Towards Analyzing Multimodality of Multiobjective Landscapes

Pascal Kerschke[1], Hao Wang[2], Mike Preuss[1], Christian Grimme[1], André Deutz[2], Heike Trautmann[1], and Michael Emmerich[2]

[1] Information Systems and Statistics Group, University of Münster, Germany
{pascal.kerschke, christian.grimme, mike.preuss, heike.trautmann}@wi.uni-muenster.de

[2] LIACS, Leiden University, The Netherlands
{fm.t.m.emmerich, h.wang, a.h.deutz}@liacs.leidenuniv.nl

Introduction

Multiobjective optimization algorithms are particularly challenged by multimodality of the underlying landscape caused by interaction of objective functions. Thus, sophisticated Exploratory Landscape Analysis (ELA) features which are able to assess the level and type of multimodality have huge potential for understanding algorithm behaviour, automated algorithm selection and algorithm design.

Here, we lay the groundwork for constructing such experimental features systematically by providing formal definitions of multimodality in terms of distinguishing between local and global efficient sets.

Multimodality

Some topological notation in $\mathbb{R}^d$ (decision space):

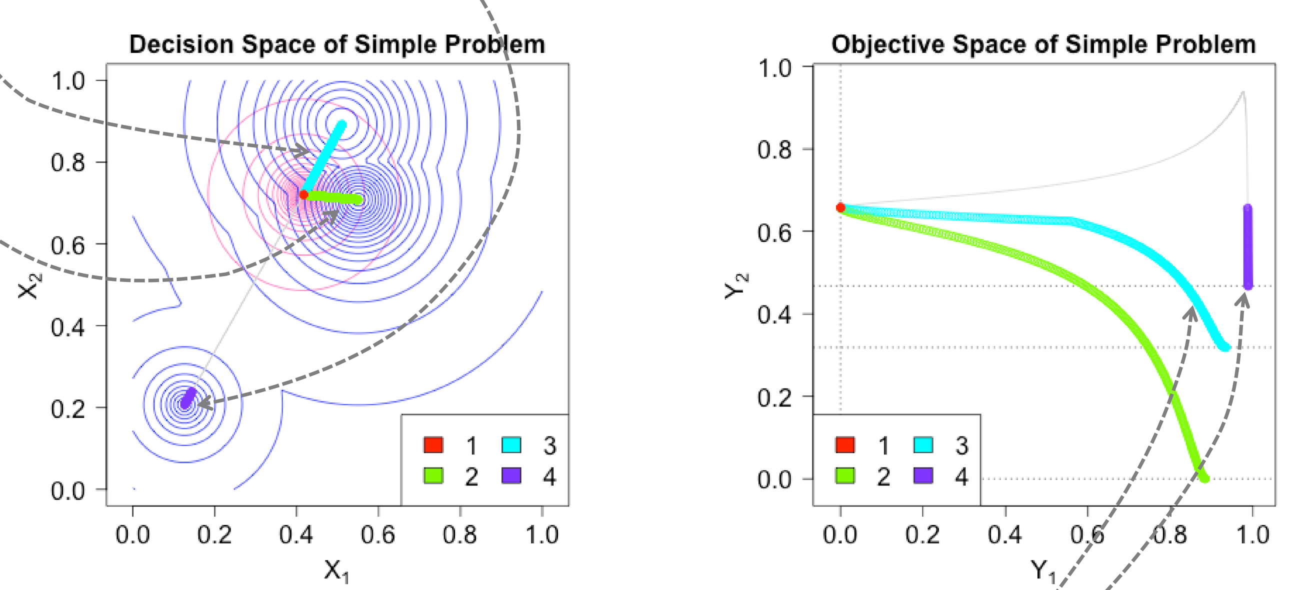
1. Let $A \subseteq \mathbb{R}^n$. The set A is called connected if and only if there do not exist two open, *disjoint* subsets U_1 and U_2 of $\mathbb{R}^n$ such that $A \subseteq U_1 \cup U_2$, $U_1 \cap A \neq \emptyset$, and $U_2 \cap A \neq \emptyset$.
2. Let $B \subseteq \mathbb{R}^n$. A subset $C \subseteq B$ is a connected component of B iff C is connected, and any subset of B which is a strict superset of C is not connected, and C is non-empty.

Pareto concepts (very brief):

Let $\mathbf{f} : \mathcal{X} \to \mathbb{R}^m$ be a multiobjective function where $\mathcal{X} \subseteq \mathbb{R}^d$ is the decision space. We will denote the component functions of $\mathbf{f}$ by $f_i : \mathcal{X} \to \mathbb{R}$, $i = 1, \ldots, m$. A point $\mathbf{x} \in \mathcal{X}$ is called *Pareto efficient* or *global efficient* iff there does not exist $\tilde{\mathbf{x}} \in \mathcal{X}$ such that $\mathbf{f}(\tilde{\mathbf{x}}) \prec \mathbf{f}(\mathbf{x})$. The subset of $\mathcal{X}$ consisting of all the efficient points of $\mathcal{X}$ is denoted by $\mathcal{X}_E$ and is called the *efficient subset* of $\mathcal{X}$ The image of $\mathcal{X}_E$ under $\mathbf{f}$ is called the Pareto front of $\mathbf{f}$.

Efficient points in $\mathcal{X}$ and local efficient sets in the multiobjective case:

- 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 $U \subseteq \mathbb{R}^d$ with $\mathbf{x} \in \mathbf{U}$ such that there is no point $\tilde{\mathbf{x}} \in U \cap \mathcal{X}$ such that $\mathbf{f}(\tilde{\mathbf{x}}) \prec \mathbf{f}(\mathbf{x})$. The subset of all the local efficient points of $\mathcal{X}$ is denoted by $\mathcal{X}_{LE}$.
- A point $\mathbf{x} \in \mathcal{X}$ is called a *global efficient point* of $\mathcal{X}$ (or of $\mathbf{f}$) if there is no point $\tilde{\mathbf{x}} \in \mathbb{R}^d \cap \mathcal{X}$ such that $\mathbf{f}(\tilde{\mathbf{x}}) \prec \mathbf{f}(\mathbf{x})$. The subset of all the global efficient points of $\mathcal{X}$ is termed *efficient set* of $\mathbf{f}$ and denoted by $\mathcal{X}_E$.
- A subset $A \subseteq \mathcal{X}$ is a *local efficient set* of $\mathbf{f}$ if A is a connected component of $\mathcal{X}_{LE}$ (= the subset of $\mathcal{X}$ which consists of the local efficient points of $\mathcal{X}$).



Local and Global Pareto Front:

- A subset P of the image of $\mathbf{f}$ is a *local Pareto front* of $\mathbf{f}$, if there exists a local efficient set E such that $P = \mathbf{f}(E)$.
- The (global) *Pareto front* (PF) of $\mathbf{f}$ is obtained by taking the image under $\mathbf{f}$ of the union of the connected components of the set of global efficient points of $\mathcal{X}$. If $\mathcal{X}_E$ is connected, then the (global) Pareto front of $\mathbf{f}$ is also connected, provided $\mathbf{f}$ is continuous on $\mathcal{X}_E$.

Analytics on Simple Mixed Sphere Problems

$$f(\mathbf{x}) = 1 - \max_{1 \le i \le N} \{g_i(\mathbf{x})\}, \quad \mathbf{x} \in \mathbb{R}^d \tag{1}$$

$$g_i(\mathbf{x}) = H_i \left(1 + \left(\sqrt{(\mathbf{x} - \mathbf{c}_i)^T \mathbf{D} (\mathbf{x} - \mathbf{c}_i)}\right)^{s_i} / R_i \right)^{-1}, \quad i = 1, \ldots, N \tag{2}$$

Functions g_i define peaks with center $\mathbf{c}_i$, depth H_i, radius R_i and shape s_i. $\mathbf{D}$ is the covariance matrix inverse. We focus on the simplest bi-objective case where each objective function consists of only one spherical peak. The efficient set is the line segment connecting $\mathbf{c}$ to $\mathbf{c}'$, for $N = 2$.

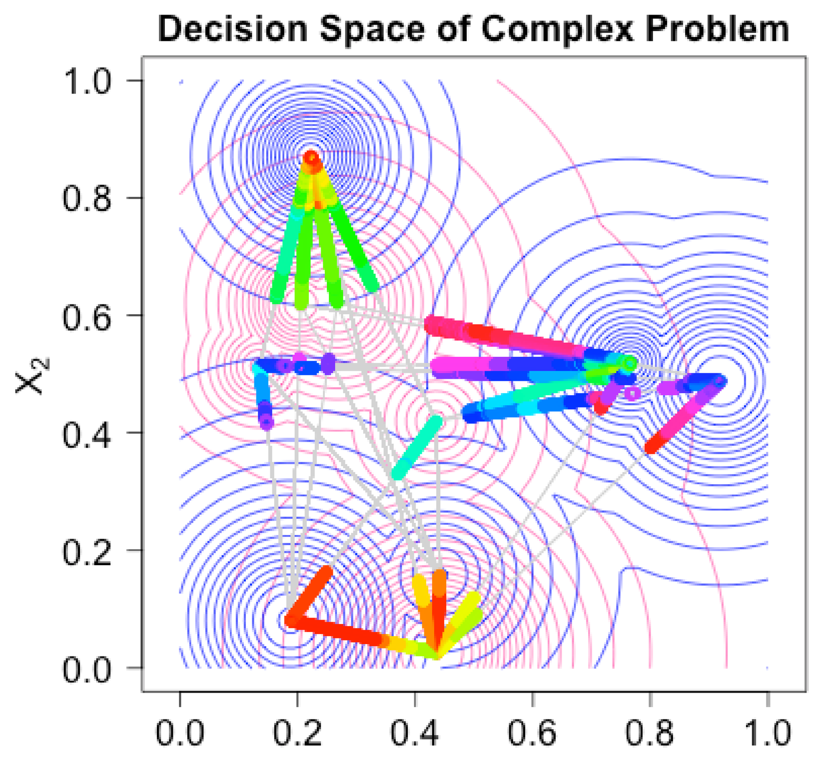
$$f_2 = 1 - H' \left(1 + \left(d(\mathbf{c}, \mathbf{c}'; \mathbf{D}') \left(1 - \frac{R^{1/s}}{d(\mathbf{c}, \mathbf{c}'; \mathbf{D})} \left(\frac{H}{1 - f_1} - 1\right)^{1/s}\right)\right)^{s'} / R' \right)^{-1}$$

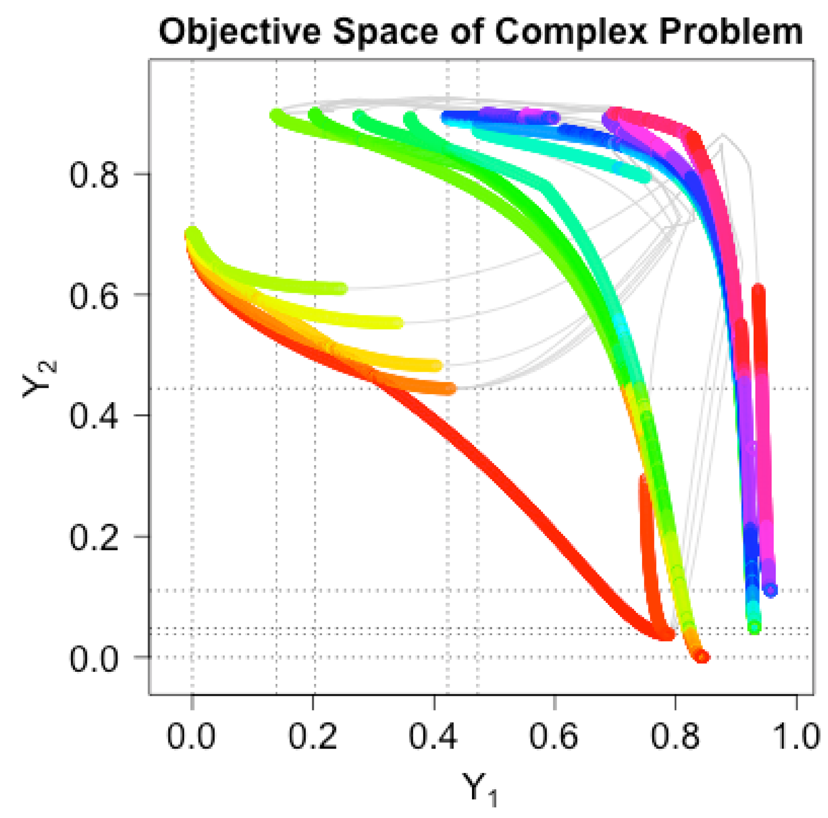
The range of f_1 is $[\min\{f_1(\mathbf{c}), f_1(\mathbf{c}')\}, \max\{f_1(\mathbf{c}), f_1(\mathbf{c}')\}]$.

Experiments

- Evaluation of problem and algorithm characteristics
- MPM-2 Generator used to build simple and complex problems (examples presented here for m = 2)

Count Characteristic	Scenario Simple	Scenario Complex
optima (obj. 1 vs. obj. 2)	1 vs. 3	5 vs. 5
domination layers	3	19
connected components	3	30
sets connected to global efficient set	2	66
fronts connected to global efficient front	2	12
local (global) efficient sets	4 (2)	167 (7)





- Apply two different local search strategies to the given problems:

Hypervolume Indicator Gradient Ascent (HIGA-MO)

Initialize the search points $\mathbf{X}$ uniformly in the search space
while the termination criteria are not satisfied **do**
  Evaluation: $\mathbf{Y} \leftarrow \mathbf{X}$
  $\{L_i\}_{i=1}^q \leftarrow$ non-dominated-sorting($\mathbf{R}$)
  for $i = 1$ to q **do**
    for every $\mathbf{s}_j$ in L_i **do**
      Compute the subgradient $\partial \mathcal{H}(\mathbf{X}) / \partial \mathbf{x}_j$
      $\mathbf{x}_j \leftarrow \mathbf{x}_j + \sigma \cdot [\partial \mathcal{H}(\mathbf{X}) / \partial \mathbf{x}_j]$
    end for
  end for
end while
return $\{L_i\}_{i=1}^q, \mathbf{X}, \mathbf{Y}$
  with $\mathbf{y}_i = [f_1(\mathbf{x}_i),\ f_2(\mathbf{x}_i)]^T,\ i = 1, \ldots, \mu$

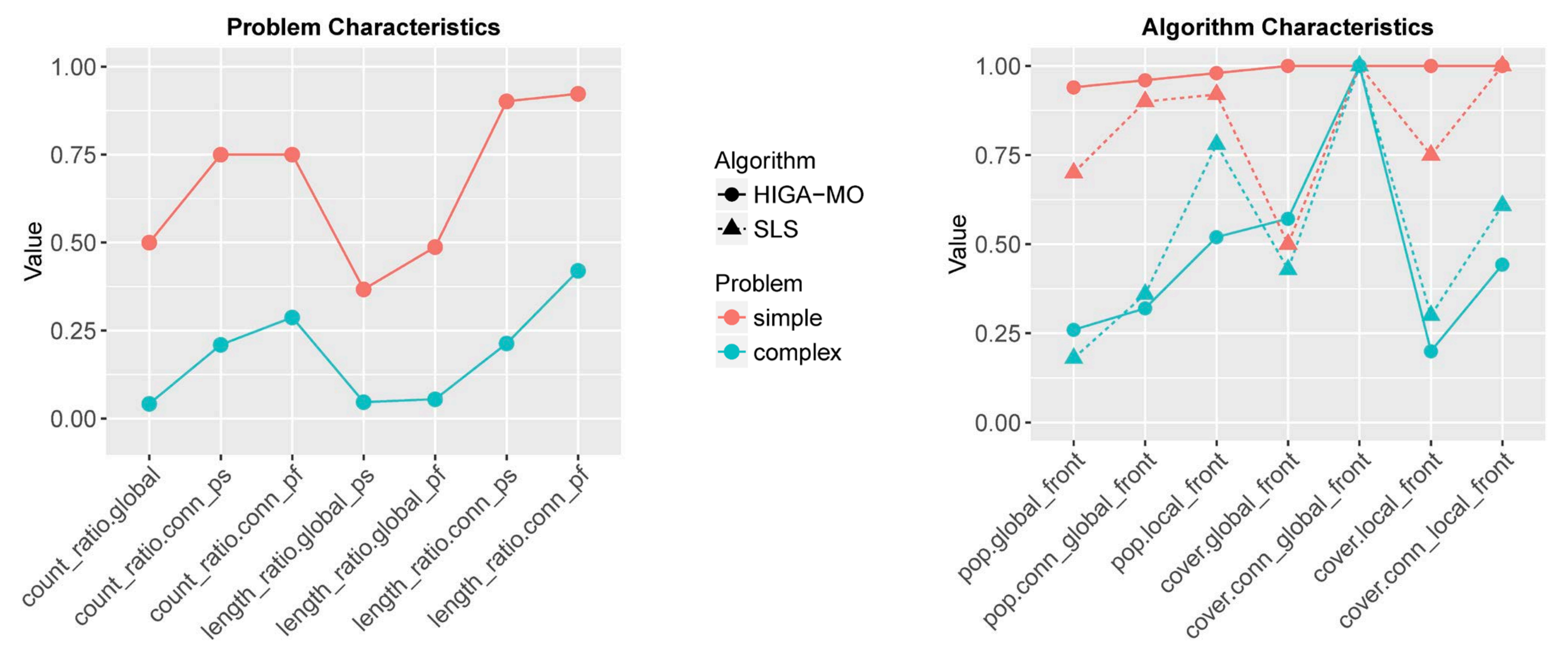
Stochastic Local Search (SLS)

Initially, μ independently random solutions (LHD).

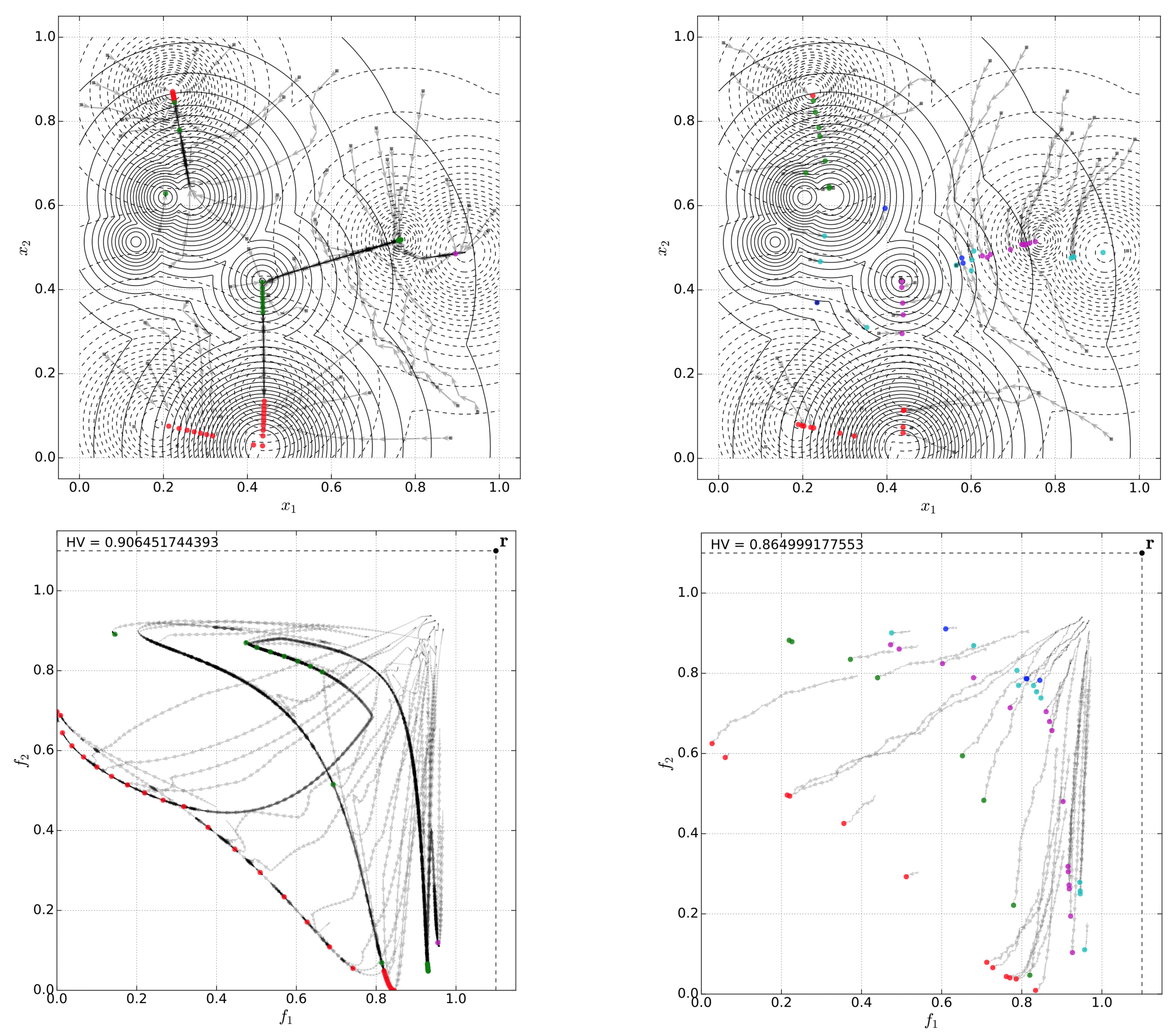
In every iteration, each solution undergoes upper bounded normal distributed perturbation with maximum step size of σ. Here the step-size is fixed

For each pair of original and related perturbed solutions the dominating solution survives.

Problem and Algorithm Characteristics



Local Search Behaviour



Conclusion

- This work provides a thorough definition of multimodality for multiobjective optimization problems.
- Analytical and experimental approaches are presented which derive the global and local Pareto fronts.
- Mixed sphere test problems of different levels of multimodality are designed and the behavior of HIGA-MO and a stochastic local search variant are contrasted.
- Multimodality is a crucial factor determining the difficulty of a problem
- Indicators are derived which allow to assess algorithm behavior w.r.t. the detection of global and local Pareto fronts which can further be used for performance assessment and (later) algorithm selection.