Are Weaker Stationarity Concepts of Stochastic MPCC Problems Significant in Absence of SMPCC-LICQ?

Arnab Sur*

Abstract

In this article, we study weak stationarity conditions (A- and C-) for a particular class of degenerate stochastic mathematical programming problems with complementarity constraints (SMPCC, for short). Importance of the weak stationarity concepts in absence of SMPCC-LICQ are presented through toy problems in which the point of local or global minimum are weak stationary points rather than satisfying other stronger stationarity conditions. Finally, a well known technique to solve stochastic programming problems, namely sample average approximation (SAA) method, is studied to show the significance of the weak stationarity conditions for degenerate SMPCC problems. Consistency of weak stationary estimators are established under weaker constraint qualifications than SMPCC-LICQ.

Key Words: Degenerate Stochastic mathematical programming with complementarity constraints, A-stationarity , C-stationarity, Sample average approximation (SAA) method, Consistency.

*The University of Chicago, IL 60637, USA.
email: arnabsur2002@gmail.com.

1 Introduction

In this article we consider the following one stage stochastic mathematical programming problem with complementarity constraints (SMPCC, for short):

$$\begin{aligned} &\min_{x\in\mathbb{R}^n} \qquad E[f(x,\xi)] \\ &0 \le E[G(x,\xi)] \perp E[H(x,\xi)] \ge 0; \end{aligned} \tag{1.1}$$

where $f : \mathbb{R}^n \times \Xi \to \mathbb{R}$, $H, G : \mathbb{R}^n \times \Xi \to \mathbb{R}^l$ are continuously differentiable ω-a.e. and $\xi : \Omega \to \Xi$, where $\Xi \subset \mathbb{R}^m$ is the support set of the random variable and $a \perp b$ means that vector a is perpendicular to vector b. π is the probability distribution associated to ξ and expectations are calculated with respect to π. One-stage SMPCC model was first introduced by Birbil, Gürkan and Listes in [1] with an application related to toll pricing in a transportation network. The above problem can be considered as the generalization of the model introduced in [1].

The basic difficulty to solve stochastic optimization problems involving expectation functionals is that to evaluate these expectations in a closed form. Sample average approximation (SAA) method is a well known technique to approximate these functionals with high accuracy. the SAA method was studied by Robinson [17], Gürkan, Özge and Robinson [9] under the name of sample path optimization. The term "sample average approximation" was coined by Kleywegt, Shapiro and Homem-de-Mello [14]. Later, Shapiro [22] studied this method for stochastic optimization problems under equilibrium constraints. [11] presents an overview on how to solve stochastic optimization problems with expectation functional using Monte Carlo sampling based algorithms including SAA. A huge number of references on SAA have been included there. All of them either studied the large sample properties of the SAA estimators, or developed the mechanisms to establish the large sample properties. Consistency of the SAA estimators is one of the large sample properties which has been studied in a detailed fashion.

In past couple of decades study of mathematical programming problems with complementarity constraints (MPCC, for short) has been the subject of intense research. This class of optimization problems are inherently non-convex and the constraint qualifications, like LICQ, MFCQ fail for these problems. These posed a great challenge to the researchers to tackle those problems efficiently and develop MPCC variant of different constraint qualifications. So, they reformulated the MPCC model in different equivalent non-linear programming problems to obtain different stationary concepts, like strong, M-, C- and A-stationarity, by applying Fritz John necessary optimality conditions under various MPCC tailored constraint qualifications; for references, see [5], [6], [18], [29]. The stochastic version of mathematical programming problem with variational inequality was first studied by Patriksson and Wynter [15]. It was subsequently investigated in details by Ye [28], Evgrafov and Patriksson [4], Shapiro [22], Ralph and Xu [16], Xu and Ye [27], Xu [26] and then complementarity as a special case of stochastic VI. Fukushima and Lin [8] discussed different kinds of models in their survey paper. One can find applications of stochastic MPCC problems in economics, electricity markets, traffic networks, management sciences; for example, see Christiansen, Patriksson and Wynter [3], Werner [24], Werner and Wang [25], Chen, Wets and Zhang [2], Henrion and Römisch [10] and the references therein.

In the literature of MPCC problems, the concept of C-stationarity was introduced by Scheel and Scholtes in [18]. Scholtes [19] later introduced an algorithm which under certain moderate assumptions on MPCC problems converges to a C-stationary point. Recently, Hoheisel, Kanzow and Schwartz [12] showed that certain relaxation schemes under some assumptions on MPCC problems lead us to C-stationarity. A-stationarity was introduced by Flegel and Kanzow in [5] and in the same article authors derived strong stationarity for MPCC problems under MPCC-LICQ.

The importance of further investigating C- and A-stationarity for the deterministic MPCC problems has become a point of contention. As it has been shown that all these stationarity concepts except strong stationarity do not preclude the existence of feasible first order descent directions if MPCC-LICQ holds at those points [18]. Hence, in the presence of MPCC-LICQ, C- and A-stationary points are redundant solutions. Moreover, it has been shown that MPCC-LICQ is a generic property of MPCC problems, see [20]. That is, if MPCC-LICQ is violated for a particular MPCC problem, it can always be satisfied by small perturbations of that problem.

In this article, we study degenerate SMPCC problems where the stochastic MPCC-LICQ fails. Degenerate condition may arise often in practice. Any SMPCC with a linear component would have degeneracy as every linear programming problem in practice eventually has a degenerate solution. The objective is to show the importance of C- and A-stationarity when degeneracy is studied separately. In this article, examples show that C- or A-stationary (which are not M-stationary) points could be the global minimum of a degenerate SMPCC problem. Moreover, consistency of C- and A-stationary points will be shown, under weaker constraint qualifications, to establish the importance of these stationary conditions in absence of stochastic MPCC-LICQ.

In [23], the author showed that the strong stationary estimators are inconsistent even under the assumption of stochastic MPCC-LICQ, i.e., the limiting solution of a sequence of strong stationary points of the SAA subproblems may not satisfy the strong stationary properties. Therefore, strong stationary points are not always useful even though they preclude all first order descent directions. On the contrary, if we can pose a stochastic MPCC problem in an equivalent way where the feasible set remains the same and the strong stationary points become C- or A-stationary points of the equivalent version, then we would be able to exploit the consistency of C- or A-stationary estimators to solve the stochastic MPCC problem. In other words, some form of weak stationarity is required in case of degeneracy and this paper shows how that can be obtained. Two examples are formulated later in this paper where LICQ fails to hold and C- or A-stationary points become the global minimum. In such a situation SAA can be employed to solve the SMPCC problem as the estimators are consistent.

The paper is organized as follows. Section **2** deals with C- and A-stationarity under weaker constraint qualifications than stochastic MPCC-LICQ. Toy problems are constructed to show the significance of C- and A-stationarity when stochastic MPCC-LICQ fails. In those examples, we observe that the global minimum are C- and A-stationary point respectively which do not satisfy other stronger stationarity conditions. In Section **3**, we discuss the SAA method to approximate C- and A-stationary points of a degenerate SMPCC problem. We study the consistency of C- and A-stationary estimators to establish the importance of these stationarity conditions in the context of degenerate SMPCC problems. Section **4** concludes the paper.

2 Stationarity Concepts and Constraint Qualifications

We begin with the definitions of C- and A-stationary concepts of the SMPCC problem (1.1) and related constraint qualifications.

Definition 2.1. A feasible solution x^* of the SMPCC problem (1.1) is said to satisfy the C-type Fritz John conditions if there exist scalars $\bar{\lambda}_0 \geq 0$, $\bar{\lambda}_i^G \in \mathbb{R}, i = 1, ...l$ and $\bar{\lambda}_i^H \in \mathbb{R}, i = 1, ...l$ such that

i) $\bar{\lambda}_0 E[\nabla_x f(x^*, \xi)] = \sum_{i=1}^{l} \bar{\lambda}_i^G E[\nabla_x G_i(x^*, \xi)] + \sum_{i=1}^{l} \bar{\lambda}_i^H E[\nabla_x H_i(x^*, \xi)]$;

ii) $(\bar{\lambda}_0, \bar{\lambda}^G, \bar{\lambda}^H) \neq 0$;

iii) $\bar{\lambda}_\alpha^G \in \mathbb{R}$, $\bar{\lambda}_\gamma^H \in \mathbb{R}$, $\bar{\lambda}_\gamma^G = 0$, $\bar{\lambda}_\alpha^H = 0$ and $\bar{\lambda}_i^G \bar{\lambda}_i^H \geq 0$, $\forall\ i \in \beta$;

where

$$\alpha := \alpha(x^*) := \{i : E[G_i(x^*, \xi)] = 0, E[H_i(x^*, \xi)] > 0\},$$

$$\beta := \beta(x^*) := \{i : E[G_i(x^*, \xi)] = 0, E[H_i(x^*, \xi)] = 0\},$$

$$\gamma := \gamma(x^*) := \{i : E[G_i(x^*, \xi)] > 0, E[H_i(x^*, \xi)] = 0\}.$$

If $\bar{\lambda}_0 = 1$ (after normalizing), the above condition is known as C-stationary conditions and the corresponding Lagrange multipliers are C-stationary multipliers.
The above condition will give us A-stationarity if we replace (*iii*) by $\bar{\lambda}_\alpha^G \in \mathbb{R}$, $\bar{\lambda}_\gamma^H \in \mathbb{R}$, $\bar{\lambda}_\gamma^G = 0$, $\bar{\lambda}_\alpha^H = 0$ and $\bar{\lambda}_i^G \geq 0$ or $\bar{\lambda}_i^H \geq 0$, $\forall\ i \in \beta$; and will give strong stationarity if we replace (*iii*) by $\bar{\lambda}_\alpha^G \in \mathbb{R}$, $\bar{\lambda}_\gamma^H \in \mathbb{R}$, $\bar{\lambda}_\gamma^G = 0$, $\bar{\lambda}_\alpha^H = 0$ and $\bar{\lambda}_i^G \geq 0$ and $\bar{\lambda}_i^H \geq 0$, $\forall\ i \in \beta$. It is evident that strong stationary implies C- or A- stationary. The following constraint qualifications are motivated from the definitions of the stationarity conditions.

Definition 2.2. C-type SMPCC-No Non-zero Abnormal Multiplier constraint qualification (SMPCC-NNAMCQ for short) is said to satisfy at a feasible point x^* if there exist no non-zero scalars $\lambda^G \in \mathbb{R}^l$ and $\lambda^H \in \mathbb{R}^l$, such that

(*i*) $0 = \sum_{i=1}^{l} \lambda_i^G E[\nabla_x G_i(x^*, \xi)] + \sum_{i=1}^{l} \lambda_i^H E[\nabla_x H_i(x^*, \xi)]$.

(*ii*) $\lambda_i^G = 0$, if $i \in \gamma$; $\lambda_i^H = 0$, if $i \in \alpha$ and $\lambda_i^G \lambda_i^H \geq 0$, $\forall\ i \in \beta$.

The above definition will be called A-type SMPCC-NNAMCQ if (*ii*) is replaced by $\lambda_i^G = 0$, if $i \in \gamma$; $\lambda_i^H = 0$, if $i \in \alpha$ and $\lambda_i^G \geq 0$ or $\lambda_i^H \geq 0$, $\forall\ i \in \beta$. SMPCC-LICQ holds at x^* if no non-zero scalars satisfy (*i*). It is clear from the definitions that SMPCC-LICQ is more stringent than the others.

It is well known in the literature of deterministic MPCC problems that under the assumption of MPCC-LICQ, any solution of an MPCC problem is a strong stationary point. Hence, in presence of MPCC-LICQ other stationarities, like C- and A-stationarity are not relevant solutions as they do not preclude the existence of feasible first order descent directions, for example see [18]. The other stationarities are important when MPCC-LICQ fails. Absence of MPCC-LICQ is a result of a degenerate solution as linear independence is a generic property of MPCC problems, see [20]. C and A-stationarity play a key role in the literature on stochastic MPCC problems while the degeneracy is discussed separately. This fact is illustrated by the following examples.

Example 2.1. $\xi : \Omega \to \mathbb{R}$ is a random variable which follows the normal distribution with mean μ and variance σ^2, where $\mu > 1$. Further assume that $f(x, \xi) = (x_1 - 1)^2 + \xi$, hence $E[f(x, \xi)] = (x_1 - 1)^2 + \mu$, $g(x) = x_1^3$, $G_1(x, \xi) = x_1 - x_2\xi$, so $E[G_1(x, \xi)] = x_1 - x_2\mu$ and $H_1(x, \xi) = (x_1^2 + x_2)\xi$, then $E[H_1(x, \xi)] = (x_1^2 + x_2)\mu$. The SMPCC problem then is the following

$$\begin{aligned} &\min_{x\in\mathbb{R}^2} \quad (x_1 - 1)^2 + \mu \\ &\text{subject to} \quad x_1^3 \le 0, \\ &\qquad 0 \le (x_1 - x_2\mu) \perp (x_1^2 + x_2)\mu \ge 0. \end{aligned} \tag{2.1}$$

It is easy to show that $x^* = (0, 0)$ is the global minimizer and $\alpha(x^*) = \emptyset$, $\beta(x^*) = \{1\}$ and $\gamma(x^*) = \emptyset$. The problem is not pathological as $(-1, -1)$ is also a feasible point. SMPCC-LICQ fails at x^* and by KKT conditions we obtain

$$(-2, 0) + \lambda^g(0, 0) - \lambda^G(1, -\mu) - \lambda^H(0, \mu) = 0; \qquad \lambda^g \ge 0, \lambda^G, \lambda^H \in \mathbb{R}$$

From the above equation we obtain $-2 - \lambda^G = 0$, implies that $\lambda^G = -2 < 0$ and $(\lambda^G - \lambda^H)\mu = 0$, implies that $\lambda^H = -2 < 0$. Hence we have $\lambda^G\lambda^H > 0$. So $(1, \lambda^g, -2, -2)$ is a C-stationary multiplier associated to x^*, but x^* is not a strong stationary point, not even an M-stationary point.

The following example illustrates the importance of A-stationarity concept for SMPCC problems when SMPCC-LICQ fails.

Example 2.2. $\xi : \Omega \to \mathbb{R}$ is a random variable which follows the normal distribution with mean μ and variance σ^2, where $\mu > 0$. Further assume that $f(x, \xi) = (x_1 - 1)^2 + \xi$, hence $E[f(x, \xi)] = (x_1 - 1)^2 + \mu$, $g(x) = x_1^3$, $G_1(x, \xi) = x_1 + x_2\xi$, so $E[G_1(x, \xi)] = x_1 + x_2\mu$ and $H_1(x, \xi) = (x_1^2 + x_2)\xi$, then $E[H_1(x, \xi)] = (x_1^2 + x_2)\mu$. The SMPCC problem then is the following

$$\begin{aligned} &\min_{x\in\mathbb{R}^2} \quad (x_1 - 1)^2 + \mu \\ &\text{subject to} \quad x_1^3 \le 0, \\ &\qquad 0 \le (x_1 - x_2\mu) \perp (x_1^2 + x_2)\mu \ge 0. \end{aligned}$$

It is easy to show that $x^* = (0, 0)$ is the global minimizer and $\alpha(x^*) = \emptyset$, $\beta(x^*) = \{1\}$ and $\gamma(x^*) = \emptyset$. The problem is not pathological as $(-\mu, 1)$ is also a feasible point. SMPCC-LICQ fails at x^* and by KKT conditions we obtain

$$(-2,0)+\lambda^g(0,0)-\lambda^G(1,\mu)-\lambda^H(0,\mu)=0; \qquad \lambda^g\geq 0, \lambda^G, \lambda^H\in\mathbb{R}$$

From the above equation we obtain $-2-\lambda^G=0$, implies that $\lambda^G=-2<0$ and $(\lambda^G+\lambda^H)\mu=0$, implies that $\lambda^H=2>0$. Hence we have $\lambda^G<0$ and $\lambda^H>0$. So $(1,\lambda^g,-2,2)$ is an A-stationary multiplier associated to x^*. On the contrary, x^* does not satisfy the stronger stationary conditions.

Remark 1. At these points SMPCC-LICQ fails, hence strong stationarity can not be guaranteed. The above examples show that we could have C- or A-stationarity conditions (though not a strong, M-stationary point) at the point of global (local) minimum which preclude the existence of a feasible first order descent direction.

The above examples motivate the study of C- and A-stationarity in the context of SMPCC problems. They are extremely meaningful as one could obtain a C- or A-stationary point which does not satisfy stronger stationary conditions (strong or M-stationary) and however, they eventually turn out to be local or global minimum of that degenerate stochastic MPCC problem. Next, we study the consistency of these stationary points under suitable SMPCC-NNAMCQ when the true problem is approximated by the SAA subproblems.

3 Consistency of the SAA Estimators

SAA is a well explored technique to approximate the expectations. In this section, we study the consistency of weaker stationary SAA estimators of SMPCC problem (1.1) under suitable SMPCC-NNAMCQ. One has to discretize the underlying probability space associated to the random variable to approximate the expectation functionals through SAA method. Suppose that, the random variable ξ has a finite sample space and it can take a finite number of possible values, say $\boldsymbol{\xi}_1$, $\boldsymbol{\xi}_2$, $\ldots\boldsymbol{\xi}_N$ with respective probabilities p_1, p_2,....p_N. Then the expected value function can be evaluated in the following form:

$$E[f(x,\xi)]=\sum_{k=1}^{N}p_k f(x,\boldsymbol{\xi}_k).$$

Monte Carlo sampling technique is an approach to such a discretization. Suppose, we draw an independent and identically distributed random sample ξ^1, ξ^2,.....ξ^N from π. Then the corresponding sample average function is given by $\frac{1}{N}\sum_{k=1}^{N}f(x,\xi^k)$, as $p_k=\frac{1}{N}$, for each $k=1,....N$. $\frac{1}{N}\sum_{k=1}^{N}f(x,\xi^k)$ can either be considered as a random variable as it depends on the chosen random sample or as a numerical value for a particular set of realizations of the random sample, which is being understood from the context. Thus, the SMPCC problem (1.1) has been approximated by the following sequence of SAA problems:

$$\begin{aligned} &\min_{x\in\mathbb{R}^n} && \frac{1}{N}\sum_{k=1}^{N}f(x,\xi^k) \\ &\text{subject to} && 0\leq \tfrac{1}{N}\sum_{k=1}^{N}G(x,\xi^k)\perp\frac{1}{N}\sum_{k=1}^{N}H(x,\xi^k)\geq 0, \end{aligned} \tag{3.1}$$

where $f : \mathbb{R}^n \times \Xi \to \mathbb{R}$, $g : \mathbb{R}^n \to \mathbb{R}^p$ and $H, G : \mathbb{R}^n \times \Xi \to \mathbb{R}^l$.

We begin by stating the following assumption which are quite well known in the literature of stochastic optimization. They are essential to achieve the consistency of the SAA estimators.

Assumption 3.1. Assume that the functions $f(x, .)$, $G_i(x, .)$ and $H_i(x, .)$ are measurable and bounded by integrable functions $K_f(\omega)$, $K_i^G(\omega)$ and $K_i^H(\omega)$ such that $\| f(x, \xi(\omega)) \| \leq K_f(\omega)$, $\| G_i(x, \xi(\omega)) \| \leq K_i^G(\omega)$, and $\| H_i(x, \xi(\omega)) \| \leq K_i^H(\omega)$ ω-a.e, for $i = 1, ...l$. Further assume that there exist positive valued random variables $C_f(\omega)$, $C_i^G(\omega)$ and $C_i^H(\omega)$ such that $E[C_f(\omega)]$, $E[C_i^G(\omega)]$ and $E[C_i^H(\omega)]$ for all $i = 1, ...l$ are finite and let there exists a neighbourhood U of x^* such that for every x_1, x_2 in U we have

a) $\quad |f(x_1, \xi(\omega)) - f(x_2, \xi(\omega))| \leq C_f(\omega)||x_1 - x_2|| \qquad \omega$-a.e.,

b) $\quad |G_i(x_1, \xi(\omega)) - G_i(x_2, \xi(\omega))| \leq C_i^G(\omega)||x_1 - x_2|| \qquad \omega$-a.e.,

c) $\quad |H_i(x_1, \xi(\omega)) - H_i(x_2, \xi(\omega))| \leq C_i^H(\omega)||x_1 - x_2|| \qquad \omega$-a.e.

The results on consistency are established under suitable SMPCC-NNAMCQ. We first prove the consistency of C-stationary estimators and subsequently, the consistency of A-stationary estimators.

Theorem 3.1. Let us consider the SMPCC problem (1.1) and the corresponding sequence of SAA problems. Assume that the Assumption 3.1 holds. Let $\{x^N : N = 1, 2 \ldots\}$ be a sequence of C-type Fritz John points of SAA problems and x^* is a cluster point of the sequence $\{x^N\}$, ω-a.e. Then x^* is a C-stationary point of the SMPCC problem (1.1), provided C-type SMPCC-NNAMCQ holds at x^*.

Proof. Given that, x^N be a C-type Fritz John point of the SAA problem (3.1). Then by the definition of C-type Fritz John conditions we have scalars $(\bar{\lambda}_0)^N \geq 0$, $(\bar{\lambda}_i^G)^N \in \mathbb{R}$, $i = 1, ...l$ and $(\bar{\lambda}_i^H)^N \in \mathbb{R}$, $i = 1, \ldots l$ such that

i) $$\frac{1}{N}\sum_{k=1}^{N}(\bar{\lambda}_0)^N \nabla_x f(x^N, \xi^k) = \sum_{i=1}^{l}[\frac{1}{N}\sum_{k=1}^{N}(\bar{\lambda}_i^G)^N \nabla_x G_i(x^N, \xi^k)] + \sum_{i=1}^{l}[\frac{1}{N}\sum_{k=1}^{N}(\bar{\lambda}_i^H)^N \nabla_x H_i(x^N, \xi^k)].$$

ii) $\quad ((\bar{\lambda}_0)^N, (\bar{\lambda}^G)^N, (\bar{\lambda}^H)^N) \neq 0.$

iii) $\quad (\bar{\lambda}_{\alpha_N}^G)^N \in \mathbb{R}$, $(\bar{\lambda}_{\gamma_N}^H)^N \in \mathbb{R}$, $(\bar{\lambda}_{\gamma_N}^G)^N = 0$, $(\bar{\lambda}_{\alpha_N}^H)^N = 0$ and $\bar{\lambda}_i^G \bar{\lambda}_i^H \geq 0$, $\forall\ i \in \beta_N$.

Let us assume that

$$\lambda^N = \frac{((\bar{\lambda}_0)^N, (\bar{\lambda}_1^G)^N), ..., (\bar{\lambda}_l^G)^N, (\bar{\lambda}_1^H)^N, ...(\bar{\lambda}_l^H)^N)}{\sqrt{\{(\bar{\lambda}_0)^N\}^2 + \{(\bar{\lambda}_1^G)^N\}^2 + ...\{(\bar{\lambda}_l^G)^N\}^2 + \{(\bar{\lambda}_1^H)^N\}^2 + ...\{(\bar{\lambda}_l^H)^N\}^2}}.$$

So $\| \lambda^N \| = 1$ for all $N \in \mathbb{N}$, bounded sequence implies that there exists a convergent subsequence and hence without any loss of generality, say there exists $\{\lambda^N\}$ such that $\lambda^N \to \lambda^*$, ω-a.e. and by the continuity of the norm $\| \lambda^* \| = 1$ ω-a.e. Further, it is given that $x^N \to x^*$ ω-a.e. (without any loss of generality). Let us now denote that

$$\lambda^N = (\lambda_0^N, (\lambda_1^G)^N, ...(\lambda_l^G)^N, (\lambda_1^H)^N, ...(\lambda_l^H)^N).$$

$$\lambda^* = (\lambda_0^*, (\lambda_1^G)^*, ...(\lambda_l^G)^*, (\lambda_1^H)^*, ...(\lambda_l^H)^*).$$

We construct a compact set B around x^* such that $\frac{1}{N}\sum_{k=1}^{N} \nabla_x f(., \xi^k)$ converges uniformly to $E[\nabla_x f(., \xi)]$ on B, ω-a.e. Hence, we have $\frac{1}{N}\sum_{k=1}^{N} \nabla_x f(x^N, \xi^k)$ converges to $E[\nabla_x f(x, \xi)]$, ω-a.e; which implies $\lambda_0^N[\frac{1}{N}\sum_{k=1}^{N} \nabla_x f(x^N, \xi^k)]$ converges to $\lambda_0^* E[\nabla_x f(x^*, \xi)]$ ω-a.e. Similarly, for $i = 1, \dots l$, $(\lambda_i^G)^N[\frac{1}{N}\sum_{k=1}^{N} \nabla_x G_i(x^N, \xi^k)]$ converges to $(\lambda_i^G)^* E[\nabla_x G_i(x^*, \xi)]$ and $(\lambda_i^H)^N[\frac{1}{N}\sum_{k=1}^{N} \nabla_x H_i(x^N, \xi^k)]$ converges to $(\lambda_i^H)^* E[\nabla_x H_i(x^*, \xi)]$, ω-a.e. This implies that, $\sum_{i=1}^{l}(\lambda_i^G)^N[\frac{1}{N}\sum_{k=1}^{N} \nabla_x G_i(x^N, \xi^k)]$ converges to $\sum_{i=1}^{l}(\lambda_i^G)^* E[\nabla_x G_i(x^*, \xi)]$ and $\sum_{i=1}^{l}(\lambda_i^H)^N[\frac{1}{N}\sum_{k=1}^{N} \nabla_x H_i(x^N, \xi^k)]$ converges to $\sum_{i=1}^{l}(\lambda_i^H)^* E[\nabla_x H_i(x^*, \xi)]$, ω-a.e. For details, see the proof of Theorem 3.1 in [23].

Passing to the limit as $N \to \infty$, we have

$$\lambda_0^* E[\nabla_x f(x^*, \xi)] = \sum_{i=1}^{l}(\lambda_i^G)^* E[\nabla_x G_i(x^*, \xi)] + \sum_{i=1}^{l}(\lambda_i^H)^* E[\nabla_x H_i(x^*, \xi)]$$

Moreover, due to Assumption 3.1 we have for $i = 1, \dots l$, $\frac{1}{N}\sum_{k=1}^{N} G_i(x^N, \xi^k)$ converges to $E[G_i(x^*, \xi)]$ and $\frac{1}{N}\sum_{k=1}^{N} H_i(x^N, \xi^k)$ converges to $E[H_i(x^*, \xi)]$, ω-a.e. As a reference see the proof of Theorem 3.1 in [23].

Let us choose and then fix an $i \in \beta$. By definition of β we know that $E[G_i(x^*, \xi)] = 0$ and $E[H_i(x^*, \xi)] = 0$, i.e. $\frac{1}{N}[\sum_{k=1}^{N} G_i(x^N, \xi^k)] \to 0$ and $\frac{1}{N}[\sum_{k=1}^{N} H_i(x^N, \xi^k)] \to 0$, ω-a.e. If $i \in \beta$, then we will study the following scenarios:

Case 1: Let us assume that there exists $m \in \mathbb{N}$ such that for $N > m$, $\frac{1}{N}[\sum_{k=1}^{N} G_i(x^N, \xi^k)] > 0$, and $\frac{1}{N}[\sum_{k=1}^{N} H_i(x^N, \xi^k)] = 0$, ω-a.e., i.e. $i \in \gamma_N$, ω-a.e for $N > m$. Hence $(\lambda_i^G)^N = 0$, ω-a.e. for $N > m$, this implies $(\lambda_i^G)^* = 0$ (as $\lambda^N \to \lambda^*$, ω-a.e.). Hence $(\lambda_i^G)^*(\lambda_i^H)^* = 0$.

Case 2: Let us assume that there exists $m \in \mathbb{N}$ such that for $N > m$, $\frac{1}{N}[\sum_{k=1}^{N} G_i(x^N, \xi^k)] = 0$, and $\frac{1}{N}[\sum_{k=1}^{N} H_i(x^N, \xi^k)] > 0$, ω-a.e., i.e. $i \in \alpha_N$, ω-a.e. for $N > m$. Hence $(\lambda_i^H)^N = 0$, ω-a.e. for $N > m$, this implies $(\lambda_i^H)^* = 0$. Hence $(\lambda_i^G)^*(\lambda_i^H)^* = 0$.

Case 3: Let us assume that there exists $m \in \mathbb{N}$ such that for $N > m$, $\frac{1}{N}[\sum_{k=1}^{N} G_i(x^N, \xi^k)] = 0$, and $\frac{1}{N}[\sum_{k=1}^{N} H_i(x^N, \xi^k)] = 0$, ω-a.e., i.e. $i \in \beta_N$, ω-a.e. for $N > m$. Noting that x^N is an C-type Fritz John point of the SAA problem (3.1), ω-a.e. Hence $(\lambda_i^G)^N(\lambda_i^H)^N \geq 0$, ω-a.e. for $N > m$. Let without any loss of generality $(\lambda_i^G)^N(\lambda_i^H)^N \geq 0$, ω-a.e. for $N \in \mathbb{N}$. This leads us to the following possibility to consider:

If for some $m_1 \in \mathbb{N}$ and for any $N > m_1$, $(\lambda_i^G)^N(\lambda_i^H)^N \geq 0$, ω-a.e. then we have $(\lambda_i^G)^*(\lambda_i^H)^* \geq 0$.

Case 4: Let us assume that, *a)* $\frac{1}{N}[\sum_{k=1}^N G_i(x^N, \xi^k)] > 0$ and $\frac{1}{N}[\sum_{k=1}^N H_i(x^N, \xi^k)] = 0$, ω-a.e. for infinitely many N.
b) $\frac{1}{N}[\sum_{k=1}^N G_i(x^N, \xi^k)] = 0$ and $\frac{1}{N}[\sum_{k=1}^N H_i(x^N, \xi^k)] > 0$, ω-a.e. for infinitely many N.
c) $\frac{1}{N}[\sum_{k=1}^N G_i(x^N, \xi^k)] = 0$ and $\frac{1}{N}[\sum_{k=1}^N H_i(x^N, \xi^k)] = 0$, ω-a.e. for infinitely many N.
From *a)* and *b)* we get, $(\lambda_i^G)^*(\lambda_i^H)^* = 0$ and from *c)* we get, $(\lambda_i^G)^*(\lambda_i^H)^* \geq 0$. Among the three possibilities mentioned above *a)*, *b)* and *c)*- all of them can occur together, or only *a)* or *b)* or *c)* can occur or they can occur even in pairs *a)*, *b)* or *b)*, *c)* or *a)*, *c)*. This will also lead us to the fact that $(\lambda_i^G)^*(\lambda_i^H)^* \geq 0$.
Hence from the above four possibilities we can conclude that if $i \in \beta$, then $(\lambda_i^G)^*(\lambda_i^H)^* \geq 0$.

Similarly, the Lagrange multipliers satisfy the restrictions when $i \in \alpha$ and $i \in \beta$. Now, applying C-type SMPCC-NNAMCQ at x^*, we have the statement of the theorem. ∎

Theorem 3.2. Let us consider the SMPCC problem (1.1) and the corresponding sequence of SAA problems. Assume that the Assumption 3.1 holds. Let $\{x^N : N = 1, 2 \ldots\}$ be a sequence of A-type Fritz John points of SAA problems and x^* is a cluster point of the sequence $\{x^N\}$, ω-a.e. Then x^* is a A-stationary point of the SMPCC problem (1.1), provided A-type SMPCC-NNAMCQ holds at x^*.

Proof. To prove the first part we adopt the same procedure of the previous theorem and the combinatorial argument to prove the rest is as follows.

Case 1: Let us assume that there exists $m \in \mathbb{N}$ such that for $N > m$, $\frac{1}{N}[\sum_{k=1}^N G_i(x^N, \xi^k)] > 0$, and $\frac{1}{N}[\sum_{k=1}^N H_i(x^N, \xi^k)] = 0$, ω-a.e., i.e. $i \in \gamma_N$, ω-a.e for $N > m$. Hence $(\lambda_i^G)^N = 0$, ω-a.e. for $N > m$, this implies $(\lambda_i^G)^* = 0$ (as $\lambda^N \to \lambda^*$, ω-a.e.).

Case 2: Let us assume that there exists $m \in \mathbb{N}$ such that for $N > m$, $\frac{1}{N}[\sum_{k=1}^N G_i(x^N, \xi^k)] = 0$, and $\frac{1}{N}[\sum_{k=1}^N H_i(x^N, \xi^k)] > 0$, ω-a.e., i.e. $i \in \alpha_N$, ω-a.e. for $N > m$. Hence $(\lambda_i^H)^N = 0$, ω-a.e. for $N > m$, this implies $(\lambda_i^H)^* = 0$.

Case 3: Let us assume that there exists $m \in \mathbb{N}$ such that for $N > m$, $\frac{1}{N}[\sum_{k=1}^N G_i(x^N, \xi^k)] = 0$, and $\frac{1}{N}[\sum_{k=1}^N H_i(x^N, \xi^k)] = 0$, ω-a.e., i.e. $i \in \beta_N$, ω-a.e. for $N > m$. Noting that x^N is an A-type Fritz John point of the SAA problem (3.1), ω-a.e. Hence $(\lambda_i^G)^N \geq 0$ or $(\lambda_i^H)^N \geq 0$, ω-a.e. for $N > m$. Let without any loss of generality $(\lambda_i^G)^N \geq 0$ or $(\lambda_i^H)^N \geq 0$, ω-a.e. for $N \in \mathbb{N}$. This leads us to the following possibilities:

i) If for some $N > m$, $(\lambda_i^G)^N \geq 0$ or $(\lambda_i^H)^N \geq 0$, ω-a.e. then we have either $(\lambda_i^G)^* \geq 0$ or $(\lambda_i^H)^* \geq 0$.

ii) If *a)* $(\lambda_i^G)^N \geq 0$, ω-a.e. for infinitely many $N > m$ and *b)* $(\lambda_i^H)^N \geq 0$, ω-a.e. for infinitely many N. Thus we can construct two subsequences such that $\{N > m : (\lambda_i^G)^N \geq 0,\ \omega - a.e.\}$ and $\{(\lambda_i^H)^N \geq 0,\ \omega - a.e.\}$. From *a)* and *b)* we can conclude that either $(\lambda_i^G)^* \geq 0$ or $(\lambda_i^H)^* \geq 0$.

Case 4: Let us assume that, *a)* $\frac{1}{N}[\sum_{k=1}^{N} G_i(x^N, \xi^k)] > 0$ and $\frac{1}{N}[\sum_{k=1}^{N} H_i(x^N, \xi^k)] = 0$, ω-a.e. for infinitely many N.
b) $\frac{1}{N}[\sum_{k=1}^{N} G_i(x^N, \xi^k)] = 0$ and $\frac{1}{N}[\sum_{k=1}^{N} H_i(x^N, \xi^k)] > 0$, ω-a.e. for infinitely many N.
c) $\frac{1}{N}[\sum_{k=1}^{N} G_i(x^N, \xi^k)] = 0$ and $\frac{1}{N}[\sum_{k=1}^{N} H_i(x^N, \xi^k)] = 0$, ω-a.e. for infinitely many N.
From *a)*, *b)* and *c)* we get, $(\lambda_i^G)^* \geq 0$ or $(\lambda_i^H)^* \geq 0$. Among the three possibilities mentioned above *a)*, *b)* and *c)* all can occur, or only *a)* or *b)* or *c)* can occur or they can occur even in pairs *a)*, *b)* or *b)*, *c)* or *a)*, *c)*. This will also lead us to the fact $(\lambda_i^G)^* \geq 0$ or $(\lambda_i^H)^* \geq 0$.
Therefore, if $i \in \beta$, then $(\lambda_i^G)^* \geq 0$ or $(\lambda_i^H)^* \geq 0$. This completes the proof. ∎

4 Conclusion

The article shows that weaker stationary concepts are significant for degenerate stochastic MPCC problems. C- or A-stationary point (which is not M- or strong stationary) could be a point of local/global minimum of a degenerate SMPCC problem in the absence of SMPCC-LICQ. We need to discuss these weaker stationarities when we study degeneracy of stochastic MPCC problems separately. Moreover, C- and A-stationary estimators are consistent under weaker constraint qualification than SMPCC-LICQ. This result on consistency is necessary since evaluation of the expectation functionals are difficult and one has to approximate them by sample averages. Otherwise, one can never use a particular stationarity concept when SMPCC problems are solved via SAA method; like, strong stationary estimators, as they are inconsistent even under the assumption of SMPCC-LICQ [23]. On the contrary, weaker stationarities including M-stationary points do not preclude first order descent direction in the presence of SMPCC-LICQ and linear independence is a generic property of the SMPCC problems. This situation poses a critical question to us. Is SAA an efficient technique to solve SMPCC problems while SMPCC-LICQ holds? Although it works well for other stochastic optimization problems involving expectations. In fact, it could also play a significant role to solve degenerate SMPCC problems, that is, in the absence of SMPCC-LICQ.

References

[1] S. I. Birbil, G. Gürkan and O. Listes (2006). Solving stochastic mathematical programs with complementarity constraints using simulation. *Math. Oper. Res.* Vol. 31, pp. 739-760.

[2] X. Chen, R. J.-B. Wets and Y. Zhang (2012). Stochastic variational inequalities: residual minimization smoothing sample average approximations. *SIAM Journal on Optimization* Vol. 21, pp. 361-371.

[3] S. Christiansen, M. Patriksson and L. Wynter (2001). Stochastic bilevel programming in structural optimization. *Structual Multidisciplinary Optimization.* Vol. 22, pp. 649-673.

[4] A. Evgrafov and M. Patriksson (2004). On the Existence of Solutions to Stochastic Mathematical Programs with Equilibrium Constraints. *Journal of Optimization Theory and Applications.* Vol. 121, pp. 65-76.

[5] M. Flegel and C.Kanzow (2003). A Fritz John approach to first order optimality conditions for mathematical programs with equilibrium constraints. *Optimization.* Vol. 52, pp. 277-286.

[6] M. Flegel and C. Kanzow (2005). On M-stationary points for mathematical programs with equilibrium constraints. *J. Math. Anal. Appl.* Vol. 310, pp. 286-302.

[7] M. Florian and D. Hearn (1995). Network equilibrium models and algorithms. M. O. Ball, T. Magnanti, C. Monma, G. Nemhauser, eds. *Network Routing Handbooks in OR and MS*, Vol. 8. Elsevier Science, Amsterdam, The Netherlands, pp. 485-550.

[8] M. Fukushima and G-H. Lin (2010). Stochastic Equilibrium Problems and Stochastic Mathematical Programs with Equilibrium Constraints: A Survey. *Pacific Journal of Optimization.* Vol. 6, pp. 455-482.

[9] G. Gürkan, A. Y. Özge and S. M. Robinson (1999). Sample-path solutions of stochastic variational inequalities. *Math. Program.* Vol. 84, pp. 313-334.

[10] R. Henrion, W. Römisch (2007). On M-stationary points for a stochastic equilibrium problem under equilibrium constraints in electricity spot market modeling. *Appl. Math.* Vol. 52, pp. 473-494.

[11] T. Homem-de-Mello and G. Bayraksan (2014). Monte Carlo sampling-based methods for stochastic optimization. *Surveys of Operations Research and Management Science.* Vol. 19, pp. 56-85.

[12] T. Hoheisel, C. Kanzow and A. Schwartz (2013). Theoretical and numerical comparison of relaxation methods for mathematical programs with complementarity constraints. *Math. Program.* Vol. 1-2, Ser. A, 257-288.

[13] C. Kanzow and A. Schwartz (2015). The price of inexactness: convergence properties of relaxation methods for mathematical programs with complementarity constraints revisited. *Math. Oper. Res.* Vol. 40, pp. 253-275.

[14] A. Kleywegt, A. Shapiro and T. Homem-de-Mello (2002). The sample average approximation method for stochastic discrete optimization. *SIAM J. Optim.* Vol. 12, pp. 479-502.

[15] M. Patriksson and L. Wynter (1999). Stochastic Mathematical Programs with Equilibrium Constraints. *Operations Research Letters.* Vol. 25, pp. 159-167.

[16] D. Ralph and H. Xu (2011). Convergence of stationary points of sample average two-stage stochastic programs: a generalized equation approach. *Math. Oper. Res.* Vol. 36, pp. 568-592.

[17] S. M. Robinson(1996). Analysis of sample-path optimization. *Math. Oper. Res.* Vol. 21, pp. 513-528.

[18] H. Scheel and S. Scholtes (2000). Mathematical programs with equilibrium constraints: Stationarity, optimality and sensitivity. *Math. Oper. Res.* Vol. 25, pp. 1-22.

[19] S. Scholtes (2001). Convergence properties of a regularization scheme for mathematical programs with complementarity constraints. *SIAM J. Optim.* Vol. 4, pp. 918-936.

[20] Scholtes, S. and M. Stohr (2001). How Stringent is the Linear Independence Assumption for Mathematical Programs with Complementarity Constraints? *Math. Oper. Res.*, Vol. 26, pp. 851-863.

[21] A. Shapiro (2003). Monte Carlo Sampling Methods. *Stochastic programming.* Edited by A. Ruszczyński and A. Shapiro. pp. 353-425, Handbooks Oper. Res. Management Sci., vol.10, Elsevier Sci. B. V., Amsterdam.

[22] A. Shapiro (2006). Stochastic programming with equilibrium constraints. *Journal of Optimization Theory and Applications.* Vol. 128, pp. 223-243.

[23] A. Sur (2016). Study of M-stationarity and strong stationarity for a Class of SMPCC Problems via SAA Method. *Oper. Res. Lett.* Vol. 44, pp. 425-429.

[24] A. S. Werner (2004). *Bilevel stochastic programming problems: Analysis and application to telecommunications.* Doctoral dissertation, Norwegian University of Science and Technology, Trondheim, Norway.

[25] A. S. Werner and Q. Wang (2009). Resale in vertically separated markets: Profit and consumer surplus implications. *Proc. 20th Internat. Sympos. Math. Programming, Chicago.*

[26] H. Xu (2006). An implicit programming approach for a class of stochastic mathematical programs with complementarity constraints. *SIAM Journal on Optimization.* Vol. 16, pp. 670-696.

[27] H. Xu and J. J. Ye (2010). Necessary optimality conditions for two-stage stochastic programs with equilibrium constraints. *SIAM Journal on Optimization.* Vol. 20, pp. 1685-1715.

[28] J. J. Ye (2000). Constraint qualifications and necessary optimality conditions for optimization problems with variational inequality constraints. *SIAM J. Optim.* Vol. 10, pp. 943-962.

[29] J. J. Ye (2005). Necessary and sufficient optimality conditions for mathematical programs with equilibrium constraints. *J. Math. Anal. Appl.* Vol. 307, pp. 350-369.