THE HONG KONG POLYTECHNIC UNIVERSITY
香港理工大學
Pao Yue-kong Library
包玉剛圖書館

Copyright Undertaking

This thesis is protected by copyright, with all rights reserved.

By reading and using the thesis, the reader understands and agrees to the following terms:

1. The reader will abide by the rules and legal ordinances governing copyright regarding the use of the thesis.
2. The reader will use the thesis for the purpose of research or private study only and not for distribution or further reproduction or any other purpose.
3. The reader agrees to indemnify and hold the University harmless from and against any loss, damage, cost, liability or expenses arising from copyright infringement or unauthorized usage.

IMPORTANT

If you have reasons to believe that any materials in this thesis are deemed not suitable to be distributed in this form, or a copyright owner having difficulty with the material being included in our database, please contact lbsys@polyu.edu.hk providing details. The Library will look into your claim and consider taking remedial action upon receipt of the written requests.

Pao Yue-kong Library, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong

http://www.lib.polyu.edu.hk

QUANTIFICATION AND CONVERGENCE ANALYSIS OF TWO-STAGE STOCHASTIC VARIATIONAL INEQUALITY PROBLEMS

JIE JIANG

PhD

The Hong Kong Polytechnic University

This programme is jointly offered by The Hong Kong Polytechnic University and Xi'an Jiaotong University

2019

The Hong Kong Polytechnic University

Department of Applied Mathematics

Xi'an Jiaotong University

School of Mathematics and Statistics

Quantification and Convergence Analysis of Two-stage Stochastic Variational Inequality Problems

Jie Jiang

A thesis submitted in partial fulfilment of the requirements

for the degree of Doctor of Philosophy

March 2019

Certificate of Originality

I hereby declare that this thesis is my own work and that, to the best of my knowledge and belief, it reproduces no material previously published or written, nor material that has been accepted for the award of any other degree or diploma, except where due acknowledgement has been made in the text.

_____________________(Signature)

______JIANG Jie______(Name of student)

Dedicate to my family.

Abstract

The thesis is concerned with two-stage stochastic variational inequality problems. Then two topics are considered:

1. Quantitative analysis for a class of two-stage stochastic linear variational inequality problems.

2. Regularized two-stage stochastic variational inequality problems for Cournot-Nash equilibrium under uncertainty.

For topic 1, we consider a class of two-stage stochastic linear variational inequality problems whose first stage problems are stochastic linear box-constrained variational inequality problems and second stage problems are stochastic linear complementary problems owning a unique solution. We first give several conditions for the existence of solutions to both the original problem and its perturbed problem. Next we derive quantitative stability assertions of this two-stage stochastic problem under the total variation metric via the corresponding residual function. After that, we study its discrete approximation problem. The convergence and the exponential rate of convergence of optimal solution sets are obtained under moderate assumptions respectively. Finally, through solving a noncooperative two-stage stochastic game of multi-player, we numerically illustrate the obtained theoretical results.

In view of the strong monotonicity of the second stage problem in topic 1, we relax this requirement to the monotonicity situation in topic 2. Specifically, for topic 2, we reformulate a convex two-stage non-cooperative multi-player game under uncertainty as a two-stage stochastic variational inequality problem where the second stage problem is just a monotone stochastic linear complementarity problem. Under standard assumptions, we provide sufficient conditions for the existence of solutions of the two-stage stochastic variational inequality problem and propose a regularized sample average approximation method for solving it. We prove the convergence of the method as the regularization parameter tends to zero and the sample size tends to infinity. Moreover, our approach is applied to a two-stage stochastic production and supply planning problem with homogeneous commodity in an oligopolistic market. Numerical results based on randomly generated data are presented to demonstrate the effectiveness of our convergence results.

Publications Arising from the Thesis

- J. Jiang, X. Chen and Z. Chen. Quantitative analysis for a class of two-stage stochastic linear variational inequality problems. *Submitted*, 2018.
- J. Jiang, Y. Shi, X. Wang and X. Chen. Regularized two-stage stochastic variational inequalities for Cournot-Nash equilibrium under uncertainty. *To appear on Journal of Computational Mathematics*, 2019.

Acknowledgements

The last several years were critical, memorable, difficult as well as rewarding for me. I went through many things and met many nice persons who helped me a lot during these years. They make me mature and lay a foundation for my further career. So I want to express my gratitude to these people around me during my joint PhD programmes in The Hong Kong Polytechnic University (PolyU). Without their help, I cannot finish this PhD thesis successfully.

First and foremost, I would like to express my heartfelt appreciation to my supervisor Professor Xiaojun Chen in PolyU. She gave me a lot of help in my school life, especially when I just came to Hong Kong. She gave me the professional guidance in my research study through time and time again academic discussion. Her expansive knowledge of optimization, inexhaustible enthusiasm and prudent scientific attitude for research have impressed me profoundly. She is undoubtedly my role model in my academic career.

Moreover, I would like to express my deep thanks to my supervisor Professor Zhiping Chen in Xi'an Jiaotong University (XJTU) for his special care, encouragement and recommendation that enabled me to come to PolyU for the joint PhD programmes. I also appreciated that he gave me the initial enlightenment in my research study and revised my first research paper word by word.

Especially, I would like to thank Professor Xiaoqi Yang to host my PhD confirmation of registration. I also thank Professor Ting Kei Pong and Professor Zaikun

Zhang for their participation and recording.

Furthermore, I would like to thank everyone I met in Professor Xiaojun Chen's group: Professor Congpei An, Professor Wei Bian, Professor Xin Liu, Professor Dingtao Peng, Professor Shulin Wu, Professor Hailin Sun, Professor Zaikun Zhang, Professor Bo Jiang, Professor Sitian Qin, Professor Chengjing Wang, Dr. Yang Zhou, Dr. Hong Wang, Dr. Lei Yang, Dr. Lili Pan, Dr. Tianxiang Liu, Dr. Liaoyuan Zeng, Dr. Lin Chen, Dr. Liang Chen, Dr. Qiyu Wang, Dr. Xiaoliang Song, Mr. Andy Yat Ming Cheung, Dr. Bo Wen, Dr. Yun Shi, Dr. Guidong Liu, Mr. Rui Zhang, Mr. Zhehua Peng, Mr. Xiaozhou Wang, Miss. Chao Li, Miss Fang He, Mr. Jianfeng Luo for their help and companion. I have benefited a lot from them during the past several years. Special thanks go to Dr. Lei Yang and Professor Hailin Sun for their kind helps in my school life and some good suggestions for my research. I also thank Dr. Bo Wen and Dr. Guidong Liu for their help when I just came to PolyU.

I would also like to thank all the friends I have met in PolyU for their friendships and help. In particular, I would like to especially acknowledge Professor Minghua Li, Professor Kaiwen Meng, Dr. Jin Yang for their kind help.

Finally, I would like to express my deepest gratitude and love to my families: my father Daowei Jiang, my mother Qiyan Luo, my wife Min Liu, my newborn son Jiayan Jiang, my sister Han Jiang for their unconditional love, encouragement, trust and support throughout my PhD study.

Contents

Bibliography 87

List of Figures

List of Tables

List of Notations

$\mathbb{N}$	set of natural numbers
$\mathbb{R}^n$	set of n-dimensional real vectors
$\mathbb{R}^{m\times n}$	set of $m \times n$ real matrices
$\mathbb{R}^n_+$	the nonnegative orthant of $\mathbb{R}^n$
$\mathbb{R}^n_{++}$	the positive orthant of $\mathbb{R}^n$
$x \geq y$	the (usual) partial ordering: $x_i \geq y_i,\ i = 1, \cdots n$
$x > y$	the (usual) partial ordering: $x_i > y_i,\ i = 1, \cdots n$
$x \perp y$	x and y are perpendicular
$\|x\|$	the ℓ_2-norm of $x \in \mathbb{R}^n$
x^T	transpose of matrix/vector x
$\langle x, y\rangle$	The inner product of vectors x, y, i.e., $x^T y$
I_n	identity matrix of dimension n
$d(a, B)$	Euclidean distance function from vector a to set B
$d(A, B)$	Euclidean deviation distance function from set A to set B, i.e., $\sup_{a\in A} d(a, B)$
$\mathrm{LCP}(q, M)$	LCP defined by the matrix M and vector q
$\mathcal{N}_X(x)$	The normal cone X at x
$\rightrightarrows$	multifunction or set-valued mapping
$\mathbb{E}$	expectation operator

$\mathbb{B}$	the closed ball centred at zero with radius one
$\mathcal{P}(\Xi)$	the set of probability distributions on the support set Ξ
$\mathcal{P}_k(\Xi)$	the set of probability measures such that $\mathbb{E}_P[\|\xi\|^k] < +\infty$

Chapter 1

Introduction

1.1 Background

Complementarity problems (CPs) and variational inequality (VI) problems (collectively known as equilibrium problems) construct an important branch in modern optimization research field, which has a wide range of important applications in engineering and economics, such as all kinds of economic/traffic equilibrium problems, as well as is a highly valued theoretical research topic, especially in describing the optimality conditions of optimization problems. Therefore, the systematic study for equilibrium problems is crucial and necessary for both practical applications and theoretical development in modern optimization. Although diverse examples of the linear CP, abbreviated LCP, can be dated back to as far as 1940s, it is believed that the systematic and concentrated study of the CP and VI started in the middle of 1960s [1, 2]. In a span of six decades, equilibrium problems are extensively investigated and tons of excellent papers, monographs are sprung out, see for example [1, 2, 3, 4, 5, 6, 7, 8, 9, 10] to name a few.

The CP consists of finding a vector that satisfies a certain system of inequalities. Let $F : \mathbb{R}^n \to \mathbb{R}^n$. Then the CP is to find a vector $x \in \mathbb{R}^n$ such that

$$0 \leq x \perp F(x) \geq 0.$$

Specifically, if $F(x) = Mx + q$, where $q \in \mathbb{R}^n$ and $M \in \mathbb{R}^{n\times n}$, we obtain the simplest class of CP, that is, the LCP, denoted by LCP(q, M). Obviously, the LCP(q, M) can be equivalently expressed as the optimal solution set of the following quadratic programming problem:

$$\begin{array}{ll} \min\limits_{x} & x^T(Mx+q) \\ \text{s.t.} & x \geq 0, Mx+q \geq 0. \end{array}$$

On the other hand, a well known context for LCPs that it can be read as the first order optimality conditions of quadratic programming problems.

Given a subset X of the Euclidean n-dimensional space $\mathbb{R}^n$ and a mapping $F : X \to \mathbb{R}^n$, the VI problem, denoted VI(X, F), is to find a vector $x \in X$ such that

$$\langle y - x, F(x)\rangle \geq 0, \text{ for all } y \in X.$$

We straightforwardly know from the definition of CP and VI that $0 \leq x \perp F(x) \geq 0$ is equivalent to VI$(\mathbb{R}^n_+, F)$. Thus, a CP can be viewed as a special case of VI problem with X being $\mathbb{R}^n_+$.

Generally, subset X is restricted to be closed and convex [1]. Especially, when X is convex, we have the following definition of normal cone:

Definition 1.1. *The normal cone to X at x, denoted by $\mathcal{N}_X(x)$, is defined by*

$$\mathcal{N}_X(x) = \{\bar{x} \in X : \langle \bar{x}, y - x\rangle \leq 0 \text{ for all } y \in X\}.$$

In this case, we can rewrite VI(X, F) as

$$-F(x) \in \mathcal{N}_X(x) \text{ or } 0 \in F(x) + \mathcal{N}_X(x).$$

Of course, by the discussion above, $0 \leq x \perp F(x) \geq 0$ can be written as

$$0 \in F(x) + \mathcal{N}_{\mathbb{R}^n_+}(x).$$

In recent decades, with more and more complex decision-making environment or to describe possibly unknown parameters in CPs and VI problems, stochastic CPs and VI problems (abbreviated SCPs and SVI problems respectively) are proposed. To present this, let $\xi : \Omega \to \Xi \subseteq \mathbb{R}^s$ denote a random vector with the probability space being $(\Omega, \mathcal{F}, \mathbb{P})$ and support set being $\Xi \subseteq \mathbb{R}^s$ which is the smallest closed subset such that $P(\Xi) = 1$ and $P = \mathbb{P} \circ \xi^{-1}$. Redefine F and X being a response of random vector ξ, i.e., $F : X \times \Xi \to \mathbb{R}^n$ and $X : \Xi \rightrightarrows \mathbb{R}^n$. Then, we obtain the following SVI problem:

$$-F(x, \xi) \in \mathcal{N}_{X(\xi)}(x),$$

as well as SCP:

$$0 \leq x \perp F(x, \xi) \geq 0.$$

When it comes to the SVI or SCP, the concept of solutions is crucial. Generally, there are two kinds of solutions: the here and now solution and the wait and see solution. For SVI problem $\mathrm{VI}(X(\xi), F(x, \xi))$, we obtain a solution $x(\xi)$ when $\xi \in \Xi$ is fixed. That is, x is a response of the random vector ξ or x is a measurable function of ξ. We call this kind of solution the wait and see solution. Intuitively, this type of solution waits the realization of random vector ξ and then we can see a solution. However, this kind of solution seems to be helpless or hardly used by decision makers because they usually make a decision when the realization of ξ is unknown. If a decision maker has a wait and see solution and he or she must make a decision before a realization, the quite possible approach is that: he or she will choose a decision $x(\xi_0)$ with realization ξ_0 such that $P(\xi_0)$ reaches its maximum. Obviously, this kind of decision is quite unstable and even meaningless when ξ has some continuous distribution. Moreover, how to obtain a response $x(\cdot)$ in a general setting is also an intractable problem. Comparatively, the here and now solution is a solution made before the realization of ξ. Usually, it is assumed that we know the distribution of ξ and we make a decision by using the comprehensive distribution information. To clarify this as well as an example, we consider the expected value

(EV) form of $\mathrm{VI}(X(\xi), F(x,\xi))$. To this end, we let $X(\xi) = X$. Then, the EV form can be formulated as $\mathrm{VI}(X, \mathbb{E}[F(x,\xi)])$, or

$$-\mathbb{E}[F(x,\xi)] \in \mathcal{N}_X(x).$$

We can see from the EV form that x has nothing to do with the specific realization of ξ, but the distribution information. This kind of solution employ a comprehensive information of ξ. So it can be relatively stable. In recent years, more and more scholars argue that the true distribution cannot be known, which leads to a very hot topic: the distributionally robust optimization problem. This is out of the range of this thesis. For more information, we refer to [11] for the distributionally robust LCP.

The shortfalls of the EV form are obvious: Firstly, it loses too much information of random vector ξ; Secondly, the here and now solution can hardly be a true solution for any realization of ξ; Thirdly, it fails to deal with the case when $X(\xi)$ is a random set-valued mapping. Considering these shortcomings, the expected residual minimization (ERM) approach is put forward, see [8]. The main idea of ERM approach is to find a solution such that it can solve the SVI problem best according to minimizing residual functions. First of all, we need to introduce residual functions (see also merit functions, gap functions). $\theta : \mathbb{R} \to \mathbb{R}^n$ is a residual function for $\mathrm{VI}(X, F(x))$, if $\theta(x) \geq 0$ for all $x \in X$ and $\theta(x) = 0$ if and only if $x \in X$ and x solves the original VI problem, see [9, Definition 1.1]. Thus, we can rewrite the original VI problem as

$$\min_{x \in X} \theta(x).$$

The residual function for VI problem is a powerful tool to transfer the VI problem to a minimization problem. Then we can employ the existing results in minimization problem to discuss the VI problem, especially for the numerical treatment aspect.

The examples of commonly-used residual functions are as follows:

$$\theta(x) := \max_{y}\{\langle F(x), x-y\rangle : y \in X\},$$

$$\theta(x) := \max_{y}\{\langle F(x), x-y\rangle - \frac{1}{2}\langle y-x, G(y-x)\rangle : y \in X\},$$

where G is some positive definite matrix. Then by adopting the residual function, we can obtain the ERM problem of VI$(X, F(x,\xi))$ as below:

$$\min_{x\in X} \mathbb{E}[\theta(x,\xi)].$$

Specially, residual functions of CPs are also known as nonlinear CP (abbreviated NCP) functions. A function $\phi : \mathbb{R}\times\mathbb{R} \to \mathbb{R}_+$ is called the NCP function, if it satisfies

$$\phi(a,b) = 0 \Leftrightarrow a \geq 0, b \geq 0, ab = 0.$$

Thus, $0 \leq x \perp F(x) \geq 0$ is equivalent to the following optimization problem:

$$\min_{x} \|\mathcal{R}(x, F(x))\|^2,$$

where $\mathcal{R} : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ with $\mathcal{R}_i(x, F(x)) := \phi(x_i, F_i(x))$ for some NCP function ϕ, and x_i and $F_i(x)$ are the ith component of x and $F(x)$, respectively.

Two well-known NCP functions are the Fischer-Burmeister (FB) function and natural NCP function denoted by ϕ_{FB} and ϕ_N respectively. Specifically, we have

$$\phi_{FB}(a,b) = a + b - \sqrt{a^2+b^2},$$
$$\phi_N(a,b) = \min\{a,b\}.$$

Moreover, the relationship between FB function and natural NCP function reads:

$$\frac{2}{2+\sqrt{2}}\,|\phi_N(a,b)| \leq |\phi_{FB}(a,b)| \leq (2+\sqrt{2})\,|\phi_N(a,b)|$$

for all $a, b \in \mathbb{R}$.

There is rich literature on algorithms to solve equilibrium problems. Usually, they employ the residual function to rewrite the VI problem or CP to a minimization problem and then use existing algorithm to solve it, such as, trust region [12], Douglas–Rachford splitting method [10], Newton method [7] and so on.

More recently, the concept of multistage SVI problems was proposed by Rockafellar and Wet in [6] when the support set are discrete. Different from the single-stage SVI problem, its solution is a sequence of vectors depending on the information flow available at each stage. To solve the multistage SVI, the Progressive Hedging Method (PHM) was adopted in [13]. However, it is still worth considering the two-stage case in the continuous distribution case. In view of this, Chen and her cooperators has done many pioneering works on the two-stage SVI problem, see for example [10, 11, 14].

In the following section, we will make a brief literature review on SVI problems.

1.2 Literature review

The deterministic variational inequality problem has been extensively investigated, see monographs [1, 15, 16] and the references therein. Recently, to describe uncertainty in the complex decision process or make a long-term decision, SVI problems have been put forward and studied increasingly. There are usual two methods to deal with the uncertainty: One is the EV form; the other one is the ERM form. For the EV form, [3] considered a sample-path solution of the EV form SVI problem and the convergence assertions were established. Much later, Xu discussed in [17] the sample average approximation (SAA) approach for the SVI problem, and under certain conditions, the convergence conclusion between the SAA problem and the original SVI problem was derived. For the ERM form, Chen and Fukushima considered in [8] the stochastic LCP by the ERM procedure. The quasi-Monte Carlo method was adopted to generate scenarios of observation and thus to obtain the discrete approximation problem. Chen, Wets and Zhang [9] investigated SVI problems by the ERM procedure, and the SAA method was employed to approximate the expected smoothing residual function. Chen, Pong and Wets [10] first investigated the ERM process for the two-stage SVI problem and solved it by Douglas-Rachford splitting method.

More recently, as an extension from single-stage case to multi-stage case, Rockafellar and Wets did a pioneering work in [18] for the multistage SVI problem when the support set is discrete, which lays a theoretical foundation for numerical solution by reformulating the multistage SVI problem in an extensive form. Closely following this work, in [13], Rockafellar and Sun employed the well-known PHM to solve the multistage SVI problem. It is worth pointing out that PHM was first introduced by Rockafellar and Wets in [19] to solve multistage stochastic programs. Chen, Pong and Wets [10] first introduced the two-stage SVI problem and an expected residual minimization procedure for solving it. Chen, Sun and Xu [11] proposed a discretization scheme for the two-stage stochastic linear complementarity problem (SLCP) with continuous random variables. Moreover, they studied the distributionally robust counterpart of the two-stage SLCP when the ambiguity set is constructed with the first order moment information. More recently, Chen, Shapiro and Sun [14] generalized the two-stage SVI problem to the two-stage stochastic generalized equation. They studied the convergence of its SAA without the relatively complete recourse assumption. As a special case, they also considered a mixed two-stage stochastic nonlinear variational inequality problem and examined the uniqueness of its solution and the exponential convergence of its discrete approximation. Considering the risk-averse players, Pang, Sen and Shanbhag made a comprehensive discussion about the two-stage non-cooperative multi-player game under uncertainty in [5].

From the perspective of the numerical calculation, the equilibrium problems are usually rewritten as a minimization optimization problem, mostly nonsmooth and nonconvex. For this class of problems, the smoothing techniques (see [20]) usually are employed so that differentiable methods, e.g., Newton's method, become applicable in solving the smoothing problem, see for instance [7, 21, 22, 23]. Another different avenue is adopted by stochastic approximation (SA) schemes which were first introduced by Robbins and Monro in [24]. There has been a surge of interest in the solution of SVI problems via SA schemes. Amongst the earliest work was conducted by Jiang and Xu [25], who considered the stochastic approach for SVI problems with strongly monotone and Lipschitz continuous assumptions. They proved that the sequence of solution iterates converge to the unique solution in the sense of almost sure. As an extension of this study, motivated by Tikhonov regularization scheme, a regularized SA method was developed for solving SVI problems with a merely monotone but Lipschitz continuous mapping in [26]. Further, in [4], the authors focused on developing asymptotic statements for the SVI problem, where the map is not necessarily Lipschitz continuous, through the SA method. For numerical implementation of the multistage SVI, Rockafellar and Sun extended the well-known PHM for multistage stochastic programming problems to multistage SVI problems in [13]. As for the monotone two-stage SVI or SLCP, PHM has been employed to give numerical results, see [11, 14].

1.3 Summary of contributions of the thesis

The main contributions of this thesis are summarized as follows:

- Firstly, we investigate different sufficient conditions for the existence of solutions to a class of two-stage SVI problem, and stability assertions between it and its perturbed problems. Under the assumption that the solution to the second stage problem is unique, we carry out quantitative stability analysis of this class problems with respect to suitable probability metrics. Moreover, we further consider the discrete approximation scheme, and derive both the convergence and exponential rate of convergence of the optimal solution set of the discrete approximation problem to that of the original problem. Finally, to confirm these theoretical results as well as their applications, we consider a multi-player noncooperative two-stage stochastic game problem and present numerical results by using PHM.

- Since the first work makes an assumption that the second stage problem has a unique solution, we further weaken this assumption. To this end, a two-stage stochastic Nash equilibrium problem is proposed to model the production and supply competition of a homogenous product under uncertainty in an oligopolistic market. The model is recast as two-stage SVI problems whose solutions characterize a Nash-Cournot equilibrium. Considering the possible multiple solutions of the second stage problem, which may be cost expensive or time consuming by numerical results, a regularized sample average approximation method is proposed to solve the two-stage SVI problem. The second stage problem of the regularized problem and its SAA problem have the unique solution. To approximate the original two-stage SVI problem, we establish convergence properties under mild assumptions. Finally, we use some numerical results to test its effectiveness.

1.4 Organization of the thesis

The thesis lays out as follows.

- Chapter 2 gives the models and prerequisite knowledge of this thesis. Specifically, two models are presented: One is a class of two-stage SVI problems; the other one is a convex two-stage non-cooperative multi-player game problem but we can reformulate it as a two-stage SVI problem.

- Chapter 3 considers a class of two-stage stochastic linear variational inequality problems whose first stage problems are stochastic linear box-constrained variational inequality problems and second stage problems are stochastic linear complementary problems having a unique solution. We first present the existence of solutions and quantitative stability results. Next, we consider the discrete approximation and conduct convergence analyses. In numerical part, we study a multi-player noncooperative two-stage stochastic game problem and its numerical tractability by PHM.

- Chapter 4 concentrates on a convex two-stage non-cooperative multi-player game under uncertainty, which can be formulated as a two-stage SVI problem. First of all, the two-stage stochastic Cournot-Nash equilibrium problem is developed and recast into a two-stage SVI problem. Then, a regularized method is proposed and structural results of two-stage regularized SVI problem are presented. The convergence assertions of our regularized SAA problem to the original SVI problem are shown in the sequel. We finally conduct some numerical results based on randomly generated data, which verify our convergence analysis well.

- Chapter 5 concludes the whole thesis and gives some possible future work.

Chapter 2

Models and preliminaries

In this chapter, we give models of this thesis and some preliminaries which are useful in the further discussion.

2.1 A class of two-stage stochastic linear variational inequality problems

We consider a class of two-stage stochastic linear variational inequality problems in the following form [11, 14, 27]:

$$\begin{cases} 0 \in Ax + \mathbb{E}_P[B(\xi)y(\xi)] + q_1 + \mathcal{N}_{[l,u]}(x), \\ 0 \leq y(\xi) \perp M(\xi)y(\xi) + N(\xi)x + q_2(\xi) \geq 0, \text{ for } P\text{-a.e. } \xi \in \Xi, \end{cases} \tag{2.1}$$

where $l, u \in \mathbb{R}^n$ and $l < u$ in the sense of componentwise; $\xi : \Omega \to \Xi \subseteq \mathbb{R}^s$ with the probability space being $(\Omega, \mathcal{F}, \mathbb{P})$, $P \in \mathcal{P}(\Xi)$ and $\mathcal{P}(\Xi)$ denotes the set of probability distributions on the support set Ξ, $A \in \mathbb{R}^{n\times n}, q_1 \in \mathbb{R}^n$; $B(\cdot) : \mathbb{R}^s \to \mathbb{R}^{n\times m}$, $M(\cdot) : \mathbb{R}^s \to \mathbb{R}^{m\times m}$, $N(\cdot) : \mathbb{R}^s \to \mathbb{R}^{m\times n}$ and $q_2(\cdot) : \mathbb{R}^s \to \mathbb{R}^m$ are all matrix-valued or vector-valued mappings. The mathematical expectation $\mathbb{E}_P$ is taken in componentwise with respect to (w.r.t.) the corresponding probability distribution $P := \mathbb{P} \circ \xi^{-1}$. Problem (2.1) aims to find a pair $(x, y(\cdot)) \in [l, u] \times \mathcal{Y}$ satisfying (2.1), where $\mathcal{Y}$ is the collection of measurable functions from Ξ to $\mathbb{R}^m$ such that the expectation in the first stage problem of model (2.1) is well-defined. $\mathcal{N}_{[l,u]}(x)$ denotes the normal cone to the box

$[l, u]$ at x. We say that problem (2.1) satisfies the relatively complete recourse if the second stage problem of (2.1) has a solution $y^*(x, \xi)$ for any $x \in [l, u]$ and almost every (a.e.) $\xi \in \Xi$.

When it comes to variational inequality problems, nonlinear complementarity functions or residual functions have been widely employed. There is rich literature about this issue, see for example [8, 9, 27] and the references therein. Of particular interest in this thesis, we consider the residual function of the box-constrained variational inequality problem. Usually, for the box-constrained variational inequality problem VI$([l, u], g)$, where $l, u \in \mathbb{R}^n$ with $l < u$ and $g : \mathbb{R}^n \to \mathbb{R}^n$, the commonly used residual function for VI$([l, u], g)$ is the so-called "mid" function. It is easy to verify that the first stage problem of (2.1) can be equivalently rewritten as

$$x - \mathrm{mid}\{l, u, x - Ax - \mathbb{E}_P[B(\xi)y(\xi)] - q_1\} = 0,$$

where the "mid" function is defined componentwise as follows:

$$\mathrm{mid}\{l_i, u_i, z_i\} = \begin{cases} l_i, & z_i < l_i, \\ z_i, & l_i \le z_i \le u_i, \\ u_i, & z_i > u_i, \end{cases} \quad \text{for } i = 1, \dots, n.$$

Assume that for any pair $(x, \xi) \in X \times \Xi$, the second stage SLCP of problem (2.1) has a unique solution $y^*(x, \xi)$. Then submitting it into the first stage problem, we obtain

$$0 \in Ax + \mathbb{E}_P[B(\xi)y^*(x, \xi)] + q_1 + \mathcal{N}_{[l,u]}(x),$$

where the right-hand side only depends on x. This inspires us to consider a residual function $f_P : \mathbb{R}^n \to \mathbb{R}_+$ as follows:

$$f_P(x) := \|x - \mathrm{mid}\{l, u, x - Ax - \mathbb{E}_P[B(\xi)y^*(x, \xi)] - q_1\}\|^2. \tag{2.2}$$

If there is $x \in \mathbb{R}^n$ such that $f_P(x) = 0$, then x must be a solution to problem (2.1). For the convenience of further discussion in the sequel, we equivalently consider the following box-constrained minimization problem

$$\min_{x\in[l,u]} \quad f_P(x). \tag{2.3}$$

In Chapter 3, we analyze the quantitative stability of problem (2.1) by employing the minimization problem (2.3). It is noteworthy that recasting the stochastic variational inequality problem (2.1) as a stochastic (nonconvex) optimization problem, such as (2.3), provides a vehicle for conducting the analysis. It is not a necessary avenue to compute an approximation solution, see for example [5, 11, 18, 28].

Probability metrics are distance functions on the space of probability measures or probability distributions. In this thesis, we need the so-called pseudo metric between two probability measures/distributions. We call them pseudo metrics because they usually do not satisfy the axioms of distance. In pseudo metrics, there is a large class of probability metrics called ζ-structure metrics.

Definition 2.1 (probability metric with ζ-structure, see [29])**.** *Let $\mathcal{G}$ be a collection of real-valued measurable functions on support set Ξ. Then, for any two probability measures $P, Q \in \mathcal{P}(\Xi)$, we call*

$$\mathbb{D}_{\mathcal{G}}(P,Q) = \sup_{g\in\mathcal{G}} |\mathbb{E}_P[g(\xi)] - \mathbb{E}_Q[g(\xi)]|$$

the ζ-structure probability metric between P and Q induced by $\mathcal{G}$.

$\mathbb{D}_{\mathcal{G}}(P,Q)$ is a pseudo metric because $\mathbb{D}_{\mathcal{G}}(P,Q) = 0$ does not imply $P = Q$ unless $\mathcal{G}$ is rich enough. Obviously, we have the symmetry and triangle inequality for $\mathbb{D}_{\mathcal{G}}$. It is known from Definition 2.1 that different ζ-structure metrics can be derived through choosing different $\mathcal{G}$s. For example, if we take

$$\mathcal{G}_{TV} := \{g : \Xi \to \mathbb{R} : g \text{ is measurable and } \sup_{\xi\in\Xi} |g(\xi)| \le 1\},$$

the resulting ζ-structure metric

$$\mathbb{D}_{TV}(P,Q) := \sup_{g \in \mathcal{G}_{TV}} |\mathbb{E}_P[g(\xi)] - \mathbb{E}_Q[g(\xi)]|$$

is called the total variation metric. If

$$\mathcal{G}_{FM_p} := \Big\{ g : \Xi \to \mathbb{R} : |g(\xi_1) - g(\xi_2)| \leq \max\left\{1, \|\xi_1\|, \|\xi_2\|\right\}^{p-1} \|\xi_1 - \xi_2\| \Big\},$$

the corresponding ζ-structure metric

$$\zeta_p(P,Q) := \sup_{g \in \mathcal{G}_{FM_p}} |\mathbb{E}_P[g(\xi)] - \mathbb{E}_Q[g(\xi)]|$$

is called the pth order Fortet-Mourier metric, which is often used in the stability analysis of stochastic programs. Usually, which probability metric to select depends on the properties of the stochastic optimization problem. For example, to employ the total variation metric, some boundedness properties of the objective function are needed. This can be easily observed from its definition. The Fortet-Mourier metric requires some locally Lipschitz continuity conditions for the objective function, which is widely used in the quantitative stability analysis of two-stage stochastic linear programming problems. One can refer to [30] and [31] and references therein for more details. Here we employ these two ζ-structure metrics due to the boundedness and locally Lipschitz continuity of the corresponding objective functions. As for their equivalence, weak convergence and discrete approximations of above pseudo metrics, the reader is referred to [30].

In what follows, we give some useful properties about the solution to the second stage SLCP problem. Recall that: A matrix $M \in \mathbb{R}^{m\times m}$ is a $\mathbb{P}$-matrix if all principal minors of M are positive.

Proposition 2.1 ([23]). *Let $M(\xi)$ be a $\mathbb{P}$-matrix for every $\xi \in \Xi$. The following assertions hold for problem* (2.1).

(i) For any given $x \in [l, u]$ and $\xi \in \Xi$, the second stage problem of (2.1) has a unique solution $y^*(x, \xi)$, which can be implicitly written as

$$y^*(x, \xi) = -W(x, \xi)(N(\xi)x + q_2(\xi)),$$

where $W(x, \xi) := [I - D(x, \xi)(I - M(\xi))]^{-1}D(x, \xi)$ and $D(x, \xi)$ is the m-dimensional diagonal matrix defined by

$$D_{jj}(x, \xi) = \begin{cases} 1, & \text{if } (M(\xi)y^*(x, \xi) + N(\xi)x + q_2(\xi))_j \leq y_j^*(x, \xi), \\ 0, & \text{otherwise} \end{cases}$$

for $j = 1, \cdots, m$;

(ii) $y^*(\cdot, \xi)$ is Lipschitz continuous, i.e.,

$$\|y^*(x_1, \xi) - y^*(x_2, \xi)\| \leq \max_{J \in \mathcal{J}} \left\| M_{J \times J}^{-1}(\xi) \right\| \|N(\xi)\| \|x_1 - x_2\|,$$

where $M_{J \times J}(\xi)$ is the sub-matrix of $M(\xi)$, whose entries are indexed by $J \times J$, and $\mathcal{J}$ denotes the power set of $\{1, 2, \cdots, m\}$.

For further discussion, we need the following assumption (see, for example, [14, 11]).

Assumption 2.1. *Let $M(\xi)$ be a $\mathbb{P}$-matrix for every $\xi \in \Xi$. Moreover, there exists a continuous function $\kappa_M : \Xi \to \mathbb{R}_{++}$, such that*

$$\max_{J \in \mathcal{J}} \left\| M_{J \times J}^{-1}(\xi) \right\| \leq \frac{1}{\kappa_M(\xi)}$$

for any $\xi \in \Xi$.

A sufficient condition for Assumption 2.1 is $y^T M(\xi) y \geq \kappa_M(\xi) \|y\|^2$ for any $y \in \mathbb{R}^m$ and $\xi \in \Xi$, from which we can deduce from [11, Lemma 2.1] that $M(\xi)$ is a $\mathbb{P}$-matrix and in addition $\left\| M_{J \times J}^{-1}(\xi) \right\| \leq \frac{1}{\kappa_M(\xi)}$ for any $J \in \mathcal{J}$. A stronger assumption is adopted in [11, Assumption 2.1] (see Assumption 3.1 below).

2.2 Two-stage stochastic Cournot-Nash game

In this section, we consider a two-stage stochastic Cournot-Nash (CN) J-agent(or player) game problem, which extends the classical deterministic CN equilibrium problem in [32]. All stochastic models involve inherently "ordered" components over decision horizons based on the available information at corresponding stage. In particular, the decisions in a strategy may respond to the information that is available only at the present stage. Here, we consider a two-stage case. A strategy pair of agent $j \in \mathcal{J} := \{1, ..., J\}$ is denoted as

$$(x_j \in \mathbb{R},\ y_j : \Xi \to \mathbb{R}), \tag{2.4}$$

where x_j is a first stage decision vector and $y_j \in \mathcal{Y}$ represents a second stage response function with $\mathcal{Y}$ being the space of measurable functions defined on Ξ. Let $\mathfrak{L}_n$ be the Lebesgue space of $\mathbb{R}^n$-valued functions with $\mathfrak{L}_n^\infty$ denoteing the class of measurable essentially bounded functions. Following a similar treatment as that in [10], we further require the second stage response function of random variable to be essentially bounded, i.e., $y_j \in \mathfrak{L}_1^\infty$. Collectively, the vector of strategy pairs of all agents can be written as

$$(x \in \mathbb{R}^J,\ y : \Xi \to \mathbb{R}^J). \tag{2.5}$$

A strategy pair $\left(x_j^*, y_j^*\right) \in \mathbb{R} \times \mathfrak{L}_1^\infty$ is said to be an equilibrium of our stochastic model if it solves the following problem for all agents $j \in \mathcal{J}$.

$$\begin{array}{llr} \max\limits_{(x_j, y_j(\cdot))} & W_j^1(x_j, x_{-j}^*) + \mathbb{E}\big[W_j^2\big(\xi, y_j(\xi), y_{-j}^*(\xi)\big)\big], & \text{(objective function)} \\ \text{s.t.} & x_j \in X_j, & \text{(first stage constraints)} \\ & y_j(\xi) \in_{a.s.} Y_j, \quad g_j\big(\xi, x_j, y_j(\xi)\big) \leq_{a.s.} 0 & \text{(second stage constraints)} \end{array} \tag{2.6}$$

where

$$\begin{aligned} x^*_{-j} &= (x^*_1, \dots, x^*_{j-1}, x^*_{j+1}, \dots, x^*_J), \\ y^*_{-j}(\xi) &= (y^*_1(\xi), \dots, y^*_{j-1}(\xi), y^*_{j+1}(\xi), \dots, y^*_J(\xi)), \end{aligned}$$

and $y_j(\xi)$ denotes the value of response $y_j(\cdot)$ to realization ξ[1], with

- $W^1_j : \mathbb{R} \times \mathbb{R}^{J-1} \to \mathbb{R}$ being a first stage wealth function of agent j, which is concave and continuously differentiable w.r.t. x_j;
- $W^2_j : \Xi \times \mathbb{R} \times \mathbb{R}^{J-1} \to \mathbb{R}$ being a second stage wealth function of agent j, which is concave, well-defined and finite;
- X_j, Y_j being nonempty, closed and convex subsets of $\mathbb{R}$ and the second stage constraints holding almost surely ($a.s.$);
- $g_j : \Xi \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ being a continuously differentiable function w.r.t. $\big(x_j, y_j(\xi)\big)$ for almost every (a.e.) $\xi \in \Xi$.

In this subsection, the model (2.6) is formulated under the assumption that the uncertainty can be described by a random variable ξ with known distribution. From the perspective of the entire system, the market "chooses" $x \in \mathbb{R}^J$ before a realization $\xi \in \Xi$ is revealed and later "selects" $y(\xi) \in \mathbb{R}^J$ with known realization. Or equivalently, x is a here and now solution and y is a wait and see solution.

The application of commodity production and supply in an oligopolistic market serves as a motivation as well as the practical application problem of interest. Presented as a stochastic game, the strategy of each agent in supply-side of the market can be described as the solution of a stochastic optimization problem (2.6). The decision process follows that agent $j \in \mathcal{J}$ decides an optimal production quantity x_j of the commodity at the production stage. At the second stage, each agent decides a

[1] For ease the exposition, the same notation ξ is used for both a random variable and its realization in $\mathbb{R}^s$ without causing confusion on the context.

supply quantity $y_j(\xi)$ after ξ is observed, and a total quantity $T(y(\xi)) := \sum_{j=1}^{J} y_j(\xi)$ is supplied to the market. Our focus on oligopolistic markets requires that the price is dominantly affected by the total supplied quantity in the market $T(y(\xi))$. Therefore, all the trading occurs at the price $p : \Xi \times \mathbb{R} \to \mathbb{R}_+$, determined by a stochastic inverse demand curve $p(\xi, T(y(\xi))$. In practice, production and supply quantities are subject to physical restrictions, for example, capability of production plant, logistic restriction, etc., i.e., $x_j \in X_j$ and $y_j(\xi) \in_{a.s.} Y_j$. More specifically, we have non-negative requirements for both production and supply, $X_j = \mathbb{R}_+$ and $Y_j = \mathbb{R}_+$. The relations between stage-wise decision variables x_j and $y_j(\xi)$ are captured by constraints $g_j\big(\xi, x_j, y_j(\xi)\big) \leq_{a.s.} 0$ in (2.6). Essentially, every agent needs to formulate and solve a two-stage stochastic programming problem with recourse in the sense of achieving equilibria of a J-agents non-cooperative game of the market. We further require that agent j's supply to the market cannot exceed his/her production quantity, i.e., $y_j(\xi) - x_j \leq_{a.s.} 0$. This can be interpreted as the fact that agents may have no stock to start with, or they need to preserve certain reserved quantities prior to each decision process.

The problem can then be viewed from a slight different perspective than that of problem (2.6). As seen from the first stage, agent $j \in \mathcal{J}$ wants to find a production quantity $x_j \geq 0$ to

$$\max_{x_j \geq 0} \; W_j^1(x_j, x_{-j}^*) + \mathbb{E}\big[\Phi_j(x_j, x_{-j}^*, \xi)\big], \tag{2.7}$$

where,

$$\Phi_j(x_j, x_{-j}^*, \xi) = \sup_{y_j(\xi) \geq 0} \{W_j^2(\xi, y_j(\xi), y_{-j}^*(\xi)) \mid x_j \geq y_j(\xi), \text{ for a.e. } \xi \in \Xi\}. \tag{2.8}$$

Objective function in (2.7) is regarded as the expected profit of agent j, and problems (2.7)-(2.8) are termed intrinsic first stage problem following that of a related treatment in [33]. In particular, the analysis of intrinsic first stage problem and the stochastic programming problem with recourse in convex case were carried out in a series of studies by Rockafellar and Wets [34, 33, 35, 36] and more recently in [6]. The key feature of intrinsic first stage problem as well as formulation (2.6) is the requirement on precise orders of decision execution, commonly known as the constraints of nonanticipativity. In problems (2.7) and (2.8), the second stage decisions are explicitly determined after the first stage decision, provided for each x_j the second stage problem is well-defined [37]. However, the study of optimality condition of (2.7)-(2.8), in the case of a general probability space $(\Xi, \mathcal{F}, P)^2$, is very complicated since one needs to characterize the order of the decision process explicitly. For ease of analysis, we assume that there exists a multiplier $\lambda_j \in \mathfrak{L}_1^1$ corresponds to second stage constraint and study the saddle-point condition of the Lagrangian formulation of problem (2.6). It is worth mentioning that the Karush-Kuhn-Tucker (KKT) condition of problem (2.6) (see [34]) introduces a second stage multiplier $\tilde{\lambda}_j \in (\mathfrak{L}_1^\infty)^*$ for every $j \in \mathcal{J}$ which incorporates the two-stage decision making process. This can be seen from the fact that any element of the dual space $(\mathfrak{L}_1^\infty)^*$ can be decomposed into a component of $\mathfrak{L}_1^1$ and a "singular" component, corresponding to the multiplier of nonanticipativity. The saddle-point condition is shown to be sufficient and "almost" necessary for optimality of problem (2.6), and we refer the interested readers to [6, 33, 34, 35, 36] for more details.

The Lagrangian formulation of problem (2.6) associated with agent j is of the following form:

$$L_j(x_j, x^*_{-j}, y_j, y^*_{-j}, \lambda_j) = L^1_j(x_j, x^*_{-j}) + \mathbb{E}\big[L^2_j\big(\xi, x_j, y_j(\xi), y^*_{-j}(\xi), \lambda_j(\xi)\big)\big],$$

[2] In cases of finitely supported distribution, the equivalence between intrinsic first stage problem and the original recourse problem can be established, and the optimality condition of the recourse problem can be applied, see for example [6].

where

$$L_j^1(x_j, x_{-j}^*) = W_j^1(x_j, x_{-j}^*),$$

$$L_j^2\big(\xi, x_j, y_j(\xi), y_{-j}^*(\xi), \lambda_j(\xi)\big) = W_j^2(\xi, y_j(\xi), y_{-j}^*(\xi)) + \lambda_j(\xi)(x_j - y_j(\xi)).$$

The constraints $y_j(\xi) \leq_{a.s.} x_j$ can be interpreted as the situation under which the profit maximizing supply $y_j^*(\xi)$ of agent j is not necessarily equal to the total production quantity x_j. This feature of our model differs from conventional requirement on production-clearing condition, i.e., all the produced goods are expected to supply to the market.

In order to make further progress in characterizing the CN equilibrium, we need to specify the structures of our wealth functions, W_j^1 and W_j^2, suitable for our application. We assume that the production cost for jth agent is quadratic, i.e., for each $j \in \mathcal{J}$ the cost of producing x_j amount of production is

$$\frac{1}{2}c_j x_j^2 + a_j x_j,$$

for some $c_j > 0, a_j > 0$. In the second stage, the cost function of the supply or second stage is assumed to be linear and of stochastic nature, i.e., for each $j \in \mathcal{J}$ the cost of supplying $y_j(\xi)$ amount of commodity is $h_j(\xi)y_j(\xi)$ for a.e. $\xi \in \Xi$. Concretely, $h_j(\xi)$ can be regarded as the unit transport cost for agent j. We adopt a classic stochastic inverse demand curve, see for example [38], that takes the expression $p\big(\xi, T(y(\xi))\big) = p_0(\xi) - \gamma(\xi)T(y(\xi))$ for the spot price. In practice, the stochastic benchmark price excluding the effect of supply to the market $p_0 : \Xi \mapsto \mathbb{R}_+$ can be estimated via statistical approaches based on real data. The supply discount $\gamma : \Xi \to \mathbb{R}_+$ acts as a market mechanism to adjust and reflect uncertainty in quantity in the market. In order to respect the market mechanism of supply-demand relation, we make the following assumption throughout our study.

Assumption 2.2. *There exists a $\gamma_0 > 0$ such that $\gamma(\xi) \geq \gamma_0$ for a.e. $\xi \in \Xi$.*

Thus, agent j's stage-wise wealth functions are,

$$W_j^1(x_j, x_{-j}^*) = -\frac{1}{2}c_j x_j^2 - a_j x_j,$$

and

$$W_j^2(\xi, y_j(\xi), y_{-j}(\xi)) = \big(p_j(\xi) - \gamma(\xi)T(y(\xi))\big)y_j(\xi),$$

where the short-handed notation of the risk-adjusted spot price of agent j's is denoted by $p_j(\xi) := p_0(\xi) - h_j(\xi)$.

We are now ready to consider the specific stochastic programming problem for every agent $j \in \mathcal{J}$:

$$\begin{aligned} \max_{x_j} \quad & \mathbb{E}[\Phi_j(\xi, x)] - \frac{1}{2}c_j x_j^2 - a_j x_j \\ \text{s.t.} \quad & 0 \leq x_j, \end{aligned} \tag{2.9}$$

where

$$\begin{aligned} \Phi_j(\xi, x) = \quad & \max_{y_j(\xi)} \quad \Big(p_j(\xi) - \gamma(\xi)\Big(\sum_{i \neq j}^{J} y_i^*(\xi) + y_j(\xi)\Big)\Big)y_j(\xi) \\ & \text{s.t.} \quad 0 \leq y_j(\xi) \leq x_j, \text{ for a.e. } \xi \in \Xi. \end{aligned} \tag{2.10}$$

Note that the requirements in problem (2.10) hold almost surely in accordance with the a.s. constraints of the second stage in problem (2.6). However, (2.9)-(2.10) is not easy to solve, especially in a stochastic Nash equilibrium problem with $J \geq 2$, see [14]. The complication arises since the jth agent's problem contains that of the other agents' strategy, not yet known at the decision horizon. A commonly used method is to recast problem (2.9)-(2.10) of each agent as a stochastic equilibrium problem. Then, obtaining an equilibrium of the convex J-player game (2.9)-(2.10) is equivalent to finding its solution for all agents. Stochastic equilibrium has been shown to be an effective method to study and to solve two-stage multi-player stochastic game problems, see for instance [11, 5, 39, 6, 13]. We study the saddle-point condition of the problem (2.9)-(2.10), rewritten in the form of problem (2.6). More specifically, for all $j \in \mathcal{J}$, there exists $\bar{\lambda}_j(\xi) \in \mathfrak{L}_1^1$ with $\bar{\lambda}(\xi) \geq_{a.s.} 0$ so that a strategy $\big(\bar{x}_j, \bar{y}_j\big) \in \mathbb{R}_+ \times \mathfrak{L}_+^\infty$ solves the following system.

$$
\begin{aligned}
&- c_j\bar{x}_j - a_j + \mathbb{E}\big[\bar{\lambda}_j(\xi)\big] \in \mathcal{N}_{[0,\infty)}(\bar{x}_j), && \\
&p_j(\xi) - \gamma(\xi)\sum_{i\neq j}^{J} \bar{y}_i(\xi) - 2\gamma(\xi)\bar{y}_j(\xi) - \bar{\lambda}_j(\xi) \in_{a.s.} \mathcal{N}_{[0,\infty)}(\bar{y}_j(\xi)), && \text{(stationality)} \\
&\bar{x}_j \geq 0, \bar{y}_j(\xi) \geq_{a.s.} 0, \bar{x}_j - \bar{y}_j(\xi) \geq_{a.s.} 0, && \text{(feasibility)} \\
&\bar{\lambda}_j(\xi) \geq_{a.s.} 0, && \text{(dual feasibilty)} \\
&\bar{\lambda}_j(\xi) \perp_{a.s.} (\bar{x}_j - \bar{y}_j(\xi)). && \text{(complementarity)}
\end{aligned}
$$

In particular, stationarity comes from the first order necessary optimality condition under the assertion $\partial\mathbb{E}\big[\Phi_j(\xi,x)\big] \subseteq \mathbb{E}\big[\partial_x\Phi(\xi,x)\big]$. The assertion is discussed in [40], and the above system can be viewed as a weaker condition for optimality.

Rewritten in a compact form as SVI, the optimal strategy-multiplier pair $(x_j, y_j, \lambda_j) \in \mathbb{R}_+ \times \mathfrak{L}_+^\infty \times \mathfrak{L}_+^1$ must satisfy,

$$
\begin{array}{lllll}
0 \leq & x_j & \perp & c_j x_j + a_j - \mathbb{E}\big[\lambda_j(\xi)\big] & \geq \quad 0, \\
0 \leq_{a.s.} & y_j(\xi) & \perp_{a.s.} & - p_j(\xi) + \gamma(\xi)\sum_{i\neq j}^{J} y_i(\xi) + 2\gamma(\xi)y_j(\xi) + \lambda_j(\xi) & \geq_{a.s.} 0, \\
0 \leq_{a.s.} & \lambda_j(\xi) & \perp_{a.s.} & x_j - y_j(\xi) & \geq_{a.s.} 0.
\end{array}
\tag{2.11}
$$

It follows that since all agents in oligopolistic market act non-cooperatively, we write down the equilibrium interpreted as that of the whole system. More specifically, let $x = (x_1, \ldots, x_J)^T$ be the first stage decision vectors of the system, and for almost every $\xi \in \Xi$, second stage decision vector $y(\xi) = \big(y_1(\xi), \ldots, y_J(\xi)\big)^T$ and the corresponding multiplier vector $\lambda(\xi) = \big(\lambda_1(\xi), \ldots, \lambda_J(\xi)\big)^T$ are denoted respectively. Similarly, parameter vectors can be written collectively as $a = (a_1, \ldots, a_J)^T$,

$p(\xi) = (p_1(\xi), \ldots, p_J(\xi))^T$. Then, we can treat the SVI for all agents as a two-stage stochastic complementarity problem:

$$\begin{aligned}
&0 \leq x && \perp\; Cx - \mathbb{E}[\lambda(\xi)] + a && \geq 0, \\
&0 \leq \begin{pmatrix} y(\xi) \\ \lambda(\xi) \end{pmatrix} && \perp \begin{pmatrix} \Pi(\xi) & I \\ -I & 0 \end{pmatrix} \begin{pmatrix} y(\xi) \\ \lambda(\xi) \end{pmatrix} + \begin{pmatrix} -p(\xi) \\ x \end{pmatrix} && \geq 0, \quad \text{for a.e. } \xi \in \Xi,
\end{aligned} \tag{2.12}$$

where

$$C = \mathrm{diag}(c_1, c_2, ..., c_J), \quad \Pi(\xi) = \gamma(\xi)(ee^T + I).$$

It follows that for the whole system, a J-tuple of strategies

$$\left(x^*, y^*, \lambda^*\right) = \left((x_1^*, y_1^*, \lambda_1^*), \ldots, (x_J^*, y_J^*, \lambda_J^*)\right) \in \mathbb{R}^J \times \mathfrak{L}_J^\infty \times \mathfrak{L}_J^1$$

is called a solution of the two-stage stochastic complementarity problem (2.12).

Chapter 3

Quantitative analysis for problem (2.1)

3.1 Quantitative stability

Stability analysis of stochastic optimization problems is important for not only theoretical study but also numerical approximation. When we handle a stochastic optimization problem numerically, usually the first step is the discrete or empirical approximation to the included high dimensional integrals. Then some critical questions arise: what is the quantitative relationship between the original continuous problem and its discrete approximation? Do the optimal value and/or optimal solution set of the approximation problem converge to those of the original problem? All these questions can be answered through stability analysis. In view of this, we carry out the quantitative stability analysis of problem (2.1) in this section.

3.1.1 Existence of solutions

The existence statement for stochastic variational inequality problems is a relatively sparse subarea. There are some results, see for example [41, 42, 43, 44, 45]. Specially, Ravat and Shanbhag considered in [42] the stochastic Nash game where the expectation of each player's cost function is minimized. Conditions to admit an equilibrium for both smooth and nonsmooth (but continuous) objective functions were investigated. More recently, the same authors discussed in [43] some verifiable sufficiency conditions for the existence of solutions to stochastic (quasi-)variational inequality problems which extended the results in [42] from single-valued stochastic variational inequality problems to multi-valued stochastic quasi-variational inequality problems.

The existing works mainly concentrate on the deterministic case or the single stage case. Here, we adopt these pioneering works or concepts to give some assertions about the existence of solutions to the two-stage stochastic ones. In the two-stage case, Chen, Sun and Xu employed the strong monotonicity in terms of a redefined inner product on the product space of the first stage and second stage variables in [11], under which the existence and uniqueness assertion of solutions to the two-stage stochastic linear complementarity problem were derived. Under Assumption 2.1, we know that there always exists a unique solution $y^*(x,\xi)$ to the second stage SLCP problem for any given pair $(x,\xi) \in [l,u] \times \Xi$. Namely, problem (2.1) satisfies the relatively complete recourse condition. However, this does not necessarily ensure the existence of a solution to problem (2.1). Therefore, in the sequel, we will introduce several conditions such that problem (2.1) has at least one solution under probability distribution P, and so does its perturbed problem under Q, i.e.,

$$\begin{cases} 0 \in Ax + \mathbb{E}_Q[B(\xi)y(\xi)] + q_1 + \mathcal{N}_{[l,u]}(x), \\ 0 \leq y(\xi) \perp M(\xi)y(\xi) + N(\xi)x + q_2(\xi) \geq 0, \text{ for } Q\text{−}a.e.\ \xi \in \Xi. \end{cases} \tag{3.1}$$

To introduce the first sufficient condition, we make the following assumption which was first used in [11] to study the two-stage SLCP.

Assumption 3.1. *There exists a continuous function $\kappa(\cdot) : \Xi \to \mathbb{R}_{++}$, such that*

$$(x^T, y^T) \begin{pmatrix} A & B(\xi) \\ N(\xi) & M(\xi) \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \geq \kappa(\xi)(\|x\|^2 + \|y\|^2) \tag{3.2}$$

P-a.e. $\xi \in \Xi$, for any $x \in \mathbb{R}^n$ and $y \in \mathbb{R}^m$, where $\mathbb{E}_P[\kappa(\xi)] < +\infty$.

It is easy to see that Assumption 3.1 implies Assumption 2.1 by letting $x = 0$. Then, under Assumption 3.1, Chen, Sun and Xu [11] gave the following conclusion.

Proposition 3.1. *Suppose that Assumption 3.1 holds. Then problem* (2.1) *has a unique solution.*

Assumption 3.1 is sufficient for problem (2.1) to have a unique solution. In this thesis, we give a weaker condition for the existence of solutions to problem (2.1) without uniqueness. For this purpose, we introduce the following notations and the concept of pseudomonotonicity.

Define the mapping $\Phi_P : \mathbb{R}^n \to \mathbb{R}^n$ as

$$\Phi_P(x) = Ax + \mathbb{E}_P[B(\xi)y^*(x, \xi)] + q_1.$$

Recall that Φ_P is pseudomonotone [1, Definition 2.3.1] if

$$\langle x_1 - x_2, \Phi_P(x_2)\rangle \geq 0 \Rightarrow \langle x_1 - x_2, \Phi_P(x_1)\rangle \geq 0.$$

Immediately, based on [1], we have the following proposition.

Proposition 3.2. *Suppose that Assumption 2.1 holds and the following integral*

$$\int_\Xi \|B(\xi)\| \max_{J\in\mathcal{J}} \left\|M^{-1}_{J\times J}(\xi)\right\| \|N(\xi)\| \, P(d\xi)$$

is finite. Then the optimal solution set of problem (2.1) *is nonempty and its projection on the first stage variable is compact. If, in addition,* Φ_P *is pseudomonotone on* $[l, u]$, *this projection is convex too.*

Proof. We first verify that $\Phi_P(x)$ is continuous w.r.t. x. Note that

$$\|\Phi_P(x_1) - \Phi_P(x_2)\| \leq \|A\| \|x_2 - x_1\| + \|\mathbb{E}_P[B(\xi)(y^*(x_2, \xi) - y^*(x_1, \xi))]\| . \tag{3.3}$$

For the second term of the right-hand side of (3.3), we have estimation

$$\begin{aligned}
&\|\mathbb{E}_P[B(\xi)(y^*(x_2, \xi) - y^*(x_1, \xi))]\| \\
&\leq \mathbb{E}_P[\|B(\xi)(y^*(x_2, \xi) - y^*(x_1, \xi))\|] \\
&\leq \mathbb{E}_P[\|B(\xi)\| \max_{J\in\mathcal{J}} \left\|M^{-1}_{J\times J}(\xi)\right\| \|N(\xi)\|] \|x_1 - x_2\| .
\end{aligned}$$

Due to the finiteness of $\mathbb{E}_P[\|B(\xi)\| \max_{J\in\mathcal{J}} \left\|M_{J\times J}^{-1}(\xi)\right\| \|N(\xi)\|]$, we know that Φ_P is Lipschitz continuous, which is obviously continuous. Then, we derive from [1, Proposition 2.2.3] that

$$-\Phi_P(x) \in \mathcal{N}_{[l,u]}(x) \tag{3.4}$$

has a solution. Since Assumption 2.1 hold, there always exists a solution for the second stage problem for any $x \in [l,u]$. To summarize, problem (2.1) has a solution.

Corollary 2.2.5 in [1] tells us that: if $X \subseteq \mathbb{R}^n$ is compact and convex, and $F : X \to \mathbb{R}^n$ is continuous, the solution set of $-F(x) \in \mathcal{N}_X(x)$ is nonempty and compact. If, in addition, F is pseudomonotone, it is known from [1, Theorem 2.3.5] that the solution set is convex.

Due to the boundedness and convexity of interval $[l,u]$, we know from [1, Corollary 2.2.5] that the solution set of (3.4) is nonempty and compact. Moreover, if Φ_P is pseudomonotone, based on [1, Theorem 2.3.5], the solution set of (3.4) is convex. $\square$

Remark 3.1. *We have the following observations about the assumptions in Proposition 3.2.*

(i) In Proposition 3.2, Assumption 2.1 is easy to check by examining $M(\xi)$. The integrability requirement of $\|B(\xi)\| \max_{J\in\mathcal{J}} \left\|M_{J\times J}^{-1}(\xi)\right\| \|N(\xi)\|$ seems to be implicit. Usually, the restrictive integrability conditions on $B(\xi), M(\xi), N(\xi)$ are imposed for easier verifiability, see [42]. For example, if Assumptions 2.1 and 3.2 (in the following) hold, and $\kappa_M(\xi) \geq \kappa$ for some positive constant κ, we have estimation

$$\|B(\xi)\| \max_{J\in\mathcal{J}} \left\|M_{J\times J}^{-1}(\xi)\right\| \|N(\xi)\| \leq \frac{C^2 \max\{1, \|\xi\|\}^{2\alpha}}{\kappa}.$$

Then a sufficient condition for the integrability of the left-hand side is simply $P \in \mathcal{P}_{2\alpha}(\Xi)$, which can be verified easily.

(ii) As for the pseudomonotonicity of Φ_P over $[l,u]$, we can verify the monotonicity of Φ_P instead of the pseudomonotonicity if possible. This might be easier to implement, which is only necessary to examine the monotonicity of $Ax + B(\xi)y^*(x,\xi) + q_1$ for almost everywhere $\xi \in \Xi$. It is known from Proposition 2.1 that, for any $x_1, x_2 \in [l,u]$, we have

$$\begin{aligned}
&\langle x_1 - x_2, Ax_1 + B(\xi)y^*(x_1,\xi) + q_1 - (Ax_2 + B(\xi)y^*(x_2,\xi) + q_1)\rangle \\
&\quad \geq \langle x_1 - x_2, A(x_1 - x_2)\rangle - \max_{J\in\mathcal{J}} \left\|M^{-1}_{J\times J}(\xi)\right\| \|B(\xi)\| \, \|N(\xi)\| \, \|x_1 - x_2\|^2 \\
&\quad \geq \left(\lambda_{\min}(A) - \max_{J\in\mathcal{J}} \left\|M^{-1}_{J\times J}(\xi)\right\| \|B(\xi)\| \, \|N(\xi)\|\right) \|x_1 - x_2\|^2,
\end{aligned}$$

where $\lambda_{\min}(A)$ is the minimal eigenvalue of A. Then a sufficient condition for the monotonicity is that

$$\lambda_{\min}(A) - \max_{J\in\mathcal{J}} \left\|M^{-1}_{J\times J}(\xi)\right\| \|B(\xi)\| \, \|N(\xi)\| \geq 0$$

holds for a.e. $\xi \in \Xi$. This can be further simplified under some specific settings. For example, if we have $y^*(x,\xi) = -M(\xi)^{-1}(N(\xi)x + q_2(\xi)) \geq 0$ for almost everywhere $\xi \in \Xi$ and each $x \in [l,u]$, that is, $A - B(\xi)W(x,\xi)N(\xi) = A - B(\xi)M(\xi)^{-1}N(\xi)$, the monotonicity condition holds when $A - B(\xi)M(\xi)^{-1}N(\xi)$ is positive semidefinite for a.e. $\xi \in \Xi$, which can be easily verified.

In what follows, we consider the existence of solutions to the perturbed problem (3.1) under certain conditions. To ease the statement, we define the multifunction $\Theta_P : [l,u] \rightrightarrows \mathbb{R}^n$ as

$$\Theta_P(x) = Ax + \mathbb{E}_P[B(\xi)y^*(x,\xi)] + q_1 + \mathcal{N}_{[l,u]}(x)$$

and its inverse is

$$\Theta_P^{-1}(y) := \{x \in [l,u] : y \in \Theta_P(x)\}.$$

Proposition 3.3. *Under the same assumption of Proposition 3.2, $z \in \Theta_P(x)$ is solvable for any $z \in \mathbb{R}^n$.*

Proof. Note that, for any $x \in \mathbb{R}^n$, $z \in \Theta_P(x)$ is equivalent to $-\Phi_P(x) + z \in \mathcal{N}_{[l,u]}(x)$. We know from the proof of Proposition 3.2 that $-\Phi_P(x)$ is continuous w.r.t x, so is $-\Phi_P(x) + z$. Then, by the same argument as that in Proposition 3.2, we know that $z \in \Theta_P(x)$ is solvable for any $x \in \mathbb{R}^n$, which completes the proof. $\square$

Before establishing the existence of solutions to the perturbed problem, we need the following boundedness assumption.

Assumption 3.2. *There exist constants $\alpha \geq 0$ and $C > 0$, such that the random parameters included in problem (2.1) can be bounded above as*

$$\|\Lambda(\xi)\| \leq C \max\{1, \|\xi\|\}^{\alpha}, \text{ for a.e. } \xi \in \Xi,$$

where $\Lambda(\xi) = B(\xi), M(\xi), N(\xi)$ or $q_2(\xi)$.

This kind of assumption is commonly adopted in the quantitative analysis of stochastic programming problems. Actually, it is usually assumed that these parametric mappings are affine w.r.t. ξ. This would imply not only that Assumption 3.2 holds with $\alpha = 1$, but also that these mappings are Lipschitz continuous w.r.t. ξ. With Assumption 3.2, we have the following lemma.

Lemma 3.1. *Suppose that Assumptions 2.1 and 3.2 hold, $\kappa_M(\xi) \geq \kappa > 0$ and $P, Q \in \mathcal{P}_{2\alpha+1}(\Xi)$. Then there exists a positive number L such that*

$$\|\mathbb{E}_P[B(\xi)y^*(x,\xi)] - \mathbb{E}_Q[B(\xi)y^*(x,\xi)]\| \leq L\mathbb{D}_{TV}(P,Q)^{\frac{1}{2\alpha+1}}, \tag{3.5}$$

when $\mathbb{D}_{TV}(P,Q) + \zeta_{2\alpha+1}(P,Q) \leq 1$.

Proof. It is known from [23, Theorem 2.1] and Assumption 2.1 that $\|W(x,\xi)\| \leq \max_{J\in\mathcal{I}} \left\|M^{-1}_{J\times J}(\xi)\right\| \leq \frac{1}{\kappa_M(\xi)}$. Under Assumptions 2.1 and 3.2, we have from (i) of

Proposition 2.1 that

$$\|y^*(x,\xi)\| \le \frac{1}{\kappa_M(\xi)}\|N(\xi)x + q_2(\xi)\| \le \frac{(R+1)C}{\kappa}\max\{1,\|\xi\|\}^{\alpha}, \tag{3.6}$$

where $R := \max_{x\in[l,u]}\|x\|$. Thus,

$$\|B(\xi)y^*(x,\xi)\| \le \frac{(R+1)C^2}{\kappa}\max\{1,\|\xi\|\}^{2\alpha}. \tag{3.7}$$

Meanwhile, we have

$$\begin{aligned}\|\mathbb{E}_P[B(\xi)y^*(x,\xi)] - \mathbb{E}_Q[B(\xi)y^*(x,\xi)]\| \le& \int_{\{\xi\in\Xi:\|\xi\|>\Gamma\}}\|B(\xi)y^*(x,\xi)\|(P+Q)(d\xi)\\ &+\left\|\int_{\{\xi\in\Xi:\|\xi\|\le\Gamma\}}B(\xi)y^*(x,\xi)(P-Q)(d\xi)\right\|.\end{aligned}$$

Here, we select $\Gamma \ge 1$. For the second term at the right-hand side, we have

$$\begin{aligned}&\left\|\int_{\{\xi\in\Xi:\|\xi\|\le\Gamma\}}B(\xi)y^*(x,\xi)(P-Q)(d\xi)\right\|\\ &\quad= \frac{(R+1)C^2\Gamma^{2\alpha}}{\kappa}\left\|\int_{\{\xi\in\Xi:\|\xi\|\le\Gamma\}}\frac{B(\xi)y^*(x,\xi)}{(R+1)C^2\Gamma^{2\alpha}/\kappa}(P-Q)(d\xi)\right\|.\end{aligned}$$

It is known from (3.7) that

$$\|B(\xi)y^*(x,\xi)\| \le \frac{(R+1)C^2\Gamma^{2\alpha}}{\kappa}$$

for any ξ with $\|\xi\| \le \Gamma$. This implies

$$\frac{(B(\xi)y^*(x,\xi))_i}{(R+1)C^2\Gamma^{2\alpha}/\kappa} \le 1$$

because $|(B(\xi)y^*(x,\xi))_i| \le \|B(\xi)y^*(x,\xi)\|$, for $i = 1,2,\cdots,n$. Define $g_i(x,\xi)$ by

$$g_i(x,\xi) = \begin{cases}\frac{(B(\xi)y^*(x,\xi))_i}{(R+1)C^2\Gamma^{2\alpha}/\kappa}, & \|\xi\| \le \Gamma;\\ 0, & \text{otherwise.}\end{cases}$$

Obviously, we have $g_i(x,\xi) \in \mathcal{G}_{TV}$ which indicates that

$$\left| \int_{\Xi} g_i(x,\xi)(P-Q)(d\xi) \right| \leq \mathbb{D}_{TV}(P,Q)$$

for $i = 1, 2, \cdots, n$. Denote by $g = (g_1, \cdots, g_n)^T$. Then, by the definition of total variation metric, we have

$$\begin{aligned}
\left\| \int_{\{\xi \in \Xi : \|\xi\| \leq \Gamma\}} \frac{B(\xi) y^*(x,\xi)}{(R+1)C^2\Gamma^{2\alpha}/\kappa}(P-Q)(d\xi) \right\| &= \left\| \int_{\Xi} g(x,\xi)(P-Q)(d\xi) \right\| \\
&= \left(\sum_{i=1}^{n} \left| \int_{\Xi} g_i(x,\xi)(P-Q)(d\xi) \right|^2 \right)^{\frac{1}{2}} \\
&\leq \sqrt{n}\mathbb{D}_{TV}(P,Q).
\end{aligned}$$

Finally, we obtain

$$\left\| \int_{\{\xi \in \Xi : \|\xi\| \leq \Gamma\}} B(\xi) y^*(x,\xi)(P-Q)(d\xi) \right\| \leq \sqrt{n}\frac{(R+1)C^2}{\kappa}\Gamma^{2\alpha}\mathbb{D}_{TV}(P,Q).$$

Note that $\|\xi\|^p / p \in \mathcal{G}_{FM_p}$ for any $p \geq 1$, which means

$$\int_{\Xi} \|\xi\|^p Q(d\xi) - \int_{\Xi} \|\xi\|^p P(d\xi) \leq p\zeta_p(P,Q).$$

Thus,

$$\begin{aligned}
&\int_{\{\xi \in \Xi : \|\xi\| > \Gamma\}} \|B(\xi) y^*(x,\xi)\| (P+Q)(d\xi) \\
&\quad \leq \frac{(R+1)C^2}{\kappa\Gamma} \int_{\{\xi \in \Xi : \|\xi\| > \Gamma\}} \|\xi\|^{2\alpha+1} (P+Q)(d\xi) \\
&\quad \leq \frac{(R+1)C^2}{\kappa\Gamma} \left(2\mathbb{E}_P[\|\xi\|^{2\alpha+1}] + (2\alpha+1)\zeta_{2\alpha+1}(P,Q) \right).
\end{aligned}$$

To summarize the above estimation, we obtain that

$$\begin{aligned}
&\|\mathbb{E}_P[B(\xi)y^*(x,\xi)] - \mathbb{E}_Q[B(\xi)y^*(x,\xi)]\| \leq \\
&\quad \sqrt{n}\frac{(R+1)C^2}{\kappa}\Gamma^{2\alpha}\mathbb{D}_{TV}(P,Q) + \frac{(R+1)C^2}{\kappa\Gamma}(2\mathbb{E}_P[\|\xi\|^{2\alpha+1}] + 2\alpha + 1),
\end{aligned}$$

which comes from the assumption that $\mathbb{D}_{TV}(P,Q)+\zeta_{2\alpha+1}(P,Q)\le 1$. Specially, we define

$$\Gamma=\mathbb{D}_{TV}(P,Q)^{-1/(2\alpha+1)}\ge 1.$$

Finally, we derive that

$$\|\mathbb{E}_P[B(\xi)y^*(x,\xi)]-\mathbb{E}_Q[B(\xi)y^*(x,\xi)]\|\le L\mathbb{D}_{TV}(P,Q)^{\frac{1}{2\alpha+1}},$$

where $L=\left((R+1)C^2(\sqrt{n}+2\mathbb{E}_P[\|\xi\|^{2\alpha+1}]+2\alpha+1)\right)/\kappa$. $\square$

Proposition 3.4. *Suppose that Assumptions 2.1 and 3.2 hold, $\kappa_M(\xi)\ge\kappa>0$ and $P,Q\in\mathcal{P}_{2\alpha}(\Xi)$. Then the perturbed problem* (3.1) *is solvable.*

Proof. We have from Assumptions 2.1 and 3.2, and $\kappa_M(\xi)\ge\kappa>0$ that

$$\begin{aligned}\|B(\xi)\|\max_{J\in\mathcal{J}}\left\|M^{-1}_{J\times J}(\xi)\right\|\|N(\xi)\| &\le \frac{C^2\max\{1,\|\xi\|^{2\alpha}\}}{\kappa_M(\xi)}\\ &\le \frac{C^2\max\{1,\|\xi\|^{2\alpha}\}}{\kappa}.\end{aligned}$$

This and $P\in\mathcal{P}_{2\alpha}(\Xi)$ imply that Proposition 3.3 holds.

Moreover, it is known from (3.7) that

$$\|B(\xi)y^*(x,\xi)\|\le\frac{(R+1)C^2}{\kappa}\max\{1,\|\xi\|\}^{2\alpha}.$$

Since $P,Q\in\mathcal{P}_{2\alpha}(\Xi)$, we obtain that both $\mathbb{E}_P[B(\xi)y^*(x,\xi)]$ and $\mathbb{E}_Q[B(\xi)y^*(x,\xi)]$ are well-defined and have finite value. Let

$$z=\mathbb{E}_P[B(\xi)y^*(x,\xi)]-\mathbb{E}_Q[B(\xi)y^*(x,\xi)]\in\mathbb{R}^n.$$

According to Proposition 3.3,

$$\mathbb{E}_P[B(\xi)y^*(x,\xi)]-\mathbb{E}_Q[B(\xi)y^*(x,\xi)]\in\Theta_P(x)$$

is solvable, that is,

$$0\in Ax+\mathbb{E}_Q[B(\xi)y^*(x,\xi)]+q_1+\mathcal{N}_{[l,u]}(x)$$

or the perturbed problem (3.1) is solvable. $\square$

King and Rockafellar in [41] put forward the concept of subinvertibility to investigate the existence of the solution to perturbed generalized equations, which can be applied to the situation without the differentiability assumption. In the following, we employ the concept of subinvertibility to establish the existence assertion. The subinvertibility of a multifunction is defined on its graph. For more details about the graph of a multifunction, one can refer to [46]. Specifically, we have the following definition of subinvertibility.

Definition 3.1 (subinvertibility, [41])**.** *$\Theta_P(x)$ is said to be subinvertible at $(x^*, 0)$, if $0 \in \Theta_P(x^*)$ and there exist a compact neighborhood U of x^*, a positive scalar ϵ and a nonempty convex-valued multifunction $G : \epsilon\mathbb{B} \to U$, such that the graph of G, denoted by $\mathrm{gph}G$, is closed, $x^* \in G(0)$ and $G(y)$ is contained in $\Theta_P^{-1}(y)$ for each $y \in \epsilon\mathbb{B}$.*

As for more discussion of subinvertibility, one can refer to [41] for details. Then, based on the concept of subinvertibility and [41, Proposition 3.1], we have the following proposition.

Proposition 3.5. *Suppose that all assumptions in Lemma 3.1 hold and $\Theta_P(x)$ is subinvertible at $(x^*, 0)$. Then there exist a compact and convex neighborhood U of x^* and a positive scalar ϵ, such that*

$$0 \in Ax + \mathbb{E}_Q[B(\xi)y^*(x,\xi)] + q_1 + \mathcal{N}_{[l,u]}(x)$$

has at least one solution in U when $\mathbb{D}_{TV}(P,Q) + \zeta_{2\alpha+1}(P,Q) \le \epsilon$.

Proof. According to Lemma 3.1, we have that

$$\|\mathbb{E}_Q[B(\xi)y^*(x,\xi)] - \mathbb{E}_P[B(\xi)y^*(x,\xi)\| \le L\mathbb{D}_{TV}(P,Q)^{\frac{1}{2\alpha+1}}.$$

From [41, Proposition 3.1], we know that there exists an $\epsilon_0 > 0$ satisfying

$$L\mathbb{D}_{TV}(P,Q)^{\frac{1}{2\alpha+1}} \le \epsilon_0,$$

such that the perturbed problem

$$\begin{aligned}
0 \in & Ax + \mathbb{E}_P[B(\xi)y^*(x,\xi)] + q_1 + \mathcal{N}_{[l,u]}(x) \\
& + Ax + \mathbb{E}_Q[B(\xi)y^*(x,\xi)] + q_1 - (Ax + \mathbb{E}_P[B(\xi)y^*(x,\xi)] + q_1) \\
= & Ax + \mathbb{E}_Q[B(\xi)y^*(x,\xi)] + q_1 + \mathcal{N}_{[l,u]}(x)
\end{aligned}$$

has at least one solution in a compact and convex neighborhood U of x^*. Then, letting $\epsilon := \left(\frac{\epsilon_0}{L}\right)^{2\alpha+1}$ completes the proof. $\square$

The subinvertibility of $\Theta_P(x)$ can be verified under some typical cases, see [41]. The following remark tells us that our conditions are not limiting compared with those in [11, 13].

Remark 3.2. *As we mentioned before, in [11], the authors required Assumption 3.1, which is stronger than Assumption 2.1. On the other hand, in [13], the authors directly assumed that problem* (2.1) *is solvable, and the coefficient matrix*

$$\begin{pmatrix} A & B(\xi) \\ N(\xi) & M(\xi) \end{pmatrix}$$

is positive semidefinite for any $\xi \in \Xi$, where their support set Ξ is assumed to be finite. In this case, the positive semidefinite assumption is equivalent to monotonicity. Our conditions are weaker than those in [13]. To clarify this, we consider the following coefficient matrix:

$$\begin{pmatrix} A & 0 \\ N \cdot \xi & M \cdot \xi^2 \end{pmatrix},$$

where $\xi \in \Xi := [\frac{1}{2}, 1]$, $A \in \mathbb{R}^{n\times n}$ is negative definite, $N \in \mathbb{R}^{m\times n}$ and $M \in \mathbb{R}^{m\times m}$ is positive definite. Obviously, due to the negative definiteness of A, this kind of coefficient matrix is not positive semidefinite. When the coefficient matrix takes the above form, the first stage problem is always solvable if $x = -A^{-1}q_1 \in [l,u]$. Moreover, the positive semidefiniteness of $M(\xi)$ ensures that the second stage problem is always solvable. However, this situation still fails to satisfy the requirement in [13].

3.1.2 Quantitative stability

In this subsection, we consider the quantitative stability analysis of problem (2.1). Denote by $S(P)$ and $v(P)$ the optimal solution set and optimal value of problem (2.3). Note the fact that

$$\left| \|a\|^2 - \|b\|^2 \right| = \left| (a-b)^T (a+b) \right| \leq \|a-b\| \left(\|a\| + \|b\| \right)$$

for any $a, b \in \mathbb{R}^n$. By this fact and (2.2), we have the following estimation:

$$\begin{aligned}
&|f_P(x) - f_Q(x)| \\
&\leq \|\mathbb{E}_P[B(\xi)y^*(x,\xi)] - \mathbb{E}_Q[B(\xi)y^*(x,\xi)]\| \\
&\quad \cdot \big(2\|x\| + \|\mathrm{mid}\{l, u, x - (Ax + \mathbb{E}_P[B(\xi)y^*(x,\xi)] + q_1)\}\| \\
&\quad + \|\mathrm{mid}\{l, u, x - (Ax + \mathbb{E}_Q[B(\xi)y^*(x,\xi)] + q_1)\}\| \big).
\end{aligned} \tag{3.8}$$

Firstly, we assume that the support set Ξ is a compact subset in $\mathbb{R}^s$. Then we have

$$\|\mathbb{E}_P[B(\xi)y^*(x,\xi)] - \mathbb{E}_Q[B(\xi)y^*(x,\xi)]\| = \left\| \int_\Xi B(\xi)y^*(x,\xi)(P-Q)(d\xi) \right\|.$$

It is known from (3.7) that

$$|B_i(\xi)y^*(x,\xi)| \leq \|B(\xi)y^*(x,\xi)\| \leq \frac{(R+1)C^2}{\kappa_M(\xi)} \max\{1, \|\xi\|\}^{2\alpha}, \ i = 1, 2, \cdots, n,$$

where $R = \max_{x \in [l,u]}\{1, \|x\|\}$. Moreover, we have

$$0 < \max_{\xi \in \Xi} \left\{ \frac{(R+1)C^2}{\kappa_M(\xi)} \max\{1, \|\xi\|\}^{2\alpha} \right\} < +\infty$$

because of the compactness of Ξ and the positivity and continuity of $\kappa_M(\xi)$. Then, we continue

$$\left| \int_\Xi B_i(\xi)y^*(x,\xi)(P-Q)(d\xi) \right| \leq \max_{\xi \in \Xi} \left\{ \frac{(R+1)C^2}{\kappa_M(\xi)} \max\{1, \|\xi\|\}^{2\alpha} \right\} \mathbb{D}_{TV}(P, Q),$$

for $i = 1, 2, \cdots, n$. Thus we obtain

$$\left\| \int_{\Xi} B(\xi) y^*(x,\xi)(P-Q)(d\xi) \right\| \leq \sqrt{n}(R+1)C^2 \max_{\xi \in \Xi} \left\{ \frac{\max\{1, \|\xi\|\}^{2\alpha}}{\kappa_M(\xi)} \right\} \mathbb{D}_{TV}(P,Q).$$

For the second term of the right-hand side of (3.8), we can bound it by

$$2\left(4R + \|A\| R + (R+1)C^2 \max_{\xi \in \Xi} \left(\frac{\max\{1, \|\xi\|\}^{2\alpha}}{\kappa_M(\xi)} \right) + \|q_1\| \right) := \eta.$$

To sum up, we have the following quantitative estimation.

Lemma 3.2. *Let Assumptions 2.1 and 3.2 hold and Ξ be a compact set. Then there exists a positive constant L_1, such that*

$$\sup_{x \in [l,u]} |f_P(x) - f_Q(x)| \leq L_1 \mathbb{D}_{TV}(P,Q),$$

where $L_1 := \eta\sqrt{n}(R+1)C^2 \max_{\xi \in \Xi} \left(\frac{\{1, \|\xi\|\}^{2\alpha}}{\kappa_M(\xi)} \right)$.

Before establishing the relationship between $S(Q)$ and $S(P)$, we introduce the growth function and its inverse. We call $\psi_P : \mathbb{R}_+ \to \mathbb{R}$ the growth function of problem (2.3) if

$$\psi_P(\tau) := \min\{f_P(x) : d(x, S(P)) \geq \tau, x \in [l,u]\}.$$

It is not difficult to verify from its definition that $\psi_P(\cdot)$ is nondecreasing and lower semicontinuous. Its inverse function is defined by

$$\psi_P^{-1}(t) = \sup\{\tau \in \mathbb{R}_+ : \psi_P(\tau) \leq t\}, \tag{3.9}$$

which, of course, is nondecreasing too. For more information, we refer to [46, Example 7.63] and [30].

As a sequence of Lemma 3.2, we immediately obtain the following quantitative description of optimal solution sets.

Theorem 3.1. *Let Assumptions 2.1 and 3.2 hold and the support set Ξ be compact. Then*

$$S(Q) \subseteq S(P) + \psi_P^{-1}(L_1 \mathbb{D}_{TV}(P,Q))\mathbb{B},$$

where L_1 is defined in Lemma 3.2 and $\mathbb{B}$ is the closed unit ball centered at 0.

Proof. A similar proof can be found in [30, Theorem 9]. To keep the thesis self-contained, we provide a brief proof. If $S(Q) = \emptyset$, the assertion obviously holds. In the following, we assume $S(Q) \neq \emptyset$. For any $\tilde{x} \in S(Q)$, we have $\upsilon(Q) = f_Q(\tilde{x}) = 0$ and $\upsilon(P) = 0$. Then we have

$$\begin{aligned} L_1 \mathbb{D}_{TV}(P,Q) &= L_1 \mathbb{D}_{TV}(P,Q) + f_Q(\tilde{x}) - \upsilon(P) \\ &\geq f_P(\tilde{x}) - f_Q(\tilde{x}) + f_Q(\tilde{x}) - \upsilon(P) \\ &= f_P(\tilde{x}) - \upsilon(P) \\ &\geq \psi_P(d(\tilde{x}, S(P))). \end{aligned}$$

Thus, we have

$$d(\tilde{x}, S(P)) \leq \psi_P^{-1}(L_1 \mathbb{D}_{TV}(P,Q)).$$

Since $\tilde{x} \in S(Q)$ is selected arbitrarily, we have actually shown that

$$S(Q) \subseteq S(P) + \psi_P^{-1}(L_1 \mathbb{D}_{TV}(P,Q))\mathbb{B}.$$

□

In what follows, we derive the corresponding conclusions without compactness of the support set Ξ by utilizing the conclusion in Lemma 3.1.

Theorem 3.2. *Suppose that Assumptions 2.1 and 3.2 hold, $P, Q \in \mathcal{P}_{2\alpha+1}(\Xi)$ and $\kappa_M(\xi) \geq \kappa > 0$. Then there exists a positive constant L_2, such that*

$$\sup_{x \in [l,u]} |f_P(x) - f_Q(x)| \leq L_2 \mathbb{D}_{TV}(P,Q)^{\frac{1}{2\alpha+1}}, \tag{3.10}$$

$$S(Q) \subseteq S(P)) + \psi_P^{-1}(L_2 \mathbb{D}_{TV}(P,Q)^{\frac{1}{2\alpha+1}})\mathbb{B}, \tag{3.11}$$

when $\mathbb{D}_{TV}(P,Q)+\zeta_{2\alpha+1}(P,Q)\le 1$.

Proof. We know from Lemma 3.1 that there exists a positive constant $L>0$, such that

$$\|\mathbb{E}_P[B(\xi)y^*(x,\xi)]-\mathbb{E}_Q[B(\xi)y^*(x,\xi)]\|\le L\mathbb{D}_{TV}(P,Q)^{\frac{1}{2\alpha+1}}, \tag{3.12}$$

when $\mathbb{D}_{TV}(P,Q)+\zeta_{2\alpha+1}(P,Q)\le 1$.

For the second term in the right-hand side of (3.8), we can bound it above by

$$\begin{aligned}
& 8R+2\|A\|R+2\|q_1\|+\frac{(R+1)C^2\big(\mathbb{E}_P[\|\xi\|^{2\alpha}]+\mathbb{E}_Q[\|\xi\|^{2\alpha}]+2\big)}{\kappa} \\
\le\ & 8R+2\|A\|R+2\|q_1\|+\frac{2(R+1)C^2}{\kappa}(\mathbb{E}_P[\|\xi\|^{2\alpha}]+\alpha\zeta_{2\alpha}(P,Q)+1) \\
\le\ & 8R+2\|A\|R+2\|q_1\|+\frac{2(R+1)C^2}{\kappa}(\mathbb{E}_P[\|\xi\|^{2\alpha}]+\alpha+1):=C_1, \qquad (3.13)
\end{aligned}$$

where R is defined as that in Lemma 3.1 and the second inequality comes from (see [30])

$$\left|\mathbb{E}_Q[\|\xi\|^{2\alpha}]-\mathbb{E}_P[\|\xi\|^{2\alpha}]\right|\le 2\alpha\zeta_{2\alpha}(P,Q).$$

Combining (3.12) and (3.13), and letting $L_2=LC_1$, we obtain (3.10). We can derive (3.11) by using a similar proof as that of Theorem 3.1, and thus omit the proof. □

Theorems 3.1 and 3.2 assert that the solution set of the perturbed problem can be somehow bounded by that of the original problem under specific conditions. In order to quantify it, we adopt a general growth function, instead of imposing a specific growth condition, on the objective function of the original problem. Since the general growth function will vanish at 0, see [30] for details, a sufficiently small perturbation will not change the solution set too much. This stability property is important for both theoretical research and practical calculation. Recall that we say the general growth function ψ_P has the kth order growth for some scalar $k\ge 1$ if

$\psi_P(\tau) \geq C\tau^k$ for small $\tau \in \mathbb{R}_+$ and positive constant C. If ψ_P has kth order growth, Theorems 3.1 and 3.2 would establish the Hölder continuity of $S(\cdot)$ at P with rate $1/k$.

3.2 Exponential rate of convergence

In this section, we consider the discrete approximation to problem (2.1). Assume that, according to the probability distribution P, we have independent and identically distributed samples $\xi^1, \xi^2, \cdots, \xi^K$. Then, for each fixed positive integer K, we have the following discrete approximation to problem (2.1) with the sample size K, i.e.,

$$\begin{cases} 0 \in Ax + \frac{1}{K}\sum_{i=1}^{K}(B(\xi^i)y(\xi^i)) + q_1 + \mathcal{N}_{[l,u]}(x), \\ 0 \leq y(\xi^i) \perp M(\xi^i)y(\xi^i) + N(\xi^i)x + q_2(\xi^i) \geq 0, \text{ for } i = 1, 2, \cdots, K. \end{cases} \tag{3.14}$$

In the sequel, we investigate the approximation properties between problems (2.1) and (3.14) as K tends to infinity. To this end, we define the discrete approximation distribution P_K with the sample size K by

$$P_K(\xi) = \frac{1}{K}\sum_{i=1}^{K}\delta_{\xi^i}(\xi), \text{ for } \xi \in \Xi,$$

where $\delta_{\xi^i}(\cdot)$ are indicator functions, that is, $\delta_{\xi^i}(\xi) = 1$ if $\xi = \xi^i$; otherwise $\delta_{\xi^i}(\xi) = 0$ for $i = 1, 2, \cdots, K$. Under Assumption 3.1, we can equivalently rewrite (3.14) as a minimization problem as follows:

$$\min_{x \in [l,u]} f_{P_K}(x), \tag{3.15}$$

where f_{P_K} is defined in (2.2) by substituting P with P_K.

Different from the usual convergence analysis about stochastic variational inequality problems (see for instance [14, 11, 47]) which does not adopt the residual function, we consider the convergence and exponential rate of convergence between problems (2.3) and (3.15).

To investigate the convergence of the optimal solution set of problem (3.15) to that of problem (2.3), we need to consider the convergence between $f_{P_K}(x)$ and $f_P(x)$. For this purpose, we first derive the uniform convergence of term $\|\mathbb{E}_P[B(\xi)y^*(x,\xi)] - \mathbb{E}_{P_K}[B(\xi)y^*(x,\xi)]\|$ on $[l,u]$. Thus, we have the following proposition.

Proposition 3.6. *Suppose that Assumptions 2.1 and 3.2 hold, and $P \in \mathcal{P}(\Xi)$ satisfies*

$$\mathbb{E}_P\left[\frac{\|\xi\|^{2\alpha}}{\kappa_M(\xi)}\right] < +\infty. \tag{3.16}$$

Then

$$\sup_{x\in[l,u]} \|\mathbb{E}_P[B(\xi)y^*(x,\xi)] - \mathbb{E}_{P_K}[B(\xi)y^*(x,\xi)]\| \to 0$$

as $K \to \infty$, with probability one.

Proof. It is easy to see from Proposition 2.1 that $B(\xi)y^*(\cdot,\xi)$ is continuous in $[l,u]$. Moreover, we know from (3.6) and Assumption 3.2 that

$$\begin{aligned}\|B(\xi)y^*(x,\xi)\| &\le \frac{1}{\kappa_M(\xi)}\|N(\xi)x + q_2(\xi)\|\,\|B(\xi)\| \\ &\le \frac{(R+1)C^2}{\kappa_M(\xi)}\max\{1,\|\xi\|\}^{2\alpha} \\ &\le \frac{(R+1)C^2}{\kappa_M(\xi)}\left(1+\|\xi\|^{2\alpha}\right). \end{aligned} \tag{3.17}$$

By (3.16), we have that the right-hand side of (3.17) is integrable under probability distribution P. All these arguments ensure the uniform convergence by [37, Theorem 7.53]. $\square$

Based on Proposition 3.6, we immediately obtain the following corollary.

Corollary 3.1. *Under the same assumptions as Proposition 3.6, we have that*

$$\lim_{K\to\infty} \sup_{x\in[l,u]} \|\mathbb{E}_{P_K}[B(\xi)y^*(x,\xi)]\| \leq \sup_{x\in[l,u]} \|\mathbb{E}_P[B(\xi)y^*(x,\xi)]\| + 1$$

with probability one.

Lemma 3.3. *Let ψ_P^{-1} be defined in* (3.9)*. Then for any $\epsilon > 0$, there exists a sufficiently small scalar $\delta > 0$ such that $\psi_P^{-1}(\delta) \leq \epsilon$, namely, $\psi_P^{-1}(\delta) \to 0$ as $\delta \to 0$.*

Proof. Recall that

$$\psi_P^{-1}(t) = \sup\{\tau \in \mathbb{R}_+ : \psi_P(\tau) \leq t\}.$$

For any $\epsilon > 0$, there exists a sufficiently small $\delta > 0$ with $\delta \leq \psi_P(\epsilon)$, which implies $\epsilon \geq \psi_P^{-1}(\delta)$. □

Corollary 3.2. *Let Assumption 2.1 hold. Then $S(P_K) \neq \emptyset$.*

Proof. Since P_K is the empirical distribution with finite support set $\{\xi^1, \cdots, \xi^K\}$, we have that

$$\begin{aligned}
&\int_\Xi \|B(\xi)\| \max_{J\in\mathcal{J}} \left\|M_{J\times J}^{-1}(\xi)\right\| \|N(\xi)\| \, P_K(d\xi) \\
&= \frac{1}{K}\sum_{i=1}^K \left\|B(\xi^i)\right\| \max_{J\in\mathcal{J}} \left\|M_{J\times J}^{-1}(\xi^i)\right\| \left\|N(\xi^i)\right\| \\
&\leq \frac{1}{K}\sum_{i=1}^K \frac{\|B(\xi^i)\| \, \|N(\xi^i)\|}{\kappa_M(\xi^i)} \\
&< +\infty
\end{aligned}$$

for any positive integer K, where the last inequality comes from Assumption 2.1. Then according to Proposition 3.2 with $Q = P_K$, $S(P_K)$ is nonempty. □

Theorem 3.3. *Under the same assumptions as Proposition 3.6, we have*

$$d(S(P_K), S(P)) \to 0$$

as $K \to \infty$, with probability one.

Proof. We know that

$$d(S(P_K), S(P)) \leq \psi_P^{-1}\left(\sup_{x\in[l,u]} |f_P(x) - f_{P_K}(x)|\right). \tag{3.18}$$

Therefore, to establish the assertion, we only need to prove

$$\sup_{x\in[l,u]} |f_P(x) - f_{P_K}(x)| \to 0$$

with probability one as $K \to \infty$. We have from (3.8) that

$$\sup_{x\in[l,u]} |f_P(x) - f_{P_K}(x)| \leq$$

$$\sup_{x\in[l,u]} \|\mathbb{E}_P[B(\xi)y^*(x,\xi)] - \mathbb{E}_{P_K}[B(\xi)y^*(x,\xi)]\| \cdot \theta(P_K, P),$$

where

$$\begin{aligned}\theta(P_K, P) =& 8R + 2R\|A\| + 2\|q_1\| \\ &+ \sup_{x\in[l,u]} \|\mathbb{E}_{P_K}[B(\xi)y^*(x,\xi)]\| + \sup_{x\in[l,u]} \|\mathbb{E}_P[B(\xi)y^*(x,\xi)]\|.\end{aligned} \tag{3.19}$$

Then, we obtain

$$\lim_{K\to\infty} \sup_{x\in[l,u]} |f_P(x) - f_{P_K}(x)| \leq$$

$$\lim_{K\to\infty} \sup_{x\in[l,u]} \|\mathbb{E}_P[B(\xi)y^*(x,\xi)] - \mathbb{E}_{P_K}[B(\xi)y^*(x,\xi)]\| \cdot \lim_{K\to\infty} \theta(P_K, P).$$

It can be deduced from Proposition 3.6 and Corollary 3.1 that

$$\lim_{K\to\infty} \sup_{x\in[l,u]} \|\mathbb{E}_P[B(\xi)y^*(x,\xi)] - \mathbb{E}_{P_K}[B(\xi)y^*(x,\xi)]\| = 0$$

and

$$\lim_{K\to\infty} \sup_{x\in[l,u]} \|\mathbb{E}_{P_K}[B(\xi)y^*(x,\xi)]\| \leq \sup_{x\in[l,u]} \|\mathbb{E}_P[B(\xi)y^*(x,\xi)]\| + 1$$

with probability one, respectively. The second assertion above indicates that

$$\lim_{K\to\infty} \theta(P_K, P) \leq \lambda(P)$$

with probability one, where

$$\lambda(P) = 8R + 2R\|A\| + 2\|q_1\| + 2\sup_{x\in[l,u]} \|\mathbb{E}_P[B(\xi)y^*(x,\xi)]\| + 1. \tag{3.20}$$

All these imply that

$$\lim_{K\to\infty} \sup_{x\in[l,u]} |f_P(x) - f_{P_K}(x)| = 0$$

with probability one. Due to Lemma 3.3, we obtain

$$\lim_{K\to\infty} \psi_P^{-1}\left(\sup_{x\in[l,u]} |f_P(x) - f_{P_K}(x)|\right) = 0$$

with probability one, which completes the proof. □

Furthermore, under ordinary conditions, the exponential rate of convergence can be derived. It is noteworthy that an earlier paper about the exponential rate of convergence is [48]. The authors studied the uniformly exponential convergence of the SAA for stochastic mathematical programs with variational constraints through the Cramér's Large Deviation Theorem.

We can derive from the above discussion that $[B(\xi)y^*(x,\xi)]_i$ is Lipschitz continuous w.r.t. x for $i = 1, 2, \cdots, n$ under Assumption 2.1. Concretely,

$$\begin{aligned} |[B(\xi)y^*(x_1,\xi)]_i - [B(\xi)y^*(x_2,\xi)]_i| &= \|B(\xi)y^*(x_1,\xi) - B(\xi)y^*(x_2,\xi)\| \\ &\leq \|B(\xi)\|\,\|N(\xi)\|\,\|x_1 - x_2\|/\kappa_M(\xi) \\ &= C(\xi)\,\|x_1 - x_2\| \end{aligned}$$

for $i = 1, 2, \cdots, n$, where $C(\xi) = \|B(\xi)\|\,\|N(\xi)\|/\kappa_M(\xi)$.

To establish the exponential rate of convergence, similar to that in [48], we need the following assumptions.

Assumption 3.3. *Let the following assertions hold:*

(i) For each $x \in [l,u]$, the moment generating functions of random variables $[B(\xi)y^*(x,\xi)]_i - (\mathbb{E}_P[B(\xi)y^*(x,\xi)])_i$, i.e.,

$$\mathbb{E}_P[\exp(t([B(\xi)y^*(x,\xi)]_i - (\mathbb{E}_P[B(\xi)y^*(x,\xi)])_i))]$$

for $i = 1, 2, \cdots, n$, are finite valued for each t in a neighborhood of zero.

(ii) The moment generating function of $C(\xi)$, i.e.,

$$\mathbb{E}_P[\exp(tC(\xi))]$$

is finite valued for each t in a neighborhood of zero.

Proposition 3.7. *Let Assumptions 2.1 and 3.3 hold. Then for any $\epsilon > 0$, there exist two positive scalars $L(\epsilon)$ and $\beta(\epsilon)$ which depend only on ϵ, such that*

$$\mathbb{P}\left\{\sup_{x\in[l,u]} \|\mathbb{E}_{P_K}[B(\xi)y^*(x,\xi)] - \mathbb{E}_P[B(\xi)y^*(x,\xi)]\| \geq \epsilon\right\} \leq L(\epsilon)\exp(-K\beta(\epsilon)).$$

This proposition can be directly obtained from [48, Theorem 5.1]. We thus omit the proof here.

From Proposition 3.7, we can immediately obtain the following exponential rate of convergence about the optimal solution set.

Theorem 3.4. *Let Assumptions 2.1 and 3.3 hold. Then, for any $\epsilon > 0$, there exist two positive scalars $\bar{L}(\epsilon)$ and $\bar{\beta}(\epsilon)$, such that*

$$\mathbb{P}\left\{d(S(P_K), S(P)) \geq \epsilon\right\} \leq \bar{L}(\epsilon)\exp(-K\bar{\beta}(\epsilon)).$$

Proof. We have from (3.18) the following estimation:

$$\begin{aligned}\mathbb{P}\{d(S(P_K), S(P)) \geq \epsilon\} &\leq \mathbb{P}\left\{\psi_P^{-1}\left(\sup_{x\in[l,u]} |f_P(x) - f_{P_K}(x)|\right) \geq \epsilon\right\} \\ &\leq \mathbb{P}\left\{\sup_{x\in[l,u]} |f_P(x) - f_{P_K}(x)| \geq \psi_P(\epsilon)\right\}.\end{aligned}$$

The second inequality follows from the nondecreasing property of ψ_P.

We know from Proposition 3.7 that

$$\mathbb{P}\left\{\sup_{x\in[l,u]} \|\mathbb{E}_{P_K}[B(\xi)y^*(x,\xi)] - \mathbb{E}_P[B(\xi)y^*(x,\xi)]\| < 1\right\} \geq 1 - L(1)\exp(-K\beta(1)),$$

which implies

$$\mathbb{P}\left\{\sup_{x\in[l,u]} \|\mathbb{E}_{P_K}[B(\xi)y^*(x,\xi)]\| < \sup_{x\in[l,u]} \|\mathbb{E}_P[B(\xi)y^*(x,\xi)]\| + 1\right\}$$
$$\geq 1 - L(1)\exp(-K\beta(1)).$$

In addition, we have that

$$\sup_{x\in[l,u]} |f_P(x) - f_{P_K}(x)| \leq$$
$$\sup_{x\in[l,u]} \|\mathbb{E}_P[B(\xi)y^*(x,\xi)] - \mathbb{E}_{P_K}[B(\xi)y^*(x,\xi)]\| \cdot \theta(P_K, P),$$

where $\theta(P_K, P)$ is defined in (3.19). Therefore, we obtain

$$\mathbb{P}\left\{\theta(P_K, P) < \lambda(P)\right\} \geq 1 - L(1)\exp(-K\beta(1)),$$

where $\lambda(P)$ is defined in (3.20). Thus, we continue

$$\mathbb{P}\left\{\sup_{x\in[l,u]} |f_P(x) - f_{P_K}(x)| \geq \psi_P(\epsilon)\right\}$$
$$\leq \mathbb{P}\left\{\sup_{x\in[l,u]} \|\mathbb{E}_P[B(\xi)y^*(x,\xi)] - \mathbb{E}_{P_K}[B(\xi)y^*(x,\xi)]\| \cdot \theta(P_K, P) \geq \psi_P(\epsilon)\right\}$$
$$\leq L(1)\exp(-K\beta(1)) +$$
$$\mathbb{P}\left\{\sup_{x\in[l,u]} \|\mathbb{E}_P[B(\xi)y^*(x,\xi)] - \mathbb{E}_{P_K}[B(\xi)y^*(x,\xi)]\| \cdot \lambda(P) \geq \psi_P(\epsilon)\right\}$$
$$\leq L(1)\exp(-K\beta(1)) + L\left(\frac{\psi_P(\epsilon)}{\lambda(P)}\right)\exp\left(-K\beta\left(\frac{\psi_P(\epsilon)}{\lambda(P)}\right)\right),$$

where the third inequality comes from Proposition 3.7.

Letting

$$\bar{L}(\epsilon) := L(1) + L\left(\frac{\psi_P(\epsilon)}{\lambda(P)}\right)$$

and

$$\bar{\beta}(\epsilon) := \min\left\{\beta(1), \beta\left(\frac{\psi_P(\epsilon)}{\lambda(P)}\right)\right\},$$

we completes the proof. □

In this section, we study the discrete approximation properties of problem (2.1) under mild conditions. The convergence of the SAA is derived in Theorem 3.3. However, this result did not address an important issue which is interesting from both the theoretical and computational points of view. That is, what is the rate of convergence or how large should the sample size be to achieve a desired accuracy of SAA estimators? We supplement it in Theorem 3.4 under ordinary assumptions. These estimates provide an important insight into the theoretical complexity and practical application of the considered problem (2.1).

3.3 Numerical results

To illustrate the application of the two-stage stochastic linear variational inequality problem (2.1) and to verify the obtained convergence results, we consider in this section a multi-player noncooperative two-stage game problem (see also [14, 11] for the two-players case) and its numerical solution. There is a significant amount of recent research on this topic. For example, [5] investigated the two-stage game wherein each player is risk-averse and solved a rival-parameterized stochastic program with quadratic recourse. The convergence results for different versions of the best-response schemes are discussed. [28] considered a stochastic Nash game where each player minimizes a parameterized expectation-valued convex objective function by proposing three inexact proximal best-response schemes. Different from those in [28, 5] where the Nash equilibrium point is determined by (inexact) best-response schemes, we employ the PHM to solve the discrete two-stage stochastic variational inequality problem (3.14).

3.3.1 A multi-player noncooperative two-stage game

Two-stage stochastic variational inequality problems have many practical applications (see [10]). Here we consider the multi-player (say $\mathcal{I}$ players) noncooperative two-stage game. It can be described in the form of the two-stage stochastic variational inequality problem (2.1). Let $(x_1, y_1(\cdot))$, $(x_2, y_2(\cdot)), \cdots$, $(x_{\mathcal{I}}, y_{\mathcal{I}}(\cdot)) \in \mathbb{R}^n \times \mathcal{Y}$ denote the decisions of player 1 to player $\mathcal{I}$ in the two-stage stochastic game, respectively. We use x_{-i} to denote all x_js for $j \neq i$ and so does y_{-i}. $\theta_i : \mathbb{R}^{n\mathcal{I}} \to \mathbb{R}$ is the cost function of player i in the first stage and $\phi_i : \mathbb{R}^{n\mathcal{I}} \times \mathcal{Y}^{\mathcal{I}} \times \Xi \to \mathbb{R}$ is the cost function of player i in the second stage. Then, to minimize his total cost, the player i $(i = 1, 2, \cdots, \mathcal{I})$ will make a decision through solving the following two-stage stochastic optimization problem:

$$\min_{x_i \in [l_i, u_i]} \quad \theta_i(x_i, x_{-i}) + \mathbb{E}_P[\varphi_i(x_i, x_{-i}, y_{-i}(\xi), \xi)], \tag{3.21}$$

for $l_i < u_i$ and $l_i, u_i \in \mathbb{R}^n$, $i = 1, 2, \cdots, \mathcal{I}$, where $\varphi_i(x_i, x_{-i}, y_{-i}(\xi), \xi)$ is defined by

$$\varphi_i(x_i, x_{-i}, y_{-i}(\xi), \xi) = \min_{y_i \geq 0} \quad \phi_i(x_i, x_{-i}, y_i, y_{-i}(\xi), \xi). \tag{3.22}$$

We know that a two-stage stochastic programming problem can be equivalently reformulated as a two-stage variational inequality problem from the first order optimality necessary conditions. Therefore, we consider the optimality condition of the two-stage stochastic program (3.21)-(3.22). To simplify the formulation, we assume that $\theta_i(\cdot, x_{-i})$ is differentiable w.r.t. x_i and $\phi_i(x_i, x_{-i}, \cdot, y_{-i}, \xi)$ is differentiable w.r.t. y_i. In addition, $\varphi_i(\cdot, x_{-i}, y_{-i}(\xi), \xi)$ is differentiable and Lipschitz continuous with some integrable Lipschitz constant w.r.t. x_i. Then, we know from [37, Theorem 7.49] that

$$\nabla_{x_i}\mathbb{E}_P[\varphi_i(x_i, x_{-i}, y_{-i}(\xi), \xi)] = \mathbb{E}_P[\nabla_{x_i}\varphi_i(x_i, x_{-i}, y_{-i}(\xi), \xi)].$$

Finally, we obtain the equivalent form of problem (3.21)-(3.22) as

$$\begin{cases} 0 \in \nabla_{x_i}\theta_i(x_i, x_{-i}) + \mathbb{E}_P[\nabla_{x_i}\varphi_i(x_i, x_{-i}, y_{-i}(\xi), \xi)] + \mathcal{N}_{[l_i,u_i]}(x_i) \\ 0 \le y_i \perp \nabla_{y_i}\phi_i(x_i, x_{-i}, y_i, y_{-i}, \xi) \ge 0, \text{ for a.e. } \xi \in \Xi, \end{cases} \tag{3.23}$$

for $i = 1, 2, \cdots, \mathcal{I}$.

To satisfy the above conditions and to obtain concrete numerical results, we consider a two-stage stochastic quadratic programming problem. Specifically, we define

$$\theta_i(x_i, x_{-i}) = \frac{1}{2}x_i^T H_i x_i + b_i^T x_i + \sum_{j \neq i} x_i^T P_j x_j$$

and

$$\phi_i(x_i, x_{-i}, y_i, y_{-i}, \xi) = \frac{1}{2}y_i^T Q_i(\xi) y_i + c_i(\xi)^T y_i + \sum_{j=1}^{I} y_i^T S_{ij}(\xi) x_j + \sum_{j \neq i} y_i^T O_j(\xi) y_j(\xi),$$

where $H_i, P_i \in \mathbb{R}^{n\times n}$, $S_{ij} : \Xi \to \mathbb{R}^{m\times n}$, $O_i : \Xi \to \mathbb{R}^{m\times m}$, $Q_i : \Xi \to \mathbb{R}^{m\times m}$, $b_i \in \mathbb{R}^n$, $c_i : \Xi \to \mathbb{R}^m$ for $i, j = 1, 2, \cdots, \mathcal{I}$.

With the above notation, we can rewrite problem (3.23) as the following large-scale two-stage stochastic linear variational inequality problem (see [14]):

$$\begin{cases} 0 \in Ax + \mathbb{E}_P[B(\xi)y(\xi)] + q_1 + \mathcal{N}_{[l,u]}(x), \\ 0 \le y(\xi) \perp M(\xi)y(\xi) + N(\xi)x + q_2(\xi) \ge 0, \text{ for a.e. } \xi \in \Xi, \end{cases} \tag{3.24}$$

where

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_{\mathcal{I}} \end{pmatrix}, \qquad y(\xi) = \begin{pmatrix} y_1(\xi) \\ \vdots \\ y_{\mathcal{I}}(\xi) \end{pmatrix}, \quad q_1 = \begin{pmatrix} b_1 \\ \vdots \\ b_{\mathcal{I}} \end{pmatrix},$$

$$q_2(\xi) = \begin{pmatrix} c_1(\xi) \\ \vdots \\ c_{\mathcal{I}}(\xi) \end{pmatrix}, \qquad l = \begin{pmatrix} l_1 \\ \vdots \\ l_{\mathcal{I}} \end{pmatrix}, \qquad u = \begin{pmatrix} u_1 \\ \vdots \\ u_{\mathcal{I}} \end{pmatrix},$$

$$A = \begin{pmatrix} H_1 & P_2 & \cdots & P_{\mathcal{I}} \\ P_1 & H_2 & \cdots & P_{\mathcal{I}} \\ \vdots & \vdots & \ddots & \vdots \\ P_1 & P_2 & \cdots & H_{\mathcal{I}} \end{pmatrix}, \quad B(\xi) = \begin{pmatrix} S_{11}^T(\xi) & 0 & \cdots & 0 \\ 0 & S_{22}^T(\xi) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & S_{\mathcal{I}\mathcal{I}}^T(\xi) \end{pmatrix},$$

$$M(\xi) = \begin{pmatrix} Q_1(\xi) & O_2(\xi) & \cdots & O_{\mathcal{I}}(\xi) \\ O_1(\xi) & Q_2(\xi) & \cdots & O_{\mathcal{I}}(\xi) \\ \vdots & \vdots & \ddots & \vdots \\ O_1(\xi) & O_2(\xi) & \cdots & Q_{\mathcal{I}}(\xi) \end{pmatrix}, N(\xi) = \begin{pmatrix} S_{11}(\xi) & S_{12}(\xi) & \cdots & S_{1\mathcal{I}}(\xi) \\ S_{21}(\xi) & S_{22}(\xi) & \cdots & S_{2\mathcal{I}}(\xi) \\ \vdots & \vdots & \ddots & \vdots \\ S_{\mathcal{I}1}(\xi) & S_{\mathcal{I}2}(\xi) & \cdots & S_{\mathcal{I}\mathcal{I}}(\xi) \end{pmatrix}.$$

A well-known algorithm for solving two-stage stochastic variational inequality problems is PHM, see [13, 19, 18]. The main idea of this algorithm is to construct a nonanticipative first stage solution through solving several discrete problems corresponding to individual scenarios. Let $\xi^1, \xi^2, \cdots, \xi^K$ be K samples or scenarios, and PHM can be stated as follows.

Algorithm: PHM to solve (3.24)
Step 0: Choose initial points: $\bar{x}_0$ and for $k = 1, 2, \cdots, K$, $x_0^k = \bar{x}_0$, y_0^k, w_0^k with $\sum_{k=1}^K w_0^k = 0$, $r > 0$ and set $i = 0$;
Step 1: If the termination criterion is satisfied, STOP. Otherwise, go to **Step 2**;
Step 2: For $k = 1, 2, \cdots, K$, solve the following deterministic two-stage mixed problem w.r.t. (x^k, y^k):

$$\begin{cases} 0 \in Ax^k + B(\xi^k)y^k + q_1 + w_i^k + r(x^k - x_i^k) + \mathcal{N}_{[l,u]}(x^k), \\ 0 \le y^k \perp M(\xi^k)y^k + N(\xi^k)x^k + q_2(\xi^k) + r(y^k - y_i^k) \ge 0. \end{cases} \tag{3.25}$$

The obtained solution is denoted by $(\hat{x}_i^k, \hat{y}_i^k)$;
Step 3: Let $\bar{x}_{i+1} = \frac{1}{K}\sum_{k=1}^K \hat{x}_i^k$. Then, for $k = 1, 2, \cdots, K$, set $x_{i+1}^k = \bar{x}_{i+1}$, $y_{i+1}^k = \hat{y}_i^k$ and $w_{i+1}^k = w_i^k + r(\hat{x}_i^k - \bar{x}_{i+1})$. Let $i = i + 1$ and go back to **Step 1**.

Very often, the termination criterion can be chosen as: The residual becomes sufficiently small, i.e.,

$$\frac{1}{K}\sum_{k=1}^{K}\left\|x_i^k - \mathrm{mid}\left\{l, u, x_i^k - \left(Ax_i^k + \frac{1}{K}\sum_{k=1}^{K}B(\xi^k)y_i^k + q_1\right)\right\}\right\|^2 \tag{3.26}$$

is sufficiently small.

To obtain concrete numerical results and ensure the convergence of the above PHM, we consider the following specific setting.

3.3.2 Parameter settings and numerical results

We consider a 3-player two-stage noncooperative game with $n = 4$, $m = 4$. We adopt the following stopping criterion for PHM: Either the residual in (3.26) is less than or equal to 10^{-5} or the iteration number i attains 6000. Arbitrarily generate $\hat{H}_i \in \mathbb{R}^{n\times n}$, $\hat{Q}_i \in \mathbb{R}^{m\times m}$, $\hat{P}_i \in \mathbb{R}^{n\times n}$, $\hat{S}_{ij} \in \mathbb{R}^{m\times n}$, $\hat{O}_i \in \mathbb{R}^{m\times m}$ with entries choosing from $[-1, 1]$ and $b_i \in \mathbb{R}^n$, $c_i \in \mathbb{R}^m$, for $i, j = 1, 2, 3$. Let $\xi = (\xi_1, \xi_2, \cdots, \xi_{18})$ be the random vector which follows a uniform distribution on the support set $[0, 1]^{18}$. Then we set $\hat{S}_{11}(\xi) = \xi_1\hat{S}_{11}$, $\hat{S}_{12}(\xi) = \xi_2\hat{S}_{12}$, $\hat{S}_{13}(\xi) = \xi_3\hat{S}_{13}$, $\hat{S}_{21}(\xi) = \xi_4\hat{S}_{21}$, $\hat{S}_{22}(\xi) = \xi_5\hat{S}_{22}$, $\hat{S}_{23}(\xi) = \xi_6\hat{S}_{23}$, $\hat{S}_{31}(\xi) = \xi_7\hat{S}_{31}$, $\hat{S}_{32}(\xi) = \xi_8\hat{S}_{32}$, $\hat{S}_{33}(\xi) = \xi_9\hat{S}_{33}$, $\hat{O}_1(\xi) = \xi_{10}\hat{O}_1$, $\hat{O}_2(\xi) = \xi_{11}\hat{O}_2$, $\hat{O}_3(\xi) = \xi_{12}\hat{O}_3$, $\hat{Q}_1(\xi) = \xi_{13}\hat{Q}_1$, $\hat{Q}_2(\xi) = \xi_{14}\hat{Q}_2$, $\hat{Q}_3(\xi) = \xi_{15}\hat{Q}_3$, $c_1(\xi) = \xi_{16}c_1$, $c_2(\xi) = \xi_{17}c_2$ and $c_3(\xi) = \xi_{18}c_3$. The main reason to choose the above random parameters is to satisfy Assumption 3.2, which is needed in Theorems 3.1 and 3.3. Meanwhile, there are plenty of existing works and applications where the parameters are assumed to be affinely linear w.r.t. ξ, see for example [30].

To ensure the positive definiteness of coefficient matrices in problem (3.24), we construct those matrices as follows:

$$
A = \begin{pmatrix} \hat{H}_1 & \hat{P}_2 & \hat{P}_3 \\ \hat{P}_1 & \hat{H}_2 & \hat{P}_3 \\ \hat{P}_1 & \hat{P}_2 & \hat{H}_3 \end{pmatrix} + \gamma I, \qquad B(\xi) = \begin{pmatrix} \hat{S}_{11}^T(\xi) & 0 & 0 \\ 0 & \hat{S}_{22}^T(\xi) & 0 \\ 0 & 0 & \hat{S}_{33}^T(\xi) \end{pmatrix},
$$
$$
M(\xi) = \begin{pmatrix} \hat{Q}_1(\xi) & \hat{O}_2(\xi) & \hat{O}_3(\xi) \\ \hat{O}_1(\xi) & \hat{Q}_2(\xi) & \hat{O}_3(\xi) \\ \hat{O}_1(\xi) & \hat{O}_2(\xi) & \hat{Q}_3(\xi) \end{pmatrix} + \gamma I, \quad N(\xi) = \begin{pmatrix} \hat{S}_{11}(\xi) & \hat{S}_{12}(\xi) & \hat{S}_{13}(\xi) \\ \hat{S}_{21}(\xi) & \hat{S}_{22}(\xi) & \hat{S}_{23}(\xi) \\ \hat{S}_{31}(\xi) & \hat{S}_{32}(\xi) & \hat{S}_{33}(\xi) \end{pmatrix},
$$
$$
q_1 = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}, \qquad q_2(\xi) = \begin{pmatrix} c_1(\xi) \\ c_2(\xi) \\ c_3(\xi) \end{pmatrix},
$$

where $\gamma = 3(m+n)$, I_{3n} and I_{3m} stand for the identity matrices in $\mathbb{R}^{3n\times 3n}$ and $\mathbb{R}^{3m\times 3m}$, respectively. Obviously, the above setting guarantees that Assumption 3.1 holds for any $\xi \in [0,1]^{18}$, which is sufficient for the convergence of the PHM. Due to the affine linearity of all the above coefficients, Assumption 3.2 holds with $\alpha = 1$. Moreover, we adopt the uniform distribution here that must satisfy (3.16) in Proposition 3.6. Therefore, Theorems 3.1 and 3.3 hold in our specific settings.

From (i) of Proposition 2.1, the solution of the second stage satisfies

$$
\|y^*(x,\xi)\| \le \|W(x,\xi)\| \left(\|N(\xi)\|\,\|x\| + \|q_2(\xi)\|\right) \le \Gamma
$$

uniformly for any $\xi \in [0,1]^{18}$ and some positive number Γ. This implies that we can employ the homotopy-smoothing method for box-constrained variational inequalities (see [27]) to solve the two-stage mixed problem (3.25) in **Step 2**.

With the above detailed parameter selection and the solution method in **Step 2**, we can then solve the concrete 3-player two-stage non-cooperative game problem. We show in Figures 3.1-3.2 the box plot for each component of the first stage decision variable x w.r.t. the number of samples. Since our parameter setting satisfies Assumption 3.1, there exists a unique solution for both the original problem and its SAA problem (see [11]). As we discussed before, Theorem 3.3 holds in our setting. For each sample size $K = 10, 50, 200, 500, 1000, 2000, 4000$, we solve 100 randomly generated problems and draw the empirical distribution of the solutions in Figures 3.1-3.2. The 12 plots in Figures 3.1-3.2 show the convergence of the SAA problem (3.14) by adopting the hybrid algorithm combining PHM [19] and the homotopy-smoothing method [27] as the sample size goes to infinity. Actually, we know from Theorem 3.3 and the uniqueness of solution that the SAA solutions will converge to the true solution with probability one.

Now we numerically verify the quantitative stability results in section 3.1 to this example. For this purpose, we assume that the original probability distribution P is the uniform distribution on interval $[0,1]^{18}$. The perturbed distribution Q_ν ($\nu \in \mathbb{N}$) is the uniform distribution with the support set being $\left[0, \frac{\nu}{\nu+1}\right]^{18}$, that is, the probability for taking values in $[0,1]^{18} \backslash \left[0, \frac{\nu}{\nu+1}\right]^{18}$ is zero. Then, we have

$$\begin{aligned}\mathbb{D}_{TV}(P, Q_\nu) &= \sup_{h \in \mathcal{G}_{TV}} \left\{ \int_{\left[0,\frac{\nu}{\nu+1}\right]^{18}} h(\xi) \left(\left(\frac{\nu+1}{\nu} \right)^{18} - 1 \right) d\xi - \int_{[0,1]^{18} \backslash \left[0,\frac{\nu}{\nu+1}\right]^{18}} h(\xi) d\xi \right\} \\ &= 2 \left[1 - \left(\frac{\nu}{\nu+1} \right)^{18} \right]. \end{aligned} \tag{3.27}$$

Here the optimal element in $\mathcal{G}_{TV}$ is

$$h(\xi) = \begin{cases} 1, & \xi \in \left[0, \frac{\nu}{\nu+1}\right]^{18}; \\ -1, & \xi \in [0,1]^{18} \backslash \left[0, \frac{\nu}{\nu+1}\right]^{18}. \end{cases}$$

Therefore, $\mathbb{D}_{TV}(P, Q_\nu) \to 0$ as $\nu \to +\infty$. In what follows, we fix the number of scenarios at $K = 5000$ and use the sample approximation problem to approximate the original problem. Let $\nu = 1, 2, 3, 4, 5, 6, 7$, we use PHM to solve the original problem under P and the corresponding problem under perturbed distribution Q_ν, respectively. Since Assumption 3.1 holds, there always exist a unique solution for the original problem under P, as well as the problem under the perturbation Q_ν.

We calculate the distance between the unique solution x^* under probability distribution P and the unique solution x^*_ν under probability distribution Q_ν. It is known from Theorem 3.1 that

$$\|x^* - x_\nu^*\| \leq \psi_P^{-1}(L_1 \mathbb{D}_{TV}(P, Q_\nu)) \tag{3.28}$$

for some positive constant L_1. Note that ψ_P^{-1} is lower semicontinuous and nondecreasing, and vanishes at 0. Specially, under our specific setting, we know from

$$\psi_P(\tau) = \min\{f_P(x) = f_P(x) - f_P(x^*) : d(x, x^*) \geq \tau, x \in [l, u]\},$$

where $f_P(x^*) = 0$, and the continuity of $f_P(x)$ w.r.t. x that ψ_P is continuous at 0. Moreover, $\psi_P(\tau) > 0$ for any $\tau > 0$ due to the uniqueness of solutions. Its inverse function is defined by (3.9), that is,

$$\psi_P^{-1}(t) = \sup\{\tau \in \mathbb{R}_+ : \psi_P(\tau) \leq t\}$$

is continuous at $t = 0$, see Lemma 3.3.

Based on the above discussion, we have from (3.28) that $\|x^* - x_\nu^*\|$ should converge to 0 as $\nu \to \infty$. Table 3.1 shows this kind of convergence. We can see from Table 3.1 that the distance between x^* and x_ν^* monotonically decreases with the increase of ν. These results perfectly illustrate and support the quantitative analysis results in section 3.1.

Table 3.1: The distance between the pairing solutions under P and Q_ν

ν	1	2	3	4	5	6	7
$\|x^* - x_\nu^*\|$	1.33e-2	1.00e-2	0.77e-2	0.64e-2	0.54e-2	0.49e-2	0.41e-2

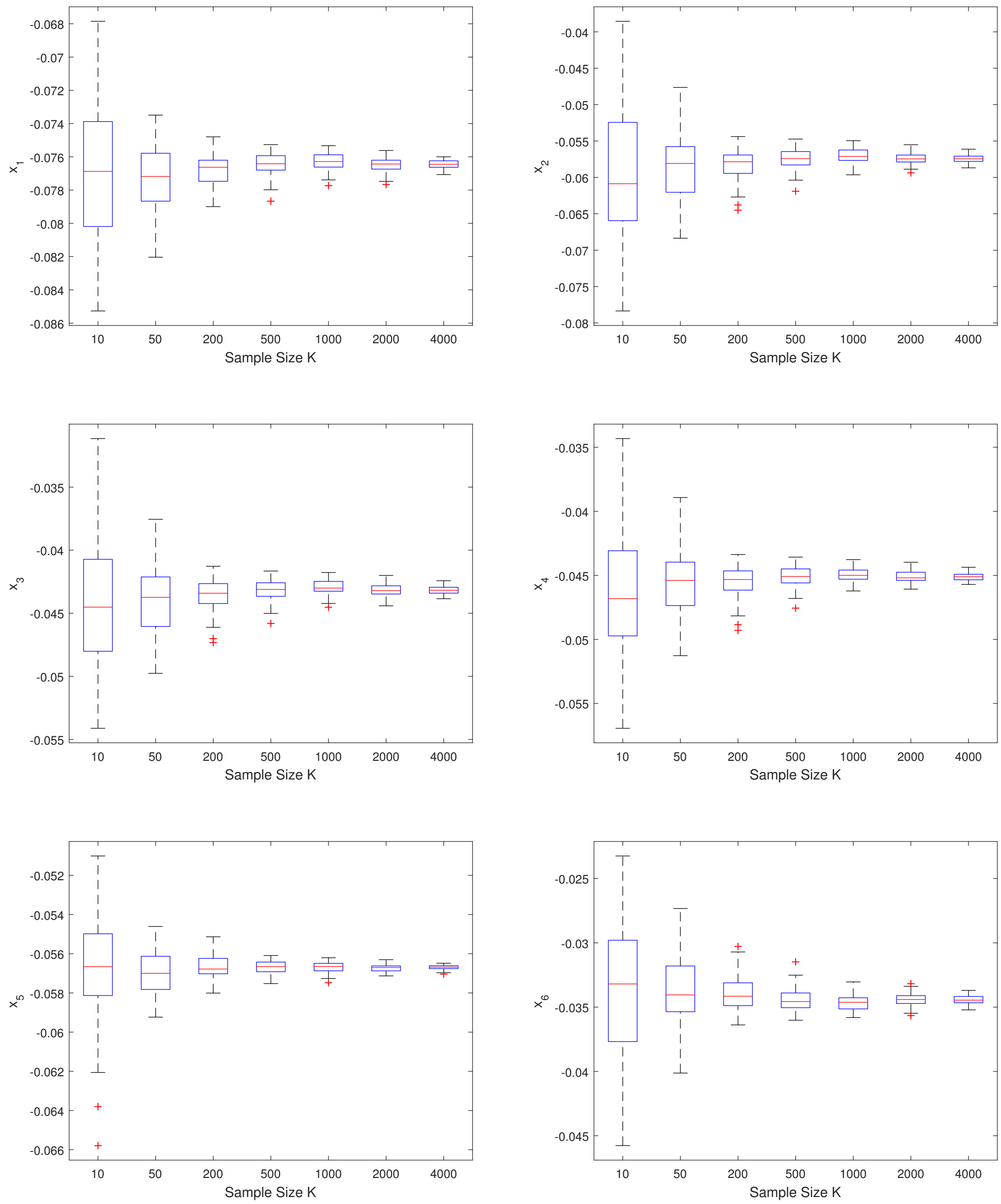

Figure 3.1: The box plots for x_1 to x_6 with different sample size

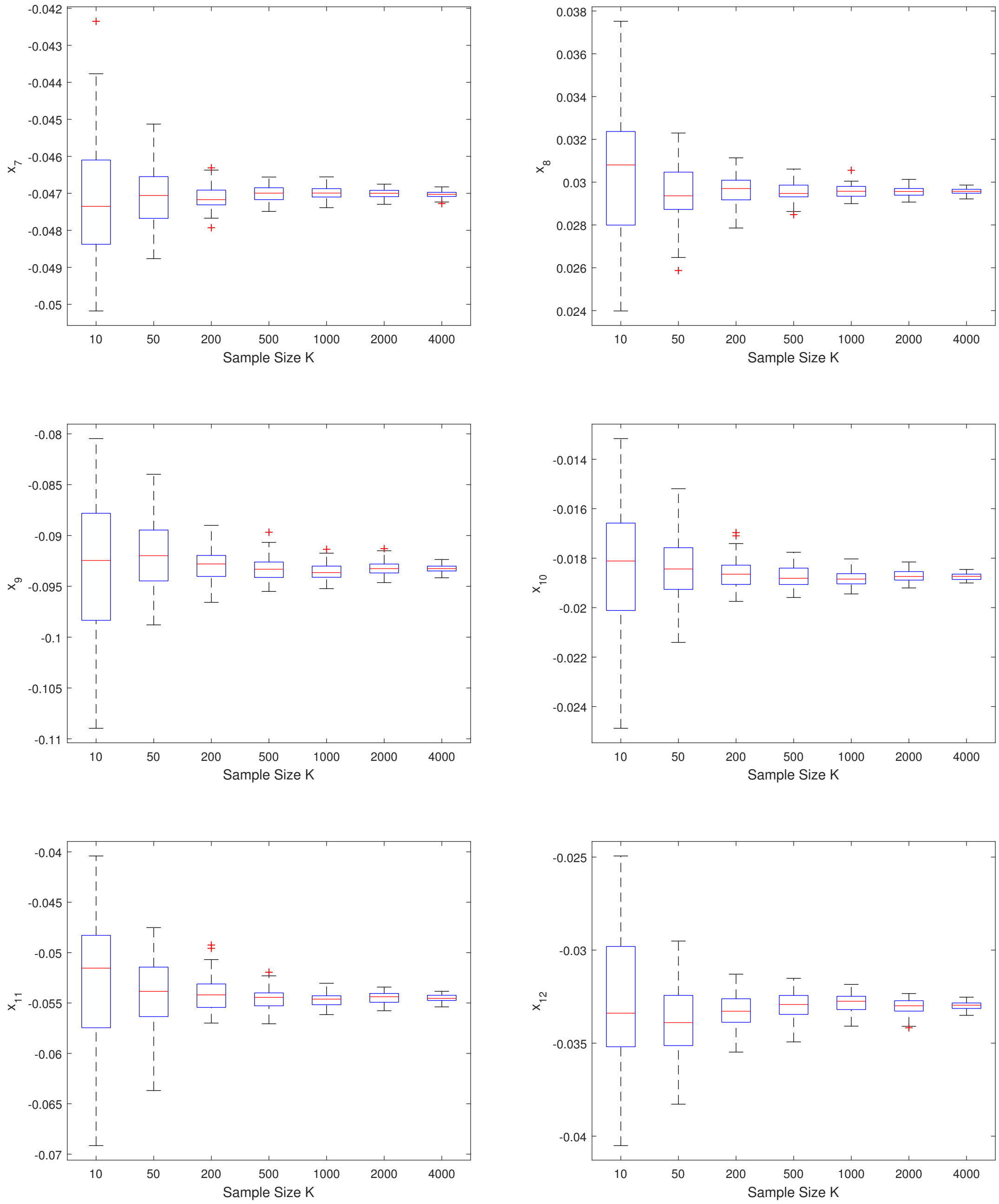

Figure 3.2: The box plots for x_7 to x_{12} with different sample size

Chapter 4

Regularization and convergence of (2.12)

Since the discussion in Chapter 3 is based on the assumption that the second stage problem has a unique solution, we weaken this assumption in this chapter. More specifically, we consider a two-stage stochastic Cournot-Nash game problem (2.12) where the second stage problem is monotone. The monotonicity of the second stage problem cannot ensure the uniqueness of solution. Thus, we propose a regularized SAA method to handle it. Corresponding convergence analysis is studied.

4.1 Structure of the regularized two-stage SLCP

In this section, we focus on characterizing solutions of two-stage stochastic linear complementarity problem (2.12). From the derivation of first-order necessary optimality conditions of problem (2.9)-(2.10) and the monotonicity of problem (2.12), we have the following results on existence of solutions.

Proposition 4.1 (Theorem 2, [49])**.** *For any fixed pair* $(x, \xi) \in \mathbb{R}^J_+ \times \Xi$*, the second stage problem* (2.10) *has a unique solution.*

Thus, for the two-stage stochastic linear complementarity problem (2.12), the following proposition holds.

Proposition 4.2. *The two-stage stochastic linear complementarity problem* (2.12) *has relatively complete recourse, i.e., for any $x \in \mathbb{R}^J_+$ and a.e. $\xi \in \Xi$ the second stage problem of* (2.12) *is solvable.*

Proof. The coefficient matrix of the second stage part of (2.12)

$$M(\xi) = \begin{pmatrix} \Pi(\xi) & I \\ -I & 0 \end{pmatrix}$$

is positive semidefinite for a.e. $\xi \in \Xi$.

For any given $x \in \mathbb{R}^J_+$, it follows that there always exists a pair $\left(\hat{y}(\xi), \hat{\lambda}(\xi)\right) \in \mathbb{R}^J \times \mathbb{R}^J$, such that

$$\begin{pmatrix} \hat{y}(\xi) \\ \hat{\lambda}(\xi) \end{pmatrix} \geq 0, \quad \begin{pmatrix} \Pi(\xi) & I \\ -I & 0 \end{pmatrix} \begin{pmatrix} \hat{y}(\xi) \\ \hat{\lambda}(x) \end{pmatrix} + \begin{pmatrix} -p(\xi) \\ x \end{pmatrix} \geq 0, \qquad \text{for a.e. } \xi \in \Xi.$$

In detail, we consider a special choice $\hat{y}(\xi) = 0$ and $\hat{\lambda}(\xi) = \max\{0, p(\xi)\}$, where the max function is taken componentwise. Thus, the corresponding quadratic programming problem of the linear complementarity problem is feasible. It follows from [15, Lemma 3.1.1, Theorem 3.1.2] that there must exist at least a solution which solves the second stage problem for any given pair (x, ξ). □

Although the second stage problem (2.10) has a unique equilibrium for any given (x, ξ) (see Proposition 4.1), the system (2.12) may admit multiple solutions. To see this, we give an illustrative example.

Example 4.1. *Consider a duopoly game, with given $x = (x_1, x_2)^T \geq 0$, and $-p(\xi) \geq_{a.s.} 0$. Then, the corresponding second stage part of complementarity system* (2.12) *reads*

$$0 \leq \begin{pmatrix} y_1(\xi) \\ y_2(\xi) \\ \lambda_1(\xi) \\ \lambda_2(\xi) \end{pmatrix} \perp \begin{pmatrix} 2\gamma(\xi) & \gamma(\xi) & 1 & 0 \\ \gamma(\xi) & 2\gamma(\xi) & 0 & 1 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix} \begin{pmatrix} y_1(\xi) \\ y_2(\xi) \\ \lambda_1(\xi) \\ \lambda_2(\xi) \end{pmatrix} + \begin{pmatrix} -p_1(\xi) \\ -p_2(\xi) \\ x_1 \\ x_2 \end{pmatrix} \geq 0, \quad \textit{for a.e. } \xi \in \Xi. \tag{4.1}$$

Then, the solution set of (4.1) *is of the following form*

$$\left\{(0,0,\tilde{\lambda}_1(\xi),\tilde{\lambda}_2(\xi)) : \tilde{\lambda}_1(\xi) = \begin{cases} 0, & x_1 > 0 \\ \lambda_1(\xi), & x_1 = 0 \end{cases},\right.$$

$$\left.\tilde{\lambda}_2(\xi) = \begin{cases} 0, & x_2 > 0 \\ \lambda_2(\xi), & x_2 = 0 \end{cases}, \quad \textit{for a.e. } \xi \in \Xi\right\},$$

where $\lambda_1(\xi) \geq_{a.e.} 0, \lambda_2(\xi) \geq_{a.e.} 0$.

In Example 4.1, the "equilibrium price" λ may admit multiple values when there exist some zero-valued components of x.

Technically, the multiple solutions of the second stage problem will cause trouble when we handle the two-stage stochastic complementarity system (2.12), both in computation and analysis [37]. The assumption ensuring the uniqueness of second stage solution is usually made, see for instance [14, 11]. Moreover, interpreted as "equilibrium price" associated with agents' production clearing, different values of λ would have ambiguous economical interpretations. Motivated by these, we propose a regularized method to seek for one particular choice of "equilibrium price". Similar approach can be found in for example [50].

For any $\epsilon > 0$, let

$$M^\epsilon(\xi) = \begin{pmatrix} \Pi(\xi) & I \\ -I & \epsilon I \end{pmatrix} \text{ and } q(x,\xi) = \begin{pmatrix} -p(\xi) \\ x \end{pmatrix}$$

and thus, we can write the regularized second stage SCP as

$$0 \leq \begin{pmatrix} y(\xi) \\ \lambda(\xi) \end{pmatrix} \perp M^\epsilon(\xi) \begin{pmatrix} y(\xi) \\ \lambda(\xi) \end{pmatrix} + q(x,\xi) \geq 0, \quad \text{for a.e. } \xi \in \Xi. \tag{4.2}$$

Then, we have the regularized SCP of (2.12) as follows:

$$\begin{aligned} &0 \leq x && \perp Cx - \mathbb{E}[\lambda(\xi)] + a && \geq 0, \\ &0 \leq \begin{pmatrix} y(\xi) \\ \lambda(\xi) \end{pmatrix} && \perp \begin{pmatrix} \Pi(\xi) & I \\ -I & \epsilon I \end{pmatrix} \begin{pmatrix} y(\xi) \\ \lambda(\xi) \end{pmatrix} + q(x,\xi) && \geq 0, \quad \text{for a.e. } \xi \in \Xi. \end{aligned} \tag{4.3}$$

For a given pair $(x, \xi) \in \mathbb{R}^J_+ \times \Xi$, the second stage problem of (2.12) and the regularized second stage problem (4.2) are denoted by $\mathrm{LCP}(q(x,\xi), M(\xi))$ and $\mathrm{LCP}(q(x,\xi), M^\epsilon(\xi))$ respectively. Their solution functions are chosen from the respective solution sets and expressed by $z(q(x,\xi))$ and $z^\epsilon(q(x,\xi))$. In the sequel, we omit the ξ and x without causing confusion, i.e., $\mathrm{LCP}(q, M) := \mathrm{LCP}(q(x,\xi), M(\xi))$ and $\mathrm{LCP}(q, M^\epsilon) := \mathrm{LCP}(q(x,\xi), M^\epsilon(\xi))$.

For clearer demonstration, recall our illustrative Example 4.1, and consider its regularization approach. Thus, the second stage of the regularized problem takes the following form

$$0 \le \begin{pmatrix} y_1(\xi) \\ y_2(\xi) \\ \lambda_1(\xi) \\ \lambda_2(\xi) \end{pmatrix} \perp \begin{pmatrix} 2\gamma(\xi) & \gamma(\xi) & 1 & 0 \\ \gamma(\xi) & 2\gamma(\xi) & 0 & 1 \\ -1 & 0 & \epsilon & 0 \\ 0 & -1 & 0 & \epsilon \end{pmatrix} \begin{pmatrix} y_1(\xi) \\ y_2(\xi) \\ \lambda_1(\xi) \\ \lambda_2(\xi) \end{pmatrix} + \begin{pmatrix} -p_1(\xi) \\ -p_2(\xi) \\ x_1 \\ x_2 \end{pmatrix} \ge 0, \quad \text{for a.e. } \xi \in \Xi. \tag{4.4}$$

Under the same condition as in Example 4.1, we can obtain the unique solution of (4.4), which $\tilde{y}_1, \tilde{y}_2, \tilde{\lambda}_1, \tilde{\lambda}_2$ equal to 0 for a.e. $\xi \in \Xi$. Due to the positive definiteness of C, it follows that we obtain the unique solution of the first stage problem is $x_1 = 0, x_2 = 0$. Then, we have obtained one particular solution of the original problem, the trivial solution in this example. The key feature of our regularized method is that it promises the existence and uniqueness of solution due to the strongly monotone of regularized two-stage problem.

In the remaining of this section, we concern ourselves with the solution z^ϵ of $\mathrm{LCP}(q, M^\epsilon)$ and explore the structure of the second stage solution.

Proposition 4.3. *For any fixed $\epsilon > 0$, the regularized problem* (4.3) *has a unique solution $(x^\epsilon, y^\epsilon, \lambda^\epsilon) \in \mathbb{R}^J \times \mathcal{Y} \times \mathcal{Y}$.*

Proof. The result can be obtained via a similar procedure as in [11, Proposition 2.1 (i)] and we only need to show that the condition in [11, Assumption 1] holds. Recall that Assumption 2.2 holds, then for a.e. $\xi \in \Xi$

$$\begin{pmatrix} x \\ u(\xi) \\ v(\xi) \end{pmatrix}^T \begin{pmatrix} C & 0 & -I \\ 0 & \Pi(\xi) & I \\ I & -I & \epsilon I \end{pmatrix} \begin{pmatrix} x \\ u(\xi) \\ v(\xi) \end{pmatrix} \geq \tau(\|x\|^2 + \|u(\xi)\|^2 + \|v(\xi)\|^2),$$

where $\tau = 2\min\{\bar{c}, \gamma_0(J+1), \epsilon\}$ with $\bar{c}$ denoting the minimum diagonal element of C. □

Theorem 4.1. *For any fixed $\epsilon > 0$, $x \geq 0$ and a.e. $\xi \in \Xi$, the jth component of the solution of problem (4.2) $\big((y^\epsilon)_j, (\lambda^\epsilon)_j\big)$ is either $(0,0)$, or one of the following two forms:*

$$\begin{aligned} &-\left(\frac{\gamma(\xi)T^\epsilon - p_j(\xi)}{\gamma(\xi)}, \qquad 0\right), \\ &-\left(\frac{\epsilon(\gamma(\xi)T^\epsilon - p_j(\xi)) - x_j}{\epsilon\gamma(\xi)+1}, \quad \frac{\gamma(\xi)(T^\epsilon + x_j) - p_j(\xi)}{\epsilon\gamma(\xi)+1}\right) \end{aligned} \tag{4.5}$$

for $j \in \mathcal{J}$, where

$$T^\epsilon := \sum_{i=1}^{J} (y^\epsilon)_i = \frac{\gamma(\xi)\sum_{i\in\mathcal{I}_3} x_i + \epsilon\gamma(\xi)\sum_{i\in\mathcal{I}_2\cup\mathcal{I}_3} p_i(\xi) + \sum_{i\in\mathcal{I}_2} p_i(\xi)}{\big(\epsilon\gamma(\xi)(|\mathcal{I}_2| + |\mathcal{I}_3| + 1) + |\mathcal{I}_2| + 1\big)\gamma(\xi)} \tag{4.6}$$

with

$$\begin{aligned} \mathcal{I}_2 &= \{j \in \mathcal{J} : \gamma(\xi)T^\epsilon + (\lambda^\epsilon)_j - p_j(\xi) < 0, \quad (y^\epsilon)_j - x_j \leq 0\}, \\ \mathcal{I}_3 &= \{j \in \mathcal{J} : \gamma(\xi)T^\epsilon + (\lambda^\epsilon)_j - p_j(\xi) < 0, \quad (y^\epsilon)_j - x_j > 0\}, \end{aligned}$$

where $|\mathcal{I}_2|$ and $|\mathcal{I}_3|$ denote the cardinality of $\mathcal{I}_2$ and $\mathcal{I}_3$ respectively.

Proof. By direct computation, we have that

$$\begin{aligned} &\left(M^\epsilon(\xi)\begin{pmatrix} y^\epsilon \\ \lambda^\epsilon \end{pmatrix} + q(x,\xi)\right)_j \\ &\qquad = \begin{cases} \gamma(\xi)(y^\epsilon)_j + \gamma(\xi)T^\epsilon + (\lambda^\epsilon)_j - p_j(\xi), \ j = 1,\ldots,J; \\ x_{j-J} - (y^\epsilon)_{j-J} + \epsilon(\lambda^\epsilon)_{j-J}, \ j = J+1,\ldots,2J. \end{cases} \end{aligned}$$

Then, we can rewrite problem (4.2) as below:

$$\begin{cases} 0 \le (y^\epsilon)_j \perp \gamma(\xi)(y^\epsilon)_j + \gamma(\xi)T^\epsilon + (\lambda^\epsilon)_j - p_j(\xi) \ge 0, \\ 0 \le (\lambda^\epsilon)_j \perp x_j - (y^\epsilon)_j + \epsilon(\lambda^\epsilon)_j \ge 0, \end{cases} \tag{4.7}$$

for $j \in \mathcal{J}$. From the first complementarity condition in (4.7), we have $(y^\epsilon)_j$ as follows:

$$(y^\epsilon)_j = \begin{cases} -\dfrac{\gamma(\xi)T^\epsilon + (\lambda^\epsilon)_j - p_j(\xi)}{\gamma(\xi)}, & \gamma(\xi)T^\epsilon + (\lambda^\epsilon)_j - p_j(\xi) < 0; \\ 0, & \gamma(\xi)T^\epsilon + (\lambda^\epsilon)_j - p_j(\xi) \ge 0 \end{cases} \tag{4.8}$$

for $j \in \mathcal{J}$. Similarly, we can derive that

$$(\lambda^\epsilon)_j = \begin{cases} \dfrac{(y^\epsilon)_j - x_j}{\epsilon}, & (y^\epsilon)_j - x_j > 0; \\ 0, & (y^\epsilon)_j - x_j \le 0 \end{cases} \tag{4.9}$$

for $j \in \mathcal{J}$. Note that $(y^\epsilon)_j = 0$ implies $(y^\epsilon)_j = 0 \le x_j$, and we have $(\lambda^\epsilon)_j = 0$. Then, based on (4.9), we have for all three cases:

$$\begin{cases} (y^\epsilon)_j = 0, \ (\lambda^\epsilon)_j = 0 & \text{for } j \in \mathcal{I}_1; \\ (y^\epsilon)_j = -\dfrac{\gamma(\xi)T^\epsilon + (\lambda^\epsilon)_j - p_j(\xi)}{\gamma(\xi)}, \ (\lambda^\epsilon)_j = 0 & \text{for } j \in \mathcal{I}_2; \\ (y^\epsilon)_j = -\dfrac{\gamma(\xi)T^\epsilon + (\lambda^\epsilon)_j - p_j(\xi)}{\gamma(\xi)}, \ (\lambda^\epsilon)_j = \dfrac{(y^\epsilon)_j - x_j}{\epsilon} & \text{for } j \in \mathcal{I}_3, \end{cases}$$

where

$$\mathcal{I}_1 := \left\{ j \in \mathcal{J} : \gamma(\xi)T^\epsilon + (\lambda^\epsilon)_j - p_j(\xi) \ge 0, (y^\epsilon)_j - x_j \le 0 \right\},$$

$$\mathcal{I}_2 := \left\{ j \in \mathcal{J} : \gamma(\xi)T^\epsilon + (\lambda^\epsilon)_j - p_j(\xi) < 0, (y^\epsilon)_j - x_j \le 0 \right\},$$

$$\mathcal{I}_3 := \left\{ j \in \mathcal{J} : \gamma(\xi)T^\epsilon + (\lambda^\epsilon)_j - p_j(\xi) < 0, (y^\epsilon)_j - x_j > 0 \right\}.$$

It follows that,

$$((y^\epsilon)_j, (\lambda^\epsilon)_j) = \begin{cases} (0,0), & j \in \mathcal{I}_1; \\ \left(-\dfrac{\gamma(\xi)T^\epsilon - p_j(\xi)}{\gamma(\xi)}, 0 \right), & j \in \mathcal{I}_2; \\ \left(-\dfrac{\epsilon\gamma(\xi)T^\epsilon - x_j - \epsilon p_j(\xi)}{\epsilon\gamma(\xi) + 1}, -\dfrac{\gamma(\xi)(T^\epsilon + x_j) - p_j(\xi)}{\epsilon\gamma(\xi) + 1} \right), & j \in \mathcal{I}_3, \end{cases}$$

which verifies (4.5). For the remaining of the proof, let $j \in \mathcal{I}_2$, we have

$$-\gamma(\xi)(y^\epsilon)_j = \gamma(\xi)T^\epsilon - p_j(\xi)$$

and thus

$$-\gamma(\xi)\sum_{i\in\mathcal{I}_2}(y^\epsilon)_i = |\mathcal{I}_2|\gamma(\xi)T^\epsilon - \sum_{i\in\mathcal{I}_2} p_i(\xi). \tag{4.10}$$

Analogously, we can derive from

$$(y^\epsilon)_j = -\frac{\gamma(\xi)T^\epsilon + (\lambda^\epsilon)_j - p_j(\xi)}{\gamma(\xi)} \text{ and } (\lambda^\epsilon)_j = \frac{(y^\epsilon)_j - x_j}{\epsilon} \text{ for } j \in \mathcal{I}_3$$

that

$$-\gamma(\xi)\sum_{i\in\mathcal{I}_3}(y^\epsilon)_i = |\mathcal{I}_3|\gamma(\xi)T^\epsilon + \frac{1}{\epsilon}\sum_{i\in\mathcal{I}_3}(y^\epsilon)_i - \frac{1}{\epsilon}\sum_{i\in\mathcal{I}_3} x_i - \sum_{i\in\mathcal{I}_3} p_i(\xi). \tag{4.11}$$

Combining (4.10) and (4.11), we obtain

$$-\gamma(\xi)T^\epsilon = (|\mathcal{I}_2| + |\mathcal{I}_3|)\gamma(\xi)T^\epsilon + \frac{1}{\epsilon}\sum_{i\in\mathcal{I}_3}(y^\epsilon)_i - \frac{1}{\epsilon}\sum_{i\in\mathcal{I}_3} x_i - \sum_{i\in\mathcal{I}_2\cup\mathcal{I}_3} p_i(\xi).$$

Therefore, we have

$$\frac{1}{\epsilon}\sum_{i\in\mathcal{I}_3}(y^\epsilon)_i = -(|\mathcal{I}_2| + |\mathcal{I}_3| + 1)\gamma(\xi)T^\epsilon + \frac{1}{\epsilon}\sum_{i\in\mathcal{I}_3} x_i + \sum_{i\in\mathcal{I}_2\cup\mathcal{I}_3} p_i(\xi). \tag{4.12}$$

Substituting (4.12) into (4.11), we have

$$\begin{aligned}
|\mathcal{I}_3|\,\gamma(\xi)T^\epsilon =& -\left(\gamma(\xi) + \frac{1}{\epsilon}\right)\sum_{i\in\mathcal{I}_3}(y^\epsilon)_i + \frac{1}{\epsilon}\sum_{i\in\mathcal{I}_3} x_i + \sum_{i\in\mathcal{I}_3} p_i(\xi) \\
=& -(\epsilon\gamma(\xi)+1)\left(-(|\mathcal{I}_2| + |\mathcal{I}_3| + 1)\gamma(\xi)T^\epsilon + \frac{1}{\epsilon}\sum_{i\in\mathcal{I}_3} x_i + \sum_{i\in\mathcal{I}_2\cup\mathcal{I}_3} p(\xi)_i\right) \\
&+ \frac{1}{\epsilon}\sum_{i\in\mathcal{I}_3} x_i + \sum_{i\in\mathcal{I}_3} p_i(\xi) \\
=& (\epsilon\gamma(\xi)+1)(|\mathcal{I}_2| + |\mathcal{I}_3| + 1)\gamma(\xi)T^\epsilon - \gamma(\xi)\sum_{i\in\mathcal{I}_3} x_i - \epsilon\gamma(\xi)\sum_{i\in\mathcal{I}_2\cup\mathcal{I}_3} p_i(\xi) - \sum_{i\in\mathcal{I}_2} p_i(\xi).
\end{aligned}$$

Then, we get

$$(\epsilon\gamma(\xi)(|\mathcal{I}_2|+|\mathcal{I}_3|+1)+|\mathcal{I}_2|+1)\,\gamma(\xi)T^\epsilon = \gamma(\xi)\sum_{i\in\mathcal{I}_3} x_i + \epsilon\gamma(\xi)\sum_{i\in\mathcal{I}_2\cup\mathcal{I}_3} p_i(\xi) + \sum_{i\in\mathcal{I}_2} p_i(\xi),$$

that is,

$$T^\epsilon = \frac{\gamma(\xi)\sum_{i\in\mathcal{I}_3} x_i + \epsilon\gamma(\xi)\sum_{i\in\mathcal{I}_2\cup\mathcal{I}_3} p_i(\xi) + \sum_{i\in\mathcal{I}_2} p_i(\xi)}{(\epsilon\gamma(\xi)(|\mathcal{I}_2|+|\mathcal{I}_3|+1)+|\mathcal{I}_2|+1)\,\gamma(\xi)}.$$

This completes the proof. □

Note that the above theorem gives the forms of the unique solution of the second stage regularized problem (4.2). However, it cannot be used to assist numerical calculation since the partition of the index set is not known in advance. Nevertheless, it is suffice for our purposes of deriving additional properties of the solutions. Due to the positive definiteness of M^ϵ and special structure of problem (4.2), we first obtain the following Lipschitz continuous property, following [23, Corollary 2.1].

Lemma 4.1. *There exists $L(\xi) > 0$ such that for any fixed $\epsilon \in (0, 1]$, we have*

$$\|z^\epsilon(q(x_1,\xi)) - z^\epsilon(q(x_2,\xi))\| \le L(\xi)\|x_1 - x_2\|, \quad \textit{for} \;\; x_1, x_2 \in \mathbb{R}^J_+ \text{ and } \xi \in \Xi.$$

Lemma 4.2. *For any fixed $\epsilon > 0$ and $(x, \xi) \in \mathbb{R}^J_+ \times \Xi$, T^ϵ has the following upper bound:*

$$T^\epsilon \le \|x\|_1 + \left(\epsilon + \frac{1}{\gamma(\xi)}\right)\|p(\xi)\|_1.$$

Proof. We have the following derivation from (4.6) that

$$\begin{aligned}
T^\epsilon =& \frac{\gamma(\xi)\sum_{i\in\mathcal{I}_3} x_i + \epsilon\gamma(\xi)\sum_{i\in\mathcal{I}_2\cup\mathcal{I}_3} p_i(\xi) + \sum_{i\in\mathcal{I}_2} p_i(\xi)}{(\epsilon\gamma(\xi)(|\mathcal{I}_2|+|\mathcal{I}_3|+1)+|\mathcal{I}_2|+1)\,\gamma(\xi)} \\
\le& \frac{\gamma(\xi)\sum_{i=1}^J x_i + (\epsilon\gamma(\xi)+1)\sum_{i=1}^J |p_i(\xi)|}{(\epsilon\gamma(\xi)(|\mathcal{I}_2|+|\mathcal{I}_3|+1)+|\mathcal{I}_2|+1)\,\gamma(\xi)} \\
\le& \frac{\gamma(\xi)\sum_{i=1}^J x_i + (\epsilon\gamma(\xi)+1)\sum_{i=1}^J |p_i(\xi)|}{\gamma(\xi)} \\
=& \|x\|_1 + \left(\epsilon + \frac{1}{\gamma(\xi)}\right)\|p_1(\xi)\|_1.
\end{aligned}$$

$\square$

We end this section by establishing the convergence result of the second stage LCP(q, M^ϵ) solutions as $\epsilon \downarrow 0$ for any given pair $(x, \xi) \in \mathbb{R}^J \times \Xi$.

Proposition 4.4. *For any fixed $\epsilon > 0$ and $(x, \xi) \in \mathbb{R}^J_+ \times \Xi$, let $z^\epsilon(\xi) = (y^\epsilon(\xi), \lambda^\epsilon(\xi))$ denote the unique solution of the regularized problem LCP(q, M^ϵ). Then*

$$\lim_{\epsilon \downarrow 0} \|z^\epsilon(\xi) - \bar{z}(\xi)\| = 0,$$

where $\bar{z}(\xi) = (\bar{y}(\xi), \bar{\lambda}(\xi))$ denotes the unique least ℓ_2-norm solution of the LCP(q, M). Moreover, the jth component of the least ℓ_2-norm solution of problem (2.12) has one of the following three forms:

$$\left\{(0,0), \left(-\frac{\gamma(\xi)\bar{T} - p_j(\xi)}{\gamma(\xi)}, 0\right), \left(x_j, -\gamma(\xi)(\bar{T} + x_j) + p_j(\xi)\right)\right\} \tag{4.13}$$

for $j \in \mathcal{J}$, where

$$\bar{T} := \lim_{\epsilon \downarrow 0} T^\epsilon = \sum_{i=1}^{J} \bar{y}_i.$$

Furthermore, for a.e. $\xi \in \Xi$ there exists $\bar{\kappa}(\xi) > 0$, such that

$$\|\lambda^\epsilon(\xi) - \bar{\lambda}(\xi)\| \le \bar{\kappa}(\xi)\epsilon. \tag{4.14}$$

Proof. Let $\hat{z} = (\hat{y}, \hat{\lambda})$ be any solution of LCP(q, M) and we have the derivation:

$$\begin{aligned} 0 &\ge (z^\epsilon - \hat{z})^T (M^\epsilon z^\epsilon + q - (M\hat{z} + q)) \\ &= (z^\epsilon - \hat{z})^T (M^\epsilon z^\epsilon - M\hat{z}) \\ &= (z^\epsilon - \hat{z})^T M (z^\epsilon - \hat{z}) + (z^\epsilon - \hat{z})^T \begin{pmatrix} 0 \\ \epsilon\lambda^\epsilon \end{pmatrix} \\ &\ge (z^\epsilon - \hat{z})^T \begin{pmatrix} 0 \\ \epsilon\lambda^\epsilon \end{pmatrix} \\ &= \epsilon(\lambda^\epsilon - \hat{\lambda})^T \lambda^\epsilon, \end{aligned}$$

where the second inequality follows from the positive semidefiniteness of M. Then, we have

$$\|\lambda^\epsilon\|^2 \leq \hat{\lambda}^T \lambda^\epsilon \leq \|\hat{\lambda}\| \|\lambda^\epsilon\|,$$

which implies the boundedness of λ^ϵ,

$$\|\lambda^\epsilon\| \leq \|\hat{\lambda}\|. \tag{4.15}$$

It follows from (4.15) that any accumulation point of $\{\lambda^\epsilon\}$ as $\epsilon \downarrow 0$ is the least ℓ_2-norm solution. Since M is positive semidefinite, we know from [15, Theorem 5.6.2] that there is a unique least ℓ_2-norm solution. On the other hand, we know from Proposition 4.1, for any fixed (x, ξ), $\hat{y}$ is unique. Therefore, the limit of z^ϵ exists as $\epsilon \downarrow 0$ and converges to the least ℓ_2-norm solution of LCP(q, M).

Due to the existence of limit for z^ϵ as $\epsilon \downarrow 0$, (4.13) can be derived directly from (4.5). In what follows, we focus on deriving the expression (4.14). To this end, for each $j \in \mathcal{J}$, three cases are discussed:

$$\gamma(\xi)T^\epsilon + (\lambda^\epsilon)_j - p_j(\xi) \geq 0, \ \ (y^\epsilon)_j - x_j \leq 0, \tag{4.16}$$

$$\gamma(\xi)T^\epsilon + (\lambda^\epsilon)_j - p_j(\xi) < 0, \ \ (y^\epsilon)_j - x_j \leq 0, \tag{4.17}$$

$$\gamma(\xi)T^\epsilon + (\lambda^\epsilon)_i - p_j(\xi) < 0, \ \ (y^\epsilon)_j - x_j > 0. \tag{4.18}$$

Case 1: If there exists a sequence $\{\epsilon_k\}_{k=1}^\infty$ converging to 0 such that (4.16) holds, we have

$$\lim_{k \to \infty} \big((y^{\epsilon_k})_j, (\lambda^{\epsilon_k})_j\big) = (0, 0).$$

Thus, $\big|(\lambda^{\epsilon_k})_j - \bar{\lambda}_j\big| = 0$.

Case 2: If there exists a sequence $\{\epsilon_k\}_{k=1}^\infty$ converging to 0 such that (4.17) holds, we have an estimation

$$
\begin{aligned}
\lim_{k\to\infty}\left((y^{\epsilon_k})_j,(\lambda^{\epsilon_k})_j\right) &= \lim_{k\to\infty}\left(-\frac{\gamma(\xi)T^{\epsilon_k}-p_j(\xi)}{\gamma(\xi)},0\right)\\
&= \left(-\frac{\gamma(\xi)\lim_{k\to\infty}T^{\epsilon_k}-p_j(\xi)}{\gamma(\xi)},0\right)\\
&= \left(-\frac{\gamma(\xi)\bar{T}-p_j(\xi)}{\gamma(\xi)},0\right).
\end{aligned}
$$

Thus, $\left|(\lambda^{\epsilon_k})_j-\bar{\lambda}_j\right|=0$.

Case 3: If there exists a sequence $\{\epsilon_k\}_{k=1}^{\infty}$ converging to 0 such that (4.18) holds, we have

$$
\begin{aligned}
\lim_{k\to\infty}\left((y^{\epsilon_k})_j,(\lambda^{\epsilon_k})_j\right) &= \lim_{k\to\infty}\left(-\frac{\epsilon_k\gamma(\xi)T^{\epsilon_k}-x_j-\epsilon_k p_j(\xi)}{\epsilon_k\gamma(\xi)+1},-\frac{\gamma(\xi)(T^{\epsilon_k}+x_j)-p_j(\xi)}{\epsilon_k\gamma(\xi)+1}\right)\\
&= \left(x_j,-\gamma(\xi)(\bar{T}+x_j)+p_j(\xi)\right).
\end{aligned}
$$

Thus, we have

$$
\begin{aligned}
&\left|(\lambda^{\epsilon_k})_j-\bar{\lambda}_j\right|\\
&=\left|(\lambda^{\epsilon_k})_j+\gamma(\xi)(\bar{T}+x_j)-p_j(\xi)\right|\\
&=\left|-\frac{\gamma(\xi)(T^{\epsilon_k}+x_j)-p_j(\xi)}{\epsilon_k\gamma(\xi)+1}+\gamma(\xi)(\bar{T}+x_j)-p_j(\xi)\right|\\
&=\frac{\left|-\gamma(\xi)(T^{\epsilon_k}+x_j)+p_j(\xi)+\gamma(\xi)(\bar{T}+x_j)-p_j(\xi)+\epsilon_k\gamma(\xi)(\gamma(\xi)(\bar{T}+x_j)-p_j(\xi))\right|}{\epsilon_k\gamma(\xi)+1}\\
&\le\frac{\gamma(\xi)\left|T^{\epsilon_k}-\bar{T}\right|+\left|\gamma(\xi)(\gamma(\xi)(\bar{T}+x_j)-p_j(\xi))\right|\epsilon_k}{\epsilon_k\gamma(\xi)+1}.
\end{aligned}
$$

Collectively, we know from **Case 1**, **Case 2** and **Case 3** that

$$(y^{\epsilon_k})_j-\bar{y}_j=0, \tag{4.19}$$

$$(y^{\epsilon_k})_j-\bar{y}_j=-(T^{\epsilon_k}-\bar{T}), \tag{4.20}$$

$$(y^{\epsilon_k})_j-\bar{y}_j=\frac{-\gamma(\xi)T^{\epsilon_k}+p_j(\xi)-\gamma(\xi)x_j}{\epsilon_k\gamma(\xi)+1}\cdot\epsilon_k. \tag{4.21}$$

Furthermore, we have that $T^{\epsilon_k} - \bar{T} \geq 0$ always holds. For the purpose of arriving at a contradiction, we assume $T^{\epsilon_k} - \bar{T} < 0$. Then (4.20) implies that

$$(y^{\epsilon_k})_j - \bar{y}_j > 0.$$

Moreover, (4.19) and (4.21) induce

$$\begin{aligned}
&(y^{\epsilon_k})_j - \bar{y}_j = 0, \\
&(y^{\epsilon_k})_j - \bar{y}_j \geq x_j - \bar{y}_j \geq 0,
\end{aligned}$$

respectively. Clearly, we have $T^{\epsilon_k} - \bar{T} \geq 0$, which contradicts our assumption. In addition, we have

$$T^{\epsilon_k} - \bar{T} \leq \frac{-\gamma(\xi) T^{\epsilon_k} + \|p(\xi)\|_1 + \gamma(\xi)\|x\|_1}{\epsilon_k \gamma(\xi) + 1} \cdot \epsilon_k \leq \left(\|p(\xi)\|_1 + \gamma(\xi)\|x\|_1\right)\epsilon_k.$$

Then, it follows that

$$\begin{aligned}
&\left|(\lambda^{\epsilon_k})_j - \bar{\lambda}_j\right| \\
\leq\ & \frac{\gamma(\xi)\left|T^{\epsilon_k} - \bar{T}\right| + \left|\gamma(\xi)(\gamma(\xi)(\bar{T} + x_j) - p_j(\xi))\right|\epsilon_k}{\epsilon_k \gamma(\xi) + 1} \\
\leq\ & \frac{\gamma(\xi)\left(\|p(\xi)\|_1 + \gamma(\xi)\|x\|_1\right) + \left|\gamma(\xi)(\gamma(\xi)(\bar{T} + x_j) - p_j(\xi))\right|}{\epsilon_k \gamma(\xi) + 1} \cdot \epsilon_k \\
\leq\ & \left(\gamma(\xi)\left(\|p(\xi)\|_1 + \gamma(\xi)\|x\|_1\right) + \gamma(\xi)^2\left(\|x\|_1 + \frac{\|p(\xi)\|_1}{\gamma(\xi)} + \|x\|_1\right) + \gamma(\xi)\|p(\xi)\|_1\right)\epsilon_k \\
\leq\ & 3\left(\gamma(\xi)^2\|x\|_1 + \gamma(\xi)\|p(\xi)\|_1\right)\epsilon_k,
\end{aligned}$$

where the third inequality follows Lemma 4.2 and the continuity of T^{ϵ} that

$$\bar{T} \leq \|x\|_1 + \frac{\|p(\xi)\|_1}{\gamma(\xi)}.$$

To summarize, for each $j \in \mathcal{J}$, we always have

$$\left|(\lambda^{\epsilon})_j - \bar{\lambda}_j\right| \leq 3\left(\gamma(\xi)^2\|x\|_1 + \gamma(\xi)\|p(\xi)\|_1\right)\epsilon.$$

Then, according to the definition of ℓ_2-norm, for any given $x \in \mathbb{R}^J_+$ one can compute

$$\bar{\kappa}(\xi) := 3\sqrt{J}\left(\gamma(\xi)^2 \|x\|_1 + \gamma(\xi)\|p(\xi)\|_1\right).$$

$\square$

4.2 Convergence analysis

In this section, we first prove the convergence of the unique solution of the regularized problem (4.3) to the solution set of the original problem as the regularized parameter ϵ decreases to zero. Next, we will study the SAA to solve the regularized problem, see [50]. Combined with our regularization approaches, we demonstrate the convergence property of the solution of our regularized SAA model as the number of samples goes to infinity. More specifically, the convergence analysis in this section is divided into two parts: the convergence analysis of the regularized problem as the regularized parameter ϵ tends to zero, and the analysis of regularized SAA. We finally build up the convergence relationship between the regularized SAA approach and the original problem.

4.2.1 Convergence of the regularized model

In this subsection, we only need to consider the convergence properties of the first stage decision vector, i.e., $x^\epsilon \in \mathbb{R}^J_+$ that solves problem (4.3), when the regularized parameter ϵ tends to zero. The convergence property of the solution $(x^\epsilon, y^\epsilon, \lambda^\epsilon)$ then follows by combining the result of section 3. From Proposition 4.3, we know that for fixed $\epsilon > 0$ problem (4.3) admits a unique first stage solution x^ϵ. In the following, we concern about the sequence of accumulation points of $\{x^\epsilon\}$ as $\epsilon \downarrow 0$.

For the existence of accumulation points, we have the following result.

Proposition 4.5. *Suppose there exists $p_0 > 0$ such that for all $j \in \mathcal{J}$, $p_j(\xi) \leq_{a.s.} p_0$. Then, with $\epsilon \downarrow 0$, $\{x^\epsilon\}$ is bounded.*

Proof. From the condition on $p_j(\xi)$, there must exist a sufficiently large $\alpha > 0$ such that for any $j \in \mathcal{J}$

$$\gamma_0 \alpha - p_j(\xi) > 0, \quad \text{for a.e. } \xi \in \Xi.$$

Then we have

$$-\frac{\gamma(\xi)(T^\epsilon + \alpha) - p_j(\xi)}{\epsilon\gamma(\xi) + 1} < 0, \quad \text{for a.e. } \xi \in \Xi.$$

Assume that $\{x^\epsilon\}$ is unbounded for the purpose of arriving at a contradiction. Then, it follows that there exist some indices $j \in \mathcal{J}$, such that $(x^\epsilon)_j \geq \alpha$. Then we consider the jth component of the first stage complementarity relation,

$$0 \leq (x^\epsilon)_j \perp c_j (x^\epsilon)_j - \mathbb{E}\big[\big(\lambda^\epsilon(\xi)\big)_j\big] + a_j \geq 0,$$

which can be expressed, from (4.5), as

$$0 \leq (x^\epsilon)_j \perp c_j (x^\epsilon)_j + a_j \geq 0.$$

However, this complementarity relation cannot be obtained because $(x^\epsilon)_j > 0$ and $c_j (x^\epsilon)_j + a_j > 0$. This completes our proof. □

Note that the conditions $p_j(\xi) \leq_{a.s.} p_0$ can be easily satisfied in many practical applications. For example, with given data sets of $p(\xi)$ we can always find an upper bound $p_0 := \max_j \{p_j(\xi)\}$.

Lemma 4.3. *Suppose there exists a constant $p_0 > 0$ such that for all $j \in \mathcal{J}$, $p_j(\xi) \leq_{a.s.} p_0$. Then, there exists a subsequence $\{\epsilon_k\}_{k=1}^\infty$ with $\epsilon_k \downarrow 0$ as $k \to \infty$, such that $x^{\epsilon_k} \to \hat{x}$ as $k \to \infty$ for some $\hat{x}$, and*

$$\lim_{k \to \infty} \mathbb{E}[\lambda^k(\xi)] = \mathbb{E}[\bar{\lambda}(\xi)],$$

where x^{ϵ_k} and $\lambda^k(\xi)$ is part of the unique solution of problem (4.3) for $x = x^{\epsilon_k}$ and $\xi \in \Xi$, and $\bar{\lambda}(\xi)$ is part of the least norm solution of the second stage problem (4.2) for $x = \hat{x}$ and $\xi \in \Xi$.

Proof. From the results of Proposition 4.5, there must exist an accumulation point of $\{x_\epsilon\}$ as $\epsilon \downarrow 0$, denoted by $\hat{x}$. Take a subsequence $\{\epsilon_k\}_{k=1}^{\infty}$ with $\epsilon_k \downarrow 0$ as $k \to \infty$, such that $x^{\epsilon_k} \to \hat{x}$ as $k \to \infty$. Denote by $\bar{\lambda}^k(\xi)$ the least norm solution of problem (2.12) for $x = x^{\epsilon_k}$ and $\xi \in \Xi$. We derive from Proposition 4.5 and (4.14) that

$$\begin{aligned}&\left\|\mathbb{E}[\lambda^k(\xi)] - \mathbb{E}[\bar{\lambda}(\xi)]\right\| \\ \leq &\left\|\mathbb{E}[\lambda^k(\xi)] - \mathbb{E}[\bar{\lambda}^k(\xi)]\right\| + \left\|\mathbb{E}[\bar{\lambda}^k(\xi)] - \mathbb{E}[\bar{\lambda}(\xi)]\right\| \\ \leq &\mathbb{E}[\bar{\kappa}(\xi)]\epsilon_k + \left\|\mathbb{E}[\bar{\lambda}^k(\xi)] - \mathbb{E}[\bar{\lambda}(\xi)]\right\|.\end{aligned}$$

Since Lemma 4.1 and (4.14), then for a.e. $\xi \in \Xi$

$$\left\|\bar{\lambda}^k(\xi) - \bar{\lambda}(\xi)\right\| \to 0 \text{ as } k \to \infty.$$

Furthermore, we derive that the estimation

$$\begin{aligned}\left\|\bar{\lambda}^k(\xi) - \bar{\lambda}(\xi)\right\| &\leq \left\|\bar{\lambda}^k(\xi)\right\| + \left\|\bar{\lambda}(\xi)\right\| \\ &\leq 4\sqrt{J}\big(\gamma(\xi)\left\|w\right\|_1 + \left\|p(\xi)\right\|_1\big), \quad \text{for a.e. } \xi \in \Xi,\end{aligned}$$

where the last term comes from Lemma 4.1 with some vector w with $\{x^{\epsilon_k}\}_{k=1}^{\infty}, \hat{x} \subseteq [0, w]$. It follows from the Lebesgue Dominated Convergence Theorem, we have

$$\left\|\mathbb{E}[\bar{\lambda}^k(\xi)] - \mathbb{E}[\bar{\lambda}(\xi)]\right\| \to 0 \text{ as } k \to \infty.$$

Then we complete all the proof. □

Theorem 4.2. *Any accumulation point of* $\{x^\epsilon, y^\epsilon, \lambda^\epsilon\}$ *as* $\epsilon \downarrow 0$ *is a solution of problem* (2.12).

Proof. We only need to verify that for any $\epsilon_k \downarrow 0$, the accumulation point $\hat{x}$ of subsequence $\{x^{\epsilon_k}\}$ is a first stage solution of (2.12). Since x^{ϵ_k} is the first stage solution of problem (4.3) for any $\epsilon_k > 0$, we have with $x^k = x^{\epsilon_k}$

$$0 \leq x^k \perp Cx^k - \mathbb{E}[\lambda^k(\xi)] + a \geq 0,$$

which, by using the 'min' NCP function (see, for example, [15]), can be rewritten as

$$\min\{x^k, Cx^k - \mathbb{E}[\lambda^k(\xi)] + a\} = 0.$$

By Lemma 4.3, we have

$$0 = \lim_{k\to\infty} \min\{x^k, Cx^k - \mathbb{E}\big[\lambda^k(\xi)\big] + a\} = \min\{\hat{x}, C\hat{x} - \mathbb{E}[\bar{\lambda}(\xi)] + a\}$$

as $k \to \infty$. Thus we obtain that

$$\min\{\hat{x}, C\hat{x} - \mathbb{E}[\bar{\lambda}(\xi)] + a\} = 0.$$

The statement then follows from Proposition 4.4. $\square$

4.2.2 Convergence of the regularized SAA model

In this subsection, we study the SAA scheme for solving the regularized problem (4.3) and focus on the convergence of the regularized SAA approach. More specifically, we focus on the SAA convergence analysis and the solution of the first stage problem. It is noteworthy that Chen, Sun and Xu considered a discrete approximation scheme in [11], which also leads to an approximation of the response variable in the second stage problem.

Let $\xi_1, \xi_2, \ldots, \xi_\nu$ denote ν independent identically distributed (i.i.d.) samples. Then, with slight abuse of notation, we can obtain the following formulation of problem (4.3) with SAA:

$$\begin{aligned}
&0 \le x \perp Cx - \frac{1}{\nu}\sum_{\ell=1}^{\nu} \lambda(\xi_\ell) + a \ge 0,\\
&0 \le \begin{pmatrix} y(\xi_\ell) \\ \lambda(\xi_\ell) \end{pmatrix} \perp \begin{pmatrix} \Pi(\xi_\ell) & I \\ -I & \epsilon I \end{pmatrix} \begin{pmatrix} y(\xi_\ell) \\ \lambda(\xi_\ell) \end{pmatrix} + \begin{pmatrix} -p(\xi_\ell) \\ x \end{pmatrix} \ge 0,\ \ell = 1, \ldots, \nu.
\end{aligned} \tag{4.22}$$

Or, we can write the problem collectively for all ν samples

$$0 \leq \begin{pmatrix} x \\ v_1 \\ \vdots \\ v_\nu \end{pmatrix} \perp \begin{pmatrix} C & \frac{1}{\nu}B & \dots & \frac{1}{\nu}B \\ -B^T & D_1^\epsilon & & \\ \vdots & & \ddots & \\ -B^T & & & D_\nu^\epsilon \end{pmatrix} \begin{pmatrix} x \\ v_1 \\ \vdots \\ v_\nu \end{pmatrix} + \begin{pmatrix} a \\ q_1 \\ \vdots \\ q_\nu \end{pmatrix} \geq 0, \tag{4.23}$$

where $C \in R^{J \times J}$, $B = \begin{pmatrix} 0 & -I \end{pmatrix} \in R^{J \times 2J}$, and for $\ell = 1, \dots, \nu$, $D_\ell^\epsilon = \begin{pmatrix} \Pi(\xi_\ell) & I \\ -I & \epsilon I \end{pmatrix} \in R^{2J \times 2J}$, $v_\ell = (y(\xi_\ell), \lambda(\xi_\ell))^T, q_\ell = (-p(\xi_\ell), 0)^T$. Thus, (4.23) is treated as a large-scale deterministic linear complementarity problem:

$$0 \leq z \perp H^\epsilon z + \bar{q} \geq 0, \tag{4.24}$$

where $z = (x, v_1, ..., v_\nu)^T$, $\bar{q} = (a, q_1, \cdots, q_\nu)^T$, and H^ϵ denotes the coefficient matrix in (4.23).

We have the following assertion of existence and uniqueness of problem (4.22) by [15, Theorem 3.1.6].

Proposition 4.6. *For any fixed $\epsilon > 0$ and positive integer ν, there exists a unique solution of problem* (4.22).

Recall the result of Lemma 4.1 and the following proposition can be shown in a similar way as in [50, Proposition 3.7].

Proposition 4.7. *Let $(y^\epsilon(\xi), \lambda^\epsilon(\xi))$ be the unique solution of the regularized second stage problem (4.3) for any $(x, \xi) \in \mathbb{R}_+^J \times \Xi$. Then,*

$$\frac{1}{\nu} \sum_{\ell=1}^{\nu} \lambda^\epsilon(\xi_\ell) \to \mathbb{E}[\lambda^\epsilon(\xi)]$$

with probability (w.p.) 1 as $\nu \to \infty$ uniformly on $\mathbb{B}(x, \delta) \cap \mathbb{R}_+^J$ for any $\delta > 0$,

Let x_ν^ϵ denote the first J-components of the unique solution of problem (4.22), and we have the following assertion.

Lemma 4.4. *Suppose there exists $p_0 > 0$ such that for all $j \in \mathcal{J}$, $p_j(\xi) \leq_{a.s.} p_0$. Then, with $\epsilon \downarrow 0$, $\{x_\nu^\epsilon\}$ is bounded.*

We omit the proof since it can be shown analogously as in Proposition 4.5.

Theorem 4.3. *Suppose there exists $p_0 > 0$ such that for all $j \in \mathcal{J}$, $p_j(\xi) \leq_{a.s.} p_0$. Then, for any fixed $\epsilon > 0$, $x_\nu^\epsilon \to x^\epsilon$ w.p. 1 as $\nu \to \infty$.*

Proof. From Propositions 4.3, Proposition 4.6, and Lemma 4.4, for any fixed $\epsilon > 0$, both the regularized problem (4.3) and its SAA-regularized problem (4.22) have solutions and contained in some compact subset in $\mathbb{R}_+^J$. We know from Proposition 4.7 that

$$\frac{1}{\nu}\sum_{\ell=1}^{\nu} \lambda^\epsilon(\xi_\ell) \to \mathbb{E}[\lambda^\epsilon(\xi)]$$

as $\nu \to \infty$, uniformly with respect to x on any compact set. Then, we have $x_\nu^\epsilon \to x^\epsilon$ w.p. 1 as $\nu \to \infty$ by [47, Proposition 19]. □

Combining Theorem 4.2 with Theorem 4.3, we have the following convergence result.

Theorem 4.4. *Suppose there exists $p_0 > 0$ such that for all $j \in \mathcal{J}$, $p_j(\xi) \leq_{a.s.} p_0$. Then,*

$$\limsup_{\epsilon \downarrow 0} \lim_{\nu \to \infty} x_\nu^\epsilon \subseteq S^*$$

w.p. 1, where S^* denotes the optimal solution set of the first stage problem of (2.12).

Proof. Let $\{\epsilon_k\}_{k=1}^\infty$ be a sequence with $\epsilon_k \downarrow 0$ as $k \to \infty$. Then, let $x^k = x^{\epsilon_k}$ denote the first stage of the solutions of (4.3) with $\epsilon = \epsilon_k$. Suppose $x^k \to \hat{x}$ as $k \to \infty$, i.e., $\hat{x}$ is the accumulation point. From Theorem 4.3,

$$\lim_{\nu \to \infty} x_\nu^k = x^k,$$

for any fixed ϵ^k, w.p.1. Thus it follows that

$$\lim_{k\to\infty}\left(\lim_{\nu\to\infty} x_\nu^k\right) = \lim_{k\to\infty} x^k = \hat{x}, \quad \text{w.p.1.}$$

Recall Theorem 4.2, we have $\hat{x} \in S^*$, and by the definition of outer limit,

$$\limsup_{\epsilon\downarrow 0} \lim_{\nu\to\infty} x_\nu^\epsilon \subseteq S^* \quad \text{w.p.1.}$$

Then we complete the proof. □

4.3 Numerical tests

In this section, we first give some details about the algorithm. After that, we carry out numerical experiments using randomly generated data to illustrate the effectiveness of our model and its solution approach. All the tests run in MATLAB 2016b on a personal computer with 32GB RAM and 8-core processor ($3.6 \times 8GHz$).

4.3.1 Progressive hedging method and smoothing Newton sub-algorithm

Recall that the model we interests is the form of a scenario-based linear complementarity problem (4.22) or its equivalent expression (4.23) with sufficiently small ϵ. The solution process adapts the well-known PHM. PHM was first proposed by Rockafellar and Wets [19] for solving multistage stochastic programming. Recently, it was extended to solving the multistage SVI in [13]. The PHM is globally convergent and the convergence rate is linear for SAA problem (4.23), see [13] for details.

Algorithm: PHM for (4.23)
Step 0. Given an initial point $x^0 \in \mathbb{R}^J$, let $x_i^0 = x^0 \in \mathbb{R}^J, v_i^0 \in \mathbb{R}^{2J}$ and $w_i^0 \in \mathbb{R}^J$, for $i = 1, \ldots, \nu$, such that $\frac{1}{\nu}\Sigma_{i=1}^{\nu} w_i^0 = 0$. Set the initial point $z_\nu^0 = (x^0, v_1^0, \ldots, v_\nu^0)^T$. Choose a step size $r > 0$. Set $k = 0$.
Step 1. If the point z_ν^k satisfies the condition

$$\| \min(z_\nu^k, M_\nu z_\nu^k + q_\nu)\| \leq 10^{-6},$$

output the solution z_ν^k and terminate the algorithm; otherwise, go to **Step 2**.
Step 2. For $i = 1, \ldots, \nu$, find $(\hat{x}_i^k, \hat{v}_i^k)$ that solves linear complementarity problems

$$\begin{aligned} &0 \leq x_i \perp C x_i + B v_i + a + w_i^k + r(x_i - x_i^k) \geq 0, \\ &0 \leq v_i \perp - B^T x_i + D^\epsilon(\xi_i) v_i + q_{2i}(\xi_i) + r(v_i - v_i^k) \geq 0. \end{aligned} \tag{4.25}$$

Then let $\bar{x}^{k+1} = \frac{1}{\nu}\sum_{i=1}^{\nu} \hat{x}_i^k$, and for $i = 1, \ldots, \nu$, update

$$x_i^{k+1} = \bar{x}^{k+1}, v_i^{k+1} = \hat{v}_i^k, w_i^{k+1} = w_i^k + r(\hat{x}_i^k - x_i^{k+1}),$$

to get point $z_\nu^{k+1} = (\bar{x}^{k+1}, v_1^{k+1}, \ldots, v_\nu^{k+1})^T$.
Step 3. Set $k := k + 1$; go back to **Step 1**.

The subproblem (4.25) is a deterministic linear complementarity problem. It is well-defined, since for all $i = 1, \ldots, \nu$, matrices $\begin{pmatrix} C + rI & B \\ -B^T & D^\epsilon(\xi_i) + rI \end{pmatrix} \in \mathbb{R}^{3J \times 3J}$ are positive definite for any $\epsilon > 0$. Thus, it has a unique solution.

Notice that the main computation cost of the PHM is in the **Step 2**. To improve the efficiency of the PHM, we use the warm-start technique suggested in [13] to choose an initial point for the PHM subproblem (4.25). More specifically, the solution z_ν^k of the subproblem (4.25) at the kth iteration is used as a starting point for the $(k+1)$th iteration.

We then focus on solving subproblem (4.25) of PHM. Denote the subproblem (4.25) by

$$0 \leq z_i \perp M_i^\epsilon z_i + q_i \geq 0, \tag{4.26}$$

where $z_i = \begin{pmatrix} x_i \\ v_i \end{pmatrix}, M_i^\epsilon = \begin{pmatrix} C + rI & B \\ -B^T & D^\epsilon(\xi_i) + rI \end{pmatrix}, q_i = \begin{pmatrix} a + w_i^k - r x_i^k \\ q_{2i}(\xi_i) - r v_i^k \end{pmatrix}, i =$

$1, 2, \ldots, \nu$. In order to take advantage of the spare structure of the subproblem (4.26), we apply the smoothing Newton method proposed by Chen and Ye [7] to solve the PHM subproblem (4.26). In what follows, we give a brief introduction to the smoothing Newton method.

It is well-known that solving (4.26) is equivalent to solving the nonsmooth equation

$$F(z_i) = \min(z_i, M_i^\epsilon z_i + q_i) = 0. \tag{4.27}$$

The main idea of the smoothing Newton method is to use a smooth approximation function to approximate the nonsmooth function F and then solve the corresponding linear system. We use the smooth Gariel-Moré approximation function $f : \mathbb{R}^{3J} \times \mathbb{R}_{++} \to \mathbb{R}^{3J}$ to approximate the nonsmooth function F. The jth component of f is defined as

$$f_j(z_i, \delta) = (z_i)_j - \int_{-\infty}^{\infty} \max\big(0, (z_i - M_i^\epsilon z_i - q_i)_j - \delta s\big)\rho(s)ds, \;\; j = 1, \ldots, 3J,$$

where $\rho(s)$ is a density function satisfying $\int_{-\infty}^{\infty} \rho(s)ds = 1$, $\delta > 0$ is a smoothing parameter. It is not difficult to see that $F_j(z_i) = \lim_{\delta \to 0} f_j(z_i, \delta), j = 1, 2, \ldots, 3J$. $f(z_i, \delta)$ is continuously differentiable with respect to z_i for any $\delta > 0$. For any $(z_i, \delta) \in \mathbb{R}^{3J} \times \mathbb{R}_{++}$, the Jacobian of the smoothing function $f(z_i, \delta)$ is

$$\nabla_{z_i} f(z_i, \delta) = I - \bar{D}(z_i)(I - M_i^\epsilon),$$

where $\bar{D}(z_i)$ is a diagonal matrix with diagonal elements

$$\bar{D}_{jj}(z_i) = \int_{-\infty}^{(z_i - M_i^\epsilon z_i - q_i)_j/\delta} \rho(s)ds, \quad j = 1, \ldots, 3J. \tag{4.28}$$

One key criterion in determining the performance of the smoothing method is the choice of density function $\rho(s)$, as have been explored in [7], and we use

$$\rho(s) = \frac{2}{(s^2 + 4)^{\frac{3}{2}}}$$

for our problem based on preliminary numerical tests. Then, the jth component of the smooth approximation function f reads

$$f_j(z_i,\delta) = (z_i)_j - \frac{1}{2}\left(\sqrt{(M_i^\epsilon z_i + q_i - z_i)_j^2 + 4\delta^2} + (z_i - M_i^\epsilon z_i - q_i)_j\right), \quad j = 1,\ldots,3J,$$

where the corresponding jth diagonal element of the Jacobian $\bar{D}(z_i)$ is

$$\bar{D}_{jj} = \frac{1}{2}\left(\frac{(z_i - M_i^\epsilon z_i - q_i)_j}{\sqrt{(z_i - M_i^\epsilon z_i - q_i)_j^2 + 4\delta^2}} + 1\right), \quad j = 1,\ldots,3J.$$

Then, the smoothing Newton method for solving subproblem (4.27) requires to solve a linear equation to determine d^k at each iteration, namely

$$\nabla_{z_i} f(z_i^k, \delta_k) d^k + F(z_i^k) = 0, \tag{4.29}$$

where δ_k decreases to 0 according to the criterion in [7]. To guarantee the well-defineness of the (4.29), we make use of the following result.

Theorem 4.5 ([51]). *For any diagonal matrix $\tilde{D} = diag(\tilde{D}_{jj}) \in \mathbb{R}^{J\times J}$ with $0 \leq \tilde{D}_{jj} \leq 1, j = 1, 2, \ldots, J$, the matrix $I - \tilde{D}(I - A)$ is nonsingular if and only if A is a $\mathbb{P}$-matrix.*

It is known that for any $i = 1,\ldots,\nu$, M_i^ϵ is positive definite and hence a $\mathbb{P}$-matrix. Moreover, combining the fact that $\int_{-\infty}^{\infty} \rho(s)ds = 1$ and (4.28), then the $\bar{D}(z_i)$ is a diagonal matrix with its element on the interval $[0, 1]$ for any $(z_i, \delta) \in \mathbb{R}^{3J} \times R_{++}$. Therefore, using Theorem 4.5, the Jacobian $\nabla_{z_i} f(z_i, \delta)$ is nonsingular for any $(z_i, \delta) \in \mathbb{R}^{3J} \times R_{++}$. Thus, the linear equation (4.29) is well-defined.

Denoting the matrix $\bar{D}(z_i) = diag(\bar{D}_1(z_i), \bar{D}_2(z_i), \bar{D}_3(z_i))$, the Jacobian $\nabla_{z_i} f(z_i, \delta)$

at the point z_i is of the following structure

$$\begin{aligned} & \nabla_{z_i} f(z_i, \delta) & (4.30) \\ = & \ (I - \bar{D}(z_i)) + \bar{D}(z_i) M_i^{\epsilon}) & \\ \triangleq & \begin{pmatrix} \Lambda_1(z_i) & 0 & -\bar{D}_1(z_i) \\ 0 & u_1(z_i)e^T + \Lambda_2(z_i) & \bar{D}_2(z_i) \\ \bar{D}_3(z_i) & -\bar{D}_3(z_i) & \Lambda_3(z_i) \end{pmatrix}, & (4.31) \end{aligned}$$

where $\Lambda_1(z_i) = \bar{D}_1(z_i)(C + (r-1)I) + I$, $\Lambda_2(z_i) = (\gamma(\xi) + (r-1))\bar{D}_2(z_i) + I$, $\Lambda_3(z_i) = (\epsilon + (r-1))\bar{D}_3(z_i) + I$ are all diagonal matrices, and $u_1(z_i) = \gamma(\xi)\bar{D}_2(z_i)e$. We can take advantage of the sparse structure of (4.31) in its inverse computation. Specifically, $\nabla_{z_i} f(z_i, \delta)$ consists of only diagonal sub-matrix and the matrix $u_1(z_i)e^T + \Lambda_2(z_i)$, where the later is a sum of a diagonal sub-matrix and a rank-one matrix.

Noticing from (4.31), the linear equation (4.29) is of the following form

$$\begin{pmatrix} \Lambda_1 & 0 & \Lambda_2 \\ 0 & u_1 u_2^T + \Lambda_3 & \Lambda_4 \\ \Lambda_5 & \Lambda_6 & \Lambda_7 \end{pmatrix} \begin{pmatrix} s_1 \\ s_2 \\ s_3 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}, \tag{4.32}$$

where $\Lambda_i \in \mathbb{R}^J, i = 1, 2, \ldots, 7$ are diagonal matrices with Λ_1 and Λ_3 being non-singular, $s_i \in \mathbb{R}^J, b_i \in \mathbb{R}^J, i = 1, 2, 3, u_1, u_2 \in \mathbb{R}^J$. For ease of notation, we use $\hat{\Lambda} = diag(1/\Lambda_{11}, \ldots, 1/\Lambda_{JJ})$ to represent the inverse of any invertible diagonal matrix $\Lambda = diag(\Lambda_{11}, \ldots, \Lambda_{JJ})$. Given the sparse structure of the coefficient matrix, we can solve the (4.32) efficiently. More exactly, by the first two equations of the (4.32), we can get

$$s_1 = \hat{\Lambda}_1(b_1 - \Lambda_2 s_3), \tag{4.33}$$

$$s_2 = (u_1 u_2^T + \Lambda_3)^{-1}(b_2 - \Lambda_4 s_3). \tag{4.34}$$

Directly substituting the (4.33) and (4.34) into the third equation of (4.32), we have

$$(\Lambda_7 - \Lambda_5 \hat{\Lambda}_1 \Lambda_2 - \Lambda_6 (u_1 u_2^T + \Lambda_3)^{-1} \Lambda_4) s_3 = b_3 + const, \tag{4.35}$$

where $const = -(\Lambda_5\hat{\Lambda}_1 b_1 + \Lambda_6(u_1u_2^T + \Lambda_3)^{-1}b_2)$. For computing the inverse matrix of $(u_1u_2^T + \Lambda_3)$, the Sherman-Morrison formula is useful. We have by the Sherman-Morrison formula

$$(u_1u_2^T + \Lambda_3)^{-1} = \hat{\Lambda}_3 - \frac{\hat{\Lambda}_3 u_1 u_2^T \hat{\Lambda}_3}{1 + u_2^T\hat{\Lambda}_3 u_1}. \tag{4.36}$$

Substituting the (4.36) into (4.35), we get

$$(\Lambda_0 + \alpha\tilde{u}_1\tilde{u}_2^T)s_3 = b_3 + const,$$

where $\alpha = 1/(1 + u_2^T\hat{\Lambda}_3 u_1), \Lambda_0 = \Lambda_7 - \Lambda_5\hat{\Lambda}_1\Lambda_2 - \Lambda_6\hat{\Lambda}_3\Lambda_4, \tilde{u}_1 = \Lambda_6\hat{\Lambda}_3 u_1, \tilde{u}_2 = \Lambda_4\hat{\Lambda}_3 u_2$. Then, if Λ_0 is nonsingular, using the Sherman-Morrison formula again, we can immediately get the solution of s_3

$$s_3 = \left(\hat{\Lambda}_0 - \frac{\alpha\hat{\Lambda}_0\tilde{u}_1\tilde{u}_2^T\hat{\Lambda}_0}{1 + \alpha\tilde{u}_2^T\hat{\Lambda}_0\tilde{u}_1}\right)(b_3 + const). \tag{4.37}$$

Then, substituting the s_3 into the (4.33) and (4.34), we get the solution of s_1 and s_2, respectively.

From (4.37), one can know that the computation cost of s_3 is trivial, since we only need to compute inverse of several diagonal matrix, namely, Λ_0, Λ_1, and Λ_3. Once s_3 is obtained, the calculation of s_1 and s_2 just needs to perform matrix-vector production. Therefore, the linear equation (4.32) can be solved efficiently.

4.3.2 Randomly generated problems

For the first part of numerical test, we randomly generated the problem of the form (4.23). More specifically, we generate a set of i.i.d. samples $\{\xi_\ell\}_{\ell=1}^{\nu}$ from a uniformly distribution over the interval $[0, 1]$. For $\ell = 1, \ldots, \nu$, set

$$p(\xi_\ell) = ((\xi_\ell)_1, (\xi_\ell)_2, \ldots, (\xi_\ell)_J)^T, \qquad \Pi(\xi_\ell) = \gamma(\xi_\ell)(ee^T + I) \triangleq (\xi_\ell)_1(ee^T + I).$$

Diagonal matrix $C \in \mathbb{R}^{J\times J}$ and $a \in \mathbb{R}^J$ are generated with its elements uniformly distributed over the interval $[1, 2]$. All the numerical results are based on the average of 10 independent runs.

To show the feasibility of the solution of the regularized problem compared to that of the original problem, we compute the following residual value

$$\mathbf{Res} = \| \min(z, Hz + \bar{q})\|,$$

where H denotes the coefficient matrix (4.23) with $\epsilon = 0$.

Table 4.1: Numerical results for different ϵ and sample size ν, $J = 10$ with individual sample

ν	$J(1+2\nu)$	Iter	CPU time/s	**Res**	Iter	CPU time/s	**Res**
			$\epsilon = 10^{-3}$			$\epsilon = 10^{-6}$	
10	210	146.30	0.26	4.42e-01	176.20	0.32	3.88e-04
50	1010	194.70	1.81	9.35e-01	197.40	1.83	9.32e-04
500	10010	208.70	26.72	3.00e+00	212.20	27.21	2.99e-03
2000	40010	222.60	154.97	5.93e+00	220.50	153.54	6.00e-03
5000	100010	224.70	623.53	9.49e+00	226.40	627.53	9.48e-03
			$\epsilon = 10^{-9}$			$\epsilon = 10^{-12}$	
10	210	152.70	0.27	1.08e-06	169.40	0.30	9.49e-07
50	1010	197.20	1.83	1.41e-06	194.40	1.80	9.75e-07
500	10010	212.70	27.21	3.21e-06	209.70	26.85	9.59e-07
2000	40010	220.30	153.34	6.16e-06	220.70	153.73	9.51e-07
5000	100010	226.70	628.89	9.60e-06	226.20	627.58	9.60e-07

Selected numerical results for $J = 10$ were listed in the Table 4.3.2. The average number of iterations, the average cpu time, and the average value of **Res** were recorded in this table. For the same value of ϵ, the number of iterations increases slightly when the sample size ν increases. In cases where the sample size ν is kept constant and the values of regularization parameter ϵ are chosen from $\epsilon = 10^{-3}$ to $\epsilon = 10^{-12}$, the iteration numbers are barely influenced as well as the cpu time. Furthermore, we observe the convergence of our regularization approach with decreasing values of ϵ, as have been proved in previous sections. Also notice that, the value of **Res** decreases when the ϵ diminishes from 10^{-3} to 10^{-12}. Numerically, it shows that the solution of the regularized problem is also that of the original problem when $\epsilon = 10^{-12}$.

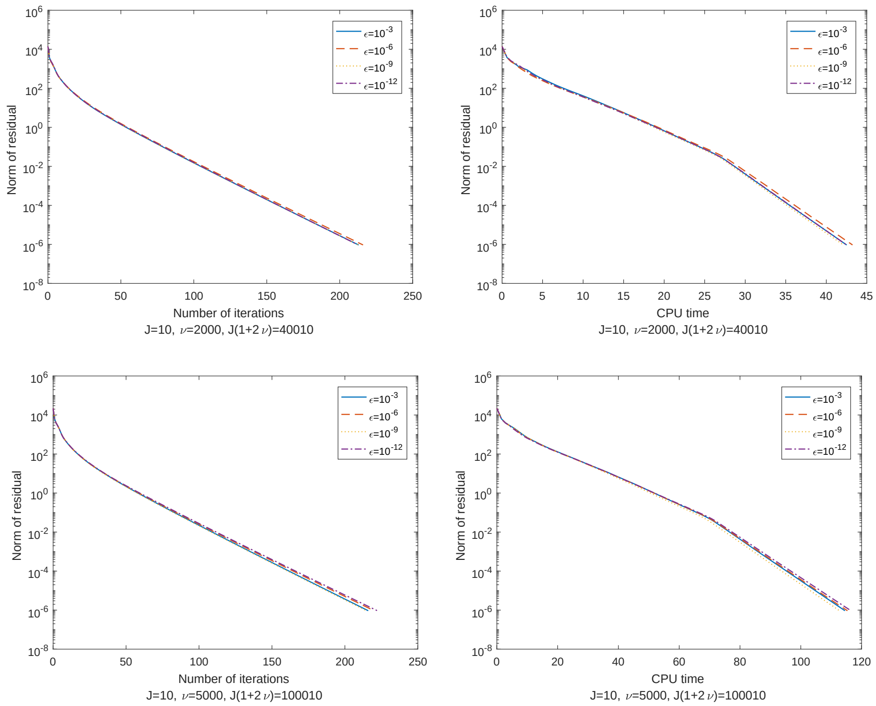

Figure 4.1: Numerical comparisons among different ϵ, $J = 100$.

Figure 4.1 illustrates the performance of the PHM measured by the number of iterations and cpu time for 10 players. It is also worth mentioning that although one might expect the problem to be more difficult to solve for a small ϵ, the numerical performance in our experiments remain roughly unaffected with decreasing values of ϵ. This is a good news from the viewpoint of approximation which requires a very small ϵ.

Figure 4.2 demonstrates the convergence property of the first stage solution x when the sample size gets large for the case $J = 10$. The convergence trend can be seen component-wisely as the solution x converges when the sample size ν gets large.

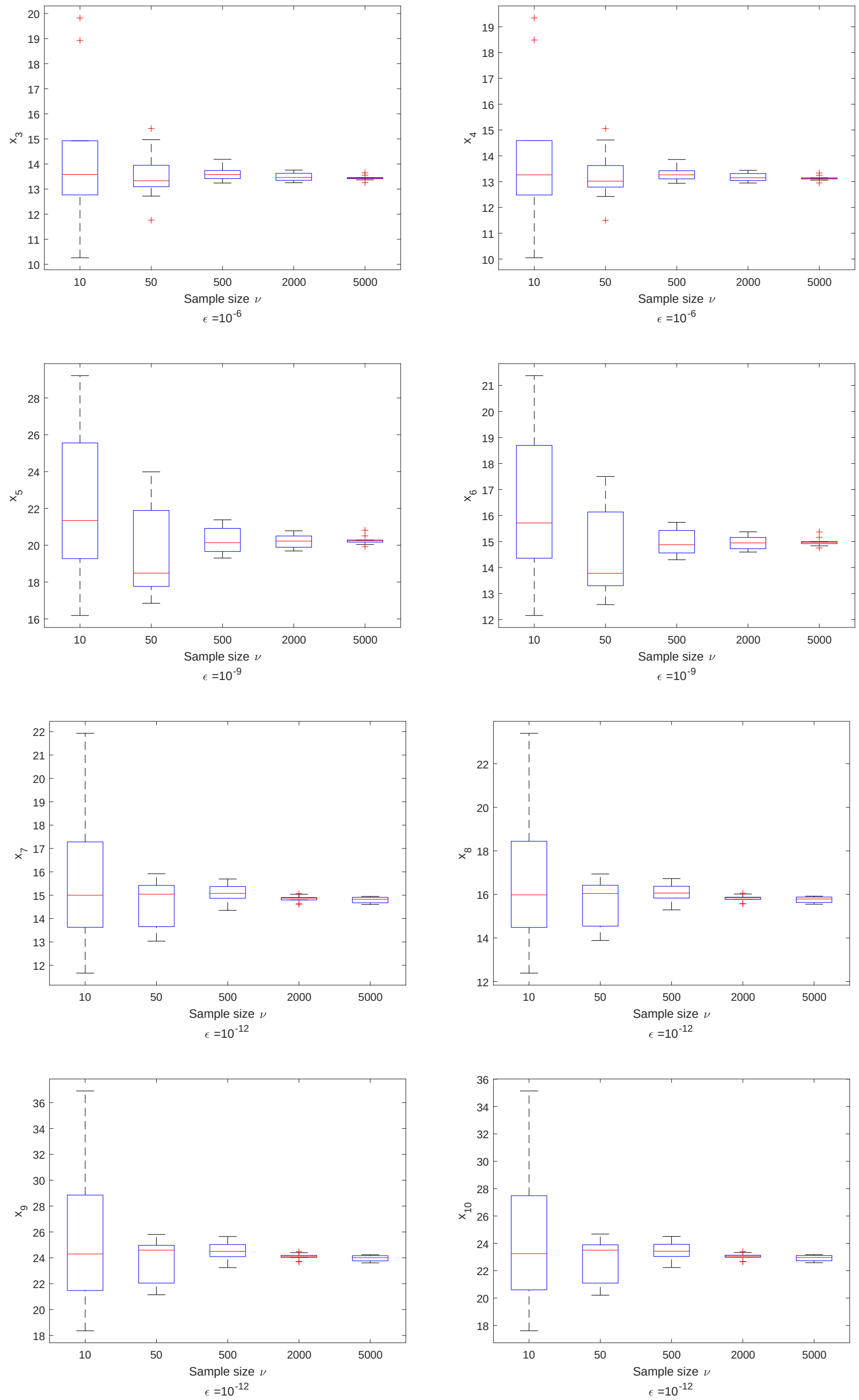

Figure 4.2: Convergence property of x with increasing ν, $J = 10$.

Chapter 5

Conclusions and future work

In this chapter, we conclude the contents of this thesis, and point out some possible work which we can do in the future.

5.1 Conclusions

The focus of the thesis has been placed on the two-stage SVI problem, applications and algorithms. Specifically, two research problems have been investigated in detail.

1. We study a class of two-stage stochastic linear variational inequality problems through the residual minimization problem (2.3). The quantitative stability and convergence analysis are conducted with respect to problem (2.3). Specifically, we first provide sufficient conditions for the existence of solutions of both the original problem and the perturbed problems. Next we conduct the quantitative stability analysis under the total variation metric, and further investigate the convergence of discrete approximations of the two-stage linear stochastic variational inequality problem. Finally, by a 3-player two-stage noncooperative game problem, we numerically illustrate our convergence conclusion and quantitative stability results.

2. A two-stage stochastic variational inequality is formulated as to describe the

equilibrium of a convex two-stage non-cooperative multi-agent game under uncertainty. Sufficient conditions for the existence of solutions of the variational inequality is provided under conventional assumptions. The numerical implementation is constructed by proposing a regularized sample average approximation and the solution concepts are given. Furthermore, we prove the convergence of the method as the regularization parameter tends to zero and the sample size tends to infinity. Numerical results are presented based on randomly generated data, which verify the effectiveness of our two-stage SVI approach.

5.2 Future work

Related topics for the future research work are listed below.

1. In our work, we assume that the distribution of the random vector is known. However, it is quite possible that we do not know the exact information of probability distribution. But we can construct a set of probability distributions that contains the true one, which leads to the so-called distributionally robust SVI problem. The numerical treatability of the distributionally robust SVI problem or the convergence between its data-driven approximated problem and the original distributionally robust SVI problem are interesting work to continue.

2. There are still few works on the quantitative stability analysis for multistage stochastic linear/nonlinear variational inequality problems. Moreover, although the PHM is designed to solve the multistage SVI problems, there exist few numerical results for the multistage case. The existing works mostly focus on the two-stage case. Therefore, these two topics would be promising in the continuing research.

Bibliography

[1] F. Facchinei, J.-S. Pang, Finite-dimensional Variational Inequalities and Complementarity Problems, Springer-Verlag, New York, 2003.

[2] J.-S. P. Richard W. Cottle, R. E. Stone, The Linear Complementarity Problem, Academic Press, Boston, 1992.

[3] G. Gürkan, A. Y. Özge, S. M. Robinson, Sample-path solution of stochastic variational inequalities, Math. Program. 84 (2) (1999) 313–333.

[4] F. Yousefian, A. Nedić, U. V. Shanbhag, On smoothing, regularization, and averaging in stochastic approximation methods for stochastic variational inequality problems, Math. Program. 165 (1) (2017) 391–431.

[5] J.-S. Pang, S. Sen, U. V. Shanbhag, Two-stage non-cooperative games with risk-averse players, Math. Program. 165 (1) (2017) 235–290.

[6] R. T. Rockafellar, R. J. Wets, Stochastic variational inequalities: single-stage to multistage, Math. Program. 165 (1) (2017) 331–360.

[7] X. Chen, Y. Ye, On homotopy-smoothing methods for box-constrained variational inequalities, SIAM J. Control Optim. 37 (2) (1999) 589–616.

[8] X. Chen, M. Fukushima, Expected residual minimization method for stochastic linear complementarity problems, Math. Oper. Res. 30 (4) (2005) 1022–1038.

[9] X. Chen, R. J. Wets, Y. Zhang, Stochastic variational inequalities: residual minimization smoothing sample average approximations, SIAM J. Optim. 22 (2) (2012) 649–673.

[10] X. Chen, T. K. Pong, R. J. Wets, Two-stage stochastic variational inequalities: an erm-solution procedure, Math. Program. 165 (1) (2017) 71–111.

[11] X. Chen, H. Sun, H. Xu, Discrete approximation of two-stage stochastic and distributionally robust linear complementarity problems, To appear in Math. Program.

[12] H. Jiang, M. Fukushima, L. Qi, D. Sun, A trust region method for solving generalized complementarity problems, SIAM J. Optim. 8 (1) (1998) 140–157.

[13] R. T. Rockafellar, J. Sun, Solving monotone stochastic variational inequalities and complementarity problems by progressive hedging, To appear in Math. Program.

[14] X. Chen, A. Shapiro, H. Sun, Convergence analysis of sample average approximation of two-state stochastic generalized equations, To appear in SIAM J. Optim.

[15] R. W. Cottle, J.-S. Pang, R. E. Stone, The Linear Complementarity Problem, SIAM, Philadelphia, 2009.

[16] Z.-Q. Luo, J.-S. Pang, D. Ralph, Mathematical Programs with Equilibrium Constraints, Cambridge University Press, Cambridge, 1996.

[17] H. Xu, Sample average approximation methods for a class of stochastic variational inequality problems, Asia-Pacific Journal of Operational Research 27 (01) (2010) 103–119.

[18] R. T. Rockafellar, R. J. Wets, Stochastic variational inequalities: single-stage to multistage, Math. Program. 165 (1) (2016) 331–360.

[19] R. T. Rockafellar, R. J. Wets, Scenarios and policy aggregation in optimization under uncertainty, Math. Oper. Res. 16 (1) (1991) 119–147.

[20] X. Chen, Smoothing methods for nonsmooth, nonconvex minimization, Math. Program. 134 (1) (2012) 71–99.

[21] X. Chen, Y. Ye, On smoothing methods for the p_0 matrix linear complementarity problem, SIAM J. Optim. 11 (2) (2000) 341–363.

[22] X. Chen, M. Fukushima, A smoothing method for a mathematical program with p-matrix linear complementarity constraints, Comput. Optim. Appl. 27 (3) (2004) 223–246.

[23] X. Chen, S. Xiang, Newton iterations in implicit time-stepping scheme for differential linear complementarity systems, Math. Program. 138 (1-2) (2013) 579–606.

[24] H. Robbins, S. Monro, A stochastic approximation method, Annals of Mathematical Statistics 22 (3) (1951) 400–407.

[25] H. Jiang, H. Xu, Stochastic approximation approaches to the stochastic variational inequality problem, IEEE Transactions on Automatic Control 53 (6) (2008) 1462–1475.

[26] J. Koshal, A. Nedic, U. V. Shanbhag, Regularized iterative stochastic approximation methods for stochastic variational inequality problems, IEEE Transactions on Automatic Control 58 (3) (2013) 594–609.

[27] X. Chen, Y. Ye, On homotopy-smoothing methods for box-constrained variational inequalities, SIAM J. Control Optim. 37 (2) (1999) 589–616.

[28] J. Lei, U. V. Shanbhag, J.-S. Pang, S. Sen, On synchronous, asynchronous, and randomized best-response schemes for computing equilibria in stochastic nash games, To appear in Math. Oper. Res.

[29] S. T. Rachev, L. Klebanov, S. V. Stoyanov, F. Fabozzi, The methods of distances in the theory of probability and statistics, Springer, New York, 2013.

[30] W. Römisch, Stability of stochastic programming problems, Handb. Oper. Res. Manag. Sci. 10 (2003) 483–554.

[31] J. Zhang, H. Xu, L. Zhang, Quantitative stability analysis for distributionally robust optimization with moment constraints, SIAM J. Optim. 26 (3) (2016) 1855–1882.

[32] F. H. Murphy, H. D. Sherali, A. L. Soyster, A mathematical programming approach for determining oligopolistic market equilibrium, Math. Program. 24 (1) (1982) 92–106.

[33] R. Rockafellar, R. Wets, Stochastic convex programming: basic duality, Pacific Journal of Mathematics 62 (1) (1976) 173–195.

[34] R. Rockafellar, R.-B. Wets, Stochastic convex programming: Kuhn-tucker conditions, Journal of Mathematical Economics 2 (3) (1975) 349–370.

[35] R. Rockafellar, R. Wets, Stochastic convex programming: singular multipliers and extended duality singular multipliers and duality, Pacific Journal of Mathematics 62 (2) (1976) 507–522.

[36] R. T. Rockafellar, R. J. Wets, Stochastic convex programming: relatively complete recourse and induced feasibility, SIAM J. Control Optim. 14 (3) (1976) 574–589.

[37] A. Shapiro, D. Dentcheva, A. Ruszczyński, Lectures on Stochastic Programming: Modeling and Theory, SIAM, Philadelphia, 2014.

[38] J. Huang, M. Leng, M. Parlar, Demand functions in decision modeling: A comprehensive survey and research directions, Decision Sciences 44 (3) (2013) 557–609.

[39] H. Sun, C.-L. Su, X. Chen, Saa-regularized methods for multiproduct price optimization under the pure characteristics demand model, Math. Program. 165 (1) (2017) 361–389.

[40] J. V. Burke, X. Chen, H. Sun, Subdifferentiation and smoothing of nonsmooth integral functionals, `https://www.polyu.edu.hk/ama/staff/xjchen/paperSIAM.pdf` (2017).

[41] A. J. King, R. T. Rockafellar, Sensitivity analysis for nonsmooth generalized equations, Math. Program. 55 (1) (1992) 193–212.

[42] U. Ravat, U. V. Shanbhag, On the characterization of solution sets of smooth and nonsmooth convex stochastic nash games, SIAM J. Optim. 21 (3) (2011) 1168–1199.

[43] U. Ravat, U. V. Shanbhag, On the existence of solutions to stochastic quasi-variational inequality and complementarity problems, Math. Program. 165 (1) (2017) 291–330.

[44] S. M. Robinson, Generalized equations and their solutions, part i: Basic theory, in: Point-to-Set Maps and Mathematical Programming, Springer, 1979, pp. 128–141.

[45] S. M. Robinson, Strongly regular generalized equations, Math. Oper. Res. 5 (1) (1980) 43–62.

[46] R. T. Rockafellar, R. J. Wets, Variational Analysis, Springer-Verlag, Berlin, 2009.

[47] A. Shapiro, Monte carlo sampling methods, Handb. Oper. Res. Manag. Sci. 10 (2003) 353–425.

[48] A. Shapiro, H. Xu, Stochastic mathematical programs with equilibrium constraints, modelling and sample average approximation, Optimization 57 (3) (2008) 395–418.

[49] J. B. Rosen, Existence and uniqueness of equilibrium points for concave n-person games, Econometrica: Journal of the Econometric Society (1965) 520–534.

[50] X. Chen, H. Sun, R. J.-B. Wets, Regularized mathematical programs with stochastic equilibrium constraints: estimating structural demand models, SIAM J. Optim. 25 (1) (2015) 53–75.

[51] S. A. Gabriel, J. J. Moré, Smoothing of mixed complementarity problems, Complementarity and Variational Problems: State of the Art (1997) 105–116.