: Solutions for DM corresponding to Scenarios I, II or III.

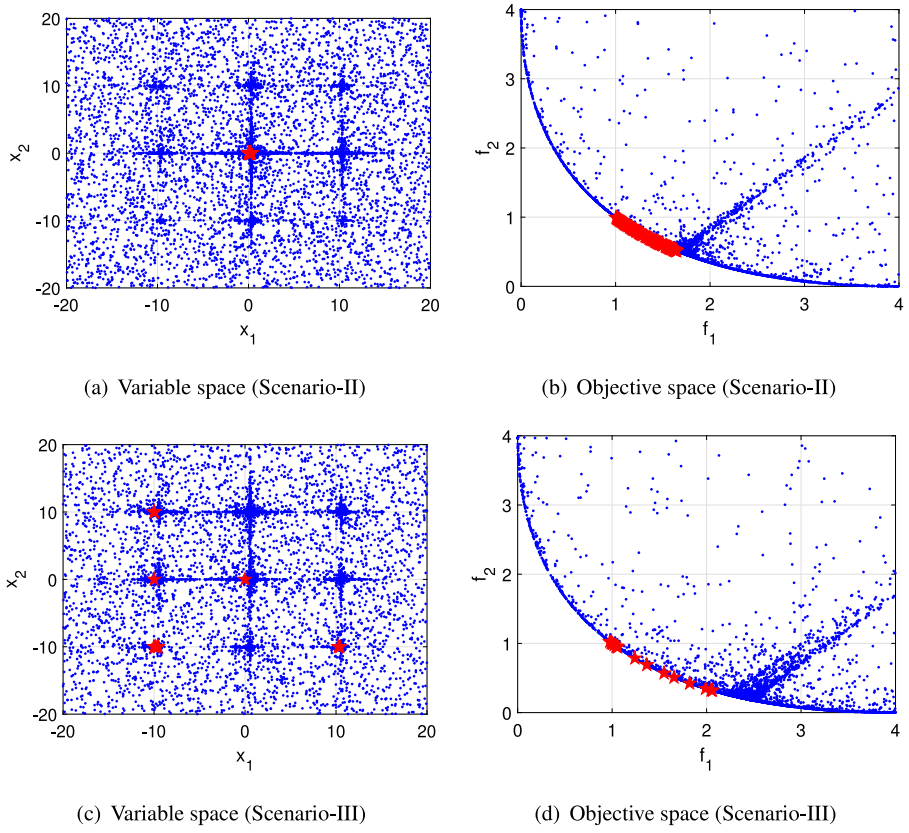
1: Initialize function evaluation count $g = 0$, archive of evaluated solutions $A = \emptyset$, generation counter $t = 1$
2: Generate initial population P_1 using LHS and evaluate it. Update g, A
3: Rank the archive A as per Scenario I, II or III as presented in Algo. 3.
4: **while** $g < G$ **do**
5: Generate child population C_t from P_t using evolutionary operators
6: Evaluate the solutions in child population C_t. Update A (append evaluated C_t to it), Update g.
7: Rank the archive A as per Scenario I, II or III as presented in Algo. 3.
8: Keep the top N_p solutions from ranked A as members of the parent population of the next generation.
9: $t = t + 1$
10: **end while**
11: Output the top N_s solutions from the A as the *SOIs*.

Algorithm 3 Scheme for complete ordering of feasible solutions

Input: Set of solutions/archive A, spacing parameter (h) for DRS removal; number of SOI required (N_s)
Output: Ordered list of solutions.

1: Order solutions based on non-dominated sorting in the objective space. ▷ Q denotes solutions in the first front
2: Normalize objective values of A using Ideal and Nadir points (Z^I, Z^N with DRS removed)
Phase 1: Identification of a set of S_o containing N_s solutions for Scenario-I
3: **if** $|Q| \geq N_s$ **then**
4: Compute $L1$ and Φ for all solutions in Q.
5: Select N_s solutions from Q in S_o in the descending order of Φ .
6: **else**
7: Select all solutions of Q in S_o.
8: Select the remainder of $N_s - |Q|$ solutions from the subsequent fronts progressively using sparse selection [50] to obtain total N_s solutions in S_o.
9: **end if**
10: The set S_o of N_s solutions is sent to Phase 2.
Phase 1: Identification of a set of S_o containing N_s solutions for Scenario-II and III
11: Find the non-dominated set Q' of $L1$ and $T1$ (for Scenario-II) or $L1$ and $T2$ (for Scenario-III) among the solutions in the non-dominated solution set Q.
12: **if** $|Q'| > N_s$ **then**
13: Use sparse selection [50] to choose top N_s solutions from Q'. The set is denoted as S_o.
14: **else**
15: Select all solutions of Q' in S_o.
16: Select the remainder of $N_s - |Q'|$ solutions from the subsequent fronts progressively using sparse selection [50] to obtain total N_s solutions in S_o.
17: **end if**
Phase 2: Ranking of the remaining solutions
18: For each solution in $A \setminus S_o$, identify its closest point in S_o based on Euclidean distance in the normalized objective space.
19: Sort the above solutions in the ascending order of the minimum Euclidean distance computed.
20: Identify extreme solutions (minimum in each objective) $Q_e \in Q$
21: Construct the final order as: [$S_o > Q_e >$ order based on line 19] .

The principle of ordering solutions is at the core of the online search process and comprises of two phases: Phase 1 which selects the required number of N_s solutions and Phase 2 where the remaining solutions are ordered. In Scenario-I, if $Q \geq N_s$, then N_s solutions are identified from Q in the descending order of Φ (Algo. 3 line 5). In case $Q < N_s$, then all solutions in Q are included in the selected set. The remaining $N_s - |Q|$



(a) Variable space (Scenario-II) (b) Objective space (Scenario-II)

(c) Variable space (Scenario-III) (d) Objective space (Scenario-III)

Fig. 6. Trade-off solutions obtained during online identification of SOIs for SYM-PART problem for Scenario-II and III. Blue dots show the archive of solutions explored during the search. The red stars denote the non-dominated solutions with respect to $L1$ and $T1$ (Scenario-II) and $L1$ and $T2$ (Scenario-III).

solutions are selected from the subsequent fronts progressively by applying a sparse selection technique called distance-based subset selection (DSS) [50] (Algo. 3 line 8).

For Scenario-II and III, the non-dominated set of solutions Q' with respect to $L1$ and $T1$ or $L1$ and $T2$ is first identified (line 11); where $Q' \subseteq Q$. Then, N_s solutions are identified from Q' using sparse selection (line 13) in $L1 - T1$ or $L1 - T2$ space. Note that DSS is a sequential process which starts by selecting the solution with best value of each objective considered. When we apply it for selection in $L1 - T1$ space (for Scenario-II) or $L1 - T2$ space (for Scenario-III), we start with the selecting the best $T1$ or best $T2$ respectively instead of the best $L1$ (line 13). This order affects the outcome only if $N_s = 1$. In effect, it is assumed that if $N_s = 1$ and best $L1$ solution is desired, then the chosen scenario will be Scenario-I. Once the set of N_s solutions, denoted as S_o, has been identified in Phase-1, the next step is to order the remaining solutions. The process relies on ensuring that solutions close to S_o are ranked higher than the ones that are far away from these. For a given solution, this proximity is calculated using its Euclidean distance to the closest solution in S_o in the objective space.

In the final ordering (line 21), the N_s solutions are given the first priority, since they need to be promoted as the SOIs. The next priority is given to the extreme solutions in the non-dominated front, identified as those that have the lowest value in any objective. This is done in order to maintain a good estimate of the range of the PF approximation for normalization during the search. Following that are all the remaining solutions, ordered based on proximity as discussed above.

Last but not the least, practical problems often involve constraints, and hence the solutions undergoing ranking may consist of both feasible and infeasible solutions. If infeasible solutions are present, we adopt the widely used parameter-less feasibility-first scheme [49] to handle them. Under this scheme, the feasible and infeasible solutions are first segregated. The feasible solutions undergo the ranking process as discussed above (Algo. 3), while the infeasible solutions are ranked simply based on their constraint violations. In the overall ranking, the ranked list of feasible solutions appear above the ranked list of infeasible solutions.

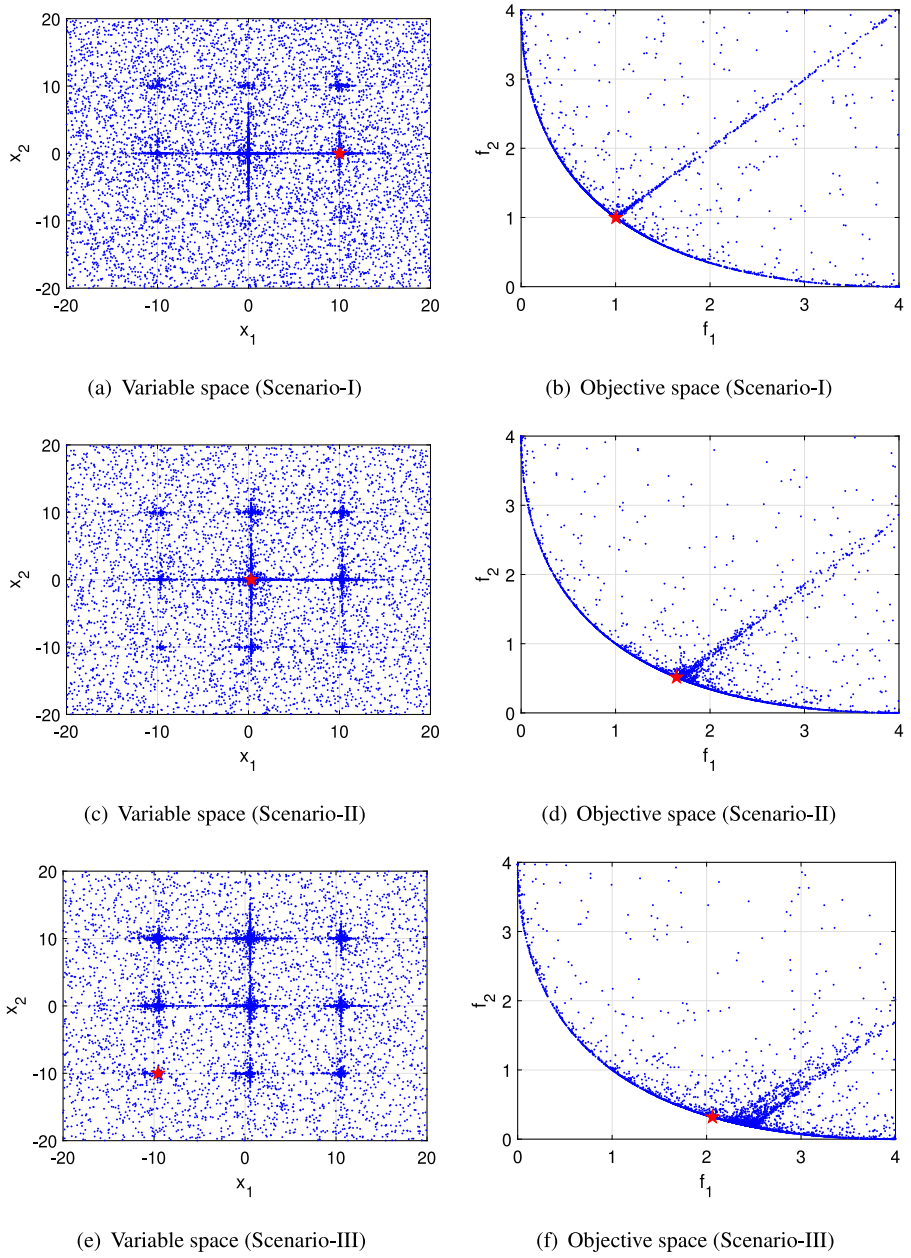
For research and benchmarking, the code for the above method can be obtained from the authors' website: http://www.mdolab.net/Ray/Research-Data/ASOC2021SOICode.zip.

4. Numerical experiments

4.1. Simple illustrative example

We use the multi-modal multi-objective optimization test problem referred to as SYM-PART [51] to highlight the solutions that would be delivered by the proposed algorithm for the different scenarios considered. SYM-PART is a two-variable, two-objective problem with a convex PF spanning [0,4] in both the objectives. The corresponding Pareto set (PS) in the variable space has 9 segments. Importantly, the solutions in any one of these segments are sufficient to generate the entire PF. The PF and PS of the problem are shown in Fig. 3.

We start with the offline identification first. A set of Pareto optimal solutions is first constructed by generating 45 points uniformly generated in each of the 9 segments of the PS, which translates to a total of 405 solutions on the PF. For Scenario-I, the solutions are ordered based on angle of influence. There are 9 solutions with the equal highest value of Φ, shown in Figs. 4(a) and 4(b). For Scenario-II and Scenario-III, since there are two measures involved, we first present the set of non-dominated solutions in $L1$ and $T1$ (Scenario-II) and $L1$ and $T2$ (Scenario-III) in Figs. 4(c) to 4(f).



(a) Variable space (Scenario-I) (b) Objective space (Scenario-I)

(c) Variable space (Scenario-II) (d) Objective space (Scenario-II)

(e) Variable space (Scenario-III) (f) Objective space (Scenario-III)

Fig. 7. Online identification of $N_s = 1$ solution for Scenario-I, II and III for SYM-PART problem.

If the DM is interested in $N_s = 1$ solution for Scenario-I, the solution identified by the proposed approach is marked with a red star in Fig. 5(a) in variable space and Fig. 5(b) in objective space, respectively. It is important to take note that the proposed algorithm (for Scenario-I), EvoKnee^r [25] or EMU^r [9] would all identify the same solution in the objective space for this scenario. However, since none of them consider variable space, the corresponding variable values could be from any of the 9 corresponding solutions in the PS.

If the DM is interested in Scenario-II, the solution in PS would come from the central segment as shown in Figs. 5(c) and 5(d), since it has the best $L1$ and is also the most robust based on $T1$ measure (as seen from Fig. 4(b)). Note that due to this, the solution selected for Scenario-I and Scenario-II will be the same (Figs. 5(b), 5(d)). For scenario-III, there is a trade-off between $L1$ and $T2$ (as seen from Fig. 4(e)) and the solution with the best $T2$ will be selected as shown in Fig. 5(e) and Fig. 5(f). In this case, the intent was that the selected designs be very different in the variable space.

For online search with $N_s = 1$, the non-dominated solutions in $L1-T1$ space for Scenario-II and $L1-T2$ space for Scenario-III are presented in red in Figs. 6(a)–6(d). The final solution identified using the proposed online approach is shown in red in Fig. 7. The online method for Scenario-I identified the solution with $\mathbf{x} = (10, 0)$, which is one among 9 possible solutions with the highest Φ in the objective space. For Scenario-II, the solution in the center of the variable space is identified inline with the observations in Fig. 4(c). Similarly, for Scenario-III, one of the solutions close to the one presented in Fig. 4(e) is identified, as shown in Fig. 7(e). Note that the range of non-dominated solutions (in terms of $L1 - T2$) is different in Fig. 4(e) and Fig. 6(d) due to different sampling (limited sampling precisely on the PS for offline case vs scattered large number of solutions for the online search). The selected solution (in Fig. 7(e)) is the one that has best $T2$ out of

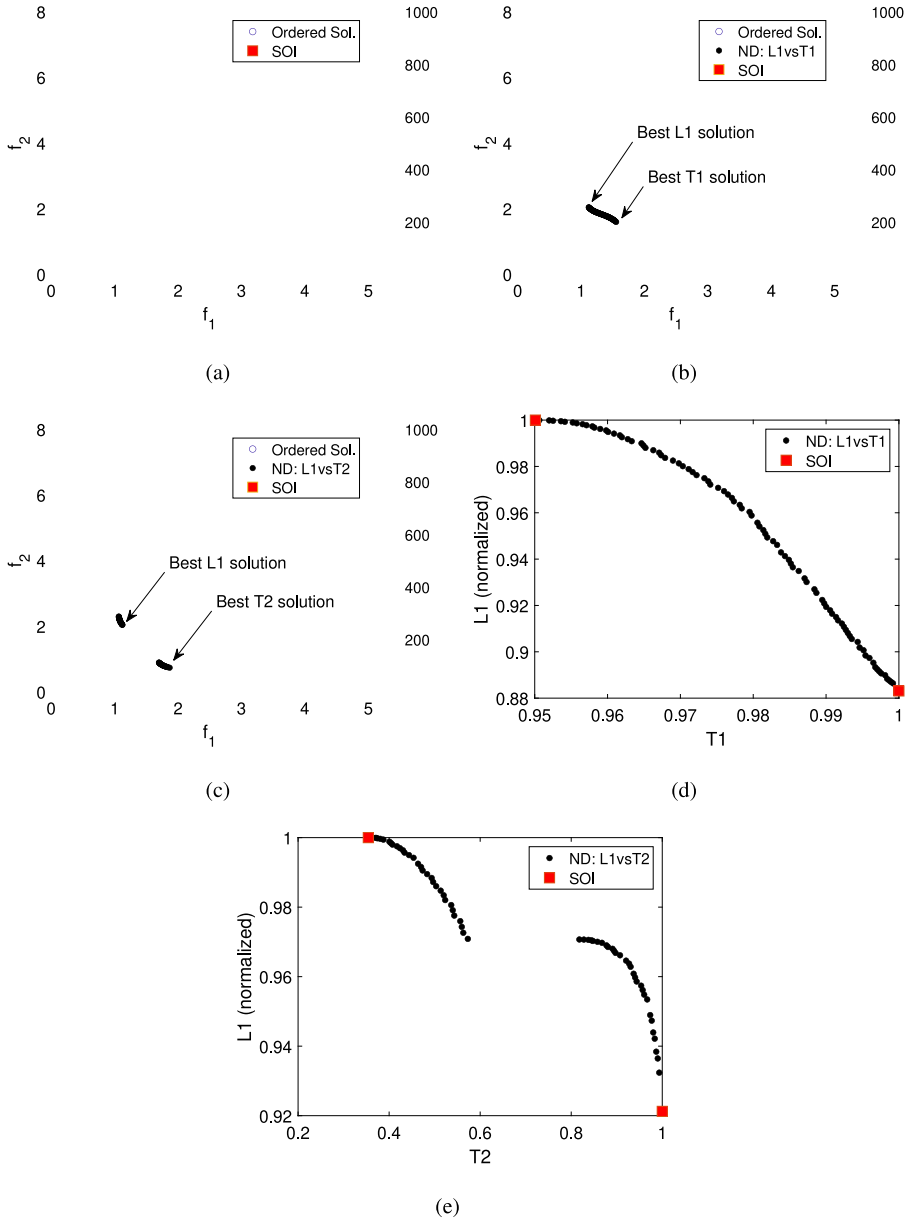


Fig. 8. Offline identification of $N_s = 2$ solutions for DO2DK (a–c) two SOIs and trade-off solutions identified for Scenarios-I, II and III respectively shown in the original objective space, (d) trade-off solutions between normalized $L1$ and $T1$ (e) trade-off solutions between normalized $L1$ and $T2$.

the red solutions shown in Fig. 6(d). Additionally, it can be seen that there is a high density of solutions around the chosen SOI; i.e., the best $L1$ (Φ), $T1$ and $T2$ solution for Scenario-I, II and III respectively. The observations confirm that the proposed online and offline selection delivers solutions as per the expectations for different scenarios. Additional observations regarding SYM-PART problem are provided in Section 3 of the supplementary material.

4.2. Offline identification of SOIs for benchmark problems

Having illustrated the working principles using SYM-PART, we now present the results obtained for offline identification of SOIs for a few benchmark problems. This is to assess the ability of the proposed measures in identifying the prescribed number of SOIs from a given set of solutions for each of the scenarios. This is relevant when the PF approximation has been already obtained through external means (e.g. conventional multi-objective algorithms) and the main intent is to help make an informed decision regarding the final design(s) to be implemented.

To illustrate the identification of SOIs, we consider one-variable versions of the two-objective problems (DO2DK with $N_s = 2$ and DEB2DK4 with $N_s = 4$), and two-variable version of three-objective problems (DEB3DK1 with $N_s = 1$ and DEB3DK4 with $N_s = 4$). For the one-variable problems, we sample 1000 solutions uniformly in the variable space, while for the two-variable problems we sample 10201 solutions uniformly in the variable space to construct the PF approximation. Given these sets of solutions, we proceed to identity the SOIs in an offline setting.

One can observe from Fig. 8(a) that the approach correctly identified two SOIs for the DO2DK problem for Scenario-I. The trade-off set of solutions for Scenario-II and III are presented in Figs. 8(b) and 8(c) respectively, of which the selected 2 SOIs are shown in red. The shift in the selection relative to Scenario-I is

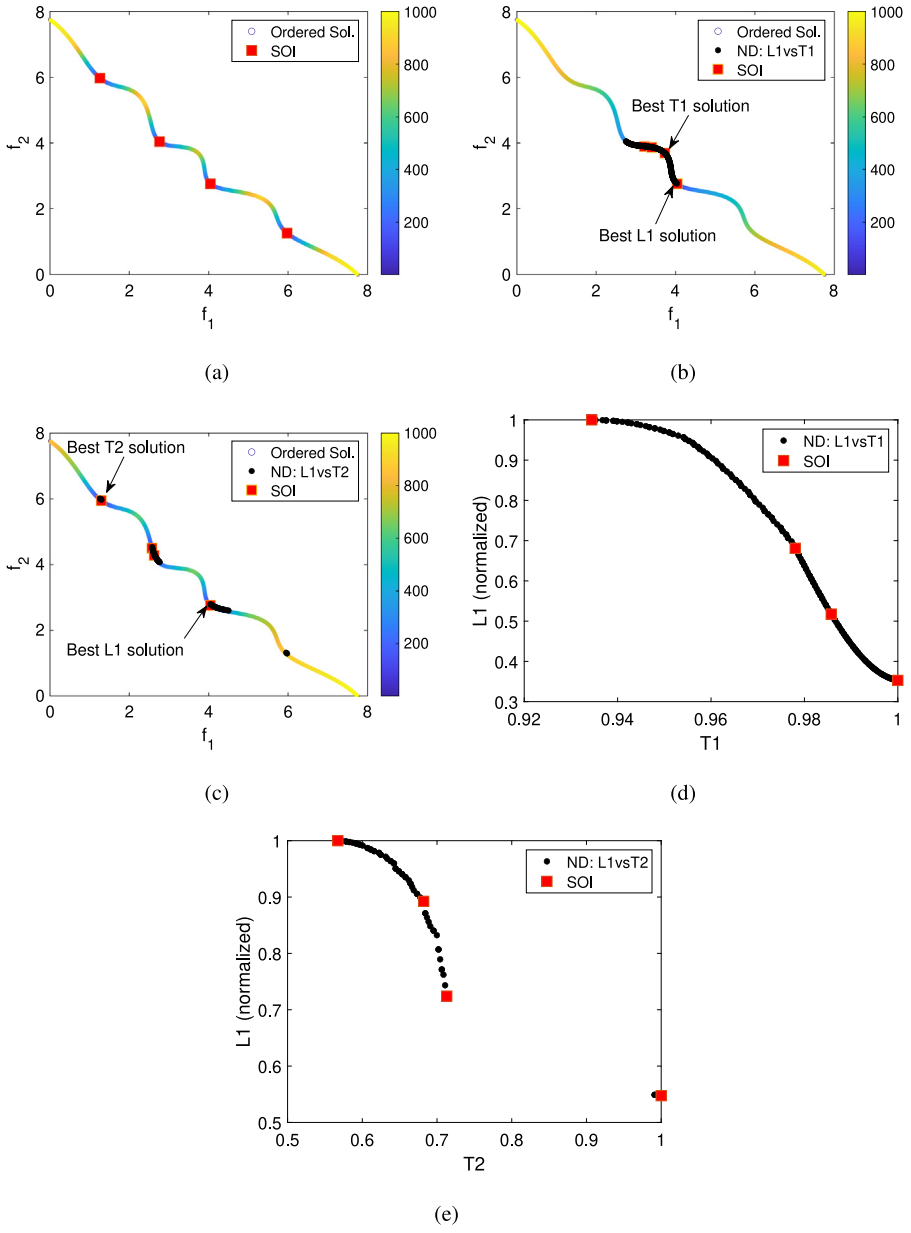


Fig. 9. Offline identification of $N_s = 4$ solutions for DEB2DK4 (a–c) four SOIs and trade-off solutions identified for Scenarios-I, II and III respectively shown in the original objective space, (d) trade-off solutions between normalized $L1$ and $T1$ (e) trade-off solutions between normalized $L1$ and $T2$.

evident here, emphasizing that the inclusion of variable space information will result in a different SOI compared to the conventional approaches wherein only the objective space is considered. One can observe that the solution with the best $T1$ measure (Scenario-II) lies in between the two SOIs (Scenario-I), while the solution with the best $T2$ measure (Scenario-III) is located close to (but not exactly on) the second SOI of Scenario-I. The trade-offs between $L1 - T1$ and between $L1 - T2$ measures are presented in Figs. 8(d) and 8(e) respectively. Similar and consistent observations as above can also be derived for DEB2DK4 with $N_s = 4$ from Fig. 9. Another notable observation from Fig. 9(d) is that some of the selected solutions are relatively closer to each other than others. This type of distribution can occur when an even number of solutions are selected using DSS. This limitation has been recently discovered in [52] and some advanced approaches to improve the DSS have been proposed that involve additional optimization of diversity criteria. While we have used DSS for the study presented herein, alternative subset selection measures such as those proposed in [52] can also be explored for potential improvements.

For the three-objective test case (DEB3DK1 with $N_s = 1$), the solutions identified are shown in Fig. 10 for each of the scenarios. Similarly for DEB3DK4, $N_s = 4$, the solutions identified for all scenarios and the trade-off solutions between $L1-T1$ and $L1-T2$ are shown in Fig. 11. It is also important to take note that the above observations are for a set of solutions that are uniformly distributed over the variable space. Since the measures are set-based, the locations of SOIs may vary if a different distribution of the solutions is used.

4.3. Online identification of SOIs for benchmark problems

Having demonstrated the offline selection, we now move on to the results obtained using the proposed approach for online identification of SOIs. This is relevant when the PF approximation is not available a priori and the aim is to bias the search to obtain

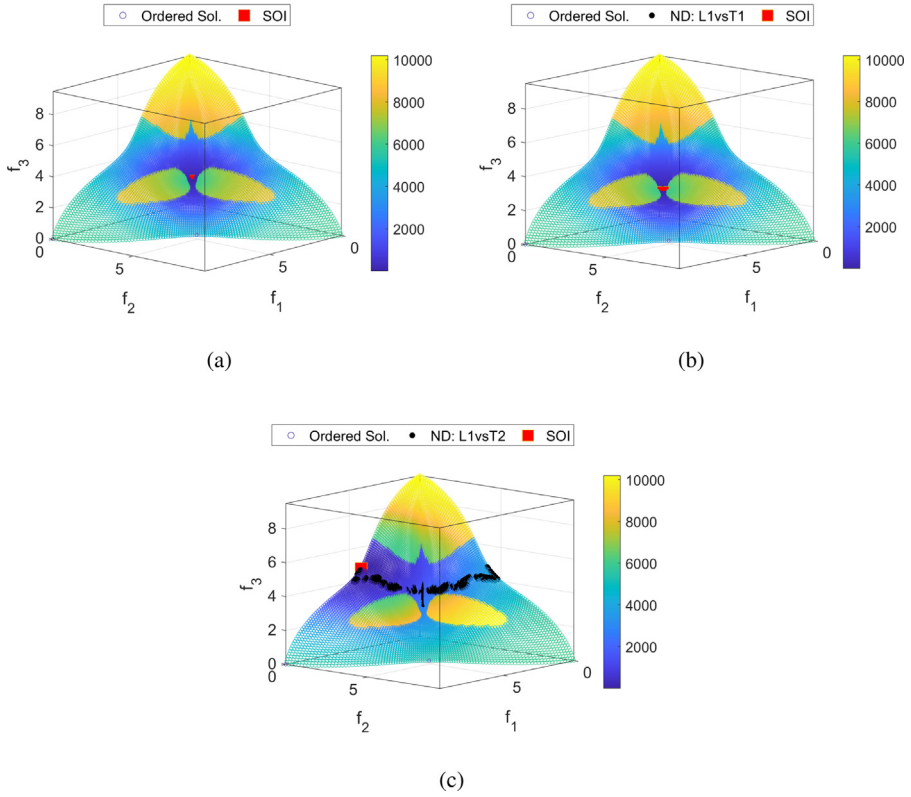


Fig. 10. Offline identification of $N_s = 1$ solution for DEB3DK1 (a) SOI for Scenario-I (b) SOI selected and trade-off solutions between normalized $L1$ and $T1$ for Scenario-II (c) SOI selected and trade-off solutions between normalized $L1$ and $T2$ for Scenario-III.

the SOIs and a high density of solutions around them rather than searching for the entire PF.

For all illustrations of online search, a population size of $N_p = 100$ was used and the computational budget was set to $G = 10{,}000$ function evaluations for the evolutionary search. The probability of simulated binary crossover was set to $P_c = 1.0$ and the probability of polynomial mutation was set to $P_m = 0.1$. The distribution indices of crossover (I_c) and mutation (I_m) were both set to 20. For DO2DK and DEB2DK4 problems, the number of variables is set as 30 and problem-specific parameters K and s are set to $\{4, 1\}$ and $\{4, 0\}$ respectively. These settings correspond to $N_s = 2$ and 4 knee points by design, respectively. Likewise, we use number of variables as 30 for the three-objective problems (DEB3DK1). Its parameters K and s are set to $\{1, 0\}$, which correspond to $N_s = 1$ knee point by design. For the welded beam design problem [7] and the wind turbine design optimization problem [53], N_s was set to 1. The spacing parameter used for generation of the weight vectors were set as 99, 21 and 15 for two-, three- and five-objective optimization problems, respectively.

For DO2DK, the approach delivered the correct set of SOIs for Scenario-I as presented in Fig. 12(a). The trade-off set of solutions for Scenario-II and III along with the selected N_s solutions with best $L1$, $T1$ and $T2$ measure are presented in Fig. 12(b) and Fig. 12(c) respectively. Furthermore, the trade-offs between $L1, T1$ and $L1, T2$ are shown in Figs. 13(a)–13(b), and the corresponding distributions of solutions around the best $L1 - T1$ and $L1 - T2$ are shown in Figs. 13(c)–13(d). It can be noted that the solution with the best $L1$, $T1$ and $T2$ are different solutions and in particular the solution with the best $L1$ measure has a poor $T2$ measure. This effectively means that if the DM selects best $L1$ solution for implementation, there are no other equivalent designs that are significantly different in the variable space as can be seen from Fig. 13(d).

As for the next problem, i.e. DEB2DK4 ($Ns = 4$), the results for Scenario-I, II and III are presented in Fig. 14. All four SOIs are identified correctly for Scenario-I in the objective space (Fig. 14(a)). Four SOIs obtained for Scenario-II and III are shown in Figs. 14(b)–14(c). For Scenario-III, the solution with the best $L1$ is also the one with the best $T2$.

It is worth noting that the number of knee points in a PF is usually not known a priori, and the number specified by DM may differ from it. For example, in this case there are four knee points by design, but the N_s specified may be different. For completeness, we have included further results in the supplementary material for the cases of $N_s = 2, 3, 5, 6$, for DEB2DK4 to illustrate the SOIs identified (both offline and online search in Figs. 2 to 9). Moreover, even if a higher N_s is desired in certain practical scenarios, the expected outputs from the proposed approach can be similarly visualized from the trade-offs between $L1 - T1$ or $L1 - T2$, since the N_s solutions will be identified using sparse selection from them.

For the three-objective problem DEB3DK1, the results are presented in Fig. 15. The observations are consistent with those from the two-objective examples.

4.4. Practical examples

While we have presented the performance of the approach for a number of unconstrained benchmark optimization problems, in this section we present its performance on two practical problems with constraints. These are the welded beam design and [54] and wind turbine design optimization problems [53].

The evolutionary search parameters for online search were kept the same as those for the benchmark problems in the previous section. The only change was for the case of wind turbine design optimization problem [53] for which the population size was set to 330 as suggested in [14].

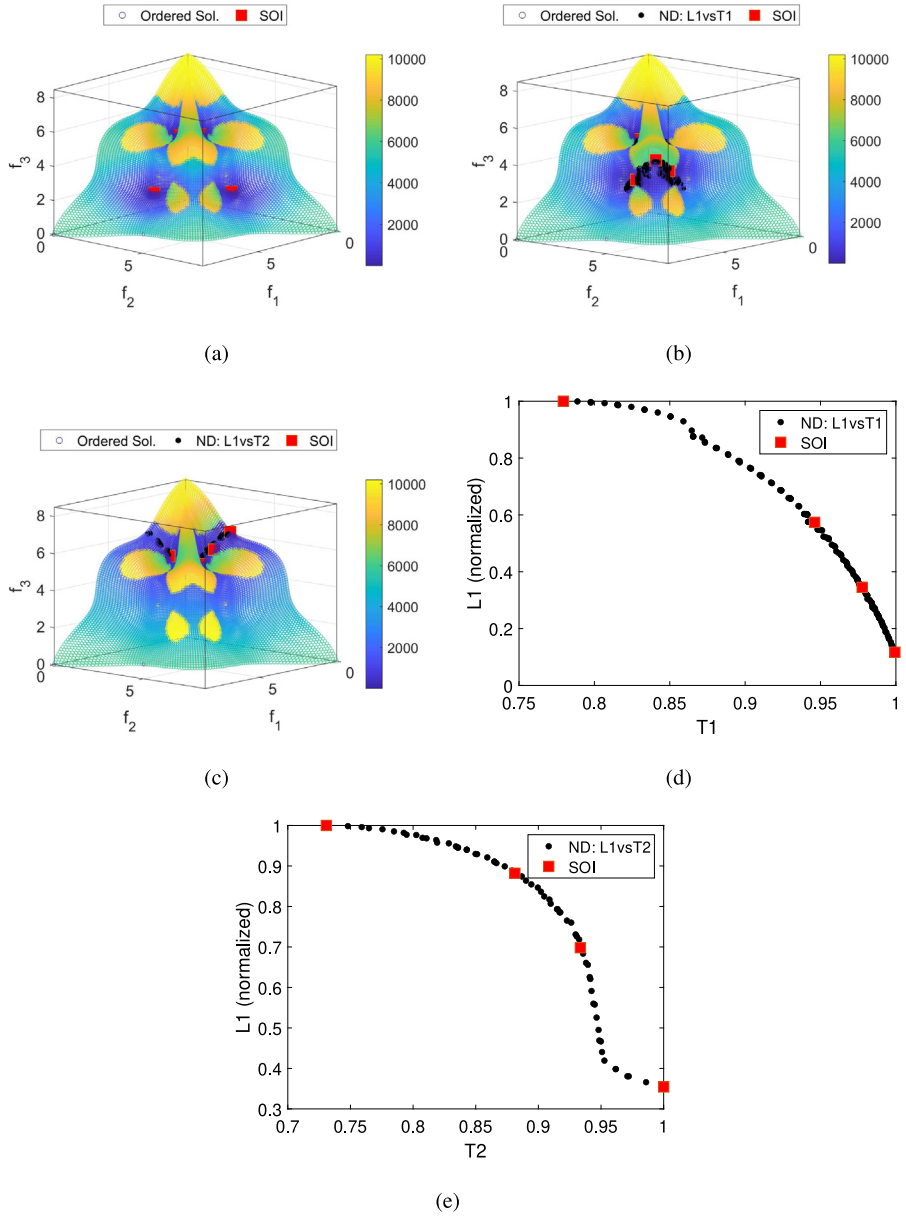


Fig. 11. Offline identification of solutions for DEB3DK4 ($N_s = 4$) (a–c) four SOIs and trade-off solutions identified for Scenarios-I, II and III respectively shown in the original objective space, (d) trade-off solutions between normalized $L1$ and $T1$ (e) trade-off solutions between normalized $L1$ and $T2$.

4.4.1. Welded beam design

The first problem is the welded beam design optimization problem which has previously been used for performance and robust optimization studies [54]. The problem involves four variables, two objectives and four constraints. The results obtained for all the three scenarios are presented in Fig. 16. The solution identified for Scenario-I in our study is in-line with solution identified in past studies [54]. Some of the previous studies have reported that the robust solutions for this problem lie in the dominated region [54]. In the current context however, the solutions for Scenario-II are selected from within the non-dominated front. Therefore, the obtained solution lies on the non-dominated front in Fig. 16(b). It is also important to take note that since the approach has the ability to intensify search around certain solutions during the course of search, robustness of such solutions can be intrinsically computed without the need for additional sampling and design evaluations to estimate the expected values. This offers an advantage over other robust optimization strategies that employ neighborhood sampling to estimate robustness of solutions. The solution with the best $T2$ measure on the non-dominated front is shown in Fig. 16(c).

4.4.2. Wind turbine design optimization

Next, we investigate the performance of the approach using a practical wind turbine design optimization problem. This problem was made available as part of a competition (Third Evolutionary Computation Competition 2019) by the Japanese Society of Evolutionary Computation (JSEC)), the detailed version of which can be found on competition website [53]. This five-optimization problem involves 32 variables and 22 constraints. For evaluating the objectives and constraints of the problem, multiple simulation

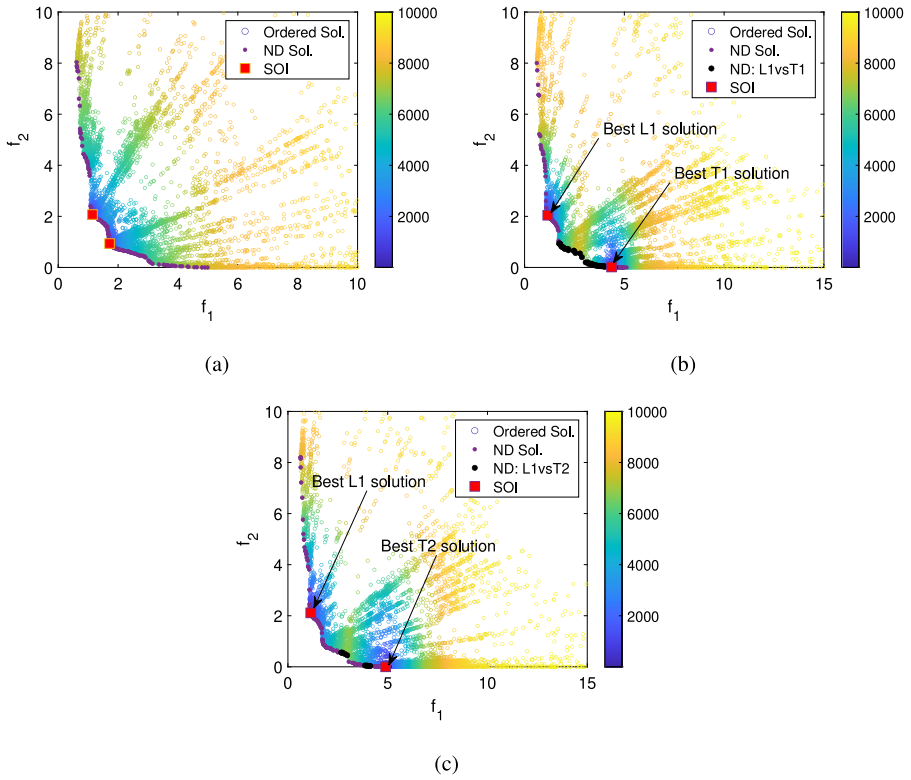


Fig. 12. Online identification for $N_s = 2$ solutions for DO2DK (a) Scenario-I, (b) Scenario-II and (c) Scenario-III.

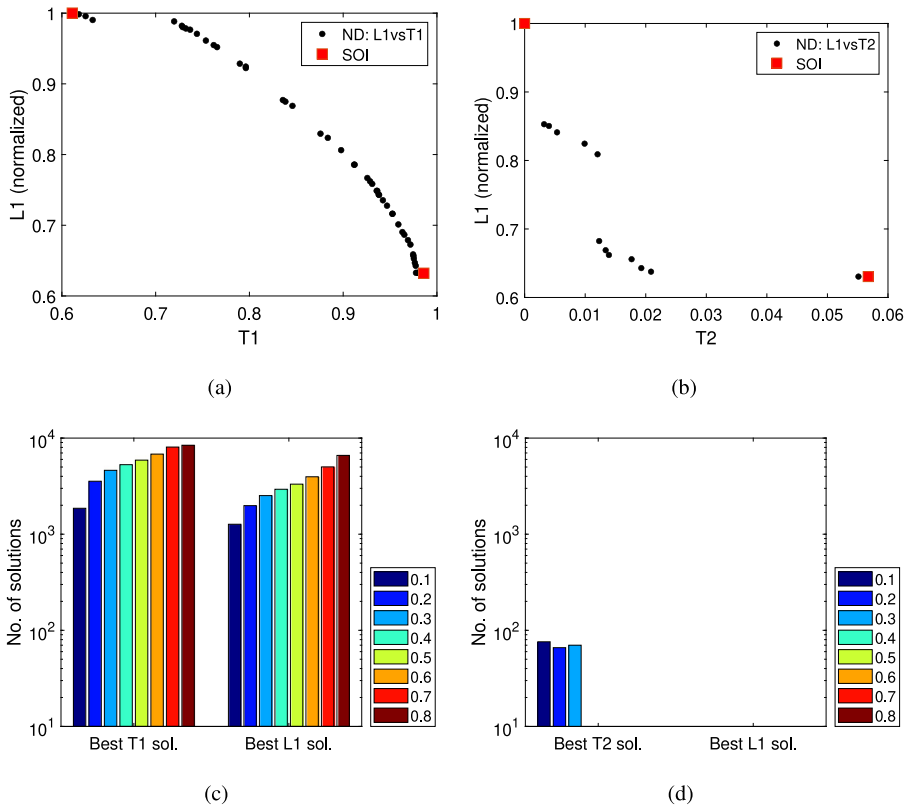


Fig. 13. Online identification of $N_s = 2$ solutions for DO2DK contd. (a) trade-offs between normalized $L1$ and $T1$, (b) trade-offs between normalized $L1$ and $T2$, (c) distribution of solutions in the archive around the solutions with best $L1$ and best $T1$ (Scenario-II), (d) distribution of solutions in the archive around the solutions with best $L1$ and best $T2$ (Scenario-III).

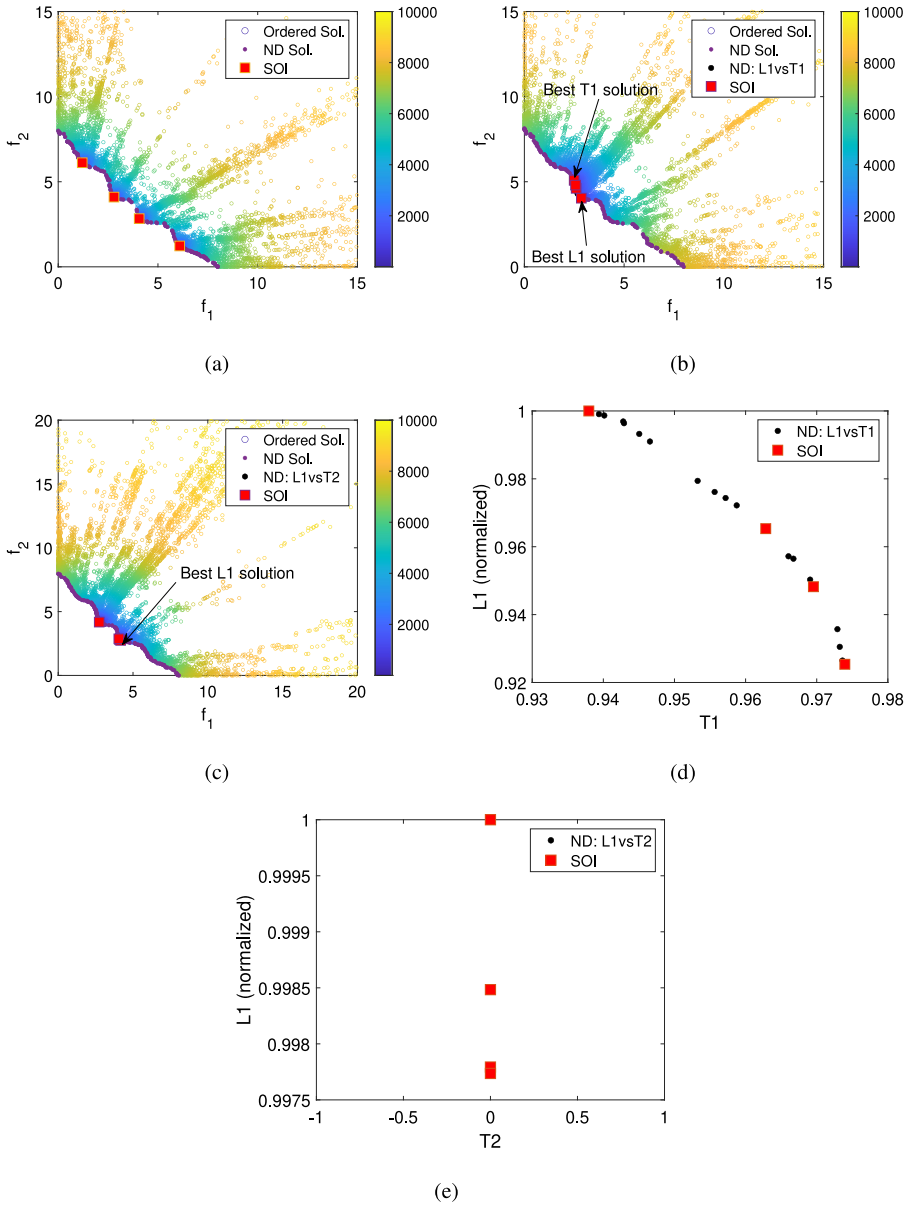


Fig. 14. Online identification of $N_s = 4$ solutions for DEB2DK4 (a–c) four SOIs and trade-off solutions identified for Scenarios-I, II and III respectively shown in the original objective space, (d) trade-off solutions between normalized $L1$ and $T1$ (e) trade-off solutions between normalized $L1$ and $T2$ (trade-off consists of only one solution in this case and therefore three SOIs are selected from subsequent fronts).

tools such as 'WISDEM', 'OpenMDAO' are used. The design constraints involve the aspects such as material strength and natural frequency of the body parts, ground clearance and maximum blade deflection, noise restrictions, production and transportation constraints, etc. It takes about three seconds to evaluate one candidate design, and therefore it takes more than eight hours to complete one optimization run; making the problem computationally expensive. For this problem, the performance of the solution with the best $L1$ measure is presented in Fig. 17(a). The distribution of the solutions around this solution is presented in Figs. 17(b) and 17(c), which is used for computing its $T1$ and $T2$ measures.

For Scenario-II, the spider plot of the solution with best $L1$ and $T1$ is presented in Fig. 18(a). For Scenario-III, the spider plot of the solution with best $L1$ and $T2$ measure is presented in Fig. 18(b). While the solution with the best $L1$ obtained for Scenario-II and Scenario-III are not identical, they have similar nature and such variations are expected due to stochastic nature of the underlying optimization algorithm.

5. Summary and future work

This paper presents a study towards identifying solutions of interest based on different scenarios that a decision-maker might be interested in. Unlike majority of existing works that consider only objective space while suggesting potential SOIs, this work considers three different scenarios. The Scenario-I seeks prescribed number (N_s) of solutions based on the normalized net gain $L1$ and angle of influence Φ in the objective space. Next, to incorporate variable space information in decision-making, we proposed new measures $T1$ and $T2$. These measures were targeted towards situations where the DM is interested in N_s solutions with a preference for robustness (Scenario-II) or finding diverse equivalent designs in the variable space (Scenario-III).

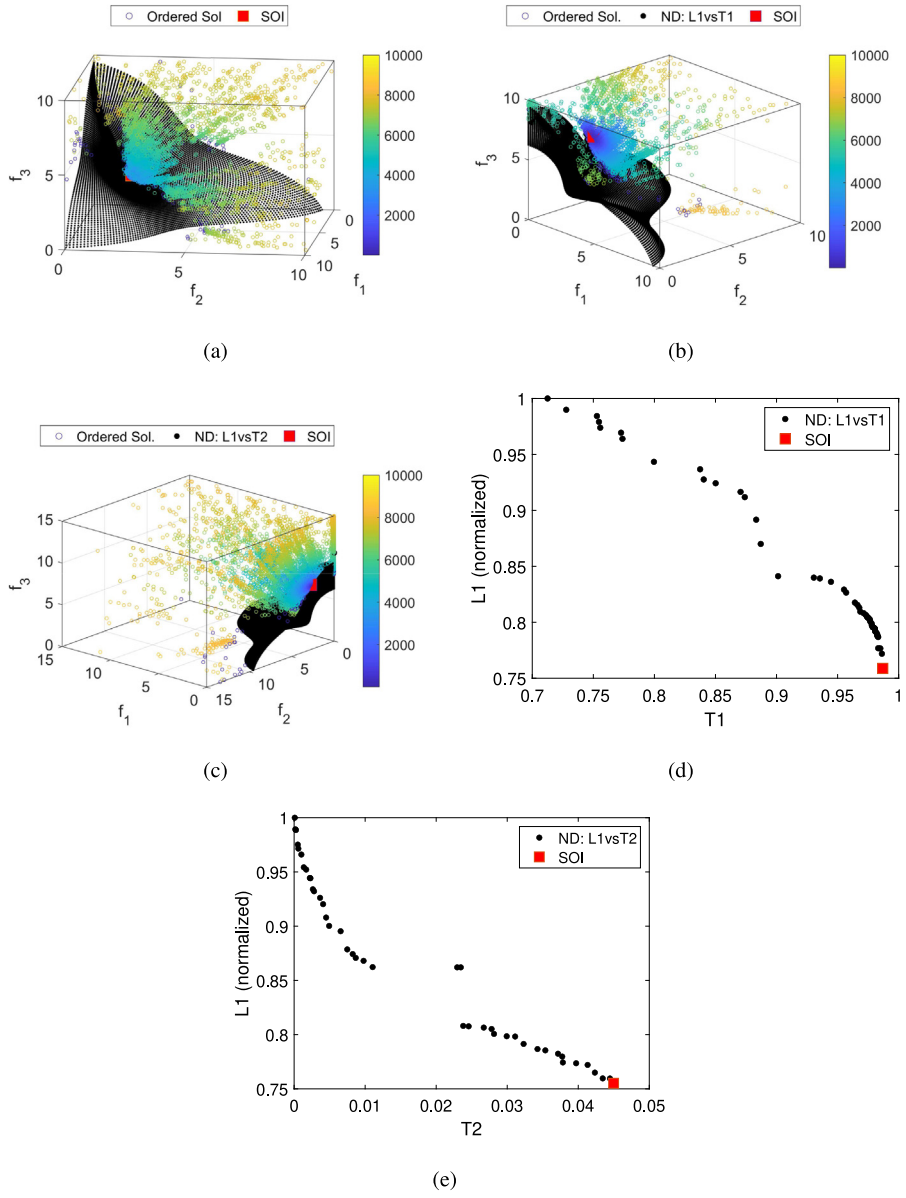


Fig. 15. Online identification of $N_s = 1$ solution for DEB3DK1 (a–c) Solutions obtained for Scenarios-I–III, (c) trade-off set of solutions between $T1$ and normalized $L1$ for Scenario-II, (d) trade-off set of solutions between $T2$ and normalized $L1$ for Scenario-III.

The working principles of the above measures were demonstrated using offline identification of N_s solutions from a large set of solutions sampled on the PF/PS for range of benchmark problems. Moreover, using these measures as the basis, an algorithmic framework was constructed for online search for SOIs. The results on a number of benchmark problems demonstrated the capability of the proposed approach to identify SOIs as per the chosen scenario and intensify the search around them. Since real-world problems often involve constraints, we further demonstrated the performance of the proposed algorithm for such classes of problems. To this end, we solved a two-objective constrained welded beam design problem and a five objective wind turbine design optimization problem with twenty-two constraints using simulation-based evaluations.

While this work has introduced an approach to identify SOIs with both objective and variable space considerations, there are a number of areas for further development. For example, robustness in industry is often represented using sigma levels (e.g. six sigma) and while $T1$ measure resembles it in a simplified form, there is an opportunity for closer alignment of the $T1$ measure with sigma levels. The levels 0.1 through to 0.8 in the normalized space has been used to represent variations in both variable and objective spaces which can be mapped to sigma levels for easier comprehension and interpretation. Likewise, further improvements could be made in the $T2$ measure by considering other density estimation methods. Secondly, the use of axial weight vectors to remove DRS in the current form inherently assumes the trade-off surfaces are of a particular form and the scheme also requires the number of weight vectors as an input. Both of these choices can affect the performance as the $L1$ measure relies on normalized objective vectors. The sparse selection could also be improved by using more advanced versions recently proposed [52] that overcome some limitations of DSS [50].

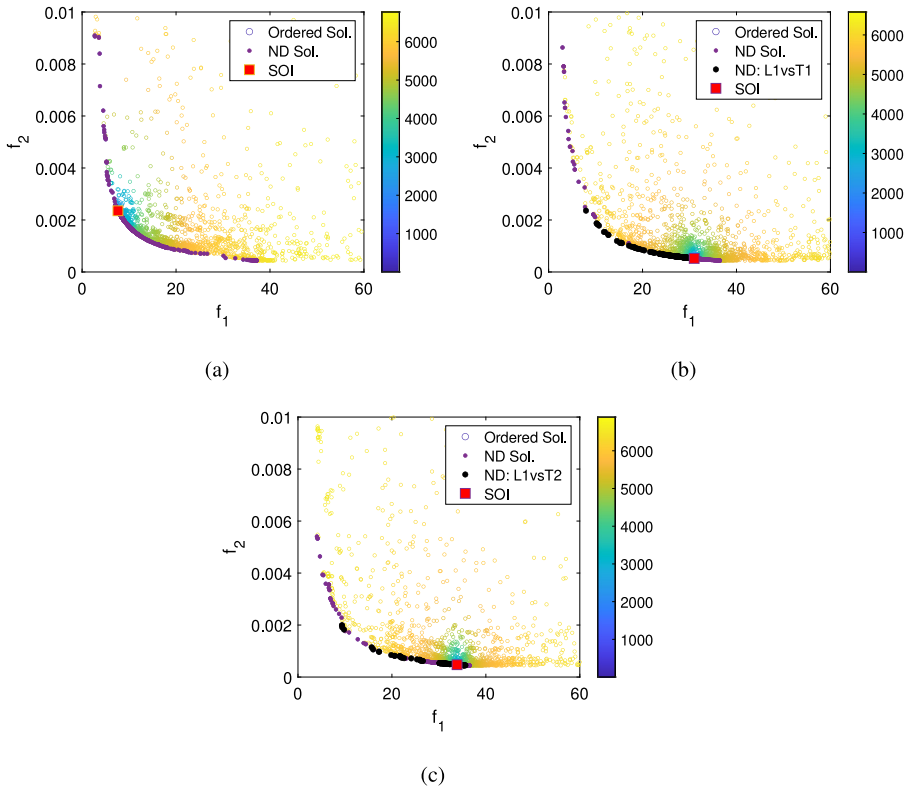


Fig. 16. Online identification of $N_s = 1$ solution for the welded beam design problem (a) Scenario-I (b) Scenario-II (c) Scenario-III.

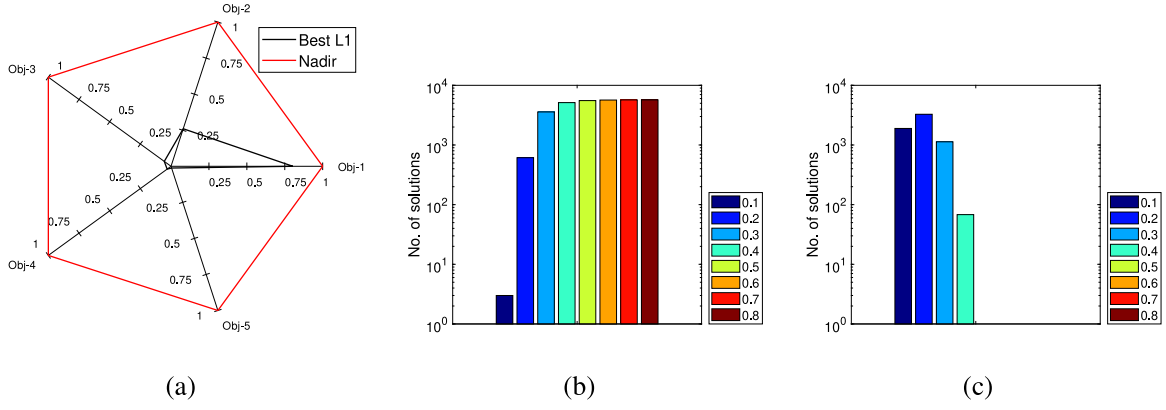


Fig. 17. Wind turbine design problem for Scenario-I (a) objective values (normalized) of best $L1$ solution (b–c) distribution of solutions around the best $L1$ solution to calculate its $T1$ and $T2$.

Finally, it would be interesting to apply and evaluate the proposed approach on a range of other practical problems.

CRediT authorship contribution statement

Tapabrata Ray: Conceptualization, Methodology, Software, Writing – original draft. **Hemant Kumar Singh:** Conceptualization, Methodology, Writing – original draft. **Kamrul Hasan Rahi:** Software, Writing – original draft. **Tobias Rodemann:** Conceptualization, Writing – review & editing. **Markus Olhofer:** Conceptualization, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

The authors acknowledge the financial support from Honda Research Institute Europe.

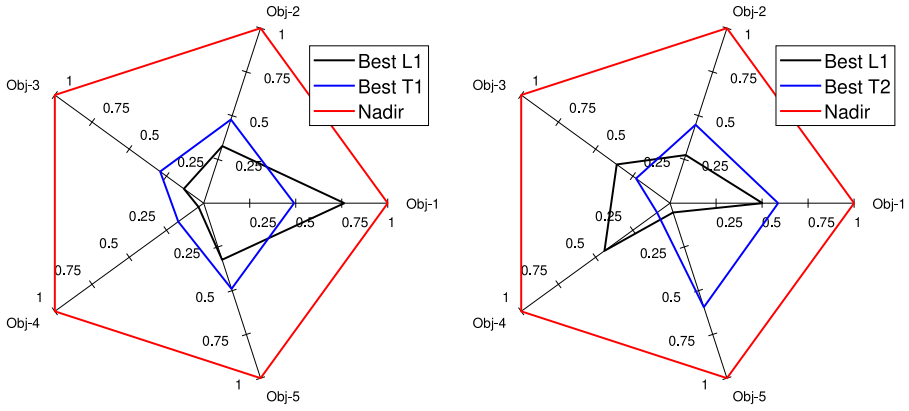


Fig. 18. Wind turbine design problem (a) objective values (normalized) of best $L1$ and $T1$ solutions for Scenario-II (b) objective values (normalized) of best $L1$ and $T2$ solutions for Scenario-III.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.asoc.2022.108505.

References

[1] K.S. Bhattacharjee, H.K. Singh, T. Ray, Multiple surrogate-assisted many-objective optimization for computationally expensive engineering design, J. Mech. Des. 140 (5) (2018).
[2] J. Branke, B. Scheckenbach, M. Stein, K. Deb, H. Schmeck, Portfolio optimization with an envelope-based multi-objective evolutionary algorithm, Eur. J. Oper. Res. 199 (3) (2009) 684–693.
[3] A. Alarcon-Rodriguez, G. Ault, S. Galloway, Multi-objective planning of distributed energy resources: A review of the state-of-the-art, Renew. Sustain. Energy Rev. 14 (5) (2010) 1353–1366.
[4] H. Ishibuchi, N. Tsukamoto, Y. Nojima, Evolutionary many-objective optimization: A short review, in: IEEE Congress on Evolutionary Computation, CEC, IEEE, 2008, pp. 2419–2426.
[5] C.A.C. Coello, Evolutionary multi-objective optimization: A critical review, in: Evolutionary Optimization, Springer, 2003, pp. 117–146.
[6] B. Li, J. Li, K. Tang, X. Yao, Many-objective evolutionary algorithms: A survey, ACM Comput. Surv. 48 (1) (2015) 1–35.
[7] K. Deb, Multi-objective evolutionary optimization: Past, present, and future, in: Evolutionary Design and Manufacture, Springer, 2000, pp. 225–236.
[8] R.K. Ursem, P.D. Justesen, Multi-objective distinct candidates optimization: Locating a few highly different solutions in a circuit component sizing problem, Appl. Soft Comput. 12 (1) (2012) 255–265.
[9] K.S. Bhattacharjee, H.K. Singh, M. Ryan, T. Ray, Bridging the gap: Many-objective optimization and informed decision-making, IEEE Trans. Evol. Comput. 21 (5) (2017) 813–820.
[10] J. Branke, K. Deb, H. Dierolf, M. Osswald, Finding knees in multi-objective optimization, in: International Conference on Parallel Problem Solving from Nature, Springer, 2004, pp. 722–731.
[11] H.K. Singh, T. Ray, T. Rodemann, M. Olhofer, Identifying solutions of interest for practical many-objective problems using recursive expected marginal utility, in: Proceedings of the Genetic and Evolutionary Computation Conference Companion, 2019, pp. 1734–1741.
[12] M. Torres, D.A. Pelta, M.T. Lamata, R.R. Yager, An approach to identify solutions of interest from multi and many-objective optimization problems, Neural Comput. Appl. 33 (7) (2021) 2471–2481.
[13] G. Rudolph, M. Preuss, A multiobjective approach for finding equivalent inverse images of Pareto-optimal objective vectors, in: IEEE Symposium on Computational Intelligence in Multi-Criteria Decision-Making, MCDM, IEEE, 2009, pp. 74–79.
[14] T. Ray, H.K. Singh, A. Habib, T. Rodemann, M. Olhofer, Online intensification of search around the solutions of interest for multi/many -objective optimization, in: IEEE Congress on Evolutionary Computation, CEC, 2020, pp. 1–8.
[15] I. Das, On characterizing the "knee" of the Pareto curve based on normal-boundary intersection, Struct. Optim. 18 (2–3) (1999) 107–115.
[16] K.S. Bhattacharjee, H.K. Singh, T. Ray, A study on performance metrics to identify solutions of interest from a trade-off set, in: Australasian Conference on Artificial Life and Computational Intelligence, ACALCI, in: Lecture Notes in Computer Science, vol. 9592, 2016, pp. 66–77.
[17] H. Wang, M. Olhofer, Y. Jin, A mini-review on preference modeling and articulation in multi-objective optimization: current status and challenges, Complex Intell. Syst. 3 (4) (2017) 233–245.
[18] O. Schütze, M. Laumanns, C.A.C. Coello, Approximating the knee of an MOP with stochastic search algorithms, in: International Conference on Parallel Problem Solving from Nature, Springer, 2008, pp. 795–804.
[19] K. Deb, S. Gupta, Understanding knee points in bicriteria problems and their implications as preferred solution principles, Eng. Optim. 43 (11) (2011) 1175–1204.
[20] M. Braun, P. Shukla, H. Schmeck, Angle-based preference models in multi-objective optimization, in: International Conference on Evolutionary Multi-Criterion Optimization, Springer, 2017, pp. 88–102.
[21] K. Li, H. Nie, H. Gao, X. Yao, Knee point identification based on trade-off utility, 2020, arXiv preprint arXiv:2005.11600.
[22] G. Yu, Y. Jin, M. Olhofer, A method for a posteriori identification of knee points based on solution density, in: IEEE Congress on Evolutionary Computation, CEC, IEEE, 2018, pp. 1–8.
[23] W.-Y. Chiu, G.G. Yen, T.-K. Juan, Minimum manhattan distance approach to multiple criteria decision making in multiobjective optimization problems, IEEE Trans. Evol. Comput. 20 (6) (2016) 972–985.
[24] Y. Zhou, G.G. Yen, Z. Yi, A knee-guided evolutionary algorithm for compressing deep neural networks, IEEE Trans. Cybern. 51 (3) (2019) 1626–1638.
[25] K. Zhang, G.G. Yen, Z. He, Evolutionary algorithm for knee-based multiple criteria decision making, IEEE Trans. Cybern. 51 (2) (2021) 722–735.
[26] L. Rachmawati, D. Srinivasan, Multiobjective evolutionary algorithm with controllable focus on the knees of the Pareto front, IEEE Trans. Evol. Comput. 13 (4) (2009) 810–824.
[27] S. Bechikh, L. Ben Said, K. Ghédira, Searching for knee regions in multi-objective optimization using mobile reference points, in: Proceedings of the 2010 ACM Symposium on Applied Computing, ACM, 2010, pp. 1118–1125.
[28] S. Sudeng, N. Wattanapongsakorn, A decomposition-based approach for knee solution approximation in multi-objective optimization, in: IEEE Congress on Evolutionary Computation, CEC, IEEE, 2016, pp. 3710–3717.
[29] H.K. Singh, A. Isaacs, T. Ray, A Pareto corner search evolutionary algorithm and dimensionality reduction in many-objective optimization problems, IEEE Trans. Evol. Comput. 15 (4) (2011) 539–556.
[30] S. Sudeng, N. Wattanapongsakorn, Adaptive geometric angle-based algorithm with independent objective biasing for pruning pareto-optimal solutions, in: 2013 Science and Information Conference, IEEE, 2013, pp. 514–523.
[31] C. Ramirez-Atencia, S. Mostaghim, D. Camacho, A knee point based evolutionary multi-objective optimization for mission planning problems, in: Proceedings of the Genetic and Evolutionary Computation Conference, 2017, pp. 1216–1223.
[32] K.S. Bhattacharjee, H.K. Singh, T. Ray, Q. Zhang, Decomposition based evolutionary algorithm with a dual set of reference vectors, in: IEEE Congress on Evolutionary Computation, CEC, IEEE, 2017, pp. 105–112.

[33] H. Ishibuchi, Y. Setoguchi, H. Masuda, Y. Nojima, Performance of decomposition-based many-objective algorithms strongly depends on Pareto front shapes, IEEE Trans. Evol. Comput. 21 (2) (2016) 169–190.
[34] X. Zhang, Y. Tian, Y. Jin, A knee point-driven evolutionary algorithm for many-objective optimization, IEEE Trans. Evol. Comput. 19 (6) (2015) 761–776.
[35] J. Zou, C. Ji, S. Yang, Y. Zhang, J. Zheng, K. Li, A knee-point-based evolutionary algorithm using weighted subpopulation for many-objective optimization, Swarm Evol. Comput. 47 (2019) 33–43.
[36] F. Zou, G.G. Yen, L. Tang, A knee-guided prediction approach for dynamic multi-objective optimization, Inform. Sci. 509 (2020) 193–209.
[37] J. Maltese, B.M. Ombuki-Berman, A.P. Engelbrecht, Pareto-based many-objective optimization using knee points, in: IEEE Congress On Evolutionary Computation, CEC, IEEE, 2016, pp. 3678–3686.
[38] G. Yu, Y. Jin, M. Olhofer, Benchmark problems and performance indicators for search of knee points in multiobjective optimization, IEEE Trans. Cybern. 50 (8) (2019) 3531–3544.
[39] Y. Jin, J. Branke, Evolutionary optimization in uncertain environments-a survey, IEEE Trans. Evol. Comput. 9 (3) (2005) 303–317.
[40] H.-G. Beyer, B. Sendhoff, Robust optimization–a comprehensive survey, Comput. Methods Appl. Mech. Engrg. 196 (33–34) (2007) 3190–3218.
[41] B. Torstenfelt, A. Klarbring, Conceptual optimal design of modular car product families using simultaneous size, shape and topology optimization, Finite Elem. Anal. Des. 43 (14) (2007) 1050–1061.
[42] K. Willcox, S. Wakayama, Simultaneous optimization of a multiple-aircraft family, J. Aircr. 40 (4) (2003) 616–622.
[43] M. Asafuddoula, H.K. Singh, T. Ray, Six-sigma robust design optimization using a many-objective decomposition-based evolutionary algorithm, IEEE Trans. Evol. Comput. 19 (4) (2015) 490–507.
[44] R. Tanabe, H. Ishibuchi, A review of evolutionary multimodal multiobjective optimization, IEEE Trans. Evol. Comput. 24 (1) (2019) 193–200.
[45] H. Ishibuchi, T. Matsumoto, N. Masuyama, Y. Nojima, Effects of dominance resistant solutions on the performance of evolutionary multi-objective and many-objective algorithms, in: Proceedings of the 2020 Genetic and Evolutionary Computation Conference, 2020, pp. 507–515.
[46] Q. Yang, Z. Wang, H. Ishibuchi, It is hard to distinguish between dominance resistant solutions and extremely convex Pareto optimal solutions, in: International Conference on Evolutionary Multi-Criterion Optimization, EMO, 2021, pp. 3–14.
[47] I. Das, J.E. Dennis, Normal-boundary intersection: A new method for generating the Pareto surface in nonlinear multicriteria optimization problems, SIAM J. Optim. 8 (3) (1998) 631–657.
[48] J. Blank, K. Deb, P.C. Roy, Investigating the normalization procedure of NSGA-III, in: International Conference on Evolutionary Multi-Criterion Optimization, Springer, 2019, pp. 229–240.
[49] K. Deb, A. Pratap, S. Agarwal, T. Meyarivan, A fast and elitist multiobjective genetic algorithm: NSGA-II, IEEE Trans. Evol. Comput. 6 (2) (2002) 182–197.
[50] H.K. Singh, K.S. Bhattacharjee, T. Ray, Distance-based subset selection for benchmarking in evolutionary multi/many-objective optimization, IEEE Trans. Evol. Comput. 23 (5) (2018) 904–912.
[51] G. Rudolph, B. Naujoks, M. Preuss, Capabilities of EMOA to detect and preserve equivalent Pareto subsets, in: International Conference on Evolutionary Multi-Criterion Optimization, EMO, Springer, 2007, pp. 36–50.
[52] K. Shang, H. Ishibuchi, Y. Nan, Distance-based subset selection revisited, in: Proceedings of the Genetic and Evolutionary Computation Conference, 2021, pp. 439–447.
[53] T.J.S. of Evolutionary Computation, The 3rd evolutionary computation competition, 2019, http://www.jpnsec.org/files/competition2019/EC-Symposium-2019-Competition-English.html. (Online; Accessed 21 January 2020).
[54] K. Deb, H. Gupta, Introducing robustness in multi-objective optimization, Evol. Comput. 14 (4) (2006) 463–494.