

This version of the article has been accepted for publication, after peer review (when applicable) and is subject to Springer Nature's AM terms of use (<https://www.springernature.com/gp/open-research/policies/accepted-manuscript-terms>), but is not the Version of Record and does not reflect post-acceptance improvements, or any corrections. The Version of Record is available online at: <http://dx.doi.org/10.1007/s10107-018-1266-4>.

# Discrete Approximation of Two-Stage Stochastic and Distributionally Robust Linear Complementarity Problems

Xiaojun Chen\*    Hailin Sun<sup>†</sup>    Huifu Xu<sup>‡</sup>

March 21, 2018

**Abstract.** In this paper, we propose a discretization scheme for the two-stage stochastic linear complementarity problem (LCP) where the underlying random data are continuously distributed. Under some moderate conditions, we derive qualitative and quantitative convergence for the solutions obtained from solving the discretized two-stage stochastic LCP (SLCP). We explain how the discretized two-stage SLCP may be solved by the well-known progressive hedging method (PHM). Moreover, we extend the discussion by considering a two-stage distributionally robust LCP (DRLCP) with moment constraints and proposing a discretization scheme for the DRLCP. As an application, we show how the SLCP and DRLCP models can be used to study equilibrium arising from two-stage duopoly game where each player plans to set up its optimal capacity at present with anticipated competition for production in future.

**Key Words.** Two-stage stochastic linear complementarity problem, discrete approximation, error bound, distributionally robust linear complementarity problem, ex post equilibrium

## 1 Introduction

Let  $\xi : \Omega \rightarrow \mathbb{R}^l$  be a random variable defined in the probability space  $(\Omega, \mathcal{F}, P)$  with support set  $\Xi \subset \mathbb{R}^l$  and  $\mathcal{Y}$  be the space of measurable functions defined on  $\Xi$ . We consider the following two-stage stochastic linear complementarity problem: find  $(x, y(\cdot)) \in \mathbb{R}^n \times \mathcal{Y}$  which solves

$$(SLCP) \quad \begin{cases} 0 \leq x \perp Ax + \mathbb{E}[B(\xi)y(\xi)] + q_1 \geq 0, \\ 0 \leq y(\xi) \perp M(\xi)y(\xi) + N(\xi)x + q_2(\xi) \geq 0, \quad \text{for a.e. } \xi \in \Xi, \end{cases} \quad (1.1)$$

where  $A \in \mathbb{R}^{n \times n}$ ,  $q_1 \in \mathbb{R}^n$ ,  $B(\cdot) : \mathbb{R}^l \rightarrow \mathbb{R}^{n \times m}$ ,  $M(\cdot) : \mathbb{R}^l \rightarrow \mathbb{R}^{m \times m}$ ,  $N(\cdot) : \mathbb{R}^l \rightarrow \mathbb{R}^{m \times n}$ , and  $q_2(\cdot) : \mathbb{R}^l \rightarrow \mathbb{R}^m$  are continuous matrix valued mappings,  $\mathbb{R}^n \ni a \perp b \in \mathbb{R}^n$  means  $a^T b = 0$ ,

---

\*Department of Applied Mathematics, The Hong Kong Polytechnic University, Hong Kong, (maxjchen@polyu.edu.hk). This author's work is supported in part by Hong Kong Research Grant Council PolyU153016/16p

<sup>†</sup>School of Economics and Management, Nanjing University of Science and Technology, Nanjing, 210094, China, (hlsun@njust.edu.cn). This author's work is supported in part by the National Natural Science Foundation of China (Grant No. 11401308)

<sup>‡</sup>School of Mathematics, University of Southampton, Southampton, SO17 1BJ, UK, (h.xu@soton.ac.uk). The work of this author is supported in part by EPSRC Grant EP/M003191/1

the mathematical expectation is taken componentwise w.r.t. probability distribution of  $\xi$ , and abbreviation ‘‘a.e.’’ stands for almost everywhere. In the case when  $\xi$  follows a finite discrete distribution, the problem above reduces to a deterministic complementarity problem which has been extensively investigated over the past few decades, see monographs [12, 15]. Note that if we consider  $(\Xi, \mathcal{B})$  as a measurable space equipped with Borel sigma algebra  $\mathcal{B}$ , then  $P$  may be viewed as probability measure defined on  $(\Xi, \mathcal{B})$  induced by the random variable  $\xi$  and consequently  $P$  is called the probability distribution of  $\xi$ . Throughout the paper, we use terminology probability measure and probability distribution interchangeably.

In the first stage of SLCP (1.1), one is supposed to find  $x \in \mathbb{R}^n$  here and now before the random data  $\xi(\omega)$  is available. At the second stage when  $x$  is fixed and a realization of  $\xi(\omega) = \xi$  becomes known,  $y \in \mathbb{R}^m$  is sought to satisfy the second stage of SLCP (1.1). The two-stage model arises naturally from first order optimality conditions of a two-stage stochastic linear program with recourse (see [4, 32]). It can also be used to characterize equilibrium arising from a two-stage stochastic game where at the first stage players compete for capacity expansion before realization of uncertainty and at the second stage they bid for producing goods or services after production capacity is developed and uncertainty (i.e. market demand) is observed.

The two-stage SLCP (1.1) may be regarded as a special case of two-stage stochastic nonlinear complementarity problem (NCP) and two-stage stochastic variational inequalities (SVI) considered by Chen, Pong and Wets [6] and multi-stage SVI developed by Rockafellar and Wets [28], which synthesize and extend the well investigated Expected Residual Minimization (ERM) models [5, 7] and expected value models [16, 19] for one-stage SVI. In a more recent development, Rockafellar and Sun [29] apply the well-known PHM to solve multi-stage SVI and stochastic NCP (SNCP). However, the application is restricted to the case where the distribution of the underlying uncertainty is discrete and finite.

In this paper, we discuss discrete approximation of the two-stage SLCP which is partly aimed to fill out the gap on application of PHM to the continuously distributed SLCP. Moreover, a discretized two-stage SLCP is virtually a deterministic SLCP which makes it readily solvable by existing code such as PATH. We are focusing on a two-stage SLCP so that we may concentrate on the key ideas in our approach and leave the potential extension to multi-stage and/or non-linear case for future research.

Extending from the two-stage SLCP (1.1), we consider a situation where the true probability distribution  $P$  is unknown but it is possible to use partial information to construct an ambiguity set  $\mathcal{P}$  of distributions which contains the true distribution. Consequently, we propose a two-stage distributionally robust LCP as follows:

$$(\text{DRLCP}) \quad \begin{cases} 0 \leq x \perp Ax + \mathbb{E}_P[B(\xi)y(\xi)] + q_1 \geq 0, & \forall P \in \mathcal{P}, \\ 0 \leq y(\xi) \perp M(\xi)y(\xi) + N(\xi)x + q_2(\xi) \geq 0, & \text{for } P\text{-a.e. } \xi \in \Xi. \end{cases} \quad (1.2)$$

The DRLCP requires every solution to satisfy the first stage complementarity condition for all  $P \in \mathcal{P}$  to mitigate the risk arising from ambiguity of the true distribution. Obviously the new model is more demanding on its solution than the two-stage SLCP (1.1) and as a result, it might not have a solution if the ambiguity set is too large. Like the two-stage SLCP (1.1) which stems from two-stage stochastic linear programming or two-stage stochastic games with continuous actions, the two-stage DRLCP (1.2) can be linked to distributionally robust optimization

and games. Indeed, the two-stage DRLCP (1.2) may be used to characterize the first order optimality conditions of the so-called ex post optimal solution to a two-stage distributionally robust optimization problem and the ex post equilibrium of two-stage distributionally robust games. The notion of ex post equilibrium is a well-known distribution-free solution concept and is especially prevalent in the auction theory literature, see [2, 13, 18].

The paper has three main contributions.

- (i) We provide sufficient conditions for the existence and uniqueness of the solution of the two-stage SLCP (Proposition 2.1) and propose a discrete approximation scheme for the problem. We then present some qualitative and quantitative convergence analysis of the solutions obtained from solving the discretized two-stage SLCP to their true counterparts (Theorems 3.1 and 3.2). Application of PHM to the discretized problem is outlined (Section 3.3).
- (ii) In the absence of complete information on the true probability distribution, we propose a distributionally robust model for the two-stage SLCP, the model is parallel to the ex post equilibrium of robust games studied by Aghassi and Bertsimas [2]. We derive a dual formulation of the DRLCP model and discuss a discretization approach for the latter (Section 4).
- (iii) As an application as well as a motivation, we propose a two-stage distributionally robust game in a duopoly market where two players need to make strategic decisions on capacity for future production with anticipation of Nash-Cournot type competition after demand uncertainty in future is observed. Under some standard conditions, we reformulate the problem as a two-stage SLCP when the true distribution of the demand uncertainty is known and a two-stage DRLCP otherwise. We give an academic example to show existence of a solution to the two-stage DRLCP (Section 5.1).

Throughout the paper, we use the following notation.  $\mathbb{R}_+^n$  denotes the non-negative orthant and  $\mathbb{R}_{++}^n$  the interior of  $\mathbb{R}_+^n$ . For a vector  $a \in \mathbb{R}^n$ , we write  $(a)_+$  for  $\max(0, a)$ , where the maximum is taken componentwise. For matrices  $A, B \in \mathbb{R}^{n \times n}$ , we write  $A \succeq B$  and  $A \succ B$  to indicate  $A - B$  being positive semidefinite and positive definite respectively. Differing from the convention in semi-definite programming, here  $A$  and  $B$  are not necessarily symmetric. We use  $\|\cdot\|$  to denote the 2-norm for both vectors and matrices and indicate any other norms by a subscript such as  $\|\cdot\|_\infty$  for the infinity norm. Finally, to ease the exposition, we write  $\bar{K} := \{1, \dots, K\}$  and use the same notation  $\xi$  for both a random vector and a deterministic vector of variables in  $\mathbb{R}^l$  depending on the context.

## 2 Structure of the two-stage SLCP

Although our main emphasis, later on in this paper, will rest on the case where the second stage of SLCP (1.1) has a unique solution, we believe that it will be helpful to discuss a precise meaning of the model in a general setting where the second stage of SLCP (1.1) has multiple solutions for each fixed  $x$  and  $\xi$ . Let  $\mathcal{S}(x, \xi)$  denote the set of solutions of the second stage of

SLCP (1.1). Then the two-stage SLCP can be written as: find  $x \in \mathbb{R}^n$  and  $y(\cdot) \in \mathcal{S}(x, \xi(\cdot))$  which solve

$$0 \leq x \perp Ax + \mathbb{E}[B(\xi)y(\xi)] + q_1 \geq 0. \quad (2.1)$$

Here  $\mathcal{S}(x, \xi(\cdot)) : \Omega \rightarrow \mathcal{Y}$  is a random set-valued mapping, and  $y(\cdot)$  is a measurable selection such that  $\mathbb{E}[B(\xi)y(\xi)]$  is finite-valued. In the case when  $\mathcal{S}(x, \xi)$  is a singleton for each  $x \in \mathbb{R}^n$  and  $\xi \in \Xi$ ,  $\mathcal{S}(x, \xi) = \{\bar{y}(x, \xi)\}$ .

In what follows, we investigate conditions for the existence and uniqueness of a solution to (1.1). For this purpose, we make the following technical assumptions.

**Assumption 1** *There exists a positive continuous function  $\kappa(\xi)$  such that  $\mathbb{E}[\kappa(\xi)] < +\infty$  and for almost every  $\xi$ ,*

$$\begin{pmatrix} z^T, u^T \end{pmatrix} \begin{pmatrix} A & B(\xi) \\ N(\xi) & M(\xi) \end{pmatrix} \begin{pmatrix} z \\ u \end{pmatrix} \geq \kappa(\xi)(\|z\|^2 + \|u\|^2), \quad \forall z \in \mathbb{R}^n, u \in \mathbb{R}^m. \quad (2.2)$$

Moreover,  $\mathbb{E}[\|B(\xi)\|] < \infty$ ,  $\mathbb{E}[\|M(\xi)\|] < \infty$ ,  $\mathbb{E}[\|N(\xi)\|] < \infty$  and  $\mathbb{E}[\|q(\xi)\|] < \infty$ .

Let  $\mathcal{D}$  denote the set of  $m \times m$  diagonal matrices  $D$  with diagonal components  $D_{jj} \in \{0, 1\}$ , for  $j \in \bar{m}$ , and  $M$  be an  $m \times m$  positive definite matrix. Let  $\mathcal{J}$  denote the power set of  $\{1, \dots, n\}$  and  $J \in \mathcal{J}$ . Let  $D_J \in \mathcal{D}$  with

$$(D_J)_{jj} := \begin{cases} 1, & \text{if } j \in J, \\ 0, & \text{otherwise.} \end{cases}$$

It is known that  $I - D_J(I - M)$  is invertible when  $M \succ 0$ , see for instance [10]. Let

$$U_J(M) := (I - D_J(I - M))^{-1}D_J.$$

By permutation if necessary, we assume for the simplicity of exposition that  $J = \{1, 2, \dots, |J|\}$ , where  $|J|$  denotes the cardinality of set  $J$ . Consequently, we know from [10] that

$$U_J(M) = \begin{cases} 0_{n \times n}, & \text{if } J = \emptyset, \\ \begin{pmatrix} M_J^{-1} & 0 \\ 0 & 0 \end{pmatrix}, & \text{otherwise,} \end{cases}$$

where  $M_J$  is the  $|J| \times |J|$  sub-matrix of  $M$  whose entries of  $M$  are indexed by the set  $J \in \mathcal{J}$ .

From time to time in the follow-up discussions, we need to look into positive definiteness of  $A - B(\xi)U_J(M(\xi))N(\xi)$  and its inverse. To this end, we state the following intermediate technical result.

**Lemma 2.1** *Under Assumption 1, the following assertions hold.*

- (i)  $z^T Az \geq \text{ess sup}_{\xi \in \Xi} \kappa(\xi) \|z\|^2$  for all  $z \in \mathbb{R}^n$ , and  $u_J^T M_J(\xi) u_J \geq \kappa(\xi) \|u_J\|^2$ , for a.e.  $\xi \in \Xi$ , all  $u_J \in \mathbb{R}^{|J|}$  and  $J \in \mathcal{J}$ .

(ii)  $A - B(\xi)U_J(M(\xi))N(\xi)$  is well defined and  $z^T(A - B(\xi)U_J(M(\xi))N(\xi))z \geq \kappa(\xi)\|z\|^2$ , for a.e.  $\xi \in \Xi$  and all  $z \in \mathbb{R}^n$ .

(iii)  $\|M_J(\xi)^{-1}\| \leq \frac{1}{\kappa(\xi)}$  and  $\|(A - B(\xi)U_J(M(\xi))N(\xi))^{-1}\| \leq \frac{1}{\kappa(\xi)}$ . for a.e.  $\xi \in \Xi$ .

**Proof.** We only prove Part (ii) since Part (i) follows straightforwardly from (2.2) and Part (iii) follows from Parts (i) and (ii). By setting  $u = -U_J(M(\xi))N(\xi)z$  in (2.2), and using  $U_J(M(\xi))M(\xi)U_J(M(\xi)) = U_J(M(\xi))$ , we have

$$z^T Az - z^T B(\xi)U_J(M(\xi))N(\xi)z \geq \kappa(\xi)(\|z\|^2 + \|U_J(M(\xi))N(\xi)z\|^2) \geq \kappa(\xi)\|z\|^2$$

for any  $z \in \mathbb{R}^n$  and a.e.  $\xi \in \Xi$ . ■

We are now ready to state existence of solutions to SLCP (1.1) and the structure of the second stage solution.

**Proposition 2.1** *Let Assumption 1 hold. For any given  $x$  and  $\xi \in \Xi$ , let  $D(x, \xi) \in \mathcal{D}$  be an  $m$ -dimensional diagonal matrix with*

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

Let

$$W(x, \xi) := [I - D(x, \xi)(I - M(\xi))]^{-1}D(x, \xi) \quad (2.3)$$

and

$$J(x, \xi) := \{j : (M(\xi)y(\xi) + N(\xi)x + q_2(\xi))_j \leq y_j(\xi)\}.$$

Then the following assertions hold.

(i) The two-stage SLCP (1.1) has a unique solution  $(x^*, y^*(\cdot)) \in \mathbb{R}^n \times \mathcal{Y}$ .

(ii) The solution to the second stage of SLCP (1.1) can be written as

$$\bar{y}(x, \xi) = -W(x, \xi)(N(\xi)x + q_2(\xi)) \quad (2.4)$$

and  $\bar{y}(\cdot, \xi)$  is globally Lipschitz continuous w.r.t  $x$ .

(iii) The first equation of SLCP (1.1) can be reformulated as

$$0 \leq x \perp (A - \mathbb{E}[B(\xi)W(x, \xi)N(\xi)])x - \mathbb{E}[B(\xi)W(x, \xi)q_2(\xi)] + q_1 \geq 0, \quad (2.5)$$

where

$$\|(A - \mathbb{E}[B(\xi)W(x, \xi)N(\xi)])^{-1}\| \leq \frac{1}{\mathbb{E}[\kappa(\xi)]} < +\infty.$$

(iv) Let

$$F(x) := \min(x, (A - \mathbb{E}[B(\xi)W(x, \xi)N(\xi)])x - \mathbb{E}[B(\xi)W(x, \xi)q_2(\xi)] + q_1). \quad (2.6)$$

Then  $F$  is Lipschitz continuous and every matrix  $V_x$  in the Clarke generalized Jacobian  $\partial F(x)$  (see definition in [11, Section 2.6]) is nonsingular with  $\|V_x^{-1}\| \leq \bar{d}$  for some constant  $\bar{d} > 0$  which is independent of  $x$ .

**Proof.** We only prove Part (i) and Part (iv) as the other parts follow straightforwardly from [10], Lemma 2.1 and the implicit function theorem [38, Lemma 2.2].

Part (i). For this purpose, we prove monotonicity of the infinite complementarity system (1.1). Let  $\langle \cdot, \cdot \rangle$  denote the scalar product in the Hilbert space of  $\mathbb{R}^n \times \mathcal{Y}$  equipped with  $\mathcal{L}_2$ -norm, that is, for  $x, z \in \mathbb{R}^n$  and  $y, u \in \mathcal{Y}$ ,

$$\langle (x, y), (z, u) \rangle := x^T z + \int_{\Xi} y(\xi)^T u(\xi) P(d\xi).$$

Under Assumption 1, for any  $(x, y(\cdot)), (z, u(\cdot)) \in \mathbb{R}^n \times \mathcal{Y}$ , we have

$$\begin{aligned} & \left\langle \begin{pmatrix} A(x-z) + \mathbb{E}[B(\xi)(y(\xi) - u(\xi))] \\ M(\xi)(y(\xi) - u(\xi)) + N(\xi)(x-z) \end{pmatrix}, \begin{pmatrix} x-z \\ y(\xi) - u(\xi) \end{pmatrix} \right\rangle \\ &= \mathbb{E} \left[ \begin{pmatrix} x-z \\ y(\xi) - u(\xi) \end{pmatrix}^T \begin{pmatrix} A & B(\xi) \\ N(\xi) & M(\xi) \end{pmatrix} \begin{pmatrix} x-z \\ y(\xi) - u(\xi) \end{pmatrix} \right] \\ &\geq \mathbb{E}[\kappa(\xi)(\|x-z\|^2 + \|y(\xi) - u(\xi)\|^2)]. \end{aligned}$$

The existence and uniqueness of  $(x^*, y^*(\cdot))$  follow from the monotonicity of problem (1.1), see [17, Theorem 12.2 and Lemma 12.2].

Part (iv). Let us rewrite  $F(x)$  in (2.6) as

$$F(x) = \min(x, Ax - \mathbb{E}[B(\xi)\bar{y}(x, \xi)] + q_1),$$

where  $\bar{y}(x, \xi)$  is defined in (2.4). By [10, Theorem 2.1],

$$\begin{aligned} \partial_x \bar{y}(x, \xi) &= \text{conv} \{ \lim_{z \rightarrow x} \nabla_z \bar{y}(z, \xi) = -[I - D(z, \xi)(I - M(\xi))]^{-1} D(z, \xi) N(\xi) : \bar{y}(z, \xi) \text{ is nondegenerate} \} \\ &\subseteq \text{conv} \{ -U_J(M(\xi))N(\xi) : J \in \mathcal{J} \}, \end{aligned}$$

where ‘‘conv’’ denotes convex hull of a set,  $\partial_x \bar{y}(x, \xi)$  denotes the Clarke generalized Jacobian with respect to  $x$  and non-degeneration of  $\bar{y}(z, \xi)$  means that  $\{i | (M(\xi)\bar{y}(z, \xi) + N(\xi)z + q_2(\xi))_i = \bar{y}_i(z, \xi)\} = \emptyset$ . To ease the exposition, let  $\Gamma_J(\xi) := A - B(\xi)U_J(M(\xi))N(\xi)$ . Then

$$\partial_x \mathbb{E}[Ax - B(\xi)\bar{y}(x, \xi) + q_1] \subseteq \mathbb{E}[\text{conv}\{\Gamma_J(\xi) : J \in \mathcal{J}\}],$$

where the expectation is taken in the sense of Aumann [3], and hence

$$\partial F(x) \subseteq \{\mathbb{E}[I - D + D\Upsilon(\xi)] : \Upsilon(\xi) \in \text{conv}\{\Gamma_J(\xi)\}, J \in \mathcal{J}, D \in \mathcal{D}, \xi \in \Xi\}, \quad (2.7)$$

where  $\mathcal{D}$  is defined immediately after Assumption 1, and  $\Upsilon(\xi(\cdot))$  represents integrable selections from the convex hull of random set-valued mapping  $\{\Gamma_J(\xi(\cdot))\}, J \in \mathcal{J}, D \in \mathcal{D}, \xi(\cdot) : \Omega \rightarrow \Xi$ . On the other hand, under Assumption 1,  $\Gamma_J(\xi)$  is positive definite for all  $\xi$  and it follows by Lemma 2.1 (ii)

$$z^T \mathbb{E}[\Gamma_J(\xi)]z = \sum_{i=1}^n z_i (\mathbb{E}[\Gamma_J(\xi)]z)_i \geq \mathbb{E}[\kappa(\xi)] \|z\|^2, \quad \forall z \in \mathbb{R}^n.$$

This implies

$$\max_{i \in \bar{n}} z_i (\mathbb{E}[\Gamma_J(\xi)]z)_i \geq \frac{\mathbb{E}[\kappa(\xi)]}{n} \|z\|^2, \quad \forall z \in \mathbb{R}^n.$$

For a fixed  $z \in \mathbb{R}^n$ , let  $i_0 := \arg \max_{i \in \bar{n}} z_i (\mathbb{E}[\Gamma_J(\xi)]z)_i$ . Then

$$\frac{\mathbb{E}[\kappa(\xi)]}{n} \|z\|^2 \leq z_{i_0} (\mathbb{E}[\Gamma_J(\xi)]z)_{i_0} \leq |z_{i_0}| \|\mathbb{E}[\Gamma_J(\xi)]\| \|z\| \quad (2.8)$$

from which we deduce

$$|z_{i_0}| \geq \frac{\mathbb{E}[\kappa(\xi)]}{n \|\mathbb{E}[\Gamma_J(\xi)]\|} \|z\|. \quad (2.9)$$

Moreover

$$z_{i_0} ((I - D)z)_{i_0} + z_{i_0} (D\mathbb{E}[\Gamma_J(\xi)]z)_{i_0} = \begin{cases} |z_{i_0}|^2, & \text{if } D_{i_0 i_0} = 0, \\ z_{i_0} (\mathbb{E}[\Gamma_J(\xi)]z)_{i_0}, & \text{if } D_{i_0 i_0} = 1. \end{cases} \quad (2.10)$$

Let  $\Theta := \min \left\{ \left( \frac{\mathbb{E}[\kappa(\xi)]}{n \|\mathbb{E}[\Gamma_J(\xi)]\|} \right)^2, \frac{\mathbb{E}[\kappa(\xi)]}{n} \right\}$ . Then we obtain by combining (2.8)-(2.10) that

$$z_{i_0} ((I - D)z)_{i_0} + z_{i_0} (D\mathbb{E}[\Gamma_J(\xi)]z)_{i_0} \geq \Theta \|z\|^2.$$

Hence for any  $z \in \mathbb{R}^n$ , we have

$$\max_{i \in \bar{n}} z_i ((I - D + D\mathbb{E}[\Gamma_J(\xi)]z)_i) = \max_{i \in \bar{n}} z_i ((I - D)z)_i + z_i (D\mathbb{E}[\Gamma_J(\xi)]z)_i \geq \Theta \|z\|^2,$$

which implies that  $I - D + D\mathbb{E}[\Gamma_J(\xi)]$  is a P-matrix<sup>1</sup> for any  $D \in \mathcal{D}$ . Let

$$\theta(\mathbb{E}[\Gamma_J(\xi)]) := \min_{\|z\|_\infty=1} \left\{ \max_{i \in \bar{n}} z_i (\mathbb{E}[\Gamma_J(\xi)]z)_i \right\}.$$

It follows from (2.8) and (2.9) that  $\theta(\mathbb{E}[\Gamma_J(\xi)]) \geq \Theta$  and by [8, formula (1.6)],

$$\max_{D \in \mathcal{D}} \|(I - D + D\mathbb{E}[\Gamma_J(\xi)])^{-1}\|_\infty \leq \frac{\max\{1, \|\mathbb{E}[\Gamma_J(\xi)]\|_\infty\}}{\theta(\mathbb{E}[\Gamma_J(\xi)])} \leq \frac{1}{\Theta} \max\{1, \|\mathbb{E}[\Gamma_J(\xi)]\|_\infty\}.$$

Through (2.7), this implies that every matrix in  $\partial_x F(x)$  is nonsingular and the infinity norm of its inverse is bounded by  $\max\{1, \|\mathbb{E}[\Gamma_J(\xi)]\|_\infty\} / \Theta$ .  $\blacksquare$

Recall that one of the main objectives of this paper is to develop a discretization scheme for the two-stage SLCP (1.1) and we will do so in the next section. While it is not necessarily a prerequisite, we find it is much more convenient, at least from presentational perspective, to discuss the approach when the support set  $\Xi$  is compact. If we made a direct assumption on the compactness of  $\Xi$ , it would exclude a number of practically interesting cases where the support set is unbounded. In what follows, we try to address the dilemma under some moderate conditions.

Let  $\epsilon$  be a small positive number and  $\Xi_\epsilon$  be a compact subset of  $\Xi$ . Since  $B(\xi)U_J(M(\xi))N(\xi)$  and  $B(\xi)U_{J(\xi)}(M(\xi))q_2(\xi)$  are integrable for all  $J(\xi) \subseteq \{1, \dots, n\}$ , then we can find  $\Xi_\epsilon \subset \Xi$  such that

$$\left\| \int_{\Xi \setminus \Xi_\epsilon} B(\xi)U_{J(\xi)}(M(\xi))N(\xi)P(d\xi) \right\| \leq \epsilon \quad \text{and} \quad \left\| \int_{\Xi \setminus \Xi_\epsilon} B(\xi)U_{J(\xi)}(M(\xi))q_2(\xi)P(d\xi) \right\| \leq \epsilon. \quad (2.11)$$

<sup>1</sup>Recall that  $A \in \mathbb{R}^{n \times n}$  is a P-matrix if  $\max_{i \in \bar{n}} z_i (Az)_i > 0$  for all nonzero  $z \in \mathbb{R}^n$ .

We consider the complementarity problem

$$\begin{cases} 0 \leq x \perp Ax + \mathbb{E}_{\Xi_\epsilon}[B(\xi)y(\xi)] + q_1 \geq 0, \\ 0 \leq y(\xi) \perp M(\xi)y(\xi) + N(\xi)x + q_2(\xi) \geq 0, \quad \text{for a.e. } \xi \in \Xi_\epsilon, \end{cases} \quad (2.12)$$

where  $\mathbb{E}_{\Xi_\epsilon}[H(\xi)] := \int_{\xi \in \Xi_\epsilon} H(\xi)P(d\xi)$  with a compact set  $\Xi_\epsilon$ . Note that  $\mathbb{E}_{\Xi_\epsilon}$  may be regarded as the expected value of  $H(\xi)$  conditional on  $\xi \in \Xi_\epsilon$ . However, we are a little unwilling to write this as a conditional expectation since it would make the exposition more complex.

Under Assumption 1, it follows by Proposition 2.1 that (2.12) has a unique solution (see [12, 15]). Let us denote the solution by  $(x_\epsilon, \bar{y}(x_\epsilon, \cdot))$  and substitute  $\bar{y}(x_\epsilon, \cdot) = -W(x_\epsilon, \xi)(N(\xi)x_\epsilon + q_2(\xi))$  obtained from the second equation of (2.12) into the first equation, we obtain

$$0 \leq x \perp (A - \mathbb{E}_{\Xi_\epsilon}[B(\xi)W(x, \xi)N(\xi)])x - \mathbb{E}_{\Xi_\epsilon}[B(\xi)W(x, \xi)q_2(\xi)] + q_1 \geq 0, \quad (2.13)$$

where  $W(x, \xi)$  is defined as in (2.3).

The following proposition quantifies the difference between  $x_\epsilon$  and the true solution of the two-stage SLCP (1.1).

**Proposition 2.2** *Let Assumption 1 hold. Then there exists a positive constant  $\epsilon_0$  such that the two-stage SLCP (2.12) has a unique solution  $x_\epsilon$  for all  $\epsilon \in (0, \epsilon_0]$ . Moreover, there exists a positive constant  $C$  such that*

$$\|x^* - x_\epsilon\| \leq C\epsilon, \quad \forall \epsilon \in (0, \epsilon_0]. \quad (2.14)$$

**Proof.** Note that NCP (2.5) and (2.13) can be rewritten respectively as  $F(x) = 0$  and

$$F_\epsilon(x) := \min(x, (A - \mathbb{E}_{\Xi_\epsilon}[B(\xi)W(x, \xi)N(\xi)])x - \mathbb{E}_{\Xi_\epsilon}[B(\xi)W(x, \xi)q_2(\xi)] + q_1) = 0, \quad (2.15)$$

where  $F(x)$  is defined in (2.6). By (2.11), we have

$$\begin{aligned} \|F(x) - F_\epsilon(x)\| &\leq \|(\mathbb{E}[B(\xi)W(x, \xi)N(\xi)] - \mathbb{E}_{\Xi_\epsilon}[B(\xi)W(x, \xi)N(\xi)])x\| \\ &\quad + \|\mathbb{E}[B(\xi)W(x, \xi)q_2(\xi)] - \mathbb{E}_{\Xi_\epsilon}[B(\xi)W(x, \xi)q_2(\xi)]\| \\ &\leq \epsilon(1 + \|x\|). \end{aligned}$$

By Lemma 2.1,

$$\|(A - \mathbb{E}[B(\xi)W(x, \xi)N(\xi)])^{-1}\|, \|(A - \mathbb{E}_{\Xi_\epsilon}[B(\xi)W(x, \xi)N(\xi)])^{-1}\|,$$

$\|\mathbb{E}[B(\xi)W(x, \xi)q_2(\xi)] + q_1\|$  and  $\|\mathbb{E}_{\Xi_\epsilon}[B(\xi)W(x, \xi)q_2(\xi)] + q_1\|$  are uniformly bounded for  $\epsilon \in (0, \epsilon_0]$ , thus, there exists a positive number  $\rho > 0$  such that the solutions of NCP (2.5) and (2.13) lie in the ball  $B(0, \rho)$  of  $\mathbb{R}^n$ , we have

$$\sup_{\|x\| \leq \rho} \|F(x) - F_\epsilon(x)\| \leq \epsilon(1 + \rho).$$

By Lemma 2.1 (ii),  $A - \mathbb{E}[B(\xi)W(x, \xi)N(\xi)]$  is positive definite and

$$\|(A - \mathbb{E}[B(\xi)W(x, \xi)N(\xi)])^{-1}\| \leq \frac{1}{\kappa(\xi)}.$$

Thus, there exists a positive constant  $\epsilon_0$  such that for all  $\epsilon \in (0, \epsilon_0]$ , we have

$$A - \mathbb{E}_{\Xi_\epsilon}[B(\xi)W(x, \xi)N(\xi)] \succ 0 \quad \text{and} \quad \|(A - \mathbb{E}_{\Xi_\epsilon}[B(\xi)W(x, \xi)N(\xi)])^{-1}\| \leq \frac{2}{\kappa(\xi)}.$$

Since  $\|\mathbb{E}_{\Xi_\epsilon}[B(\xi)W(x, \xi)q_2(\xi)]\|$  is also bounded, there is a positive number  $\alpha_1$  such that  $\|x_\epsilon\| \leq \alpha_1$  for all  $\epsilon \in (0, \epsilon_0]$ . This enables us to bound  $\rho$  by a positive constant independent of  $x_\epsilon$ .

By Proposition 2.1 (iv), there exists a positive constant  $\alpha_2$  such that

$$\|x_\epsilon - x^*\| \leq \alpha_2 \|F(x_\epsilon) - F(x^*)\| = \alpha_2 \|F(x_\epsilon) - F_\epsilon(x_\epsilon)\| \leq \alpha_2 \epsilon (1 + \alpha_1).$$

We obtain the conclusion by setting  $C = \alpha_2(1 + \alpha_1)$ . ■

The proposition says that the solution obtained from solving (2.12) is close to the solution of (1.1) when  $\epsilon$  is set sufficiently small. This means that we can trim off the tail of the probability distribution  $P$ .

### 3 A discretization scheme

In this section, we move on to discuss discretization approaches for the two-stage SLCP (1.1). The key challenge is that the second stage of SLCP (1.1) comprises an infinite number of complementarity problems when  $\xi$  is continuously distributed. Our idea here is to divide the support set  $\Xi$  of  $\xi$  into small subsets  $\{\Xi_i\}$  and set  $y(\xi) \equiv \mathbf{y}_i$  for  $\xi \in \Xi_i$ . This will effectively reduce the infinite number of complementarity problems at the second stage to a finite number. We then attach a probability to each of the subset and consequently discretize the continuous probability distribution  $P$  and the two-stage SLCP (1.1). Throughout this section, we assume that  $\Xi$  is a compact and convex set.

#### 3.1 Description of the discretization scheme

Let  $\{\Xi_i^K\}$  be a partition of the support set  $\Xi$ , that is,  $\Xi_i^K$  is a compact and convex subset of  $\Xi$  such that,

$$\bigcup_{i=1}^K \Xi_i^K = \Xi, \quad \text{int}\Xi_i^K \cap \text{int}\Xi_j^K = \emptyset, \quad \forall i \neq j, \quad i, j \in \bar{K},$$

where  $\text{int}S$  denotes the interior of  $S$ . Note that since  $\Xi$  is assumed to be a compact set, each  $\Xi_i^K$  is also a compact set. Let

$$\mathbb{E}_{\Xi_i^K}[H(\xi)] := \frac{1}{p_i^K} \int_{\xi \in \Xi_i^K} H(\xi) P(d\xi) \quad \text{with} \quad p_i^K = P(\Xi_i^K) \quad (3.1)$$

for  $H(\xi) = M(\xi), N(\xi), B(\xi)$  or  $q_2(\xi)$ . Let

$$\Delta(\Xi_i^K) := \max_{\xi_1, \xi_2 \in \Xi_i^K} \|\xi_1 - \xi_2\| \quad (3.2)$$

denote the diameter of  $\Xi_i^K$ . We require  $\max_{i \in \bar{K}} \Delta(\Xi_i^K) \rightarrow 0$  as  $K \rightarrow \infty$ .

Let  $x \in \mathbb{R}^n$  be fixed. For  $i \in \bar{K}$ , we consider the linear complementarity problem

$$0 \leq \mathbf{y}_i \perp \mathbb{E}_{\Xi_i^K} [M(\xi)] \mathbf{y}_i + \mathbb{E}_{\Xi_i^K} [N(\xi)] x + \mathbb{E}_{\Xi_i^K} [q_2(\xi)] \geq 0, \quad (3.3)$$

$\mathbf{y}_i$  is defined at the beginning of Section 3. Under Assumption 1,  $M(\xi) \succ 0$  for all  $\xi \in \Xi$  and hence  $\mathbb{E}_{\Xi_i^K} [M(\xi)] \succ 0$ . This ensures problem (3.3) has a unique solution. By [10] and Lemma 2.1, we can write the solution of (3.3) as

$$\bar{\mathbf{y}}_i^K(x) := -W_i^K(x) (\mathbb{E}_{\Xi_i^K} [N(\xi)] x + \mathbb{E}_{\Xi_i^K} [q_2(\xi)]), \quad (3.4)$$

where

$$W_i^K(x) := [I - D_i^K(x) (I - \mathbb{E}_{\Xi_i^K} [M(\xi)])]^{-1} D_i^K(x),$$

$D_i^K(x) \in \mathcal{D}$  is an  $m \times m$ -dimension diagonal matrix with

$$(D_i^K(x))_{jj} := \begin{cases} 1, & \text{if } (\mathbb{E}_{\Xi_i^K} [N(\xi)] x + \mathbb{E}_{\Xi_i^K} [M(\xi)] \bar{\mathbf{y}}_i^K(x) + \mathbb{E}_{\Xi_i^K} [q_2(\xi)])_j \leq (\bar{\mathbf{y}}_i^K(x))_j, \\ 0, & \text{otherwise.} \end{cases}$$

Let  $\mathbf{1}_A(a)$  denote the indicator function with  $\mathbf{1}_A(a) = 1$  for  $a \in A$  and 0 otherwise. Let

$$\bar{y}^K(x, \xi) := \sum_{i=1}^K \bar{\mathbf{y}}_i^K(x) \mathbf{1}_{\Xi_i^K}(\xi), \quad (3.5)$$

that is,  $\bar{y}^K(x, \xi) = \bar{\mathbf{y}}_i^K(x)$ ,  $\forall \xi \in \Xi_i^K$ .

We use  $\bar{y}^K(x, \xi)$  as a piecewise step-like approximation to the true solution  $\bar{y}(x, \xi)$  of the second stage of SLCP (1.1). Substituting  $\bar{y}^K(x, \xi)$  into the first stage of SLCP (1.1), we obtain

$$0 \leq x \perp Ax + \mathbb{E}[B(\xi) \bar{y}^K(x, \xi)] + q_1 \geq 0. \quad (3.6)$$

**Remark 3.1** Problems (3.3)–(3.6) provide a discrete approximation of two-stage SLCP (1.1). To see this, let  $p_i^K := P(\Xi_i^K)$ . Then we can write (3.3)–(3.6) as

$$\begin{cases} 0 \leq x \perp Ax + \sum_{i=1}^K p_i^K \mathbb{E}_{\Xi_i^K} [B(\xi)] \mathbf{y}_i + q_1 \geq 0, \\ 0 \leq \mathbf{y}_i \perp \mathbb{E}_{\Xi_i^K} [M(\xi)] \mathbf{y}_i + \mathbb{E}_{\Xi_i^K} [N(\xi)] x + \mathbb{E}_{\Xi_i^K} [q_2(\xi)] \geq 0, \quad i \in \bar{K}. \end{cases} \quad (3.7)$$

*This is a two-stage SLCP with discrete distribution. The discretization scheme should be distinguished from the well-known sample average approximation scheme where the second stage solution is restricted only to the sample points each of which is attached with equal probability. Our approach is more accurate by exploiting the information of problem data in each set  $\Xi_i^K$  albeit at the cost of calculating  $p_i^K$ .*

## 3.2 Qualitative and quantitative convergence analysis

Let  $(x^K, \mathbf{y}^K)$  denote the solution of (3.7) whereby we write  $\mathbf{y}^K$  for  $(\mathbf{y}_1^K, \dots, \mathbf{y}_K^K)$ . Let

$$y^K(\xi) := \sum_{i=1}^K \mathbf{y}_i^K \mathbf{1}_{\Xi_i^K}(\xi). \quad (3.8)$$

We investigate convergence of  $(x^K, y^K(\cdot))$  to  $(x^*, y^*(\cdot))$ , the true solution of the two-stage SLCP (1.1) as  $\max_{k \in \bar{K}} \Delta(\Xi_i^K)$  goes to zero. At this point, it might be helpful to emphasize the difference between  $y^K(\xi)$  and  $\bar{y}^K(x, \xi)$  defined in (3.5): the latter depends on each fixed  $x$  whereas the former does not depend on  $x$ . Using (3.4) and (3.5), we have

$$\bar{y}^K(x, \xi) = - \sum_{i=1}^K W_i^K(x) (\mathbb{E}_{\Xi_i^K}[N(\xi)]x + \mathbb{E}_{\Xi_i^K}[q_2(\xi)]) \mathbf{1}_{\Xi_i^K}(\xi). \quad (3.9)$$

Substituting the explicit form of  $\bar{y}^K(x, \xi)$  above into the first equation of (1.1), we obtain

$$0 \leq x \perp \left( A - \mathbb{E} \left[ B(\xi) \left( \sum_{i=1}^K \mathbf{1}_{\Xi_i^K}(\xi) W_i^K(x) \mathbb{E}_{\Xi_i^K}[N(\xi)] \right) \right] \right) x - Q^K(x) \geq 0, \quad (3.10)$$

where

$$Q^K(x) := \mathbb{E} \left[ B(\xi) \left( \sum_{i=1}^K \mathbf{1}_{\Xi_i^K}(\xi) W_i^K(x) \mathbb{E}_{\Xi_i^K}[q_2(\xi)] \right) \right] - q_1.$$

The following theorem states the convergence of  $(x^K, y^K(\xi))$  to  $(x^*, y^*(\xi))$  as  $K \rightarrow \infty$ .

**Theorem 3.1** *Under Assumption 1, the following assertions hold.*

- (i) *The complementarity problem (3.7) has a unique solution  $(x^K, \mathbf{y}^K)$ .*
- (ii) *If, in addition,  $\max_{i \in \bar{K}} \Delta(\Xi_i^K) \rightarrow 0$ , then  $\{(x^K, y^K(\cdot))\}$  is bounded on  $\mathbb{R}^n \times \mathcal{Y}$ , where the boundedness of  $y^K(\cdot)$  is in the sense of the norm topology of  $\mathcal{L}_1(\mathcal{Y})$ .*
- (iii)  *$\{x^K, y^K(\cdot)\}$  converges to the true solution  $(x^*, y^*(\cdot))$  of problem (1.1), where the convergence of  $\{y^K(\cdot)\} \rightarrow y^*(\cdot)$  is in the sense of the norm topology of  $\mathcal{L}_2(\mathcal{Y})$ .*

**Proof.** Part (i). By Assumption 1,

$$(z^T, u^T) \begin{pmatrix} A & \mathbb{E}_{\Xi_i^K}[B(\xi)] \\ \mathbb{E}_{\Xi_i^K}[N(\xi)] & \mathbb{E}_{\Xi_i^K}[M(\xi)] \end{pmatrix} \begin{pmatrix} z \\ u \end{pmatrix} \geq \mathbb{E}_{\Xi_i^K}[\kappa(\xi)] (\|z\|^2 + \|u\|^2).$$

Analogous to Proposition 2.1, we can demonstrate that the discretized two-stage SLCP (3.7) has a unique solution.

Part (ii). By definition,  $x^K$  is a solution to the first stage of SLCP (3.7) and it is the unique solution for all  $K$ . Moreover

$$W_i^K(x^K) = (I - D_i^K(x^K)(I - \mathbb{E}_{\Xi_i^K}[M(\xi)]))^{-1} D_i^K(x^K),$$

where  $D_i^K(x^K) \in \mathcal{D}$ ,  $i \in \bar{K}$ . For any given  $\bar{D}_i \in \mathcal{D}$ , let

$$\bar{W}_i^K := (I - \bar{D}_i(I - \mathbb{E}_{\Xi_i^K}[M(\xi)]))^{-1} \bar{D}_i \quad \text{and} \quad \tilde{W}_i^K(\xi) := (I - \bar{D}_i(I - M(\xi)))^{-1} \bar{D}_i.$$

Under Assumption 1,  $\|\tilde{W}_i^K(\xi)\| = \|U_J(M(\xi))\| \leq \frac{1}{\kappa(\xi)}$  for some subset  $J \in \mathcal{J}$ . Thus  $\|\bar{W}_i^K\| \leq \frac{1}{\kappa(\xi)} + 1$  for all  $K$ . Let

$$R^K := \mathbb{E} \left[ B(\xi) \left( \sum_{i=1}^K \mathbf{1}_{\Xi_i^K}(\xi) \left( \bar{W}_i^K \mathbb{E}_{\Xi_i^K}[N(\xi)] - \tilde{W}_i^K(\xi) N(\xi) \right) \right) \right]. \quad (3.11)$$

For  $\bar{D}_i = D_i^K(x^K)$ ,  $x^K$  satisfies

$$0 \leq x \perp \left( A - \mathbb{E} \left[ B(\xi) \left( \sum_{i=1}^K \mathbf{1}_{\Xi_i^K}(\xi) \bar{W}_i^K N(\xi) \right) \right] \right) x - \bar{Q}^K \geq 0, \quad (3.12)$$

where

$$\bar{Q}^K = \mathbb{E} \left[ B(\xi) \left( \sum_{i=1}^K \mathbf{1}_{\Xi_i^K}(\xi) \bar{W}_i^K \mathbb{E}_{\Xi_i^K}[q_2(\xi)] \right) \right] - q_1.$$

Moreover, (3.12) can be written as

$$0 \leq x \perp \left( A - \mathbb{E} \left[ B(\xi) \left( \sum_{i=1}^K \mathbf{1}_{\Xi_i^K}(\xi) \tilde{W}_i^K(\xi) N(\xi) \right) \right] + R^K \right) x - \bar{Q}^K \geq 0. \quad (3.13)$$

Note that for any  $\xi \in \Xi_i^K$ ,

$$\bar{W}_i^K \mathbb{E}_{\Xi_i^K}[N(\xi)] - \tilde{W}_i^K(\xi) N(\xi) = \bar{W}_i^K (\mathbb{E}_{\Xi_i^K}[N(\xi)] - N(\xi)) + N(\xi) (\bar{W}_i^K - \tilde{W}_i^K). \quad (3.14)$$

Since  $\max_{i \in \bar{K}} \Delta(\Xi_i^K) \rightarrow 0$ , and both  $M(\cdot)$  and  $N(\cdot)$  are continuous over  $\Xi$ , we have

$$\sup_{\xi \in \Xi_i^K} \|M(\xi) - \mathbb{E}_{\Xi_i^K}[M(\xi)]\| \rightarrow 0 \quad \text{and} \quad \sup_{\xi \in \Xi_i^K} \|N(\xi) - \mathbb{E}_{\Xi_i^K}[N(\xi)]\| \rightarrow 0.$$

Moreover, under Assumption 1, it follows by Lemma 2.1,  $\|\bar{W}_i^K\| \leq \frac{1}{\min_{\xi \in \Xi} \kappa(\xi)} + 1$  (note that  $\kappa(\xi)$  is positive continuous over the compact set  $\Xi$ ,  $\min_{\xi \in \Xi} \kappa(\xi) > 0$ ). Thus, the first term at the right hand side of (3.14) goes to zero as  $K \rightarrow \infty$ . Likewise, we can show that the second term at the right hand side of (3.14) goes to zero as  $K \rightarrow \infty$ . Summarizing the discussions above, we are able to claim, by the Lebesgue Dominated Convergence Theorem that,

$$\lim_{K \rightarrow \infty} \mathbb{E} \left[ B(\xi) \left( \sum_{i=1}^K \mathbf{1}_{\Xi_i^K}(\xi) \bar{W}_i^K(\xi) N(\xi) \right) \right] = \mathbb{E} \left[ B(\xi) \left( \sum_{i=1}^K \mathbf{1}_{\Xi_i^K}(\xi) \tilde{W}_i^K(\xi) N(\xi) \right) \right]$$

and  $R^K \rightarrow 0$  as  $K \rightarrow \infty$ . By Lemma 2.1,

$$x^T \left( A - \mathbb{E} \left[ B(\xi) \left( \sum_{i=1}^K \mathbf{1}_{\Xi_i^K}(\xi) \bar{W}_i^K \mathbb{E}_{\Xi_i^K}[N(\xi)] \right) \right] \right) x \geq \frac{1}{2} \mathbb{E}[\kappa(\xi)] \|x\|^2$$

and

$$\left\| \left( A - \mathbb{E} \left[ B(\xi) \left( \sum_{i=1}^K \mathbf{1}_{\Xi_i^K}(\xi) \bar{W}_i^K \mathbb{E}_{\Xi_i^K}[N(\xi)] \right) \right] \right)^{-1} \right\| \leq \frac{2}{\mathbb{E}[\kappa(\xi)]}$$

for any  $\{\bar{D}_i\} \subset \mathcal{D}$ . This entails

$$\left\| \left( A - \mathbb{E} \left[ B(\xi) \left( \sum_{i=1}^K \mathbf{1}_{\Xi_i^K}(\xi) \bar{W}_i^K(x) \mathbb{E}_{\Xi_i^K}[N(\xi)] \right) \right] \right)^{-1} \right\| \leq \frac{2}{\mathbb{E}[\kappa(\xi)]} \quad (3.15)$$

for all  $x \in \mathbb{R}_+^n$ . On the other hand, the boundedness of  $\bar{Q}^K$  implies that  $\|Q^K(x)\|$  is uniformly bounded w.r.t.  $x$ . Together with (3.15), we are able to claim the boundedness of  $\{x^K\}$ . This

together with integrable boundedness of  $M(\xi)^{-1}$ ,  $N(\xi)$  and  $q_2(\xi)$ , ensures that  $\mathbf{y}^K$  is bounded and so is  $y^K(\cdot)$ .

Part (iii). Let  $\hat{x}$  be a cluster point of the sequence  $\{x^K\}$  and assume without loss of generality that  $x^K \rightarrow \hat{x}$ . For any fixed  $\xi \in \Xi$ , there exists  $\{i_K\}$  such that  $\xi \in \bigcap_K \Xi_{i_K}^K$ . By the implicit function theorem [38, Lemma 2.2],  $y^K(\xi)$  converges to  $\hat{y}(\xi)$  which satisfies

$$0 \leq \hat{y}(\xi) \perp M(\xi)\hat{y}(\xi) + N(\xi)\hat{x} + q_2(\xi) \geq 0, \text{ for a.e. } \xi \in \Xi. \quad (3.16)$$

Moreover, since  $y^K(\cdot)$  is bounded, by the Lebesgue Dominated Convergence Theorem,

$$\lim_{K \rightarrow \infty} \mathbb{E}[B(\xi)y^K(\xi)] = \mathbb{E}\left[\lim_{K \rightarrow \infty} B(\xi)y^K(\xi)\right] = \mathbb{E}[B(\xi)\hat{y}(\xi)].$$

Through (3.5) and (3.6), we have

$$0 \leq \hat{x} \perp A\hat{x} + \mathbb{E}[B(\xi)\hat{y}(\xi)] + q_1 \geq 0. \quad (3.17)$$

The equation above coincides with the first stage of SLCP (1.1). Since the second stage problem has a unique solution,  $(\hat{x}, \hat{y}(\xi))$  coincides with  $(x^*, y^*(\xi))$  a.e. The proof is complete.  $\blacksquare$

In what follows, we take a step further to quantify the discrepancy between  $(x^K, y^K(\xi))$  and  $(x^*, y^*(\xi))$  as  $K \rightarrow \infty$ . For this purpose, we require the underlying coefficient matrices and vectors to be Lipschitz continuous w.r.t.  $\xi$ .

**Assumption 2**  $M(\cdot)$ ,  $N(\cdot)$ ,  $q_2(\cdot)$  and  $B(\cdot)$  are Lipschitz continuous over a compact set containing  $\Xi$  with Lipschitz constant  $L$ .

Under Assumptions 2, it is easy to show that

$$\|H(\xi) - \mathbb{E}_{\Xi_i^K}[H(\xi)]\| \leq L\Delta(\Xi_i^K), \forall \xi \in \Xi_i^K, \quad (3.18)$$

for  $H(\xi) = M(\xi), N(\xi), B(\xi)$  and  $q_2(\xi)$ .

**Theorem 3.2** Under Assumptions 1 and 2, there exist a positive number  $\gamma \geq 0$  and nonnegative integrably bounded functions  $c(\xi)$  and  $h(\xi)$  such that

$$\|x^K - x^*\| \leq \gamma \mathbb{E}[\|B(\xi)\|c(\xi)]L \max_{i \in \bar{K}} \Delta(\Xi_i^K) \quad (3.19)$$

and

$$\|y^K(\xi) - y^*(\xi)\| \leq h(\xi)L \max_{i \in \bar{K}} \Delta(\Xi_i^K), \text{ for } \forall \xi \in \Xi. \quad (3.20)$$

**Proof.** We first prove (3.19) and proceed it in two steps.

Step 1. By Theorem 3.1, there exists a compact set  $X \subset \mathbb{R}^n$  which encompasses  $\{x^K\}$  and  $x^*$ . For any fixed  $\xi \in \Xi$  and  $x \in X$ , we consider the second stage complementarity problem

$$0 \leq y \perp M(\xi)y + N(\xi)x + q_2(\xi) \geq 0. \quad (3.21)$$

As we discussed in the previous section, (3.21) has a unique solution  $y$  under Assumption 1. Moreover, we can write (3.21) equivalently as

$$\Phi(y, x, \xi) := \min(y, M(\xi)y + N(\xi)x + q_2(\xi)) = 0.$$

Under Assumption 2,  $\Phi$  is locally Lipschitz continuous near  $(y, x, \xi)$ . By the implicit function theorem [38, Lemma 2.2], (3.21) has a unique locally Lipschitz continuous function  $\bar{y}(\tilde{x}, \tilde{\xi})$  such that  $\bar{y}(x, \xi) = y$  and  $\Phi(\bar{y}(\tilde{x}, \tilde{\xi}), \tilde{x}, \tilde{\xi}) = 0$  for all  $(\tilde{x}, \tilde{\xi})$  close to  $(x, \xi)$ . Likewise, we can show that there is a unique locally Lipschitz continuous  $\bar{\mathbf{y}}_i^K(x)$  which satisfies

$$0 \leq \mathbf{y}_i \perp \mathbb{E}_{\Xi_i^K} [M(\xi)] \mathbf{1}_{\Xi_i^K}(\xi) \mathbf{y}_i + \mathbb{E}_{\Xi_i^K} [N(\xi)] \mathbf{1}_{\Xi_i^K}(\xi) x + \mathbb{E}_{\Xi_i^K} [q_2(\xi)] \mathbf{1}_{\Xi_i^K}(\xi) \geq 0, \quad (3.22)$$

for  $\xi \in \Xi_i^K, i \in \bar{K}$ . Note that the solution to (3.22) may be represented as in (3.4).

With (3.18), we may regard (3.22) as a perturbation of (3.21). Let  $\eta \in (0, 1)$  be a constant and  $K$  sufficiently large such that

$$\mathbb{E}_{\Xi_i^K} [M(\xi)] \in \{Q | \beta(\xi) \|M(\xi) - Q\| \leq \eta\}, \quad \text{for } i \in \bar{K},$$

where

$$\beta(\xi) := \max_{D \in \mathcal{D}} \|(I - D + DM(\xi))^{-1} D\|.$$

Let  $\alpha(\xi) := \beta(\xi)/(1 - \eta)$ . By [9, Theorem 2.8], for  $\xi \in \Xi_i^K$ ,

$$\begin{aligned} \|\bar{\mathbf{y}}_i^K(x) - \bar{y}(x, \xi)\| &\leq \alpha^2(\xi) \|(-N(\xi)x - q_2(\xi))_+\| \|\mathbb{E}_{\Xi_i^K} [M(\xi)] - M(\xi)\| \\ &\quad + \alpha(\xi) \|\mathbb{E}_{\Xi_i^K} [N(\xi)]x - N(\xi)x + \mathbb{E}_{\Xi_i^K} [q_2(\xi)] - q_2(\xi)\| \\ &\leq L(\alpha^2(\xi) (\|N(\xi)\| \|x\| + \|q_2(\xi)\|) + \alpha(\xi) (\|x\| + 1)) \Delta(\Xi_i^K) \\ &= Lh_1(x, \xi) \Delta(\Xi_i^K), \end{aligned} \quad (3.23)$$

where  $h_1(x, \xi) := \alpha^2(\xi) (\|N(\xi)\| \|x\| + \|q_2(\xi)\|) + \alpha(\xi) (\|x\| + 1)$ .

Under Assumption 1,  $\mathbb{E}[h_1(x, \xi)] < \infty$ . The error bound holds for all  $i \in \bar{K}$ .

Step 2. By substituting  $\bar{y}(x, \xi)$  and  $\bar{\mathbf{y}}_i^K(x)$  into the first equation of (1.1) and (3.7) respectively, we have

$$0 \leq x \perp Ax + \sum_{i=1}^K p_i^K \mathbb{E}_{\Xi_i^K} [B(\xi) \bar{y}(x, \xi)] + q_1 \geq 0 \quad \text{and} \quad 0 \leq x \perp Ax + \sum_{i=1}^K p_i^K \mathbb{E}_{\Xi_i^K} [B(\xi)] \bar{\mathbf{y}}_i^K(x) + q_1 \geq 0.$$

We write them equivalently as

$$F(x) := \min \left\{ x, Ax + \sum_{i=1}^K p_i^K \mathbb{E}_{\Xi_i^K} [B(\xi) \bar{y}(x, \xi)] + q_1 \right\} = 0$$

and

$$F^K(x) := \min \left\{ x, Ax + \sum_{i=1}^K p_i^K \mathbb{E}_{\Xi_i^K} [B(\xi)] \bar{\mathbf{y}}_i^K(x) + q_1 \right\} = 0.$$

Let  $X$  be defined as in Step 1. Since  $\Xi$  is compact and  $\bar{y}(x, \xi)$  is continuous over  $X \times \Xi$ , there is a positive constant  $\sigma$  such that  $\|\bar{y}(x, \xi)\| \leq \sigma$  for all  $(x, \xi) \in X \times \Xi$ . Using (3.18), we have

$$\begin{aligned}
\|F(x) - F^K(x)\| &\leq \left\| \sum_{i=1}^K p_i^K \mathbb{E}_{\Xi_i^K} [B(\xi) \bar{y}(x, \xi)] - \sum_{i=1}^K p_i^K \mathbb{E}_{\Xi_i^K} [B(\xi)] \bar{y}_i^K(x) \right\| \\
&\leq \sum_{i=1}^K p_i^K \mathbb{E}_{\Xi_i^K} [\|B(\xi)\| \|\bar{y}(x, \xi) - \bar{y}_i^K(x)\|] \\
&\leq L \max_{i \in \bar{K}} \Delta(\Xi_i^K) \left( \sum_{i=1}^K p_i^K \mathbb{E}_{\Xi_i^K} [\|B(\xi)\| h_1(x, \xi)] \right) \quad (\text{by 3.23}) \\
&\leq L \mathbb{E}[\|B(\xi)\| c(\xi)] \max_{i \in \bar{K}} \Delta(\Xi_i^K),
\end{aligned}$$

where

$$c(\xi) := \alpha^2(\xi)(\|N(\xi)\| \|X\| + \|q_2(\xi)\|) + \alpha(\xi)(\|X\| + 1)$$

and  $\|X\| := \max\{\|x\| : x \in X\}$ . By Proposition 2.1 (iv) and [37, Lemma 2.2],

$$\|x^K - x^*\| \leq \gamma \sup_{x \in X} \|F(x) - F^K(x)\| \leq L\gamma \mathbb{E}[\|B(\xi)\| c(\xi)] \max_{i \in \bar{K}} \Delta(\Xi_i^K)$$

for any positive number  $\gamma \geq \frac{1}{\mathbb{E}[\kappa(\xi)]}$ . This completes the proof of (3.19).

Next we prove (3.20). Using the established error bound (3.19) for the  $x$ -component of the SLCP solutions, we can again use [9, Theorem 2.8] for the  $y$ -component of the solutions,

$$\begin{aligned}
\|y^K(\xi) - y^*(\xi)\| &= \|\bar{y}^K(x^K, \xi) - \bar{y}(x^*, \xi)\| \\
&\leq \alpha^2(\xi) \|(-N(\xi)x^* - q_2(\xi))_+\| \|\mathbb{E}_{\Xi_i^K} [M(\xi)] - M(\xi)\| \\
&\quad + \alpha(\xi) \|\mathbb{E}_{\Xi_i^K} [N(\xi)] x^K - N(\xi) x^* + \mathbb{E}_{\Xi_i^K} [q_2(\xi)] - q_2(\xi)\| \\
&\leq Lh(\xi) \max_{i \in \{1, \dots, K\}} \Delta(\Xi_i^K),
\end{aligned} \tag{3.24}$$

where  $h(\xi) := \alpha^2(\xi)(\|N(\xi)\| C_0 + \|q_2(\xi)\|) + \alpha(\xi) C_0$  with  $C_0 = (\|X\| + 1) + \gamma \mathbb{E}[\|B(\xi)\| c(\xi)]$ . It is easy to see that  $\mathbb{E}[h(\xi)] < +\infty$  under Assumption 1.  $\blacksquare$

It might be helpful to discuss how the partition of  $\Xi$  is made. If  $\xi$  is a single random variable, then we may divide the interval of the support set  $\Xi$  evenly into  $K$  subintervals. However, if  $\xi$  is a random vector which has several components, then  $K$  might have to be very large in order to reduce the size of  $\Xi_i^K$ . In that case, it would be sensible to use Monte Carlo sampling to generate a set of points  $\Xi^K := \{\xi^1, \dots, \xi^K\}$  and use them to develop the Voronoi partition of  $\Xi$ , that is,

$$\Xi_i^K \subseteq \left\{ \xi \in \Xi : \|\xi - \xi^i\| = \min_{k \in \bar{K}} \|\xi - \xi^k\| \right\} \quad \text{for } i \in \bar{K}, \tag{3.25}$$

are pairwise disjoint subsets forming a partition of  $\Xi$ .

### 3.3 Progressive hedging method (PHM)

The discretized two-stage SLCP (3.7) is a deterministic LCP which may be solved by any existing solvers. However, when  $K$  is large, it might be more efficient to solve (3.7) with the well known

PHM [27] which exploits the two-stage structure. Note that PHM is an iterative approach which solves finite scenario multi-stage stochastic programming problems at each scenario and then average them to get a feasible solution at each iterate. The main advantage of the approach is that the scenario based solutions can be obtained in parallel computation. Recently, Rockafellar and Sun [29] extend the method to finite scenario multi-stage stochastic variational inequalities. Here we describe how to apply PHM to solve (3.7).

Let  $\Xi^K := \{\Xi_1^K, \dots, \Xi_K^K\}$  and  $\Omega_i^K = \xi^{-1}(\Xi_i^K)$  for  $i \in \bar{K}$ . Let  $\tilde{\Omega} := \{\Omega_1^K, \dots, \Omega_K^K\}$  and  $\tilde{\mathcal{B}}$  the sigma algebra over  $\tilde{\Omega}$ . Let  $\tilde{P}$  be a probability measure over the measurable space  $\{\tilde{\Omega}, \tilde{\mathcal{B}}\}$  and  $U : \tilde{\Omega} \rightarrow 2^{\Xi^K}$  be a random variable (set-valued mapping indeed) with  $\tilde{P}(U = \Xi_i^K) := p_i^K$ , where  $p_i^K$  is defined as in (3.1). By slightly abusing the notation, we also regard  $\tilde{P}$  as a probability measure over  $\{\Xi^K, \mathcal{B}^K\}$ , where  $\mathcal{B}^K$  is Borel sigma algebra over  $\Xi^K$ , with  $P(\Xi_i^K) = p_i^K$  for  $i \in \bar{K}$ . Let  $(x(\cdot), y(\cdot))$  be a measurable mapping from  $\{\Xi^K, \mathcal{B}^K\}$  to  $\mathbb{R}^{m+n}$ , where  $x(U) = \sum_{i=1}^K x_i \mathbf{1}_{\Xi_i^K}(U)$  and  $y(U) = \sum_{i=1}^K y_i \mathbf{1}_{\Xi_i^K}(U)$ . The linear space  $\mathcal{L}$  consisting of all such mappings  $z(\cdot)$  from  $\Xi^K$  to  $\mathbb{R}^{n+m}$  is given with the expectational inner product. Let  $\mathcal{Z}$  denote the space of all measurable functions defined as such. Define the inner product

$$\langle (x(\cdot), y(\cdot)), (z(\cdot), u(\cdot)) \rangle = \mathbb{E}_{\tilde{P}}[(x(U), y(U))^T (z(U), u(U))] = \sum_{i=1}^K p_i^K (x_i^T z_i + y_i^T u_i),$$

where  $x_i, z_i \in \mathbb{R}^n$  and  $y_i, u_i \in \mathbb{R}^m$ . Let  $\mathcal{N}$  be the space of all measurable functions of form  $(x, y(U))$ , where  $x$  is independent of  $U$ . Then we may view  $\mathcal{N}$  as a subspace of  $\mathcal{Z}$  where the  $x$ -component is made deterministic (scenario free). Let

$$\mathcal{M} := \mathcal{N}^\perp = \{w(\cdot) = (w_1(\cdot), w_2(\cdot)) \in \mathcal{Z} \mid \langle (x, y), (w_1, w_2) \rangle = 0, \forall (x, y(\cdot)) \in \mathcal{N}\}.$$

Then  $w(\cdot) \in \mathcal{M}$  implies that  $\langle x, w_1 \rangle = \mathbb{E}_{\tilde{P}}[x^T w_1(U)] = 0$  and  $\mathbb{E}_{\tilde{P}}[y(U)^T w_2(U)] = 0$  for all  $(x, y(\cdot)) \in \mathcal{N}$  and therefore  $\mathbb{E}_{\tilde{P}}[w_1(U)] = 0, w_{2i} = 0, i \in \bar{K}$ . In what follows, we describe PHM for solving (3.7) as follows.

For  $i \in \bar{K}$ , let  $\tilde{B}(\Xi_i^K) := \mathbb{E}_{\Xi_i^K}[B(\xi)], \tilde{M}(\Xi_i^K) = \mathbb{E}_{\Xi_i^K}[M(\xi)], \tilde{N}(\Xi_i^K) = \mathbb{E}_{\Xi_i^K}[N(\xi)]$  and  $\tilde{q}_2(\Xi_i^K) = \mathbb{E}_{\Xi_i^K}[q_2(\xi)]$ . Then the discrete two-stage SLCP (3.7) can be reformulated as:

$$\begin{cases} 0 \leq x \perp Ax + \mathbb{E}_{\tilde{P}}[\tilde{B}(U)y(U)] + q_1 \geq 0, \\ 0 \leq y(U) \perp \tilde{M}(U)y(U) + \tilde{N}(U)x + \tilde{q}_2(U) \geq 0, \quad U \in \Xi^K. \end{cases} \quad (3.26)$$

**Algorithm 3.1 (PHM)** Given initial points  $(x^0, y^0) \in \mathcal{N}$  and  $w^0 \in \mathcal{M}$  with  $x^0(U) = \sum_{i=1}^N x_i^0 \mathbf{1}_{\Xi_i^K}(U)$ ,  $y^0(U) = \sum_{i=1}^N y_i^0 \mathbf{1}_{\Xi_i^K}(U)$  and  $w^0(U) = \sum_{i=1}^N w_i^0 \mathbf{1}_{\Xi_i^K}(U)$ ,  $i \in \bar{K}$ . Let  $r > 0$  fixed and  $\nu = 0$ .

**Step 1.** For  $i \in \bar{K}$ , solve the LCP

$$\begin{cases} 0 \leq x_i \perp Ax_i + \tilde{B}(\Xi_i^K)y_i + q_1 + w_{1i}^\nu + r(x_i - x_i^\nu) \geq 0, \\ 0 \leq y_i \perp \tilde{M}(\Xi_i^K)y_i + \tilde{N}(\Xi_i^K)x_i + \tilde{q}_2(\Xi_i^K) + r(y_i - y_i^\nu) \geq 0, \end{cases} \quad (3.27)$$

and obtain a solution  $(\hat{x}_i^\nu, \hat{y}_i^\nu), i \in \bar{K}$ .

**Step 2.** For  $i \in \bar{K}$ , let

$$\bar{x}^{\nu+1} = \sum_{i=1}^K p_i \hat{x}_i^\nu, \quad x_i^{\nu+1} = \bar{x}^{\nu+1}, \quad y_i^{\nu+1} = \hat{y}_i^\nu, \quad w_{1i}^{\nu+1} = w_{1i}^\nu + r(\hat{x}_i^\nu - x_i^{\nu+1}).$$

Set  $\nu = \nu + 1$ , go to Step 1.

Step 1 solves the SLCP per scenario and Step 2 corrects the  $x$ -component by averaging the obtained scenario based solutions.  $w_1$  is a vector of auxiliary variables corresponding to multipliers. Note that for any fixed  $r > 0$ , the algorithm terminates when

$$\left\| \min \left\{ \bar{x}^{v+1}, A\bar{x}^{v+1} + \frac{1}{N} \sum_{i=1}^N \bar{B}(\Xi_i^K) y_i^{v+1} + q_1 \right\} \right\| \leq \epsilon$$

for some specified  $\epsilon > 0$ .

## 4 Distributionally robust formulation of two-stage SLCP

In this section, we revisit the two-stage SLCP (1.1) by considering a situation where the true probability distribution  $P$  is unknown but it is possible to use partial information such as empirical data, computer simulation or subjective judgements to construct an ambiguity set of distributions which contains the true distribution with certain confidence. In such circumstances, it might be sensible to consider a robust solution to the two-stage DRLCP (1.2). In the case that  $\mathcal{P}$  encompasses all probability measures in the support set of  $\xi$ , i.e.,  $\mathcal{P}(\Xi)$ , the first stage DRLPC (1.2) reduces to

$$0 \leq x \perp Ax + B(\xi)y(\xi) + q_1 \geq 0, \quad \forall \xi \in \Xi. \quad (4.1)$$

The system might not have a solution and consequently one may consider the ERM model by replacing the complementarity system with

$$\min_{x, y(\cdot)} \mathbb{E}_P [\| \min(x, Ax + B(\xi)y(\xi) + q_1) \|^2],$$

where  $P$  is any continuous distribution with support set  $\Xi$ , see [5]. Our focus here is on the case that  $\mathcal{P}$  is only a subset of  $\mathcal{P}(\Xi)$ . We make a blanket assumption that the two-stage DRLCP (1.2) has a solution as it is difficult to figure out simple sufficient conditions for the generic model, we will come back to this in the next section when we discuss ex post equilibrium of two-stage distributionally robust games. In what follows, we discuss computational schemes for solving DRLCP (1.2). To this end, we write the first stage of (1.2) equivalently as

$$\begin{cases} -x \leq 0, \\ -Ax - \mathbb{E}_P[B(\xi)y(\xi)] - q_1 \leq 0, & \forall P \in \mathcal{P}, \\ x^T(Ax + \mathbb{E}_P[B(\xi)y(\xi)] + q_1) \leq 0, & \forall P \in \mathcal{P}. \end{cases} \quad (4.2)$$

It is easy to verify that the system of inequalities above can be equivalently written as

$$\begin{cases} -x \leq 0, \\ \max_{P \in \mathcal{P}} [-Ax - \mathbb{E}_P[B(\xi)y(\xi)] - q_1]_i \leq 0, & \text{for } i \in \bar{n}, \\ \max_{P \in \mathcal{P}} x^T(Ax + \mathbb{E}_P[B(\xi)y(\xi)] + q_1) = 0, \end{cases} \quad (4.3)$$

where we write  $[a]_i$  for the  $i$ -th component of vector  $a$ . Observe that under the first and the second equations of (4.3), the third equation of (4.3) is equivalent to

$$\sum_{i=1}^n [x]_i \max_{P \in \mathcal{P}} [Ax + \mathbb{E}_P[B(\xi)y(\xi)] + q_1]_i \leq 0. \quad (4.4)$$

To see the equivalence, we note that (4.4) implies the third equation of (4.3). Conversely, by (4.2), for every  $i \in \bar{n}$ ,

$$[x]_i[Ax + \mathbb{E}_P[B(\xi)y(\xi)] + q_1]_i \leq 0, \quad \forall P \in \mathcal{P},$$

which implies (4.4). Thus system (4.3) can be written as

$$\begin{cases} -x \leq 0, \\ \max_{P \in \mathcal{P}} [-Ax - \mathbb{E}_P[B(\xi)y(\xi)] - q_1]_i \leq 0, & \text{for } i \in \bar{n}, \\ \sum_{i=1}^n [x]_i \max_{P \in \mathcal{P}} [Ax + \mathbb{E}_P[B(\xi)y(\xi)] + q_1]_i \leq 0. \end{cases} \quad (4.5)$$

Note that if  $\mathcal{P}$  is a convex combination of a finite number of known distributions, then the maximum in  $P$  is achieved at the vertices of  $\mathcal{P}$  and consequently, the problem above reduces to the two-stage SLCP (1.1).

In what follows, we consider the case when  $\mathcal{P}$  is constructed through moment conditions:

$$\mathcal{P} := \left\{ P \in \mathcal{D} : \begin{array}{l} \mathbb{E}_P[\psi_j(\xi)] = b_j \quad \text{for } j = 1, \dots, s \\ \mathbb{E}_P[\psi_j(\xi)] \leq b_j \quad \text{for } j = s+1, \dots, t \end{array} \right\},$$

where all  $\psi_i$  are Lipschitz continuous function of  $\xi$  with Lipschitz constants  $L$  defined in Assumption 2. Our purpose is to get rid of the maximum operations w.r.t.  $P$  in (4.5) when  $\mathcal{P}$  has the specific structure. To this end, we need to assume as in the previous sections that the second stage of DRLCP (1.2) defines a unique solution  $\bar{y}(x, \xi)$  for each fixed  $x$  and  $\xi$ . Substituting  $\bar{y}(x, \xi)$  to the second equation of (4.5), we obtain

$$\max_{P \in \mathcal{P}} [-Ax - \mathbb{E}_P[B(\xi)\bar{y}(x, \xi)] - q_1]_i \leq 0, \quad \text{for } i \in \bar{n}.$$

We also need to consider

$$\max_{P \in \mathcal{P}} [Ax + \mathbb{E}_P[B(\xi)\bar{y}(x, \xi)] + q_1]_i, \quad \text{for } i \in \bar{n}$$

in the third equation of (4.5). To ease the exposition, we consider

$$\max_{P \in \mathcal{P}} \mathbb{E}_P[f_i(x, \xi)], \quad \text{for } i \in \bar{n}, \quad (4.6)$$

where  $f_i(x, \xi)$  represents the  $i$ -th component of  $-Ax - B(\xi)\bar{y}(x, \xi) - q_1$  when  $i \leq n$  and the  $(i - n)$ -th component of  $Ax + B(\xi)\bar{y}(x, \xi) + q_1$  when  $i > n$ . Define the Lagrange function

$$L_i(x, \Lambda_i, P) := \int_{\Xi} f_i(x, \xi) P(d\xi) + \lambda_0 \left( 1 - \int_{\Xi} P(d\xi) \right) + \sum_{j=1}^t \lambda_j^i \left( b_j - \int_{\Xi} \psi_j(\xi) P(d\xi) \right), \quad (4.7)$$

and

$$\bar{\Lambda}_i := \{ \lambda^i = (\lambda_0^i, \lambda_1^i, \dots, \lambda_t^i)^T : \lambda_j^i \geq 0, \text{ for } j = s+1, \dots, t \}, \quad i \in \bar{n}.$$

The Lagrange dual of problem (4.6) can be written as

$$\min_{\lambda^i \in \bar{\Lambda}_i} \max_{P \in \mathcal{M}} L_i(x, \lambda^i, P), \quad (4.8)$$

where  $\mathcal{M}$  denotes the set of all positive measures over  $\Xi$ . Conditions for strong duality can be easily established. For instance, we can consider the following Slater type condition

$$(1, 0_s, 0_{t-s}) \in \text{int}\{(\langle P, 1 \rangle, \langle P, \psi_E \rangle, \langle P, \psi_I \rangle) + \mathcal{K}_1 : P \in \mathcal{M}_+\}, \quad (4.9)$$

where  $\mathcal{K}_1 := \{0\} \times \{0_s\} \times \mathbb{R}_+^{t-s}$  and  $0_s$  denotes the zero vector in  $\mathbb{R}^s$ ,  $\psi_E = (\psi_1, \dots, \psi_s)$  and  $\psi_I = (\psi_{s+1}, \dots, \psi_t)$ . The following proposition comes straightforwardly from Xu, Liu and Sun [36, Proposition 2.1].

**Proposition 4.1** *The following assertions hold.*

(i) *Condition (4.9) is equivalent to*

$$(\mu_E, \mu_I) \in \text{int}\{(\langle P, \psi_E \rangle, \langle P, \psi_I \rangle) + \mathcal{K}_2 : P \in \mathcal{P}(\Xi)\}, \quad (4.10)$$

where  $\mathcal{K}_2 := \{0_s\} \times \mathbb{R}_+^{t-s}$ .

(ii) *Condition (4.10) is fulfilled if*

$$\mu_E \in \text{int} \{ \langle P, \psi_E(\xi) \rangle : P \in \mathcal{P}(\Xi) \} \quad (4.11)$$

and there exists  $P_E \in \mathcal{P}(\Xi)$  with  $\langle P, \psi_E(\xi) \rangle = \mu_E$  such that

$$0_{s-t} \in \text{int} \{ \langle P_E, \psi_I(\xi) \rangle - \mu_I - \mathbb{R}_-^{s-t} \}. \quad (4.12)$$

In the case when  $s = t$ , i.e., there is no inequality constraint, condition (4.11) coincides with condition (4.10). Likewise, when  $s = 0$ , i.e., there is no equality constraint, (4.13) reduces to existence of  $P \in \mathcal{P}(\Xi)$  such that

$$0_t \in \text{int} \{ \langle P, \psi_I(\xi) \rangle - \mu_I - \mathbb{R}_-^t \}, \quad (4.13)$$

which coincides with (4.10).

(iii) *Condition (4.11) holds naturally in the case when*

$$\{ \langle P, \psi_E(\xi) \rangle : P \in \mathcal{P}(\Xi) \} = \mathbb{R}^s \quad (4.14)$$

whereas condition (4.13) is fulfilled if there exists  $P_E \in \mathcal{P}(\Xi)$  with  $\langle P, \psi_E(\xi) \rangle = \mu_E$  such that

$$\langle P_E, \psi_I(\xi) \rangle - \mu_I < 0. \quad (4.15)$$

Throughout this section, we make a blanket assumption that the strong duality holds.

We now return to discuss (4.8). Through standard analysis of Lagrange duality (see i.e. discussions at page 308 of [32]), we have

$$\max_{P \in \mathcal{M}} L_i(x, \lambda^i, P) = \sum_{j=1}^t \lambda_j^i b_j + \max_{\xi \in \Xi} \left( f_i(x, \xi) - \sum_{j=1}^t \lambda_j^i \psi_j(\xi) \right). \quad (4.16)$$

Consequently, (4.5) can be written as

$$\begin{cases} -x \leq 0, \mathbf{\Lambda}_1, \mathbf{\Lambda}_2 \in \bar{\Lambda}, \\ -Ax - B(\xi)y(\xi) - q_1 - \mathbf{\Lambda}_1(\psi(\xi) - b) \leq 0, & \forall \xi \in \Xi, \\ x^T [(Ax + B(\xi)y(\xi) + q_1) - \mathbf{\Lambda}_2(\psi(\xi) - b)] \leq 0, & \forall \xi \in \Xi, \\ 0 \leq y(\xi) \perp M(\xi)y(\xi) + N(\xi)x + q_2(\xi) \geq 0, & \forall \xi \in \Xi, \end{cases} \quad (4.17)$$

where  $\bar{\Lambda} := \{\mathbf{\Lambda} = (\lambda^1, \dots, \lambda^n)^T \in \mathbb{R}^{n \times t} : -\lambda_j \leq 0, \text{ for } j = s+1, \dots, t\}$ . We consider the discrete approximation of (4.17). As what we did for the two-stage SLCP in Section 3, our first step is to develop a discretize approximation of the infinite inequality system. At this point, it might be helpful to point out the difference between the two-stage SLCP and (4.17): the former depends on the true probability distribution of  $\xi$  whereas the latter is independent of the probability distribution and it is entirely determined by the support set  $\Xi$ . This motivates us to adopt a slightly different discretization approach by using Monte Carlo sampling.

Let  $\{\xi^i\}_{i=1}^K$  be i.i.d samples of  $\xi$  generated by any probability distribution with support set  $\Xi$ . We consider the following discretization scheme for (4.17):

$$\begin{cases} -x \leq 0, \mathbf{\Lambda}_1, \mathbf{\Lambda}_2 \in \bar{\Lambda}, \\ (-Ax - B(\xi^i)\mathbf{y}_i - q_1) - (\mathbf{\Lambda}_1(\psi(\xi^i) - b))_i \leq 0, & i \in \bar{K}, \\ x^T [Ax + B(\xi^i)\mathbf{y}_i + q_1 - \mathbf{\Lambda}_2(\psi(\xi^i) - b)] \leq 0, & i \in \bar{K}, \\ 0 \leq \mathbf{y}_i \perp M(\xi^i)\mathbf{y}_i + N(\xi^i)x + q_2(\xi^i) \geq 0, & i \in \bar{K}. \end{cases} \quad (4.18)$$

Here  $\mathbf{y}_i$  is determined by the second stage LCP at sampled point  $\xi^i$ . This is in contrast to (3.3) where  $\mathbf{y}_i$  is determined by the average value of the second stage problem data over set  $\Xi_i^K$ . The underlying reason is that the first stage inequality system (the first three inequalities of (4.18)) here does not involve any probability distribution. Of course, (4.18) depends on the iid samples and hence the probability distribution which generates them. We should also point out that problem (4.18) is a dual formulation of (1.2) with the ambiguity

$$\mathcal{P}_K := \left\{ P \in \mathcal{P}(\Xi_K) : \begin{array}{l} \sum_{i=1}^K \psi_j(\xi^i) = b_j \quad \text{for } j = 1, \dots, s \\ \sum_{i=1}^K \psi_j(\xi^i) \leq b_j \quad \text{for } j = s+1, \dots, t \end{array} \right\},$$

where  $\Xi_K := \{\xi^1, \dots, \xi^K\}$ .

Problem (4.18) comprises  $Km$  complementarity problems and  $K(2n+1)$  inequalities with  $(1+2t)n + Km$  variables. It is easy to see that a solution to the true problem (4.17) is also a solution to the discretized problem (4.18). In what follows, we analyze convergence of the latter as  $K \rightarrow \infty$ .

**Theorem 4.1** *Let  $\{(x^K, \mathbf{\Lambda}_1^K, \mathbf{\Lambda}_2^K, \mathbf{y}^K)\}$  be a sequence of solutions of (4.18) with different size of samples. Assume: (a)  $\Xi$  is a compact set, (b) Assumption 2 holds, and (c) the iid samples  $\xi^1, \dots, \xi^K$  are generated by randomizing  $\xi$  and attaching to it with a continuous probability distribution  $P$  over  $\Xi$  such that*

$$P(\|\xi - \xi_0\| \leq \delta) > C\delta^\nu$$

*for any fixed point  $\xi_0 \in \Xi$  and  $\delta \in (0, \delta_0)$ , where  $C, \nu$  and  $\delta_0$  are some positive constants. Then every cluster point of the sequence  $\{(x^K, \mathbf{\Lambda}_1^K, \mathbf{\Lambda}_2^K)\}$  is the solution of (1.2) w.p.1.*

**Proof.** By taking a subsequence if necessary, we assume for the simplicity of notation that  $\{(x^K, \Lambda_1^K, \Lambda_2^K)\}$  converges to  $(\hat{x}, \hat{\Lambda}_1, \hat{\Lambda}_2)$ . This means that there exists a positive number  $\rho$  such that  $\|(x^K, \Lambda_1^K, \Lambda_2^K)\| \leq \rho$  for all  $K$ . Let  $\Xi_1^K, \dots, \Xi_K^K$  be the Voronoi partition of  $\Xi$  centred at  $\xi^1, \dots, \xi^K$ . Thus for every  $\xi \in \Xi$ , there exists a Voronoi cell  $\Xi_K^i$  centred at  $\xi^{i_K}$  such that  $\xi \in \Xi_K^i$ . Moreover, under condition (c), we can easily use [36, Lemma 3.1] to show that  $\max_{j \in \bar{K}} \Delta_j^K \rightarrow 0$  at exponential rate as  $K \rightarrow \infty$ . See [20, Proposition 8].

On the other hand, under condition (b), we can show, following a similar analysis to Step 1 in the proof of Theorem 3.2, that there exists a positive constant  $C$  such that

$$\|\bar{y}(x, \xi^{i_K}) - \bar{y}(x, \xi)\| \leq CL\Delta(\Xi_i^K). \quad (4.19)$$

Let

$$G(x, \Lambda_1, \Lambda_2, \xi) := \begin{pmatrix} -Ax - B(\xi)\bar{y}(x, \xi) - q_1 - \Lambda_1(\psi(\xi) - b) \\ x^T(Ax + B(\xi)\bar{y}(x, \xi) + q_1 - \Lambda_2(\psi(\xi) - b)) \end{pmatrix},$$

$$F_j(x, \Lambda_1, \Lambda_2) = \max_{\xi \in \Xi} G_j(x, \Lambda_1, \Lambda_2, \xi),$$

and

$$F_j^K(x, \Lambda_1, \Lambda_2) = \max_{i \in \bar{K}} G_j(x, \Lambda_1, \Lambda_2, \xi^i), \quad \text{for } j \in \overline{n+1}.$$

Then we can write the first three equations in (4.17) equivalently as

$$\begin{cases} \Lambda_1, \Lambda_2 \in \bar{\Lambda}, \\ x \geq 0, \\ F_j(x, \Lambda_1, \Lambda_2) \leq 0, \quad \text{for } j \in \overline{n+1}, \end{cases} \quad (4.20)$$

and

$$\begin{cases} \Lambda_1, \Lambda_2 \in \bar{\Lambda}, \\ x \geq 0, \\ F_j^K(x, \Lambda_1, \Lambda_2) \leq 0, \quad \text{for } j \in \overline{n+1}. \end{cases} \quad (4.21)$$

Let

$$G^K(x, \Lambda_1, \Lambda_2, \xi) := \sum_{i=1}^K \mathbf{1}_{\Xi_i^K}(\xi) G(x, \Lambda_1, \Lambda_2, \xi^i).$$

Then  $F_j^K(x, \Lambda_1, \Lambda_2) = \max_{\xi \in \Xi} G_j^K(x, \Lambda_1, \Lambda_2, \xi)$ . Under Assumption 2, for any  $(x, \Lambda_1, \Lambda_2)$  such that  $\|(x, \Lambda_1, \Lambda_2)\| \leq \rho$ , we have the boundedness of  $\bar{y}(x, \xi)$  and then there exists positive constant  $C_2$  such that

$$\begin{aligned} & \max_{i \in \overline{n+1}} |F_i(x, \Lambda_1, \Lambda_2) - F_i^K(x, \Lambda_1, \Lambda_2)| \\ & \leq \max_{i \in \overline{n+1}} \max_{\xi \in \Xi} |G_i(x, \Lambda_1, \Lambda_2, \xi) - G_i^K(x, \Lambda_1, \Lambda_2, \xi)| \\ & \leq \max_{i \in \bar{K}, k=1,2} \max_{\xi \in \Xi_i^K} (\rho + 1) \|B(\xi)\bar{y}(x, \xi) + \Lambda_k\psi(\xi) - B(\xi^k)\bar{y}(x, \xi^i) - \Lambda_k\psi(\xi^i)\| \\ & \leq (\rho + 1)LC_2\Delta(\Xi_i^K). \end{aligned}$$

This shows

$$\lim_{K \rightarrow \infty} \sup_{\|(x, \Lambda_1, \Lambda_2)\| \leq \rho} \max_{k \in \overline{n+1}} |F_k(x, \Lambda_1, \Lambda_2) - F_k^K(x, \Lambda_1, \Lambda_2)| = 0, \quad \text{w.p.1.}$$

By [34, Lemma 4.2 (i)], any cluster point of the sequence of solutions  $\{(x^K, \Lambda_1^K, \Lambda_2^K)\}$  obtained from solving system (4.21) is a solution of (4.20) almost surely.  $\blacksquare$

## 5 Two-stage distributionally robust game

In this section, we move on to discuss a two-stage distributionally robust game whose ex post equilibrium can be represented in the form of DRLCP (1.2). The notion of ex post equilibrium was introduced by Holmström and Myerson [18] under the term of “uniform incentive compatibility” and first used by Crémer and McLean [13] to provide an important avenue for computing a Bayesian equilibrium.

We consider a duopoly market<sup>2</sup> where two firms compete to supply a homogeneous product (or service) noncooperatively in future. Neither of the firms has an existing capacity and thus must make a decision at the present time on their capacity for future supply of quantities in order to allow themselves enough time to build the necessary facilities.

The market demand in future is characterized by a random inverse demand function  $p(q, \xi(\omega))$ , where  $p(q, \xi(\omega))$  is the market price,  $q$  is the total supply to the market, and  $\xi : \Omega \rightarrow \mathbb{R}$  is a continuous random variable. Specifically, for each realization of the random variable  $\xi : \Omega \rightarrow \mathbb{R}$ , we obtain a different inverse demand function  $p(q, \xi(\omega))$ . The uncertainty in the inverse demand function is then characterized by the distribution of the random variable  $\xi$ .

Firm  $i$ 's cost function for building up capacity  $x_i$  is  $C_i(x_i)$  and the cost of producing (supplying) a quantity of  $y_i$  in future is  $H_i(y_i, \xi)$ ,  $i = 1, 2$ . Assuming each firm aims to maximize the expected profit, we can then develop a mathematical model for their decision making: for  $i = 1, 2$ , find  $(x_i^*, y_i^*(\cdot))$  such that it solves the following two-stage stochastic programming problem

$$\begin{aligned} \max_{x_i, y_i(\cdot)} \quad & \mathbb{E}_P[p(y_i(\xi) + y_{-i}^*(\xi), \xi)y_i(\xi) - H_i(y_i(\xi), \xi)] - C_i(x_i) \\ \text{s.t.} \quad & 0 \leq y_i(\xi) \leq x_i, \end{aligned} \tag{5.1}$$

where the mathematical expectation is taken w.r.t. the distribution of  $\xi$  and by convention we write  $y_{-i}$  for decision variable of the firm(s) other than  $i$ . This is a closed loop two-stage stochastic Nash-Cournot game where each player (firm) needs to make a decision on capacity before realization of uncertainty anticipating competition in future (second stage). At this point, we refer readers to Wongrin et al. [33] for a deterministic model with application in electricity markets, and a more sophisticated two-stage stochastic model by Luna, Sagastizábal and Solodov [22] where each player is risk-averse and all players share an identical constraint in the second stage. Similar models can also be found in Ralph and Smeers [24] for stochastic endogenous equilibrium in asset pricing. Here we concentrate on reformulation of problem (5.1) as a two-stage SLCP under some moderate conditions and investigate the latter under this particular context.

Let us now consider a situation where each player does not have complete information on the true probability distribution  $P$ . However, each player may use available partial information to construct an ambiguity set of probability distributions, denoted respectively by  $\mathcal{P}_1, \mathcal{P}_2$ . Assuming both players base their decision on the worst probability distribution, then we may

---

<sup>2</sup>The model can be easily extended to an oligopoly, we consider a duopoly for simplicity of exposition so that we can concentrate on the main ideas.

consider a distributionally robust game: for  $i = 1, 2$ , find  $(x_i^*, y_i^*(\cdot))$  such that

$$(x_i^*, y_i^*(\cdot)) \in \arg \max_{x_i, y_i(\cdot)} \min_{P_i \in \mathcal{P}_i} \mathbb{E}_{P_i} [p(y_i(\xi) + y_{-i}^*(\xi), \xi) y_i(\xi) - H_i(y_i(\xi), \xi)] - C_i(x_i) \quad (5.2)$$

s.t.  $0 \leq y_i(\xi) \leq x_i.$

To see the structure of (5.2) clearly, we write down the optimal decision making problems at the second stage game after the market demand is observed by both firms, that is, for  $i = 1, 2$ , find  $y_i^*$  such that

$$y_i^*(\xi) \in \arg \max_{y_i} p(y_i + y_{-i}^*(\xi), \xi) y_i - H_i(y_i, \xi) \quad (5.3)$$

s.t.  $0 \leq y_i \leq x_i,$

where  $x_i$  is fixed. To analyse (5.3), we need to make some assumption on the cost functions  $H_i(y_i, \xi)$  and the inverse demand function  $p(q, \xi)$ .

**Assumption 3** For  $i = 1, 2$ ,  $H_i(y_i, \xi)$  is twice continuously differentiable,  $H_i'(y_i, \xi) \geq 0$  and  $H_i''(y_i, \xi) \geq 0$  for  $y_i \geq 0$ .

This assumption is standard. It requires that the production cost function of each firm be convex and sufficiently smooth, see [35] and references therein.

**Assumption 4** The inverse demand function  $p(q, \xi)$  satisfies the following conditions.

- (i)  $p(q, \xi)$  is twice continuously differentiable in  $q$  and  $p'_q(q, \xi) < 0$  for  $q \geq 0$  and  $\xi \in \Xi$ .
- (ii)  $p'_q(q, \xi) + qp''_{qq}(q, \xi) \leq 0$ , for  $q \geq 0$  and  $\xi \in \Xi$ .

This assumption is similar to an assumption used by Sherali, Soyster and Murphy [31] and De Wolf and Smeers [14]. Consider a monopoly market with an extraneous supply  $\bar{c} \geq 0$ . If the monopoly's output is  $q$ , then its revenue at demand scenario  $\epsilon(\omega) = \xi$  is  $q(p(q + \bar{c}, \xi))$ . The marginal revenue is  $p(q + \bar{c}, \xi) + qp'_q(q + \bar{c}, \xi)$ . The rate of change of this marginal revenue with respect to the increase in the extraneous supply  $\bar{c}$  is  $p'_q(q + \bar{c}, \xi) + qp''_{qq}(q + \bar{c}, \xi)$ . Assumption 4 (ii) implies that this rate is not positive when  $\bar{c} = 0$  for any  $\xi \in \Xi$ . In other words, any extraneous supply will potentially reduce the monopoly's marginal revenue in any demand scenario. See [31] for a similar explanation for a deterministic leader-followers' market. The following result is established by Xu [35].

**Proposition 5.1** *Under Assumption 4, the following assertions hold.*

- (i) For fixed  $\bar{c} \geq 0$ ,

$$p'_q(q + \bar{c}, \xi) + qp''_{qq}(q + \bar{c}, \xi) \leq 0, \text{ for } q \geq 0, \xi \in \Xi. \quad (5.4)$$

- (ii)  $qp(q + \bar{c}, \xi)$  is strictly concave in  $q$  for  $q \geq 0, \xi \in \Xi$ .

By Proposition 3, we know that problem (5.3) has a unique optimization solution (we are short of claiming unique equilibrium at this point) for each  $i$ . Moreover, we can write down the Karush-Kuhn-Tucker (KKT) conditions for (5.3) as follows:

$$0 \leq \begin{pmatrix} y_i \\ \mu_i \end{pmatrix} \perp \begin{pmatrix} -p(y_i + y_{-i}, \xi) - y_i p'_q(y_i + y_{-i}, \xi) + H'_i(y_i, \xi) + \mu_i \\ x_i - y_i \end{pmatrix} \geq 0, \quad (5.5)$$

where  $\mu_i$ ,  $i = 1, 2$ , are Lagrange multipliers of constraints  $y_i(\xi) \leq x_i$ ,  $i = 1, 2$ . Moreover, by Rosen [30, Theorem 1], the second stage Nash-Cournot game has an equilibrium which means the second stage complementarity problem (5.5) has a solution. The solution depends on  $x_1, x_2$  and  $\xi$ , we denote it by  $\bar{y}(x, \xi)$  and write  $x$  for  $(x_1, x_2)$ .

With the second stage equilibrium  $\bar{y}(x, \xi)$ , we are ready to write down the first stage decision making problem for player  $i$ :

$$\begin{aligned} \max_{x_i, y_i(\cdot)} \min_{P_i \in \mathcal{P}_i} \mathbb{E}_{P_i}[v_i(x, \xi)] - C_i(x_i) \\ \text{s.t.} \quad x_i \geq 0, \end{aligned} \quad (5.6)$$

where

$$v_i(x, \xi) := p(\bar{y}_i(x, \xi) + \bar{y}_{-i}(x, \xi), \xi) \bar{y}_i(x, \xi) - H_i(\bar{y}_i(x, \xi), \xi).$$

A 4-tuple  $(x_1^*, x_2^*, y_1^*(\cdot), y_2^*(\cdot))$  with  $(y_1^*(\cdot), y_2^*(\cdot)) = (\bar{y}_1(x^*, \cdot), \bar{y}_2(x^*, \cdot))$  is called a *two-stage distributionally robust equilibrium* if  $(x_i^*, x_{-i}^*)$  solves (5.6). Following Agassi and Bertsimas [2], we may introduce so-called *ex post equilibrium*  $(x_i^*, x_{-i}^*)$  which satisfies

$$x_i^* \in \arg \max \mathbb{E}_{P_i}[v_i(x_i, x_{-i}^*, \xi)] - C_i(x_i), \quad \forall P_i \in \mathcal{P}_i, \quad i = 1, 2. \quad (5.7)$$

The latter describes a situation where each player's response is optimal regardless of probability distributions (from the ambiguity sets). Ex post equilibrium may not exist but if it does, then it is a distributionally robust equilibrium. Given that ex post equilibrium is relatively easy to compute, it provides an attractive avenue for computing a distributionally robust equilibrium. In the forthcoming discussions, we show that the ex post equilibrium can be represented by the DRLCP model (1.2).

Assuming that  $C_i(x_i)$  is continuously differentiable, we may write down the first order optimality condition of (5.7):

$$0 \in \mathbb{E}_{P_i}[\partial_{x_i} v_i(x, \xi)] - C'_i(x_i) + \mathcal{N}_{[0, \infty)}(x_i), \quad \forall P_i \in \mathcal{P}_i, \quad i = 1, 2, \quad (5.8)$$

where  $\mathcal{N}_{[0, \infty)}(x_i)$  denotes the normal cone of interval  $[0, +\infty)$  at  $x_i$  and  $\partial v_i$  denotes the Clarke subdifferential of  $v_i$  with respect to  $x_i$ . Note that under Assumption 4, the second stage problem (5.3) has a unique solution and the set of Lagrange multipliers defined by the KKT system (5.5) is a singleton, it follows by Ralph and Xu [25, Lemma 5.2] that  $v(x, \xi)$  is continuously differentiable w.r.t.  $x_i$  for  $x_i > 0$  and

$$\nabla_{x_i} v_i(x, \xi) = \frac{L(y_i(\xi), \lambda_i(\xi), \mu_i(\xi), x_i)}{dx_i} = \mu_i(\xi),$$

where

$$L(y_i(\xi), \lambda_i(\xi), \mu_i(\xi), x_i) := p(y_1(\xi) + y_2(\xi), \xi) y_i(\xi) - H_i(y_i(\xi), \xi) + \lambda_i(\xi) y_i(\xi) - \mu_i(\xi) (y_i(\xi) - x_i).$$

Consequently, we can rewrite (5.8) as

$$0 \leq x_i \perp \mathbb{E}_{P_i}[\mu_i(\xi)] - C'_i(x_i) \geq 0, \quad \forall P_i \in \mathcal{P}_i, \quad i = 1, 2. \quad (5.9)$$

Let

$$g_i(y_i, y_{-i}, \xi) := -p(y_i + y_{-i}, \xi) - y_i p'_q(y_i + y_{-i}, \xi) + H'_i(y_i, \xi), \quad \text{for } i = 1, 2.$$

Note that in the case when  $x_i = 0$ ,  $y_i(\xi) \equiv 0$ , we have

$$\partial_{x_i} v_i(x, \xi) = \{\mu_i(\xi) : \mu_i(\xi) \geq (-g_i(0, 0, \xi))_+, \forall \xi \in \Xi \text{ and } \mathbb{E}_{P_i}[\mu_i(\xi)] \geq C'_i(x_i), \forall P_i \in \mathcal{P}_i\}.$$

Summarizing the discussions above, we can derive the following two-stage ex post complementarity problem

$$\begin{cases} 0 \leq x_i \perp \mathbb{E}_{P_i}[\mu_i(\xi)] - C'_i(x_i) \geq 0, & \forall P_i \in \mathcal{P}_i, \quad i = 1, 2, \\ 0 \leq \begin{pmatrix} y_i(\xi) \\ \mu_i(\xi) \end{pmatrix} \perp \begin{pmatrix} g_i(y_i(\xi), y_{-i}(\xi), \xi) + \mu_i(\xi) \\ x_i - y_i(\xi) \end{pmatrix} \geq 0, & \text{for } P_i\text{-a.e. } \xi \in \Xi, \quad i = 1, 2. \end{cases} \quad (5.10)$$

## 5.1 An example

To explain how to reformulate a two-stage duopoly game as a two-stage ex post complementarity problem (5.10), we consider a simple example where  $p(q, \xi) = a(\xi_1) - b(\xi_1)q + \xi_2$ ,

$$C_i(x_i) = \alpha_i + \beta_i x_i - \frac{1}{2} \gamma_i x_i^2 \quad \text{and} \quad H_i(y_i, \xi_1) = s_i(\xi_1) + \zeta_i(\xi_1) y_i + \frac{1}{2} \eta_i(\xi_1) y_i^2 \quad \text{for } i = 1, 2,$$

where  $\xi = (\xi_1, \xi_2)^T$  is a random vector with support set  $\Xi := [-1, 1] \times [-1, 1]$ ,  $\xi_1, \xi_2$  are independent,  $a(\xi_1), b(\xi_1), s_i(\xi_1), \zeta_i(\xi_1)$  and  $\eta_i(\xi_1)$  map from  $[-1, 1]$  to  $\mathbb{R}_{++}$ , and  $\alpha_i, \beta_i, \zeta_i, \eta_i$ , and  $\gamma_i$  are positive constants. The cost functions  $C_i(x_i)$ ,  $i = 1, 2$ , are concave which means the marginal cost for capacity set-up decreases for both players as capacity increases. We assume that the ratio  $\beta_i/\gamma_i$  is large enough so that the marginal capital cost does not become negative within any possible capacity that the two players may install. To ease the exposition, let

$$A := \begin{pmatrix} \gamma_1 & 0 \\ 0 & \gamma_2 \end{pmatrix}, \quad B := \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad N := B^T,$$

$$\Pi(\xi_1) := \begin{pmatrix} 2b(\xi_1) + \eta_1(\xi_1) & b(\xi_1) \\ b(\xi_1) & 2b(\xi_1) + \eta_2(\xi_1) \end{pmatrix} \quad \text{and} \quad M(\xi_1) := \begin{pmatrix} \Pi(\xi_1) & I_2 \\ -I_2 & 0 \end{pmatrix}.$$

We can write (5.10) in the following matrix-vector form:

$$\begin{cases} 0 \leq \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \perp A \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \mathbb{E}_P \left[ B \begin{pmatrix} y_1(\xi) \\ y_2(\xi) \\ \mu_1(\xi) \\ \mu_2(\xi) \end{pmatrix}^T \right] - \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} \geq 0, \quad \forall P \in \mathcal{P}, \\ 0 \leq \begin{pmatrix} y_1(\xi) \\ y_2(\xi) \\ \mu_1(\xi) \\ \mu_2(\xi) \end{pmatrix} \perp M(\xi_1) \begin{pmatrix} y_1(\xi) \\ y_2(\xi) \\ \mu_1(\xi) \\ \mu_2(\xi) \end{pmatrix} + N \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} - \begin{pmatrix} a(\xi_1) + \xi_2 - \zeta_1(\xi_1) \\ a(\xi_1) + \xi_2 - \zeta_2(\xi_1) \\ 0 \\ 0 \end{pmatrix} \geq 0, \quad \forall \xi \in \Xi. \end{cases} \quad (5.11)$$

Note that matrix  $M(\xi_1)$  is positive semidefinite and nonsingular for all  $\xi \in \Xi$ .

In the stochastic case when  $\mathcal{P}_i$  reduces to a singleton, that is, the true probability distribution, the two-stage distributionally robust game (5.2) collapses to two-stage stochastic game (5.1) and consequently (5.11) can be written as a two-stage SLCP,

$$\begin{cases} 0 \leq \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \perp A \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \mathbb{E} \left[ B(y_1(\xi), y_2(\xi), \mu_1(\xi), \mu_2(\xi))^T \right] - \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} \geq 0, \\ 0 \leq \begin{pmatrix} y_1(\xi) \\ y_2(\xi) \\ \mu_1(\xi) \\ \mu_2(\xi) \end{pmatrix} \perp M(\xi_1) \begin{pmatrix} y_1(\xi) \\ y_2(\xi) \\ \mu_1(\xi) \\ \mu_2(\xi) \end{pmatrix} + N \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} - \begin{pmatrix} a(\xi_1) + \xi_2 - \zeta_1(\xi_1) \\ a(\xi_1) + \xi_2 - \zeta_2(\xi_1) \\ 0 \\ 0 \end{pmatrix} \geq 0, \forall \xi \in \Xi. \end{cases} \quad (5.12)$$

By [30, Theorem 2], the second stage game (5.3) has a unique equilibrium at each scenario  $\xi \in \Xi$ , which means that the  $(y_1(\xi), y_2(\xi))$ -components of the solutions to the second stage complementarity problem is unique for each  $\xi$ . The Lagrange multiplies  $\mu_i(\xi)$ ,  $i = 1, 2$  are also unique when  $x_1, x_2 > 0$ .

We then consider the distributionally robust case that the true probability distribution is unknown and a distributionally robust game described as in (5.2) is played with the ambiguity set being defined through moment conditions  $\mathcal{P} := \{P \in \mathcal{P} : \mathbb{E}_P[\xi] = 0\}$ .

For the simplicity of analysis, we assume that both players use the same ambiguity set, that is  $\mathcal{P}_1 = \mathcal{P}_2 = \mathcal{P}$ . Moreover, we set  $\beta_1 = \beta_2 = \beta$  and  $\zeta_1(\xi_1) = \zeta_2(\xi_1) = 1$  in (5.11). Observe that (5.11) may have a trivial solution with  $x_1 = x_2 = 0$ ,  $y_1(\xi) = y_2(\xi) = 0$  for all  $\xi \in \Xi$  and

$$\mu_i(\xi) \geq (a(\xi_1) + \xi_2 - 1)_+, \quad \forall \xi \in \Xi \text{ and } \mathbb{E}_P[\mu_i(\xi)] \geq \beta, \quad \forall P \in \mathcal{P}, \text{ for } i = 1, 2.$$

In what follows, we concentrate on non-trivial solutions with both  $x_1$  and  $x_2$  being positive. We endeavour to obtain an analytical solution to (5.11). For this purpose, we simply assume that  $a, b, s_i$  are deterministic positive numbers and  $\eta_i(\xi_1) = \bar{\eta}_i + \xi_1$ , where  $\bar{\eta}_i \geq 1$ .

By (4.17), the dual formulation of (5.11) can be written as

$$\begin{cases} x_1, x_2 \geq 0, \\ \left[ -A \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} - B(y_1(\xi), y_2(\xi), \mu_1(\xi), \mu_2(\xi))^T + \begin{pmatrix} \beta \\ \beta \end{pmatrix} - \mathbf{\Lambda}_1 \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} \right] \leq 0, \forall \xi \in \Xi, \\ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}^T \left[ A \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + B(y_1(\xi), y_2(\xi), \mu_1(\xi), \mu_2(\xi))^T - \begin{pmatrix} \beta \\ \beta \end{pmatrix} - \mathbf{\Lambda}_2 \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} \right] \leq 0, \forall \xi \in \Xi, \\ 0 \leq \begin{pmatrix} y_1(\xi) \\ y_2(\xi) \\ \mu_1(\xi) \\ \mu_2(\xi) \end{pmatrix} \perp M(\xi_1) \begin{pmatrix} y_1(\xi) \\ y_2(\xi) \\ \mu_1(\xi) \\ \mu_2(\xi) \end{pmatrix} + N \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} - \begin{pmatrix} a + \xi_2 - 1 \\ a + \xi_2 - 1 \\ 0 \\ 0 \end{pmatrix} \geq 0, \forall \xi \in \Xi, \end{cases} \quad (5.13)$$

where  $\mathbf{\Lambda}_i = \begin{pmatrix} \lambda_{11}^i & \lambda_{12}^i \\ \lambda_{21}^i & \lambda_{22}^i \end{pmatrix} \in \mathbb{R}^{2 \times 2}$ . Since  $\Xi$  is a compact set and the underlying functions are continuous, problems (5.11) and (5.13) are equivalent in that there is no dual gap in deriving the Lagrange dual of maximization with respect to  $P$ , see [32, page 208]. Let  $\tilde{y}(\xi) := (\tilde{y}_1(\xi), \tilde{y}_2(\xi))^T = \Pi(\xi_1)^{-1} \begin{pmatrix} a + \xi_2 - 1 \\ a + \xi_2 - 1 \end{pmatrix}$ . Note that  $\Pi(\xi_1)^{-1}$  is positive definite and diagonally dominant, therefore  $\tilde{y}(\xi) > 0$  when  $a + \xi_2 - 1 > 0$ . The following example proposes a way to choose  $(a, b, \bar{\eta}_1, \bar{\eta}_2, \gamma_1, \gamma_2, \beta)$  such that (5.13) has a solution.

**Example 5.1** Choose  $(a, b, \bar{\eta}_1, \bar{\eta}_2, \gamma_1, \gamma_2, \beta)$  with  $a > 2$  such that  $z = \hat{\Pi}^{-1} \begin{pmatrix} a - \beta - 1 \\ a - \beta - 1 \end{pmatrix}$  satisfies  $0 < z_i \leq \inf_{\xi \in \Xi} \tilde{y}_i(\xi)$ , for  $i = 1, 2$ , where

$$\hat{\Pi} := \begin{pmatrix} 2b + \bar{\eta}_1 - \gamma_1 & b \\ b & 2b + \bar{\eta}_2 - \gamma_2 \end{pmatrix}.$$

Then we can show that problem (5.13) has a solution  $(x_1, x_2, y_1(x_1, \cdot), y_2(x_2, \cdot), \mu_1(x, \cdot), \mu_2(x, \cdot), \mathbf{\Lambda})$  with  $x_i = y_i(z_i, \xi) = z_i$ ,

$$\mu_i(z, \xi) = a + \xi_2 - 1 - (2b + \bar{\eta}_i + \xi_1)z_i - bz_{-i} \quad \text{and} \quad \mathbf{\Lambda}_1 = -\mathbf{\Lambda}_2 = \begin{pmatrix} \lambda_{11} & \lambda_{12} \\ \lambda_{21} & \lambda_{22} \end{pmatrix} = \begin{pmatrix} z_1 & -1 \\ z_2 & -1 \end{pmatrix} \quad (5.14)$$

for all  $\xi_1, \xi_2 \in \Xi$  and  $i = 1, 2$ .

To see this, we consider second stage complementarity problem of (5.13) (forth equation of (5.13)). It is not difficult to verify that when  $x_i \leq \tilde{y}_i(\xi)$ , for all  $\xi \in \Xi$  and  $i = 1, 2$ , the forth equation of (5.13) has a solution  $y_i(x_i, \xi) = x_i$ , for  $i = 1, 2$  with

$$\begin{aligned} \begin{pmatrix} \mu_1(x, \xi) \\ \mu_2(x, \xi) \end{pmatrix} &= -\Pi(\xi_1) \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} a + \xi_2 - 1 \\ a + \xi_2 - 1 \end{pmatrix} = \Pi(\xi_1)(\tilde{y}(\xi) - x) \\ &= \begin{pmatrix} a + \xi_2 - 1 - (2b + \bar{\eta}_i + \xi_1)x_1 - bx_2 \\ a + \xi_2 - 1 - (2b + \bar{\eta}_i + \xi_1)x_1 - bx_2 \end{pmatrix} \geq 0, \quad \forall \xi \in \Xi. \end{aligned}$$

In what follows, we show that  $(z_1, z_2, y_1(z_1, \xi), y_2(z_2, \xi), \mu_1(z, \xi), \mu_2(z, \xi))$  satisfies the first equation of (5.13) with  $\mathbf{\Lambda}_1 = -\mathbf{\Lambda}_2 = \begin{pmatrix} z_1 & -1 \\ z_2 & -1 \end{pmatrix}$ . With the explicit formulation of  $\mu_i(x, \xi)$  from the second stage of (5.13), the first three equations in (5.13) can be rewritten as

$$\begin{cases} x_1, x_2 \geq 0, \\ \begin{bmatrix} -(\gamma_1 - (2b + \bar{\eta}_1 + \xi_1) & -b \\ -b & \gamma_2 - (2b + \bar{\eta}_2 + \xi_1) \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} - \begin{pmatrix} a + \xi_2 - \beta - 1 \\ a + \xi_2 - \beta - 1 \end{pmatrix} - \mathbf{\Lambda}_1 \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} \leq 0 \\ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}^T \begin{bmatrix} \gamma_1 - (2b + \bar{\eta}_1 + \xi_1) & -b \\ -b & \gamma_2 - (2b + \bar{\eta}_2 + \xi_1) \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} a + \xi_2 - \beta - 1 \\ a + \xi_2 - \beta - 1 \end{pmatrix} - \mathbf{\Lambda}_2 \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} \leq 0 \end{cases} \quad (5.15)$$

for all  $\xi \in \Xi$ . When  $x_1, x_2 > 0$ , we can obtain a solution to (5.15) by solving

$$\begin{cases} \hat{\Pi} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} a - \beta - 1 \\ a - \beta - 1 \end{pmatrix}, \\ \xi_2 - x_1 \xi_1 + \lambda_{11} \xi_1 + \lambda_{12} \xi_2 = 0, \quad \forall \xi \in \Xi, \\ \xi_2 - x_2 \xi_1 + \lambda_{21} \xi_1 + \lambda_{22} \xi_2 = 0, \quad \forall \xi \in \Xi, \end{cases} \quad (5.16)$$

whereby from the first equation we have  $x = \hat{\Pi}^{-1} \begin{pmatrix} a - \beta - 1 \\ a - \beta - 1 \end{pmatrix} = z$  and  $\lambda_{i1} = z_i$ ,  $\lambda_{i2} = -1$  for  $i = 1, 2$  for the second equation.

It is worth noting that it is easy to choose  $(a, b, \bar{\eta}_1, \bar{\eta}_2, \gamma_1, \gamma_2, \beta)$  such that  $0 < z_i \leq \inf_{\xi \in \Xi} \tilde{y}_i(\xi)$ , for  $i = 1, 2$ . For example, when  $\bar{\eta}_i \geq 1$ ,  $i = 1, 2$ , we have

$$\begin{aligned} \begin{pmatrix} \inf_{\xi \in \Xi} \tilde{y}_1(\xi) \\ \inf_{\xi \in \Xi} \tilde{y}_2(\xi) \end{pmatrix} &= \Pi(1)^{-1} \begin{pmatrix} a-2 \\ a-2 \end{pmatrix} = \frac{1}{\det(\Pi(1))} \begin{pmatrix} 2b + \bar{\eta}_2 + 1 & -b \\ -b & 2b + \bar{\eta}_1 + 1 \end{pmatrix} \begin{pmatrix} a-2 \\ a-2 \end{pmatrix} \\ &= \frac{1}{\det(\Pi(1))} \begin{pmatrix} (b + \bar{\eta}_2 + 1)(a-2) \\ (b + \bar{\eta}_1 + 1)(a-2) \end{pmatrix} \end{aligned}$$

and

$$z = \frac{1}{\det(\hat{\Pi})} \begin{pmatrix} 2b + \bar{\eta}_2 - \gamma_2 & -b \\ -b & 2b + \bar{\eta}_1 - \gamma_1 \end{pmatrix} \begin{pmatrix} a - \beta - 1 \\ a - \beta - 1 \end{pmatrix} = \frac{1}{\det(\hat{\Pi})} \begin{pmatrix} (b + \bar{\eta}_2 - \gamma_2)(a - \beta - 1) \\ (b + \bar{\eta}_1 - \gamma_1)(a - \beta - 1) \end{pmatrix}.$$

Moreover, when

$$\frac{(a-2)(\bar{\eta}_i + 1 + b)}{(a-\beta-1)(\bar{\eta}_i + b - \gamma_i)} \geq \frac{\det(\Pi(1))}{\det(\hat{\Pi})}, \quad \text{for } i = 1, 2, \quad (5.17)$$

we have  $z_i \leq \inf_{\xi \in \Xi} \tilde{y}_i(\xi)$ ,  $i = 1, 2$ .

We now move on to verify the discretization scheme discussed in Section 4 for (5.13). To this end, we set particular values for the underlying parameters in the table below:

Parameters	$\alpha_1$	$\beta_1$	$\gamma_1$	$s_1$	$\zeta_1$	$\bar{\eta}_1$	$a$	$\alpha_2$	$\beta_2$	$\gamma_2$	$s_2$	$\zeta_2$	$\bar{\eta}_2$	$b$
Values	3	5	1	2	1	1	10	3	5	0.5	2	1	2	5

Let  $\Xi_K := \{\xi^1, \dots, \xi^K\}$  be i.i.d. samples of  $\xi$ , where  $\xi = (\xi_1, \xi_2)$ ,  $\xi_i$  follow truncated normal distribution over  $[-1, 1]$  which is constructed from normal distribution with mean 0 and standard deviation  $\sigma$  independently,  $i = 1, 2$ . We carry out numerical experiments with different values of  $\sigma$ . With the specified parameter values, condition (5.17) is satisfied and consequently all conditions in Example 5.1 are fulfilled. This means we are able to obtain a solution of (5.13) with  $x_1 = y_1(\xi) = 0.2844$ ,  $x_2 = y_2(\xi) = 0.2222$  and

$$\mu_1(\xi) = 4.7111 + \xi_2 - 0.2844\xi_1, \quad \text{and} \quad \mu_2(\xi) = 4.8889 + \xi_2 - 0.2222\xi_1, \quad \forall \xi \in \Xi.$$

Let us now apply the discrete scheme to (5.13). The solution of the discretized problem is  $x_1^K = (\mathbf{y}_j^K)_1 = 0.2844$ ,  $x_2^K = (\mathbf{y}_j^K)_2 = 0.2222$  and

$$\mu_1^K(\xi^j) = 4.7111 + \xi_2^j - 0.2844\xi_1^j, \quad \text{and} \quad \mu_2^K(\xi^j) = 4.8889 + \xi_2^j - 0.2222\xi_1^j, \quad \text{for } j \in \bar{K}.$$

Since  $x^K = x$  and  $y^K(\xi) = y(\xi)$  for all  $\xi \in \Xi$ , where  $y^K(\xi) = \sum_{j=1}^K \mathbf{y}_j^K \mathbf{1}_{\Xi_j^K}(\xi)$ ,  $\Xi_j^K$  is defined by the Voronoi partition in (3.25). We only need to investigate the error of  $\mu^K(\xi^j)$ , for  $j \in \bar{K}$ . Define the errors of two components of  $\mu^K$  obtained from the discretized problem by

$$\text{error}_i^K = \mathbb{E}[|\mu_i(\xi) - \bar{\mu}_i^K(\xi)|], \quad \text{where} \quad \bar{\mu}_i^K(\xi) = \sum_{j=1}^K \mu_i^K(\xi^j) \mathbf{1}_{\Xi_j^K}(\xi), \quad \text{for } i = 1, 2,$$

and the mathematical expectation is taken with respect to the distribution of  $\xi$  (truncated normal distribution).

Note that it is not easy to calculate  $\text{error}_i^K$  directly. Therefore we propose to use sample average approximation method to estimate the quantity, that is, generate iid samples  $\hat{\xi}^1, \dots, \hat{\xi}^N$  with sample size  $N = 5000$ , and calculate

$$\text{error}_i^K \approx \text{error}_{iN}^K = \frac{1}{N} \sum_{k=1}^N |\mu_i(\hat{\xi}^k) - \bar{\mu}_i^K(\hat{\xi}^k)|, \text{ for } i = 1, 2.$$

Here we are using notation  $\hat{\xi}^k$  to distinguish the samples from those in  $\Xi_K$ . We carried out tests with sample sizes  $K = 5, 10, 20, 40, 60, 100$  and the standard deviation  $\sigma = 0.1, 0.5, 1, 10$  of the normal distribution. For each fixed  $K$  and  $\sigma$ , we generate  $K$  samples  $\Xi_K$ , calculate the  $\text{error}_{iN}^K$  100 times and average them. Figures 1-2 depict the decreasing tendency of  $\text{error}_{iN}^K$  as  $K$  increases and  $\sigma$  decreases.

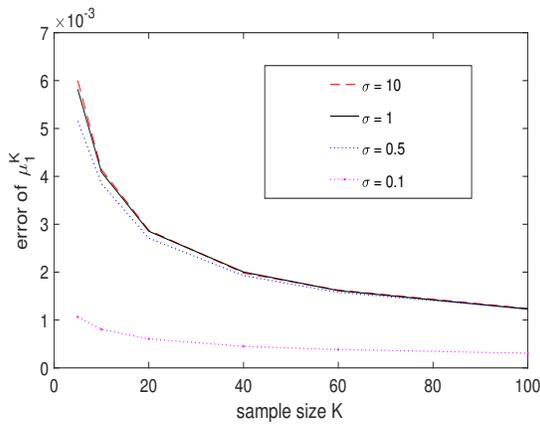


Figure 1: error of  $\mu_1^K$

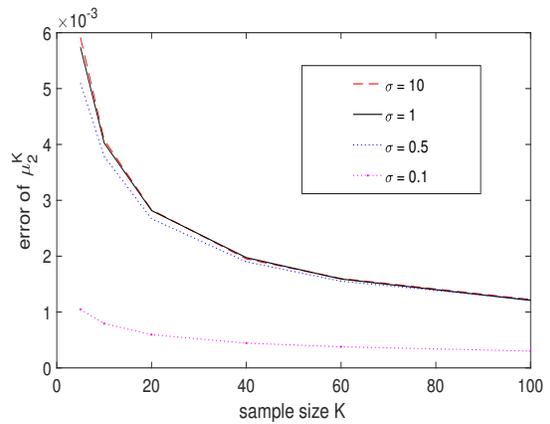


Figure 2: error of  $\mu_2^K$

## References

- [1] E. Anderson, H. Xu and D. Zhang, Varying confidence levels for CVaR risk measures and minimax limits, manuscript (2014)
- [2] M. Aghassi and D. Bertsimas, Robust game theory, *Math. Program.*, 107, 231-273 (2006)
- [3] R. J. Aumann, Integrals of set-valued functions, *J. Math. Anal. Appl.*, 12, 1-12 (1965)
- [4] J.R. Birge and F Louveaux, *Introduction to Stochastic Programming*, Springer Science & Business Media (2011)
- [5] X. Chen and M. Fukushima, Expected residual minimization method for stochastic linear complementarity problems, *Math. Oper. Res.*, 30, 1022-1038 (2005)
- [6] X. Chen, T.K. Pong and R. Wets, Two-stage stochastic variational inequalities: an ERM-solution procedure, *Math. Program.*, 165, 71-111 (2017)
- [7] X. Chen, R. Wets and Y. Zhang, Stochastic variational inequalities: residual minimization smoothing/sample average approximations, *SIAM J. Optim.*, 22, 649-673 (2012)
- [8] X. Chen and S. Xiang, Computation of error bounds for P-matrix linear complementarity problem, *Math. Program.*, 106, 513-525 (2006)

- [9] X. Chen and S. Xiang, Perturbation bounds of P-matrix linear complementarity problems, *SIAM J. Optim.*, 18, 1250-1265 (2007)
- [10] X. Chen and S. Xiang, Newton iterations in implicit time-stepping scheme for differential linear complementarity systems, *Math. Program.*, 138, 579-606 (2013)
- [11] F. H. Clarke, *Optimization and nonsmooth analysis*, Wiley, New York (1983)
- [12] R.W. Cottle, J.-S. Pang and R. E. Stone, *The Linear Complementarity Problem*, Academic Press, New York (1992)
- [13] J. Crémer and R. McLean, Optimal selling strategies under uncertainty for a discriminating monopolist when demands are interdependent, *Econometrica*, 53, 345-362 (1985)
- [14] D. De Wolf and Y. Smeers, A stochastic version of a Stackelberg-Nash-Cournot equilibrium model, *Manage. Sci.*, 43, 190-197 (1997)
- [15] F. Facchinei and J.-S. Pang, *Finite-Dimensional Variational Inequalities and Complementarity Problems*, Springer-Verlag, New York (2003)
- [16] G. Gürkan, A.Y. Özge and S.M. Robinson, Sample-path solution of stochastic variational inequalities, *Math. Program.*, 84, 313-333 (1999)
- [17] N. Hadjisavvas, S. Komlósi and S. Schaible, *Handbook of Generalized Convexity and Generalized Monotonicity*, Springer, Boston (2005)
- [18] B. Holmström and R. Myerson, Efficient and durable decision rules with incomplete information, *Econometrica*, 51, 1799-1820 (1983)
- [19] H. Jiang and H. Xu, Stochastic approximation approaches to the stochastic variational inequality problem, *IEEE Trans. Autom. Control*, 53, 1462-1475 (2008)
- [20] Y. Liu, A. Pichler and H. Xu, Discrete approximation and quantification in distributionally robust optimization, to appear in *Math. Oper. Res.*, (2017)
- [21] Y. Liu, H. Xu, S. Yang and J. Zhang, Distributionally robust equilibrium for continuous games: Nash and stackelberg models, *Eur. J. Oper. Res.*, 265, 631-643 (2018)
- [22] J.P. Luna, C. Sagastizbal and M. Solodov, An approximation scheme for a class of risk-averse stochastic equilibrium problems, *Math. Program.*, 157, 451-481 (2016)
- [23] G. Ch. Pflug and A. Pichler. *Multistage Stochastic Optimization*, *Springer Series in Operations Research and Financial Engineering*, Springer, New York (2014)
- [24] D. Ralph and Y. Smeers, Pricing risk under risk measures: an introduction to stochastic-endogenous equilibria, Social Science Research Network (2011) URL: [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=1903897](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1903897)
- [25] D. Ralph and H. Xu, Convergence of stationary points of sample average two stage stochastic programs: a generalized equation approach, *Math. Oper. Res.*, 36, 568-592 (2011)
- [26] R.T. Rockafellar, *Convex Analysis*, Princeton University Press, Princeton (1970)
- [27] R.T. Rockafellar and R. J-B Wets, Scenarios and policy aggregation in optimization under uncertainty, *Math. Oper. Res.*, 16, 119-147 (1991)
- [28] R.T. Rockafellar and R. J-B Wets, Stochastic variational inequalities: single-stage to multistage, *Math. Program.*, 165, 331-360 (2017)
- [29] R.T. Rockafellar and J. Sun, Solving monotone stochastic variational inequalities and complementarity problems by progressive hedging, to appear in *Math. Program.* (2017)

- [30] J.B. Rosen, Existence and uniqueness of equilibrium points for concave  $N$ -person games, *Econometrica*, 33, 520–534 (1965)
- [31] H.D. Sherali, A. L. Soyster and F.H. Murphy, Stackelberg-Nash-Cournot equilibria: characterizations and computations, *Oper. Res.*, 31, 253-276 (1983)
- [32] A. Shapiro, D. Dentcheva and A. Ruszczyński, *Lectures on Stochastic Programming: Modeling and Theory*, SIAM, Philadelphia (2009)
- [33] S. Wogrin, B.F. Hobbs, D. Ralph, E. Centeno and J. Barquin, Open versus closed loop capacity equilibria in electricity markets under perfect and oligopolistic competition, *Math. Program.*, 140, 295-322 (2013)
- [34] H. Xu, Uniform exponential convergence of sample average random functions under general sampling with applications in stochastic programming, *J. Math. Anal. Appl.*, 368, 692-710 (2010)
- [35] H. Xu, An MPCC approach for stochastic Stackelberg-Nash-Cournot Equilibrium, *Optimization*, 54, 27-57 (2005)
- [36] H. Xu, Y. Liu and S. Sun, Distributionally robust optimization with matrix moment constraints: Lagrange duality and cutting plane methods, to appear in *Math. Program.* (2017)
- [37] H. Xu, Sample average approximation methods for a class of stochastic variational inequality problems, *Asia Pac. J. Oper. Res.*, 27, 103-119 (2010)
- [38] H. Xu, An implicit programming approach for a class of stochastic mathematical programs with complementarity constraints, *SIAM J. Optim.*, 16, 670-696 (2006)