Computers & Operations Research 136 (2021) 105489

Contents lists available at ScienceDirect

Computers and Operations Research

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

Peeking beyond peaks: Challenges and research potentials of continuous multimodal multi-objective optimization

Christian Grimme [a,*], Pascal Kerschke [b], Pelin Aspar [a,c], Heike Trautmann [a,e], Mike Preuss [c], André H. Deutz [c], Hao Wang [c,d], Michael Emmerich [c]

[a] Data Science: Statistics and Optimization Group, University of Münster, Germany
[b] Big Data Analytics in Transportation, TU Dresden, Germany
[c] Leiden Institute of Advanced Computer Science, University of Leiden, The Netherlands
[d] Sorbonne Université - LIP6, Paris, France
[e] Data Management & Biometrics Group, University of Twente, The Netherlands

ARTICLE INFO

Keywords:
Multimodal optimization
Multi-objective continuous optimization
Landscape analysis
Visualization
Benchmarking
Theory
Algorithms

ABSTRACT

Multi-objective (MO) optimization, i.e., the simultaneous optimization of multiple conflicting objectives, is gaining more and more attention in various research areas, such as evolutionary computation, machine learning (e.g., (hyper-)parameter optimization), or logistics (e.g., vehicle routing). Many works in this domain mention the structural problem property of *multimodality* as a challenge from two classical perspectives: (1) finding all globally optimal solution sets, and (2) avoiding to get trapped in local optima. Interestingly, these streams seem to transfer many traditional concepts of single-objective (SO) optimization into claims, assumptions, or even terminology regarding the MO domain, but mostly neglect the understanding of the structural properties as well as the algorithmic search behavior on a problem's landscape. However, some recent works counteract this trend, by investigating the fundamentals and characteristics of MO problems using new visualization techniques and gaining surprising insights.

Using these visual insights, this work proposes a step towards a unified terminology to capture multimodality and locality in a broader way than it is usually done. This enables us to investigate current research activities in multimodal continuous MO optimization and to highlight new implications and promising research directions for the design of benchmark suites, the discovery of MO landscape features, the development of new MO (or even SO) optimization algorithms, and performance indicators. For all these topics, we provide a review of ideas and methods but also an outlook on future challenges, research potential and perspectives that result from recent developments.

1. Introduction

Multi-objective optimization (MOO) (Miettinen, 1998) has become one of the most important research areas in evolutionary computation over the last decades (van Veldhuizen, 1999; Deb, 2001; Coello Coello et al., 2007). As MOO aims at the simultaneous optimization of multiple (at least two) and contradicting objectives, multi-objective (MO) optimization problems (MOPs) mainly challenge the selection mechanism of the original single-objective (SO) evolutionary loop (Beyer, 2001). Instead of a single optimum or few equivalent optima in case of multimodal optimization, MOPs have incomparable optimal solutions, i.e., a set of optimal trade-off solutions, which is called Pareto set (also known as global efficient set). Thus, evolutionary multi-objective algorithms (EMOAs) – i.e., evolutionary algorithms (EAs) that aim at the optimization of MOPs – are required to converge to optimal solutions and simultaneously need to ensure, that all – or at least a diverse set of – solutions are found.

For many years, the research on solving MOPs by using evolutionary approaches focused on the development of algorithms and their experimental validation (Schaffer, 1985; Deb et al., 2002; Emmerich et al., 2005; Zhang and Li, 2007; Emmerich and Deutz, 2018; Coello Coello et al., 2019). Specifically in the domain of *continuous multi-objective optimization* and after the early steps in algorithm development, the problems themselves became an important branch of research. Inspired by the methodological advances in the SO domain of evolutionary

* Corresponding author.
E-mail addresses: christian.grimme@uni-muenster.de (C. Grimme), pascal.kerschke@tu-dresden.de (P. Kerschke), asparp@uni-muenster.de (P. Aspar), trautmann@uni-muenster.de (H. Trautmann), m.preuss@liacs.leidenuniv.nl (M. Preuss), a.h.deutz@liacs.leidenuniv.nl (A.H. Deutz), hao.wang@lip6.fr (H. Wang), m.t.m.emmerich@liacs.leidenuniv.nl (M. Emmerich).

https://doi.org/10.1016/j.cor.2021.105489
Received 23 September 2020; Received in revised form 17 May 2021; Accepted 20 July 2021
Available online 31 July 2021

computation over the last decades, several benchmark suites (Zitzler et al., 2000; Deb et al., 2005; Tušar et al., 2016; Yue et al., 2019) were developed. Based on those benchmarks, the fine tuning of algorithmic parameters (see also algorithm configuration (Blot et al., 2016; López-Ibáñez et al., 2016; Belkhir et al., 2017; Eggensperger et al., 2019)) became an important field of investigation. During these steps in the community's advancement, many notions, insights, challenges, and strategies from the traditional field of SO optimization (SOO) diffused into the field of continuous MOO (as it also did for the discrete domain).

One of these aspects and considered challenges is *multimodality* (Preuss, 2015), i.e., the existence of multiple global and/or local optima. Multiple *global* optima are solutions, which occur in different positions in the decision space but whose objective values are of optimal quality. *Local* optima are solutions that are best within a certain neighborhood but do not necessarily have optimal objective values. Consequently in SOO, the challenge of multimodality can be categorized into two (not necessarily disjoint) streams: (1) due to the existence of multiple basins of attractions, multimodality is regarded as an obstacle for algorithms trying to reach a global optimum; (2) due to the existence of multiple (globally) optimal solutions, it is associated with the search for diverse solutions (in decision space) that are of similar quality (measured in objective space). While the first aspect – avoiding to get stuck – is a difficulty for algorithm design in general, the latter aspect – finding all optimal solutions – is essential for practical applications. For instance, when a globally optimal solution cannot be realized, another global solution with a different parametrization but similar quality, may become of major interest.

Similar streams can be observed for multimodality in continuous MOO. Most of the algorithmic development dedicated to multimodality is rather recent (Tanabe and Ishibuchi, 2020) and focuses on finding many or all Pareto optimal solutions in decision space. The main challenge considered in these approaches is to find all globally optimal solutions in decision space, even if they map to the same solutions, see Fig. 1 (left side). This could be termed *multiglobal* optimization as the challenge is to find and preserve multiple global optima in a multimodal environment.

The other perspective on multimodality – not getting trapped – is implicitly considered in all MOO tasks.[1] All approaches – whether they strive for preserving diversity in decision space or not – assume that locally optimal solutions may pose a challenge for algorithmic convergence to global optimality (Deb, 1999) (see also right side of Fig. 1 for a visual notion of this perspective). Whereas the existence of traps due to multimodality is often inherent in discrete optimization problems (Liefooghe et al., 2018c; Paquete et al., 2004, 2007a; Liefooghe et al., 2018a), it is less clear in the continuous case. In fact, for continuous problems in (evolutionary) MO optimization, the properties (and challenges) of multimodality are rather transferred from SO optimization than understood in the context of MOO.

Many designers of early continuous MO benchmarks and/or algorithms assumed similarities to known and widely accepted effects from SO optimization. They thus transferred the common understanding of multimodality – local optima are traps for local search methods and sometimes also major obstacles for global optimizers such as evolutionary algorithms (EAs) – to MOO. As a result, benchmark designers regularly integrate multimodality into test problems (with the aim of posing challenging problems) and algorithm engineers either rely on the global search behavior of EAs or try to tune their algorithms to avoid these traps. Our impression is, however, both groups often consider multimodality as major problem without knowing what it exactly means in the context of continuous MO landscapes.

[1] If it is not mentioned explicitly in this context, it is integral part of the decision to tackle a problem using evolutionary algorithms. Otherwise, local optimizers would suffice for solving such problems.

Recently, some works have tried to close this gap using visual methods for investigating continuous multimodal landscapes (Kerschke and Grimme, 2017; Schäpermeier et al., 2020). Similarly, first theoretical formulations of localness (Kerschke et al., 2016, 2019b), or investigations of algorithm behavior and new design principles (Wang et al., 2017a,b; Grimme et al., 2019a,b) appeared and started to become a research "streamlet" in continuous MOO. Literally, these new visualizations of localness in continuous MOO can provide fundamental insights into continuous MO landscapes and thereby enable the extraction of algorithmic implications for multimodal and multiglobal MOO.

The contribution of this work is twofold: on the one hand, it provides a comprehensive overview on the research activities in the field of evolutionary optimization for continuous multimodal (and multiglobal) MOPs and highlights algorithmic development, visualization approaches, and benchmarking. On the other hand, it details new insights into the structure of continuous multimodal problem landscapes, points to possible misconceptions, disruptive insights, as well as existing gaps.

However, beyond that, the paper should be considered as a position paper, which discusses the misperceptions and challenges, as well as the perspective research potential of multimodality in the continuous MOO domain. Based on very recent – and for many people probably still surprising – visual insights into the structure and properties of continuous multimodal problem landscapes, we highlight some promising directions for future research. We are convinced that what is known of multimodality (in the broader sense) in MOO today is only a starting point for new research in benchmarking, landscape analysis, algorithm development (notably in MO *and* SO optimization), as well as in theory. In order to enable future discussion, we first propose to extend the terminology of solution efficiency for continuous multimodal MOO derived from observations in problem landscapes. This terminology enables the community to address and distinguish *local* from *global* efficient sets (the latter are usually called Pareto set) in a more precise way. Thereafter, we address the research areas of benchmarking, landscape analysis, algorithms, and performance indicators to identify possible future trends in research on MOO, as well as potentially promising paths to follow.

Specifically, Section 2 summarizes the basic and well-known concepts of MOO using current terminology, while Section 3 provides an extended terminology to capture localness and related properties for MOO in a formal way. This is complemented by Section 4, which dives into available methods for visualizing the landscapes of continuous MO problems and provides an overview of very recent visualization techniques that allow us to demonstrate the previously defined local properties of MO landscapes. Section 5 details the promising areas of research that are directly related to the new terminology and visual insights. Section 5.1 revisits the available MO test suites and, at the same time, offers the opportunity to (also visually) reflect on their MOPs (along with their structural characteristics). This paves the ground for a discussion of MOP landscapes and perspectives on landscape analysis in MOO in Section 5.2. Section 5.3 addresses modern trends and perspectives in algorithm development, while Section 5.4 considers the implications of multimodality on performance measurement. Note that each subsection in Section 5 comprehensively investigates the literature and subsequently highlights possible new directions. Finally, Section 6 provides a concluding argument, why multimodality is able to push the entire (evolutionary) multi-objective optimization (EMO) community forward.

2. Preliminaries on multi-objective optimization

In this section, we briefly introduce the classical and fundamental terminologies, which commonly describe multi-objective optimization problems in the evolutionary computation literature (Coello Coello

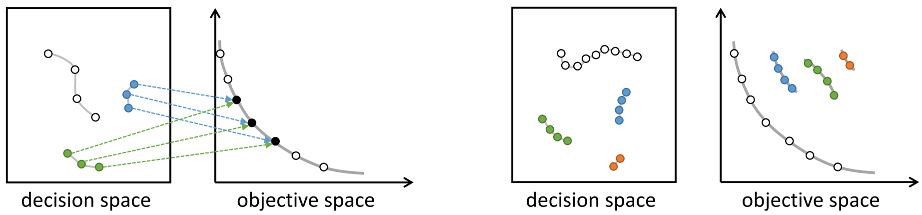

Fig. 1. Different perceptions of multimodality in the MO context. Left: A multiglobal perspective, where diverse solutions from the decision space map to the same image in objective space (i.e., a *surjection*). Right: A second perspective on multimodality with locally optimal solutions in decision space that correspond to (separate) dominated fronts.

et al., 2007), and which have been adapted from classical operations research literature (Miettinen, 1998). Its usage is rooted in the early days of algorithm development, when the focus of MOO was on transferring the single-objective evolutionary loop to the domain of multiple objectives. Back then, the central problem was that of modeling the selection operator, which was required to allow (if possible direct) comparison between solutions to select the best and reject the dominated ones.

Definition 1 (*Multi-objective Optimization Problem*). A *multi-objective optimization problem (MOP)* is commonly denoted as a vector-valued function

$$\mathbf{f}: \mathcal{X} \to \mathbb{R}^m, \quad \mathbf{x} \mapsto \left(f_1(\mathbf{x}), f_2(\mathbf{x}), \ldots, f_m(\mathbf{x})\right)^\top,$$

with m real-valued single-objective functions $f_i : \mathcal{X} \to \mathbb{R}$, $i = 1, \ldots, m$, which are to be optimized simultaneously.

As a consequence, the domination of solutions over others is not as simple to realize as in SO optimization. Further, in the herein considered scenario of *continuous* MOO, the MOP's search space $\mathcal{X}$ is also real-valued, i.e., $\mathcal{X} \subseteq \mathbb{R}^n$.

In contrast to a totally ordered set $(\mathbb{R}, \leq)$ in SO optimization, with its natural total order $\leq$ on $\mathbb{R}$, the set of solution candidates of a MOP follows the *weak Pareto order* $\preceq$ on $\mathbb{R}^m$, which is defined as follows:

Definition 2 (*Pareto Order or Pareto Dominance*). Let $\mathbf{a} = (a_1, \ldots, a_m) \in \mathbb{R}^m$ and $\mathbf{b} = (b_1, \ldots, b_m) \in \mathbb{R}^m$. We say **a** weakly dominates **b** (written as $\mathbf{a} \preceq \mathbf{b}$), if and only if $a_i \leq b_i$ for all $i = 1, \ldots, m$. Then the *Pareto order* $\prec$ on $\mathbb{R}^m$ is defined as follows: **a** dominates **b** ($\mathbf{a} \prec \mathbf{b}$), if and only if $\mathbf{a} \preceq \mathbf{b}$ and $\mathbf{a} \neq \mathbf{b}$. The weak Pareto order can also be extended to sets of points: for $A, B \subseteq \mathbb{R}^m$, A weakly dominates B, if and only if $\forall \mathbf{b} \in B\ \exists \mathbf{a} \in A$ such that $\mathbf{a} \preceq \mathbf{b}$. The order $\prec$ can be generalized similarly: A dominates B, if and only if $\forall \mathbf{b} \in B\ \exists \mathbf{a} \in A$ such that $\mathbf{a} \prec \mathbf{b}$.

As a consequence of this order, the solution of a MOP is not a single (or equally performing set of) globally optimal solution(s), but a set of incomparable, i.e., globally non-dominated solutions. They form a set of efficient solutions in decision space and are called Pareto set, while the image of this set represents the optimal trade-off between objectives and is called Pareto front.

Definition 3 (*Pareto Set and Pareto Front*). A point $\mathbf{x} \in \mathcal{X}$ is called a *globally efficient point* of $\mathcal{X}$ (or of **f**) if there is no point $\tilde{\mathbf{x}} \in \mathcal{X}$ such that $\mathbf{f}(\tilde{\mathbf{x}}) \prec \mathbf{f}(\mathbf{x})$. The set of all globally efficient points of $\mathcal{X}$ is termed *Pareto set* (or globally efficient set) of **f** and denoted by $\mathcal{X}_\mathrm{E}$. The image of $\mathcal{X}_\mathrm{E}$ under **f** is called the *Pareto front* of **f** and denoted by $\mathbf{f}(\mathcal{X}_\mathrm{E})$.

Essentially, Definition 3 states that a point is globally optimal or efficient, if it is not dominated by any other point in $\mathcal{X}$ w.r.t. the Pareto dominance. For many years, in the evolutionary multi-objective optimization (EMO) community the "hunt" for the best Pareto front approximation was a predominant topic. A plethora of algorithms, selection mechanisms, benchmarks, and performance measures were designed to get closer to the true Pareto front, achieve the best diversity, and measure whether real-world as well as artificial problems could be adequately solved. Also challenges induced by localness were considered.

However, in order to define *local* optimality, a definition of localness based on the concept of neighborhood in the search space is needed. For isolated points, the only relevant neighborhood coincides with the point itself, and thus an isolated point is trivially a local efficient point (defined in Section 3). As $\mathcal{X} \subseteq \mathbb{R}^n$, we define *localness* by means of distances in the Euclidean space. Intuitively, a neighborhood of a point is any set containing the point such that you can walk from the point "some distance" in any direction without leaving the set. Classically, a neighborhood of a point in spaces equipped with a distance function is modeled by (a) an ϵ-ball containing the point, or, more generally, (b) any set which contains (or is equivalent to) an open set containing the point. Similarly, in the MO sense, a point can be called *locally* optimal or efficient, if it is not dominated by any of the points from its neighborhood.

Given additional restrictions on the class of considered functions, local optimality can be related to various other conditions. In case of differentiable objective functions, which are the focus of this work, the so-called *Fritz John conditions* (John, 2014) imply a necessary condition for local optimality. In the unconstrained case they read $\exists \lambda : \lambda \geq 0 \wedge \lambda \neq 0$, and w.l.o.g. $\sum_{i=1}^m \lambda_i = 1$, such that $\sum_{i=1}^m \lambda_i \nabla f_i(x) = 0$. In case of locally convex objective functions, (i) these conditions are also sufficient, and (ii) these necessary and sufficient conditions for local optimality are termed *Karush Kuhn Tucker (KKT) conditions* (Miettinen, 1998; Hillermeier, 2001).

Efficient points are interpreted as solutions for which there exists a weighted linear scalarization with non-negative weights (and at least one positive weight) such that the point under investigation is a local minimum. In most practical cases the conditions need to be extended by inequality constraints, such as box-constraints, as many local optima occur at constraint boundaries. For the full KKT conditions and their interpretation we refer the interested reader to the standard literature (Miettinen, 1998).

Interestingly, although the concept of localness is present in EMO and many algorithms implicitly try to solve benchmarks that explicitly integrate localness, until recently almost no insight into the actual MO landscapes – and thus hardly any interpretation of their shapes and characteristics – existed. As a consequence, visualization, algorithmic development, benchmark design and other areas concentrated on the "classical" Pareto front (and sometimes the Pareto set) for investigating the "landscape" of MOPs.

3. Refining definitions and notions of localness in continuous MOO

As a starting point for our following discussion, we argue that the classically used terminology in the EMO domain is currently not capable of capturing all expected properties of functional landscapes that we are used to from analogies in the context of single-objective optimization. If we speak of "Pareto optimality", "Pareto front", or "Pareto set" we usually refer to *global* optimality. Transferring these terms directly to the local structures – as, e.g., done in Liu et al. (2019) – and thus speaking of a "local Pareto set", lacks mathematical rigor and thereby would lead to potential confusion. For instance, in Liu et al. (2019), the notion of localness is understood via the set dominance relation in the objective space, i.e., the image of the global Pareto set (strictly) dominates that of a local Pareto set. Plausible as it seems, this notion brings in extra confusion when, for instance, if the pre-image of the (global) Pareto set is disconnected, which contradicts what we typically conceive. While plausible, this notion introduces additional confusion when, for instance, the pre-image of the (global) Pareto set is disconnected. This is contrary to what we normally conceive. Likewise, it is not (yet) accepted throughout the community to speak of efficiency instead.

In order to facilitate a concise wording on continuous multimodal MOO for the following sections – and maybe also as consolidation of available notions (Custódio and Madeira, 2018; Liefooghe et al., 2018c; Kerschke et al., 2016) for future discussions in the EMO domain – we provide some definitions and terminology that pick up the term of *efficiency* in the global and local sense, and use those to extend the current notion of optimality in the MO domain. Thereby, we specifically address multimodality in a broader sense than with a mere focus on multiple *global* optima (as done in Liu et al., 2018b). Nonetheless, the *multi-global* view is certainly a practically important aspect, even if not explicitly defined as part of our formal consideration. We will address this important area in Section 5.3 in more detail and from an algorithmic perspective. Note further that we do not strive for reinventing notions of localness — those are already discussed in early theoretical works that deal with local efficiency in MOO and aim for first and second order optimality criteria (e.g., refer to Wan, 1975; Van Geldrop, 1980; Jiménez, 2002).

In general, multimodal optimization problems are defined as problems that have more than one locally and/or globally efficient point (or both). The localness of the latter can also be defined by means of an open set in $\mathcal{X}$.

Definition 4 (*Locally Efficient Point*)**.** A point $\mathbf{x} \in \mathcal{X}$ is called a *locally efficient point* of $\mathcal{X}$ (or of $\mathbf{f}$) if there is an open set $U \subseteq \mathcal{X}$ with $\mathbf{x} \in U$ and there is no point $\tilde{\mathbf{x}} \in U$ such that $\mathbf{f}(\tilde{\mathbf{x}}) \prec \mathbf{f}(\mathbf{x})$. The set of all locally efficient points of $\mathcal{X}$ is denoted by $\mathcal{X}_{\mathrm{LE}}$.

In contrast to defining locally efficient points, the localness of an efficient set additionally requires the notion of connectedness.

Definition 5 (*Connectedness and Connected Component*)**.** The subset $A \subseteq \mathcal{X}$ is called *connected*, if and only if there do not exist two open subsets U_1 and U_2 of $\mathcal{X}$ such that $A \subseteq (U_1 \cup U_2)$, $(U_1 \cap A) \neq \emptyset$, $(U_2 \cap A) \neq \emptyset$, and $(U_1 \cap U_2 \cap A) = \emptyset$; or equivalently, there do not exist two non-empty subsets A_1 and A_2 of A which are open in the relative topology of A such that $(A_1 \cup A_2) = A$ and $(A_1 \cap A_2) = \emptyset$. Let B be a non-empty subset of $\mathcal{X}$. A subset C of B is a *connected component* of B, if and only if C is non-empty, connected, and there exists no strict superset of C that is connected.

Definition 6 (*Locally Efficient set and Locally Efficient Front*)**.** A subset $A \subseteq \mathcal{X}$ is a *locally efficient set* of $\mathbf{f}$, if A is a connected component of $\mathcal{X}_{\mathrm{LE}}$. The image $f(A)$ under $\mathbf{f}$ is called a *locally efficient front* of $\mathbf{f}$.

Of course, solutions in locally efficient sets, which are not dominated by any other solution are contained in the Pareto set. This implies $\mathcal{X}_E \subseteq \mathcal{X}_{LE}$. Following our distinction of globally and locally efficient points, we may also speak of a globally efficient set of solutions and a globally efficient front of solutions instead of a Pareto set or a Pareto front, respectively.

To further sharpen the notion of localness, we define the ε-neighborhood of a set in analogy to the ε-ball, which surrounds a single solution (see Definition 4).

Definition 7 (ε*-neighborhood of a Set*)**.** Let $A \subseteq \mathcal{X}$ and $\varepsilon > 0$. The set $A^{(\varepsilon)} := \{x \in \mathcal{X} \mid \exists a \in A \text{ with } \|x - a\| < \varepsilon\}$ is the ε*-neighborhood* of A, where $\|\cdot\|$ is the Euclidean norm.

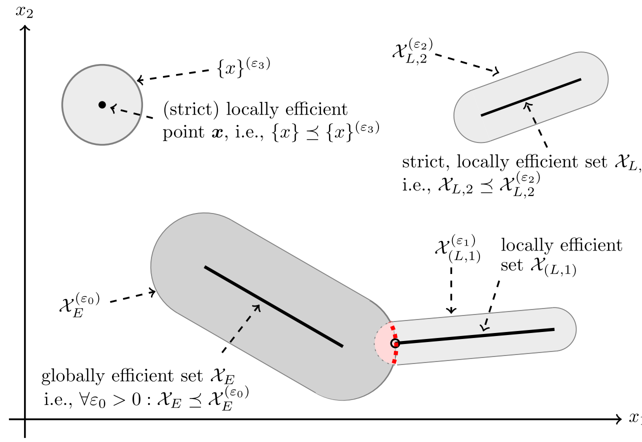

Fig. 2. Illustration of some of the introduced concepts for decision space $\mathcal{X} = \mathbb{R}^2$.

Definition 8 (*Strict, Locally Efficient Set*)**.** Let $C \subseteq \mathcal{X}_{\mathrm{LE}}$ be a locally efficient set. Then C is a ***strict, locally efficient set***, if and only if $\exists \varepsilon > 0$ such that $C \preceq C^{(\varepsilon)}$.

Clearly, for a locally efficient set C not to be strict, the following must hold. For each $\varepsilon > 0$ there are some points in the ε-neighborhood of C which are not dominated by C. That is, $\exists p \in C^{(\varepsilon)}$ such that $\forall c \in C : p \prec c$ or $p \parallel c$, where $p \parallel c$ means that p and c are incomparable. For instance, in case $\mathcal{X}_E$ is connected, then $\mathcal{X}_E$ is a strict globally (and thus also locally) efficient set, in fact $\forall \varepsilon : \mathcal{X}_E \preceq \mathcal{X}_E^{(\varepsilon)}$. If C is a connected component of $\mathcal{X}_{LE}$, then both cases can occur: C is strict or C is not strict; here a crucial role is played by incomparability. For an illustration on how a locally efficient set can fail to be a strict locally efficient set see Fig. 2 for decision space $\mathcal{X} = \mathbb{R}^2$. Therein, the set $\mathcal{X}_{LE}$ is the union of the connected components $\{x\}$, $\mathcal{X}_{L,1}$, $\mathcal{X}_{L,2}$ and $\mathcal{X}_E$, where $\mathcal{X}_E$ is the globally efficient set. We further assume that $\{x\}$ and $\mathcal{X}_{L,2}$ are *strict*, locally efficient sets, i.e., $\exists \varepsilon_3 > 0 : \{x\} \preceq \{x\}^{(\varepsilon_3)}$ and $\exists \varepsilon_2 > 0 : \mathcal{X}_{L,2} \preceq \mathcal{X}_{L,2}^{(\varepsilon_2)}$. The set $\mathcal{X}_{L,1}$ is not strict for the following reason: it is a half-open segment of a straight line, whose leftmost end point is excluded (indicated by an open circle in the image). Moreover, it is an accumulation point of $\mathcal{X}_E^{(\varepsilon_0)}$, i.e., for a fixed $\varepsilon_0 > 0$ there exists $\varepsilon_1 > 0$ such that *none* of the points of $\mathcal{X}_E^{(\varepsilon_0)} \cap \mathcal{X}_{L,1}^{(\varepsilon_1)}$ (the red intersecting area in Fig. 2) is dominated by $\mathcal{X}_{L,1}$. This gives an illustration of a locally efficient set which is not strict.[2] Note that there exist many other reasons for a locally efficient set not to be strict. Also, if for a point $p \in \mathcal{X}$ it holds true that $\{p\}$ is a strict locally efficient set, then there exists a neighborhood of p which will not contain any points that are incomparable to p.

Note that definitions for locally and strictly locally efficient sets are also given, for example, in Liefooghe et al. (2018c) and Paquete et al. (2004). However, these works are rooted in the combinatorial domain and they therefore consider different search spaces and different neighborhood relations. As a result, their understanding of (strictly) locally efficient sets differs slightly from our perspective (as presented in Definitions 6 and 8) and their definitions thus do not consider the concept of locally efficient sets as defined in Definition 6 of our work.

A further difference is that we require locally efficient sets to be connected. Furthermore, the Pareto front of $\mathbf{f}$ is obtained by taking the image under $\mathbf{f}$ of the union of connected components of $\mathcal{X}_{\mathrm{E}}$. If $\mathcal{X}_{\mathrm{E}}$ is connected and $\mathbf{f}$ is continuous on $\mathcal{X}_{\mathrm{E}}$, the Pareto front is also connected.

[2] For a numerical example, we refer to Section 4, where a simple problem (denoted as Aspar's problem) is visualized in Fig. 4. Therein, a non-strict locally efficient set is shown, which is open at one end. There, we find points $p \in C^{(\varepsilon)}$ such that $\forall c \in C : p \prec c$. In fact, these points are part of the "attraction basin" of the globally efficient set.

With a view towards algorithms, which compute approximations to (locally) efficient sets and/or (locally) efficient fronts, one needs to be able to identify whether a finite approximation set is a subset of a locally efficient set. To cope with such finite sets, we need to adopt the concept of ε-connectedness.

Definition 9 (*ε-connectedness*)**.** Let $\varepsilon > 0$. A subset $S \subseteq \mathcal{X}$ is *ε-connected* if for any two points $s, s' \in S$ there exists a finite subset $\{s_1, \dots, s_k\} \subseteq S$ with $s_1 := s$ and $s_k := s'$ such that $\|s_i - s_{i+1}\| \leq \varepsilon$ for all $i = 1, 2, \dots, k-1$.

Definition 10 (*Finite ε-locally Efficient Set*)**.** Let $\varepsilon > 0$ and let S be a *finite* subset of $\mathcal{X}_{\mathrm{LE}}$. Then S is an *ε-locally efficient set*, if $S \neq \emptyset$, and S is ε-connected.

Note that the points of an ε-locally efficient set could in theory belong to *multiple* locally efficient sets, if the magnitude of $\varepsilon > 0$ is larger than the smallest distance between the points from two adjacent locally efficient sets. Thus, a finite ε-locally efficient set bridges the gap between a MOP's locally efficient sets and the approximation sets that are usually generated by MOO algorithms. Practically, it still remains to develop an algorithm to choose a proper value for ε.

Considering the previous definitions and terminology, we finally provide a definition for a multimodal MOP.

Definition 11 (*Multimodal Multi-objective Optimization Problem, MMMOP*)**.** Let $f : \mathcal{X} \to \mathbb{R}^m$ be an MOP as defined in Definition 1 and $g : \mathcal{X}_{LE} \to f(\mathcal{X}_{LE})$, where $\mathcal{X}_{LE}$ is the set of locally efficient points of f according to Definition 4. We denote f *multimodal*, if at least one of the following two conditions is satisfied:

1. g is a surjection, but no injection (i.e., it is not bijective).
2. $\mathcal{X}_E \subset \mathcal{X}_{LE}$.

Otherwise, i.e., if g is a bijection and all efficient points in $\mathcal{X}$ are globally efficient ($\mathcal{X}_{LE} = \mathcal{X}_E$), the MOP f is denoted *unimodal*.

This definition unifies the previously described different perspectives on multimodality: multiglobality and the wider perspective including locally efficient sets. To improve comprehensibility, we depict the conditions of Definition 11 in a matrix view (see Fig. 3). It is clear that merely MOPs, for which g is bijective and which comprise a Pareto set, but no other local efficient set are considered to be unimodal. All other problems are multimodal and either comprise multiple global efficient points in $\mathcal{X}_E$ that map to the same image in objective space (*multiglobal*), comprise additional locally efficient sets (*multilocal*), or both.

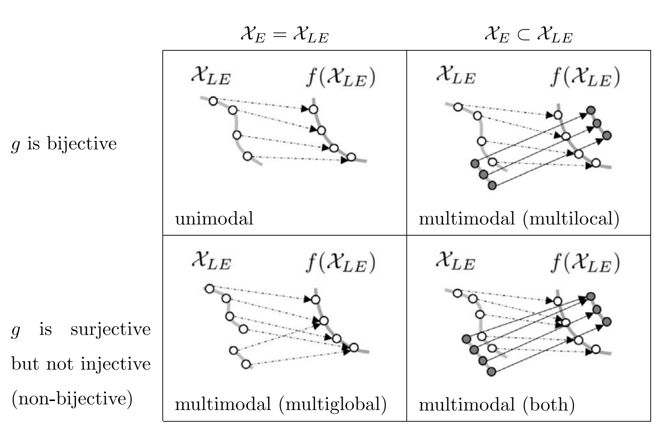

Fig. 3. Schematic depiction of the definition and "types" of uni- and multimodal MOPs.

4. Visualization of multi-objective landscapes

In order to complement the theoretical definitions and terminology of the previous section, we next provide a brief discussion of current visualization techniques and available recent means to show the aforementioned properties of MOPs in landscape depictions, which resemble the intuition known from SO.

Considering that one can depict at most three dimensions, illustrating multiple decision variables *and* objectives simultaneously is nearly impossible. In consequence, most publications in the MO community only display the problems' Pareto fronts (or approximation sets thereof) in two- or three-dimensional scatter plots. That is, the entire structure of a MOP is ignored and instead, the problem is reduced to its Pareto front or set. Similarly, the quality of a MO optimizer is defined solely based on its Pareto front approximation. The SO analog to this perspective would be reducing the whole landscape of the SO problem to its global optimal point(s) in decision *or* objective space — and assessing algorithmic search behavior solely based on its distance to the respective globally optimal point(s). This neglects many structures algorithm developers are interested in. One of the most interesting (and also most intuitive) properties of each point is its corresponding *basin of attraction*. An attraction basin is defined as the set of points from which a gradient descent leads to the same locally efficient solution (the attractor). Specifically, in a multimodal problem, multiple of these basins exist and for deterministic gradient descent (and many other local optimizers that follow an approximated gradient path) these basins may resemble local traps. But also other structural properties like discontinuities along with the imposed challenges well-known for SO optimization algorithms (Schwefel, 1993), would be (visually) neglected.

This very limited – though far too often used – representation of MOPs is particularly surprising, since some suitable approaches have been developed over the years, which reveal at least some characteristics of a problem landscape's structure. For instance, the *cost landscapes* proposed in Fonseca's PhD thesis (Fonseca, 1995) constitute one of the first approaches, which are capable of illustrating structural relationships in a MOP's decision space. The height of a point in such a cost landscape is defined by its (normalized) Pareto rank, i.e., the number of solutions, which dominate the respective point. In consequence, this approach not only depicts globally optimal regions in the decision space, but also reveals related characteristics such as symmetry w.r.t. global optima — due to the ranking of the points. The only potential drawback of this approach is its focus on the Pareto rank and thereby on the *global* optimality of the problem. Therefore, as illustrated in the first row of Fig. 4, this method is incapable of illustrating *local* optima and/or the corresponding basins of attraction (as multi-objective equivalents to the SO case). However, both are important aspects for investigations of *multimodal* landscapes.

Nonetheless, depicting MOPs using Fonseca's visualization approach is much more informative than a pure reduction to the Pareto front, and already reveals interactions among the objectives. In consequence, it has been used for the depiction of various MOPs – including the benchmark suite of CEC 2019 (Yue et al., 2019) – despite its multiglobal emphasis.

The up-to-now most comprehensive overview of visualization approaches for MO landscapes can be found in the PhD thesis of Tušar (2014). Alternatively, one can study the succeeding publication by Tušar and Filipič (2015), which constitutes a compact version of the aforementioned thesis. Within both works, a variety of visualization methods – ranging from scatter plot matrices and parallel coordinate plots, via sammon mapping and principal components, to hyper-space diagonal counting and hyper-radial visualization – are reviewed. However, as discussed in more detail in both works, all of the considered approaches have certain limitations. Therefore, the authors proposed an alternative visualization technique named *prosection* (Tušar, 2014; Tušar and Filipič, 2015). This approach allows to depict a user-specified

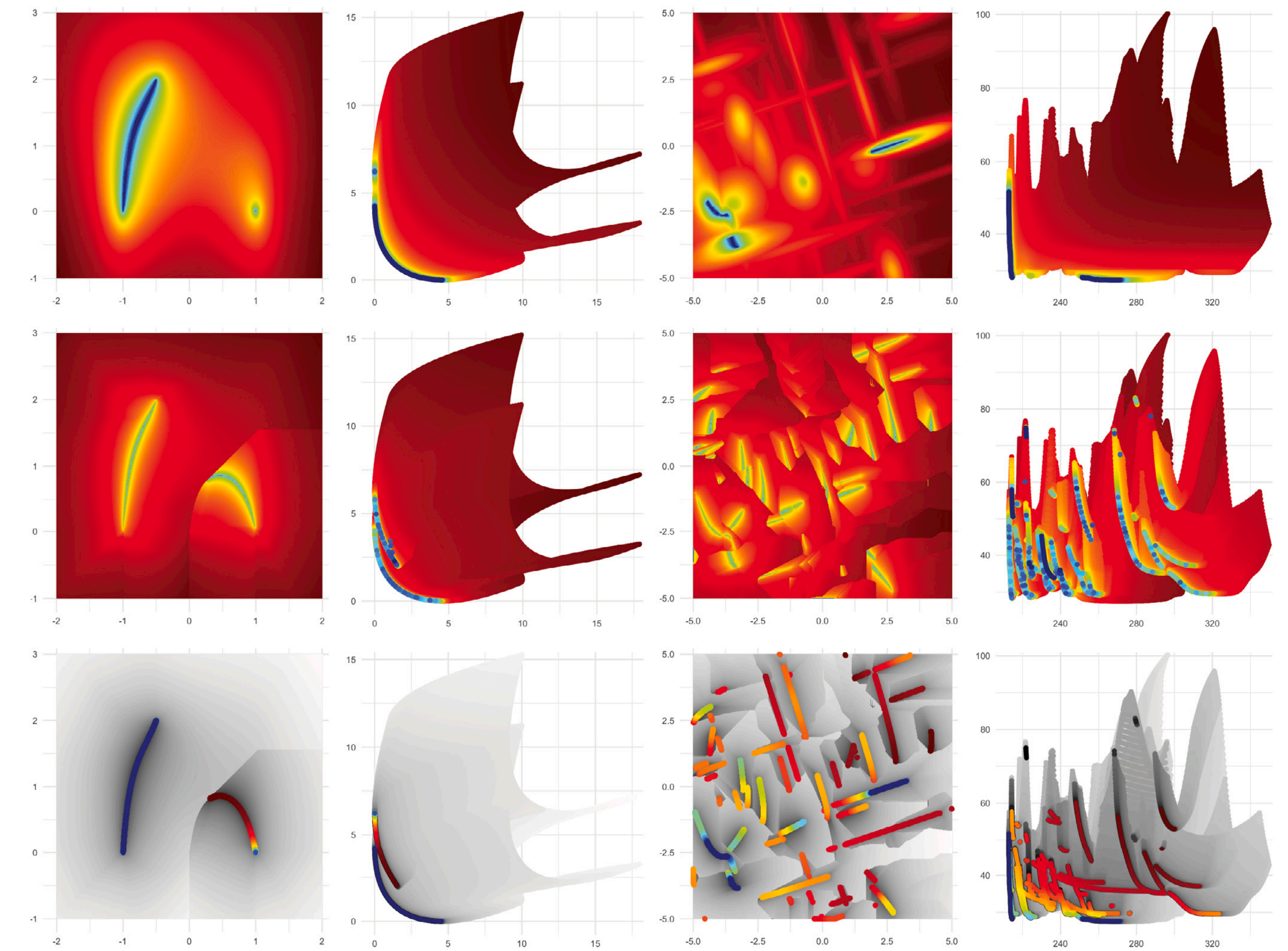

Fig. 4. Comparison of the three most sophisticated techniques for visualizing MOPs in decision (odd columns) and objective space (even columns), respectively. Shown images are based on two exemplarily MOPs: left two columns correspond to the bi-objective polynomial $f(x_1, x_2) = (x_1^4 - 2 \cdot x_1^2 + x_2^2 + 1, (x_1 + 0.5)^2 + (x_2 - 2)^2)$, which we denote *Aspar Function*, whereas the six images on the right are based on an instance (FID: 10, IID: 7) from the bi-objective BBOB test suite (Tušar et al., 2016). The three visualization methods are distinguished by row (top to bottom): the Pareto ranking based on the *cost landscape* method of Fonseca (1995), the cumulated gradient field heatmaps (Kerschke and Grimme, 2017), PLOT technique (Schäpermeier et al., 2020). All 'heights' (i.e., colors) are shown on a logarithmic scale to put more emphasis on the (locally) optimal solutions in decision and objective space, respectively.

section of a higher-dimensional space in a lower-dimensional scatter plot via projections of points. More precisely, points that are located in a certain section of the considered high-dimensional space (i.e., in the proximity of a user-specified hyperplane) are (orthogonally) projected onto the associated hyperplane. Unfortunately, from a practical point of view, this method is not very feasible for illustrating MO landscapes as each projection comes with a strong loss of information — all points that are not in the "close proximity" of the cutting hyperplane will be discarded. In consequence, this method is at most beneficial for the very special case of $m = 4$ objectives; for $m \leq 3$ one can depict all objectives without any projections and for $m > 4$, one would need to perform multiple projections (and thereby sequentially reduce the dimensionality by one) until one reduced the original data to a three-dimensional space. Moreover, all discussed methods have been investigated w.r.t. their ability of displaying points based on their representation in objective space, i.e., independent of the corresponding decision space. For multimodal investigations, this is again not beneficial. Yet, neglecting the high loss of information, these methods could nonetheless provide useful insights as one could simply apply them to the combination of search and objective space.

For the continuous domain, the first method that combines all of the aforementioned desired properties – i.e., simultaneous visualization of decision *and* objective space, as well as depiction of local optima – has been proposed by Kerschke and Grimme (2017). It depicts the interaction effects (of the different optima) of the problem's *objectives* in the *decision* space using the notion of a gradient descent direction (Cauchy, 1847). Based on a grid of points that is spanned across the whole decision space, the bi-objective gradient is computed per point. This MO gradient is defined as the sum of normalized gradients of both individual objectives. Each of the combined gradients – one gradient per point of the grid – is directed towards the largest simultaneous improvement w.r.t. both objectives. The length of such a MO gradient – i.e., a value between zero and two – indicates the steepness of the *gradient landscape* in that particular point. By "following" the path of gradients downhill towards a point whose gradient is zero (i.e., a locally efficient point, see Definition 4) one reaches (a part of) a MO local optimum (i.e., a locally efficient set, see Definition 6). Note that analogous to SO optimization, points whose gradients lead to the same (set of) locally efficient point(s) form the corresponding *basin of attraction*. Noticeably, the proposed approach not only allows to depict locally optimal sets, as well as their basins of attraction in the search space, but can also be used to depict the corresponding images in the objective space (Kerschke et al., 2019b) (see images in the second row of Fig. 4). And although this method has originally been introduced

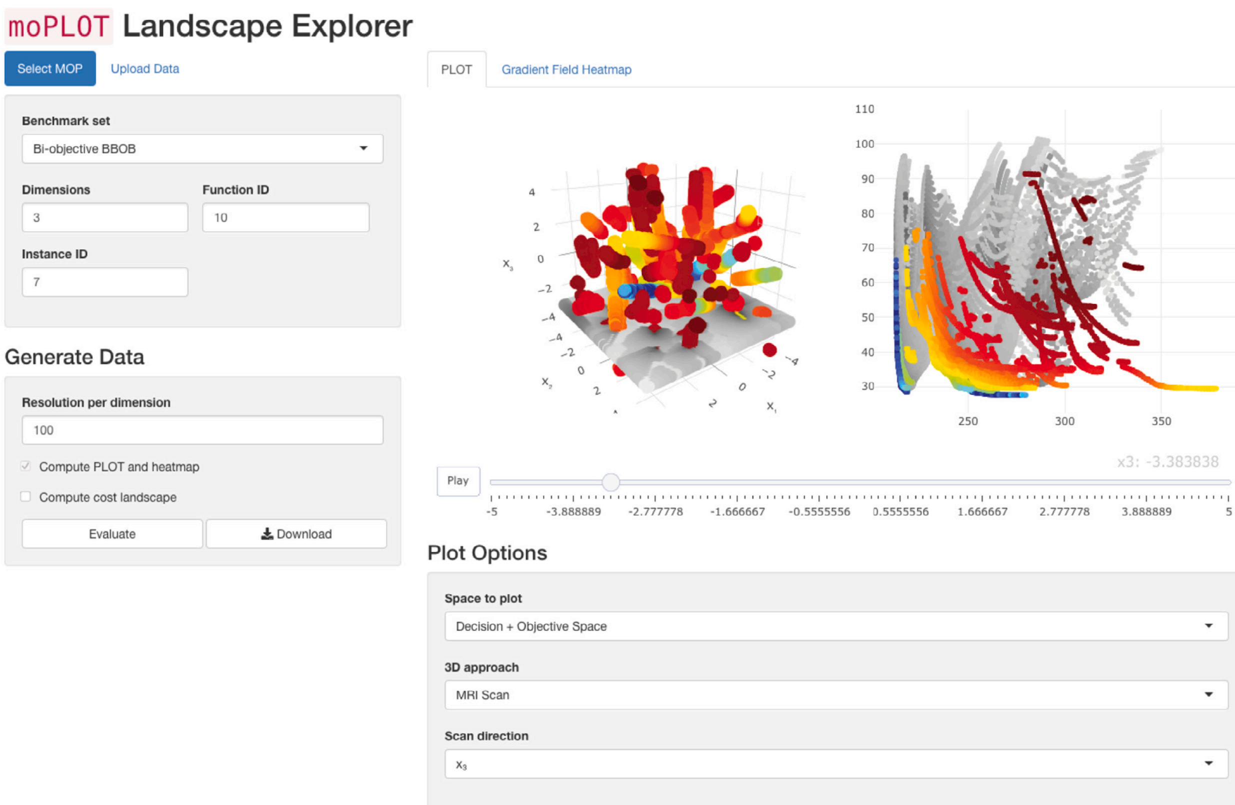

Fig. 5. An exemplary MOP from bi-objective BBOB (Tušar et al., 2016) (FID: 10, IID: 7), shown in its 3D search (left) and 2D objective space (right) using the interactive moPLOT dashboard (Schäpermeier et al., 2021). The colored points indicate the location and dominance relationship of the MOP's efficient sets and their corresponding fronts. Further, the MRI scan method used herein illustrates slices of the search space and their corresponding objective values by means of gray shaded points.

for MOPs with two-dimensional search and objective spaces, it can actually be used to visualize any MOP, as long as either the search or the objective space is two-dimensional.

Noticeably, the heatmap's strong emphasis on localness is not purely beneficial. Of course, it is advantageous to easily identify local optima and their surrounding attraction basins. However, only looking at the locally efficient sets in decision space does not help in identifying a quality ranking of the different sets in the decision space. This approach is lacking a global perspective.

The recently proposed *Plot of Landscapes with Optimal Trade-offs (PLOT)* (Schäpermeier et al., 2020) combines the advantages of both aforementioned methods — the global perspective of cost landscapes (based on Pareto ranking) and the local view of the gradient-based heatmaps. As shown in the bottom row of Fig. 4, PLOT displays the same attraction basins as the aforementioned heatmap approach — but uses a less flamboyant (gray) color scheme. Here, darker values indicate points that are closer to their respective local optimum. On top of those gray-colored attraction basins, PLOT displays the corresponding local optima. But in contrast to the heatmap approach, the sets of local optima are colored according to their dominance relationship: dark blue segments correspond to global optima, whereas red segments correspond to inferior local optima. As visually confirmed in the bottom row of Fig. 4, PLOT is even capable of revealing the global structure of highly multimodal MOPs, such as the bi-objective BBOB problem (see right half of Fig. 4).

In addition, the visualization methods based on MO gradients are capable of depicting locally efficient sets as defined in Section 3 (and shown schematically in Fig. 2). As exemplified in Fig. 4, the landscapes consist not only of color (or gray) shaded basins of attractions surrounding the locally or globally efficient sets. The images also show locally efficient sets that are either completely surrounded by a basin of attraction or intersected by a basin with dominating solutions resulting in a ridge-like structure in decision space. This fits exactly with what was shown in Fig. 2: a globally (and also strictly locally) efficient set dominates and overlaps another locally efficient set and its basin of attraction. Later, we will return to precisely this problem to get a first idea of how to exploit these structures for algorithm design (see Section 5.3).

In an attempt to consolidate recent developments in MOP visualization and to facilitate the use of the underlying techniques, Schäpermeier et al. (2021) recently published a user-friendly dashboard that enables interactive and platform-independent exploration of MOP landscapes.[3] The dashboard also provides two methods for interactively exploring three-dimensional search spaces of MOPs: one based on the idea of MRI scans (see Fig. 5 for an exemplary visualization of a bi-objective BBOB problem) – primarily known from the medical field – and one based on a layered visualization of so-called isosurfaces (i.e., 3D level sets).

Building on our understanding of multimodality in continuous MOO (Section 3), we will use the visualization methods presented in this section to examine the structures of various widely used MOPs (especially in terms of their multimodality) in Section 5.1. Subsequently, we will discuss the resulting implications for characterizing these problems, as well as for algorithmic search behavior in Sections 5.2 and 5.3, respectively.

5. Implications for multi-objective optimization

With the previously consolidated terminology of localness in MOO (Section 3) and the visual insights into MOP landscapes (Section 4), we examine the potentials of considering multimodality in EMO more

[3] The dashboard is, e.g., available at https://schaepermeier.shinyapps.io/moPLOT.

strongly. In particular, we focus on benchmarking, automated characterization of (multimodal) MOP landscapes, the related impact on algorithm development (even in SO optimization), and the associated challenges in performance assessment.

5.1. Benchmarks

When dealing with MO benchmark problems, a rather poor understanding of the landscapes' structural properties and their associated impact on an algorithm's search behavior is observable. In fact, various MOPs have been created based on principles that were initially designed for *single-objective* test problems (Whitley et al., 1995; Bäck et al., 1997) – without ever questioning their applicability in the context of MOO.

Deb (1999) discussed the difficulties that a MO *genetic* algorithm (GA) might face when solving a MOP and explicitly listed multimodality as one of those challenges, because it could hinder the GA's convergence to the Pareto front. Unfortunately, this statement has oftentimes been quoted outside of its original context and thereby developed into a general claim about potential obstacles a MO search algorithm is facing. In combination with the design principles mentioned above, this led to the common – yet false – belief that multimodality is a challenging structural property in general.

Within the following two subsubsections, we give an overview of various benchmark suites that are frequently used for the comparison of MO optimizers, and address their relationship to multimodality (according to our understanding). While Section 5.1.1 lists test suites that are used for general benchmark comparisons, Section 5.1.2 focuses on benchmark suites that have explicitly been designed from a multimodal perspective.

5.1.1. General MO benchmarks

Kursawe (1990) and Viennet et al. (1996) were among the first who proposed MO test functions. Although multimodality was not of particular interest back then – instead one was mainly looking for MOPs in general – their test functions already contained a variety of local optima as shown in their colorful PLOTs in Fig. 6. The actual awareness for multimodality has increased around the millennium, when the nowadays well-known MO benchmark suites MOP (van Veldhuizen, 1999) and ZDT (Zitzler et al., 2000) were proposed. Within the corresponding works, the respective authors stated that multimodality poses a challenge in SO optimization and stochastic algorithms usually perform better on such problems. They thus concluded that their MO test suites should also contain multimodal MOPs. In line with the previous argumentation, the DTLZ test problems (Deb et al., 2005) were constructed such that they ensure "controllable hindrance to converge to the true Pareto front" (for EMOAs). Emmerich and Deutz (2007), who introduced a set of MOPs based on Lamé superspheres, also approached MO multimodality via local convergence of an algorithm; yet, at the same time, they also emphasized the "vital importance" of this property for the assessment and comparison of MO algorithms.

More than a decade ago, Huband et al. (2006) comprehensively reviewed various MO test suites and identified a lack of multimodal problems. Noticeably, within their work they reduced MO multimodality to the SO case by arguing that MOPs are a superset of SO problems, and therefore design principles of MO test suites should be a superset of guidelines from the SO domain. According to their definition, the MOPs in their proposed WFG test suite (Huband et al., 2006) are multimodal, if one of its underlying objectives is a multimodal function.

The ten (unconstrained) test functions UF1 to UF10, which Zhang et al. (2008) proposed for CEC 2009, were the first MOPs designed for and used within a conference's competition. However, they did not explicitly address the aspect of multimodality; instead, their main motivation apparently was providing additional MOPs for benchmarking purposes in the community.

In 2016, Tušar et al. proposed the bi-objective BBOB test suite (Tušar et al., 2016), which essentially is a collection of 55 bi-objective problems resulting from an exhaustive, pairwise concatenation of a subset of ten SO functions from BBOB (Hansen et al., 2009) – which is the gold standard in SO optimization. Although the authors did not explicitly discuss the issue of MO multimodality, they support the concept of knowledge transfer from the SO to the MO domain by claiming that MOPs inherit the properties of their SO components. According to our visualizations many of the bi-objective BBOB problems look indeed highly multimodal as exemplarily depicted (for its 52nd problem) in Fig. 6.

In addition to the aforementioned test suites, a variety of (specialized) MOPs have been proposed over the last decade. Noticeably, for none of them multimodality is the primary focus. For instance, the main objective of the difficult-to-approximate (DtA) test problem generator (Wang et al., 2018) was the generation of MOPs for which the extreme solutions of the respective Pareto fronts (i.e., the global optima of the underlying objectives) are difficult to approximate. However, the authors were at least aware of multimodality as they explicitly stated their generator's capability to create multimodal MOPs. The framework proposed by Saxena et al. (2011) also enables users to generate MOPs with many and/or complex Pareto sets. Yet, the primary focus of the Saxena–Zhang–Duro–Tieari (SZDT) problems are many-objective problems. Complementing the aforementioned SZDT problems, which aimed at a high-dimensional objective space, the large-scale multi-objective problems (LSMOPs) proposed by Cheng et al. (2016) emphasize high-dimensional decision spaces. According to the review of scalable MOPs by Zapotecas-Martínez et al. (2018), seven out of the nine LSMOPs are multimodal. Probably the most recent collection of (multimodal) MOPs are the distance-based multi/many-objective point problems (DBMOPP) proposed by Fieldsend et al. (2019). Although the authors specifically advertise attributes such as the generation of MOPs with disjoint Pareto sets or varying objective scales as strengths of their framework, their generated MOPs also revealed multimodal structures.

Further, Cheng et al. (2017) developed MaF as suite of many-objective test functions, Glasmachers (2019) analyzed the challenges of convex quadratic bi-objective problems, and the games-related benchmark GBEA (Volz et al., 2019) by Volz et al. aims at making the interesting properties of games-related problems available to the EC community. All three works strongly focus on the emphasized characteristics (i.e., many-objective, convex quadratic bi-objective, and relation to games), but do not explicitly address multimodality.

5.1.2. Focus on multimodal landscapes

Despite the plethora of MO benchmarks – and the multimodal problems occasionally found therein – the set of MOPs considered in Kerschke et al. (2016, 2019b) potentially is the first MO test suite with a particular focus on multimodality. Similar to bi-objective BBOB, the problems contained therein are concatenations of several multimodal SO problems — and each of them results from superpositioning multiple unimodal functions. The respective problem instances can be easily created using the scalable *multiple peaks model 2 (MPM2)* generator by Wessing (2015a,b), which is an enhanced version of the MPM generator utilized in Preuss (2015). Implementations of MPM2 are, e.g., available for python (`optproblems`, Wessing, 2016) and R (`smoof`, Bossek, 2017). As demonstrated in Fig. 6, these MOPs can be kept simple (e.g., only few basins and a mostly smooth landscape) to facilitate the understanding of interactions among the objectives, but on the other hand can easily be extended to highly multimodal problems.

One of the currently most prominent test suites for multimodal MOO is the collection of Yue et al. (2019). It comprises a diverse set of scalable test problems (see, e.g., the third and sixth row of Fig. 6) and provided the basis for the *multimodal MOO competition*[4] at CEC

[4] http://www5.zzu.edu.cn/ecilab/info/1036/1171.htm.

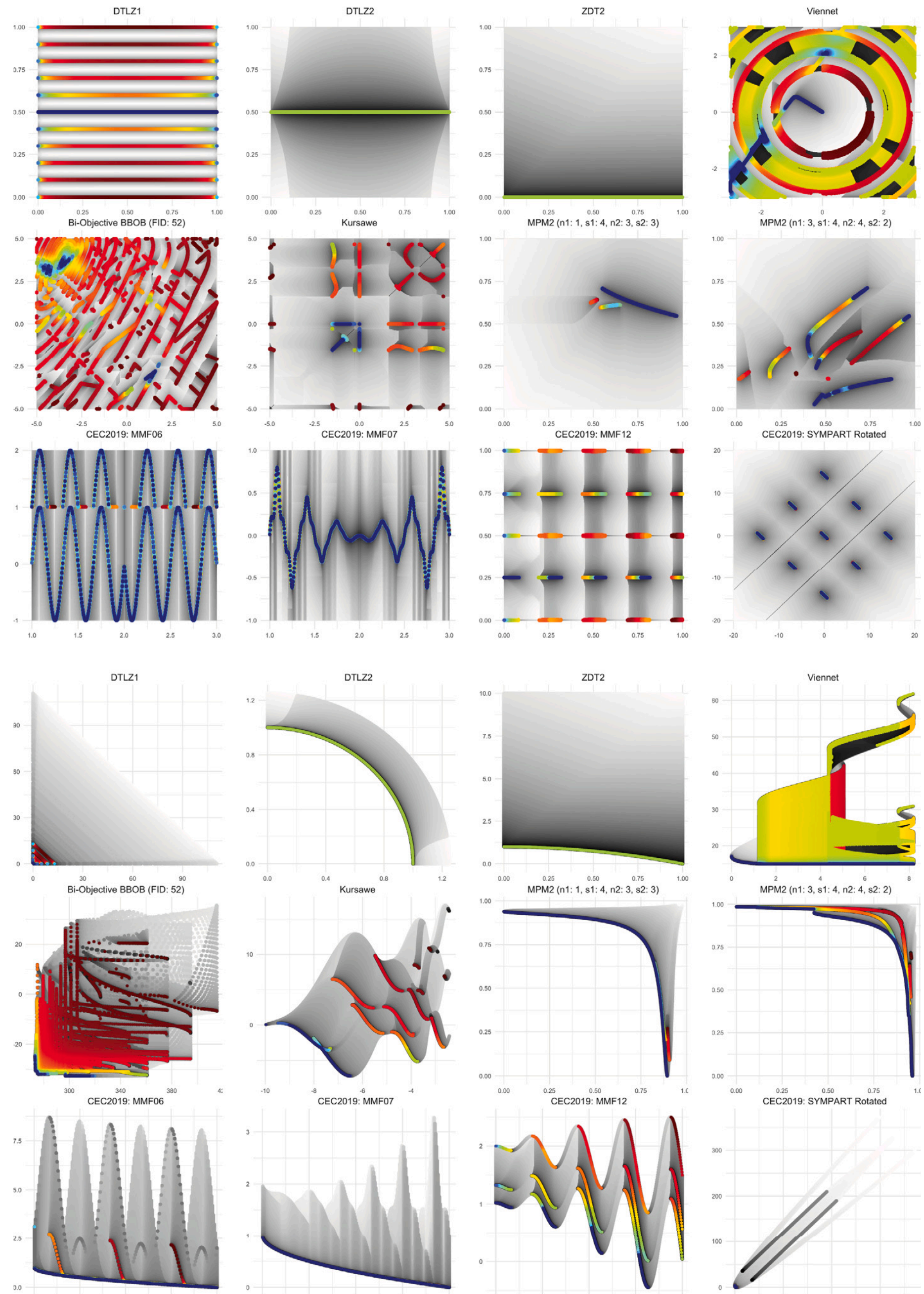

Fig. 6. Exemplary, yet representative MOPs depicted in decision (top three rows) and objective space (bottom three rows), respectively, using *PLOTs*. While some problems (e.g., DTLZ2, ZDT2 and MMF7) possess only a single set of locally (and thus globally) efficient points, other MOPs like the depicted bi-objective BBOB problem are highly multimodal.

2019 (Liang et al., 2019). Although the posed problems indeed possess challenging (and for gradient-based approaches even deceptive) traits, it should be noted that the vast majority of them is either *unimodal* or *multiglobal* – but not *multilocal* (see Definition 11). That is, many of these MOPs contain multiple efficient sets, however, they usually all map to the (same part of the) Pareto front. Therefore, this test suite addresses only a specific type of multimodality.

The most recent collection of multimodal MOPs are so-called multi-polygon test problems (Ishibuchi et al., 2019; Peng et al., 2019), whose objectives are the minimal distances to different polygons. However, they defined the localness of Pareto sets in terms of set dominance in the objective space, which is different from the perspective that we propose here.

5.1.3. Perspectives for benchmarks with multimodal MO test problems

Despite the plethora of MO test suites, we are still lacking a truly multimodal benchmark suite. The collection by Yue et al. (2019) – which in addition to their proposed MMF problems also contains multimodal MOPs from other works, e.g., SYMPART Rotated (Rudolph et al., 2007) – already points in the right direction. However, their understanding of *multimodality* is rather related to the diversity in decision space (and similar quality in objective space) than to the challenges resulting from efficient sets whose associated fronts are of different quality. According to their function definitions, and as confirmed by Fig. 6, the majority of their MOPs possess only very few (and occasionally disconnected) local *fronts*.

Another perspective for future work in this research area is a systematic (and visually supported) assessment of the aforementioned test suites — with an emphasis on their structural properties like multimodality. Compared to the comprehensive overview by Huband et al. (2006), this overview could (a) make use of the new visualization methods, (b) consider the refined terminologies (e.g., multimodality, localness), (c) integrate benchmarks from the last one-and-a-half decades (Huband et al., 2006 has been published in 2006), and thereby (d) reveal properties that are, if at all, only poorly captured by the current MO test suites.

5.2. Features for characterizing MOPs

When confronted with the task of optimizing a continuous problem, having at least a vague idea of the landscape's structure is usually highly beneficial. Already a slightly better understanding of the landscape, including the associated structural challenges for algorithms acting on it, facilitates the selection and/or configuration of a suitable, powerful optimization algorithm (Kerschke et al., 2019a; Kerschke and Trautmann, 2019a; Eggensperger et al., 2019). However, the characterization of problem landscapes – a research area that is usually termed *Exploratory Landscape Analysis (ELA)* or sometimes also *Fitness Landscape Analysis (FLA)* – already poses various challenges in case of SO optimization as elaborated in detail in the corresponding surveys of Malan and Engelbrecht (2013), Muñoz Acosta et al. (2015), or Kerschke and Trautmann (2019b).

However, all of these features have explicitly been designed for the characterization of SO landscapes and thus are incapable of capturing effects caused by the interaction of the different objectives. For instance, in Kerschke and Trautmann (2016) the authors tested the suitability of some of the aforementioned ELA features for characterizing bi-objective instances of the well-known benchmark suites DTLZ (Deb et al., 2005) and ZDT (Zitzler et al., 2000). Within their study, they first computed the features per pair of problem and objective, and then computed a feature vector for the bi-objective problems by taking the ratio of the two feature vectors of the SO components. The associated correlation matrix has then been investigated w.r.t. relations and/or differences among the problems. Surprisingly, some (dis-)similarities could be observed, yet, in the end, the proposed method is insufficient of characterizing MO problems, as interactions between the objectives were only weakly respected (via the ratio of the SO features).

Another step towards measuring features of MO landscapes has been made in a previous work (Kerschke et al., 2019b) of this manuscript's authors. Therein, we followed a white-box approach – i.e., we assumed full knowledge of the MOP's function – and designed a first set of characteristics that quantify information that is contained within the problems' landscapes. This set of pseudo-features contains measures such as the number of locally efficient sets, the relative length of the Pareto front (w.r.t. the cumulated length of all local fronts) or the number of local fronts that are connected to (parts of) the Pareto front. Although these features definitely are capable of respecting interactions among the (local) optima of the different objectives, we want to emphasize that they are mainly of exploratory character (due to their white-box approach).

Moreover, recent works on the visualization of MOPs (Grimme et al., 2019a,b) revealed very strong patterns among common MO benchmark suites. While problems of historic and long-established benchmarks such as DTLZ (Deb et al., 2005) and ZDT (Zitzler et al., 2000) possess a rather smooth structure, landscapes of concatenated (multimodal) SO problems, e.g., the ones from bi-objective BBOB (Tušar et al., 2016), have a highly scattered appearance. Similar findings can be made for the other MO benchmarks. Thus, from a structural point of view, MOPs appear to be rather homogeneous within a benchmark suite, but heterogeneous between them. Therefore, features that actually quantify (size and number of) the basins of attraction, or measure the landscape's smoothness or ruggedness, could potentially be very useful for feature-based approaches such as automated algorithm configuration and/or selection. For instance, one could adapt features that extract the corresponding information in the SO domain (Lunacek and Whitley, 2006; Malan and Engelbrecht, 2009; Mersmann et al., 2011; Kerschke et al., 2014).

There are some features and visualization techniques used in multimodal SO optimization that cannot be easily generalized to multimodal MO optimization. Obviously, the *number of local optima* cannot be generalized. For discrete landscapes, however, we may well utilize this feature. More advanced features in landscape analysis are derived from tree- and network-based descriptions of multimodal landscapes, such as barrier trees (Stadler and Flamm, 2003) and local optima networks (Fieldsend and Alyahya, 2019). In both cases, it is possible to generalize them to the discrete MO case, but in the continuous case there would be infinitely many nodes in such trees or networks.

In addition to the development of features that capture the information of the basins of attraction – and thereby measure the problem's degree of multimodality – there is a great need for research on how the interaction of a problem's objectives can be measured. For instance, one could consider techniques that have initially been designed for measuring the interaction among variables in general (Reshef et al., 2011; Sun et al., 2017) – i.e., independent of the space (search, objective, etc.) they belong to – and adapt them to capture the respective information among the objectives.

In the end, multimodality not only affects the automated feature-based characterization of the MOPs' landscapes, but also the behavior of search algorithms operating on them. Therefore, we will in the following investigate how algorithm engineers utilized this additional information.

5.3. Algorithmic ideas

In our opinion, the potential of investigating localness for MOPs can build on the previous aspects of visualization, benchmarking, as well as feature identification. It offers a good opportunity to push research even beyond current endeavors in what is called multimodal MOO. In the following, we will discuss this corpus of literature and then propose possible and promising directions of research that may advance EMO and broaden development, application and transfer of new insights.

Today, the algorithmic consideration of multimodality in (continuous) MOO is mainly focusing on what we considered being a subset of our multimodal perspective, denoted as *multiglobal* optimization (see Definition 11). In most cases, this refers solely to what was shown in Fig. 2 (left) in Section 1: the challenge to find all parts of the globally efficient set — even if those solutions map to the same representation in objective space.

In fact, finding multiple realizations (in decision space) with equivalent quality (in objective space) is of major practical importance. In case a found globally efficient solution is not realizable (e.g., in production processes, where only a certain precision can be achieved, or when secondary restrictions hinder the production in principle), the availability of another (and sometimes very different) solution of equivalent quality is very welcome.

The main goal of all approaches in that stream of finding all parts of the globally optimal set is to preserve solution diversity not only on the Pareto front, but also on the Pareto set. As different (and potentially very diverse) globally efficient solutions in the Pareto set may map to the same image in objective space (see, e.g., SYMPART Rotated (Rudolph et al., 2007; Yue et al., 2019) in the bottom right of Fig. 6), standard EMO approaches (like NSGA-II Deb et al., 2002 or SMS-EMOA Beume et al., 2007 to name only two of the most famous ones) tend to lose diversity in decision space due to diversity preservation mechanisms that solely operate in the objective space.

5.3.1. A summary of current multimodal MO optimization algorithms

From a pure algorithmic point of view, approximating as many as possible globally optimal solutions demands (a) preservation of (near) efficient solutions during (b) an optimization process that covers the whole decision space as good as possible. Therefore, various approaches have been proposed: One of the earliest approaches is the inclusion of an additional archive to preserve decision space diversity. SPEA2+ (Kim et al., 2004; Hiroyasu et al., 2005) extends the original SPEA2 (Zitzler et al., 2001) approach by considering two archive sets for objective and solution space, respectively. Also by archiving, 4D-Miner (Sebag et al., 2005; Krmicek and Sebag, 2006) keeps diverse global (and good local) solutions. A related approach based on multiple populations has been proposed by Ulrich et al. (2010). The authors design the Diversity Integrating Optimizer (DIOP), which incorporates an archive population A for diversity preservation and a target population T for convergence. The offsprings are then generated out of their union $A \cup T$. The $P_{Q,\epsilon}$-MOEA by Schütze et al. (2011b) is a steady-state archive-based MOEA, which differs from the ϵ-MOEA (Deb et al., 2003) in two ways: the archiver now focuses on decision space diversity, and the implementation omits the population in favor of the archive.

Another idea is to integrate existing diversity preservation mechanisms into existing algorithms. Deb and Tiwari (2005, 2008) set the stage for this by extending NSGA-II and considering decision space diversity in the original crowding distance measure. In addition, they use ϵ-dominance as ranking mechanism to preserve diverse solutions. They find that this approach can also 'degenerate' to a competitive SO optimizer and hence call it Omni Optimizer. Other authors adapt this approach with slight modifications or extend some internal mechanisms. They apply them in the context of artificial immune systems (Coelho and Von Zuben, 2006), as well as in particle swarm-based approaches (Yue et al., 2018).

Strongly related to these diversity measures, niching methods from the SO domain are transferred and adapted to MOO. Shir et al. (2009) adapt an existing CMA-ES niching framework and apply it to the MO domain for boosting decision space diversity. The underlying idea is to aggregate diversity in decision and objective space in a rather naïve way — which works surprisingly well.

Zechman et al. (2013) designed the MO Niching Co-evolutionary Algorithm (MNCA), which identifies "distinct sets of non-dominated solutions which are maximally different in their decision vectors and are located in the same non-inferior regions of a Pareto front". Liang et al. (2016) deal with the problem of finding all globally optimal sets and define a so-called decision space based niching MOEA (DN-NSGAII). NIMMO (Tanabe and Ishibuchi, 2019) is a niching indicator-based multimodal multi- and many-objective optimization algorithm that has some commonalities with MOEA/D-AD (Tanabe and Ishibuchi, 2018) and TriMOEA-TA&R (Liu et al., 2018b). It uses a niching method, which is similar to the deterministic crowding method, but the environmental selection in NIMMO is based on the indicator values of the child and closest individuals. Liu et al. (2018a) introduce the Double-Niched EA (DNEA), which adopts a niche sharing method in objective *and* decision space.

Two very recent studies also focus on restricting diversity estimators to niches in decision space. Peng and Ishibuchi (2021) define a niche by the k closest solutions of a point in decision space, and it is shown based on incorporation into NSGA-II and SPEA2 that the loss of equivalent Pareto optimal solutions can be extremely diminished. NxEMMO by Javadi and Mostaghim (2021) builds on the same underlying idea. Essentially, it is a modified version of NSGA-II using a new environment selection method based on density-based indicators. Specifically, a procedure denoted as truncation incorporates either the nearest neighbor distances or the harmonic average distance (HAD) which is applied to the k nearest neighbors of each solution. This is shown to have potential to significantly increase decision space diversity.

Further niching-related approaches are pursued by Kramer and Danielsiek (2010) and Maree et al. (2019). Both propose the application of clustering approaches for diversity preservation. While the former apply DBSCAN as density-based method, the latter transfer the idea of hill-valley clustering from the SO domain. Their clusters cover roughly similar areas in decision space, which then were claimed to relate to globally and locally efficient sets. However, other visualization techniques (see Section 4) do not confirm the localness findings.

With two different and opposing perspectives, other authors address the challenge of diverse solution preservation or production: Ishibuchi et al. (2012) create a bi-objective solution set optimization problem by maximizing the decision space diversity and the objective space hypervolume. In solving this problem, they find multiple non-dominated sets in the sense of multiglobality. From another decomposition-related perspective, Rudolph et al. (2007), Rudolph and Preuss (2009) search for multiple globally efficient solution sets by using a multi-start approach that first clusters test solutions to gain adequate starting solutions for a SO optimizer that aims for one specific subset.

Tanabe and Ishibuchi (2018) propose a decomposition-based EA for multimodal MOO extending MOEA/D (Zhang and Li, 2007), named MOEA/D-AD, by allowing multiple similar solutions per sub-problem, if they are diverse in decision space. As a consequence (in order to consider archived individuals) the authors allow dynamic adaptation of the population size μ. This approach is related to an approach by Hu and Ishibuchi (2018) which identifies the same problems for MOEA/D but proposes to integrate decision space distance and neighborhood distances into selection. It should be noted that this also works the other way around, namely by *multiobjectivization* in order to detect very hard to find optima in the SO domain, e.g., by means of MOAMO (Preuss et al., 2015).

Finally, we should mention some special mechanisms for diversity preservation that stem from natural or self-organizational paradigms: Coelho and Von Zuben (2011) apply concentration-based immune networks for (multimodal) MOO inspired by an extension of a previous SO approach, whereas Liang et al. propose SMPSO_MM (Liang et al., 2018), a particle swarm approach that integrates a self-organized network to build a neighborhood relationship in decision space.

The extensive discussion of the literature on what we call "multi-global optimization" shows, that localness is usually not playing a central role in this area. As in original methods that aim at finding "some" part of the Pareto front, the perspective of localness is neglected. Recent work in this area, however, discovered localness as additional important aspect: The DNEA-L approach (Liu et al., 2019) integrates an additional archive to also store good locally efficient sets, which may contain interesting solutions for practical realization. This is in line with the classical argument in SO optimization to also keep near optimal local solutions (Preuss, 2015). Although localness is addressed here explicitly, there is no deeper analysis of local structures in decision space. The mechanisms of "multi-global optimization" are slightly extended to preserving good local solutions.

5.3.2. New perspectives on multimodality by exploiting localness

Instead of actively maintaining the population diversity in both decision and objective space, some recent algorithms approach the problem of multimodality and localness from a different perspective. This research direction is inspired by works that, in case of a continuous decision space, extend gradient-based SO optimization methods to the MO scenario (Fliege and Svaiter, 2000; Schütze et al., 2011a). Although these methods are (in analogy to the SO case) considered as local search methods, their application inside EMOAs, and also as stand-alone approaches, suggest that they can contribute to good solution approximations (Schütze et al., 2008). Consequently, a more detailed investigation of gradient-based local search in the context of MOO is promising. In contrast to the previous discussion of algorithmic approaches, here, the main challenge is to understand and exploit local search mechanisms in the context of multimodal and MO landscapes instead of discovering all optimal solution sets. This approach focuses on investigating, whether local efficient sets and their basins of attraction in MOO should be considered as ‘traps’ as in the SO case, or whether local search can possibly benefit from local structures and properties that are special to MOO.

In Emmerich and Deutz (2012), a set of non-dominated points has been found by using the gradient of its hypervolume indicator. In tradition of the SO understanding, it has been assumed that a naïvely implemented gradient ascending method can only act as local search algorithm and will possibly get stuck in locally efficient sets. Hence, Wang et al. (2017b) proposed an improved gradient-based algorithm, called Hypervolume Indicator Gradient Ascent MOO (HIGA-MO), to search for global and local optima of the hypervolume indicator simultaneously, thereby approximating the globally and locally efficient sets (in the sense of ϵ-connectedness; see Section 3) due to Pareto compliance property of the hypervolume indicator. HIGA-MO explicitly addresses the multimodality perspective by partitioning its search population into layers of locally non-dominated points (via non-dominated sorting (Srinivas and Deb, 1994)) and refining (pre-images of) each layer using the hypervolume indicator gradient, which is defined solely on this layer. Aiming to find as many locally efficient set as possible, HIGA-MO also tries to avoid wasting budget by restarting runs where it stagnates (in locally efficient sets). Already found locally efficient sets are archived.

A very recent algorithmic idea (Grimme et al., 2019b,a) is based on insights from the visualization of globally and locally efficient sets (as described in Section 4) across a variety of MOPs (see Fig. 6). A noticeable observation in the visualization of MOPs and a general characteristic of them is that the problem’s basins of attraction superpose each other. That is, basins of dominating locally efficient sets superpose basins of dominated locally efficient sets, leading to so-called ‘ridges’. These ridges are observed in decision space and indicate the boundary between basins of attraction for the MO gradient. Moving from a dominated basin across such a ridge and following the MO gradient leads into the superposing basin and from there towards the dominating efficient set. It is worthwhile to point out that moving from a dominated basin across the ridge is feasible because the locally efficient set in one basin is ‘touching’ a superior basin of attraction through the ridge. As illustrated in Fig. 2, a locally efficient set that touches its dominating basin of attraction is not strict according to Definition 8.

The Multi-Objective Gradient Sliding Algorithm (MOGSA) (Grimme et al., 2019b,a) utilizes and exploits these findings. From an initial point in decision space it follows the direction of the MO gradient towards the attracting locally efficient set. The local search phase terminates, when the (normalized) gradients of the considered objectives cancel each other out (see Fritz John condition in Section 3). Then, MOGSA moves along the efficient set in order to detect a ridge towards a superposing basin of attraction. When such a ridge is passed, MOGSA switches back to local search and descents to the next locally efficient set. Once MOGSA detects an efficient set without a cutting ridge, it has reached a strict locally efficient set (see Definition 8). Empirically in many cases, strict locally efficient sets are also globally efficient — at least for commonly used benchmarks. Note that MOGSA is technically a deterministic local search mechanism. Interestingly, first experiments of the authors show that this algorithm can be competitive or even superior to MOEAs on current multimodal MOPs (Grimme et al., 2019b).

To go beyond current research, we are convinced that the insights gained from visualization of multimodal MO landscapes can contribute to many innovations in algorithms for solving MOPs more efficiently. Like in the MOGSA approach, knowledge of the landscape may facilitate simple mechanisms to direct search for a certain (and better) region in decision space. Features of landscapes (as discussed in Section 5.2) may even contribute to an automatic detection of structural challenges in the landscape, and help in selecting solvers from a portfolio of algorithms.

We even noticed the potential to transfer knowledge gained on MO landscapes back to the SO case. As recently proposed in Steinhoff et al. (2020), Aspar et al. (2021), aspects of MOGSA may be utilized

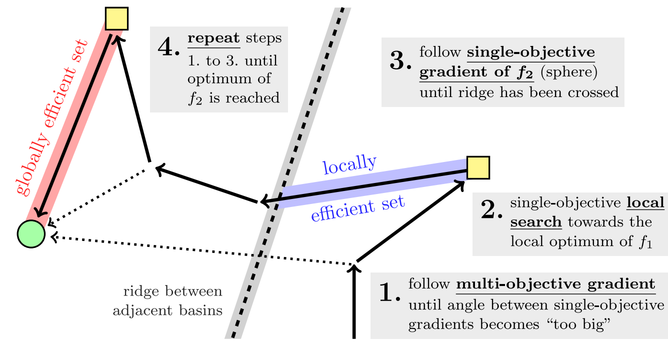

Fig. 7. Schema of (multiobjectivized) single-objective local search based on transferred MOGSA principles.

for local optimization in the SO domain. As schematically depicted in Fig. 7, a SO variant of MOGSA may be applied to any SO problem after the original problem with objective f_1 has been extended to a MOP by adding a simple but known second objective f_2 (e.g., the sphere function as proposed in Steinhoff et al., 2020). This *multiobjectivization* comes with virtually no costs for evaluating the additional objective f_2, as we can select f_2 such that it is unimodal, the gradient can be computed directly, and the global optimum is known. As conceptually shown in Fig. 7, the MO gradient descent of MOGSA could then be used to reach the vicinity of a (locally) efficient set of f_1 and f_2 (step 1). From the perimeter of this set, we can try to reach the most efficient solution w.r.t f_1 by single objective local search (step 2). Thereafter, we can move along the (locally) efficient set towards f_2 (step 3), and once the ridge to the superposing (MO) basin of attraction has been passed, MOGSA is started again to reach the next (local) efficient set (step 1). This loop can be repeated (in a first naïve approach), until the (known) optimum of f_2 is reached.

This conceptual proposal exemplarily demonstrates the high potential of new insights into MO landscapes and thus the wider consideration of MO multimodality including local structures. We are able to propose rather simple techniques that can bypass problems of SO multimodality by using MO landscape properties like local efficient sets and ridges between basins of attraction. Certainly, in the hope to find better solutions faster, the extensions of SO problems to MOPs has been proposed multiple times before, e.g., refer to Brockhoff et al. (2007), Tran et al. (2013), Garza-Fabre et al. (2015) and Segura et al. (2016). However, these approaches often assumed advantages in multiobjectivization but had no conclusive insight into the MO landscape. As a consequence, the step of multiobjectivization was merely complemented by the application of standard MOEAs — not by dedicated algorithmic concepts that exploited local structures and landscape properties in a sophisticated and direct way.

We are convinced, that landscape insights open up future algorithmic research perspectives – also in this area – and enable the community to provide enhancements that can advance MO and SO problem solving capabilities.

5.4. Performance assessment

Performance assessment of multimodal MO optimizers highly depends on the taken perspective (see Section 1). If multimodality is only perceived as hindrance to be overcome in order to approximate the Pareto front as fast and as accurately as possible, one ends up with the classical unary performance indicators (see below) that are widely used in MOO such as the Dominated Hypervolume Indicator (HV), (Inverted) Generational Distance, Δ_p, etc. (Zitzler and Thiele, 1998; Zitzler et al., 2003).

Definition 12 (*Performance Indicator and Pareto Compliance*)**.** A unary *performance indicator* $I : \mathcal{P}(\mathbb{R}^m) \to \mathbb{R}$ is a function on the power set $\mathcal{P}(\mathbb{R}^m)$ of $\mathbb{R}^m$, which w.l.o.g. is to be maximized. Further, the performance indicator I is called *Pareto compliant*, if and only if the following implication holds: $\forall A, B \in \mathcal{P}(\mathbb{R}^m)$ with $A \preceq B$ and $B \npreceq A \Rightarrow I(A) > I(B)$. Similarly, an indicator I is called *weakly Pareto compliant*, if the relation $I(A) > I(B)$ is relaxed to $I(A) \geq I(B)$.

Yet, it should be noted that when maximizing a weakly Pareto compliant indicator (see above) with a finite approximation set, the pre-image of a maximum might not even be locally efficient at all since it might occur that $I(A) = I(B)$ and $A \prec B$ while both A and B maximize I (locally).

However, in algorithms such as Omni optimizer and the Niching-based SMS-EMOA by Shir et al. (2009) decision space diversity is considered as a secondary quality indicator via crowding distance or an indicator based on pairwise distances between points. Strictly speaking, such algorithms use mechanisms to increase diversity in the decision space but do not explicitly measure the performance of sets. Most algorithmic approaches discussed in Section 5.3 utilize IGD or IGDX (Zhou et al., 2009). In contrast to IGD, which measures both convergence and diversity in *objective* space, the latter measures how well the Pareto optimal set is approximated in the *decision* space. Also, Pareto set proximity (PSP) (Yue et al., 2018) and the ratio of nondominated individuals (RNI) (Kim et al., 2004) are frequently used.

Recently, the discussion of diversity and coverage indicators for set-oriented optimization has yielded many new insights and potentially useful indicators. Although these indicators have not all been considered in MO multimodal optimization they offer interesting options. For instance, the Riesz s-energy (Falcón-Cardona et al., 2019) is the sum of inverse distances of the distance matrix $d_{ij}, i = 1, \ldots, \mu$, where μ is the size of the set of which the diversity has to be measured. It is computed as $\sum_{i=1} \sum_{j>i} 1/d_{ij}^q$ where q is an integer constant that is proportional to the dimension of the embedding decision space. The Riesz s-energy has the property that its minimization distributes points uniformly across manifolds and its computation is scalable in terms of number of decision variables.

Indicators designed for level set approximation have been discussed in Emmerich et al. (2013). Level set approximation is the task of approximating a level set $\{x \in X : f(x) \leq \tau\}$ of a function $f : X \to \mathbb{R}$. In this context decision space diversity is of paramount importance, but Emmerich et al. (2013) argued that it needs to be distinguished from representativeness (or set coverage). Diverse approximation sets maximize gaps between points, and diversity indicators have the tendency of distributing points at the boundary of a feasible domain, whereas coverage or representedness metrics seek to have a close representative in the approximation set for every point in the covered set. As a consequence, sets that obtain optimum values have more points in the interior of the level sets as compared to diversity metrics. In this context diversity indicators (Solow Polaski, gap metrics (cf. Wessing and Preuss, 2016), Weitzman diversity) have been contrasted to coverage metrics such as a version of the average Hausdorff distance (Schütze et al., 2012).

The challenge of finding all optima, however, requires sophisticated and specifically designed performance indicators that simultaneously take (i) the extent of coverage of all local Pareto fronts (in objective space), and (ii) an assessment of diversity and coverage in decision space into account. Ideally, such an indicator shall be Pareto compliant – which still has to be explicitly defined when simultaneously considering all local fronts – and also offer a possibility for setting the weights of components, i.e., including experts' preferences or increasing importance w.r.t. the rank of the respective local front. Moreover, it shall explicitly take decision space coverage or diversity into account.

So far, however, such an indicator does not exist. In fact, it also does not exist for multimodal SO optimization, despite decades of research in this direction. The reason for this is presumably that it is not at all trivial to define what "all" optima means in this context. In practice, benchmarks like the SO niching competition problem set (Li et al., 2013), which started in the early 2010's and still is in use as of 2020, circumvent this problem by focusing on global optima only. Preuss (2015) also discusses this issue extensively and suggests that the indicator problem in the multimodal case cannot be boiled down to one measure. Instead, the fit of a measure depends on the use case, and the definition of a suitable compromise between diversity and quality depends on the circumstances.

6. Summary and conclusions

We have found a wide variety of works that deal with multimodal MO continuous optimization. However, the majority of these works followed a perspective that we have summarized as "finding all optima". Therein, the goal is essentially to find all global (and eventually local) optima. In contrast to this line of research, a second stream ("not getting trapped") has emerged more recently, which explicitly focuses on understanding the characteristics of (multimodal) MO optimization problems. This second stream enables the integration of problem-specific knowledge into various methods and algorithms, which in turn have a great potential to improve the state of the art in MO optimization. Beyond the scope of this work, but definitely important for future work in the continuous domain – particularly with respect to landscape analysis and local search methods – is the incorporation of existing insights and concepts from the multi-objective (and often inherently multimodal) discrete and combinatorial domains (e.g., Basseur et al., 2013; Paquete et al., 2007b; Lust and Teghem, 2010; Verel et al., 2011; Liefooghe et al., 2018b).

Benchmarks On the basis of the insights gained from sophisticated visualization techniques such as gradient-based heatmaps and algorithmic approaches that efficiently exploit multimodality (e.g., MOGSA), we can conclude that almost none of the existing benchmarks contain MO landscapes which pose severe difficulties in terms of traps which result in algorithms getting stuck locally. Therefore, there is an urgent need for a comprehensive and practically relevant benchmark set which comprises a large and diverse set of MO test problems featuring different kinds of multimodal structures. Apart from designing (ideally feature-based) generators for artificial multimodal MO test problems, also real-world problems should be thoroughly inspected and analyzed for multimodal structures. In this regard, the research expertise bundled by the Benchmarking Network[5] has huge potential.

Landscape analysis and visualization However, in order to construct such a benchmark, various fundamental work is required, especially w.r.t. problem understanding, visualization and feature design. Capturing multimodality in terms of specifically designed exploratory landscape features requires a deep understanding of multimodal structures and landscape characteristics. This improved problem understanding can be achieved, for example, with the help of suitable visualization techniques. Therefore, a crucial prerequisite for feature design would be an extension of visualization methods in terms of scalability with objective, as well as the decision space. Here, the challenge of not loosing too much information along with the involved dimensionality reduction techniques has to be faced. Moreover, the sensitivity w.r.t. the underlying grid resolutions has to be investigated with the aim of becoming as independent as possible from the resolution level. This should include the option to adapt the latter locally, if desired, in order to explicitly focus on specific areas of interest in the search and/or objective space.

Building on the knowledge gained from the visual inspection of various kinds of multimodal structures and challenges, specific landscape features can be designed with the long-term goal of serving as

[5] https://sites.google.com/view/benchmarking-network/.

efficient input for automated algorithm selection and configuration models (Kerschke et al., 2019a). This includes investigating the potential for generalization of the underlying ideas of features typically used in single-objective optimization to capture a MOP's structural high-level properties. The computation of single-objective features on a per-objective-basis is certainly not sufficient, as interaction effects of the objectives lead to specific (possibly also multimodal) structures in the MO space. In the MO context, synergy effects can also occur through the transfer of concepts from combinatorial optimization and through the additional consideration of so-called probing features, which characterize algorithm behavior along initial stages of the optimization runs. Also, archiving strategies within MO metaheuristics and insights gained from analyzing paths of local search strategies might be used for the same purpose.

Algorithm design Similarly, the development of new, as well as the enhancement of existing algorithms will substantially benefit from landscape insights, informative features and comprehensive benchmark sets. With reference to the two different perspectives adopted along this paper, multiple research angles can be pursued. To overcome local traps, algorithm design can rely on profound knowledge of local structures (incl. optima) and will ideally exploit ridges and gradient paths. Appropriate local search approaches – such as MOGSA and/or improved variants – can be hybridized perspectively with state-of-the art MO meta-heuristics and niching techniques. Of course, in-depth knowledge of the landscape's characteristics and local structures enables specially designed techniques to approach all (or a desired subset of) local and global optima simultaneously. Most probably, multi-objective approaches will have to be focused first, which will pave the way to specifically designed techniques for many-objective optimization problems.

Knowledge transfer to SO Interestingly, the single-objective optimization domain might also substantially benefit from high-performing MO optimization algorithms which efficiently exploit multimodal structures. Multiobjectivization of single-objective problems, i.e., adding well-suited additional (artificial) objectives and afterwards solving the MO (multimodal) problem, could potentially outperform state-of-the art SO optimizers. Initial experiments based on the MOGSA variant SOMOGSA already showed very promising results (Steinhoff et al., 2020; Aspar et al., 2021). Also, a hybridization with state-of-the art SO optimizers in order to overcome stochasticity induced by the initial search point should be a matter of future research. However, there is a lack of theory on how to choose the most appropriate additional objective function(s) for the SO problem at hand, in order to make the multimodal MOP as easy to handle as possible. The resulting Pareto front of course comprises the single-objective optimum as well.

Algorithm performance assessment A crucial remaining open question of current research, which has a significant impact on the design of algorithms, is the appropriate assessment of optimization quality w.r.t. multimodality and thus multiple optima. Neither exists a commonly accepted performance indicator, nor have the theoretical properties of the existing ones been investigated satisfactorily. Of course, the optimal choice of the indicator heavily depends on the taken perspective, i.e., the "find them all" case demands for other indicator properties than the "do not get trapped" perspective. The latter stream basically leads to common MO performance indicators such as the Dominated Hypervolume, while the former essentially represents an unexplored field of research. In any case, Pareto dominance and the possibility to intuitively integrate the preferences of decision makers are most desirable properties which poses substantial challenges to the design of suitable performance indicators. Meeting all these different aspects of performance assessment in this domain might even lead to performance indicators which themselves are multi-objective in nature providing trade-offs between different performance criteria (Bossek et al., 2020).

Theoretical analyses All these practice-oriented developments are closely linked to progress in the theory-driven part of this research field. In the latter, a potential objective for future work is a comparison and possible unification of the different definitions of locally efficient sets between the domains of continuous and combinatorial optimization – e.g., the definitions given in this survey and the ones given in Liefooghe et al. (2018c). If we find a common terminology, we might be able to join and exploit synergies between the two areas and ultimately foster research in both domains. Another interesting line of research at the intersection of theory and practice are investigations of the choice of a suitable, problem-specific value ε (for the ε-neighborhood). Progress in this area is of great importance as it links the MO problem's (true) locally efficient sets and the ε-local efficient sets that were approximated by the MO algorithm(s). Noticeably, (the numbers of) these locally efficient sets can differ significantly, depending on the choice of ε.

Scalability As obvious from the discussion of the current state of multimodal multi-objective optimization, detailed insight into the structures and challenges of (multimodal) multi-objective landscapes is still at an early stage, of course offering very promising perspectives due to recent research developments. However, objective and decision space dimensionality are assumed to be manageable, i.e., visualization as an important ingredient reaches its natural limits at 3D. Thus, multimodal *many-objective* optimization as a very important research field, either in terms of large number of objectives or decision variables (or even both) will have to remain a subsequent step after substantial progress on the classical multi-objective scenarios. In our view, however, the ability of MO algorithms to efficiently detect and exploit localness will even have a larger impact on algorithm performance improvement in many-objective scenarios.

CRediT authorship contribution statement

Christian Grimme: Conceptualization, Project administration, Formal analysis, Writing – original draft, Writing – review & editing. **Pascal Kerschke:** Conceptualization, Formal analysis, Visualization, Writing – original draft, Writing – review & editing, Validation. **Pelin Aspar:** Conceptualization, Formal analysis, Validation. **Heike Trautmann:** Conceptualization, Writing – original draft, Writing – review & editing, Validation. **Mike Preuss:** Conceptualization, Writing – original draft, Validation. **André H. Deutz:** Conceptualization, Formal analysis, Writing – original draft, Writing – review, Validation. **Hao Wang:** Conceptualization, Formal analysis, Writing – original draft, Writing – review, Validation. **Michael Emmerich:** Conceptualization, Formal analysis, Writing – original draft, Validation.

Acknowledgments

The authors acknowledge support by the *European Research Center for Information Systems (ERCIS)*. In addition, Hao Wang acknowledges support from the Paris Île-de-France Region.

References

Aspar, P., Kerschke, P., Steinhoff, V., Trautmann, H., Grimme, C., 2021. Multi[3]: Optimizing multimodal single-objective continuous problems in the multi-objective space by means of multiobjectivization. In: Proceedings of the 11th International Conference on Evolutionary Multi-Criterion Optimization (EMO). pp. 311–322.

Bäck, T.H.W., Fogel, D.B., Michalewicz, Z., 1997. Handbook of Evolutionary Computation. CRC Press.

Basseur, M., Goëffon, A., Liefooghe, A., Verel, S., 2013. On set-based local search for multiobjective combinatorial optimization. In: Proceeding of the Fifteenth Annual Conference on Genetic and Evolutionary Computation (GECCO). ACM Press, pp. 471–478.

Belkhir, N., Dréo, J., Savéant, P., Schoenauer, M., 2017. Per instance algorithm configuration of CMA-ES with limited budget. In: Proceedings of the 19th Annual Conference on Genetic and Evolutionary Computation (GECCO). ACM, pp. 681–688.

Beume, N., Naujoks, B., Emmerich, M.T.M., 2007. SMS-EMOA: Multiobjective selection based on dominated hypervolume. Eur. J. Oper. Res. (EJOR) 181 (3), 1653–1669.
Beyer, H.-G., 2001. The Theory of Evolution Strategies. Springer.
Blot, A., Hoos, H.H., Jourdan, L., Kessaci-Marmion, M.-E., Trautmann, H., 2016. MO-ParamILS: A multi-objective automatic algorithm configuration framework. In: Proceedings of the 10th International Conference on Learning and Intelligent Optimization (LION). Springer, pp. 32–47.
Bossek, J., 2017. Smoof: Single- and multi-objective optimization test functions. R J..
Bossek, J., Kerschke, P., Trautmann, H., 2020. A multi-objective perspective on performance assessment and automated selection of single-objective optimization algorithms. Appl. Soft Comput. 2020 (88), 105901.
Brockhoff, D., Friedrich, T., Hebbinghaus, N., Klein, C., Neumann, F., Zitzler, E., 2007. Do additional objectives make a problem harder? In: Proceedings of the 9th Annual Conference on Genetic and Evolutionary Computation (GECCO). pp. 765–772.
Cauchy, A.-L., 1847. Méthode générale pour la résolution des systemes d'équations simultanées. C. R. Hebd. Séances Acad. Sci. 25 (1), 536–538.
Cheng, R., Jin, Y., Olhofer, M., Sendhoff, B., 2016. Test problems for large-scale multiobjective and many-objective optimization. IEEE Trans. Cybern. (TCYB) 47 (12), 4108–4121.
Cheng, R., Li, M., Tian, Y., Zhang, X., Yang, S., Jin, Y., Yao, X., 2017. A benchmark test suite for evolutionary many-objective optimization. Complex Intell. Syst. 3 (1), 67–81.
Coelho, G.P., Von Zuben, F.J., 2006. Omni-ainet: An immune-inspired approach for omni optimization. In: Artificial Immune Systems. Springer, pp. 294–308.
Coelho, G.P., Von Zuben, F.J., 2011. A concentration-based artificial immune network for multi-objective optimization. In: Proceedings of the 6th International Conference on Evolutionary Multi-Criterion Optimization (EMO). Springer, pp. 343–357.
Coello Coello, C.A., González Brambila, S., Figueroa Gamboa, J., Castillo Tapia, M., Hernández Gómez, R., 2019. Evolutionary multiobjective optimization: Open research areas and some challenges lying ahead. Complex Intell. Syst. 1–16.
Coello Coello, C.A., van Veldhuizen, D.A., Lamont, G.B., 2007. Evolutionary Algorithms for Solving Multi-Objective Problems, second ed. Springer.
Custódio, A.L., Madeira, J.F.A., 2018. MultiGLODS: Global and local multiobjective optimization using direct search. J. Glob. Optim. 72 (2), 323–345.
Deb, K., 1999. Multi-objective genetic algorithms: Problem difficulties and construction of test problems. Evol. Comput. (ECJ) 7 (3), 205–230.
Deb, K., 2001. Multi-Objective Optimization using Evolutionary Algorithms. Wiley, Chichester.
Deb, K., Mohan, M., Mishra, S., 2003. Towards a quick computation of well-spread Pareto-optimal solutions. In: Proceedings of the 2nd International Conference on Evolutionary Multi-Criterion Optimization (EMO). Springer, pp. 222–236.
Deb, K., Pratap, A., Agarwal, S., Meyarivan, T., 2002. A fast and elitist multiobjective genetic algorithm: NSGA-II. IEEE Trans. Evol. Comput. (TEVC) 6 (2), 182–197.
Deb, K., Thiele, L., Laumanns, M., Zitzler, E., 2005. Scalable test problems for evolutionary multiobjective optimization. In: Evolutionary Multiobjective Optimization. Springer, pp. 105–145.
Deb, K., Tiwari, S., 2005. Omni-optimizer: A procedure for single and multi-objective optimization. In: Proceedings of the 9th International Conference on Evolutionary Multi-Criterion Optimization (EMO). Springer, pp. 47–61.
Deb, K., Tiwari, S., 2008. Omni-optimizer: A generic evolutionary algorithm for single and multi-objective optimization. Eur. J. Oper. Res. (EJOR) 185, 1062–1087.
Eggensperger, K., Lindauer, M., Hutter, F., 2019. Pitfalls and best practices in algorithm configuration. J. Artif. Intell. Res. (JAIR) 64, 861–893.
Emmerich, M.T.M., Beume, N., Naujoks, B., 2005. An EMO algorithm using the hypervolume measure as selection criterion. In: Proceedings of the 3rd International Conference on Evolutionary Multi-Criterion Optimization (EMO). Springer, pp. 62–76.
Emmerich, M.T.M., Deutz, A.H., 2007. Test problems based on Lamé superspheres. In: Proceedings of the 4th International Conference on Evolutionary Multi-Criterion Optimization (EMO). Springer, pp. 922–936.
Emmerich, M.T.M., Deutz, A.H., 2012. Time complexity and zeros of the hypervolume indicator gradient field. In: EVOLVE–A Bridge between Probability, Set Oriented Numerics, and Evolutionary Computation III. Springer, pp. 169–193.
Emmerich, M.T.M., Deutz, A.H., 2018. A tutorial on multiobjective optimization: Fundamentals and evolutionary methods. Nat. Comput. 17 (3), 585—609.
Emmerich, M.T.M., Deutz, A.H., Kruisselbrink, J.W., 2013. On quality indicators for black-box level set approximation. In: EVOLVE — A Bridge between Probability, Set Oriented Numerics, and Evolutionary Computation. Springer, pp. 157–185.
Falcón-Cardona, J.G., Coello Coello, C.A., Emmerich, M.T.M., 2019. CRI-EMOA: A Pareto-front shape invariant evolutionary multi-objective algorithm. In: Proceedings of the 10th International Conference on Evolutionary Multi-Criterion Optimization (EMO). Springer, pp. 307–318.
Fieldsend, J.E., Alyahya, K., 2019. Visualising the landscape of multi-objective problems using local optima networks. In: Proceedings of the 21st Annual Conference on Genetic and Evolutionary Computation (GECCO). pp. 1421–1429.
Fieldsend, J.E., Chugh, T., Allmendinger, R., Miettinen, K., 2019. A feature rich distance-based many-objective visualisable test problem generator. In: Proceedings of the 21st Annual Conference on Genetic and Evolutionary Computation (GECCO). ACM, pp. 541–549.
Fliege, J., Svaiter, B.F., 2000. Steepest descent methods for multicriteria optimization. Math. Methods Oper. Res. 51 (3), 479–494.
Fonseca, C.M.M.d., 1995. Multiobjective Genetic Algorithms with Application to Control Engineering Problems (Ph.D. thesis). Department of Automatic Control & Systems Engineering, University of Sheffield.
Garza-Fabre, M., Toscano-Pulido, G., Rodriguez-Tello, E., 2015. Multi-objectivization, fitness landscape transformation and search performance: A case of study on the HP model for protein structure prediction. Eur. J. Oper. Res. (EJOR) 243 (2), 405–422.
Glasmachers, T., 2019. Challenges of convex quadratic bi-objective benchmark problems. In: Proceedings of the 21st Annual Conference on Genetic and Evolutionary Computation (GECCO). ACM, pp. 559–567.
Grimme, C., Kerschke, P., Emmerich, M.T.M., Preuss, M., Deutz, A.H., Trautmann, H., 2019a. Sliding to the global optimum: How to benefit from non-global optima in multimodal multi-objective optimization. In: AIP Conference Proceedings. AIP Publishing, pp. 020052–1–020052–4.
Grimme, C., Kerschke, P., Trautmann, H., 2019b. Multimodality in multi-objective optimization — More boon than bane? In: Evolutionary Multi-Criterion Optimization. Springer, pp. 126–138.
Hansen, N., Finck, S., Ros, R., Auger, A., 2009. Real-Parameter Black-Box Optimization Benchmarking 2009: Noiseless Functions Definitions. Tech. Rep. RR-6829, INRIA.
Hillermeier, C., 2001. Generalized homotopy approach to multiobjective optimization. J. Optim. Theory Appl. 110 (3), 557–583.
Hiroyasu, T., Nakayama, S., Miki, M., 2005. Comparison study of SPEA2+, SPEA2, and NSGA-II in diesel engine emissions and fuel economy problem. In: Proceedings of the IEEE Congress on Evolutionary Computation (CEC), Vol. 1. pp. 236–242.
Hu, C., Ishibuchi, H., 2018. Incorporation of a decision space diversity maintenance mechanism into MOEA/D for multi-modal multi-objective optimization. In: Proceedings of the 20th Annual Conference on Genetic and Evolutionary Computation (GECCO). ACM, pp. 1898–1901.
Huband, S., Hingston, P., Barone, L., While, L., 2006. A review of multiobjective test problems and a scalable test problem toolkit. IEEE Trans. Evol. Comput. (TEVC) 10 (5), 477–506.
Ishibuchi, H., Peng, Y., Shang, K., 2019. A scalable multimodal multiobjective test problem. In: Proceedings of the IEEE Congress on Evolutionary Computation (CEC). IEEE, pp. 310–317.
Ishibuchi, H., Yamane, M., Akedo, N., Nojima, Y., 2012. Two-objective solution set optimization to maximize hypervolume and decision space diversity in multiobjective optimization. In: Proceedings of the 6th International Conference on Soft Computing and Intelligent Systems (SCIS), and the 13th International Symposium on Advanced Intelligence Systems (ISIS). pp. 1871–1876.
Javadi, M., Mostaghim, S., 2021. Using neighborhood-based density measures for multimodal multi-objective optimization. In: Proceedings of the 11th International Conference on Evolutionary Multi-Criterion Optimization (EMO). Springer, pp. 335–345.
Jiménez, B., 2002. Strict efficiency in vector optimization. J. Math. Anal. Appl. 265 (2), 264–284.
John, F., 2014. Extremum problems with inequalities as subsidiary conditions. In: Traces and Emergence of Nonlinear Programming. Springer, pp. 197–215.
Kerschke, P., Grimme, C., 2017. An expedition to multimodal multi-objective optimization landscapes. In: Proceedings of the 9th International Conference on Evolutionary Multi-Criterion Optim. (EMO). Springer, pp. 329–343.
Kerschke, P., Hoos, H.H., Neumann, F., Trautmann, H., 2019a. Automated algorithm selection: Survey and perspectives. Evol. Comput. (ECJ) 27, 3–45.
Kerschke, P., Preuss, M., Hernández, C., Schütze, O., Sun, J.-Q., Grimme, C., Rudolph, G., Bischl, B., Trautmann, H., 2014. Cell mapping techniques for exploratory landscape analysis. In: EVOLVE — A Bridge between Probability, Set Oriented Numerics, and Evolutionary Computation V. Springer, pp. 115–131.
Kerschke, P., Trautmann, H., 2016. The R-package FLACCO for exploratory landscape analysis with applications to multi-objective optimization problems. In: Proceedings of the IEEE Congress on Evolutionary Computation (CEC). IEEE, pp. 5262–5269.
Kerschke, P., Trautmann, H., 2019a. Automated algorithm selection on continuous black-box problems by combining exploratory landscape analysis and machine learning. Evol. Comput. (ECJ) 27, 99–127.
Kerschke, P., Trautmann, H., 2019b. Comprehensive feature-based landscape analysis of continuous and constrained optimization problems using the R-package flacco. In: Applications in Statistical Computing. Springer, pp. 93–123.
Kerschke, P., Wang, H., Preuss, M., Grimme, C., Deutz, A.H., Trautmann, H., Emmerich, M.T.M., 2016. Towards analyzing multimodality of multiobjective landscapes. In: Proceedings of the 14th International Conference on Parallel Problem Solving from Nature (PPSN XIV). Springer, pp. 962–972.
Kerschke, P., Wang, H., Preuss, M., Grimme, C., Deutz, A.H., Trautmann, H., Emmerich, M.T.M., 2019b. Search dynamics on multimodal multi-objective problems. Evol. Comput. (ECJ) 27, 577–609.
Kim, M., Hiroyasu, T., Miki, M., Watanabe, S., 2004. SPEA2+: Improving the performance of the strength Pareto evolutionary algorithm 2. In: Proceedings of the 8th International Conference on Parallel Problem Solving from Nature (PPSN VIII). Springer, pp. 742–751.
Kramer, O., Danielsiek, H., 2010. DBSCAN-based multi-objective niching to approximate equivalent Pareto-subsets. In: Proceedings of the 12th Annual Conference on Genetic and Evolutionary Computation (GECCO). ACM, pp. 503–510.

Krmicek, V., Sebag, M., 2006. Functional brain imaging with multi-objective multi-modal evolutionary optimization. In: Proceedings of the 9th International Conference on Parallel Problem Solving from Nature (PPSN IX). Springer, pp. 382–391.

Kursawe, F., 1990. A variant of evolution strategies for vector optimization. In: Proceedings of the 1st International Conference on Parallel Problem Solving from Nature (PPSN I). Springer, pp. 193–197.

Li, X., Engelbrecht, A.P., Epitropakis, M.G., 2013. Benchmark Functions for CEC'2013 Special Session and Competition on Niching Methods for Multimodal Function Optimization. Tech. rep., Evolutionary Computation and Machine Learning Group, RMIT Univ., Australia, pp. 1–10.

Liang, J., Guo, Q., Yue, C., Qu, B., Yu, K., 2018. A self-organizing multi-objective particle swarm optimization algorithm for multimodal multi-objective problems. In: International Conference on Swarm Intelligence (ICSI). Springer, pp. 550–560.

Liang, J., Qu, B., Gong, D., Yue, C., Problem Definitions and Evaluation Criteria for the CEC 2019 Special Session on Multimodal Multiobjective Optimization.

Liang, J., Yue, C., Qu, B., 2016. Multimodal multi-objective optimization: A preliminary study. In: Proceedings of the IEEE Congress on Evolutionary Computation (CEC). pp. 2454–2461.

Liefooghe, A., Derbel, B., Verel, S., López-Ibáñez, M., Aguirre, H.E., Tanaka, K., 2018a. On Pareto local optimal solutions networks. In: Proceedings of the 15th International Conference on Parallel Problem Solving from Nature (PPSN XV). Springer, pp. 232–244.

Liefooghe, A., Derbel, B., Verel, S., López-Ibáñez, M., Aguirre, H., Tanaka, K., 2018b. On Pareto local optimal solutions networks. In: Auger, A., Fonseca, C.M., Lourenço, N., Machado, P., Paquete, L., Whitley, D. (Eds.), Parallel Problem Solving from Nature (PPSN XV), Vol. 11102. Springer, pp. 232–244.

Liefooghe, A., López-Ibáñez, M., Paquete, L., Verel, S., 2018c. 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. ACM, pp. 324–331.

Liu, Y., Ishibuchi, H., Nojima, Y., Masuyama, N., Han, Y., 2019. Searching for local Pareto optimal solutions: A case study on polygon-based problems. In: 2019 IEEE Congress on Evolutionary Computation (CEC). IEEE, pp. 896–903.

Liu, Y., Ishibuchi, H., Nojima, Y., Masuyama, N., Shang, K., 2018a. A double-niched evolutionary algorithm and its behavior on polygon-based problems. In: Proceedings of the 15th International Conference on Parallel Problem Solving from Nature (PPSN XV). Springer, pp. 262–273.

Liu, Y., Yen, G., Gong, D., 2018b. A multi-modal multi-objective evolutionary algorithm using two-archive and recombination strategies. IEEE Trans. Evol. Comput. (TEVC) 23 (4), 660–674.

López-Ibáñez, M., Dubois-Lacoste, J., Pérez Cáceres, L., Birattari, M., Stützle, T., 2016. The irace package: Iterated racing for automatic algorithm configuration. Oper. Res. Perspect. 3, 43–58.

Lunacek, M., Whitley, L.D., 2006. The dispersion metric and the CMA evolution strategy. In: Proceedings of the 8th Annual Conference on Genetic and Evolutionary Computation (GECCO). ACM, pp. 477–484.

Lust, T., Teghem, J., 2010. Two-phase Pareto local search for the biobjective traveling salesman problem. J. Heuristics 16 (3), 475–510.

Malan, K.M., Engelbrecht, A.P., 2009. Quantifying ruggedness of continuous landscapes using entropy. In: Proceedings of the IEEE Congress on Evolutionary Computation (CEC). IEEE, pp. 1440–1447.

Malan, K.M., Engelbrecht, A.P., 2013. A survey of techniques for characterising fitness landscapes and some possible ways forward. Inf. Sci. (JIS) 241, 148–163.

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

Mersmann, O., Bischl, B., Trautmann, H., Preuss, M., Weihs, C., Rudolph, G., 2011. Exploratory landscape analysis. In: Proceedings of the 13th Annual Conference on Genetic and Evolutionary Computation (GECCO). ACM, pp. 829–836.

Miettinen, K., 1998. Nonlinear Multiobjective Optimization. In: International Series in Oper. Res. & Management Science, vol. 12, Springer.

Muñoz Acosta, M.A., Sun, Y., Kirley, M., Halgamuge, S.K., 2015. Algorithm selection for black-box continuous optimization problems: A survey on methods and challenges. Inf. Sci. (JIS) 317, 224–245.

Paquete, L., Chiarandini, M., Stützle, T., 2004. Pareto local optimum sets in the biobjective traveling salesman problem: An experimental study. In: Gandibleux, X., Sevaux, M., Sörensen, K., T'kindt, V. (Eds.), Metaheuristics for Multiobjective Optimisation. Springer, pp. 177–199.

Paquete, L., Schiavinotto, T., Stützle, T., 2007a. On local optima in multiobjective combinatorial optimization problems. Ann. Oper. Res. 156, 83–97.

Paquete, L., Schiavinotto, T., Stützle, T., 2007b. On local optima in multiobjective combinatorial optimization problems. Ann. Oper. Res. 156 (1), 83–97.

Peng, Y., Ishibuchi, H., 2021. Niching diversity estimation for multi-modal multi-objective optimization. In: Proceedings of the 11th International Conference on Evolutionary Multi-Criterion Optimization (EMO). Springer, pp. 323–334.

Peng, Y., Ishibuchi, H., Shang, K., 2019. Multi-modal multi-objective optimization: Problem analysis and case studies. In: Proceedings of the IEEE Symposium Series on Computational Intelligence (SSCI). IEEE, pp. 1865–1872.

Preuss, M., 2015. Multimodal Optimization by Means of Evolutionary Algorithms. Springer.

Preuss, M., Wessing, S., Rudolph, G., Sadowski, G., 2015. Solving phase equilibrium problems by means of avoidance-based multiobjectivization. In: Kacprzyk, J., Pedrycz, W. (Eds.), Springer Handbook of Computational Intelligence. Springer, pp. 1159–1171.

Reshef, D.N., Reshef, Y.A., Finucane, H.K., Grossman, S.R., McVean, G., Turnbaugh, P.J., Lander, E.S., Mitzenmacher, M., Sabeti, P.C., 2011. Detecting novel associations in large data sets. Science 334 (6062), 1518–1524.

Rudolph, G., Naujoks, B., Preuss, M., 2007. Capabilities of EMOA to detect and preserve equivalent Pareto subsets. In: Proceedings of the 4th International Conference on Evolutionary Multi-Criterion Optimization (EMO). Springer, pp. 36–50.

Rudolph, G., Preuss, M., 2009. A multiobjective approach for finding equivalent inverse images of Pareto-optimal objective vectors. In: 2009 IEEE Symposium on Computational Intelligence in Multi-Criteria Decision-Making (MCDM). IEEE, pp. 74–79.

Saxena, D.K., Zhang, Q., Duro, J.A., Tiwari, A., 2011. Framework for many-objective test problems with both simple and complicated Pareto-set shapes. In: Proceedings of the 6th International Conference on Evolutionary Multi-Criterion Optimization (EMO). Springer, pp. 197–211.

Schaffer, J.D., 1985. Multiple objective optimization with vector evaluated genetic algorithms. In: Proceedings of the 1st International Conference on Genetic Algorithms. L. Erlbaum Associates Inc., pp. 93–100.

Schäpermeier, L., Grimme, C., Kerschke, P., 2020. One PLOT to show them all: Visualization of efficient sets in multi-objective landscapes. In: Proceedings of the 16th International Conference on Parallel Problem Solving from Nature (PPSN XVI). Springer, pp. 154–167.

Schäpermeier, L., Grimme, C., Kerschke, P., 2021. To boldly show what no one has seen before: A dashboard for visualizing multi-objective landscapes. In: Proceedings of the 11th International Conference on Evolutionary Multi-Criterion Optimization (EMO). pp. 632–644.

Schütze, O., Esquivel, X., Lara, A., Coello Coello, C.A., 2012. Using the averaged hausdorff distance as a performance measure in evolutionary multiobjective optimization. IEEE Trans. Evol. Comput. (TEVC) 16 (4), 504–522.

Schütze, O., Lara, A., Coello Coello, C.A., 2011a. The directed search method for unconstrained multi-objective optimization problems. In: Proceedings of the EVOLVE–A Bridge between Probability, Set Oriented Numerics, and Evolutionary Computation. pp. 1–24.

Schütze, O., Sanchez, G., Coello Coello, C.A., 2008. A new memetic strategy for the numerical treatment of multi-objective optimization problems. In: Proceedings of the 10th Annual Conference on Genetic and Evolutionary Computation (GECCO). ACM, pp. 705–712.

Schütze, O., Vasile, M., Coello Coello, C.A., 2011b. Computing the set of epsilon-efficient solutions in multiobjective space mission design. J. Aerosp. Comput. Inf. Commun. 8, 53–70.

Schwefel, H.-P., 1993. Evolution and Optimum Seeking: The Sixth Generation. John Wiley & Sons, Inc..

Sebag, M., Tarrisson, N., Teytaud, O., Lefevre, J., Baillet, S., 2005. A multi-objective multi-modal optimization approach for mining stable spatio-temporal patterns. In: Proceedings of the 19th International Joint Conference on Artificial Intelligence (IJCAI). Morgan Kaufmann Publishers Inc., pp. 859–864.

Segura, C., Coello Coello, C.A., Miranda, G., León, C., 2016. Using multi-objective evolutionary algorithms for single-objective constrained and unconstrained optimization. Ann. Oper. Res. 240 (1), 217–250.

Shir, O.M., Preuss, M., Naujoks, B., Emmerich, M.T.M., 2009. Enhancing decision space diversity in evolutionary multiobjective algorithms. In: Proceedings of the 5th International Conference on Evolutionary Multi-Criterion Optimization (EMO). Springer, pp. 95–109.

Srinivas, N., Deb, K., 1994. Multiobjective optimization using nondominated sorting in genetic algorithms. Evol. Comput. (ECJ) 2 (3), 221–248.

Stadler, P.F., Flamm, C., 2003. Barrier trees on poset-valued landscapes. Genet. Program. Evol. Mach. 4 (1), 7–20.

Steinhoff, V., Kerschke, P., Aspar, P., Trautmann, H., Grimme, C., 2020. Multiobjectivization of local search: Single-objective optimization benefits from multi-objective gradient descent. In: Proceedings of the IEEE Symposium Series on Computational Intelligence (SSCI). pp. 2445–2452.

Sun, Y., Kirley, M., Halgamuge, S.K., 2017. Quantifying variable interactions in continuous optimization problems. IEEE Trans. Evol. Comput. (TEVC) 21 (2), 249–264.

Tanabe, R., Ishibuchi, H., 2018. A decomposition-based evolutionary algorithm for multi-modal multi-objective optimization. In: Proceedings of the 15th International Conference on Parallel Problem Solving from Nature (PPSN XV). Springer, pp. 249–261.

Tanabe, R., Ishibuchi, H., 2019. A niching indicator-based multi-modal many-objective optimizer. Swarm Evol. Comput. (SWEVO) 49, 134–146.

Tanabe, R., Ishibuchi, H., 2020. A review of evolutionary multi-modal multi-objective optimization. IEEE Trans. Evol. Comput. (TEVC) 24 (1), 193–200.

Tran, T.-D., Brockhoff, D., Derbel, B., 2013. Multiobjectivization with NSGA-II on the noiseless BBOB testbed. In: Proceedings of the 15th Annual Conference on Genetic and Evolutionary Computation (GECCO) Companion. ACM, pp. 1217–1224.

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

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

Tušar, T., Filipič, B., 2015. 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.

Ulrich, T., Bader, J., Thiele, L., 2010. Defining and optimizing indicator-based diversity measures in multiobjective search. In: Proceedings of the 11th International Conference on Parallel Problem Solving from Nature (PPSN XI). Springer, pp. 707–717.

Van Geldrop, J., 1980. A note on local Pareto optima. J. Math. Econom. 7 (1), 51–54.

van Veldhuizen, D.A., 1999. 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.

Verel, S., Liefooghe, A., Jourdan, L., Dhaenens, C., 2011. Pareto local optima of multiobjective NK-landscapes with correlated objectives. In: Merz, P., Hao, J.-K. (Eds.), Evolutionary Computation in Combinatorial Optimization, Vol. 6622. Springer, pp. 226–237.

Viennet, R., Fonteix, C., Marc, I., 1996. Multicriteria optimization using a genetic algorithm for determining a Pareto set. Int. J. Syst. Sci. 27 (2), 255–260.

Volz, V., Naujoks, B., Kerschke, P., Tušar, T., 2019. Single- and multi-objective game-benchmark for evolutionary algorithms. In: Proceedings of the 21st Annual Conference on Genetic and Evolutionary Computation (GECCO). ACM, pp. 647–655.

Wan, Y.-H., 1975. On local Pareto optima. J. Math. Econom. 2 (1), 35–42.

Wang, H., Deutz, A.H., Bäck, T.H.W., Emmerich, M.T.M., 2017a. Hypervolume indicator gradient ascent multi-objective optimization. In: Proceedings of the 9th International Conference on Evolutionary Multi-Criterion Optimization (EMO). Springer, pp. 654–669.

Wang, Z., Ong, Y.-S., Sun, J., Gupta, A., Zhang, Q., 2018. A generator for multiobjective test problems with difficult-to-approximate Pareto front boundaries. IEEE Trans. Evol. Comput. (TEVC) 23 (4), 556–571.

Wang, H., Ren, Y., Deutz, A.H., Emmerich, M.T.M., 2017b. On steering dominated points in hypervolume indicator gradient ascent for bi-objective optimization. In: NEO 2015: Results of the Numerical and Evolutionary Optimization Workshop. Springer, pp. 175–203.

Wessing, S., 2015a. The Multiple Peaks Model 2. Tech. Rep. TR15-2-001, TU Dortmund University, Germany, pp. 104–108.

Wessing, S., 2015b. Two-Stage Methods for Multimodal Optimization (Ph.D. thesis). Technische Universität Dortmund, http://dx.doi.org/10.17877/DE290R-7804.

Wessing, S., 2016. Optproblems: Infrastructure to define optimization problems and some test problems for black-box optimization.

Wessing, S., Preuss, M., 2016. On multiobjective selection for multimodal optimization. Comput. Optim. Appl. 63 (3), 875–902.

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

Yue, C., Qu, B., Liang, J., 2018. A multiobjective particle swarm optimizer using ring topology for solving multimodal multiobjective problems. IEEE Trans. Evol. Comput. (TEVC) 22 (5), 805–817.

Yue, C., Qu, B., Yu, K., Liang, J., Li, X., 2019. A novel scalable test problem suite for multimodal multiobjective optimization. Swarm Evol. Comput. 48, 62–71.

Zapotecas-Martínez, S., Coello Coello, C.A., Aguirre, H.E., Tanaka, K., 2018. A review of features and limitations of existing scalable multiobjective test suites. IEEE Trans. Evol. Comput. (TEVC) 23 (1), 130–142.

Zechman, E.M., Giacomoni, M.H., Shafiee, M.E., 2013. An evolutionary algorithm approach to generate distinct sets of non-dominated solutions for wicked problems. Eng. Appl. Artif. Intell. 26 (5), 1442–1457.

Zhang, Q., Li, H., 2007. MOEA/D: A multiobjective evolutionary algorithm based on decomposition. IEEE Trans. Evol. Comput. (TEVC) 11 (6), 712–731.

Zhang, Q., Zhou, A., Zhao, S., Suganthan, P.N., Liu, W., Tiwari, S., 2008. Multiobjective Optimization Test Instances for the CEC 2009 Special Session and Competition, Vol. 264. Tech. rep., Univ. of Essex, Colchester, UK and Nanyang Tech. Univ., Singapore, Special Session on Performance Assessment of Multi-Objective Optimization Algorithms.

Zhou, A., Zhang, Q., Jin, Y., 2009. Approximating the set of Pareto-optimal solutions in both the decision and objective spaces by an estimation of distribution algorithm. IEEE Trans. Evol. Comput. 13 (5), 1167–1189.

Zitzler, E., Deb, K., Thiele, L., 2000. Comparison of multiobjective evolutionary algorithms: Empirical results. Evol. Comput. (ECJ) 8 (2), 173–195.

Zitzler, E., Laumanns, M., Thiele, L., 2001. SPEA2: Improving the strength pareto evolutionary algorithm for multiobjective optimization. In: Evolutionary Methods for Design Optimization and Control with Applications to Industrial Problems. International Center for Numerical Methods in Engineering, Athens, Greece, pp. 95–100.

Zitzler, E., Thiele, L., 1998. Multiobjective optimization using evolutionary algorithms – a comparative case study. In: Proceedings of the 5th International Conference on Parallel Problem Solving from Nature (PPSN V). Springer, pp. 292–304.

Zitzler, E., Thiele, L., Laumanns, M., da Fonseca, C.M.M., da Fonseca, V.G., 2003. Performance assessment of multiobjective optimizers: An analysis and review. IEEE Trans. Evol. Comput. (TEVC) 7 (2), 117–132.