CONVERGENCE ANALYSIS OF A SMOOTHING SAA METHOD
FOR A STOCHASTIC MATHEMATICAL PROGRAM WITH
SECOND-ORDER CONE COMPLEMENTARITY CONSTRAINTS

LI CHU

School of Mathematical Sciences, Dalian University of Technology
Dalian 116024, China
City Institute, Dalian University of Technology
Dalian 116600, China

BO WANG*

Key Laboratory of Operations Research and Control of Universities in Fujian
College of Mathematics and Computer Science, Fuzhou University
Fuzhou 350116, China

JIE ZHANG

School of Mathematics, Liaoning Normal University
Dalian 116029, China

HONG-WEI ZHANG

School of Mathematical Sciences, Dalian University of Technology
Dalian 116024, China

(Communicated by Jinyan Fan)

2010 Mathematics Subject Classification. Primary: 90C46, 90C26.

Key words and phrases. Stochastic mathematical programming, second-order cone, complementarity constraints.

The second author’s research is supported in part by the National Natural Science Foundation of China under Project No. 11701091, and Fujian Education and Research Program for Young Teachers under Project No. JAT170096. The third author’s research is supported by the National Natural Science Foundation of China under Project No. 11671183 and No. 11671184, Program for Liaoning Excellent Talents in University under Project No. LR2017049, Scientific Research Fund of Liaoning Provincial Education Department under Project No. L201783638, Liaoning BaiQianWan Talents Program, and Project of Liaoning Provincial Natural Science Foundation of China No. 2019MS-217.

* Corresponding author: Bo Wang.
Abstract. A stochastic mathematical program model with second-order cone complementarity constraints (SSOCMPCC) is introduced in this paper. It can be considered as a non-trivial extension of stochastic mathematical program with complementarity constraints, and could arise from a hard-to-handle class of billevel second-order cone programming and inverse stochastic second-order cone programming. By introducing the Chen-Harker-Kanzow-Smale (CHKS) type function to replace the projection operator onto the second-order cone, a smoothing sample average approximation (SAA) method is proposed for solving the SSOCMPCC problem. It can be shown that with proper assumptions, as the sample size goes to infinity, any cluster point of global solutions of the smoothing SAA problem is a global solution of SSOCMPCC almost surely, and any cluster point of stationary points of the former problem is a C-stationary point of the latter problem almost surely. C-stationarity can be strengthened to M-stationarity with additional assumptions. Finally, we report a simple illustrative numerical test to demonstrate our theoretical results.

1. Introduction. Stochastic mathematical program with symmetric cone complementarity constraints (SMPSCCC) is as follows:

$$\begin{align*}
\min & \quad E[f(x,y,\xi(\omega))] \\
\text{s.t.} & \quad K \ni y \perp E[F(x,y,\xi(\omega))] \in K
\end{align*}$$

where $\xi : \Omega \to \Xi \subseteq \mathbb{R}^k$ is a random vector defined on a probability space $(\Omega, \mathcal{F}, P)$, $f : \mathbb{R}^m \times \mathcal{A} \times \Xi \to \mathbb{R}$ and $F : \mathbb{R}^m \times \mathcal{A} \times \Xi \to \mathcal{A}$ are continuously differentiable mappings, $\mathcal{A}$ is an $m$-dimension real Euclidean space with $S = (\mathcal{A}, \langle \cdot, \cdot \rangle, \circ)$ being a Euclidean Jordan algebra, $K$ is a symmetric cone in $S$ and $E$ denotes the mathematical expectation. It can be considered as a special mathematical program with symmetric cone complementarity constraints, and the expectation is introduced to capture the “average state” in real-world problems when unpredictable or uncertain factors exist.

Taking $K = \mathbb{R}^m_+$, SMPSCCC reduces to the following stochastic mathematical program with complementarity constraints (SMPCC):

$$\begin{align*}
\min & \quad E[f(x,y,\xi(\omega))] \\
\text{s.t.} & \quad 0 \leq y \perp E[F(x,y,\xi(\omega))] \geq 0.
\end{align*}$$

The SMPCC is known as “here-and-now” model where both upper-level and lower-level decisions need to be chosen before the realization of random variable $\xi$, and it has been widely applied in engineering design [4] and economics [1, 3, 11, 17].

Because the famous Mangasarian-Fromovitz constraint qualification (MFCQ) never holds at any feasible point of SMPCC, the theory of classical sample average approach can not be applied directly, see [9, 20]. Birbil et al. [1] proposed a sample-path method for solving SMPCC. Lin et al. [11] suggested a Monte Carlo sampling method. Liu et al. [13] solved SMPCC via a partial exact penalization approach. Liu and Lin [12] introduced the regularization sample average approximation (SAA) method to solve SMPCC. Zhang et al.[24] proposed a smoothing SAA method to approximate the SMPCC.

In this paper, we consider the case that $K$ takes the Cartesian product of second-order cones. We refer to this model as the stochastic mathematical program with second-order cone complementarity constraints (SMPSOCCC or SSOCMPCC). Precisely, this model can be stated as follows:

$$\begin{align*}
\min & \quad E[f(x,y,\xi(\omega))] \\
\text{s.t.} & \quad \mathcal{K}_m \ni \Psi(x,y) \perp y \in \mathcal{K}_m,
\end{align*}$$

(1)
where $\Psi(x, y) = \mathbb{E}[F(x, y, \xi(\omega))]$ and cone $K_m := K_{m_1} \times K_{m_2} \times \cdots \times K_{m_J}$ with $m = \sum_{j=1}^{J} m_j$, which is a Cartesian product of second-order cones. The second-order cone (or ice-cream cone, or Lorentz cone) of dimension $m_j$ is defined by

$$K_{m_j} := \{ (x_1; x_2) \in \mathbb{R} \times \mathbb{R}^{m_j - 1} \mid x_1 \geq \|x_2\| \}.$$ 

Throughout this paper, we assume that $\mathbb{E}[f(x, y, \xi(\omega))]$ and $\mathbb{E}[F(x, y, \xi(\omega))]$ are all well defined, finite and differentiable for any $(x, y) \in \mathbb{R}^2$. For simplicity, $\xi(\omega)$ is denoted by $\xi$. In particular, SSOCMPCC with all $m_j = 1$ for $j = 1, 2, \cdots, J$ coincides with the SMPCC.

It worths to note that the optimization problem (1) is different from but closely related to the following stochastic second-order cone complementarity problem: Find vectors $x, y \in \mathbb{R}^n$ and $z \in \mathbb{R}^l$ such that $x \in K, y \in K, x^T y = 0, F(x, y, z, \xi) = 0$ a.e. $\xi \in \Xi$, where $F : \mathbb{R}^m \times \mathbb{R}^m \times \mathbb{R}^l \times \Xi \rightarrow \mathbb{R}^n \times \mathbb{R}^l$, cf. [10, 18].

The SSOCMPCC has many applications. It could arise in a stochastic bilevel program when the lower level stochastic second-order cone program is replaced by corresponding KKT condition. This corresponds to a class of difficult Stackelberg game. The “inverse problem” of stochastic quadratic program over the second-order cone is also an SSOCMPCC, e.g., in Weber’s problem (cf. [25]) which is to find a best location for the warehouse of a company, the number of customers may depend on uncertain market price and random population movement in regions. This leads to an SSOCMPCC model.

Compared with SMPCC, SSOCMPCC is essentially more difficult due to the lack of the polyhedral property of the second-order cone. It is well known that the classical KKT condition is equivalent to the S-stationary condition for SMPCC. However, the classical KKT condition implies but not equivalent to the S-stationary condition when the dimension of some second-order cone $K_j$ is more than two [22]. The local error bound property holds according to Robinson’s result on polyhedral multifunctions for SMPCC. However, for SSOCMPCC, the local error bound property may not hold without additional assumptions if the dimension of one of the second-order cones $K_j$ is larger than two [23]. Above facts remind us that not all results for SMPCC can be trivially extended to SSOCMPCC.

Consider Problem (1). If the distribution of random variables are known, theoretically it is possible to evaluate all function values via integration. Thus the techniques for deterministic counterpart of SSOCMPCC can be directly applied. However, as has been shown in [6], this approach fails due to the prohibitively expensive calculation. This fact motivates us to consider the sample average approach, which has been extensively studied by many authors (e.g., [6, 15, 19]).

Let $\xi^1, \xi^2, \ldots, \xi^N$ be i.i.d. samples, then Problem (1) can be approximated by the following sample average approach (SAA) problem

$$\min \; \hat{f}_N(x, y)$$
$$\text{s.t.} \; K_m \ni \hat{F}_N(x, y) \perp y \in K_m,$$

where $\hat{f}_N(x, y) := \frac{1}{N} \sum_{i=1}^{N} f(x, y, \xi^i), \hat{F}_N(x, y) := \frac{1}{N} \sum_{i=1}^{N} F(x, y, \xi^i)$. Let $S_{K_m}(x)$ be the projection to cone $K_m$, the SAA problem above is equivalent to

$$\min \; \hat{f}_N(x, y)$$
$$\text{s.t.} \; y - S_{K_m}(y - \hat{F}_N(x, y)) = 0.$$

Thus, the SAA problem is a deterministic counterpart of Problem (1).
Nonsmoothness of above reformulation inspires us to consider its corresponding smoothed problem

\[
\min \quad \tilde{f}_N(x, y) \\
\text{s.t.} \quad \tilde{\Phi}_{\varepsilon N}(x, y) = 0,
\]

where

\[
F(x, y, \xi(\omega)) = \\
\begin{pmatrix}
F_1(x, y, \xi(\omega)) \\
F_2(x, y, \xi(\omega)) \\
\vdots \\
F_J(x, y, \xi(\omega))
\end{pmatrix},
\]

\[
\tilde{\Phi}_{\varepsilon N}(x, y) = \\
\begin{pmatrix}
\varphi_{\varepsilon N}(\tilde{F}_{1}^N(x, y), y_1) \\
\varphi_{\varepsilon N}(\tilde{F}_{2}^N(x, y), y_2) \\
\vdots \\
\varphi_{\varepsilon N}(\tilde{F}_{J}^N(x, y), y_J)
\end{pmatrix},
\]

\[
\varphi_{\varepsilon N}(\tilde{F}_{j}^N(x, y), y_j) := \tilde{F}_{j}^N(x, y) + y_j - \sqrt{(y_j - \tilde{F}_{j}^N(x, y))^2 + 4\varepsilon^2 N e_j},
\]

\[
\tilde{F}_{j}^N(x, y) := \frac{1}{N} \sum_{i=1}^{N} F_j(x, y, \xi^i),
\]

\(F_j\) is the \(j\)-th part of \(F\) corresponding to \(K_{m_j}\), constant vector \(e_j = (1; 0; \cdots; 0) \in \mathbb{R}^{m_j}\) is the identity element corresponding to the Jordan product, sequence \(\varepsilon_N \to 0\) as \(N \to +\infty\). The function \(\varphi_{\varepsilon}\) is the Chen-Harker-Kanzow-Smale (CHKS) function [2, 7, 16], which has been applied to the deterministic counterpart of Problem (1) in [26].

In this paper, we focus on the detailed analysis on the relationship between the smoothed problem (2) and the original problem (1). With more complicated techniques, we verified the following desired properties hold as the sample size tends to infinity under suitable assumptions:

- Any cluster point of global optimal solutions of the smoothed SAA problem is almost surely a global optimal solution of the origin problem (1).
- Any accumulation point of stationary points of the smoothed SAA problem (2) is a C-stationary point of origin problem (1) almost surely. M-stationarity can be achieved with proper additional assumptions.

This paper is organized as follows: Section 2 provides some useful preliminary results. In Section 3, by introducing the subinvertibility and other tools in variational analysis, the almost sure convergence of optimal solutions of the smoothed SAA problem (2) is achieved. Almost sure convergence of stationary points is then established under certain assumptions as the sample size tends to infinity. Some simple numerical result is reported in Section 4.

Some relatively new notations used in this paper need some attention. For a vector \(x = (x_1; x_2) \in \mathbb{R} \times \mathbb{R}^{m-1}\), define \(\tilde{x} := (x_1; -x_2)\). Let \([a]_+ = \max\{a, 0\}\) for \(a \in \mathbb{R}\), and \(S_C(\cdot)\) be the projection onto convex set \(C\). In the sequel, \(S(z)\) is short for \(S_{K_m}(z)\), and \(S_j(z)\) is short for \(S_{K_{m_j}}(z)\). It holds that \(S(z) = S_1(z) \times S_2(z) \times \cdots \times S_J(z)\).

The rest are “standard” notations. Denote by \((A_1; A_2)\) the matrix constructed by joining two parts \(A_1\) and \(A_2\) vertically. For a point \(z\) and \(\varepsilon > 0\), \(B_\varepsilon(z)\) denotes the closed ball with centre \(z\) and radius \(\varepsilon\). Specifically, \(B\) denotes the closed unit ball centred at zero. For a differentiable mapping \(H : \mathbb{R}^m \to \mathbb{R}^n\) and a vector \(z \in \mathbb{R}^m\), we denote by \(JH(z)\) the Jacobian matrix of \(H\) at \(z\) and \(\nabla H(z) := JH(z)^T\). If \(H\) is twice continuously differentiable at \(z \in \mathbb{R}^m\), we denote \(\nabla^2 H(z) := J(\nabla H)(z)\). Inner product is denoted by \(\langle \cdot, \cdot \rangle\) and the norm \(\| \cdot \|\) is the usual 2-norm.
To simplify our analysis, for a given positive decreasing sequence \( \{\varepsilon_N\} \) such that \( \varepsilon_N \to 0 \) as \( N \to +\infty \), we introduce the following notations:

\[
\mathcal{O}_0 := \{(\alpha, \beta) \in \mathbb{R}^m \times \mathbb{R}^m \mid \beta_j - S\mathbb{R}_n (\beta_j - \alpha_j) = 0, j = 1, 2, \cdots, J\},
\]

\[
\mathcal{O}_N := \{(\alpha, \beta) \in \mathbb{R}^m \times \mathbb{R}^m \mid \varphi_{\varepsilon_N}(\alpha_j, \beta_j) = 0, j = 1, 2, \cdots, J\},
\]

\[
Z_0 := \{(x, y) \in \mathbb{R}^m \times \mathbb{R}^m \mid y - S\mathbb{R}_n (y - \Psi(x, y)) = 0\},
\]

\[
Z_N := \{(x, y) \in \mathbb{R}^m \times \mathbb{R}^m \mid \hat{\Phi}_{\varepsilon_N} (x, y) = 0\},
\]

\[
\kappa_0 := \inf\{E[f(x, y, \xi)] \mid (x, y) \in Z_0\},
\]

\[
\mathcal{F}(x, y) := E[f(x, y, \xi)] + \delta_{Z_0}(x, y),
\]

\[
\mathcal{F}_N(x, y) := \hat{f}_N(x, y) + \delta_{Z_N}(x, y).
\]

2. Preliminaries. In this section, we introduce some basic notions in variational analysis and properties about second-order cone.

\( \mathbb{N} \) denotes the set of all positive integers,

\[
\mathcal{N}^\# := \{N \subseteq \mathbb{N} \mid N \text{ infinite}\}, \quad \mathcal{N}_\infty := \{N \subseteq \mathbb{N} \mid \mathbb{N}\setminus N \text{ finite}\}.
\]

For a sequence of sets \( \{C^\nu\}_{\nu \in \mathbb{N}} \), where every set is a subset of \( \mathbb{R}^n \), the outer limit and inner limit are defined by

\[
\limsup_{\nu \to \infty} C^\nu := \{x \mid \exists N \in \mathcal{N}^\#, \exists x^\nu \in C^\nu (\nu \in \mathbb{N}) \text{ with } x^\nu \overset{N}{\to} x\},
\]

\[
\liminf_{\nu \to \infty} C^\nu := \{x \mid \exists N \in \mathcal{N}_\infty, \exists x^\nu \in C^\nu (\nu \in \mathbb{N}) \text{ with } x^\nu \overset{N}{\to} x\},
\]

and the sequence \( \{C^\nu\}_{\nu \in \mathbb{N}} \) is said to converge to \( C \), denoted by \( \lim_{\nu \to \infty} C^\nu = C \), if

\[
\limsup_{\nu \to \infty} C^\nu \subseteq C \subseteq \liminf_{\nu \to \infty} C^\nu.
\]

Consider a set-valued mapping \( M : \mathbb{R}^m \rightrightarrows \mathbb{R}^n \). The graph of \( M \) is a subset of \( \mathbb{R}^m \times \mathbb{R}^n \) defined by

\[
gph M := \{(x, u) \mid u \in M(x)\}.
\]

Set-valued mapping \( M(\cdot) \) is called continuous at \( \bar{x} \), if

\[
\limsup_{\bar{x} \to \bar{x}^+} M(x) \subseteq M(\bar{x}) \subseteq \liminf_{\bar{x} \to \bar{x}^-} M(x).
\]

For set \( C \subseteq \mathbb{R}^m \) and \( \bar{x} \in C \),

\[
\bar{N}_C(\bar{x}) = \{v \mid \langle v, x - \bar{x} \rangle \leq o(\|x - \bar{x}\|), \forall x \in C\}
\]

is the regular normal cone of the set \( C \) at point \( \bar{x} \), and

\[
N_C(\bar{x}) = \limsup_{x \overset{C}{\to} \bar{x}} \bar{N}_C(x)
\]

is the (limiting) normal cone of set \( C \) at point \( \bar{x} \).

Let \( M : \mathbb{R}^m \rightrightarrows \mathbb{R}^n \) be a set-valued mapping and \( (x^*, y^*) \in gph M \). The regular coderivative and the limiting (Mordukhovich) coderivative of \( M \) at \( (x^*, y^*) \) are the set-valued mappings defined by

\[
\hat{D}^* M(x^*, y^*)(v) := \{u \in \mathbb{R}^m \mid (u, -v) \in \bar{N}_{gph M}(x^*, y^*)\},
\]

\[
\check{D}^* M(x^*, y^*)(v) := \{u \in \mathbb{R}^m \mid (u, -v) \in N_{gph M}(x^*, y^*)\}.
\]
For a single-valued Lipschitz continuous map $H : \mathbb{R}^m \to \mathbb{R}^n$, the B(ouligand)-subdifferential $\partial_B H$ is defined as
\[
\partial_B H(x) = \{ \lim_{k \to \infty} JH(x_k) | x_k \to x, H \text{ is differentiable at } x_k \},
\]
and the Clake generalized Jacobian of $H$ at $x$ is defined as the convex hull of $\partial_B H(x)$ which denoted by $\partial H(x)$, cf. [5].

Suppose $\{f^\nu\}$ be any sequence of functions on $\mathbb{R}^m$. By Proposition 7.2 in [15], we have $\text{epi} f^\nu \to \text{epi} f$ if and only if at each point $x \in \mathbb{R}^m$, one has
\[
\liminf_{\nu \to \infty} f^\nu(x^\nu) \geq f(x), \forall x^\nu \to x \text{ as } \nu \to \infty,
\]
\[
\limsup_{\nu \to \infty} f^\nu(x^\nu) \leq f(x), \exists x^\nu \to x \text{ as } \nu \to \infty.
\]

For all $x \in \mathbb{R}^m$ and $i = 1$ or 2, let
\[
\lambda_i(x) = x_1 + (-1)^i \|x_2\|, \\
c_i(x) = \begin{cases} \frac{1}{2} (1; (-1)^i \frac{x_2}{\|x_2\|}) & \text{if } x_2 \neq 0 \\ \frac{1}{2} (1; (-1)^i v) & \text{if } x_2 = 0 \end{cases},
\]
where $v$ is an arbitrarily selected unit vector, $\lambda_1(x)$ and $\lambda_2(x)$ are spectral values of $x$, $c_1(x)$ and $c_2(x)$ are the associated spectral vectors of $x$. Value $\det(x) = \lambda_1(x)\lambda_2(x) = x_1^2 - \|x_2\|^2$ is called the determinant of $x$. Obviously, for all $x \in \mathbb{R}^m$ the following spectral factorization holds,
\[
x = \lambda_1(x)c_1(x) + \lambda_2(x)c_2(x). \tag{3}
\]

The Jordan product of $x = (x_1; x_2) \in \mathbb{R} \times \mathbb{R}^{m-1}$ and $y = (y_1; y_2) \in \mathbb{R} \times \mathbb{R}^{m-1}$ is defined by
\[
x \circ y := (x^Ty; x_1y_2 + y_1x_2).
\]

Jordan product is closely related to second-order cone, or more generally, closely related to all symmetric cones. It holds that $\mathcal{K} = \{ x \circ x \mid x \in \mathbb{R}^m \}$. Vector $e = (1; 0; \cdots; 0) \in \mathbb{R}^m$ is the identity element corresponding to Jordan product in the sense that for all $x \in \mathbb{R}^m$,
\[
e \circ x = x \circ e = x.
\]

For simplicity, “$x \circ x$” is often written as “$x^2$”. For $x = (x_1; x_2) \in \mathbb{R} \times \mathbb{R}^{m-1}$, define matrix
\[
L_x := \begin{pmatrix} x_1 & x_2^T \\ x_2 & x_1I \end{pmatrix}.
\]

Obviously $L_{(-x)} = -L_x, x \in \mathcal{K} \Leftrightarrow L_x \succeq 0, x \in \text{int} \mathcal{K} \Leftrightarrow L_x > 0$. If $L_x$ is invertible, then
\[
L_x^{-1} = \frac{1}{\det(x)} \begin{pmatrix} x_1 & -x_2^T \\ -x_2 & \det(x)x_1I + x_2x_2^T \end{pmatrix}.
\]

For a bounded sequence $\{x_k\} \subseteq \mathbb{R}^m$, $\{L_{x_k}\}$ is also a bounded matrix sequence. $L_x$ is closely related to Jordan product. It holds that
\[
x \circ y = L_x y = L_y x.
\]

Many operations to $x$ can be easily expressed or defined via its spectral decomposition. For projection operator onto the second-order cone $\mathcal{K}$, it holds that
\[
S_\mathcal{K}(x) = [\lambda_1(x)]_+c_1(x) + [\lambda_2(x)]_+c_2(x).
\]
Square and square root of $x$ are equal to or defined by certain spectral decompositions.

\[ x^2 = \lambda_1^2(x)c_1(x) + \lambda_2^2(x)c_2(x), \]
\[ \sqrt{x} := \sqrt{\lambda_1(x)c_1(x) + \sqrt{\lambda_2(x)c_2(x)}}. \]

Next we introduce some useful results and definitions from [8], [22], [24] and [26].

**Definition 2.1.** A multifunction $\mathcal{M} : \mathbb{R}^m \rightrightarrows \mathbb{R}^n$ is subinvertible at $(\bar{x}, 0)$, if one has that $0 \in \mathcal{M}(\bar{x})$ and there exists a compact convex neighbourhood $U$ of $\bar{x}$, a positive constant $\varepsilon > 0$, and a non-empty convex-valued mapping $G : \varepsilon B \rightrightarrows U$ such that $\text{gph}G$ is closed, the point $\bar{x}$ belongs to $G(0)$, and $G(y)$ is contained in $\mathcal{M}^{-1}(y)$ for all $y \in \varepsilon B$.

Subinvertible multifunctions have the following desirable property.

**Lemma 2.2.** Let $\mathcal{M} : \mathbb{R}^m \rightrightarrows \mathbb{R}^n$ be a multifunction that is subinvertible at $(\bar{x}, 0)$. Then there is a compact convex subset $U$ and a real number $\varepsilon > 0$ such that for the solution mapping

\[ J(Q) = \{ x \in \mathbb{R}^m \mid 0 \in \mathcal{M}(x) + Q(x) \}, \]

one has that $U \cap J(Q)$ is non-empty for every perturbation function $Q : U \rightrightarrows \mathbb{R}^n$ satisfying

\[ \sup_{x \in U} \sup_{y \in Q(x)} \| y \| \leq \varepsilon. \]

Let $(x, y)$ be a feasible point of Problem (1). Define the following index sets for convenience,

\[ \mathcal{N}_1 := \{ j \mid \Psi_j(x, y) = 0, y_j \in \text{int} \mathcal{K}_{m_j} \} = \{ j \mid y_j - \Psi_j(x, y) \in \text{int} \mathcal{K}_{m_j} \}, \]
\[ \mathcal{N}_2 := \{ j \mid \Psi_j(x, y) \in \text{int} \mathcal{K}_{m_j}, y_j = 0 \} = \{ j \mid y_j - \Psi_j(x, y) \in -\text{int} \mathcal{K}_{m_j} \}, \]
\[ \mathcal{N}_3 := \{ j \mid \Psi_j(x, y) \in \text{bd} \mathcal{K}_{m_j} \setminus \{ 0 \}, y_j \in \text{bd} \mathcal{K}_{m_j} \setminus \{ 0 \} \} = \{ j \mid y_j - \Psi_j(x, y) \in \mathbb{R}^{m_j} \setminus (\mathcal{K}_{m_j} \cup -\mathcal{K}_{m_j}) \}, \]
\[ \mathcal{N}_4 := \{ j \mid \Psi_j(x, y) = 0, y_j \in \text{bd} \mathcal{K}_{m_j} \setminus \{ 0 \} \} = \{ j \mid y_j - \Psi_j(x, y) \in \text{bd} \mathcal{K}_{m_j} \setminus \{ 0 \} \}, \]
\[ \mathcal{N}_5 := \{ j \mid \Psi_j(x, y) \in \text{bd} \mathcal{K}_{m_j} \setminus \{ 0 \}, y_j = 0 \} = \{ j \mid y_j - \Psi_j(x, y) \in -\text{bd} \mathcal{K}_{m_j} \setminus \{ 0 \} \}, \]
\[ \mathcal{N}_6 := \{ j \mid \Psi_j(x, y) = 0, y_j = 0 \} = \{ j \mid y_j - \Psi_j(x, y) = 0 \}, \]
\[ \mathcal{N} := \mathcal{N}_3 \cup \mathcal{N}_4 \cup \mathcal{N}_5 \cup \mathcal{N}_6. \]

The following regularization properties are necessary in construction of convergence theory, which has been introduced in [26].

**Definition 2.3.** Let $(\bar{x}, \bar{y})$ be a feasible point of Problem (1) and $\Psi$ is continuously differentiable at $(\bar{x}, \bar{y})$. We say the SOCMPCC linear independence constraint qualification (SOCPCC-LICQ) holds at $(\bar{x}, \bar{y})$, if the matrix

\[ \tilde{C}(\bar{x}, \bar{y}) := \begin{pmatrix} \mathcal{J}^{\Psi \mathcal{N}_1 \cup \mathcal{N}}(\bar{x}, \bar{y}) & \mathcal{J}\hat{y}\mathcal{N}_2 \cup \mathcal{N} \end{pmatrix} \]

has full row rank, where $\mathcal{J}\hat{y}\mathcal{J}_j = \begin{pmatrix} 0 & 0 & I_{m_j} & 0 \end{pmatrix}$, $j \in \mathcal{N}_2 \cup \mathcal{N}$.

The Lagrangian function of Problem (1) is defined as

\[ L(x, y, u, v) = \mathbb{E}[f(x, y, \xi(\omega))] + \sum_{j=1}^{J} \Psi_j(x, y)^T u_j + \sum_{j=1}^{J} y_j^T v_j, \]
where \( u = (u_1; u_2; \cdots; u_J) \), \( v = (v_1; v_2; \cdots; v_J) \) with \( u_j, v_j \in \mathbb{R}^{m_j}, j = 1, 2, \cdots, J \). For a feasible point \((\bar{x}, \bar{y})\) of Problem (1), stationary concepts can be defined similarly to those complementarity involved problems (cf. [21, 26]) as follows.

**Definition 2.4.** Feasible point \((\bar{x}, \bar{y})\) is called a W-stationary point of Problem (1), if there exist multiplier vectors \( \bar{u} = (\bar{u}_1; \bar{u}_2; \cdots; \bar{u}_J) \) and \( \bar{v} = (\bar{v}_1; \bar{v}_2; \cdots; \bar{v}_J) \), such that

\[
\nabla \mathbb{E}[f(\bar{x}, \bar{y}, \xi(\omega))] + \sum_{j=1}^{J} \nabla \Psi_j(\bar{x}, \bar{y}) \bar{u}_j + \sum_{j=1}^{J} \nabla \bar{y}_j \bar{v}_j = 0,
\]

\( \bar{v}_j = 0, \quad j \in N_1, \) \( \bar{u}_j = 0, \quad j \in N_2, \) \( -\bar{u}_j = \partial B S_j(\bar{y}_j - \Psi_j(\bar{x}, \bar{y}))(\bar{y}_j - \bar{v}_j), \quad j \in N_3. \) \( \tag{7} \)

**Definition 2.5.** Feasible point \((\bar{x}, \bar{y})\) is called a C-stationary point of Problem (1) if (4)-(7) hold, and

\[ -\bar{u}_j \in \partial S_j(\bar{y}_j - \Psi_j(\bar{x}, \bar{y}))(\bar{y}_j - \bar{v}_j), \quad j \in N_4 \cup N_5 \cup N_6. \] \( \tag{8} \)

**Definition 2.6.** Feasible point \((\bar{x}, \bar{y})\) is called an M-stationary point of Problem (1) if (4)-(7) hold, and

\[ -\bar{u}_j \in D^+ S_j(\bar{y}_j - \Psi_j(\bar{x}, \bar{y}))(\bar{y}_j - \bar{v}_j), \quad j \in N_4 \cup N_5 \cup N_6. \]

**Definition 2.7.** Feasible point \((\bar{x}, \bar{y})\) is called an S-stationary point of Problem (1) if (4)-(7) hold, and

\[ -\bar{u}_j \in \hat{D}^+ S_j(\bar{y}_j - \Psi_j(\bar{x}, \bar{y}))(\bar{y}_j - \bar{v}_j), \quad j \in N_4 \cup N_5 \cup N_6. \]

The C-stationary concept defined in Ye and Zhou [22] is different from the definition above. It is referred to as weak C-stationary in this paper.

**Definition 2.8.** Feasible point \((\bar{x}, \bar{y})\) is called a weak C-stationary point of Problem (1) if (4)-(6) hold, and

\[ \bar{u}_j \perp \Psi_j(\bar{x}, \bar{y}), \quad \bar{v}_j \perp \bar{y}_j, \quad \Psi_j(\bar{x}, \bar{y}) \bar{u}_j + \bar{y}_j \bar{v}_j \in \mathbb{R} \Psi_j(\bar{x}, \bar{y}), \quad j \in N_3, \]

\[ \bar{v}_j \in \mathbb{R} \bar{y}_j, \quad j \in N_4, \]

\[ \bar{u}_j \in \mathbb{R} \Psi_j(\bar{x}, \bar{y}), \quad j \in N_5, \]

\[ \bar{v}_j^T \bar{v}_j \geq 0, \quad j = 1, 2, \cdots, J. \]

Note that Ye and Zhou [22] also defined the W-, C-, M-, S-stationarity in different forms. It can be verified that W-, M- and S-stationarity by Ye and Zhou coincide with ours. However, the C-stationarity implies the weak C-stationarity. The following example illustrates the difference between C- and weak C-stationarity.

**Example 1.** Suppose \( x \in \mathbb{R}, \ y = (y_1; y_2) \in \mathbb{R}^2 \), and random variable \( \xi \) obeys some distribution. Consider the following problem

\[
\begin{align*}
\min & \quad \mathbb{E}[f(x, y, \xi)] = x^2 - 8x + y_1^2 - 4y_1 + y_2^2 - 7y_2 \\
\text{s.t.} & \quad K_2 \ni y \perp \Psi(x, y) = (x + y_1; x + y_2) \in K_2
\end{align*}
\]

It is obvious that \((\bar{x}, \bar{y}) = (0; 0)\) is a feasible point of the above problem. In this case, \( j \in N_6 \). Direct calculation shows that
\[ \nabla E[f(\bar{x}, \bar{y}, \xi)] = (-8; -4; -7), \quad \nabla \Psi(\bar{x}, \bar{y}) = \begin{pmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \nabla \bar{y} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}. \]

Taking \( \bar{u} = (5; 3) \), \( \bar{v} = (-1; 4) \), then we have (4) and \( \bar{u}^T \bar{v} > 0 \), which means that the feasible point \((0; 0; 0)\) is a weak C-stationary point. Next we show that \((0; 0; 0)\) is \( \lambda = 2.12 \) that will be introduced later, there exists \( 0 \leq \lambda \leq 2 \). Addition, it holds that

\[ \alpha \leq \lambda \leq 1, \quad i = 1, 2, 3; \sum_{i=1}^{3} \lambda_i = 1 \] and \( |w| = 1 \), such that

\[ \begin{cases} \bar{u}_1 = \lambda_2 (\bar{u}_1 + \bar{v}_1) + \alpha \lambda_3 (\bar{u}_1 + \bar{v}_1) + \frac{\lambda_3}{\lambda_2} ((\bar{u}_1 + \bar{v}_1)(1 - 2\alpha) + w(\bar{u}_2 + \bar{v}_2)), \\ \bar{u}_2 = \lambda_2 (\bar{u}_2 + \bar{v}_2) + \alpha \lambda_3 (\bar{u}_2 + \bar{v}_2) + \frac{\lambda_3}{\lambda_2} ((\bar{u}_1 + \bar{v}_1)w + (1 - 2\alpha)w^2(\bar{u}_2 + \bar{v}_2)). \end{cases} \]

By solving the system above, we find that \( \alpha \) plays no role in above equations. In addition, it holds that

\[ \begin{cases} \lambda_3 = \frac{46}{33} > 1 & \text{if } w = 1, \\ \lambda_3 = \frac{46}{33} < -1 & \text{if } w = -1, \end{cases} \]

which contradicts the condition “\( 0 \leq \lambda_3 \leq 1 \)”.

Based on the relationships among \( \bar{D}^* S_j(\cdot) \), \( D^* S_j(\cdot) \) and \( \partial S_j(\cdot) \) (the explicit expressions of \( \bar{D}^* S_j(\cdot) \) and \( D^* S_j(\cdot) \) are provided in [26]), the following relationship holds for a feasible point \((x, y)\):

\[ \text{S-stationarity } \Rightarrow \text{M-stationarity } \Rightarrow \text{C-stationarity} \]

\[ \Rightarrow \text{weak C-stationarity } \Rightarrow \text{W-stationarity}. \]

When the following regularization conditions hold, Problem (1) will have better properties, which will be discussed later in the following section.

**Definition 2.9.** Let \((\bar{x}, \bar{y})\) be a feasible point of Problem (1). The strict complementary condition is said to hold at \((\bar{x}, \bar{y})\), if

\[ \mathcal{N}_5 = \{ j | \Psi_j(\bar{x}, \bar{y}) = 0, \ \bar{y}_j = 0 \} = \emptyset. \]

**Definition 2.10.** The second-order cone upper level strict complementarity (SOC-ULSC) condition for Problem (1) is said to hold at a W-stationary point \((\bar{x}, \bar{y})\) if there exist multipliers \( \bar{u} \) and \( \bar{v} \) satisfying (4)-(7) and

(i) \( \langle \bar{u}_j + \bar{v}_j, c_1(\bar{y}_j - \Psi_j(\bar{x}, \bar{y})) \rangle \leq 0, j \in \mathcal{N}_4, \)

(ii) \( \langle \bar{u}_j + \bar{v}_j, c_2(\bar{y}_j - \Psi_j(\bar{x}, \bar{y})) \rangle \leq 0, j \in \mathcal{N}_5, \)

(iii) \( \bar{u}_j \notin bd\mathcal{K}_m \cup (-bd\mathcal{K}_m) \) and \( \bar{v}_j \notin bd\mathcal{K}_m \cup (-bd\mathcal{K}_m), j \in \mathcal{N}_6, \)

where \( c_i(\cdot) \) is the eigenvector in spectral decomposition for elements in the second-order cone.

The following three lemmas from [26] provide some basic knowledge needed in further discussion. For simplicity, define

\[ a_{jN} := y_j^N - \bar{F}_j^N(x_N, y_N), \]

\[ \bar{u}_j := \bar{y}_j - \Psi_j(\bar{x}, \bar{y}), \]

\[ W_{jN} := \sqrt{(y_j^N - \bar{F}_j^N(x_N, y_N)) + 4\varepsilon_N^2}. \]
Lemma 2.11. For any \((x_N, y_N) \in Z_N\), if \((x_N, y_N)\) converges to \((\bar{x}, \bar{y})\) as \(N \to +\infty\), then the following conclusion holds:

\[
\lim_{N \to +\infty} L_{W_j/N}^1 L_{\alpha_j/N} = \lim_{N \to +\infty} L_{\alpha_j/N} L_{W_j/N}^{-1}
\]

\[
\begin{cases}
I & j \in \mathcal{N}_1, \\
-I & j \in \mathcal{N}_2, \\
\frac{a_j}{\|a_j\|} I + \left( \frac{a_j}{\|a_j\|} \frac{\bar{a}_j}{\|a_j\|} \right) & j \in \mathcal{N}_3,
\end{cases}
\]

\[
= \begin{cases}
I + (1 - r) I + \left( \frac{r_1 + r_2 - 2r}{\|a_j\|} \frac{\bar{a}_j}{\|a_j\|} \right), r \in [0, 1] & j \in \mathcal{N}_4, \\
-I + r \left( \frac{1}{\|a_j\|} \frac{\bar{a}_j}{\|a_j\|} \right), r \in [0, 1] & j \in \mathcal{N}_5,
\end{cases}
\]

\[
0 \leq r_1 \leq r \leq r_2 \leq 1, v \in \mathbb{R}^{m-j-1}, \|v\| = 1 \\
& j \in \mathcal{N}_6.
\]

Lemma 2.12. Let \(x \in \mathbb{R}^m\) have the spectral decomposition as in (3). It holds that

(i) If \(x \in \text{int} \mathcal{K}_m\), then \(\partial_B S_K(x) = \{I\}\).

(ii) If \(x \in -\text{int} \mathcal{K}_m\), then \(\partial_B S_K(x) = \{0\}\).

(iii) If \(x \in \mathbb{R}^m \setminus (-\mathcal{K}_m \cup \mathcal{K}_m)\), then

\[
\partial_B S_K(x) = \left\{ \frac{1}{2} \left( 1 + \frac{x_1}{\|x_2\|} \right) I + \frac{1}{2} \left( \frac{x_1}{\|x_2\|} - \frac{x_2^T}{\|x_2\|} \right) \right\}.
\]

(iv) If \(x \in \text{bd} \mathcal{K}_m \setminus \{0\}\), then

\[
\partial_B S_K(x) = \left\{ I, I + \frac{1}{2} \left( -1 \frac{x_2}{\|x_2\|} \frac{x_2^T}{\|x_2\|^2} \right) \right\}.
\]

(v) If \(x \in -\text{bd} \mathcal{K}_m \setminus \{0\}\), then

\[
\partial_B S_K(x) = \left\{ 0, \frac{1}{2} \left( 1 \frac{x_2}{\|x_2\|^2} \frac{x_2^T}{\|x_2\|^2} \right) \right\}.
\]

(vi) If \(x = 0\), then

\[
\partial_B S_K(x) = \{0, I\} \cup \left\{ \alpha I + \frac{1}{2} \left( 1 - 2\alpha \frac{v}{\|v\|} \frac{v^T}{(1 - 2\alpha)\|v\|^2} \right) \mid \alpha \in [0, 1], \|v\| = 1 \right\}.
\]

Some useful properties which have been used implicitly in [26] are summarized in the following lemma. It shows that the generalized Jacobian of the project operator can be approximated via the Jacobian of the generalized square root root operator. This implies that the generalized CHKS function is a smoothing function for the non-smooth equation corresponding to the complementarity. The proof is omitted as it can be easily checked by the results in [26].

Lemma 2.13. Consider the function \(\psi_e(z) := \sqrt{z^2 + 4\varepsilon^2 e} \) and \(e = (1; 0; \cdots; 0)\) is the identity element of Jordan product second-order cone \(\mathcal{K}\), then we have
(i) $\partial_B S_k(z) \leq \frac{1}{2}(I + \partial_B \psi_0(z)) \subseteq \partial S_k(z)$, where

$$
\partial_z \psi_0(z) := \lim_{N \to +\infty} L_{\psi_N(z)}^{-1} L_{z_N} \begin{cases}
\psi_1(\cdot) \text{ is differentiable at } (z_N, z) \text{ with } z_N \to z, z_N \to 0 \text{ as } N \to +\infty.
\end{cases}
$$

(ii) $J\psi_1(z) = L_{\psi_1(z)}^{-1} L_z = \begin{cases}
(2\tilde{g}'(\frac{z}{\varepsilon}) - 1)I & \text{if } z_2 = 0
\end{cases}$ where

$$
\tilde{g}(y) := \sqrt{\frac{y^2 + 4 + y}{2}}, \quad a := 2\tilde{g}'(\frac{\lambda_2(\varepsilon)}{\varepsilon}) - \tilde{g}'(\frac{\lambda_1(\varepsilon)}{\varepsilon}), \quad b := \tilde{g}'(\frac{\lambda_2(\varepsilon)}{\varepsilon}) + \tilde{g}'(\frac{\lambda_1(\varepsilon)}{\varepsilon}), \quad c := \tilde{g}'(\frac{\lambda_2(\varepsilon)}{\varepsilon}) - \tilde{g}'(\frac{\lambda_1(\varepsilon)}{\varepsilon}).
$$

(iii) $I \pm J\psi_1(z)$ is invertible.

3. Convergence of the smoothing SAA method. In this section, the almost sure convergence of the optimal values, optimal solutions and stationary points of the smoothed SAA problem (2) to the corresponding ones of the original problem (1) are obtained as the sample size tends to infinity. This convergence will be illustrated later by the numerical tests in the following section.

In the rest of this paper, assume that samples $\xi^1, \xi^2, \ldots, \xi^N$ of the random vector $\xi$ are i.i.d., and some more proper assumptions are needed in the following discussion.

**Assumption 1.** The mappings $f(\cdot, \cdot, \xi)$ and $F(\cdot, \cdot, \xi)$ are twice continuously differentiable on $\mathbb{R}^m$ a.e. $\xi \in \Xi$.

**Assumption 2.** For any $(\bar{x}, \bar{y}) \in \mathbb{R}^{2m}$, there exists a neighbourhood $D$ of $(\bar{x}, \bar{y})$ and a non-negative measurable function $g(\xi)$ such that $\mathbb{E}[g(\xi)] < +\infty$, it holds that for all $(x, y) \in D$ and for almost every $\xi \in \Xi$,

$$
\max\{\|\nabla f(x, y, \xi)\|, \|\nabla^2 f(x, y, \xi)\|, \|JF(x, y, \xi)\|, \|J^2F(x, y, \xi)\|\} \leq g(\xi).
$$

According to Theorem 7.44 in [15], we have that $\mathbb{E}[f(x, y, \xi)]$ and $\mathbb{E}[F(x, y, \xi)]$ are twice continuously differentiable on $\mathbb{R}^{2m}$ under Assumptions 1 and 2. In particular,

$$
\nabla \mathbb{E}[f(x, y, \xi)] = \mathbb{E}[\nabla_{xy} f(x, y, \xi)] \quad \text{and} \quad J\mathbb{E}[F(x, y, \xi)] = \mathbb{E}[J_{xy}F(x, y, \xi)].
$$

The following results are from the Uniform Laws of Large Numbers, see Theorem 7.48 in [15].

**Lemma 3.1.** Suppose that Assumptions 1 and 2 are satisfied. Let $(x_N, y_N) \in Z_N$, if the sequence $(x_N, y_N) \to (\bar{x}, \bar{y})$ w.p.1 as $N \to +\infty$, then the expectation and differential operators are commutative, and the following convergences hold with probability one:

$$
\hat{f}_N(x_N, y_N) \to \mathbb{E}[f(\bar{x}, \bar{y}, \xi)],
$$

$$
\nabla \hat{f}_N(x_N, y_N) \to \mathbb{E}[\nabla_{xy} f(\bar{x}, \bar{y}, \xi)] = \nabla \mathbb{E}[f(\bar{x}, \bar{y}, \xi)],
$$

$$
\nabla^2 \hat{f}_N(x_N, y_N) \to \mathbb{E}[\nabla^2_{xy,xy} f(\bar{x}, \bar{y}, \xi)] = \nabla^2 \mathbb{E}[f(\bar{x}, \bar{y}, \xi)],
$$

$$
\hat{F}_N(x_N, y_N) \to \Psi(\bar{x}, \bar{y}),
$$

$$
J\hat{F}_N(x_N, y_N) \to \mathbb{E}[J_{xy} F(\bar{x}, \bar{y}, \xi)] = J\mathbb{E}[F(\bar{x}, \bar{y}, \xi)],
$$

$$
J^2\hat{F}_N(x_N, y_N) \to \mathbb{E}[J^2_{xy} F(\bar{x}, \bar{y}, \xi)] = J^2 \mathbb{E}[F(\bar{x}, \bar{y}, \xi)].
$$
The following lemma characterizes the convergence properties of $O_N$ as $N$ tends to infinity.

**Lemma 3.2.** Suppose that $\{\varepsilon_N\}$ is a decreasing sequence of positive real numbers converging to 0. Then it holds that

$$O_0 = \lim_{N \to +\infty} O_N.$$  

**Proof.** For arbitrary vector $w \in \mathbb{R}^m$, construct $(\alpha_N, \beta_N) \in \mathcal{O}_N$ in the following way. For every $j = 1, 2, \cdots, J$, find the spectral decomposition of its $j$-th part $w_j$

$$w_j = \lambda_1 c_1 + \lambda_2 c_2.$$  

Define the $j$-th part of $(\alpha_N, \beta_N)$ by

$$\begin{align*}
(\alpha_N)_j &= \frac{1}{2} (\sqrt{\lambda_1^2 + 4\varepsilon_N^2} - \lambda_1) c_1 + \frac{1}{2} (\sqrt{\lambda_2^2 + 4\varepsilon_N^2} - \lambda_2) c_2, \\
(\beta_N)_j &= \frac{1}{2} (\sqrt{\lambda_1^2 + 4\varepsilon_N^2} + \lambda_1) c_1 + \frac{1}{2} (\sqrt{\lambda_2^2 + 4\varepsilon_N^2} + \lambda_2) c_2.
\end{align*}$$  

Direct calculation shows that

$$\varphi_{\varepsilon_N}((\alpha_N)_j, (\beta_N)_j) = 0,$$  

i.e., $\varphi_{\varepsilon_N}((\alpha_N)_j, (\beta_N)_j) = 0$, thus $(\alpha_N, \beta_N) \in \mathcal{O}_N$. Let $N \to +\infty$, then one of the following cases occurs.

(i) if $\lambda_1 \geq 0, \lambda_2 \geq 0$, then $(\alpha_N)_j \to 0$ and $(\beta_N)_j \to \lambda_1 c_1 + \lambda_2 c_2$,

(ii) if $\lambda_1 \leq 0, \lambda_2 \leq 0$, then $(\alpha_N)_j \to -\lambda_1 c_1 - \lambda_2 c_2$ and $(\beta_N)_j \to 0$,

(iii) if $\lambda_1 \leq 0, \lambda_2 \geq 0$, then $(\alpha_N)_j \to -\lambda_1 c_1$ and $(\beta_N)_j \to \lambda_2 c_2$,

(iv) if $\lambda_1 \geq 0, \lambda_2 \leq 0$, then $(\alpha_N)_j \to -\lambda_2 c_2$ and $(\beta_N)_j \to \lambda_1 c_1$.

For arbitrary $(\alpha, \beta) \in \mathcal{O}_0$, define $w_j = \beta_j - \alpha_j$, for all $j = 1, 2, \cdots, J$. Corresponding $(\alpha_N, \beta_N) \in \mathcal{O}_N$ can be constructed according to previous discussion by equations (9)-(10). It is simple to verify that $(\alpha_N, \beta_N) \to (\alpha, \beta)$ as $N \to +\infty$. Thus inclusion

$$O_0 \subseteq \lim inf_{N \to +\infty} \mathcal{O}_N$$  

holds.

On the other hand, for arbitrary $(\bar{\alpha}, \bar{\beta})$ in $\lim sup_{N \to +\infty} \mathcal{O}_N$, w.l.o.g., suppose that for all $j = 1, 2, \cdots, J$, there exists $((\alpha_N)_j, (\beta_N)_j) \to (\bar{\alpha}_j, \bar{\beta}_j)$ as $N \to +\infty$, and it holds that

$$\varphi_{\varepsilon_N}((\alpha_N)_j, (\beta_N)_j) = 0.$$  

Observe that

$$\bar{\beta}_j - S_{\mathcal{K}_m_j}(\bar{\beta}_j - \bar{\alpha}_j) = \lim_{N \to +\infty} \varphi((\alpha_N)_j, (\beta_N)_j) = 0,$$  

we have $\mathcal{K}_m_j \ni \bar{\alpha}_j, \bar{\beta}_j \in \mathcal{K}_m_j$, so

$$\lim sup_{N \to +\infty} \mathcal{O}_N \subseteq \mathcal{O}_0.$$  

\[\square\]

**Proposition 1.** Suppose Assumptions 1 and 2 hold. Let $\{\varepsilon_N\}$ be a decreasing sequence of positive real numbers converging to 0. If $\mathcal{J}_x \Psi(\bar{x}, \bar{y})$ is of full row rank, then

(i) $\lim_{N \to +\infty} Z_N = Z_0$, w.p.1,

(ii) $\text{epi} \mathcal{J}_N \to \text{epi} \mathcal{J}$ as $N \to +\infty$, w.p.1.
Proof. (i) We first show that \( \limsup_{N \to +\infty} Z_N \subseteq Z_0 \). For arbitrary \( (\bar{x}, \bar{y}) \in \limsup_{N \to +\infty} Z_N \), w.l.o.g, suppose there exists \( x_N, y_N \in Z_N \) for all \( N \), and \( (x_N, y_N) \to (\bar{x}, \bar{y}) \), w.p.1 as \( N \to +\infty \). For all \( j = 1, 2, \cdots, J \), by Lemma 3.1 and continuity of \( \Phi(\cdot, \cdot) \), there exists \( \delta > 0 \) such that \( \limsup_{N \to +\infty} x_N \subseteq Z_0 \). As \( \beta N \) is also of full row rank, by the Clarke’s implicit theorem, there exist positive numbers \( \varepsilon, \delta \) and a Lipschitz continuous single-valued function \( z(\cdot, \cdot) : \mathbb{B}(\bar{\alpha}, \bar{\beta}) \to \mathbb{B}(\bar{\alpha}, \bar{\beta}) \) with a Lipschitz constant \( c > 0 \). For any \( (\alpha, \beta) \in \mathbb{B}_\delta(\bar{\alpha}, \bar{\beta}) \),

\[
0 = \Phi(z(\alpha, \beta)) - \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = 0
\]  

(11)

and \( \bar{z}(\bar{\alpha}, \bar{\beta}) \). Defining \( \bar{G}_N(x, y) := \begin{pmatrix} \Psi(x, y) - \bar{F}_N(x, y) + \alpha N \\ \beta N \end{pmatrix} \) and \( \delta' = \min\{\delta, (2c)^{-1} \varepsilon\} \). According to Lemma 3.1, for sufficiently large \( N \) it holds w.p.1 that

\[
\max_{(x_N, y_N) \in \mathbb{B}_\delta(\bar{x}, \bar{y})} \|G_N(x, y) - \begin{pmatrix} \bar{\alpha} \\ \bar{\beta} \end{pmatrix} \| \leq \delta',
\]

which means \( G_N(x, y) \in \mathbb{B}_\delta(\bar{\alpha}, \bar{\beta}) \) and for \( (x, y) \in \mathbb{B}_\varepsilon(\bar{x}, \bar{y}) \),

\[
\|z(G_N(x, y)) - z(\bar{\alpha}, \bar{\beta})\| \leq c\|G_N(x, y) - (\bar{\alpha}, \bar{\beta})\| \leq c\delta' < \frac{\varepsilon}{2}.
\]

Then \( z(G_N(\cdot, \cdot)) : \mathbb{B}_\varepsilon(\bar{x}, \bar{y}) \to \mathbb{B}_\varepsilon(\bar{x}, \bar{y}) \) is a continuous function from a convex compact set to itself. Applying Brouwer’s Fixed Point Theorem, there exists a fixed point \( (x_N, y_N) \in \mathbb{B}_\varepsilon(\bar{x}, \bar{y}) \) such that \( z(G_N(x_N, y_N)) = (x_N, y_N) \). By choosing \( \{x_N\} \) a decreasing sequence approaching to \( 0 \) as \( N \) increasing, it holds that \( (x_N, y_N) \to (\bar{x}, \bar{y}) \) as \( N \to +\infty \).

Finally, we show \( (x_N, y_N) \in Z_N \). For sufficiently large \( N \), \( G_N(x_N, y_N) \in \mathbb{B}_\delta(\bar{\alpha}, \bar{\beta}) \) holds. By (11),

\[
0 = \Phi(z(G_N(x_N, y_N))) - G_N(x_N, y_N) = \begin{pmatrix} \bar{F}_N(x_N, y_N) - \alpha N \\ y_N - \beta_N \end{pmatrix}.
\]
Together with \((\alpha_N, \beta_N) \in \mathcal{O}_N\), it holds that \((x_N, y_N) \in Z_N\).

(ii) According to Proposition 7.2 in [14], the conclusion holds.

\[\square\]

**Theorem 3.3.** Suppose Assumptions 1 and 2 hold, \(\kappa_0\) is finite. Let \(\{\varepsilon_N\}\) be a decreasing sequence of positive real numbers converging to 0. Let \((x_N, y_N) \in Z_N\) be an optimal solution of Problem (2) for each \(N\), \((\bar{x}, \bar{y})\) be an accumulation point of the sequence \(\{(x_N, y_N)\}\). If \(J_x \Psi(\bar{x}, \bar{y})\) is of full row rank, then \((\bar{x}, \bar{y})\) is an optimal solution of Problem (1) w.p.1.

**Proof.** Since \(\text{epi} \tilde{f}_N \to \text{epi} \tilde{f}\) w.p.1 as \(N \to +\infty\) and \(\kappa_0\) is finite, according to Theorem 7.31 in [14],

\[
\lim_{N \to +\infty} \arg \min_{\tilde{f}_N} \subseteq \arg \min_{\tilde{f}} \text{w.p.1.}
\]

As \((\bar{x}, \bar{y}) \in \lim sup_{N \to +\infty} \arg \min_{\tilde{f}_N}\), \((\bar{x}, \bar{y})\) is an optimal solution of Problem (1) w.p.1.

The LICQ condition for classical nonlinear programming is stable in the sense that if it holds at one point, then it holds in a small enough neighbourhood of the point. For SOCMPCC-LICQ, similar property holds (cf. Lemma 4.1 in [26]).

**Lemma 3.4.** Suppose Assumptions 1 and 2 hold. Let \((\bar{x}, \bar{y})\) be the accumulation point of some \((x_N, y_N) \in Z_N\), \(N = 1, 2, \cdots, +\infty\). If SOCMPCC-LICQ holds at \((\bar{x}, \bar{y}) \in Z_0\), then there exists a positive integer \(\bar{N}\) and a neighbourhood \(U(\bar{x}, \bar{y})\) of \((\bar{x}, \bar{y})\), such that for each \(N > \bar{N}\), the LICQ for Problem (2) holds at all \((x_N, y_N) \in U(\bar{x}, \bar{y}) \cap Z_N\) almost surely.

**Proof.** Since \(\psi_{\varepsilon}(y_j - \widehat{F}^N_j(x, y)) = \sqrt{(y_j - \widehat{F}^N_j(x, y))^2 + 4 \varepsilon^2 e_j}\), then

\[
\varphi_{\varepsilon}(\widehat{F}^N_j(x, y)) = \widehat{F}^N_j(x, y) + y_j - \psi_{\varepsilon}(y_j - \widehat{F}^N_j(x, y)).
\]

Similar to the proof of case (ii) in Lemma 2.13, it holds that

\[
J \varphi_{\varepsilon}(\widehat{F}^N_j(x, y), y_j) = (I + L_{\psi_{\varepsilon}(y_j - \widehat{F}^N_j(x, y))} L_{y_j - \widehat{F}^N_j(x, y)}) J \widehat{F}^N_j(x, y)
\]

\[+ (I - L_{\psi_{\varepsilon}(y_j - \widehat{F}^N_j(x, y))} L_{y_j - \widehat{F}^N_j(x, y)}) J y_j.\]

For given i.i.d. random variables \(\xi^1, \xi^2, \ldots, \xi^N\), define matrix valued mapping \(C\) on \(\cup_{N=1}^\infty (Z_N \times \{\varepsilon_N\})\) by

\[
C(x, y, \varepsilon) := \begin{pmatrix}
J \widehat{F}^N_{\mathcal{O}_1 \cup \mathcal{O}_N}(x, y) \\
J y_N \cup \mathcal{O}_N
\end{pmatrix}
\]

\[+ \begin{pmatrix}
(I + L_{\psi_{\varepsilon}(y_j - \widehat{F}^N_j(x, y))} L_{y_j - \widehat{F}^N_j(x, y)})^{-1} (I - L_{\psi_{\varepsilon}(y_j - \widehat{F}^N_j(x, y))} L_{y_j - \widehat{F}^N_j(x, y)}) J y_j & j \in \mathcal{N}_1 \\
0 & j \in \mathcal{N}_2 \\
(I - L_{\psi_{\varepsilon}(y_j - \widehat{F}^N_j(x, y))} L_{y_j - \widehat{F}^N_j(x, y)})^{-1} (I + L_{\psi_{\varepsilon}(y_j - \widehat{F}^N_j(x, y))} L_{y_j - \widehat{F}^N_j(x, y)}) J \widehat{F}^N_j(x, y) & j \in \mathcal{N}_3
\end{pmatrix}.
\]

Apparently \(C\) is continuous for variables \(x, y, \varepsilon\) over the set where it is defined. Note that as \((x, y, \varepsilon) \to (\bar{x}, \bar{y}, 0)\), the second term of \(C\) tends to zero according to Lemma 2.11, so \(C(x, y, \varepsilon) \to \hat{C}(\bar{x}, \bar{y})\), where \(\hat{C}(\bar{x}, \bar{y}) = \begin{pmatrix} J \Psi_{\mathcal{O}_1 \cup \mathcal{O}_N}(\bar{x}, \bar{y}) \\
J y_N \cup \mathcal{O}_N\end{pmatrix}\) is of full row rank w.p.1, since SOCMPCC-LICQ holds. By defining \(C(\bar{x}, \bar{y}, 0) := \hat{C}(\bar{x}, \bar{y})\), the definition of \(C\) can be extended continuously to set \(\cup_{N=1}^\infty (Z_N \times \{\varepsilon_N\}) \cup \{(\bar{x}, \bar{y}, 0)\} \).
Since the rank of matrix is stable when its elements are perturbed, there exists a small enough neighbourhood $U(\bar{x}, \bar{y})$ and a positive integer $\overline{N}$ such that for all $(x_N, y_N) \in U(\bar{x}, \bar{y}) \cap Z_N$ and $N > \overline{N}$, $C(x_N, y_N, \varepsilon_N)$ is of full row rank almost surely.

Consider $\sigma_N \in \mathbb{R}^m$ satisfying

$$\mathcal{J} \hat{\Phi}^T_{\varepsilon_N}(x_N, y_N) \sigma_N = 0.$$  

It holds that

$$\mathcal{J} \hat{\Phi}^T_{\varepsilon_N}(x_N, y_N) \sigma_N = C(x_N, y_N, \varepsilon_N)^T \begin{pmatrix} (I + L_{a_jN} L_{W_{jN}}^{-1}) \sigma_{jN} & j \in \mathcal{N}_1 \cup \mathcal{N} \\ (I - L_{a_jN} L_{W_{jN}}^{-1}) \sigma_{jN} & j \in \mathcal{N}_2 \cup \mathcal{N} \end{pmatrix}.$$  

By Lemma 2.13, $(I + L_{a_jN} L_{W_{jN}}^{-1})$, $j \in \mathcal{N}_1 \cup \mathcal{N}$ and $(I - L_{a_jN} L_{W_{jN}}^{-1})$, $j \in \mathcal{N}_2 \cup \mathcal{N}$ are invertible for sufficiently large $N$, then $\sigma_N = 0$, this implies the independence of columns of $\mathcal{J} \hat{\Phi}^T_{\varepsilon_N}(x_N, y_N)$.

In the following context, we consider the relationship of stationary points between Problems (1) and (2). Define the Lagrangian function of Problem (2) by

$$L_{\varepsilon}(x, y, \sigma) = \hat{f}_N(x, y) + \hat{\Phi}^T_{\varepsilon_N}(x, y) \sigma,$$

where $\sigma \in \mathbb{R}^m$. If with probability one, $(x_N, y_N, \sigma_N) \in Z_N \times \mathbb{R}^m$ satisfies

$$\nabla_{(x,y)} L_{\varepsilon}(x_N, y_N, \sigma_N) = \nabla_{\hat{f}_N}(x_N, y_N) + \nabla_{\hat{\Phi}_{\varepsilon_N}} (x_N, y_N) \sigma_N = 0,$$

we say that $(x_N, y_N, \sigma_N)$ is a KKT pair of Problem (2) almost surely. Note that all elements in $(x_N, y_N, \sigma_N)$ satisfying the equation above are implicit functions of $\xi$, i.e., all of them are considered as random vectors. Similarly, the accumulation point of $(x_N, y_N, \sigma_N)$, say $(\bar{x}, \bar{y}, \bar{\sigma})$, is a function of $N$ samples of $\xi$, i.e., it is a random vector.

We say that $(x_N, y_N, \sigma_N)$ satisfies the second-order necessary condition of Problem (2) almost surely if it is a KKT pair almost surely and with probability one,

$$d_N^T \nabla^2_{x,y} L_{\varepsilon}(x_N, y_N, \sigma_N) d_N \geq 0, \quad \forall d_N \in C_{\varepsilon_N}(x_N, y_N),$$

where

$$C_{\varepsilon_N}(x_N, y_N) := \{d_N \in \mathbb{R}^m \mid \mathcal{J} \hat{f}_N(x_N, y_N) d_N \leq 0, \mathcal{J} \hat{\Phi}_{\varepsilon_N}(x_N, y_N) d_N = 0\}$$

is the critical cone of Problem (2). Suppose $\sigma_N = (\sigma_{1N}; \sigma_{2N}; \cdots; \sigma_{JN})$, where $\sigma_{jN} \in \mathbb{R}^m$. Under Assumptions 1 and 2, for $j = 1, 2, \cdots, J$,

$$\lim_{N \to +\infty} a_{jN} = \bar{a}_j \quad \text{w.p.1.}$$

Similar to the proof in Lemmas 4.1 and 4.2 in [26],

$$\nabla \hat{\Phi}_{\varepsilon_N} \sigma_N = \sum_{j=1}^J \nabla \hat{P}_j^N(x_N, y_N)(I + L_{a_jN} L_{W_{jN}}^{-1}) + \nabla y_j^N(I - L_{a_jN} L_{W_{jN}}^{-1}) \sigma_{jN}, \quad (12)$$
and
\[
\nabla^2_{(x,y)} f_{\varepsilon_N}(x_N, y_N, \sigma_N) = \nabla^2 \tilde{f}_N + \sum_{j=1}^{J} N_j \nabla \tilde{f}_N + \sum_{j=1}^{J} \sum_{i=1}^{m_j} \nabla^2 \tilde{F}_j^{i}(I + L_{a_j N} L_{w_{j N}}^{-1}) \sigma_j N_i,
\]

where for \( j = 1, 2, \ldots, J \), \( \tilde{F}_j^{i} \) is the \( i \)-th component of \( \tilde{F}_j^{N}(x_N, y_N) \), other notations are defined as
\[
\begin{align*}
S_{j N} &= L_{w_{j N}}^{-1} \sigma_j N, \\
N_j &= \mathcal{J} y_j^N - \mathcal{J} \tilde{F}_j^{N}, \\
M_{e_j} &= L_{a_j N} L_{w_{j N}}^{-1} L_{S_{j N}} L_{w_{j N}}^{-1} L_{a_j N} - L_{S_{j N}}.
\end{align*}
\]

The following results show the relationship of the stationary points between Problems (1) and (2), and the proof here is similar to Theorem 4.2 in [26]. To keep the integrity of the context, we provide the proof below.

**Theorem 3.5.** Suppose Assumptions 1 and 2 hold. If \((x_N, y_N, \sigma_N)\) is a KKT pair of Problem (2) almost surely, and \((x_N, y_N, \sigma_N) \rightarrow (\bar{x}, \bar{y}, \bar{\sigma})\) as \( N \rightarrow +\infty \), then \((\bar{x}, \bar{y})\) is a C-stationary point of Problem (1) almost surely.

**Proof.** With probability 1, KKT pair \((x_N, y_N, \sigma_N)\) satisfies that
\[
\nabla \tilde{f}_N(x_N, y_N) + \nabla \Phi_{\varepsilon_N}(x_N, y_N) \sigma_N = 0,
\]

According to (12), it is proper to define that
\[
\bar{u}_j := \begin{cases}
\lim_{N \rightarrow +\infty} (I + L_{a_j N} L_{w_{j N}}^{-1}) \sigma_j N \quad j \in \mathcal{N}_1 \cup \mathcal{N}, \\
0 & j \in \mathcal{N}_2,
\end{cases}
\]

\[
\bar{v}_j := \begin{cases}
\lim_{N \rightarrow +\infty} (I - L_{a_j N} L_{w_{j N}}^{-1}) \sigma_j N \quad j \in \mathcal{N}_2 \cup \mathcal{N}, \\
0 & j \in \mathcal{N}_1,
\end{cases}
\]

then
\[
\bar{u}_j + \bar{v}_j = \begin{cases}
\lim_{N \rightarrow +\infty} (I + L_{a_j N} L_{w_{j N}}^{-1}) \sigma_j N = 2\bar{\sigma}_j \quad j \in \mathcal{N}_1, \\
\lim_{N \rightarrow +\infty} (I - L_{a_j N} L_{w_{j N}}^{-1}) \sigma_j N = 2\bar{\sigma}_j \quad j \in \mathcal{N}_2, \\
\lim_{N \rightarrow +\infty} 2\sigma_j N = 2\bar{\sigma}_j \quad j \in \mathcal{N}.
\end{cases}
\]

Consequently, we obtain that \( \bar{\sigma}_j = \frac{1}{2}(\bar{u}_j + \bar{v}_j) \) for \( j = 1, 2, \ldots, J \).

Let \( N \rightarrow +\infty \), by Lemma 2.11 and Lemma 2.13, equations (14)-(17) imply that
\[
\nabla \mathbb{E}[f(\bar{x}, \bar{y}, \xi(\omega))] + \sum_{j=1}^{J} \nabla \Psi_j(\bar{x}, \bar{y}) \bar{u}_j + \sum_{j=1}^{J} \nabla \bar{y}_j \bar{v}_j = 0 \quad \text{w.p.1},
\]

and
\[
\begin{align*}
\bar{v}_j &= 0 \quad j \in \mathcal{N}_1 \quad \text{w.p.1}, \\
\bar{u}_j &= 0 \quad j \in \mathcal{N}_2 \quad \text{w.p.1}, \\
-\bar{u}_j &= \partial_B S_j(\bar{a}_j)(-\bar{u}_j - \bar{v}_j) \quad j \in \mathcal{N}_3 \quad \text{w.p.1}, \\
-\bar{u}_j &= \partial S_j(\bar{a}_j)(-\bar{u}_j - \bar{v}_j) \quad j \in \mathcal{N}_4 \cup \mathcal{N}_5 \cup \mathcal{N}_6 \quad \text{w.p.1}.
\end{align*}
\]

This means that \((\bar{x}, \bar{y})\) is a C-stationary point w.p.1. ∎
Corollary 1. Suppose Assumptions 1 and 2 hold, and \((x_N, y_N)\) be a stationary point of Problem (2) almost surely such that \((x_N, y_N) \rightarrow (\bar{x}, \bar{y})\) as \(N \rightarrow +\infty\). If SOCMPCC-LICQ holds at \((\bar{x}, \bar{y})\), then \((\bar{x}, \bar{y})\) is a C-stationary point of Problem (1) almost surely.

Proof. It follows from the SOCMPCC-LICQ at \((\bar{x}, \bar{y})\) and Lemma 3.4, for any \(N > 0\), LICQ holds almost surely at \((x_N, y_N)\) of Problem (2). Therefore, there exists a unique Lagrange multiplier \(\sigma_N = (\sigma_{1N}; \sigma_{2N}; \ldots; \sigma_{JN}) \in \mathbb{R}^m\) with \(\sigma_{jN} \in \mathbb{R}^{m_j}\), \(j = 1, 2, \cdots, J\) satisfying

\[
\nabla \hat{f}_N(x_N, y_N) + \nabla \Phi_{\varepsilon_N}(x_N, y_N)\sigma_N = 0.
\]

By Lemma 3.4,

\[
-\nabla \hat{f}_N(x_N, y_N) = C(x_N, y_N, \varepsilon_N)^T (u_{jN})
\]

where \(u_{jN} = (I + L_{a_{jN}}L_{W_{jN}}^{-1})\sigma_{jN}, j \in N_1 \cup N\) and \(v_{jN} = (I - L_{a_{jN}}L_{W_{jN}}^{-1})\sigma_{jN}, j \in N_2 \cup N\). Since \(C(x_N, y_N, \varepsilon_N)\) is full of row rank for \(N > 0\), then sequences \(\{u_{jN}\}_{N_1 \cup N}\) and \(\{v_{jN}\}_{N_2 \cup N}\) converge. Let

\[
\bar{u}_j := \left\{ \begin{array}{ll}
\lim_{N \rightarrow +\infty} u_{jN} & j \in N_1 \cup N, \\
0 & j \in N_2,
\end{array} \right.
\]

\[
\bar{v}_j := \left\{ \begin{array}{ll}
\lim_{N \rightarrow +\infty} v_{jN} & j \in N_2 \cup N, \\
0 & j \in N_1,
\end{array} \right.
\]

then \(\bar{u}_j + \bar{v}_j = \left\{ \begin{array}{ll}
\lim_{N \rightarrow +\infty} (I + L_{a_{jN}}L_{W_{jN}}^{-1})\sigma_{jN} = \lim_{N \rightarrow +\infty} 2\sigma_{jN} & j \in N_1, \\
\lim_{N \rightarrow +\infty} (I - L_{a_{jN}}L_{W_{jN}}^{-1})\sigma_{jN} = \lim_{N \rightarrow +\infty} 2\sigma_{jN} & j \in N_2, \\
\lim_{N \rightarrow +\infty} 2\sigma_{jN} & j \in N.
\end{array} \right.\)

Thus sequence \(\{\sigma_N\}\) converges and \(\lim_{N \rightarrow +\infty} \sigma_{jN} = \bar{\sigma}_j = \frac{1}{2}(\bar{u}_j + \bar{v}_j), j = 1, 2, \cdots, J\).

The rest is similar to the proof of Theorem 3.5. \(\square\)

Theorem 3.6. Suppose Assumptions 1 and 2 hold. Let \((x_N, y_N, \sigma_N)\) be a KKT pair of Problem (2) almost surely for each \(N\), and \((x_N, y_N, \sigma_N) \rightarrow (\bar{x}, \bar{y}, \bar{\sigma})\) as \(N \rightarrow +\infty\). If in addition it holds that

(i) point \((x_N, y_N, \sigma_N)\) satisfies the second-order necessary condition of Problem (2) almost surely for each \(N\),

(ii) SOCMPCC-LICQ holds at \((\bar{x}, \bar{y})\),

(iii) the strictly complementary condition of Problem (1) holds at \((\bar{x}, \bar{y})\), i.e., \(\bar{y} - \Psi(\bar{x}, \bar{y}) \neq 0\),

then \((\bar{x}, \bar{y})\) is an M-stationary point of Problem (1) almost surely.

Proof. For \((\bar{x}, \bar{y})\), define multipliers by (15) and (16). By Theorem 3.5, \((\bar{x}, \bar{y})\) is a C-stationary point of Problem (1) w.p.1. Suppose that with probability \(\eta \neq 0\), \((\bar{x}, \bar{y})\) is not an M-stationary point of Problem (1), i.e.,

\[
-\bar{u}_{j0} \in \partial S_{j0}(\bar{y}_{j0} - \Psi_{j0}(\bar{x}, \bar{y}))\left[\bar{u}_{j0} + \bar{v}_{j0}\right],
\]

\[-\bar{u}_{j0} \notin D^* S_{j0}(\bar{y}_{j0} - \Psi_{j0}(\bar{x}, \bar{y}))\left[\bar{u}_{j0} + \bar{v}_{j0}\right].
\]

Comparing the exact expression of \(\partial S_{j0}\) and \(D^* S_{j0}\) (the explicit expression of \(D^* S_{j0}\) is provided in [26]), the only possible difference happens when \(j_0 \in N_4 \cup N_5\). W.l.o.g., for sequence \(\{(x_N, y_N)\}\), we can assume that at least one of the following two cases happens with probability \(\eta \neq 0\):
Case 1: For $j_0 \in \mathcal{N}_4$, $\bar{a}_{j_0} \in bd K_m \setminus \{0\}$, $\sigma_{j_0}^T c_j(\bar{a}_{j_0}) > 0$,
\[
\lim_{N \to +\infty} L_{W_j^N}^{-1} L_{a_j^N} = I + (1 - r) \left( -\frac{\sigma_{j_0}^T}{\|\sigma_{j_0}\|^2} \right), \quad 0 < r < 1.
\]

Case 2: For $j_0 \in \mathcal{N}_5$, $\bar{a}_{j_0} \in -bd K_m \setminus \{0\}$, $\sigma_{j_0}^T c_j(\bar{a}_{j_0}) > 0$,
\[
\lim_{N \to +\infty} L_{W_j^N}^{-1} L_{a_j^N} = -I + r \left( \frac{\sigma_{j_0}^T}{\|\sigma_{j_0}\|^2} \right), \quad 0 < r < 1.
\]

We are going to show that the second-order condition is violated when either of the above two cases happens. By (12), the critical cone can be expressed as
\[
C_{\epsilon_N}(x_N, y_N) = \left\{ d_N \in \mathbb{R}^{2m} \mid \begin{array}{l}
\mathcal{J}\hat{f}_N(x_N, y_N)d_N \leq 0, \\
((I + L_{W_j^N}^{-1} L_{a_j^N}) \mathcal{J}\hat{f}_N(x_N, y_N) \\
+ (I - L_{W_j^N}^{-1} L_{a_j^N}) \mathcal{J}y_j^N)d_N = 0,
\end{array} \right\}.
\] (18)

A bounded non-zero sequence $\{d_N\} \in C_{\epsilon_N}(x_N, y_N)$ can be constructed in the following way. Let
\[
A_N := \begin{pmatrix}
\mathcal{J}\hat{f}_N(x_N, y_N) \\
\mathcal{J}\hat{y}_j(x_N, y_N)
\end{pmatrix}
= \begin{pmatrix}
(I + L_{W_j^N}^{-1} L_{a_j^N})^{-1} (I - L_{W_j^N}^{-1} L_{a_j^N}) \mathcal{J}y_j^N & j \in \mathcal{N}_1 \\
(I - L_{W_j^N}^{-1} L_{a_j^N})^{-1} (I + L_{W_j^N}^{-1} L_{a_j^N}) \mathcal{J}\hat{f}_N(x_N, y_N) & j \in \mathcal{N}_2
\end{pmatrix},
\]

Select some $j_0 \in \mathcal{N}_4 \cup \mathcal{N}_5$, and let
\[
b_N = \begin{pmatrix}
0 & j \in \mathcal{N}_1 \\
0 & j \in \mathcal{N}_2 \\
-(I - L_{W_j^N}^{-1} L_{a_j^N}) \tau_N & j = j_0 \\
0 & j \in \mathcal{N}_2 \\
0 & j \in \mathcal{N}_2 \\
(I + L_{W_j^N}^{-1} L_{a_j^N}) \tau_N & j = j_0
\end{pmatrix},
\]

where $\{\tau_N\} \subseteq \mathbb{R}^{m_0}$ is a bounded sequence and $\tau_N$ converges to $\bar{r} \neq 0$ as $N \to +\infty$. According to Lemma 2.11 and Lemma 3.1, $A_N$ converges to $C(\bar{x}, \bar{y})$ almost surely as $N \to +\infty$. Because SOCMPCC-LICQ holds, i.e., $C(\bar{x}, \bar{y})$ has full rank, when $N$ is large enough, $A_N$ is also of full row rank almost surely. Therefore, w.p.1, there exists $d_N$ satisfying $A_N d_N = b_N$, which means $\mathcal{J}\hat{f}_N(x_N, y_N)d_N = 0$, then $\mathcal{J}\hat{f}_N(x_N, y_N)d_N = -\sigma_{j_0}^T \mathcal{J}\hat{f}_N(x_N, y_N)d_N = 0$. Thus $d_N \in C_{\epsilon_N}(x_N, y_N)$. Note that $d_N \neq 0$ as $b_N \neq 0$. We will demonstrate that sequence $\{d_N\}$ can be bounded.
Sequence \( \{ b_N \} \) converges to \( b \) as \( N \to +\infty \) where
\[
b = \begin{pmatrix}
0 & j \in N_1 \\
0 & j \in N \backslash j_0 \\
-(I - \lim_{N \to +\infty} L_{W_jN}^{-1} L_{a_jN})^\top & j = j_0 \\
0 & j \in N_2 \\
0 & j \in N \backslash j_0 \\
(I + \lim_{N \to +\infty} L_{W_jN}^{-1} L_{a_jN})^\top & j = j_0
\end{pmatrix}
\]
As \( \tilde{C}(\bar{x}, \bar{y}) \) is of full row rank, there exists \( \bar{d} \) with \( \tilde{C}(\bar{x}, \bar{y})\bar{d} = b \). Let \( T(z) = \tilde{C}(\bar{x}, \bar{y})z - b \), \( G(z, p) = T(z) - p \), then \( G(\bar{d}, 0) = 0 \) and \( JG(z, p) = (\tilde{C}(\bar{x}, \bar{y}), -I) \) is of full row rank due to \( \tilde{C}(\bar{x}, \bar{y}) \) of full row rank. Similar to the proof in Proposition 1, there exist positive numbers \( \varepsilon, \delta \) and a Lipschitz continuous function \( z(\cdot) : \mathbb{B}_\delta(\bar{d}) \) almost surely. Applying Lemma 2.2, there exists \( d_N \in U \) satisfying \( A_N d_N = b_N \) almost surely. Second-order necessary condition holds at \( (x_N, y_N, \sigma_N) \) almost surely, one has that
\[
d_N^T \nabla^2_{(x_N, y_N)} L_{\varepsilon, N}(x_N, y_N, \sigma_N) d_N \geq 0. \tag{19}
\]
As \( d_N \in C_{\varepsilon}(x_N, y_N) \), it holds that for all \( j = 1, 2, \ldots, J \),
\[
(J \tilde{F}_j^N + J y_j^N) d_N = L_{W_jN}^{-1} L_{a_jN} (J y_j^N - J \tilde{F}_j^N) d_N.
\]
Direct calculation shows that
\[
d_N^T (J y_j^N - J \tilde{F}_j^N)^T M_j (J y_j^N - J \tilde{F}_j^N) d_N
= d_N^T (J y_j^N - J \tilde{F}_j^N)^T (-L_{S_jN}) (J y_j^N - J \tilde{F}_j^N) d_N
+ d_N^T (J y_j^N - J \tilde{F}_j^N)^T L_{S_jN} (J y_j^N - J \tilde{F}_j^N) d_N
= 4d_N^T (J \tilde{F}_j^N)^T L_{S_jN} J y_j^N d_N.
\]
By Combining above results to (13), the left hand side of inequality (19) is equal to
\[
d_N^T \nabla f_N(x_N, y_N) d_N
- 4 \sum_{j \in N_1} d_N^T (J y_j^N)^T (I - L_{a_jN} L_{W_jN}^{-1}) (I + L_{a_jN} L_{W_jN}^{-1})^{-1} L_{S_jN} J y_j^N d_N
- 4 \sum_{j \in N_2} d_N^T (J \tilde{F}_j^N)^T L_{S_jN} (I + L_{W_jN}^{-1} L_{a_jN}) (I - L_{W_jN}^{-1} L_{a_jN})^{-1} J \tilde{F}_j^N d_N
- 4 \tau_N^2 (I - L_{a_{jN}} L_{W_{jN}}^{-1}) L_{S_jN} (I + L_{W_{jN}}^{-1} L_{a_{jN}}) \tau_N
+ \sum_{j=1}^m \sum_{i=1}^{m_j} [(I + L_{a_{jN}} L_{W_{jN}}^{-1}) \sigma_{jN}] d_N^T \nabla^2 \tilde{F}_j^N d_N, \tag{20}
\]
where \( S_{jN} = L_{W_{jN}}^{-1} \sigma_{jN}, j = 1, 2, \ldots, J \).
Because of the convergence of sequences \( \{x_N\}, \{y_N\}, \{\sigma_N\} \) and \( \{\tau_N\} \), together with Assumptions 1 and 2, sequences \( \{\nabla f_N\}, \{a_{jN}\}, \{W_{jN}\}, \{Jy_N^j\}, \{\tilde{F}_j\} \), \( \{\nabla^2 F_j^N\}, \{\tau_N\} \) and \( \{\sigma_N\} \), \( j = 1, 2, \ldots, J \) are all bounded almost surely. According to Lemma 2.11, \( L_{a_{jN}} L_{W_{jN}^{-1}}^{-1} \) and \( L_{W_{jN}^{-1}}^{-1} L_{a_{jN}} \), \( j = 1, 2, \ldots, J \), are also bounded almost surely.

For \( j \in \mathcal{N}_1, \tilde{a}_j \in \text{int} \mathcal{K}_{m_j} \), \( a_{jN} \) converges to \( \tilde{a}_j \) as \( N \to +\infty \), then

\[
\lim_{N \to +\infty} L_{W_{jN}}^{-1} = L_{\tilde{a}_j}^{-1} = L_{\tilde{a}_j}^{-1}.
\]

Similarly, for \( j \in \mathcal{N}_2 \),

\[
\lim_{N \to +\infty} L_{W_{jN}}^{-1} = L_{\tilde{a}_j}^{-1} = -L_{\tilde{a}_j}^{-1}.
\]

Then we have

\[
\lim_{N \to +\infty} S_{jN} = \begin{cases} L_{\tilde{a}_j}^{-1} \bar{\sigma}_j & j \in \mathcal{N}_1 \\ -L_{\tilde{a}_j}^{-1} \bar{\sigma}_j & j \in \mathcal{N}_2 \end{cases} \quad \text{w.p. 1}
\]

which means the sequence \( \{S_{jN}\} \), \( j \in \mathcal{N}_1 \cup \mathcal{N}_2 \) is bounded almost surely. As a result, all terms in (20) except term \( q_N \) are all bounded almost surely.

In the following discussion, we focus on term \( q_N \) in (20). For simplicity, use the following notations: \( a_N := a_{j_0N} = (a_{1N}; a_{2N}) = y_{j_0N} - \tilde{F}_{j_0N}(x_N, y_N), \bar{\sigma} := \tilde{y}_{j_0} - \Psi_{j_0}(\bar{x}, \bar{y}) = (\bar{a}_1; \bar{a}_2), w_N := W_{j_0N} = (w_{1N}; w_{2N}), s_N := S_{j_0N} = (s_{1N}; s_{2N}), \sigma_N := \sigma_{j_0N} = (\sigma_1; \sigma_2), \]

\[
L_{w_N}^{-1} = \frac{1}{\det(w_N)} \begin{pmatrix} w_{1N} & -w_{2N}^T \\ w_{2N} & w_{1N} \end{pmatrix},
\]

\[
s_N = L_{w_N}^{-1} \sigma_N = \frac{1}{\det(w_N)} \begin{pmatrix} w_{1N} \sigma_1 - w_{2N}^T \sigma_2 \\ \sigma_1 w_{2N} + \frac{\det(w_N)}{w_{1N}} \sigma_2 + \frac{w_{2N} w_{1N}^T \sigma_2}{w_{1N}} \end{pmatrix},
\]

\[
\begin{pmatrix} z_{1N} & z_{2N}^T \\ z_{3N} & z_{4N}^T \end{pmatrix} := L_{a_N} L_{w_N}^{-1}, L_{a_N} L_{a_N}^{-1} = \begin{pmatrix} z_{1N} & z_{2N}^T \\ z_{3N} & z_{4N}^T \end{pmatrix},
\]

then

\[
q_N = -\tau_N^T (I - L_{a_N} L_{w_N}^{-1}) L_{s_N} (I + L_{w_N}^{-1} L_{a_N}) \tau_N
\]

\[
= -\tau_N^T L_{s_N} \tau_N + \tau_N L_{a_N} L_{w_N}^{-1} L_{s_N} L_{w_N}^{-1} L_{a_N} \tau_N. \quad (21)
\]

Suppose \( a_N \) has the spectral decomposition

\[
a_N = \lambda_1(a_N) c_1(a_N) + \lambda_2(a_N) c_2(a_N)
\]
and let \( w_N = \sqrt{\alpha_2^2 + 4\epsilon_N^2} \), then
\[
w_N = \sqrt{\lambda_1(a_N)^2 + 4\epsilon_N^2} c_1(a_N) + \sqrt{\lambda_2(a_N)^2 + 4\epsilon_N^2} c_2(a_N),
\]
\[
w_{1N} = \frac{1}{2} \left( \sqrt{\lambda_1(a_N)^2 + 4\epsilon_N^2} + \sqrt{\lambda_2(a_N)^2 + 4\epsilon_N^2} \right) > 0,
\]
\[
w_{2N} = \begin{cases} \frac{1}{2} \left( \sqrt{\lambda_1(a_N)^2 + 4\epsilon_N^2} + \sqrt{\lambda_2(a_N)^2 + 4\epsilon_N^2} \right) \frac{\sigma_{2N}}{\sigma_{2N}} & \|a_{2N}\| \neq 0 \\ 0 & \|a_{2N}\| = 0 \end{cases},
\]
\[
det(w_N) = w_N^2 - \|w_{2N}\|^2 = \sqrt{\lambda_1(a_N)^2 + 4\epsilon_N^2} \sqrt{\lambda_2(a_N)^2 + 4\epsilon_N^2} > 0,
\]
\[
\lim_{N \to +\infty} w_{1N} = \frac{1}{2} (|\lambda_1(\overline{a})| + |\lambda_2(\overline{a})|),
\]
\[
\lim_{N \to +\infty} w_{2N} = \begin{cases} \frac{1}{2} (|\lambda_1(\overline{a})| + |\lambda_2(\overline{a})|) \frac{\sigma_{2N}}{\sigma_{2N}} & \|a_{2N}\| \neq 0 \\ 0 & \|a_{2N}\| = 0 \end{cases},
\]
\[
\lim_{N \to +\infty} \det(w_N) = |\lambda_1(\overline{a})| |\lambda_2(\overline{a})|.
\]

**Case 1.** If \( j_0 \in N_4, \overline{a} \in \text{bd}K_{m_{j_0}} \setminus \{0\}, \overline{\alpha}_1 = \|\overline{\alpha}_2\| \neq 0, \lambda_1(\overline{a}) = 0, \lambda_2(\overline{a}) = 2\overline{a}_1 > 0, \sigma^T c_1(\overline{a}) > 0 \) (i.e., \( \overline{a}_1 \overline{\sigma}_1 - \overline{a}_2^T \overline{\sigma}_2 > 0 \)),
\[
\lim_{N \to +\infty} L_{w_N}^{-1} L_{a_N} = I + (1 - r) \begin{pmatrix} -1 & \frac{\overline{\alpha}_2^T}{\|\overline{\alpha}_2\|^2} \\ \frac{\overline{\alpha}_2}{\|\overline{\alpha}_2\|^2} & \frac{\overline{\alpha}_2^T \overline{\alpha}_2}{\|\overline{\alpha}_2\|^4} \end{pmatrix}, \quad 0 < r < 1,
\]
then \( \lim_{N \to +\infty} z_{1N} = r, \lim_{N \to +\infty} z_{2N} = \lim_{N \to +\infty} z_{3N} = (1 - r) \frac{\overline{\alpha}_2}{\|\overline{\alpha}_2\|} \),
\( \lim_{N \to +\infty} w_{1N} = \overline{a}_1, \lim_{N \to +\infty} w_{2N} = \overline{a}_2, \lim_{N \to +\infty} \det(w_N) = 0 \).

Take \( r_N = \epsilon \) in \( q_N \) in (21), then
\[
q_N = -e^T \begin{pmatrix} s_{1N}^T & s_{2N}^T \end{pmatrix} e + e^T \begin{pmatrix} z_{1N} & z_{2N}^T \\ z_{3N}^T & z_{4N} \end{pmatrix} \begin{pmatrix} s_{1N} & s_{2N}^T \\ s_{2N}^T & s_{1N} \end{pmatrix} e
\]
\[
= (z_{1N}^2 + \|z_{2N}\|^2 - 1)s_{1N} + 2z_{1N} z_{2N}^T s_{2N}^T
\]
\[
= \frac{1}{\det(w_N)} D_N,
\]
where
\[
D_N = (z_{1N}^2 + \|z_{2N}\|^2 - 1)(w_{1N} \sigma_{1N} - w_{2N}^T \sigma_{2N})
\]
\[
+ 2z_{1N} z_{2N}^T (-\sigma_{1N} w_{2N} + \frac{\det(w_N)}{w_{1N}} + \frac{w_{2N} w_{2N}^T \sigma_{2N}}{w_{1N}}).
\]

Let \( N \to +\infty \), then we get
\[
\lim_{N \to +\infty} D_N = [r^2 + (1 - r)^2 - 1](\overline{a}_1 \overline{\sigma}_1 - \overline{a}_2^T \overline{\sigma}_2) + 2r(1 - r)(-\overline{a}_1 \overline{\sigma}_1 + \overline{a}_2^T \overline{\sigma}_2)
\]
\[
= 4r(1 - r)(\overline{a}_1 \overline{\sigma}_1 - \overline{a}_2^T \overline{\sigma}_2) < 0,
\]
which means \( \lim_{N \to +\infty} q_N = -\infty \).

**Case 2.** If \( j_0 \in N_5, \overline{a} \in -\text{bd}K_{m_{j_0}} \setminus \{0\}, -\overline{\alpha}_1 = \|\overline{\alpha}_2\| \neq 0, \lambda_1(\overline{a}) = 2\overline{a}_1 > 0, \lambda_2(\overline{a}) = 0, \sigma^T c_2(\overline{a}) > 0 \) (i.e., \( \overline{a}_1 \overline{\sigma}_1 - \overline{a}_2^T \overline{\sigma}_2 < 0 \)),
\[
\lim_{N \to +\infty} L_{w_N}^{-1} L_{a_N} = -I + r \begin{pmatrix} 1 & \frac{\overline{\alpha}_2}{\|\overline{\alpha}_2\|} \\ \frac{\overline{\alpha}_2}{\|\overline{\alpha}_2\|^2} & \frac{\overline{\alpha}_2^T \overline{\alpha}_2}{\|\overline{\alpha}_2\|^4} \end{pmatrix}, \quad 0 < r < 1,
then \( \lim_{N \to +\infty} z_{1N} = r - 1 \), \( \lim_{N \to +\infty} z_{2N} = \lim_{N \to +\infty} z_{3N} = r \frac{\bar{a}_2}{\|\bar{a}_2\|} \),
\( \lim_{N \to +\infty} w_{1N} = -\bar{a}_1 \), \( \lim_{N \to +\infty} w_{2N} = -\bar{a}_2 \), \( \lim_{N \to +\infty} \det(w_N) = 0 \).

We take \( r_N = e \) in \( q_N \) in (21), then
\( q_N = \frac{1}{\det(w_N)} D_N \),
with
\( \lim_{N \to +\infty} D_N = 4r(r - 1)(-\bar{a}_1 \bar{\sigma}_1 + \bar{a}_2^T \bar{\sigma}_2) < 0 \),
which means \( \lim_{N \to +\infty} q_N = -\infty \).

One of the two cases above that violates the second-order necessary condition, i.e., (19), happens with probability \( \eta \neq 0 \), which contradicts the assumption that second-order condition holds w.p.1. We conclude that \((\bar{x}, \bar{y})\) is an M-stationary point of Problem (1) almost surely.

Corollary 2. Suppose Assumptions 1 and 2 hold. Consider the case that w.p.1, \((x_N, y_N, \sigma_N)\) satisfies the second-order necessary condition of Problem (2) for each \( N \), and \((x_N, y_N) \to (\bar{x}, \bar{y})\) as \( N \to +\infty \). If in addition, SOCPCC-LICQ, strict complementarity and SOC-ULSC conditions hold at \((\bar{x}, \bar{y})\), then \((\bar{x}, \bar{y})\) is an \( S \)-stationary point of Problem (1) almost surely.

Proof. According to Lemma 3.1 in [26], if w.p.1, \((\bar{x}, \bar{y})\) is an M-stationary point of Problem (1) and SOC-ULSC condition holds at \((\bar{x}, \bar{y})\), then \((\bar{x}, \bar{y})\) is an S-stationary point of Problem (1) w.p.1, while M-stationary is achieved by above Theorem 3.6.

4. Numerical illustration. In this section, we apply our algorithm to the following problem:
\[
\min \quad E_{\xi, \zeta} \left[ \frac{1}{2} \| A^T u - c_0 + \xi \|^2 + \frac{1}{2} \| v - v_0 + \zeta \|^2 \right]
\]
\[
s.t. \quad E_{\xi, \zeta} (u + \zeta)^T v = 0
\]
\[
E_{\xi, \zeta} (u + \zeta) \in K_m
\]
\[
v \in K_m.
\]

where \( \xi \) is an \( n \)-dimensional random vector obeys the standard normal distribution, \( \zeta \) is an \( m \)-dimensional random vector obeys the standard normal distribution.

Note that Problem (22) shares the same solutions and stationary points with the following deterministic problem since the expectation in (22) can be obtained explicitly:
\[
\min \quad \frac{1}{2} \| A^T u - c_0 \|^2 + \frac{1}{2} \| v - v_0 \|^2
\]
\[
s.t. \quad u^T v = 0
\]
\[
u \in K_m
\]
\[
v \in K_m.
\]

The problem above is an inverse problem for the following linear second-order cone programming problem
\[
\min \quad c^T x
\]
\[
s.t. \quad Ax - b \in K_m.
\]

The deterministic version, i.e. Problem (23) has been discussed by Zhang [25].
Table 1. Numerical result for Problem (22)

| N  | \( \bar{f} \) | \( \bar{\varepsilon}_u \) | \( \bar{\varepsilon}_v \) | infea | time(s) |
|----|----------------|-----------------|-----------------|-------|--------|
| 1000 | 1.53           | 8.88E-02        | 4.07E-02        | 5.43E-06 | 0.02   |
| 10000 | 1.49           | 5.14E-02        | 2.82E-02        | 3.97E-05 | 0.02   |
| 100000 | 1.54          | 5.74E-02        | 6.77E-02        | 4.74E-03 | 0.02   |
| 1000000 | 1.44          | 3.74E-04        | 5.22E-04        | 6.23E-06 | 0.13   |
| 10000000 | 1.44          | 1.82E-04        | 1.89E-04        | 7.35E-06 | 1.23   |

In our example, we take \( K_m = K_4 \times K_4 \), \( A = \begin{pmatrix} A_1 \\ A_2 \end{pmatrix} \), \( b = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \), \( b_1 = (-1; 1; 0; 0) \), \( b_2 = (-1; 1; 0; 0) \), \( c_0 = (1; 2; 1; 0; 0) \).

\[
A_1 = \begin{pmatrix} 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 1 & -1 & -1 & 0 & 1 \\ 1 & 1 & 0 & 0 & 1 \\ 1 & -1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{pmatrix},
\]

\( v_0 \) is the least square solution of \( A^Tv_0 = c_0 \).

All the experiments are done with MATLAB R2016b running on a PC with a Intel 3.5GHz CPU and 16GB RAM. Smoothing parameter is chosen as a constant \( \varepsilon = 1e^{-7} \). Let \( N \) be the sample size that varies in \( \{10^3, 10^4, 10^5, 10^6, 10^7\} \). For each \( N \), we repeat 30 times with different samplings. \( \bar{f} \) denotes the average of approximate objective function value; \( \bar{\varepsilon}_u \) and \( \bar{\varepsilon}_v \) denote the average 2-norm distance between the corresponding stationary points of the true and approximate solutions for \( u \) and \( v \), respectively; “infea” is used to denote the average infeasibility; “time” represents the runtime of the 30 runs measured by seconds. All the results are presented in Table 1.

This simple example illustrates the convergence of our approach: while \( N \) gets bigger enough, it is observed that the stationary points of the approximate problem converge to the stationary point of the true problem.

5. Conclusion. In this paper, the new SSOCMPCC problem and corresponding smoothing SAA method are proposed. Employ the concept of epi-convergence in variational analysis, the subinvertibility condition and Euclidean Jordan algebra, the almost sure convergence of optimal solutions of the smoothed SAA sub-problem is obtained as the sample size tends to infinity. Under second-order condition, SOCMPCC-LICQ and strict complementarity, we demonstrate that any accumulation point of stationary points of the smoothed SAA sub-problem is an M-stationary point almost surely. S-stationarity can be achieved if in addition assume that the SOC-ULSC condition holds. One simple numerical test is reported to illustrate the applicability of our smoothing SAA approach.

REFERENCES

[1] Ş. İ. Birbil, G. Gürkan and O. Listeş, Solving stochastic mathematical programs with complementarity constraints using simulation, Math. Oper. Res., 31 (2006), 739–760.
[2] B. T. Chen and P. T. Harker, A non-interior-point continuation method for linear complementarity problems, SIAM J. Matrix Anal. Appl., 14 (1993), 1168–1190.
[3] X. J. Chen, H. L. Sun and R. J.-B. Wets, Regularized mathematical programs with stochastic equilibrium constraints: Estimating structural demand models, SIAM J. Optim., 25 (2015), 53–75.
[4] S. Christiansen, M. Patriksson and L. Wynter, Stochastic bilevel programming in structural optimization, *Struct. Multidiscip. Optim.*, **21** (2001), 361–371.

[5] F. H. Clarke, *Optimization and Nonsmooth Analysis*, Second edition, Classics in Applied Mathematics, 5. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1990.

[6] H. Y. Jiang and H. F. Xu, Stochastic approximation approaches to the stochastic variational inequality problem, *IEEE Trans. Autom. Control*, **53** (2008), 1462–1475.

[7] C. Kanzow, Some noninterior continuation methods for linear complementarity problems, *SIAM J. Matrix Anal. Appl.*, **17** (1996), 851–868.

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

[9] G.-H. Lin, M.-J. Luo and J. Zhang, Smoothing and SAA method for stochastic programming problems with non-smooth objective and constraints, *J. Global Optim.*, **66** (2016), 487–510.

[10] G.-H. Lin, M.-J. Luo, D. L. Zhang and J. Zhang, Stochastic second-order-cone complementarity problems: expected residual minimization formulation and its applications, *Math. Program.*, **165** (2017), 197–233.

[11] G.-H. Lin, H. F. Xu and M. Fukushima, Monte Carlo and quasi-Monte Carlo sampling methods for a class of stochastic mathematical programs with equilibrium constraints, *Math. Method Oper. Res.*, **67** (2008), 423–441.

[12] Y. C. Liu and G.-H. Lin, Convergence analysis of a regularized sample average approximation method for stochastic mathematical programs with complementarity constraints, *Asia Pac. J. Oper. Res.*, **28** (2011), 755–771.

[13] Y. C. Liu, H. F. Xu and J. J. Ye, Penalized sample average approximation methods for stochastic mathematical programs with complementarity constraints, *Math. Oper. Res.*, **36** (2011), 670–694.

[14] R. T. Rockafellar and R. J.-B. Wets, *Variational Analysis*, Grundlehren der mathematischen Wissenschaften, 317. Springer-Verlag, Berlin, 1998.

[15] A. Shapiro, D. Dentcheva and A. Ruszczynski, *Lectures on Stochastic Programming*, Society for Industrial and Applied Mathematics, 2009.

[16] S. Smale, Algorithms for solving equations, *Proceedings of the International Congress of Mathematicians, Amer. Math. Soc.*, Providence, RI, **1,2** (1986), 172–195.

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

[18] G. X. Wang, J. Zhang, B. Zeng and G.-H. Lin, Expected residual minimization formulation for a class of stochastic linear second-order cone complementarity problems, *Eur. J. Oper. Res.*, **265** (2018), 437–447.

[19] H. F. Xu, Uniform exponential convergence of sample average random functions under general sampling with applications in stochastic programming, *J. Math. Anal. Appl.*, **368** (2010), 692–710.

[20] H. F. Xu and D. L. Zhang, Smooth sample average approximation of stationary points in nonsmooth stochastic optimization and applications, *Math. Program. Ser. A*, **119** (2009), 371–401.

[21] J. J. Ye, The exact penalty principle, *Nonlinear Anal.*, **75** (2012), 1642–1654.

[22] J. J. Ye and J. C. Zhou, First-order optimality conditions for mathematical programs with second-order cone complementarity constraints, *SIAM J. Optim.*, **26** (2016), 2820–2846.

[23] J. J. Ye and J. C. Zhou, Verifiable sufficient conditions for the error bound property of second-order cone complementarity problems, *Math. Program. Ser. A*, **171** (2018), 361–395.

[24] J. Zhang, L.-W. Zhang and S. Lin, A class of smoothing SAA methods for a stochastic mathematical program with complementarity constraints, *J. Math. Anal. Appl.*, **387** (2012), 201–220.

[25] Y. Zhang, Y. Jiang, L. W. Zhang and J. Z. Zhang, A perturbation approach for an inverse linear second-order cone programming, *J. Ind. Manag. Optim.*, **9** (2013), 171–189.

[26] Y. Zhang, L. W. Zhang and J. Wu, Convergence properties of a smoothing approach for mathematical programs with second-order cone complementarity constraints, *Set-Valued Var. Anal.*, **19** (2011), 609–646.

Received May 2019; revised October 2019.
