A COARSE SOLUTION OF GENERALIZED SEMI-INFINITE OPTIMIZATION PROBLEMS VIA ROBUST ANALYSIS OF MARGINAL FUNCTIONS AND GLOBAL OPTIMIZATION.

vorgelegt von

Dipl.-Math. Abebe Geletu W. Selassie

DISSERTATION

zur Erlangung des akademischen Grades

Dr. rer. nat.

Eingereicht bei der Technischen Universität Ilmenau

Juni, 2004

Tag der wissenschaftlichen Aussprache: 17.12.2004

1. Gutachter: Prof. Dr. rer. nat. habil. Armin Hoffmann.
2. Gutachter: Prof. Dr. rer. nat. habil. Gerhard-Wilhelm Weber
3. Gutachter: Prof. Dr. rer. nat. habil. Rainer Tichatschke

urn:nbn:de:gbv:ilm1-2004000167

March 22, 2005

To my wife Genet, and my son Abenezer and my daughter Bethel whose love, patience and understanding carried me so far.

March 22, 2005

Table of Contents

Acknowledgment

First and foremost, my greatest indebtedness goes to my academic advisor Prof. Armin Hoffmann. He has spared none of his time in extending his guidance and counselling - be it academic or otherwise. Without his constant supervision and encouragement this work would not have been a reality.

I would also like to express my deepest gratitude to Dr. Rüdiger Reinhard and Prof. Silvia Vogel for their profound amicability, sociability and sympathy. In the same vain, I also express my acknowledgment to Prof. Reinhardt Deumlich (of Addis Abeba University, Dept. of Mathematics), who is always behind me wherever I go. My thanks are also due to Dr. Tibor Csendes, for facilitating my stay at the Department of Computer Science, University of Szeged, Hungary, in the summers of 2001 & 2002, through the DAAD-MÖB scholarship.

I also take this opportunity to thank Dr. Axel Wolf, Mr. Hans Müller, Mrs. Erika Voigt and several other administrative and academic members of TU-Ilmenau, for their good will, kindness and for setting me up with a good social and academic atmosphere.

At last, there should always be a means to an end. The means has been graciously provided to me by DAAD. Thus, my utmost appreciation also goes to the German Academic Exchange Service (DAAD) for the financial, material and moral support I received for the successful completion of my study.

Abebe Geletu.

March 22, 2005

List of Symbols

General Symbols and Conventions

$\mathbb{N}$	set of the natural numbers
$\mathbb{Q}$	set of the rational numbers
$\mathbb{R}$	set of the real numbers
$\mathbb{R}^n$	space of real vectors of dimension n
$A, \dots, Z$	sets in a topological space
$\emptyset$	empty set
$a, \dots, z$	real numbers, vectors
(a_k)	real number sequences or nets
a_j	j-th component of a vector a or
	j-th element of a number sequence (a_k)
a^i	i-th vector (of a set of vectors)
$\Pi_{i=1}^m A_i$	Cartesian product of the sets $A_1 \dots A_m$
$\|x\|$	absolute value of x
$\\|x\\|$	the Euclidean norm of x
$x^\top$	the transpose of a vector x
$\{\}$	set braces
$(\,), \{\,\}, [\,]$	parenthesis for alternative expressions
(a, b), $[a, b)$, $[a, b]$	open, half-open and closed intervals
$\geq$, $>$, etc.	component wise comparative relations for vectors in $\mathbb{R}^n$
$f : A \to B$	function f from the set A into the set B
f^{-1}	inverse of the function f

$a \in A$	a belongs to (an element of) the set A
$A \cup B, A \cap B$	union, intersection of the sets A and B
$A \setminus B$	set A without the elements of set B
$int\,(A)$	interior of the set A
$\partial\, A$	boundary of the set A
$cl\,(A)$	the closure of the set A
$\min A$, $\max A$	minimum and maximum element of the set A
$\inf A$, $\sup A$	infimum and supremum of the set A
$dist(x, A)$	distance from the point x to the set A
$x \to y$	convergence of x to y, or x tends to y
$dom\,(\varphi)$	domain of the mapping φ
$Graph\,(\varphi)$	graph of the mapping φ
M^{+1}, M^{-1}	upper and lower inverse of a set-valued map, resp.
$\partial\varphi$	sub-differential of the function φ
$D_xG(x, y, t)$	a row (column) vector for the partial derivative
	of a function of several variables w.r.t. x
$\Box$	end of a proof

Abbreviations

cf.	confer (meaning refer, compare or see)
et al.	and others
w.r.t.	with respect to
e.t.c.	and so on
resp.	respectivley
l.s.c.	lower semi-continuous
u.s.c.	upper semi-continuous
l.r.	lower robust
u.r.	upper robust
l.a.	lower approximatable

$u.a.$	upper approximatable
SCQ	Slater Constraint Qualification
SSC	Strong Slater Condition
$MFCQ$	Mangasarian-Fromovitz constraint qualification
$EMFCQ$	Extended Mangasarian-Fromovitz Constraint Qualification
SNH	Semi-neighbourhood
$SV-map$	Set-Valued Map
(PSIP),(GSIP),(SIP)	parametric, Generalized, semi-infinite optimization problems, resp.
(BL)	Bi-level optimization problem
$IGOM$	the Integral Global Optimization Method

Reserved Symbols and Notations

$\mathcal{M}$	feasible set of a generalized semi-infinite programming problem
X, Y, T	topological spaces
$M: X \rightrightarrows Y$	a set valued map from set X to set Y
$M(\cdot)$	a set valued map
$int_B A$	relative interior of the set A in the set B
$\mathbf{B}_\varepsilon$	the ball of radius $\varepsilon > 0$ and center at 0
$ess \inf f$	essential infimum of a function f
$h = (h_1, \ldots, h_p)$	a vector of real valued functions
$[f \leq \alpha]$	lower level set of the function f at the level α
μ	Lebesgue-measure in the space $\mathbb{R}^n$
$\alpha, \beta, \delta, \theta, \gamma, \varepsilon, \tau, \mu$	real parameters
φ	marginal function of a (PSIP)
Λ, Ω	sets of parameters or indices
$(\text{GSIP})_{red}$	finite reduced (GSIP)

Introduction

In recent years, several practical engineering, approximation, optimal control and discrete optimization problems have been found to lead into a class of optimization problems known as *generalized semi-infinite optimization* (GSIP) problems. Among such problems are: maneuverability of a robot Grettinger-Krogh [20], time-optimal control Krabs [40], reverse Chebychev approximation Hoffmann & Reinhard [29], terminal variational problems Kaplan & Tichatschke [35], optimal control of discrete structures Weber [83, 87] etc. Hence, there arose the need for both theoretical investigation and numerical solution methods for (GSIP)s.

It seems, currently, that the theoretical aspects of (GSIP) are maturely developed. Thus, we find results on optimality conditions (both first and second order), studies on the structure of the feasible set, regularity and stability, local reducibility and connection of (GSIP) with other optimization models. It is observed that, under general assumptions, the feasible set of a (GSIP) may not be closed, not convex, not connected, even could exhibit some disjunctive structures, may also have re-interant corner points, etc; several of which are strange to the feasible set of a *standard semi-infinite optimization (SIP) problem*. These are sticking points for existing numerical methods of (SIP) (or their generalizations) to be applied for the treatment of (GSIP). But, recently, we have a successful approach being proposed, by Stein & Still [74, 75, 77, 78], in tackling a class of (GSIP) with convexity structures. However, so much work still remains to be done to come up with computational methods that could handel a general class of (GSIP)s. Hence, research is still in progress in tackling problems of somehow generalized nature. Therefore, the presentation made in this manuscript might be taken as one of such attempts.

The problem that is to be considered here has the form:

$$
(GSIP) \qquad \begin{aligned} & f(x) \to inf \\ & s.t. \\ & \quad G(x,t) \geq 0,\ \forall t \in B(x) \\ & \quad x \in X, \end{aligned}
$$

where

- $X \subset \mathbb{R}^n$ and $T \subset \mathbb{R}^m$ are compact sets in their respective topologies;
- the functions $f : X \to \mathbb{R}$; $G : X \times T \to \mathbb{R}$ are at least continuous;
- the index $B : X \rightrightarrows T$ is a set-valued map.

We will also take X and T as being topological or metric spaces as required and further specific assumptions on the problem data will be given as the discussion progresses. To avoid certain computational difficulties associated with the method suggested to solve (GSIP) and for simpler theoretical presentations (both from literature), we exclusively consider (GSIP) with inequality constraints. Moreover, the index set-valued map may have the structure:

$$B(x) := \{t \in T \mid h_i(x,t) \leq 0, i \in I\},$$

where I is a finite index set. The (GSIP) will reduce to a *standard semi-infinite optimization* (SIP) problem if the index set-valued map $B(\cdot)$ is a constant map, i.e. if $B(x) \equiv B \neq \emptyset, \forall x \in X$. What actually makes (GSIP) a generalization of (SIP) is that it is an optimization problem over a finite dimension (since $x \in \mathbb{R}^n$) w.r.t. infinite number of constraints, where *the (usually) infinite index set $B(\cdot)$ of the constraints varies w.r.t. x*. A comprehensive coverage of issues related to standard (SIP) could be found in

[19, 25, 51, 52].

In [33] it has been shown that the feasible set of (GSIP) is closed if the index set-valued map $B(\cdot)$ is lower semi-continuous. This has been one of the major assumptions in the numerical algorithms proposed in [47, 49, 74, 75, 77, 78, 79, 80]. However, in this manuscript two approaches are presented without exclusively making such an assumption. At the same time, the proposed approach will be seen to include problems with the 'nice' structures.

The two approaches discussed in Chap. 4 are: *a conceptual penalty method* (cf. Abebe and Hoffmann [1]) and *an exact penalty discretization method*. In the first approach, there is defined a discontinuous penalty function for the (GSIP) based on the marginal function of a certain auxiliary parametric semi-infinite optimization problem (PSIP) . The solution of the resulting penalized problem is shown to yield a *generalized minimizer* of the (GSIP). In the second approach, we define two penalty functions: one based on the marginal function of the lower level problem and, a second, based on the feasible set of (GSIP). By introducing a descretization of the index map $B(x)$, it has been verified that the discretized penalty problems could give an upper and a lower bound for the optimal value of the (GSIP), in the limit.

In both the proposed approaches we need to be able to solve discontinuous optimization problems. For this, the aim is to use the *Integral Global Optimization Method* (IGOM). The (IGOM) was first initiated by Chew and Zheng [13, 88] and was further developed by Hoffmann, Phú and Hichert [26, 27, 28, 48]. In fact, Hichert [26] wrote a program code, which he called (BARLO), for a generalized version of (IGOM) and (BARLO) is found to be computationally efficient for global optimization problems with *fairly discontinuous* data. Where the fairness of the discontinuity is characterized by the notions of *robustness*. Since the (IGOM) has its theoretical root in measure theory and robust analysis, to use (BARLO), we needed to verify some robustness and measurability properties of the involved marginal functions and that of the feasible set of the (GSIP). To this end, Chap. 3 presents relatively new results on robustness of marginal functions and of set-valued maps with given structures. The results obtained in Chap. 3 have been effectively applied for the proposed computational approaches in Chap. 4. Furthermore, to show the validity of the methods, certain computational experiments are also presented.

The material of the document has been organized into five chapters. Chap. 1 presents a review of ideas from set-valued analysis, emphasizing on set-valued maps with given structures; Chap. 2 contains a review on the current state of the art of (GSIP). Chaps. 3 & 4 make up the bulk of the manuscript. Specifically, in chapter 3, we will extend the theory of robust analysis and, in Chap. 4, we show how the ideas of Chap. 3 could be applied to solve some class of (GSIP)s. Chapter 5 presents some computational experiments which are carried through the descretization method.

Chapter 1

Facts from Set-Valued Analysis

Set-valued maps (SV-maps)[1] are indispensable tools in optimizations problems whose problem data depend on some parameter(s). They are basically used in characterizing the variation of feasible sets and/or stationary points, of optimization problems, with respect to a chosen or a prescribed set of parameters. This is particularly the case in parametric optimization; hence, in generalized semi-infinite optimization as well. In fact, both theoretical and numerical investigations of (GSIP) are based on properties and assumptions made on the set-valued map $B(\cdot)$. We, thus, find here a brief review of those definitions and properties which are relevant to the forthcoming discussions. Leaving out the detailed features to the specialized literature, we are mainly interested here in the basic issues of continuity. Furthermore, special emphasis is given to those maps which are defined using a parametric family of functions. These are known as *set-valued maps with given structures*. For details on general issues of set-valued maps one is referred to the books of Aubin and Cellina [4], Aubin and Frankowska [5], Berge [8], Hu and Papagorgious [32], Rockafellar and Wets [57], etc. And set-valued maps with given structures are also given due treatment in the book of Bank *et al.* [7], in the paper of Hogan [31], etc.

[1]The terms 'set-valued map'Aubin and Frankowska [5],'point-to-set map'Hogan [31],'correspondences' Aliprantis and Boder [3] and 'multi-valued maps'Robinso [53, 54] are usually used interchangeably; while the first being frequently used in current literature.

1.1 Set-Valued Maps - General Definitions and Properties

In the following, unless explicitly specified, the spaces X and T are taken to be Hausdorff topological spaces.

Definition 1.1.1. A set-valued map $B(\cdot)$ from X into T, written $B : X \rightrightarrows T$, is a relation that associates with every $x \in X$ a subset $B(x)$ of T. The *domain* of $B(\cdot)$, denoted by $Dom(B)$, is defined as:

$$Dom(B) := \{x \in X | B(x) \neq \emptyset\};$$

and the *graph* of $B(\cdot)$, dented by $Graph(B)$, is defined as

$$Graph(B) := \{(x,t) \mid t \in B(x), x \in Dom(B)\}.$$

Definition 1.1.2 (lower inverse of a SV-map, [5]). Let $B : \ X \rightrightarrows T$.

1. For any $\bar{t} \in T$, the *(lower) inverse image* of $\bar{t}$ under $B(\cdot)$ is defined as:

$$B^{-1}(\bar{t}) := \{x \in X \mid \bar{t} \in B(x)\}.$$

2. For any $V \subset T$ the (lower) inverse image of V under $B(\cdot)$ is denoted by $B^{-1}(V)$ and is defined as:

$$B^{-1}(V) := \{x \in X \mid B(x) \cap V \neq \emptyset\} = \bigcup_{t \in V} B^{-1}(t).$$

Definition 1.1.3 (upper inverse of a SV-map, [5]). Let $B : \ X \rightrightarrows T$. Then for any $S \subset T$, the upper inverse of S by $B(\cdot)$, denoted by $B^{+1}(S)$, is defined as:

$$B^{+1}(S) := \{x \in X \mid B(x) \subset S\}.$$

The terminologies *lower*- and *upper*-inverse are from Berge [8]; while in the book of Aubin and Frankowska [5] the former is simply termed *inverse*, and $B^{+1}(S)$ is termed the *core* of the set S under $B(\cdot)$. Moreover, the above two definitions of inverses of a SV-map lead into two types of continuities - upper and lower semi-continuity.

Definition 1.1.4 (upper semi-continuous SV-map).

Let $B:\ X \rightrightarrows T$ and $Dom(B) \neq \emptyset$. Then $B(\cdot)$ is said to be upper semi-continuous (u.s.c) at $x^0 \in X$ iff for any open set $V \subset T$, where $B(x^0) \subset V$, there exists a neighborhood $U \subset X$ of x^0 such that

$$\forall x \in U:\ B(x) \subset V,\ i.e.\ U \subset B^{+1}(V).$$

The map $B(\cdot)$ is said to be u.s.c. on X if it is u.s.c. at every $x \in X$.

Definition 1.1.5 (lower semi-continuous SV-map). Let $B:\ X \rightrightarrows T$. Then $B(\cdot)$ is said to be lower semi-continuous (l.s.c.) at $x^0 \in X$ iff for any $t^0 \in B(x^0)$ and any neighborhood $V \subset T$ of t^0, there exists a neighborhood $U \subset X$ of x^0 such that

$$\forall x \in U:\ B(x) \cap V \neq \emptyset;\ i.e.\ \ U \subset B^{-1}(V).$$

The map $B(\cdot)$ is said to be l.s.c. on X if $B(\cdot)$ is l.s.c. at every $x \in X$.

A set-valued map which is both lower and upper semi-continuous is called continuous. Furthermore, Def. 1.1.5 has an equivalent formulation using sequences, if both T and X are metric spaces.

Definition 1.1.6. Let X and T be metric spaces, $B:\ X \rightrightarrows T$ and $x^0 \in Dom(B)$. $B(\cdot)$ is said to be lower semi-continuous (l.s.c.) at x^0 iff for any $t^0 \in B(x^0)$ and for any sequence $\{x^k\} \subset Dom(B)$ such that $x^k \to x^0$, there exists a sequence $\{t^k\} \subset T$, where $t^k \in B(x^k)$ and $t^k \to t^0$.

Proposition 1.1.1 (Thm. 2.9, Kisielewicz [37]). *If both X and T are metric spaces, then Def. 1.1.5 and Def. 1.1.6 are equivalent.*

Definition 1.1.7 (closed SV-map, Hogan [31]). Let X and T be metric spaces and $B:\ X \rightrightarrows T$ be a set-valued map. Then $B(\cdot)$ is said to be closed at $x^0 \in X$ iff for sequences $\{x^k\}_{k\in\mathbb{N}}$ and $t^k \in B(x^k)$ such that $x^k \to x^0$ and $t^k \to t^0$ implies $t^0 \in B(x^0)$. Moreover, $B(\cdot)$ is called a *closed set-valued map* if it is closed at every point $x \in X$.

For $B : X \rightrightarrows T$, if, for each $t \in T$, $B(t)$ is a closed (compact or convex) set in X, then $B(\cdot)$ is called *closed (compact or convex) valued*. However, the notions: a *closed valued* SV-map and *a closed SV-map* are different. In the former, *closed* qualifies the values of the SV-map; while, in the latter, the SV-map itself.

Hence, a similar sequential characterization of an u.s.c. SV-map also exists if $B(\cdot)$ is assumed to be compact valued. In fact, such a characterization is of interest when the set-valued map is defined using a parametric system of functions, such as in the index set of a (GSIP).

Proposition 1.1.2 (Thm. 2.2, Kisielewicz [37]). *Let X and T be metric spaces. $B : X \rightrightarrows T$ is u.s.c. and compact valued iff for every $x^0 \in X$ and every sequence $\{x^n\} \subset X$, $x^n \rightarrow x^0$ and every sequence $\{t^n\} \subset T$, with $t^n \in B(x^n)$, there is a convergent subsequence $\{t^{n_k}\}$ of $\{t^n\}$ such that $t^{n_k} \rightarrow t^0 \in B(x^0)$.*

In accordance with the sequential characterization of upper semi-continuity given in Prop. 1.1.2, the following statement relates the closedness and the upper semi-continuity of a set-valued map.

Proposition 1.1.3. *Let X and T be metric spaces and $B : X \rightrightarrows T$ be a set-valued map. If $B(\cdot)$ is u.s.c. at x^0 and $B(x^0)$ is a closed set in X, then $B(\cdot)$ is a closed SV-map at x^0.*

Proof. Follows from Prop. 2.22, Hu and Papageorgiou [32] and Lem. 1, Hogan [31]. □

Remark 1.1.1. Let X and T be metric spaces such that $B : X \rightrightarrows T$. Hence, $B(\cdot)$ is a closed set-valued map iff $Graph(B)$ is a closed set in $X \times T$. Consequently, Prop. 1.1.3 implies that, an upper semi-continuous closed valued SV-map has a closed graph.

Nevertheless, the converse of Prop. 1.1.3 may not hold; i.e. the closedness of $B(\cdot)$ at x^0 may not imply its upper semi-continuity at x^0, even if $B(x^0)$ is a compact set (cf. for instance Rem. 2.1 of Kisielewicz [37] for an example). Hence, more is required to conclude that a closed set-valued map is also upper semi-continuous. For this the following definition is needed.

Definition 1.1.8 (local uniform boundedness, see also Hogan[31]).
Let $B: X \rightrightarrows T$. Then $B(\cdot)$ is called *locally uniformly bounded* at $x^0 \in X$ iff there is a neighborhood U of x^0 such that the set

$$\bigcup_{x \in U} B(x)$$

is a bounded set in T. And $B(\cdot)$ is called locally uniformly bounded iff it is locally uniformly bounded at every $x \in X$.

Note that, $B(\cdot)$ is locally uniformly bounded implies that, for each $x \in X$, there is a neighborhood U of x such that

$$cl\left(\bigcup_{x \in U} B(x)\right)$$

is bounded - thus, compact in T; when T is a compact metric space. That is, $B(\cdot)$ is locally uniformly compact. Consequently, the following statement is nothing but that of Hogan [31].

Proposition 1.1.4 (see also Thm. 3 in Hogan [31]). *Let X be a metric space and T be a compact metric space. Then $B: X \rightrightarrows T$ is u.s.c. and compact valued iff $B(\cdot)$ is closed and locally uniformly bounded.*

Corollary 1.1.5. *Let X and T be metric spaces. If $B: X \rightrightarrows T$ is an u.s.c. SV-map with compact values, then the set*

$$\{x \in X \mid B(x) = \emptyset\}$$

is open in $\mathbb{R}^m$; equivalently, the set

$$\{x \in X \mid B(x) \neq \emptyset\}$$

is a closed set.

Proof. Follows from Prop. 1.1.4 and Cor. 3.2 of Hogan [31]. □

Note that, if $B: X \rightrightarrows T$ is a closed valued SV-map and T is compact, then the upper semi-continuity of $B(\cdot)$ follows automatically from Prop. 1.1.4. This will be so, as we are going to consider, later on, a (GSIP) with the index set-valued map $B : X \rightrightarrows T$ defined w.r.t. T being a compact set.

1.2 Set-Valued Maps with given Structures

Set-valued maps which are defined using a parametric family of functions play a vital role in parametric optimization, in sensitivity and perturbation analysis of optimization problems. The main issue, behind set-valued maps with such given structures, is to characterize them through the topological properties of their defining functions. As such, one obtains u.s.c property under weaker assumptions, while the l.s.c. requires stronger ones.

Of interest are set-valued maps $B: X \rightrightarrows T$ and $M: X \rightrightarrows Y$ with structures:

$$B(x) := \{t \in T \mid g_j(x,t) = 0, j \in J;\ h_i(x,t) \leq 0, i \in I\};$$

and

$$M(x) := \{y \in Y \mid f_k(x,y) = 0, k \in K; G(x,y,t) \leq 0, t \in B(x)\}, x \in X,$$

under the following general assumptions

- $I = \{1,\dots,p\}$, $J := \{1,\dots,q\}$ and $K := \{1,\dots,r\}$ are finite index sets;
- $X \subset \mathbb{R}^n, T \subset \mathbb{R}^m$ and $Y \subset \mathbb{R}^l$;
- $h_i : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$, $i \in I$; $f_k : \mathbb{R}^n \times \mathbb{R}^l \to \mathbb{R}, k \in K$; and $G : \mathbb{R}^n \times \mathbb{R}^l \times \mathbb{R}^m \to \mathbb{R}$ are continuous functions.

The map $B(\cdot)$ appears as an index set of a GSIP; while, $M(\cdot)$ usually appears as a feasible set of some parametric semi-infinite optimization problem (PSIP). Frequently, one finds in the literature that $B(x) \equiv B$, i.e. $B(\cdot)$ is a constant SV-map. SV-maps with finite number of constraints has been extensively studied by Bank *et al.* [7]; while the latter form is a subject of current research interest (cf. Henrion and Klatte [22, 39], Cánovas *et al.* [11]).

Proposition 1.2.1 (Thm. 3.1.1 Bank *et al.* [7]). *Let*

$$B(x) = \{t \in T \mid g_j(x,t) = 0, j \in J; h_i(x,t) \le 0, i \in I\}.$$

If the sets X and T are closed and the functions $g_j, j \in J; h_i, i \in I$ are continuous, then $B(\cdot)$ is a closed set valued map.

Corollary 1.2.2. *Let X and T be closed sets and $B(x) = \{t \in T \mid h_i(x,t) \le 0, i \in I;\ g_j(x,t) = 0, j \in J\}$ and the functions $h_i, i \in I;\ g_j, j \in J$ are continuous. Then $B(\cdot)$ is locally uniformly bounded iff $B(\cdot)$ is u.s.c. and compact valued.*

Proof. Follows directly from Prop. 1.1.4 and Prop.1.2.1. □

Remark 1.2.1.

(i) Obviously, by the continuity of the $g_j, j \in J;\ h_i, i \in I,\ B(\cdot)$ is a closed valued SV-map. Moreover, Prop. 1.2.1 indicates that, if T is a compact set, then

$$B(x) = \{t \in T \mid g_j(x,t) = 0, j \in J;\ h_i(x,t) \le 0, i \in I\}$$

will be locally uniformly bounded. Hence, applying Prop. 1.1.4, $B(\cdot)$ is an u.s.c. map with compact values.

(ii) Recently, Cánovas et.al. [11] have established the upper semi-continuity of a set-valued map of the form

$$M(x) := \{y \in Y \mid G(x,y,t) \le 0, t \in B\};\ i.e.\ B(x) \equiv B \subset T,$$

where the $G : X \times Y \times T \to \mathbb{R}$ is a continuous function and, for each fixed $x \in X$ and $t \in T$, $G(x,\cdot,t) : Y \to \mathbb{R}$ is a convex function. The assumptions on G yield that $M(\cdot)$ is a closed-convex-valued SV-map. This map has been used in the stability analysis of linear and convex semi-infinite optimization problems. Let

$$M^{\varepsilon}(x^0) := cl\left(conv\left(\bigcup_{x\in B_{\varepsilon}(x^0)} M(x)\right)\right)$$

represent the ε-reinforced map associated with $M(\cdot)$ (where $B_\varepsilon(x^0)$ is an $\varepsilon-$neighborhood of x^0). Then Prop. 2 in [11] states that $M(\cdot)$ will be u.s.c. at $x^0 \in X$ if M is a closed SV-map at x^0 and, for some $\varepsilon > 0$, $M^\varepsilon(x^0) \setminus M(x^0)$ is a bounded set. Moreover, if Y is a compact set, then the boundedness assumption on $M^\varepsilon(x^0) \setminus M(x^0)$ is trivially satisfied. And the result follows by part (i) of this remark.

That $B(\cdot)$ has compact values could also be algebraically enforced, for instance, by modifying $B(\cdot)$ as

$$B(x) = \{t \in \mathbb{R}^m \mid g_j(x,t) = 0, j \in J;\; h_i(x,t) \leq 0, i \in I; l \leq t \leq u\}$$

for $l, u \in \mathbb{R}^m$. Such a map $B(\cdot)$ has been used for an index SV-map of a (GSIP) by Stein and Still in [77].

Thus, the above discussion reveals that the upper semi-continuity of a SV-map with a given structure could be seen to hold true under somehow weaker assumptions. However, to guarantee the lower semi-continuity one may need certain regularity conditions, like *Metric regularity* and *constraint qualifications*, etc.

In case when $B(\cdot)$ is defined by a linear inequality system, lower semi-continuity could be easily obtained. Such a map has also been used by Stein and Still [77] w.r.t. a linear (GSIP).

Corollary 1.2.3 (Prop. 3.2, Thoai [81]). *Let $B : X \rightrightarrows T$, $h(x,t) := Rx + Qt + b$, where the matrices $R \in \mathbb{R}^{r\times n}, Q \in \mathbb{R}^{r\times m}$, the vector $b \in \mathbb{R}^r$ and*

$$B(x) := \{t \in T \mid h(x,t) \leq 0\}.$$

If T is a closed set, then $B(\cdot)$ is a l.s.c. SV-map on Dom(B).

1.2.1 Regularity Conditions and their Consequences

In order to make sure that the lower semi-continuity of the maps $B(\cdot)$ and $M(\cdot)$ holds true, we may require stronger conditions known as *regularity conditions*. For this, the minimal assumption is that the functions $h_i, i \in I; g_j, j \in J$ and G to be continuous, in their respective domains. We suppose here that the sets X, Y and T to be normed spaces, each with a metric defined through its corresponding norm.

Definition 1.2.1 (r-Regularity - the finite case). Let $t^0 \in \mathbb{R}^m$ and $x^0 \in \mathbb{R}^n$. Then the defining system of $B(\cdot)$

$$g_j(x,t) = 0, j \in J; h_i(x,t) \leq 0, i \in I$$

is called *r-regular* at (x^0, t^0) iff there are open neighborhoods $U(x^0)$ of x^0 and $V(t^0)$ of t^0, and a non-decreasing continuous function $r : \mathbb{R}_+ \to \mathbb{R}_+$ (here $\mathbb{R}_+ := [0, +\infty))$ such that

$$dist(t, B(x)) \leq r\left(\max\left\{\max_{1\leq j\leq q} |g_j(x,t)|, \max_{1\leq i\leq p} [h_i(x,t)]^+\right\}\right), \forall x \in U(x^0), \forall t \in V(t^0),$$

where $h_i^+(x,t) = \max\Big\{0, h_i(x,t)\Big\}$, for each $i \in I$.

Similarly, we can give an $r-$regularity condition for the parametric semi-infinite system of functions defining $M(\cdot)$.

Definition 1.2.2 ($r-$Regularity - the semi-infinite case). Suppose that $B(x) \neq \emptyset$, for each $x \in X$ and $B(\cdot)$ is compact valued. Let $x^0 \in X$ and let $y^0 \in M(x^0)$. We say that the system

$$f_k(x,y) = 0, k \in K; G(x,y,t) \leq 0, t \in B(x) \tag{1.2.1}$$

satisfies the $r-$*regularity* condition at (x^0, y^0) if there exist open neighborhoods $U(x^0)$ of x^0 and $W(y^0)$ of y^0, a non-decreasing continuous function $r : \mathbb{R}_+ \to \mathbb{R}_+$ such that

$$dist(y, M(x)) \leq r\left(\max\left\{\max_{k\in K} |f_k(x,y)|, \max_{t\in B(x)} \Big[G(x,y,t)\Big]^+\right\}\right), \forall x \in U(x^0), \forall y \in W(y^0).$$

If we take the function $r(t) = Ct, C > 0$, then Def. 1.2.1 is the same as the regularity condition given in Def. 1.2 by Auslender [6]. At same time, Def. 1.2.2 is similar to the *metric regularity* (MR) condition given by Klatte and Henrion [22, 39].

Remark 1.2.2. Observe that

1. from Def. 1.2.2, $r-$regularity implies that the set $M(x) \neq \emptyset$ for x sufficiently close to x^0, since, by definition, we have that $dist(y, \emptyset) =: +\infty$;

2. if, for all $x \in X$, $B(x) \equiv B$ is a constant and B is a finite set, then the metric regularity given in Def. 1.2.2 is equivalent to the one in Def. 1.2.1. Therefore, the finite case could be taken as a special instance.

It is worth mentioning that the definition of metric regularity given by Klatte and Henrion [22, 39] are w.r.t. some right hand-side perturbations of the parametric systems 1.2.1. However, it is believed here that the perturbations are not required for the discussion at hand.

Theorem 1.2.4. *Let $x^0 \in X$ and $B(\cdot)$ is compact valued and u.s.c. If for every $y^0 \in M(x^0)$ the system (1.2.1) is $r-$regular at (x^0, y^0), then for every sequence $\{x^n\}_{n\in\mathbb{N}} \subset X$, such that $x^n \to x^0$ there is a sequence $\{y^n\}_{n\in\mathbb{N}} \subset Y$ where $y^n \in M(x^n)$ such that $y^n \to y^0$, i.e. $M(\cdot)$ is l.s.c. at x^0.*

Proof. Let $U(x^0)$ and $W(y^0)$ are neighborhoods which exist according to the r-regularity of (1.2.1). Since, $x^k \to x^0$, there is a sufficiently large positive integer k_0 such that $x^n \in U(x^0), \forall n \geq n_0$. Hence,

$$dist(y, M(x^n)) \leq r\left(\max\left\{\max_{k\in K} |f_k(x^n, y)|, \max_{t\in B(x^n)} [G(x^n, y, t)]^+\right\}\right), \forall y \in W(y^0), n \geq n_0$$

This implies that

$$dist(y^0, M(x^n)) \leq r\left(\max\left\{\max_{k\in K} |f_k(x^n, y^0)|, \max_{t\in B(x^n)} \left[G(x^n, y^0, t)\right]^+\right\}\right), \forall n \geq n_0$$

That means that, for each $n \geq n_0$, there is an $y^n \in M(x^n)$ such that

$$\|y^0 - y^n\| \leq r\left(\max\left\{\max_{k\in K}|f_k(x^0,y^n)|, \max_{t\in B(x^n)}\left[G(x^0,y^n,t)\right]^+\right\}\right) + d(x^n,x^0), \forall n \geq n_0.$$

If we define

$$\psi(x^0,y) := \max_{t\in B(x^0)}\left[G(x^0,y,t)\right]^+$$

then, using the assumption made on the function G and the SV-map $B(\cdot)$, we have $\psi(x^0,\cdot)$ is u.s.c. (cf. Thm. 5, p.52, Aubin and Cellina [4]). Hence, we have that

$$\|y^0 - y^n\| \leq r\left(\max\left\{\max_{k\in K}|f_k(x^0,y^n)|, \psi(x^0,y^n)\right\}\right) + d(x^n,x^0), \forall n \geq n_0.$$

Then taking the limit, and using the continuity of the functions $f_k, k \in K$, the upper semi-continuity of $\psi(x^0,\cdot)$, and the properties of the function r, we obtain that $\lim_{n\to\infty}\|y^0 - y^n\| = 0$. Therefore, according to Def. 1.1.6, $M(\cdot)$ is l.s.c. at x^0. □

Corollary 1.2.5. *Let $x^0 \in X$. If for $t^0 \in B(x^0)$ the system*

$$g_j(x,t) = 0, j \in J; h_i(x,t) \leq 0, i \in I$$

is $r-$regular at (x^0,t^0), then the set-valued map $B(\cdot)$ is lower semi-continuous at x^0.

Remark 1.2.3.

1. Observe that, if $B(x^0) = \emptyset$ and the equality constraints are deleted from the definition of $M(\cdot)$, then $M(x^0) = Y$. Hence, the SV-map $M(\cdot)$ may fail to be lower semi-continuous at such a point x^0. This could be verified, if one considers semi-continuity properties in the $\varepsilon-$sense (cf. pp. 45 & 46 of Aubin and Cellina [4]).

2. We can, in fact, weaken the r-regularity given in Def. 1.2.2 as: the system (1.2.1) is r-regular at (x^0,y^0) iff there is a neighborhood $U_0(x^0)$ such that

$$dist(y^0,M(x)) \leq r\left(\max\left\{\max_{k\in K}|f_k(x,y^0)|, \max_{t\in B(x)}\left[G(x,y^0,t)\right]^+\right\}\right), \forall x \in U_0(x^0).$$

Hence, if $y^0 \in M(x^0)$ and the system (1.2.1) is $r-$regular at (x^0,y^0), then $M(\cdot)$ is l.s.c. at x^0. To see this, let

$$g(x,y^0) := \max_{t\in B(x)}\left[G(x,y^0,t)\right]^+ \text{ and } f(x) := \max_{k\in K} f_k(x,y^0).$$

Then $g(\cdot, y^0)$ and f are u.s.c at x^0. Thus, given $\varepsilon > 0$, there is a neighborhood $U_1(x^0)$ such that, for each $x \in U_1(x^0)$

$$g(x, y^0) \leq g(x^0, y^0) + \varepsilon \Rightarrow g(x, y^0) \leq \varepsilon.$$

Similarly, there exists a neighborhood $U_2(x^0)$ such that

$$f(x) \leq \varepsilon, \forall x \in U_2(x^0).$$

Consequently, by $r-$regularity, for each $x \in U_3(x^0) := U_0(x^0) \cap U_1(x^0) \cap U_2(x^0)$ we have

$$dist(y^0, M(x)) \leq r(\max\{\varepsilon, \varepsilon\}) = r(\varepsilon).$$

Let $W(y^0)$ be any neighborhood of y^0. Then there is an $\varepsilon > 0$ such that the open ball $B_{2r(\varepsilon)}(y^0) \subset W(y^0)$ and a corresponding $U_\varepsilon(x^0)$ such that

$$M(x) \cap B_{2r(\varepsilon)}(y^0) \neq \emptyset, \forall x \in U_3(x^0) \cap U_\varepsilon(x^0).$$

From which follows that

$$M(x) \cap V(y^0) \neq \emptyset, \forall x \in U_3(x^0) \cap U_\varepsilon(x^0).$$

Which implies that $M(\cdot)$ is l.s.c. at x^0.

1.2.2 Constraint Qualifications and their Consequences

In many cases regularity conditions are difficult to verify[2]. Thus, if the defining systems of $B(\cdot)$ or $M(\cdot)$ satisfy conditions known as *constraint qualifications*, then lower semi-continuity properties could be guaranteed. In general, constraint qualifications require certain differentiability (or convexity) properties of the functions defining $B(\cdot)$ and $M(\cdot)$.

[2]For this reason, much effort has been invested by various authors to characterize regularity conditions through relatively simpler equivalent conditions; for instance, using constraint qualifications. Such work have been undertaken, for instance, by Auslender [6], Klatte and Henrion [22, 39], etc.

For the sake of simplicity, we consider in this section the set-valued maps $B(\cdot)$ and $M(\cdot)$ without equality constrains. That is

$$B(x) = \{t \in T \mid h_i(x,t) \leq 0, i \in I\}$$

and

$$M(x) = \{y \in Y \mid G(x,y,t) \leq 0, \forall t \in B(x)\}.$$

Definition 1.2.3 (Slater constraint qualification (SCQ)). Let $T \subset \mathbb{R}^m$ be convex and bounded and, for each $i \in I$, $h_i(x,\cdot) : \mathbb{R}^m \to \mathbb{R}$ is convex. We say that the system

$$h_i(x,t) \leq 0, i \in I.$$

satisfies the *Slater constraint qualification (SCQ)* at $x^0 \in \mathbb{R}^m$ if there exists $t^* \in T$ such that

$$h_i(x^0,t^*) < 0, \forall i \in I.$$

Proposition 1.2.6 (Thm. 12, Hogan [31]). *Let T be a convex subset of some normed space and let $x^0 \in X$. If, for each $i \in I$, $h_i(\cdot,\cdot)$ is continuous on $\{x^0\} \times B(x^0)$, $h_i(x^0,\cdot), i \in I$, is convex and the (SCQ) is satisfied at x^0, then the map*

$$B(x) = \{t \in T \mid h_i(x,t) \leq 0, i \in I\}.$$

is l.s.c. at x^0.

Proof. See Hogan [31].(cf. also Thm. 3.1.5 of Bank *et al.* [7] for a similar discussion). □

In the presence of differentiability and convexity properties of the defining functions of $B(\cdot)$, the SCQ can be shown to follow from the the *Mangasarian-Fromovitz Constraint Qualification (MFCQ).* Thus, we have the following well known constraint qualifications.

Definition 1.2.4 (Mangasarian-Fromovitz Constraint Qualification (MFCQ)). Let $B(x) = \{t \in T \mid h_i(x,t) \leq 0, i \in I\}$; the functions $h_i, i \in I$ be continuous in $\mathbb{R}^n \times \mathbb{R}^m$ and, for each $x \in \mathbb{R}^n$, $h_i(x,\cdot) : \mathbb{R}^m \to \mathbb{R}$ is continuously differentiable; and for $x^0 \in \mathbb{R}^n$, let $t^0 \in B(x^0)$. The Mangasarian-Fromovitz constraint qualification (MFCQ) is said to hold at (x^0, t^0) iff there exists a vector $\xi \in \mathbb{R}^m$ such that

$$\xi \nabla_t h_i(x^0, t^0) < 0, \forall i \in I(x^0, t^0);$$

where $I(x^0, t^0) = \{i \in I \mid h_i(x^0, t^0) = 0\}$. The vector ξ with the above property is known as an *(MFCQ) vector.*

Auslender [6] indicated that the validity of (MFCQ) at (x^0, t^0), $t^0 \in B(x^0)$, implies the metric regularity (Def. 1.2.1 with $r(t) = Ct$) of the defining system of $B(\cdot)$ at (x^0, t^0) (cf. Thm. 1.1. in Auslender [6]). However, the lower semi-continuity of $B(\cdot)$ at x^0 could also be directly verified under the satisfaction of (MFCQ).

Definition 1.2.5 (Linear Independence Constraint Qualification (LICQ)). Let $B(x) = \{t \in T \mid h_i(x, t) \leq 0, i \in I\}$; the functions $h_i, i \in I$ be continuous in $\mathbb{R}^n \times \mathbb{R}^m$ and, for each $x \in \mathbb{R}^n$, $h_i(x, \cdot) : \mathbb{R}^m \to \mathbb{R}$ is continuously differentiable; and for $x^0 \in \mathbb{R}^n$, let $t^0 \in B(x^0)$. The Linear Independence constraint qualification (LICQ) is said to hold at (x^0, t^0) iff the system

$$\left\{\nabla_t h_i(x^0, t^0) \mid i \in I(x^0, t^0)\right\}$$

is linearly independent.

Remark 1.2.4. The satisfaction of (LICQ) at $t^0 \in B(x^0)$ implies that of (MFCQ). In fact, for $\varepsilon > 0$, the system

$$\xi^\top \nabla_t h_i(x^0, t^0) = -\varepsilon, i \in I(x^0, t^0)$$

has a solution $\xi \in \mathbb{R}^m$; implying that (MFCQ) is satisfied. However, the converse is not always true; i.e. (LICQ) is stronger than (MFCQ).

Proposition 1.2.7. *If $t^0 \in B(x^0)$ and (MFCQ) holds at (x^0, t^0), then the map $B(\cdot)$ is lower semi-continuous at x^0.*

Proof. Follows using standard arguments (see Sec. 3.6.4 for a more general discussion). □

For the system (1.2.1) (without equality constraints) we define the set of active constraints as

$$E(x, y) := \{t \in B(x) \mid G(x, y, t) = 0\}.$$

Definition 1.2.6 (Extended Mangasarian-Fromowitz Constraint Qualification). Let $(x^0, y^0) \in \mathbb{R}^n \times \mathbb{R}^q$, the function $G : \mathbb{R}^n \times \mathbb{R}^q \times T \to \mathbb{R}$ ($T \subset \mathbb{R}^m$) is continuous, and $G(x, \cdot, t)$ differentiable w.r.t. y and $\nabla_y G(\cdot, \cdot, t)$ is continuous for each $t \in B(x)$. Then the *extended Mangasarian-Fromowitz constraint qualification* (EMFCQ) is said to be satisfied at (x^0, y^0) if there exists a vector $\xi \in \mathbb{R}^q$ such that

$$\nabla_y G(x^0, y^0, t)^\top \xi < 0, \forall t \in E(x^0, y^0).$$

Proposition 1.2.8. *Let $B(\cdot)$ be u.s.c. and compact valued and $y^0 \in M(x^0)$. Then if (EMFCQ) is satisfied at (x^0, y^0), then $M(\cdot)$ is l.s.c. at x^0.*

Actually under the assumptions of Prop. 1.2.8, the system (1.2.1) is metrically regular (cf. Klatte and Henrion [39]).

All in all, the (MFCQ) and (LICQ) play a pivotal role in the characterization of a (GSIP). In several literature of (GSIP), (MFCQ) is usually assumed to hold. Optimality conditions, local reducibility conditions and numerical methods for (GSIP) are usually based on (MFCQ). Therefore, constraint qualifications do not only guarantee the lower semicontinuity of the index map, they are also the tickets for a "well-behaving" (GSIP).

In Chap. 3, the *lower robustness* (*upper robustness*) of a set-valued map has been introduced, which is weaker than l.s.c. (u.s.c.) and has been shown to follow from a weaker regularity condition. Further particular definitions and properties of set-valued maps are also given later on, as is relevant.

Chapter 2

A Review of Generalized Semi-infinite Optimization

2.1 Introduction

While problems of type (GSIP) might have appeared, in some form or other, elsewhere in the mathematical literature, the first well established (GSIP) model appeared in 1987 (Krabs [40]); particularly, in 1988 Graettinger & Krogh [20] coined the term *generalized semi-infinite programming* (GSIP). Beginning the 1990's the number of practical problems of type (GSIP) started to pick up. For instance, in 1991 Hettich & Still formulated a robot maneuverability problem into a (GSIP); in 1994 Hoffmann & Reinhardt [29] came up with a (GSIP) model out of an approximation problem; in 1997 Kaplan & Tichatschke [35] published a (GSIP) model from a terminal variational problem, etc.

The first publication that directly deals with theory of (GSIP) was that of Hettich & Still in 1995. In [23] Hettich & Still established first and second order optimality condition based on the local reducibility of the lower level problem of (GSIP). Actually, a study on the local reducibility of the lower level problem of (GSIP) was, by then, already published by Klatte [38] in 1991. In the same year, 1995, Rückmann, Jonge & Stein [34] recognized the special structure of (GSIP) in a preprint (cf. also Levitin [42]). There they gave a summery of practical (GSIP) models, established theoretical examples indicating the difficult structure of (GSIP) and they also set up a first order optimality condition which culminated into a publication in 1998. After 1998 we find a quick succession of publications dealing on both the theoretical investigation and computational algorithms of (GSIP)s.

Given the structural difficulty of (GSIP), it is only recently (Stein [74, 75], Still & Stein [77, 78]) that we have witnessed a viable and practical method that could successfully tackle certain class of these problems. The method of Still & Stein transforms a given (GSIP) into an equivalent Bi-level optimization(BL for short) problem, so that well established algorithms of (BL) could be used for the computation. However, for such a transformation to work properly, the index map $B(\cdot)$ of (GSIP) needs to be lower semi-continuous and certain convexity properties are also expected of the problem data. There are also computational algorithms being proposed by other authors, but we are still waiting for the news of their computational experiments.

The literature on (GSIP) could be roughly put into four categories:

- special (GSIP) models arising from practical problems [2, 20, 29, 35, 36, 40, 83];
- topological and stability issues [43, 62, 70, 71, 72, 85];
- optimality conditions [21, 23, 34, 38, 59, 60, 62, 73, 87, 84];
- reducibility of (GSIP) into problems of somehow manageable types; for instance, into: finite, semi-infinite, bi-level optimization problems, [38, 43, 42, 71, 74, 77, 78, 80, 84];
- numerical solution methods [47, 49, 74, 75, 77, 78, 79, 80]

2.2 Structure of the Feasible Set

We consider the problem

$$(GSIP) \qquad f(x) \to inf$$
$$s.t.$$
$$G(x,t) \geq 0, \ \forall t \in B(x)$$
$$x \in X,$$

where we make

Assumption (A1) The sets $X \subset \mathbb{R}^n, T \subset \mathbb{R}^m$ are compact nonempty, the function $f : X \to \mathbb{R}$ and $G : X \times T \to \mathbb{R}$ are continuous and upper semi-continuous(u.s.c.) on X and $X \times T$; respectively. The set-valued map (SV-map) $B : X \rightrightarrows T$ is at least compact valued, but may have empty values for some $x \in X$ and may also be given by

$$B(x) = \{t \in T \mid h_i(x,t) \leq 0, i \in I\},$$

in which the functions $h_i : X \times T \to \mathbb{R}$ are at least upper semi-continuous and $I = \{1, \ldots, p\}$.

Hence, the the feasible(admissible) set of (GSIP) is given by

$$\mathcal{M} := \{x \in X \mid G(x,t) \geq 0, t \in B(x)\}.$$

The parametric optimization problem

$$(GO(x)) \qquad G(x,t) \to \inf$$
$$h_i(x,t) \leq 0, i \in I, t \in T.$$

is known as the *lower level problem* associated with the (GSIP).

Thus, we could also write the feasible set of (GSIP) as

$$\mathcal{M} = \{x \in X \mid \inf_{t \in B(x)} G(x,t) \geq 0\} = \{x \in X \mid v(x) \geq 0\}$$

where

$$v(x) := \begin{cases} \inf_{t \in B(x)} G(x,t) & \text{if } B(x) \neq \emptyset \\ +\infty & \text{else.} \end{cases}$$

is the optimal value function of the (GO(x)).

The following example shows that (GSIP) may not posses a solution

Example 2.2.1.

$$f(x) := (x_1 - 2)^2 + (x_2 + 3)^2 \rightarrow min$$

$$s.t.$$

$$x_1 + x_2 - t \geq 0, \;\; \forall t \in B(x)$$

$$x \in X = [-5, 5]^2,$$

where

$$B(x) := \{t \in T \mid t + 2 \leq x_1, -2t - 3 \leq x_2\}, T = [-10, 10].$$

Obviously, the minimum of f over $\mathbb{R}^2$ occurs at $x^0 = (2, -3)$. A trivial calculation reveals that $\mathcal{M} = \{x \in X \mid 2x_1 + x_2 < 1\} \cup \{x \in X \mid x_2 + 2 \geq 0, 2x_1 + x_2 \geq 1\}$, $x^0 \in cl(\mathcal{M}) \setminus \mathcal{M}$ and $\inf_{x \in \mathcal{M}} f(x) = f(x^0) = 0$.

However, if $B(x) \equiv B$ (a constant SV-map), then the (GSIP) reduces to a (SIP) and we have

$$\mathcal{M} = \bigcap_{t \in B} \{x \in X \mid G(x,t) \geq 0\}$$

which is a closed set, because of the u.s.c. assumption on G. Moreover, $\mathcal{M}$ will be convex if, for each fixed $t \in T$, $G(\cdot, t) : X \rightarrow \mathbb{R}$ is a convex function and X is a convex set.

Therefore, in contrast to the feasible set of a (SIP), the feasible set $\mathcal{M}$ of a (GSIP) may not be closed; not be convex (cf. Jongen et al. [33], Levitin [43]); even if all its defining functions are affine) and it may also have disjunctive structures (cf. Rückmann & Stein [61, 62]). Hence, various attempts have been made to characterize the global and local structure of the set $\mathcal{M}$ (see Stein *et al.* [34, 70, 71, 72] and Weber [87, 85, 86]).

Obviously, the nature of (GSIP) mainly depends on the properties of the lower level problem (GO(x)). Thus, the theory of finite parametric optimization plays a major role in the characterization of (GSIP).

2.2.1 Global Structure of the Feasible Set

Recall that using the marginal function v of (GO(x)) we had

$$\mathcal{M} = \{x \in X \mid v(x) \geq 0\};$$

which indicates that $\mathcal{M}$ is the *upper level set* of a marginal function. Characterization of $\mathcal{M}$ as a level set of the marginal function v was carried out by Stein [70, 72].

Thus, if G is u.s.c. on $X \times T$ and $B(\cdot)$ is a l.s.c. SV-map, then $v(\cdot)$ will be u.s.c. on X (cf. Thm. 4, p 51, Aubin & Cellina [4]) yielding that $\mathcal{M}$ is a closed set. However, if $B(\cdot)$ is not a l.s.c. SV-map, the closedness of $\mathcal{M}$ may not be sure.

The above result have also been verified by Jongen et al. [34].

Proposition 2.2.2 (Prop. 2.1., [34]). *If $B(\cdot)$ is lower semi-continuous, then $\mathcal{M}$ is closed.*

That is, the lower semi-continuity of $B(\cdot)$ is a sufficient condition for $\mathcal{M}$ to be closed.

Corollary 2.2.3. *Let G be an u.s.c. function on $X \times T$. If x^0 belongs to the boundary $\partial\mathcal{M}$ of $\mathcal{M}$ but $x^0 \notin \mathcal{M}$, then $B(\cdot)$ is not lower semi-continuous at x^0.*

Proof. Let $\{x^k\} \in int(\mathcal{M})$ such that $x^k \to x^0$. Since $x^0 \notin \mathcal{M}$, there is $\bar{t} \in B(x^0)^*$ such that

$$G(x^0, \bar{t}) < 0$$

*W.l.o.g. we assume that $B(x^0) \neq \emptyset$. But, obviously, if $B(x^0) = \emptyset$, then $x^0 \in \mathcal{M}$.

But $x^k \in \mathcal{M}$ implies that $G(x^k, t) \geq 0, \forall t \in B(x^k)$. Let $\{t^k\}$ be any sequence where $t^k \in B(x^k)$ for each k. If we assume that $t^k \to \bar{t}$, then by the u.s.c. of G we have that $0 \leq \lim sup G(x^k, t^k) \leq G(x^0, \bar{t})$, which is a contradiction. Consequently, every such sequence $\{t^k\}$ either diverges or converges to an element other than $\bar{t}$. Which concludes that $B(\cdot)$ will not be lower semi-continuous at x^0. □

We easily deduce that: among the points on the boundary of $\mathcal{M}$, but which do not belong to $\mathcal{M}$, are those at which $B(\cdot)$ fails to be lower semi-continuous.

Remark 2.2.1. Let G be a continuous function. If $B(\cdot)$ is assumed to be u.s.c. and compact valued (which is often a standard assumption in (GSIP)), then the continuity of G, implies that $v(\cdot)$ is l.s.c. (see Thm. 5, p. 52, Aubin & Cellina [4]). Moreover, $B(\cdot)$ is u.s.c. implies that $\{x \in \mathbb{R}^n \mid B(x) = \emptyset\}$ is, by Cor. 1.1.5, an open set and we have

$$\{x \in \mathbb{R}^n \mid B(x) = \emptyset\} \subset \{x \in \mathbb{R}^n \mid v(x) > 0\} \subset int(\mathcal{M}). \tag{2.2.1}$$

Note that, by definition, $v(x) = \inf_{t \in B(x) = \emptyset} G(x, t) = +\infty$. Hence, $v(x) > 0$ when $B(x) = \emptyset$. In other words, $[Dom B]^c \subset int(\mathcal{M})$.

However, in general, $\mathcal{M}$ would be closed if v is u.s.c., which is assured if $B(\cdot)$ is l.s.c.; i.e. the relation $\{x \in \mathbb{R}^n \mid v(x) = 0\} \subset \mathcal{M}$ always holds if $B(\cdot)$ is l.s.c.

2.2.2 Disjunctive Structures in GSIP

The following simple example further elaborates what bad global structure $\mathcal{M}$ could have.

Example 2.2.4 (Stein [72]). *If*

$$G(x, t) = t \text{ and } B(x) = \{x \in \mathbb{R}^2 \mid t \geq x_1,\ t \leq x_2\}$$

then

$$\mathcal{M} = \{x \in \mathbb{R}^2 \mid t \geq 0,\ \forall t \in B(x)\}$$

that is

$$\mathcal{M} = \{x \in \mathbb{R}^2 \mid x_1 \leq x_2,\ x_1 \geq 0\} \cup \{x \in \mathbb{R}^2 \mid x_1 > x_2\}$$

which is not-closed and not-convex. A simple sketch of $\mathcal{M}$ also reveals that, it has a re-interant corner point *at $x = (0,0)^\dagger$. Observe that, G is a linear function depending only on the index parameter t.*

Considering the given structure of $B(\cdot)$; i.e.

$$B(x) = \{t \in \mathbb{R}^m \mid h_i(x,t) \leq 0, i \in I\},$$

there is a second global characterization of $\mathcal{M}$, by setting:

$$\begin{aligned} \mathcal{G} &:= \{(x,t) \in \mathbb{R}^n \times \mathbb{R}^m \mid G(x,t) \geq 0\} \\ \mathcal{B} &:= \{(x,t) \in \mathbb{R}^n \times \mathbb{R}^m \mid h_i(x,t) \leq 0, i \in I\}. \end{aligned}$$

The following statement (by Rückmann & Stein [62]) shows the direct role played by the index SV-map $B(\cdot)$ on the structure of the feasible set $\mathcal{M}$ of (GSIP).

Proposition 2.2.5 (Lemma 2.1, [62]). *The feasible set $\mathcal{M}$ of (GSIP) is given by*

$$\mathcal{M} = \Big[Pr_x(\mathcal{B} \cap \mathcal{G}^c)\Big]^c,$$

where $Pr_x : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ represents the canonical projection map w.r.t. the component x of (x,t).

Prop. 2.2.5 shows that disjunctive structures are inherent in the feasible set of a (GSIP) (see Jongen *et al.* [33]). This could be more obvious if all the defining functions of G and $h_i, i \in I$, are affine linear w.r.t. (x,t). In such a case, $\mathcal{M}$ will be the union of a finite number of closed and open half spaces.

†*A re-interant corner point x^0 of $\mathcal{M}$ is a point where the set $\mathcal{M}$ fails to be locally approximatable (around x^0) using convex sets.*

To further elucidate the nature of the feasible set $\mathcal{M}$, Rückmann and Stein [62] considered very simple forms of affine linear function for the lower level problem (GO(x)).

$$G(x,t) := c^\top t + d(x)$$

and

$$h_i(x,t) := a_i^\top t + b_i(x), i \in I,$$

where in both cases the variations in x are attached only to the constant terms. Set

$$A = \begin{pmatrix} \vdots \\ a_i^\top \\ \vdots \end{pmatrix}, b(x) = \begin{pmatrix} \vdots \\ b_i(x) \\ \vdots \end{pmatrix}$$

Then the Lagrange function of the lower level problem (GO(x)) will be

$$\mathcal{L}(x,t,\alpha,\gamma) := \alpha\Big[c^\top t + d(x)\Big] + \gamma^\top\Big[At + b(x)\Big]$$

which is a continuous function. Consequently, Rückmann and Stein [62] stated and proved that

Proposition 2.2.6 (Cor. 2.5, [62]).

$$\mathcal{M} = \bigcup_{\gamma\in\mathcal{V}_1}\{x \in \mathbb{R}^n \mid \mathcal{L}(x,t,1,\gamma) \geq 0\} \cup \bigcup_{\gamma\in\mathcal{V}_0}\{x \in \mathbb{R}^n \mid \mathcal{L}(x,t,0,\gamma) > 0\}$$

where $\mathcal{V}_0$ and $\mathcal{V}_1$ are some finite index sets.

Prop. 2.2.6 stress the prevalence of a disjunctive structure in $\mathcal{M}$ even in the simplest possible cases. Detailed discussion on this and related issue are found in [62] and also in Stein [72, 74, 75].

2.3 Convexity Structures in GSIP

The discussion in the last two sections reveal that, even a (GSIP) (with all) linear constraints may not have a convex feasible set. Consequently, it would be advantageous, both theoretically and numerically, to identify and characterize convexity structures that may exist in a given (GSIP). Thus, if the objective function f is convex and the feasible set $\mathcal{M}$ is a convex set, then we have a convex (GSIP). This is what is known as the *completely convex* case. Such considerations have been made, for instance, by Still [79]. Moreover, the convexity of the lower level problem (GO(x)) also plays an important role in the derivation of optimality conditions, development of numerical algorithms and for a simple description of the feasible set (see for instance Prop. 2.2.6).

2.3.1 Convexity of the feasible set $\mathcal{M}$

Observe that, if the marginal function v of (GO(x)) is a concave function then $\mathcal{M}$ will be a convex set; hence, connected. We next have conditions on G and the set-valued map $B(\cdot)$ that yield the convexity of the feasible set $\mathcal{M}$.

Definition 2.3.1 (Fiacco & Kyparisis [16]). Let $G : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ be any function. G is said to be *bi-convex* iff it is convex w.r.t. $(x,t) \in \mathbb{R}^n \times \mathbb{R}^m$; i.e., for $(x_1,t_1),(x_2,t_2) \in \mathbb{R}^n \times \mathbb{R}^m$ and $\lambda \in [0,1]$ we have $G(\lambda(x_1,t_1)+(1-\lambda)(x_2,t_2)) \leq \lambda G(x_1,t_1)+(1-\lambda)G(x_2,t_2)$.

Definition 2.3.2 (concave set-valued map, Fiacco & Kyparisis [16]). Let $B : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ be a set-valued map. If for each $x_1, x_2 \in \mathbb{R}^n$ and $\lambda \in [0,1]$

$$B(\lambda x_1 + (1-\lambda)x_2) \subset \lambda B(x_1) + (1-\lambda)B(x_2),$$

then $B(\cdot)$ is said to be concave.

A trivial example of a concave set-valued map is given by

Lemma 2.3.1. *Let $B : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ and*

$$B(x) := \prod_{i=1}^{m} [a_i(x), b_i(x)],$$

where for each $i \in \{1,\ldots,m\}$, $a_i, b_i : \mathbb{R}^n \to \mathbb{R}$ are continuous functions and $a_i(x) \leq b_i(x), \forall x \in \mathbb{R}^n$. If, for each $i \in \{1,\ldots,m\}$, a_i is a concave while b_i is a convex function, then $B(\cdot)$ is a concave map.

Proposition 2.3.2 (Lem.2, Still [79]; Prop. 3.1, Fiacco & Kyparisis [16]). *Consider the function G and the marginal function v of (GSIP). If $-G$ is bi-convex and B is a concave SV-map, then v is concave function; hence, $\mathcal{M}$ is a convex set .*

The proposition next gives another situation in which the set-valued map $B(\cdot)$ is concave. Below, for a function s, $\partial s(\cdot)$ represents the sub-differential in the sense of Rockafellar [55].

Proposition 2.3.3 (see Prop. 3.4, p112, [16]). *Let $h : \mathbb{R}^n \to \mathbb{R}$ and $s : \mathbb{R}^m \to \mathbb{R}$ such that both h and s are convex and for all $t \in \mathbb{R}^m$, $0 \notin \partial s(t)$, then the set-valued map*

$$B(x) = \{t \in \mathbb{R}^m \mid h(x) + s(t) \geq 0\}$$

is concave.

Proof. The idea is taken from the proof of Prop. 3.4. of [16] in a modified sense. Choose arbitrary $x_1, x_2 \in Dom(B)$, $\lambda \in [0, 1]$ and $\bar{t} \in B(\lambda x_1+(1-\lambda)x_2)$. Set $x_\lambda := \lambda x_1+(1-\lambda)x_2$. Hence, $h(x_\lambda) + s(\bar{t}) \geq 0$. By assumption, there exists $\xi \in \partial s(\bar{t})$ such that $\xi \neq 0$ and

$$s(t) - s(\bar{t}) \geq \xi^\top (t - \bar{t}), \forall t \in \mathbb{R}^m.$$

Let $b := s(\bar{t}) - \xi^\top \bar{t}$, from which follows that $s(t) \geq \xi^\top t + b, \forall t \in \mathbb{R}^m$ (where equality holds when $t = \bar{t}$). Observe also that $h(x_\lambda) + \xi^\top \bar{t} + b \geq 0$. After using the convexity of h, we set

$$\triangle_\lambda = \lambda h(x_1) + (1 - \lambda)h(x_2) + \xi^\top \bar{t} + b \geq 0.$$

One can find a vector $\eta \in \mathbb{R}^m$ such that the following holds

$$\begin{cases} \xi^\top \eta = 0 \\ \bar{t} = \eta + \frac{\xi^\top \bar{t}}{\xi^\top \xi}\xi \end{cases}$$

since $\xi \neq 0$. The latter is a sort of (scaled) decomposition of $\bar{t}$ along the orthogonal vectors η and ξ.

Let

$$\begin{aligned} t_1 &= \eta + (\triangle_\lambda - h(x_1) - b)\frac{\xi}{\xi^\top \xi} \\ t_2 &= \eta + (\triangle_\lambda - h(x_2) - b)\frac{\xi}{\xi^\top \xi}. \end{aligned}$$

It then easily follows that

$$\begin{aligned} h(x_1) + s(t_1) &\geq 0 \text{ and} \\ h(x_2) + s(t_2) &\geq 0. \end{aligned}$$

These mean, $t_1 \in B(x_1)$ and $t_2 \in B(x_2)$; furthermore,

$$\bar{t} = \lambda t_1 + (1-\lambda)t_2.$$

Hence, $\bar{t} \in \lambda B(x_1) + (1-\lambda)B(x_2)$. Therefore, $B(\cdot)$ is a concave set-valued map.

□

The assumption $0 \notin \partial s(t), t \in \mathbb{R}^m$ is satisfied if, for instance, s is taken to be a non-trivial affine linear function.

2.3.2 Convex Lower Level Problem in a GSIP

As stated earlier convexity structures in the lower level problem are advantageous. Thus, for

$$(GO(x)) \qquad G(x,t) \to \inf$$
$$h_i(x,t) \leq 0, i \in I,$$

(GO(x)) is a *convex lower level problem* if, for each fixed x, the functions $G(x,\cdot) : \mathbb{R}^m \to \mathbb{R}$ and $h_i(x,\cdot) : \mathbb{R}^m \to \mathbb{R}, i \in I$, are convex.

Note that, the convexity of the lower level problem may not imply the convexity of the feasible set $\mathcal{M}$.

Example 2.3.4. *Let the feasible set of a (GSIP) be defined as*

$$\mathcal{M} = \{x \in \mathbb{R}^2 \mid x_1 + x_2 \geq t, \forall t \in B(x)\}$$

with $B(x) = \{t \in \mathbb{R} \mid t + 2 \leq x_1, -2t - 3 \leq x_2\}$*. Then, the lower level problem will be*

$$\begin{aligned} (GO(x)) \qquad\qquad & x_1 + x_2 - t \to \inf \\ & -x_1 + t + 2 \leq 0 \\ & -x_2 - 2t - 3 \leq 0. \end{aligned}$$

Obviously, (GO(x)) is a convex problem, while the feasible set could be rewritten as

$$\mathcal{M} = \{x \in \mathbb{R}^2 \mid 2x_1 + x_2 < 1\} \cup \{x \in \mathbb{R}^2 \mid x_2 + 2 \geq 0, 2x_1 + x_2 \geq 1\}$$

which is a non-convex set.

Remark 2.3.1. The following is one simple example which shows the existence of convexity in both the lower and upper problems.

$$\begin{aligned} (GSIP) \qquad\qquad & f(x) \to \inf \\ & s.t. \;\; x \in \{x \in \mathbb{R}^n \mid a^\top x + b^\top t \geq 0, \forall t \in B(x)\}; \end{aligned}$$

where $B(x) = \{t \in \mathbb{R}^m \mid h(x) + d^\top t \geq 0\}$ with $a \in \mathbb{R}^n$; $b, d \in \mathbb{R}^m$ are vectors and $f, h : \mathbb{R}^n \to \mathbb{R}$ being convex functions (Prop. 2.3.3 has been used here). Note that, for total convexity to be attained G needs to be a linear function in both variables.

In any case, the convexity of (GO(x)) yields simpler descriptions for the directional derivative of the marginal function $v(\cdot)$ (cf. Hogan [30]). Such simpler descriptions of the directional derivatives of $v(\cdot)$, in turn, are helpful to drive first and second optimality conditions. Issues based on the convexity of the lower level problem have been taken, for instance, by Rückmann & Stein [61], Stein [73, 74], Stein & Still [77, 78], etc.

2.4 First Order Optimality Conditions

First order optimality conditions for (GSIP) are recently given by Hettich & Still [23], Jongen *et al.* [34], Rückmann & Shapiro [60], Stein & Still [76], Stein[73] and Weber [84, 87]. In [73], the derivation of first order conditions not only follows a similar line of argument as for standard (SIP) (see Hettich & Zencke [25]), but it is also a generalization of similar results obtained in [21, 23, 34, 60, 76, 84, 87]. Thus, we follow the approach of Stein [73].

We suppose that the following standard assumption holds true for the rest of this chapter.

Assumption (USC): The index map $B : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ of (GSIP) is locally uniformly bounded (cf. also Cor. 1.2.2 in chapter 1).

Recall that, using the marginal value function v of (GO(x)) the feasible set $\mathcal{M}$ is written as

$$\mathcal{M} = \{x \in \mathbb{R}^n \mid v(x) \geq 0\}$$

where v is given by

$$v(x) = \inf_{t \in B(x)} G(x, t).$$

In general, the function v is not differentiable. In other words, under general assumptions, (GSIP) is a non-smooth optimization problem. Thus, first and second order optimality conditions require certain differentiability properties of the function v. Hence, the study of differentiability properties of marginal functions of parametric optimization problems plays a significant role here. The idea is to use approximates of the first and second order directional derivatives of the function v (at a given $x \in \mathcal{M}$) in terms of the derivatives of the Lagrange function of (GO(x)); as required by first and second order optimality conditions, respectively. Differentiability of marginal functions are discusses, for instance, by Gauvin & Debeau [18], Bonans & Shapiro [9, 10], Hogan [30], Levitin [44, 45, 46] and Shapiro [63, 64, 65], etc. However, we bypass the differentiability properties of v here, and concentrate on assumptions and results that lead to first and second order optimality conditions.

First order optimality conditions for (GSIP) are also based on the following basic assumption :

Assumption (FO): The functions f, G and $h_i, i \in I$, are at least one time continuously differentiable w.r.t. x and t, correspondingly.

Definition 2.4.1 (local minimizer). Let $x^0 \in \mathcal{M}$. Then x^0 is a local minimum point of (GSIP) iff there exists a neighborhood U of x^0 such that

$$f(x) \geq f(x^0), \forall x \in \mathcal{M} \cap U.$$

Definition 2.4.2 (strict local minimizer). Let $x^0 \in \mathcal{M}$. Then x^0 is called a strict local minimizer of (GSIP) of order $\kappa = 1$ or $\kappa = 2$ iff there exist a constant $\gamma > 0$ and a neighborhood U of x^0 such that

$$f(x) \geq f(x^0) + \gamma \|x - x^0\|^{\kappa}, \forall x \in \mathcal{M} \cap U.$$

Remark 2.4.1.

(i) If a local optimizer x^0 of (GSIP) lies in the interior $int\mathcal{M}$, then (GSIP) could be considered locally, around x^0, as an unconstrained problem. Thus, the well known first order optimality conditions, i.e. $Df(x^0) = 0$, holds. In reality, optimality conditions (both first and second order) are considered for local optimal points that lie on *the feasible boundary* of $\mathcal{M}$; i.e., when a local minimizer $x^0 \in \partial\mathcal{M} \cap \mathcal{M}$. Such points are called *feasible boundary* points Stein[72]. Moreover, if x^0 is a feasible boundary point, then we have $B(x^0) \neq \emptyset$. (Recall that, the feasible set $\mathcal{M}$ of (GSIP) may not contain all of its boundary points, unless $B(\cdot)$ is l.s.c.).

(ii) If $x^0 \in int\mathcal{M}$ is a local minimizer of (GSIP), then x^0 cannot be a strict local minimizer of order one (cf. Stein & Still [76]). Usually, strict local minimizers of order one occur w.r.t. Chebychev approximation problems; for instance, w.r.t. reverse Chebychev approximation problems (cf. Hoffmann & Reinhardt [29]).

Accordingly, we proceed with the following definition as was given in [73].

Definition 2.4.3 (contingent cone). Let $\mathcal{M}$ be any set and $x^0 \in \mathcal{M}$. The contingent cone of $\mathcal{M}$ at x^0 is denoted by $\Gamma^*(x^0, \mathcal{M})$ and is defined as: $d^0 \in \Gamma^*(x^0, \mathcal{M})$ iff there exists a sequence $\{\tau^k\}_{k\in\mathbb{N}} \subset \mathbb{R}$ and $\{d^k\}_{k\in\mathbb{N}} \subset \mathbb{R}^n$ such that

$$\tau^k \searrow 0, d^k \rightarrow d^0 \textit{ and } x^0 + \tau^k d^k \in \mathcal{M}, \forall k \in \mathbb{N}.$$

Definition 2.4.4 (inner tangent cone). Let $\mathcal{M}$ be any set and $x^0 \in \mathcal{M}$. The inner tangent cone of $\mathcal{M}$ at x^0 is denoted by $\Gamma(x^0, \mathcal{M})$ and is defined as: $d^0 \in \Gamma(x^0, \mathcal{M})$ iff there exists $\overline{\tau} > 0$ and a neighborhood $U(d^0)$ such that

$$x^0 + \tau d \in \mathcal{M}, \forall \tau \in (0, \overline{\tau}], d \in U(d^0).$$

For alternative definitions of inner and contingent cones (*outer tangent cones*) consult p. 33. of Hettich & Zencke [25].

The following statement has been a forerunner for first order conditions in [25] for (SIP) problems. And this same statement has been effectively used in [73].

Lemma 2.4.1 (Lemma 2.4 in [73], also Thm. 2.1.3 & pp. 38 - 39 in Hettich & Zencke [25]).

(i) If x^0 is a local minimizer of (GSIP), then

$$\{d \in \mathbb{R}^n \mid Df(x^0)d < 0\} \cap \Gamma^*(x^0, \mathcal{M}) = \emptyset. \tag{2.4.1}$$

(ii) x^0 is a strict local minimizer of order one for (GSIP) iff

$$\{d \in \mathbb{R}^n \mid Df(x^0)d \leq 0\} \cap \Gamma^*(x^0, \mathcal{M}) = \{0\} \tag{2.4.2}$$

The basic work done in [73] is in (approximately) describing the cones $\Gamma^*(x^0, \mathcal{M})$ and $\Gamma(x^0, \mathcal{M})$ in terms of the lower level problem of (GO(x)) of (GSIP). These have been achieved through the Lagrange function of the lower level problem (GO(x^0)) at the local optimal point x^0.

Given a point $x^0 \in \mathcal{M}$, define the indices of the *active constraints* of (GSIP) at x^0 by

$$E(x^0) := \{t \in B(x^0) \mid G(x^0, t) = 0\}.$$

Then $E(x^0)$ contains the set of global minimizer of the problem

$$(GO(x^0)) \qquad G(x^0, t) \to \inf$$
$$s.t.\ h_i(x^0, t) \leq 0, i \in I.$$

Thus, for a point $\bar{t} \in E(x^0)$, let $I(x^0, \bar{t})$ denote the set of active constraints of the lower level problem $GO(x^0)$ at $\bar{t}$ and be define as

$$I(x^0, \bar{t}) := \{i \in I \mid h_i(x^0, \bar{t}) = 0\}.$$

Furthermore, let the Lagrange function corresponding to x^0 and $\bar{t} \in E(x^0)$ be given as

$$\mathcal{L}(x^0, \bar{t}, \lambda) := \lambda_0 G(x^0, \bar{t}) + \sum_{i \in I(x^0, \bar{t})} \lambda_i h_i(x^0, \bar{t}), \tag{2.4.3}$$

where $\lambda \in \mathbb{R}^{|I(x^0, \bar{t})|+1}$.

Consequently, the well known Fritz-John Optimality condition, that $\bar{t} \in E(x^0)$ is a local minimizer of $GO(x^0)$, states that:

(i) there is a multiplier vector $\overline{\lambda} \in \mathbb{R} \times \mathbb{R}^{|I(x^0, t)|+1}$ with the property that

$$\overline{\lambda}_0 + \sum_{i \in I(x^0, \bar{t})} \overline{\lambda}_i = 1; \quad \lambda_0 \geq 0,\ \lambda_i \geq 0, i \in I(x^0, \bar{t}), \tag{2.4.4}$$

such that

(ii)

$$D_t\mathcal{L}(x^0,\bar{t},\bar{\lambda}) = 0. \tag{2.4.5}$$

Definition 2.4.5. The *set of Fritz-John multipliers* at $\bar{t} \in E(x^0)$ for $GO(x^0)$ is given by

$$\Lambda(x^0,\bar{t}) := \{\lambda \in \mathbb{R}^{|I(x^0,\bar{t})|+1} \mid \lambda \text{ satisfies (2.4.4) and (2.4.5) }\}.$$

Lemma 2.4.2 (Lem. 2.8., Stein[73]). *Let $x^0 \in \partial\mathcal{M} \cap \mathcal{M}$. Then for each $\bar{t} \in E(x^0)$ the set $\Lambda(x^0,\bar{t})$ is non-empty and compact.*

Remark 2.4.2. With the satisfaction of constraint qualifications w.r.t. the lower level problem (GO(x^0)), the set of multipliers $\Lambda(x^0,\bar{t})$ displays special structures. For instance, under the validity of (MFCQ) at $(x^0,\bar{t}), \bar{t} \in E(x^0)$, one can choose $\lambda_0 > 0$ in the definition of $\Lambda(x^0,\bar{t})$ (because of the relation (2.4.4)); there by obtaining a Kuhn-Tucker condition for the optimality of $\bar{t}$ w.r.t. $GO(x^0)$. Furthermore, under (MFCQ), Gauvin[17] proves that $\Lambda(x^0,\bar{t})$ is a convex Polytope (hence a compact set; in fact, regardless of the satisfaction of 2.4.4). If one further assumes the stronger constraint qualification (LICQ), then the set of multipliers $\Lambda(x^0,\bar{t})$ will be a singleton (cf. Kyparisis [41]).

Using the assumption (USC) on the SV-map $B(\cdot)$, we have, for each x^0, $E(x^0)$ is a compact set. This fact along with the compactness of $\Lambda(x^0,\bar{t})$ (Lem. 2.4.2) at $\bar{t} \in E(x^0)$ yield that

Proposition 2.4.3 (Lem. 3.3, Jongen *et al.* [33]).
Let $x^0 \in \partial(\mathcal{M}) \cap \mathcal{M}$. Then the set

$$V(x^0) := \left\{D_x\mathcal{L}(x^0,\bar{t},\bar{\lambda}) \mid \bar{t} \in E(x^0), \bar{\lambda} \in \Lambda(x^0,\bar{t})\right\} \tag{2.4.6}$$

is compact in $\mathbb{R}^n$.

Corollary 2.4.4. *The set $\{-Df(x^0)\} \cup V(x^0)$ is compact.*

To state algebraic equivalents to the local optimality conditions given in Lem. 2.4.1, Stein [70] used the description of the directional derivatives of the marginal function $v(\cdot)$ of (GO(x)) through Lagrange function as defined in (2.4.3).

Theorem 2.4.5 (Thm. 2.15(i), Stein [70]). *Let $x^0 \in \mathbb{R}^n$. The following inclusions holds true*

$$\left\{ d \in \mathbb{R}^n \;\middle|\; \max_{\bar{t} \in E(x^0)} \max_{\overline{\lambda} \in \Lambda(x^0,\bar{t})} D_x \mathcal{L}(x^0, \bar{t}, \overline{\lambda}) d > 0 \right\} \subset \Gamma(x^0, \mathcal{M})$$
$$\subset \Gamma^*(x^0, \mathcal{M}) \subset \left\{ d \in \mathbb{R}^n \;\middle|\; \max_{\bar{t} \in E(x^0)} \min_{\overline{\lambda} \in \Lambda(x^0,\bar{t})} D_x \mathcal{L}(x^0, \bar{t}, \overline{\lambda}) d \geq 0 \right\}$$

Corollary 2.4.6 (Thm. 2.15(iii), Stein [70]). *Let $x^0 \in \mathbb{R}^n$. If, for each $\bar{t} \in E(x^0)$, the set $\Lambda(x^0, \bar{t}) = \{\lambda(\bar{t})\}$ is a singleton, then the following inclusion holds*

$$\left\{ d \in \mathbb{R}^n \;\middle|\; \max_{\bar{t} \in E(x^0)} D_x \mathcal{L}(x^0, \bar{t}, \lambda(\bar{t})) d > 0 \right\} \subset \Gamma(x^0, \mathcal{M})$$
$$\subset \Gamma^*(x^0, \mathcal{M}) \subset \left\{ d \in \mathbb{R}^n \;\middle|\; \max_{\bar{t} \in E(x^0)} D_x \mathcal{L}(x^0, \bar{t}, \lambda(\bar{t})) d \geq 0 \right\}$$

Trivially, the assumption of Cor. 2.4.6 is satisfied, if the (LICQ) holds true w.r.t. each $\bar{t} \in E(x^0)$ (see Rem. 2.4.2).

Definition 2.4.6 (EMFCQ, Jongen. *et al.* [33], see also Stein & Still [76]). Let $x^0 \in \mathcal{M} \cap \mathcal{M}$. Then the Extended Mangasarian-Fromovitz Constraint Qualification (EMFCQ) is said to hold at x^0 if there is a vector $\xi \in \mathbb{R}^n$ such that

$$D_x \mathcal{L}(x^0, \bar{t}, \overline{\lambda}) \xi > 0, \forall \bar{t} \in E(x^0), \forall \overline{\lambda} \in \Lambda(x^0, \bar{t}).$$

Lemma 2.4.7 (see also Guerra & Rückmann[21]). *Let $x^0 \in \partial\mathcal{M} \cap \mathcal{M}$. Then the following hold true*

(i)

$$\left\{ d \in \mathbb{R}^n \;\middle|\; D_x \mathcal{L}(x^0, \bar{t}, \overline{\lambda}) d > 0, \text{for all } \bar{t} \in E(x^0), \text{ and all } \overline{\lambda} \in \Lambda(x^0, \bar{t}) \right\} \subset \Gamma(x^0, \mathcal{M})$$

(ii) If the (EMFCQ) holds at x^0, then

$$\left\{ d \in \mathbb{R}^n \;\middle|\; D_x \mathcal{L}(x^0, \bar{t}, \overline{\lambda}) d \geq 0, \text{for all } \bar{t} \in E(x^0), \text{ and all } \overline{\lambda} \in \Lambda(x^0, \bar{t}) \right\} \subset \Gamma^*(x^0, \mathcal{M})$$

Proof. (i) follows directly from the left hand-side inclusion in Thm. 2.4.5.

(ii) has been shown, by Guerra & Rückmann [21], to follow from a constraint qualification weaker than (EMFCQ).

□

In the standard semi-infinite case, the inclusions (i) and (ii) of Lem. 2.4.7 are actually equalities (cf. Hettich & Zencke [25]), but this is not always the case with (GSIP)‡. For instance, Stein [73] takes (EMFCQ) along with an additional assumption to verify equality in Lem. 2.4.7(ii). In any case, using (EMFCQ), one can easily reckon that

$$\left\{d \in \mathbb{R}^n \;\middle|\; D_x\mathcal{L}(x^0,\bar{t},\overline{\lambda})d \geq 0, \text{for all } \bar{t} \in E(x^0), \text{ and all } \overline{\lambda} \in \Lambda(x^0,\bar{t})\right\} \subset cl\Gamma(x^0,\mathcal{M}).$$

Nevertheless, according to Lem. 2.4.1, the inclusions given in Lem. 2.4.7(i) & (ii) are enough to set up algebraic primal optimality conditions.

Theorem 2.4.8 (Primal Optimality Conditions, Thm. 3.1 Stein & Still [76]). *Let $x^0 \in \partial\mathcal{M} \cap \mathcal{M}$.*

(i) If x^0 is a local minimizer of (GSIP), then the system

$$Df(x^0)d < 0; D_x\mathcal{L}(x^0,\bar{t},\overline{\lambda})d > 0, \text{ for all } \bar{t} \in E(x^0), \overline{\lambda} \in \Lambda(x^0,\bar{t})$$

has no solution.

(ii) If (EMFCQ) is satisfied at x^0 and x^0 is a strict local minimizer of order $\kappa = 1$ of (GSIP), then the system

$$Df(x^0)d \leq 0; D_x\mathcal{L}(x^0,\bar{t},\overline{\lambda})d \geq 0, \text{ for all } \bar{t} \in E(x^0), \overline{\lambda} \in \Lambda(x^0,\bar{t})$$

posses only the trivial solution.

The following corollary is a reiteration of Thm. 2.4.8(i) (in light of Lem. 2.4.7(i))

Corollary 2.4.9. *Let $x^0 \in \partial\mathcal{M} \cap \mathcal{M}$. If x^0 is a local minimizer of of (GSIP), then*

$$\left\{d \in \mathbb{R}^n \mid Df(x^0)d < 0\right\} \cap \left\{d \in \mathbb{R}^n \mid d^\top s > 0, \forall s \in V(x^0)\right\} = \emptyset.$$

‡This in fact refers to the general case. For instance, in the special case of Cor. 2.4.6 we obtain equality in Lem 2.4.7(ii) through (EMFCQ).

In other words, there is no $d \in \mathbb{R}^n$ such that

$$d^\top s > 0, \forall s \in \{-Df(x^0)\} \cup V(x^0).$$

The following statements are important in connecting primal and dual optimality conditions in the theory of (SIP).

Lemma 2.4.10 (Cheney [12], also Hettich & Zencke [25], Stein [70]).

(i) Let $S \subset \mathbb{R}^n$ be a non-empty and compact set. Then the inequality system

$$d^\top s < 0, \forall s \in S$$

is inconsistent for $d \in \mathbb{R}^n$ iff $0 \in conv(S)$.

(ii) Let $S \subset \mathbb{R}^n$ be arbitrary and $s_0 \in \mathbb{R}^n$. If s_0 posses a representation

$$s_0 = \sum_{i=1}^{n} \gamma_i s_i$$

with $s_i \in S$ linearly independent and $\gamma_i > 0, i = 1, \ldots, n$, then the inequality system

$$d^\top s_0 \leq 0,\ d^\top s \geq 0, \forall s \in S$$

has only the trivial solution $d = 0$. Moreover, if $|S| \leq n$, then the converse of this statement also holds true.

Thus, using Cor. 2.4.9 and Lem. 2.4.10 we obtain

Theorem 2.4.11 (Thm. 3.3(i), Stein [70], also [33, 60, 76]). *If $x^0 \in \partial\mathcal{M} \cap \mathcal{M}$ is a local minimizer of (GSIP), then there exist $\overline{t}^i \in E(x^0), \overline{\lambda}^i \in \Lambda(x^0, \overline{t}^i)$ and non-trivial multipliers $\alpha \geq 0, \mu_i \geq 0, i = 1, \ldots, n$ such that*

$$\alpha Df(x^0) - \sum_{i=1}^{n} \mu_i D_x \mathcal{L}(x^0, \overline{t}^i, \overline{\lambda}^i) = 0.$$

Proof. Using the compactness of $\{-Df(x^0)\} \cup V(x^0)$, by Cor. 2.4.4, the claim follows from Cor. 2.4.9 and Lem. 2.4.10 by applying Carathoédory's theorem. □

Corollary 2.4.12 (Thm. 3.3(i), Stein [70], also [33, 60, 76]). *If $x^0 \in \partial\mathcal{M} \cap \mathcal{M}$ is a local minimizer of (GSIP) and the set $\Lambda(x^0, \bar{t}) = \{\lambda(\bar{t})\}$ for each $\bar{t} \in E(x^0)$, then there exist $\bar{t}^i \in E(x^0)$ and non-trivial multipliers $\alpha \geq 0, \mu_i \geq 0, i = 1, \ldots, n$ such that*

$$\alpha Df(x^0) - \sum_{i=1}^{n} \mu_i D_x \mathcal{L}(x^0, \bar{t}^i, \lambda(\bar{t}^i)) = 0.$$

Kuhn-Tucker type optimality condition for (GSIP) are also obtained under the satisfaction of the (EMFCQ) at the local optimal point x^0.

Theorem 2.4.13 (Kuhn-Tucker Optimality Conditions). *If $x^0 \in \partial\mathcal{M} \cap \mathcal{M}$ is local minimizer of (GSIP) and the (EMFCQ) holds at x^0, then the multiplier α in Thm. 2.4.11 and Cor. 2.4.12 can be chosen to be equal to 1; i.e. $\alpha = 1$.*

Remark 2.4.3. In case when the index set map $B(\cdot)$ does not depend on x; i.e. $B(x) \equiv B$, the (GSIP) reduces to a (SIP). That is

$$B(x) \equiv B = \{t \in \mathbb{R}^n \mid h_i(t) \leq 0, i \in I\}.$$

Then, for $t \in E(x)$, the Lagrange function

$$\mathcal{L}(x, t, \lambda) = \lambda_0 G(x, t) + \sum_{i \in I_0(x,t)} \lambda_i h_i(x, t)$$

reduces to

$$\mathcal{L}(x, t, \lambda) = \lambda_0 G(x, t) + \sum_{i \in I_0(t)} \lambda_i h_i(t).$$

Consequently,

$$D_x \mathcal{L}(x, t, \lambda) = \lambda_0 D_x G(x, t).$$

In other words, the optimality conditions in Thms. 2.4.11 and 2.4.13 also contain first order optimality conditions for (SIP).

Considering strict local minimizers, one finds the following necessary and sufficient condition.

Theorem 2.4.14 (necessary and sufficient condition, Thm. 3.4. Stein [73], Thm. 3.3 [76]). *Let $x^0 \in \partial\mathcal{M} \cap \mathcal{M}$, (EMFCQ) holds at x^0 and $|E(x^0)| = q_0 \leq n$. Then for each $\overline{t}^l \in E(x^0)$, there exists $\overline{\lambda}^l \in \Lambda(x^0, \overline{t}^l)$ and multipliers $\mu_l > 0, l = 1, \ldots, q_0$, such that the vectors $D_x\mathcal{L}(x^0, \overline{t}^l, \overline{\lambda}^l), l = 1, \ldots, q_0$, are linearly independent and*

$$Df(x^0) - \sum_{l=1}^{q_0} \mu_l D_x\mathcal{L}(x^0, \overline{t}^l, \overline{\lambda}^l) = 0$$

iff x^0 is a strict local minimizer of (GSIP) of order $\kappa = 1$.

Proof. Uses Lem. 2.4.10(ii) (cf. Stein [70]). □

2.5 Second Order Optimality Conditions

Second order optimality conditions for (GSIP) were first given by Hettich & Still [23]. Recently, Rückmann & Shapiro [59] also came up with a generalization of those optimality conditions of Hettich & Still [23]. Basically, there are two major approaches in setting up second order conditions of optimality. One is based on reduction assumption, meaning that, when (GSIP) could be reduced to an equivalent finite problem in some neighborhood of a given local minimizer (cf. [23]); and the second, with out requiring local reducibility(cf. [23] & [59]).

2.5.1 Second Order Optimality with Local Reducibility

The reducibility of (GSIP) into a finite programming problem is of paramount importance from both numerical and theoretical point of view. In particular, knowledge of the existence of a finite non-linear optimization problem (NLP) which is equivalent to (GSIP), in some neighborhood of a given point $x^0 \in \mathcal{M}$, allows one to use suitable algorithms of (NLP) to solve (GSIP). In [23], such local reducibility, in a neighborhood of a local minimizer, has also been used to prove conditions of optimality.

Let $x^0 \in \mathcal{M}$. If $x^0 \in int\mathcal{M}$ (note that this also includes the case $E(x^0) = \emptyset$), then (GSIP) will be reduced locally (in some neighborhood of x^0) to an optimization problem without constraints. In this case, for x^0 to be a local minimizer the necessary second order condition takes the well known form: $Df(x^0) = 0$ and $D^2f(x^0)$ is positive semi-definite. Thus, for local reducibility of (GSIP), usually it is assumed that $E(x^0) \neq \emptyset$ and $x^0 \in \partial\mathcal{M} \cap \mathcal{M}$; i.e.local reducibility of (GSIP) to a finite problem in a neighborhood of a feasible boundary point. However, for $x^0 \in \partial\mathcal{M} \cap \mathcal{M}$, $B(\cdot)$ may or may not be l.s.c. at x^0. Hence, Hettich & Still [23] and Klatte [38] consider local reducibility with the satisfaction of (LICQ) at each point $\bar{t} \in E(x^0)$ w.r.t. $GO(x^0)$. Which implies that $B(\cdot)$ is lower semi-continuous at x^0. In this case, the ideas of reducibility are mainly generalizations of the corresponding issues for standard (SIP) (cf. Hettich & Zencke [25]). Conditions for local reducibility of (GSIP) without lower semi-continuity assumption on $B(\cdot)$ are considered by Stein [71].

Given $x^0 \in \partial\mathcal{M} \cap \mathcal{M}$ and $E(x^0) \neq \emptyset$, there are two fundamental ideas behind local reducibility. These ideas constitute what is known as the *Reduction Ansatz* (or the Reduction Assumption, cf. [71]) in the literature of (SIP) and (GSIP).

RA1. For each $\bar{t}^l \in E(x^0)$, there is a neighborhood $V(\bar{t}^l)$ such that $\bar{t}^l$ is a unique local minimizer of $GO(x^0)$ on $B(x^0) \cap V(\bar{t}^l)$. Consequently, the family $\{V(\bar{t}^l) \mid \bar{t}^l \in E(x^0)\}$ will generate an open covering of $E(x^0)$. By the compactness of $E(x^0)$ (due to the continuity of G and compactness of $B(x^0)$), there is a finite covering $\{V(\bar{t}^l) \mid l = 1, \ldots, q_0\}$ of $E(x^0)$; i.e. $E(x^0) \subset \bigcup_{l=1}^{q_0} V(\bar{t}^l)$. Since, each $\bar{t}^l$ is a unique local minimizer of $GO(x^0)$, it follows that $E(x^0) = \{\bar{t}^1, \ldots, \bar{t}^{q_0}\}$. Which is a finite set.

RA2. Let $E(x^0) = \{\bar{t}^1, \ldots, \bar{t}^{q_0}\}$ be a finite set. There exist open neighborhoods $U(x^0)$ of x^0 and $\widetilde{V}(\bar{t}^l)$ of $\bar{t}^l$, $l = 1, \ldots, q_0$; and (at least) continuous functions $t^l(\cdot) : U(x^0) \to \widetilde{V}(\bar{t}^l)$ with $t^l(x^0) = \bar{t}^l$ such that for each $l \in \{1, \ldots, q_0\}$ and for all $x \in U(x^0)$, $t^l(x)$ is the unique local minimizer of (GO(x)) in $B(x) \cap \widetilde{V}(\bar{t}^l)$. So that the following hold true:

- The functions $v^l(x) = G(x, t^l(x)), l = 1, \ldots, q_0$, are well defined and continuous on $U(x^0)$.
- The following equality holds

$$\mathcal{M} \cap U(x^0) = \{x \in U(x^0) \mid v^l(x) \geq 0, l = 1, \ldots, q_0\}.$$

  Meaning that, the feasible set $\mathcal{M}$ of (GSIP) is locally representable with finite number of constraints.

Definition 2.5.1 (reducibility). Let $x^0 \in \partial\mathcal{M} \cap \mathcal{M}$. If both (RA1) and (RA2) hold true in some neighborhood $U(x^0)$ of x^0, then (GSIP) is said to be *locally reducible* at x^0.

Consequently, if one could guarantee (RA1) and (RA2), then (GSIP) will be locally reducible. Thus, the following general assumption is required.

Assumption (SO): The functions f, G and $h_i, i \in I$ are twice continuously differentiable.

Recall that the Lagrange function (2.4.3)

$$\mathcal{L}(x, t, \lambda) = \lambda_0 G(x, t) + \sum_{i \in I(x,t)} \lambda_i h_i(x, t)$$

with $\lambda \in \mathbb{R}^{|I(x,\bar{t})|+1}$. For $x^0 \in \partial\mathcal{M} \cap \mathcal{M}$, each $\bar{t} \in E(x^0)$ is (global) minimizer of $(GO(x^0))$ with $I(x^0, \bar{t}) = \{i \in I \mid h_i(x^0, \bar{t}) = 0\}$. If the (LICQ) holds at $\bar{t} \in E(x^0)$, then there is a unique multiplier $\lambda(\bar{t}) \in \mathbb{R}_+^{|I(x^0,\bar{t})|}$ such that the KT-condition

$$D_t\mathcal{L}(x^0, \bar{t}, \lambda) = D_t G(x^0, \bar{t}) + \sum_{i \in I(x^0,\bar{t})} \lambda_i(\bar{t}) D_t h_i(x^0, \bar{t}) = 0$$

holds for $(GO(x^0))$.

Theorem 2.5.1 (Thm. 2.2 Hettich & Still [23], Prop. 2.2. Klatte [38]). *Let $x^0 \in \partial\mathcal{M} \cap \mathcal{M}$. Suppose the functions G and $h_i, i \in I$, are twice continuously differentiable. Furthermore, the following assumptions hold true*

(LICQ): *The (LICQ) holds at each* $\bar{t} \in E(x^0)$. *(Hence, there is a unique multiplier* $\lambda(\bar{t})$ *corresponding to each* $\bar{t} \in E(x^0)$*).*

(SSOSC): *For each* $\bar{t} \in E(x^0)$, *the* strong second order sufficient optimality condition *(SSOSC)*

$$\xi^\top D_{tt}\mathcal{L}(x^0, \bar{t}, \lambda(\bar{t}))\xi > 0, \forall \xi \in T(x^0, \bar{t}) \setminus \{0\} \tag{2.5.1}$$

holds at $(\bar{t}, \lambda(\bar{t}))$ *w.r.t.* $GO(x^0)$; *where*

$$T(x^0, \bar{t}) := \{\xi \in \mathbb{R}^m \mid \xi^\top D_t h_i(x^0, \bar{t}) = 0, i \in I_+(x^0, \bar{t})\}, \tag{2.5.2}$$

known as the tangent space *of* $B(x^0)$ *at* $\bar{t}$ *w.r.t.* $(GO(x^0))$, *and* $I_+(x^0, \bar{t}) := \{i \in I(x^0, \bar{t}) \mid \lambda_i(\bar{t}) > 0\}$.

Then $E(x^0) = \{\bar{t}^1, \ldots, \bar{t}^{q_0}\}$; *i.e.* $E(x^0)$ *is a finite set (with corresponding unique set of multiplier vectors* $\{\overline{\lambda}^l \mid l = 1, \ldots, q_0\}$*) and there exists a neighborhood* $U(x^0)$ *of* x^0 *and there are functions*

$$t^l : U(x^0) \to \mathbb{R}^m, l \in \{1, \ldots, q_0\};$$

such that for each $\bar{t}^l, l \in \{1, \ldots, q_0\}$,

1. $t^l(x^0) = \bar{t}^l$;

2. $t^l(x) \in B(x)$ *for* $x \in U(x^0)$;

3. $t^l(\cdot)$ *is Lipschitz-continuous and directionally differentiable in every direction* $\xi \in \mathbb{R}^n$ *on* $U(x^0)$;

4. *For each* $l \in \{1, \ldots, q_0\}$ *and each* $x \in U(x^0)$, $t^l(x)$ *is a unique local minimizer of (GO(x)) with a unique Lagrange multiplier* $\lambda^l(x)$. *Furthermore, the multiplier function* $\lambda^l(\cdot) : U(x^0) \to \mathbb{R}^{|I(x^0, \bar{t}^l)|}$ *is Lipschitz continuous and directionally differentiable in every direction* $\xi \in \mathbb{R}^n$ *on* $U(x^0)$;

5. *For each $l \in \{1, \dots, q_0\}$, the function*

$$v^l(x) := G(x, t^l(x))$$

is continuously differentiable and twice directionally differentiable in every direction $\xi \in \mathbb{R}^n$ and $Dv^l(x)$ is Lipschitz continuous on $U(x^0)$.

Theorem 2.5.2 (reducibility theorem, Thm. 2.2. Hettich & Still [23]). *Let $x^0 \in \partial\mathcal{M} \cap \mathcal{M}$. If the (LICQ) and the (SSOSC) assumptions in Thm. 2.5.1 hold true, then*

1. *The marginal value function v of (GO(x)) is the maximum of a finite number of continuously differentiable functions on $U(x^0)$; in particular,*

$$v(x) = \max_{1 \le l \le q_0} v^l(x), x \in U(x^0).$$

2. *The feasible set $\mathcal{M}$ of (GSIP) could be described in terms of a finite number of constraints in a neighborhood $U(x^0)$ of x^0; i.e.*

$$\begin{aligned}\mathcal{M} \cap U(x^0) &= \{x \in U(x^0) \mid v^l(x) \ge 0, k = 1, \dots, q_0\} \\ &= \{x \in U(x^0) \mid G(x, \bar{t}^l(x)) \ge 0, k = 1, \dots, q_0\}.\end{aligned}$$

In other words, (GSIP) is locally reducible at x^0 into the finite optimization problem

$$(GSIP_{red}) \qquad f(x) \to \min$$
$$s.t.\ x \in \{x \in U(x^0) \mid G(x, t^l(x)) \ge 0, l = 1, \dots, q_0\}.$$

Remark 2.5.1.

(i) In Thms. 2.5.1 and 2.5.2, the assumptions (LICQ) and (SSOSC) imply the lower semi-continuity of the map $B(\cdot)$ at x^0 (cf. Chap. 1, Prop. 1.2.7). Thus, the local reducibility given in Thm. 2.5.2 is based on an implicit lower semi-continuity assumption of $B(\cdot)$ at x^0. However, in Stein [72] one finds a local reduction approach at a feasible boundary point without making an apriori assumption on the lower semi-continuity of $B(\cdot)$ at x^0.

(ii) If, instead of the tangent space given by (2.5.2), one considers the cone

$$Z(x^0,\bar{t}) = \{\xi \in \mathbb{R}^m \mid D_t h_i(x^0,\bar{t})\xi = 0, i \in I_+(x^0,\bar{t}); D_t h_i(x^0,\bar{t})\xi \leq 0, i \in I(x^0,\bar{t})\backslash I_+(x^0,\bar{t})\},$$

then $Z(x^0,\bar{t}) \subset T(x^0,\bar{t})$. Hence, Rückmann and Shapiro [59] assumed the satisfaction of the sufficient second order condition

$$\xi^\top D_{tt}\mathcal{L}(x^0,\bar{t},\lambda(\bar{t})) > 0, \forall \xi \in Z(x^0,\bar{t}) \setminus \{0\},$$

which is obviously weaker than the (SSOSC). In fact, in [59] examples are given showing that, if in Thms. 2.5.1 and 2.5.2 (SSOSC) is replaced by the above weaker condition, then the local reducibility of (GSIP) may not hold true. Hence, the second order optimality conditions given for (GSIP), in [59], could be taken as a generalization of the one given by Hettich and Still [23].

(iii) Besides the reducibility of (GSIP) into a finite NLP, Weber [87, 84], Still [80] and Levitin [43] considered conditions for local and global reducibility of (GSIP) into a standard (SIP) (cf. also Levitin [42]). Such reduction approach has been successfully used by Weber [85, 86] in the study of structural and topological stability of the feasible set $\mathcal{M}$ of (GSIP) and, in Weber [84], for the derivation of first order optimality conditions for (GSIP).

Thm. 2.5.2(2) indicates that if the required differentiability assumptions are satisfied, then both first and second order optimality conditions, for local optimality of $x^0 \in \partial\mathcal{M} \cap \mathcal{M}$, could be obtained through the corresponding conditions of the finite NLP $(GSIP_{red})$. This is what has been done by Hettich & Still [23]. (see also Hettich & Still [24] for a recent review).

Theorem 2.5.3 (Thms. 3.2 & 3.3 in Hettich & Still [23], Thm. 4.2 in Rückmann & Sapiro [59]). *Suppose that the functions f, G and $h_i, i = 1,\ldots,p$ of (GSIP) are twice continuously differentiable. Let $x^0 \in \partial\mathcal{M} \cap \mathcal{M}$ and the (LICQ) and the (SSOSC) assumptions of Thm. 2.5.1 hold true w.r.t. $G(x^0)$.*

1. *Then if x^0 is a local minimizer of (GSIP), then for every $\xi \in \mathcal{K}(x^0)$, where*

$$\mathcal{K}(x^0) := \{\xi \in \mathbb{R}^n \mid Df(x^0)\xi \le 0; D_x\mathcal{L}(x^0, \overline{t}^l, \overline{\lambda}^l)\xi \ge 0, l = 1, \dots, q_0\},$$

there exists a vector $\mu \in \mathbb{R}^{q_0+1}$ of multipliers with $\mu_0 := \mu(\xi) \ge 0$ and $\mu_l := \mu_l(\xi) \ge 0, l = 1, \dots q_0$ not all equal to zero such that

$$\mu_0 Df(x^0) - \sum_{i=1}^{q_0} \mu_l D_x\mathcal{L}(x^0, \overline{t}^l, \overline{\lambda}^l) = 0 \tag{2.5.3}$$

and

$$\mu_0 \xi^\top D^2 f(x^0)\xi - \Theta(x^0; \xi, \mu) \ge 0, \tag{2.5.4}$$

where

$$\Theta(x^0; \xi, \mu) := \sum_{l=1}^{q_0} \mu_l \xi^\top D_{xx} G(x^0, \overline{t}^l)\xi + \sum_{l=1}^{q_0} \mu_l Dt^l(x^0; \xi) D_{tt}\mathcal{L}(x^0, \overline{t}^l, \overline{\lambda}^l) Dt^l(x^0; \xi)$$
$$+ \sum_{l=1}^{q_0} \mu_l \left[\sum_{i \in I(x^0, \overline{t}^l)} \left(\overline{\lambda}^l \xi^\top D_{xx} h_i(x^0, \overline{t}^l)\xi + 2D\lambda^l(x^0; \xi) D_x h_i(x^0, \overline{t}^l)\xi \right) \right].$$

Furthermore, if, for $l \in \{1, \dots, q_0\}$, $D_x\mathcal{L}(x^0, \overline{t}^l, \overline{\lambda}^l)\xi > 0$, then the multipliers μ_l can be chosen to be equal zero; and μ_0 could also be chosen as equal to zero if $Df(x^0)\xi > 0$.

2. *If for every $\xi \in \mathcal{K}(x^0)$, there exists $\mu(\xi) \in \mathbb{R}^{q_0+1}, \mu^\top(\xi) := \mu = (\mu_0, \dots, \mu_{q_0})$ with $\mu_0 \ge 0, \mu_l \ge 0, l = 1, \dots, q_0$, such that*

$$\mu_0 Df(x^0) - \sum_{i=1}^{q_0} \mu_l D_x\mathcal{L}(x^0, \overline{t}^l, \overline{\lambda}^l) = 0 \tag{2.5.5}$$

and

$$\mu_0 \xi^\top D^2 f(x^0)\xi - \Theta(x^0; \xi, \mu) > 0, \tag{2.5.6}$$

then x^0 is a local minimizer of (GSIP).

2.6 Numerical Solution Methods for GSIP

Owing to the existence of infinite number of constraints coupled with the variability of the index set of these constraints, the development of computational algorithms for (GSIP)s is not a straightforward matter. Consequently, a number of authors have attempted to transform or approximate (GSIP)s by problems of more manageable nature, for which there are well developed computational algorithms. To date, we have the following three major approaches:

- (global) transformation of a (GSIP) into a bi-level optimization problem (Stein [74, 75], Stein & Still [77, 78]);
- successive local linearization approach (Pickl & Weber [49], Weber [87]);
- a branch-and-bound method for a class of (GSIP) (Levitin & Tichatschke[47]).

Almost all of the above three approaches assume convexity structures in the lower level problem and also take the index set-valued $B(\cdot)$ as lower semi-continuous (thereby the feasible set $\mathcal{M}$ is a closed set). Among the three, we have computational experiments being reported by Stein [74, 75] and Stein & Still [77, 78]. Thus, here is given a brief review leaning towards the approaches of Stein & Still.

2.6.1 GSIP as a Bi-level Optimization Problem

In [74, 75, 77, 78] the possibility transforming a (GSIP) into an equivalent *Bi-level* optimization (BL) problem has been considered (for a more general and detailed discussions cf. Stein [74, 75]).

Consider the (GSIP)

$$(GSIP) \qquad f(x) \quad \to \quad inf$$
$$s.t. \quad G(x,t) \geq 0 \quad , \forall t \in B(x).$$

where

$$B(x) := \{t \in T \mid h_i(x,t) \leq 0, i \in I\}, I := \{1, \ldots, p\};$$

under the following usual assumption

Assumption (USC): $B(\cdot)$ is u.s.c. and compact valued on $\mathbb{R}^n$.

Now, define

$$F(x,t) := f(x)$$

and consider the *bi-level optimization problem* (BL)

$$(BL) \qquad \min_{x,t} F(x,t)$$
$$s.t. \quad G(x,t) \geq 0,$$

where t is a solution of the problem

$$(GO(x)) \qquad \min_t G(x,t)$$
$$s.t. \qquad h_i(x,t) \leq 0, i \in I.$$

The following notations are used:

- $\mathcal{M} = \{x \in \mathbb{R}^n \mid G(x,t) \geq 0, \forall t \in B(x)\}$;
- $E(x) := \{t \in \mathbb{R}^m \mid t \text{ is a global solution of } (GO(x))\}$;
- $S := \{(x,t) \mid t \in E(x)\} = GraphE(\cdot)$;
- $\mathcal{M}_G := \{(x,t) \in \mathbb{R}^n \times \mathbb{R}^m \mid G(x,t) \geq 0\}$;
- $\mathcal{M}_{BL} := \{(x,t) \mid (x,t) \in \mathcal{M}_G, t \in E(x)\} \subset \mathbb{R}^n \times \mathbb{R}^m$.

Remark 2.6.1. According to the definition of the problem (BL) (see Dempe [15]), the lower level problem (also known as the *follower's problem*) (GO(x)) need not have a unique solution for each feasible x; i.e. $E(x)$ need not be a singleton. However, it should be the case that $E(x) \neq \emptyset$, so that (BL) is well defined.

Thus, one needs to know the relations that do exist between the feasible sets and solutions sets of (GSIP) and (BL)[§]. Recall that, if $B(x) = \emptyset$, then $x \in int\mathcal{M}$. However, if $B(x) = \emptyset$, then $x \notin \mathcal{M}_{BL}$, by definition of (BL). In general, $Pr_x(\mathcal{M}_{\mathcal{BL}}) \subset \mathcal{M}$, where Pr_x represents the (canonical) projection into $\mathbb{R}^n$. Nevertheless, the equality of $Pr_x(\mathcal{M}_{\mathcal{BL}})$ and $\mathcal{M}$ is important if one intends to solve (GSIP) through (BL).

Proposition 2.6.1 (see Stein & Still [77]). *If for all $x \in \mathbb{R}^n$, $B(x) \neq \emptyset$ (i.e. $DomB(\cdot) = \mathbb{R}^n$), then $\mathcal{M} = pr_x(\mathcal{M}_{BL})$.*

Proof. Let $x \in \mathcal{M}$ be given. Since $G(x, \cdot)$ is continuous and $B(x)$ is compact and non-empty, there is $\bar{t} \in B(x)$ which is a solution of (GO(x)); i.e., $\bar{t} \in E(x)$ and $G(x, \bar{t}) \geq 0$. Hence, $(x, \bar{t}) \in \mathcal{M}_{BL}$. From which follows that $x \in Pr_x(\mathcal{M}_{BL})$. Hence $\mathcal{M} \subset Pr_x(\mathcal{M}_{BL})$. Similarly, if $x \in Pr_x(\mathcal{M}_{BL})$, then $(x, t^0) \in \mathcal{M}_{BL}$ for some $t^0 \in E(x)$ and $G(x, t^0) \geq 0$. Hence, t^0 is a global solution of (GO(x)) and $G(x, t^0) \geq 0$. That is

$$0 \leq G(x, t^0) \leq G(t, x), \forall t \in B(x).$$

Which implies that $x \in \mathcal{M}$. Hence, $Pr_x(\mathcal{M}_{BL}) \subset \mathcal{M}$, which completes the proof. □

Remark 2.6.2. Properly speaking, the problem (BL) is equivalent to the problem

$$(BL) \qquad \min_x F(x, t(x)) \qquad s.t. \quad G(x, t(x)) \geq 0, t(x) \in E(x),$$

where $E(x)$ is the set of global solutions of the problem

$$(GO(x)) \qquad \min_t G(x, t) \qquad s.t. \qquad h_i(x, t) \leq 0, i \in I.$$

In fact, this is the standard formulation of a bi-level optimization problem (see Dempe [15] and Shimizu *et al.* [69]). Hence, $Pr_x(\mathcal{M}_{BL}) = \{x \in \mathbb{R}^n \mid G(x, t(x)) \geq 0, t(x) \in E(x)\}$.

[§]For foundations on bi-level optimization problems consult the books of Dempe [15] and Shimizu *et al.* [69].

Accordingly, if $B(x) \neq \emptyset$, for all $x \in \mathbb{R}^n$, we have $\mathcal{M} = \{x \in \mathbb{R}^n \mid G(x, t(x)) \geq 0, t(x) \in E(x)\}$. In this sense, we can say that (GSIP) could be written into an equivalent bi-level optimization problem.

Rem. 2.6.2 reveals that (GSIP) could be taken as a special case of a bi-level optimization problem. Thus, all those solution methods for (BL) could generate solutions for (GSIP). It then remains to know : under what conditions does (BL) have a solution? Nevertheless, the feasible set $\mathcal{M}_{BL}$ may, in general, fail to be closed (see Examples 1 & 2 in [77]). The lack of closedness of $\mathcal{M}_{BL}$ may cause certain difficulties for the numerical methods designed for (BL). Thus, for $\mathcal{M}_{BL}$ to be closed the map $B(\cdot)$ needs to be lower semi-continuous. Note that in Prop. 2.6.1, the lower semi-continuity of $B(\cdot)$ is not required.

Proposition 2.6.2. *Let (USC) be satisfied and $B(x) \neq \emptyset$ for all $x \in \mathbb{R}^n$. If $B(\cdot)$ is a lower semi-continuous SV-map, then the feasible sets $\mathcal{M}$ and $\mathcal{M}_{BL}$ are closed.*

Proof. Note that $\mathcal{M}_{BL} = \mathcal{M}_G \cap S$. By the continuity of G on $\mathbb{R}^n \times \mathbb{R}^m$, the set $\mathcal{M}_G$ is closed. Moreover, by the assumptions (USC) and the lower semi-continuity of $B(\cdot)$, we have that $B(\cdot)$ is a continuous map. Which implies that the map $E(\cdot)$ is a closed set-valued map (Thm. 6.3.5, Shimizu *et al.* [69]). Consequently, $S = GraphE(\cdot)$ is a closed set (cf. Def. 1.1.7 and Rem. 1.1.1 in Chap. 1). Hence, $\mathcal{M}_{BL}$ is a closed set. Furthermore, $\mathcal{M}$ is closed set, since it is a continuous projection of a closed set. (cf. also Prop. 2.2.2 for closedness of $\mathcal{M}$). □

It is to be recalled that the map $B(\cdot)$ is lower semi-continuous if (MFCQ) holds (Prop. 1.2.7) w.r.t. (GO(x)), for each $x \in \mathbb{R}^n$.

Remark 2.6.3. Supposing $DomB(\cdot) = \mathbb{R}^n$, the set

$$\{(x, t) \in \mathbb{R}^n \times \mathbb{R}^m \mid h_i(x, t) \leq 0, i \in I\}$$

is compact; the (MFCQ) holds true at all points (x, t) where $t \in B(x)$; and the lower level problem (GO(x)) has a unique solution for each $x \in \mathbb{R}^n$, then the problem (BL) has global optimal solution (cf. Thm. 5.1, Dempe [15]).

For each fixed x, the requirement of the existence of a unique solution of the lower level problem (GO(x)) in Rem 2.6.3 could be actually guaranteed if (GO(x)) is a convex lower level problem and the (SSOSC) holds true (cf. Thm. 2.5.1). Furthermore, with the satisfaction the (LICQ) (along with other additional assumptions), Thm. 5.5 in [15] claims the existence of a local optimal solution for (BL). These are actually what have been assumed by Stein & Still in [74, 75, 77, 78].

Consequently, given the assumptions made above hold true, currently, one could apply any one of available algorithms for the treatment of (BL); so that one could determine (approximate) solutions for (GSIP). Stein & Still [78] introduced an interior point algorithm by transforming the (BL) into mathematical programming problems with equilibrium constraints. Convergence properties of the proposed algorithm is also given in Stein [74].

In this connection, Levitin & Tichatschke [47] have also suggested a Branch-and-bound method for a class of (GSIP), whose objective function is a *generalized supremum* function. In particular, the objective function is taken as the supremum function of a linear parametric optimization problem; where the lower level problem is also a linear parametric optimization problem. Their problem has the general form

$$\begin{array}{rcl} f_0(x) + h_0(x) & \to & inf \\ s.t. & & x \in \mathcal{M} := \left\{ x \in X \;\middle|\; \begin{array}{l} h_i(x) + \frac{1}{2}\left(A_i t_i\right)^\top t_i + \\ \left(b_i + G_i x\right)^\top t_i \le 0, \forall t_i \in B_i(x), i \in I \end{array} \right\} \end{array}$$

with

$$h_0(x^0) := \max_{t_0} \left\{ \frac{1}{2}\left(A_0 t_0\right)^\top t_0 + \left(b_0 + G_0 x\right)^\top t_0 \le 0, \forall t_0 \in B_0(x) \right\},$$

where

$$B_i(x) := \left\{ t_i \in T_i \;\middle|\; \begin{array}{l} p_{ij}^\top t_i + q_{ij}(x) \le 0, j \in J_{i1}; \\ p_{ij}^\top t_i + q_{ij}(x) \le 0, j \in J_{i2} \end{array} \right\}, i \in I \cup \{0\};$$

so that

- $X \subset \mathbb{R}^n$ is closed and convex;

- I; $J_{ij}, i \in I$ *and* $j \in \{1,2\}$, are finite index sets;
- $A_i : T_i \to T_i$ are symmetric negative definite linear operators in the Euclidean space T_i for each $i \in I$;
- $b_i, p_{ij} \in T_i, i \in I$ *and* $j \in \{1,2\}$, are vectors;
- $G_i : X \to T_i, i \in I \cup \{0\}$, are arbitrary linear operators;
- for each $i \in I$, the functions f_0, h_i and $-q_{ij}, j \in J_{i1}$ are convex on X and continuous on $\mathbb{R}^n$; and the functions $q_{ij}, j \in J_{i2}$ are affine on $\mathbb{R}^n$.

Hence, the objective function is, in general, non-convex (since $B_0(\cdot)$ is dependent on x and h_0 may fail to be convex). So that, the above (GSIP) could be treated as a non-convex (GSIP) with a convex lower level problem. Under certain duality assumptions, this (GSIP) has been shown in [47] to be re-written, equivalently, as a finite number of convex quadratic semi-infinite problems (see Thm. 4.1. p. 310 in [47]), which is rather a global transformation than a local one. Furthermore, a Branch-and-Bound algorithm is suggested to approximately determine a global optimal solution of the (GSIP) through these convex problems. Again it is to be observed that, the index set-valued maps $B_i(\cdot), i \in I$, are also continuous here (and have compact values) with in their respective spaces.

2.6.2 Iterative Linear Approximation of a GSIP

Pickl and Weber [49] considered a (GSIP) with index set-valued map $B(\cdot)$ defined in terms of affine equality and convex inequality constraints. Hence, for $x \in X$, let

$$E(x) := \{t \in B(x) \mid G(x,t) = 0\}$$

be the *active set-valued map* of (GSIP). Given $\overline{x} \in \mathcal{M}$ and a bounded open neighborhood U of $\overline{x}$, (GSIP) is considered under the assumption of the satisfaction of a variant of (EM-FCQ) at all points $x \in \mathcal{M} \cap cl(U)$. Furthermore, they gave a linearization of the function f at $\overline{x}$ and a corresponding linearization of G (i.e. a linearization of the feasible set $\mathcal{M}$)

locally at points $(\overline{x}, t)$ assuming (LICQ) is satisfied at each $t \in E(x)$ (i.e. at each solution t of the lower level problem (GO(x)), for each $x \in cl(U)$. By doing so they generated a sequence of locally approximative finite dimensional linear optimization problems to the (GSIP) constrained to U. The sequence of optimal solutions obtained from these linear programs has been also shown to converge to a global solution of the (GSIP) restricted to $cl(U)$. The theoretical foundation of this iterative procedure is deeply embedded in the study of the topological structure of the (GSIP). Observe, that in this approach the assumption of (LICQ) on $cl(U)$ leads to the lower semi-continuity of $B(\cdot)$ on $cl(U)$, which implies (with local uniform boundedness) that $B(\cdot)$ is continuous on $cl(U)$. Hence, it follows that $\mathcal{M} \cap cl(U)$ is a closed set. In other words, implicitly, the feasible set $\mathcal{M}$ has been taken to be locally closed.

Summing up, the theoretical and practical algorithms proposed so far for (GSIP), assume (or make assumptions that yield): the lower semi-continuity of the index set-valued map $B(\cdot)$ on the whole problem space; and the existence of a convex lower level problem . However, there are even practical problems that do not have convex lower level problems, such as the reverse-Chebychev Approximation Problem (Hoffmann & Reinhardt [29], Kaplan & Tichatschke [35]), as indicated by Stein [74]. Furthermore, if one drops the lower semi-continuity assumption on $B(\cdot)$, then one needs to have a different theoretical basis than the one traditionally accepted - at least to guarantee an approximate, in some sense, generalized solution of the (GSIP). Clearly then such an approximate solution could be infeasible, but might be assumed to belong to $cl(\mathcal{M})$. Hence, as presented in the Chapters 3 & 4, one such an approach would be to use the theory of robust analysis.

Chapter 3

Robustness of Set-Valued Maps and Marginal Value Functions

3.1 Introduction

The concept of robust sets and functions was first initiated by Chew & Zheng [13, 88] as a weakening of the continuity requirements of certain global optimization methods. Later on this theory was elaborated and extended by Shi, Zheng & Zhuang [66, 67, 68], Hoffmann, Phú, and Hichert [26, 27, 28, 48]. Furthermore, Zheng *et al.* [66] have also introduced robustness of general set-valued maps with the same purpose of weakening set-valued continuity - a concept which is tantamount to an *almost (semi-) continuity* property.

As was observed in chapter 2, the feasible set $\mathcal{M}$ of a (GSIP) may not be closed, as well as, the marginal function v of (GO(x)) may not be continuous if the index map $B(\cdot)$ is not lower semi-continuous. Furthermore, lack of convexity in the upper level problem may also entail a disjunctive structure in $\mathcal{M}$. Taking these difficult structures in a (GSIP) for granted, we would like to characterize them in this chapter.

Moreover, this chapter lays the theoretical background for the numerical approaches proposed for a class of (GSIP) in Chapter 4. The major aim here is to extend the theory of robust analysis to that of *robust analysis of marginal value functions* and *set-valued maps with given structures*. Consequently, relatively new results on

- approximatable functions (Sec. 3.3);
- robustness of marginal value functions (Sec. 3.5); and
- robustness of set-valued maps with given structures (Sec. 3.6)

have been presented. Moreover, attempt has been made to give the robust versions of some well-known and standard results of set-valued maps; there by pointing out connections, differences and similarities between robustness and continuity of such maps.

3.2 Preliminaries

We begin with basic definitions and results from robust analysis. The results mentioned in this section are mainly taken from Zheng *et al.* [13, 67, 88]. At the same time, minor complementary results have been supplied for the sake of later discussions.

Definition 3.2.1 (robust set, Zheng [88]).
Let X be a topological space and let $D \subset X$. Then D is called a *robust set* iff $clD = cl(intD)$.

Where clD and $intD$ denote the topological *closure* and the *interior* of D, resp., in the topology of X.

Remark 3.2.1. In Zheng [88] we find that $\emptyset$, X and open sets are robust, the union of an arbitrary collection of robust sets is again robust; the intersection of an open and a robust set is again robust. However, the intersection of two robust sets may not be robust.

Corollary 3.2.1 (see also Zheng [88]).
Let $D \subset X$. If D is convex (or star-shaped) and $intD \neq \emptyset$, then D is a robust set.

Definition 3.2.2 (robust point, Zheng [88]).

Let $D \subset X$. A point $x \in clD$ is said to be a *robust point to* D if $N(x) \cap intD \neq \emptyset$ for each neighborhood $N(x)$ of x. If, further $x \in D$, then x is said to be a *robust point of* D.

Proposition 3.2.2 (Zheng [88]).

1. *A set D is a robust subset of X if and only if each point $x \in D$ is a robust point of D.*

2. *Any accumulation point of a set of robust points to D is also a robust point to D.*

Moreover, an open set is a neighborhood of each of its points. Hence, robustness of a set is connected with a weaker notion of a neighborhood.

Definition 3.2.3 (semi-neighborhood (SNH), Zheng [88]). A set D is called a *semi-neighborhood* of a point x iff x is a robust point of the set D.

Corollary 3.2.3 (Zheng [88]). *A robust set D is a semi-neighborhood of each of its points.*

We also have the following properties, which we may frequently make use of:

Proposition 3.2.4 (Zheng [88]).

1. *If D is a semi-neighborhood of x and $intD \subset A$, then A is also a semi-neighborhood of x.*

2. *If D is a semi-neighborhood of x and $x \in O$, where O is an open set, then $D \cap O$ is also a semi-neighborhood of x.*

Remark 3.2.2. The union of a family of semi-neighborhoods of x is again a semi-neighborhood of x; whereas, intersection of two semi-neighborhoods of x may not be again a semi-neighborhood of x. Consequently, the collection of all semi-neighborhoods of a point x (or of robust sets) cannot define a topology.

Definition 3.2.4 (upper robust (u.r.) function, [88]).

A function $f : X \to \mathbb{R}$ is called *upper robust* (*u.r.*) *[upper semi-continuous]* on X iff for all $c \in \mathbb{R}$ the set

$$F_c := \{x \in X \mid f(x) < c\} =: [f < c]$$

is a robust [open] set.

The upper robustness of a function can also be defined pointwise in the traditional way.

Definition 3.2.5 (upper robustness at a point).

Let X be a topological space, $f : X \to \mathbb{R}$ and $x^0 \in X$. If for each given $\varepsilon > 0$ there is a semi-neighborhood $SNH_\varepsilon(x^0)$ of x^0 such that

$$f(x) \leq f(x^0) + \varepsilon, \forall x \in SNH_\varepsilon(x^0)$$

then f is said to be upper robust at x^0.

Proposition 3.2.5.

Let X be a topological space and $f : X \to \mathbb{R}$. Then f is upper robust at each $x \in X$ iff f is an upper robust function.

Proof. a) Suppose f is upper robust at each $x \in X$. Let $c \in \mathbb{R}$ be arbitrary, then we show that $F_c = \{x \in X \mid f(x) < c\}$ is a robust set. If $F_c = \emptyset$, then we are done. Thus, let $F_c \neq \emptyset$ and $x^0 \in F_c$ be any. Then $f(x^0) < c$. Choose ε such that $0 < \varepsilon < c - f(x^0)$. Then, by assumption, there is a semi-neighborhood $SNH_\varepsilon(x^0)$ such that

$$\forall x \in SNH_\varepsilon(x^0) : f(x) < f(x^0) + \varepsilon.$$

This implies that

$$x^0 \in SNH_\varepsilon(x^0) \subset F_c.$$

Since x^0 is a robust point of $SNH_\varepsilon(x^0)$, x^0 is a robust point of F_c (cf. Prop. 3.2.4(i)). Since $x^0 \in F_c$ is arbitrary, we conclude that F_c is a robust set. Hence, f is an upper robust function.

b) Suppose that f is an upper robust function. Let $x^0 \in X$ and $\varepsilon > 0$ be given. Then the set

$$SNH_\varepsilon(x^0) := \{x \in X \mid f(x) < f(x^0) + \varepsilon\}$$

contains x^0 and, by assumption, $SNH_\varepsilon(x^0)$ is a robust set. Consequently, $SNH_\varepsilon(x^0)$ is a semi-neighborhood of x^0 and

$$\forall x \in SNH_\varepsilon(x^0) : f(x) < f(x^0) + \varepsilon.$$

This yields that

$$\forall x \in SNH_\varepsilon(x^0) : f(x) \leq f(x^0) + \varepsilon.$$

Hence, f is upper robust at x^0. Since $x^0 \in X$ is arbitrary, we conclude that f is upper robust at each $x \in X$.

□

Among lots of properties of upper robust functions we find the following statements.

Corollary 3.2.6 (Zheng [88]). *Let $D \subset X$ and $f : D \to \mathbb{R}$. If D is a robust set and f is u.s.c., then f is u.r.*

The converse of Cor. 3.2.6 is not always true (cf. Zheng [88] for an example).

Definition 3.2.6 (lower robust (l.r.) function, Zheng [88]). A function $f : X \to \mathbb{R}$ is called *lower robust* (*l.r.*) on X iff $-f$ upper robust on X.

Proposition 3.2.7 (upper robustness of composition). *Let X be a topological space. If $f : X \to \mathbb{R}$ be an u.r. and $r : \mathbb{R} \to \mathbb{R}$ is a strictly increasing functions, then the composite function $r \circ f$ (i.e. $(r \circ f)(x) = r(f(x))$) is upper robust.*

Proof. Trivial! □

Remark 3.2.3. It is to be noted that, the concept of robustness (like openness) of sets depends on the topology of the underlying space. Likewise, robustness of a function is also dependent on the topologies of both its domain and images spaces. Hence, the definitions of robustness here are w.r.t. the relative topology on the set X, when X assumed to be a subset of some topological space. Such issues of *relative robustness* are discussed in Section 2.2.4 of [88].

3.3 Approximatable Functions

The notion of approximatability has been discussed in relation with robustness by Shi, Zheng & Zhuang [66, 67]. This concept would be seen to reveal the possibility of numerical approximation of the values of a robust function or SV-map at a given point. Roughly spoken, when a function is approximatable at a point x^0, then $f(x^0)$ could be approximated by those values of f at which it is continuous. The same holds true of SV-maps (cf. Sec. 3.4.5). In fact, the concept of approximatability reveals the practical usability of robustness for computational purposes (at least in optimization); especially, when robustness is guaranteed to be equivalent to approximatability. Hence, in this section, the definition of approximatable functions (of [67]) will be extended to that of *upper approximatable functions*. And a statement of equivalence between upper approximatable and upper robustness is also stated and proved.

We proceed by citing relevant definitions and results.

Definition 3.3.1 (robust function, Zheng *et al.* [67, 88]). Let $f : X \to Y$ and let $x \in X$. Then f is called robust at x iff for any neighborhood $U \subset Y$ of $y = f(x)$, x is a robust point of $f^{-1}(U)$.

Clearly,

Corollary 3.3.1. *If $f : X \to Y$ is continuous, then f is robust.*

Definition 3.3.2 (approximatable functions, [67]). Let X and Y be topological spaces and let $f : X \to Y$ be a function. Suppose that $S \subset X$ is the set of points of continuity of f. Then f is said to be *approximatable* iff

1. S is dense in X; and

2. for any $x^0 \in X$, there **exists** a net (a Moor-Smith sequence) $\{x_\alpha\}_{\alpha \in \Lambda} \subset S$ such that

$$\lim_{\alpha} x_\alpha = x^0 \quad \text{and} \quad \lim_{\alpha} f(x_\alpha) = f(x^0).$$

Definition 3.3.2 requires only the existence of a net to guarantee the approximatability, in contrast to its continuity counter part, where property 2. is expected to hold true for every net.

Corollary 3.3.2 (Thm. 2.1. in Zheng *et al.* [67]). *Any approximatable function is robust.*

The converse of Cor. 3.3.2, in general, may not be true. However, if X is a Baire space and Y is second countable, then approximatability is equivalent to robustness.

Proposition 3.3.3 (Thm 3.1. Zheng *et al.* [67]). *Let X be a Baire and Y be a second countable topological spaces and $f : X \to Y$. Then f is robust iff f is approximatable.*

Remark 3.3.1. If X is a complete metric space and $Y = \mathbb{R}$, then assumption of Prop. 3.3.3 will be easily satisfied. However, it should be stressed that, the result in Prop. 3.3.3 is based on general topological spaces. That is what to be exploited next.

Hence, we give below a generalization of Def. 3.3.2 in case when X is a metric space and $Y = \mathbb{R}$.

Definition 3.3.3 (upper approximatable functions). Let X be a metric space and $f : X \to \mathbb{R}$. Suppose that $S \subset X$ be the set of points where f is u.s.c. Then f is upper approximatable (u.a.) iff

1. S is dense in X; and

2. for any $x^0 \in X$, there is a sequence $\{x^k\} \subset S$ such that

$$\lim_{k} x^k = x^0 \quad \textit{and} \quad \limsup_{k} \; f(x^k) \le f(x^0).$$

We show next that upper approximatability implies upper robustness.

Proposition 3.3.4. *Let X be a metric space and $f : X \to \mathbb{R}$. If f is upper approximatable, then f is upper robust.*

Proof. Let $x \in \mathbb{R}$ be any and

$$[f, c] = \{x \in X \mid f(x) < c\}.$$

Take an arbitrary $x^0 \in [f, c]$; i.e. $f(x^0) < c$. We show that x^0 is a robust point of $[f, c]$. There are two cases to consider:

Case a: If f is u.s.c. at x^0, then for every $\varepsilon > 0$, there is a neighborhood $N(x^0)$ in X, such that

$$f(x) \leq f(x^0) + \varepsilon, \forall x \in N(x^0).$$

In particular taking ε with $0 < \varepsilon < c - f(x^0)$ we have

$$\forall x \in N(x^0) : f(x) < c.$$

This implies that $x^0 \in int([f, c])$. Hence, x^0 is a robust point of $[f, c]$.

Case b: If f is not u.s.c. at x^0, then there is a sequence $\{x^k\} \subset S$ such that

$$\lim_k x^k = x^0 \quad \text{and} \quad \limsup_k f(x^k) \leq f(x^0)$$

and $\limsup_k f(x^k)$ implies (by definition of lim sup) that there exists $k_0(\varepsilon)$

$$f(x^k) \leq f(x^0) + \varepsilon, \forall k \geq k_0(\varepsilon).$$

Choosing $0 < \varepsilon < c - f(x^0)$ it then follows that

$$x^k \in [f, c], \forall k \geq k_0(\varepsilon).$$

But then for each $k \geq k_0(\varepsilon)$, by Case(a), x^k is a robust point of $[f, c]$. Moreover, $\lim_k x^k = x^0$. Consequently, by Prop. 3.2.2(2), x^0 is a robust point of $[f, c]$. This completes the proof.

□

A statement of equivalence between upper robustness and upper approximatability could be given if X is assumed to be a complete metric space. For this we recall the standard definition of topological spaces.

Definition 3.3.4 (topological space, Royden [58]).
A topological space $< X, \tau >$ is a non-empty set X together with a family of subsets τ of X having the properties:

1. $X \in \tau$, $\emptyset \in \tau$;
2. $O_1, O_2 \in \tau$ implies $O_1 \cap O_2 \in \tau$; and
3. for any family $\{O_\alpha \mid \alpha \in \Omega\} \subset \tau : \bigcup_{\alpha \in \Omega} O_\alpha \in \tau$.

Proposition 3.3.5. *Let X be a complete metric space and $f : X \to \mathbb{R}$. Then f is upper robust iff f is upper approximatable.*

Proof. It remains to show only the forward implication (the revers implication is already contained in Prop. 3.3.4). Take the following family of subsets of $\mathbb{R}$

$$\beth := \{(-\infty, q) \mid q \in \mathbb{Q} \cup (-\infty, \infty)\}.$$

Obviously, $< \mathbb{R}, \beth >$ is a topological space with countable basis of open sets (but it is not Hausdorff). Moreover, $< \mathbb{R}, \beth >$ is a separable topological space (i.e. it satisfies the first axiom of countability). We now consider,

$$f :< X, \tau > \to < \mathbb{R}, \beth > .$$

Then the upper robustness of f in the usual topology of $\mathbb{R}$ is now the robustness of f w.r.t. the topology $\beth$ on $\mathbb{R}$ (cf. Def. 3.3.1). Hence, Thm. 3.1. of Zheng *et al.* [67] yields that f is approximatable w.r.t. $\beth$ in $\mathbb{R}$. That is, there exists a set $S \subset X$ such that

(i) f is continuous at each $x \in S$ w.r.t. $\beth$ on $\mathbb{R}$;

(ii) S is dense in X; and

(iii) for each $x^0 \in X$, there is a sequence $\{x^k\} \subset S$ such that

$$\lim_k x^k = x^0 \quad \text{and} \quad (\sqsupset) \lim_k f(x^k) = f(x^0).$$

(Observe that the limit w.r.t. $\sqsupset$ is not unique). We next formulate $(\sqsupset) lim_k f(x^k) = f(x^0)$ in the traditional notation. Thus, for any $\varepsilon > 0$, there exists $q(\varepsilon) \in \mathbb{Q}$, $f(x^0) < q(\varepsilon) < f(x^0) + \varepsilon$, such that $(-\infty, q(\varepsilon)) \in \sqsupset$ and $f(x^0) \in (-\infty, q(\varepsilon))$. Consequently, $(\sqsupset) lim_k f(x^k) = f(x^0)$ implies that there is $k_0(\varepsilon)$ such that

$$f(x^k) \in (-\infty, q(\varepsilon)) \subset (-\infty, f(x^0) + \varepsilon), \forall k \geq k_0(\varepsilon).$$

Hence,

$$f(x^k) < f(x^0) + \varepsilon, \forall k \geq k_0(\varepsilon).$$

From this follows that

$$\limsup_k f(x^k) \leq f(x^0) + \varepsilon.$$

It remains now to show that S contains the set of points of X where f is upper semi-continuous w.r.t. *the usual topology* on $\mathbb{R}$. Thus, let $x^0 \in S$, then by the continuity of $f :< X, \tau > \to < \mathbb{R}, \sqsupset >$ it follows that for any $\varepsilon > 0, \exists U(x^0) \subset X$ such that

$$f(x) \in (-\infty, f(x^0) + \varepsilon), \forall x \in U(x^0).$$

This concludes that

$$f(x) \leq f(x^0) + \varepsilon, \forall x \in U(x^0).$$

This is the usual upper semi-continuity of f at x^0. Hence, the claim is justified. □

3.4 Robustness of Set-Valued Maps

As an open set is to a robust set; a continuous map corresponds to a robust map. Thus, beginning with the basic definitions of robustness of SV-maps, one may like to find out: *how much of continuity could be re-writable in terms of robustness.* Furthermore, Section 3.6 refines this concept to set-valued maps with given structures.

3.4.1 Definitions and Results

For a set-valued map $M : X \rightrightarrows Y$ and $U \subset Y$, recall the definitions of upper inverse $M^{+1}(U)$ and lower inverse $M^{-1}(U)$ of the set U, as were given in Sec. 1.1 of Chap. 1.

Definition 3.4.1 (lower robust SV-map, Zheng *et al.* [66, 88]). Let X and Y be topological spaces and $M : X \rightrightarrows Y$ be a set-valued map. Then $M(\cdot)$ is *lower robust* [l.s.c.] at $x \in X$ iff for each $y \in M(x)$ and each neighborhood $U(y) \subset Y$ of y, $M^{-1}(U(y))$ is a semi-neighborhood [neighborhood] of x in X. The map $M(\cdot)$ is *lower robust* (l.r.) [l.s.c.] iff $M(\cdot)$ is lower robust [lower semi-continuous] at x, for all $x \in X$.

Corollary 3.4.1. *$M(\cdot)$ is l.r. iff $M^{-1}(U)$ is a robust set in X for every (nonempty) open set $U \subset Y$.*

Corollary 3.4.2. *If $M : X \rightrightarrows Y$ is l.s.c., then $M(\cdot)$ is l.r.*

But, the converse of Cor. 3.4.2 is not always true.

Example 3.4.3. *The set-valued map*

$$M(x) := \begin{cases} [1,4] & \text{if } x > 0, \\ \{4\} & \text{if } x = 0, \\ [2,3] & \text{if } x < 0, \end{cases}$$

is a simple example of a map which is l.r., but not l.s.c. at $x = 0$.

Definition 3.4.2 (upper robust SV-map, Zheng *et al.* [66, 88]).
Let X and Y be topological spaces and $M : X \rightrightarrows Y$ be a set-valued map.

1. The map $M(\cdot)$ is said to be upper robust (u.r.) [u.s.c.] at $x \in X$ iff for any neighborhood U of $M(x)$; *i.e.* $M(x) \subset U$, $M^{+1}(U)$ is a semi-neighborhood [neighborhood] of x. (i.e. x is a robust point of $M^{+1}(U)$).

2. The map $M(\cdot)$ is said to be upper robust [u.s.c.] iff $M(\cdot)$ is upper robust [u.s.c.] at every $x \in X$.

Correspondingly, we have

Corollary 3.4.4.

1. *The map $M(\cdot)$ is u.r. iff for any open set $U \subset Y$, $M^{+}(U)$ is a robust set in X.*
2. *If $M(\cdot)$ is an u.r. SV-map, then the set*

$$E := \{x \in X \mid M(x) = \emptyset\}$$

is robust in X. (cf. Chap. 1, Cor. 1.1.5 that E is an open set if $M(\cdot)$ is u.s.c.)

Corollary 3.4.5 (Zheng [88]). *If $M(\cdot)$ is u.s.c., then $M(\cdot)$ is u.r.*

Example 3.4.3 demonstrates that there is a l.r. set-valued map which is not l.s.c. A similar example could be set up for upper robustness. Furthermore, the SV-map in Example 3.4.3 is lower robust, but not upper robust. To see this, for $\varepsilon > 0$, we find that

$$M^{+1}((-\varepsilon, \varepsilon) + 4) = \{0\}.$$

Which shows that $M(\cdot)$ is not upper robust.

At this juncture one may pose the question: "How much discontinuous is a l.r. (u.r.) set-valued map?" One crude answer could be "a l.r. (u.r.) set-valued map is almost l.s.c. (u.s.c.)". A formal answer has been supplied by Zheng *et al.* [66]. (cf. sec. 3.4.4).

3.4.2 $\varepsilon-$Robustness of set-valued Maps

In the following, we would like to see how far the notions of Hausdorff or ε-semi-continuity could be carried over to that of robustness.

Hence, let X and Y be normed linear spaces, $M : X \rightrightarrows Y$ be a set-valued map, and denote by $\mathbf{B}_\varepsilon$ the open ball of radius ε at the zero element of Y, with $\varepsilon > 0$.

Definition 3.4.3 (ε-upper robust SV-map). We say that $M(\cdot)$ is ε-upper robust [ε-upper semi-continuous] at x^0 if given $\varepsilon > 0$, there exists a semi-neighborhood [neighborhood] $SNH_\varepsilon(x^0)$ such that

$$\forall x \in SNH_\varepsilon(x^0) : M(x) \subset M(x^0) + B_\varepsilon.$$

And $M(\cdot)$ is called an ε-upper robust [ε-upper semi-continuous] map, if it is ε-upper robust [ε-upper semi-continuous] at every $x^0 \in X$.

Proposition 3.4.6. *If $M(\cdot)$ is upper robust, then $M(\cdot)$ is ε-upper robust.*

Proof. Given $\varepsilon > 0$ and $x^0 \in X$, let $U := M(x^0) + B_\varepsilon$ (which is an open set). Hence, $M(x^0) \subset U$. By assumption $M^{+1}(U)$ is a semi-neighborhood of x^0. Set $SNH_\varepsilon(x^0) := M^{+1}(U)$. Thus

$$\forall x \in SNH_\varepsilon(x^0) : M(x) \subset U = M(x^0) + B_\varepsilon.$$

And that completes the proof. □

Proposition 3.4.7. *If $M(\cdot)$ is compact valued and ε-upper robust, then $M(\cdot)$ is upper robust.*

Proof. Let $U \subset Y$ be an open set. We need to show that $M^{+1}(U)$ is a robust set in X. Let $x^0 \in M^{+1}(U)$. Hence, $M(x^0) \subset U$. This implies that, for each $y \in M(x^0)$, there exists $\varepsilon(y) > 0$ such that $B_\varepsilon(y) \subset U$. But, since $M(x^0)$ is compact, there are $y_1, \dots, y_m \in M(x^0)$ and

$$M(x^0) \subset \bigcup_{i=1}^{m} B_{\varepsilon(y_i)}(y_i) \subset U.$$

Let $\varepsilon_0 := \min_{1 \le i \le m} \varepsilon(y_i)$. From this follows that

$$M(x^0) + B_{\varepsilon_0} \subset U.$$

By ε-upper robustness, there is a semi-neighborhood $SNH(x^0)$ such that

$$\forall x \in SNH(x^0) : M(x) \subset M(x^0) + B_{\varepsilon_0}.$$

Consequently,

$$SNH(x^0) \subset M^{+1}(M(x^0) + B_{\varepsilon_0}) \subset M^{+1}(U).$$

Therefore, x^0 is a robust point of $M^{+1}(U)$. Since $x^0 \in M^{+1}(U)$ is arbitrary, we conclude that $M(\cdot)$ is upper robust. □

Similarly, we define

Definition 3.4.4 (ε-lower robust SV-map).

We say that $M(\cdot)$ is *ε-lower robust at* x^0 iff for any $\varepsilon > 0$ there exists a semi-neighborhood $SNH_\varepsilon(x^0)$ such that

$$\forall x \in SNH_\varepsilon(x^0) : M(x^0) \subset M(x) + B_\varepsilon.$$

We say that $M(\cdot)$ is ε-lower robust, if it is ε-lower robust at every $x^0 \in X$.

Proposition 3.4.8. *If $M(\cdot)$ is $\varepsilon-$lower robust, then $M(\cdot)$ is lower robust.*

Proof. Let $U \subset Y$ be an open set. We show that $M^{-1}(U)$ is a robust set in X; i.e., we show for arbitrary $x^0 \in M^{-1}(U)$, x^0 is a robust point of $M^{-1}(U)$. But then, $M(x^0) \cap U \neq \emptyset$. Which implies, there is $y^0 \in M(x^0) \cap U$. Hence, for some $\varepsilon > 0$, we have

$$B_\varepsilon(y^0) \subset U \textit{ and } M(x^0) \cap B_\varepsilon(y^0) \neq \emptyset.$$

By ε-lower robustness, there is a semi-neighborhood $SNH_\varepsilon(x^0)$ of x^0 such that

$$\forall x \in SNH_\varepsilon(x^0) : M(x^0) \subset M(x) + B_\varepsilon.$$

From this follows that

$$\forall x \in SNH_\varepsilon(x^0) : y^0 \in M(x) + B_\varepsilon.$$

Consequently,

$$\forall x \in SNH_\varepsilon(x^0) : y^0 \in M(x) \cap B_\varepsilon(y^0) \neq \emptyset.$$

Since $B_\varepsilon(y^0) \subset U$, we have

$$\forall x \in SNH_\varepsilon(x^0) : y^0 \in M(x) \cap U \neq \emptyset.$$

That is

$$SNH_\varepsilon(x^0) \subset M^{-1}(U).$$

Hence, $M^{-1}(U)$ is also a semi-neighborhood of x^0. Since x^0 is arbitrary, it follows that $M^{-1}(U)$ is a robust set; therefore, $M(\cdot)$ is a lower robust map. □

The converse of prop. 3.4.8 may not hold true even if $M(\cdot)$ is compact valued. Hence, a similar statement of equivalence, as in the case of l.s.c set-valued maps with compact values (see p. 45, paragraph 3, of Aubin & Cellina [4]), fails to exist between lower robust and ε-lower robust set-valued maps.

Example 3.4.9. *Consider the set-valued map $M: \mathbb{R} \rightrightarrows \mathbb{R}$ given by*

$$M(x) := \begin{cases} [2,5] & \text{if } x < 0, \\ [1,2] \cup [3,5] & \text{if } x = 0, \\ [1,3] & \text{if } x > 0. \end{cases}$$

Let $\varepsilon = \frac{1}{2}$. For any semi-neighborhood $SNH(0)$ and neighborhood $N(0)$ of 0, there is $x \in SNH(0) \cap N(0)$. Hence, if $x > 0$, we have $M(x) = [1,3]$, but then $M(0) = [1,2] \cup [3,5] \not\subset [1,3] + (-\frac{1}{2}, \frac{1}{2})$. Similarly, if $x < 0$, we have $M(x) = [2,5]$, so that $M(0) = [1,2] \cup [3,5] \not\subset [2,5] + (-\frac{1}{2}, \frac{1}{2})$. Consequently, $M(\cdot)$ is both not l.s.c. and not ε-lower robust at $x = 0$.

Obviously, $M(\cdot)$ is lower robust (also l.s.c) at x, for either $x < 0$ or $x > 0$. And, if $y^0 \in M(0)$, then either $y^0 \in [1,2]$ or $y^0 \in [3,5]$. Hence, for any neighborhood $B_\varepsilon(y^0)$, we have

- *if $y^0 \in [1,2]$, then $M^{-1}(B_\varepsilon(y^0)) = [0, \infty)$; or*
- *if $y^0 \in [3,5]$, then $M^{-1}(B_\varepsilon(y^0)) = (-\infty, 0]$.*

In both cases, $M^{-1}(B_\varepsilon(y^0))$ is a semi-neighborhood of $x = 0$. Consequently, $M(\cdot)$ is a lower robust SV-map with compact values.

3.4.3 Further Examples of Robust set-valued Maps

Additional examples are given below, to make the differences and similarities of robustness and continuity of set-valued maps more transparent. The point of interest in the examples is $x = 0$.

<table>
<tr>
<td>Example 3.4.10.
$M(x) = \begin{cases} [1,4] \text{ if } x < 0, \\ [2,3] \text{ if } x \geq 0. \end{cases}$
Then $M(\cdot)$ is
• not u.s.c., but u.r.;
• l.s.c.; hence l.r.</td>
<td>Example 3.4.11.
$M(x) = \begin{cases} [1,4] \text{ if } x \leq 0, \\ [2,3]\,, \text{ if } x > 0. \end{cases}$
Then $M(\cdot)$ is
• u.s.c.; hence u.r.;
• not l.s.c., but l.r.</td>
<td>Example 3.4.12.
$M(x) = \begin{cases} [1,4] \text{ if } x \neq 0, \\ [2,3]\,, \text{ if } x = 0. \end{cases}$
Then $M(\cdot)$ is
• not u.s.c.; while not u.r.;
• l.s.c.; hence l.r.</td>
</tr>
<tr>
<td>Example 3.4.13.
$M(x) = \begin{cases} [1,4] \text{ if } x = 0, \\ [2,3] \text{ if } x \neq 0. \end{cases}$
Then $M(\cdot)$ is
• u.s.c.; hence u.r.;
• not l.s.c.; while not l.r.</td>
<td>Example 3.4.14.
$M(x) = \begin{cases} [2,4] \text{ if } x \leq 0, \\ [1,3] \text{ if } x > 0. \end{cases}$
Then $M(\cdot)$ is
• not u.s.c., but u.r.;
• not l.s.c., but l.r.</td>
<td></td>
</tr>
</table>

3.4.4 Piecewise Semi-continuous set-valued Maps

Again following Zheng *et al.* [66] we define piecewise semi-continuity. Analogously, as a sort of extension, we also consider *piecewise robustness* properties for set-valued mappings (and of functions in Sec. 3.5.3). Here we have the property that piecewise robustness implying robustness, which is not true of semi-continuity. Thus some suitable decomposition of the domain space is possible under the weaker robustness assumptions. Such a decomposition will be important in characterizing disjunctive structures that could crop in a (GSIP), due to the arising and vanishing of components of the index map $B(\cdot)$.

Definition 3.4.5. Let X and Y be two topological spaces. We say that $X_1, X_2, \ldots, X_r$ is a *partition of* X iff the sets X_i are pairwise disjoint and X is the union of all X_i. The *partition is called robust* iff each X_i is robust w.r.t. X.

Definition 3.4.6. A set-valued map $M : X \rightrightarrows Y$ is said to be *piecewise l.s.c. (l.r.) [u.s.c.] {u.r.}* iff there exists a robust partition $X_1, X_2, \ldots, X_r$ of X such that for all $i \in \{1, \ldots, r\}$ the restriction of $M(\cdot)$ to X_i is *l.s.c.* (*l.r.*) [*u.s.c.*] {*u.r.*} with respect to the relative topology of X_i induced by the topological space X.

The proofs of the following two theorems are not available in their original source Zheng *et al.* [66]. Hence, they are supplied here because of their simplicity.

Theorem 3.4.15 (Zheng *et al.* [66]). *If $M(\cdot)$ is piecewise l.s.c., then $M(\cdot)$ is l.r.*

Proof. Let $U \subset Y$ be any open set. We want to show that $M^{-1}(U)$ is a robust set in X. Since, for each $i = 1, \ldots, r$, $M|_{X_i} : X_i \rightrightarrows Y$ is l.s.c. with respect to the relative topology of X_i, we have that $M^{-1}(U) \cap X_i$ is a relatively open set in X_i. Hence, there is an open set $V \subset X$ such that $X_i \cap V = M^{-1}(U) \cap X_i$. But then $X_i \cap V$ is a robust set in X by Rem. 3.2.1. Hence, for each $i \in \{1, \ldots, r\}$, $M^{-1}(U) \cap X_i$ is a robust set. From which follows that

$$\bigcup_{i=1}^{r} M^{-1}(U) \cap X_i = M^{-1}(U)$$

is a robust set in X. □

Similarly, we have

Theorem 3.4.16 (Zheng *et al.* [66]). *If $M(\cdot)$ is piecewise u.s.c., then $M(\cdot)$ is u.r.*

Proof. Let $V \subset Y$ be any open set. We show that $M^{+1}(V) = \{x \in X \mid M(x) \subset V\}$ is a robust set in X. Since, $M : X_i \rightrightarrows Y$ is u.s.c. in the relative topology of X_i, we have, for each i,

$$(M|_{X_i})^{+1}(V) = \{x \in X_i \mid M(x) \subset V\}$$

is a relatively open set in X_i. The rest of the proof is as in Thm. 3.4.15. □

Lemma 3.4.17. *Let X be a topological space and A be a non-empty robust subset of X. If $B \subset A$ is such that $intB_A \neq \emptyset$, then $intB_X \neq \emptyset$, where $intB_A$ is interior of B relative to the topology of A induced by X.*

Proof. Clearly, $intB_A$ is an open set in A. Hence, there exists $O \subset X$ open in X such that $B \supset intB_A = O \cap A$. Since A is robust in X and $O \cap A \neq \emptyset$ (while $intB_A \neq \emptyset$ and $A \neq \emptyset$) we have that $O \cap intA_X \neq \emptyset$. This yields $B \supset intB_A = O \cap A \supset O \cap intA_X \neq \emptyset$ and $intB_X \supset O \cap intA_X \neq \emptyset$ which completes the proof. □

If the set $A \subset X$ is not assumed to be robust, then the above implication fails to be true. Take for e.g. $X = \mathbb{R},\ A = B = \mathbb{Q}$ ($\mathbb{Q}$ - set of rational numbers). Observe that $intB_A \neq \emptyset$. However, $intB_X = \emptyset$ and $A = \mathbb{Q}$ is not robust in $X = \mathbb{R}$.

Lemma 3.4.18. *Let X_0 be a robust subset of a topological space X and assume that $\hat{X} \subset X_0$. If $\hat{X}$ is robust in X_0 in the relative topology of X_0 w.r.t. X, then $\hat{X}$ is a robust set in X.*

Proof. Let $x \in \hat{X}$ and $N(x)$ be any open neighborhood of x w.r.t. X. Then $N(x) \cap X_0$ is neighborhood of x in the relative topology of X_0. Since $\hat{X}$ is robust in the relative topology of X_0 we have

$$int\left(N(x) \cap X_0 \cap \hat{X}\right)_{X_0} \neq \emptyset$$

(note that $X_0 \cap \hat{X} = \hat{X}$) and

$$int\left(N(x) \cap X_0 \cap \hat{X}\right)_{X_0} \subset N(x) \cap X_0.$$

Since $N(x)$ is open in X, $N(x) \cap X_0$ is robust in X (cf. Rem. 3.2.1). Hence, we get, by Lem. 3.4.17, that

$$int(N(x) \cap X_0 \cap \hat{X}) \neq \emptyset.$$

Since $N(x)$ is arbitrary, it follows that x is a robust point of $\hat{X}$. Therefore, using Prop. 3.2.2, X is robust set. □

Theorem 3.4.19. *If $M(\cdot)$ is piecewise l.r. [u.r.], then $M(\cdot)$ is l.r. [u.r.].*

Proof. Taking $M(\cdot)$ piecewise-l.r. and $U \subset Y$ as an open set, we have to show that

$$M^{-1}(U) = \{x \in X \mid M(x) \cap U \neq \emptyset\}$$

is a robust set. Since, $M : X_i \rightrightarrows Y$ is l.r. in the relative topology of X_i, we have, for each i

$$(M|_{X_i})^{-1}(U) = \{x \in X_i \mid M(x) \cap U \neq \emptyset\}$$

is a robust subset of X_i in the relative topology of X_i w.r.t. X. The rest of the proof follows by a similar argument as in Thm. 3.4.15 using Lem. 3.4.18. □

3.4.5 Approximatable Set-Valued-maps

Corresponding to approximatability of functions (cf. Sec. 3.3), there are results related with approximatability of SV-maps.

Definition 3.4.7 (lower approximatable SV-map, Def. 2.4. Zheng *et al.* [66]).

Let X and Y be topological spaces and $M : X \rightrightarrows Y$ be a SV-map. Suppose that S is the set of points of lower semi-continuity of $M(\cdot)$. Then $M(\cdot)$ is called *lower approximatable* iff

1. S is dense in X; and

2. for any $x^0 \in X$ and $y^0 \in M(x^0)$, there exist a net $\{x_\alpha\}_{\alpha\in\Lambda} \subset S$ and a net $\{y_\alpha\}_{\alpha\in\Lambda}$ with $y_\alpha \in M(x_\alpha)$ for every $\alpha \in \Lambda$ such that

$$\lim_\alpha x_\alpha = x^0 \quad \textit{and} \quad \lim_\alpha y_\alpha = y^0.$$

Comparing the definition for lower semi-continuous SV-map (in Chap. 1, Def. 1.1.6) with lower approximatability in metric spaces X and Y, we observe that Def. 3.4.7 requires only the existence of sequences. It then immediately follows that, a lower semi-continuous SV-map is lower approximatable. Furthermore, we have

Proposition 3.4.20 (Thm. 2.1 Zheng *et al.* [66]). *Any lower approximatable set-valued map is lower robust.*

Definition 3.4.8 (upper approximatable SV-map, Zheng *et al.* [66]). Let X and Y be topological spaces and $M: X \rightrightarrows Y$ be a SV-map. Suppose that S is the set of points of upper semi-continuity of $M(\cdot)$. Then $M(\cdot)$ is called *upper approximatable* iff

1. S is dense in X;

2. for any $x^0 \in X$ there is a net $\{x_\alpha\}_{\alpha\in\Lambda}$ in S such that for each neighborhood U of x^0 there is some $\alpha(U)$ with

$$\forall \alpha \in \alpha(U) : M(x_\alpha) \subset U.$$

Proposition 3.4.21. *Any upper approximatable set-valued map is upper robust.*

Proof. Follows by a similar argument as for Prop. 3.4.20 (see the proof of Thm. 2.1. in [66]). □

It has been indicated by Zheng *et al.* [66, 67] that robustness of SV-maps is weaker than approximatability. However, if a lower or upper robust map has a dense set of upper or lower semi-continuity, then it will be approximatable (cf. Prop. 2.3. in [66]).

3.5 Marginal Value Functions

3.5.1 Upper Robustness of Infimum

Let us next come to the investigation of the behavior of marginal functions w.r.t. robustness properties of its defining data. Hence, we first consider the marginal function φ defined by

$$\varphi(x) := \inf_{y \in M(x)} \psi(x, y). \tag{3.5.1}$$

Theorem 3.5.1 (upper robustness of infimum). *Let $\psi : X \times Y \to \mathbb{R}$ be u.s.c. on $\{x^0\} \times M(x^0)$, where $M : X \rightrightarrows Y$ is a l.r. set-valued map, then φ is u.r. at x^0.*

Proof. Let $c \in \mathbb{R}$ and $x^0 \in \Phi_c := \{x \mid \varphi(x) < c\}$. We have to show that Φ_c is a semi-neighborhood of x^0. Let $\varepsilon > 0$ be arbitrary, then there exists $\overline{y}_\varepsilon \in M(x^0)$ such that $\psi(x^0, \overline{y}_\varepsilon) < \varphi(x^0) + \varepsilon$. Since ψ is u.s.c. on $\{x^0\} \times M(x^0)$, there exist open neighborhoods $N(x^0)$ of x^0 and $N(\overline{y}_\varepsilon)$ of $\overline{y}_\varepsilon$ such that

$$\forall x \in N(x^0), \forall y \in N(\overline{y}_\varepsilon) : \psi(x, y) \leq \psi(x^0, \overline{y}_\varepsilon) + \varepsilon.$$

Moreover, $M(\cdot)$ is l.r., $\overline{y}_\varepsilon \in M(x^0)$ and $N(\overline{y}_\varepsilon)$ is a neighborhood of $\overline{y}_\varepsilon$ imply that $M^{-1}(N(\overline{y}_\varepsilon))$ is a semi-neighborhood of x^0. Hence, $Q := N(x^0) \cap M^{-1}(N(\overline{y}_\varepsilon))$, by Proposition 3.2.4(2), is a semi-neighborhood of x^0, too. Thus, we have for all $x \in Q$, $\widetilde{y} \in M(x) \cap N(\overline{y}_\varepsilon)$

$$\varphi(x) = inf_{y \in M(x)} \psi(x, y) \leq \psi(x, \widetilde{y}) \leq \psi(x^0, \overline{y}_\varepsilon) + \varepsilon < \varphi(x^0) + 2\varepsilon.$$

Choosing $\varepsilon > 0$ such that $0 < 2\varepsilon < c - \varphi(x^0)$, we have $Q \subset \Phi_c$. Hence Φ_c is a semi-neighborhood of x^0. □

Corollary 3.5.2 (lower robustness of supremum). *Let $\psi : X \times Y \to \mathbb{R}$ be l.s.c. on $\{x^0\} \times M(x^0)$, where $M : X \rightrightarrows Y$ is a l.r. set-valued map, then the marginal value function*

$$\phi(x) := \sup_{y \in M(x)} \psi(x, y)$$

is l.r. at x^0.

Proof. The claim follows trivially if we write

$$-\phi(x) := \inf_{y \in M(x)} -\psi(x,y)$$

and observe that $-\phi$ is upper robust by Thm. 3.5.1. From which follows that ϕ is lower robust. □

Corollary 3.5.3. *Let (X,ρ) be a metric space and $M : X \rightrightarrows Y$ be a set-valued map. If $M(\cdot)$ is l.r., $r : \mathbb{R}_+ \to \mathbb{R}_+$ is continuous and strictly increasing on $\mathbb{R}_+$, then the function*

$$\varphi(x) = r\left(dist(x, M(x))\right) := \inf_{\xi \in M(x)} r\left(\rho(x,\xi)\right)$$

is u.r.

Proof. The functions ρ and r are continuous and $\inf_{\xi \in M(x)} r\left(\rho(x,\xi)\right) = r\left(\inf_{\xi \in M(x)} \rho(x,\xi)\right)$. Then, using Thm. 3.5.1 and Prop. 3.2.7 the claim follows. □

This corollary guarantees that, when r is as above, X is normed space, $\psi : X \times X \to \mathbb{R}_+$, $\psi(x,\xi) := r\left(\|x - \xi\|\right)$ and the map $M(\cdot)$ is l.r., then the marginal function $\varphi(x) = dist(x, M(x))$ is u.r. In general (see Remark 3.2.1), the upper semi-continuity assumption on ψ cannot be replaced by upper robustness.

Hu & Papageorgiou [32] stated and proved the following.

Proposition 3.5.4 (Prop. 2.26, Hu & Papageorgiou [32]).
If Y is a metric space, $M : X \rightrightarrows Y$ and $M(x) \neq \emptyset, \forall x \in X$. Then $M(\cdot)$ is l.s.c. if and only if for every fixed $\xi \in Y$, the function $\varphi_\xi : X \to \mathbb{R}$, $\varphi_\xi(x) := dist(\xi, M(x))$ is u.s.c.,

where

$$\varphi_\xi(x) = \inf_{y \in M(x)} \rho(\xi, y).$$

A similar statement of equivalence will be

Proposition 3.5.5. *Let X and Y be metric spaces, with a metric ρ on X, $M : X \rightrightarrows Y$ and $M(x) \neq \emptyset, \forall x \in X$. Then $M(\cdot)$ is l.r. if and only if for every fixed $\xi \in Y$, the function $\varphi_\xi : X \to \mathbb{R}$, $\varphi_\xi(x) := dist(\xi, M(x))$ is u.r.*

Proof. The forward implication follows from Thm. 3.5.1 with $\psi(x, y) = \rho(x, y)$. To show the backward implication (which follows with some modification of the proof of Prop. 2.26, in [32]), we let $V \subset Y$ be any open set. We need to show that $M^{-1}(V)$ is a robust set in X; i.e. if $\overline{x} \in M^{-1}(V)$, we have to show that $\overline{x}$ is a robust point of $M^{-1}(V)$. Hence, $M(\overline{x}) \cap V \neq \emptyset$. Let $\overline{\xi} \in M(\overline{x}) \cap V$ and $B_\epsilon(\overline{\xi}) \subset V$, for some $\epsilon > 0$. Hence, we have

$$\{x \in X \mid \varphi_{\overline{\xi}}(x) < \varphi_{\overline{\xi}}(\overline{x}) + \epsilon\}$$

is a non-empty robust set, since $\overline{x} \in \{x \in X \mid \varphi_{\overline{\xi}}(x) < \varphi_{\overline{\xi}}(\overline{x}) + \epsilon\} = \{x \in X \mid \varphi_{\overline{\xi}}(x) < \epsilon\}$ as $\varphi_{\overline{\xi}}(\overline{x}) = 0$. Thus, for any neighborhood $N(\overline{x})$ of $\overline{x}$ we have

$$N(\overline{x}) \cap int\{x \in X \mid \varphi_{\overline{\xi}}(x) < \epsilon\} \neq \emptyset.$$

It then follows that, there is $x^* \in N(\overline{x}) \cap int\{x \in X \mid \varphi_{\overline{\xi}}(x) < \epsilon\}$ and an open neighborhood $U(x^*) \subset N(\overline{x}) \cap int\{x \in X \mid \varphi_{\overline{\xi}}(x) < \epsilon\}$ of x^*. Consequently,

$$\forall x \in U(x^*) : \varphi_{\overline{\xi}}(x) < \epsilon.$$

That is

$$\forall x \in U(x^*) : dist(\overline{\xi}, M(x)) < \epsilon.$$

Subsequently,

$$\forall x \in U(x^*), \exists y \in M(x) : \|\overline{\xi} - y\| < \epsilon.$$

From this follows that

$$\forall x \in U(x^*) : M(x) \cap B_\epsilon(\overline{\xi}) \subset M(x) \cap V.$$

Hence, $U(x^*) \subset M^{-1}(V)$. Which yields $N(\overline{x}) \cap intM^{-1}(V) \neq \emptyset$. However, since $N(\overline{x})$ is an arbitrary neighborhood, we conclude that $\overline{x}$ is a robust point of $M^{-1}(V)$. As $\overline{x} \in M^{-1}(V)$ was chosen arbitrarily, we have that $M^{-1}(V)$ is a robust set in X. □

3.5.2 Upper Robustness of Supremum

Let again

$$\phi(x) = \sup_{y \in M(x)} \psi(x, y).$$

Theorem 3.5.6 (upper robustness of supremum).

Let X and Y be topological spaces and $x^0 \in X$. If $M : X \rightrightarrows Y$ is compact valued and u.r. at x^0; and $\psi : X \times Y \to \mathbb{R}$ u.s.c. on $\{x^0\} \times M(x^0)$, then ϕ is upper robust at x^0.

Proof. (follows word for word from Thm. 2, p. 52 in [4]). Let $c \in \mathbb{R}$ be arbitrary and define

$$[\phi, c] := \{x \in X \mid \phi(x) < c\}.$$

Let $x^0 \in [\phi, c]$ be any. Then we show that x^0 is a robust point of $[\phi, c]$ or, equivalently, $[\phi, c]$ is a semi-neighborhood of x^0. Since, ψ is u.s.c. on $\{x^0\} \times M(x^0)$, we have, for each $\overline{y} \in M(x^0)$ and $\varepsilon > 0$ neighborhoods $N^\varepsilon(\overline{y})$ and $N^\varepsilon_{\overline{y}}(x^0)$ of $\overline{y}$ and x^0, respectively, such that

$$\forall y \in N^\varepsilon(\overline{y}), \forall x \in N^\varepsilon_{\overline{y}}(x^0) : \psi(x, y) \le \psi(x^0, \overline{y}) + \varepsilon.$$

The compactness of $M(x^0)$ implies the existence of $\{\overline{y}_1, \ldots, \overline{y}_{n(\varepsilon)}\} \subset M(x^0)$ such that:

$$M(x^0) \subset \bigcup_{i=1}^{n(\varepsilon)} N^\varepsilon(\overline{y}_i).$$

We define the open set:

$$N := \bigcup_{i=1}^{n(\varepsilon)} N^\varepsilon(\overline{y}_i).$$

Since, $M(\cdot)$ is upper robust at x^0, $M(x^0) \subset N$ and N is open, there is a semi-neighborhood $S(x^0)$ of x^0 such that

$$\forall x \in S(x^0) : M(x) \subset N$$

(as $M(\cdot)$ is u.r. at x^0 and $M(x^0) \subset N$, we may take the set $S(x^0) := \{x \in X \mid M(x) \subset N\}$). Hence, the set

$$N_\varepsilon(x^0) := S(x^0) \cap \bigcap_{i=1}^{n(\varepsilon)} N^\varepsilon_{\overline{y}_i}(x^0)$$

is a semi-neighborhood of x^0 (see Prop. 3.2.4(2)). If $D(x^0)$ is any open neighborhood of x^0 in X, then $D(x^0) \cap N_\varepsilon(x^0)$ is also a semi-neighborhood of x^0. Let $x \in D(x^0) \cap N_\varepsilon(x^0)$ be arbitrarily chosen. Then it follows that

$$x \in S(x^0), x \in \bigcap_{i=1}^{n(\varepsilon)} N^\varepsilon_{\overline{y}_i}(x^0) \quad \textit{and} \quad y \in M(x) \subset N.$$

Hence, for some $i_0, 1 \leq i_0 \leq n_\varepsilon(x)$, $y \in N^\varepsilon(\overline{y}_{i_0})$ and $x \in \bigcap_{i=1}^{n(\varepsilon)} N^\varepsilon_{\overline{y}_i}(x^0) \subset N^\varepsilon_{\overline{y}_{i_0}}(x^0)$. Which implies that $\psi(x,y) \leq \psi(x^0, \overline{y}_{i_0}) + \varepsilon$. Moreover, since $x \in D(x^0) \cap N_\varepsilon(x^0)$ and $y \in M(x)$ are arbitrary, we have that

$$\begin{aligned}\forall x \in D(x^0) \cap N_\varepsilon(x^0) : \sup_{y \in M(x)} \psi(x,y) \leq \\ \leq \psi(x^0, \overline{y}_{i_0}) + \varepsilon \leq \phi(x^0) + \varepsilon.\end{aligned}$$

This yields

$$\forall x \in D(x^0) \cap N_\varepsilon(x^0) : \phi(x) \leq \phi(x^0) + \varepsilon.$$

Now, since $x^0 \in [\phi, c]$ and $\varepsilon > 0$ are arbitrary, we can choose $0 < \varepsilon < c - \phi(x^0)$. It then follows that

$$\begin{aligned}\forall x \in D(x^0) \cap N_\varepsilon(x^0) : \phi(x) \leq \phi(x^0) + \varepsilon < \\ < \phi(x^0) + c - \phi(x^0) = c.\end{aligned}$$

From this follows that

$$\forall x \in D(x^0) \cap N_\varepsilon(x^0) : \phi(x) < c.$$

Hence, $D(x^0) \cap N_\varepsilon(x^0) \subset [\phi, c]$. Therefore, $[\phi, c]$ is a semi-neighborhood of x^0 and the claim follows from Prop 3.2.4(2). □

Corollary 3.5.7 (lower robustness of infimum). *Let X and Y be topological spaces and $x^0 \in X$. If $M : X \rightrightarrows Y$ is compact valued and u.r. at x^0; and $\psi : X \times Y \rightarrow \mathbb{R}$ l.s.c. on $\{x^0\} \times M(x^0)$, then the marginal value function*

$$\varphi(x) = \inf_{y \in M(x)} \psi(x,y)$$

is lower robust at x^0.

Proof. Use

$$-\varphi(x) = \sup_{y \in M(x)} [-\psi(x,y)]$$

and apply Thm. 3.5.6; i.e. $-\varphi$ will be u.r. at x^0. Which implies that φ is lower robust x^0. □

3.5.3 Upper Robustness over a Robust Partition

Definition 3.5.1. We call the function $\varphi : X \to \mathbb{R}$ *piecewise u.r. (l.r.)* iff there exists a robust partition $X_1, X_2, \ldots, X_r$ of X such that for all $i \in \{1, \ldots, r\}$ the restriction of φ to X_i is u.r. (l.r.) with respect to the relative topology of X_i induced by the topological space X.

Theorem 3.5.8. *Let X be a topological space and $\varphi : X \to \mathbb{R}$. If φ is piecewise u.r. (l.r.), then φ is u.r. (l.r.).*

Proof. Let $c \in \mathbb{R}$, such that $F_c := \{x \in X \mid \varphi(x) < c\}$. Then $F_c = \bigcup_{i \in I} (X_i \cap \{x \in X \mid \varphi(x) < c\})$. Assume now $x \in F_c$ and $N(x)$ be any open neighborhood of x w.r.t. X. Then $x \in X_i \cap \{x \in X \mid \varphi(x) < c\}$ for some $i \in I$. Hence, $N(x) \cap X_i$ is a neighborhood of x relative to X_i. Since φ is u.r. w.r.t. the relative topology on X_i, $int[X_i \cap \{x \in X \mid \varphi(x) < c\} \cap N(x)]_{X_i} \neq \emptyset$ and $int[X_i \cap \{x \in X \mid \varphi(x) < c\} \cap N(x)]_{X_i} \subset N(x) \cap X_i$. Since $N(x)$ is open and $N(x) \cap X_i$ is robust in X, Lemma 3.4.17 yields:

$$int[\{x \in X \mid \varphi(x) < c\} \cap N(x)]_X \supset int[X_i \cap \{x \in X \mid \varphi(x) < c\} \cap N(x)]_X \neq \emptyset.$$

As a result

$$N(x) \cap int\,(\{x \in X_i \mid \varphi(x) < c\}) \neq \emptyset.$$

From this follows that x is a robust point of $\{x \in X_i \mid \varphi(x) < c\}$. Consequently, by Rem. 3.2.1, we have that the set $\{x \in X \mid \varphi(x) < c\}$ is robust. Therefore, φ is u.r. on X. The proof for l.r. follows the same line of argument. □

Theorem 3.5.9. *Let $\psi : X \times Y \to \mathbb{R}$ be an u.s.c. function and let $M : X \rightrightarrows Y$ be a piecewise l.s.c. (l.r.) SV-map on X. Then the marginal function φ is u.r. on X.*

Proof. Since $M(\cdot)$ is piecewise l.s.c. (piecewise l.r.), there is a robust partition $X_1, ..., X_r$ of X such that for each $i \in I := \{1, ..., r\}$, X_i is robust in X and the restriction of $M(\cdot)$ to X_i is l.s.c. (l.r.). Thus, using Thm. 4 of Aubin & Cellina [4] (or Theorem 3.5.1), we see that φ is u.s.c. (φ is u.r.) on X_i, which implies that φ is u.r. on X_i for each $i \in I$. Therefore, by Thm. 3.5.8, φ is u.r. on X. □

Similarly,

Theorem 3.5.10. *If $\psi : X \times Y \to \mathbb{R}$ is u.s.c. and $M : X \rightrightarrows Y$ is a piecewise u.s.c. (u.r.) compact valued SV-map, then the marginal function ϕ is u.r.*

Furthermore, Cor. 3.5.2 and Cor. 3.5.7 could be reformulated to provide lower robustness properties of marginal functions, based on the corresponding piece-wise semi-continuity of $M(\cdot)$.

Remark 3.5.1. Obviously, to verify the upper robustness of φ and ϕ in Thms. 3.5.9 and 3.5.10 we need only the upper semi-continuity of ψ on $X_i \times Y$ w.r.t. relative topology.

3.5.4 Approximatable Marginal Function

Proposition 3.5.11 (upper approximatability of infimum). *Let X and Y be metric spaces. Suppose that the function $\psi : X \times Y \to \mathbb{R}$ is u.s.c. and $M : X \rightrightarrows Y$ is a lower approximatable SV-map, then the function φ is upper approximatable.*

Proof. Let $x^0 \in X$ be any. Since

$$\varphi(x^0) = \inf_{y \in M(x^0)} \psi(x^0, y)$$

given $\varepsilon > 0$, choose $\overline{y} \in M(x^0)$ such that

$$\psi(x^0, \overline{y}) < \varphi(x^0) + \frac{\varepsilon}{2}.$$

By the u.s.c. of ψ, there are neighborhoods $U(x^0)$ and $V(\overline{y})$ of x^0 and $\overline{y}$, respectively, such that

$$\forall x \in U(x^0), \forall y \in V(\overline{y}) : \psi(x^0, y) \leq \psi(x^0, \overline{y}) + \frac{\varepsilon}{2}.$$

Since $M(\cdot)$ is lower approximatable, there is a set $S \subset X$ such that

(i) $M(\cdot)$ is lower semi-continuous on S and S is dense in X; and

(ii) there are sequences $\{x^k\} \subset S$ and $\{y^k\}$ with $y^k \in M(x^k)$, such that

$$\lim_k x^k = x^0 \quad \textit{and} \quad \lim_k y^k = \overline{y}.$$

Using Thm. 4, Aubin & Cellina [4], φ is u.s.c. at each of the points in S. Moreover, since $x^k \to x^0$ and $y^k \to \overline{y}$, there exists a $k_0 \in \mathbb{N}$ (*i.e.*, $k_0 = k_0(\varepsilon)$) such that

$$x^k \in U(x^0) \;\; and \;\; y^k \in V(\overline{y}), \forall k \geq k_0.$$

Hence, $\psi(x^k, y^k) \leq \psi(x^0, \overline{y}) + \frac{\varepsilon}{2}, \forall k \geq k_0$, which implies that

$$\varphi(x^k) = \inf_{y \in M(x^k)} \psi(x^k, y) \leq \psi(x^k, y^k) \leq \psi(x^0, \overline{y}) + \frac{\varepsilon}{2}, \forall k \geq k_0.$$

By passing to the limit we find that

$$\limsup_k \varphi(x^k) \leq \psi(x^0, \overline{y}) + \frac{\varepsilon}{2} \leq \varphi(x^0) + \varepsilon.$$

As $\varepsilon > 0$ was chosen arbitrarily, we conclude that $\limsup_k \varphi(x^k) \leq \varphi(x^0)$. Therefore, φ is upper approximatable. □

To guarantee the upper approximatability of a supremum we require the following lemma.

Lemma 3.5.12. *Let X be a metric space, $\phi : X \to \mathbb{R}$ be a function and $S \subset X$ be the set of points of upper semi-continuity of ϕ. If ϕ is upper robust and S is dense in X, then ϕ is upper approximatable.*

Proof. (The idea of the proof is identical to the one given by Zheng *et al.* for a robust function (cf. Prop. 3.4 in [67])). Take an arbitrary $x^0 \in X$. For each $n \in \mathbb{N}$ consider the set

$$\Phi_n := \left\{ x \in X \mid \phi(x) < \phi(x^0) + \frac{1}{n} \right\}.$$

Then for all $n \in \mathbb{N}, x^0 \in \Phi_n$ and Φ_n is a robust set in X. Moreover, if $B_{\frac{1}{n}}(x^0)$ is any open neighborhood of x^0, then $x^0 \in B_{\frac{1}{n}}(x^0) \cap \Phi_n, \forall n \in \mathbb{N}$, and $B_{\frac{1}{n}}(x^0) \cap \Phi_n$ is a robust set in X. Hence, $int \left(B_{\frac{1}{n}}(x^0) \cap \Phi_n \right) \neq \emptyset$. Since S is dense in X, for each $n \in \mathbb{N}$, there is $x^n \in S \cap \left(B_{\frac{1}{n}}(x^0) \cap \Phi_n \right)$. Consequently, there is a sequence $\{x^n\} \subset S$ such that

$$\lim_{n \to \infty} x^n = x^0 \text{ and}$$
$$\phi(x^n) < \phi(x^0) + \frac{1}{n}, \forall n \in \mathbb{N}.$$

The latter implies that $\limsup_n \phi(x^n) \leq \phi(x^0)$. Therefore, ϕ is upper approximatable by Def. 3.3.3. □

Remark 3.5.2. In contrast to Prop. 3.3.5, we do not need here that the set X to be a complete metric space, but we have to assume the existence of a dense set S of upper semi-continuity of ϕ.

Theorem 3.5.13 (upper approximatability of supremum). *Let X and Y be metric spaces. If $M : X \rightrightarrows Y$ is upper approximatable and compact valued; and $\psi : X \times Y \to \mathbb{R}$ is u.s.c., then ϕ is upper approximatable.*

Proof. (i) Since $M(\cdot)$ is upper approximatable, $M(\cdot)$ is upper robust. Moreover, since $M(\cdot)$ is compact valued, Thm. 3.5.6 assures that ϕ is upper robust.

(ii) At the same time, $M(\cdot)$ is upper approximatable implies that $M(\cdot)$ has a dense set S of upper semi-continuity. From this follows that $\phi(x) = \sup_{y \in M(x)} \psi(x, y)$ is u.s.c. on S (Thm. 5, Aubin & Cellina [4]); i.e. ϕ has a dense set of upper semi-continuity.

Consequently, using (i) and (ii), the claim follows from Lem. 3.5.12. □

3.6 Robustness of SV-maps with given structures

Subsequently, we consider the robustness of set-valued maps with given structures. Accordingly, it is necessary to characterize the robustness of such maps through the properties of their defining functions. For similar issues related with continuity properties of these maps see the review given in chapter 1.

3.6.1 The Finite Parametric Case

Lower Robustness

Let X and T be topological spaces and $B : X \rightrightarrows T$ be a SV-map given by

$$B(x) = \{t \in T \mid h_i(x, t) \leq 0, i \in I\},$$

where $h_i : X \times T \to \mathbb{R}, i \in I$; and $I := \{1, \ldots, p\}$. We would like to characterize robustness properties of $B(\cdot)$ through the properties of the functions $h_i, i \in I$. This has been done below under three sets of conditions. Two of them are discussed in this subsection and a third one is found in Sec. 3.6.4.

Conditions-1

Proposition 3.6.1. *Let $B(\cdot)$ be as given above and suppose that for each fixed $t \in T$ the set*

$$\triangle(t) := \bigcap_{i=1}^{p} \{x \in X \mid h_i(x,t) \leq 0\}$$

is non-empty and robust in X. Then for every open set $U \subset T$, the set $B^{-1}(U)$ is robust; i.e., $B(\cdot)$ is a lower robust SV-map.

Proof. Let $U \subset T$ be any open set. It suffices to show that $B^{-1}(U)$ is a robust set in X. Then

$$\begin{aligned} B^{-1}(U) &= \bigcup_{t \in U} B^{-1}(t) \\ &= \bigcup_{t \in U} \{x \in X \mid t \in B(x)\} \\ &= \bigcup_{t \in U} \{x \in X \mid h_i(x,t) \leq 0, i \in I\} \\ &= \bigcup_{t \in U} \bigcap_{i=1}^{p} \{x \in X \mid h_i(x,t) \leq 0\}. \end{aligned}$$

By assumption $\triangle(t) = \bigcap_{i=1}^{p} \{x \in X \mid h_i(x,t) \leq 0\}$ is robust; hence, $B^{-1}(U)$ is a union of robust sets. Therefore, Remark 3.2.1 implies that $B^{-1}(U)$ is a robust set. □

Corollary 3.6.2. *Let $B(\cdot)$ and $\triangle(t)$ be as given in Prop. 3.6.1. If*

(i) for every fixed $t \in T$ and for each fixed $i \in I$, the functions $h_i(\cdot,t)$ are quasi-convex; equivalently, the set

$$N_i(t) := \{x \in X \mid h_i(x,t) \leq 0\}$$

is convex;

(ii) for all $t \in T, h(\hat{x},t) < 0$ for at least one $\hat{x} \in X$;

(iii) $x \mapsto h_{i_0}(x,t)$ is upper robust;

(iv) $x \mapsto h_i(x,t)$ *is upper semi-continuous w.r.t.* x *for each* $i \in I \setminus \{i_0\}$*;*

then $\triangle(t)$ *will be a convex set with a non-empty interior, which is a robust set (cf. Cor. 3.2.1).*

Conditions-2

Proposition 3.6.3. *Let* X *be a topological space,* T *be a non-empty subset of a normed linear space, and* $B: \ X \rightrightarrows T$ *be given according to*

$$B(x) = \{t \in T \mid h_i(x,t) \leq 0, \forall i \in I\},$$

where $I = \{1, \ldots, p\}$*. If*

(i) for every pair $(x^0, t^0) \in X \times T$*, and every neighborhood* $V(t^0)$ *of* t^0*, there exists* $\widetilde{t} \in V(t^0)$ *such that*

$$h_i(x^0, t^0) \leq 0 \ \ \textit{implies} \ \ h_i(x^0, \widetilde{t}) < 0, \forall i \in I;$$

(ii) $h_{i_0}(\cdot, t)$ *is u.r. on* X *for each* $t \in T$*;*

(iii) $h_i(\cdot, t)$ *u.s.c. on* X *for each* $t \in T, i \in I \setminus \{i_0\}$*;*

then $B(\cdot)$ *is l.r. on* X*.*

Proof. Let $x^0 \in X$ and $t^0 \in B(x^0)$ and $V(t^0)$ is a neighborhood of t^0. Then we want to show that $B^{-1}(V)$ is a semi-neighborhood of x^0, i.e. x^0 is a robust point of $B^{-1}(V)$. By (i) we have some $\widetilde{t} \in V(t^0)$ with $h_i(x^0, \widetilde{t}) < 0$. And using the upper semi-continuity, there is some neighborhood $U(x^0)$ such that for all $x \in U(x^0)$

$$h_i(x, \widetilde{t}) < 0, i \in I \setminus \{i_0\}$$

and we know, by the upper robustness of $h_{i_0}(\cdot, \widetilde{t})$, that $x^0 \in \{x \in X \mid h_{i_0}(x, \widetilde{t}) < 0\} =: H$ and that H is robust. Hence, $U(x^0) \cap H$ is a robust set containing x^0 (cf. Rem. 3.2.1). Furthermore, $B^{-1}(V) \supset U(x^0) \cap H \ni x^0$, i.e. $B^{-1}(V)$ is a semi-neighborhood of x^0. □

Note that assumption (i) of Prop. 3.6.3 is Slater type condition. Nevertheless, convexity is not demanded from the functions. In contrast to Cor. 3.6.2, we require next quasi-convexity w.r.t. to t.

Corollary 3.6.4 (cf. also Thm. 3.1.6., Bank *et al.* [7]).
Let X be a topological space, T be a normed linear space and $B : X \rightrightarrows T$ be given by

$$B(x) = \{t \in T \mid h_i(x,t) \leq 0, i \in I\}.$$

If the following hold true:

(i) for each fixed $x \in X$, $B(x) \neq \emptyset$ and is not a singleton;

(ii) for each fixed $x \in X$ and each $i \in I$, $h_i(x,\cdot) : T \to \mathbb{R}$ is strictly quasi-convex;

(iii) for each fixed $t \in T$ and one $i_0 \in I$, $h_{i_0}(\cdot,t)$ is upper robust on X;

(iv) for each fixed $t \in T$ and each $i \in I \setminus \{i_0\}$, $h_i(\cdot,t)$ is u.s.c.;

then $B(\cdot)$ is lower robust on X.

Proof. It suffices to show that assumptions (i) & (ii) imply assumption (i) of Prop. 3.6.3. Let $x^0 \in X$ be any. By (i), $B(x^0) \neq \emptyset$. Hence, there is some $t^0 \in B(x^0)$. Suppose $t \in B(x^0)$ be arbitrary and let $U(t)$ be any neighborhood of t. Thus, $h_i(x^0,t^0) \leq 0$ and $h_i(x^0,t) \leq 0$ for each $i \in I$. Then we could find an α with $0 < \alpha < 1$ such that $t_\alpha := (1-\alpha)t^0 + \alpha t \in U(t)$. Since, the $h_i(x,\cdot)$'s are strictly quasi-convex and u.s.c, for each $i \in I \setminus \{i_0\}$, there is a neighborhood $V_i(x^0)$ such that

$$h_i(x,t_\alpha) < 0, \forall x \in V_i(x^0).$$

Set $V(x^0) := \bigcap_{i \in I \setminus \{i_0\}} V_i(x^0)$. (Clearly, $V(x^0)$ is an open neighborhood of x^0). Hence, for each $x \in V(x^0)$

$$h_i(x,t_\alpha) < 0, \forall i \in I \setminus \{i_0\}.$$

The rest of the proof is as in the proof of Prop. 3.6.3. □

Remark 3.6.1. Let X be a topological space, T be a normed linear space, $x^0 \in X$ and $t^0 \in T$. If $h_i(x^0, \cdot) : T \to \mathbb{R}, i \in I$ are convex and there is some $\widetilde{t} \in T \setminus \{t^0\}$ such that for all $i \in I$:

$$h_i(x^0, t^0) \leq 0 \Rightarrow h_i(x^0, \widetilde{t}) < 0, \text{(Slater's Condition)},$$

then condition (i) of Prop. 3.6.3 is satisfied.

Proof. Let $h_i(x^0, t^0) \leq 0, \forall i \in I$. Hence, by assumption, there is $\widetilde{t} \neq t^0$ such that $h_i(x^0, \widetilde{t}) < 0, \forall i \in I$. Accordingly,

$$t_n = \frac{1}{n}\widetilde{t} + (1 - \frac{1}{n})t^0 \to t^0 \quad for \quad n \to \infty.$$

Hence, for a given neighborhood $V(t^0)$ and sufficiently large n we have that $t_n \in V(t^0)$. Furthermore,

$$h_i(x^0, t_n) \leq \frac{1}{n} h_i(x^0, \bar{t}) + (1 - \frac{1}{n}) h_i(x^0, \widetilde{t}) < 0.$$

□

To relate lower robustness to a well-known result of Bank *et al.* [7], we consider a continuous function $h : X \to \mathbb{R}$ and define its level set map as

$$\mathcal{L}_{h,X}(\alpha) := \{x \in X \mid h(x) \leq \alpha\}.$$

Thm. 3.1.7 of Bank *et al.* [7] claims that $L_{h,X}(\cdot)$ is l.s.c. on X if and only if h has only global minima on X. However, the statement next indicates that the lower robustness of $\mathcal{L}_{h,X}(\cdot)$ does not preclude the existence of local minima of h; furthermore, the continuity of $h : T \to \mathbb{R}$ is not even required.

Proposition 3.6.5. *Let* $X \subset \mathbb{R}^n$, $T \subset \mathbb{R}^m$, X *a robust set in* $\mathbb{R}^n$ *and* $B(x) := \{t \in T \mid h_i(t) \leq x_i, \ for \ all \ i, \ 1 \leq i \leq n\}$, *where* $h_i : X \to \mathbb{R}, 1 \leq i \leq n$, *are functions. If, for each fixed* $t \in T$, $int\{x \in X \mid h_i(t) \leq x_i, for \ all \ i, \ 1 \leq i \leq n\} \neq \emptyset$, *then the set-valued map* $B : X \rightrightarrows T$ *is lower robust on* X.

Proof. Given $x^0 \in X$ and a $t^0 \in B(x^0)$, observe that

$$\prod_{i=1}^{n} \left[h_i(t^0), +\infty\right) \cap X = B^{-1}(t^0)$$
$$= \{x \in X \mid h_i(t^0) \leq x_i, i = 1, \ldots, n\}.$$

Hence, $x^0 \in \prod_{i=1}^{n} \left[h_i(t^0), +\infty\right) \cap X$ and $int \left[\prod_{i=1}^{n} \left[h_i(t^0), +\infty\right) \cap X\right] \neq \emptyset$. Thus, x^0 is a robust point of $\prod_{i=1}^{n} \left[h_i(t^0), +\infty\right) \cap X$. Consequently, $intB^{-1}(t^0) \neq \emptyset$ and x^0 is a robust point of $B^{-1}(t^0)$ (cf. Prop. 3.2.4(1)). Since, x^0 is arbitrary, $B(\cdot)$ will be a lower robust SV-map. □

In the special cases when X is a compact set, X is an open set or $X = \mathbb{R}^n$, the assumption $int\{x \in X \mid h_i(t) \leq x_i, for\ all\ i,\ 1 \leq i \leq n\} \neq \emptyset$ of Prop. 3.6.5 is obviously satisfied. Note also that, in Prop. 3.6.5, the functions $h_i, i \in I$, are not required to be continuous.

Upper Robustness

Once again, reiterating Def. 3.4.2, we have that $B:\ X \rightrightarrows T$ is upper robust at $x^0 \in X$ if for each neighborhood U of $B(x^0)$ there is a semi-neighborhood $SNH(x^0)$ of x^0 such that

$$\forall x \in SNH(x^0) : B(x) \subset U.$$

In contrast to the lower robustness of $B(\cdot)$, its upper robustness could follow from relatively weaker assumptions. One standard result (Thm 3.1.2, Bank *et al.* [7]) is that: if X is a closed and T is a compact sets and $h_i, i \in I$, are lower semi-continuous, then $B : X \rightrightarrows T$ with $B(x) = \{t \in T \mid h_i(x,t) \leq 0\}$ is u.s.c. (cf. also Chap. 1, Cor. 1.2.2).

To give further results of upper robustness, we introduce the following definition:

Definition 3.6.1. Let X and T be topological spaces and $h : X \times T \to \mathbb{R}^p$. Then $h(\cdot, t)$ is called *lower robust* [*l.s.c.*] *at* x^0 *uniformly* for all $t \in T$ iff for all $\varepsilon > 0$ there exists a semi-neighborhood $SNH_\varepsilon(x^0)$ [a neighborhood $U(x^0)$] of x^0 such that

$$h_i(x,t) > h_i(x^0,t) - \varepsilon, \forall x \in SNH_\varepsilon(x^0), \forall t \in T, \forall i \in I$$
$$\left[h_i(x,t) > h_i(x^0,t) - \varepsilon, \forall x \in U(x^0), \forall t \in T, \forall i \in I\right].$$

Moreover, we need the regularity condition given by (see also Def. 1.2.1)

Definition 3.6.2. Let X be a topological space and T be a metric space. The function $h : X \times T \to \mathbb{R}^p$ is called *strictly r-regular* at $x^0 \in X$, for all $t \in T$, iff there is a **strictly increasing** function $r : \mathbb{R}_+ \to \mathbb{R}_+,\ r(0) = 0$, such that

$$dist(t, B(x^0)) \leq r\left(\max_{i=1,\dots,p}\left[h_i(x^0, t)^+\right]\right).$$

Proposition 3.6.6. *Let X be a topological space, T be a compact metric space, $h : X \times T \to \mathbb{R}^p$, $h := (h_1, \dots, h_p)$, and let $B : X \rightrightarrows T$ be a set-valued map, such that for each $x \in X, B(x)$ is bounded and is given by*

$$B(x) = \{t \in T \mid h(x,t) \leq 0\}.$$

If

(i) $h(\cdot, t)$ is lower robust [l.s.c.] at x^0 uniformly for all $t \in T$; and

(ii) h is strictly r-regular at x_0 for all $t \in T$,

then $B(\cdot)$ is upper robust [u.s.c.] at x^0.

Proof. (In the following, to prove the upper semi-continuity, replace SNH by neighborhood). Thus, for any neighborhood U of $B(x^0)$ we have to find some semi-neighborhood $SNH_U(x^0)$ such that

$$\forall x \in SNH_U(x^0) : \{t \in T \mid h(x,t) \leq 0\} \subset U.$$

But, since $B(x^0)$ is bounded, there is $\varepsilon > 0$ such that

$$U_\varepsilon := \{t \in T \mid dist(t, B(x^0)) < \varepsilon\} \subset U.$$

Consequently, we need only to show that: there is a semi-neighborhood SNH_{U_ε} of x^0 such that

$$\forall x \in SNH_{U_\varepsilon} : \{t \in T \mid h(x,t) \leq 0\}$$
$$\subset \{t \in T \mid dist(t, B(x^0)) < \varepsilon\}.$$

From the lower robustness of $h(\cdot, t)$ at x^0 uniformly for $t \in T$, we get, for any $\tau > 0$, a semi-neighborhood $SNH_\tau(x_0)$ such that, for each $x \in SNH_\tau(x_0)$ and each $t \in \{t \in T \mid h_i(x,t) \leq 0, i = 1, \dots, p\}$, the following holds

$$h_i(x,t) > h_i(x^0,t) - \tau.$$

Hence,

$$h_i(x^0,t) < \tau, \qquad \forall i = 1, \dots, p.$$

This implies

$$\max_{i=1,\dots,p} \left[h_i(x^0,t)\right]^+ < \tau.$$

Using strict r-regularity and the monotonicity of r, we obtain

$$dist(t, B(x^0)) < r(\tau).$$

Taking $SNH_{U_\varepsilon} := SNH_\varepsilon(x^0)$ (i.e. $\varepsilon := r(\tau)$) the proof is complete. □

3.6.2 A Semi-infinite Case

Consider the following SV-map defined by using a semi-infinite system of constraints

$$M(x) := \{\xi \in Y \mid G(\xi, x, t) \leq 0, t \in B(x)\},$$

where $G : Y \times X \times T \to \mathbb{R}$ and $B : X \rightrightarrows T$ is a SV-map. Define the marginal function

$$m(\xi, x) := \begin{cases} \sup_{t \in B(x)} G(\xi, x, t), & \text{if } B(x) \neq \emptyset; \\ -\infty, & \text{if } B(x) = \emptyset. \end{cases}$$

Obviously, we have that

$$M(x) = \{\xi \in Y \mid m(\xi, x) \leq 0\}.$$

The following is a consequence of Prop. 3.6.3.

Corollary 3.6.7. *Let $M(x) := \{\xi \in X \mid m(\xi, x) \leq 0\}$. If*

(i) for every $(\xi^0, x^0) \in Y \times X$, and every neighborhood $U(\xi^0)$ of ξ^0, there is $\hat{\xi}$ such that

$$m(\xi, x^0) \leq 0 \;\; \textit{implies} \;\; m(\hat{\xi}, x^0) < 0;$$

(ii) for each fixed $\xi \in Y$, $m(\xi, \cdot)$ is upper robust;

then $M(\cdot)$ is a lower robust SV-map.

Proof. See Prop. 3.6.3. □

In particular, from Cor. 3.6.4 we obtain

Corollary 3.6.8. *Let $M(x) := \{\xi \in X \mid m(\xi, x) \leq 0\}$. If*

(i) for each fixed $x \in X, m(\cdot, x)$ is strictly quasi-convex and $M(x)$ is not a singleton; and

(ii) for each fixed $\xi \in Y, m(\xi, \cdot)$ is upper robust;

then $M(\cdot)$ is lower robust.

Proof. See Cor. 3.6.4. □

Assumption (ii) of both Cor. 3.6.7 and Cor. 3.6.8 is guaranteed by the following proposition.

Proposition 3.6.9. *If, for each fixed $\xi \in Y, G(\xi, \cdot, \cdot)$ is u.s.c. and $B(\cdot)$ is upper robust and compact valued, then $m(\xi, \cdot)$ is upper robust.*

Proof. For a fixed $\overline{\xi} \in Y$ we could write

$$m(\overline{\xi}, x) := \sup_{t \in B(x)} G(\overline{\xi}, x, t).$$

Thus, if we let $\phi(x) := m(\overline{\xi}, x)$ and $\psi(x, t) := G(\overline{\xi}, x, t)$, then the claim follows from Thm. 3.5.6. □

Furthermore, assumption (i) of Cor. 3.6.7 follows if we suppose that: $G(\cdot, \cdot, \cdot)$ is upper semi-continuous, $B(\cdot)$ is compact valued and, for each fixed $x \in Dom(B)$, there is $\widetilde{\xi} \in Y$ such that

$$G(\widetilde{\xi}, x, t) < 0, \forall t \in B(x).$$

Next we try to guarantee the assumption (i) of the Cor. 3.6.8 through the properties of G and $B(\cdot)$. We use the following notions of quasi-convexity of functions.

Definition 3.6.3 (γ-strong convex function). Let Y be a linear space and $f : Y \to \mathbb{R}$. If for any $\xi_1, \xi_2 \in Y$ and $\alpha \in (0,1)$ there are some fixed $c > 0$ and $0 < \gamma \leq 2$ such that

$$f(\alpha\xi_1 + (1-\alpha)\xi_2) \leq \\ \alpha f(\xi_1) + (1-\alpha)f(\xi_2) - \frac{1}{2}c\alpha(1-\alpha)\|\xi_1 - \xi_2\|^\gamma,$$

then f is called $\gamma-$strongly convex.

In Def. 3.6.3, when $\gamma = 2$, f is called *strongly convex* (cf. Urruty & Lemaréchal [82]).

Definition 3.6.4 (γ-strongly quasi-convex function). Let Y be a linear space and $f : Y \to \mathbb{R}$. If for any $\xi_1, \xi_2 \in Y$ and $\alpha \in (0,1)$ there are some fixed $c > 0$ and $0 < \gamma \leq 2$ such that

$$f(\alpha\xi_1 + (1-\alpha)\xi_2) \leq \\ max\Big\{f(\xi_1), f(\xi_2)\Big\} - \frac{1}{2}c\alpha(1-\alpha)\|\xi_1 - \xi_2\|^\gamma,$$

then f is called *γ-strongly quasi-convex.*

One may wonder if there is any γ-strongly quasi-convex function.

Theorem 3.6.10. *If $f : [a,b] \to \mathbb{R}, f'(x) \geq d > 0$, for all $x \in [a,b]$, then f is γ-strongly quasi-convex*

(i) for $\gamma = 1$; or

(ii) for $\gamma > 0$, whenever $[a,b]$ is a bounded interval.

Proof. Given $x, x^0 \in [a,b]$, w.lo.g. $x < x_0$, we have to show that

$$f(\alpha x + (1-\alpha)x_0) \leq f(x^0) - \frac{1}{2}x\alpha(1-\alpha)|x - x_0|^\gamma$$

for some $c > 0$. Denote by $x_\alpha := \alpha x + (1-\alpha)x_0$. By Mean-Value-Theorem, there is $\xi \in (x_\alpha, x_0)$ such that:

$$f(x_\alpha) = f(x^0) + f'(\xi)(x_\alpha - x_0) \leq f(x^0) + d(x_\alpha - x_0).$$

Thus we look for c and d for which the following holds true:

$$(f(x_\alpha) \le) f(x^0) + d(x_\alpha - x_0) \le f(x^0) - \frac{1}{2}c\alpha(1-\alpha)|x - x_0|^\gamma \tag{3.6.1}$$

$$\begin{aligned}
&\Leftrightarrow \alpha d(x - x_0) \le -\frac{1}{2}c\alpha(1-\alpha)|x - x_0|^\gamma \\
&\Leftrightarrow 2d(x_0 - x) \ge c(1-\alpha)|x - x_0|^\gamma \\
&\Leftrightarrow 2d|x_0 - x| \ge c(1-\alpha)|x - x_0|^\gamma \\
&\Leftrightarrow 2d|x_0 - x|^{1-\gamma} \ge c(1-\alpha), \forall \alpha \in (0,1) \\
&\Leftrightarrow 2d|x_0 - x|^{1-\gamma} \ge c.
\end{aligned}$$

Hence,

(i) if $\gamma = 1$, then (3.6.1) is satisfied for $2d \ge c$;

(ii) else if $\gamma > 0$, then (3.6.1) is satisfied for $2d|b-a|^{1-\gamma} \ge c$.

But the satisfaction of (3.6.1) implies that $f(x_\alpha) \le f(x^0) - \frac{1}{2}c\alpha(1-\alpha)|x - x_0|^\gamma$, which shows the γ-strong quasi-convexity of f. □

To give a concrete example

Example 3.6.11. *Consider $f(x) = x^3 + x$. Then f is γ-strongly quasi-convex*

(i) for $\gamma = 1$ and $x \in \mathbb{R}$; or

(ii) for $\gamma > 0$ and $x \in [a,b]$, where $[a,b]$ is a bounded interval,

for appropriately chosen c and d.

Obviously, a $\gamma-$strongly quasi-convex function is strictly quasi-convex. Moreover, we have

Lemma 3.6.12. *γ-strong convexity implies γ-strong quasi-convexity.*

Proposition 3.6.13. *Let X and T be topological spaces, Y be a linear space and $B : X \rightrightarrows T$ be a SV-map with compact values. If, for any fixed $x \in X$ and $t \in T$, $G(\cdot, x, t)$ is $\gamma-$strongly (quasi-) convex, then $m(\cdot, x)$ is $\gamma-$strongly quasi convex; hence, strictly quasi-convex.*

Proof. For $\xi_1, \xi_2 \in Y$ and $\lambda_1, \lambda_2 \in [0, 1]$ we have

$$
\begin{aligned}
m(\lambda_1\xi_1 + \lambda_2\xi_2, x) &= \sup_{t\in B(x)} \Big[G(\lambda_1\xi_1 + \lambda_2\xi_2, x, t)\Big] \\
&\le \sup_{t\in B(x)} \Big[\max\Big\{G(\xi_1, x, t), G(\xi_2, x, t)\Big\} \\
&\qquad - \frac{1}{2}c\|\xi_1 - \xi_2\|^\gamma\Big] \\
&= max\Big\{\sup_{t\in B(x)} G(\xi_1, x, t), \sup_{t\in B(x)} G(\xi_2, x, t)\Big\} \\
&\qquad - \frac{1}{2}c\|\xi_1 - \xi_2\|^\gamma \\
&= max\Big\{m(\xi_1, x), m(\xi_2, x)\Big\} - \frac{1}{2}c\|\xi_1 - \xi_2\|^\gamma.
\end{aligned}
$$

Therefore, $m(\cdot, x)$ $\gamma-$strongly quasi-convex; hence, it is strictly quasi-convex. □

Summing up, given a linear space Y and

$$M(x) = \{\xi \in Y \; | G(\xi, x, t) \le 0, t \in B(x)\},$$

$M(\cdot)$ will be a lower robust SV-map

(i) if

- (a) $G(\cdot, \cdot, \cdot)$ is u.s.c. and $M(x)$ is not a singleton;
- (b) $G(\cdot, x, t)$ is $(\gamma-)$ strong $(quasi-)$ convex, w.r.t $\xi \in Y$ for each $t \in B(x)$;
- (c) $B(\cdot)$ is upper robust and compact valued;

(ii) or

- (a) $G(\cdot, \cdot, \cdot)$ is u.s.c.;
- (b) for each $x \in X$, there exists $\hat{\xi}$ such that $G(\hat{\xi}, x, t) < 0$ for all $t \in B(x)$ (note that G u.s.c. implies there is $U(\hat{\xi}) : G(\xi, x, t) < 0$ for all $\xi \in U(\hat{\xi})$);

(c) $B(\cdot)$ is upper robust and compact valued.

For a further characterization of robustness the map $M(\cdot)$ through regularity and constraint qualifications see sections 3.6.3 and 3.6.4.

3.6.3 Piecewise Semi-continuity of a SV-map with a Structure

In Sec. 3.4.4 we have considered piecewise semi-continuity properties of a general SV-map. Correspondingly, we would like to characterize piecewise semi-continuity for set-valued maps with given structures.

Recall that $M(x) = \{\xi \in Y \mid G(\xi, x, t) \leq 0, \ \forall t \in B(x)\}$. We give now a second characterization of lower robustness of $M(\cdot)$, besides the ones in Sec. 3.6.2, based on piecewise upper semi-continuity of $B(\cdot)$, joint upper semi-continuity of G and some weaker regularity condition of the system defining $M(x)$. Let X be a metric space, with metric ρ, let Y be a robust subset of some topological space, and T be a topological space.

Assumption (A): The set X has a robust partition $(X_i)_{i\in I}$, $I = \{0, 1, 2, ..., r+1\}$, where $X_0 := \{x \in X \,|B\,(x) = \emptyset\}$ and $X_{r+1} := \{x \in X \mid M(x) = \emptyset\}$ are among the robust partitions.

Assumption (B): The map $B: X \rightrightarrows T$ is compact valued and $B|_{X_i}, i = 0, 1, 2, ..., r, r+1$ is u.s.c. w.r.t. the relative topology on X_i.

Definition 3.6.5 (local r-regularity). The system

$$\begin{aligned} G(\xi, x, t) &\leq 0, \forall t \in B(x) \\ \xi &\in Y \end{aligned}$$

is r-regular at $(\xi^0, x^0) \in Y \times X_i$ if there is a semi-neighborhood $SNH(x^0) \subset X_i$ of x^0 w.r.t. the relative topology of X_i and a non-decreasing function $r_{\xi^0, x^0, SNH(x^0)} : \mathbb{R}_+ \to \mathbb{R}_+$

continuous at 0, with $r(0) = 0$ such that

$$\forall x \in SNH(x^0) : dist(\xi^0, M(x)) \leq r\left(\max_{t \in B(x)} \left[G(\xi^0, x, t)\right]^+\right).$$

The r-regularity given in Def. 3.6.5 is quite weaker than the metric regularity condition given by Klatte & Henrion [39] (see also Chap. 1, Def. 1.2.2). In fact, from the metric regularity follows the lower semi-continuity of $M(\cdot)$ (cf. Chap. 1, Thm. 1.2.4).

Theorem 3.6.14. *Let X and Y be normed spaces. If G is u.s.c. on $Y \times X \times T$, Assumptions (A), (B) are satisfied; and for all $i \in \{0, 1, \ldots, r, r+1\}$ and all $x^0 \in X_i$ the system defining $M(x)$ is r-regular at each (ξ^0, x^0), for each $\xi^0 \in M(x^0)$, then $M(\cdot)$ is l.r. on X.*

Proof. We show that $M(\cdot)$ is piecewise lower robust. That is, we show that for each $i \in \{0, 1, \ldots, r+1\}$, $M(\cdot)$ is lower robust on X_i. If $x \in X_0$, then it follows that $M(x) = X$, i.e. $M(\cdot)$ is continuous on X_0 in the relative topology. For all $x \in X_{r+1}$ we get, from $M(x) = \emptyset$, that $M(\cdot)$ is l.r. on X_{r+1}. Thus, it remains to discuss the case $1 \leq i \leq r$.

Thus, let $i \in \{1, \ldots, r\}$ and $x^0 \in X_i$. Let also $\xi^0 \in M(x^0)$.

a) By definition of $M(\cdot)$ we have that

$$G(\xi^0, x^0, t) \leq 0, \forall t \in B(x^0).$$

If we let $g(\xi, x) := \max_{t \in B(x)}[G(\xi, x, t)]^+$, then $g(\xi^0, x^0) \leq 0$. Moreover, $g(\xi^0, x) := \max_{t \in B(x)} G(\xi^0, x, t)$ and $B(\cdot)$ is u.s.c on X_i w.r.t the relative topology of X_i. Consequently, $g(\xi, \cdot)$ is u.s.c. at x^0 in the topology of X_i. Hence, given $\varepsilon > 0$, there a neighborhood $V_\varepsilon(x^0)$ such that

$$g(\xi^0, x) \leq \varepsilon, \forall x \in V_\varepsilon(x^0) \cap X_i. \tag{3.6.2}$$

b) By the r-regularity at (ξ^0, x^0), we obtain that

$$\forall x \in SNH(x^0) \cap \left[V_\varepsilon(x^0) \cap X_i\right] : dist(\xi^0, M(x)) \leq r\left(\max_{t \in B(x)} \left[G(\xi^0, x, t)\right]^+\right) = r(g(\xi^0, x)).$$

Using (3.6.2) and the property of the function $r(\cdot)$, we find that

$$\forall x \in SNH(x^0) \cap \left[V_\varepsilon(x^0) \cap X_i\right] : dist(\xi^0, M(x)) \leq r(\varepsilon).$$

Now, given an arbitrary neighborhood $U(\xi^0) \subset Y$ of ξ^0, there is $\varepsilon > 0$ (and a corresponding $V_\varepsilon(x^0)$) such that the open ball $\mathbf{B}_{2r(\varepsilon)}(\xi^0)$ is contained in $U(\xi^0)$. Accordingly, for each fixed $x \in SNH(x^0) \cap [V_\varepsilon(x^0) \cap X_i]$, we deduce that

$$\mathbf{B}_{2r(\varepsilon)}(\xi^0) \cap M(x) \neq \emptyset.$$

From this follows that

$$\forall x \in SNH(x^0) \cap \left[V_\varepsilon(x^0) \cap X_i\right] : M(x) \cap U(\xi^0) \neq \emptyset.$$

In other words

$$SNH(x^0) \cap \left[V_\varepsilon(x^0) \cap X_i\right] \subset M^{-1}(U(\xi^0)).$$

Since $SNH(x^0)$ is robust set w.r.t. the topology of X_i, we have $[int_{X_i} SNH(x^0)] \cap [V_\varepsilon(x^0) \cap X_i] \neq \emptyset$. Hence, $int_{X_i} M^{-1}(U(\xi^0)) \neq \emptyset$. Moreover, x^0 is a robust point of $SNH(x^0) \cap [V_\varepsilon(x^0) \cap X_i]$; there by, x^0 is robust point of $M^{-1}(U(\xi^0))$ in the relative topology of X_i. Since $x^0 \in X_i$ is arbitrary, we conclude that $M(\cdot)$ is lower robust on X_i in the relative topology. Therefore, $M(\cdot)$ is piece-wise lower robust; and hence, it is lower robust (cf. Thm. 3.4.19).

□

Observe that **the upper semi-continuity of $B(\cdot)$ is not assumed on the whole of** X, except on each of the partitioning sets X_i of X.

Remark 3.6.2. Let X and Y be normed spaces. If the regularity condition given by Def. 3.6.5 holds at $(\xi^0, x^0) \in Y \times X$, where $\xi^0 \in M(x^0)$, for a neighborhood $V(x^0)$ of x^0 w.r.t. X and $B(\cdot)$ is u.s.c. on X, then $M(\cdot)$ will be lower semi-continuous at x^0. The verification of this follows with a slight modification of the proof of Thm. 3.6.14.

For a related result of upper robustness, we make the following assumption:

Assumption(C): $B|_{X_i}, i = 1, \ldots, r$ is l.s.c. with respect to the relative topology of X_i.

Theorem 3.6.15. *Let Y be a compact set. If G is l.s.c. on $Y \times X \times T$ and assumptions (A) and (C) are satisfied, then $M(\cdot)$ is u.r. on X.*

Proof. Considering

$$m(\xi, x) = \max_{t \in B(x)} G(\xi, x, t),$$

we get $M(x) = \{\xi \in Y \mid m(\xi, x) \leq 0\}$. Then Assumption(C) and the lower semi-continuity of G yield, by Thm. 4, Aubin & Cellina [4], that m is l.s.c. on $Y \times X_i, 1 \leq i \leq r$, in relative topology. Furthermore, since Y is compact, we have that $M(\cdot)$ is u.s.c. on X_i in the relative topology of X_i, for each $i \in \{1, \ldots, r\}$ (cf. Chap. 1, Cor. 1.2.2) .

Putting all together, we obtain that $M(\cdot)$ is piecewise-u.s.c. on X. Applying Thm. 3.4.16, we conclude that $M(\cdot)$ is upper robust. □

3.6.4 Characterization of Robustness through Constraint Qualifications

In this section we try to find out some results connecting certain Mangasarian-Fromvitz type constraint qualifications with the robustness of set-valued maps.

Lemma 3.6.16. *Suppose that X and O are robust and open sets, respectively. Then $X \cap O \neq \emptyset$ implies $intX \cap O \neq \emptyset$*

Proof. Obvious. □

Definition 3.6.6. In a normed linear space W we define the tangential cone $\mathcal{T}(X, \bar{x})$ of $X \subset W$ at $\bar{x} \in clX$ by

$$\mathcal{T}(X, \bar{x}) = \{\xi \in W \mid \forall \varepsilon > 0, \forall \rho > 0 : (\bar{x} + [cone\,(\xi + \rho \mathsf{B}) \cap \varepsilon \mathsf{B}]) \cap X \neq \emptyset\}$$

which is known to be closed and which has the well-known properties

$$\mathcal{T}(intX, \bar{x}) \subset \mathcal{T}(X, \bar{x}) = \mathcal{T}(clX, \bar{x}) .$$

We put, for $\xi \in W, \lambda, \rho > 0$, the convex set

$$\mathcal{K}_\lambda(\xi, \rho) = cone\,(\xi + \rho \mathsf{B}) \cap \lambda \mathsf{B}$$

which has a nonempty interior, where **B** represents the unit-ball of center at 0 of W.

Lemma 3.6.17. *If X is a robust set and $\bar{x} \in clX$ then $\mathcal{T}(X, \bar{x}) = \mathcal{T}(intX, \bar{x})$.*

Proof. Observe that $\mathcal{T}(intX, \bar{x}) \subset \mathcal{T}(X, \bar{x}) = \mathcal{T}(clX, \bar{x}) = \mathcal{T}(cl\,(intX), \bar{x}) = \mathcal{T}(intX, \bar{x})$. □

Lemma 3.6.18. *Suppose X is a robust subset of W, $\bar{x} \in X$ (or $\bar{x} \in clX$). If $\xi_0 \in \mathcal{T}(X, \bar{x})$ and $\lambda > 0$, then $int\,([\bar{x} + \mathcal{K}_\lambda(\xi_0, \rho)] \cap X) \neq \emptyset$ and $\bar{x}$ is a robust point of [robust point to] the set $[\bar{x} + \mathcal{K}_\lambda(\xi_0, \rho)] \cap X$.*

Proof. Since $\mathcal{T}(X, \bar{x}) = \mathcal{T}(intX, \bar{x})$ and $int\mathcal{K}_\lambda(\xi_0, \rho) \neq \emptyset$ and convex (hence robust) we get, from Lemma. 3.6.16, that $\forall \varepsilon > 0, \forall \rho > 0 : \emptyset \neq int\,[\bar{x} + \mathcal{K}_\varepsilon(\xi_0, \rho)] \cap intX \subset int\,([\bar{x} + \mathcal{K}_\varepsilon(\xi_0, \rho)] \cap X)$. For each such $\varepsilon < \lambda$, we have $\emptyset \neq int\mathcal{K}_\varepsilon(\xi_0, \rho) \subset \mathcal{K}_\varepsilon(\xi_0, \rho) \subset \mathcal{K}_\lambda(\xi_0, \rho)$ and $int\mathcal{K}_\varepsilon(\xi_0, \rho) \subset \varepsilon \mathsf{B}$. Hence, in each $\varepsilon-$ball $\bar{x} + \varepsilon \mathsf{B}$ of $\bar{x}$ there are interior points of $[\bar{x} + \mathcal{K}_\lambda(\xi_0, \rho)] \cap X$; i.e., $\bar{x}$ is a robust point of $[\bar{x} + \mathcal{K}_\lambda(\xi_0, \rho)] \cap X$ if $\overline{x} \in X$; or a robust point to $[\bar{x} + \mathcal{K}_\lambda(\xi_0, \rho)] \cap X$ if $\overline{x} \in clX \setminus X$. □

Theorem 3.6.19. *Suppose X and T are nonempty subsets of normed spaces, X is robust, $X \times T \subset W$, W is an open set, $h : W \to \mathbb{R}^p$ and $B : X \rightrightarrows T$ is a set-valued map defined by*

$$B(x) = \{x \in X \mid h_i(x, t) \leq 0, \forall i \in I := \{1, 2, ..., p\}\}.$$

*If the **(MFCQ)** is satisfied at all $(\overline{x}, \overline{t}) \in \{\bar{x}\} \times B(\bar{x})$, i.e. for each $\bar{t} \in B(\bar{x})$ with the active index set*

$$I_0 := \{i \in I \mid h_i(\bar{x}, \bar{t}) = 0\},$$

i) h_i is Frechet-differentiable at $(\bar{x}, \bar{t})$ for each $i \in I_0$ and h_i is continuous at $(\bar{x}, \bar{t})$ for each $i \in I \setminus I_0$;

ii) there are vectors $\xi_0 \in \mathcal{T}(X,\bar{x})$ and $\eta_0 \in \mathcal{T}(T,\bar{t})$ such that, for each $i \in I_0$,

$$D_x h_i(\bar{x},\bar{t})\,\xi_0 + D_t h_i(\bar{x},\bar{t})\,\eta_0 < 0;$$

then $B(\cdot)$ is lower robust at $\bar{x}$.

Proof. We show that for an arbitrary $\varepsilon > 0$ the pre-image $B^{-1}(V_\varepsilon(\bar{t}))$ of the neighborhood $V_\varepsilon(\bar{t}) = (\bar{t} + \varepsilon \mathsf{B}) \cap T$ is a semi-neighborhood of $\bar{x}$.

$$B^{-1}(V_\varepsilon(\bar{t})) = \bigcup_{t \in V_\varepsilon(\bar{t})} \{x \in X \mid h_i(x,t) \leq 0, \forall i \in I\}.$$

By the continuity and linearity of the derivative $(D_x h(\bar{x},\bar{t}), D_t h(\bar{x},\bar{t}))$, there are positive radii ρ_x and ρ_t such that, for all $i \in I_0$,

$$D_x h_i(\bar{x},\bar{t})\,\xi + D_t h_i(\bar{x},\bar{t})\,\eta < 0$$

for each $\lambda > 0$ and each $(\xi,\eta) \in \mathcal{K}_\lambda(\xi_0,\rho_x) \times \mathcal{K}_\lambda(\eta_0,\rho_t)$. The Taylor approximation of h_i at $(\bar{x},\bar{t})$ for $i \in I_0$

$$h_i(x,t) = h_i(\bar{x},\bar{t}) + D_x h_i(\bar{x},\bar{t})\,(x-\bar{x}) + D_t h_i(\bar{x},\bar{t})\,(t-\bar{t}) + o(x-\bar{x},t-\bar{t})$$

yields radii $\varepsilon > \gamma_x, \gamma_t > 0$ such that for all $(\xi,\eta) \in \mathcal{K}_{\gamma_x}(\xi_0,\rho_x) \times \mathcal{K}_{\gamma_t}(\eta_0,\rho_t)$

$$h_i(\bar{x}+\xi, \bar{t}+\eta) < 0$$

holds and the continuity of h_i, for $i \in I \setminus I_0$, yields radii $\varepsilon > \beta_x, \beta_t > 0$ such that, for all $(\xi,\eta) \in \beta_x \mathsf{B} \times \beta_t \mathsf{B}$, again the inequality

$$h_i(\bar{x}+\xi, \bar{t}+\eta) < 0$$

is satisfied. It then follows that

$$\begin{aligned} B^{-1}(V_\varepsilon(\bar{t})) &\supset B^{-1}((\bar{t} + \min(\gamma_t,\beta_t)\,\mathsf{B}) \cap T) \\ &\supset (\bar{x} + [\mathcal{K}_{\gamma_x}(\xi_0,\rho_x) \cap \beta_x \mathsf{B}]) \cap X \\ &= \left(\bar{x} + \mathcal{K}_{\min(\gamma_x,\beta_x)}(\xi_0,\rho_x)\right) \cap X. \end{aligned}$$

Hence, by Lemma 3.6.18, $\bar{x}$ is a robust point of $\left(\bar{x} + \mathcal{K}_{\min(\gamma_x,\beta_x)}(\xi_0,\rho_x)\right) \cap X$. Which implies that, $B^{-1}(V_\varepsilon(\bar{t}))$ is a semi-neighborhood of $\bar{x}$. $\square$

Remark 3.6.3. If the (MFCQ) is satisfied separately at $\overline{x} \in X$, for all $\overline{t} \in B(\overline{x})$; i.e., there is $\xi \in \mathbb{R}^n$ such that

$$D_x h_i(\overline{x}, \overline{t})\xi < 0, \forall i \in I_0, \forall \overline{t} \in B(\overline{x}),$$

then this implies again only the robustness of $B(\cdot)$ at $\overline{x}$. However, if the (MFCQ) is satisfied separately for all $t \in B(\bar{x})$, then , as is well-known, $B(\cdot)$ turns out to be lower-semi-continuous at $\bar{x}$, since $\mathcal{K}_{\gamma_x}(\xi_0, \rho_x)$ can be replaced by the full neighborhood $\bar{x} + \gamma_x \mathsf{B}$.

Next, we find a similar characterization for set-valued maps defined with a semi-infinite system. Thus, in the following we suppose that X, Y, T are nonempty subsets of normed spaces, $B : X \rightrightarrows T$ is a set-valued map and the set-valued map $M : X \rightrightarrows Y$ is defined by

$$M(x) = \{y \in Y \mid G(y, x, B(x)) \leq 0\},$$

where $G(y, x, Q) \leq 0$ means that $G(y, x, t) \leq 0$ for all $t \in Q$ for a subset Q of T. We use further the active index set

$$E(x, y) = \{t \in T \mid G(y, x, t) = 0\} \subset T.$$

Definition 3.6.7. We say the (EMFCQ) is satisfied for the System

$$G(y, x, B(x)) \leq 0$$

w.r.t. $Y \times X$ at $(\bar{y}, \bar{x})$ if

1. there is some $\tau > 0$ such that $G(\cdot, \cdot, \cdot)$ is F-differentiable at $(\bar{y}, \bar{x}, t)$ w.r.t. (y, x), the remainder property is satisfied uniformly in t on a compact subsets of T and $G(\bar{y}, \bar{x}, \cdot), D_y G(\bar{y}, \bar{x}, \cdot), D_x G(\bar{y}, \bar{x}, \cdot)$ are continuous for all $t \in (E(\bar{y}, \bar{x}) + \tau \mathsf{B}) \cap (B(\bar{x}) + \tau \mathsf{B}) \cap T$ and

2. there are directions $\eta_0 \in \mathcal{T}(Y, \bar{y}), \xi_0 \in \mathcal{T}(X, \bar{x})$ such that for all $t \in E(\bar{y}, \bar{x}) \cap B(\bar{x})$

$$D_y G(\bar{y}, \bar{x}, t)\eta_0 + D_x G(\bar{y}, \bar{x}, t)\xi_0 < 0.$$

Theorem 3.6.20. *Suppose the robust subset X, the nonempty subset Y and the nonempty, compact subset T are supplied with respective induced topologies of their including normed spaces and $G : Y \times X \times T \to \mathbb{R}$ is continuous. If $B : X \rightrightarrows T$ is upper semi-continuous on X and the defining system*

$$G(y, x, B(x)) \leq 0, x \in X, y \in Y$$

of the set-valued map $M : X \rightrightarrows Y$ satisfies the (EMFCQ) w.r.t. $Y \times X$ at $(y, \bar{x})$ for all $y \in M(\bar{x})$, then $M(\cdot)$ is lower robust at $\bar{x}$.

Proof. Let first $B(\bar{x}) \neq \emptyset$. We show that for an arbitrary $\varepsilon > 0$ the pre-image $M^{-1}(V_\varepsilon(\bar{y}))$ of the neighborhood $V_\varepsilon(\bar{y}) = (\bar{y} + \varepsilon\mathsf{B}) \cap Y$, for an arbitrary $\bar{y} \in M(\bar{x})$, is a semi-neighborhood of $\bar{x}$. We have

$$M^{-1}(V_\varepsilon(\bar{y})) = \bigcup_{y \in V_\varepsilon(\bar{y})} \{x \in X \mid G(y, x, B(x)) \leq 0\}.$$

The (EMFCQ) implies the existence of directions $\eta_0 \in \mathcal{T}(Y, \bar{y}), \xi_0 \in \mathcal{T}(X, \bar{x})$ such that for all $t \in E(\bar{y}, \bar{x}) \cap B(\bar{x})$

$$D_y G(\bar{y}, \bar{x}, t)\eta_0 + D_x G(\bar{y}, \bar{x}, t)\xi_0 < 0$$

holds.

By the continuity and linearity of the derivative $(D_y G(\bar{y}, \bar{x}, t), D_x G(\bar{y}, \bar{x}, t))$, the continuity of $(D_y G(\bar{y}, \bar{x}, \cdot), D_x G(\bar{y}, \bar{x}, \cdot))$ and the compactness of $E(\bar{x}, \bar{y}) \cap B(\bar{x})$ there are positive radii ρ_y, ρ_x, $\delta < \tau$ such that

$$D_y G(\bar{y}, \bar{x}, t)\eta + D_x G(\bar{y}, \bar{x}, t)\xi < 0$$

for each $\lambda > 0$, each $(\eta, \xi) \in \mathcal{K}_\lambda(\xi_0, \rho_y) \times \mathcal{K}_\lambda(\eta_0, \rho_x)$ and each $t \in ((E(\bar{y}, \bar{x}) + \delta\mathsf{B}) \cap (B(\bar{x}) + \delta\mathsf{B})) \cap T$. Because of the upper semi-continuity of $B(\cdot)$ and the compactness of T and $B(x)$ there is a $\sigma(\delta) > 0$ such that

$$B(x) \subset (B(\bar{x}) + \delta\mathsf{B}) \cap T \tag{3.6.3}$$

for all $x \in X \cap (\bar{x} + \sigma \mathsf{B})$ (which is a relative open set in X). The Taylor approximation of G at $(\bar{y}, \bar{x}, t)$

$$G(y,x,t) = G(\bar{y},\bar{x},t) + D_y G(\bar{y},\bar{x},t)(y-\bar{y}) + D_x G(\bar{y},\bar{x},t)(x-\bar{x}) + o(y-\bar{y}, x-\bar{x}, t)$$

and the continuity properties w.r.t t and the uniform remainder property in t yields radii $\varepsilon > \gamma_y, \gamma_x > 0$ such that for all $(\eta, \xi) \in \mathcal{K}_{\gamma_y}(\eta_0, \rho_y) \times \mathcal{K}_{\gamma_x}(\xi_0, \rho_x)$ and all

$$t \in ((E(\bar{y},\bar{x}) + \delta \mathsf{B}) \cap (B(\bar{x}) + \delta \mathsf{B})) \cap T$$

the inequality

$$G(\bar{y} + \eta, \bar{x} + \xi, t) < 0$$

holds true. The set-valued map $(y, x) \longmapsto E(y, x)$ is closed because of the continuity of G on $Y \times X \times T$ and the compactness of T implies its upper semi-continuity (cf. Chap. 1, Prop. 1.1.4). Hence, there is $\varepsilon > \mu(\delta) > 0$ such that for all $(y, x) \in ((\bar{y} + \mu \mathsf{B}) \cap Y) \times ((\bar{x} + \mu \mathsf{B}) \cap X)$

$$E(y,x) \subset E(\bar{y},\bar{x}) + \delta \mathsf{B}.$$

Thus, using (3.6.3), we have

$$[(E(\bar{y},\bar{x}) + \delta \mathsf{B}) \cap (B(\bar{x}) + \delta \mathsf{B})] \cap T \supset E(y,x) \cap B(x)$$

for all $(y, x) \in ((\bar{y} + \mu \mathsf{B}) \cap Y) \times ((\bar{x} + \mu \mathsf{B}) \cap X)$. So far we proved the inverse map of the active constraints contains the intersection of the semi-neighborhood $\mathcal{K}_{\gamma_x}(\xi_0, \rho_x)$ and the neighborhood $[(\bar{x} + \min\{\sigma, \mu\} \mathsf{B}) \cap X]$ of $\bar{x}$.

The complement C of $((E(\bar{y},\bar{x}) + \delta \mathsf{B}) \cap (B(\bar{x}) + \delta \mathsf{B})) \cap T$ w.r.t. $cl\,(B(\bar{x}) + \delta \mathsf{B}) \cap T$ is a compact set in T (note that $cl\,(B(\bar{x}) + \delta \mathsf{B}) \cap T \supset B(x)$). Here is $G(\bar{y}, \bar{x}, t) < 0$, i.e. non active. Hence, for all $t \in C$ and some $\beta_y > 0, \sigma > \beta_x > 0$ we have, by the continuity of G, that

$$G(y,x,t) < 0$$

for all $(y, x) \in ((\bar{y} + \beta_y \mathsf{B}) \cap X) \times ((\bar{x} + \beta_x \mathsf{B}) \cap Y)$. It follows

$$\begin{aligned} M^{-1}(V_\varepsilon(\bar{y})) &\supset M^{-1}((\bar{y} + \min(\gamma_y, \beta_y, \mu)\,\mathsf{B}) \cap Y) \\ &\supset (\bar{x} + [\mathcal{K}_{\gamma_x}(\xi_0, \rho_x) \cap \beta_x \mathsf{B} \cap \mu \mathsf{B}]) \cap X \\ &= \left(\bar{x} + \mathcal{K}_{\min(\gamma_x, \beta_x, \mu)}(\xi_0, \rho_x)\right) \cap X. \end{aligned}$$

Hence, by Lemma 3.6.18, $\overline{x}$ is a robust point of $\left(\bar{x} + \mathcal{K}_{\min(\gamma_x, \beta_x, \mu)}(\xi_0, \rho_x)\right) \cap X$ which implies that $M^{-1}(V_\varepsilon(\bar{y}))$ is a semi-neighborhood of $\bar{x}$.

Furthermore, if $B(\bar{x}) = \emptyset$, then there is a neighborhood U of $\bar{x}$ such that $B(x) = \emptyset$ for all $x \in U$. It follows immediately that $M(\bar{x}) \equiv Y$ on U. This even implies the continuity of $M(\cdot)$ at $\bar{x}$. □

Remark 3.6.4. In the same manner as for the finite case, we get again lower robustness if we have the (EMFCQ) being satisfied separately w.r.t. $\overline{x} \in X$, for all $y \in M(\bar{x})$; and the lower semi-continuity if we have the (EMFCQ) separately w.r.t. y for all $y \in M(\bar{x})$. In both proofs, the compactness of $B(x)$ plays an important role.

Remark 3.6.5. Both Theorems 3.6.19 and 3.6.20 can be extended to a piecewise upper semi-continuous set-valued map $B(\cdot)$. Taking that $\{X_k\}_{k \in J}$ is a robust partition of the robust set X, we demand the assumptions of the theorems to hold true for each component X_k with respect to its relative topology. Naturally, we have to take the tangential cones w.r.t. X_k and not w.r.t. X. For instance, in Thm. 3.6.20, the piecewise upper semi-continuity of $B(\cdot)$ with the validity of the regularity condition (EMFCQ) on each X_k imply that $M(\cdot)$ is piece-wise lower robust. Hence, $M(\cdot)$ will be lower robust (see Thm. 3.4.19). Note that, in this case the regularity separately in t (Thm. 3.6.19) or in y (Thm. 3.6.20) do not yield lower semi-continuity but lower robustness, since a piece-wise lower semi-continuous set-valued map is at least lower robust.

Klatte & Henrion [39] have shown that the (EMFCQ) (w.r.t. a neighborhood) implies metric-regularity (see Def. 3.6.5). Under such instances, $M(\cdot)$ will be lower semi-continuous (cf. Chap. 1, Thm. 1.2.4). However, for us the upper semi-continuity of $B(\cdot)$

along with a weaker form (EMFCQ) (w.r.t. a semi-neighborhood, Def. 3.6.7) is enough to derive the lower robustness of $M(\cdot)$. Indeed, it would have been very interesting to find out the relation between r-regularity (of Def. 3.6.5) and the (EMFCQ) (Def. 3.6.7). But, this has been left out for a future research activity.

Before passing to the next chapter it might be worthwhile to foretell that, in chapter 4, we are going to consider a set-valued map of the form

$$M(x) = \{x \in X \mid G(x,t) \geq 0, t \in B(x)\}.$$

where the sets X and Y collapse into one. Hence, the discussion in Sec. 3.6.2 needs to be carried over in light of the form

$$M(x) = \{x \in X \mid -G(x,t) \leq 0, t \in B(x)\}.$$

Furthermore, for set-valued maps with given structure we have explicitly considered inequality constraints. However, when equality constraints are assumed to be present one may need certain stronger regularity conditions to guarantee the corresponding robustness properties.

In many instances, the upper semi-continuity (upper robustness) of a set-valued map of the form

$$B(x) = \{t \in T \mid h_i(x,t) \leq 0\}$$

follows easily, if we demand that T is a compact set and $h_i, i = 1, \ldots, p$, are continuous. In this respect, the uniform lower robustness (Def. 3.6.1) and the r-regularity (Def. 3.6.2) assumptions of Prop. 3.6.6 will become superfluous. In any case, one may need to note that: the validity of continuity properties on partitioning sets implying robustness on the whole.

Chapter 4

A Coarse Solution of GSIP via a Global Optimization Method

4.1 Introduction

We consider a generalized semi-infinite optimization problem (GSIP) with one semi-infinite inequality constraint, *with no prior assumption on the lower semi-continuity of the index set valued map*. Such a (GSIP) is known to be *ill-behaved*. In particular, the lack of lower semi-continuity of the index map indicates that the feasible set $\mathcal{M}$ of (GSIP) might not be closed. Consequently, (GSIP) may not possess a solution in $\mathcal{M}$. Nevertheless, one may be interested in determining a generalized solution of (GSIP); i.e., a solution that lies in $cl(\mathcal{M})$. Hence, the intention here is to determine a coarse approximation of such a generalized solution of (GSIP) through the *integral global optimization method* (IGOM).

The IGOM was first proposed by Chew & Zheng [13], further elaborated and extended by Zheng [88], Hoffmann and Phú [48], Hichert [26]. It has been found out that the IGOM is computationally efficient when the data of the optimization problem possess some relevant discontinuity properties, which are characterized by the notions of *robustness* [88]. In fact, Hichert [26] have designed a more general version of (IGOM) into a software routine called BARLO.

Thus, the main objective is to verify the validity of the theoretical assumptions of the (IGOM) in order to apply BARLO to the (GSIP). Since the (IGOM) has its roots in robust analysis and measure theory, we mainly make use of results of robustness from Chapter 3 and certain standard results of measurability from the literature.

Subsequently, two penalty approaches have been proposed:

Approach-1: a pure penalty approach; and

Approach-2: a penalty approach coupled with discretization.

In the first approach,

- an auxiliary *parametric semi-infinite optimization problem* (PSIP) has been set up in order to characterize the admissible set $\mathcal{M}$ of the (GSIP); and
- then a discontinuous penalty function is defined using the marginal function of the (PSIP), so that, (GSIP) could be re-written as, an equivalent, global optimization problem;
- finally, relations between the minimizing sequences of the penalty problem and that of the (GSIP) have been investigated.

In Approach-2,

- the marginal function of the lower level problem of (GSIP) and the well known distance function are used to define two penalty problems related with the (GSIP); then
- relations between the minimizing sequences of the penalty problems and that of the (GSIP) have been examined; finally
- convergence of the considered discretization method have been analyzed.

In both approaches, to guarantee the satisfaction of the assumptions of (IGOM), we assume *piecewise semi-continuity* properties (cf. Chap. 3, sec. 3.4.4) of the index map $B(\cdot)$ - an idea which fits the disjunctive nature of the feasible set $\mathcal{M}$ of (GSIP). Thus, *we need to ensure the upper robustness and measurability of the penalty* terms and robustness and measurability properties of the feasible set $\mathcal{M}$.

This chapter has been organized as follows:

- In Section 4.2, the problem has been stated along with a perspective of the index set-valued map.
- Section 4.4.1 presents an obvious characterization of the feasible set of (GSIP) to be used in Section 4.4.2 for the construction of a global optimization problem which is equivalent to (GSIP) under mild conditions.
- In Section 4.4.3 is to be found a brief review of the IGOM and a discussion on the main assumptions which are necessary for its application.
- Sections 4.4.4 - 4.4.5 contain results that ensure the robustness and measurability properties of the penalty function and the feasible set of (GSIP).
- Section 4.5 discusses a second type of penalty method; where Sec. 4.5.1 presents two penalty problems and discusses their relation with the (GSIP). Moreover, Sec. 4.5.2 presents the discretization method used and investigates the convergence of the proposed penalty methods.

4.2 Problem and Motivation

We consider the problem

$$(GSIP) \qquad f(x) \to \inf$$

$$\text{s.t.}$$

$$G(x,t) \geq 0, \ \forall\, t \in B(x),$$

$$x \in X.$$

And make the assumption

Assumption (A1): The sets $X \subset \mathbb{R}^n$, $T \subset \mathbb{R}^m$ are compact and non empty, the functions $f : X \to \mathbb{R}$ and $G : X \times T \to \mathbb{R}$ are upper semi-continuous (u.s.c.) on X and continuous on $X \times T$, respectively. The set-valued map $B : X \rightrightarrows T$ is at least compact valued but may have empty values for some $x \in X$.

Thus, the *admissible (feasible) set* of the (GSIP) is given by

$$\mathcal{M} = \{x \in X \mid G(x,t) \geq 0, \ \forall\, t \in B(x)\}$$

may possess strange properties [70, 71, 72, 80], as was briefly summarized in Chapter 2. Once more, the following example gives an impression about the situation.

Example 4.2.1. *(*Jongen et. al. [34] *) Let $X = [-10, 10]^2$ and put*

$$B(x) := \begin{cases} [-\sqrt{x_1}, \sqrt{x_1}] & \textit{if } x_1 \geq 0, \\ \emptyset & \textit{if } x_1 < 0. \end{cases}$$

Obviously, $B(\cdot)$ is compact valued, and ***continuous**** *w.r.t. the* ***relative*** *topology when restricted to each of the sets $X_1 := \{x \in X \mid x_1 \geq 0\}$ and $X_2 := \{x \in X \mid x_1 < 0\}$.*

**From the usual definitions of the semi-continuity (see e.g. [4, 5]) it follows that a set-valued map is l.s.c. and u.s.c. at those points where it has empty image, whenever the image is empty in some neighborhood of such points.*

However, $B(\cdot)$ is u.s.c but not l.s.c. on the whole set X. Even with the single linear semi-infinite constraint

$$G(x,t) = t + x_2 \geq 0, \forall t \in B(x)$$

we easily find that

$$\mathcal{M} = \{x \in X \mid x_1 < 0\} \cup \{x \in X \mid x_1 \geq 0 \wedge \sqrt{x_1} \leq x_2\},$$

which is not a closed set.

Such a situation is not specific to Example 4.2.1. Actually, several standard examples of (GSIP) are observed to possess index maps $B(\cdot)$ of this nature. Nevertheless, recall that Rückmann *et al.* [34] have shown that if $B(\cdot)$ is a l.s.c. set-valued map, then $\mathcal{M}$ will be a closed set (Prop. 2.2.2). Instead, if $B(\cdot)$ is only piecewise semi-continuous (cf. Chap. 3, Sec. 3.4.4), then $B(\cdot)$ may not be semi-continuous. Consequently, the (GSIP) may not possess a solution.

In general, we may claim that there is a *minimizing sequence* for (GSIP), the existence of which may be assured, for e.g., by assuming boundedness of $\mathcal{M}$ and continuity of the objective function f, etc. Actually, the use of minimizing sequences in optimization is usually justified when the feasible set is not known to be closed. Such approaches are frequently used in optimization problems arising from engineering design (see Polak & Wardi [50]).

4.3 Minimizing Sequence and Generalized Minimizers - Definitions

Definition 4.3.1 (generalized minimizer). We say $\overline{x}$ is *generalized minimizer* of (GSIP) if $\overline{x} \in cl(\mathcal{M})$ and $f(\overline{x}) = \inf_{x \in \mathcal{M}} f(x)$. We denote this by

$$\overline{x} \in \overline{arg}(GSIP).$$

If a generalized minimizer $\overline{x}$ belongs to $\mathcal{M}$, then we call $\overline{x}$ a *minimizer* of (GSIP). We also designate this by writing

$$\overline{x} \in arg(GSIP).$$

Since $X \subset \mathbb{R}^n$ is assumed to be a compact set, the continuity of f implies that $\overline{arg}(GSIP) \neq \emptyset$, whenever $\mathcal{M} \neq \emptyset$.

Definition 4.3.2 (minimizing sequence).

A sequence $\{x_n\}$ is called a *minimizing sequence* of the problem (GSIP) iff

(i) $\lim_{n\to\infty} f(x_n) = \inf_{x\in\mathcal{M}} f(x)$;

(ii) there exists $n_0 \in \mathbb{N}$, such that $x_n \in \mathcal{M}$ for all $n \geq n_0$ (i.e., feasibility for $n \geq n_0$).

When $\{x_n\}$ is a minimizing sequence of (GSIP) we write

$$\{x_n\} \in MS(GSIP).$$

Moreover, we call $\{x^n\}$ a *generalized minimizing sequence* of (GSIP) if instead of (ii) we have

(iii) $\lim_{n\to\infty} dist(x_n, \mathcal{M}) = 0$ is satisfied.

In the latter case we use the notation

$$\{x_n\} \in GMS(GSIP).$$

Furthermore, let $\psi : X \to \mathbb{R}$ be given and take the global optimization problem

$$(P_\psi) \qquad\qquad \psi(x) \to inf$$
$$x \in X.$$

Then

Definition 4.3.3. Let $\overline{x} \in X$. Then

(i) $\overline{x}$ is called a minimizer of (P_ψ) iff

$$\psi(\overline{x}) = \inf_{x \in X} \psi(x); \text{ and}$$

(ii) $\overline{x}$ is called a generalized minimizer of (P_ψ) iff

$$\liminf_{\substack{x \to \overline{x} \\ x \in X.}} \psi(x) = \inf_{x \in X} \psi(x)$$

(iii) A sequence $\{x_n\} \subset X$ is called a minimizing sequence of (P_ψ) iff $\lim_{n\to\infty} \psi(x_n) = \inf_{x \in X} \psi(x)$.

We use the following corresponding notations:

$$\overline{x} \in arg(P_\psi), \quad \overline{x} \in \overline{arg}(P_\psi) \text{ and } \{x_n\} \in MS(P_\psi).$$

Consequently, we have

Remark 4.3.1. Each accumulation point $\overline{x}$ of a minimizing sequence $\{x_n\} \in MS(P_\psi)$ belongs to $\overline{arg}(P_\psi)$.

4.4 A Conceptual Penalty Method

4.4.1 Problem of Feasibility

To define a *penalty function* for the (GSIP), take some kind of distance function p from some open superset W of $X - X$ into the nonnegative reals with the following additional properties.

Assumptions (A2):

i) The function $p : W \to \mathbb{R}_+$ is continuous on W.

ii) $p\,(x) = 0$ if and only if $x = 0$.

Furthermore, for each parameter $x \in X$, consider the *problem of feasibility*

$$(PSIP) \qquad \begin{array}{rcl} p(x-\xi) & \to & inf \\ \text{s.t.} & & \\ G(\xi,t) & \geq & 0,\ \forall\, t \in B(x), \\ \xi & \in & X. \end{array}$$

Problem (PSIP) is a parametric semi-infinite optimization problem, in which, if we fix $x \in X$, the resulting problem is an ordinary semi-infinite optimization problem (SIP). If we let

$$M(x) := \{\xi \in X \mid G(\xi,t) \geq 0,\ \forall t \in B(x)\}, \tag{4.4.1}$$

it can be seen that the sets $M(x),\ x \in X$, of (PSIP) and $\mathcal{M}$ of (GSIP) possess, in general, entirely different structures. Actually, the set-valued map $M : X \rightrightarrows X$ has closed values, due to the upper semi-continuity of the function G. For instance, for the example given above (Exa. 4.2.1), we have

$$M(x) = \{\xi \in X \mid G(\xi,t) = t + \xi_2 \geq 0,\ \forall\, t \in B(x)\},$$

which yields

$$M(x) = \begin{cases} \{\xi \in X \mid \sqrt{x_1} \leq \xi_2\} & \text{if } x_1 \geq 0, \\ X & \text{if } x_1 < 0. \end{cases}$$

That is, $M(\cdot)$ is a closed valued map while the set $\mathcal{M}$ is not closed.

The *marginal value function* φ of (PSIP)

$$\varphi(x) := \begin{cases} \inf\limits_{\xi \in M(x)} p(x-\xi) & \text{if } M(x) \neq \emptyset, \\ +\infty & \text{if } M(x) = \emptyset. \end{cases} \tag{4.4.2}$$

is generally discontinuous.

An obvious but important property of the function φ is given in the following proposition.

Proposition 4.4.1. *Assume that (A1) and (A2) are satisfied. Then we obtain the equivalence*

$$x \in \mathcal{M} \Longleftrightarrow x \in M(x) \Longleftrightarrow \varphi(x) = 0.$$

Proof. Trivially, without using the assumptions, we find, by the definitions of $\mathcal{M}$ and $M(x)$, the implications $x \in \mathcal{M} \Longleftrightarrow G(x,t) \geq 0, \forall\, t \in B(x) \Longleftrightarrow x \in M(x) \Longrightarrow \varphi(x) = 0$. Now let $\varphi(x) = 0$. Then $0 = \inf_{\xi \in M(x)} p(x-\xi)$ yields by (A1) a convergent sequence $(\xi_n)_{n\in\mathbb{N}}$ with limit $\bar{x} \in M(x)$ and, by (A2), $0 = \lim_{n\to\infty} p(x-\xi_n) = p(x-\bar{x})$. Hence $x \in M(x)$. □

Remark 4.4.1.

1. We have various options for the function p. Some possible choices for p could be

$$p(x-\xi) := \|x-\xi\|,$$

or

$$p(x-\xi) := \|x-\xi\|^2,$$

or

$$p(x-\xi) := r\left(\|x-\xi\|\right),$$

where $r : \mathbb{R}_+ \to \mathbb{R}_+$ is continuous on $\mathbb{R}_+$ and zero only at zero. Consequently, we have, in the first case (for $M(x) \neq \emptyset$), that

$$\varphi(x) = \inf_{\xi \in M(x)} \|x-\xi\| = dist(x, M(x)),$$

where $dist\,(x, M(x))$ is the well known *distance from a point x to the set $M(x)$*. In this case, *the function φ "measures the extent to which x does not satisfy the relation $x \in M(x)$"; equivalently, $x \in \mathcal{M}$.* In other words, the set $\mathcal{M}$ is the set of all *fixed points* of the SV-map $M(\cdot)$.

2. From Prop. 4.4.1 we also observe that: the feasible points of (GSIP) are the zeros of the function φ, and vice-versa.

4.4.2 Exact Penalty Approach

Next, we introduce a discontinuous penalty function and verify, under some assumptions, the *penalized problem* and (GSIP) possess the same set of *minimizing sequences.* Following Proposition 4.4.1 we have, for a fixed $d > 0$, the *discontinuous penalty function*

$$\varphi_d(x) := \begin{cases} 0 & \text{if } x \in \mathcal{M}, \\ \varphi(x) + d & \text{if } x \notin \mathcal{M}, \end{cases}$$

of the admissible set $\mathcal{M}$ of (GSIP) and consider the associated penalty problem

$$(\mathrm{PP}_{\lambda d}) \qquad \begin{array}{rcl} f(x) + \lambda\varphi_d(x) & \rightarrow & inf, \\ x & \in & X. \end{array}$$

Theorem 4.4.2. *Let f be Lipschitz-continuous modulo L, let D be the diameter of X and let $d > 0$. If $\{x_n\}$ in X is a minimizing sequence of ($PP_{\lambda d}$) and $\lambda d > DL$, then $\{x_n\}$ is a minimizing sequence of (GSIP).*

Proof. Let $\{x_n\}$ be a minimizing sequence of ($PP_{\lambda d}$) and $\varepsilon > 0$ be given. Then, there is some $n_0(\varepsilon)$ such that for all $n \geq n_0(\varepsilon)$

$$\beta_* \leq f(x_n) + \lambda\varphi_d(x_n) \leq \beta_* + \varepsilon \tag{4.4.3}$$

and corresponding $x_{n\varepsilon} \in \mathcal{M}$ such that

$$\|x_{n\varepsilon} - x_n\| \geq dist(x_n, \mathcal{M}) \geq \|x_{n\varepsilon} - x_n\| - \varepsilon. \tag{4.4.4}$$

It follows that

$$f(x_n) + \lambda\varphi_d(x_n) - \varepsilon \leq \beta_* \leq f(x_{n\varepsilon}) + \lambda\varphi_d(x_{n\varepsilon}) = f(x_{n\varepsilon}) \tag{4.4.5}$$

and

$$\begin{aligned} \varphi_d(x_n) &\leq \frac{|f(x_{n\varepsilon}) - f(x_n)| + \varepsilon}{\lambda} \\ &\leq \frac{L}{\lambda}\|x_{n\varepsilon} - x_n\| + \frac{\varepsilon}{\lambda} \\ &\leq \frac{L}{\lambda}(dist(x_n,\mathcal{M}) + \varepsilon) + \frac{\varepsilon}{\lambda}\ . \end{aligned} \tag{4.4.6}$$

Hence for $x_n \notin \mathcal{M},\ n \geq n_0(\varepsilon)$ and

$$\varepsilon < \frac{(d - \frac{L}{\lambda}D)}{L+1}\lambda$$

we get the contradiction

$$0 < d - \frac{L}{\lambda}D \leq \varphi(x_n) + d - \frac{L}{\lambda}D \leq \frac{(L+1)}{\lambda}\varepsilon < d - \frac{L}{\lambda}D.$$

Thus for all $n \geq n_0(\varepsilon_0/2)$ we have $x_n \in \mathcal{M}$ and finally

$$\beta_* = \inf_{x\in X}\{f(x) + \lambda\varphi_d(x)\} \leq \inf_{x\in\mathcal{M}}\{f(x) + \lambda\varphi_d(x)\} = \inf_{x\in\mathcal{M}} f(x) = \alpha_*, \qquad (4.4.7)$$

$$\beta_* = \lim_{n\to\infty}\{f(x_n) + \lambda\varphi_d(x_n)\} = \lim_{n\to\infty} f(x_n) \geq \alpha_*. \qquad (4.4.8)$$

Therefore, $\{x_n\}$ is a minimizing sequence for (GSIP). □

Theorem 4.4.3. *Let f be Lipschitz-continuous modulo L, let D be the diameter of X and let $d > 0$. If $\{x_n\}$ in X is a minimizing sequence of ($PP_{\lambda d}$) with $\lambda, d > 0$ and assume that there is some $\gamma > 0$ such that*

$$\varphi(x) \geq \gamma\, dist(x, \mathcal{M}), \forall x \in X,$$

and $\gamma > \frac{L}{\lambda}$, then $\{x_n\}$ is a minimizing sequence of (GSIP). Furthermore, $\{x_n\}$ is a minimizing sequence of (GSIP) for each other $d > 0$ with the above parameter λ.

Proof. Let $\{x_n\}$ be a minimizing sequence of $(PP_{\lambda d})$ and $\varepsilon > 0$ be given. Then, there is some $n_0(\varepsilon)$ such that for all $n \geq n_0(\varepsilon)$ and associated $x_{n\varepsilon} \in \mathcal{M}$ we have again (4.4.3) – (4.4.6) and we get

$$\gamma dist(x_n, \mathcal{M}) + d \leq \varphi_d(x_n) \leq \frac{L}{\lambda}(dist(x_n, \mathcal{M}) + \varepsilon) + \frac{\varepsilon}{\lambda},$$

which implies that

$$d \leq (\frac{L}{\lambda} - \gamma)dist(x_n, \mathcal{M}) + (\frac{L+1}{\lambda})\varepsilon.$$

Moreover, for $x_n \notin \mathcal{M},\ n \geq n_0(\varepsilon)$ and choosing

$$\varepsilon < \frac{d}{L+1}\lambda$$

we obtain that

$$0 < d - (\frac{L+1}{\lambda})\varepsilon.$$

Consequently, we arrive at the contradiction

$$\begin{aligned} 0 &< d - (\frac{L+1}{\lambda})\varepsilon \leq (\frac{L}{\lambda} - \gamma)dist(x_n, \mathcal{M}) \\ &= (\frac{L}{\lambda} - \gamma)dist(x_n, \mathcal{M}) < 0. \end{aligned}$$

Thus for all $n \geq n_0(\varepsilon_0/2)$ we have $x_n \in \mathcal{M}$ and again (4.4.7), (4.4.8) hold true. Therefore, $\{x_n\}$ is a minimizing sequence for (GSIP). Since for an arbitrary $d > 0$, except for a finite number of elements, the sequence $\{x_n\}$ belongs to $\mathcal{M}$, by a similar argument (as in Thm. 4.4.2) we conclude that $\{x_n\}$ is a minimizing sequence of (GSIP) for the arbitrarily selected parameter $d > 0$. □

Note that, if we take the set X as compact, then each minimizing sequence of (GSIP) has a convergent subsequence which converges to a generalized solution of the (GSIP).

Standard Statement - in case when $d = 0$ (see Clark [14])

If f is Lipshitz continuous, $\varphi(x) \geq \gamma dist(x, \mathcal{M}), \forall x \in \mathcal{M}$ and $\gamma > \frac{L}{\lambda}$, then each minimizing sequence of $(PP_{\lambda 0})$, i.e. $d = 0$, has the property that

1) $\lim_{n\to\infty} dist(x_n, \mathcal{M}) = 0$;

2) $\lim_{n\to\infty} f(x_n) = \alpha_* = \beta_*$;

3) there exists a convergent subsequence which converges to a generalized solution $\overline{x} \in cl(\mathcal{M})$.

4.4.3 Coarse Global Optimization Approach

Assuming that (PSIP) could be handled, for fixed x, by known algorithms of semi-infinite optimization problems (see, for instance, Part II of Reemtsen and Rückmann [52] for a recent review of such algorithms), the idea is to determine a coarse approximation to the solution of (GSIP) by solving ($PP_{\lambda d}$) using the integral global optimization method IGOM, with suitable parameters λ and d. Thus, we require here the robustness properties which we have discussed in Chapter 3.

Essential Infimum and Integral Global Optimization

Let $X \subset \mathbb{R}^n$ be a Lebesgue measurable, bounded set and let μ be the Lebesgue measure on $\mathbb{R}^n$. Let also $f : X \rightarrow \mathbb{R}$ be Lebesgue measurable. Then $\alpha \in \mathbb{R}$ is an *essential lower bound* of f iff $f(x) \geq \alpha$ almost everywhere (a.e.) in X, i.e. $\mu\{x \in X \mid f(x) < \alpha\} = 0$. The *essential infimum* of f over X is the supremum over all essential lower bounds of f (cf. Def. 2.1, Phú & Hoffmann [48]), i.e.

$$ess \inf f = \sup \{\alpha \in \mathbb{R} \mid \mu\{x \in X \mid f(x) < \alpha\} = 0\}.$$

The IGOM theoretically uses iterations of the form

$$\alpha_{k+1} := \frac{\int_{[f \leq \alpha_k]} f(x) d\mu(x)}{\mu[f \leq \alpha_k]},$$

where $[f \leq \alpha_k] := \{x \in X \mid f(x) \leq \alpha_k\}$ are the *level sets* of f at the level α_k. The IGOM determines the essential infimum of f over X (cf. Hichert [26]). Indeed we have

Theorem 4.4.4. *(Chew & Zheng [13]) If $f \in L_\infty(X)$ and $\mu(X) < \infty$, then* $\lim_{k\to\infty} \alpha_k = ess \inf_{x\in X} f(x)$.

As indicated earlier, Barlo (Hichert [26]) is a software implementation of the IGOM. The coded routines of BARLO include: Monte-Carlo Sampling methods and Mean-value/Riemann sum methods, which are developed by Zheng [88] for computation of integrals; and branch and bound methods for level set approximation (Hichert [26]). Furthermore, the algorithms suggested by Chew & Zheng [13] are improved and sped up using some duality and Newton techniques (Hichert *et al.* [27, 28]) through the *volume* *function* introduced by Phú & Hoffmann [48].

Hence, according to Thm. 4.4.4, BARLO could be used to determine essential infimum. Thus, to apply BARLO for the purpose at hand, it is required to ensure that *infimum and essential infimum are equal.* Hence, we need to answer the question: When is min = inf = ess inf?

Relation between min, inf, and ess inf

Theorem 4.4.5. *(Prop. 3.1, p. 176, Phú & Hoffmann [48]) Let* $f \in L_\infty(X)$*. If* X *is robust, Lebesgue measurable and* f *is u.r., then ess* $\inf_{x\in X} f(x) = \inf_{x\in X} f(x)$.

Corollary 4.4.6. *Let* $f \in L_\infty(X)$*. If* X *is compact, robust, and Lebesgue measurable; and* f *is u.r. and l.s.c., then ess* $\inf_{x\in X} f(x) = \min_{x\in X} f(x)$.

Thm 4.4.5 indicates that the concept of robustness allows us to minimize special kinds of discontinuous functions using the IGOM. However, the set of all u.r. functions is not a linear space (cf. Example 3.1, Phú & Hoffmann [48]), since the sum of two or more u.r. functions, generally, may not be u.r. Nevertheless, we have

Proposition 4.4.7. *(Zheng [88]) Let* $f : X \to \mathbb{R}$ *and* φ_d *be as defined in Sec. 4.4.2. If* f *is u.s.c. and* φ_d *is upper robust, then* $f + \lambda\varphi_d$ *is upper robust, for every* $\lambda > 0$.

Proposition 4.4.8. *Let* φ *and* φ_d *be as defined in sections 4.4.1 and 4.4.2, respectively. If* φ *is an upper robust and measurable function and* $\mathcal{M}$ *is a robust and measurable set, then* $\forall d > 0 : \varphi_d$ *is also upper robust and measurable.*

Proof. Given $c \in \mathbb{R}$

$$\{x \in X \mid \varphi_d(x) < c\} = \begin{cases} \emptyset & \text{if } c < 0 \\ \mathcal{M} & \text{if } 0 < c \leq d \\ \{x \in X \mid \varphi(x) < c - d\} & \text{if } c > d. \end{cases} \tag{4.4.9}$$

□

Consequently, it remains to guarantee the upper robustness and measurability of the marginal function φ of (PSIP) and the robustness and measurability of the admissible set $\mathcal{M}$ of (GSIP). While the upper robustness of φ follows from the discussion in Chapter 3 (Sec. 3.5), the rest could be verified by using simple arguments.

4.4.4 Robustness of the Admissible Set $\mathcal{M}$ of GSIP

Consider the lower level problem associated with the (GSIP)

$$(GO) \qquad\qquad v(x) := \inf_{t \in B(x)} G(x,t)$$

Recall that $\mathcal{M} = \{x \in X \mid G(x,t) \geq 0, \forall\, t \in B(x)\}$. Then using $v(\cdot)$ we write $\mathcal{M}$ as $\mathcal{M} = \{x \in X \mid v(x) \geq 0\}$. When the defining functions of the lower level problem (GO) are affine linear, Props. 2.2.5 and 2.2.6 yield the robustness of $\mathcal{M}$ (using Cor. 3.2.1 and Rem. 3.2.1). Along with this observation, the aim here is to provide some supplementary results based on certain topological assumptions.

Take the function $v : X \to \mathbb{R}$ and consider the set $\mathcal{M} := \{x \in X \mid v(x) \geq 0\}$ and let $\mathcal{M}_0 := \{x \in \mathbb{R}^n \mid v(x) = 0\}$.

Strong Slater condition (SSC): For each $x^0 \in \mathcal{M}_0$ and each neighborhood $N(x^0)$, in the relative topology of X, there is some $x \in N(x^0)$ such that $v(x) > v(x^0)$.

Lemma 4.4.9. *If v is l.s.c. and (SSC) holds on X, then $\mathcal{M}$ is a robust set.*

Proof. Clearly $\{x \in X \mid v(x) > 0\} \subset int(\mathcal{M})$ and $\mathcal{M} = \{x \in X \mid v(x) > 0\} \cup \{x \in X \mid v(x) = 0\}$. It then follows $\mathcal{M} = int(\mathcal{M}) \cup \{x \in X \mid v(x) = 0\} = int(\mathcal{M}) \cup \mathcal{M}_0$. Obviously, every element of $int(M)$ is a robust point of $\mathcal{M}$. And if $x^0 \in \mathcal{M}_0$, by assumption, for every neighborhood $N(x^0)$ we have $N(x^0) \cap int(\mathcal{M}) \neq \emptyset$. Then x^0 is a robust point of $\mathcal{M}$. Therefore, the set $\mathcal{M}$ is robust. $\square$

Note that under the assumption (A1), if $B(\cdot)$ is u.s.c. and $G(\cdot,\cdot)$ is l.s.c., then $v(\cdot)$ will be l.s.c. (cf. Aubin & Cellina [4]). Furthermore, the upper semi-continuity of $B(\cdot)$ follows from the continuity of its defining functions $h_i, i \in I$, and compactness of the set T (cf. Rem. 1.2.1(i), Chap. 1). In other words, the l.s.c. assumption in Lem. 4.4.9 on the function $v(\cdot)$ is somehow natural.

Proposition 4.4.10. *If $B(\cdot)$ is piece-wise u.s.c. and compact valued, G l.s.c., $X_0 = \{x \in X \mid B(x) = \emptyset\}$ is robust* [†]*, and $\{X_i, X_0\}_{i\in I}$ is a robust partition of X, and $v : X\backslash X_0 \to \mathbb{R}$ is given with*

$$\text{(GO)} \qquad v(x) := \inf_{t\in B(x)} G(x,t)$$

fulfils the (SSC) on each partition $X_i, i = 1, 2, .., r$ with respect to the relative topology on X_i, then the admissible set $\mathcal{M}$ of (GSIP) is robust.

Proof. $\mathcal{M}$ can be equivalently written as

$$\mathcal{M} = \{x \in X \mid v(x) \geq 0\} \cup X_0 = \bigcup_{i=1}^{r} \{x \in X_i \setminus X_0 \mid v(x) \geq 0\} \cup X_0.$$

By assumption we have the robustness of X_0. Since $B(\cdot)$ is piece-wise u.s.c. and compact valued and G is l.s.c., then v is l.s.c. on the partitioning sets X_i w.r.t. the relative topology (cf. Thm. 5, p. 52, Aubin & Cellina [4]). The (SSC) on each X_i yields that $\mathcal{M}_i := \{x \in X_i \mid v(x) \geq 0\}$ is a robust subset of X_i in the relative topology of X_i (by Lemma 4.4.9). By Lemma 3.4.17, $\mathcal{M}_i$ is robust in X. Therefore, $\mathcal{M}$ is a robust set, since it is a union of robust sets. □

Sufficient for the (SSC) may be conditions of the type *extended Mangasarian Fromowitz Constraint qualification (EMFCQ)*, whenever some differentiability assumptions made on G and the constraint functions describing the set-valued map $B(\cdot)$. Actually, it needs to be stressed that, the known strong conditions for nice behavior of a (GSIP) are expected to hold on each component of a suitable robust partition.

[†] *The robustness of the set $X_0 = \{x \in X \mid B(x) = \emptyset\}$ follows from the upper robustness of $B(\cdot)$, since $B(\cdot)$ is piecewise u.s.c. (cf. Chap. 3, Cor. 3.4.4)*

4.4.5 Measurability of Marginal Functions

The measurability of the marginal function φ follows from standard arguments from the literature. Consequently, we require the concept of measurability of set-valued mappings with given structures. For this, we mainly consult the book of Rockafellar & Wets [57].

Again, only results concerning the partitioning sets $X_i, i \in I$, of X need to be verified. With respect to functions and sets in $\mathbb{R}^n$, the ordinary notion of Lebesgue measurability is used. Thus, we simply say measurable instead of Lebesgue measurable.

Assumption (A3): $B\,|_{X_i}$ is l.s.c. on X_i in the relative topology of X_i for $i = 1, 2, ..., r$.

Basic Definitions and Results

Let $X, Y \subset \mathbb{R}^n$ be closed and measurable sets. The SV-map $M : X \rightrightarrows Y$ is called *measurable* iff $M(\cdot)$ is closed valued and $M^{-1}(C)$ is measurable for each closed set $C \subset Y$.

Let $\psi : X \times Y \to \mathbb{R} \cup \{-\infty, +\infty\} =: \overline{\mathbb{R}}$. Then the map $E_\psi(x) := \{(y, \alpha) \in Y \times \mathbb{R} \mid \psi(x, y) \leq \alpha\}$ is called the *epigraphical map* associated with ψ.

An extended real-valued function ψ is a *normal integrand* iff for each $x \in X$ the function $y \to \psi(x, y)$ is l.s.c. and its epigraphical map $E_\psi : X \rightrightarrows Y \times \mathbb{R}$ is measurable.

Theorem 4.4.11 (Thm. 2K, Rockafellar [56]).
If $\psi : X \times Y \to \overline{\mathbb{R}}$ is a normal integrand and $M : X \rightrightarrows Y$ is a measurable SV-map, then the marginal function $\varphi(x) = \inf_{y \in M(x)} \psi(x, y)$ is measurable.

Recalling the marginal value function $\varphi(\cdot)$ of (PSIP), if $M(\cdot)$ is measurable, then, by Thm. 4.4.11, φ will be measurable, since the distance function $p(\cdot)$ is measurable (due to its continuity). However, the measurability of the map $M(\cdot)$ still remains to be verified.

Special Structure of $M(x)$

The feasible SV-map $M(\cdot)$ of (PSIP) has a special structure defined by the constraint function G and the index map $B(\cdot)$. Thus, we give conditions on the map $B(\cdot)$ and the function G to guarantee the measurability of $M(\cdot)$, of the marginal and of the penalty functions φ and φ_d, and of the feasible set $\mathcal{M}$ of (GSIP). We say that $X_0, X_1, ..., X_r$ is a (robust and measurable) measurable partition of X iff $X_0, X_1, ..., X_r$ is a partition of X and all parts X_i are (robust and measurable) measurable. We assume further that $X_0 := \{x \in X \,|B\,(x) = \emptyset\}$ belongs to this partition.

Remark 4.4.2. Note that for an upper robust measurable SV-map $B(\cdot)$, the set $X_0 = \{x \in X \mid B(x) = \emptyset\}$ is both robust and measurable. The measurability of X_0 follows from the measurability of the set $X \setminus X_0 = Dom(B) = \{x \in X \mid B(x) \neq \emptyset\}$. The robustness has been guaranteed by Cor. 3.4.4 in Chap. 3.

Proposition 4.4.12. *Let X be some measurable subset of $\mathbb{R}^n$ and let $\varphi : X \to \mathbb{R}$ be a function. Suppose also that $X_0, X_1, ..., X_r$ is a measurable partition of X. If, for each $i \in \{0, 1, \ldots, r\}$, φ is measurable on X_i, then φ is measurable on X.*

Proof. Let $\{O_\alpha\}_{\alpha\in\Lambda}$ be the family of measurable sets in $\mathbb{R}^n$. Then, for each $i \in \{0, 1, \ldots, r\}$, the family of sets $\{X_i \cap O_\alpha\}_{\alpha\in\Lambda}$ is the family of measurable sets w.r.t. the relative topology on X_i. Thus, the σ-algebra $\sigma(\{X_i \cap O_\alpha\}_{\alpha\in\Lambda})$ defines the induced measure on X_i. As φ is measurable w.r.t. X_i, then for any measurable set $D \subset \mathbb{R}$ we have that $\varphi^{-1}(D) \cap X_i$ is measurable in X_i. Since X_i's are measurable in X, we have $\varphi^{-1}(D) \cap X_i$ is measurable in X. Since the set $\{0, 1, \ldots, r\}$ is finite (countable infinite is also possible here), we conclude that

$$\bigcup_{i\in I} (\varphi^{-1}(D) \cap X_i) = \varphi^{-1}(D)$$

is measurable in X. Consequently, φ is measurable on X. □

Hence, based on Prop. 4.4.12, we are required to verify the measurability of the φ of (PSIP) only on each of the partitioning sets X_i of X. For this, we need to assure the measurability of $M(\cdot)$. Hence, using the function $m(\xi, x) = \inf_{t \in B(x)} G(\xi, t)$, we write $M(x) = \{\xi \in X \mid m(\xi, x) \geq 0\}$.

Proposition 4.4.13. *Let $(X_i)_{i \in I}$ a measurable partition of X. If assumptions (A1) & (A3) hold true, $B(\cdot)$ is u.s.c. with compact values and $G : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ is continuous, then $-m$ is a normal integrand on $\mathbb{R}^n \times X_i$, for each $i \in I$. In this case, the set-valued map $M(x) = \{\xi \in X \mid -m(\xi, x) \leq 0\}$ is measurable on X_i, for each $i \in I$.*

Proof. We use the redefinition $\widetilde{G}(\xi, x, t) := G(x, t)$ and $\widetilde{B}(\xi, x) := B(x)$. Hence, we can write

$$-m(\xi, x) = \sup_{t \in \widetilde{B}(\xi, x)} -\widetilde{G}(\xi, x, t).$$

By assumption $\widetilde{B}|_{\mathbb{R}^n \times X_i}$ is continuous and compact valued; and $\widetilde{G}(\xi, x, t)$ is also continuous. These imply that $-m$ is continuous on $\mathbb{R}^n \times X_i$ (cf. Thm. 6, p. 53, Aubin & Cellina [4]). Since, X_i is measurable we conclude that $-m$ is a normal integrand on $\mathbb{R}^n \times X_i$ (cf. Exa. 14.31 in Rockafellar & Wets [57]). Moreover, using Prop. 14.33 of [57] and the closedness of the set X, the map

$$M(x) = \{\xi \in \mathbb{R}^n \mid -m(\xi, x) \leq 0\} \cap X = \{\xi \in X \mid -m(\xi, x) \leq 0\}$$

is measurable on X_i, for each $i \in I$. □

Theorem 4.4.14. *(Measurability of the marginal function) Let φ be the marginal function of (PSIP) and let $(X_i)_{i \in I}$ be a (robust) measurable partition of X. If ψ is a normal integrand on $X \times X$ and assumptions (A1), (A2) and (A3) hold true, $B(\cdot)$ is u.s.c. with compact values and $G : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ is continuous, then φ is measurable on each X_i; hence, measurable on X.*

Proof. Clearly ψ is a normal integrand and $M(\cdot)$ is closed valued and measurable, by Prop. 4.4.13, on X_i. Moreover, by Thm. 4.4.11, φ is measurable w.r.t. X_i. Therefore, by Prop. 4.4.12, φ is measurable. □

Corollary 4.4.15. *If assumptions (A1), (A2) & (A3) hold true, $B(\cdot)$ is u.s.c. with compact values and $G : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ is continuous, then the feasible set $\mathcal{M}$ of (GSIP) is measurable. Hence, φ_d is also measurable.*

Proof. By Prop. 4.4.1 $\mathcal{M} = \varphi^{-1}(0)$. Morevoer, under assumptions (A1), (A2) and (A3), φ is measurable, by Thm. 4.4.14. Consequently, $\mathcal{M}$ is measurable. Furthermore, for the function φ_d (see Section 4.4.2), given $c \in \mathbb{R}$, we have (Eqn. 4.4.9) from which follows the measurability of φ_d due to the measurability of φ and that of $\mathcal{M}$. □

Before winding up this section, it needs to be stressed that, the known strong conditions for nice behavior of (GSIP) are expected to hold on each component of a suitable robust measurable partition in the associated relative topology. What is really needed apriori, in any case, is the robustness of the set of all x where the image of $B(\cdot)$ is empty - which has been actually guaranteed by Cor. 3.4.4. In the known examples of (GSIP), given by [34, 70, 72], with ill behavior, the assumptions (A1)-(A3) can be principally satisfied. However, for a few of these examples the free choice of the functions $(G, h_i, \in I)$ must be properly done, so that the assumptions made here hold true.

Furthermore, the method suggested is mainly conceptual, waiting a practical implementation. However, it could serve as a starting point for the development of some fixed point proximal type algorithm. At the same time, here it is assumed that the function G to depend on both variables x and t. If G does not depend on x, then we will have

$$\mathcal{M} = \{x \in X \mid G(t) \geq 0, \forall t \in B(x)\}$$

$$M(x) = \{\xi \in X \mid G(t) \geq 0, t \in B(x)\}$$

and using, for instance, $\psi(x, \xi) = \|x - \xi\|$ it follows that

$$\varphi(x) = \inf_{\xi \in M(x)} \psi(x, \xi) = \begin{cases} 0 & \text{if } x \in \mathcal{M}, \\ +\infty & \text{if } x \notin \mathcal{M}. \end{cases} \tag{4.4.10}$$

which is the indicator function of the set $\mathcal{M}$. Under such instances the proposed approach might not be interesting. Consequently, practical problems like the *Design Centering* problems (see Stein [75]) have feasible sets of the above type and they are out of the considerations of this approach. However, the second approach can be used for the treatment of Design Centering problems.

4.5 Penalty Methods with Discretization

In this section we discuss a second variant of penalty method. There are two penalty problems presented here. One defined using the (marginal) value function of the lower level problem; and a second defined through the *proximity* function of the feasible set $\mathcal{M}$ of (GSIP). To justify the relevance of the approach, relations between (GSIP) and the penalty problems have been studied. The discussion uses mainly minimizing sequences (cf. 4.3).

The second half of this section indicates how one of the penalty problems could be used in a numerical computation, to determine a coarse approximation to the optimal value of a (GSIP). Specifically, we take up the penalty problem defined through the lower level problem (GO) of (GSIP), introduce a discretization of the underlying set T and show that the values of the discretized penalty problem provide bounds to the optimal value of (GSIP) in the limit.

4.5.1 Two Penalty Problems

We consider again

$$(GSIP) \qquad \begin{array}{l} f(x) \to inf \\ G(x,t) \geq 0, \forall t \in B(x) \\ x \in X \end{array}$$

and make the stronger assumptions

Assumption (A4):

- $B(x) = \{t \in T \mid h_i(x,t) \leq 0, i = 1, \ldots, p\}$;
- $f : X \to \mathbb{R}$ a Lipschitz continuous function modulo L_f;
- $G : X \times T \to \mathbb{R}$ l.s.c. on X;
- $h_i : X \times T \to \mathbb{R}, i = 1, \ldots, p$, are continuous on $X \times T$.

At times X and T may be treated as metric spaces. We also suppose that $X \subset \mathbb{R}^n$, $T \subset \mathbb{R}^m$. In view of the intention to solve (GSIP) with BARLO, we let X and T to be compact sets. Hence, if X is a compact set, then each sequence $\{x^n\}$ in X has at least one accumulation point $\overline{x}$ in X. Moreover, with the compactness of T, we have $B(x) := \{t \in \mathbb{R}^m \mid h_i(x,t) \leq 0, i = 1, \ldots, p\} \cap T$; i.e., $B(\cdot)$ is compact valued and upper semi-continuous.

Once more, consider the marginal function $v : X \to \mathbb{R}$ of the lower level problem (GO) of (GSIP)

$$v(x) = \inf_{t \in B(x)} G(x,t),$$

and the feasible set $\mathcal{M}$ of (GSIP) being written as

$$\mathcal{M} = \{x \in X \mid v(x) \geq 0\}.$$

We require here the strict $r-$regularity of $[-v]^+$(cf. also Def. 3.6.2); meaning that, there is a strictly increasing continuous function r from $\mathbb{R}_+$ into $\mathbb{R}_+$ such that

$$dist(x, \mathcal{M}) =: \inf_{z \in \mathcal{M}} d(x,z) \leq r([-v(x)]^+), \ \forall x \in X.$$

Specifically, we may take

$$r(t) = C \cdot t, \ C > 0,$$

to obtain simpler forms.

Assumption (A5): In the following we suppose that $B(\cdot)$ is piece-wise l.s.c. on X, $\mathcal{M}$ is a robust subset of X and $(-v)^+$ is $r-$regular on X w.r.t. $\mathcal{M}$.

We consider the following two penalty functions:

$$p(x) = f(x) + \lambda dist(x, \mathcal{M}) + \alpha \cdot \chi(x), \tag{4.5.1}$$

$$q(x) = f(x) + \lambda r([-v(x)]^+) + \alpha \cdot \chi(x), \tag{4.5.2}$$

in which the function $\chi(\cdot)$ is the *indicator function* of $X \setminus \mathcal{M}$ with

$$\chi(x) := \begin{cases} 1, & \text{if } x \notin \mathcal{M} \\ 0, & \text{if } x \in \mathcal{M}. \end{cases}$$

We also use the short form

$$s(x) := r([-v(x)]^+).$$

Thus, we consider the following two penalty problems:

$$(P_{\lambda\alpha}) \qquad p(x) \to inf,\ x \in X,$$

and

$$(Q_{\lambda\alpha}) \qquad q(x) \to inf,\ x \in X.$$

Since the functions $\chi(\cdot)$ and $s(\cdot)$ are generally not continuous, we use the notions of minimizer, generalized minimizer and minimizing sequences for discontinuous optimization problems as was introduced in Sec. 4.3.

Proposition 4.5.1 (cf. also Clark [14]).

Let $\lambda > L_f$, $\alpha \geq 0$ and $\inf_{x\in\mathcal{M}} f(x) > -\infty$. Then the following hold true:

i)

$$\inf_{x\in X} p(x) = \inf_{x\in X} q(x) = \inf_{x\in\mathcal{M}} f(x). \tag{4.5.3}$$

ii)

$$MS(GSIP) \subset MS(Q_{\lambda\alpha}) \subset MS(P_{\lambda\alpha}) \subset GMS(GSIP). \tag{4.5.4}$$

iii) If, additionally, $\alpha > 0$, then

$$MS(Q_{\lambda\alpha}) \subset MS(P_{\lambda\alpha}) \subset MS(GSIP)$$

Proof. (i) If $x \in \mathcal{M}$, then $dist(x, \mathcal{M}) = s(x) = \chi(x) = 0$. Further, for all $x \in X: \; dist(x, \mathcal{M}) \leq s(x)$. Hence,

$$\begin{aligned} \inf_{x \in X} p(x) &\leq \inf_{x \in X} q(x) \\ &\leq \inf_{x \in \mathcal{M}} q(x) = \inf_{x \in \mathcal{M}} p(x) = \inf_{x \in \mathcal{M}} f(x). \end{aligned}$$

Therefore, it suffices to show that $\inf_{x \in \mathcal{M}} f(x) \leq \inf_{x \in X} p(x)$. Given an arbitrary $x \in X$ and a sufficiently small $\eta > 0$, there is $z_\eta \in \mathcal{M}$ such that

$$dist(x, \mathcal{M}) \geq dist(x, z_\eta) - \eta$$

which implies that

$$L_f \, dist(x, \mathcal{M}) \geq L_f \, dist(x, z_\eta) - L_f \eta.$$

Using the Lipschitz continuity of f, we have

$$L_f \, dist(x, \mathcal{M}) + L_f \, \eta \geq f(z_\eta) - f(x).$$

From this follows that

$$f(x) + L_f \, dist(x, \mathcal{M}) + L_f \, \eta \geq f(z_\eta).$$

Hence,

$$\begin{aligned} &f(x) + \lambda dist(x, \mathcal{M}) + L_f \, \eta \\ &\qquad \geq f(z_\eta) + (\lambda - L_f) dist(x, \mathcal{M}). \end{aligned}$$

Moreover,

$$f(x) + \lambda dist(x, \mathcal{M}) + \alpha\chi(x) + L_f\eta \tag{4.5.5}$$
$$\geq f(z_\eta) + (\lambda - L_f)dist(x, \mathcal{M}) + \alpha\chi(x).$$

But $\lambda > L_f, \alpha \geq 0$ and $\chi(x) \geq 0$, for all $x \in X$, imply

$$f(x) + \lambda dist(x, \mathcal{M}) + \alpha\chi(x) + L_f\eta \geq f(z_\eta).$$

Thus,

$$p(x) + L_f\eta \geq \inf_{x\in cl(\mathcal{M})} f(x), \forall x \in X \; and \; \eta > 0.$$

Since $\eta > 0$ has been chosen arbitrarily, we have

$$p(x) \geq \inf_{x\in\mathcal{M}} f(x), \forall x \in X,$$

where we used the fact that $\inf_{x\in cl(\mathcal{M})} f(x) = \inf_{x\in\mathcal{M}} f(x)$, for a Lipschitz-continuous function f. Consequently,

$$\inf_{x\in X} p(x) \geq \inf_{x\in\mathcal{M}} f(x),$$

which was what we intended to verify. Therefore,

$$\inf_{x\in X} p(x) = \inf_{x\in X} q(x) = \inf_{x\in\mathcal{M}} f(x).$$

(ii) Let $\{x^n\} \subset X$ and $\{x^n\} \in MS(P_{\lambda\alpha})$. Then using the relation (4.5.5), we obtain

$$p(x_n) + L_f\,\eta \geq \inf_{x\in\mathcal{M}} f(x) + (\lambda - L_f)dist(x_n, \mathcal{M}) + \alpha\chi(x_n), \forall n \in \mathbb{N}.$$

Hence,

$$\lim_{n\to\infty} p(x_n) + L_f\eta \geq \inf_{x\in\mathcal{M}} f(x) + (\lambda - L_f) \lim_{n\to\infty} dist(x_n, \mathcal{M}) + \lim_{n\to\infty} \alpha\chi(x_n).$$

Since $\{x^n\} \in MS(P_{\lambda\alpha})$, we have

$$\inf_{x\in X} p(x) + L_f\,\eta \geq \inf_{x\in\mathcal{M}} f(x) + (\lambda - L_f) \lim_{n\to\infty} dist(x_n, \mathcal{M}) + \alpha\chi(x_n).$$

Applying part (i), we get

$$0 \leq (\lambda - L_f) \lim_{n\to\infty} dist(x_n, \mathcal{M}) + \alpha\chi(x_n) \leq L_f\,\eta, \forall \eta > 0.$$

But this, using $\lambda > L_f$, implies that

$$\lim_{n\to\infty} dist(x_n, \mathcal{M}) = 0, \tag{4.5.6}$$

and

$$\lim_{n\to\infty} \chi(x_n) = 0.$$

It then follows that: if $\alpha = 0$, then $\{x_n\} \in GMS(GSIP)$; and, in case $\alpha > 0$, then $\{x_n\} \in MS(GSIP)$.

Let now $\{x_n\} \in MS(Q_{\lambda\alpha})$. By $r-$regularity we have that

$$p(x) \leq q(x), \forall x \in X \Rightarrow p(x_n) \leq q(x_n), \forall n \in \mathbb{N}$$

We then observe that

$$\inf_{x\in X} p(x) \leq \lim_{n\to\infty} p(x_n) \leq \lim_{n\to\infty} q(x_n) = \inf_{x\in X} q(x).$$

Applying part (i) again, we conclude that

$$\inf_{x\in X} p(x) = \lim_{n\to\infty} p(x_n).$$

Consequently, $\{x_n\} \in MS(P_{\lambda\alpha})$. Therefore,

$$MS(Q_{\lambda\alpha}) \subset MS(P_{\lambda\alpha}).$$

Furthermore, if $\{x_n\} \in MS(GSIP)$, then there is $n_0 \in \mathbb{N}$ such that $x_n \in \mathcal{M}$ for all $n \geq n_0$. This implies that

$$p(x_n) = q(x_n) = f(x_n), \forall n \geq n_0.$$

Along with part i) we have

$$\lim_{n\to\infty} q(x_n) = \lim_{n\to\infty} f(x_n) = \inf_{x\in\mathcal{M}} f(x) = \inf_{x\in X} q(x).$$

From this follows that $MS(GSIP) \subset MS(Q_{\lambda\alpha})$, and that completes the proof.

□

Remark 4.5.1. We make the following observation.

Let $\{x_n\} \in MS(P_{\lambda\alpha}) \subset X$ or $\{x_n\} \in MS(Q_{\lambda\alpha}) \subset X$ and X is a compact set. Then $\{x_n\}$ has a limit point in X. Suppose $\overline{x} \in X$ is any limit point of $\{x_n\}$; meaning that, there is a subsequence $\{x_n^k\} \subset \{x_n\}$ such that $x_n^k \to \overline{x}$. Assuming $\lambda > L_f$, we know, from (4.5.6), that

$$\lim_{n\to\infty} dist(x_n, \mathcal{M}) = 0.$$

Using the continuity of the distance function $\phi(x) := dist(x, \mathcal{M})$ (even Lipschitz continuous, see Clark [14]), we obtain that

$$dist(\overline{x}, \mathcal{M}) = \lim_{k\to\infty} dist(x_n^k, \mathcal{M}) = \lim_{n\to\infty} dist(x_n, \mathcal{M}) = 0.$$

From which follows that $\overline{x} \in cl(\mathcal{M})$. At the same time,

$$\lim_{k\to\infty} (f(x_n^k) + \lambda dist(x_n^k, \mathcal{M}) + \alpha\chi(x_n^k)) =$$
$$\lim_{k\to\infty} p(x_n^k) = \lim_{n\to\infty} p(x_n) = \inf_{x\in X} p(x) = \inf_{x\in\mathcal{M}} f(x).$$

Hence, trivially,

$$\lim_{k\to\infty} f(x_n^k) \leq \inf_{x\in\mathcal{M}} f(x) \Rightarrow f(\overline{x}) \leq \inf_{x\in\mathcal{M}} f(x).$$

Then, using the continuity of f and $\overline{x} \in cl(\mathcal{M})$,

$$\inf_{x\in cl(\mathcal{M})} f(x) \leq f(\overline{x}) \leq \inf_{x\in\mathcal{M}} f(x).$$

Once again, for a continuous function f, $\inf_{x\in cl(\mathcal{M})} f(x) = \inf_{x\in\mathcal{M}} f(x)$, yielding

$$f(\overline{x}) = \inf_{x\in\mathcal{M}} f(x)$$

and $\overline{x} \in cl(\mathcal{M})$. Therefore, $\overline{x}$ is a generalized minimizer of (GSIP); i.e. $\overline{x} \in \overline{arg}(GSIP)$. We then deduce that, given the assumptions made above hold true, every minimizing sequence of $(P_{\lambda\alpha})$ or that of $(Q_{\lambda\alpha})$ would give us a generalized minimizer of (GSIP). Furthermore, if we know that $\mathcal{M}$ is a closed set, then the result will be a minimizer.

Next, we find a special instances of Prop. 4.5.1.

Corollary 4.5.2. *Let X be a compact set, $\lambda > L_f$, $\alpha = 0$ and $\inf_{x\in\mathcal{M}} f(x) > -\infty$. Then*

i)

$$MS(Q_{\lambda 0}) \subset MS(P_{\lambda 0}) = GMS(GSIP).$$

ii) If $s(\cdot)$ is a continuous function and $\mathcal{M}$ is a closed set, then for every $\{x_n\} \in GMS(GSIP)$, there is a subsequence $\{x_n^k\} \subset \{x_n\}$ such that $\{x_n^k\} \in MS(Q_{\lambda 0})$. Specifically, each convergent generalized minimizing sequence of (GSIP), is a minimizing sequence of $(Q_{\lambda 0})$.

Proof. (i) Follows trivially from Prop. 4.5.1(i) using $\alpha = 0$.

(ii) Let $\{x_n\} \in GMS(GSIP)$. Then $\lim_{n\to\infty} dist(x_n, \mathcal{M}) = 0$. In particular, using the discussion in Rem. 4.5.1, there is a subsequence $\{x_n^k\} \subset \{x_n\}$ such that $x_n^k \to \overline{x}$, for some $\overline{x} \in X$ and $\overline{x} \in cl(\mathcal{M})$. But $cl(\mathcal{M}) = \mathcal{M}$ implies $\overline{x} \in \mathcal{M}$. Consequently, $s(\overline{x}) = 0$. Then, by the continuity of $s(\cdot)$, we have $\lim_{k\to\infty} s(x_n^k) = s(\overline{x}) = 0$. Moreover,

$$\begin{aligned}
\inf_{x\in X} q(x) \leq q(x_n^k) \Rightarrow \inf_{x\in X} q(x) \\
\leq \lim_{k\to\infty} q(x_n^k) = \lim_{k\to\infty} \left[f(x_n^k) + \lambda \underbrace{s(x_n^k)}_{\to 0} \right] \\
= \lim_{k\to\infty} f(x_n^k) = \lim_{n\to\infty} f(x_n) = \inf_{x\in\mathcal{M}} f(x).
\end{aligned}$$

Using Prop. 4.5.1(i), we have that

$$\inf_{x\in X} q(x) = \lim_{k\to\infty} q(x_n^k).$$

Therefore, $\{x_n^k\} \in MS(Q_{\lambda 0})$.

□

Remark 4.5.2.

i) In general, for $\alpha = 0$, equality in (4.5.4) may not hold even if $s(\cdot)$ is a continuous function and $\mathcal{M}$ is a closed set, as Cor. 4.5.2(ii) indicates.

ii) When $\mathcal{M}$ is a closed set, then the most $\chi(\cdot)$ can be is l.s.c. However, this does not help so much, even to prove part of the equality in (4.5.4).

4.5.2 A Discretization Method

For computational reason, we consider only the penalty problem $(Q_{\lambda\alpha})$ with $\alpha = 0$. Thus, to determine the values of the function $[-v]^+$ from the lower level problem (GO) of (GSIP), we need to work with a discretization of the set T. In every step of discretization, the discretized problem $(Q_{\lambda\alpha})$ will be solved by IGOM.
Thus, to make sure that IGOM is applicable for the purpose at hand, we require to verify additional properties of robustness. One of these is

Proposition 4.5.3. *If $B(\cdot)$ is upper robust, $B(x)$ is compact for each $x \in X$, G is l.s.c., then $[-v]^+$ is upper robust.*

Proof. Since,

$$[-v(x)]^+ = \sup_{t\in B(x)} ([-G(x,t)]^+),$$

then $[-v]^+$ is upper robust.(cf. Thm. 3.5.6, Chap. 3). □

Now we use the following penalty approach under the satisfaction of strict $r-$regularity of $\mathcal{M}$ with $r(t) = Ct$.

Assumption (A5):

$$dist(x, \mathcal{M}) \leq C([-v(x)]^+).$$

For $\lambda > C \cdot L_f$, take the problem

$$q(x) = f(x) + \lambda \cdot [-v(x)]^+ \quad \rightarrow \quad inf$$
$$x \quad \in \quad X$$

and try to find coarse minimizers using the global optimization routine - BARLO.

Each function value of $v(x)$ needs the global solution of the non-linear programming problem

$$(P^l) \qquad \begin{array}{rcl} G(x,t) & \to & min \\ t & \in & B(x). \end{array}$$

In the first phase of using BARLO, we need only a coarse solution of P^l. Since, $B(x)$ is embedded in T, $B(\cdot)$ is upper robust, i.e. $B^{-1}(U)$ a robust set for each open set in X. Therefore, we use a grid $T_n = \{t_1, \ldots, t_n\} \subset T$ and consider the following simplified lower level problem

$$(P^l_n) \qquad \begin{array}{rcl} G(x,t) & \to & min \\ t & \in & B(x) \cap T_n. \end{array}$$

The solution of this discretized problem could be realized by simpler routines. Under upper robustness (or upper semi-continuity) of $B(\cdot)$, the function

$$[-v_n(x)]^0 := \sup_{t \in B(x) \cap T_n} [-G(x,t)]^+$$

remains upper robust (upper semi-continuous) as the statement below claims.

Proposition 4.5.4. *If $B(\cdot)$ is an u.r. [u.s.c.] SV-map with compact values and G l.s.c., then $[-v_n]^+$ is u.r. [u.s.c.]*

Proof. Let $x^0 \in X$. Since $B(\cdot)$ is compact valued and u.r., by Prop. 3.4.6, for every $\varepsilon > 0$, there is a semi-neighborhood $SNH_\varepsilon(x^0)$ such that

$$\forall x \in SNH_\varepsilon(x^0) : B(x) \subset B(x^0) + B_\varepsilon.$$

We can choose a sufficiently small $\varepsilon > 0$, by the closedness of $B(x^0)$, so that

$$\Big[B(x^0) + B_\varepsilon\Big] \cap T_n = B(x^0) \cap T_n. \tag{4.5.7}$$

Hence,

$$B(x) \cap T_n \subset B(x^0) \cap T_n, \forall x \in SNH_\varepsilon(x^0).$$

Moreover, by assumption, $-G(\cdot, t_i)$ is u.s.c. for every fixed $t_i \in T_n$. Hence, there is a neighborhood $U^i_\varepsilon(x^0)$ of x^0 such that

$$\forall x \in U^i_\varepsilon(x^0): \quad -G(x, t_i) \leq -G(x^0, t_i) + \varepsilon$$

Consequently,

$$N(x^0) := \bigcap_{t_i \in T_n \cap B(x^0)} U^i_\varepsilon(x^0) \cap SNH_\varepsilon(x^0)$$

is a semi-neighborhood of x^0 (by Rem. 3.2.1). Hence, for $x \in N(x^0)$ we have

$$\begin{aligned} [-v_n(x)]^+ &= \max_{t_i \in B(x) \cap T_n} \Big\{ -G(x, t_i), 0 \Big\} \\ &\leq \max_{t_i \in B(x^0) \cap T_n} \Big\{ -G(x, t_i), 0 \Big\} \\ &\leq \max_{t_i \in B(x^0) \cap T_n} \Big\{ -G(x^0, t_i) + \varepsilon, 0 \Big\} \\ &\leq \max_{t_i \in B(x^0) \cap T_n} \Big\{ -G(x^0, t_i), 0 \Big\} + \varepsilon \\ &= [-v_n(x^0)]^+ + \varepsilon. \end{aligned}$$

Consequently,

$$\forall x \in N(x^0): \quad [-v_n(x)]^+ \leq [-v_n(x^0)]^+ + \varepsilon.$$

Hence, $[-v_n(\cdot)]^+$ is upper robust (cf. Prop. 3.2.5). To prove the upper semi-continuity, replace the semi-neighborhood $SNH(x^0)$ by a neighborhood $U(x^0)$. □

In prop. 4.5.4, the upper robustness of $B(\cdot)$ is important. In fact, if $B(\cdot)$ is not u.r., $[-v_n(\cdot)]^+$ may not be u.r.

Example 4.5.5. *Let $G : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ be given as $G(x, t) = -t$ and let $B : X \rightrightarrows T$ be*

$$B(x) := \begin{cases} [1, 4]\,, \text{ if } x \neq x_0 \\ [2, 3]\,, \text{ if } x = x_0. \end{cases}, \quad x_0 = 1.$$

Hence, G is a continuous function and $B(\cdot)$ is compact valued and l.s.c. (hence l.r.), but $B(\cdot)$ is not u.r. Take $T = [1, 4]$ and a fine discretization T_n with $3, 4 \in T_n$. Consequently, we have

$$[-v_n(x)]^+ = \max_{t_i \in B(x) \cap T_n} \Big\{ - G(x,t), 0 \Big\} = \begin{cases} 4, \text{ if } x \neq x_0 \\ 3, \text{ if } x = x_0. \end{cases}$$

Hence, $[-v_n(x)]^+$ is not upper robust.

Remark 4.5.3. However, a similar statement as in Prop. 4.5.4 fails to exist for the lower robustness of $[-v_n(x)]^+$ if we assume that $G(\cdot, t)$ is u.s.c. and $B(\cdot)$ is lower robust.

Remark 4.5.4. Along with Prop. 4.5.3 it is worth to note that, assuming the function $G(\cdot, t) : \mathbb{R}^n \to \mathbb{R}$ is continuous, the function v_n is the infimum of a finite number of measurable functions; hence, both v_n and $[-v_n]^+$ will be measurable, for each $n \in \mathbb{N}$. Thus, measurability will not be an issue of further discussion.

The simplified problem (P_n^l) can now be solved using BARLO to find a coarse approximation to the optimal value of the (GSIP). Starting with a coarse grid, in each subsequent iteration of BARLO, we refine the discretization in the neighborhood of approximate minimizers of (P_n^l).

Since $v_n(x) \geq v(x)$ and BARLO works with Monte-Carlo sampling and level-set shrinking methods, we get an upper estimate of the infimum of (GSIP). For problems with lower dimension in t and fast function evaluation of $G(x, t)$, $f(x)$ and $h(x, t)$, BARLO gives acceptable approximations in a moderate CPU time.

The original intention was to use the penalty problem (see Sec. 4.4.2)

$$(PP_{\lambda d}) \qquad\qquad f(x) + \lambda \varphi_d(x) \to inf.$$

in determining an approximate solution for the (GSIP). However, initial computational experiments indicated that the computation of this penalty problem is too expensive. Furthermore, the more or less exact evaluation of the problem (P^l) does the same job in a much shorter time.

If the level sets around the expected generalized solution is determined, one can continue with (P^l) to determine values of $v(\cdot)$ by using fast local methods, in the neighborhood of the last minima in t.

If the grid density goes to zero, then we get a formal convergence (see next section). But this might be hard to realize numerically.

4.5.3 Convergence of the discretization method for the penalization approach

Let T' be a grid of T, T compact. We call

$$\Delta T' = \max_{t \in T} \min_{\tau \in T'} \|t - \tau\| \tag{4.5.8}$$

the *density* of T' in T and say that grid T'' is finer than T' if $\Delta T'' < \Delta T'$. From the triangle inequality, it follows immediately that each grid point $t' \in T'$ has at least one neighboring grid point $t'' \in T''$ with

$$\|t' - t''\| \leq 2\Delta T'$$
$$\|t' - t''\|_2 \leq \frac{2\Delta T'}{\sqrt{m}} \text{ in } \mathbb{R}^m \text{ with Euclidean norm.}$$

Actually, this is true if we consider equidistant grid points (cf. p. 140 Hettich & Zencke [25]).

We consider a sequence T_n of grids of T with

$$\lim_{n \to \infty} \Delta_n = 0, \quad \text{where } \Delta_n := \Delta T_n.$$

We further assume that $\bar{x}$ is a generalized solution of the problem

$$\begin{aligned} f(x) &\rightarrow \inf \\ G(x,t) &\geq 0 \;\; \forall t \in B(x) \\ x &\in X, \end{aligned}$$

where

$$B(x) = \{t \in T \mid h_i(x,t) \leq 0,\ i = 1,\dots,p\}.$$

And $f : X \to \mathbb{R}, G : X \times T \to \mathbb{R}$ and $h : X \times T \to \mathbb{R}^p$ (i.e. $h = (h_1,\dots,h_p)$) are functions with the following properties:

- f is a Lipschitzian function with Lipschitz constant L_f;
- G is *Lipschitz continuous in* t *with constant* L_G uniformly on $S \subset X$; i.e.

$$\exists L_G > 0, \forall x \in S, \forall t, t^{'} \in T :$$
$$|G(x,t) - G(x,t^{'})| \leq L_G \|t - t^{'}\|;$$

- $h_i : X \times T \to \mathbb{R}, i = 1,\dots,p$, are Lipschitz continuous at t with constant L_h uniformly on $\widetilde{S} \subset X$.

Furthermore, let for each $\varepsilon \geq 0$

$$\begin{aligned} B^{\varepsilon}(x) &= \{t \in T \,|h(x,t) \leq \varepsilon\}, \\ B^{\varepsilon}_n(x) &= B^{\varepsilon}(x) \cap T_n, \end{aligned}$$

and let v^{ε}_n be the discrete version of v with respect to $B(x)$ defined as

$$\begin{aligned} v^{\varepsilon}(x) &:= \min_{t \in B^{\varepsilon}(x)} G(x,t) \text{ and} \\ v^{\varepsilon}_n(x) &:= \min_{t \in B^{\varepsilon}_n(x)} G(x,t). \end{aligned}$$

Remark 4.5.5. Given an $\varepsilon > 0$, if we assume that the SV-map $B^{\varepsilon}(\cdot)$ is upper robust with compact values, then we could guarantee the upper robustness of the marginal value function $[-v^{\varepsilon}_n(\cdot)]^+$ in the same manner as in Cor. 4.5.4.

Subsequently, we have the functions

$$\begin{aligned} q(x) &= f(x) + \lambda[-v(x)]^+ \\ q^{\varepsilon}_n(x) &= f(x) + \lambda[-v^{\varepsilon}_n(x)]^+,\ \varepsilon \geq 0, n \in \mathbb{N} \end{aligned}$$

on X.

Lemma 4.5.6. *Let* B *be the unite ball of* $\mathbb{R}^m$.

i) The Lipschitz continuity of $h := (h_i)_{i\in I}$ *uniformly on* $\widetilde{S} \subset X$ *implies*

$$B(x) + \frac{\varepsilon}{L_h}\mathsf{B} \subset B^\varepsilon(x)$$

on $\widetilde{S}$ *for each* $\varepsilon \geq 0$.

ii) Given $\varepsilon \geq 0$, *if the defining system of inequalities of* $B(\cdot)$ *is r-regular (cf. Chap. 1, Def. 1.2.1*[‡]*) for* $x \in \widetilde{S}$, *then*

$$B^\varepsilon(x) \subset B(x) + r(\varepsilon)\mathsf{B}.$$

Proof. i) For $t \in B(x)$ the statement is obvious. Let $t \in \left(B\left(x\right) + \frac{\varepsilon}{L_h}\mathsf{B}\right) \setminus B\left(x\right)$. Then we find some $t' \in B\left(x\right)$, i.e $h_i(x,t') \leq 0$, such that $\|t - t'\| \leq \frac{\varepsilon}{L_h}$. From $h_i\left(x,t\right) \leq h_i\left(x,t\right) - h_i\left(x,t'\right) \leq L_h\|t-t'\|$ for all $i \in I$, follows that $h_i(x,t) \leq \varepsilon, \forall i \in I$; i.e. $t \in B^\varepsilon(x)$.

ii) Let $t \in B^\varepsilon(x)$ be arbitrary; i.e. $h_i\left(x,t\right) \leq \varepsilon, i \in I$. Which in turn implies that $h_i\left(x,t\right)^+ \leq \varepsilon, \forall i \in I$. Hence, $\|h(x,t)^+\| \leq \varepsilon$. Consequently, by r-regularity, we have that

$$dist(t, B(x)) \leq r(\|h(x,t)^+\|) \leq r(\varepsilon).$$

□

The inclusions in Lem. 4.5.6 are applicable in the convergence analysis of the envisaged discretization approach, where we use $B^\varepsilon(x)$ instead of $B(x)$. The significance of the inclusions, particularly the one in Lem. 4.5.6(ii), lies in the fact that the set $B^\varepsilon(x)$ could help not to miss some important discretization points of T_n which might fall out of $B(x)$; in case $B(x)$ has some difficult structures. Furthermore, given $\varepsilon > 0$, for a grid density $\Delta_n \leq r(\varepsilon)$, the points that lie in $B^\varepsilon(x)$ are with in the $r(\varepsilon)$ neighborhood of $B(x)$.
We need the following elementary Lemma.

[‡] *The defining system of* $B(\cdot)$ *is r-regular if* $dist(t, B(x)) \leq r(\|h(x,t)^+\|), \forall x \in X$, *for a continuous non-decreasing function* $r : \mathbb{R}_+ \to \mathbb{R}_+$

Lemma 4.5.7. *Let $\alpha, \beta, \gamma \in \mathbb{R}$. If $\alpha \leq \beta + \gamma$ and $\gamma \geq 0$, then $[\alpha]^+ \leq [\beta]^+ + \gamma$.*

Proof. We identify the following four cases

Case-1: $\alpha, \beta \geq 0$,

Then $\alpha, \beta \geq 0$ implies $\alpha = [\alpha]^+ \leq \beta + \gamma = [\beta]^+ + \gamma$.

Case-2: if $\alpha < 0 \leq \beta$,

then $[\alpha]^+ = 0 \leq \beta + \gamma = [\beta]^+ + \gamma$.

Case-3: if $\alpha, \beta < 0$,

then $[\alpha]^+ = 0 = [\beta]^+ \leq [\beta]^+ + \gamma$.

Case-4: if $\alpha \geq 0, \beta < 0$,

then $[\alpha]^+ = \alpha \leq \beta + \gamma \leq [\beta]^+ + \gamma$.

□

To prepare the ground for the statement of convergence, we also require the following two lemmas on the various marginal functions.

Lemma 4.5.8.

i) For each $\varepsilon \geq 0$

$$\begin{aligned} v^{\varepsilon}(x) &\leq v(x); \\ [-v(x)]^+ &\leq [-v^{\varepsilon}(x)]^+. \end{aligned}$$

ii) Let $\varepsilon \geq 0$. If the defining system of $B(\cdot)$ is r-regular for each $x \in \widetilde{S}$, $t \in B^{\varepsilon}(x)$ and G is uniformly Lipschitz continuous on $\widetilde{S}$, then

$$\begin{aligned} v^{\varepsilon}(x) &\geq v(x) - r(\varepsilon) L_G; \\ [-v(x)]^+ + r(\varepsilon) L_G &\geq [-v^{\varepsilon}(x)]^+. \end{aligned}$$

on $\widetilde{S}$.

Proof. i) is trivial!!

ii)

$$\begin{aligned} v^{\varepsilon}(x) &= \min_{t\in B^{\varepsilon}(x)} G(x,t) \\ &\geq \min_{t\in B(x)+r(\varepsilon)\mathsf{B}} G(x,t) \qquad \text{(using Lem. 4.5.6(ii))} \\ &\geq \min_{t\in B(x)} G(x,t) - r(\varepsilon)L_G \qquad \text{(using the Lip. property of G)} \\ &= v(x) - r(\varepsilon)L_G. \end{aligned}$$

The rest follows from Lem. 4.5.7.

□

The convergence of the values of the function $q_n(x)$ needs to take the grid density Δ_n into account. Accordingly, we have

Lemma 4.5.9.

i) For each $\varepsilon \geq 0$

$$\begin{aligned} v^{\varepsilon}(x) &\leq v_n^{\varepsilon}(x); \\ \left[-v_n^{\varepsilon}(x)\right]^+ &\leq \left[-v^{\varepsilon}(x)\right]^+. \end{aligned}$$

ii) If h is Lipschitz and G is Lipschitz, both uniformly on $\widetilde{S}$, then for $\delta \geq \Delta_n L_h$

$$\begin{aligned} v(x) &\geq v_n^{\delta}(x) - \Delta_n L_G; \\ \left[-v_n^{\delta}(x)\right]^+ + \Delta_n L_G &\geq \left[-v(x)\right]^+. \end{aligned}$$

Proof. i) is obvious.

ii) Since $B(x) + \triangle_n \mathsf{B} \subset B(x) + \frac{\delta}{L_h}\mathsf{B} \subset B^{\delta}(x)$ (Lem. 4.5.6(i)) and using (4.5.8), we have, for each $t \in B(x)$, some $t' \in B_n^{\delta}(x)$ such that $\|t - t'\| \leq \Delta_n$ (actually, $B(x) \cap T_n \subset B_n^{\delta}(x)$). Since $B(\cdot)$ is compact valued,

$$\min_{t\in B(x)} G(x,t) = G(x,\bar{t}), \text{ for some } \bar{t} \in B(x).$$

Hence, there is some $t_i \in B_n^\delta(x)$ such that $\|\bar{t} - t_i\| \leq \Delta_n$. Thus, using the Lip. continuity of G, we obtain

$$\min_{t \in B(x)} G(x,t) =\geq G(x,t_i) - L_G\|\bar{t} - t_i\| \geq G(x,t_i) - L_G\Delta_n$$
$$\geq \min_{t_i \in B_n^\delta(x)} G(x,t_i) - L_G\Delta_n.$$

This implies

$$v(x) \geq v_n^\delta(x) - L_G\triangle_n.$$

The rest follows from Lem. 4.5.7.

□

In Lem. 4.5.9(ii), the δ controls the density of the discretization for which convergence could be guaranteed. A simple observation also reveals the importance of the Lipschitz continuity assumptions on the functions G and h. With this remark, we are now ready to meet the statement of convergence.

Theorem 4.5.10. *Let $\lambda > 0$. Then*

i) if h Lipschitz and G is Lipschitz uniformly on X and $\delta_n = \Delta_n L_h$, then for each $n \in \mathbb{N}$

$$\inf_{x \in X} q_n^{\delta_n}(x) + \lambda\,\Delta_n L_G \geq \inf_{x \in X} q(x) \geq \inf_{x \in X} q_n^0(x)\,;$$

ii) if, additionally, the defining system of $B(\cdot)$ is r-regular and G is Lipschitz continuous uniformly on X, then for each $n \in \mathbb{N}$, and $\varepsilon \geq 0$

$$\inf_{x \in X} q(x) \geq \inf_{x \in X} q_n^\varepsilon(x) - \lambda\, r(\varepsilon) L_G.$$

Proof. We make use of the results in Lem. 4.5.8 and Lem. 4.5.9.

i) If h Lipschitz, G is Lipschitz continuous uniformly on X and $\delta_n = \Delta_n L_h$, we have

$$\left[-v_n^{\delta_n}(x)\right]^+ + \Delta_n L_G \geq \left[-v(x)\right]^+ \geq \left[-v_n^0(x)\right]^+.$$

Thus

$$\inf_{x \in X} q_n^{\delta_n}(x) + \lambda\,\Delta_n L_G \geq \inf_{x \in X} q(x) \geq \inf_{x \in X} q_n^0(x)\,.$$

ii) If the defining system of $B(\cdot)$ is $r-$regular and G is Lipschitzian, then for each $\varepsilon \geq 0$ and Δ_n

$$[-v(x)]^+ + r(\varepsilon) L_G \geq [-v^\varepsilon(x)]^+ \geq [-v_n^\varepsilon(x)]^+ . \qquad \text{(Lem. 4.5.8)(ii)}$$

Hence

$$\inf_{x\in X} q(x) \geq \inf_{x\in X} q_n^\varepsilon(x) - \lambda\, r(\varepsilon)\, L_G.$$

□

Corollary 4.5.11. *Let h be Lipschitz and G be Lipschitz uniformly on X. With the absence of regularity we get, by using $\varepsilon = \delta_n = \Delta_n L_h$ and both penalty functions $q_n^{\delta_n}$ and q_n^0, an upper and lower estimation of the optimal value of (GSIP), i.e.*

$$\liminf_{n\to\infty}\left[\inf_{x\in X} q_n^{\delta_n}(x)\right] \geq \inf_{x\in X} q(x) \geq \limsup_{n\to\infty}\left[\inf_{x\in X} q_n^0(x)\right].$$

When, additionally, the r-regularity holds, we obtain

$$\lim_{n\to\infty}\left[\inf_{x\in X} q_n^{\delta_n}(x)\right] = \inf_{x\in X} q(x) \geq \limsup_{n\to\infty}\left[\inf_{x\in X} q_n^0(x)\right].$$

Thm. 4.5.10 and Cor. 4.5.11 provide an upper and a lower estimates for the value of $\inf_{x\in X} q(x)$. Using Prop. 4.5.1, these are actually estimates for the optimal value of the (GSIP).

Remark 4.5.6. Observe that the r-regularity assumption in Lem. 4.5.6(ii), Lem. 4.5.8(ii) and Thm. 4.5.10(ii) leads to the lower semi-continuity of $B(\cdot)$ (cf. Chap. 1 Cor. 1.2.5). But with out this assumption we are still able to estimate $\inf_{x\in X} f(x)$ as is shown in Cor. 4.5.11.

4.6 Conclusion

It is now obvious that, the lack of continuity in the index map $B(\cdot)$ of a (GSIP) might entail difficult structures on the feasible set. Thus, the approaches presented makes use of piecewise continuity properties of $B(\cdot)$ to characterize the feasible set. But, it needs to be stressed that, the known strong conditions for nice behavior of a (GSIP) are expected to hold on each component of a suitable robust partition. However, for the application of the (IGOM) such a partition need not be explicitly known. Only the existence of it is important. Actually, the following two basic facts of robust measurable partitions have been used:

- robustness on the parts implies robustness on whole; and
- measurability on the parts implies measurability on the whole.

Consequently, the proposed approaches could be used to tackle (GSIP)s with "ill-behavior". At the same time "well-behaved" (GSIP)s are part and parcel of the discussion. However, problems of nicer structures (like convexity, etc) better be numerically treated by methods that exploit their structures.

The discussion considers (GSIP)s only with inequality constraints. It then still remains to find out the validity of the proposed method in the presence of equality constraints. Nevertheless, one might imbed equality constraints into the objective as penalties.

Moreover, the approach through an auxiliary (PSIP) could lead to certain fixed point and proximal point like algorithm. Therefore, this might be taken as a future research direction.

Chapter 5

Some Remarks on Numerical Experiments

The numerical experiments made here are done through BARLO (Hichert[26]). But, the details of the procedures of BARLO are left to [26].

As explained earlier, the second approach has been used for the computation of some representative examples of (GSIP). From the theoretical investigations we know that, the upper robustness of the index set-valued map, the continuity of the semi-infinite constraint function and the upper semi-continuity of the objective function of (GSIP) imply the upper robustness of the objective function of ($Q_{\lambda 0}$) (second penalty approach). In the examples considered, we have additionally the upper semi-continuity of the index set valued map. Thus, the penalty objective is upper semi-continuous. Hence, the essential infimum is equal to the infimum (Thm. 4.4.5) and we can principally use a stochastic based method for the determination of the infimum of the objective over the compact set X. If the penalty parameter is large enough, then the penalty problem is exact and its solution or minimizing sequence is at least a generalized minimizing sequence of (GSIP) (see Prop. 4.5.1). Nevertheless, the numerical computation using the global optimization method is somehow numerically expensive. Thus, we can only expect to get coarse approximations of the solution or generalized solution or some elements of a generalized minimizing sequence of (GSIP) in an acceptable CPU-time for problems of lower dimensions. This, in fact, has been the main aim of the proposed approach. After determining such coarse guesses it is necessary to apply local methods for further refinements. But this has not been the subject of the investigation. Concerning local methods for (GSIP) or (SIP) one can refer e.g. the latest book of Stein [75] and several papers of Stein, Still and Tichatschke for details.

The GO procedure BARLO uses **B**ranch **A**nd **R**educed **L**evel set **O**ptimization method, which was developed by Zheng and modified by Hichert and Hoffmann. Roughly spoken, this method uses a uniformly distributed sampling of boxes, decides wether a point does or does not belong to some level set and determines successful points of the new level, which yield a new mean value. Then after, the size of the search box is reduced by using some stochastically based ideas. If the size of the reduced box is too small or the cost to find a point in a level set is too large, then the box is divided in two boxes. Mainly, the following five criteria have been used as stopping rules. The first satisfied rule stops the procedure.

1. maximal number of function evaluation is exceeded;
2. maximal number of iterations is exceeded;
3. maximal length of the box sides relative to the first box X is smaller than Δb;
4. the variance of the function values in the current level; set relative to the variance of the first level set is smaller than some $V(f)$;
5. the function values in the level set are all close enough to the current level (Δf).

Whenever the parameters $\Delta b, V(F)$ and Δt are too small, then, on the one hand, the costs become extremely high and, on the other hand, the results might not be more exact, since there is a danger of cutting off a possible global optimum. For the purpose at hand, we use $\Delta b, V(f) \in [10^{-5}, 10^{-3}]$ and $\Delta f \in [10^{-3}, 10^{-1}]$. The larger the sample size the more sure is the global search. Since we need an average of 3 - 12 function evaluations, to find a point in a new level set we use a sampling of 15 - 40 new points. Points of earlier searches which also belong to the new level set are used again. After finding an initial guess with the procedure BARLO, we repeat it several times - each time by decreasing the box-neighborhood around the current guess. The diameter of the box-neighborhoods depend on the variance of the previous step is about 1/(10*iter) of the diameter of X. This strategy has moderate costs, is more successful than high sampling and prevents the possible cut off of global optima.

Now, let us begin with the first example which shows the major properties of the approach.

Example 5.0.1.

$$\begin{aligned} f(x) &= (x_1 + 0.5 - 1/(1+\sqrt{5}))^2 + (x_2 - 2.5)^2 \\ \mathcal{M} &= \left\{x \in [-5,5]^2 \mid x_1 - x_2 - t \geq 0, \forall t \in B(x)\right\} \\ B(x) &= \{t \in [-5,5] \mid t \leq x_1 - 2, 2t \geq -x_2 - 3\} \end{aligned}$$

The first two figures (Figure 1) show the nature of the admissible set $\mathcal{M}$ as determined by G and $B(\cdot)$. The broken line characterizes the boundary of $\mathcal{M}$, where G plays no role and is mainly determined by $B(x)$. Only on the feasible boundary (the solid line) we have the interplay of G and B. We also observe that $x = (-0.5, 2)$ is an inward corner point, $\mathcal{M}$ is connected, robust, but not closed.

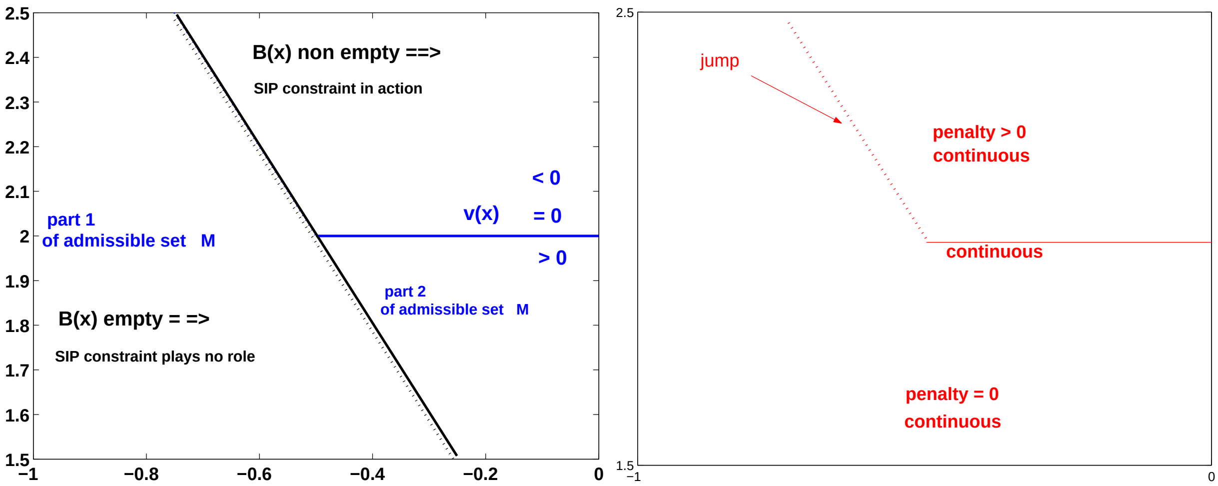


Figure 1: Construction of admissible set, discontinuity of the penalty term.

It is easy to show that $B(\cdot)$ is u.s.c. and lower robust, piecewise continuous (with 2 components); and G is linear in x and t. In Figure 2., the figure on the left shows the behavior of the penalty term. It is discontinuous along the boundary of $\mathcal{M}$ which does not belong to $\mathcal{M}$. Such discontinuities accelerate the procedure BARLO to find admissible points and principally do not create a problem.

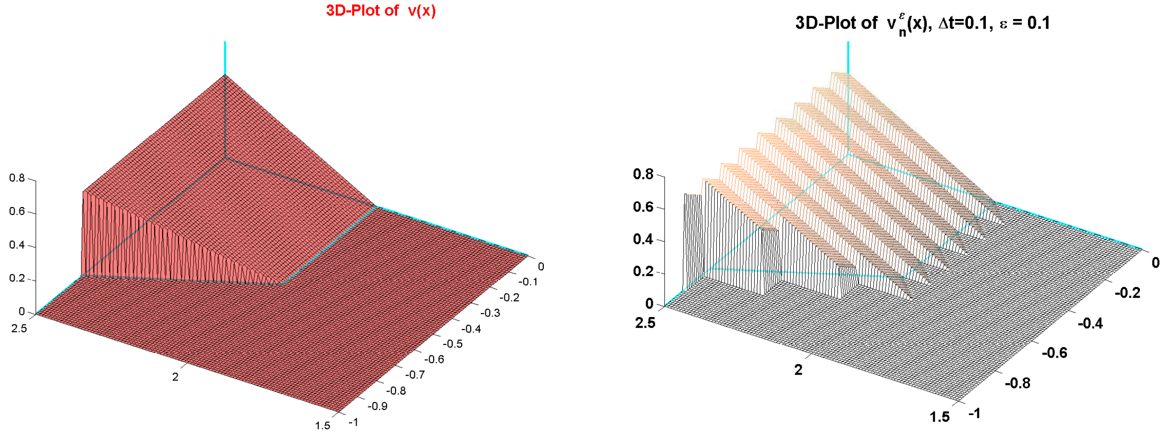


Figure 2: Discontinuity of the penalty term v, discontinuity of the penalty term v_n^ε.

In contrast to (SIP), some difficulties arise in (GSIP) by the descretization of $B(x)$. It also causes additional discontinuities for the max-penalty function (Figure 2., right). This plays no significant role in BARLO, but the boundaries of the corresponding approximations $\mathcal{M}_n$ and $\mathcal{M}_n^\varepsilon$ of $\mathcal{M}$ are 'rugged' as is illustrated in Figure 3. Despite this fact, BARLO finds the global optima with respect to this worse structured boundary.

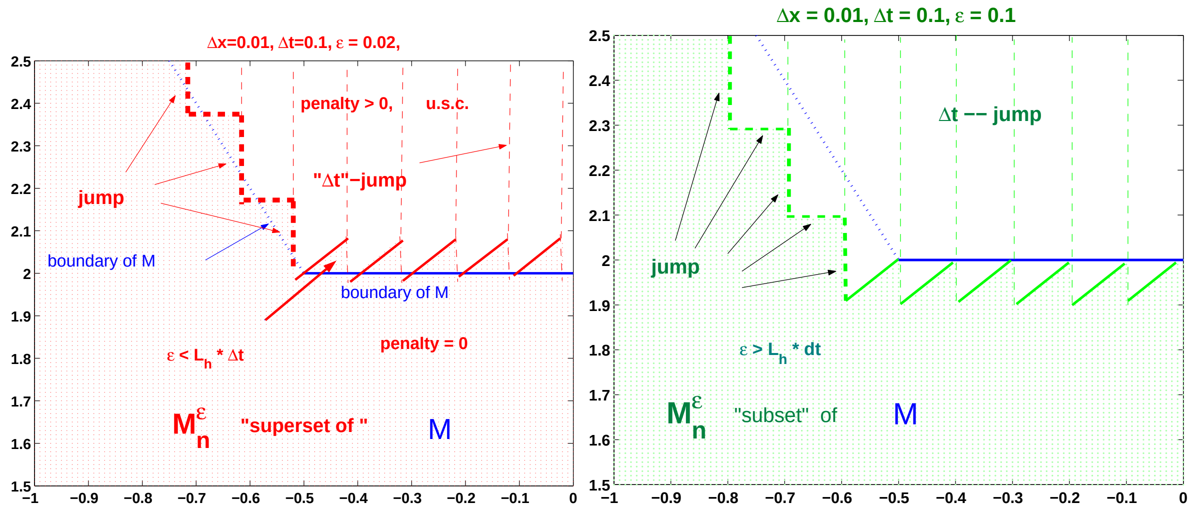


Figure 3: Structure of $\mathcal{M}_n^\varepsilon$ for $\varepsilon < L_h\Delta t$ (left), $\varepsilon \geq L_h\Delta t$ (right).

Figure 4. shows the results for different sampling and accuracy without further re-starts of BARLO. The comments that have been made earlier are really observable in this example. The exact optimal solutions are indicated by a blue star.

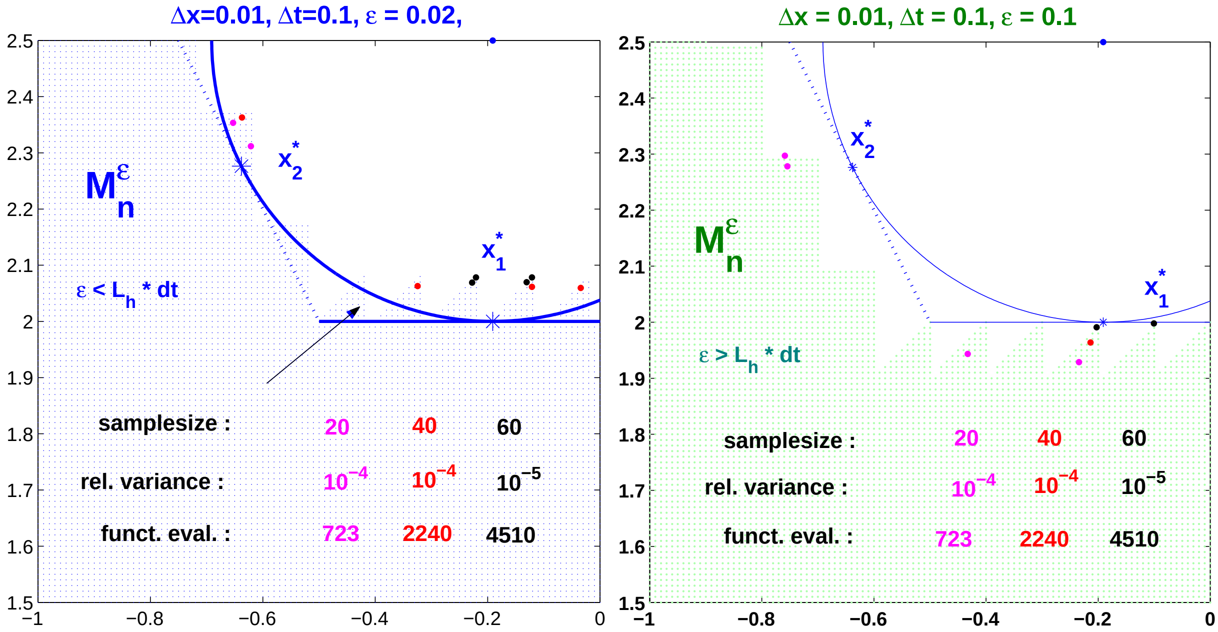
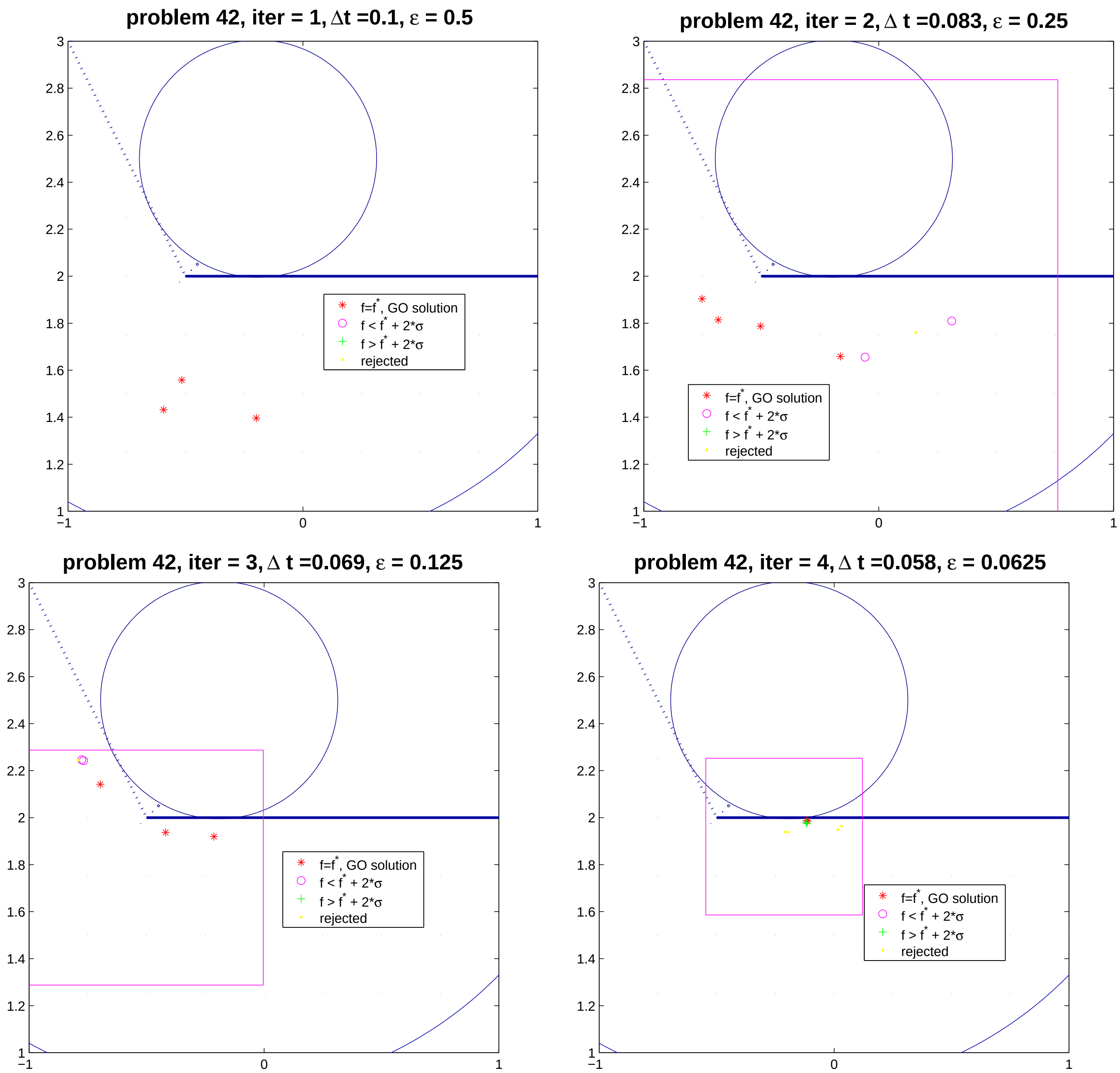


Figure 4: A coarse solution for several parameters, cf. corresponding colors.

Next, we re-start BARLO several times by using the small sample size of 15 and the low accuracy of $\Delta b = 10^{-1}, V(F) = 10^{-3}$ and $\Delta f = 10^{-2}$. At the same time we use $\varepsilon = 0.5$ and $\Delta t = 0.1$ and reduce ε by half, but not smaller than Δt. Divide Δt by 1.2 at each BARLO re-start (Figure 5.). This variant has about the same computational cost, exactness and seems to work more stable. In fact, after three or four BARLO re-starts, one could proceed with a local method.


problem 42, iter = 1, Δt =0.1, ε = 0.5
problem 42, iter = 2, Δ t =0.083, ε = 0.25
problem 42, iter = 3, Δ t =0.069, ε = 0.125
problem 42, iter = 4, Δ t =0.058, ε = 0.0625
f=f*, GO solution
f < f* + 2*σ
f > f* + 2*σ
rejected

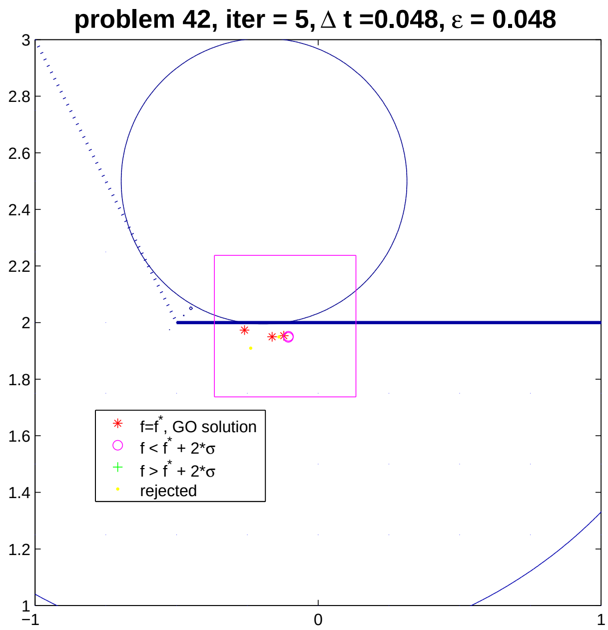


Figure 5: Function evaluations: 267, 459, 800, 1255, 404; CPU-time:1.05, 2.04, 4.23, 7.75, 3.13.

Now we consider a strange example. Here there is only one feasible point of $\mathcal{M}$ for which $B(x)$ is non-empty and the semi-infinite constraint is satisfied. All other feasible points are those at which $B(x)$ is empty.

Example 5.0.2. *The point $x = (-2, 0)$ is an inward corner of cl$\mathcal{M}$ which is connected. The feasible set $\mathcal{M}$ is the union of two open sets and the point 0, which belongs to the interior of cl$\mathcal{M}$. Moreover, $\mathcal{M}$ is robust and $B(\cdot)$ is u.s.c.*

$$\begin{aligned} \mathcal{M} &= \left\{x \in [-5,5]^2 \mid -x_1^2 + x_2 t \geq 0, \forall t \in B(x)\right\} \\ B(x) &= \left\{t \in [-5,5] \mid x_2 \leq -t^2, t^2 x_1 - 2x_2 \leq 0\right\} \end{aligned}$$

The following are 3D-representations of the penalty term v_n^ε (of v) when $\varepsilon = 0.5,\ \Delta t = 0.02$ and $\varepsilon = 0.0,\ \Delta t = 0.5$ (Figure 6.).

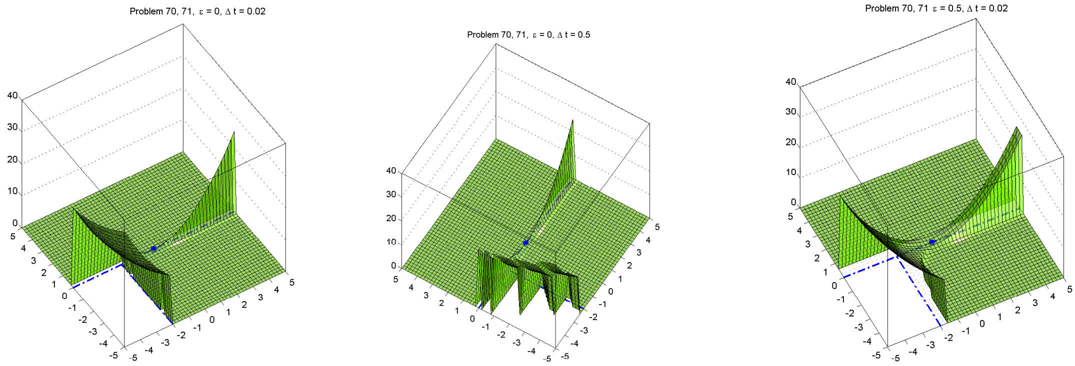
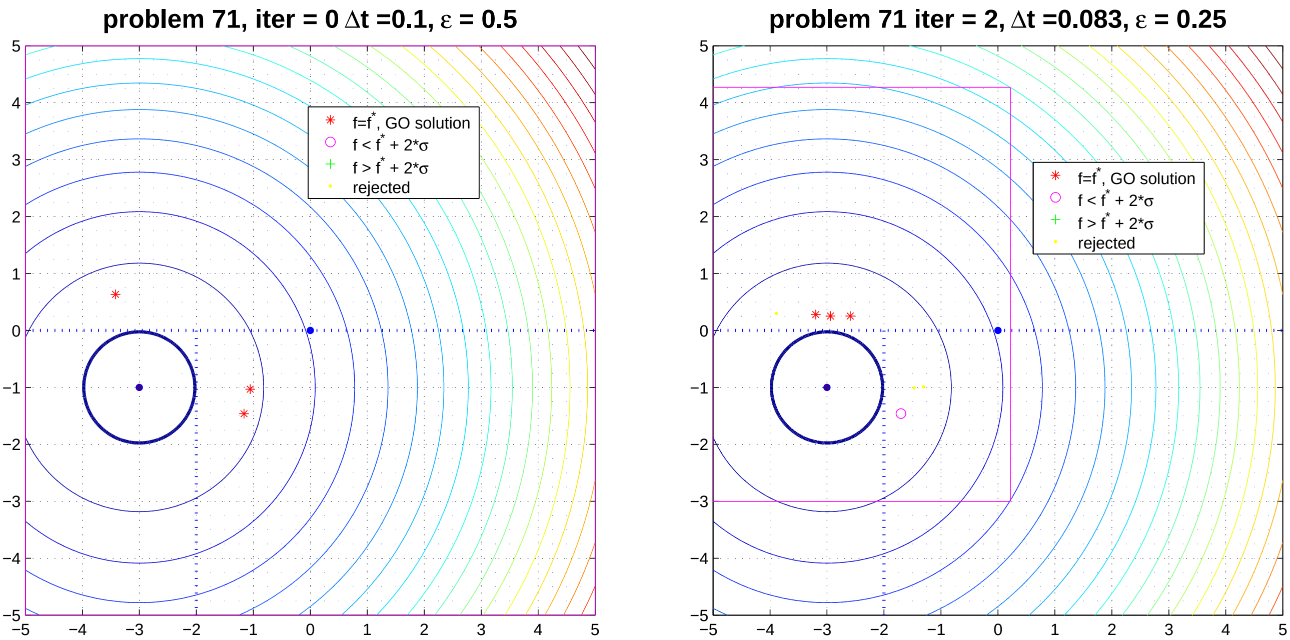


Figure 6: v (left), $\mathcal{M}_n^\varepsilon \underset{\approx}{\supset} \mathcal{M}$ (middle), $\mathcal{M}_n^\varepsilon \underset{\approx}{\subset} \mathcal{M}$ (right).

Under the same set of parameters as above, we get for the quadratic objective function $f(x) = (x_1 + 3)^2 + (x_2 + 1)^2$ the following (Figure 7).



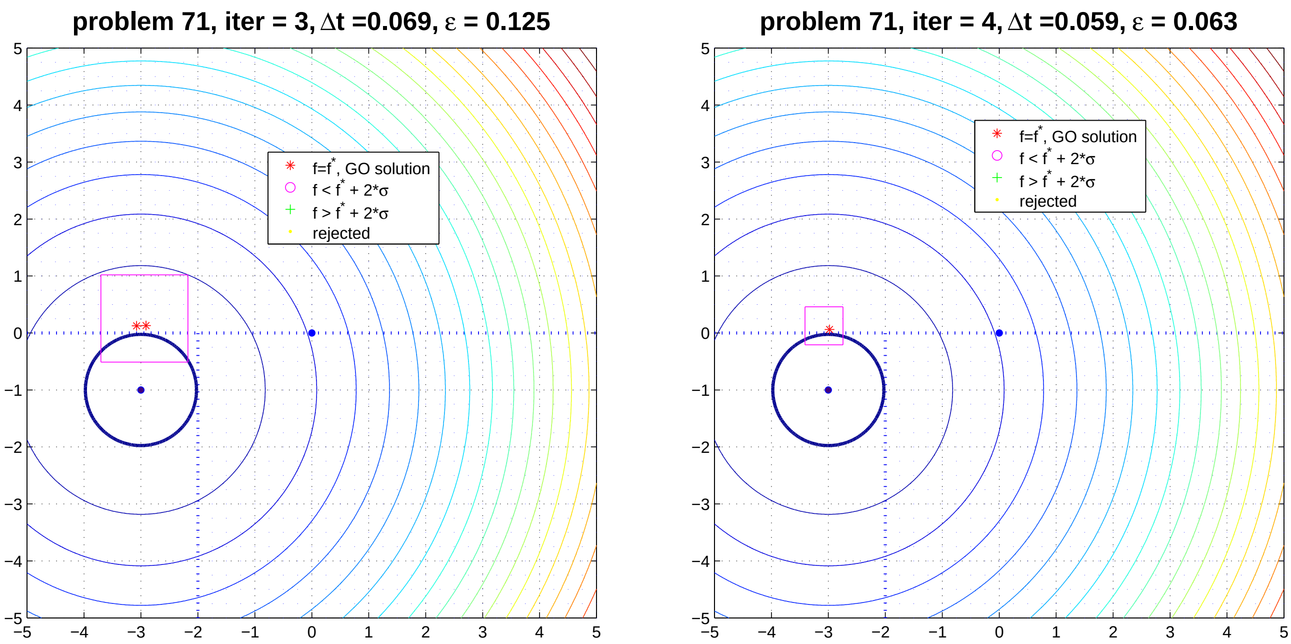


Figure 7: Function evaluations: 193, 989, 370, 314; CPU-time: 1.7, 4.62, 1.98, 1.92.

Example 5.0.3. *Now we consider a design centering problem with non convex non simply connected container. Here, we look for the largest ball* $B(x)$ *which is contained in the set*

$$\mathcal{G} = \{t \mid G_j(t) \geq 0, j \in J = \{1, 2, ..., 7\}\}$$

where

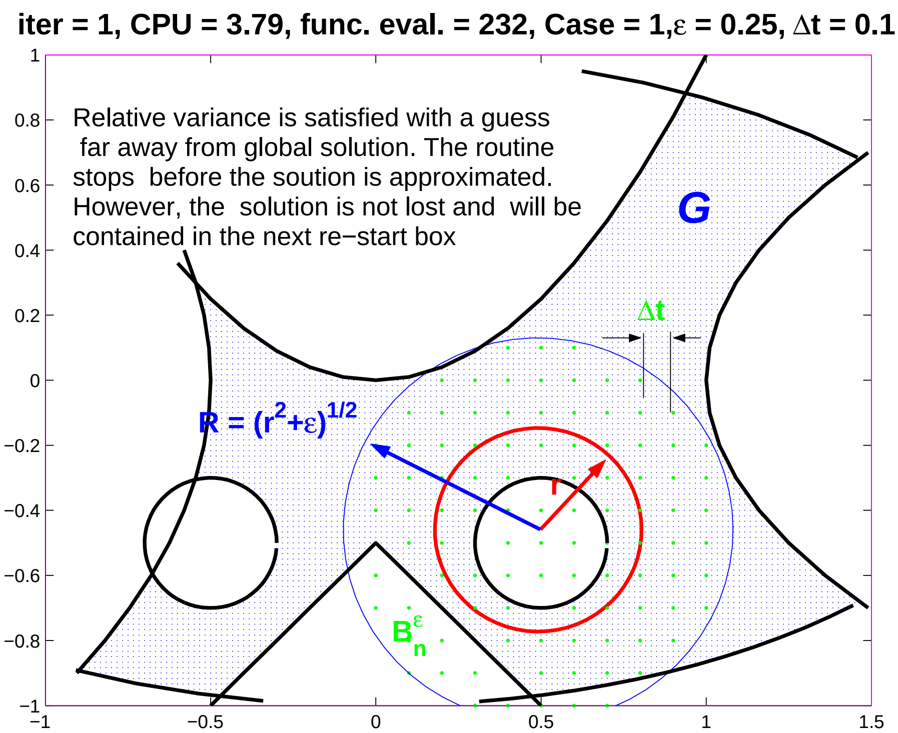
$$\begin{aligned} G_1(t) &= t_1^2 - t_2 \\ G_2(t) &= -t_1 + t_2^2 + 1 \\ G_3(t) &= 1 - \frac{t_1^2}{4} - t_2^2 \\ G_4(t) &= 2t_1 + t_2^2 + 1 \\ G_5(t) &= (t_1 + 0.5)^2 + (t_2 + 0.5)^2 - 0.04 \\ G_6(t) &= (t_1 - 0.5)^2 + (t_2 + 0.5)^2 - 0.04 \\ G_7(t) &= |t_1| + t_2 + 0.5. \end{aligned}$$

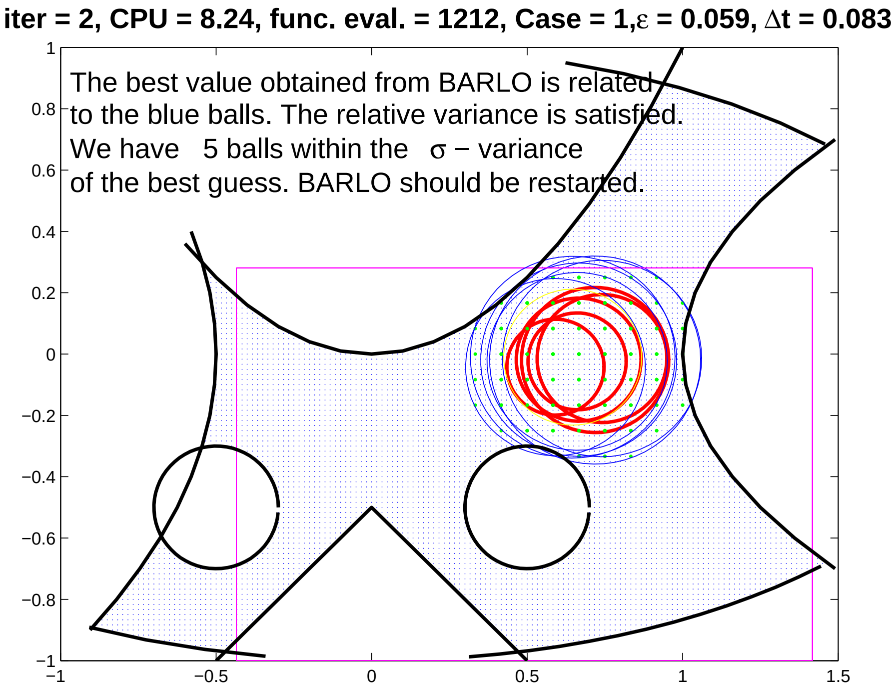
This problem is restated as a GSIP as follows

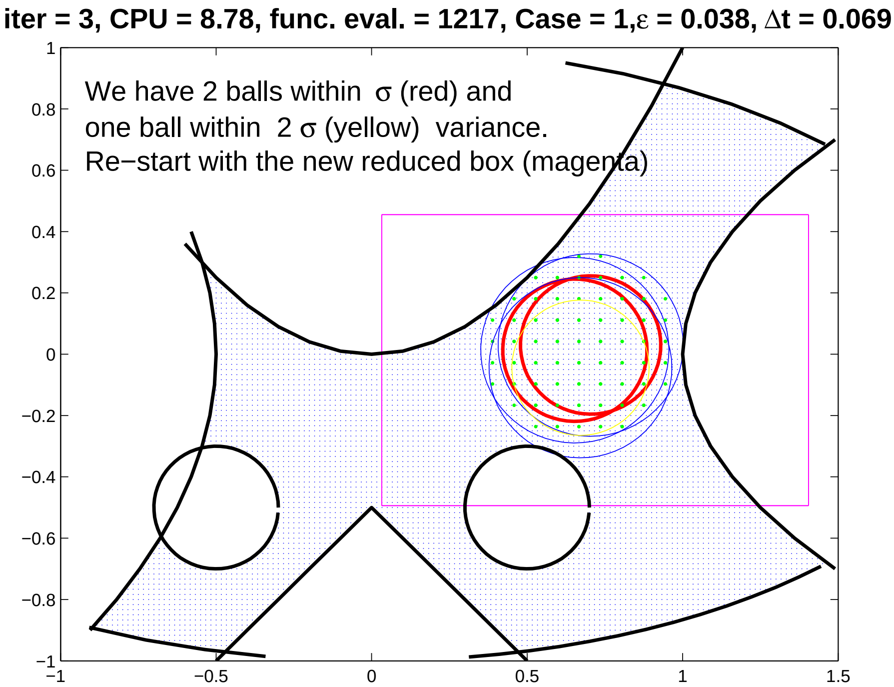
$$\begin{aligned} f(x) &= -\pi x_3^2 \to \min \\ x &\in \mathcal{M} = \{x \in [-1.5, 1.5] \times [-1, 1] \times [\Delta t, 1] \mid G_j(t) \geq 0, \forall t \in B(x), j \in J\} \\ B(x) &= \left\{t \in [-1.5, 1.5] \times [-1, 1] \mid (t_1 - x_1)^2 + (t_2 - x_2)^2 \leq x_3^2\right\} \end{aligned}$$

Since B is continuous we have no strange behavior. But, the GO - approach does not identify wether a problem has strange behavior or not. It works the same way in both cases. Naturally, for problems with l.s.c index map $B(\cdot)$, after having a coarse global guess for the solution, a local refinement can be done by using super linearly convergent methods along with local reduction.

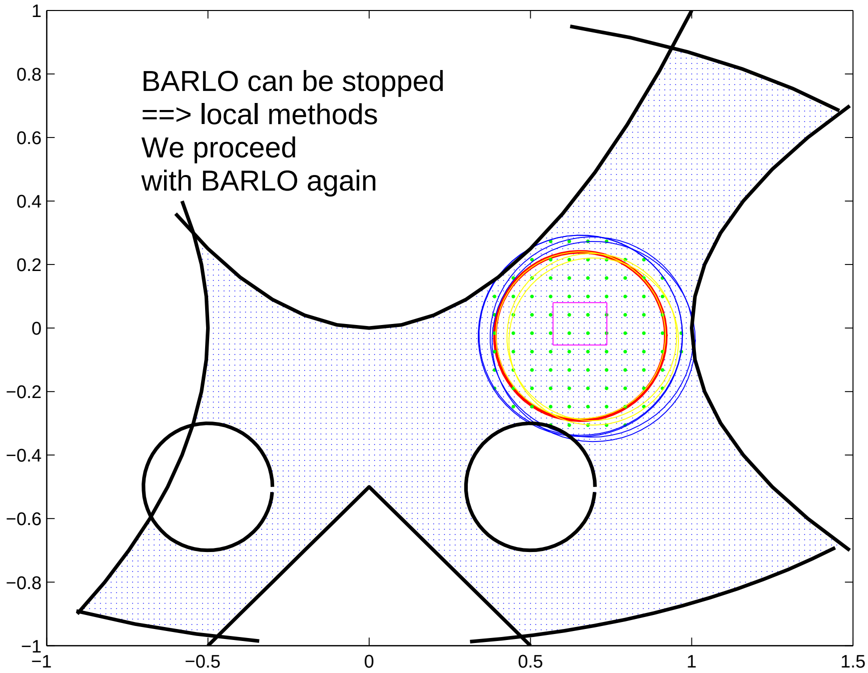
First (see Figure 8.), during each call of the BARLO routine, we use a fixed descretization of T and extract, for each evaluation of the penalty objective, those points of the T-grid which belong to $B^\varepsilon(x)$. If we use the 'find' routine of Matlab with respect to the whole grid in T, then this takes so much time (this was, actually, used for the examples with one dimensional T, see above). Instead, we choose the nearest point of the T-grid with respect to the midpoint of the ball and generate a grid-box whose convex hull contains the ball $B_n^\varepsilon(x)$. At the same time, Δt and ε are chosen so that $B(x)$ is contained in the container $\mathcal{G}$. We also reduce Δt by the factor 1.2. In order to avoid too early cut off of global optima, the size of re-start boxes should not be less than a lower bound which is about 1/25 of the size of X. If a guess is near the boundary of a box, then BARLO is re-started with a new box having the found guess as a center.



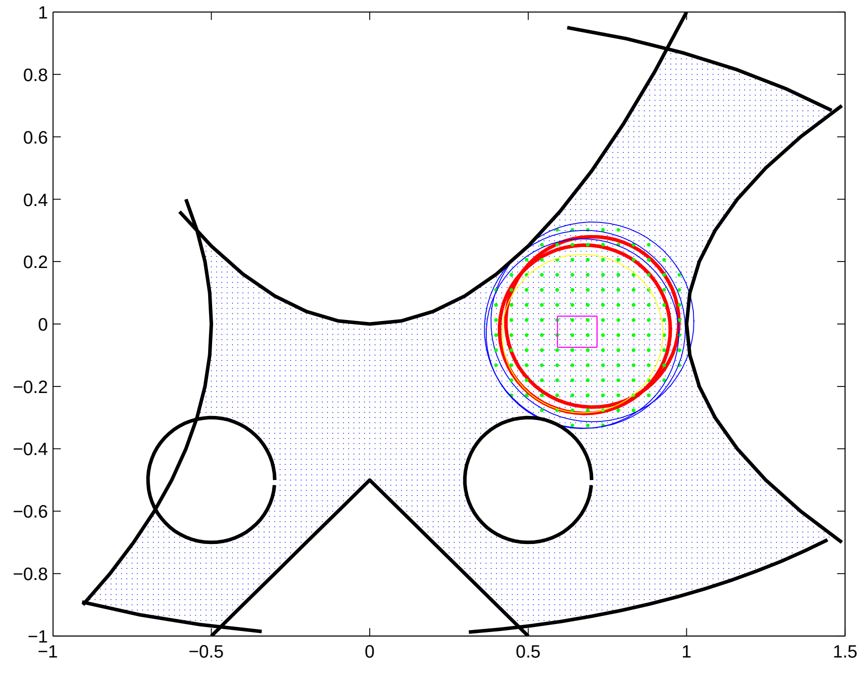

iter = 2, CPU = 8.24, func. eval. = 1212, Case = 1,ε = 0.059, Δt = 0.083
The best value obtained from BARLO is related
to the blue balls. The relative variance is satisfied.
We have 5 balls within the σ − variance
of the best guess. BARLO should be restarted.
1
0.8
0.6
0.4
0.2
0
−0.2
−0.4
−0.6
−0.8
−1
−1
−0.5
0
0.5
1
1.5


iter = 3, CPU = 8.78, func. eval. = 1217, Case = 1,ε = 0.038, Δt = 0.069
We have 2 balls within σ (red) and
one ball within 2 σ (yellow) variance.
Re−start with the new reduced box (magenta)
1
0.8
0.6
0.4
0.2
0
−0.2
−0.4
−0.6
−0.8
−1
−1
−0.5
0
0.5
1
1.5

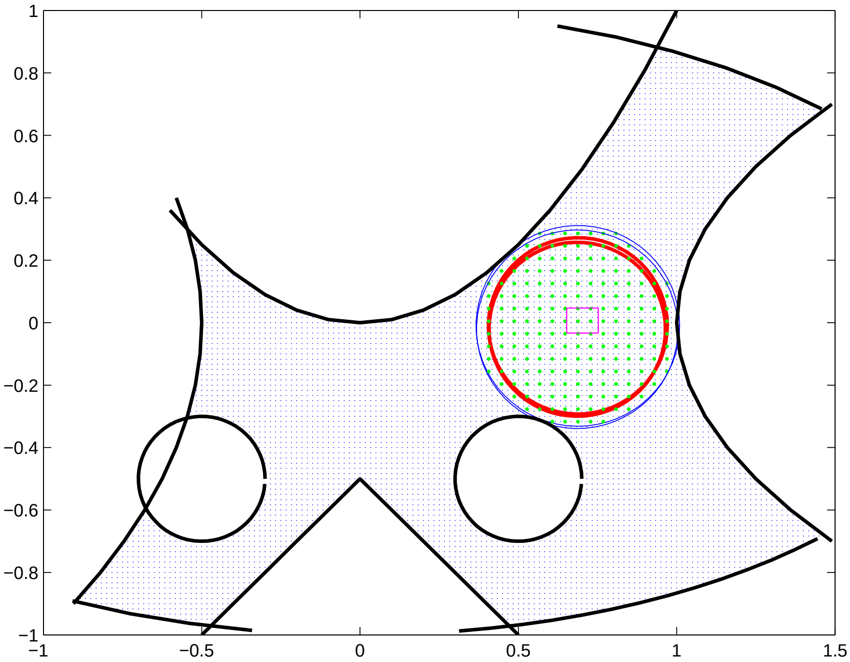
iter = 4, CPU = 6.7, func. eval. = 708, Case = 1,ε = 0.030, Δt = 0.058



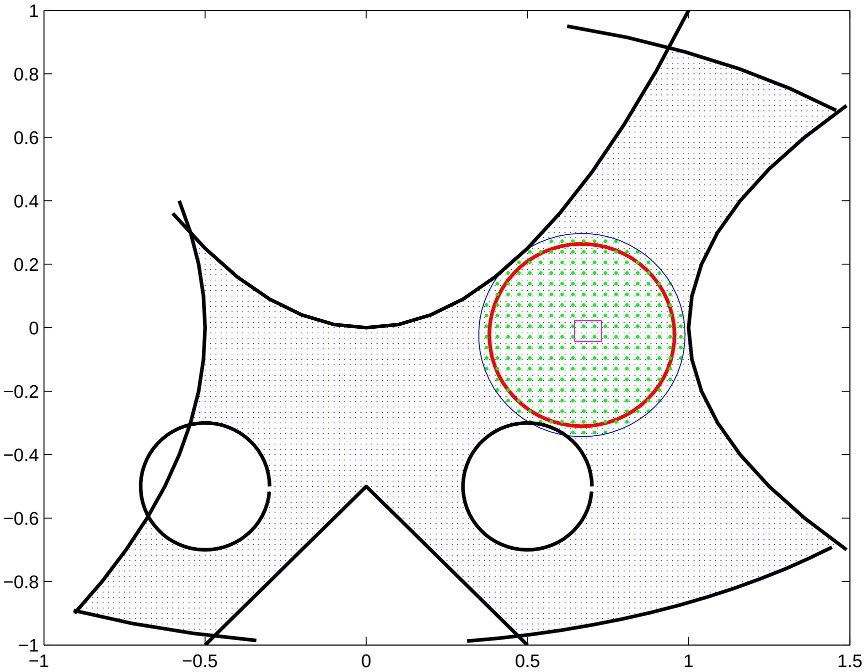
iter = 5, CPU = 3.02, func. eval. = 192, Case = 1,ε = 0.028, Δt = 0.048



iter = 6, CPU = 6.21, func. eval. = 305, Case = 1,ε = 0.024, Δt = 0.040



iter = 7, CPU = 3.02, func. eval. = 60, Case = 1,ε = 0.020, Δt = 0.033



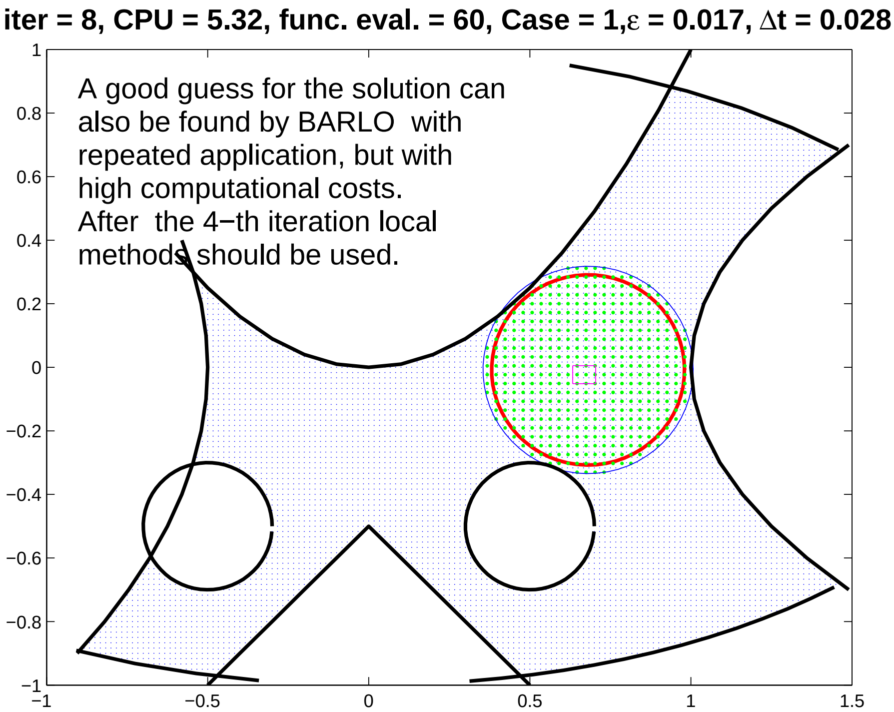


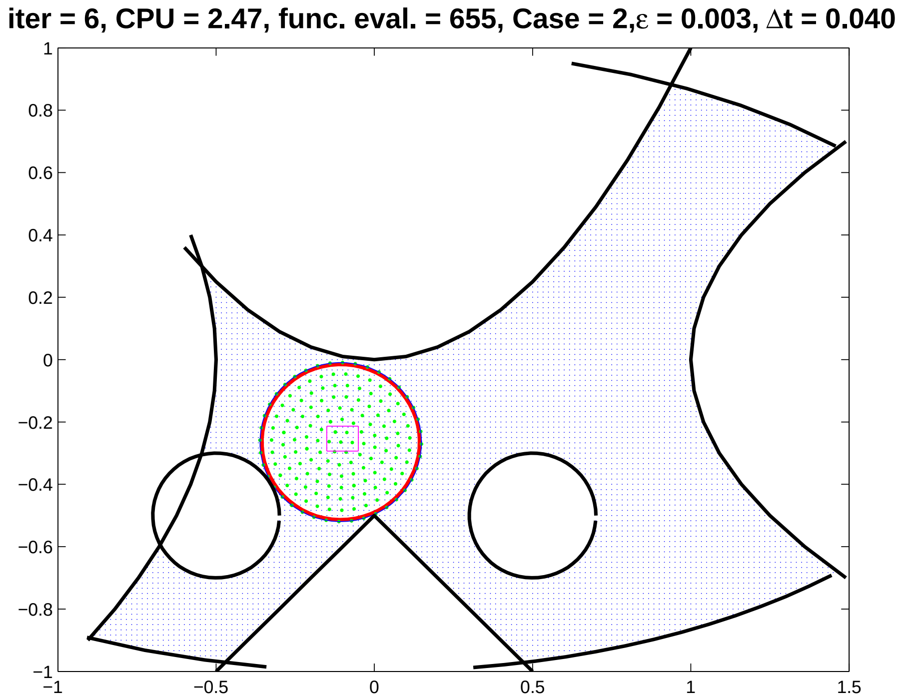
Figure 8: Design centering with one ball in the container $\mathcal{G}$.

Second (Figure 9.), we use a direct descretization of $B^{\varepsilon}(x)$ by taking polar coordinates and by generating an approximate grid of Δt along boundary circles of the ball. The starting angle of the polar coordinates is randomly chosen at each function evaluation. This procedure is faster than the first one. Thus we can increase the sample size (30) and sharpen the stopping criteria ($V(f) = 10^{-4}, \Delta b = 10^{-2}, \Delta f = 10^{-3}$). This shows, that the effectiveness of the GO approach depends enormously on how fast we determine the points of the grid which belong to $B^{\varepsilon}(x)$.

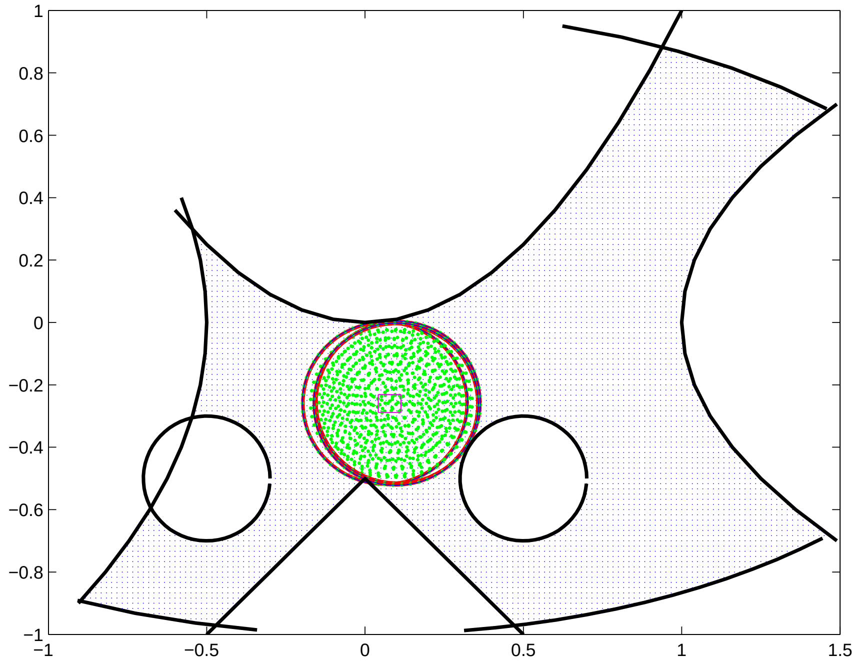
Remark 5.0.1. The effectiveness of the evaluation of the penalty term depends on the possibility of fast determination of a grid with density Δt which is surely contained in $B^{\varepsilon}_n(x)$. A simple algorithm has been used to test all grid points of T_n wether they belong to $B^{\varepsilon}_n(x)$ (along with the "find" routine of Matlab) and to compute the penalty value by using the "maximum" function of Matlab.

In higher dimensions, such a task requires a lot of computational time for the determination of one single penalty value. It is believed that, tools of interval mathematics could create more effective box inclusions of $B_n^\varepsilon(x)$. In the case of the design centering problem, we can directly discretize the set $B_n^\varepsilon(x)$. In this case the routine BARLO is about 10 to 20 times faster, when T is of dimension 2. However, for the approach to function properly the discretization procedure must be able to recognize arising and vanishing components of $B_n^\varepsilon(x)$, which is true for the above trivial but expensive approach.

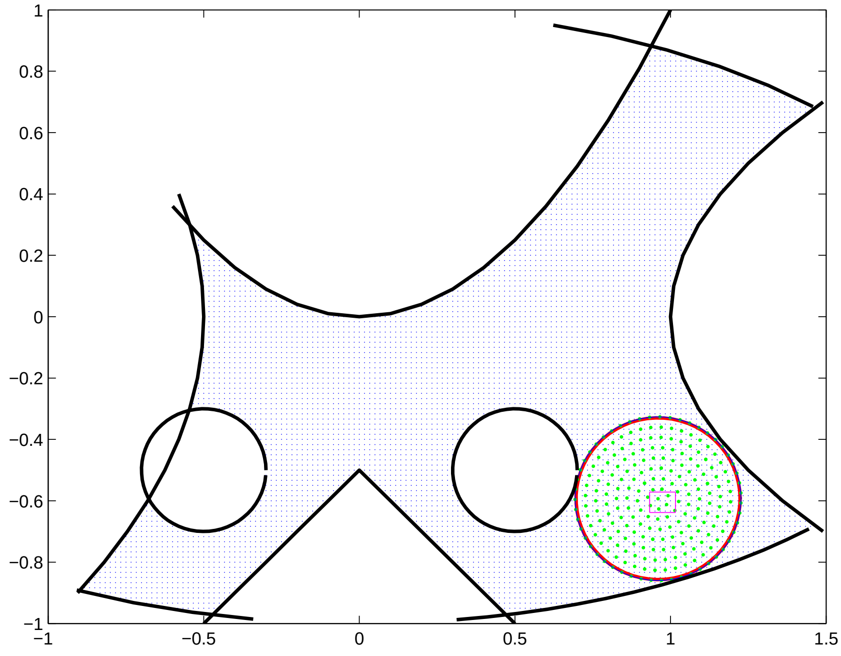
A repeated call of BARLO with respect to the box X under different seeds of the random generator is usefull to ensure the final result. Different seeds can produce different local minima, because the box strategy with a repeated call of BARLO prefers the neighborhood of the last guess. The figures next show the design centering using the second descretization method.



iter = 8, CPU = 4.5, func. eval. = 272, Case = 2,ε = 0.0008, Δt = 0.028



iter = 7, CPU = 4.39, func. eval. = 1088, Case = 2,ε = 0.002, Δt = 0.033



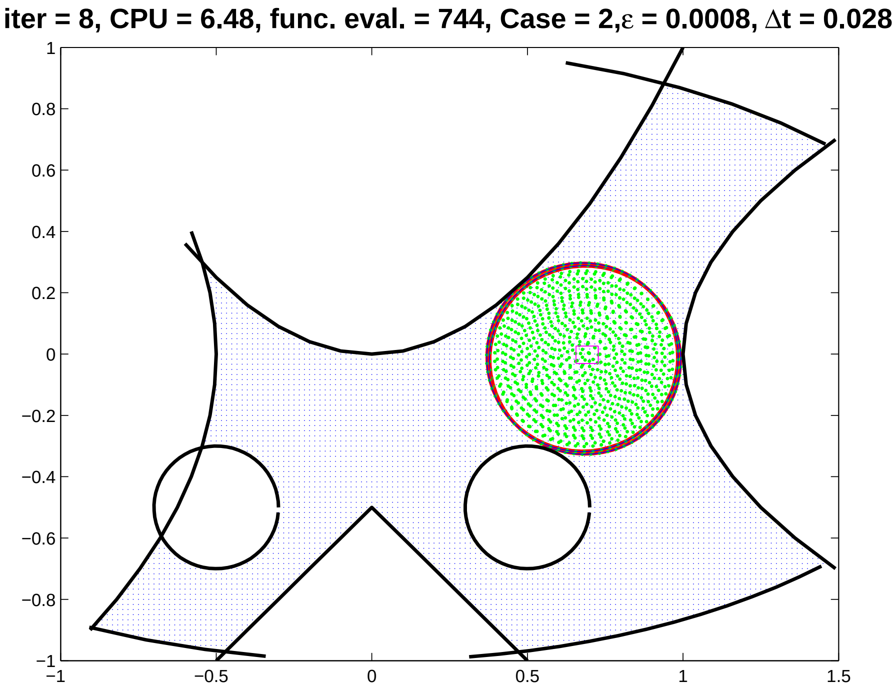


Figure 9: Design centering with one ball in the container $\mathcal{G}$, several local solutions.

Bibliography

[1] G. Abebe and A. Hoffmann. A conceptual method for solving generalized semi-infinite programming problems via global optimization by exact discontinuous penalization. *Euro. J. of OR*, 157:3–15, 2004.

[2] W. Achtziger. A numerical approach to optimization w.r.t. variable-dependent constraint-indices. In *Operations Research Proceedings 1998, Zürich, August 31 - September 3, 1998, P. Kall and H.-J. Lüthi (eds)*, pages 83–92. Springer-Verlag, 1999.

[3] D. Aliprantis and C. Border. *Infinite Dimensional Analysis: A Hitchhicker's Guide, Studies in Economic Theory V. 4*. Springer Verlag, 1999.

[4] J.-P. Aubin and A. Cellina. *Differential Inclusions*. Springer Verlag, 1984.

[5] J.-P. Aubin and H. Frankowska. *Set-Valued Analysis*. Birkhaüser, 1990.

[6] A. Auslender. Stability in mathematical programming with non-differentiable data. *SIAM J. Control and Optimization*, 22:239–254, 1984.

[7] B. Bank, J. Guddat, D. Klatte, B. Kummer, and K. Tammer. *Non-linear Parametric Optimization*. Akademie-Verlag, Berlin, 1982.

[8] C. Berge. *Topological Spaces*. Oliver & Boyd, Edinburg, London, 1963.

[9] J. F. Bonnans and A. Shapiro. Optimization problems with perturbations: A guided tour. *SIAM Rev.*, 40(2):228–264, 1998.

[10] J. F. Bonnans and A. Shapiro. *Perturbation Analysis of Optimization Problems*. Springer, New York, 2000.

[11] J. Cánovas, A. López, E.-M. Ortega, and J. Parraa. Upper semicontinuity of closed-convex-valued multifunctions. *Math. Meth. of OR*, 57:409–425, 2003.

[12] E. W. Cheney. *Introduction to Approximaiton Theory.* McGraw-Hill, New York, 1966.

[13] S. Chew and Q. Zheng. *Integral Global Optimization.* Springer-Verlag, 1988.

[14] F. H. Clarke. *Optimization and Nonsmooth Analysis.* John Wiley & Sons, 1983.

[15] S. Dempe. *Foundations of Bilevel Programming.* Kluwer Academi Publishers, 2002.

[16] A. V. Fiacco and J. Kyparisis. Convexity and concavity properties of the optimal value function in parametric non-linear programming. *JOTA*, 48(1), 1986.

[17] J. Gauvin. A necessary and sufficient regularity condition to have bounded multipliers. *Math. Prog.*, 12:136–138, 1977.

[18] J. Gauvin and F. Debeau. Differential properties of the marginal function in mathematical programming. *Math. Prog. Stud.*, 19:101–119, 1985.

[19] M. A. Goberna and M. A. Lopez. *Semi-infinite Programming - Recent Advances.* Kluwer Academic Publishers, 2001.

[20] T. J. Graettinger and B. H. Krogh. The acceleration radius: a global performance measure for robotic manuplators. *IEEE J. of Robotics and Automation*, 4:60–69, 1988.

[21] F. Guerra and J.-J. Rückmann. A Karush-Kuhn-Tucker condition for a generalized semi-infinite optimization problem under an extended Kuhn-Tucker constraint qualification. 2001. Preprint, University of Las Americas, School of Science, Dept. Physics and Math.

[22] R. Henrion and D. Klatte. Metric regualrity of the feasible set map in semi-infinite optimization. *Appl. Math. Optim.*, 30:103–109, 1994.

[23] R. Hettich and G. Still. Second order optimality conditions for generalized semi-infinite programming problems. *Optimization*, 34:195–211, 1995.

[24] R. Hettich and G. Still. Semi-infinite programming: second order optimality conditions. In *Encyclopedia of Optimization, Volume 5, C. Floudas & P. M. Pardalos(eds.)*, pages 117–122. Kluwer Academic Publishers, 2001.

[25] R. Hettich and P. Zencke. *Numerische Methoden der Approximation und Semi-infiniten Optimierung.* Teubner, Stuttgart, 1982.

[26] J. Hichert. *Methoden zur Bestimmung des wesentlichen Supremums mit Anwendung in der globalen Optimierung.* PhD thesis, TU-Ilmenau, Fakultät für Mathematik und Naturwissenschaften, 1999.

[27] J. Hichert, A. Hoffmann, and H. X. Phu. Convergence speed of an integral method for computing essential supremum. In *In Developments in Global Optimization.* Kluwer Academic Publishers, 1997.

[28] J. Hichert, A. Hoffmann, H. X. Phu, and R. Reinhardt. A primal-dual integral method in global optimization. *Disscussiones Mathematicae: Differential Inclusions, Control and Optimization*, 20(2):257–170, 2000.

[29] A. Hoffmann and R. Reinhard. On reverse chebychev approximation problems. Faculty of Mathematics and Natural Sciences, Technical Univeristy of Ilmenau, Preprint, No. M 08/94, 1994.

[30] W. W. Hogan. Directional derivative for the extremal-value funcitons with applications to the completely convex case. *Oper. Res.*, 21:188–209, 1973.

[31] W. W. Hogan. Point-to-set maps in mathematical programming. *SIAM Review*, 15(3):591–603, 1973.

[32] S. Hu and N. S. Papageorgiou. *Handbook of Multivalued Analysis: Volume I.* Kluwer Academic Publishers, 1997.

[33] H. T. Jongen, J.-J. Rückmann, and O. Stein. Disjunctive optimization: critical point theory. *JOTA*, 9(2):321–336, 1997.

[34] H. T. Jongen, J.-J. Rückmann, and O. Stein. Generalized semi-infinite optimization: a first order optimality condition and examples. *Math. Prog.*, 83:145–158, 1998.

[35] A. Kaplan and R. Tichatschke. On a class of terminal variational problems. In *Parametric Programming and Related Topics IV, J. Guddat, et al (eds.)*, pages 185 – 199. Peter Lang Verlag, Frankfurt a.M., 1997.

[36] A. Kaplan and R. Tichatschke. On the numerical treatment of a class of semi-infinite terminal problems. *Optimization*, 41:1–36, 1997.

[37] M. Kisielewicz. *Differential Inclusions and Optimal Control*. Polish Scientific Publishers, Kluwer Academic Publishers, 1991.

[38] D. Klatte. Stability of stationary points in semi-infinite optimization via the reduction approach. In *Advances in Optimization, Lecture Note in Economics and Mathematical Systems, V. 382, W. Otteli & D. Pallaschke (eds.)*, pages 155–170. 1991.

[39] D. Klatte and R. Henrion. Regularity and stability in non-linear semi-infinite optimization. In *Semi-infinite Programming, R. Reemtsen and J.-J. Rückmann (eds.)*, pages 69–102. Kluwer Academic Press, 1998.

[40] W. Krabs. On time minimal heating or cooling of a ball. In *International Series Numer. Math.*, volume 81, pages 121–131. Birkhäuser Verlag, Basel, 1987.

[41] J. Kyparisis. On uniqueness of Kuhn-Tucker multipliers in non-linear programming. *Math. Prog.*, 32:242–246, 1985.

[42] E. Levitin. On local convex majorizing approximation of generalized semi-infinite programming problems. Forshungsbericht, Nr. 96-35, Department of Mathematics, University of Trier, 1996.

[43] E. Levitin. Reduction of generalized semi-infinite programming problems to semi-infinite or piece-wise smooth programming problems. Forschungsbericht, Nr. 01-08, Department of Mathematics, University of Trier, 2001.

[44] E. Levitin. On differential properties of the optimal value of parametric problems of mathematical programming. *Dokl. Akad. Nauk SSSR*, 224:1354–1358, 1975.

[45] E. Levitin. *Perturbation Theory in Mathematical Programming and its Applications*. John Wiley & Sons Ltd, 1994.

[46] E. Levitin. Differential properties of parametric minimum funcitons and extremal mappings. In *Parametric Optimization and Related Topics IV, Gudat et al. (eds.)*, pages 213–226. Peter Lang Verlag, 1997.

[47] E. Levitin and R. Tichatschke. A branch-and-bound approach for solving a class of generalized semi-infinite programming problems. *J. Glob. Opt.*, 13:299–315, 1998.

[48] H. X. Phu and A. Hoffmann. Essential supremum and supremum of summable functions. *Numer. Funct. Anal. and Optim.*, 17(1&2):167–180, 1996.

[49] S. Pickl and G.-W. Weber. An algorithmic approach by linear programming problems in generalized semi-infinite optimization. *J. Comput. Technol.*, 5:62, 2000.

[50] E. Polak and Y. Y. Wardi. A study of minimizing sequences. *SIAM J. Ctrl. and Optim.*, 22(4):599–609, 1984.

[51] H. R. and K. Kortanek. Semi-infinite programming: theory, methods and applications. *SIAM Review*, 35:380–429, 1993.

[52] R. Reemtsen and J.-J. Rückmann. *Semi-infinite Programming*. Kluwer Academic Publishers, 1998.

[53] S. M. Robinson. Regularity and stability for convex multivalued functions. *Math. of OR*, 1(2):130–143, 1976.

[54] S. M. Robinson. Some continuity properties of polyhedral multifunctions. *Math. Prog. Study*, 14:206–214, 1991.

[55] R. T. Rockafellar. *Convex Analysis*. Princton University Press, Princeton, 1970.

[56] R. T. Rockafellar. Integral functionals, normal integrals and measurable selections. In *Nonlinear Operators and Calculus of Variations, Lecture Notes in Mathematics, V. 543, A. Dold and B. Eckmann (eds.)*, pages 157–207. Springer Verlag, 1976.

[57] R. T. Rockafellar and R. J.-B. Wets. *Varaitional Analysis*. Springer Verlag, 1998.

[58] H. L. Royden. *Real Analysis*. Macmillan Publ. Comp., 3rd. edition, 1988.

[59] J.-J. Rückmann and A. Shapiro. Second order optimality conditions in generalized semi-infinite programming. to appear.

[60] J.-J. Rückmann and A. Shapiro. First order optimality conditions in generalized semi-infinite programming. *JOTA*, 101(3):677–691, 1999.

[61] J.-J. Rückmann and O. Stein. On convex lower level problems in generalized semi-infinite optimization. In *Semi-Infinite Programming - Recent Advances, M. A. Goberna, M.A. Lopez (eds.)*, pages 121–134. Kluwer, Dordrecht, 2001.

[62] J.-J. Rückmann and O. Stein. On linear and liearized generalized semi-infinite optimization problems. *Annals of Operations Research*, 101:191–208, 2001.

[63] A. Shapiro. Second-order derivatives of extremal-value functions and optimality conditions for semi- infinite programs. *Math. of OR*, 10(2):207–219, 1985.

[64] A. Shapiro. Perturbation theory of nonlinear programs when the set of optimal solutions is not a singleton. *Appl. Math. Optim*, 18:215–229, 1988.

[65] A. Shapiro. Directional differentiability of the optimal value funciton in convex semi-infinite programming. *Math. Prog.*, 70:149–157, 1995.

[66] S. Shi, Q. Zheng, and D. Zhuang. Set valued robust mappings and approximatable mappings. *J. Math. Anal. Appl.*, 183:706–726, 1994.

[67] S. Shi, Q. Zheng, and D. Zhuang. Discontinuous robust mappings are approximatable. *American Math. Soc., Trans.*, 347(12):4943–4957, 1995.

[68] S. Shi, Q. Zheng, and D. Zhuang. On existence of robust minimizers. In *The state of the art in global optimization, C. A. Fouldas and P. M. Pardalos (eds.)*, pages 47–56. Kluwer Academic Publishers, 1996.

[69] K. Shimizu, I. Y., and J. Bard. *Nondifferentiable and Two-Level Mathematical Programming*. Kluwer Academic Publishers, Boston, 1997.

[70] O. Stein. On level sets of marginal functions. *Optimization*, 48:43–67, 2000.

[71] O. Stein. The reduction ansatz in the absence of lower semi-continuity. In *Parametric Optimization and Related Topics Vol. V, J. Guddat, R. Hirabayashi and H. Th. Jongen (eds.)*, pages 165–178. Peter Lang Verlag, 2000.

[72] O. Stein. The feasible set in generalized semi-infinite optimization. In *Approximation, Optimization and Mathematical Economics, M. Lasdon (ed.)*, pages 309–327. Physica, Heidelberg, 2001.

[73] O. Stein. First order optimality conditions for degenerate index sets in generalized semi-infinite optimization. *Math. of OR*, 6(3):565–582, 2001.

[74] O. Stein. Bi-level strategies in semi-infinite programming, post-doctoral thesis, 2002.

[75] O. Stein. *Bi-level Strategies in Semi-infinite Programming*. Kluwer Academic Publishers, 2003.

[76] O. Stein and G. Still. On optimality conditions for generalized semi-infinite programming. *JOTA*, 104(2):443–458, 2000.

[77] O. Stein and G. Still. On generalized semi-infinite and bi-level optimization. *European Journal of OR*, 142:444–462, 2002.

[78] O. Stein and G. Still. Solving semi-infinite optimization problems with interior point techniques. *SIAM J. Ctrl. and Optim.*, 42:769–788, 2003.

[79] G. Still. Generalized semi-infinite programming: Numerical aspects. University of Twente, Faculty of Mathematical Sciences, Memorandum No. 1470, 1998.

[80] G. Still. Generalized semi-infinite programming: Theory and methods. *European Journal of OR*, 119:301–313, 1999.

[81] N. V. Thoai. Convergence and applications of a decomposition method using duality bounds for nonconvex global optimization. *JOTA*, 113(1):165–193, 2002.

[82] J.-B. Urruty and H. C. Lemaréchal. *Convex Analysis and Minimization Algorithms I*. Springer Verlag, 1993.

[83] G.-W. Weber. Optimal control theory: on the global structure and connections with optimization, part 2. Preprint, Dept. of Mathematics, Darmstadt University of Technology, 1998.

[84] G.-W. Weber. Generalized semi-infinite optimization: on some foundations. *J. Comput. Technol.*, 4(3):41–61, 1999.

[85] G.-W. Weber. Structural stability in generalized semi-infinite optimization. *Comput. Techn.*, 6(2):25–46, 2001.

[86] G.-W. Weber. On the topology of generalized semi-infinite optimization. *J. Convex Analysis*, 9(2):665–691, 2002.

[87] G.-W. Weber. *Generalized Semi-infinite Optimization and Related Topics*. Heldermann Verlag, Research and Expostions in Mathematics 29, Lemgo, eds.: K. H. Hoffmann and R. Wille, 2003.

[88] Q. Zheng. *Integral Global Optimization of Robust Discontinuous Functions*. PhD thesis, The Graduate School of Clemson University, Clemson, 1992.

Index