Marcin Łoś
Maciej Smołka
Robert Schaefer
Jakub Sawicki

MISFIT LANDFORMS IMPOSED BY ILL-CONDITIONED INVERSE PARAMETRIC PROBLEMS

Abstract

In this paper, we put forward a new topological taxonomy that allows us to distinguish and separate multiple solutions to ill-conditioned parametric inverse problems appearing in engineering, geophysics, medicine, etc. This taxonomy distinguishes the areas of insensitivity to parameters called the landforms of the misfit landscape, be it around minima (lowlands), maxima (uplands), or stationary points (shelves). We have proven their important separability and completeness conditions. In particular, lowlands, uplands, and shelves are pairwise disjoint, and there are no other subsets of the positive measure in the admissible domain on which the misfit function takes a constant value. The topological taxonomy is related to the second, "local" one, which characterizes the types of ill-conditioning of the particular solutions. We hope that the proposed results will be helpful for a better and more precise formulation of ill-conditioned inverse problems and for selecting and profiling complex optimization strategies used in solving these problems.

Keywords　taxonomy of ill-conditioned problems, ill-posed global optimization problems, fitness landscapes

1. Introduction

Formulating and solving Global Optimization Problems (GOPs) is one of the fundamental ways for modeling, planning, and optimizing important human activities. Roughly speaking, we first identify the decision variables that affect a given process, then define a numerical criterion that describes the process; e.g., as objective function f of the real-valued decision variables, in such a way that the better the controlled process is evaluated, the smaller the value of f is. Finally, we seek variable values that minimize this function. Unfortunately, in many real-world settings, the natural objective function has large regions (surrounding the global solutions) where they lose sensitivity to the parameters. Such GOPs belong to the group of **ill-conditioned** problems.

If admissible set of decision variables $\mathcal{D}$ is equipped with a metric and a vector structure (e.g., $\mathcal{D}$ is a regular domain in $\mathbb{R}^\ell$), the regions of insensitivity can be studied by topological methods. Such methods may also be applicable if $\mathcal{D}$ is discrete with a topology imposed by the inner graph structure. In both cases, finding the regions of insensitivity surrounding the global objective minimizers seems to be the primary task when solving ill-conditioned GOPs. If there are multiple non-intersecting regions of low sensitivity surrounding different global minimizers, the problem becomes **multimodal**.

Traditional approaches to handling multimodality and insensitivity in GOP solving rely on regularization methods [24]. These methods proceed by supplementing the objective function with a regularization term, making it globally convex. Unfortunately, such methods may lead to many undesirable artifacts as well as to the loss of information regarding the modeled process. Indeed, it can even lead to outright false solutions, forced by the regularization supplement.

A different and perhaps more reliable approach to handling multimodality and insensitivity is based on finding approximations of the insensitivity regions that contain the global objective minimizers.

The effective handling of the insensitivity and multimodality of the objective functions is particularly essential for solving various tasks in technology formulated as **inverse parametric problems** (IPPs) (e.g., regarding lens design [12]), in geophysics (e.g., in the calibration of conceptual Rainfall-Runoff Models [10], or for investigating oil and gas resources [22]) as well as in medical diagnoses (e.g., in tumor tissue identification [18]).

The important introductory step to such an analysis needs a precise definition of the set of solutions adequate to the particular problem and the taxonomy of other important sets contained in the admissible domain affecting the problem difficulty, as the sets and basins of attractions and diverse kinds of "plateaus" can trap the searching process.

Various definitions of such sets might be found in the book of Mike Preuss [19] and Bernadette Addis's PhD Thesis [1] as well as in earlier papers by Aimo Törn [25, 26], Dixon and Szegö [8, 9], and Boender, RinnoyKan, Stougie, Timmer [5].

Unfortunately, the taxonomies found in the literature do not fit all of the natural needs formulated in the analysis of ill-posed inverse problems. In particular, there is no precise definition of the sets of solutions of the positive Lebesgue measure (plateaus, lowlands) nor a proper definition of their sets of attractions and basins of attractions. In particular, the current definitions of basins of attraction do not guarantee the separation of the sets composed of local and global minima.

The main contribution of this paper is an exhaustive topological taxonomy of the sets mentioned above, handled as the "landforms" in the global optimization landscape; i.e., the subsets of the admissible domain playing the special role in the objective function's chart. The landforms of our particular interest are lowlands, uplands, shelves, and their sets and basins of attraction. It was proven that lowlands and their basins of attractions are pairwise disjoint, which allows for precise defining and finding separate sets of solutions to multimodal ill-conditioned problems by stochastic methods. Moreover, the uplands and shelves with their sets of attraction that constitute potential traps for searching algorithms are separated.

Moreover, it was possible to analyze the completeness of the introduced topological taxonomy. We have proven lowlands, uplands and shelves to be the only subsets of a positive measure of the admissible domain on which objective takes a constant value.

The topological taxonomy is related to the so-called "local" taxonomy of solutions to ill-posed inverse parametric problems formulated as global optimization ones. This taxonomy tries to characterize the types of ill-conditioning of a particular solution or a set of solutions.

We suggest to avoid the term "plateau", as it is less concrete than the proposed terms of lowland, upland, and shelf. Even though the term is used in the literature (see [16, 17, 19]), it introduces some ambiguity.

We hope that the proposed taxonomies and definitions will be helpful for a precise formulation of the ill-posed problems appearing in engineering. Moreover, the taxonomies might be helpful for selecting and profiling complex optimization strategies used to solve these problems.

2. Taxonomy of ill-conditioned inverse problems – the local approach

Following [15], we present a simple taxonomy of ill-posed IPPs. For introducing abstract mathematical definitions, we consider a sample *forward* problem represented by the elliptic variational equation:

$$\text{Given } \omega \in \mathcal{P}, \text{ find } u \in U; \ b(\omega; u, v) = l(v) \ \forall v \in U, \tag{1}$$

where U is a Hilbert solution space (space of states of the system), which typically is infinite-dimensional, $\mathcal{P}$ is a Banach space of the problem's parameters,

$b : \mathcal{P} \times U \times U \to \mathbb{R}$ is a bilinear continuous form with respect to the second and third variables, and $l : U \to \mathbb{R}$ is a continuous linear functional.

If form b is uniformly coercive (i.e., $b(\omega; u, u) > c_0 \|u\|_U^2 \ \forall u \in U$, where constant $c_0 > 0$ does not depend on $\omega \in \mathcal{P}$), then (1) is well-posed in the sense of Hadamard, uniformly with respect to the parameters. The forward solution frequently continuously depends on the parameter, so we can establish a continuous mapping:

$$A_1 : \mathcal{P} \ni \omega \to u \in U; \quad \omega, u \ \text{satisfy (1).} \tag{2}$$

Typically, the forward solution (state) is not totally observable, and the state measurements being the input of the IPP belong to $\mathcal{O}$ – a finite dimensional subspace of U. Therefore, that we can establish the linear continuous projection:

$$A_2 : U \ni u \to \overline{u} \in \mathcal{O}. \tag{3}$$

An abstract IPP, a problem inverse to (1), consists in finding the following:

$$\omega \in \mathcal{D} \subset \mathcal{P}; \ \ A(\omega) = u_0, \tag{4}$$

where $u_0 \in \mathcal{O}$ is the given state (forward solution) observation and $A = A_2 \circ A_1 : \mathcal{P} \to \mathcal{O}$ is assumed to be continuous in the appropriate topologies. Moreover, $\mathcal{D} \subset \mathcal{P}$ stands for the nonempty set of admissible parameters restricted by some particular constraints. We also assume that $\mathcal{D}$ has a positive measure in space $\mathcal{P}$.

Problem (4) is *well posed in the sense of Hadamard* if it possesses a unique stable solution that is always satisfied if A is one-to-one and A^{-1} is continuous (see, e.g., [4, 13]). It is possible to consider a wider class of IPPs *conditionally well-posed in the sense of Tikhonov* for which there exists a *set of well posedness $M \subset \mathcal{D} \subset \mathcal{P}$* so that $u_0 \in A(M)$ and:

1. $A(\omega) = \overline{u}$ has a unique solution for all $\overline{u} \in A(M)$,

2. A^{-1} is continuous on $A(M)$.

(see Definitions 4.3 and 4.4 in [13]). This is the broadest class of IPPs feasible to be solved by regularization, which iteratively "smoothes" operator A^{-1} to finally reach its global continuity (see, e.g., [24]).

Definition 1. *(see [15], Definition 1) An inverse parametric problem that is not conditionally well-posed in the sense of Tikhonov and cannot be successfully solved by regularization will be called* irremediable.

Most frequently, IPPs appearing in real-world engineering applications are irremediable ill-posed because of at least two reasons:

(i) An ambiguous mathematical model (which allows multiple solutions; see, e.g., [6]) such that A_1 is not an injection. Moreover, the exact value of $A_1(\omega)$ may not be explicitly reachable – it can only be approximated by numerical simulation (as in the case of (1)).

(ii) Measurements $\overline{u} \in \mathcal{O}$ do not deliver the complete information necessary to reconstruct forward solution $u \in U$ with sufficient accuracy. Moreover, A_2 is rarely an injection, as it is a projection of an infinitely-dimensional space U on a finitely-dimensional space $\mathcal{O}$.

Summing up, A is rarely an injection and IPPs may allow multiple solutions, even forming infinite dense subsets contained in $A^{-1}(u_0) \subset \mathcal{D}$.

In the remaining part of this section, we will introduce a taxonomy of irremediable ill-posed IPPs based on their formulation as a GOP that is not as general as (4), suggesting a particular way of finding its solution:

$$\text{arg } \min_{\omega \in \mathcal{D} \subset \mathcal{P}} \{f(u_0, u_\omega); \; A_1(\omega) = u_\omega\}, \tag{5}$$

where $f : \mathcal{O} \times U \to \mathbb{R}_+ \cup \{0\}$ stands for a *misfit functional* evaluating the gap between an observation and a simulated forward solution.

The transition from abstract formulation (4) to the above (5) is not unique and depends on the misfit setting. Moreover, the misfit selection strongly affects the conditioning of the inverse problem. The simplest setting might be $f(u_0, u) = \|u_0 - A_2(u)\|_{\mathcal{O}}^2$ or $f(u_0, u) = \|u_0 - u\|_U^2$ if $\mathcal{O} \subset U$.

Let us denote by $\mathcal{S} \subset \mathcal{D} \subset \mathcal{P}$ a set of solutions to problem (5) for given observation $u_0 \in \mathcal{O}$ and assume that it is nonempty ($\mathcal{S} \neq \emptyset$); then:

C1. Problem (5) is *unimodal* if $\text{card}(\mathcal{S}) = 1$.

C2. Problem (5) is *multimodal* if $\text{card}(\mathcal{S}) > 1$.

C3. Solution $\omega \in \mathcal{S}$ is *sensitive and stable* on open connected set $B \subset \mathcal{D}$ with positive Lebesgue measure $\text{meas}(B) > 0$ if $\omega \in B$ and

$$\forall x \in B, \; x \neq \omega \quad c_0 \leq \frac{|f(u_0, u_\omega) - f(u_0, u_x)|}{\|\omega - x\|_{\mathcal{P}}} \leq c_1,$$

for some reasonable real constants $c_0, c_1 \in \mathbb{R}$; $0 < c_0 < c_1 < +\infty$.

C4. Solution $\omega \in \mathcal{S}$ is *g-unstable* on open connected set $B \subset \mathcal{D}$ $\text{meas}(B) > 0$ if $\omega \in B$ and

$$\exists n \in \mathcal{P}; \; \|n\|_{\mathcal{P}} = 1; \quad \forall \alpha \neq 0, \; x = \omega + \alpha n \in B, \; g \leq \frac{|f(u_0, u_\omega) - f(u_0, u_x)|}{\|\omega - x\|_{\mathcal{P}}}$$

for some sufficiently large constant $g > 0$.

C5. Solution $\omega \in \mathcal{S}$ is *γn-insensitive* on open connected set $B \subset \mathcal{D}$ $\text{meas}(B) > 0$ if $\omega \in B$ and

$$\exists n \in \mathcal{P}; \; \|n\|_{\mathcal{P}} = 1; \quad \forall \alpha \neq 0, \; x = \omega + \alpha n \in B, \; \frac{|f(u_0, u_\omega) - f(u_0, u_x)|}{\|\omega - x\|_{\mathcal{P}}} \leq \gamma,$$

for some sufficiently small constant $\gamma > 0$.

C6. Solution $\omega \in \mathcal{S}$ is *γ-insensitive* on open-connected set $B \subset \mathcal{D}$ meas$(B) > 0$ if $\omega \in B$ and w is *γn-insensitive* on set B, $\forall n \in \mathcal{P}$; $\|n\|_{\mathcal{P}} = 1$.

C7. Set of solutions $A \subseteq \mathcal{S}$ to problem (5) is *γ-insensitive* if there exist disjoint connected sets $B_1, \ldots, B_p \subset \mathcal{D}$, $p \geq 1$ such that meas$(B_i) > 0$, $i = 1, \ldots, p$ and $\forall \omega \in A$ $\exists i \in \{1, \ldots, p\}$ so that ω is *γ-insensitive* on set B_i.

C8. Problem (5) is *γ-unspecified* if there exists a set $A \subseteq \mathcal{S}$ of *γ-insensitive* solutions so that meas$(A) > 0$.

C9. Problem (5) is *γ-unspecified, multimodal* if it is *γ-unspecified* and $p > 1$.

Remark 2.

(a) *Conditions C1, C2, and C3 correspond in some way to the conditional well-posedness in the sense of Tikhonow, but it depends strongly on the particular misfit setting.*

(b) *The above taxonomy is computationally-oriented, so the values of constants g, c_0, c_1, and γ may depend on the accuracy of the algorithm applied for solving (5) and the available arithmetic accuracy.*

(c) *The conditioning of problem (5) might be additionally worsened by the errors of an approximate solution of forward problem (1) and the global optimization strategies of finding all misfit minimizers.*

(d) *The g-unstable solutions satisfying C4 (e.g. needle minimizers) are most likely artifacts and can be ignored.*

3. Topological taxonomy of GOP's landforms

3.1. Global optimization problems in continuous domains

For many important IPPs, we use the approximate version of (5) in which the admissible set of parameters $\omega \in \Sigma$ will be represented by the closed, bounded, and connected domain $\mathcal{D} \subset \mathbb{R}^{\ell}$, $1 \leq \ell < +\infty$ with a nonempty interior and sufficiently regular boundary (e.g., Lipschitz boundary [28]). The dependency on observation u_0 will be further omitted for the sake of simplicity. Assuming observation $u_0 \in \mathcal{O}$, the misfit will be represented by continuous function $f \in C(\mathbb{R}^{\ell} \to \mathbb{R})$.

It is obvious that f attains its minimum and maximum values over $\mathcal{D}$; i.e.,

$$\exists x_{min}, x_{max} \in \mathcal{D}; \ \ \forall x \in \mathcal{D} \ \ -\infty < f(x_{min}) \leq f(x) \leq f(x_{max}) < +\infty. \tag{6}$$

Of course, we can have many global minimizers and maximizers. Moreover, neither of the $f(x_{min})$ and $f(x_{max})$ values are necessarily known a priori. In some IPPs, $f(x_{min}) = 0$, which represents the perfect parameters' fit. It other cases, $f(x_{min}) > 0$ because of the irreducible errors of the mathematical model of the real-world problem as well as imperfection of the applied approximation.

Following (5), we intend to solve the GOP:

$$\arg \min_{x \in \mathcal{D}} \{f(x)\}, \ \mathcal{D} \subset \mathbb{R}^{\ell}, \ f : \mathcal{D} \to \mathbb{R}, \tag{7}$$

The goal of next sections is to establish a taxonomy of solution sets for (7) that corresponds to the local taxonomy of the IPP's ill-conditioning. Such a taxonomy will be convenient for studying the features of complex solving strategies (e.g., memetic strategies [7]) that combine a global stochastic search with various algorithms that improve the search results locally. The latter can be both stochastic or deterministic and can either speed up finding isolated local minimizers or approximate the shapes of areas on which the objective attains a locally minimal constant value.

3.2. Minimum manifolds and lowlands

Definition 3.

1. *Each* $y \in \mathcal{D}$ *will be called the local minimizer to* f *in* $\mathcal{D}$ *if*

$$\exists B_y \subset \mathcal{D}; \ y \in B_y, \ B_y \ is \ connected, \ f(y) = f(\xi) \ \forall \xi \in B_y \ \ and$$
$$\exists A \in \text{top}(\mathbb{R}^{\ell}); \ \overline{B_y} \subset A, \ f(y) < f(\xi) \ \forall \xi \in (A \cap \mathcal{D}) \setminus B_y. \tag{8}$$

2. *Each local minimizer* y *to* f *in* $\mathcal{D}$ *will be called a global minimizer if* $f(y) = f(x_{min})$.

3. *Set* B_y *will be called the minimum manifold to* f *in* $\mathcal{D}$. *We will denote by* $\text{MManifolds}_{f,\mathcal{D}}$ *the family of all minimum manifolds to* f *in* $\mathcal{D}$.

4. *Each minimizer* $y \in \mathcal{D}$ *to* f *in* $\mathcal{D}$ *will be called isolated if* $B_y = \{y\}$.

Lemma 4. *Disjoint open sets* U *and* V *are separated* – $\overline{U} \cap V = U \cap \overline{V} = \emptyset$ *(The upper bar over the set symbol denotes its topological closure; i.e.,* $\overline{V}$ *is the closure of* V*).*

Proof. Suppose $x \in U \cap \overline{V}$. By definition of closure $U \cap V \neq \emptyset$, which makes the contradiction. The second case follows by symmetry. $\square$

Theorem 5. *Let* y *be a local minimizer of* f *in* $\mathcal{D}$, $f(y) = a$. *Then,* B_y *is a connected component of* $(f|_{\mathcal{D}})^{-1}(a)$ *that contains* y.

Proof. Let $D_a = (f|_{\mathcal{D}})^{-1}(a)$, then $D_a = f^{-1}(a) \cap \mathcal{D}$. Since f is continuous, $f^{-1}(a)$ is closed. Therefore, D_a is closed as an intersection of two closed sets. Let C_y denote the connected component of D_a containing y. Since the connected components of D_a are closed in D_a and D_a is itself closed in $\mathbb{R}^n$, then C_y is closed as well. Our goal is to prove that $B_y = C_y$.

By definition of B_y and by the continuity of f, we have $\overline{B_y} \subset D_a$. Moreover, $\overline{B_y}$ is connected because B_y is connected.

This together with $y \in B_y$ implies $\overline{B_y} \subset C_y$. It is enough to show that $C_y \subset A$, since if $\xi \in C_y \setminus B_y$, we have $\xi \in (A \cap D) \setminus B_y$ and so $f(\xi) > a$, which contradicts $\xi \in C_y$.

Since $\overline{B_y} \subset A$, sets $\overline{B_y}$ and $\mathbb{R}^n \setminus A$ are disjoint and closed. $\mathbb{R}^n$ as a metric space is normal, so they have disjoint open neighborhoods; i.e., there are open sets $U \supset \overline{B_y}$, $V \supset \mathbb{R}^n \setminus A$ with $U \cap V = \emptyset$. By Lemma 4, $\overline{U} \cap V = U \cap \overline{V} = \emptyset$, so the same is true for $S = U \cap C_y$, $T = V \cap C_y$. Furthermore, $S \cup T = C_y$, for $C_y \subset \mathcal{D}$ and if $\xi \in \mathcal{D} \setminus (S \cup T)$, then $\xi \in \mathcal{D} \setminus (U \cup V) \subset \mathcal{D} \setminus (B_y \cup \mathbb{R}^n \setminus A) = (A \cap \mathcal{D}) \setminus B_y$, and so $f(\xi) > a$, thus $\xi \notin C_y$.

Recall that, by the definition of connectedness, either $S = \emptyset$ or $T = \emptyset$. Since $y \in B_y \subset C_y \cap U = S$, we have $T = \emptyset$; so, $C_y \setminus A = C_y \cap (\mathbb{R}^n \setminus A) \subset C_y \cap V = T = \emptyset$ and $C_y \subset A$. $\qquad\square$

Observation 6.

1. $\forall y' \in B_y \ B_{y'} = B_y$.
2. $B_y = \overline{B_y}$.

Proof. The first part follows immediately from the above theorem and the definition of the connected component. The second assertion follows from the above proof – using the notation from the proof, $B_y = C_y$, and C_y is shown to be closed. $\qquad\square$

Remark 7. *Note that the above definition of the local minimizer is strictly stronger than the classical definition of a (weak) local minimum. In some cases, points that we would intuitively like to consider as minimizers fail to meet the requirements of the definition. Consider $D = [-1, 1]$ and $f\colon D \to \mathbb{R}$ given by*

$$f(x) = \begin{cases} 0 & \text{for } x = 0 \\ x^2 \sin^2\left(\frac{1}{x}\right) & \text{for } x \neq 0 \end{cases} \tag{9}$$

It is a continuous function with minimal value 0 assumed at $M = \{1/k\pi\colon k \in \mathbb{Z},\ k \neq 0\} \cup \{0\}$, but $x = 0$ is not a local minimizer. Indeed, if it were, then $B_0 = \{0\}$ by Theorem 5; however, since $x = 0$ is a limit point of M, no neighborhood A of 0 can satisfy the condition that $f(\xi) > 0$ for $\xi \in A \setminus \{0\}$.

This can be seen as a limitation of the proposed definition. Note, however, that the nature of the above example makes it unlikely for any numerical method to handle in a satisfactory manner.

Theorem 5 tells us that minimum manifolds are connected components of level sets of f. The example from Remark 7 shows that the converse in general is false; i.e., the connected components of (even globally) minimal level sets of f need not be minimum manifolds. The following theorem may be seen as a partial converse:

Theorem 8. *Let f satisfy the additional condition: for each $a \in \mathbb{R}$, the connected components of $(f|_\mathcal{D})^{-1}(a)$ form a locally finite family. Then, the following is true: let $y \in \mathcal{D}$, $f(y) = a$ and let C_y be a connected component of $(f|_\mathcal{D})^{-1}(a)$ containing y. If there exists an open set A such that $C_y \subset A$ and $f(\xi) \geq a$ for each $\xi \in A \cap \mathcal{D}$, then y is a local minimizer of f.*

Proof. By the assumption, the family of connected components of $(f|_{\mathcal{D}})^{-1}(a)$ is locally finite. For each $x \in C_y$, let us choose open neighborhood U_x of x that intersects finitely many such components $C_1, \ldots, C_k$ apart from C_y. We can assume $U_x \subset A$ (otherwise, we can pick $U_x \cap A$). Then, $\tilde{U}_x = U_x \setminus (C_1 \cup, \ldots, \cup C_k)$ is open, since C_i are closed, and $x \in \tilde{U}_x$, since the connected components are disjoint. Let

$$U = \bigcup_{x \in C_y} \tilde{U}_x.$$

Then U is an open neighborhood of C_y contained in A that does not intersect any connected components of $(f|_{\mathcal{D}})^{-1}(a)$ except C_y.

Let $\xi \in (U \cap \mathcal{D}) \setminus C_y$. By assumptions, $f(\xi) \geq a$. However, if $f(\xi) = a$, then $\xi \in f^{-1}(a)$; so, $\xi \in C_y$, since U does not intersect any other component of $(f|_{\mathcal{D}})^{-1}(a)$. Thus, $f(\xi) > a = f(y)$; so, y is a local minimizer. □

Let us consider the local minimizer z contained in some minimal manifold B_y satisfying Definition 3.

Definition 9. *Lowland $\mathcal{P}_z$ to f in $\mathcal{D}$ associated with local minimizer $z \in B_y$ is the maximum open-connected set (in the sense of inclusion) so that $\mathcal{P}_z \subset B_y$ and $z \in \mathcal{P}_z$ or the empty set. Let us denote by* Lowlands$_{f,\mathcal{D}}$ *the family of all lowlands to f in $\mathcal{D}$.*

Remark 10.

1. *definition 9 is well-posed; i.e., for each $z \in B_y$, set $\mathcal{P}_z$ is unambiguously defined (if it exists);*
2. $\forall z' \in \mathcal{P}_z \; f(z') = f(z) = f(y)$;
3. $\forall z' \in \mathcal{P}_z \; \mathcal{P}_z = \mathcal{P}_{z'}$.

There is no simple relationship between the number of minimum manifolds and the number of lowlands (see Figure 1).
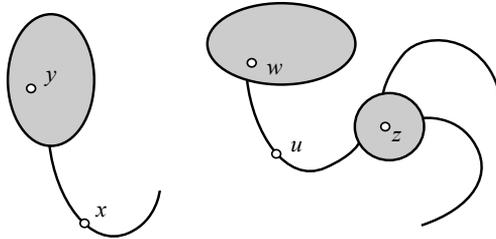
Figure 1. Minimum manifold B_x contains only one lowland ($\mathcal{P}_y$), while minimum manifold B_u contains two lowlands ($\mathcal{P}_w$ and $\mathcal{P}_z$)

Observation 11. *If f is a differentiable function, then it is well-known that, for every minimum manifold B_y and every $x \in B_y$, we have*

$$\nabla f(x) = 0.$$

If, moreover, f is twice differentiable, then Hessian matrix $Hf(x)$ is positive semi-definite.

3.3. Maximum manifolds and uplands

Definition 12.

1. Each $y \in \mathcal{D}$ will be called the local maximizer to f in $\mathcal{D}$ if

$$\exists G_y \subset \mathcal{D};\ y \in G_y,\ G_y\ is\ connected,\ \ f(y) = f(\xi)\ \forall \xi \in G_y\ \ and$$
$$\exists A \in \mathrm{top}(\mathbb{R}^\ell);\ \ \overline{G_y} \subset A,\ \ f(y) > f(\xi)\ \ \forall \xi \in (A \cap \mathcal{D}) \setminus G_y. \tag{10}$$

2. Each local maximizer y to f in $\mathcal{D}$ will be called a global maximizer if $f(y) = f(x_{max})$.

3. Set G_y will be called the maximum manifold to f in $\mathcal{D}$. The family of all maximum manifolds to f and $\mathcal{D}$ will be denoted by $\mathrm{MaxManifolds}_{f,\mathcal{D}}$.

4. Each maximizer $y \in \mathcal{D}$ to f in $\mathcal{D}$ will be called isolated if $G_y = \{y\}$.

Maximal manifolds share a lot of properties of minimal manifolds due to the symmetry of both definitions.

Theorem 13.

(a) Let y be a local maximizer of f in $\mathcal{D}$, $f(y) = a$. Then, B_y is a connected component of $(f|_\mathcal{D})^{-1}(a)$ that contains y.

(b) Assume the connected components of $(f|_\mathcal{D})^{-1}(a)$ form a locally finite family for each $a \in \mathbb{R}$. Let $y \in \mathcal{D}$, $f(y) = a$ and let C_y be a connected component of $(f|_\mathcal{D})^{-1}(a)$ containing y. If there exists an open set $A \supset C_y$ with $f(\xi) \leq a$ for each $\xi \in A \cap \mathcal{D}$, then y is a local maximizer of f.

Proof. This is the same as the proofs of Theorem 5 and 8 with flipped inequalities between $f(\xi)$ and $f(y)$ (since such inequalities are only used to conclude that $\xi \notin C_y$). $\square$

Let us consider local maximizer z that belongs to some maximum manifold G_y.

Definition 14. *Upland $\mathcal{U}_z$ to f in $\mathcal{D}$, associated with local maximizer $z \in G_y$ is the maximum open-connected set (in the sense of inclusion) so that $\mathcal{U}_z \subset G_y$ and $z \in \mathcal{U}_z$ or the empty set. Let us denote by $\mathrm{Uplands}_{f,\mathcal{D}}$ the family of all uplands to f in $\mathcal{D}$.*

Observation 15. *If f is a differentiable function, then (as in the case of minimum manifolds), for every maximum manifold B_y and for every $x \in B_y$ we have*

$$\nabla f(x) = 0.$$

If, moreover, f is twice differentiable, then Hessian matrix $Hf(x)$ is negative semi-definite.

3.4. Stationary manifolds and shelves

Let $x \in \mathcal{D}$ and $c \in [f(x_{min}), f(x_{max})]$ so that $f(x) = c$.

Definition 16. $H_x \subset \mathcal{D}$ *will be called the stationary manifold associated with* $x \in \mathcal{D}$ *if it is the maximum connected set (in the sense of inclusion) satisfying* $H_x \subset (f|_{\mathcal{D}})^{-1}(c)$, $x \in H_x$ *and* $H_x \setminus \{x\} \neq \emptyset$.

Definition 17. $\mathcal{S}_x \subset \mathcal{D}$ *will be called the shelf to* f *in* $\mathcal{D}$ *associated with* $x \in \mathcal{D}$ *if it is the maximum open-connected set (in the sense of inclusion) satisfying* $\mathcal{S}_x \subset H_x$, $x \in \mathcal{S}_x$, *and* $\mathcal{S}_x \notin (\mathrm{Lowlands}_{f,\mathcal{D}} \cup \mathrm{Uplands}_{f,\mathcal{D}})$. *Let us denote by* $\mathrm{Shelves}_{f,\mathcal{D}}$ *the family of all shelves to* f *in* $\mathcal{D}$.

Observation 18. *It is easy to show that, if* f *is differentiable, then for every shelf* $\mathcal{S}_y$ *and every* $x \in \mathcal{S}_y$, *we have*

$$\nabla f(x) = 0.$$

The observation above is generally not true for an arbitrary stationary manifold. If, however, a stationary manifold that is neither a minimum nor maximum manifold contains a point with a vanishing gradient (i.e., a critical point), then the Hessian matrix at this point can have both positive and negative eigenvalues, which means that it is a saddle point *for* f.

Remark 19. *Observations 11, 15, and 18 state that minimum manifolds, maximum manifolds, and shelves are composed of critical points for* f.

Observation 20. *Stationary manifolds are disjoint since they are connected components of the sets of form* $(f|_{\mathcal{D}})^{-1}(c)$.

Observation 21. *In particular, minimum and maximum manifolds are disjoint.*

3.5. Steepest- and strictly-descent local optimization methods

Definition 22. *Local optimization method* loc *will be called* steepest-descent *if it generates sequence* $\{x_i\}_{i=0,1,2,\ldots} \subset \mathcal{D}$ *so that*

$$x_{i+1} = x_i + \xi_i \mathrm{n}_i; \quad i = 0, 1, 2, \ldots, \tag{11}$$

where x_0 *is the starting point,* $\mathrm{n}_i \in \mathbb{R}^\ell$, $\|\mathrm{n}_i\| = 1$ *the unit direction vector of the maximum slope of function* f *at* x_i, *and* $\xi_i \geq 0$ *is the i-th step length. Unit vector* n_i *does not exist if, for all* $y \in A$, $f(y) \geq f(x_i)$, *where* $A \in \mathrm{top}(\mathcal{D})$ *is open set* $A \in \mathrm{top}(\mathcal{D})$ *containing* x_i. *In this case,* $\xi_i = 0$, *and the method is stopped. We will denote by* $\mathrm{loc}(x_0)$ *the limit of approximate sequence* $\{x_i\}_{i=0,1,2,\ldots}$.

Remark 23. *The above definition implies that, if a steepest-descent method is started from* x_0 *such that neighborhood* A *of* x_0 *exists such that*

$$f(y) \geq f(x_0) \quad \text{for all } y \in A, \tag{12}$$

then the method stops immediately. Alternatively, we can say that it generates a constant minimizing sequence. Anyway, in such a case, we have that

$$\mathrm{loc}(x_0) = x_0.$$

Note that condition (12) holds for the local minimizers but also for any element of a lowland, shelf, or upland.

If $f \in C^1(\mathcal{D})$, then we can set $\mathrm{n}_i = -\|\nabla f(x_i)\|^{-1} \nabla f(x_i)$. The gradient operator might be replaced by the Gâteaux derivative if function f is only differentiable in the Gâteaux sense. In the case of continuous functions $f \in C^0(\mathcal{D})$ that are not Gâteaux differentiable, the selection of the maximum slope unit vector would be difficult or even impossible. For such cases, we suggest the class of local methods that are approximately steepest-decent.

Definition 24. *Local method* loc *will be called* α-*steepest-descent if it generates the minimizing sequence of form (11) and the unit direction vector of the maximum slope of function f at x_i will be set as*

$$\mathrm{n}_i = \alpha^{-1}(w - x_i); \;\; w \in O_{\alpha,x_i} = \{y \in \mathcal{D}; \|y - x_i\| = \alpha\} \; where$$

$$f(x_i) - f(w) = \max_{y \in O_{\alpha,x_i}} \{f(x_i) - f(y)\} \; and \tag{13}$$

$$\forall \beta, \gamma; \;\; 0 \le \beta < \gamma \le 1 \;\; f(x_i + \beta(w - x_i)) > f(x_i + \gamma(w - x_i)).$$

If such a n_i does not exist, then the method is stopped. Real constant $\alpha > 0$ is the parameter of this definition.

Remark 25. *The unit direction vector of maximum slope n_i of function f at x_i given by formula (13) is well-defined because continuous function $f(x_i) - f(\cdot)$ reaches its maximum on compact O_{α,x_i}.*

We will distinguish a special subclass of steepest-descent local optimization methods (see Rinnooy, Kan, Timmer [20], Dixon Gomulka Szegö [8]).

Definition 26. *Steepest-descent (α-steepest-descent) local optimization method* loc *will be called* strictly-steepest-descent *(α-strictly-steepest-descent) on $\mathcal{D}$ if, for each starting point $x_0 \in \mathcal{D}$, it generates minimizing sequence $\{x_i\}_{i=0,1,2,\dots} \subset \mathcal{D}$ according to rule (11) so that*

$$\forall i = 0, 1, 2 \dots \forall \beta, \gamma; \; 0 \le \beta < \gamma \le \xi_i \Rightarrow f(x_i + \gamma \mathrm{n}_i) < f(x_i + \beta \mathrm{n}_i). \tag{14}$$

We will denote by loc(x_0) the limit of the minimizing sequence generated by the local strictly-steepest-descent method (α-strictly-steepest-descent method) loc started from $x_0 \in \mathcal{D}$.

Remark 27. *Let us note that, if starting point x_0 satisfies assumption (12), then a strictly-steepest-descent method stops at the first step. Therefore, Remark 23 remains true in this case.*

3.6. Sets and basins of attraction

Definition 28. *Let* loc *be the strictly-steepest-descent (α-strictly-steepest-descent) local optimization method on $\mathcal{D}$ and y the local minimizer of f in $\mathcal{D}$. Set*

$$R_y^{\mathrm{loc}} = \{x \in \mathcal{D}; y = \mathrm{loc}(x)\} \tag{15}$$

will be called the set of attraction of y with respect to method loc. *We will frequently simplify its notation as R_y.*

Remark 29. *Local method* loc *is a parameter of the definition. Although inconvenient, it generally cannot be omitted. One important particular situation when, in fact, we can get rid of this parameter is the case of a C^1 function f. Then, instead of the discrete dynamic system generated by* loc, *we can use the antigradient flow for f; i.e., the family of solutions of equation*

$$\gamma'(t) = -\nabla f(\gamma(t)). \tag{16}$$

Namely, we say that $x \in R_y$ if a positive constant ε exists and curve

$$\gamma : (-\varepsilon, +\infty) \longrightarrow \mathcal{D}$$

such that γ is a solution of (16) *with*

$$\gamma(0) = x, \quad \lim_{t \to +\infty} \gamma(t) = y.$$

Remark 30. *Remark 27 implies that, if y satisfies condition* (12), *then*

$$y \in R_y.$$

Proposition 31. *If f is C^1 and y is located in a lowland, shelf, or maximum manifold, then*

$$R_y = \{y\}.$$

Proof. First, assume that y is an element of a lowland or shelf. Then, neighborhood $A \ni y$ exists which is contained in the landform. For all $x \in A$, we have

$$\nabla f(x) = 0.$$

If there was a solution of (16) starting from a point different than y, then this solution would have common points with A; however, this is of course impossible since all elements of A are stationary points for (16) and any solution would get stuck in any of them. Therefore, R_y does not contain points other than y.

If y is an element of a maximum manifold, then it either has a neighborhood contained in that manifold or any its neighborhoods has common points with the exterior of the manifold. In the first case, we can proceed exactly as in the case of lowlands and shelves. In the latter case, there are points in a neighborhood of y with a value of f less than $f(y)$. But those points cannot belong to R_y because the solutions of (16) have decreasing values of f for increasing times. The other elements of the neighborhood are stationary points; hence, also in this case,

$$R_y = \{y\}.$$

□

Definition 32. *The set*

$$R_{B_y} = \bigcup_{x \in B_y} R_x$$

will be called the set of attraction of minimum manifold B_y.

Remark 33. *From Remark 30, it follows that, if B_y is a minimum manifold, then*

$$B_y \subset R_{B_y}.$$

Definition 34. *Basin of attraction $\mathcal{B}_y$ of an isolated minimizer y to function $f : \mathcal{D} \to \mathbb{R}$ is the connected part of set $\{x \in \mathcal{D}; f(x) < h_y\} \cap (R_y \cup \{y\})$ that contains y $(y \in \mathcal{B}_y)$, where $h_y = \inf\{f(z), \ z \in \partial \overline{R_y} \setminus \partial \mathcal{D}\}$.*

Definition 35. *Basin of attraction $\mathcal{B}_{B_y}$ of minimum manifold B_y to function $f : \mathcal{D} \to \mathbb{R}$ is the connected part of set $\{x \in \mathcal{D}; f(x) < h_y\} \cap (R_{B_y} \cup B_y)$ that includes B_y, where $h_y = \inf\{f(z), z \in \partial \overline{R_{B_y}} \setminus \partial \mathcal{D}\}$.*

Definition 36. *Set*

$$R_{\mathcal{P}_y} = \bigcup_{x \in \overline{\mathcal{P}_y}} R_x$$

will be called the set of attraction of lowland $\mathcal{P}_y$.

Remark 37. *From Remark 30, it follows that, if $\mathcal{P}_y$ is a lowland, then*

$$\mathcal{P}_y \subset R_{\mathcal{P}_y}.$$

Definition 38. *Basin of attraction $\mathcal{B}_{\mathcal{P}_y}$ of lowland $\mathcal{P}_y$ to function $f : \mathcal{D} \to \mathbb{R}$ is the connected part of set $\{x \in \mathcal{D}; f(x) < h_y\} \cap (R_{\mathcal{P}_y} \cup \overline{\mathcal{P}_y})$ that includes $\overline{\mathcal{P}_y}$, where $h_y = \inf\{f(z), z \in \partial \overline{R_{\mathcal{P}_y}} \setminus \partial \mathcal{D}\}$.*

Definition 39. *Let $\mathcal{S}_z \in \mathrm{Shelves}_{f,\mathcal{D}}$ be the shelf of graph f in $\mathcal{D}$; then,*

$$R_{\mathcal{S}_z} = \bigcup_{\xi \in \overline{\mathcal{S}_z}} \{x \in \mathcal{D}; \xi = \mathrm{loc}(x)\} \tag{17}$$

will be called the set of attraction of shelf $\mathcal{S}_z$, with respect to strictly-steepest-descent (α-strictly-steepest-descent) local optimization method loc.
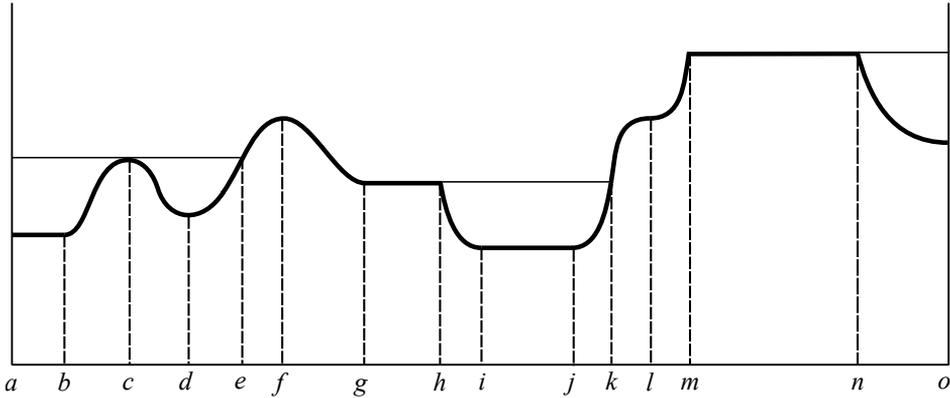
Definition 40. *In the case of C^1 function f, we can define the attraction set independently upon any local method. Namely, the attraction set of shelf $\mathcal{S}_z$ is the following set*

$$R_{\mathcal{S}_z} = \left\{ x \in \mathcal{D} \mid \exists \xi \in \overline{\mathcal{S}_z}, \gamma \text{ solution of } (16) : \gamma(0) = x, \ \lim_{t \in +\infty} \gamma(t) = \xi \right\}.$$

Remark 41. *From Remark 30, it follows that, if $\mathcal{S}_y$ is a shelf, then*

$$\mathcal{S}_y \subset R_{\mathcal{S}_y}.$$

The illustration of the various landforms in $\mathcal{D} \subset \mathbb{R}$ is shown in Figure 2.

Figure 2. Landforms in 1D: d, o are local minimizers, $(c, e) = \mathcal{B}_d$, $(c, f) = R_d$, $(n, o] = \mathcal{B}_o = R_o$; $(a, b) = \mathcal{P}_{(b-a)/2}$, $[a, c) = \mathcal{B}_{\mathcal{P}_{(b-a)/2}}$, $[b, c) = R_{\mathcal{P}_{(b-a)/2}}$; $(i, j) = \mathcal{P}_{(j-i)/2}$, $[h, k] = \mathcal{B}_{\mathcal{P}_{(j-i)/2}}$, $[h, i] \cup [j, l) = R_{\mathcal{P}_{(j-i)/2}}$; $(g, h) = \mathcal{S}_{(h-g)/2}$, $(f, g] = R_{\mathcal{S}_{(h-g)/2}}$; $(m, n) = \mathcal{U}_{(n-m)/2}$

4. Landform separability and completeness of topological taxonomy

4.1. Separability of landforms

Lemma 42. *Shelves, lowlands, and uplands are maximal open-connected subsets of the connected components of the sets of form $f^{-1}(c)$.*

Proof. Lowlands (uplands) are, by definition, maximal open-connected subsets of the minimal (maximal) manifolds, and shelves are maximal open-connected subsets of the stationary manifolds. Minimal (maximal) manifolds are connected components of the sets of form $(f|_{\mathcal{D}})^{-1}(c)$ by Theorem 5. Stationary manifolds are such connected components by their definition.

$\square$

Theorem 43. *All shelves, lowlands, and uplands are pairwise disjoint.*

Proof. Let each U, V be a shelf, lowland, or upland. Suppose $U \cap V \neq \emptyset$. Then, f is constant on $O = U \cup V$, $f(x) = c$ for each $x \in O$. Both U and V are maximal open subsets of the connected components of $f^{-1}(c)$. Since O is open and connected, contains both U and V, and $f|_O \equiv c$, then by maximality $U = V = O$. $\square$

Lemma 44. *If $\mathcal{D}$ is connected and f is not constant on $\mathcal{D}$, no subset of $\mathcal{D}$ can be an upland and lowland simultaneously.*

Proof. Let U be lowland with $f|_U = c$ and B be the associated minimum manifold. By the definition of minimum manifold, open set $V_+ \supset B$ exists such that $f(\xi) > c$

for $\xi \in (V_+ \cap \mathcal{D}) \setminus B$. We know that B is a connected component of $(f|_{\mathcal{D}})^{-1}(c)$ for some value of c. If U is also an upland, the same reasoning implies that B is also the associated maximum manifold and open set $V_- \supset B$ with $f(\xi) < c$ exists for $\xi \in (V_- \cap \mathcal{D}) \setminus B$. Let $V = V_+ \cap V_-$. Then, V is open, and for each $\xi \in (V \cap \mathcal{D}) \setminus B$, we have both $f(\xi) > c$ and $f(\xi) < c$; so, $V \cap \mathcal{D} = B$. Since $B = \bar{B}$ (Observation 6), $A_1 = V \cap \mathcal{D}$ is closed. Since V is open, $A_2 = \mathcal{D} \setminus V$ is closed as well. Thus, $\mathcal{D} = A_1 \cup A_2$, and since $\mathcal{D}$ is connected and $B \subset A_1 \neq \emptyset$, we must have $A_2 = \emptyset$, $\mathcal{D} \subset V$, and consequently, $B = V \cap \mathcal{D} = \mathcal{D}$, so f is constant. $\qquad\square$

Theorem 45. *Let f be non-constant and U be a maximal open-connected subset of $\mathcal{D}$ such that $f|_U$ is constant. Then, U is either a lowland, shelf, or upland, and these possibilities are mutually exclusive.*

Proof. Since U is a maximal open-connected subset of a stationary manifold, if it is neither a lowland nor upland, it is necessarily a shelf (by definition). It cannot be both a shelf and a lowland or upland (again, by definition), and by Theorem 44, it cannot be both a lowland and an upland. $\qquad\square$

Theorem 46. *Let U be a lowland, upland, or shelf and H be a stationary manifold. If $U \cap H \neq \emptyset$, then $U \subset H$.*

Proof. By definition, U is a maximal open subset of some stationary manifold K. Since $K \cap H \supset U \cap H \neq \emptyset$, H and K are non-disjoint stationary manifolds, then, by Observation 20, we have $K = H$. Therefore $U \subset H$. $\qquad\square$

4.2. Separability of attraction sets and basins of attraction

Let us assume in this section that objective f is a C^1 function with the Lipschitz--continuous gradient in admissible domain $\mathcal{D}$.

Theorem 47. *Let A and B be different sets, each belonging to one of the following categories (the same or different):*

- *minimum manifolds,*
- *stationary manifolds,*
- *maximum manifolds.*

Then, $R_A \cap R_B = \emptyset$. Notice that, for $A \in \mathrm{MaxManifolds}_{f,\mathcal{D}}$, we may identify R_A with A.

Proof. First, from the Picard–Lindelöf theorem, we know that, in our case, every point from $\mathcal{D}$ belongs to exactly one trajectory of (16). Of course, this trajectory can converge to one point at most. Therefore, for each $y, y' \in \mathcal{D}$ such that $y \neq y'$, we have

$$R_y \cap R_{y'} = \emptyset. \tag{18}$$

If C is a minimum or stationary manifold, then the definition of attraction set implies that

$$R_C = \bigcup_{y \in C} R_y.$$

Maximum manifolds consist of stationary points for (16) that do not have nontrivial attraction sets. Thus, for any z belonging to a maximum manifold, we have

$$R_z = \{z\}.$$

Hence, if C is a maximum manifold, then we have

$$R_C = \bigcup_{z \in C} \{z\} = \bigcup_{z \in C} R_z.$$

From the previous sections, we know that, if A and B satisfy the assumptions of this theorem, then

$$A \cap B = \emptyset.$$

Therefore, we obtain

$$R_A \cap R_B = \left(\bigcup_{y \in A} R_y \right) \cap \left(\bigcup_{z \in B} R_z \right) = \emptyset.$$

$\square$

Remark 48. *The above theorem is generally not true if A and B are lowlands, shelves, or uplands because the closures of these landforms can have common points as in Figure 3, which in turn can provide quite large common parts of the attraction sets. Nevertheless, using the same reasoning as in the proof of the theorem, we can obtain the following results:*
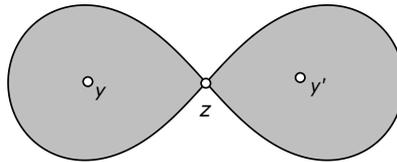
Figure 3. Two separate lowlands $\mathcal{P}_y, \mathcal{P}_{y'}$ in 2D sharing point z in their closures

Proposition 49. *Let both A and B be a lowland or shelf. Assume additionally that $\overline{A} \cap \overline{B} = \emptyset$. Then, $R_A \cap R_B = \emptyset$.*

Proof. First, observe that the respective definitions imply that, if C is a lowland or shelf, then

$$R_C = \bigcup_{y \in \overline{C}} R_y.$$

Therefore, similar to the proof of the preceding theorem, using assumption $\overline{A} \cap \overline{B} = \emptyset$, we obtain

$$R_A \cap R_B = \left(\bigcup_{y \in \overline{A}} R_y \right) \cap \left(\bigcup_{z \in \overline{B}} R_z \right) = \emptyset.$$

$\square$

Proposition 50. *Let $A, B \subset \mathcal{D}$, then $\mathcal{B}_A \cap \mathcal{B}_B = \emptyset$ if one of the following conditions hold:*

1. *$A, B \in \text{MManifolds}_{f,\mathcal{D}}$ and $A \neq B$,*

2. *$A, B \in \text{Lowlands}_{f,\mathcal{D}}$ and $\overline{A} \cap \overline{B} = \emptyset$,*

3. *$A \in \text{MManifolds}_{f,\mathcal{D}}$ and $B \in \text{Lowlands}_{f,\mathcal{D}}$ and $A \cap \overline{B} = \emptyset$.*

Notice that category $\text{MManifolds}_{f,\mathcal{D}}$ also contains isolated local minimizers to f on $\mathcal{D}$.

Proof. Since we know that $\mathcal{B}_A \subset R_A$ and $\mathcal{B}_B \subset R_B$, the first part of the thesis follows from Theorem 47. The second part of the thesis follows from Proposition 49 in a similar way. If B is a lowland such that $A \cap \overline{B} = \emptyset$, then a minimum manifold C exists for which $C \neq A$ and $B \subset C$. Moreover, from Definitions 38 and 36, 32, we have

$$\mathcal{B}_B \subset R_B = \bigcup_{x \in \overline{B}} R_x \subset \bigcup_{x \in C} R_x = R_C.$$

The last part of the thesis follows then again from Theorem 47. $\square$

5. Conclusions and perspective of applications

This paper delivers exhaustive taxonomies of both the objective characteristics (local approach) as well as the specific subsets of the candidate solutions (landform taxonomy) associated with ill-posed GOPs. The taxonomies are mainly motivated and devoted (but not restricted) to the problems appearing in the inverse parametric analysis, which consists of restoring the parameters of the partial differential equations and variational partial differential equations. Such problems are formulated as global optimization ones, where the objective is to minimize the misfit between the measured and simulated forward solutions.

Special attention was paid to the problems in which the objective (misfit) is insensitive with respect to at least one decision variable (γ-insensitive problems C6 and C7) and the problems in which the objective (misfit) is insensitive with respect to all of the decision variables on the dense subsets of the positive measure (γ-unspecified problems C8 and C9).

We introduced the topological definitions of some interesting sets appearing in the objective landscape imposed by the problems mentioned above as minimum manifolds, lowlands, stationary manifolds, shelves, and their sets of attraction and basins of attraction as well as maximum manifolds and uplands. Both cases of the continuous and continuously differentiable objective function are simultaneously considered.

We have proven several important features that ensure the quality of the topological taxonomy. In particular, lowlands, uplands, and shelves are pairwise disjoint (see theorem 43). Moreover, theorem 45 shows the completeness of this taxonomy; i.e., lowlands, uplands, and shelves are the only maximal open-connected proper subsets in the admissible domain interior for which the objective takes a constant value.

The separability of the sets and basins of attraction were studied for the case of continuously differentiable objectives only. In particular, it was proven that the local minimizers as well as the minimum and stationary manifolds have disjoint sets of attractions (see Theorem 47). A similar result was obtained for the sets of attraction of the lowlands and shelves (see Proposition 49). Also, the basins of attraction of the local minimizers, lowlands, and shelves are pairwise disjoint under some mild assumptions (see Proposition 50). All of these sets are disjoint with uplands.

The taxonomies are helpful for the more precise definition of ill-conditioned GOPs and inverse parametric problems as well, being the necessary introductory part to each multimodal analysis [19]. In particular, they allow for defining lowlands as separate admissible subregions of low sensitivity, which might represent the common behavior of the analyzed system, different to those represented by the parameter vectors belonging to some other lowland.

Moreover, the number and size of the uplands and shelves (being deceptive regions) are important factors that allow us to evaluate the problem difficulty. Such information might be helpful for creating and profiling the complex strategies that solve such problems, boosting the recognition of minimum manifolds, lowlands, and their basins of attraction as well as allowing us to leave and omit uplands and shelves.

The introduced class of the α-steepest-descent local optimization methods (see Definition 24) can be helpful for defining the sets and basins of attractions in an approximated way (with the accuracy controlled by parameter α). It might be utilized in the case of objectives for which the maximum slope could not be explicitly determined; e.g., by computing the objective gradient in an exact or approximated way. This new idea seems to be disputable and needs further research (in particular, to study the a priori evaluation of the landform identification error in the Hausdorff norm).

The above theory supports the stochastic strategies of multimodal analysis gathered in [14, 19]. It may be also used to study the features of a Clustered Genetic Search (CGS) [21, 27] and Hierarchic Memetic Strategy (HMS) dedicated for approximating lowlands and their basins of attraction [15, 16]. The solving of real-world engineering ill-posed inverse problems from Classes C.6 – C.9 by the methods mentioned above is presented in papers [2, 3, 11, 22, 23].

Acknowledgements

References

[1] Addis B.: *Global Optimization Using Local Searches*. Ph.D. thesis, Universitá Degli Studi di Firenze, 2004.

[2] Barabasz B., Gajda-Zagórska E., Migórski S., Paszyński M., Schaefer R., Smołka M.: A hybrid algorithm for solving inverse problems in elasticity, *International Journal of Applied Mathematics and Computer Science*, vol. 24(4), pp. 865–886, 2014.

[3] Barabasz B., Migórski S., Schaefer R., Paszyński M.: Multi-deme, twin adaptive strategy hp-HGS, *Inverse Problems in Science and Engineering*, vol. 19(1), pp. 3–16, 2011.

[4] Beilina L., Klibanov M.V.: *Approximate Global Convergence and Adaptivity for Coefficient Inverse Problems*. Springer, 2012. `http://dx.doi.org/10.1007/978-1-4419-7805-9`.

[5] Boender C.G.E., Rinnooy-Kan A.H.G., Timmer G.T., Stougie L.: A stochastic method for global optimization, *Mathematical Programming*, vol. 22, pp. 125–140, 1982.

[6] Cabib E., Davini C., Chong-Quing R.: A problem in the optimal design of networks under transverse loading, *Quarterly of Applied Mathematics*, vol. 48(2), pp. 251–263, 1990.

[7] Cotta C., Schaefer R.: Complex metaheuristics, *Journal of Computational Sciences*, vol. 17, pp. 171–173, 2016. `http://dx.doi.org/doi:10.1016/j.jocs.2016.06.001`.

[8] Dixon L.C.W., Szegö G.P. (eds.): *Toward Global Optimization*, North Holland, 1975.

[9] Dixon L.C.W., Szegö G.P. (eds.): *Towards Global Optimisation 2*, North-Holland, Amsterdam, 1978.

[10] Duan Q., Sorooshian S., Gupta V.: Effective and efficient global optimization for conceptual Rainfall–Runoff models, *Water Resource Research*, vol. 28(4), pp. 1015–1031, 1992.

[11] Gajda-Zagórska E., Schaefer R., Smołka M., Paszyński M., Pardo D.: A hybrid method for inversion of 3D DC logging measurements, *Natural Computing*, vol. 14(3), pp. 355–374, 2014. `http://dx.doi.org/10.1007/s11047-014-9440-y`.

[12] Isshiki M., Sinclair D., Kaneko S.: Lens design: Global optimization of both performance and tolerance sensitivity. In: *International Optical Design*, p. TuA5. Optical Society of America, 2006. `http://dx.doi.org/10.1364/IODC.2006.TuA5`.

[13] Kabanikhin S.I.: Definitions and examples of inverse ill-posed problems, *Journal of Inverse and Ill-Posed Problems*, vol. 16, pp. 317–357, 2008. `http://dx.doi.org/10.1515/JIIP.2008.069`.

[14] Koper K., Wysession M., Wiens D.: Multimodal function optimization with a niching genetic algorithm: A seismological example, *Bulletin of the Seismological Society of America*, vol. 89(4), pp. 978–988, 1999.

[15] Łoś M., Sawicki J., Smołka M., Schaefer R.: Memetic approach for irremediable ill-conditioned parametric inverse problems. In: *Procedia Computer Science*, vol. 108, pp. 867–876, Elsevier, 2017. `http://dx.doi.org/10.1016/j.procs.2017.05.007`.

[16] Łoś M., Schaefer R., Sawicki J., Smołka M.: Local misfit approximation in memetic solving of ill-posed inverse problems. In: Squillero G., Sim K. (eds.), *Applications of Evolutionary Computation. EvoApplications 2017*, Lecture Notes in Computer Science, vol 10199, pp. 297–309. Springer, 2017.

[17] Pardalos P., Romeijn H, (eds.): *Handbook of global optimization*, vol. 2, Springer US, 2002. `http://dx.doi.org/10.1007/978-1-4757-5362-2`.

[18] Paruch M., Majchrzak E.: Identification of tumor region parameters using evolutionary algorithm and multiple reciprocity boundary element method, *Engineering Applications of Artificial Intelligence*, vol. 20(5), pp. 647–655, 2007.

[19] Preuss M.: *Multimodal optimization by means of evolutionary algorithms*, Natural Computing. Springer, 2015.

[20] Rinnooy-Kan A.H.G., Timmer G.T.: Stochastic global optimization methods. Part 1: Clustering methods, *Mathematical Programming*, vol. 39, pp. 27–56, 1987.

[21] Schaefer R., Adamska K., Telega H.: Genetic clustering in continuous landscape exploration, *Engineering Applications of Artificial Intelligence (EAAI)*, vol. 17, pp. 407–416, 2004.

[22] Smołka M., Gajda-Zagórska E., Schaefer R., Paszyński M., Pardo D.: A hybrid method for inversion of 3D AC logging measurements, *Applied Soft Computing*, vol. 36, pp. 422–456, 2015.

[23] Smołka M., Schaefer R., Paszyński M., Pardo D., Álvarez-Aramberri J.: An agent-oriented hierarchic strategy for solving inverse problems, *International Journal of Applied Mathematics and Computer Science*, vol. 25(3), pp. 483–498, 2015. `http://dx.doi.org/10.1515/amcs-2015-0036`.

[24] Tikhonov A.N., Goncharskii A., Stepanov V.V., Yagola A.G.: *Numerical methods for the solution of ill-posed problems*, Springer-Verlag, 1995.

[25] Törn A.: A search clustering approach to global optimization. In: Dixon L.C.W., Szegö G.P. (eds.), *Towards Global Optimization 2*, North-Holland, Amsterdam, pp. 49–62, 1978.

[26] Törn A., Ali M.M., Viitanen S.: Stochastic Global Optimization: Problem Classes and Solution Techniques, *Journal of Global Optimization*, vol. 14, pp. 437–447, 1999.

[27] Wolny A., Schaefer R.: Improving population-based algorithms with fitness deterioration, *Journal of Telecommunications and Information Technology*, vol. 4, pp. 31–44, 2011.

[28] Zeidler E.: *Nonlinear functional analysis and its application. II/A: Linear monotone operators*, Springer, 2000.

Affiliations

Marcin Łoś
   AGH University of Science and Technology, Krakow, Poland, marcin.los.91@gmail.com

Maciej Smołka
   AGH University of Science and Technology, Krakow, Poland, smolka@agh.edu.pl

Robert Schaefer
   AGH University of Science and Technology, Krakow, Poland, schaefer@agh.edu.pl

Jakub Sawicki
   AGH University of Science and Technology, Krakow, Poland, jsawicki@agh.edu.pl