Abstract

ÖZÇAM, BURCU. Solving Semi-Infinite Variational Inequalities (Under the direction of Dr. Shu-Cherng Fang and Dr. Henry Lee Nuttle).

The variational inequality problem arises in numerous contexts. In this dissertation, we consider solving a semi-infinite variational inequality problem, which is a variational inequality problem defined on a domain described by infinitely many constraints. We present characterization and the solution analysis for semi-infinite variational inequalities. After introducing the solution analysis, three solution methodologies, namely a discretization-based smoothing method, an exchange method and an entropic analytic center cutting plane method are proposed. A comprehensive computational results with the comparison of the algorithms is provided.

Solving Semi-infinite Variational Inequalities

by

Burcu Özçam

A dissertation submitted to the Graduate Faculty of

North Carolina State University

in partial fulfillment of the requirements for the

Degree of Doctor of Philosophy

Industrial Engineering

2006

Raleigh

Approved by:

Dr. Shu-Cherng Fang Chair, Advisory Committee	Dr. Henry L.W. Nuttle Co-Chair, Advisory Committee
Dr. Elmor L. Peterson Committee Member	Dr. Xiuli Chao Committee Member

Dedicated to

my parents for their endless love..

Biography

Burcu Özçam was born in 1978 in Izmir, Turkey. She received her Bachelor of Science degree in Industrial Engineering from Bilkent University in 1999. She joined North Carolina State University to pursue her graduate study in industrial engineering in August 1999. In 2001, she received her M.S. degree, and subsequently continued working toward a Ph.D. degree. During her graduate studies in the Edward P. Fitts Department of Industrial Engineering, she worked as a teaching assistant for several undergraduate and graduate courses and as a research assistant for the North Carolina Furniture Manufacturing and Management Center. From May 2003 – March 2006, she has worked as an optimization specialist in the Operations Research R&D group, Analytical Solutions Department at SAS Institute, Inc., located in Cary, NC. Upon graduation, she plans to take up an assistant professor position in Izmir University of Economics.

Acknowledgements

I would like to express my appreciation to many people who shared their knowledge, wisdom, and support with me during the course of my graduate study at North Carolina State University. First of all, I am very grateful to my advisor Dr. Shu-Cherng Fang for his invaluable advice, constant encouragement and guidance. I am also grateful to Dr. Henry L. W. Nuttle, Co-chair of my advisory committee, for his continuous assistance and supervision throughout my study. Special thanks are also due to other members of the committee, including Dr. Elmor L. Peterson and Dr. Xiuli Chao, for expressing interest in and reviewing my dissertation.

I appreciate all the help and good company I have had in FANGROUP. I especially thank Hao Cheng for spending his time in having endless discussions with me on my research and for helping me in so many different ways. I would like to extend my sincere thanks to dear friends Yi Liao, Yue Dai, Shunmin Wang, Xiaoli Ling, Xiang Zhou, Yu-min Lin, Wei Zhang, Carin Lightner, Chian-Fen Ding and Aysegul Peker. The support and friendship of Girish Ramachandra will always be appreciated.

I am grateful to SAS Institute for providing financial support for the latter half of my doctoral study. I also would like to express my special thanks to Trevor Kearney, Radhika Kulkarni, Ed Hughes, and all the people in the OR group at SAS Institute, for making it a wonderful place to gain valuable industrial experience.

Finally, I would like to thank all my family members especially to my mother Şengün and brother Ali Evren Özçam for always being there and for pushing me to complete this thesis. I am grateful to my friends Anil, Yusuf, Funda and Erdogan, among others, for listening to me so many times.

Table of Contents

List of Tables **vii**

List of Figures **viii**

1 Introduction **1**

1.1 Semi-infinite Variational Inequalities 2

1.2 Organization of Dissertation 3

2 Literature Review **4**

2.1 Nonlinear Programming 5

2.1.1 Structure of the Feasible Set 5

2.1.2 Constraint Qualifications 8

2.1.3 First-order Conditions 10

2.1.4 Second-order Conditions 11

2.2 Semi-Infinite Programming 13

2.2.1 Structure of the Feasible Set 13

2.2.2 Constraint Qualifications 23

2.2.3 First-order Conditions 26

2.2.4 Second-order Conditions 28

2.3 Variational Inequalities 29

2.3.1 Equivalent Reformulations of Variational Inequalities 30

2.3.2 Constraint Qualifications 36

2.3.3 Solution Analysis 37

2.4 Semi-infinite Variational Inequality Problems 41

3 Solution Analysis for Semi-infinite Variational Inequalities **43**

3.1 Structure of the Domain 44

3.2 Equivalent Reformulations of Semi-infinite Variational Inequalities 46

3.2.1 Projection Reformulation 46
3.2.2 Mathematical Programming Formulation with Merit Functions . . . 47
3.3 Existence Results 49
3.4 Constraint Qualifications 51
3.4.1 Extended Abadie Constraint Qualification 52
3.4.2 Extended Linear Independence Constraint Qualification 53
3.4.3 Extended Mangasarian Fromowitz Constraint Qualification 54
3.4.4 Extended Kuhn-Tucker Constraint Qualification 55
3.5 Solution Analysis 55
3.6 Duality 59
3.6.1 Dual with Continuous Borel Measure 59
3.6.2 Dual with Discrete Borel Measure 60

4 Solution Methods **62**
4.1 SIVI(X, F) and Subproblems 63
4.2 Solving VI(X_i, F) 64
4.2.1 Smoothing Method 67
4.3 A Discretization-based Smoothing Method 71
4.3.1 Algorithm 1 73
4.4 Exchange Method 76
4.4.1 Algorithm 2 78
4.5 Entropic Analytic Center Cutting Plane Method 80
4.5.1 Algorithm 3 81
4.6 Numerical Results 83
4.6.1 Test Problems 83
4.6.2 Results for Discretization-based Smoothing Method 85
4.6.3 Results for Exchange Method 89
4.6.4 Results for Entropic Analytic Center Cutting Plane Method 92
4.6.5 Observations 95

5 Summary and Future Research Directions **98**
5.1 Summary of Dissertation 98
5.2 Future Research Directions 99

List of References **101**

A Computational Results **109**

List of Tables

A.1 Results for Problems (1-3) by implementing discretization-based smoothing method 110
A.2 Results for Problems (4-6) by implementing discretization-based smoothing method 111
A.3 Results for Problems (7-9) by implementing discretization-based smoothing method 112
A.4 Results for Problems (1-3) by implementing exchange method 113
A.5 Results for Problems (4-6) by implementing exchange method 114
A.6 Results for Problems (7-9) by implementing exchange method 115
A.7 Results for Problems (1-3) by implementing entropic analytic center cutting plane method 116
A.8 Results for Problems (4-6) by implementing entropic analytic center cutting plane method 117

List of Figures

2.1 Feasible set structure in finite and semi-infinite programming 14
2.2 Geometric interpretation of VI(X, F) 30

4.1 Max-plus function and smooth approximation $p_\varepsilon(x)$ with $\varepsilon = 10^{-18}$ 68
4.2 Algorithm 1: (a) Number of Newton iterations vs. the grid size N for problems 1-3, (b) Deviation from the best known solution for problems 1-3. . . . 87
4.3 Algorithm 1: (a) Number of Newton iterations vs. the grid size N for problems 4-6, (b) Deviation from best known solution for problems 4-6. 88
4.4 Algorithm 2: (a) Constraint violation vs. the number of iterations for problems 1-3, (b) Deviation from best known solution at each iteration for problems 1-3. 90
4.5 Algorithm 2: (a) Constraint violation vs. the number of iterations for problems 4-6, (b) Deviation from best known solution at each iteration for problems 4-6. 91
4.6 Algorithm 3: (a) Number of function evaluations vs. the number of iterations k for problems 1-3, (b) Deviation from best known solution at each iteration for problems 1-3. 93
4.7 Algorithm 3: (a) Number of function evaluations vs. the number of iterations k for problems 4-6, (b) Deviation from best known solution at each iteration for problems 4-6. 94
4.8 Comparison of the computational effort required by the proposed algorithms 95
4.9 Comparison of the solution accuracy among the proposed algorithms 96

Chapter 1

Introduction

Let R^n be the n-dimensional Euclidean space. Given the continuous functions $F : R^n \to R^n$ and $g(x,t) : R^n \times R^l \to R^p$, consider the following semi-infinite variational inequality problem:

$$\begin{aligned} \text{SIVI}(X,F): \qquad & \text{Find a vector } x^* \in X \text{ such that} \\ & F(x^*)^\intercal (x - x^*) \geq 0 \qquad \text{for all } x \in X. \end{aligned} \tag{1.1}$$

where

$$X = \{x \in R^n \mid \ g(x,t) \geq 0, \quad \text{for all} \ \ t \in T\}, \tag{1.2}$$

and $T \subset R^l$ is an index set that may contain infinitely many elements. Our objective is to solve (1.1) when $|T| = \infty$.

The contributions of semi-infinite variational inequalities are two-fold. First, they broaden the application area of variational inequalities. Secondly, they link semi-infinite programming to a new set of problems through a variational inequality formulation. By means of semi-infinite variational inequalities, factors involving infinitely many constraints can be introduced into equilibrium and complementary problems.

1.1 Semi-infinite Variational Inequalities

Problem (1.1) becomes an "ordinary" variational inequality problem if T has finite cardinality, say $|T| = m$. As a special case of the variational inequality problem, SIVI(X, F) arises in numerous contexts such as economics, transportation, telecommunications and supply chain management. Infinitely many equalities and inequalities make it a challenging task to find a solution to variational inequalities since general convex programming methods are unable to tackle infinitely many constraints all at one time.

Originating in [22], the theory, algorithms and applications of variational inequalities have been well studied for over four decades. A comprehensive manuscript is given in [12]. However, most of the algorithms for solving variational inequalities in the literature apply only for cases in which X has a certain geometric structure (such as a compact polyhedral set) or is defined by a finite number of equality or inequality constraints. Despite the numerous work on variational inequalities, the study of theory and algorithms for solving semi-infinite variational inequality problems is relatively recent [16].

Basically, there are two approaches for solving semi-infinite variational inequalities. The first approach involves the reduction of the infinite number of constraints that defines the domain X to a finite number. The second and more direct approach is an analytic center cutting plane method which provides an outer approximation of the domain described by infinitely many constraints.

The focus of this dissertation is to analyze the solutions of problem (1.1) and to extend both the reduction-based and cutting plane algorithms. The computational results for each algorithm are studied with a comparative analysis.

1.2 Organization of Dissertation

Throughout this dissertation, the term *nonlinear programming* refers to the class of optimization problems with a finite number of constraints, *semi-infinite programming* refers to the class of optimization problems with an infinite number of constraints and the term *ordinary* or *finite variational inequalities* refers to variational inequality problems whose domain is defined by a finite number of equalities and inequalities.

The remainder of the dissertation is organized as follows. In the next chapter, we review the existing methods for solving problem (1.1) in relation to nonlinear programming, semi-infinite programming and ordinary variational inequalities. In Chapter 3, the solution analysis for semi-infinite variational inequalities is presented. Three different algorithms corresponding to the discretization-based smoothing method, exchange method and entropic analytic center cutting plane method are proposed and the numerical results are presented in Chapter 4. Finally, Chapter 5 concludes our study with a summary and future research directions.

Chapter 2

Literature Review

The semi-infinite variational problem inherits certain structure from nonlinear programming, semi-infinite programming and finite variational inequalities. In nonlinear programming, the feasible set is determined by a finite number of constraints that satisfy certain "constraint qualifications". The constraint qualification requirements in nonlinear programming extends to semi-infinite variational inequalities. The relation of semi-infinite variational inequalities to semi-infinite programming is due to the number of constraints in defining the feasible solution set. We explore the structural similarity and differences between finite variational inequalities and semi-infinite variational inequalities.

This chapter consists of four sections. The first three sections review the related fundamental concepts in nonlinear programming, semi-infinite programming and finite variational inequalities. The fourth section reviews the existing methods and the underlying assumptions for solving semi-infinite variational inequalities.

2.1 Nonlinear Programming

Consider a nonlinear optimization problem in the form of

$$\begin{aligned} \text{minimize} \quad & f(x) \\ \text{subject to} \quad & h_1(x) \geq 0 \\ & h_2(x) \geq 0 \\ & \vdots \\ & h_m(x) \geq 0 \\ & x \in \Omega, \end{aligned} \tag{2.1}$$

where $\Omega \subseteq R^n$, $f : R^n \to R$ and $h_i : R^n \to R, i = 1, \ldots, m$, with m being a finite integer, are C^2 continuous. For brevity, we introduce a vector-valued function $h : R^n \to R^m$ with $h(x) = (h_1(x), h_2(x), \ldots, h_m(x))$ and rewrite the problem (2.1) as

$$\begin{aligned} \text{minimize} \quad & f(x) \\ \text{subject to} \quad & h(x) \geq 0 \\ & x \in \Omega. \end{aligned} \tag{2.2}$$

Any point $x \in \Omega$ satisfying the constraint $h(x) \geq 0$ is a *feasible point* and the set defined by all such points is called *feasible set* and denoted by

$$M = \{x \in R^n \mid h(x) \geq 0, x \in \Omega\}.$$

Next, we focus on the structure of the feasible set and introduce some fundamental concepts for theoretical development.

2.1.1 Structure of the Feasible Set

The structure of the feasible set plays an important role in the existence, stability and computation of an optimal solution. We are interested in the minimization (maximization) of a convex (concave) function over a convex set. When the feasible set has a convex structure

and the existence of a solution is guaranteed, the necessary and sufficient conditions for optimality are provided and stability analysis can be performed.

In consideration of a system of constraints, the first important issue is to have a *nonempty* feasible set. Given a nonempty set defined by equality or inequality constraints, we need to provide specific conditions for describing the feasible directions properly. For this purpose we need the following definitions:

Definition 2.1.1. An inequality constraint $h_i(x) \geq 0$ is said to be *active* at a feasible point $\bar{x}$, if $h_i(\bar{x}) = 0$, and *inactive* at $\bar{x}$, if $h_i(\bar{x}) > 0$.

Since the inactive constraints at a given feasible point $\bar{x}$ have no influence in the neighborhood of $\bar{x}$, the set of active constraints becomes more restrictive in defining feasible directions at $\bar{x}$. Although the information as to which constraints will be active is not known a priori, the geometry of the active constraints still plays an essential role in optimization. Let $\mathcal{I}(\bar{x})$ denote the set of indices of active constraints at a feasible point $\bar{x}$, i.e.,

$$\mathcal{I}(\bar{x}) = \{i \mid h_i(\bar{x}) = 0\}.$$

If the function $h_i(x)$, for $i = 1, \ldots, m$, is at least C^1 continuous, then the hypersurface defined by the active constraints will be smooth. This helps to deduce an important property of the active set $\mathcal{I}(\bar{x})$ in a neighborhood $\mathcal{N}(\bar{x})$ of point $\bar{x}$. For a small $\varepsilon > 0$, moving away from the point $\bar{x}$ may cause a change of the active set to $\mathcal{I}(\bar{x}+\varepsilon)$. As a result, the structure of the feasible set is characterized by feasible directions and tangent planes, which are described by active constraints.

Definition 2.1.2. Given $\bar{x} \in M \subseteq R^n$, the *cone of feasible directions* at $\bar{x}$ is defined by

$$\mathcal{D}(\bar{x}; M) = \{d \in R^n \mid d \neq 0, \bar{x} + \alpha d \in M \text{ for all } \alpha \in (0, \delta) \text{ for some } \delta > 0\}.$$

Given a function $f(x)$, the *cone of improving directions* at $\bar{x} \in M$ is given by

$$\mathcal{F}(\bar{x}; M) = \{d \in R^n \mid f(\bar{x} + \alpha d) < f(\bar{x}), \text{ for all } \alpha \in (0, \delta) \text{ for some } \delta > 0\}.$$

If $\nabla f(\bar{x})^\intercal d < 0$, then d is an improving direction and the *open half space* containing the improving directions at $\bar{x} \in M$ is defined by

$$\mathcal{F}_0(\bar{x}; M) = \{d \in R^n \mid \nabla f(\bar{x})^\intercal d < 0\}.$$

By definition, we have $\mathcal{F}_0(\bar{x}; M) \subseteq \mathcal{F}(\bar{x}; M)$. However, $\mathcal{F}_0(\bar{x}; M) = \mathcal{F}(\bar{x}; M)$ only if the function f is pseudo-convex.

Given a set of active constraints $\mathcal{I}(\bar{x})$, the *cone of interior directions* of M at $\bar{x}$ is defined by

$$G_0(\bar{x}; M) = \{d \in R^n \mid \nabla h_i(\bar{x})^\intercal d > 0 \quad \text{for each } i \in \mathcal{I}(\bar{x})\},$$

and the *polar normal cone* at $\bar{x} \in M$ is given by

$$G(\bar{x}; M) = \{d \in R^n \mid \nabla h_i(\bar{x})^\intercal d \geq 0 \quad \text{for each } i \in \mathcal{I}(\bar{x})\}.$$

Note that the relation $G_0(\bar{x}; M) \subseteq \mathcal{D}(\bar{x}; M)$ is always true. Only when $h_i, i \in \mathcal{I}(\bar{x})$ is pseudo-convex do we have $\mathcal{D}(\bar{x}; M) = G_0(\bar{x}; M)$. If $h_i, i \in \mathcal{I}(\bar{x})$ is strictly pseudo-convex then $\mathcal{D}(\bar{x}; M) = G(\bar{x}; M) \backslash \{0\}$.

When optimality is reached, geometrically we cannot find a feasible improving direction. Thus, if $\bar{x}$ is a local optimal solution then the following geometric condition holds true

$$\mathcal{F}_0(\bar{x}; M) \cap \mathcal{D}(\bar{x}; M) = \emptyset. \tag{2.3}$$

In addition to the geometric condition (2.3), more algebraic optimality conditions are defined based on the active constraint gradients.

If the geometric condition (2.3) is satisfied, that is, if there are no improving directions in the open half space $\mathcal{F}_0(\bar{x}; M)$, the relation $G_0(\bar{x}; M) \subseteq \mathcal{D}(\bar{x}; M)$ implies the following

$$\mathcal{F}_0(\bar{x}; M) \cap G_0(\bar{x}; M) = \emptyset. \tag{2.4}$$

Condition (2.4) is known as the Fritz John optimality condition [40]. Geometrically, it means that at a given point, the objective function gradient ∇f and the constraint gradients ∇h_i have to be on different hyperplanes, separated by a tangent plane.

Note that the Fritz John optimality condition $\mathcal{F}_0(\bar{x}; M) \cap G_0(\bar{x}; M) = \emptyset$ can be satisfied by any feasible solution $\bar{x}$ at which $G_0(\bar{x}; M) = \emptyset$. However, such a point $\bar{x}$ can be nonoptimal since the Fritz John condition (2.4) is satisfied whenever M has no interior, or active constraint gradients at $\bar{x}$ cancel out each other. It can be shown that condition $\mathcal{F}_0(\bar{x}; M) \cap G_0(\bar{x}; M) = \emptyset$ is not sufficient for optimality even for linear problems.

Therefore, rather than $\mathcal{F}_0(\bar{x}; M) \cap \mathcal{D}(\bar{x}; M) = \emptyset$, one should seek $G_0(\bar{x}; M) \neq \emptyset$ as a sufficient condition for $\bar{x}$ to be local minimum. By definition, the condition $G_0(\bar{x}; M) \neq \emptyset$ is closely related to the following definition of the tangent cone.

Definition 2.1.3. Let M be a nonempty set in R^n and $\bar{x} \in M$. The *tangent cone* at $\bar{x} \in M$, denoted by $\mathcal{T}(\bar{x}; M)$, is the set of all directions d with $d = \lim_{k\to\infty} \alpha_k(x_k - \bar{x})$, where $\alpha_k > 0, x_k \in M$ for each k, and $x_k \to \bar{x}$.

For a single constraint $h_1(x)$, the tangent cone is the set of all vectors perpendicular to the surface gradient at a particular point $\bar{x}$, i.e., $\{y \mid \nabla h_1(\bar{x})^\intercal y = 0\}$. However, if there are multiple constraints, the definition of a tangent cone requires certain assumptions to hold at $\bar{x}$. These assumptions (sometimes called regularity assumptions) coincide with the constraint qualifications explored in the next section.

2.1.2 Constraint Qualifications

For problem (2.1), a necessary condition for $\bar{x}$ to be optimal is having a tangent cone not containing any improving direction, i.e., $\mathcal{F}_0(\bar{x}; M) \cap \mathcal{T}(\bar{x}; M) = \emptyset$. However, as is well known, this condition by itself is not sufficient for optimality since it is also true when $\mathcal{F}_0(\bar{x}; M) = \emptyset$.

Therefore, a more general condition involving the tangent cone is the following

$$\mathcal{T}(\bar{x}; M) = G(\bar{x}; M). \tag{2.5}$$

Only when condition (2.5) holds is $\mathcal{F}_0(\bar{x}; M) \cap \mathcal{T}(\bar{x}; M) = \emptyset$ sufficient to obtain first-order (optimality) conditions.

There exist several definitions of constraint qualifications based on Definition 2.1.2 including (2.5), see [40]. We present them here for completeness.

Slater's Constraint Qualification

Given that each h_i for $i \notin \mathcal{I}(\bar{x})$ is continuous at $\bar{x}$, there exists $x' \neq \bar{x}$ in M such that $h_i(x') > 0$ for all $i \in \mathcal{I}(\bar{x})$. In other words, there exists an interior feasible solution for the problem. Slater's constraint qualification is satisfied at point $\bar{x}$, if and only if $G_0(\bar{x}; M) \neq \emptyset$.

Linear Independence Constraint Qualification

Linear independence constraint qualification (LICQ) is said to hold at $\bar{x}$ if the gradient vectors $\nabla h_i(\bar{x})$ of all active constraints $i \in \mathcal{I}(\bar{x})$ are linearly independent. With LICQ being satisfied at $\bar{x}$, constraint gradients cannot cancel out each other. Therefore, $G_0(\bar{x}; M) \neq \emptyset$.

Cottle's Constraint Qualification

Let $cl\ X$ denote the *closure* of set X. Given that each h_i for $i \notin \mathcal{I}(\bar{x})$ is continuous at $\bar{x}$, Cottle's constraint qualification (CCQ) holds if

$$cl\ G_0(\bar{x}; M) = G(\bar{x}; M).$$

Mangasarian-Fromowitz Constraint Qualification

This constraint qualification is also known as MFCQ and it requires the existence of some interior directions described by the active constraint gradients, i.e., $G_0(\bar{x}; M) \neq \emptyset$. By

definition, both Slater's condition and LICQ imply MFCQ.

Kuhn-Tucker's Constraint Qualification

Let $A(\bar{x}; M)$ be the *cone of attainable directions*, i.e., $d \neq 0$ belongs to $A(\bar{x}; M)$ if there exists a feasible arc starting from $\bar{x}$ that is tangential to $d \in R^n$. Given the closure $cl\ A(\bar{x}; M)$, Kuhn-Tucker's constraint qualification (KTCQ) is satisfied if $cl\ A(\bar{x}; M) = G(\bar{x}; M)$.

Abadie's Constraint Qualification

As the weakest among all, Abadie's constraint qualification requires the tangent cone to be defined by the gradient of the active constraints in such a way that feasible directions attained will be descent directions, i.e., $\mathcal{T}(\bar{x}; M) = G(\bar{x}; M)$.

The relations among these constraint qualifications is the following:

$$cl\ G_0(\bar{x}; M) \subseteq cl\ \mathcal{D}(\bar{x}; M) \subseteq cl\ A(\bar{x}; M) \subseteq \mathcal{T}(\bar{x}; M) \subseteq G(\bar{x}; M).$$

2.1.3 First-order Conditions

Under constraint qualifications, the first-order optimality conditions for nonlinear programming problems are derived by Karush, and by Kuhn and Tucker [40] as follows.

First-order necessary conditions

Given a nonempty open set $\Omega \in R^n$ and functions $f, h_i : R^n \to R, f, h_i \in C^1$, for $i = 1, \ldots, m$, consider problem (2.1). At a feasible solution x^*, let $\mathcal{I}(x^*)$ be the set of active constraints satisfying one of the constraint qualifications above. If x^* is a local minimum

for problem (2.1), then there exist scalars $\mu_i \in R, i = 1, \ldots, m$ such that

$$\nabla f(x^*) - \sum_{i=1}^{m} \mu_i \nabla h_i(x^*) = 0, \tag{2.6}$$

$$\mu_i h_i(x^*) = 0, \qquad i = 1, \ldots, m \tag{2.7}$$

$$\mu_i \geq 0, h_i(x^*) \geq 0, \qquad i = 1, \ldots .m. \tag{2.8}$$

Any point x^* with corresponding multiplier μ^* that satisfies conditions (2.6)-(2.8) is called a *KKT-point* or *critical solution*. An optimal solution is a KKT-point but not every KKT-point is optimal. In order for a KKT-point to be optimal we need additional conditions to be satisfied.

First-order sufficient conditions

Let x^* be a critical solution, let M be the nonempty feasible set and let $\mathcal{N}(x^*)$ be some neighborhood of x^* in M. If f is pseudo-convex over $\mathcal{N}(x^*) \cap M$, and $h_i \in C^1, i \in \mathcal{I}(x^*)$, are quasi-convex over $\mathcal{N}(x^*) \cap M$, then x^* is a local minimum for problem (2.1). Moreover, if f is pseudo-convex at x^* and if $h_i \in C^1, i \in \mathcal{I}(x^*)$, are quasi-convex at x^*, then x^* is a global minimum solution to problem (2.1).

Since the convexity property of f and h_i is a stronger condition than pseudo- and quasi-convexity, the KKT conditions are sufficient under the convexity assumption as well.

2.1.4 Second-order Conditions

Both the necessary and sufficient second-order conditions are derived based on the restricted constraint set, which contains only the active constraints. Basically, the second-order conditions restrict the direction d based on the curvature of the feasible set so that the gradient of the Lagrangian function vanishes at the optimal solution point x^*.

Let f, h_i for $i = 1, \ldots, m$, be C^2 continuous. The *Lagrangian* function for problem (2.1)

is defined as

$$L(x) \equiv f(x) - \sum_{i=1}^{m} \mu_i h_i(x). \tag{2.9}$$

Second-order necessary conditions

Given $f, h \in C^2$ and a point $x^* \in M$ at which the linear independence constraint qualification is satisfied, if x^* is a local minimum for problem (2.1), then x^* is a KKT-point and we have

$$\nabla^2 L(x^*) \equiv \nabla^2 f(x^*) - \sum_{i=1}^{m} \mu_i \nabla^2 h_i(x^*).$$

Moreover, the matrix $\nabla^2 L(x^*)$ is positive semi-definite on the tangent cone of the active constraints at x^*, i.e.,

$$d^\intercal \nabla^2 L(x^*) d \geq 0 \quad \forall d \in \mathcal{T}(x^*; M).$$

In fact, these necessary conditions would also be sufficient in the absence of degenerate inequality constraints (having a zero multiplier while active). In the presence of degenerate constraints, following sufficient conditions are required.

Second-order sufficient conditions

Let $f, h \in C^2$. A feasible solution x^* is a local optimum for problem (2.1) if there exist multipliers $\mu_i, i = 1, \ldots, m$, such that

$$\nabla f(x^*) - \sum_{i=1}^{m} \mu_i \nabla h_i(x^*) = 0 \tag{2.10}$$

$$\mu_i h_i(x^*) = 0 \tag{2.11}$$

$$\mu_i \geq 0, h_i(x^*) \geq 0, \tag{2.12}$$

and the Hessian matrix $\nabla^2 L(x^*)$ is positive definite on the subspace C, i.e.,

$$d^\intercal \nabla^2 L(x^*) d > 0 \quad \forall d \in C,$$

where $C \equiv \{d \mid \nabla h_i(x^*)^\intercal d = 0 \quad \forall i \in \mathcal{I}(x^*) \text{ and } \nabla h_i(x^*)^\intercal d \geq 0 \quad \forall i \notin \mathcal{I}(x^*)\}$.

2.2 Semi-Infinite Programming

A semi-infinite programming problem, SIP for short, is an optimization problem with a finite number of variables and infinitely many constraints in the following form:

$$\begin{array}{lll} \text{(SIP)} \quad \text{minimize} & f(x) & \\ \text{subject to} & g(x,t) \geq 0, \ t \in T & \\ & x \in \Omega, & \end{array} \tag{2.13}$$

where $T \subset R^l$ is a compact index set, $f : R^n \to R$ and $g : R^n \times R^l \to R^p$ are continuous functions.

Then, the feasible set of problem (2.13) is defined as

$$M(g) = \{x \in \Omega \mid g(x,t) \geq 0, \quad \text{for all } t \in T\}.$$

The infinite number of inequality constraints brings several difficulties, especially in that they may change the structure of the feasible set. Within the past four decades, considerable research has been directed to overcoming these difficulties, in particular, for linear semi-infinite programs. A comprehensive literature survey and a manuscript on semi-infinite programming is given by Hettich and Kortanek [24] and by Reemtsen [50] with the references therein. In this section, we review the feasible set structure, extended constraint qualifications, the first and second order conditions for semi-infinite programs.

2.2.1 Structure of the Feasible Set

The feasible set $M(g)$ possesses different structural properties depending on the topology of the index set and the properties of the constraint functions. For a feasible set M that consists of finitely many inequalities, the active set $\mathcal{I}(\bar{x})$ at $\bar{x}$ is nonempty and contains a finite number of active constraint indices. However for an arbitrary T, the feasible set

$$M(g) = \{x \in \Omega \mid \ g(x,t) \geq 0, \quad \text{for all } \ t \in T\}, \tag{2.14}$$

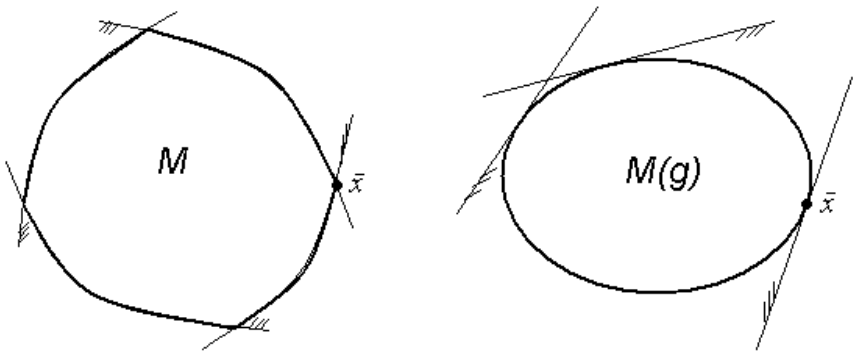

Figure 2.1: Feasible set structure in finite and semi-infinite programming

the set of active constraints at a particular point $\bar{x}$, which is denoted by

$$\mathcal{I}_g(\bar{x}) \equiv \{t \in T \mid g(\bar{x}, t) = 0\},$$

can be empty or can have an infinite cardinality. The feasible set M in Figure 2.1 is described by four inequality constraints of which two are active at $\bar{x}$. By continuity, we have $\mathcal{I}(\bar{x}+\varepsilon) \subset \mathcal{I}(\bar{x})$ for all $\varepsilon > 0$ that satisfy $\bar{x}+\varepsilon \in M \cap \mathcal{N}(\bar{x})$. However, for the feasible set $M(g)$ depicted as an intersection of the tangent halfspaces described by infinitely many linear inequalities (only three of them shown), we have

$$\mathcal{I}_g(\bar{x}+\varepsilon) \not\subset \mathcal{I}_g(\bar{x}).$$

Clearly for semi-infinite programs, the active set $\mathcal{I}_g(\bar{x})$ changes from one point another along the boundary of the feasible set. Hence, the control of the active set becomes a main challenge in studying semi-infinite programs. One way to achieve control of the active set is to locally reduce the feasible set $M(g)$ to a set described by means of a finite number of constraints. This is a well-known approach introduced by Hettich and Jongen [23] and is known as *Reduction Ansatz* or *local reduction*.

Moreover, without any restrictions on the index set T, the tangent cone $\mathcal{T}(\bar{x}, M(g))$

cannot be described by constraint gradients and hence $\mathcal{T}(\bar{x}, M(g)) \neq G(\bar{x}, M(g))$ in semi-infinite programming. The following results by Li et al. [37] provide additional conditions to extend nonlinear programming results to the semi-infinite programming.

Proposition 2.2.1. *Let $\bar{x}$ be a point on the boundary of $M(g)$. If the constraint set $\sigma(\bar{x}) \equiv \{g(\bar{x}, t) \mid t \in T\}$ is closed in R, then $\mathcal{I}_g(\bar{x}) \neq \emptyset$.*

Definition 2.2.1. Let $co(A)$ denote the convex hull of set A and $\partial g(\bar{x})$ denote the subdifferential of a convex function g at $\bar{x}$, i.e.,

$$\partial g(\bar{x}) \equiv \{y \in R^n \mid \langle y, z - \bar{x} \rangle \leq g(z) - g(\bar{x}), \quad z \in R^n\}.$$

Then, the convex hull of the subdifferentials of the active constraints, denoted by $\mathcal{V}(\bar{x})$ is given by

$$\mathcal{V}(\bar{x}) = \text{co} \left(\bigcup_{t \in \mathcal{I}_g(\bar{x})} \partial g(\bar{x}, t) \right).$$

Given the function $G(\bar{x}) = \inf_{t \in T} g(\bar{x}, t)$, an important property due to Danskin [9] is

$$\partial G(\bar{x}) = \mathcal{V}(\bar{x}). \tag{2.15}$$

In semi-infinite programming, in general property (2.15) is violated for an arbitrary index set T. The following result states conditions under which property (2.15) holds for semi-infinite programming problems.

Proposition 2.2.2. *If T is a compact set and the function $g(x, t)$ is convex and continuous in t over T, then (2.15) holds for all $\bar{x} \in R^n$.*

In light of Propositions 2.2.1 and 2.2.2, we make the following assumptions throughout this manuscript:

Assumption 2.2.1. *The index set $T \subset R^l$ is compact.*

Assumption 2.2.2. *The function $g(x, t)$ for $t \in T$ is convex and continuous in t.*

Given a compact index set T and the function $g(x,t) \in C^k, k \geq 2$, we will present the topological structure of the feasible set $M(g)$ with and without local reduction, which allows to represent $M(g)$ by means of finite number of constraints that are active in some neighborhood of a given point $\bar{x}$.

Local Reduction

Let $\phi(x)$ denote the marginal function defined as

$$\phi(x) = \min_{t \in T} g(x,t).$$

Then the feasible set $M(g)$ takes the following form:

$$M(g) = \{x \in \Omega \mid \phi(x) \geq 0\}.$$

From Assumption 2.2.1, it follows that $\phi(x)$ is continuous but not differentiable. If $\bar{x} \in M(g)$ and $\phi(\bar{x}) > 0$, then $\phi(x) > 0$ for all x in some neighborhood of $\mathcal{N}(\bar{x})$. Hence, $\mathcal{N}(\bar{x}) \subset M(g)$ and the local structure of $M(g)$ at $\bar{x}$ is not of interest because the active set is empty, i.e. $\mathcal{I}_g(\bar{x}) = \emptyset$. The important case to study is $\bar{x} \in M(g)$ and $\phi(\bar{x}) = 0$, i.e., $\mathcal{I}_g(\bar{x}) \neq \emptyset$ and the feasible points at which $\phi(\bar{x})$ vanishes. The required regularity conditions for this case can be satisfied by local reduction, if applicable.

Applicability of local reduction is more easily explained for the case in which the index set T is defined by a finite number of inequalities as follows:

$$T \equiv \{t \in R^l \mid c_r(t) \leq 0, r = 1, \ldots, \bar{r}\},$$

where $c_r : T \to R, r = 1, \ldots, \bar{r}$, are C^2 functions. Consider the problem of minimizing $g(\bar{x}, t)$ over T. Associated with this problem, the Lagrangian function at a given point (x,t) is

$$L(x,t,v) \equiv g(x,t) - \sum_{r=1}^{\bar{r}} v_r c_r(t), \quad (x,t,v) \in R^n \times T \times R^{\bar{r}}.$$

Then, $\bar{t} \in T$ is called a *nondegenerate* minimizer of $g(\bar{x}, t)$ for $\bar{x} \in R^n$ if

(a) LICQ is satisfied at $\bar{x}$,

(b) there exists $\bar{v}$ such that $\nabla_t L(\bar{x}, \bar{t}, \bar{v}) = 0$ and $\bar{v}^\intercal c(\bar{t}) = 0$,

(c) $\bar{v}_r \neq 0$ for the active constraints,

(d) $\eta^\intercal \nabla_t^2 L(\bar{x}, \bar{t}, \bar{v})\eta \neq 0$ on the tangent cone of the active constraints at $\bar{x}$.

Conditions (a) and (b) are the first-order necessary conditions and (c) and (d) are the second-order sufficiency conditions with strict complementary slackness. This characterization of nondegeneracy can be carried over to the case in which T is a compact interval and minimizing $g(\bar{x}, t)$ over T is an unconstrained optimization problem.

It is well-known that when g is infinitely differentiable and $T \subset R$, the necessary and sufficient optimality conditions for minimizing $g(\bar{x}, t)$ over T can be expressed in terms of higher order derivatives as follows.

Theorem 2.2.1. *Let $g(\bar{x}, t) \in C^\infty$. Then, $\bar{t}$ is a local minimum if and only if either of the following holds:*

(1) $(\frac{\partial}{\partial t})^i g(\bar{x}, \bar{t}) = 0,\ i = 1, 2, \ldots,$

(2) there exists an even $n \geq 2$ such that $(\frac{\partial}{\partial t})^n g(\bar{x}, \bar{t}) > 0$ while $(\frac{\partial}{\partial t})^i g(\bar{x}, \bar{t}) = 0$ for all $1 \leq i < n$.

The local reduction approach is said to hold at $\bar{x} \in M(g)$, if each $t \in \mathcal{I}_g(\bar{x})$ is the nondegenerate local minimizer of $g(\bar{x}, \cdot)$ over T. If each $t \in \mathcal{I}_g(\bar{x})$ is the only strict local minimizer, then the active set $\mathcal{I}_g(\bar{x})$ is a discrete closed subset of the compact set T and hence finite, say $|\mathcal{I}_g(\bar{x})| = m$. Suppose

$$\mathcal{I}_g(\bar{x}) = \{\bar{t}_1, \ldots, \bar{t}_m\}.$$

Applying the implicit function theorem around each point $(\bar{x}, \bar{t}_i)$, the local C^1- mappings $t_1(x), \ldots, t_m(x)$ (with each $t_i(x)$ as a local minimum for $g(x, \cdot)$ over T) is uniquely determined. Thus, we obtain the following representation as a new local coordinate system in a

neighborhood $\mathcal{N}(\bar{x})$ of $\bar{x}$:

$$\phi(x) = \min_{i=1,\ldots,m} \phi^i(x), \qquad \phi^i(x) = g(x, t_i(x)). \tag{2.16}$$

Since $t_i(x)$ is a continuously differentiable function, $\phi^i(x) = g(x, t_i(x))$ is also continuously differentiable with derivative

$$\nabla\phi^i(x) = \nabla_x g(x, t_i(x)) + \nabla_t g(x, t_i(x))\nabla t_i(x).$$

With $\bar{t}_i = t_i(x)$ being a local minimum for $g(x, \cdot)$, $\nabla_t g(x, t_i(x)) = 0$. Hence,

$$\nabla\phi^i(x) = \nabla_x g(x, t_i(x))|_{(\bar{x},\bar{t})} = \nabla_x g(\bar{x}, \bar{t}). \tag{2.17}$$

Assuming that local reduction is possible at $\bar{x} \in M(g)$, the following result due to [23] says that in a neighborhood $\mathcal{N}(\bar{x})$ the feasible set $M(g)$ of SIP can be described by finitely many inequality constraints.

Proposition 2.2.3. *Assume local reduction is possible at $\bar{x} \in M(g)$ and let $\phi^i(x) = g(x, t_i(x)), \quad x \in \mathcal{N}(\bar{x})$. Then there is an open neighborhood $\mathcal{N}_0(\bar{x}) \subseteq \mathcal{N}(\bar{x})$ of $\bar{x}$ with the following properties:*

(a) for $x \in \mathcal{N}_0(\bar{x}), \quad x \in M(g) \iff \phi^i(x) \geq 0, \ i = 1, \ldots, m.$

(b) $x^* \in \mathcal{N}_0(\bar{x})$ is (strict) local minimizer of SIP if and only if x^* is a (strict) local minimizer of the reduced problem

$$\begin{aligned} (SIP_{red}) \quad & \textit{minimize} && f(x) \\ & \textit{subject to} && \phi^i(x) \geq 0, \ i = 1, \ldots, m \\ & && x \in \mathcal{N}_0(\bar{x}). \end{aligned} \tag{2.18}$$

The local feasible set $M(g) \cap \mathcal{N}_0(\bar{x}) = \{x \in \mathcal{N}(\bar{x}) \mid \phi^i(x) \geq 0, \ i = 1, \ldots, m\}$ makes it possible to use nonlinear programming algorithms to solve semi-infinite problems. However, the local reduction approach may not be applicable at all feasible points. In this case, the semi-infinite problem cannot locally be reduced to a finite problem. When the reduction

approach fails, the number of active constraints at the considered point $\bar{x}$ can be infinite due to degeneracy. This makes the local description of the feasible set in terms of new coordinates (2.16) extremely complicated. Consideration of degenerate local minima and their unfolding has been accomplished only for simple index sets.

A detailed topological analysis of the feasible set $M(g)$ with an index set in the form of interval $[a, b]$ is given by Jongen and Zwier [32, 33]. By using coordinate transformation and Whitney topology, they locally define the feasible set in special new coordinates and give conditions under which the active set has finite number of elements. Jongen and Zwier [33] consider the objective function also and introduce a "critical point" concept to show that the local optimality criteria of "*local reduction*" by means of the implicit function theorem is generic. In addition, Jongen et al. [31] extend their analysis of the feasible set by considering both equality and inequality constraints on a more general compact index set $T \subset R^l$. Again, they validate that the conditions required for optimality are satisfied generically.

Description of the feasible set without local reduction

Local reduction is not possible if there exists a degenerate minimum $\bar{t}$ for $g(\bar{x}, \cdot)$ over T. Based on Theorem 2.2.1, a definition of the order of degeneracy for $T \subset R$ is given in [32, 29] as follows.

Definition 2.2.2. For $k \geq 1$, a point $t \in T$ is called a *zero-point* for $g(\bar{x}, \cdot)$ of *order* k, if the following conditions hold at $(\bar{x}, \bar{t})$:

$$(\frac{\partial}{\partial t})^i g(\bar{x}, \cdot) = 0, \ \ i = 0, 1, \ldots, r,$$

$$\text{and } (\frac{\partial}{\partial t})^{r+1} g(\bar{x}, \cdot) \neq 0,$$

where $r = k$ if $\bar{t} \in \text{int}(T)$ and $r = k - 1$ if $\bar{t} \in \text{bdry}(T)$.

According to Definition 2.2.2, an interior point $\bar{t}$ is of order k (k is odd) if

$$g(\bar{x},\bar{t}) = (\frac{\partial}{\partial t})g(\bar{x},\cdot)|_{\bar{t}} = (\frac{\partial}{\partial t})^2 g(\bar{x},\cdot)|_{\bar{t}} = \ldots = (\frac{\partial}{\partial t})^k g(\bar{x},\cdot)|_{\bar{t}} = 0 \quad \text{and}$$

$$(\frac{\partial}{\partial t})^{k+1} g(\bar{x},\cdot)|_{\bar{t}} > 0.$$

Similarly, a boundary point $\bar{t}$ is of order k, if

$$g(\bar{x},\bar{t}) = (\frac{\partial}{\partial t})g(\bar{x},\cdot)|_{\bar{t}} = (\frac{\partial}{\partial t})^2 g(\bar{x},\cdot)|_{\bar{t}} = \ldots = (\frac{\partial}{\partial t})^{k-1} g(\bar{x},\cdot)|_{\bar{t}} = 0 \quad \text{and}$$

$$(\frac{\partial}{\partial t})^{k} g(\bar{x},\cdot)|_{\bar{t}} > 0.$$

The next definition states that a zero-point is of a finite order and introduces a set of vectors with some important properties.

Definition 2.2.3. If $\bar{t} \in T$ is a zero-point for $g(\bar{x},\cdot)$ of order $k, k \geq 1$, then we say $\bar{t}$ is a zero-point of finite order. In that case, denote $Ord(\bar{t}) = k$ and introduce $S_k(\bar{t})$ as a set consisting of k vectors, i.e.,

$$S_k(\bar{t}) = \{(\frac{\partial}{\partial t})^i \nabla_x g(\bar{x},\bar{t}),\ i = 0, 1, \ldots, k-1\}.$$

Proposition 2.2.4. *The point $\bar{t} \in T$ is nondegenerate if and only if $k = 1$, i.e., $Ord(\bar{t})$=1.*

Note that for $k = 1$, $S_1(\bar{t})$ is equal to $\{\nabla_x g(\bar{x},\bar{t})\}$ as in the nondegenerate case given by equation (2.17). In general, the linear independence of the set $S_k(\bar{t})$ guarantees the stability of the degeneracy. The stable case of degeneracy is given based on the following definition of regularity:

Definition 2.2.4. The function $g(\bar{x},t)$ is said to be *regular* with respect to T if the following conditions hold whenever $\phi(\bar{x}) = 0$:

(R1) Every $t \in \mathcal{I}_g(\bar{x})$ is a zero-point for $g(\bar{x},\cdot)$ of finite order.

(R2) The dimension of $span\big(\bigcup_{t\in\mathcal{I}_g(\bar{x})} S_k(t)\big)$ is equal to $\sum_{t\in\mathcal{I}_g(\bar{x})} Ord(t)$.

Let $\mathcal{G}$ denote the set of regular functions with respect to T, i.e.,

$$\mathcal{G} = \{g \in C^\infty \mid g(\bar{x}, t) \text{ regular w.r.t. } T\}.$$

In Definition 2.2.4, (R1) implies the set $\mathcal{I}_g(\bar{x})$ is finite.(R2) implies that for every boundary point $\bar{x}$ of $M(g)$, $\bigcup_{t\in\mathcal{I}_g(\bar{x})} S_k(t)$ defines the normal cone, which is orthogonal to the tangent cone of active constraints at $\bar{x}$. Hence, in a generic stable case, the active set $\mathcal{I}_g(\bar{x})$ is finite. Moreover, $\bar{t}_i, i = 1, \ldots, m$, is of finite order and condition (R2) is satisfied. Using singularity theory and universal unfoldings [4], it has been shown that there exists a local C^∞-coordinate transformation, sending $\bar{x}$ onto the origin such that $M(g)$ locally takes the form

$$\{x \in R^n \mid \eta_{k_l}(\cdot) \geq 0, l = 1, \ldots, m\}$$

where η_k is defined by Ψ_k^1 or Ψ_k^2 depending on whether $\bar{t}$ is an interior or a boundary point, where by [32] and [29],

$$\Psi_k^1(x_1, \ldots, x_k) = \min_{t\in[-1,1]} t^{k+1} + x_1 t^{k-1} + \ldots + x_{k-1}t + x_k,$$

$$\Psi_k^2(x_1, \ldots, x_k) = \min_{t\in[0,1]} t^{k} + x_1 t^{k-1} + \ldots + x_{k-1}t + x_k.$$

These results are summarized in the subsequent theorem by Jöngen and Zwier [33].

Theorem 2.2.2. *Given $g \in \mathcal{G}$ and $M(g) \neq \emptyset$, the feasible set $M(g)$ is a topological manifold with boundary of dimension n.*

Proof. Refer to p.57 of [33]. □

The above results can be specialized to a case in which the reduction principle is applicable. Let the feasible set $M(g)$ consist of two components, the relative interior Σ_0 and the boundary Σ_r, i.e.,

$$\Sigma_0 = \{x \in M(g) \mid \phi(x) > 0\}, \tag{2.19}$$

$$\Sigma_r = \{x \in bdry(M(g)) \mid Ord(t) = 1 \text{ for all } t \in \mathcal{I}_g(x)\}. \tag{2.20}$$

Then, using the fact that each $\bar{t} \in \mathcal{I}_g(\bar{x})$ is a nondegenerate minimum for $g(\bar{x}, \cdot)$ over T, the following result is a direct application of the implicit function theorem.

Proposition 2.2.5. *Let $g \in \mathcal{G}$ and $M(g) \neq \emptyset$. Then,*

(a) $\Sigma_0 \cup \Sigma_r$ is a C^∞ manifold with boundary.

(b) $\Sigma_0 \cup \Sigma_r$ is open and dense in $M(g)$.

(c) Σ_r is open and dense in $bdry(M(g))$.

Moreover, for $i \geq 1$, define

$$\Sigma_i = \{x \in R^n \mid \phi(x) = 0 \text{ and } \sum_{t \in \mathcal{I}_g(x)} Ord(t) = i\}. \tag{2.21}$$

Then, the feasible set $M(g)$ becomes a disjoint union of the sets $\Sigma_0, \Sigma_1, \ldots, \Sigma_n$ and the following result provided in [32] characterizes the structure of $M(g)$.

Proposition 2.2.6. *Let $g \in \mathcal{G}$ and $M(g) \neq \emptyset$. Then,*

(a) Σ_i is a smooth submanifold of R^n of dimension i (or empty).

(b) The representation of $M(g)$ by means of the sets Σ_i is Whitney regular.

(c) If $\Sigma_i \neq \emptyset$, then the closure of Σ_i is $\bigcup_{j \geq i} \Sigma_j$.

(d) Let $\bar{x} \in \Sigma_i$ for $i \geq 1$. Then, $\bigcup_{t \in \mathcal{I}_g(\bar{x})} S_k(t)$ is a basis for the normal space $\mathcal{T}(\bar{x}; \Sigma_i)^\perp$, where $\mathcal{T}(\bar{x}; \Sigma_i)$ denotes the tangent space of Σ_i at $\bar{x}$.

Although the structure of the feasible set for semi-infinite problems is quite complicated, the analysis and results presented in this section help determine the requirements for algorithm development. Given that the feasible set $M(g)$ can be reduced to a topological manifold with boundary of finite dimension, we are able to impose certain constraint qualifications in order to study the optimality conditions.

2.2.2 Constraint Qualifications

It is well known that constraint qualifications are required for obtaining first-order necessary conditions. In Section 2.2.1, we reviewed several constraint qualifications for nonlinear programming problems. Some of these constraint qualifications can be generalized to semi-infinite programming. In particular, extensions of Mangasarian-Fromovitz and Abadie's constraint qualifications have been proposed with corresponding Karush-Kuhn-Tucker theorem and the first-order optimality conditions in [31, 24, 30].

In semi-infinite programming, the description of the feasible set in differentiable local coordinates can be quite complicated when the local reduction is not possible. With some constraint qualifications, the generic description of $M(g)$ in local Lipschitzian coordinates becomes possible. Moreover, the fulfillment of the constraint qualifications leads to characterization of some desired properties (e.g., continuity, compactness) of the feasible set. We next define some extensions of the well-known constraint qualifications and present their interrelations for semi-infinite programming.

Let Assumptions 3.2.1 and 3.2.2 hold and recall that

$$\mathcal{V}(\bar{x}) = \operatorname{co}\left(\bigcup_{t\in\mathcal{I}_g(\bar{x})} \partial g(\bar{x},t)\right).$$

For $\bar{x} \in M(g)$, define

$$\begin{aligned}
&\mathcal{L}_0(\bar{x}; M(g)) = \{\xi \in R^n \mid w^\top \xi > 0,\ w \in \mathcal{V}(\bar{x})\},\\
&\mathcal{L}(\bar{x}; M(g)) = \{\xi \in R^n \mid w^\top \xi \geq 0,\ w \in \mathcal{V}(\bar{x})\},\\
&\tilde{\mathcal{L}}(\bar{x}; M(g)) = \{\xi \in R^n \mid \inf_{t\in\mathcal{I}_g(\bar{x})} \nabla g(\bar{x},t)^\top \xi \geq 0\}.
\end{aligned} \tag{2.22}$$

Observe the following relation among the cones in (2.22):

$$\mathcal{L}_0(\bar{x}; M(g)) \subseteq \mathcal{L}(\bar{x}; M(g)) \subseteq \tilde{\mathcal{L}}(\bar{x}; M(g)).$$

For optimization problems with a finite number of explicit constraints, we have

$$\mathcal{L}(\bar{x}; M(g)) = \tilde{\mathcal{L}}(\bar{x}; M(g)),$$

whereas for semi-infinite programs, $\mathcal{L}(\bar{x}; M(g))$ can be much smaller than $\tilde{\mathcal{L}}(\bar{x}; M(g))$ and it need not contain the tangent cone. Another major difference is related to the convex hull $\mathcal{V}(\bar{x})$. For nonlinear programming problems, $\mathcal{V}(\bar{x})$ is closed and

$$\mathcal{L}(\bar{x}; M(g)) = \mathcal{V}(\bar{x}).$$

For semi-infinite programs, $\mathcal{V}(\bar{x})$ is not necessarily closed and it is not equal to $\mathcal{L}(\bar{x}; M(g))$. Therefore, we need the following extended constraint qualifications:

Extended Abadie constraint qualification

The extended Abadie constraint qualification (EACQ) is said to hold at $\bar{x} \in M(g)$ if $\mathcal{T}(\bar{x}; M(g)) = \tilde{\mathcal{L}}(\bar{x}; M(g))$.

Extended linear independence constraint qualification

The extended linear independence (ELICQ) is said to hold at $\bar{x} \in M(g)$ if the vectors $\nabla_x g(\bar{x}, t)$, for $t \in \mathcal{I}_g(\bar{x})$, are linearly independent. Note that this condition implies that active set $\mathcal{I}_g(\bar{x})$ is finite. Also, if the active set is known to be finite, then by the regularity conditions of Definition 2.2.4, we know ELICQ holds.

As in nonlinear programming, ELICQ implies the following constraint qualification, which is an extended version of MFCQ and is less restrictive.

Extended Mangasarian Fromowitz constraint qualification

The EMFCQ is said to hold at $\bar{x} \in M(g)$ if $\mathcal{L}_0(\bar{x}; M(g)) \neq \emptyset$ or if there exists a vector $\xi \in R^n$ such that

$$\nabla_x g(\bar{x}, t)^\intercal \xi > 0, \quad \text{for all } t \in \mathcal{I}_g(\bar{x}). \tag{2.23}$$

Note that EMFCQ implies the closedness of $\mathcal{V}(\bar{x})$, but the converse is not true.

The following result of Jongen et al. [31] shows that under the assumption of EMFCQ, the feasible set $M(g)$ is a Lipschitzian manifold with boundary. They also show that EMFCQ is a generic property with respect to $g \in C^1$.

Theorem 2.2.3. *Suppose EMFCQ is satisfied at all points of the feasible set $M(g)$. Then, $M(g)$ is a Lipschitzian manifold with boundary of dimension $n - |\mathcal{I}_g(\bar{x})|$, and the boundary is given by*

$$bdry(M(g)) = \{x \in R^n \mid \phi(x) = 0\}.$$

Proof. The proof of can be found on p.540-542 in [31]. □

Another property of the feasible set as a result of EMFCQ is its stability. A set $M(g)$ is called *stable* if there exists a C^1-neighborhood of g such that for every $\tilde{g}$, the corresponding set $M(\tilde{g})$ is topologically equivalent to (homeomorphic with) $M(g)$. The following stability theorem and Theorem 2.2.3 imply that the feasible set $M(g)$ does not bifurcate.

Theorem 2.2.4. *Assume $M(g)$ is compact. Then, the feasible set $M(g)$ is stable if and only if EMFCQ holds at every point $x \in M(g)$.*

Proof. A lengthy proof can be found on p.547-551 in [31]. □

Extended Kuhn-Tucker constraint qualification

The extended Kuhn-Tucker constraint qualification (EKTCQ) is said to hold at $\bar{x} \in M(g)$ if $\mathcal{L}'(\bar{x}; M(g)) \subset cl\ A(\bar{x}; M(g))$. Since we always have $cl\ A(\bar{x}; M(g)) \subset \mathcal{L}'(\bar{x}; M(g))$, EKTCQ holds if $\mathcal{L}'(\bar{x}; M(g)) = cl\ A(\bar{x}; M(g))$.

Note that EKTCQ by itself is not sufficient to imply the closedness of $\mathcal{V}(\bar{x})$. Therefore, the results and the first-order conditions obtained by EMFCQ can be obtained only if EKTCQ and the closedness of $\mathcal{V}(\bar{x})$ hold simultaneously [57].

For nonlinear programming problems, constraint qualifications have the following implication relations at a point $\bar{x}$:

$$\text{(LICQ), (SCQ)} \Rightarrow \text{(MFCQ)} \Rightarrow \text{(KTCQ)} \Rightarrow \text{(ACQ)}.$$

This well-known interrelation cannot be extended to semi-infinite programming. A somewhat more complicated implication relations are given in [57], namely

$$\begin{gathered}\text{(ELICQ)} \Rightarrow \text{(EMFCQ)} \Rightarrow \text{(EKTCQ)},\\ \text{(EMFCQ)} \not\Rightarrow \text{(EACQ)},\\ \text{(EKTCQ)} \not\Rightarrow \text{(EACQ)}.\end{gathered}$$

In the subsequent section, the first-order conditions are defined assuming EMFCQ. If $\mathcal{V}(\bar{x})$ is closed, then the same conditions are valid under EKTCQ as well.

2.2.3 First-order Conditions

This section presents the first-order necessary and sufficient conditions for semi-infinite optimization problems. Note that the first-order conditions vary with respect to the structure of the index set T and consequently the structure of the feasible set. That is, depending on the degeneracy and regularity of the active set, optimality conditions are defined differently. The first-order conditions for degenerate index sets as a compact interval have been studied by Jongen and Zwier [32]. Similar results for more generalized compact index set have been obtained by Stein and Still [53], and Stein [52]. For more details on the first-order optimality conditions for reduced problems (nondegenerate case) and related literature, refer to Hettich and Kortanek [24], Jongen et al. [30] and Shapiro [51].

Given that T is a compact set and the functions f and g are continuously differentiable with respect to x, we introduce the first-order conditions that apply to both degenerate and nondegenerate index sets as long as $g \in \mathcal{G}$ (cf. Definition 2.2.4).

Theorem 2.2.5. *Let $\bar{x}$ be a local minimum for problem (SIP). If $\mathcal{I}_g(\bar{x}) \neq \emptyset$ and $\mathcal{I}_g(\bar{x}) = \{\bar{t}_1, \ldots, \bar{t}_m\}$, for some m, then there exist real numbers λ_j such that*

$$\begin{aligned} &\nabla f(\bar{x}) = \textstyle\sum_{j=1}^{m} \lambda_j \nabla_x g(\bar{x}, \bar{t}_j), \\ &\lambda_j \geq 0, \ j = 1, \ldots, m. \end{aligned} \tag{2.24}$$

A point $\bar{x} \in M(g)$ at which condition (2.24) holds is called a critical solution or KKT-point of SIP. An optimal solution to SIP satisfies condition (2.24) and therefore is a critical solution, but not every critical solution is optimal. The sufficient conditions for a given critical point to be optimal can be obtained only under constraint qualifications.

Note that the first-order condition (2.24) is applicable only to the locally reduced semi-infinite problems.

Next, we present the first-order condition under certain constraint qualifications.

Theorem 2.2.6. *Assume $\mathcal{I}_g(\bar{x}) \neq \emptyset$ and EMFCQ is satisfied at $\bar{x}$. If $\bar{x}$ is an optimal solution to the semi-infinite problem (2.13), then there is no vector $\xi \in R^n$ satisfying*

$$\nabla f(x)^\intercal \xi < 0 \quad \textit{and} \quad \nabla_x g(\bar{x}, t)^\intercal \xi > 0, \quad \forall t \in T. \tag{2.25}$$

Condition (2.25) also defines a sufficient condition for ξ to be an improving feasible direction. Imposing the constraint qualifications ensures that every vector satisfying $\nabla_x g(\bar{x}, t)^\intercal \xi \geq 0, \quad \forall t \in T$, is a feasible direction.

Another set of first-order conditions is defined by Shapiro [51] and Hettich [24] in the space of nonnegative Borel measures on T.

Theorem 2.2.7. *Assume $\mathcal{I}_g(\bar{x}) \neq \emptyset$ and EMFCQ is satisfied at $\bar{x}$. Then there exists a nonnegative Borel measure μ on T such that*

$$F(\bar{x}) - \int_T \nabla_x g(\bar{x}, t) \mu(t) dt = 0 \quad \textit{and } supp(\mu) \subset \mathcal{I}_g(\bar{x}) \tag{2.26}$$

and the set $\mathcal{B}(\bar{x})$ of such nonnegative Borel measures μ, satisfying condition (2.26), is bounded.

If $M(g)$ has a nonempty interior, the converse is also true. In other words, if the set of Borel measures satisfying (2.26) is nonempty and bounded, then the EMFCQ holds.

2.2.4 Second-order Conditions

The second-order optimality conditions for semi-infinite problems are developed for locally reduced problems by applying the results obtained from nonlinear programming.

Notice that the reduced problem can be written as

$$\begin{array}{lll} (\text{SIP}_{red}) \quad \text{minimize} & f(x) & \\ \text{subject to} & g(x, t_i(x)) \geq 0, \quad i = 1, \ldots, m & \\ & x \in \mathcal{N}(\bar{x}), & \end{array} \tag{2.27}$$

where $t_i(\bar{x})$ is a global minimizer of $g(\bar{x}, \cdot)$ over T. Let the cone of feasible directions at $\bar{x}$, $\mathcal{C}(\bar{x})$, be defined as

$$\mathcal{C}(\bar{x}) = \{\xi \in R^n \mid \nabla f(\bar{x})^\intercal \xi \leq 0 \ \text{ and } \nabla_x g(\bar{x}, t_i(x))^\intercal \xi \geq 0, \ \text{ for } i = 1, \ldots, m\}$$

and the Lagrangian as follows

$$L(\bar{x}, \lambda) = \lambda_0 f(\bar{x}) - \sum_{i=1}^{m} \lambda_i g(\bar{x}, t_i(\bar{x})).$$

Then the second order conditions are given by the following theorem of Still [54].

Theorem 2.2.8. *Suppose that the local reduction is applicable to SIP and that SIP can be reduced to problem (2.27) in a neighborhood of $\bar{x}$. Given,*

$$\nabla_x L(\bar{x}, \bar{\lambda}) = \bar{\lambda}_0 \nabla_x f(\bar{x}) - \sum_{i=1}^{m} \bar{\lambda}_i \nabla_x g(\bar{x}, \bar{t}_i(\bar{x}))$$

and

$$\nabla_x{}^2 L(\bar{x}, \bar{\lambda}) \bar{\lambda}_0 \nabla_x{}^2 f(\bar{x}) - \sum_{i=1}^{m} \bar{\lambda}_i \nabla_x{}^2 g(\bar{x}, \bar{t}_i(\bar{x})) + \sum_{i=1}^{m} \bar{\lambda}_i \nabla^\intercal t_i(\bar{x}) \nabla_t^2 g(\bar{x}, \bar{t}_i(\bar{x})) \nabla t_i(\bar{x}),$$

we have the following necessary and sufficient conditions:

(a) Let $\bar{x}$ be a local minimizer of SIP given in (2.13). Then, for any $\xi \in \mathcal{C}(\bar{x})$ there exists a multiplier $\bar{\lambda} \geq 0$ such that

$$\nabla_x L(\bar{x}, \bar{\lambda}) = 0 \quad \textit{and} \quad \xi^\intercal \nabla_x{}^2 L(\bar{x}, \bar{\lambda}) \xi \geq 0. \tag{2.28}$$

(b) Suppose for any $\xi \in \mathcal{C}(\bar{x}) \backslash \{0\}$, there exists a multiplier $\bar{\lambda} \geq 0$ such that $\nabla_x L(\bar{x}, \bar{\lambda}) = 0$ and $\xi^\intercal \nabla_x{}^2 L(\bar{x}, \bar{\lambda}) \xi > 0$. Then $\bar{x}$ is a (strict) local minimizer of SIP in (2.13).

The second-order necessary and sufficient conditions under weaker assumptions can be found in [51] and [25].

2.3 Variational Inequalities

Variational inequalities are considered to be a powerful tool for modelling a variety of problems including optimization problems, equilibrium problems, complementarity problems and systems of equations. The variational inequality problem in infinite-dimensional metric spaces was first introduced by Hartman and Stampacchia [22]. During the past four decades, a considerable amount of work has been dedicated to so called *finite dimensional variational inequalities*; that is, variational inequality problems defined by a finite dimensional mapping on a finite dimensional domain.

Definition 2.3.1. Let X be a nonempty subset of R^n and let F be a function from R^n into itself. The variational inequality problem, denoted by $\mathrm{VI}(X, F)$, is to find a vector $x^* \in X$ such that

$$F(x^*)^\intercal (y - x^*) \geq 0 \quad \text{for all} \quad y \in X. \tag{2.29}$$

The geometric interpretation of variational inequality $\mathrm{VI}(X, F)$ is depicted in Figure 2.2. As illustrated, x^* is a solution of $\mathrm{VI}(X, F)$ if and only if vector $F(x^*)$ makes an acute angle with all vectors $y - x^*$ such that $y \in X$.

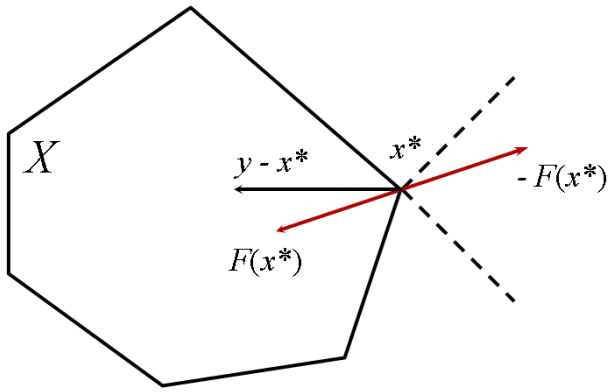

Figure 2.2: Geometric interpretation of VI(X, F)

Among the several published approaches to study variational inequalities, there are three commonly seen methods: merit function based optimization methods [17, 36, 18, 48, 55]; reformulating VI(X, F) as a nonlinear and nonsmooth system of equations [45, 46, 61, 62]; and applying continuation methods [34, 28, 63, 13] or interior-point methods [56, 44, 10], see also in [12, 20].

2.3.1 Equivalent Reformulations of Variational Inequalities

This section explores equivalent reformulations of variational inequalities including a nonlinear programming formulation, a projection formulation and a mixed nonlinear complementarity formulation.

Nonlinear programming reformulation

A merit function for VI(X, F) on set $K \supseteq X$ is a nonnegative function $\theta : K \to R_+$ such that x^* is a solution to VI(X, F) if and only if $x^* \in K$ and $\theta(x^*) = 0$. In other words, the function θ is a merit function for VI(X, F) if and only if the solutions of VI(X, F) coincide

with the global solutions of the problem

$$\begin{aligned} \text{minimize} \quad & \theta(y) \\ \text{subject to} \quad & y \in K \end{aligned} \tag{2.30}$$

and the optimal objective value of this problem is zero. The first and most commonly used merit functions are the class of functions known as "gap function".

It is well known that if the $F(x)$ is the gradient of a real-valued differentiable function $f(x) : R^n \to R$ and X is convex, $\mathrm{VI}(X, F)$ is equivalent to the first-order necessary conditions of an optimization problem

$$\text{minimize}_{x \in X} \ f(x).$$

Hence, every nonlinear program defined on a convex set corresponds to a variational inequality via its first-order necessary conditions. The question of whether the variational inequality $\mathrm{VI}(X, F)$ can be formulated as an optimization problem had been an open one until Fukushima [18] proposed an equivalent constrained optimization problem formulation. His reformulation is based on the gap function θ defined as follows.

Definition 2.3.2. The gap function $\theta(x)$ associated with $\mathrm{VI}(X, F)$ is defined by

$$\theta(x) \equiv \max_{y \in X} \big\{ F(x)^{\mathsf{T}}(x - y) \big\}. \tag{2.31}$$

The gap function $\theta(x)$ is nonnegative on the feasible set and $\theta(x) = 0$ if and only if $x \in X$ solves $\mathrm{VI}(X, F)$. Based on the definition of the gap function, the variational inequality problem can be formulated as a constrained minimization problem

$$\begin{aligned} \text{minimize} \quad & \theta(x) \\ \text{subject to} \quad & x \in X. \end{aligned} \tag{2.32}$$

Although the gap function gives an equivalent optimization problem (2.32), $\theta(x)$ is not differentiable in general. Furthermore, if X is unbounded, θ may not be of finite value or

well-defined. Therefore, this reformulation has not been of much use until Fukushima [18] and Auchmuty [1] independently introduced the *regularized gap function*

$$\theta_\alpha(x) \equiv \max_{y \in X} \left\{ F(x)^\top (x - y) - \frac{\alpha}{2} ||x - y||^2 \right\}, \tag{2.33}$$

where α is a positive parameter. An important property of the regularized gap function that it is continuously differentiable. Also, since θ_α is strongly concave with respect to y, there is a unique maximizer and therefore, the regularized gap function is finite-valued everywhere. This gives rise to following equivalent formulation of VI(X, F):

$$\begin{aligned} &\text{minimize} \;\; \theta_\alpha(x) \\ &\text{subject to } x \in X. \end{aligned} \tag{2.34}$$

As an extension, Peng [48] proposed an equivalent unconstrained minimization formulation,

$$\begin{aligned} &\text{minimize} \;\; g_{\alpha\beta}(x) \\ &\text{subject to } x \in R^n, \end{aligned} \tag{2.35}$$

by using the *D-gap function*, defined as the difference of two regularized gap functions, i.e.,

$$g_{\alpha\beta}(x) \equiv \theta_\alpha(x) - \theta_\beta(x), \quad 0 < \alpha < \beta. \tag{2.36}$$

It is easy to see the nice properties (e.g. differentiability and finite-valued) of the regularized gap function carry over to the D-gap function. Suppose x^* is the optimal solution for problem (2.34) (or (2.35)). Then, x^* solves VI(X, F) provided that the function F is monotone. The monotonicity requirement is considered to be the drawback of using gap functions for variational inequalities.

The use of merit functions can be very attractive for theoretical solution analysis and development for variational inequality problems. However, the evaluation of a merit function e.g., $\theta(x)$ can be a difficult task. Merit function based algorithms can be implemented effectively only if X has a relatively simple structure.

Projection Reformulation

Variational inequality problem can be reformulated as a classical fixed-point problem by using the notion of projection.

Definition 2.3.3. Let S be a nonempty, closed and convex subset of R^n. The Euclidean projection of a vector $x \in R^n$ onto the set S, denoted by $\Pi_S(x)$, is defined as the solution to the following optimization problem:

$$\begin{aligned} &\text{minimize } \frac{1}{2}||y-x||^2 \\ &\text{subject to } \ y \in S. \end{aligned} \tag{2.37}$$

The well-known properties of the Euclidean projection are given by the following proposition [12].

Proposition 2.3.1. *Let S be a nonempty closed convex subset of R^n.*

(a) For each vector $x \in R^n$, $\Pi_S(x)$ exists and is unique.

(b) For each vector $x \in R^n$,

$$(y - \Pi_S(x))^\intercal(\Pi_S(x) - x) \geq 0 \ \ for\ all \ \ y \in S. \tag{2.38}$$

(c) For any two vectors $v \in R^n$ and $z \in R^n$,

$$(\Pi_S(v) - \Pi_S(z))^\intercal(v - z) \geq ||\Pi_S(v) - \Pi_S(z)||_2^2.$$

(d) As a function of x, $\Pi_S(x)$ is nonexpansive, that is, for any two vectors $v, z \in R^n$,

$$||\Pi_S(v) - \Pi_S(z))||_2 \leq ||v - z||_2;$$

making Π_S globally Lipschitz continuous on R^n.

(e) The squared distance function

$$\rho(x) \equiv \frac{1}{2}||x - \Pi_S(x)||_2^2, \quad x \in R^n$$

is continuously differentiable in x and $\nabla\rho(x) = x - \Pi_S(x)$.

Using these properties, we obtain the fixed-point formulation of a variational inequality problem $\mathrm{VI}(X, F)$.

Proposition 2.3.2. *Let X be a nonempty closed convex subset of R^n. Then x^* solves $VI(X, F)$ if and only if*

$$x^* - \Pi_X(x^* - F(x^*)) = 0,$$

i.e., x^* is a fixed point of the mapping $H : R^n \to R^n$

$$H(x) = \Pi_X(x - F(x)).$$

Projection methods are applicable only for monotone variational inequality problems. The projection-type algorithms are usually based on basic fixed-point iteration, which requires strongly monotone and Lipschitz continuous F for convergence [42]. An extragradient method (and its variants) that evaluates two projections at each iteration has been proposed for pseudo-monotone but still Lipschitz continuous F with the knowledge of Lipschitz constant. The requirements for Lipschitz continuity and the knowledge of the Lipschitz constant of F were finally removed in [26] by the introduction of hyperplane projection method, which uses the projection on a separating hyperplane determined by a line-search.

Efficient implementation of all projection methods depends on efficient calculation of the projection onto a convex set, which might be computationally difficult. Since no derivative information is required, these methods are reported to be extremely simple when the projection is easily computable. A review of other projection-type algorithms for monotone variational inequality problems can be found in [12].

KKT conditions and mixed nonlinear complementarity formulation

It is well known that mixed nonlinear complementarity problem (MiCP) is a special case of variational inequalities. Moreover, the variational inequality VI(X, F) defined on domain

$$X = \{x \in R^n : g_i(x) \leq 0 \;\; i = 1, 2, \ldots, m; \; h_j(x) = 0 \;\; j = 1, 2, \ldots, p\} \tag{2.39}$$

can be converted to a mixed complementarity problem if the functions $g : R^n \to R^m$ and $h : R^n \to R^p$ in (2.39) satisfy certain constraint qualifications.

The following result is due to [20] and is known as *Karush-Kuhn-Tucker conditions* for VI(X, F).

Proposition 2.3.3. *Let $g : R^n \to R^m$ and $h : R^n \to R^p$ be continuously differentiable and let X be defined by (2.39).*

(a) If x^* solves VI(X, F) and the constraint qualifications hold for X at x^*, then for some $\lambda^* \in R^m$ and $\mu^* \in R^p$, (x^*, λ^*, μ^*) solves MiCP$(R^n \times R^m_+ \times R^p, H)$ where $H : R^{n+m+p} \to R^{n+m+p}$ is defined by

$$H\begin{pmatrix} x \\ \lambda \\ \mu \end{pmatrix} = \begin{pmatrix} F(x) + \sum_{i=1}^{m} \lambda_i \nabla g_i(x) + \sum_{j=1}^{p} \mu_j \nabla h_j(x) \\ -g(x) \\ h(x) \end{pmatrix}. \tag{2.40}$$

(b) Conversely, if g_i is convex, for $i = 1, \ldots, m$, and h_j is affine, for $j = 1, \ldots, p$, and (x^*, λ^*, μ^*) solves MiCP$(R^n \times R^m_+ \times R^p, H)$, then x^* solves VI(X, F).

Based on this result, a variational inequality problem becomes a problem of solving the system of equations

$$H(z) = 0,$$

where $z = (x, \lambda, \mu)$.

Methods based on the KKT conditions (2.40) have been well-studied in nonlinear programming [41, 59] and offer a convenient approach for solving variational inequality problems. In contrast to merit function and projection type methods, the complexity of the domain X is not an issue for approaches based on the KKT conditions. In particular, the mixed complementarity reformulation allows the implementation of equation-based, interior point and smoothing-based solution algorithms, which usually provide sharp convergence results. Among the several algorithms for computing the KKT triple (x^*, λ^*, μ^*) are the quasi-Newton method of Qi and Jiang [49], the B-differentiable Newton method of Pang [45, 46], Xiao and Harker [61, 62] and the trust region method of Yang et al. [63]. The smoothing methods and their fast convergence results can be found in Chen, Qi and Sun's work [7, 12]. Therefore, KKT condition based equation reformulation is the chosen approach for solving finite dimensional variational inequalities within our solution methodology for semi-infinite variational inequality problem SIVI(X, F) in Chapter 4.

2.3.2 Constraint Qualifications

The constraint qualification requirements for variational inequality problems depend on the structure of the domain X. Recall that if X has a special geometric structure such as polyhedral or conic, then the regularity of active constraints is guaranteed by the specific properties of the corresponding geometry. However, when X is defined in terms of nonlinear equality and inequality constraints, then we need to impose certain constraint qualifications on the domain.

Suppose that the domain X is defined as below

$$X = \{x \in R^n \mid g_i(x) \geq 0 \quad i = 1, 2, \ldots, m\}. \tag{2.41}$$

Note that the domain X has the same topological structure as the feasible set M in nonlinear programming. Therefore, the constraint qualifications (Slater, LICQ, MFCQ, KTCQ, and

Abadie) presented in Section 2.2 apply to problem VI(X, F).

The constraint qualifications in the context of variational inequalities play an important role especially in obtaining equivalent reformulations of VI(X, F). Without constraint qualifications, VI(X, F) cannot be easily reformulated as a system of equations since the KKT conditions (2.40) will not be well-defined. Similarly, since the optimal solution to a nonlinear optimization problem cannot be characterized without certain constraint qualifications, it is difficult to obtain an equivalent reformulation of VI(X, F).

The existence of constraint qualifications for VI(X, F) also guarantees that a solution to VI(X, F) is nondegenerate and the corresponding multiplier set is nonempty. In most cases, having a nondegenerate solution implies that the Jacobian of the mapping F is nonsingular, which is a required property in obtaining a global error bound and performing sensitivity analysis. For more details, reader can refer to [12].

2.3.3 Solution Analysis

The existence and uniqueness of a solution to variational inequalities and nonlinear complementarity problems have been well studied. Before stating the key results (without proofs), we need the following definitions [21].

Definition 2.3.4. The mapping $F : R^n \to R^n$ is said to be

(a) monotone over a set X if

$$\big[F(x) - F(y)\big]^\intercal (x - y) \geq 0, \quad \forall x, y \in X.$$

(b) pseudo-monotone over X if

$$F(y)^\intercal (x - y) \geq 0 \quad \text{implies} \quad F(x)^\intercal (x - y) \geq 0, \quad \forall x, y \in X.$$

(c) strictly monotone over X if

$$\big[F(x) - F(y)\big]^\intercal (x - y) > 0, \quad \forall x, y \in X,\ x \neq y.$$

(d) strongly monotone over X if there exists an $\alpha > 0$ such that

$$\big[F(x) - F(y)\big]^\top (x - y) \geq \alpha \|x - y\|^2, \quad \forall x, y \in X.$$

(e) coercive with respect to X if there exists a vector $x^0 \in X$ such that

$$\lim_{x \in X, \|x\| \to \infty} F(x)^\top (x - x^0)/\|x\| = \infty,$$

where $\|\cdot\|$ denotes any vector norm in R^n.

Clearly, strong monotonicity implies strict monotonicity, which implies monotonicity, which implies pseudo-monotonicity.

Theorem 2.3.1. *Let X be a nonempty, compact and convex subset of R^n and let F be a continuous mapping from X into R^n. Then there exists a solution to VI(X, F).*

Proposition 2.3.4. *Let X be a nonempty, closed and convex subset of R^n and let F be a continuous, pseudo-monotone mapping from X into R^n. Then x^* solves VI(X, F) if and only if $x^* \in X$ and*

$$F(y)^\top (y - x^*) \geq 0 \quad \forall y \in X.$$

Moreover, the solution set of VI(X,F) is convex when it is nonempty.

Proposition 2.3.5. *If F is strictly monotone on X, then VI(X, F) has at most one solution.*

Note that according to Propositions 2.3.4 and 2.3.5, the solution set maybe empty. In order to ensure that solution set is not empty when X is unbounded, the following result is needed.

Theorem 2.3.2. *Let X be a nonempty, closed convex subset of R^n and F be a continuous mapping from X into R^n. If F is coercive with respect to X, then VI(X, F) has a nonempty compact solution set.*

If X is not assumed to be bounded, a necessary and sufficient condition for VI(X, F) to have a solution is given in Harker [20] with its proof in [12].

Theorem 2.3.3. *Let $X \subseteq R^n$ be closed and convex with $int(X) \neq \emptyset$, i.e., nonempty topological interior, and let $F : R^n \to R^n$ be continuous. The VI(X, F) has a nonempty solution set if and only if there exists a closed set $E \subseteq R^n$ with $int(E) \neq \emptyset$ such that the restricted $VI(X \cap E, F)$ has a solution in $int(E)$.*

Generalizing the concepts for complementarity problems, Zhao [64] introduced the concept of "exceptional family of elements" for a weaker existence condition for variational inequalities over polyhedral sets. In a later work, Zhao and Han [65] considered the set

$$K = \{x \in R^n \mid g_i(x) \leq 0 \;\; i = 1, \ldots, m, h_j(x) \leq 0 \;\; j = 1, \ldots, l\}$$

and showed that the nonexistence of an exceptional family is a sufficient condition for the existence of a solution to variational inequality VI(K, F). In [65], an exceptional family of elements for VI(K, F) is defined as follows.

Definition 2.3.5. Let $\hat{x}$ be an arbitrary feasible point, i.e., $\hat{x} \in K$. A sequence $\{x^r\}_{r\to\infty} \subset K$ is said to be an exceptional family of elements for VI(K, F) if the sequence satisfies following two conditions:

(i) $||x^r|| \to +\infty$ as $r \to +\infty$.

(ii) There exist two vector sequences $\{\lambda_r\} \subset R^m_+, \{\mu_r\} \subset R^l$ and a positive scalar sequence $\{\alpha_r\}$, where $\alpha_r > 1$ for all r, such that the sequence $\{\alpha_r x^r + (1 - \alpha_r)\hat{x}\} \subset K$, and the following equations hold for all r:

$$F(x^r) = -(\alpha_r - 1)(x^r - \hat{x}) - \frac{1}{2}\big[\nabla g(\alpha_r x^r + (1 - \alpha_r)\hat{x})^\intercal \lambda_r + \nabla h(\alpha_r x^r + (1 - \alpha_r)\hat{x})^\intercal \mu_r\big],$$

$$\lambda_r{}^\intercal g(\alpha_r x^r + (1 - \alpha_r)\hat{x}) = 0.$$

Among the most recent work on the exceptional family of elements, Han, Huang and Fang [19] give an alternative definition of exceptional family for VI(X, F) over a general unbounded closed convex subset of R^n and generalize all well-known existence results for variational inequalities. We restate their definition followed by their main existence results.

Definition 2.3.6. Let $\hat{x} \in R^n$ be any given point, i.e., $\hat{x} \in K$. A family of elements for $\{x^r\}_{r>0} \subset K$ is said to be an exceptional family for VI(K, F) with respect to $\hat{x}$ if:

(i) $||x^r|| \to +\infty$ as $r \to +\infty$.

(ii) for any $r > ||\Pi_K(\hat{x})||$, there exists a real number $\alpha_r > 0$ such that

$$-[F(x^r) + \alpha_r(x^r - \hat{x})] \in N(\hat{x}; K),$$

where $N(\hat{x}; K)$ denotes the normal cone of K at $\hat{x}$, i.e.,

$$N(\hat{x}; K) = \begin{cases} \{z \in R^n \mid z^\top(y - x^r) \leq 0,\ \forall y \in K\}, & \text{if } x^r \in K, \\ \emptyset, & \text{otherwise.} \end{cases}$$

Theorem 2.3.4. *Let K be a nonempty closed convex set in R^n, let $F : R^n \to R^n$ be a continuous function, and let $\hat{x} \in R^n$ be a vector. For any family of elements $\{x^k\} \subset K$ with $||x^k|| \to \infty$ as $k \to \infty$, if there exist $x^{k_0} \in \{x^k\}$ and $y \in K$ such that*

$$||y - \hat{x}|| < ||x^{k_0} - \hat{x}|| \text{ and } (x^{k_0} - y)^\top F(x^{k_0}) \geq 0, \tag{2.42}$$

then VI(K, F) has no exceptional family with respect to $\hat{x}$; hence, VI(K, F) has a solution. Furthermore, if (2.42) is satisfied as

$$||y - \hat{x}|| < ||x^{k_0} - \hat{x}|| \text{ and } (x^{k_0} - y)^\top F(x^{k_0}) > 0, \tag{2.43}$$

then the solution set of VI(K, F) is nonempty and compact.

Based on this result, they give both necessary and sufficient existence conditions for pseudo-monotone variational inequality problems and discuss the solvability of VI(K, F) with quasi-monotone functions.

2.4 Semi-infinite Variational Inequality Problems

With the purpose of generalizing the structure of the domain X, the semi-infinite variational inequality problems were recently introduced by Fang et al. [16]. They proposed an analytic center cutting plane algorithm for solving SIVI(X, F) and provided the complexity analysis of their algorithm.

Monotone variational inequality problems can be solved in polynomial time by ellipsoid algorithms [39] and in pseudo-polynomial time by certain path-following algorithms [43]. A polynomial time analytic center cutting plane method was introduced by Goffin et al. [27] for solving pseudo-monotone variational inequalities when X is a bounded polyhedron. Later, Nesterov and Vial [44] introduced a homogeneous analytic center cutting plane method with pseudo-polynomial time complexity for solving monotone variational inequalities embedded in a cone structure.

An analytic center cutting plane method (ACCPM) is an interior point method that generates a sequence of analytic centers which are supposed to converge towards a solution of the problem. For a polytope of the form

$$\Omega = \{x \in R^n \mid A^\intercal x \leq c\},$$

with the corresponding slack variables $s = c - A^\intercal x, s \in R^m$ and a full rank $m \times n$ matrix A, the *analytical center* of Ω is the unique solution of the problem

$$\max \left\{ \sum_{i=1}^{n} \ln s_i \mid A^\intercal x + s = c,\ s > 0 \right\}.$$

As a generalization of [27], Fang et al. [16] proposed an $O^*(n^2/\rho^2)$ analytic center cutting plane algorithm to solve semi-infinite variational inequalities under the assumptions that X has an nonempty interior, F is Lipschitz continuous and pseudo-co-coercive on X, where n is the dimension of X and ρ is the radius of a ball contained in the approximate solution set of SIVI(X, F). In a more recent work, Fang et al. [15] consider semi-infinite

variational inequality problem with a domain described by linear constraints and propose a discretization method followed by an analytic center based inexact cutting plane method.

Depending on the implementation of the analytic center cutting plane method for solving variational inequalities there may be centering issues and possibly non-convergence of the sequence of analytic centers, especially when X is not a polyhedral set. Additionally, there is considerable amount of computational effort required to find the analytic center and to evaluate the gap function as the stopping criteria.

Since the analytic center cutting plane approach is the only existing methodology that addresses semi-infinite variational inequalities, there is a need to study the underlying theory and solution structure and to explore other possible applications.

Chapter 3

Solution Analysis for Semi-infinite Variational Inequalities

An important distinctive feature of a semi-infinite variational inequality problem stems from the fact that its domain involves an infinite number of constraints. As a consequence, the solution analysis has to be done on a domain which need not behave as a manifold with boundary. We need to explore the structural similarity and differences between finite variational inequalities and semi-infinite variational inequalities. In this chapter, we present characterization and the solution analysis for semi-infinite variational inequalities.

This chapter consists of five sections. The first section analyzes the structure of the domain for semi-infinite variational inequalities. This is followed in Sections 3.2 and 3.3 by equivalent reformulations and existence results under appropriate conditions. Section 3.4 discusses constraint qualifications preceding the solution analysis in Section 3.5. Finally, Section 3.6 introduces dual forms of semi-infinite variational inequalities.

Given the continuous functions $F : R^n \to R^n$ and $g(x,t) : R^n \times R^l \to R^p$, we consider the following as the semi-infinite variational inequality problem:

$$\begin{aligned} \text{SIVI}(X,F): \quad & \text{Find a vector } x^* \in X \text{ such that} \\ & F(x^*)^\intercal (x - x^*) \geq 0 \ \text{ for all } x \in X, \end{aligned} \tag{3.1}$$

where

$$X = \{x \in R^n \mid \ g(x,t) \geq 0, \quad \text{for all} \ \ t \in T\}, \tag{3.2}$$

and $T \subset R^l$ is an index set that may contain infinitely many elements. For any $t \in T$, we assume that $g(x,t) : R^n \times R^l \to R^p$ is concave and C^k-continuous for some $k \geq 2$. Since g is concave, X is a convex set in R^n.

3.1 Structure of the Domain

The distinguishing feature of the semi-infinite variational inequality problem relative to the finite variational inequalities is the structure of its domain. For the latter, the feasible domain X (in Figure 2.2) is a topological manifold. However, as illustrated in Figure 2.1, when it consists of infinitely many constraints, the domain X of SIVI(X, F) is not necessarily a manifold with boundary. The well-known nonlinear programming methods can be extended to the semi-infinite variational inequality problem only if the set X can be described as a manifold with boundary. We will study the structure of X by considering its similarity to the feasible set in semi-infinite programming.

Let $\mathcal{E}(\bar{x}; X)$ denote the extremal set of X at a given point $\bar{x} \in X$, i.e.,

$$\mathcal{E}(\bar{x}; X) \equiv \{t \in T \mid g(\bar{x}, t) = 0\}.$$

Similar to the active set $\mathcal{I}_g(x)$ in semi-infinite programming, $\mathcal{E}(\bar{x}; X)$ is useful for locally describing the domain X as a manifold with boundary. The local boundary of X is characterized at the point $\bar{x}$ with $\mathcal{E}(\bar{x}; X) \neq \emptyset$.

Proposition 3.1.1. *The extremal set $\mathcal{E}(\bar{x}; X) \neq \emptyset$ if and only if the set T is compact, $g(\bar{x}, \cdot) \in C^1$ and*

$$\min_{t \in T} g(\bar{x}, t) = 0.$$

Proof. If $\mathcal{E}(\bar{x}; X) \neq \emptyset$, then, by definition, there exists a $\hat{t} \in T$ such that $g(\bar{x}, \hat{t}) = 0$. Since $g(x,t) \geq 0$ for all $t \in T$, with $g(\bar{x}, \hat{t}) = 0$ we have $\min_{t \in T} g(\bar{x}, t) = 0$. For the reverse implication, consider the unconstrained problem $\min_{t \in T} g(\bar{x}, t) = 0$. When the set T is compact and $g(\bar{x}, \cdot) \in C^1$, this unconstrained problem has a global minimum of zero. The result follows. □

Since infinitely many inequality constraints are involved in defining the domain, the extremal set of X at a particular point $\bar{x}$ can contain infinitely many elements. This can happen at $\bar{x}$ only if $\bar{t} \in \mathcal{E}(\bar{x}; X)$ is a degenerate minimizer of min $g(x, \cdot)$ over T.

Proposition 3.1.2. *If $\mathcal{E}(\bar{x}; X) \neq \emptyset$ and $Ord(\bar{t}) = 1$ for all $\bar{t} \in \mathcal{E}(\bar{x}; X)$, then the domain X of SIVI(X, F) can locally be reduced to a set described by a finite number of inequalities.*

Proof. By Definition 2.2.2, if $Ord(\bar{t}) = 1$ then

$$g(\bar{x}, \bar{t}) = \nabla_t g(\bar{x}, \bar{t}) = 0 \text{ and } \nabla_t^2 g(\bar{x}, \bar{t}) > 0.$$

The resulting gradient set is

$$S_1(\bar{t}) = \{\nabla_x g(\bar{x}, \bar{t})\}.$$

Since the dimension of *span* $\left(\bigcup_{t \in \mathcal{E}(\bar{x}, X)} S_1(t)\right)$ is finite and equal to $\sum_{t \in \mathcal{E}(\bar{x}, X)} Ord(t)$, the extremal set $\mathcal{E}(\bar{x}, X)$ consists of finite number of elements. □

When the conditions in Proposition 3.1.2 are satisfied, the local reduction is said to be applicable for semi-infinite variational inequality problem SIVI(X, F).

Proposition 3.1.3. *Assume that local reduction is applicable for SIVI(X, F). Then, the locally reduced domain at $\bar{x}$ is given by*

$$X_{red} = \{x \in \mathcal{N}(\bar{x}) \mid g(x, t_i(x)) \geq 0,\ i = 1, \ldots, m\},$$

where $m = |\mathcal{E}(\bar{x}, X)|$ and each $t_i(x), i = 1, \ldots, m$, is the local minimizer of $g(\bar{x}, \cdot)$ over T.

If the extremal set $\mathcal{E}(\bar{x}; X)$ has degenerate elements and the local reduction is not applicable, then following result helps to make the local description of the domain X.

Proposition 3.1.4. *If the index set T is a subset of R and the function $g(x,t) : R^n \times R^l \to R^p$ is regular (ref. Definition 2.2.4), then the domain $X \neq \emptyset$ of SIVI(X,F) is a topological manifold with boundary of dimension n.*

Proof. When $T \subset R$, the function $g(\bar{x}, \cdot)$ is univariate and the degeneracy of local minimizer t_j can be characterized using singularity theory and universal unfoldings [4]. Thus, by Theorem 2.2.2, the domain X, which is defined by a possibly infinite number of inequalities $g(x,t) \geq 0$, can locally be described as a topological manifold with boundary of dimension n. □

3.2 Equivalent Reformulations of Semi-infinite Variational Inequalities

As with ordinary variational inequalities, there are equivalent reformulations of a semi-infinite variational inequality problem. In this section, we consider projection and mathematical programming reformulations to attain insights for characterizing the solution of SIVI(X,F).

3.2.1 Projection Reformulation

We can reformulate a semi-infinite variational inequality problem as a classical fixed-point problem as follows.

Proposition 3.2.1. *Let X be a nonempty closed convex subset of R^n. Then x^* solves SIVI(X,F) if and only if*

$$x^* - \Pi_X(x^* - F(x^*)) = 0,$$

i.e., x^* is a fixed point of the mapping $H : R^n \to R^n$.

$$H(x) = \Pi_X(x - F(x)).$$

Proof. The defining condition of SIVI(X, F)

$$F(x)^\intercal(y - x) \geq 0 \ \ for\ all \ \ y \in X$$

can be restated as

$$(x - (x - F(x))^\intercal(y - x) \geq 0 \ \ for\ all \ \ y \in X.$$

From the projection property (b) (in Chapter 2) and the last inequality, it follows that

$$x = \Pi_X(x - F(x)) = 0. \qquad \square$$

Notice that the computation of the projection $\Pi_X(x - F(x))$ could be at least as hard to solve as solving the original problem SIVI(X, F). The main reason for this is the nonlinearity of the function $F(x)$ combined with quadratic nature of the projection problem. For the reduced problem SIVI(X_{red}, F), the projection $\Pi_{X_{red}}(x)$ can be computed by quadratic programming methods only if $F(x)$ is linear. Since linearity of F is too restrictive, the projection reformulation fails to contribute to the solution analysis of the problem.

3.2.2 Mathematical Programming Formulation with Merit Functions

In this section, we establish the connection between semi-infinite variational inequality problems and certain mathematical programming problems. Under appropriate conditions, the optimal solutions of some semi-infinite and nonlinear programming problems coincide with the solutions of the semi-infinite variational inequality problems.

Obviously, if the given function $F(x)$ is a gradient of a function $f(x)$, then the semi-infinite variational inequality problem is equivalent to the first order optimality conditions of a semi-infinite programming problem of minimizing $f(x)$ over the domain X. Moreover, when the local reduction is applicable this equivalent reformulation results in a nonlinear

programming problem. However, since in general F is not a gradient mapping, we need to use merit functions in order to obtain equivalent optimization reformulations.

The following mathematical programming formulation of a semi-infinite variational inequality is established by using the regularized gap function

$$\theta_\alpha(x) \equiv \max_{y \in X} \left\{ F(x)^\intercal (x - y) - \frac{\alpha}{2} ||x - y||^2 \right\}. \tag{3.3}$$

Theorem 3.2.1. *The vector $x^* \in X$ is a solution of SIVI(X, F) if and only if x^* solves*

$$\begin{aligned} &\textit{minimize} && \theta_\alpha(x) \\ &\textit{subject to} && x \in X. \end{aligned} \tag{3.4}$$

Proof. Using the explicit representation of X, the problem (3.4) can be rewritten as

$$\begin{aligned} &\text{minimize} && \theta_\alpha(x) \\ &\text{subject to} && g(x,t) \geq 0, t \in T \\ & && x \in R^n. \end{aligned} \tag{3.5}$$

Suppose x^* is an optimal solution to problem (3.5). Then, the optimality condition

$$\nabla \theta_\alpha(x^*)(x - x^*) \geq 0 \tag{3.6}$$

is satisfied for all $x \in X = \{x \in R^n \mid g(x,t) \geq 0 \text{ for all } t \in T\}$. By applying the chain rule for derivatives, we obtain

$$\nabla \theta_\alpha(x^*) = F(x^*) + \nabla F(x^*)(x^* - y_\alpha(x^*)) - \alpha(x^* - y_\alpha(x^*)),$$

where $y_\alpha(x^*) = \Pi_X(x^* - \frac{1}{\alpha} F(x^*))$ is the unique maximizer of $F(x^*)^\intercal (x^* - y) - \frac{\alpha}{2} ||x - y||^2$. Let $x^* = y_\alpha(x^*)$. Then by the fixed point theorem $\nabla \theta_\alpha(x^*) = F(x^*)$ and the optimality condition (3.6) is equivalent to SIVI(X, F).

Now, suppose x^* solves SIVI(X, F), i.e., $F(x^*)(x - x^*) \geq 0$. By Proposition 3.2.1, we have

$$x^* = \Pi_X(x^* - F(x^*)).$$

Thus, $x^* = y_\alpha(x^*)$ for $\alpha = 1$ and $F(x^*) = \nabla\theta_\alpha(x^*)$. As a solution of SIVI$(X, F)$, x^* satisfies the optimality condition for problem (3.5), and consequently for problem (3.4). □

The optimization problem in (3.5) is, in fact, a semi-infinite programming problem and the theory of semi-infinite optimization in Section 2.2 can be directly applied to obtain first and second order conditions.

Another differentiable optimization formulation of SIVI(X, F) is possible after implementing the local reduction to set X, if applicable. The following proposition establishes this result based on Theorem 3.2.1 and does not require proof.

Proposition 3.2.2. *Suppose local reduction is applicable at $\bar{x} \in X$ and the domain X can locally be reduced to $X_{red}(\bar{x})$. Then the equivalent nonlinear programming problem is as follows.*

$$\begin{aligned} &\text{minimize } \theta_\alpha^{red}(x) \\ &\textit{subject to } x \in X_{red}, \end{aligned} \tag{3.7}$$

where

$$\theta_\alpha^{red}(x) \equiv \max_{y \in X_{red}} \left\{ F(x)^\intercal (x - y) - \frac{\alpha}{2} ||x - y||^2 \right\}. \tag{3.8}$$

Note that if $F(x)$ is continuously differentiable, the finite optimization problem in (3.7) is continuously differentiable. Its first- and second-order optimality conditions are given in Sections 2.3 and 2.4, respectively.

3.3 Existence Results

The existence of solutions to a semi-infinite variational inequality problem depends on the properties of its domain and the function F. The well-known existence results for variational inequalities apply to SIVI(X, F). The solvability of finite dimensional variational

inequalities and consequently of semi-infinite variational inequalities is guaranteed by the following existence result in Hartman and Stampacchia [22].

Theorem 3.3.1. *Let X be a compact convex set in R^n and $F(x)$ a continuous function from R^n into itself. Then there exists a solution to variational inequality VI(X, F).*

Proof. The original proof for variational inequalities can be found in Hartman and Stampacchia [22]. For semi-infinite variational inequalities, recall the equivalent fixed point formulation of SIVI(X, F):

$$H(x) = \Pi_X(x - F(x))$$

and Brouwer's fixed point theorem, which states that every continous function has a fixed point in its domain. When $\bar{x}$ is a fixed point of $H(x)$, the result follows by Proposition 3.2.1. □

Following result adopts the existence result due to Han, Huang and Fang [19] under weaker assumptions.

Theorem 3.3.2. *Let X be a nonempty closed convex set in R^n, let F be continuous and $\hat{x} \in R^n$. If there exists a constant $\rho > 0$ such that, for any $x \in X$ with $||x - \hat{x}|| > \rho$, there exists $y \in X$ satisfying*

$$||y - \hat{x}|| < ||x - \hat{x}|| \text{ and } F^{\intercal}(x)(x - y) \geq 0,$$

then SIVI(X, F) has no exceptional family with respect to $\hat{x}$; hence it has a solution.

If the mapping F possesses certain properties, other existence and additionally uniqueness results can be established. The following results are adapted from those for variational inequalities given in Section 3.3.

Theorem 3.3.3. *Let $X \subseteq R^n$ be closed convex and let F be continuous. If F is strictly monotone on X, then the problem SIVI(X, F) has at most one solution.*

Proof. A similar proof by contradiction can be found in Pang [12]. Suppose F is strictly monotone and there exist two distinct solutions $\bar{x} \neq x_0 \in X$ to SIVI(X, F):

$$F(\bar{x})^\intercal(x - \bar{x}) \geq 0 \quad \text{and} \quad F(x_0)^\intercal(x - x_0) \geq 0 \ \ \forall x \in X.$$

Since $\bar{x} \neq x_0 \in X$, we have

$$F(\bar{x})^\intercal(x_0 - \bar{x}) \geq 0 \quad \text{and} \quad F(x_0)^\intercal(\bar{x} - x_0) \geq 0$$

which implies

$$[F(x_0) - F(\bar{x})]^\intercal(x_0 - \bar{x}) \leq 0.$$

Since the last inequality contradicts the strict monotonicity of F, the proof is complete. □

A Slater-type condition makes pseudo-monotonicity a sufficient condition for the existence of a solution to variational inequalities. Given a convex set X, the *dual cone* of X, denoted by X^*, is defined as

$$X^* = \{y \in R^n \mid y^\intercal x \geq 0 \ \ \forall x \in X\}.$$

Using the dual cone, Harker [20] gives the following result that applies to semi-infinite variational inequalities as well.

Theorem 3.3.4. *Let $X \subseteq R^n$ be closed convex and let F be continuous. If F is pseudo-monotone on X and there exists a vector $v \in X$ such that $F(v) \in int(X^*)$, then the solution set of the problem VI(X, F) is nonempty, convex and compact.*

3.4 Constraint Qualifications

The role of constraint qualifications for semi-infinite variational inequalities is important for describing the solution set and also for deriving necessary and sufficient conditions for a

solution to SIVI(X, F). The constraint qualifications given below are the similar those for semi-infinite programming.

Given a compact set $T \subset R^l$ and a function $g(x, t)$, recall the convex hull of the subdifferentials of the active constraints at $\bar{x}$

$$\mathcal{V}(\bar{x}) = \text{co} \left(\bigcup_{t \in \mathcal{E}(\bar{x}; X)} \partial g(\bar{x}, t) \right).$$

For $\bar{x} \in X$, we have

$$\mathcal{L}_0(\bar{x}; X) = \{\xi \in R^n \mid w^\intercal \xi > 0,\ w \in \mathcal{V}(\bar{x})\} \tag{3.9}$$

$$\mathcal{L}'(\bar{x}; X) = \{\xi \in R^n \mid w^\intercal \xi \geq 0,\ w \in \mathcal{V}(\bar{x})\} \tag{3.10}$$

$$\tilde{\mathcal{L}}(\bar{x}; X) = \{\xi \in R^n \mid \inf_{t \in \mathcal{E}(\bar{x}; X)} \nabla_x g(\bar{x}, t)^\intercal \xi \geq 0\}. \tag{3.11}$$

We can characterize the nondegenerate solutions to SIVI(X, F) with the following definition of nondegeneracy given by Dunn [11].

Definition 3.4.1. Let x be a solution to SIVI(X, F). Then, x is said to be *nondegenerate* if $-F(x)$ is in the relative interior of the normal cone, i.e., $-F(x) \in ri\ N(x; X)$, where

$$N(x; X) = \{d \in R^n \mid d^\intercal (y - x) \leq 0,\ \forall y \in X\}.$$

3.4.1 Extended Abadie Constraint Qualification

The extended Abadie constraint qualification (EACQ) is said to hold if the tangent cone at $\bar{x} \in X$ is equal to $\tilde{\mathcal{L}}(\bar{x}; X)$, i.e.,

$$\mathcal{T}(\bar{x}; X) = \tilde{\mathcal{L}}(\bar{x}; X).$$

An important consequence of EACQ is related to the nondegeneracy of the solution to SIVI(X, F).

Proposition 3.4.1. *If EACQ holds at the solution $\bar{x}$ of SIVI(X, F), then $\bar{x}$ is nondegenerate.*

Proof. Since $\bar{x}$ is a solution to SIVI(X, F), $0 \in F(\bar{x}) + N(\bar{x}; X)$. We know that

$$-F(\bar{x}) \in ri\ N(\bar{x}; X) \iff F(\bar{x})^{\mathsf{T}}\xi > 0 \quad \forall \xi \in \mathcal{T}(\bar{x}; X), \xi \neq 0.$$

Under EACQ, the tangent cone is a polyhedral cone generated the by active constraint gradients such that

$$F(\bar{x}) = \sum_{t \in \mathcal{E}(\bar{x}; X)} \lambda_t \nabla_x g(\bar{x}, t) \text{ for some } \lambda_t > 0.$$

Thus, the result follows. □

3.4.2 Extended Linear Independence Constraint Qualification

The extended linear independence (ELICQ) is said to hold at $\bar{x} \in X$ if the active constraint gradients $\nabla_x g(\bar{x}, t),\ t \in \mathcal{E}(\bar{x}; X)$ are linearly independent. The ELICQ has important consequences for SIVI(X, F) as given by the following results.

Proposition 3.4.2. *If ELICQ is satisfied, then the extremal set $\mathcal{E}(\bar{x}; X)$ is finite.*

Proof. If ELICQ holds, by definition the extremal set cannot contain more than n points.

□

Proposition 3.4.3. *Suppose ELICQ holds at $\bar{x} \in X$. If the solution $\bar{x}$ of SIVI(X, F) is nondegenerate, then there exists a unique set of multipliers $\lambda_t > 0, t \in \mathcal{E}(\bar{x}; X)$ such that*

$$F(\bar{x}) - \sum_{t \in \mathcal{E}(\bar{x}; X)} \lambda_t \nabla_x g(x, t) = 0.$$

Proof. When ELICQ satisfied, by Proposition 3.4.2 and the nondegeneracy property of the point $\bar{x}$, it follows that the set of multipliers satisfying $F(\bar{x}) - \sum_{t \in \mathcal{E}(\bar{x}; X)} \lambda_t \nabla_x g(x, t) = 0$ is unique and positive. □

The last result is especially important in deriving the first-order conditions for SIVI(X, F). As stated in Section 3.2.2, the ELICQ implies the following well-known constraint qualification.

3.4.3 Extended Mangasarian Fromowitz Constraint Qualification

The EMFCQ is said to hold at $\bar{x} \in X$ if $\mathcal{L}_0(\bar{x}; X) \neq \emptyset$. When the function g is continuously differentiable and the set X is finitely representable, EMFCQ holds at $\bar{x} \in X$ if there exists a vector $\xi \in R^n$ such that

$$\nabla_x g(\bar{x}, t)^\intercal \xi > 0, \quad \text{for all } t \in \mathcal{E}(\bar{x}; X). \tag{3.12}$$

Note that if X is convex, EMFCQ holding at any point in X is equivalent to the existence of a vector $\hat{x}$ satisfying $g(\hat{x}) > 0$, which is known as Slater's condition. Thus, if EMFCQ is satisfied at a given point of a convex set X, then EMFCQ holds at every point of X. The subsequent results are critical in the solution analysis of semi-infinite variational inequalities.

Proposition 3.4.4. *Assume EMFCQ holds at all points of X. Then X is a Lipschitzian manifold with boundary of $(n - |\mathcal{E}(\bar{x}; X)|)$:*

$$bdry(X) = \{x \in R^n \mid \min_{t \in T} g(\bar{x}, t) = 0\}.$$

Proof. The result follows from Theorem 2.2.2. □

Proposition 3.4.5. *If EMFCQ holds and $\bar{x}$ solves SIVI(X, F). Then, the multiplier set*

$$\Lambda(\bar{x}) \equiv \left\{ \lambda \in R_+^{|\mathcal{E}(\bar{x};X)|} \mid F(\bar{x}) - \sum_{t \in \mathcal{E}(\bar{x};X)} \lambda_t \nabla_x g(\bar{x}, t) = 0, \ \lambda_t g(\bar{x}, t) = 0, t \in \mathcal{E}(\bar{x}; X) \right\}$$

is nonempty and bounded.

Proof. If EMFCQ holds, then the following system of inequalities is solvable:

$$\nabla_x g(\bar{x}, t)^\intercal \lambda > 0, \quad t \in \mathcal{E}(\bar{x}; X).$$

Since $\bar{x}$ solves SIVI(X, F), as a minimizer of the corresponding nonlinear program it satisfies the first-order optimality conditions. Thus, $\Lambda(\bar{x}) \neq \emptyset$. □

Additional results related to the multiplier set $\Lambda(\bar{x})$ and the EMFCQ requirements can be found in the next sections.

3.4.4 Extended Kuhn-Tucker Constraint Qualification

The extended Kuhn-Tucker constraint qualification (EKTCQ) is said to hold at $\bar{x} \in X$ if $\mathcal{L}(\bar{x}; X) = cl\ A(\bar{x}; X)$, i.e., all the attainable directions are contained in the cone

$$\{\xi \in R^n \mid w^\intercal \xi \geq 0,\ w \in \mathcal{V}(\bar{x})\}.$$

If $\mathcal{V}(\bar{x})$ is closed and EKTCQ holds at $\bar{x} \in X$, then results equivalent to EMFCQ results can be obtained.

3.5 Solution Analysis

In this section, we establish the necessary and sufficient conditions for solutions to SIVI(X, F).

Theorem 3.5.1. *(Necessary Condition) Given a C^2-continuous function $g(x,t)$, let EMFCQ be satisfied at $x^* \in X$. If x^* is a solution to problem SIVI(X,F), then there exists a vector $\lambda \in R^m$ such that*

$$\begin{aligned} F(x^*) - \textstyle\sum_{j=1}^{m} \lambda_j \nabla_x g(x^*, t_j) = 0, \\ \lambda_j \geq 0, \quad j = 1, \ldots, m. \end{aligned} \tag{3.13}$$

Proof. By Theorem 3.2.1, if x^* solves SIVI(X, F), then it also solves the following optimization problem

$$\begin{aligned} \text{minimize} \quad & \gamma(x) = F(x^*)^\intercal (x - x^*) \\ \text{subject to} \quad & g(x,t) \geq 0, \text{ for all } t \in T \\ & x \in R^n. \end{aligned} \tag{3.14}$$

Since EMFCQ holds, the feasible set of (3.14) is a Lipschitzian manifold with boundary and the extremal set has finite number of elements, i.e., $\mathcal{E}(x^*; X) = \{t_1, \ldots, t_j, \ldots, t_m\}$. As

a solution of (3.14), by Theorem 2.2.5 x^* satisfies the following optimality conditions:

$$\begin{aligned} \nabla\gamma(x^*) - \textstyle\sum_{j=1}^{m} \lambda_j \nabla_x g(x^*, t_j) &= 0, \\ \lambda_j g(x^*, t_j) &= 0 \quad j = 1, \ldots, m, \\ \lambda_j \geq 0,\ g(x^*, t_j) &\geq 0 \quad j = 1, \ldots, m. \end{aligned} \tag{3.15}$$

Since $g(x^*, t) = 0$ for $t \in \mathcal{E}(x^*; X)$ and $\nabla\gamma(x^*) = F(x^*)$, the proof is complete.

□

Theorem 3.5.2. *(Sufficient Condition) Suppose the EMFCQ holds at $x^* \in X$. Given a C^2-continuous quasi-concave function $g(x, t)$, if there exists a vector $\lambda \in R^m$ satisfying (3.13), then x^* is a solution to problem SIVI(X, F).*

Proof. Given that the function $g(x, t)$ is quasi-concave in x and EMFCQ satisfied, a pair (x^*, λ) satisfying (3.13) is an optimal solution to problem (3.14). Thus,

$$0 = F(x^*)^\intercal (x^* - x^*) \leq F(x^*)^\intercal (x - x^*), \quad \text{for all } x \in X$$

and x^* is a solution to problem SIVI(X, F). □

Next, based on the results for semi-infinite programming [51], we characterize the solution set of SIVI(X, F) using Borel sets. For necessary and sufficient conditions defined over Borel sets, we need to specify the choice of spaces for the primal and dual problems. Let $C(T)$ denote the Banach space of continuous functions $g : R^n \times T \to R^p$, and let $K \equiv \{g \in C(T) \mid g(\cdot, t) \geq 0, \text{ for } t \in T\}$ be the cone of nonnegative valued continuous functions. Then, the primal variable space for the semi-infinite variational inequality problem is R^n. The dual variable space $\mathcal{B}$ of $C(T)$ is the space of finite signed Borel measures on T, with the norm given by the variation of the corresponding measure and, for $g \in C(T)$ and $\mu \in \mathcal{B}$,

$$\langle \mu, g \rangle = \int_T g(\cdot, t)\mu(t)dt.$$

Let supp(μ) denote the support of measure μ such that

$$\text{supp}(\mu) \subset \mathcal{E}(x; X).$$

Then, we have following necessary and sufficient conditions for SIVI(X, F).

Theorem 3.5.3. *(Necessary Condition) Suppose EMFCQ holds at point $x^* \in X$. Given that $g(x,t)$ is a continuously differentiable function, if x^* is a solution to problem SIVI(X,F), then there exists a nonnegative Borel measure μ on T such that*

$$F(x^*) - \int_T \nabla_x g(x^*, t)\mu(t)dt = 0, \ \ supp(\mu) \subset \mathcal{E}(x^*, X), \tag{3.16}$$

and the set $\mathcal{B}$ of such nonnegative Borel measures μ, satisfying condition (3.16), is bounded.

Proof. If x^* solves SIVI(X, F), then it also solves the following semi-infinite programming problem

$$\begin{aligned} \text{minimize} \quad & \gamma(x) = F(x^*)^\intercal (x - x^*) \\ \text{subject to} \quad & g(x,t) \geq 0, \quad \text{for all } t \in T \\ & x \in R^n. \end{aligned} \tag{3.17}$$

Then, by Theorem 2.2.7, x^* satisfies the optimality conditions for (3.17) and the result follows. □

Theorem 3.5.4. *(Sufficient Condition) Suppose EMFCQ holds at $x^* \in int(X)$. Given a continuously differentiable and quasi-concave function $g(x,t)$, if there exists a nonnegative Borel measure μ on T satisfying (3.16), then x^* is a solution to problem SIVI(X,F).*

Proof. Under the given assumptions of quasi-concavity of g and nonempty interior of X, conditions (3.16) are sufficient for x^* to be an optimal solution to (3.17). Thus,

$$0 = F(x^*)^\intercal (x^* - x^*) \leq F(x^*)^\intercal (x - x^*), \quad \text{for all } x \in X$$

and x^* is a solution to problem SIVI(X, F).

□

Note that if X has a nonempty interior and the set of Borel measures satisfying (3.16) is nonempty and bounded, then the EMFCQ holds at a given point.

Suppose that μ in (3.16) is a discrete Borel measure having an infinite support, i.e.,

$$\mu = \sum_{i=1} \lambda_i \delta(t_i)$$

with $\delta(t)$ being the Dirac measure. Then the total variation of μ is $\sum_{i=1} |\lambda_i|$. Note that this makes the index set T be countably infinite. Then, the necessary conditions (3.16) in Theorem 3.5.3 can be rewritten for this discrete measure as follows:

$$\begin{aligned} F(x^*) - \textstyle\sum_{i=1} \lambda_i \nabla_x g(x^*, t_i) = 0, \\ \lambda_i > 0, \ i = 1, 2 \ldots \\ t_i \in \mathcal{E}(x^*; X). \end{aligned} \tag{3.18}$$

Provided that the set of Borel measures satisfying (3.16) is nonempty, Shapiro [51] showed that extreme points (measures) are discrete with support points less than or equal to $n+1$. Consequently, when $m \leq n+1$, there exists a discrete measure $\mu = \sum_{i=1}^{m} \lambda_i \delta(t_i)$ satisfying (3.18). If $m > n$, then the corresponding vectors $\nabla_x g(x^*, t), i = 1, \ldots, m$ are linearly independent and hence such a discrete measure can always be found. This fact motivates us to develop an efficient discretization-based method.

When the local reduction is applicable for the domain X, the necessary and sufficient conditions for SIVI(X, F) can be derived through the equivalent nonlinear programming formulation with ordinary constraint qualifications.

Based on the necessary and sufficient conditions above, we use discretization and exchange methods to approximate the solution for SIVI(X, F) since we cannot know which constraints are active a priori.

3.6 Duality

In this section, we derive Lagrangian dual of the semi-infinite variational inequality problem by using its equivalent mathematical programming reformulation.

The point x^* is a solution of SIVI(X, F) if and only if it solves the following optimization problem:

$$\begin{aligned} &\text{minimize} && F(x^*)^\intercal(x - x^*) \\ &\text{subject to} && x \in X, \end{aligned} \tag{3.19}$$

where $X = \{x \in R^n \mid g(x,t) \geq 0, \quad \text{for all } t \in T\}$.

With EMFCQ satisfied at x^* and given the Lagrangian

$$L(x, \mu; x^*) = F(x^*)^\intercal(x - x^*) - \int_T g(x,t)\mu(t)dt, \tag{3.20}$$

we can write the dual of problem (3.19) as

$$\text{(D)} \qquad \sup_{\mu \in \mathcal{B}} \left\{ \inf_{x \in R^n} \{L(x, \mu; x^*) \right\}. \tag{3.21}$$

Given the set of nonnegative Borel measures satisfying (3.16), the dual semi-infinite variational inequality (DSIVI) problem can be written with both discrete and continuous measure.

3.6.1 Dual with Continuous Borel Measure

Let $\mathcal{B}(x^*)$ be the set of nonempty, nonnegative Borel measures satisfying (3.16). Then the dual of problem (3.19) is

$$\text{(D)} \qquad \sup_{\mu \in \mathcal{B}(x^*)} \left\{ \inf_{x \in R^n} \left[F(x^*)^\intercal(x - x^*) - \int_T g(x,t)\mu(t)dt \right] \right\}. \tag{3.22}$$

Now, for each $\mu \in \mathcal{B}(x^*)$, define

$$S(\mu) \equiv \left\{ x \in R^n \mid F(x) - \int_T \nabla_x g(x,t)\mu(t)dt = 0 \right\}.$$

Let $G(\mu)$ denote the Banach space of nonnegative valued continuous functions $g : S(\mu) \times T \to R^p$. Then, we obtain the following definition of the dual semi-infinite variational inequality problem based on the duality scheme of Auslender and Teboulle [2].

Definition 3.6.1. The dual semi-infinite variational inequality problem (DSIVI) is to find a vector $\mu^* \in \mathcal{B}(x^*)$ such that

$$\langle y^*, \mu^* - \mu \rangle \geq 0 \quad y^* \in G(\mu^*),\ \forall \mu \in \mathcal{B}(x^*). \tag{3.23}$$

3.6.2 Dual with Discrete Borel Measure

By the standard optimality theorems, we know that (x^*, λ^*) is a saddle point of L if and only if x^* solves (3.19) optimally and λ^* solves (3.21) optimally. When necessary and sufficient conditions given in Theorems 3.5.3 and 3.5.4 are satisfied for a discrete Borel measure, there exists a vector λ^* that can be interpreted as the solution of a Lagrangian dual SIVI(X, F), which is defined next.

For each $\lambda \geq 0$, set

$$M(\lambda) \equiv \{x \in R^n \mid F(x) - \sum_{i=1}^{m} \lambda_i \nabla_x g(x, t_i) = 0\ t_i \in \mathcal{E}(x; X)\}.$$

Then, the following definition of Lagrangian dual corresponds to the first order necessary conditions for SIVI(X, F).

Definition 3.6.2. Dual semi-infinite variational inequality problem (DSIVI) is to find a vector $\lambda^* \in R^m_+$ such that

$$\langle y^*, \lambda^* - \lambda \rangle \geq 0 \quad \text{for all } \lambda \in R^m_+, \tag{3.24}$$

where $y^* = \begin{pmatrix} g(x,t_1) \\ g(x,t_2) \\ \vdots \\ g(x,t_m) \end{pmatrix}$ with $t_i \in \mathcal{E}(x,X)$ and $x \in M(\lambda^*)$.

Since the dual problem DSIVI is another representation of the necessary conditions (3.15), we have the following result.

Theorem 3.6.1. *Suppose EMFCQ holds for the set X. If x^* solves the problem SIVI(X,F) then there exists a vector $\lambda^* \in R^m_+$ such that (x^*,λ^*) solves DSIVI.*

Proof. By Proposition 3.4.12, if x^* solves the problem SIVI(X,F) and EMFCQ is satisfied, then the multiplier set $\Lambda(x^*)$ is nonempty and bounded. Since $M(\lambda) \subseteq \Lambda(x^*)$, $M(\lambda)$ is bounded and the inequality (3.24) is equivalent to (3.15). Hence, if x^* solves SIVI(X,F), then by Theorem 3.5.1, there exists a vector $\lambda^* \in \Lambda(x^*)$ that satisfies (3.15) and hence (x^*,λ^*) solves DSIVI. □

Chapter 4

Solution Methods

Based on the solution analysis in Chapter 3, the semi-infinite variational inequality problem SIVI(X, F) can be approached by different solution techniques based on projection methods, merit functions and necessary and sufficient first-order conditions. There are two possible ways to handle the challenge introduced by infinitely many inequalities for solving semi-infinite variational inequalities: a static approach where all the inequalities are considered at one time after reduction, and a progressive approach where a few chosen inequalities are considered at a time. For any system of infinitely many inequalities, the discretization method is a static but rather straight forward approach whereas exchange and cutting place methods are dynamic but more involved approaches.

In this chapter, we propose three solution methodologies including a *discretization-based smoothing method*, an *exchange method* and an *entropic analytic center cutting plane method*. The proposed algorithms are tested on both linear and nonlinear problems. Computational results comparing the performance of the algorithms are also reported.

Starting with a coarse grid at iteration $i = 1$, we discretize the index set T to T_i such that $|T_i| < \infty$. With approximating the infinitely constrained set X by a series of finitely constrained set $X_i \supset X$, the first algorithm successively solves the variational inequality problem VI(X_i, F) by refining the grid at each iteration to approximate the solution of

SIVI(X, F). Instead of using fixed grid size to determine the number of constraints defining the domain X_i at each iteration, the proposed exchange algorithm progressively adds constraints to X_i until the solution to VI(X_i, F) solves SIVI(X, F). For each VI(X_i, F), its equivalent KKT conditions, which are nondifferentiable, is smoothed and solved by Newton's method. A new smoothing function to approximate $\max\{0, x\}$ is designed to serve this purpose. Note that both algorithms use the actual constraints defining the domain X to describe X_i. The entropic analytic center cutting plane method, on the other hand, starts with a polyhedron which encloses the domain X and gradually generates linear cuts to provide the outer approximation to X. We modify the analytic center cutting plane method in [16] by replacing the analytic center by entropic analytic center.

This chapter is organized as follows. Section 4.1 introduces SIVI(X, F) and a collection of associated subproblems VI(X_i, F), which are defined on set X_i. Section 4.2 describes our solution approach for solving each VI(X_i, F). The discretization-based smoothing method is presented in Section 4.3, which is followed by exchange method in Section 4.4 and entropic analytic center cutting plane method in Section 4.5. Finally, the numerical examples and computational results are reported in Section 4.6.

4.1 SIVI(X, F) and Subproblems

Recall the definition of a semi-infinite variational inequality problem:

$$\begin{aligned} \text{SIVI}(X, F): \qquad & \text{Find a vector } x^* \in X \text{ such that} \\ & F(x^*)^\intercal (x - x^*) \geq 0 \qquad \text{for all } x \in X. \end{aligned} \tag{4.1}$$

where

$$X = \{x \in R^n \mid \ g(x, t) \geq 0, \quad \text{for all} \ \ t \in T\}. \tag{4.2}$$

We assume that $|T| = \infty$ and $g(x, t) : R^n \times R^l \to R^p$ is twice continuously differentiable and concave in x. Therefore, X is implicitly assumed to be a convex set in R^n. We further

assume that X is nonempty and bounded.

When X is described by a sequence of finitely constrained sets X_i, SIVI(X, F) can be considered as a collection of finite variational inequalities VI(X_i, F). The proposed discretization-based and exchange methods involve approximating the solution of SIVI(X, F) through a sequence of solutions to VI(X_i, F). Therefore, we need first to address the new solution technique developed for solving VI(X_i, F).

4.2 Solving VI(X_i, F)

We consider the following variational inequality problem

$$\begin{aligned} \text{VI}(X_i, F) \ : \quad & \text{find } x^i \in X_i \text{ such that} \\ & F(x^i)^T(x - x^i) \geq 0, \quad \text{for all } x \in X_i, \end{aligned} \tag{4.3}$$

where $X \subset X_i$ is finitely constrained. Given that F is continuous and X_i is nonempty closed convex in R^n with no exceptional family, the solution to (4.3) is guaranteed to exist by Theorem 2.3.4.

The two basic approaches for solving VI(X_i, F) are solving its equivalent KKT conditions and using the direct methods. The basic idea of the direct methods is to find a sequence $\{x^k\}$ in X_i such that each x^{k+1} solves VI(X_i, F^k), where $F^k(x)$ is an approximation to $F(x)$. Based on this approximation, a Newton-based iterative algorithms can be found in [47]. Another direct method is the use of merit functions. Although the use of merit functions is theoretically sound, severe drawbacks arise in the evaluation of merit functions. As long as the function F is defined everywhere, methods based on the KKT conditions offer a convenient approach mainly because KKT conditions can easily be reformulated as a mixed complementarity problem. This allows the implementation of equation-based, interior point and smoothing-based solution algorithms, which usually possess sharp convergence results. The smoothing methods and their fast convergence results can be found in Chen, Qi and

Sun's work [7].

Our solution approach to VI(X_i, F) is based on the min reformulation of its KKT conditions. By smoothing the min function via smooth approximation of the max operator, we reformulate VI(X, F) as a smooth system of equations. Thus, the reformulated smoothed problem possesses good potential to satisfy constraint qualifications, allowing the direct application of existing nonlinear programming algorithms. The smoothing approach has been shown to have superior performance for solving mixed complementarity problems in the numerical results of Billups, Dirske and Ferris [3].

The following result is based on Proposition 2.3.3 and states that solving KKT conditions for VI(X_i, F) and solving problem VI(X_i, F) are equivalent.

Proposition 4.2.1. *If $g : R^n \times R^l \to R^p$ is twice continuously differentiable and the MFCQ is satisfied, then solving the problem VI(X_i, F) is equivalent to solving its Lagrangian as described below.*

(i) If x^* solves VI(X_i, F) and constraint qualification holds for set X_i at the point x^*, then there exists a vector $\lambda^* \in R^m, m = p \times |T_i|$, such that

$$\begin{aligned} F(x^*) - \sum_{j=1}^{m} \lambda_j^* \nabla g_j(x^*, t) &= 0, \\ \lambda_j^* g_j(x^*, t) &= 0 \quad j = 1, \ldots, m, \\ \lambda_j^* \geq 0,\ g_j(x^*, t) &\geq 0 \quad j = 1, \ldots, m. \end{aligned} \tag{4.4}$$

(ii) Conversely, if $g(x, t)$ is concave in x and if (x^*, λ^*) satisfies (4.4), then x^* solves VI(X_i, F).

The Lagrangian function $L : R^n \times R^m \to R^n$ of the VI(X_i, F) in the vector format can be written as

$$L(x, \lambda) \equiv F(x) - \nabla_x g(x, t)^\intercal \lambda.$$

Note that (4.4) is a mixed nonlinear complementarity problem (MiCP) that can be reformulated as a system of nonsmooth equations. For nonsmooth reformulations of the complementarity problems, there are two main solution techniques. First and the most common one is to use the Fisher-Burmeister function $\psi(a,b) \equiv \sqrt{a^2+b^2} - (a+b)$ as studied in [49]. The other one is to use smoothing functions for the min-reformulation. The min-reformulation of complementarity problems allows using a locally fast converging algorithm that requires solving a system of linear equations of reduced dimension at each iteration. Moreover, the convergence of such algorithms can be established under weaker assumptions than the ones used in the Fisher-Burmeister reformulation. Also, in the case of linear complementarity, finite convergence is attainable with min-reformulation but not with the Fisher-Burmeister reformulation [12].

Hence, mainly for reduced dimensionality and better convergence, we implement the min-reformulation of the MiCP first by rewriting the KKT conditions as follows:

$$\begin{aligned} L(x,\lambda) &= 0, \\ g(x,t) - y &= 0, \\ \lambda \geq 0,\ y \geq 0,\ \lambda_i y_i &= 0. \end{aligned} \tag{4.5}$$

The nonnegativity and orthogonality conditions in (4.5) are equivalent to having the minimum of y and λ to be zero. Therefore, we have the following nonsmooth, nonlinear operator $\Phi_{\min}(x,\lambda,y)$ to represent the KKT conditions of VI(X_i, F):

$$\Phi_{\min}(x,\lambda,y) \equiv \begin{pmatrix} L(x,\lambda) \\ g(x,t) - y \\ \min(\lambda, y) \end{pmatrix} = 0, \qquad (x,\lambda,y) \in R^{n+m+m}. \tag{4.6}$$

Note that $\Phi_{\min}$ is semi-smooth and its limiting Jacobian has an exact expression [12].

The semi-smooth Newton method might be suitable for solving the above system of equations; however, the calculation and the choice of generalized Jacobians bring some drawbacks. Therefore, the difficulty due to the non-differentiability of the min operator in (4.6) will be overcome by smoothing methods.

4.2.1 Smoothing Method

Smoothing functions are available for the scalar plus function $b_+ \equiv \max(b, 0)$, $b \in R$. Since $\min(a, b) = b - (b - a)_+$, any smoothing function for the plus function defines a smoothing function for the min operator. Several smoothing functions including Fisher-Burmeister, Chen-Harker-Kanzow-Smale (CHKS), Log-exponential, Pinar-Zenios, and Zang have been developed and used as the max operator. Consider the Pinar-Zenios function, $\hat{p}_\varepsilon(x)$, with uniform kernel on interval [0,1]

$$\hat{p}_\varepsilon(x) = \begin{cases} 0 & \text{if } x < 0 \\ \frac{x^2}{2\varepsilon} & \text{if } 0 \le x \le \varepsilon \\ x - \frac{\varepsilon}{2} & \text{if } x > \varepsilon \end{cases}$$

and Zang function, $\tilde{p}_\varepsilon(x)$, with uniform kernel on interval $[-\frac{1}{2}, \frac{1}{2}]$,

$$\tilde{p}_\varepsilon(x) = \begin{cases} 0 & \text{if } x < -\frac{\varepsilon}{2} \\ \frac{1}{2\varepsilon}(x + \frac{\varepsilon}{2})^2 & \text{if } |x| \le \frac{\varepsilon}{2} \\ x & \text{if } x > \frac{\varepsilon}{2}. \end{cases}$$

Smoothing functions like Pinar-Zenios and Zang, whose kernel functions have finite support, often result in singularity which makes the implementation of Newton-based solution algorithms unsuitable. In contrast, smoothing functions with infinite kernel supports result in nonsingular Jacobians and thus allow the implementation of Newton-based algorithms with fast convergence rates. For this reason, Fisher Burmeister, CHKS and the

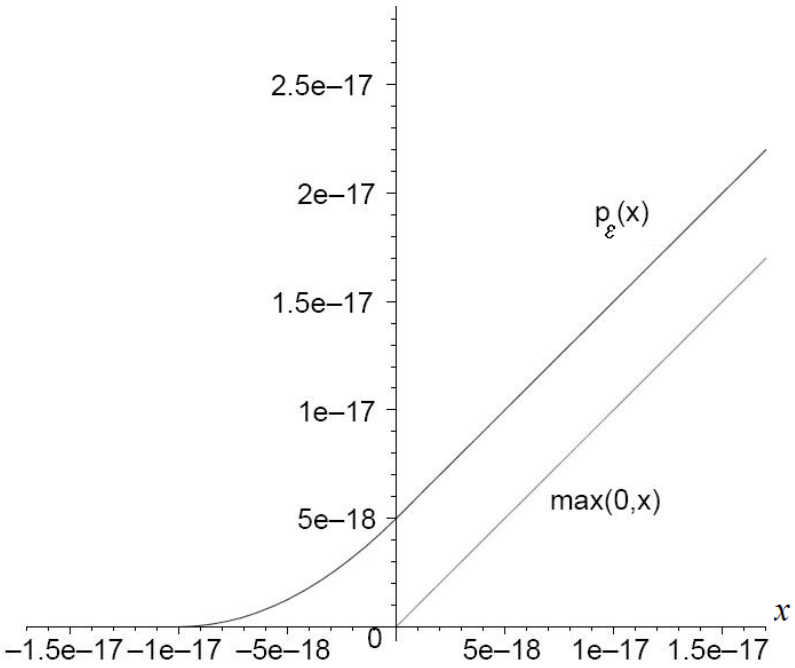

Figure 4.1: Max-plus function and smooth approximation $p_\varepsilon(x)$ with $\varepsilon = 10^{-18}$

log-exponential smoothing functions with infinite support are the ones most commonly used. The Chen-Harker-Kanzow-Smale (CHKS) function, $\bar{p}_\varepsilon(x)$, is defined as

$$\bar{p}_\varepsilon(x) = \frac{\sqrt{4\varepsilon^2 + x^2} + x}{2},$$

and the log-exponential function, $\breve{p}_\varepsilon(x)$, is given by

$$\breve{p}_\varepsilon(x) = x + \varepsilon \log(1 + e^{-x/\varepsilon}).$$

Another finite support smoothing function is the Wu-Sun smoothing function [60], which utilizes an additional smoothing parameter q as a function of the Lagrange multipliers of the corresponding optimization problem:

$$p_{\varepsilon,q}(x) = \begin{cases} 0 & \text{if } \; x < -\frac{\varepsilon}{q} \\ \frac{qx^2}{2\varepsilon} + x + \frac{\varepsilon}{2q} & \text{if } \; -\frac{\varepsilon}{q} \leq x \leq 0 \\ x + \frac{\varepsilon}{2q} & \text{if } \; x > 0 \end{cases} \tag{4.7}$$

In order to guarantee the nonsingularity property, we perturb the finite support smoothing function in (4.7) and use the following function as our smoothing function:

$$p_\varepsilon(x) = \begin{cases} \varepsilon x + \varepsilon^2, & \text{if } \ x \le -\varepsilon \\ (\frac{1}{2\varepsilon} - 1)x^2 + (1-\varepsilon)x + \frac{\varepsilon}{2}, & \text{if } \ -\varepsilon < x < 0 \\ (1-\varepsilon)x + \frac{\varepsilon}{2}, & \text{if } \ x \ge 0 \end{cases} \tag{4.8}$$

With $p_\varepsilon(x)$ being the smoothing function of the plus function as illustrated in Figure 4.1, we can easily obtain a smooth approximation of the min formulation in (4.6). Specifically, for given $\varepsilon > 0$, define

$$\Phi_\varepsilon(z) \equiv \Phi_\varepsilon(x, \lambda, y) \equiv \begin{pmatrix} L(x,\lambda) \\ g(x,t) - y \\ \lambda - p_\varepsilon(\lambda - y) \end{pmatrix}, \ \forall z \in R^{n+m+m}. \tag{4.9}$$

Theorem 4.2.1. *If $F : R^n \to R^n$ is C^1 and $g : R^n \times R^l \to R^p$ is C^2 continuous with $\nabla L(x,\lambda)$ being positive definite, then the Jacobian of $\Phi_\varepsilon(z)$, denoted by $J\Phi_\varepsilon(z)$, is nonsingular for any $\varepsilon > 0$ and $z = (x, \lambda, y)$.*

Proof. Since the functions p_ε and g are at least C^2 continuous, the operator Φ_ε is at least C^2 continuous and its Jacobian is given by

$$J\Phi_\varepsilon(z) = \begin{pmatrix} \nabla L & -\nabla_x g(x,t)^\intercal & 0 \\ \nabla_x g(x,t) & 0 & -I \\ 0 & D_\lambda & D_y \end{pmatrix},$$

where

$$\nabla L = \nabla F(x) - \sum_{i=1}^{m} \lambda_i \nabla_x^2 g_i(x,t),$$

$$D_\lambda := diag\left(\frac{\partial}{\partial\lambda_1}(\lambda_1 - p_\varepsilon(\lambda_1 - y_1)), \ldots, \frac{\partial}{\partial\lambda_m}(\lambda_m - p_\varepsilon(\lambda_m - y_m))\right),$$
$$D_y := diag\left(\frac{\partial}{\partial y_1}(\lambda_1 - p_\varepsilon(\lambda_1 - y_1)), \ldots, \frac{\partial}{\partial y_m}(\lambda_m - p_\varepsilon(\lambda_m - y_m))\right).$$

For a given $\varepsilon > 0$, the square matrix D_λ is nonsingular and positive definite since its diagonal elements are calculated by

$$\frac{\partial}{\partial\lambda_i}(\lambda_i - p_\varepsilon(\lambda_i - y_i)) = \begin{cases} 1-\varepsilon & \text{if } \lambda_i - y_i \le -\varepsilon \\ -(\frac{1}{\varepsilon} - 2)(\lambda_i - y_i) + \varepsilon & \text{if } -\varepsilon < \lambda_i - y_i < 0 \\ \varepsilon & \text{if } \lambda_i - y_i \ge 0. \end{cases}$$

Similarly, the square matrix D_y is nonsingular since all of its diagonal elements are nonzero (strictly positive for $\varepsilon > 0$) as it can be seen in

$$\frac{\partial}{\partial y_i}(\lambda_i - p_\varepsilon(\lambda_i - y_i)) = \begin{cases} \varepsilon & \text{if } \lambda_i - y_i \le -\varepsilon \\ (\frac{1}{\varepsilon} - 2)(\lambda_i - y_i) + 1 - \varepsilon & \text{if } -\varepsilon < \lambda_i - y_i < 0 \\ 1-\varepsilon & \text{if } \lambda_i - y_i \ge 0 \end{cases}$$

The nonsingularity of the matrix $J\Phi_\varepsilon(z)$ will be shown by proving that the only solution to the system $J\Phi_\varepsilon(z)v = 0$ is the trivial solution. Let $v = (v^{(1)}, v^{(2)}, v^{(3)})$ be a vector in $R^n \times R^m \times R^m$. Then we have

$$\nabla L v^{(1)} - \nabla_x g(x,t)^\intercal v^{(2)} = 0 \tag{4.10}$$
$$\nabla_x g(x,t) v^{(1)} - v^{(3)} = 0 \tag{4.11}$$
$$v^{(2)} + D_\lambda^{-1} D_y v^{(3)} = 0. \tag{4.12}$$

By substituting $v^{(3)} = \nabla_x g(x,t) v^{(1)}$ and $v^{(2)} = -D_\lambda^{-1} D_y v^{(3)}$ into (4.10), we obtain

$$\left[\nabla L + \nabla_x g(x,t)^T D_\lambda^{-1} D_y \nabla_x g(x,t)\right] v^{(1)} = 0.$$

Since $\nabla L(x, \lambda)$ is assumed to be positive definite, having D_λ and D_y as positive definite matrices results in the trivial solution, $v^{(1)} = v^{(2)} = v^{(3)} = 0$. Thus the matrix $J\Phi_\varepsilon(z)$ is nonsingular. □

Theorem 4.2.2. *Let z^* be a solution to system $\Phi_\varepsilon(z) = 0$. The Jacobian $J\Phi_\varepsilon(z)$ is Lipschitz continuous near z^*.*

Proof. By definition of Lipschitz continuity, we need to show that

$$||J\Phi_\varepsilon(z_1) - J\Phi_\varepsilon(z_2)|| \leq \gamma||z_1 - z_2||$$

holds for a constant γ and all z_1 and z_2 sufficiently close to z^*. This is true when all elements of the Jacobian are locally Lipschitz continuous. □

4.3 A Discretization-based Smoothing Method

In order to solve the SIVI(X, F), we use the discretization approach to approximate the domain

$$X = \{x \in R^n \mid g(x, t) \geq 0 \quad \text{for all } t \in T\}.$$

by constructing a nested sequence $\{T_i\}$ of finite subsets of T with the property that for each i, $T_i \subset T_{i+1}$. Since T is a compact subset of R^l, T_i with $|T_i| < \infty$ can satisfy the following assumption.

Assumption 4.3.1. *Let $\Delta : \mathbb{N} \to R$ be a positive-valued, strictly monotone decreasing function such that $\lim_{n\to\infty} \Delta(n) = 0$. Then for each $n \in \mathbb{N}$ and $t \in T$, there exists $n_0 \in \mathbb{N}$ and $t' \in T_{n_0}$ such that*

$$\|t - t'\| \leq \Delta(n).$$

We can replace the index set T by T_i and define finitely constrained domain X_i as

$$X_i = \{x \in R^n \mid g(x, t) \geq 0 \quad \text{for all } t \in T_i\}. \tag{4.13}$$

We assume that X_1 is bounded. Then each X_i is a compact set. Since $X \subset X_i$ for any $i \in \mathbb{N}$, from the nonempty assumption of X, we know X_i is a nonempty set consisting of $|T_i|$ constraints.

We consider the following variational inequality problem

$$\begin{aligned} \mathrm{VI}(X_i, F) \quad : \quad & \text{find } x^i \in X_i \text{ such that} \\ & F(x^i)^T(x - x^i) \geq 0, \quad \text{for all } x \in X_i. \end{aligned} \tag{4.14}$$

Since $T_i \subset T_{i+1}$, we have $X \subset X_{i+1} \subset X_i$. Given that F is continuous and X_i is compact and convex, the solution to (4.14) is guaranteed to exist.

Let x^i be a solution of $\mathrm{VI}(X_i, F)$. The following result guarantees the existence of a subsequence of $\{x^i\}$ converging to the solution of $\mathrm{SIVI}(X, F)$.

Theorem 4.3.1. *The sequence $\{x^i\}$ has at least one accumulation point that solves problem SIVI(X, F).*

Proof. Since $\{x^i\} \subset X_1$, a compact set, there exists a subsequence $\{x^{k_i}\}$ such that

$$\lim_{i \to \infty} \{x^{k_i}\} = \hat{x}.$$

We prove in the following that $\hat{x}$ solves $\mathrm{SIVI}(X, F)$, i.e.,

$$F(\hat{x})^T(x - \hat{x}) \geq 0 \quad \text{for all } x \in X.$$

In fact, if $\hat{x}$ is not a solution of $\mathrm{SIVI}(X, F)$, then there exists at least one $\bar{x}$ such that

$$F(\hat{x})^T(\bar{x} - \hat{x}) < 0.$$

Since F is continuous, there exists $\bar{k} \in \{k_i\}$ such that

$$F(x^{\bar{k}})^T(\bar{x} - x^{\bar{k}}) < 0.$$

However, since $\bar{x} \in X \subset X_{\bar{k}}$ and $x^{\bar{k}}$ is a solution of VI$(X_{\bar{k}}, F)$, we have

$$F(x^{\bar{k}})^T(\bar{x} - x^{\bar{k}}) \geq 0,$$

which is a contradiction. Thus, for all $x \in X$, $F(\hat{x})^T(x - \hat{x}) \geq 0$. □

4.3.1 Algorithm 1

We propose the following iterative algorithm to solve the SIVI(X, F).

Step 0 Initialize the integer discretization parameters $c, n_0 > 0$ and stopping parameter $\delta_0 > 0$. Set $i = 1$.

Step 1 Set the discretization parameter as $n_i = n_{i-1} + c$. For $T_i = \{t_1, \ldots, t_{n_i}\} \subset T$, define $X_i = \{x \in R^n \mid g(x,t) \geq 0 \text{ for all } t \in T_i\}$.

Step 2 Solve VI(X_i, F), i.e., find $x^i \in X_i$ such that

$$F(x^i)^T(x - x^i) \geq 0 \quad \text{for all } x \in X_i.$$

Step 2.0 Set $k = 0$. Initialize the smoothing parameters $\varepsilon_k > 0, a > 1$.

Step 2.1 Define equivalent KKT conditions for VI(X_i, F) as

$$\Phi_{n_i,\varepsilon_k}(z) \equiv \Phi_{n_i,\varepsilon_k}(x, \lambda, y) \equiv \begin{pmatrix} L(x, \lambda, y) \\ g(x,t) - y \\ \lambda - p_{n_i,\varepsilon_k}(\lambda - y) \end{pmatrix} = 0.$$

Step 2.2 Solve $\Phi_{n_i,\varepsilon_k}(z) = 0$. Let $z_k^i = (x_k^i, \lambda_k^i, y_k^i)$ be the solution to this system.

Step 2.3 If $F(x_k^i)^T(x - x_k^i) \geq 0$ for all $x \in X_i$, then set $x^i = x_k^i$ and go to Step 3. Otherwise, decrease the smoothing parameter ε_k by factor of a, i.e., $\varepsilon_{k+1} = \varepsilon_k / a$. Set $k = k + 1$ and go to Step 2.1.

Step 3 If $\Delta(n_i) = \max_{t \in T} \min_{t' \in T_i} ||t - t'|| \leq \delta_0$ then stop. Otherwise, set $i = i + 1$ and go to Step 1.

The overall convergence rate of this finite termination algorithm depends on the discretization process and the convergence rate of the solution algorithm used in Step 2.2. In Step 2.2, we use the standard Newton method since the Jacobian of the given nonlinear system is nonsingular. Given the nonsingularity and Lipschitz continuity properties of the Jacobian, we have the following classic convergence theorem [35].

Theorem 4.3.2. *Consider function $H : R^n \to R^n$ and system $H(z^*) = 0$. Assume $JH(z^*)$ is Lipschitz continuous and nonsingular. If z_0 is sufficiently close to z^*, then the Newton sequence exists (i.e., $JH(z_n)$ is nonsingular for all $n \geq 0$), converges to z^*, and there is $K > 0$ such that*

$$||z_{n+1} - z^*|| \geq K||z_n - z^*||^2$$

for n sufficiently large.

In the next two theorems, we show that the sequence of solutions to the problems obtained in Step 2 converges to a solution of $\mathrm{VI}(X_i, F)$.

Theorem 4.3.3. *The sequence of points $z_i^k = (x_i^k, \lambda_i^k, y_i^k)$ generated by the algorithm has a convergent subsequence.*

Proof. To simplify the notation, we assume that the convergent subsequence is the sequence under consideration itself.

For sequence $\{z_i^k\}$ to be convergent, all its elements should be convergent. Since x_i^k is an element of a compact set X_i, its convergence $x_i^k \to x$ follows. Similarly, as y_i^k is identical to $g(x,t)$, which is a continuous function defined on compact set, y_i^k is convergent. To conclude the proof, we only need to show that the sequence $\{\lambda_i^k\}$ is convergent.

If $\{\lambda_i^k\}$ is bounded then the theorem holds.

Assume $\{\lambda_i^k\}$ is unbounded. Since z_i^k solves $\Phi_{n_i,\varepsilon_k}(z) = 0$, for every k and $t \in T_i$ we have

$$F(x_i^k) + \nabla_x g(x_i^k, t)\lambda_i^k = 0. \tag{4.15}$$

For the unbounded sequence $\{\lambda_i^k\}$, we have

$$\lim_{k\to\infty}\left[\frac{F(x_i^k)}{||\lambda_i^k||}+\nabla_x g(x_i^k,t)\frac{\lambda_i^k}{||\lambda_i^k||}\right]=0. \tag{4.16}$$

Since the normalized vector $\frac{\lambda_i^k}{||\lambda_i^k||}$, lies on the unit ball, which is closed and bounded, there exists a convergent subsequence, i.e.,

$$\frac{\lambda_i^k}{||\lambda_i^k||}\to\bar{\lambda}\neq 0,\bar{\lambda}\in R^m. \tag{4.17}$$

Thus, by (4.16) and (4.17) we obtain

$$\nabla_x g(x,t)\bar{\lambda}=0. \tag{4.18}$$

Since $\bar{\lambda}_i=0$ only for the inactive constraints and $\bar{\lambda}_i\neq 0$ for active constraints, (4.18) contradicts the linear independence assumption on the active gradients of the constraints $g_i(x,t), i\in I(x,t)$ where $I(x,t)=\{i\mid g_i(x,t)=0\}$. Thus, the sequence of λ_k cannot be unbounded. This completes the proof. □

Theorem 4.3.4. *For a given i, the sequence of solutions* $\{z_i^k\}$ *generated by the algorithm to the smoothed KKT system* $\Phi_{n_i,\varepsilon_k}(z_i^k)=0$ *converges to the solution z of the equivalent KKT system* $\Phi(z)$ *of VI*(X_i,F).

Proof. We need to show that as the point z_i^k converges to point z, the function value $\Phi_{n_i,\varepsilon_k}(z^k)$ converges to $\Phi(z)$, i.e.,

$$\lim_{k\to\infty}||\Phi_{n_i,\varepsilon_k}(z_i^k)-\Phi(z)||=0. \tag{4.19}$$

Notice that

$$\begin{aligned}||\Phi_{n_i,\varepsilon_k}(z_i^k)-\Phi(z)|| &= ||\Phi_{n_i,\varepsilon_k}(z_i^k)-\Phi_{n_i,\varepsilon_k}(z)+\Phi_{n_i,\varepsilon_k}(z)-\Phi(z)||\\ &\leq ||\Phi_{n_i,\varepsilon_k}(z_i^k)-\Phi_{n_i,\varepsilon_k}(z)||+||\Phi_{n_i,\varepsilon_k}(z)-\Phi(z)||.\end{aligned}$$

Since Φ_{n_i,ε_k} is a continuous function and $z_i^k\to z$ as proved in the previous theorem, we have

$$\lim_{k\to\infty}||\Phi_{n_i,\varepsilon_k}(z_i^k)-\Phi_{n_i,\varepsilon_k}(z)||=0.$$

For a given point z, the convergence of $\Phi_{n_i,\varepsilon_k}(z)$ to $\Phi(z)$ as $k \to \infty$, i.e.,

$$\lim_{k\to\infty} ||\Phi_{n_i,\varepsilon_k}(z) - \Phi(z)|| = 0$$

follows because

$$\lim_{\varepsilon_k\to 0} p_{\varepsilon_k}(\lambda - y) = \max\{0, \lambda - y\}.$$

This establishes the desired result in equation (4.19). □

Theorem 4.3.5. *As $n_i \to \infty$ and $\varepsilon_k \to 0$, the sequence of solutions $\{x_i^k\}$ to the system $\Phi_{n_i,\varepsilon_k}(z_i^k) = 0$ converges to x^* which is a solution of SIVI(X, F).*

Proof. The result follows from Theorems 4.3.4, 4.2.1, and 4.3.1. □

4.4 Exchange Method

In this section, we introduce our second algorithm to solve semi-infinite variational inequalities by generating a sequence of finite problems by using the exchange method. By using the smoothing scheme in Section 4.2.1, finite problems are smoothed and solved by Newton's method. We prove the convergence of the proposed algorithm.

In the context of semi-infinite programming, the exchange method is one of the reduction techniques to approximate the feasible set by finitely many constraints. Basically, it is a way to approximate the infinitely constrained feasible set by exchanging the constraints to remove or minimize constraint violation. From either a local or global point of view, the exchange method takes advantage of having a better representation of the feasible set. Therefore, it is known to be more efficient than the ordinary discretization method.

Although the name refers to "exchange of the constraints", depending on the implemented rules, there may or may not be any real exchange of constraints. However, there new constraints are added at each step.

We approximate the domain

$$X = \{x \in R^n \mid g(x,t) \geq 0 \quad \text{for all } t \in T\}$$

according to the following conceptual exchange method:

(Conceptual Exchange Method)

Step 1. Let the iteration number be i. Given a finitely constrained set $X \subset X_i$ and a fixed small value $\alpha > 0$, calculate a solution x_i of VI(X_i, F).

Step 2. Given x_i, let t_i^s denote the s^{th} local solution of the subproblem

$$\min_t g(x_i, t) \quad \text{s.t. } t \in T.$$

Compute all local solutions $t_i^s, s = 1, \ldots, \bar{s}$.

Step 3. If $g(x_i, t_i^s) \geq -\alpha$, then STOP. Otherwise, update $X_{i+1} = X_i \cup \{t_i^s, s = 1, \ldots, \bar{s}\}$. Set $i = i + 1$ and go to Step 1.

For nonlinear and nonconvex problems, especially for cases where $t \in R^l, l \geq 2$, the determination of a global solution for

$$\min_t g(x_i, t) \quad \text{s.t. } t \in T$$

may be very costly. To overcome this computational difficulty, one strategy is to include all local solutions found. Another strategy is to select a violated constraint arbitrarily and cut off a feasible region containing no solution. Both strategies are reported to have remarkable success in [24] and we select the first one. The following result ensures the convergence of the exchange method.

Theorem 4.4.1. *Assume that the domain X of SIVI(X, F) is compact and that F and $g(x,t)$ are continuous. Then, the exchange method stops after a finite number of steps with a solution x^* equal to x_i from the last iteration or the sequence $\{x_i\}$ has an accumulation point that solves SIVI(X, F).*

Proof. Since $T_i \subset T_{i+1} \subset T$, we have $X \subset X_{i+1} \subset X_i$. Given that F is continuous and X_i is compact, the solution to VI(X_i, F) is guaranteed to exist.

Since $\{x_i\}$ is bounded, an accumulation point $\bar{x}$ exists. Without loss of generality, assume that $x_i \to \bar{x}$. Then it is sufficient to show that $\bar{x} \in X$ or equivalently $\upsilon(\bar{x}) \geq 0$, where

$$\upsilon(\bar{x}) = \min_t \{g(x,t),\ t \in T\}.$$

Assume $\upsilon(\bar{x}) < 0$, i.e., there is a constraint violation and $\bar{x} \notin X$. If, $\upsilon(x_i) = g(x_i, \tilde{t}^k)$, we have

$$\begin{aligned} \upsilon(\bar{x}) &= \upsilon(x_i) + [\upsilon(\bar{x}) - \upsilon(x_i)] \\ &= g(x_i, \tilde{t}^k) + [\upsilon(\bar{x}) - \upsilon(x_i)]. \end{aligned}$$

Since $\tilde{t}^k \in T_{k+1}$, we have $g(x_{i+1}, \tilde{t}^k) \leq 0$. By continuity of g and υ, we obtain

$$\upsilon(\bar{x}) \geq [g(x_i, \tilde{t}^k) - g(x_{i+1}, \tilde{t}^k)] + [\upsilon(\bar{x}) - \upsilon(x_i)].$$

Obviously, the right hand side becomes arbitrarily small as $i \to \infty$. This results in $\upsilon(\bar{x}) \geq 0$, which is a contradiction to the assumption of $\upsilon(\bar{x}) < 0$. □

4.4.1 Algorithm 2

We propose the following algorithm based on the exchange method.

Step 0 Set the value of $\alpha > 0$ and the number of constraints to N for the starting finite set T_i. Let $i = 1$.

Step 1 Define $X_i = \{x \in R^n \mid g(x,t) \geq 0 \ \text{ for all } t \in T_i\}$.

Step 2 Solve VI(X_i, F), i.e., find $x^i \in X_i$ such that

$$F(x^i)^T(x - x^i) \geq 0 \quad \text{for all } x \in X_i.$$

Step 2.0 Initialize smoothing parameters $\varepsilon_0 > 0, a > 1$. Set $k = 0$.

Step 2.1 Define the equivalent KKT conditions for VI(X_i, F) as

$$\Phi_{n_i,\varepsilon_k}(z) \equiv \Phi_{n_i,\varepsilon_k}(x,\lambda,y) \equiv \begin{pmatrix} L(x,\lambda,y) \\ g(x,t)-y \\ \lambda - p_{n_i,\varepsilon_k}(\lambda - y) \end{pmatrix} = 0.$$

Step 2.2 Solve $\Phi_{n_i,\varepsilon_k}(z) = 0$. Let $z_k^i = (x_k^i, \lambda_k^i, y_k^i)$ be the solution.

Step 2.3 If $F(x_k^i)^T(x - x_k^i) \geq 0$ for all $x \in X_i$, then set $x^i = x_k^i$ and go to Step 3. Otherwise, update the smoothing parameter, i.e., decrease $\varepsilon_{k+1} = \varepsilon_k/a$. Set $k = k+1$ and go to Step 2.1.

Step 3 Given the solution x^i, find local solutions $t_i^s, s = 1, \ldots, \bar{s}$, of the subproblem

$$\min_t g(x^i, t) \quad \text{s.t. } t \in T.$$

Step 4 If $g(x^i, t_i^s) \geq -\alpha$ for all $s = 1, \ldots, \bar{s}$, then STOP. Otherwise, set $X_{i+1} = X_i \cup \{t_i^s, s = 1, \ldots, \bar{s}\}$. Go to Step 2.

Next two theorems show that the sequence of solutions to smoothed problems, obtained in Step 2, converges to a solution of VI(X_i, F). Note that the overall convergence rate of Algorithm 2 depends on the convergence rate of the solution algorithm used in Step 2.2 and Step 3. In Step 2.2, we use the standard Newton method with quadratic convergence.

Theorem 4.4.2. *As $n_i \to \infty$ and $\varepsilon_k \to 0$, the sequence of solutions $\{x_i^k\}$ to the system $\Phi_{n_i,\varepsilon_k}(z_i^k) = 0$ converges to x^* which is a solution of SIVI(X, F).*

Proof. By Theorem 4.3.4, the sequence of solutions $\{z_i^k\}$ generated by the exchange algorithm to the smoothed KKT system $\Phi_{n_i,\varepsilon_k}(z_i^k) = 0$ converges to the solution z of the equivalent KKT system $\Phi(z)$ of VI(X_i, F). Given a solution of VI(X_i, F) at each iteration, by Theorem 4.4.1, the exchange algorithm terminates with a solution to SIVI(X, F). □

4.5 Entropic Analytic Center Cutting Plane Method

An analytic center cutting plane method (ACCPM) is an interior point method that generates a sequence of analytic centers which are supposed to converge to a solution of the problem. The analytic center is a function of the data describing the polytope and defined as follows [27].

Definition 4.5.1. The analytic center (AC) of a bounded convex polyhedral set $\Omega = \{x \in R^n \mid A^\top x \leq c\}$ with $int(\Omega) \neq \emptyset$ is the unique solution of the problem

$$\begin{aligned} \max \quad & \sum_{i=1}^{n} \ln s_i \\ \text{s.t.} \quad & A^\top x + s = c \\ & s > 0. \end{aligned} \tag{4.20}$$

Finding the analytic center can be computationally expensive. Chen [5, 6] converts problem (4.20) into an entropy optimization problem and defines an "entropic analytic center" (EAC) as follows.

Definition 4.5.2. Given a polyhedral set $\Omega_0 = \{x \in R^n \mid A^\top x \leq c, x \geq 0\}$ with $int(\Omega_0) \neq \emptyset$, the point x^* is the entropic analytic center of Ω_0 if it is a unique minimizer of the linearly constrained minimum entropy optimization problem,

$$\begin{aligned} \min \quad & \sum_{i=1}^{n} x_i \ln x_i + \sum_{j=1}^{m} s_j \ln s_j \\ \text{s.t.} \quad & A^\top x + s = c \\ & x \geq 0, s \geq 0. \end{aligned} \tag{4.21}$$

In comparison to the analytic center, the entropic analytic center does not require the containing polytope to be bounded and it helps in detecting the inconsistency of a system. Note that solving (4.21) is equivalent to solving its dual which is the following unconstrained optimization problem:

$$\max_{w \in R^m} \left\{ \gamma(w) \equiv b^\top w - \sum_{k=1}^{r} \exp\left\{ \hat{A}_k{}^\top w - 1 \right\} \right\}, \tag{4.22}$$

where $\hat{A} = [A\ I_m]$ and $\hat{A}_k$ is the k^{th} column vector of matrix $\hat{A}$.

We modify the analytic center cutting plane algorithm of Fang et al. [16] by replacing the analytic center by the entropic analytic center to observe the convergence and computational results in comparison to other algorithms for solving SIVI(X, F). Although we expect it to be similar in complexity to using Karmarkar's potential function, the complexity analysis of EAC cutting plane method has not been determined.

In order for the entropic analytic center cutting plane algorithm work, we need the following result and assumptions as in [16]:

Assumption 1. X is compact and F is continuous.

Proposition 4.5.1. *Assume that F is Lipschitz continuous with modulus L and that M is an upper bound on the norm of F over X. Let x^* be any solution to SIVI(X, F) and let $B(x^*, \delta)$ be a ball with center at x^* and radius $\delta = \epsilon/(M + Ld)$. Then every point $\bar{x} \in B(x^*, \delta) \cap X$ is an ϵ-solution of SIVI(X, F), i.e., the gap function $\theta(\bar{x}) \leq \epsilon$.*

Assumption 2. $int(X) \neq \emptyset$ and $B(x^*, \delta) \subset int(X)$.

Assumption 3. The ball $B(\bar{x}, \rho(\epsilon)) \subset B(x^*, \delta) \cap X$ is contained in $\Omega^0 = \{x \in R^n \mid A^0 x \leq c^0, x \geq 0\}$.

4.5.1 Algorithm 3

We propose the following entropic analytic center cutting plane algorithm for solving SIVI(X, F):

Step 0. (Initialization)

Set $k = 0$ and pick a sufficiently small $\epsilon > 0$. Let

$A^0 = [I\ -I]^\intercal \in R^{2n \times n}$,

$c^0 = [e\ 0]^\intercal \in R^{2n}$, $x^0 = \frac{1}{2}e$.

Step 1 *(Computation of the entropic analytic center)*

Find the entropic analytic center x^k of $\Omega^k = \{x \in R^n \mid A^k x \leq c^k, x \geq 0\}$.

If no x^k exists STOP. SIVI(X, F) has no solution.

Step 2 *(Feasibility check)*

Find $t^k = \arg\ \min_{t \in T} g(x^k, t)$. If $g(x^k, t^k) < 0$, compute the subgradient $a \in \partial_x g(x^k, t^k)$ and continue. Otherwise go to Step 4.

Step 3 *(Generation of a cutting plane)*

$$H^k = \{x \in R^n \mid a^\top (x - x^k) = 0\},$$

$$A^{k+1} = \begin{bmatrix} A^k \\ \frac{a}{||a||} \end{bmatrix}, \quad c^{k+1} = \begin{bmatrix} c^k \\ \frac{a^\top x^k}{||a||} \end{bmatrix}.$$

Step 4 *(Stopping criterion)*

Compute $\theta(x^k) = \max\limits_{y \in X} \{F(x^k)^\top (x^k - y)\}$.

If $\theta(x^k) \leq \epsilon$ then STOP. Otherwise return to Step 1.

Step 5 *(Generation of an optimality cut)*

$$H^k = \left\{x \in R^n \mid \frac{F(x^k)^\top}{||F(x^k)^\top||}(x - x^k) = 0\right\},$$

$$A^{k+1} = \begin{bmatrix} A^k \\ \frac{F(x^k)^\top}{||F(x^k)^\top||} \end{bmatrix}, \quad c^{k+1} = \begin{bmatrix} c^k \\ \frac{F(x^k)^\top x^k}{||F(x^k)^\top||} \end{bmatrix}.$$

Step 6 Set $k = k + 1$ and go to Step 1.

Under the assumption that F is pseudo-monotone, the following theorem for ACCPM by Fang et al. [16] proves the convergence of EACCPM algorithm as well.

Theorem 4.5.1. *Assume that SIVI(X, F) is given with F being nonzero, Lipschitz continuous on X with modulus L and pseudo-monotone plus on X. The proposed algorithm either terminates with an $\varepsilon-$solution of SIVI(X, F) or generates a subsequence of $\{x^k\}$ converging to a solution of SIVI(X, F).*

4.6 Numerical Results

In this section we report the results obtained when we implemented our solution algorithms on some problems commonly encountered in the literature. The test problems (1-3) are semi-infinite variational inequality problems with linear constraints and were chosen from [15]. Due to the shortage in the examples of SIVI(X, F) in the literature, we added six additional test problems from semi-infinite programming literature. Larger but linear problems (4-6) are mathematical programming reformulations of semi-infinite programming problems reported in [14, 38]. Finally, Problems (7-9) are nonlinear semi-infinite programs from SIP test library with source given as [58, 8].

All three algorithms have been implemented in Matlab 6.5 and run on a computer with 1400 MHz processor, 768 MB of RAM and Windows XP operating system. For Algorithms 1 and 2, we used standard Newton method to solve the system of equations in Step 2.2. The smoothing parameter ε is reported for each problem and updated by factor of $a = 2$ at each iteration. For Algorithm 3, the computation of exact entropic analytic center and gap function evaluation is performed by using Matlab optimization toolbox. For all computations, we used 1.0E-8 as the required accuracy. Numerical results for each method and group of problems are reported in Tables A.1-A.8 in Appendix A.

4.6.1 Test Problems

We solve the following test problems with our proposed algorithms:

Problem 1. $n = 7,\ T = [0, 1]$, and

$$X = \{x \in R^n \mid \sum_{j=1}^{n} t^{j-1} x_j \leq \sum_{l=1}^{4} t^{2l} + 1,\ t \in T \text{ and } 0 \leq x_j \leq 1,\ j = 1, \ldots, n\},$$

$$F = (F_1, \ldots, F_7) \text{ with } F_j = x_j - \frac{1}{\sqrt{x_j}},\ j = 1, \ldots, n.$$

Problem 2. $n = 7,\ T = [0, 1]$, and

$$X = \{x \in R^n \mid \sum_{j=1}^{n} t^{j-1} x_j \leq 4t^5 + 1,\ t \in T \text{ and } 0 \leq x_j \leq 1,\ j = 1, \ldots, n\},$$

$$F = (F_1, \ldots, F_7) \text{ with } F_j = 1 + 3x_j - \frac{1}{x_j^2},\ j = 1, \ldots, n.$$

Problem 3. $n = 7,\ T = [0, 1]$, and

$$X = \{x \in R^n \mid \sum_{j=1}^{n} t^{j-1} x_j \leq 3t^5 + 2t^2 + 1,\ t \in T \text{ and } 0 \leq x_j \leq 1,\ j = 1, \ldots, n\},$$

$$F = (F_1, \ldots, F_7) \text{ with } F_j = \sqrt{x_j} - \frac{1}{x_j^2},\ j = 1, \ldots, n.$$

Problem 4. $n = 50,\ T = [0, 1]$, and

$$X = \{x \in R^n \mid \sum_{j=1}^{n} t^{j-1} x_j \geq \sin(t),\ t \in T \text{ and } x_j \geq 0,\ j = 1, \ldots, n\},$$

$$F = (F_1, \ldots, F_n) \text{ with } F_j = \frac{1}{j},\ j = 1, \ldots, n.$$

Problem 5. $n = 50,\ T = [0, 1]$, and

$$X = \{x \in R^n \mid \sum_{j=1}^{n} t^{j-1} x_j \geq e^t,\ t \in T \text{ and } x_j \geq 0,\ j = 1, \ldots, n\},$$

$$F = (F_1, \ldots, F_n) \text{ with } F_j = \frac{1}{j},\ j = 1, \ldots, n.$$

Problem 6. $n = 50,\ T = [0, 1]$, and

$$X = \{x \in R^n \mid \sum_{j=1}^{n} t^{j-1} x_j \geq \frac{1}{2 - t},\ t \in T \text{ and } x_j \geq 0,\ j = 1, \ldots, n\},$$

$$F = (F_1, \ldots, F_n) \text{ with } F_j = \frac{1}{j},\ j = 1, \ldots, n.$$

Problem 7. $VI(\nabla f, X)$ formulation of the following nonlinear optimization problem [58]:

$$\begin{aligned} \min \quad & f = (x_1 - 2x_2 + 5x_2^2 - x_2^3 - 13)^2 + (x_1 - 14x_2 + x_2^2 + x_2^3 - 29)^2 \\ \text{s.t.} \quad & -x_1^2 - 2x_2t^2 + e^{x_1+x_2} + e^t \geq 0 \\ & t \in [0, 1]. \end{aligned}$$

Problem 8. $VI(\nabla f, X)$ formulation of the following nonlinear optimization problem [8]:

$$\begin{aligned} \min \quad & f = c^2e^{x_1} + e^{x_2} \\ \text{s.t.} \quad & e^{x_1+x_2} - t \geq 0 \\ & t \in [0, 1]. \end{aligned}$$

Problem 9. $VI(\nabla f, X)$ formulation of the following nonlinear optimization problem [8]:

$$\begin{aligned} \min \quad & f = \tfrac{1}{2}x_1^2 + x_2^2 + \tfrac{1}{2}x_1 \\ \text{s.t.} \quad & (1 - x_1^2t^2)^2 - x_1t^2 - x_2^2 + x_2 \leq 0 \\ & t \in [0, 1]. \end{aligned}$$

4.6.2 Results for Discretization-based Smoothing Method

We solved nine test problems using Algorithm 1 and report the results in Figures 4.2, 4.3 and in Tables A.1-A.3. Note that in each table, N corresponds to the number of discretization points, x^* is the solution obtained by our methodology, and x_b is the best known solution of the corresponding problem in the literature. In Tables A.2 and A.3, f^* is the optimal objective value obtained using Algorithm 1 and f_b is the reported optimal objective value for the given problem. In each table, the starting point x_0 and the number of Newton iterations needed to find the solution are also reported.

The convergence rate of the discretization-based smoothing method can be observed in Figure 4.2 for test problems 1-3, and in Figure 4.3 for test problems 4-6. In both the figures,

red plots on the left hand side show the number of Newton iteration for the selected grid size N at each iteration. Notice that as the grid size gets larger, the deviation of our solution from the best known solution decreases, which is expected. In other words, to obtain better solution approximation a finer grid (and correspondingly higher computational effort) is needed.

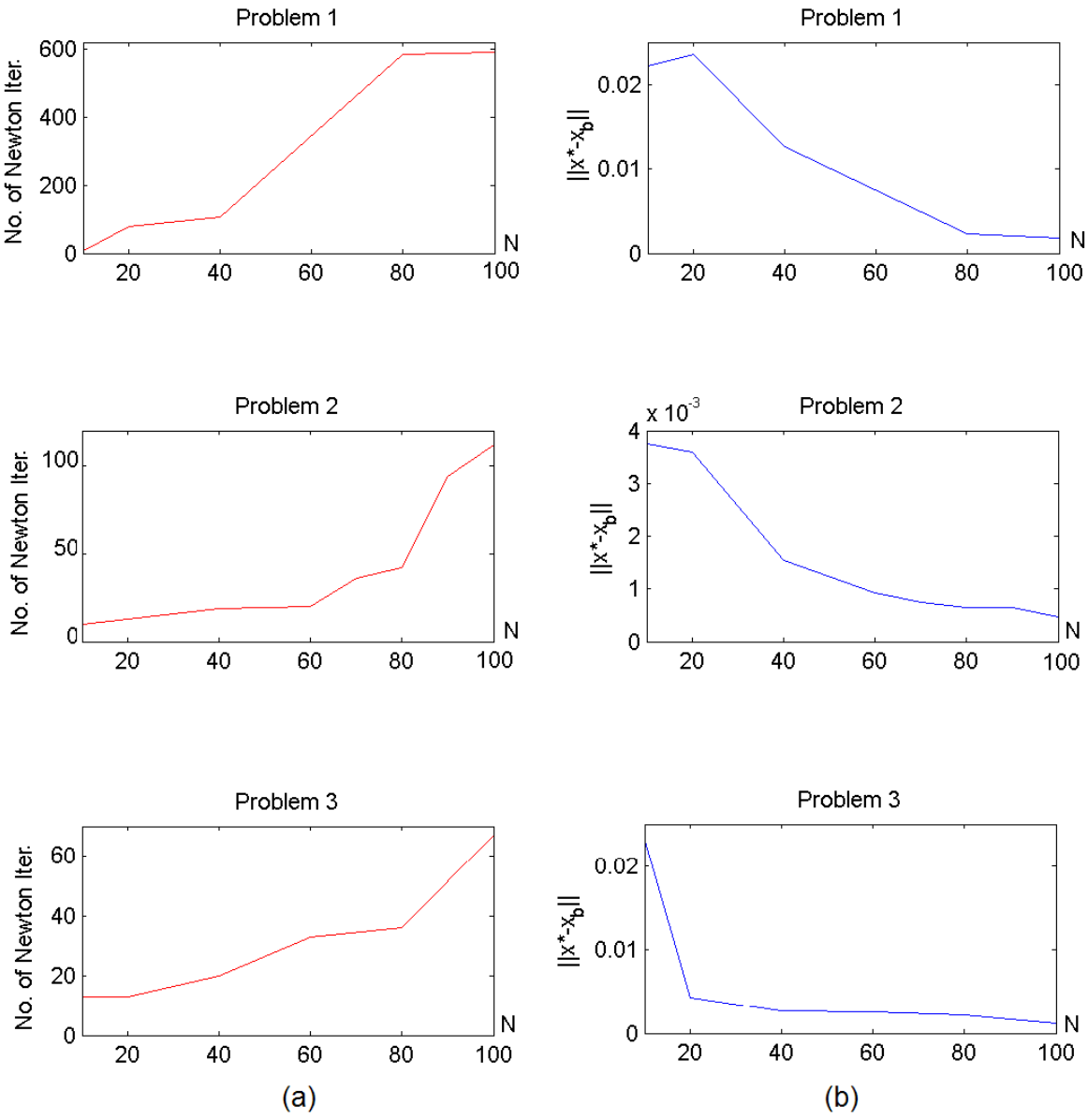

Figure 4.2: Algorithm 1: (a) Number of Newton iterations vs. the grid size N for problems 1-3, (b) Deviation from the best known solution for problems 1-3.

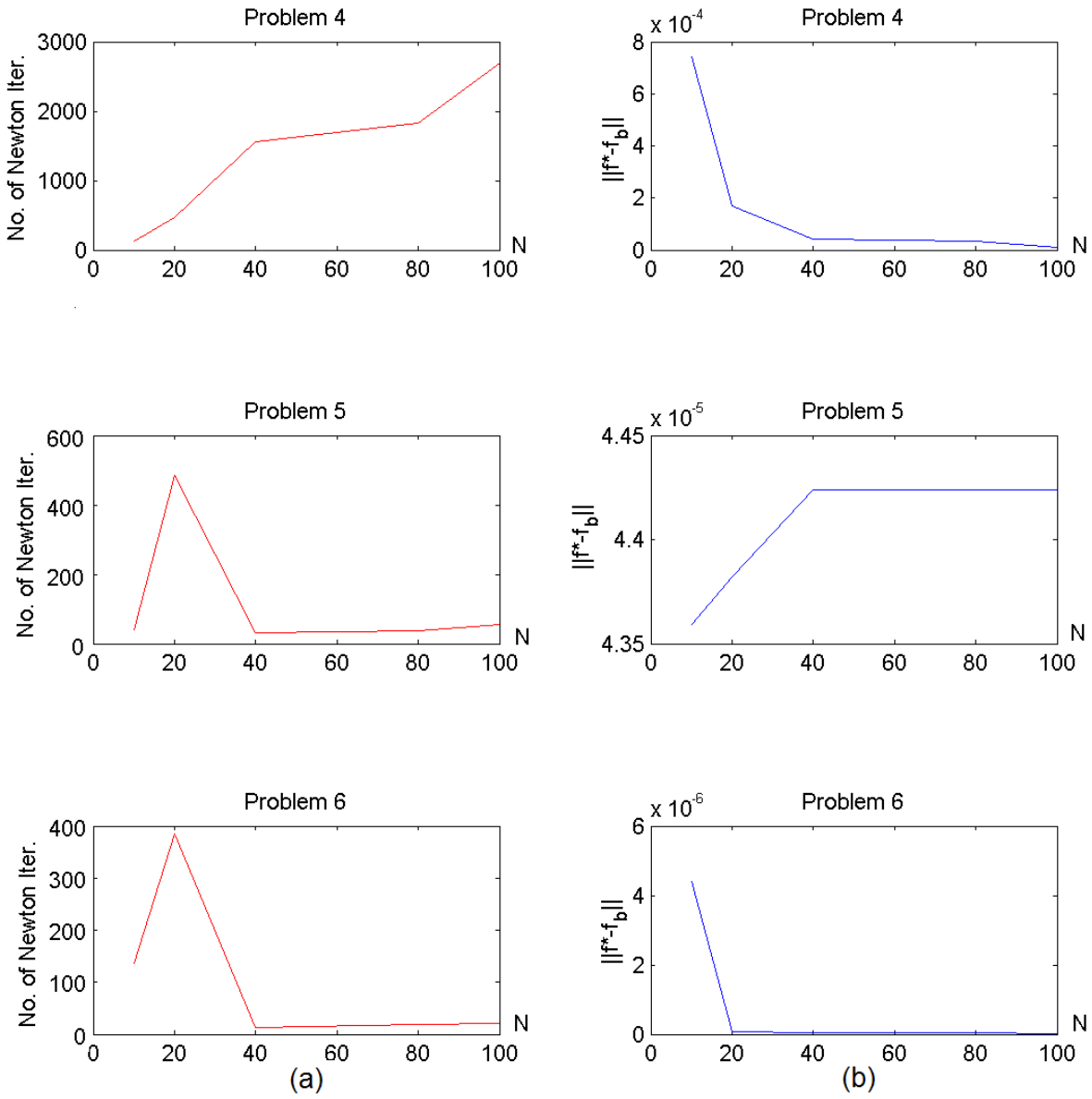

Figure 4.3: Algorithm 1: (a) Number of Newton iterations vs. the grid size N for problems 4-6, (b) Deviation from best known solution for problems 4-6.

4.6.3 Results for Exchange Method

By implementing the proposed exchange algorithm, we obtained the set of numerical results reported in Tables A.4-A.6. Specific to this exchange method, we report N as the required number of constraints defining the domain X to find the solution x^*. The column t_l indicates the local minimizers of $g(\bar{x}, t)$ over T, at which new constraints are added to define X. The constraint function $g(\bar{x}, t_l)$ evaluation is also reported. Note that the function value for $g(\bar{x}, t_l)$ is supposed to be zero for exchange algorithm to terminate and give the solution.

In addition to the starting point x_0 and the reported solution x_b in the literature, the norm difference between x^* and x_b is also reported as a performance indicator for our algorithm. Note that the lack of precision in the reported values of x_b affects the precision of the last column and should not lead to misinterpretation of our results. When the original problem is an optimization problem as in (4-9), x_b and $||x^* - x_b||$ are replaced by the best known objective value f_b and $||f^* - f_b||$, respectively.

As it can be seen in Table A.6, the test problems 6-9 do not require more than one iteration of the exchange algorithm. Therefore, we illustrate the convergence results of Algorithm 2 in Figure 4.4 for test problems 1-3 and in Figure 4.5 for test problems 4-6. The red plots on the left hand side illustrate the constraint violation as a function of number of constraints added to define the domain X. The blue plots on the right hand side shows the deviation from the best known solution as a function of number constraints. In comparison to Figures 4.2 and 4.3, the convergence rate of the exchange method for the same problems observed to be faster than of the discretization based smoothing method.

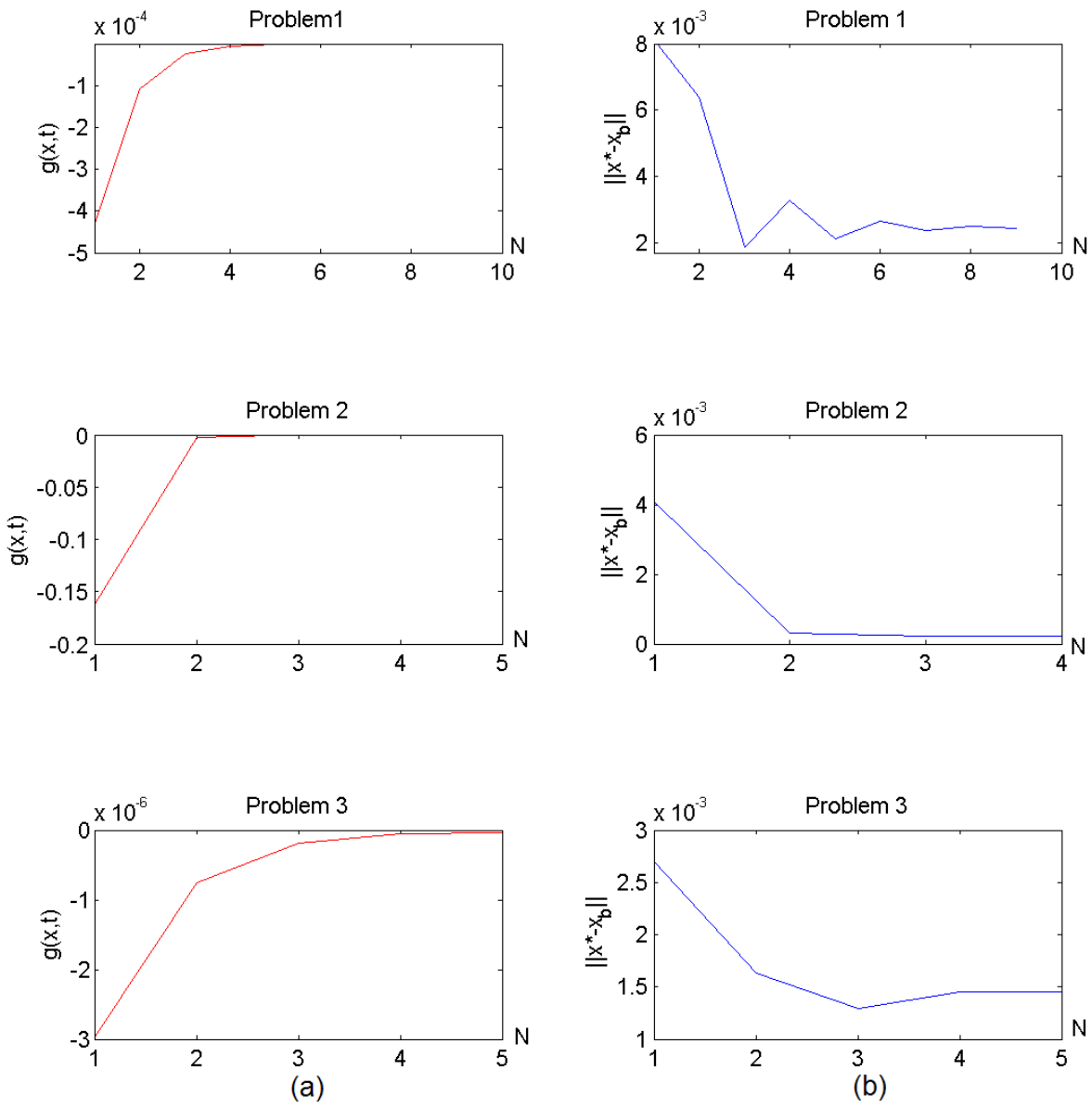

Figure 4.4: Algorithm 2: (a) Constraint violation vs. the number of iterations for problems 1-3, (b) Deviation from best known solution at each iteration for problems 1-3.

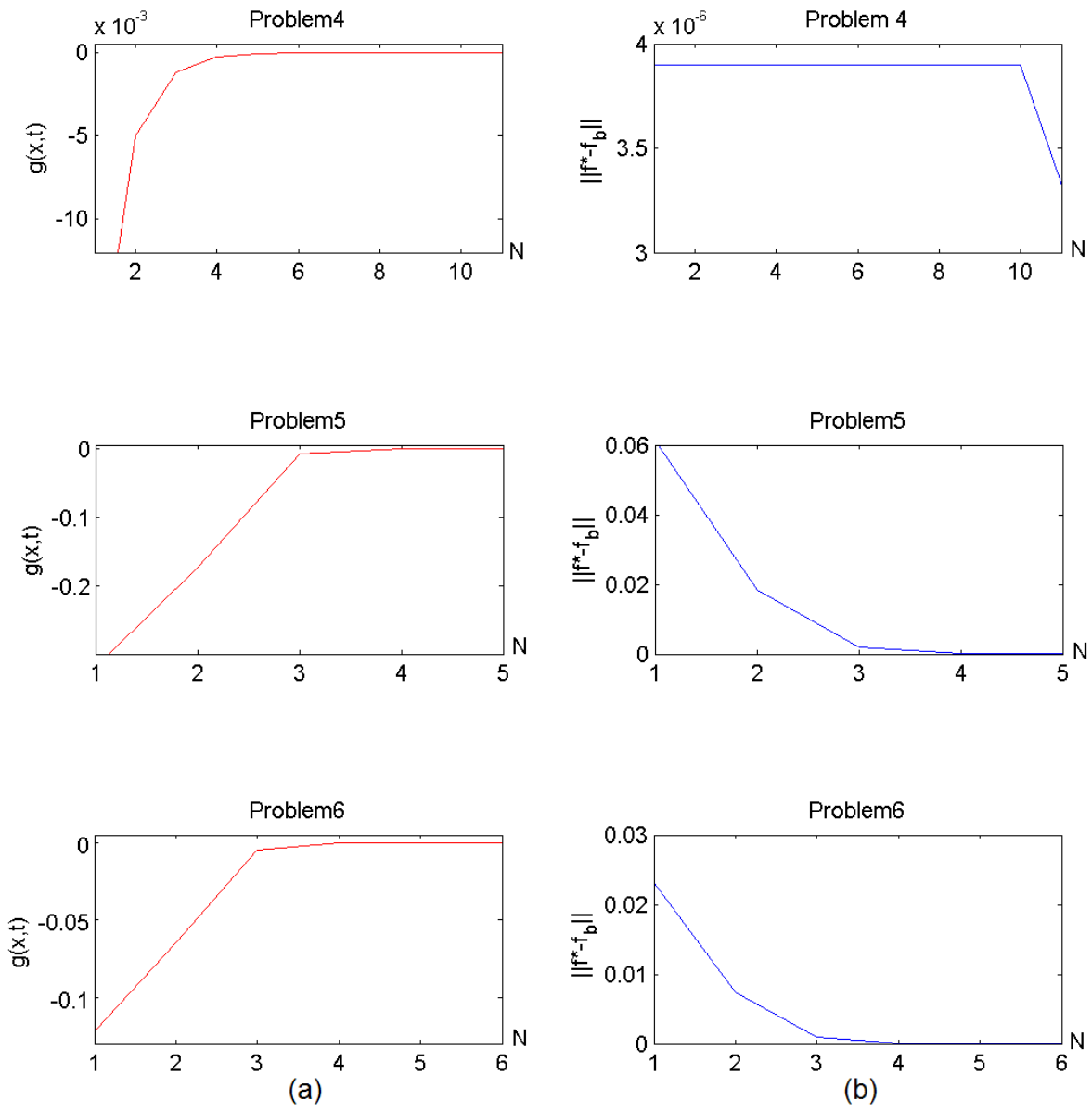

Figure 4.5: Algorithm 2: (a) Constraint violation vs. the number of iterations for problems 4-6, (b) Deviation from best known solution at each iteration for problems 4-6.

4.6.4 Results for Entropic Analytic Center Cutting Plane Method

Tables A.7 and A.8 lists the results from the implementation of the proposed entropic analytic center cutting plane method (EACCPM). The k in the first column of the tables refers to the iteration index. The x^* in the second column is the solution obtained by the entropic analytic center cutting plane algorithm. The f^* in the third column is the optimal objective function value of the corresponding optimization problem obtained and f_b is the best reported objective function value.

As performance indicators, we chose the number of function evaluations and the deviation from the best known solution. We counted the function evaluations required both to find the entropic analytic center and to compute the gap function, and reported their sum. Similarly, the total number of function iterations is obtained by adding the number of iterations at Step 1 and Step 4 of the Algorithm 3. Note that problems 1-3 require relatively many entropic analytic center computations.

The computational results of our entropic analytic center cutting plane algorithm are presented in Figure 4.6 and Figure 4.7, respectively for test problems 1-3 and 4-6. Figure 4.6 (a) shows the number of function evaluations with respect to the number of entropic analytic centers calculated. In other words, the computational effort increases as the analytical centers gets closer to the solution. The deviation from the best known solution as a function of algorithm iterate k is given in Figure 4.6 (b). As it can be seen, the required computational effort by the EACCPM algorithm is considerably higher than of both discretization-based smoothing and exchange methods.

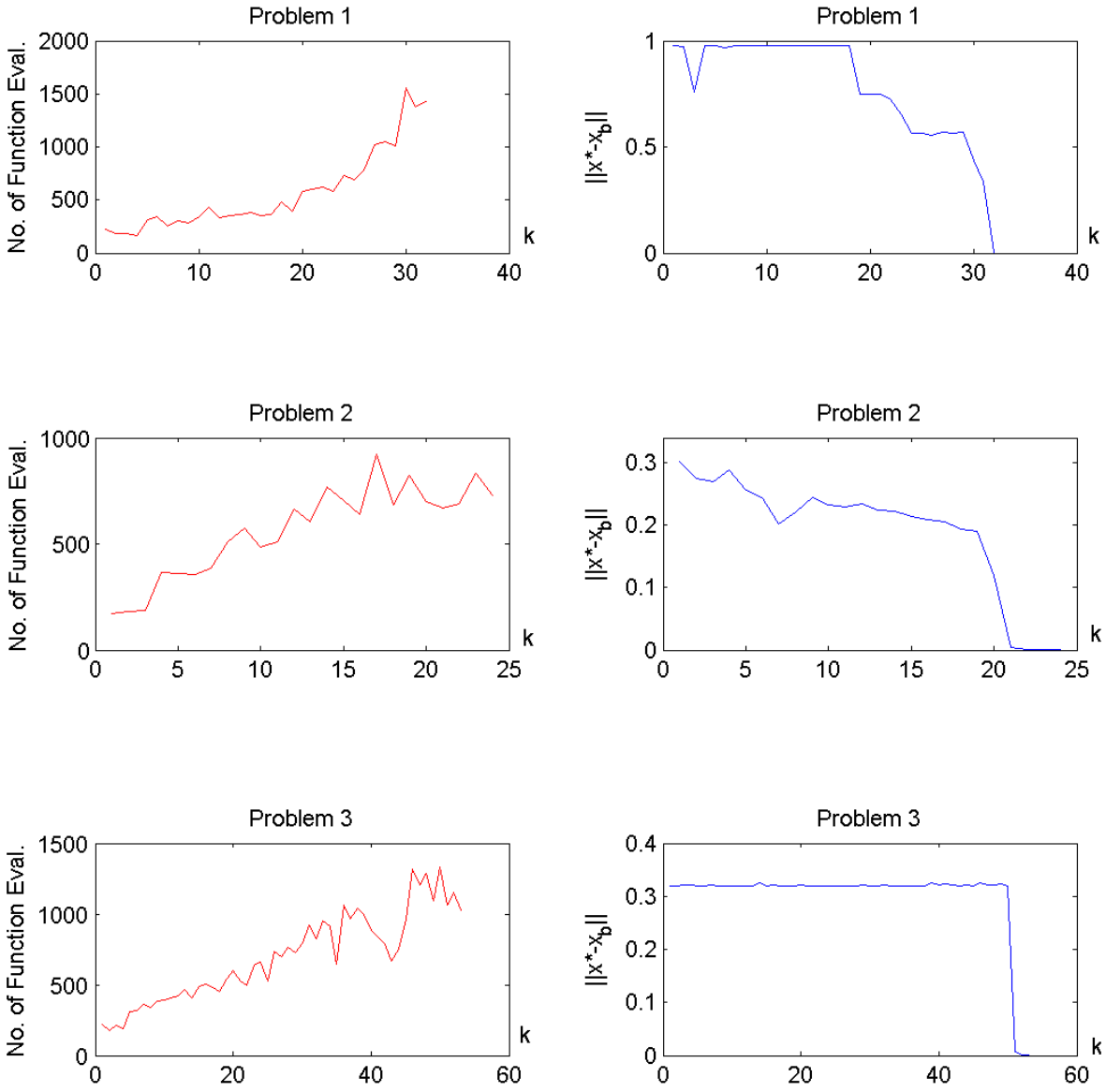

Figure 4.6: Algorithm 3: (a) Number of function evaluations vs. the number of iterations k for problems 1-3, (b) Deviation from best known solution at each iteration for problems 1-3.

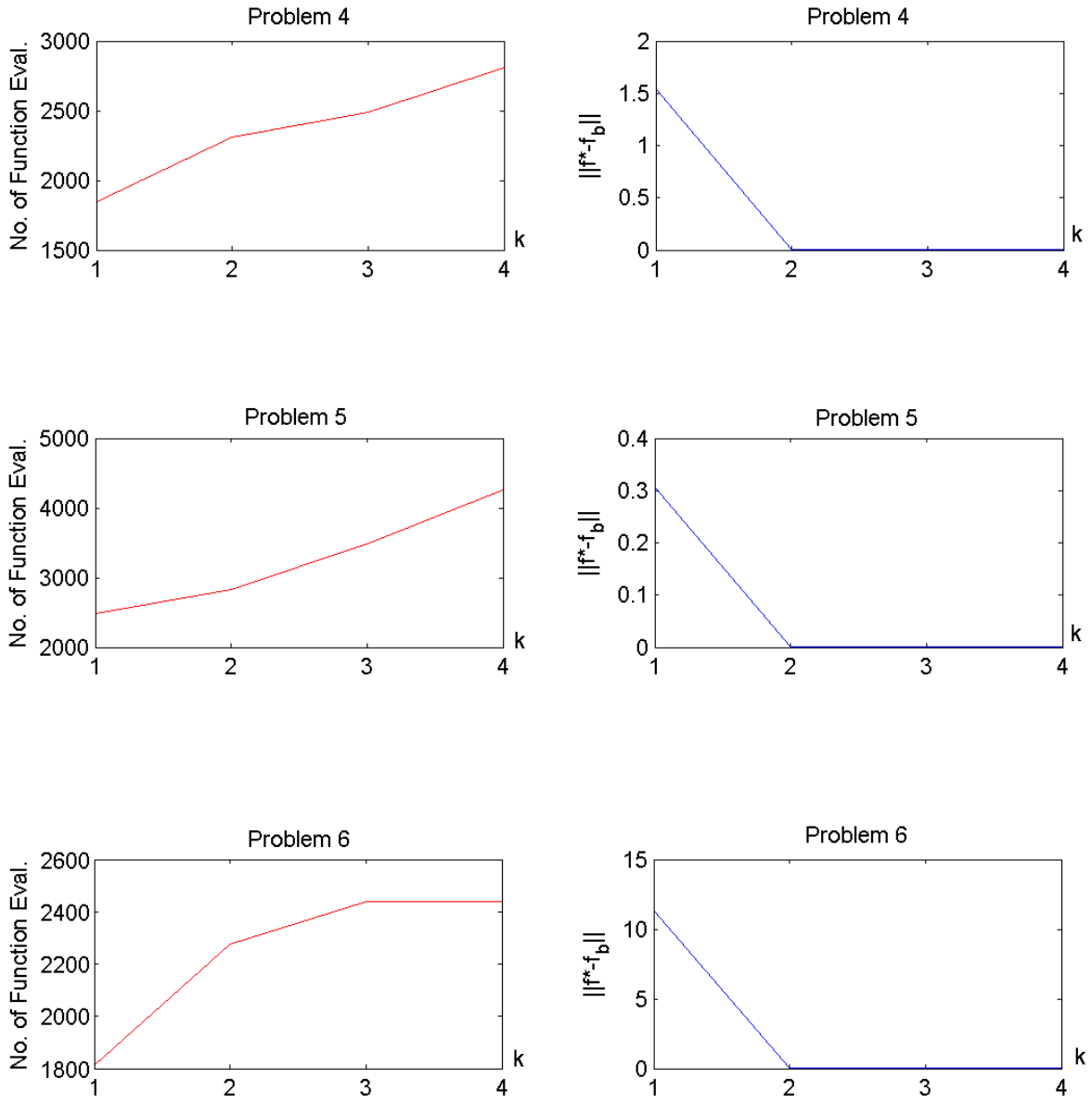

Figure 4.7: Algorithm 3: (a) Number of function evaluations vs. the number of iterations k for problems 4-6, (b) Deviation from best known solution at each iteration for problems 4-6.

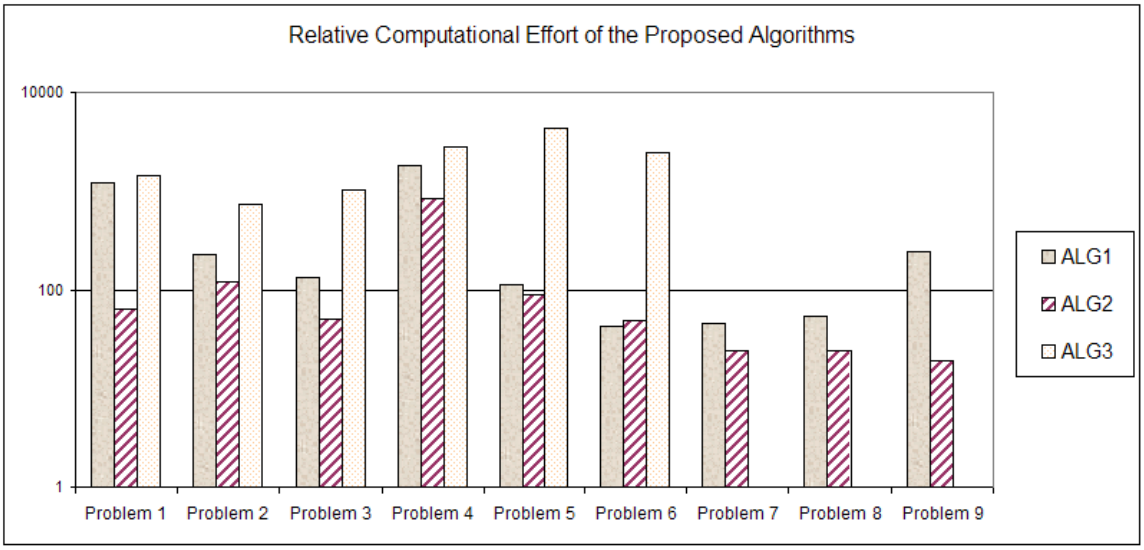

Figure 4.8: Comparison of the computational effort required by the proposed algorithms

4.6.5 Observations

The numerical results listed in Tables A.1-A.8 indicate that proposed algorithms are quite efficient in solving semi-infinite variational inequality problems. Despite the fact that we used the same parameter settings without optimizing them for specific problems, most test problems are solved after just a small number of iterations. Figure 4.8 shows the relative computational effort required for all three algorithms. According to this figure, EACCPM is the computationally most expensive algorithm compared to discretization-based smoothing method and exchange method. As Figure 4.9 illustrates, for the same problems accuracy results are comparable.

In comparison to results reported in [15, 38], our discretization-based smoothing algorithm requires fewer discretization points to find the same quality solutions to both linear and nonlinear problems. If we compare our discretization-based smoothing algorithm with the exchange method, the latter results in relatively better precision in approximating the solution at the expense of complexity in implementation. With respect to EACCPM, the

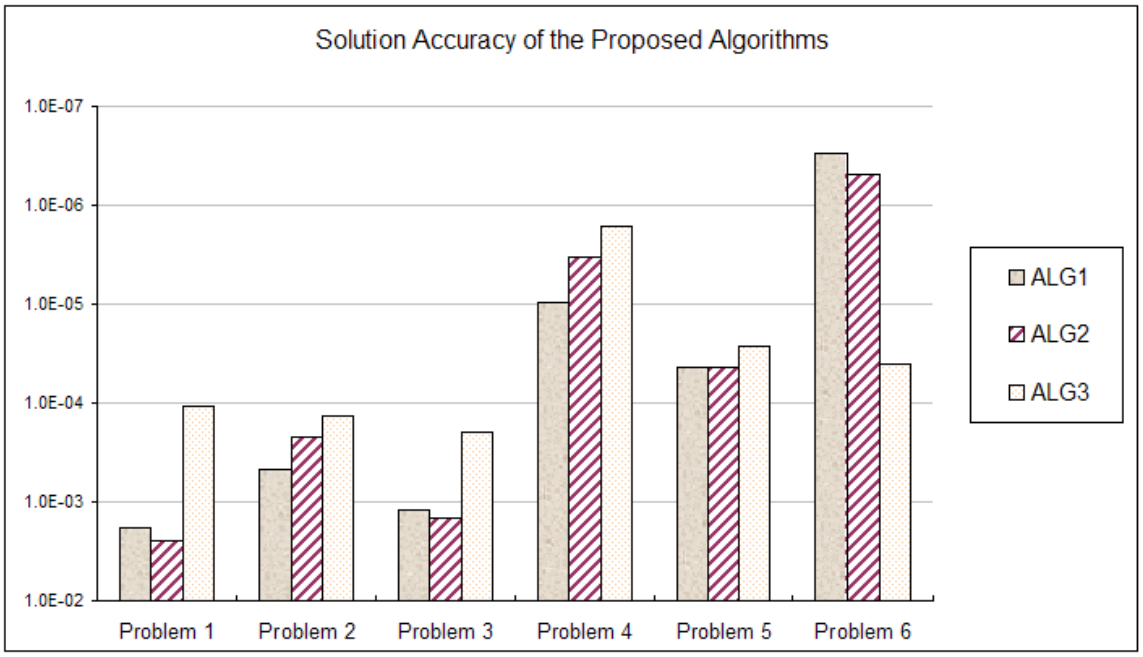

Figure 4.9: Comparison of the solution accuracy among the proposed algorithms

discretization method is computationally much cheaper and easier to implement for Problems 4-6.

The proposed exchange algorithm is computationally more advantageous than the discretization since a much smaller system of equalities are needed to be solved at each iteration. Also, the progressive way feasible domain is constructed helps to obtain better quality solutions. As for EACCPM and the exchange method based on Problems 4-6, EACCPM has obviously taken less number of iterations but more computational effort.

The EACCPM is implemented only for solving linear semi-infinite variational problems. Despite the ease due to linearity, as expected, EACCPM has required tremendous amount of computational effort compared to first two approaches. However, it has achieved high precision in approximating the optimal solution.

Considering all the results, we can conclude that for relatively small size semi-infinite

variational problems, the discretization is a viable approach. However, for any problem, the suggested approach is to use the exchange method. Utilizing EACCPM is suggested for only small size, linearly constrained and strongly monotone problems.

Chapter 5

Summary and Future Research Directions

In this chapter, we summarize the contributions of this dissertation and present future research directions.

5.1 Summary of Dissertation

In this dissertation, we have developed new approaches for solving semi-infinite variational inequality problem defined as:

$$\begin{aligned} \text{SIVI}(X, F): \quad & \text{Find a vector } x^* \in X \text{ such that} \\ & F(x^*)^\intercal (x - x^*) \geq 0 \ \text{ for all } x \in X. \end{aligned} \tag{5.1}$$

where

$$X = \{x \in R^n \mid \ g(x,t) \geq 0, \quad \text{for all } \ t \in T\}, \tag{5.2}$$

$F : R^n \to R^n$ and $g(x,t) : R^n \times R^l \to R^p$ are continuous functions and $T \subseteq R^l$ is an index set with $|T| = \infty$.

In Chapter 3, the solution analysis of SIVI(X, F) was provided in reference to nonlinear and semi-infinite programming. Equivalent reformulations that enable the implementation of the well-known solution methods have been introduced.

Based on the solution analysis and the equivalent reformulations, several solution methodologies for solving semi-infinite variational inequality problems were proposed in Chapter 4.

First, we proposed to solve a sequence of variational inequality problems to approximate the solution of SIVI(X, F). For each subproblem, its equivalent KKT formulation is smoothed and solved by Newton's method. The convergence proof for the algorithm is included. The numerical experiments have shown that the proposed algorithm is efficient in finding solutions of the same quality to both linear and nonlinear problems with by taking a few discretization points. We have also provided a new smoothing function for the $\max\{0, x\}$ operator. As our second approach we introduced an exchange method to solve semi-infinite variational inequality problems by generating a sequence of finite problems. At each iteration, nonlinear constraints generated at multiple points are added at each iteration. Our numerical experiments have shown that the proposed algorithm is more efficient than the pure discretization approach in finding solutions to both linear and nonlinear problems the same quality. Finally, we proposed an entropic analytic center cutting plane algorithm with some numerical results.

5.2 Future Research Directions

Some future directions of this research are as follows:

1. Solving VI(X, F). We focused on the KKT conditions of variational inequality problems and used standard Newton's method. The assumptions on the mapping F might be relaxed and proposed algorithms can be developed by implementing other methods for solving VI(X, F). Such methods include interior point algorithms based on potential reduction, path-following non-interior and homotopy method.

2. Generalization of the index set T. The index set can be generalized to

$$T \equiv \{t \in R^l \mid c_r(t) \leq 0, r = 1, \ldots, \bar{r}\},$$

where $c_r : T \to R, r = 1, \ldots, \bar{r}$, are C^2 functions. In this case, the domain X may

not be compact and several difficulties may arise in existence and stability of the solution. We might be able address these issues in connection to bilevel programming and generalized semi-infinite programming.

3. Parametric semi-infinite variational inequality problem. By introducing the parameter $t \in T$ into the mapping F, a parametric version of the problem SIVI$(X, F(t))$ might be studied. This extension would allow us to model many time-dependent equilibrium problems by semi-infinite variational inequality problems.

4. Duality. The Lagrangian dual for SIVI(X, F) might provide additional insights for solution analysis and algorithm development. An explicit form of Fenchel's dual has been derived variational inequalities with box constraints. This result could be extended to SIVI(X, F).

5. Discretization. The discretization method can be refined by using error-based adaptive discretization or entropy-based discretization.

6. Smoothing. We have provided a new smoothing function for the $\max\{0, x\}$ operator that results in nonsingular Jacobian. Further implementation of this approximation function for various complementarity type constraints may be investigated.

List of References

[1] Auchmuty, G. Variational principles for variational inequalities. *Numerical Functional Analysis and Optimization 10* (1989), 863–874.

[2] Auslender, A., and Teboulle, M. Lagrangian duality and related multiplier methods for variational inequality problems. *SIAM J. on Optimization 10*, 4 (2000), 1097–1115.

[3] Billups, S., Dirske, S., and Ferris, M. A comparison of algorithms for large scale mixed complementarity problems. *Computational Optimization and Applications 7* (1997), 3–26.

[4] Bröcker T., and Lander L. *Differentiable Germs and Catastrophes*, vol. 17 of *London Mathematical and Social Lecture Series*. Kluwer, 1978.

[5] Chen, S.-H. *Solving Sysytems of Linear Inequalities*. PhD thesis, North Carolina State University, Raleigh, North Carolina, 2001.

[6] Chen, S.-H., Fang, S.-C., and Tsao, H.-S. J. Entropic perturbation method for solving a system of linear inequalities. *Journal of Computational and Applied Mathematics*.

[7] Chen, X., Qi, L., and Sun, D. Global and superlinear convergence of the smoothing Newton method and its application to general box-constrained variational inequalities. *Mathematics of Computation 67* (1998), 519–540.

[8] Coope, I. D., and Watson, G. A. A projected Lagrangian algoritm for semi-infinite programming. *Mathematical Programming 32* (1985), 337–356.

[9] Danskin, J. M. The theory of max-min, with applications. *SIAM Journal on Applied Mathematics 14*, 4 (1966), 641–664.

[10] Denault, M., and Goffin, J.-L. Solving variational inequalities with a quadratic cut method: a primal-dual, jacobian-free approach. *Computational Operations Research 31*, 5 (2004), 721–743.

[11] Dunn, J. On the convergence of projected gradient processes to singular critical points. *Journal of Optimization Theory and Applications 56* (1987), 203–216.

[12] Facchinei, F., and Pang, J.-S. *Finite-Dimensional Variational Inequalities and Complementarity Problems*. Springer Verlag, New York, 2003.

[13] Fang, S.-C., Han, J., Huang, Z.-H., and Birbil, I. On the finite termination of an entropy function based non-interior continuation method for vertical linear complementarity problems. *Journal of Global Optimization 33*, 3 (2005), 369–391.

[14] Fang, S.-C., and Wu, S.-Y. An inexact approach to solving linear semi-infinite programming problems. *Optimization 28* (1994), 291–299.

[15] Fang, S.-C., Wu, S.-Y., and Birbil, S. I. Solving variational inequalities defined on a domain with infinitely many linear constraints. *Computational Optimization and Applications*, to appear.

[16] Fang, S.-C., Wu, S.-Y., and Sun, J. An analytic center cutting plane method for solving semi-infinite variational inequality problems. *Journal of Global Optimization 28* (2004), 141–152.

[17] Fischer, A., and Jiang, H. Merit functions for complementarity and related problems: A survey. *Computational Optimization and Applications 17*, 2-3 (2000), 159–182.

[18] Fukushima, M. Equivalent differentiable optimization problems and descent methods for asymmetric variational inequality problems. *Mathematical Programming 53*, 1 (1992), 99–110.

[19] Han, J., Huang, Z. H., and Fang, S.-C. Solvability of variational inequality problems. *Journal of Optimization Theory and Applications 122*, 3 (2004), 501–520.

[20] Harker, P. *Lectures on Computation of Equilibria with Equation-Based Methods: Applications to the Analysis of Service Economics and Operations*. CORE Lecture Series, Universite Catholique de Louvain, Louvain-la-Neuve, 1993.

[21] Harker, P. T., and Pang, J.-S. Finite-dimensional variational inequality and nonlinear complementarity problems: A survey of theory, algorithms and applications. *Mathematical Programming 48* (1990), 161–220.

[22] Hartman, P., and Stampacchia, G. On some nonlinear elliptic differential functional equations. *Acta Mathematica 115* (1966), 153–188.

[23] Hettich, R., and Jongen, H. *Semi-Infinite Programming: Conditions of Optimality and Applications*. Lecture Notes in Control and Information Science 7. Springer, 1978, pp. 1–11.

[24] Hettich, R., and Kortanek, K. O. Semi-infinite programming: theory, methods, and applications. *SIAM Rev. 35*, 3 (1993), 380–429.

[25] Hettich, R., and Still, G. Second-order optimality conditions for generalized semi-infinite programming problems. *Optimization 34* (1995), 195–211.

[26] Iusem, A., and Svaiter, B. A variant of Korpelevich's method for variational inequalities with a new search strategy. *Optimization 42* (1997), 309–321.

[27] J-L. Goffin, P. Marcotte, D. Z. An analytic center cutting plane method for pseudo-monotone variational inequalities. *Operations Research Letters 20* (1997), 1–6.

[28] Jian, J.-B., Mo, X.-D., and Li, J.-L. Explicit and implicit continuation algorithms for strongly monotone variational inequalities with box constraints. *Journal of Global Optimization 29*, 1 (2004), 83–25.

[29] Jongen, H. T., and Rückmann, J. *On the stability and deformation in semi-infinite optimization*, vol. 25 of *Nonconvex Optimization and Applications*. Kluwer, 1998, pp. 29–67.

[30] Jongen, H. T., Rückmann, J.-J., and Stein, O. Generalized semi-infinite optimization: a first order optimality condition and examples. *Mathematical Programming 83*, 1 (1998), 145–158.

[31] Jongen, H. T., Twilt, F., and Weber, G. W. Semi-infinite optimization structure and stability of the feasible set. *J. Optim. Theory Appl. 72*, 3 (1992), 529–552.

[32] Jongen H.T., and Zwier G. *On the Local Structure of the Feasible Set in Semi-Infinite Optimization*, vol. 72 of *International Series of Numerical Mathematics, Parametric Optimization and Approximations*. Birkhauser-Verlag, 1984, pp. 185–202.

[33] Jongen H.T., and Zwier G. *On Regular Semi-Infinite Optimization*, vol. 259 of *Lecture Notes in Economics and Mathematical Systems*. Springer-Verlag, 1985, pp. 53–64.

[34] Kanzow, C., and Jiang, H. A continuation method for (strongly) monotone variational inequalities. *Mathematical Programing 81*, 1 (1998), 103–125.

[35] Kelley, C. T. *Solving Nonlinear Equations with Newton's Method*. SIAM, 2002.

[36] Larsson, T., and Patriksson, M. A class of gap functions for variational inequalities. *Mathematical Programming 64*, 1 (1994), 53–79.

[37] Li, W., Nahak, C., and Singer, I. Constraint qualifications for semi-infinite systems of convex inequalities. *SIAM J. on Optimization 11*, 1 (2000), 31–52.

[38] Lin, C.-J., Yang, E. K., and Fang, S.-C. Implementation of an inexact approach to solving linear semi-infinite programming problems. *Journal of Computational and Applied Mathematics 61* (1995), 87–103.

[39] Lüthi, H.-J. On the solution of variational inequalities by the ellipsoid method. *Mathematics of Operations Research 10* (1985), 515–522.

[40] M. Bazaara, H. Sherali, C. *Nonlinear Programming: Theory and Algorithms*. Wiley, 1993.

[41] Mangasarian, O. Equivalence of the complementarity problem to a system of nonlinear equations. *SIAM Journal on Applied Mathematics 31* (1976), 89–92.

[42] Marcotte, P., and Wu, J. H. On the convergence of projection methods: Application to the decomposition of affine variational inequalities. *Journal of Optimization Theory and Applications 85*, 2 (1995), 347–362.

[43] Nesterov, Y., and Nemirovskii, A. *Interior-point polynomial algorithms in convex programming*. Society for Industrial and Applied Mathematics, 1994.

[44] Nesterov, Y., and Vial, J.-P. Homogeneous analytic center cutting plane methods for convex problems and variational inequalities. *SIAM Journal on Optimization 9*, 3 (1999), 707–728.

[45] Pang, J.-S. Newton's method for b-differentiable equations. *Mathematics of Operations Research 15* (1990), 311–341.

[46] Pang, J.-S. A b-differentiable equation-based, globally and locally quadratically convergent algorithm for nonlinear programs, complementarity and variational inequality problems. *Mathematical Programming 51* (1991), 101–131.

[47] Pang, J.-S., and Chan, D. Iterative methods for variational and complementarity problems. *Mathematical Programming 24* (1982), 284–313.

[48] Peng, J.-M. Equivalence of variational inequality problems to unconstrained minimization. *Mathematical Programming 78*, 3 (1997), 347–355.

[49] Qi, L., and Jiang, H. Semismooth Karush-Kuhn-Tucker equations and convergence analysis of Newton and quasi-Newton methods for solving these equations. *Mathematics of Computation 22* (1997), 301–325.

[50] R. Reemtsen, and J-J. Rückmann. *Semi-Infinite Programming*, vol. 25 of *Nonconvex Optimization and Its Applications*. Kluwer, 1998.

[51] Shapiro, A. *First and Second Order Optimality Conditions and Perturbation Analysis of Semi-Infinite Programming Problems*, vol. 25 of *Nonconvex Optimization and Applications*. Kluwer, 1998, pp. 103–133.

[52] Stein, O. First-order optimality conditions for degenerate index sets in generalized semi-infinite optimization. *Mathematics of Operations Research 26*, 3 (2001), 565–582.

[53] Stein, O., and Still, G. On optimality conditions for generalized semi-infinite programming problems. *Journal of Optimization Theory and Applications 104*, 2 (2000), 443–458.

[54] Still, G. Generalized semi-infinite programming: Theory and methods. *European Journal of Operational Research 119* (1999), 301–313.

[55] Taji, K., and Fukushima, M. A new merit function and a successive quadratic programming algorithm for variational inequality problems. *SIAM Journal on Optimization 6*, 3 (1996), 704–713.

[56] Tseng, P. Global linear convergence of a path-following algorithm for some monotone variational inequality problems. *Journal of Optimization Theory and Applications 75*, 2 (1992), 265–279.

[57] Vazquez, F. G., and Rückmann, J.-J. Extensions of the kuhn-tucker constraint qualification to generalized semi-infinite programming. *SIAM Journal on Optimization 15*, 3 (2005), 926–937.

[58] Watson, G. A. *Numerical Experiments with Globally Convergent Methods for Semi-Infinite Programming Problems*. Lecture notes in Economics and Mathematical Systems no. 215. Springer Verlag, Berlin, 1983, pp. 193–205.

[59] Wierzbicki, A. Note on the equivalence of Kuhn-Tucker complementarity conditions to an equation. *Journal Optimization Theory and Applications 37* (1982), 401–405.

[60] Wu, Z.Y., and Sun, C. R. An unconstrained convex programming approach to semi-infinite convex programming. *Personal communication* (2004).

[61] Xiao, B., and Harker, P. T. A nonsmooth newton method for variational inequalities, I: Theory. *Mathematical Programming 65* (1994), 151–194.

[62] Xiao, B., and Harker, P. T. A nonsmooth newton method for variational inequalities, II: Numerical results. *Mathematical Programming 65* (1994), 195–216.

[63] Yang, Y., Li, D., and Zhou, S. A trust region method for a semismooth reformulation to variational inequality problems. *Optimization Methods and Software 14* (2000), 139–157.

[64] Zhao, Y.-B. Exceptional families and finite-dimensional variational inequalities over polyhedral convex sets. *Applied Mathematics and Computations 87*, 2-3 (1997), 111–126.

[65] Zhao, Y. B., and Han, J. Exceptional family of elements for a variational inequality problem and its applications. *Journal of Global Optimization 14*, 3 (1999), 313–330.

Appendix A

Computational Results

Table A.1: Results for Problems (1-3) by implementing discretization-based smoothing method

Problem 1.

N	x^*	No. of Newton Iterations	$\| x^* - x_b \|$
10	(0.49208692, 0.56903430, 0.63712849, 0.69579181, 0.74546035, 0.78705650, 0.82166529)	8	0.02357289
20	(0.51062811, 0.57280822, 0.62980205, 0.68101656, 0.72637533, 0.76611237, 0.80063657)	79	0.02217797
40	(0.49202283, 0.56469563, 0.63066649, 0.68899934, 0.73961394, 0.78293770, 0.81965483)	107	0.01267899
80	(0.49885431, 0.56752794, 0.63006102, 0.68568466, 0.73432506, 0.77633042, 0.81227225)	586	0.00226087
100	(0.49838561, 0.56725780, 0.62999654, 0.68580674, 0.73460208, 0.77672756, 0.81275632)	598	0.00184365

$x_0 = (0.5, 0.5, 0.9, 0.9, 0.8, 0.9, 1)$, $\varepsilon = 1.0E-8$, $x_b = (0.4987, 0.5663, 0.6303, 0.6856, 0.7340, 0.7781, 0.8128)$

Problem 2.

N	x^*	No. of Newton Iterations	$\| x^* - x_b \|$
10	(0.51088017, 0.53642979, 0.55528904, 0.56877417, 0.57820274, 0.58469377, 0.58911555)	10	0.00374726
20	(0.50742190, 0.53680129, 0.55762495, 0.57180717, 0.58121077, 0.58733747, 0.59128402)	13	0.00359266
40	(0.50770281, 0.53620147, 0.55664611, 0.57076761, 0.58027311, 0.58656269, 0.59067783)	19	0.00154749
80	(0.50800968, 0.53604204, 0.55626314, 0.57032151, 0.57985138, 0.58620308, 0.59038960)	42	0.00063975
100	(0.50808312, 0.53602056, 0.55619460, 0.57023799, 0.57977076, 0.58613346, 0.59033325)	112	0.00047173

$x_0 =$(0.5, 0.5, 0.5, 0.9, 0.8, 1, 1), $\varepsilon = 1.0E-10$, x_b=(0.5082, 0.5360, 0.5560, 0.5700, 0.5797, 0.5859, 0.5901)

Problem 3.

N	x^*	No. of Newton Iterations	$\| x^* - x_b \|$
10	(0.27967947, 0.47840552, 0.71094465, 0.87854291, 0.95651950, 0.98534271, 0.99514082)	13	0.02317129
20	(0.27680570, 0.48023768, 0.72206741, 0.89099856, 0.96408779, 0.98888098, 0.99661103)	13	0.00427876
40	(0.27591825, 0.48099317, 0.72591038, 0.89517332, 0.96657183, 0.99002668, 0.99708229)	20	0.00267255
80	(0.27656824, 0.47969865, 0.72285949, 0.89280972, 0.96549740, 0.98963729, 0.99695483)	36	0.00218477
100	(0.27632317, 0.48009722, 0.72393830, 0.89372217, 0.96594783, 0.98981400, 0.99701657)	67	0.00119504

x_0=(0.5, 0.5, 0.5, 0.5, 0.8, 0.9, 0.9), $\varepsilon = 1.0E-10$, x_b=(0.2760,0.4802, 0.7246, 0.8930, 0.9664, 0.9902, 0.9970)

Table A.2: Results for Problems (4-6) by implementing discretization-based smoothing method

Problem 4.

N	x^*	f^*	No. of Newton Iterations	$\| f^* - f_b \|$
10	(0.04012066, 0.87713001, 0.00000042, 0.00000038, 0.00000037,...,0.00000048)	0.47868693	129	0.00074307
20	(0.04051929, 0.87748011, 0.00000042, 0.00000038, 0.00000036,...,0.00000048)	0.4792606	471	0.00016940
40	(0.04060726, 0.87755736, 0.00000042, 0.00000038, 0.00000036,...,0.00000048)	0.4793872	1557	0.00004280
80	(0.08478749, 0.77968629, 0.00000000, 0.00000000, 0.00000000,...,0.00000000)	0.47939789	1831	0.00003211
100	(0.04063040, 0.87757765, 0.00000042, 0.00000038, 0.00000036,...,0.00000048)	0.47942049	2687	0.00000951

$x_0 = (1, 1, 1, \ldots, 1)$, $\varepsilon = 1.0E-8$, f_b= 0.47943000

Problem 5.

N	x^*	f^*	No. of Newton Iterations	$\| f^* - f_b \|$
10	(1.00339404, 0.95249143, 0.67170118, 0.00000001, 0.00000001,..., 0.00967502)	1.71828248	39	0.00004359
20	(1.00000000, 1.00013531, 0.49843014, 0.17209655, 0.03527334,..., 0.07162874)	1.71828225	485	0.00004382
40	(1.00000275, 0.99989093, 0.50138652, 0.16010668, 0.05446769,..., 0.00000009)	1.71828183	31	0.00004424
80	(1.00000001, 0.99999999, 0.50000008, 0.16666583, 0.04167124,..., 0.00000007)	1.71828183	37	0.00004424
100	(1.00000001, 0.99999999, 0.50000025, 0.16666431, 0.04167722,..., 0.00000020)	1.71828183	56	0.00004424

$x_0 = (1, 1, 1, \ldots, 1)$, $\varepsilon = 1.0E-8$, f_b= 1.71832607

Problem 6.

N	x^*	f^*	No. of Newton Iterations	$\| f^* - f_b \|$
10	(0.50241134, 0.21881863, 0.21766635, 0.000000005, 0.00000000,..., 0.009780431)	0.693143294	136	0.000004406
20	(0.50000003, 0.24999868, 0.12502939, 0.062196701, 0.03285327,..., 0.000122779)	0.693147612	385	0.000000088
40	(0.50000049, 0.24999836, 0.12503788, 0.062128601, 0.03311006,..., 0.000103221)	0.693147654	12	0.000000046
80	(0.50000000, 0.25000000, 0.12500013, 0.062498419, 0.03126174,..., 0.000000183)	0.693147668	19	0.000000032
100	(0.50000000, 0.25000000, 0.12500010, 0.062498395, 0.03126168,..., 0.000000097)	0.693147671	21	0.000000029

$x_0 = (1, 1, 1, \ldots, 1)$, $\varepsilon = 1.0E-8$, f_b= 0.6931477

Table A.3: Results for Problems (7-9) by implementing discretization-based smoothing method

Problem 7.

N	x^*	f^*	No. of Newton Iterations	$\| x^* - x_b \|$	$\| f^* - f_b \|$
5	(0.71996141, –1.45048733)	97.158852	9	0.00000089	0.00004778
10	(0.71996141, –1.45048733)	97.158852	10	0.00000089	0.00004778
20	(0.71996141, –1.45048733)	97.158852	13	0.00000089	0.00004778
40	(0.71996141, –1.45048733)	97.158852	16	0.00000089	0.00004778
80	(0.71996141, –1.45048733)	97.158852	21	0.00000089	0.00004778
100	(0.71996141, –1.45048733)	97.158852	23	0.00000089	0.00004778
	$x_0 = (1, -1)$, $\varepsilon = 1.0E-10$, x_b=(0.719962, –1.450488) , f_b= 97.158900				

Problem 8.

N	x^*	f^*	No. of Newton Iterations	$\| x^* - x_b \|$	$\| f^* - f_b \|$
5	(–0.095310183, 0.095310177)	2.2	10	0.000000003	0.000000007
10	(–0.095310183, 0.095310177)	2.2	12	0.000000003	0.000000007
20	(–0.095310183, 0.095310177)	2.2	15	0.000000003	0.000000007
40	(–0.095310183, 0.095310177)	2.2	20	0.000000003	0.000000007
80	(–0.095310183, 0.095310177)	2.2	24	0.000000003	0.000000007
100	(–0.095310183, 0.095310177)	2.2	27	0.000000003	0.000000007
	$x_0 = (1, 1)$, $\varepsilon = 1.0E-10$, x_b=(–0.09531018, 0.09531018) , f_b= 2.20000				

Problem 9.

N	x^*	f^*	No. of Newton Iterations	$\| x^* - x_b \|$	$\| f^* - f_b \|$
5	(–0.74999983, –0.61803397)	0.19446598	10	0.00000040	0.000000018
10	(–0.74999961, –0.61803397)	0.19446598	10	0.00000040	0.000000018
20	(–0.74999918, –0.61803397)	0.19446598	13	0.00000083	0.000000018
40	(–0.74999832, –0.61803397)	0.19446598	16	0.00000168	0.000000018
80	(–0.74999661, –0.61803397)	0.19446598	87	0.00000339	0.000000018
100	(–0.74999575, –0.61803397)	0.19446598	121	0.00000425	0.000000018
	$x_0 = (1, -1)$, $\varepsilon = 1.0E-8$, x_b=(0.750000, –0.618034) , f_b= 0.194466				

Table A.4: Results for Problems (1-3) by implementing exchange method

Problem 1.

N	t_l	x^*	# Iter.	$g(x,t)$	$\|\|x^* - x_b\|\|$
2	-	(0.7142857, 0.7142857, 0.7142857, 0.7142857, 0.71428571, 0.71428571, 0.71428571)	6	-3.46E-01	0.3006982
3	0.82328	(0.4942454, 0.5655634, 0.6303828, 0.6878217, 0.73779711, 0.78070562, 0.81719019)	6	-4.34E-04	0.0081123
4	0.83202	(0.5012893, 0.5685142, 0.6298373, 0.6845267, 0.73249667, 0.77406133, 0.80974997)	7	-1.09E-04	0.0063912
5	0.82768	(0.4978755, 0.5670460, 0.6300357, 0.6860427, 0.73498153, 0.77720318, 0.81328896)	7	-2.37E-05	0.0018732
6	0.82971	(0.4995519, 0.5677570, 0.6299214, 0.6852779, 0.73374049, 0.77564134, 0.81153535)	7	-5.92E-06	0.0032875
7	0.82870	(0.4987303, 0.5674065, 0.6299739, 0.6856483, 0.73434410, 0.77640245, 0.81239105)	8	-1.41E-06	0.0021216
8	0.82919	(0.4991418, 0.5675816, 0.6299467, 0.6854616, 0.73404061, 0.77602015, 0.81196154)	8	-3.50E-07	0.0026480
9	0.82894	(0.4989395, 0.5674954, 0.6299598, 0.6855532, 0.73418957, 0.77620788, 0.81217252)	10	-8.86E-08	0.0023694
10	0.82906	(0.4990410, 0.5675386, 0.6299532, 0.6855072, 0.73411473, 0.77611359, 0.81206657)	12	-2.10E-08	0.0025053
11	0.82900	(0.4989938, 0.5675185, 0.6299563, 0.6855285, 0.73414952, 0.77615743, 0.81211584)	25	-5.36E-09	0.0024409

x_0=(0.4,0.5,1,0.9,0.8,0.9,1), x_b=(0.4987,0.5663,0.6303,0.6856,0.7340,0.7781,0.8128)

Problem 2.

N	t_l	x^*	# Iter.	$g(x,t)$	$\|\|x^* - x_b\|\|$
2	-	(0.5981935, 0.5981935, 0.5981935, 0.5981935, 0.5981935, 0.5981935, 0.5981935)	10	-1.63E-01	0.1228848
3	0.66819	(0.5111796, 0.5365309, 0.5552723, 0.5686987, 0.5781063, 0.5845975, 0.5890297)	12	-2.07E-03	0.0040945
4	0.64740	(0.5081641, 0.5360016, 0.5561252, 0.5701519, 0.5796869, 0.5860605, 0.5902739)	34	-2.19E-06	0.0003102
5	0.64808	(0.50822628, 0.5359902, 0.5560766, 0.5700906, 0.5796268, 0.5860082, 0.5902312)	60	-1.06E-08	0.0002216
6	0.64805	(0.5082232, 0.5359907, 0.5560788, 0.5700934, 0.5796296, 0.5860106, 0.5902332)	55	-9.73E-10	0.0002247

x_0 =(0.4,0.5,0.1,0.1,0.8,0.2,1), x_b=(0.5082, 0.5360, 0.5560, 0.5700, 0.5797, 0.5859, 0.5901)

Problem 3.

N	t_l	x^*	# Iter.	$g(x,t)$	$\|\|x^* - x_b\|\|$
2	-	(0.3333334, 0.8333333, 0.8333333, 0.8333333, 0.8333333, 0.8333333, 0.8333333)	7	-2.39E-01	0.4609826
3	0.61970	(0.3333334, 0.5692340, 0.6651175, 0.7530856, 0.8264726, 0.8827118, 0.9229933)	9	-7.32E-02	0.2661536
4	0.29156	(0.2766477, 0.4795393, 0.7224529, 0.8924686, 0.9653288, 0.9895705, 0.9969309)	10	-2.97E-06	0.0021551
5	0.28885	(0.2764521, 0.4798756, 0.7233455, 0.8932215, 0.9657012, 0.9897181, 0.9969841)	16	-7.49E-07	0.0008500
6	0.29021	(0.2763659, 0.4800223, 0.7237407, 0.8935575, 0.9658686, 0.9897848, 0.9970084)	18	-1.93E-07	0.0002985
7	0.29081	(0.2764125, 0.4799417, 0.7235264, 0.8933768, 0.9657792, 0.9897494, 0.9969956)	19	-5.94E-08	0.0006075
8	0.29048	(0.2764126, 0.4799415, 0.7235262, 0.8933767, 0.9657792, 0.9897495, 0.9969956)	20	-3.77E-08	0.0006033

x_0=(0.2, 0.4, 0.5, 0.9, 0.8, 0.9, 0.9), x_b=(.27632, 0.48009, 0.72394, 0.89372, 0.96595, 0.98981,0.99702)

Table A.5: Results for Problems (4-6) by implementing exchange method

Problem 4.

N	t_l	x^*	# Iter.	$g(x,t)$	f^*	$\|\|f^* - f_b\|\|$
3	-	(0.117400, 0.724050, 0.000000, 0.000000,..., 0.000000)	21	-2.12E-02	0.479426	0.0000039
4	0.76114	(0.076724, 0.805401, 0.000000, 0.000000,..., 0.000000)	52	-5.02E-03	0.479426	0.0000039
5	0.63444	(0.058040, 0.842769, 0.000000, 0.000000,..., 0.000000)	82	-1.21E-03	0.479426	0.0000039
6	0.56839	(0.049163, 0.860524, 0.000000, 0.000000,..., 0.000000)	108	-2.98E-04	0.479426	0.0000039
7	0.53451	(0.044842, 0.869166, 0.000000, 0.000000,..., 0.000000)	260	-7.36E-05	0.479426	0.0000039
8	0.51729	(0.042699, 0.873451, 0.000000, 0.000000,..., 0.000000)	257	-1.82E-05	0.479426	0.0000039
9	0.50855	(0.041606, 0.875637, 0.000000, 0.000000,..., 0.000000)	111	-4.43E-06	0.479426	0.0000039
10	0.50405	(0.040997, 0.876856, 0.000000, 0.000000,..., 0.000000)	303	-1.05E-06	0.479426	0.0000039
11	0.50152	(0.040486, 0.877877, 0.000000, 0.000000,..., 0.000000)	357	-5.90E-07	0.479426	0.0000039
12	0.49939	(0.040486, 0.877877, 0.000000, 0.000000,..., 0.000000)	201	-5.90E-07	0.479426	0.0000039
13	0.49939	(0.040734, 0.877383, 0.000000, 0.000000,..., 0.000000)	412	4.10E-08	0.479428	0.0000033

$x_0 = (0.4, 0.4, 0.4, \ldots, 0.4)$, $f_b = 0.47943000$

Problem 5.

N	t_l	x^*	# Iter.	$g(x,t)$	f^*	$\|\|f^* - f_b\|\|$
3	-	(1.000000, 1.297442, 0.000000, 0.000000,..., 0.420837)	17	-3.20E-01	1.657138	0.0611880
4	0.943621	(1.000000, 1.297416, 0.000000, 0.000000,..., 0.000000)	19	-1.73E-01	1.700066	0.0182600
5	0.839822	(1.000000, 0.909667, 0.775511, 0.000000,..., 0.000000)	28	-8.45E-03	1.716479	0.0018470
6	0.204683	(1.000000, 0.998976, 0.509200, 0.140886,..., 0.000060)	39	-3.39E-05	1.718276	0.0000502
9	0.07902, 0.67049, 0.80000	(1.000000, 0.999999, 0.500023, 0.166483,..., 0.000000)	45	-3.56E-08	1.718282	0.0000442

$x_0 = (1, 1, 1, 0.1, \ldots, 0.1)$, $f_b = 1.71832607$

Problem 6.

N	t_l	x^*	# Iter.	$g(x,t)$	f^*	$\|\|f^* - f_b\|\|$
3	-	(0.500000, 0.333333, 0.000000, 0.000000,..., 0.1666647)	13	-1.22E-01	0.670000	0.0231480
4	0.94595	(0.500000, 0.333328, 0.000000, 0.000000,..., 0.0000001)	15	-6.46E-02	0.685726	0.0074213
5	0.84846	(0.500000, 0.198655, 0.269284, 0.000000,..., 0.0000001)	17	-5.07E-03	0.692136	0.0010116
7	0.21182, 0.87590	(0.500000, 0.246524, 0.152961, 0.000000,..., 0.0004389)	31	-2.00E-04	0.693109	0.0000379
10	0.08475, 0.68486, 0.90162	(0.500000, 0.249943, 0.126196, 0.054829,..., 0.0000465)	42	-3.42E-06	0.693147	0.0000011
13	0.03347, 0.36268, 0.77682	(0.500000, 0.249993, 0.125121, 0.061531,..., 0.0000019)	49	2.65E-09	0.693147	0.0000005

$x_0 = (0.3, 0.3, 0.3, \ldots, 0.3)$, $f_b = 0.6931477$

Table A.6: Results for Problems (7-9) by implementing exchange method

Problem 7.

N	t_l	x^*	f^*	Iter#	$g(x,t)$	$\|\|x^* - x_b\|\|$	$\|\|f^* - f_b\|\|$
2	-	(0.7199613878, -1.4504873307)	97.15885	8	-2.22E-16	9.07E-07	-4.76E-05
	x_0 =(1,-1) , x_b=(0.719962, -1.450488) , f_b= 97.1589						

Problem 8.

N	t_l	x^*	f^*	Iter#	$g(x,t)$	$\|\|x^* - x_b\|\|$	$\|\|f^* - f_b\|\|$
2	-	(-0.09531018005, 0.09531017956)	2.2000	8	-4.88E-10	4.42478E-10	5.37E-10
	x_0 =(0,1) , x_b=(-0.09531018, 0.09531018) , f_b= 2.20000						

Problem 9.

N	t_l	x^*	f^*	Iter#	$g(x,t)$	$\|\|x^* - x_b\|\|$	$\|\|f^* - f_b\|\|$
2	-	(-0.749999999, -0.6180339878)	0.19447	6	-2.06E-09	1.22E-08	1.01E-08
	x_0 =(1,0) , x_b=(-0.7500000,-0.618034) , f_b= 0.194466						

Table A.7: Results for Problems (1-3) by implementing entropic analytic center cutting plane method

Problem 1.

k	x^*	$\| x^* - x_b \|$	No. of Func. Evaluations	No. of Iterations
1	(0.000858, 0.001533, 0.997623, 0.999988, 0.999999, 0.999999, 0.999999)	0.9775543	180	16
2	(0.000376, 0.004995, 0.994936, 0.999705, 1.000000, 0.999991, 0.999997)	0.9747031	219	19
:	:	:	:	:
30	(0.254656, 0.437275, 0.619622, 0.765617, 0.884706, 1.000000, 1.000000)	0.4355661	1560	39
31	(0.318505, 0.479410, 0.621901, 0.731793, 0.852238, 0.925027, 1.000000)	0.3359108	1377	36
32	(0.498697, 0.566312, 0.630324, 0.685642, 0.734049, 0.778157, 0.812860)	0.0001087	1423	36
	x_b = (0.4987, 0.5663, 0.6303, 0.6856, 0.7340, 0.7781, 0.8128)			

Problem 2.

k	x^*	$\| x^* - x_b \|$	No. of Func. Evaluations	No. of Iterations
1	(0.449869, 0.449869, 0.449869, 0.449869, 0.449869, 0.449869, 0.449869)	0.3025590	172	16
2	(0.490880, 0.536430, 0.555289, 0.568774, 0.578203, 0.584694, 0.589116)	0.2751671	181	17
:	:	:	:	:
22	(0.509022, 0.535986, 0.555651, 0.569504, 0.579032, 0.585478, 0.589793)	0.0013271	691	19
23	(0.508757, 0.535965, 0.555762, 0.569669, 0.579203, 0.585633, 0.589922)	0.0009098	835	21
24	(0.508213, 0.535998, 0.556034, 0.570053, 0.579722, 0.586002, 0.590153)	0.0001338	731	19
	x_b=(0.5082, 0.5360, 0.5560, 0.5700, 0.5797, 0.5859, 0.5901)	1.71832607		

Problem 3.

k	x^*	$\| x^* - x_b \|$	No. of Func. Evaluations	No. of Iterations
1	(0.449064, 0.653264, 0.897673, 1.000000, 1.000000, 1.000000, 1.000000)	0.3202174	217	17
2	(0.449292, 0.653419, 0.897531, 0.999900, 0.999956, 0.999965, 0.999938)	0.3203082	179	15
:	:	:	:	:
51	(0.278466, 0.483070, 0.727125, 0.895478, 0.968917, 0.992711, 0.999965)	0.0069468	1060	22
52	(0.276648, 0.480648, 0.725217, 0.893628, 0.967053, 0.990832, 0.997539)	0.0015846	1151	23
53	(0.276060, 0.480301, 0.724680, 0.893062, 0.966463, 0.990272, 0.997076)	0.0001978	1024	21
	x_b=(0.2760, 0.4802, 0.7246, 0.8930, 0.9664, 0.9902, 0.9970)			

Table A.8: Results for Problems (4-6) by implementing entropic analytic center cutting plane method

Problem 4.

k	x^*	f^*	$\| f^* - f_b \|$	No. of Func. Evaluations
1	(0.44986985, 0.44986985, 0.44986985, 0.44986985, 0.44986985,,0.44986793)	2.0240566	1.5446266	1846
2	(0.03995430, 0.87894634, 0.00000000, 0.00000000, 0.00000000,,0.00000000)	0.4794275	0.0000025	2312
3	(0.03995402, 0.87894689, 0.00000000, 0.00000000, 0.00000000,,0.00000000)	0.4794275	0.0000025	2490
4	(0.03981342, 0.87922986, 0.00000000, 0.00000000, 0.00000000,,0.00000000)	0.4794283	0.0000017	2810
	$f_b = 0.47943000$			

Problem 5.

k	x^*	f^*	$\| f^* - f_b \|$	No. of Func. Evaluations
1	(0.44986618, 0.44986618, 0.44986618, 0.44986618, 0.44986618,...,0.44986589)	2.0240403	0.3057142	2479
2	(1.00004687, 0.99916957, 0.50527722, 0.15191463, 0.06032012,...,0.00000000)	1.7182876	0.0000385	2822
3	(1.00020234, 0.99749829, 0.51135534, 0.14341231, 0.06293189,...,0.00000000)	1.7182993	0.0000268	3487
4	(1.00000000, 1.00064931, 0.49619007, 0.17386080, 0.03740263,...,0.00015589)	1.7182975	0.0000286	4258
	$f_b = 1.71832607$			

Problem 6.

k	x^*	f^*	$\| f^* - f_b \|$	No. of Func. Evaluations
1	(0.44986618, 0.44986618, 0.44986618, 0.44986618, 0.44986618,...,0.44986589)	2.0240403	1.33089259	1812
2	(0.50000000, 0.25104182, 0.11752075, 0.08169277, 0.00596204,...,0.08673646)	0.6949144	0.0017667	2277
3	(0.50000000, 0.25104233, 0.11751740, 0.08169922, 0.00595954,...,0.00595954)	0.6949144	0.0004061	2440
	f_b= 0.6931477			