One PLOT to Show Them All: Visualization of Efficient Sets in Multi-objective Landscapes

Lennart Schäpermeier$^{(\boxtimes)}$, Christian Grimme , and Pascal Kerschke

Statistics and Optimization, University of Münster, Münster, Germany
{schaepermeier,christian.grimme,kerschke}@uni-muenster.de

Abstract. Visualization techniques for the decision space of continuous multi-objective optimization problems (MOPs) are rather scarce in research. For long, all techniques focused on global optimality and even for the few available landscape visualizations, e.g., *cost landscapes*, globality is the main criterion. In contrast, the recently proposed *gradient field heatmaps (GFHs)* emphasize the location and attraction basins of local efficient sets, but ignore the relation of sets in terms of solution quality.

In this paper, we propose a new and hybrid visualization technique, which combines the advantages of both approaches in order to represent local and global optimality together within a single visualization. Therefore, we build on the GFH approach but apply a new technique for approximating the location of locally efficient points and using the divergence of the multi-objective gradient vector field as a robust second-order condition. Then, the relative dominance relationship of the determined locally efficient points is used to visualize the complete landscape of the MOP. Augmented by information on the basins of attraction, this *Plot of Landscapes with Optimal Trade-offs* (PLOT) becomes one of the most informative multi-objective landscape visualization techniques available.

Keywords: Multi-objective optimization · Continuous optimization · Visualization · Landscape analysis · Efficient sets

1 Introduction

Traditionally, the visualization of optimization problems in decision space is one of the basic approaches to investigate challenges of so-called functional landscapes and to design basic algorithmic principles. Hence, low dimensional visualization is used in text books [1,3] and algorithm research alike. For a single objective, each point in a continuous two-dimensional search space can be assigned with a function value which is interpreted as height. Overall, this leads to very natural notions of *mountains* and *valleys* for local maxima and minima, *ridges* for discontinuities, as well as *plateaus* for areas of equal height.

In evolutionary computation, many early theoretical results as well as later algorithmic concepts were first developed (and only successively generalized) by using low-dimensional visualizations of benchmark problems and their challenges. For continuous multi-objective optimization problems (MOPs), however, such straight-forward visualization techniques are not available. This is mainly rooted in the fact that MOPs comprise at least two contradicting objectives to be optimized simultaneously. Consequently, not a single or few global optimal solutions are sought but a set of optimal trade-off solutions – the so-called *Pareto set*. These solutions have as many objective function values (and thus height values) as objectives, which makes a standard landscape visualization infeasible.

For few (≤ 3) objectives, a classical visualization of the true or approximated set of efficient solutions, usually the Pareto front – the Pareto set's image in objective space – is used. However, by focusing purely on the objective space, one ignores all interaction effects from the MOP's input variables in the decision space. Compared to the single-objective case, this is like reducing the entire landscape to the function values of its optimal solutions, and plotting them on a one-dimensional scale. Almost no information on the structural properties of the problem landscapes can be derived from this. Consequently, only little is known on MOP landscape properties and almost no algorithmic design is based on comparable insights like in the single-objective case.

Although a straightforward visualization of MO landscapes is not available, there exist very few techniques for getting insights into these landscapes. The *cost landscapes* proposed by Fonseca [5] use a dominance ranking approach to evaluate each point in decision space w.r.t. the global optimal trade-offs. Although this delivers a kind of landscape in relation to the global optimum, it does not capture local optimal sets and their basins of attraction. An alternative visualization approach that explicitly addresses locality has been proposed by Kerschke and Grimme [10,12]. It produces (multi-objective) gradient field heatmaps (GFHs) using the Fritz-John (necessary) condition for identifying local optima [11,18]. The GFHs show local basins and locally efficient sets but have two drawbacks: they do not provide a ranking of local sets w.r.t the global set and indicate local efficiency only indirectly by the height. Within this work, we address the weaknesses of both approaches and contribute the following:

1. We propose a robust approach to determine locally efficient points explicitly. This includes a suitable second-order condition based on the divergence of the multi-objective gradient to confirm or exclude points that are considered to be locally optimal according to the first-order conditions.
2. Additionally, we combine and extend the two aforementioned state-of-the-art visualization approaches, i.e., the cost landscapes [5] and the gradient field heatmaps [12]. This leads to a far more informative visualization than any of these approaches taken by themselves offered before. We name this method *Plot of Landscapes with Optimal Trade-offs (PLOT)*. Besides locality, global relations of local optima and respective basins can now be captured in a single PLOT. For two-dimensional problems, PLOT delivers a complete picture of the problem landscape that can be interpreted almost as seamless as a single-objective landscape, merely relating to the multi-objective gradient.

The remainder of this work is organized as follows. Section 2 summarizes the background by first providing fundamental notations and definitions that are needed later, and afterwards discussing the status quo on the visualization of MO landscapes. Section 3 describes the new methodology to determine locally efficient points, while Sect. 4 describes the concept of merging the cost landscape approach and the gradient field heatmaps into PLOT. Finally, Sect. 5 evaluates our proposed PLOT approach by visualizing examples from current benchmark problems before Sect. 6 concludes the paper.

2   Background

2.1   Preliminaries on Multi-objective Optimization

For this work, we consider continuous MOPs $f : \mathbb{R}^p \rightarrow \mathbb{R}^k$ with search space parameter p, k objectives and feasible search space $\mathcal{X} := [\mathbf{l}, \mathbf{u}] \subseteq \mathbb{R}^p$:

$$f(\mathbf{x}) = (f_1(\mathbf{x}), \dots, f_k(\mathbf{x})) \overset{!}{=} \min \qquad \text{with} \qquad l_i \leq x_i \leq u_i, \ i = 1, \dots, p. \quad (1)$$

The solution of a MOP is the set $\mathcal{X}^* \subseteq \mathcal{X}$ of Pareto-optimal trade-offs, i.e., all points $\mathbf{x}^* \in \mathcal{X}$ for which there exist no $\mathbf{x}' \in \mathcal{X}$ with $f_i(\mathbf{x}') \leq f_i(\mathbf{x}^*)$ for all $i = 1, \dots, k$ and $f_i(\mathbf{x}') < f_i(\mathbf{x}^*)$ for at least one i (denoted as Pareto set). Thus, the aim of multi-objective (MO) optimization is to find all points in $\mathcal{X}$ that are not dominated by other points in the decision space. The image $f(\mathcal{X}^*)$ is called the Pareto front. *Local efficiency of a point* is defined in analogy to locality in single-objective optimization: given a non-empty ε-neighborhood $B_\varepsilon(\mathbf{x}) \subseteq \mathcal{X}$ around $\mathbf{x}$, no point $\mathbf{y} \in B_\varepsilon(\mathbf{x})$ dominates $\mathbf{x}$ [4]. Extending this definition to a set of locally efficient points, a *local efficient set* is a set of points, which is not dominated by other points in their ε-neighborhood [16]. In a somewhat differentiated view, we can further discriminate different local sets, if we consider connected subsets of points as separate local efficient sets like it is done in [13].

For the remainder of this paper, we will focus on two-dimensional bi-objective problems (i.e., $p = 2$ and $k = 2$) to enable visual representations of the MOPs. Although this may seem restrictive at first sight, it should be kept in mind that two-dimensional visualizations have substantially contributed to improving our understanding of the algorithmic search behavior in the single-objective case. Also, we will adopt the notion of locally efficient sets from [13] within this work.

The Fritz-John conditions are well known first-order conditions for continuous MOPs [18]. Given a MO function f defined as above, as well as inactive constraints, a first-order critical point $\mathbf{x}^*$ fulfills $\sum_{i=1}^{k} \lambda_i \nabla f_i(\mathbf{x}^*) = \mathbf{0}$ with $\lambda_i \geq 0$ for all $i = 1, ..., k$ and $\sum_{i=1}^{k} \lambda_i = 1$. Based on this, a MO gradient $\nabla f(\mathbf{x})$ can be defined, which is zero if these conditions are satisfied, and which points towards a common descent direction of the objectives otherwise ($-\nabla f(\mathbf{x})$ for minimization). For two objectives, a definition for the MO gradient is given by the sum of the normalized single-objective gradients:

$$\nabla f(\mathbf{x}) = \nabla f_1(\mathbf{x})/||\nabla f_1(\mathbf{x})|| + \nabla f_2(\mathbf{x})/||\nabla f_2(\mathbf{x})|| \qquad (2)$$

Following the MO gradient in a gradient-descent-like manner eventually leads into a local efficient set [9], i.e., a (possibly connected) set of locally efficient points. Note that, if for a point $\mathbf{x}$ the length of one of its single-objective gradients is zero, i.e., $||\nabla f_1(\mathbf{x})|| = 0$ or $||\nabla f_2(\mathbf{x})|| = 0$, the point fulfills the necessary condition for a local optimum in the single-objective as well as in the MO case. We therefore define $\nabla f(\mathbf{x}) := 0$ for such points.

2.2   Visualization of Continuous MOPs

Benchmark problems are usually designed to test the capabilities and limitations of a broad spectrum of (optimization) algorithms [23]. Aside from a pure performance comparison, the insights gained thereof are helpful for designing better algorithms. Here, 'better' depends on various aspects such as the application or considered performance measure. Due to the different goals, a variety of test suites have been proposed over the years – ranging from MOP [22], ZDT [25] and DTLZ [7] (which aim at posing challenges for MO evolutionary algorithms), over bi-objective BBOB [21] (which extends the gold-standard test suite from single-objective optimization to the bi-objective case), up to more recent benchmark suites like the CEC 2019 test suite [24] (which emphasizes multimodality).

Although most test suites were designed with certain properties in mind, it remains questionable whether the contained MOPs actually meet them. So far, MOPs are predominantly visualized by means of their Pareto sets and/or fronts (see, e.g., [24,25]). Obviously, this is accompanied by an enormous loss of information, since all non-optimal points, and thus the information contained therein, are ignored. The tools described in [19,20] offer a slight improvement over the very limited view at Pareto optimal points. However, the approaches described therein, such as the *prosection* method, mainly focus on a dimensionality reduction of the objective space – and thus do not permit drawing conclusions about the effects of the search space parameters on the objectives of the MOP.

To the best of our knowledge, there exist only two visualization techniques which provide a joint view at decision and objective space (and thus are helpful for engineering better algorithms and benchmark problems): the *cost landscapes* by Fonseca [5] and the *gradient field heatmaps (GFH)* by Kerschke and Grimme [12,14]. Both approaches have in common that they depict the decision space of two-dimensional MOPs and scalarize the information of the MOP's multiple objectives in a single height value (per point from the decision space). For both approaches, the decision space is discretized into a rectangular grid, and the grid's resolution naturally impacts the quality of the visualizations.

The height of a cost landscape is given by the so-called Pareto ranking, i.e., an integer value that gives the amount of points from the (discretized) decision space dominating the current point. Due to the usage of the (global) Pareto ranking, cost landscapes focus on *global* optimality and thus are only able to reveal local structures, if the local fronts are close to parts of the Pareto front.

In contrast, the height of the GFHs is based on a MOP's gradient vector field. More precisely, for each point of the grid, the single-objective gradients (pointing to the closest optimum of the respective objective) are approximated
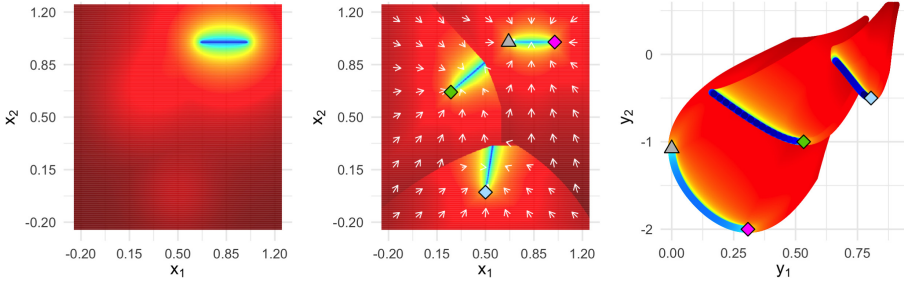
Fig. 1. Exemplary comparison of the cost landscape (left) and the gradient field heatmap (middle) based on the bi-objective SGK function. The right image shows the objective space for the GFH and thus helps identifying the superposition and relationships of the three attraction basins (incl. their associated efficient sets). The colors indicate the respective heights and change gradually from red (max.) to blue (min.). (Color figure online)

using Eq. 3, and afterwards combined into a MO gradient as defined in Eq. 2. Then, for each grid point, the gradient-based path towards the closest local efficient set is computed and the lengths of the MO gradients along the path are cumulated, defining the height of the GFH [12]. Constructing the GFHs based on the paths towards the closest *local* efficient set automatically provides insights into the local structure of the given MOP, as it depicts the attraction basins, as well as the local efficient sets contained therein. However, the focus on locality comes with the drawback that global relationships are hardly visible.

Figure 1 provides an exemplary visual comparison of the cost landscape and the GFH approach based on the bi-objective SGK function[1], which combines a unimodal and a trimodal sphere function. The single-objective local optima are depicted by a grey triangle (for f_1), as well as green, blue, and pink diamonds (for f_2), respectively. The left image shows the corresponding cost landscape, in which the MOP's Pareto set is clearly visible – in contrast to the local efficient sets, whose location can only be guessed by the shading of colors. The GFH approach, given in the second image, shows the three attraction basins formed by the three competing optima of f_2, along with the corresponding vector field of MO gradients (white arrows). Moreover, all three (local) efficient sets are visible, and one can also identify two of them being non-global as their sets – which start in the blue and green diamonds, resp. – are abruptly cut by ridges between the current and the superseding attraction basins. While such a global ranking of the three efficient sets can be derived manually for this simple scenario, it is hard to realize for more complex MOPs (e.g., see bottom row of Fig. 4).

---

[1] $f(x_1, x_2)$ with $f_1(x_1, x_2) = 1 - \left(1 + 4 \cdot \left((x_1 - 2/3)^2 + (x_2 - 1)^2\right)\right)^{-1}$ and $f_2(x_1, x_2) = 1 - \max\{g_1, g_2, g_3\}$, whose subfunctions g_1, g_2 and g_3 are defined as $g_i(x_1, x_2, h, c_1, c_2) = h / \left(1 + 4 \cdot \left((x_1 - c_1)^2 + (x_2 - c_2)^2\right)\right)$ with $h = 1.5$, $c_1 = 0.5$, $c_2 = 0$ (for g_1), $h = 2$, $c_1 = 0.25$, $c_2 = 2/3$ (for g_2), and $h = 3$, $c_1 = 1 = c_2$ (for g_3).

3   Identification of Locally Efficient Points

Locally efficient points are an important part of MOP landscapes as they indicate where local Pareto fronts (or sets) are located, and thus, where local search strategies might get stuck [6]. However, state-of-the-art visualization approaches either do not feature them at all (e.g., the cost landscapes [5]), or only show their locations indirectly (e.g., the gradient field heatmaps [12]). Only when the location of the local efficient sets is known analytically for specific test problems – like for DTLZ [7] or MMF [24] – they are represented in some visualizations.

Here, we present an approach based on the estimated gradients of the MOP and the stability of the MO gradient vector field to locate all locally efficient points for a continuous MOP. We begin by detailing our computational approach for approximating the function and its derivatives, followed by a description of first- and second-order optimality conditions for locally efficient points and how they were implemented.

3.1   Computational Approach

As continuous functions can in principle be evaluated in infinitely many different points, an approximation of the function based on a finite set of evaluated points is required. For this purpose, we evaluate points on a rectangular grid of coordinates $(x_1^{j_1}, x_2^{j_2})$. Those coordinates are aligned equidistantly with a number of steps $n_1, n_2 \in \mathbb{N}$ and step sizes $s_i = (u_i - l_i) \cdot (n_i - 1)^{-1}$ per dimension, i.e., $x_i^{j_i} = l_i + (j_i - 1) \cdot s_i$, with $i = 1, 2$ and $j_i = 1, ..., n_i$ where $x_i^{j_i}$ denotes the coordinate of the j_i-th grid point in the i-th dimension. Next, the rectangular grid of points is created by taking the cross-product of the one-dimensional coordinates $x_i^{j_i}$. The function f is then evaluated for each of the points from the grid.

In each grid point, the respective derivative is approximated (per objective) using the finite differences method based on the neighboring coordinates of the point at hand. On the decision boundary, forward- and backward-differences are taken respectively, while the interior points are evaluated using central differences. With the function values of f at grid point $(x_1^{j_1}, x_2^{j_2})$ denoted as $f(x_1^{j_1}, x_2^{j_2})$, the partial derivative of f with regard to x_1 is then estimated by:

$$\frac{\partial}{\partial x_1} f(x_1^{j_1}, x_2^{j_2}) \approx \begin{cases} \frac{1}{2s_1} \cdot \left(f(x_1^{j_1+1}, x_2^{j_2}) - f(x_1^{j_1-1}, x_2^{j_2}) \right), & \text{for } 1 < j_1 < n_i \\ \frac{1}{s_1} \cdot \left(f(x_1^{j_1+1}, x_2^{j_2}) - f(x_1^{j_1}, x_2^{j_2}) \right), & \text{for } j_1 = 1 \\ \frac{1}{s_1} \cdot \left(f(x_1^{j_1}, x_2^{j_2}) - f(x_1^{j_1-1}, x_2^{j_2}) \right), & \text{for } j_1 = n_1 \end{cases} \tag{3}$$

Derivatives for x_2 are calculated analogously. Compared to approximations with smaller step sizes, we did not observe noticeable changes in the visualization.

3.2   First-Order Conditions

Although the MO gradient is capable of capturing the local efficiency of a point in the decision space, it is not sufficient on its own to capture all critical points
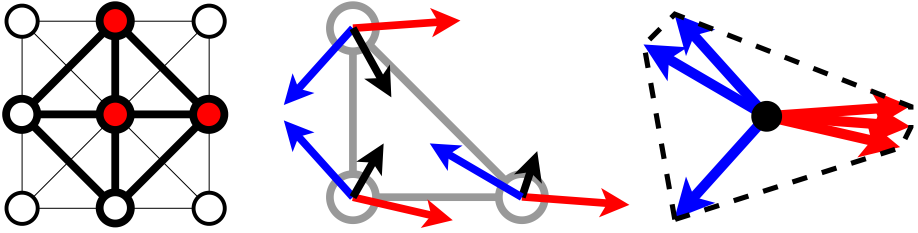
Fig. 2. Left: Example of four triangular neighborhoods considered while evaluating for critical points. The red nodes are an example for one particular neighborhood. Middle: For such a triangular neighborhood of points, considering the MO gradients (black), no critical points would be detected, as all of them point to the right as a common descent direction. However, it would be reasonable to expect that due to the changes in direction of the single-objective gradients (red and blue respectively, pointing towards descent directions), there would be a local efficient set in between. Right: Our approach jointly considering the combination of all single-objective gradients from all corner points correctly reports the neighborhood to contain a locally efficient point. (Color figure online)

while approximating them in the grid-based decision space. In particular, if one of the single-objective gradients changes much faster than the other in the neighborhood of a local efficient set, the MO gradient alone may misleadingly fail to recognize some of the critical points (as shown schematically in Fig. 2).

Thus, we present an approach that improves the status quo in detecting all critical points within the MOP's grid. We consider triangular neighborhoods of grid points and jointly consider the gradients of all constituent single-objective functions. If the convex hull of the gradients encloses the origin (see right image of Fig. 2), we presume a critical point *somewhere* in the interior of the triangle.

An illustration of the neighborhood, which is used for detecting the critical points, is given in the left image of Fig. 2. For each grid point, four triangular neighborhoods are evaluated. Each of those triangles consists of the point itself, as well as one horizontal and one vertical neighbor of that point. If a critical point is detected for a triangle, all of its corner points are considered being critical.

In addition, points along the decision boundary require special attention. So far, in the gradient field heatmaps, all (boundary) points for which the MO gradient points "out" of the feasible decision space were considered locally efficient. However, even in that case, it might still be possible to follow a descent direction for all objectives when moving along the decision boundary. We resolve this issue by considering boundary points, their adjacent boundary points and their common descent directions along the decision boundary all together. If no common descent direction for all objectives is found, the respective pair of boundary points is considered critical. Further, we rotate the MO gradient at these points to point into the common descent direction along the boundary, which is helpful for further visualizations with the gradient field heatmap (see Sect. 4).
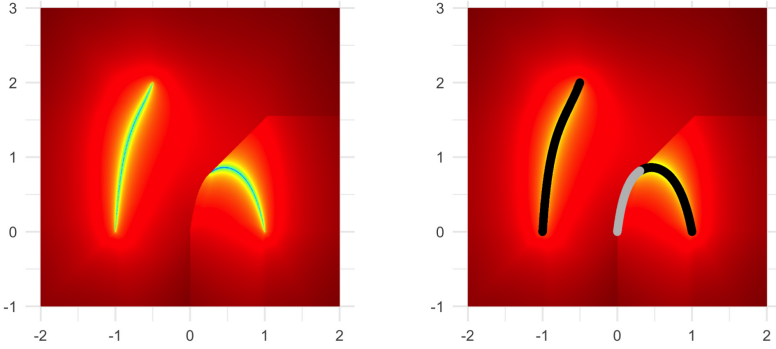
Fig. 3. GFH visualization and location of all interior critical points for the Aspar function $f(x_1, x_2) = (x_1^4 - 2x_1^2 + 2x_2^2 + 1, (x_1 + 0.5)^2 + (x_2 - 2)^2)$. Even in this very simple problem, some critical points (gray) do not belong to the efficient sets (black) but are part of the landscape's ridges, emphasizing the need for a second-order criterion.

When only aiming at the identification of locally efficient points in the decision space, one needs to be aware that not all critical points are necessarily locally efficient. An example of this is given in Fig. 3, which shows the GFH of a simple MOP and all of its critical points. Note that some of the critical points do not belong to a local efficient set, but are rather part of the ridges between two adjacent basins of attractions. To reliably extract the locally efficient points from the set of critical points, a second-order condition is required.

3.3 Second-Order Conditions

Analogously to single-objective functions satisfying first-order optimality conditions, critical points in the multi-objective sense cannot just be local minima, i.e., locally efficient points, but also local maxima, as well as saddle points.

This highlights the need for the consideration of a second-order condition to distinguish identified critical points into locally efficient points and others. There are second-order conditions for the continuous case [18], however, using the grid approximation required for our approach, these proved to be too unstable for efficient use to discern between the different types of critical points. Motivated by the gradient field heatmaps, which indicate that the MO gradient captures the local search behaviour well, we derive our second-order condition based on properties of the MO gradient vector field.

The MO gradient defines a vector field over the decision space that can be analyzed w.r.t. its stability at the critical points. A point is considered asymptotically stable, if after a small perturbation, following the vector field to the closest critical point, one stays within a small neighborhood of the original point. This property is well studied in the field of autonomous differential equations and can be analyzed by a linear approximation of the vector field at the critical

point using its Jacobian [2].[2] If the real part of all (potentially complex-valued) eigenvalues of the Jacobian is negative, the point is considered stable.

In the MO gradient field, however, we generally deal with degenerated critical points, for which (at least) one eigenvalue is zero, associated with the eigenvectors pointing along the local efficient set. This can pose problems with the numeric approximation that we require for our approach. Luckily, for the 2D case it is sufficient to consider the trace of the Jacobian, also known as the divergence in the context of vector fields, to determine asymptotic stability. Intuitively, the divergence is a measure for the "ingoingness" or "outgoingness" of the vector field at a given point, and it is numerically more stable and efficient to calculate than computing all eigenvalues of the Jacobian. Thus, the divergence of a 2D vector field $\mathbf{V} : \mathbb{R}^2 \to \mathbb{R}^2$ with $\mathbf{V}(\mathbf{x}) = (V_{x_1}(\mathbf{x}), V_{x_2}(\mathbf{x}))$ is given by:

$$\mathbf{div}(\mathbf{V}(\mathbf{x})) = \frac{\partial}{\partial x_1} V_{x_1}(\mathbf{x}) + \frac{\partial}{\partial x_2} V_{x_2}(\mathbf{x}). \tag{4}$$

In summary, for minimization as defined in Eq. 1, if an interior critical point $\mathbf{x}$ fulfills $\mathbf{div}(-\nabla f(\mathbf{x})) < 0$, it is a stable critical point in the MO gradient field, and thus locally efficient. Note that in the degenerated case, where all points in a given neighborhood are locally efficient, the divergence is zero.

Thus, to assess whether a critical point is locally efficient, we consider the divergence of each set of points that were jointly considered critical w.r.t. the first-order conditions (see Sect. 3.2). Only if all three evaluated points have non-positive divergence, we regard them as locally efficient. The divergence for each grid point is estimated using the grid-based finite differences method (see Eq. 3).

Again, the critical points along the decision boundary require special treatment. Here, we do not use the divergence to distinguish different types of critical points. Instead, only if the MO gradients of the considered boundary points are pointing "outwards" or along the decision boundary, a set of critical points is considered locally efficient.

4   Visualizing Local and Global Structures of MOPs

The previous section introduced a novel and reliable approach to approximate the location of all locally efficient points within the MOP's continuous decision space (based on a rectangular grid of evaluated points). This explicit knowledge of the location of the locally efficient points not only allows us to show them in the decision space, but also to extract information about their relative dominance relationship. This means, we utilize Pareto ranking, which also serves as the basis for the cost landscapes [5], but limit ourselves to the locally efficient points – resulting in an enormous speed-up compared to a ranking of *all* grid points.

Ultimately, this leads to a unique visualization that not only shows the locations of locally efficient solutions, but also provides information about their

---

[2] This presumes that the MO gradient field can be approximated by a linear function in the considered point. For differentiable MOPs this is a reasonable assumption, but we observed that it works well for our approach in general.
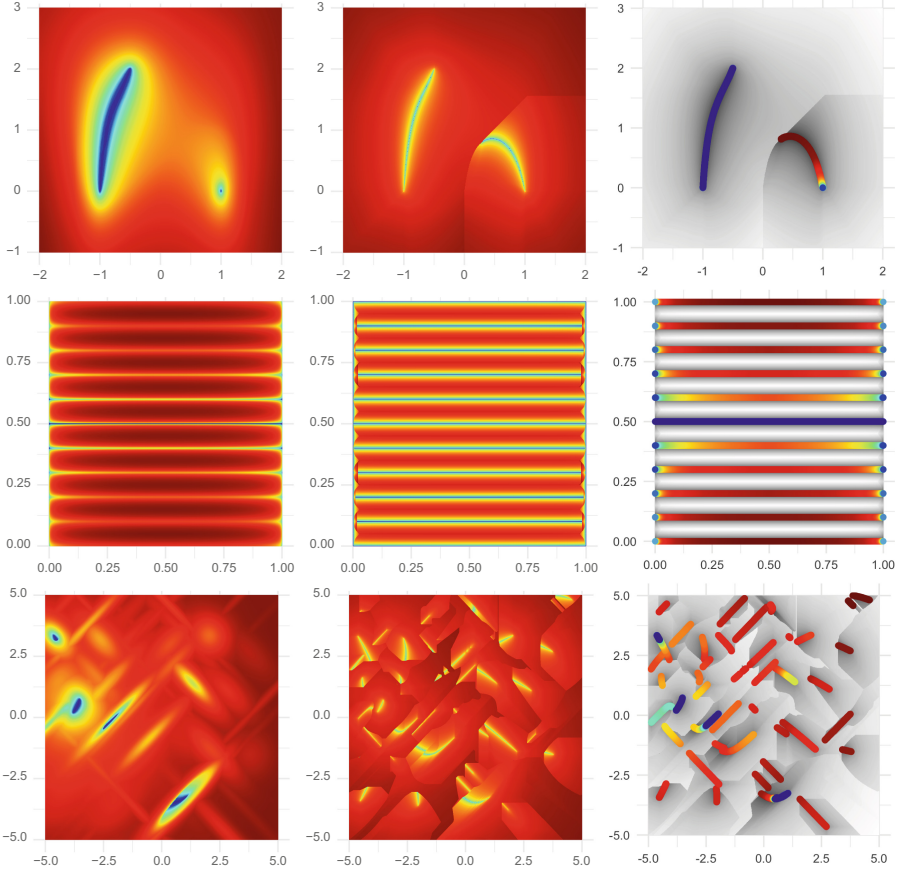
Fig. 4. Comparison of the cost landscape, GFH and PLOT (left to right) on the two-dimensional bi-objective Aspar, DTLZ1 and bi-objective BBOB (FID = 10, IID = 1) functions (top to bottom). Due to the computational overhead involved in computing the domination counts, the cost landscape is calculated with only 500 grid points per dimensions, while the GFH and PLOT use a resolution of 1,000 points per dimension.

global optimality. In addition, we enhance our visualization with a gray-scaled version of the corresponding GFH in the background, which preserves additional information about the basins of attraction (e.g., their shapes and sizes). Along the boundary points, we modify the MO gradient as described in Sect. 3.2. Also, all locally efficient points determined by our more stable detection method (see Sect. 3) can automatically be reused "for free" within the generation of the GFHs. Both modifications further improved the visualization quality of the GFHs.

Figure 4 provides a visual comparison of the current state-of-the-art visualization techniques – cost landscapes (left column) and GFHs (middle column) – and our proposed PLOT approach (right column) based on three exemplary MOPs: the simple Aspar function (top row) from Fig. 3, the well-established
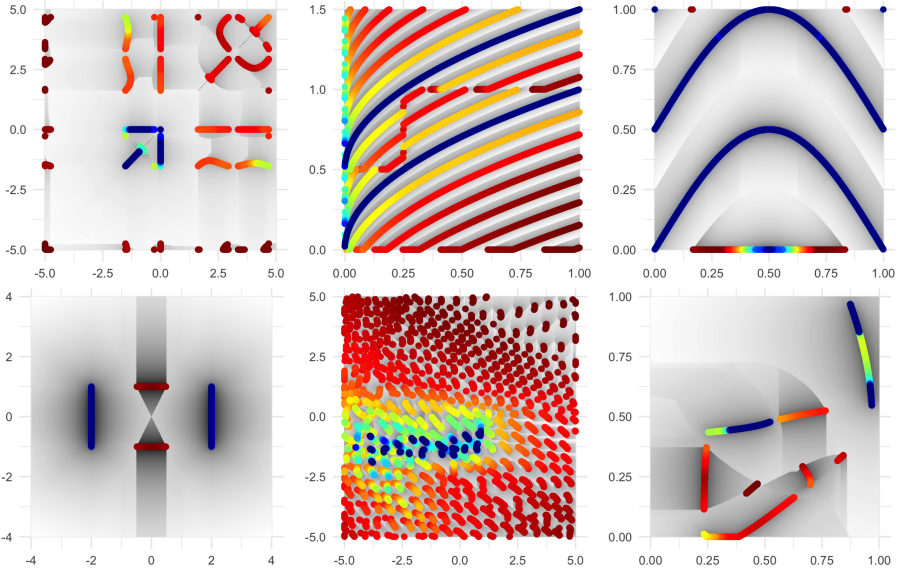
Fig. 5. Exemplary PLOTs for various continuous MOPs (left to right, top to bottom): Kursawe [15], MMF3 and MMF14a from the CEC 2019 test suite [24], a MinDist function with centers $(-2, -1), (2, 1)$ and $(-2, 1), (2, -1)$ [17], a bi-objective BBOB (FID: 42, IID: 1) [21], and a bi-objective function generated using the MPM2 generator (with parameters $(3, 2, \text{random}, 4)$ and $(3, 2, \text{random}, 8))$ [13]. All plots were generated using an equidistant grid in the decision space with a resolution of $1,000 \times 1,000$ points.

DTLZ 1 [7] (middle row), and the 10th function from the rather recent bi-objective BBOB test suite [21]. Noticeably, for the latter two problems, GFH has problems in identifying some critical points correctly and mistakenly shows some points along the boundary as locally efficient. On the other hand, the cost landscape approach has problems with *local* efficient sets and can at most identify the *location* of some of them – as long as their fronts are close to the global Pareto front(s). PLOT combines the global view of the cost landscapes with the local information of the GFH, and thus provides a much more informative depiction of the locally efficient solutions.

> Ultimately, this results in a single <u>P</u>lot of the <u>L</u>andscape with <u>O</u>ptimal <u>T</u>rade-offs (PLOT), which jointly illustrates three types of landscape characteristics: (1) *local efficient sets*, (2) the *global optimality* of their respective solutions, and (3) the *basins of attraction* associated with the respective efficient sets.

Our R-package `moPLOT`, which has been used to generate all visualizations in this paper, is available on GitHub: https://github.com/kerschke/moPLOT. Further resources and information can be found on our project's website on multimodal multi-objective optimization: https://mo-opt.github.io.

5   Observations

We provide PLOT visualizations for a selection of further benchmark functions in Fig. 5. Many MOPs that were designed with multimodality in mind reveal very simple structures in the decision space. Notably, the PLOTs show peculiarities in the definition of some functions that were designed with a focus on multiple global Pareto sets. These can lead to unintended locally and globally efficient solutions along the boundary (MMF14a) and glaring cuts in the landscape of the local efficient sets (MMF3). Otherwise, many MOPs have even simpler landscape structures only containing few local efficient sets in general (MinDist). Only few of the MOPs show interactions between the objectives, which lead to a disconnected global Pareto set (i.e., it is distributed over multiple local efficient sets). This can, e.g., be seen in the bi-objective BBOB and MPM2 functions.

Further note that the location of locally efficient points along the decision boundary implies that in general an unconstrained locally efficient set would be found outside of the feasible decision space. This can be observed in the depicted Kursawe, MMF and MPM2 functions.

The bi-objective BBOB function shows a very complex landscape with many locally efficient solutions. In fact, its sets cover the majority of the decision space and thereby reveal the limitations of PLOT. However, such extremely multimodal MOPs are challenging for any visualization method. Also, even for that very extreme problem, PLOT reliably visualized the MOP's global structure.

6   Conclusions

Visualizing an optimization problem's landscape is very useful when studying its properties, or the search behavior of the optimizers operating on it. In MO continuous optimization, however, there exist hardly any meaningful visualization methods, with the consequence that MOPs are primarily treated as black-boxes.

We present a novel approach for the numerical approximation of locally efficient points in the decision space of continuous MOPs. This new information was then integrated into PLOT – our new method for the visualization of bi-objective two-dimensional MOPs. Our approach can visualize local and global efficient sets, as well as their basins of attraction. Thereby, it enables a visualization of MOPs that encompasses information comparable to visualizations available for single-objective functions. We successfully apply our approach to a wide variety of benchmarking functions and often reveal very simple landscape properties. As with previous MO visualization techniques, we hope to inspire further progress in understanding the landscapes of existing benchmarking functions, the design of new benchmarks, as well as the development of novel algorithmic ideas.

It should be noted that the definitions for the MO gradient can be extended to an arbitrary number of dimensions and objectives [8]. Likewise, our approach for identifying critical points as well as our second-order criterion, which is based on the stability of the MO gradient field, can easily be adapted to cope well with

higher-dimensional MOPs. Thus, our proposed approach provides the fundamentals for an extension towards visualizing decision spaces of 3-dimensional MOPs. To the best of our knowledge, this is not yet available beyond the visualization of Pareto sets, or analytically known local efficient sets. An extension to 3D decision spaces would also enable more detailed investigations of the properties of 3-objective MOPs. This is currently not yet feasible, as their counterparts with 2D decision spaces in general contain degenerated critical points.

Further, it can be noted that our approach supports studying selected regions of interest in the landscape. This can be effectively achieved by *zooming* into the PLOT and supports the visualization of particularly complex MOPs. Another possible extension that aims at improving the visualization quality of our PLOT would be a dynamic resampling strategy around the identified critical points, increasing the accuracy of the approximation of the locally efficient points.

References

1. Beyer, H.G.: The Theory of Evolution Strategies. Springer, Heidelberg (2001). https://doi.org/10.1007/978-3-662-04378-3
2. Blanchard, P., Devaney, R., Hall, G.: Differential Equations. Cengage Learning, Boston (2012)
3. Coello Coello, C.A., van Veldhuizen, D.A., Lamont, G.B.: Evolutionary Algorithms for Solving Multi-Objective Problems, 2nd edn. Springer, Boston (2007). https://doi.org/10.1007/978-0-387-36797-2
4. Custódio, A.L., Madeira, J.F.A.: MultiGLODS: global and local multiobjective optimization using direct search. J. Global Optim. **72**(2), 323–345 (2018)
5. da Fonseca, C.M.M.: Multiobjective genetic algorithms with application to control engineering problems. Ph.D. Thesis, Department of Automatic Control and Systems Engineering, University of Sheffield, September 1995
6. Deb, K.: Multi-objective genetic algorithms: problem difficulties and construction of test problems. Evol. Comput. (ECJ) **7**(3), 205–230 (1999)
7. Deb, K., Thiele, L., Laumanns, M., Zitzler, E.: Scalable test problems for evolutionary multiobjective optimization. In: Abraham, A., Jain, L., Goldberg, R. (eds.) Evolutionary Multiobjective Optimization. Advanced Information and Knowledge Processing, pp. 105–145. Springer, London (2005). https://doi.org/10.1007/1-84628-137-7_6
8. Désidéri, J.A.: Multiple-gradient descent algorithm (MGDA) for multiobjective optimization. Comptes Rendus Mathematique **350**(5–6), 313–318 (2012)
9. Grimme, C., Kerschke, P., Emmerich, M.T.M., Preuss, M., Deutz, A.H., Trautmann, H.: Sliding to the global optimum: how to benefit from non-global optima in multimodal multi-objective optimization. In: AIP Conference Proceedings, pp. 020052-1-020052-4. AIP Publishing (2019)
10. Grimme, C., Kerschke, P., Trautmann, H.: Multimodality in multi-objective optimization – more boon than bane? In: Deb, K., et al. (eds.) EMO 2019. LNCS, vol. 11411, pp. 126–138. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-12598-1_11
11. John, F.: Extremum problems with inequalities as subsidiary conditions. In: Giorgi, G., Kjeldsen, T.H. (eds.) Traces and Emergence of Nonlinear Programming, pp. 197–215. Springer, Basel (2014). https://doi.org/10.1007/978-3-0348-0439-4_9

12. Kerschke, P., Grimme, C.: An expedition to multimodal multi-objective optimization landscapes. In: Trautmann, H., et al. (eds.) EMO 2017. LNCS, vol. 10173, pp. 329–343. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-54157-0_23

13. Kerschke, P., et al.: Towards analyzing multimodality of continuous multiobjective landscapes. In: Handl, J., Hart, E., Lewis, P.R., López-Ibáñez, M., Ochoa, G., Paechter, B. (eds.) PPSN 2016. LNCS, vol. 9921, pp. 962–972. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-45823-6_90

14. Kerschke, P., et al.: Search dynamics on multimodal multi-objective problems. Evol. Comput. (ECJ) **27**, 577–609 (2019)

15. Kursawe, F.: A variant of evolution strategies for vector optimization. In: Schwefel, H.-P., Männer, R. (eds.) PPSN 1990. LNCS, vol. 496, pp. 193–197. Springer, Heidelberg (1991). https://doi.org/10.1007/BFb0029752

16. Liefooghe, A., López-Ibáñez, M., Paquete, L., Verel, S.: Dominance, epsilon, and hypervolume local optimal sets in multi-objective optimization, and how to tell the difference. In: Proceedings of the 20th Annual Conference on Genetic and Evolutionary Computation (GECCO), vol. 18, pp. 324–331. ACM, Kyoto (2018)

17. Maree, S.C., Alderliesten, T., Bosman, P.A.N.: Real-valued evolutionary multimodal multi-objective optimization by hill-valley clustering. In: Proceedings of the 21st Annual Conference on Genetic and Evolutionary Computation (GECCO), pp. 568–576. ACM (2019). https://doi.org/10.1145/3321707.3321759

18. Miettinen, K.: Nonlinear Multiobjective Optimization. International Series in Operations Research & Management Science, vol. 12. Springer, Boston (1998)

19. Tušar, T.: Visualizing Solution Sets in Multiobjective Optimization. Ph.D. thesis, Jožef Stefan International Postgrad. School (2014)

20. Tušar, T., Filipič, B.: Visualization of pareto front approximations in evolutionary multiobjective optimization: a critical review and the prosection method. IEEE Trans. Evol. Comput. (TEVC) **19**(2), 225–245 (2015)

21. Tušar, T., Brockhoff, D., Hansen, N., Auger, A.: COCO: the bi-objective black box optimization benchmarking (bbob-biobj) test suite. arXiv preprint abs/1604.00359 (2016)

22. van Veldhuizen, D.A.: Multiobjective evolutionary algorithms: classifications, analyzes, and new innovations. Ph.D. thesis, Faculty of the Graduate School of Engineering of the Air Force Institute of Technology, Air University, June 1999

23. Whitley, L.D., Mathias, K.E., Rana, S.B., Dzubera, J.: Building better test functions. In: Proceedings of the 6th International Conference on Genetic Algorithms (ICGA), pp. 239–247 (1995)

24. Yue, C., Qu, B., Yu, K., Liang, J., Li, X.: A Novel Scalable Test Problem Suite for Multimodal Multiobjective Optimization. Swarm and Evolutionary Computation (2019)

25. Zitzler, E., Deb, K., Thiele, L.: Comparison of multiobjective evolutionary algorithms: empirical results. Evol. Comput. (ECJ) **8**(2), 173–195 (2000). https://doi.org/10.1162/106365600568202