NECESSARY OPTIMALITY CONDITIONS FOR TWO-STAGE STOCHASTIC PROGRAMS WITH EQUILIBRIUM CONSTRAINTS*

HUIFU XU[†] AND JANE J. YE[‡]

Abstract. Developing first order optimality conditions for two-stage stochastic mathematical programs with equilibrium constraints (SMPECs) whose second stage problem has multiple equilibria/solutions is a challenging undone work. In this paper we take this challenge by considering a general class of two-stage SMPECs whose equilibrium constraints are represented by a parametric variational inequality (where the first stage decision vector and a random vector are treated as parameters). We use the sensitivity analysis on deterministic mathematical programs with equilibrium constraints (MPECs) as a tool to deal with the challenge: First, we extend a well-known theorem in nonsmooth analysis about the exchange of the subdifferential operator with Aumann's integration from a nonatomic probability space to a general setting; second, we apply the extended result together with the existing sensitivity analysis results on the value function of the deterministic MPEC and the bilevel programming to the value function of our second stage problem; third, we develop various optimality conditions in terms of the subdifferential of the value function of the second stage problem and its relaxations which are constructed through the gradients of the underlying function at the second stage; finally we analyze special cases when the variational inequality constraint reduces to a complementarity problem and further to a system of nonlinear equalities and inequalities. The subdifferential to be used in this paper is the limiting (Mordukhovich) subdifferential, and the probability space is not necessarily nonatomic which means that Aumann's integral of the limiting subdifferential of a random function may be strictly smaller than that of Clarke's.

Key words. stochastic mathematical program with equilibrium constraints, first order necessary conditions, limiting subdifferentials, M-stationary points, random set-valued mappings, sensitivity analysis

AMS subject classifications. 90C15, 90C46, 90C30, 90C31, 90C33

DOI. 10.1137/090748974

1. Introduction. In this paper we study the following two-stage stochastic program:

$$
\begin{aligned}
\min_{x} \quad & f_1(x) + \mathbb{E}\left[v(x, \xi(\omega))\right] \\
\text{subject to (s.t.)} \quad & G(x) \leq 0, \\
& H(x) = 0, \\
& x \in Q,
\end{aligned}
\tag{1.1}
$$

where Q is a nonempty closed subset of $\mathbb{R}^n$, $f_1 : \mathbb{R}^n \to \mathbb{R}$, $G : \mathbb{R}^n \to \mathbb{R}^s$, and $H : \mathbb{R}^n \to \mathbb{R}^r$ are locally Lipschitz continuous, $\xi(\omega)$ is a random vector defined on a probability space $(\Omega, \mathcal{F}, P)$ with support set $\Xi \subset \mathbb{R}^d$, and given $x \in Q, \xi \in \Xi$, $v(x, \xi)$ is the optimal value of the following second stage problem:

$$
\mathcal{P}(x, \xi): \quad
\begin{aligned}
\min_{(y,z) \in \mathbb{R}^l \times \mathbb{R}^m} \quad & f_2(x, y, z, \xi) \\
\text{s.t.} \quad & 0 \in F(x, y, z, \xi) + \mathcal{N}_{\mathcal{C}}(z), \\
& \psi(x, y, z, \xi) \leq 0,
\end{aligned}
\tag{1.2}
$$

where $f_2 : \mathbb{R}^n \times \mathbb{R}^l \times \mathbb{R}^m \times \mathbb{R}^d \to \mathbb{R}$, $F : \mathbb{R}^n \times \mathbb{R}^l \times \mathbb{R}^m \times \mathbb{R}^d \to \mathbb{R}^m$, and $\psi : \mathbb{R}^n \times \mathbb{R}^l \times \mathbb{R}^m \times \mathbb{R}^d \to \mathbb{R}^p$, $\mathcal{C}$ is a nonempty closed subset of $\mathbb{R}^m$, $\mathcal{N}_\mathcal{C}(z)$ denotes the normal cone to $\mathcal{C}$ at $z \in \mathcal{C}$, and $\mathcal{N}_\mathcal{C}(z) := \emptyset$ if $z \notin \mathcal{C}$. The precise definition of the normal cone will be given in section 2. For the simplicity of exposition, we assume throughout this paper that the underlying functions of the second stage problem are continuously differentiable. When the functions are merely locally Lipschitz continuous, optimality conditions similar to those derived in sections 3–5 can be derived in the same manner by using [21, Theorem 3.6 and Corollary 3.7].

This is a two-stage stochastic programming framework for hierarchical decision making under uncertainty in management science and engineering. At the first stage, a decision maker needs to make a decision on x, restricted to the feasible set $X = \{x \in Q : G(x) \leq 0, H(x) = 0\}$, *before* the realization of the random data $\xi(\omega)$. At the second stage, when x is given and a realization $\xi = \xi(\omega)$ is known, an optimal decision on y and z is sought by solving (1.2) with x and ξ being treated as parameters. Since a variational inequality is often used to represent an equilibrium in economics and engineering, the second stage problem is also known as a parametric mathematical program with equilibrium constraints (MPEC), and consequently our model may be called a two-stage stochastic mathematical program with equilibrium constraints (SMPEC).

It is important to note that the second stage problem (1.2) has two decision vectors y and z. Let us use the well-known Stackelberg leader-followers problem to explain this. At the first stage, a leader needs to make an optimal decision at present on its investment or capacity expansion (denoted by x) before realization of uncertainty of market demand (represented by ξ) in the future. The leader expects that, in any future demand scenario at the time when the capacity expansion is completed, the followers will compete for the residual demand (treating the leader's capacity expansion x as given), and they will reach an equilibrium represented by the variational inequality in (1.2). Since there could be a number of possible market equilibria (that is, the equilibrium constraint has multiple solutions), the leader may wish to input some extra resources (represented by y) to influence such equilibria to improve his profit—this reflects the leader's short-term (e.g., daily operational) decision. Note that the leader's additional input (y) does not necessarily drive the followers' competition to a unique equilibrium which he prefers (the equilibrium constraint may have multiple solutions for every y); the simultaneous optimal choice of y and z means that the leader not only tries to intervene a short-term market equilibrium but also takes an optimistic attitude towards the short-term market equilibrium.

Note also that under some moderate conditions, the two-stage SMPEC (1.1)–(1.2) can be written in the following closed form:

$$
\begin{aligned}
\min_{x, y(\cdot), z(\cdot)} \quad & f_1(x) + \mathbb{E}\left[f_2(x, y(\omega), z(\omega), \xi(\omega))\right] \\
\text{s.t.} \quad & G(x) \leq 0, \\
& H(x) = 0, \\
& x \in Q, \\
& 0 \in F(x, y(\omega), z(\omega), \xi(\omega)) + \mathcal{N}_\mathcal{C}(z(\omega)) \text{ for a.e. } \omega, \\
& \psi(x, y(\omega), z(\omega), \xi(\omega)) \leq 0 \text{ for a.e. } \omega.
\end{aligned}
$$

(1.3)

This type of reformulation is well-documented for classical two-stage stochastic programming problems; see Chapters 1 and 2 in the book of Ruszczyński and Shapiro [43]. Patriksson and Wynter [33] first introduced a two-stage SMPEC model in the form (1.3) which consists of two sets of decision variables: the upper/first stage variables

(corresponding to x in our model) and the lower/second stage variables (corresponding to z in our model). They investigated a number of fundamental theoretical issues such as the existence of local and global optimal solutions, the strict convexity of the implicit upper level objective function (a sufficient condition for the uniqueness of the upper level global optimal solution), and the differentiability of the objective function (to facilitate the development of a numerical solution). Over the past few years since the first SMPEC paper, SMPEC has developed as a new area of optimization and operations research primarily driven by its potential in modeling hierarchical decision-making problems in engineering design and management science. For instance, Christiansen, Patriksson, and Wynter [5] proposed a two-stage SMPEC to model a robust and cost-optimizing structural design problem where the optimal design of a linear-elastic structure, for example, a truss topology, is considered under unilateral frictionless contact and under uncertainty in the data describing the load conditions, the material properties, and the rigid foundation. The resulting stochastic bilevel optimization model finds a structural design that responds the best to the given probability distribution in the data. Werner [48] proposed a two-stage stochastic bilevel programming model for studying competition in the Norwegian telecommunication industry [48] which can be reformulated as a two-stage SMPEC when the lower level decision-making problem is convex. During the second revision of this paper, we have seen new applications of two-stage SMPEC models in energy markets and transportation networks. See [49, 47, 60].

On the computational aspects, Shapiro [44] first applied the well-known Monte Carlo sampling method to solve a general class of two-stage SMPECs and presented a convergence analysis of the method in terms of optimal values and global optimal solutions as sample size increases. Along this direction, Shapiro and Xu [45] investigated a particular two-stage SMPEC whose underlying function in the variational constraint is uniformly strongly monotone in z. They established exponential convergence of the method to sharp local optimal solutions and explained how the discretized sample average approximate SMPEC can be solved by a nonlinear programming (NLP) code.

A particularly interesting case of the SMPEC model (1.1)–(1.2) is when the set $\mathcal{C}$ becomes $\mathbb{R}_+^n$, and consequently the equilibrium constraint reduces to a nonlinear complementarity problem and the SMPEC becomes a stochastic mathematical program with complementarity constraint (SMPCC). Lin, Chen, and Fukushima [20] first investigated the SMPCC and proposed an implicit smoothing method for solving the SMPCC where the complementarity is a P_0-linear and the random variable has a finite discrete distribution. A slightly more general SMPCC model was further studied by Xu [51], Xu and Meng [52], and Meng and Xu [24]. In the case when $\mathcal{C} = R^m$, the variational inequality constraint reduces to an equality constraint, and consequently (1.1)–(1.2) become a classical two-stage stochastic program with equality and inequality constraints.

The focus of this paper is on optimality conditions rather than numerical methods although they are essentially related to each other. Assuming that we can obtain a closed form of $\mathbb{E}[v(x, \xi(\omega))]$, then the first stage problem reduces to a deterministic minimization problem. Consequently, we may use certain subdifferentials of $\mathbb{E}[v(x, \xi(\omega))]$ to characterize the first order necessary optimality conditions. This type of *value function approach* is well-known in deterministic MPECs and bilevel programming [58, 55]. If we weaken the assumption by considering the subdifferentials of $v(x, \xi)$ instead of $\mathbb{E}[v(x, \xi)]$, then we may obtain a weaker optimality condition because $\partial \mathbb{E}[v(x, \xi(\omega))]$ is smaller than $\mathbb{E}[\partial_x v(x, \xi(\omega))]$ for many differential operators. These

type of optimality conditions date back to the earlier work by Rockafellar and Wets [41] who derived the so-called *basic Kuhn–Tucker conditions* in terms of the convex subdifferential [40] for a class of two-stage stochastic programs with convex objective and convex constraints, and to the necessary optimality condition derived by Hiriart-Urruty for the nonconvex two-stage stochastic programs in [18]. More recently, Ralph and Xu [35] derived some first order optimality conditions for the classical two-stage stochastic minimization problems in terms of Clarke subdifferentials of the value function of the second stage problem, and by using Gauvin and Dubeau's sensitivity results for the value function of parametric programming [10], they also derived a so-called *relaxed* optimality condition for the first stage problem where the Clarke subdifferential of the value function at the second stage is approximated by a collection of the gradients of the Lagrange function of the second stage problem at stationary points.

In the context of SMPECs, Xu and Meng [52] considered a weak optimality condition in terms of Clarke subdifferentials for a class of two-stage stochastic programming problems with nonsmooth equality constraints and applied it to an SMPCC which has a *unique* feasible solution in the second stage. It is well-known that the value function of a parametric MPEC is often nonconvex, and hence the Clarke subdifferential may be large under some circumstances. Over the past few decades, a number of subdifferentials smaller than the Clarke subdifferential have been developed. A popular one is called the *limiting subdifferential* (which is also known under various names such as basic subdifferential [27, 28], Mordukhovich subdifferential, and general subdifferential [42]). Using the limiting subdifferential, various first order optimality conditions for a range of deterministic MPECs and bilevel programming have been studied by a number of researchers including Henrion, Kanzow, Mordukhovich, Outrata, Treiman, Ye, Zhang, and their collaborators; see, e.g., Ye and Ye [57], Ye [54, 56], Outrata [30, 31], Mordukhovich [28], and the references therein. These optimality conditions are significantly sharper than those presented in terms of Clarke subdifferentials. In particular, when the equilibrium constraint reduces to a complementarity constraint, the optimality conditions lead to the well-known Mordukhovich stationary points (M-stationary points) in the literature of MPECs. Outrata and Römisch [32, Theorem 3.5] apparently first used the limiting subdifferential to derive first order optimality conditions for classical two-stage stochastic programming problems, and their focus is on the case when the probability space of the underlying random variables is nonatomic.

The research of this paper is inspired by the sensitivity analysis of value functions and optimality conditions in [21, 22, 54, 57] in that our second stage problem (1.2) is a parametric MPEC. Specifically, we would like to use the existing sensitivity analysis results to derive necessary optimality conditions of SMPEC (1.1)–(1.2) in terms of the limiting subdifferentials of the value function of the second stage problem (1.2). To this end, we need to tackle a number of technical challenges and complications resulting from differentiation of nonsmooth random functions including the exchange rule for the limiting subdifferential operator and Aumann's integral of a random set-valued mapping when they are both applied to a nonsmooth Lipschitz continuous random function, and the measurable selection from random set-valued mappings. We summarize our main contributions as follows:

- We derive a theorem (Theorem 2.9) which allows us to exchange the limiting subdifferential operator with the mathematical expectation operator when they are both applied to a random Lipschitz continuous function. The result generalizes a similar result established by Mordukhovich (see [28, Lemma

6.18]) to allow the measure to be atomic, and it is therefore of independent interest in variational analysis.

- We derive the first order necessary optimality conditions (Theorem 3.6) for the first stage problem (1.1) in terms of the limiting subdifferential of the value function of the second stage problem (1.2). As far as we are concerned, no such conditions (not even in terms of the Clarke subdifferentials) are available in the literature for a two-stage SMPEC whose second stage problem has *multiple* local and/or global optimal solutions. Moreover, we provide a detailed discussion on the related constraint qualifications.
- Using Filippov's measurable selection theorem, we present the optimality conditions (Theorem 3.11) in terms of the gradient of the underlying function of the second stage problem (with respect to the first stage decision vector) and a measurable selection of M-multipliers of the second stage problem. As far as we are concerned, these type of optimality conditions are first proposed for SMPECs where the second stage problem has multiple feasible solutions.
- When the SMPEC reduces to an SMPCC, we show that the established optimality conditions lead to various optimality conditions characterizing the well-known M-stationary points (Theorem 4.6) and S-stationary points (Theorem 4.7). These type of optimality conditions are sharper than the existing result of Xu and Meng [52] even when the second stage problem has a unique feasible solution.
- When the variational inequality constraint reduces to a system of equalities and inequalities, we derive optimality conditions (Theorem 5.2) which recover (when the underlying probability measure is nonatomic) and sharpen (when the underlying probability measure is atomic) their counterparts in [18, 32, 35] for the classical two-stage stochastic program. Moreover, our necessary optimality conditions are given under a very weak calmness condition which has not been used for the classical two-stage stochastic program in the literature.

The rest of this paper are organized as follows. In section 2, we present some preliminary definitions and results in variational analysis, set-valued analysis, and sensitivity analysis of value functions. In section 3, we present the main first order optimality conditions for the SMPEC (1.1)–(1.2) under various constraint qualifications. In section 4, we consider optimality conditions for SMPCCs. In section 5, we consider the special case when the equilibrium constraint is dropped; that is, we review optimality conditions derived in section 3 for the classical two-stage stochastic program with equality and inequality constraints. Finally, in section 6 we make some comments on how our optimality conditions can be possibly used for the convergence analysis when the well-known Monte Carlo sampling method or the stochastic approximation method is applied to our two-stage SMPEC.

2. Preliminary definitions and results.

2.1. Notation. Throughout this paper, we use the following notation. $\langle a, b \rangle$ denotes the scalar product of vectors a and b. $\| \cdot \|$ denotes the Euclidean norm of a vector and a compact set of vectors. If $\mathcal{M}$ is a compact set of vectors, then $\|\mathcal{M}\| := \max_{M \in \mathcal{M}} \|M\|$. $d(x, \mathcal{D}) := \inf_{x' \in \mathcal{D}} \|x - x'\|$ denotes the distance from point x to set $\mathcal{D}$.

For an m-by-n matrix A and index sets $I \subset \{1, 2, \ldots, m\}, J \subset \{1, 2, \ldots, n\}$, A_I and $A_{I,J}$ denote the submatrix of A with rows specified by I and the submatrix of A with rows and columns specified by I and J, respectively. For a vector $d \in \mathbb{R}^n$, d_i is the ith component of d and d_I is the subvector composed of the components $d_i, i \in I$.

We use $\langle a, b \rangle$ to denote the scalar product of vectors a and b, and $0 \le a \perp b \ge 0$ to denote the complementary relationship between a and b, i.e., $a_i, b_i \ge 0$ and $a_i b_i = 0$ for every pair of components. We use a^T to denote the transpose of a vector a.

For a set-valued mapping $\Phi : \mathbb{R}^m \to 2^{\mathbb{R}^q}$ (assigning to each $z \in \mathbb{R}^m$ a set $\Phi(z) \subset R^q$ which may be empty), we denote by $gph\Phi$ the graph of Φ, i.e., $gph\Phi := \{(z, v) :\in \mathbb{R}^m \times \mathbb{R}^q : v \in \Phi(z)\}$. int $\mathcal{C}$, cl $\mathcal{C}$, and $co\ \mathcal{C}$ denote the interior, the closure, and the convex hull of a set $\mathcal{C}$, respectively. We denote by $B(x, \delta)$ the open ball with radius δ and center x, that is, $B(x, \delta) := \{x' : \|x' - x\| < \delta\}$. When δ is dropped, $B(x)$ represents a neighborhood of point x.

2.2. Variational analysis. We present some background materials on variational analysis which will be used throughout the paper. Detailed discussions on these subjects can be found in [6, 7, 27, 28, 42].

Let $\Phi : \mathbb{R}^m \to 2^{\mathbb{R}^m}$ be a set-valued mapping. We denote by $\limsup_{x \to \bar{x}} \Phi(x)$ the *Painlevé–Kuratowski upper limit*,[1] i.e.,

$$\limsup_{x \to \bar{x}} \Phi(x) := \{v \in R^m : \exists \text{ sequences } x_k \to \bar{x}, v_k \to v$$
$$\text{with } v_k \in \Phi(x_k) \ \ \forall k = 1, 2, \dots\}.$$

DEFINITION 2.1 (normal cones). *Let $\mathcal{C}$ be a nonempty subset of $\mathbb{R}^m$. Given $z \in$ cl $\mathcal{C}$, the convex cone*

$$\mathcal{N}_{\mathcal{C}}^{\pi}(z) := \{\zeta \in R^m : \exists \sigma > 0 \text{ such that } \langle \zeta, z' - z \rangle \le \sigma \|z' - z\|^2 \ \forall z' \in \mathcal{C}\}$$

is called the proximal normal cone *to set $\mathcal{C}$ at point z, and the closed cone*

$$\mathcal{N}_{\mathcal{C}}(z) := \limsup_{z' \to z, z' \in \mathcal{C}} \mathcal{N}_{\mathcal{C}}^{\pi}(z')$$

is called the limiting normal cone *(also known as the Mordukhovich normal cone or basic normal cone) to $\mathcal{C}$ at point z.*

The above construction of the limiting normal cone using the proximal normal cone was given by Mordukhovich in [25]. In many publications, however, the limiting normal cone is defined by the Fréchet (also called regular) normal cones; see [27, Definition 1.1 (ii)]. The two definitions coincide in the finite dimensional space (see [27, Theorem 1.6] for a proof and [27, page 141] or [42, page 345] for a discussion). The limiting normal cone is in general smaller than the Clarke normal cone which is equal to the convex hull $co\mathcal{N}_{\mathcal{C}}(z)$, and in the case when $\mathcal{C}$ is convex, the proximal normal cone, the limiting normal cone, and the Clarke normal cone coincide with the normal cone in the sense of the convex analysis, i.e.,

$$\mathcal{N}_{\mathcal{C}}(z) := \{\zeta \in \mathbb{R}^m : \langle \zeta, z' - z \rangle \le 0 \ \ \forall z' \in \mathcal{C}\}.$$

For set-valued mappings, the definition for a limiting normal cone leads to the definition of the Mordukhovich coderivative which was first introduced in [26].

DEFINITION 2.2 (coderivatives). *Let $\Phi : \mathbb{R}^m \to 2^{\mathbb{R}^q}$ be an arbitrary set-valued mapping and $(\bar{z}, \bar{v}) \in$ cl $gph\Phi$. The coderivative of Φ at point $(\bar{z}, \bar{v})$ is defined as*

$$D^*\Phi(\bar{z}, \bar{v})(\eta) := \{\zeta \in \mathbb{R}^m : (\zeta, -\eta) \in \mathcal{N}_{gph\Phi}(\bar{z}, \bar{v})\}.$$

By convention, for $(\bar{z}, \bar{v}) \notin$ cl $gph\Phi$, $D^*\Phi(\bar{z}, \bar{v})(\eta) = \emptyset$.

---

[1] In some references, it is also called outer limit; see [42].

A particularly interesting case relevant to our discussions later on is when $\Phi(z) = \mathcal{N}_{\mathcal{C}}(z)$ and $\mathcal{C}$ is a closed convex set. By the definition of coderivatives,

$$\zeta \in D^*\mathcal{N}_{\mathcal{C}}(\bar{z}, \bar{v})(\eta) \iff (\zeta, -\eta) \in \mathcal{N}_{gph\mathcal{N}_{\mathcal{C}}}(\bar{z}, \bar{v}).$$

Hence the calculation of the coderivative $D^*\Phi(\bar{z}, \bar{v})(\eta)$ depends on the calculation of the limiting normal cone to the normal cone $\mathcal{N}_{gph\mathcal{N}_{\mathcal{C}}}(\bar{z}, \bar{v})$. In the case when $\mathcal{C} = \mathbb{R}^m_+$, the following explicit formula can be used. The proof of the formula follows easily from the formula for the proximal normal cone in [53, Proposition 2.7] and the definition of the limiting normal cones.

PROPOSITION 2.3. *For any $(\bar{z}, -\bar{v}) \in gphN_{\mathbb{R}^m_+}$, let*

$$L := L(\bar{z}, \bar{v}) := \{i \in \{1, 2, \ldots, m\} : \bar{z}_i > 0, \bar{v}_i = 0\},$$
$$I_+ := I_+(\bar{z}, \bar{v}) := \{i \in \{1, 2, \ldots, m\} : \bar{z}_i = 0, \bar{v}_i > 0\},$$
$$I_0 := I_0(\bar{z}, \bar{v}) := \{i \in \{1, 2, \ldots, m\} : \bar{z}_i = 0, \bar{v}_i = 0\}.$$

Then

$$\mathcal{N}_{gph\mathcal{N}_{\mathbb{R}^m_+}}(\bar{z}, -\bar{v}) = \{(\alpha, -\beta) \in \mathbb{R}^{2m} : \alpha_L = 0, \beta_{I_+} = 0,$$
$$\forall i \in I_0, \ \text{either } \alpha_i < 0, \beta_i < 0 \ \text{or } \alpha_i\beta_i = 0\}.$$

In the case when $\mathcal{C}$ is a polyhedral convex set, a formula for the normal cone to the graph of the standard normal cone is given in the proof of [8, Theorem 2] and also stated in [34, Proposition 4.4]. For recent results on calculating the normal cone to the graph of a standard normal cone (coderivative of the standard normal cone mapping), readers are referred to [14, 15] and [16, section 3]

DEFINITION 2.4 (subdifferentials). *Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ be a lower semicontinuous function and finite at $x \in \mathbb{R}^n$. The proximal subdifferential ([42, Definition 8.45]) of f at x is defined as*

$$\partial^\pi f(x) := \{\zeta \in \mathbb{R}^n : \exists \sigma > 0, \delta > 0 \ \text{such that } f(y) \geq f(x) + \langle \zeta, y - x \rangle - \sigma\|y - x\|^2$$
$$\forall y \in B(x, \delta)\},$$

and the limiting (Mordukhovich or basic [27]) subdifferential of f at x is defined as

$$\partial f(x) := \limsup_{x' \xrightarrow{f} x} \partial^\pi f(x'),$$

where $x' \xrightarrow{f} x$ signifies that x' and $f(x')$ converge to x and $f(x)$, respectively.

When f is Lipschitz continuous near x, the Clarke subdifferential [6] of f at x is equal to $co\partial f(x)$.

Note that in his earlier work [25] Mordukhovich defined the limiting subgradient via the limiting normal cones which was constructed by the proximal normal cones. In his later work, Mordukhovich defined the limiting subgradient via Fréchet limiting normal cones and Fréchet subgradients (also known as regular subgradients); see [27, Theorem 1.89]. The equivalence of the two definitions is well-known; see the commentary by Rockafellar and Wets [42, page 345]. The limiting subdifferential is in general smaller than the Clarke subdifferential, and in the case when f is convex and locally Lipschitz, the proximal subdifferential, the limiting subdifferential, and the Clarke subdifferential coincide with the subdifferential in the sense of convex analysis

[40]. In the case when f is continuously differentiable, these subdifferentials reduce to normal gradient $\nabla f(x)$, i.e., $\partial f(x) = \{\nabla f(x)\}$.

In what follows, we state a well-known calculus rule in Proposition 2.5 for the limiting subdifferentials of nonconvex functions. A proof of the proposition and its extension to non-Lipschitz functions can be found in [27, Theorems 2.33 and 3.36]. In subsection 2.4, we will extend Proposition 2.5 to the case when the summation is replaced by Aumann's integral in our main result of this section, Theorem 2.9.

PROPOSITION 2.5 (positive scalar multiplication and sum rule). *Let $f_i \colon \mathbb{R}^n \to \overline{\mathbb{R}}$, $i = 1, 2, \ldots, N$, be lower semicontinuous functions. Suppose that all but one of these functions are Lipschitz near $\bar{x}$ and $\lambda_i \geq 0$ are constants. Then*

$$\partial \sum_{i=1}^{N} \lambda_i f_i(\bar{x}) \subset \sum_{i=1}^{N} \lambda_i \partial f_i(\bar{x}).$$

2.3. Set-valued mappings and measurability. Let $\mathcal{X}$ be a closed subset of $\mathbb{R}^n$. A set-valued mapping $\Phi : \mathcal{X} \to 2^{\mathbb{R}^m}$ is said to be *closed* at $\bar{x}$ if for $x_k \in \mathcal{X}$, $x_k \to \bar{x}$, $y_k \in \Phi(x_k)$ and $y_k \to \bar{y}$, we have $\bar{y} \in \Phi(\bar{x})$. Φ is said to be *uniformly compact* near $\bar{x} \in \mathcal{X}$ if there is a neighborhood $B(\bar{x})$ of $\bar{x}$ such that the closure of $\bigcup_{x \in B(\bar{x})} \Phi(x)$ is compact. Φ is said to be *upper semicontinuous* at $\bar{x} \in \mathcal{X}$ if for every $\epsilon > 0$, there exists a $\delta > 0$ such that

$$\Phi(\bar{x} + \delta\mathcal{B}) \subset \Phi(\bar{x}) + \epsilon\mathcal{B},$$

where $\mathcal{B}$ denotes the closed unit ball in $\mathbb{R}^m$. The following result was known; see [10, 19].

PROPOSITION 2.6. *Let $\Phi : \mathcal{X} \to 2^{\mathbb{R}^m}$ be uniformly compact near $\bar{x}$. Then Φ is upper semicontinuous at $\bar{x}$ if and only if Φ is closed.*

Let us now consider a stochastic set-valued mapping. Let $(\Omega, \mathcal{F}, P)$ be a probability space. For fixed x, let $\mathcal{A}(x, \omega) : \Omega \to 2^{\mathbb{R}^n}$ be a set-valued mapping whose value is a closed subset of $\mathbb{R}^n$. Let $\mathfrak{B}(\mathbb{R}^n)$ or simply $\mathfrak{B}$ denote the space of closed bounded subsets of $\mathbb{R}^n$ endowed with topology $\tau_{\mathbb{H}}$ generated by the Hausdorff distance $\mathbb{H}$. We consider the Borel σ-field $\mathcal{G}(\mathfrak{B}, \tau_{\mathbb{H}})$ generated by the $\tau_{\mathbb{H}}$-open subsets of $\mathfrak{B}$. A set-valued mapping $\mathcal{A}(x, \omega) : \Omega \to 2^{\mathbb{R}^n}$ is said to be $\mathcal{F}$-measurable if, for every member $\mathcal{W}$ of $\mathcal{G}(\mathfrak{B}, \tau_{\mathbb{H}})$, one has $\mathcal{A}^{-1}(\mathcal{W}) \in \mathcal{F}$.

By a *measurable selection* of $\mathcal{A}(x, \omega)$, we refer to a vector $A(x, \omega) \in \mathcal{A}(x, \omega)$, which is measurable. Note that such measurable selections exist if $\mathcal{A}(x, \omega)$ is measurable; see [1] and references therein.

For a general set-valued mapping which is not necessarily measurable, the *expectation of $\mathcal{A}(x, \omega)$*, denoted by $\mathbb{E}[\mathcal{A}(x, \omega)]$, is defined as the collection of $\mathbb{E}[A(x, \omega)]$ where $A(x, \omega)$ is an integrable selection, and the integrability is in the sense of Aumann [4]. $\mathbb{E}[\mathcal{A}(x, \omega)]$ is regarded as *well-defined* if it is nonempty. A sufficient condition of the well definedness of the expectation is that $\mathcal{A}(x, \omega)$ is measurable and $\mathbb{E}[\|\mathcal{A}(x, \omega)\|] := \mathbb{E}[\mathbb{H}(0, \mathcal{A}(x, \omega))] < \infty$, in which case $\mathbb{E}[\mathcal{A}(x, \omega)] \in \mathfrak{B}$. See [4, Theorem 2]. In such a case, $\mathcal{A}$ is called *integrably bounded* in [4, 17].

DEFINITION 2.7 (simple set-valued mapping). *Let $\mathcal{A}(x, \omega) : \Omega \to 2^{\mathbb{R}^n}$ be a measurable set-valued mapping. $\mathcal{A}$ is said to be a* simple *set-valued mapping if it takes a finite number of $S_i \in \mathfrak{B}$ and there is an $\mathcal{F}$-measurable partition $\{\Omega_1, \ldots, \Omega_k\}$ of Ω such that for any $\omega \in \Omega_i$, $i = 1, \ldots, k$,*

$$\mathcal{A}(\omega) = \sum_{i=1}^{k} \mathbb{1}_{\Omega_i}(\omega) S_i,$$

where

$$\mathbb{1}_{\Omega_i}(\omega) := \left\{ \begin{array}{ll} 1 & \text{if } \omega \in \Omega_i, \\ 0 & \text{if } \omega \notin \Omega_i. \end{array} \right.$$

The expectation of the simple set-valued mapping $\mathcal{A}$ is

$$\mathbb{E}[\mathcal{A}(\omega)] = \sum_{i=1}^{k} P(\Omega_i) S_i.$$

The following result is well-known; see, e.g., [3, section 8.1, page 307] and [17, Lemmas 3.1–3.2].

LEMMA 2.8. *If $\mathcal{A}(x, \omega) : \Omega \to 2^{\mathbb{R}^n}$ is a closed measurable set-valued mapping, then $\mathcal{A}$ is a pointwise limit of a sequence of measurable simple set-valued mappings on Ω.*

In the case when $\mathcal{A}$ is single-valued, the above lemma indicates that a random function is a pointwise limit of a sequence of random simple functions on Ω.

2.4. The exchange rule for Aumann's integral and limiting subdifferential operator. Using Lemma 2.8[2] and Proposition 2.5, we are able to extend Proposition 2.5 to the case when the summation is replaced by an integration (mathematical expectation); that is, the integration and the limiting subdifferential operation can be exchanged when they are both applied to a random function. The result is an analogue of the exchange of an integral and the Clarke subdifferential operation in [6, Theorem 2.7.2] and will be used to establish optimality conditions of (1.1) in terms of the limiting subdifferential of the value functions of the second stage problem (1.2). Note that an exchange of Aumann's integral and the limiting subdifferential operator is established by Mordukhovich in [28, Lemma 6.18]. The proof uses the well-known Aumann's identity, that is, that the expected value of the limiting subgradient coincides with that of Clarke's subdifferential when the probability space is nonatomic. In Theorem 2.9 below, we derive an analogue of [28, Lemma 6.18] without the nonatomic condition. The two results coincide when the probability space is nonatomic.

THEOREM 2.9. *Let $\phi(x, \xi) : \mathbb{R}^n \times \Xi \to \mathbb{R}$ be a continuous function where $\xi : (\Omega, \mathcal{F}, P) \to \Xi$ is a random vector with support set $\Xi \subset \mathbb{R}^m$. Suppose* (a) *ϕ is Lipschitz continuous with respect to x in a neighborhood of $\bar{x}$ for every ξ and its Lipschitz modulus is bounded by a nonnegative integrable function $\kappa(\xi(\omega))$;* (b) *$\mathbb{E}[\phi(x, \xi(\omega))] < \infty$. Let $\psi(x) := \mathbb{E}[\phi(x, \xi(\omega))]$. Then the following conditions hold.*

(i) *$\psi(x)$ is well-defined and Lipschitz continuous near $\bar{x}$ with modulus $\mathbb{E}[\kappa(\xi(\omega))]$.*

(ii) *$\mathbb{E}[\partial_x \phi(\bar{x}, \xi(\omega))]$ is well-defined, and the following inclusion holds:*

$$(2.1) \qquad \partial\psi(\bar{x}) \subset \mathbb{E}[\partial_x \phi(\bar{x}, \xi(\omega))].$$

(iii) *The inclusion (2.1) coincides with (6.39) in [28, Lemma 6.18] when the probability space of ξ is nonatomic. In the case when $\phi(x, \xi)$ is Clarke regular [6] at $\bar{x}$, ψ is also Clarke regular and the equality holds in (2.1).*

Proof. Part (i). The well definedness of $\psi(x)$ and Lipschitz continuity of $\psi(x)$ for x close to $\bar{x}$ is well-known under conditions (a) and (b). See, for instance, [43, Proposition 2].

Part (ii). We first show the well definedness of $\mathbb{E}[\partial_x \phi(x, \xi(\omega))]$; that is, $\mathbb{E}[\partial_x \phi(x, \xi(\omega))]$ is a nonempty compact set. Following a discussion in [1, page 880] by

---

[2]In the proof, we will use an earlier counterpart of this result [29, Lemma V-2.4].

Artstein and Vitale, it suffices to show that $\partial_x \phi(x, \xi(\omega))$ is measurable and integrably bounded. The latter is implied by our condition (a). We prove the former. Let $d \in \mathbb{R}^n$ and $\xi \in \mathbb{R}^m$ be fixed. The subderivative of $\phi(x, \xi)$ with respect to x at a point x in direction d is defined as

$$\phi_x^\diamond(x, \xi; d) := \liminf_{\substack{d' \to d \\ t \to 0}} [\phi(x + td', \xi) - \phi(x, \xi)]/t.$$

By [3, Lemma 8.2.12], $\phi_x^\diamond(x, \xi; d)$ is measurable. Let

$$\hat{\partial}_x \phi(x, \xi) := \{ h : h^T d \leq \phi_x^\diamond(x, \xi; d) \quad \forall d \},$$

where $\phi_x^\diamond(x, \xi; d)$ is the support function of the set-valued mapping $\hat{\partial}_x \phi(x, \xi)$ (see, e.g., [42, Exercise 8.4]). By [3, Theorem 8.2.14], $\hat{\partial}_x \phi(x, \xi(\cdot))$ is measurable. Since $\partial_x \phi(x, \xi(\cdot))$ is the upper limit of $\hat{\partial}_x \phi(x, \xi(\cdot))$, the measurability of the former follows from that of the latter by [3, Theorem 8.2.5].

Next, we prove (2.1). By [29, Lemma V-2.4] and its proof, there exists a sequence $\{\xi^k\}_{k=1}^\infty$ which is a dense subset of Ξ such that for each k there exist $\mathcal{F}$-measurable partitions of Ω denoted by $\{\Omega_1, \ldots, \Omega_k\}$ satisfying

$$\lim_{k \to \infty} \sum_{i=1}^k \mathbb{1}_{\Omega_i}(\omega) \xi^i = \xi(\omega)$$

for every $\omega \in \Omega$. Let

$$\phi^k(x, \xi(\omega)) := \sum_{i=1}^k \mathbb{1}_{\Omega_i}(\omega) \phi(x, \xi^i)$$

and x be fixed. The continuity of ϕ in ξ implies that the sequence $\{\phi(x, \xi^k)\}_{k=1}^\infty$ is a dense subset of $\phi(x, \Xi)$. Therefore

(2.2)
$$\lim_{k \to \infty} \phi^k(x, \xi(\omega)) = \phi(x, \xi(\omega)).$$

Let $\omega \in \Omega$ be fixed and $\xi := \xi(\omega)$. By the definition of the limiting subdifferential, it is obvious that $\partial_x \phi(x, \cdot)$ is a closed set-valued mapping. By virtue of the local Lipschitz continuity of ϕ as assumed in the assumption (a) (see [27, Corollary 1.81]), it is also uniformly compact at any fixed point $\xi \in \Xi$. Hence by Proposition 2.6, $\partial_x \phi(x, \cdot)$ is upper semicontinuous at ξ. Therefore for every fixed $\omega \in \Omega$,

(2.3)
$$\lim_{k \to \infty} \sum_{i=1}^k \mathbb{1}_{\Omega_i}(\omega) \partial_x \phi(x, \xi^i) \subset \partial_x \phi(x, \xi(\omega)).$$

Since $\phi^k(x, \xi(\omega))$ is Lipschitz with respect to x with a uniform Lipschitz modulus, the limit (2.2) holds uniformly with respect to x on a compact set. Moreover,

$$\psi(x) := \mathbb{E}[\phi(x, \xi(\omega))] = \mathbb{E}\left[\lim_{k \to \infty} \phi^k(x, \xi(\omega)) \right]$$

$$= \lim_{k \to \infty} \mathbb{E}\left[\phi^k(x, \xi(\omega)) \right] = \lim_{k \to \infty} \sum_{i=1}^k \phi(x, \xi^i) P(\Omega_i).$$

The third equality is due to Lebesgue's dominated convergence theorem because $\phi^k(x, \xi(\omega))$ is bounded on any compact set of $\mathbb{R}^n$ and the above equalities hold uniformly with respect to x on any compact set of $\mathbb{R}^n$. Let

$$\psi^k(x) := \sum_{i=1}^{k} \phi(x, \xi^i) P(\Omega_i)$$

and $\zeta \in \partial^\pi \psi(x)$. Then by definition, there exist constants $\sigma > 0$, $\delta > 0$ such that

$$\lim_{k \to \infty} (\psi^k(y) - \psi^k(x)) > \langle \zeta, y - x \rangle - \sigma \|y - x\|^2 \quad \forall y \in B(x, \delta) \text{ with } y \neq x.$$

We assume without loss of generality that the strict inequality above holds for any $y \neq x$. This can be achieved by choosing a sufficiently large σ. Therefore for k sufficiently large,

$$\psi^k(y) - \psi^k(x) > \langle \zeta, y - x \rangle - \sigma \|y - x\|^2 \quad \forall y \in B(x, \delta) \text{ with } y \neq x.$$

Consequently, for all large enough k, $y = x$ is the unique local minimizer of the problem

$$\min_y \psi^k(y) + \sigma \|y - x\|^2 - \langle \zeta, y - x \rangle$$

for y restricted in a compact neighborhood of x. The optimality condition in terms of the limiting subdifferentials [42, Theorem 10.1] and the sum rule for limiting subdifferentials (Proposition 2.5) indicate that

(2.4) $$0 \in \partial \psi^k(x) - \zeta.$$

By Proposition 2.5,

(2.5) $$\partial \psi^k(x) \subset \sum_{i=1}^{k} \partial_x \phi(x, \xi^i) P(\Omega_i).$$

Since ζ is any element from set $\partial^\pi \psi(x)$, by (2.4) and (2.5),

$$\partial^\pi \psi(x) \subset \sum_{i=1}^{k} \partial_x \phi(x, \xi^i) P(\Omega_i). = \mathbb{E}\left[\sum_{i=1}^{k} \mathbb{1}_{\Omega_i}(\omega) \partial_x \phi(x, \xi^i) \right].$$

Taking the limit on both sides of the above equation and by virtue of [4, Proposition 4.1] and (2.3), we obtain that

(2.6) $$\partial^\pi \psi(x) \subset \mathbb{E}[\partial_x \phi(x, \xi(\omega))].$$

By the definition of the limiting subdifferential and [4, Proposition 4.1],

$$\partial \psi(\bar{x}) = \limsup_{x \to \bar{x}} \partial^\pi \psi(x) \subset \limsup_{x \to \bar{x}} \mathbb{E}[\partial_x \phi(x, \xi(\omega))]$$

$$\subset \mathbb{E}\left[\limsup_{x \to \bar{x}} \partial_x \phi(x, \xi(\omega)) \right] \subset \mathbb{E}\left[\partial_x \phi(\bar{x}, \xi(\omega)) \right].$$

This shows (2.1).

Part (iii). When the probability space of ξ is nonatomic, the inclusion (2.1) can be established by virtue of Aumann's identity (see [17, Theorem 5.4 (d)]); see (6.39) in [28, Lemma 6.18].

The Lipschitz continuity of the function ψ and the last assertion of the theorem follow from [6, Theorem 2.7.2] since when a function is Clarke regular, the limiting subdifferential coincides with the Clarke subdifferential. This completes the proof. $\quad\square$

2.5. Sensitivity analysis on the value function of $\mathcal{P}(x, \xi)$. We now move on to analyze the sensitivity of the value function of the second stage problem $\mathcal{P}(x, \xi)$. Recall that $v(x, \xi)$ denotes the value function of the second stage problem. We use $\Gamma(x, \xi)$ to denote the set of global optimal solutions to the second stage problem.

2.5.1. No nonzero abnormal multipliers constraint qualification (NNAMCQ) and M-multipliers. For deterministic MPECs, it is well-known that the usual NLP constraint qualifications such as the Mangasarian–Fromovitz constraint qualification (MFCQ) do not hold (see [59, Proposition 1.1]), and hence Lagrange multipliers may not exist. This leads to the introduction of the following weaker concept of multipliers (for the case of no inequality constraint, see [57], and for the case including inequality constraints, see [54]). Since the set of M-multipliers (which were called CD-multipliers in [21]) is nonempty under the MPEC variant of MFCQ, one can use the set of M-multipliers to carry out the sensitivity analysis of the value functions for MPECs.

DEFINITION 2.10 (M-multipliers). *Let $(x, \xi) \in X \times \Xi$ be fixed. Let (y, z) be a feasible solution of the second stage problem $\mathcal{P}(x, \xi)$. We say that (y, z) is an M-stationary point and $(\gamma, \eta) \in \mathbb{R}_+^p \times \mathbb{R}^m$ is an M-multiplier of $\mathcal{P}(x, \xi)$ at (y, z) if*

$$0 \in \nabla_y f_2(x, y, z, \xi) + \nabla_y \psi(x, y, z, \xi)^T \gamma + \nabla_y F(x, y, z, \xi)^T \eta,$$
$$0 \in \nabla_z f_2(x, y, z, \xi) + \nabla_z \psi(x, y, z, \xi)^T \gamma$$
$$+ \nabla_z F(x, y, z, \xi)^T \eta + D^* \mathcal{N}_{\mathcal{C}}(z, -F(x, y, z, \xi))(\eta),$$
$$0 = \psi(x, y, z, \xi)^T \gamma.$$

Here and later on, ∇F denotes the classical Jacobian of a vector-valued function F. We use $M(x, y, z, \xi)$ to denote the set of M-multipliers at stationary point (y, z).

From [54, 57], the set $M(x, y, z, \xi)$ at any local optimal solution (y, z) of the second stage problem $\mathcal{P}(x, \xi)$ is nonempty under the following constraint qualification.

DEFINITION 2.11 (NNAMCQ). *We say that NNAMCQ holds at a feasible point (y, z) of problem $\mathcal{P}(x, \xi)$ if*

$$\begin{cases} 0 \in \nabla_{y,z} \psi(x, y, z, \xi)^T \gamma \\ \quad + \nabla_{y,z} F(x, y, z, \xi)^T \eta + \{0\} \times D^* \mathcal{N}_{\mathcal{C}}(z, -F(x, y, z, \xi))(\eta), \\ 0 \le -\psi(x, y, z, \xi) \perp \gamma \ge 0 \end{cases} \implies \gamma = 0, \quad \eta = 0.$$

Here and later on we write the first order conditions in a closed form to save space. In the case when there is no equilibrium constraint, NNAMCQ reduces to the positive linear independence of the gradients of the active inequality constraints

$$\nabla_{y,z} \psi_i(x, y, z, \xi), \quad i \in \mathcal{I}(\bar{x}, \xi),$$

where $\mathcal{I}(x, \xi) := \{i : \psi_i(x, y, z, \xi) = 0\}$. By the Fakas lemma, the positive linear independence of the gradients of the active inequality constraints is equivalent to the MFCQ; i.e., there exists $(d, h) \in \mathbb{R}^l \times \mathbb{R}^m$ such that

$$\langle \nabla_{y,z} \psi_i(x, y, z, \xi), (d, h) \rangle > 0 \quad \forall i \in \mathcal{I}(x, \xi).$$

Hence NNAMCQ can be viewed as a dual form of the MFCQ. In NLP, it is well-known that the MFCQ is equivalent to the compactness of the Lagrange multiplier sets (see, e.g., [9]). This is also true for M-multipliers under NNAMCQ.

PROPOSITION 2.12. *Let $(x, \xi) \in X \times \Xi$ be fixed and*

(2.7)
$$\mathcal{M}(x, \xi) := \bigcup_{(y,z) \in \Gamma(x,\xi)} M(x, y, z, \xi).$$

Assume (a) $\Gamma(x, \xi)$ *is compact and* (b) *NNAMCQ holds at any global optimal solution point* $(y, z) \in \Gamma(x, \xi)$. *Then* $\mathcal{M}(x, \xi)$ *is nonempty and compact.*

Proof. Assume for the sake of a contradiction that $\mathcal{M}(x, \xi)$ is unbounded. Then there exists a sequence $\{(y_k, z_k)\} \subset \Gamma(x, \xi)$ and an unbounded sequence $\{(\gamma_k, \eta_k)\} \in M(x, y_k, z_k, \xi)$ with $(y_k, z_k) \in \Gamma(x, \xi)$ such that $\|\gamma_k\| + \|\eta_k\| \to \infty$ as $k \to \infty$. By definition

(2.8)
$$0 \in \nabla_{y,z} f_2(x, y_k, z_k, \xi) + \nabla_{y,z} \psi(x, y_k, z_k, \xi)^T \gamma_k$$
$$+ \nabla_{y,z} F(x, y_k, z_k, \xi)^T \eta_k + \{0\} \times D^* \mathcal{N}_\mathcal{C}(z_k, -F(x, y_k, z_k, \xi))(\eta_k).$$

Dividing the above equation on both sides by $\|(\gamma_k, \eta_k)\|$ and driving k to infinity, we have from the compactness of $\Gamma(x, \xi)$ and boundedness of $(\gamma_k, \eta_k)/\|(\gamma_k, \eta_k)\|$ that there exist a subsequence $(y_{k_j}, z_{k_j}) \to (y, z) \in \Gamma(x, \xi)$ and $(\gamma_{k_j}, \eta_{k_j})/\|(\gamma_{k_j}, \eta_{k_j})\| \to (\gamma, \eta)$ with $\|(\gamma, \eta)\| = 1$ and that

$$0 \in \nabla_{y,z} \psi(x, y, z, \xi)^T \gamma$$
$$+ \nabla_{y,z} F(x, y, z, \xi)^T \eta + \{0\} \times D^* \mathcal{N}_\mathcal{C}(z, -F(x, y, z, \xi))(\eta),$$
$$0 = \psi(x, y, z, \xi)^T \gamma, \quad \gamma \geq 0.$$

This contradicts the NNAMCQ. Similarly we can prove that the set $\mathcal{M}(x, \xi)$ is closed for each fixed (x, ξ), and hence the proof of the proposition is complete. □

The NNAMCQ plays an essential role in the sensitivity analysis of the value function of the second stage problem. It is therefore natural to consider sufficient conditions for it. The proposition below lists a few sufficient conditions for NNAMCQ, and they follow straightforwardly from [54, Theorem 4.7] and [57, Theorem 3.2].

PROPOSITION 2.13. *Let $x \in X$ and $\xi \in \Xi$. Consider the second stage problem* (1.2) *without inequality constraint $\psi \leq 0$. The following conditions suffice for NNAMCQ.*

  (i) *The strongly regular constraint qualification (SRCQ) holds at (y, z); i.e., the generalized equation*

$$0 \in F(x, y, z, \xi) + \mathcal{N}_\mathcal{C}(z)$$

  is strongly regular at (y, z) in the sense of Robinson [38].

  (ii) *$-F$ is locally strongly monotone in z uniformly with respect to y; that is, there exist a positive constant μ independent of y and neighborhoods U_1 of y, U_2 of z such that*

$$\langle -F(x, y', z', \xi) + F(x, y', z, \xi), z' - z \rangle \geq \mu \|z' - z\|^2 \quad \forall z' \in U_2 \cap \mathcal{C}, z \in \mathcal{C}, y' \in U_1.$$

  (iii) *The rank of the matrix $\nabla_y F(x, y, z, \xi)$ is m.*

2.5.2. Sensitivity analysis of the value function. To ensure the existence of a local optimal solution to the second stage problem $\mathcal{P}(x, \xi)$, we need the following inf-compact conditions.

Assumption 2.14 (inf-compactness).   Let $x \in X$ and $\xi \in \Xi$ be fixed. There exists a constant $\delta > 0$ such that the set

$$\{(y,z) : \psi(x,y,z,\xi) \leq q, r \in F(x,y,z,\xi) + \mathcal{N}_{\mathcal{C}}(z), f_2(x,y,z,\xi) \leq \alpha, (r,q) \in B(0,\delta)\}$$

is bounded for every constant α.

PROPOSITION 2.15.   *Consider the second stage problem* (1.2). *Let* $\bar{x} \in Q$ *and* $\bar{\xi} \in \Xi$ *be fixed. Suppose that* (a) *Assumption* 2.14 *holds at* $\bar{x}$ *and* $\bar{\xi}$ *and* (b) *for every* $(y,z) \in \Gamma(\bar{x}, \bar{\xi})$, *either NNAMCQ or the second stage problem* (1.2) *has no inequality constraint and one of the constraint qualifications given in Proposition* 2.13 *holds. Then*

(i) $(x,\xi) \rightarrow v(x,\xi)$ *is Lipschitz near* $\bar{x}$ *and* $\bar{\xi}$;
(ii) $\partial_x v(\bar{x}, \bar{\xi}) \subset \Psi(\bar{x}, \bar{\xi})$, *where*

$$(2.9) \qquad \Psi(x,\xi) := \bigcup_{(y,z) \in \Gamma(x,\xi)} \bigcup_{(\gamma,\eta) \in M(x,y,z,\xi)} \{\nabla_x f_2(x,y,z,\xi)$$
$$+ \nabla_x \psi(x,y,z,\xi)^{\top} \gamma + \nabla_x F(x,y,z,\xi)^{\top} \eta\};$$

(iii) $\Gamma(\bar{x}, \bar{\xi})$ *is compact.*

Proof. Parts (i)–(ii) follow from [21, Corollaries 3.7 and 3.8], and part (iii) is obvious.   ☐

THEOREM 2.16.   *Let Assumption* 2.14 *hold for* $\bar{x} \in Q$ *and every* $\xi \in \Xi$, *and let* $V(x) := \mathbb{E}[v(x,\xi(\omega))]$. *Then*

(i) $v(x,\xi(\cdot)) : \Omega \rightarrow \mathbb{R}$ *is measurable;*
(ii) $\partial_x v(\bar{x}, \xi(\cdot)) : \Omega \rightarrow 2^{\mathbb{R}^n}$ *is measurable;*
(iii) $\Gamma(x,\xi(\cdot)) : \Omega \rightarrow 2^{\mathbb{R}^n}$ *is measurable;*
(iv) *if* $\mathbb{E}[v(\bar{x}, \xi(\omega))]$ *is well-defined and the Lipschitz modulus of* $v(x,\xi)$ *in* x *is bounded by an integrable function* $\kappa(\xi)$, *then* $V(x)$ *is well-defined for all* $x \in Q$ *and it is locally Lipschitz at* $\bar{x}$. *Moreover,*

$$(2.10) \qquad\qquad \partial V(\bar{x}) \subset \mathbb{E}[\partial_x v(\bar{x}, \xi(\omega))].$$

Furthermore, if $v(x,\xi)$ *is Clarke regular in* x, *then* $V(x)$ *is Clarke regular and the equality holds.*

Proof. Parts (i) and (iii). These parts follow from the marginal map theorem in the measurability theory of set-valued mappings; see [3, Theorem 8.2.11].

Part (ii). Under Assumption 2.14, it follows from Proposition 2.15 (i) that the value function $v(x,\xi)$ is Lipschitz continuous in ξ and from Proposition 2.15 (iii) that its modulus is bounded by an integrable function. Consequently, we can show the measurability of $\partial_x v(\bar{x}, \xi(\cdot))$ in the same way as in the first part of the proof of Theorem 2.9 (ii).

Part (iv). The well definedness of $V(x)$ is obvious. The Lipschitz continuity of V follows from [43, Chapter 2, Proposition 2]. Since the Lipschitz modulus of $v(x,\xi)$ is $\kappa(\xi)$, and $\partial_x v(x,\xi)$ is contained by Clarke's generalized gradient, by [6, Proposition 2.1.2], $\|\partial_x v(x,\xi)\| \leq \kappa(\xi)$. This and the measurability of $\partial_x v(x,\xi)$ ensure the well definedness of $\mathbb{E}[\partial_x v(x,\xi)]$. Finally the inclusion (2.10) and the rest of the conclusion follow from Theorem 2.9.   ☐

3. Optimality conditions. In this section, we derive the first order necessary optimality conditions of SMPEC (1.1)–(1.2). First, we derive optimality conditions in terms of the limiting subdifferential of the expected value of the value function of

the second stage problem (1.2) under the Clarke calmness condition (Theorem 3.6 (i)); second, we sharpen the optimality condition by taking a particular measurable selection from the limiting subdifferential of the value function (Theorem 3.6 (ii)), and finally we express the measurable selection in terms of the gradients and the M-multipliers of the second stage problem (Theorem 3.7) at an optimal solution point and/or a stationary point.

3.1. Clarke calmness and pseudo upper-Lipschitz continuity of set-valued mappings. We start by considering Clarke's calmness condition [6] for problem (1.1).

DEFINITION 3.1. *We say that the problem* (1.1) *is* calm *at a local optimal solution* $\bar{x}$ *if there exists* $\mu > 0$ *such that* $\bar{x}$ *is a local optimal solution to the penalized problem*

$$(3.1) \qquad \begin{aligned} \min \quad & f_1(x) + \mathbb{E}\left[v(x, \xi(\omega))\right] + \mu[\|G(x)_+\| + \|H(x)\|] \\ \text{s.t.} \quad & x \in Q. \end{aligned}$$

The above calmness condition involves both the constraint functions and the objective function; it is therefore not a constraint qualification in the classical sense. Indeed it is a sufficient condition under which Karush–Kuhn–Tucker (KKT) type necessary optimality conditions hold. The calmness condition may hold even when the weakest constraint qualification does not hold. In practice one often uses some verifiable constraint qualifications sufficient for the calmness condition.

DEFINITION 3.2 (pseudo upper-Lipschitz continuity). *A set-valued mapping* $\Phi : \mathbb{R}^n \to 2^{\mathbb{R}^q}$ *is said to be* pseudo upper-Lipschitz continuous *at* $(\bar{z}, \bar{v}) \in gph\Phi$ *if there exist a constant* $\mu > 0$ *and a neighborhood* $B(\bar{z})$ *of* $\bar{z}$, *a neighborhood* $B(\bar{v})$ *of* $\bar{v}$ *such that*

$$\Phi(z) \cap B(\bar{v}) \subseteq \Phi(\bar{z}) + \mu\|z - \bar{z}\|\mathcal{B} \quad \forall z \in B(\bar{z}).$$

The concept of pseudo upper-Lipschitz continuity of a set-valued mapping was first introduced by Ye and Ye [57] for the purpose of providing weak and applicable constraint qualifications for the M-stationary conditions. The name "pseudo upper-Lipschitz continuity" comes from the fact that it is a combination of Aubin's pseudo Lipschitz continuity [2] and Robinson's upper-Lipschitz continuity [36, 37]. In some references (see, for example, [42, 27, 12]), the pseudo upper-Lipschitz continuity is also called calmness. Here we use the former terminology to avoid confusion with Clarke's calmness. For a recent discussion on the properties and the criterion of pseudo upper-Lipschitz continuity of a set-valued mapping, see Henrion, Jourani, and Outrata [12] and Henrion and Outrata [13]. In what follows, we consider the pseudo upper-Lipschitz continuity of the perturbed feasible region of the constraint system

$$(3.2) \qquad X(p, q) := \{x : G(x) + p \le 0, H(x) + q = 0, x \in Q\}, \quad X(0,0) := X$$

at $p = 0, q = 0$ to establish the calmness of problem (1.1). The proposition below is an easy consequence of Clarke's exact penalty principle [6, Proposition 2.4.3] and pseudo upper-Lipschitz continuity of the perturbed feasible region of the true problem. See [54, Proposition 4.2] for a proof.

PROPOSITION 3.3. *If the objective function of problem* (1.1) *is Lipschitz near* $\bar{x} \in X$ *and the perturbed feasible region of the constraint system* $X(p, q)$ *defined as in* (3.2) *is pseudo upper-Lipschitz continuous at* $(0, \bar{x})$, *then the first stage problem* (1.1) *is calm at* $\bar{x}$.

From the definition it is easy to verify that the set-valued mapping $X(p, q)$ is pseudo upper-Lipschitz continuous at $(0, \bar{x})$ if and only if there exist a constant $\mu > 0$ and $B(\bar{x})$, a neighborhood of $\bar{x}$, such that

$$d(x, X) \leq \mu(\|G(x)_+\| + \|H(x)\|) \quad \forall x \in B(\bar{x}) \cap Q.$$

See [46, Theorem 3.1] for the equivalence in a more general setting. The above property is also referred to as the existence of a local error bound for the feasible region X or metric regularity. Hence any results on the existence of a local error bound or metric regularity of the constraint system may be used as a sufficient condition for pseudo upper-Lipschitz continuity of the perturbed feasible region (see, e.g., Wu and Ye [50] for such sufficient conditions).

By virtue of Proposition 3.3, the following three constraint qualifications are stronger than the calmness condition at a local minimizer when the objective function of the problem (1.1) is Lipschitz continuous.

PROPOSITION 3.4. *Let $X(p, q)$ be defined as in (3.2) and $\bar{x} \in X$. Then $X(p, q)$ is pseudo upper-Lipschitz continuous at $(0, \bar{x})$ under one of the following constraint qualifications:*

(i) *NNAMCQ for problem (1.1) holds at $\bar{x}$,*

$$\begin{cases} 0 \in \partial\langle G, \eta^G\rangle(\bar{x}) + \partial\langle H, \eta^H\rangle(\bar{x}) + \mathcal{N}_Q(\bar{x}), \\ 0 \leq \eta^G \perp -G(\bar{x}) \geq 0 \end{cases} \implies \eta^G = 0, \quad \eta^H = 0,$$

where

$$\partial\langle G, \eta^G\rangle(\bar{x}) + \partial\langle H, \eta^H\rangle(\bar{x}) = D^*G(\bar{x})(\eta^G) + D^*G(\bar{x})(\eta^H),$$

and when G and H are differentiable at $\bar{x}$,

$$\partial\langle G, \eta^G\rangle(\bar{x}) + \partial\langle H, \eta^H\rangle(\bar{x}) = \nabla G(\bar{x})^T \eta^G + \nabla H(\bar{x})^T \eta^H.$$

(ii) *LICQ holds at $\bar{x}$:*

$$0 \in \partial\langle G, \eta^G\rangle(\bar{x}) + \partial\langle H, \eta^H\rangle(\bar{x}) \implies \eta^G = 0, \quad \eta^H = 0.$$

(iii) *$G(x), H(x)$ are affine functions, and Q is a finite union of convex polyhedral sets.*

Proof. Part (ii) is obviously stronger than part (i). Under part (i), by [54, Theorem 4.4], the perturbed feasible region of the constraint system is pseudo Lipschitz continuous. Under part (iii), the graph of the set-valued mapping X,

$$gphX(\cdot, \cdot) := \{(x, p, q) : G(x) + p \leq 0, H(x) + q = 0, x \in Q, p \in \mathbb{R}^s, q \in \mathbb{R}^r\}$$

is a union of convex polyhedral sets, and hence the perturbed feasible region of the constraint system is upper-Lipschitz by Robinson [39]. □

3.2. First order necessary optimality conditions. In order to derive the optimality conditions, we need the following assumption.

Assumption 3.5. Let $x \in X$ be fixed. There exists a nonnegative function $\sigma(\xi)$ with $\mathbb{E}[\sigma(\xi(\omega))] < \infty$ such that

$$\max(\|\nabla_x f_2(x, y, z, \xi)\|, \|\nabla_x \psi(x, y, z, \xi)\|, \|\nabla_x F(x, y, z, \xi)\|) \leq \sigma(\xi)$$

for all $(y, z) \in \Gamma(x, \xi)$.

THEOREM 3.6 (necessary optimality conditions based on the value function). *Let $\bar{x} \in X$ be a local optimal solution of problem* (1.1)*. Suppose* (a) *Assumptions* 2.14 *and* 3.5 *hold at $\bar{x}$ for every $\xi \in \Xi$;* (b) *for every $(y, z) \in \Gamma(\bar{x}, \xi)$ and every $\xi \in \Xi$, either NNAMCQ holds or* (1.1) *has no inequality constraint $\psi \leq 0$ and one of the constraint qualifications given in Proposition* 2.13 *holds;* (c) *problem* (1.1) *is calm at $\bar{x}$. Then*

(i) *there exist multipliers η^G, η^H such that*

$$(3.3) \quad \begin{cases} 0 \in \partial f_1(\bar{x}) + \partial V(\bar{x}) + \partial\langle G, \eta^G\rangle(\bar{x}) + \partial\langle H, \eta^H\rangle(\bar{x}) + \mathcal{N}_Q(\bar{x}), \\ 0 \leq \eta^G \perp -G(\bar{x}) \geq 0; \end{cases}$$

(ii) *there exist multipliers η^G, η^H such that*

$$(3.4) \quad \begin{cases} 0 \in \partial f_1(\bar{x}) + \mathbb{E}[\partial_x v(\bar{x}, \xi)] + \partial\langle G, \eta^G\rangle(\bar{x}) + \partial\langle H, \eta^H\rangle(\bar{x}) + \mathcal{N}_Q(\bar{x}), \\ 0 \leq \eta^G \perp -G(\bar{x}) \geq 0; \end{cases}$$

(iii) *there exists a measurable selection $q(\bar{x}, \omega) \in \partial_x v(\bar{x}, \xi(\omega))$ and Lagrange multipliers η^G, η^H such that*

$$(3.5) \quad \begin{cases} 0 \in \partial f_1(\bar{x}) + \mathbb{E}[q(\bar{x}, \omega)] + \partial\langle G, \eta^G\rangle(\bar{x}) + \partial\langle H, \eta^H\rangle(\bar{x}) + \mathcal{N}_Q(\bar{x}), \\ 0 \leq \eta^G \perp -G(\bar{x}) \geq 0. \end{cases}$$

Proof. Part (i). Under conditions (a) and (b) (Assumption 2.14), Proposition 2.15 states that $v(x, \xi')$ is Lipschitz near $(\bar{x}, \xi)$. By Proposition 2.15 (ii), it is easy to see that under Assumption 3.5 there exists a constant $c > 0$ such that $\|\partial_x v(\bar{x}, \xi)\| \leq c\sigma(\xi)$ which implies that the Lipschitz constant of $v(x, \xi)$ is bounded by a nonnegative integrable function $\kappa(\xi) := c\sigma(\xi)$. By Theorem 2.16 (iv), $V(x)$ is Lipschitz near $\bar{x}$. Applying the first order necessary optimality condition involving limiting subdifferentials obtained by Mordukhovich in [26, Theorem 1 (b)] (see also [42, Corollary 6.15]) to the penalized problem (3.1), we obtain (3.3).

Part (ii). By Theorem 2.16, $\mathbb{E}[\partial_x v(x, \xi)]$ is well-defined and $\partial V(\bar{x}) \subset \mathbb{E}[\partial_x v(\bar{x}, \xi(\omega))]$. The conclusion follows from part (i).

Part (iii). By part (i), there exist $\hat{q}(\bar{x}) \in \partial V(\bar{x})$ and Lagrange multipliers η^G, η^H such that

$$(3.6) \quad \begin{cases} 0 \in \partial f_1(\bar{x}) + \hat{q}(\bar{x}) + \partial\langle G, \eta^G\rangle(\bar{x}) + \partial\langle H, \eta^H\rangle(\bar{x}) + \mathcal{N}_Q(\bar{x}), \\ 0 \leq \eta^G \perp -G(\bar{x}) \geq 0. \end{cases}$$

Therefore for $\hat{q}(\bar{x})$, there exists a measurable selection $q(\bar{x}, \omega) \in \partial_x v(\bar{x}, \xi(\omega))$ such that $\hat{q}(\bar{x}) = \mathbb{E}[q(\bar{x}, \omega)]$. The conclusion follows from part (ii). $\square$

The optimality conditions derived above utilize explicitly the limiting subdifferential of the value function of the second stage problem. In Theorem 3.6 (i), we assume that $\partial V(\bar{x})$ is computable while in parts (ii)–(iii) of the theorem we assume that $\partial_x v(x, \xi)$ is computable. In some practical circumstances, calculating these subdifferentials may be difficult or impossible. Consequently, we may use the sensitivity analysis of the value function in section 2 to replace the subdifferentials with the gradients of the underlying functions of the second stage problem at optimal solution points. Specifically, we replace $\partial_x v(x, \xi)$ with set $\Psi(\bar{x}, \xi)$ defined in (2.10) although the latter is larger in general. This motivates us to derive the following more general necessary optimality conditions.

THEOREM 3.7 (general necessary optimality condition for the true problem). *Let $\bar{x}$ be a local optimal solution of the true problem* (1.1)*. Assume conditions* (a)–(c) *of Theorem* 3.6 *hold. Then*

(i) *there exists η^G, η^H such that*

$$(3.7) \begin{cases} 0 \in \partial f_1(\bar{x}) + \mathbb{E}[\Psi(\bar{x}, \xi(\omega))] + \partial\langle G, \eta^G\rangle(\bar{x}) + \partial\langle H, \eta^H\rangle(\bar{x}) + \mathcal{N}_Q(\bar{x}), \\ 0 \le \eta^G \perp -G(\bar{x}) \ge 0, \end{cases}$$

where $\Psi(x, \xi)$ is defined as in (2.10);

(ii) *there exist a selection*

$$(y(\omega), z(\omega)) \in \Gamma(\bar{x}, \xi(\omega)), \quad (\gamma(\omega), \eta(\omega)) \in M(\bar{x}, y(\omega), z(\omega), \xi(\omega))$$

and multipliers η^G, η^H such that

$$(3.8) \begin{cases} 0 \in \partial f_1(\bar{x}) + \mathbb{E}[\nabla_x f_2(\bar{x}, y(\omega), z(\omega), \xi(\omega)) \\ \quad + \nabla_x \psi(\bar{x}, y(\omega), z(\omega), \xi(\omega))^\top \gamma(\omega) \\ \quad + \nabla_x F(\bar{x}, y(\omega), z(\omega), \xi(\omega))^\top \eta(\omega)] \\ \quad + \partial\langle G, \eta^G\rangle(\bar{x}) + \partial\langle H, \eta^H\rangle(\bar{x}) + \mathcal{N}_Q(\bar{x}), \\ 0 \le \eta^G \perp -G(\bar{x}) \ge 0; \end{cases}$$

(iii) *there exist M-stationary point $(y(\omega), z(\omega))$ of* (1.2) *and corresponding M-multipliers $\gamma(\omega), \eta(\omega)$, together with the first stage Lagrange multipliers η^G, η^H, such that*

$$(3.9) \begin{cases} 0 \in \partial f_1(\bar{x}) + \mathbb{E}[\nabla_x f_2(\bar{x}, y(\omega), z(\omega), \xi(\omega)) \\ \quad + \nabla_x \psi(\bar{x}, y(\omega), z(\omega), \xi(\omega))^\top \gamma(\omega) \\ \quad + \nabla_x F(\bar{x}, y(\omega), z(\omega), \xi(\omega))^\top \eta(\omega)] \\ \quad + \partial\langle G, \eta^G\rangle(\bar{x}) + \partial\langle H, \eta^H\rangle(\bar{x}) + \mathcal{N}_Q(\bar{x}), \\ 0 \le \eta^G \perp -G(\bar{x}) \ge 0, \\ 0 \in \nabla_y f_2(\bar{x}, y(\omega), z(\omega), \xi(\omega)) + \nabla_y \psi(\bar{x}, y(\omega), z(\omega), \xi(\omega))^\top \gamma(\omega) \\ \quad + \nabla_z F(\bar{x}, y(\omega), z(\omega), \xi(\omega))^\top \eta(\omega), \\ 0 \in \nabla_z f_2(\bar{x}, y(\omega), z(\omega), \xi(\omega)) + \nabla_z \psi(\bar{x}, y(\omega), z(\omega), \xi(\omega))^\top \gamma(\omega) \\ \quad + \nabla_z F(\bar{x}, y(\omega), z(\omega), \xi(\omega))^\top \eta(\omega) \\ \quad + D^*\mathcal{N}_C(z(\omega), -F(\bar{x}, y(\omega), z(\omega), \xi(\omega)))(\eta(\omega)), \\ 0 \in F(\bar{x}, y(\omega), z(\omega), \xi(\omega)) + \mathcal{N}_C(z(\omega)), \\ 0 \le -\psi(\bar{x}, y(\omega), z(\omega), \xi(\omega)) \perp \gamma(\omega) \ge 0. \end{cases}$$

Remark 3.8. Before presenting a proof, we make a few comments on the statements of the theorem.

- First, let us compare the optimality conditions with those in Theorem 3.6. Part (i) corresponds to Theorem 3.6 (ii), and the conditions here are weaker in the sense that $\mathbb{E}[\Psi(\bar{x}, \xi)]$ contains $\mathbb{E}[\partial_x v(\bar{x}, \xi)]$. Part (ii) is equivalent to Theorem 3.6 (iii), but it is no longer described in terms of the subdifferential of the value function here. This is a significant difference from the numerical point of view in that $\mathbb{E}[\partial_x v(\bar{x}, \xi)]$ requires the calculation of the subdifferential of the optimal value function of the second stage problem which is numerically difficult particularly when the problem is nonconvex.

- Now let us compare the statements of Theorem 3.7. The condition in part (ii) is obviously sharper than that of part (i), and it uses only the derivatives of the underlying function of the second stage problem at one optimal solution and a single pair of the corresponding M-multipliers. Part (iii) is a simple relaxation from optimal solution to an M-stationary point so that the optimality condition no longer includes an implicit constraint $(y(\omega), z(\omega)) \in \Gamma(\bar{x}, \xi(\omega))$.

Proof of Theorem 3.7. Part (i). By Proposition 2.15 (ii), $\partial_x v(\bar{x}, \xi) \subset \Psi(\bar{x}, \xi)$. By Theorem 2.16, the set-valued mapping $\omega \to \partial_x v(\bar{x}, \xi(\omega))$ is measurable. Therefore $\mathbb{E}[\Psi(\bar{x}, \xi(\omega))]$ is nonempty and $\mathbb{E}[\partial_x v(\bar{x}, \xi(\omega))] \subset \mathbb{E}[\Psi(\bar{x}, \xi(\omega))]$. From part (iii) of Theorem 3.6, there exists a measurable selection $q(\bar{x}, \omega) \in \partial_x v(\bar{x}, \xi(\omega))$ such that

$$\begin{cases} 0 \in \partial f_1(\bar{x}) + \mathbb{E}[q(\bar{x}, \omega)] + \partial \langle G, \eta^G \rangle(\bar{x}) + \partial \langle H, \eta^H \rangle(\bar{x}) + \mathcal{N}_Q(\bar{x}), \\ 0 \leq \eta^G \perp -G(\bar{x}) \geq 0. \end{cases}$$

Since $q(\bar{x}, \omega) \in \Psi(\bar{x}, \xi(\omega))$, then $\mathbb{E}[q(\bar{x}, \omega)] \in \mathbb{E}[\Psi(\bar{x}, \xi(\omega))]$. This shows part (i).

Part (ii). Since $q(\bar{x}, \omega) \in \Psi(\bar{x}, \xi(\omega))$, by the definition of $\Psi(\bar{x}, \xi(\omega))$, there must be a selection $(y(\omega), z(\omega)) \in \Gamma(\bar{x}, \xi(\omega))$, $(\gamma(\omega), \eta(\omega)) \in M(\bar{x}, y(\omega), z(\omega), \xi(\omega))$ such that

$$q(\bar{x}, \xi(\omega)) = \nabla_x f_2(\bar{x}, y(\omega), z(\omega), \xi(\omega)) + \nabla_x \psi(\bar{x}, y(\omega), z(\omega), \xi(\omega))^\top \gamma(\omega)$$
$$+ \nabla_x F(\bar{x}, y(\omega), z(\omega), \xi(\omega))^\top \eta(\omega).$$

The conclusion follows.

Part (iii) follows from part (ii) because any optimal solution $(y(\omega), z(\omega))$ must be an M-stationary point. $\quad \square$

Note that in Theorem 3.7, we do not require the measurability of $\Psi(\bar{x}, \xi(\omega))$. Indeed the well definedness (nonemptiness) comes from fact that $\Psi(\bar{x}, \xi(\omega))$ contains a measurable and integrable subset $\partial_x v(\bar{x}, \xi(\omega))$. Note also that we do not claim in Theorem 3.7 (ii) the measurability of multipliers $\gamma(\omega)$ and $\eta(\omega)$. However, the measurability of $\Psi(\bar{x}, \xi(\omega))$ and the multipliers are important properties when one discusses the convergence of sample average approximation methods for solving the SMPEC (1.1)–(1.2) (see [35]). In what follows, we obtain these properties under a stronger inf-compactness condition and hence strengthen the optimality conditions of Theorem 3.7.

Assumption 3.9 (uniform inf-compactness). Let $x \in X$ fixed. For every $\xi \in \Xi$, there exists a constant $\delta > 0$ such that the set

$$\{(y, z) : \psi(x, y, z, \xi') \leq q, r \in F(x, y, z, \xi') + \mathcal{N}_C(z), f_2(x, y, z, \xi') \leq \alpha, (r, q) \in B(0, \delta)\}$$

is bounded for every constant α and every ξ' in a closed neighborhood of ξ relative to Ξ.

We need an intermediate result about the upper semicontinuity of $M(\bar{x}, \cdot, \cdot, \cdot)$.

Let $\xi \in \Xi$ and $\mathcal{B}(\xi)$ denote a small closed neighborhood of ξ relative to Ξ. Let

$$\mathcal{H} := \{\Gamma(\bar{x}, \xi') \times \{\xi'\} : \xi' \in B(\xi)\}.$$

Then $\mathcal{H}$ is a collection of certain sets in space $\mathbb{R}^l \times \mathbb{R}^m \times \Xi$. Let $(y, z) \in \Gamma(\bar{x}, \xi)$. We say $M(\bar{x}, \cdot, \cdot, \cdot)$ is upper semicontinuous at (y, z, ξ) *relative* to set $\mathcal{H}$ if for every $\nu > 0$, there exists $\delta > 0$ such that

$$M(\bar{x}, y', z', \xi') \subset M(\bar{x}, y, z, \xi) + \nu \mathcal{B}$$

for all $(y', z', \xi') \in \text{cl}B((y, z, \xi), \delta) \cap \mathcal{H}$, where $\mathcal{B}$ denotes the closed unit ball in space $\mathbb{R}^{p+m}$ and $\text{cl}B((y, z, \xi), \delta) \cap \mathcal{H}$ denotes a closed ball in $\mathbb{R}^l \times \mathbb{R}^m \times \mathbb{R}^d$ with radius δ and center (y, z, ξ).

LEMMA 3.10. *Let Assumption* 3.9 *hold, and let* $\xi \in \Xi$ *and* $(y, z) \in \Gamma(\bar{x}, \xi)$. *Then* $M(\bar{x}, \cdot, \cdot, \cdot)$ *is upper semicontinuous at* (x, y, ξ) *relative to set* $\mathcal{H}$.

Proof. Let $\{\xi_k\} \subset \mathcal{B}(\xi)$ such that $\xi_k \to \xi$ as $k \to \infty$. Consider a sequence $\{(y_k, z_k)\} \subset \Gamma(\bar{x}, \xi_k)$, $\{(\gamma_k, \eta_k)\} \subset M(\bar{x}, y_k, z_k, \xi_k)$. Under Assumption 3.9, it is easy to prove that both sequences $\{(y_k, z_k)\}$ and $\{(\gamma_k, \eta_k)\}$ are bounded.

Let $(y_k, z_k) \to (y, z)$, and assume by taking a subsequence if necessary that $(\gamma_k, \eta_k) \to (\gamma, \eta)$. Using (2.8) and driving k to infinity, we know that $(\gamma, \eta) \in M(\bar{x}, y, z, \xi)$. This shows that $M(\bar{x}, \cdot, \cdot, \cdot)$ is closed at (y, z, ξ). It also implies that $M(\bar{x}, \cdot, \cdot, \cdot)$ is uniformly compact near (y, z, ξ). The two properties give rise to the upper semicontinuity of $M(\bar{x}, \cdot, \cdot, \cdot)$ at (x, y, ξ) relative to set $\mathcal{H}$. $\quad\square$

Using Lemma 3.10, we are able to obtain a stronger version of Theorem 3.7 with multipliers $\gamma(\omega)$ and $\eta(\omega)$ being measurable.

THEOREM 3.11 (general necessary optimality conditions with measurability). *Let $\bar{x}$ be a local solution of the true problem* (1.1). *Assume conditions* (a)–(c) *of Theorem 3.6 and Assumption 3.9 hold at $\bar{x}$. Then*

(i) $\Psi(\bar{x}, \xi(\omega))$ *is integrably bounded and measurable, and there exist multipliers η^G, η^H such that* (3.7) *holds;*

(ii) *there exist $(y(\omega), z(\omega)) \in \Gamma(\bar{x}, \xi(\omega))$, a measurable selection*

$$(\gamma(\omega), \eta(\omega)) \in M(\bar{x}, y(\omega), z(\omega), \xi(\omega)),$$

and multipliers η^G, η^H such that (3.8) *holds;*

(iii) *there exists a M-stationary point $(y(\omega), z(\omega))$ of* (1.2) *and a corresponding measurable M-multiplier $(\gamma(\omega), \eta(\omega))$, together with the first stage Lagrange multipliers η^G, η^H, such that* (3.9) *holds.*

Proof. Part (i). We need only to show that $\Psi(\bar{x}, \xi(\omega))$ is measurable. To this end, we show that the set-valued mapping $\Psi(\bar{x}, \cdot)$ is upper semicontinuous on Ξ.

Let $\xi \in \Xi$ be fixed. Note that Assumption 3.9 implies the inf-compact condition in Assumption 2.14. By Proposition 2.15 (iii), $\Gamma(\bar{x}, \xi)$ is compact for every $\xi \in \Xi$. Moreover, under Assumption 3.9, $\Gamma(\bar{x}, \cdot)$ is closed at ξ. Let $\mathcal{B}(\xi)$ denote a small closed neighborhood (hence compact) of ξ relative to Ξ and $\mathcal{G}(\xi) := \bigcup_{\xi' \in \mathcal{B}(\xi)} \{\Gamma(\bar{x}, \xi')\}$. The properties of Γ stated above guarantee the boundedness of $\mathcal{G}(\xi)$ and closedness of $\mathcal{G}(\cdot)$ at ξ. This and Assumption 3.5 imply that there exists a constant positive constant C such that

$$\sup_{\xi' \in \mathcal{B}(\xi), (y,z) \in \Gamma(\bar{x}, \xi')} (\|\nabla_x f_2(\bar{x}, y, z, \xi')\|, \|\nabla_x \psi(\bar{x}, y, z, \xi')\|, \|\nabla_x F(\bar{x}, y, z, \xi')\|) \leq C.$$

On the other hand, from Proposition 2.12, we know that $\mathcal{M}(\bar{x}, \xi)$ is bounded, where $\mathcal{M}(\bar{x}, \xi)$ is defined as in (2.7). Let $\{\xi_k\} \subset \mathcal{B}(\xi)$ be such that $\xi_k \to \xi$. We show that $\bigcup_{\xi' \in \mathcal{B}(\xi)} \mathcal{M}(\bar{x}, \xi')$ is bounded. Assume for the sake of contradiction that this is not true. Then there exists a sequence $\{\xi_k\} \subset \mathcal{B}(\xi)$, $\{\xi_k\} \to \bar{\xi} \in \Xi$, $(y_k, z_k) \in \Gamma(\bar{x}, \xi_k) \subset \bigcup_{\xi' \in \mathcal{B}(\xi)} \{\Gamma(\bar{x}, \xi')\}$, and $(\gamma_k, \eta_k) \in \mathcal{M}(\bar{x}, \xi_k)$ such that $\{(\gamma_k, \eta_k)\}$ is unbounded. Since $\bigcup_{\xi' \in \mathcal{B}(\xi)} \{\Gamma(\bar{x}, \xi')\}$ is compact, we can assume by extracting a subsequence if necessary that $(y_k, z_k) \to (\bar{y}, \bar{z}) \in \Gamma(\bar{x}, \xi)$ as $\xi_k \to \xi$. Using a similar argument to that of the proof of Proposition 2.12, we can obtain a contradiction to the NNAMCQ at $(\bar{x}, \bar{y}, \bar{z}, \xi)$. This shows the boundedness of $\bigcup_{\xi' \in \mathcal{B}(\xi)} \mathcal{M}(\bar{x}, \xi')$, and together with the boundedness of $\mathcal{G}(\xi)$, this implies the boundedness of $\Psi(\bar{x}, \xi')$ over $\mathcal{B}(\xi)$.

To show the closedness of $\Psi(\bar{x}, \cdot)$ at ξ, it suffices to show the closedness of $\bigcup_{\xi' \in \mathcal{B}(\xi)} \mathcal{M}(\bar{x}, \xi')$. This can be done by considering a sequence $\{\xi_k\} \subset \mathcal{B}(\xi)$, $\{\xi_k\} \to \xi \in \Xi$, $(y_k, z_k) \in \Gamma(\bar{x}, \xi_k) \subset \bigcup_{\xi' \in \mathcal{B}(\xi)} \{\Gamma(\bar{x}, \xi')\}$ with $(y_k, z_k) \to (y, z)$ and $(\gamma_k, \eta_k) \in \mathcal{M}(\bar{x}, \xi_k)$ with $(\gamma_k, \eta_k) \to (\gamma, \eta)$ and substituting them into (2.8). Taking a

limit on both sides of the equation, we can show that $(\gamma, \eta) \in M(\bar{x}, y, z, \xi) \subset \mathcal{M}(\bar{x}, \xi)$ and hence the closedness.

Through Proposition 2.6, this gives the upper semicontinuity of $\Psi(\bar{x}, \xi)$. The measurability follows straightforwardly from [42, Corollary 14.14] because we can view $\Psi(\bar{x}, \xi(\omega))$ as a composition of an upper semincontinuous set-valued mapping $\Psi(\bar{x}, \cdot)$ and a random vector $\xi(\omega)$ (which is measurable).

Part (ii). For the given $q(\bar{x}, \omega)$ specified in part (iii) of Theorem 3.6, we know that $q(\bar{x}, \omega) \in \Psi(\bar{x}, \xi(\omega))$. Therefore there exists $(y(\omega), z(\omega)) \in \Gamma(\bar{x}, \xi(\omega))$ (which is measurable by [3, Theorem 8.2.11]) such that

$$q(\bar{x}, \omega) \in \bigcup_{(\gamma, \eta) \in M(\bar{x}, y(\omega), z(\omega), \xi(\omega))} \{\nabla_x f_2(\bar{x}, y(\omega), z(\omega), \xi(\omega)) + \nabla_x \psi(\bar{x}, y(\omega), z(\omega), \xi(\omega))^T \gamma$$
$$+ \nabla_x F(\bar{x}, y(\omega), z(\omega), \xi(\omega))^\top \eta\}.$$

We can rewrite the above inclusion as

$$(3.10) \qquad q(\bar{x}, \omega) - \nabla_x f_2(\bar{x}, y(\omega), z(\omega), \xi(\omega)) \in R(\omega, M(\bar{x}, y(\omega), z(\omega), \xi(\omega)))$$

where

$$R(\omega, u) := (\nabla_x \psi(\bar{x}, y(\omega), z(\omega), \xi), \nabla_x F(\bar{x}, y(\omega), z(\omega), \xi(\omega)))^T u.$$

Note that $R(\omega, u)$ is a Carathéodory mapping; i.e., $R(\cdot, u)$ is measurable and $R(\omega, \cdot)$ is continuous. Recall that in Lemma 3.10, we have shown that $M(\bar{x}, y, z, \xi)$ is upper semicontinuous with respect to (y, z, ξ) relative to $\mathcal{H}$. Viewing $M(\bar{x}, y(\omega), z(\omega), \xi(\omega))$ as a composition of $M(\bar{x}, \cdot, \cdot, \cdot)$ and a random vector $(y(\omega), z(\omega), \xi(\omega))$, we obtain the measurability of $M(\bar{x}, y(\omega), z(\omega), \xi(\omega))$ through [42, Corollary 14.14]. Applying Filippov's theorem [3, Theorem 8.2.10] to (3.10), we can obtain a measurable selection

$$(\gamma(\omega), \eta(\omega)) \in M(\bar{x}, y(\omega), z(\omega), \xi(\omega))$$

such that

$$q(\bar{x}, \omega) = \nabla_x f_2(\bar{x}, y(\omega), z(\omega), \xi(\omega)) + R(\omega, (\gamma(\omega), \eta(\omega)))$$
$$= \nabla_x f_2(\bar{x}, y(\omega), z(\omega), \xi(\omega)) + \nabla_x \psi(\bar{x}, y(\omega), z(\omega), \xi(\omega))^\top \gamma(\omega)$$
$$+ \nabla_x F(\bar{x}, y(\omega), z(\omega), \xi(\omega))^\top \eta(\omega).$$

The conclusion follows by combining this and (3.5).

Part (iii) is trivial[3] as it follows from part (ii). ▯

4. The case of complementarity constraints. In this section, we consider a special case when $\mathcal{C} = R_+^m$ in the second stage problem (1.2). Consequently, we can write the problem as

$$(4.1) \qquad \begin{aligned} \min_{(y,z) \in \mathbb{R}^l \times \mathbb{R}^m} \quad & f_2(x, y, z, \xi) \\ \text{s.t.} \quad & 0 \le F(x, y, z, \xi) \perp z \ge 0, \\ & \psi(x, y, z, \xi) \le 0, \end{aligned}$$

---

[3]We added the statement following a referee's comment that it might be of interest to present first order necessary conditions with the second stage part characterized for stationary points instead of global optimal solutions as in part (ii) even though the conditions are obviously weaker than those stated in the part (ii).

and the SMPEC (defined by (1.1) and (1.2)) becomes an SMPCC (defined by (1.1) and (4.1)) or equivalently

$$\min_{x,y,z(\cdot)} \quad f_1(x) + \mathbb{E}\left[f_2(x, y, z(\omega), \xi(\omega))\right]$$

(4.2)
$$\begin{aligned} \text{s.t.} \quad & G(x) \leq 0, \\ & H(x) = 0, \\ & x \in Q, \\ & 0 \leq F(x, y, z(\omega), \xi(\omega)) \perp z(\omega) \geq 0 \ \text{ for a.e. } \omega, \\ & \psi(x, y, z(\omega), \xi(\omega)) \leq 0 \ \text{ for a.e. } \omega. \end{aligned}$$

Our focus here is to derive the first order necessary optimality conditions for the SMPCC. While the optimality conditions derived in the previous section can be applied to SMPCC broadly speaking, it might be of independent interest to investigate the specific features of the optimality conditions for this problem.

Before proceeding with further discussion, we introduce some notation specific for this problem. We continue to use $v(x, \xi)$ to denote the optimal value of (4.1) and $\Gamma(x, \xi)$ its optimal solution set. Let $(x, \xi) \in X \times \Xi$ be fixed. For each feasible solution (y, z) of (4.1) we define the index sets

$$\begin{aligned} \mathcal{I}(y, z) &:= \{i : \psi_i(x, y, z, \xi) = 0\}, \\ L := L(y, z) &:= \{i : z_i > 0, F_i(x, y, z, \xi) = 0\}, \\ I_+ := I_+(y, z) &:= \{i : z_i = 0, F_i(x, y, z, \xi) > 0\}, \\ I_0 := I_0(y, z) &:= \{i : z_i = 0, F_i(x, y, z, \xi) = 0\}. \end{aligned}$$

It is important to note that these index sets depend on both x and ξ.

4.1. Constraint qualifications and stationary points. By using Proposition 2.3 to express the coderivative $D^* \mathcal{N}_{R_+^m}(z, -F(x, y, z, \xi))(\eta)$ explicitly, we can write an M-stationary point of (4.1) in the well-known form as in the following definition. Moreover as it is well-known in the literature (see, e.g., [56]) we can define the Clarke stationary point (C-stationary point) and Strong stationary point (S-stationary point).

DEFINITION 4.1 (C-, M- and S-stationary points). *Let $x \in X$ be fixed, and let (y, z) be a feasible solution of the second stage problem* (4.1). *We say that (y, z) is an M-stationary point and $(\gamma, \eta) \in \mathbb{R}_+^p \times \mathbb{R}^m$ is an M-multiplier for problem* (4.1) *if*

$$(4.3) \quad 0 = \nabla_{y,z} f_2(x, y, z, \xi) + \nabla_{y,z} \psi(x, y, z, \xi)^T \gamma + \nabla_{y,z} F(x, y, z, \xi)^T \eta + \{(0, \zeta)\},$$
$$(4.4) \quad 0 = \psi(x, y, z, \xi)^T \gamma,$$
$$\zeta_L = 0, \quad \eta_{I_+} = 0,$$
$$\forall i \in I_0, \ \text{either } \zeta_i < 0, \ \eta_i < 0, \ \text{or } \zeta_i \eta_i = 0.$$

We say that (y, z) is an S-stationary point and $(\gamma, \eta) \in \mathbb{R}_+^p \times \mathbb{R}^m$ is an S-multiplier for problem (4.1) *if* (4.3)–(4.4) *hold and*

$$\zeta_L = 0, \quad \eta_{I_+} = 0,$$
$$\forall i \in I_0, \ \zeta_i \leq 0, \ \eta_i \leq 0.$$

We say that (y, z) is a C-stationary point and $(\gamma, \eta) \in \mathbb{R}_+^p \times \mathbb{R}^m$ is a C-multiplier for problem (4.1) *if* (4.3)–(4.4) *hold and*

$$\zeta_L = 0, \quad \eta_{I_+} = 0,$$
$$\forall i \in I_0, \ \zeta_i \eta_i \geq 0.$$

It is easy to see that the following relationship between the various stationary condition holds:

S-stationary condition $\Rightarrow$ M-stationary condition $\Rightarrow$ C-stationary condition.

Moreover, under the following MPEC linear independence constraint qualification (MPEC-LICQ), a local optimal solution of an MPEC is an S-stationary point and the set of S-multipliers is a singleton; see [23].

DEFINITION 4.2 (MPEC-LICQ). *Let* $x \in X$ *be fixed and* $\mathcal{I}(y, z) := \{i : \psi(x, y, z, \xi) = 0\}$. *We say that MPEC-LICQ holds at a feasible point* (y, z) *of second stage problem* (4.1) *if*

$$0 = \nabla_{y,z}\psi(x, y, z, \xi)^T \gamma + \nabla_{y,z}F(x, y, z, \xi)^T \eta + \{(0, \zeta)\},$$
$$\gamma_i = 0 \ if \ i \in \mathcal{I}(y, z), \eta_{I_+} = 0, \zeta_L = 0$$

imply that $\gamma = 0, \eta = 0,$ *and* $\zeta = 0$.

DEFINITION 4.3 (NNAMCQ for the complementarity constraint). *Let* $(x, \xi) \in X \times \Xi$ *be fixed. We say that NNAMCQ holds at feasible point* (y, z) *of second stage problem* (4.1) *if*

$$\begin{cases} 0 = \nabla_{y,z}\psi(x, y, z, \xi)^T \gamma + \nabla_{y,z}F(x, y, z, \xi)^T \eta + \{(0, \zeta)\}, \\ 0 = \psi(x, y, z, \xi)^T \gamma, \quad \gamma \geq 0, \\ \zeta_L = 0, \quad \eta_{I_+} = 0, \\ \forall i \in I_0, \ either \ \zeta_i < 0, \ \eta_i < 0, \ or \ \zeta_i \eta_i = 0 \end{cases} \Rightarrow \gamma = 0, \quad \eta = 0.$$

It is proved in [54, Proposition 4.5] that the NNAMCQ is equivalent to an MPEC variant of MFCQ defined as follows.

DEFINITION 4.4 (MPEC-GMFCQ). *We say that MPEC generalized Mangasarian–Fromovitz constraint qualification (MPEC-GMFCQ) holds at a feasible point* (y, z) *of second stage problem* (4.1) *if one of the following holds:*

(a) *for every partition of* I_0 *into sets* P, O, R *with* $R \neq \emptyset$, *there exist vectors* $d \in \mathbb{R}^l, h \in \mathbb{R}^m$ *such that* $h_{I_+} = 0, h_O = 0, h_R \geq 0,$

$$\langle \nabla_{y,z}\psi_i(x, y, z, \xi), (d, h) \rangle \leq 0, \quad i \in \mathcal{I}(y, z),$$
$$\langle \nabla_{y,z}F_i(x, y, z, \xi), (d, h) \rangle = 0, \quad i \in L \cup P,$$
$$\langle \nabla_{y,z}F_i(x, y, z, \xi), (d, h) \rangle \leq 0, \quad i \in R,$$

and either $h_i > 0$ *or*

$$\langle \nabla_{y,z}F_i(x, y, z, \xi), (d, h) \rangle > 0 \ for \ some \ i \in R;$$

(b) *for every partition of* I_0 *into the sets* P, O, *the matrix*

$$\begin{bmatrix} \nabla_y F_{L \cup P}(x, y, z, \xi) & \nabla_z F_{L \cup P, L \cup P}(x, y, z, \xi) \end{bmatrix}$$

has full row rank and there exist vectors $d \in \mathbb{R}^l, h \in \mathbb{R}^m$ *such that*

$$h_{I_+} = 0, \quad h_O = 0,$$
$$\langle \nabla_{y,z}\psi_i(x, y, z, \xi), (d, h) \rangle < 0, \quad i \in \mathcal{I}(y, z),$$
$$\langle \nabla_{y,z}F_i(x, y, z, \xi), (d, h) \rangle = 0, \quad i \in L \cup P.$$

DEFINITION 4.5 (NNAMCQ for the complementarity constraint with C-multipliers). *Let $(x, \xi) \in X \times \Xi$ be fixed. We say that NNAMCQ with C-multipliers holds at feasible point (y, z) of second stage problem (4.1) if*

$$
\left\{
\begin{array}{l}
0 = \nabla_{y,z}\psi(x, y, z, \xi)^T\gamma + \nabla_{y,z}F(x, y, z, \xi)^T\eta + \{(0, \zeta)\}, \\
0 = \psi(x, y, z, \xi)^T\gamma, \quad \gamma \geq 0, \\
\zeta_L = 0, \quad \eta_{I_+} = 0, \\
\forall i \in I_0, \zeta_i\eta_i \geq 0
\end{array}
\right.
\quad \Rightarrow \gamma = 0, \quad \eta = 0.
$$

It is easy to see that the following relationship between the various constraint qualification hold:

$$\text{MPEC-LICQ} \Rightarrow \text{NNAMCQ} \Rightarrow \text{NNAMCQ with C-multipliers.}$$

4.2. First order necessary optimality conditions. We revisit the necessary optimality conditions established in Theorems 3.7 and 3.11 for the two-stage SMPCC defined by (1.1) and (4.1). Note that Assumptions 2.14 and 3.9 can be a bit more specific by writing the variational inequality

$$r \in F(x, y, z, \xi) + \mathcal{N}_C(z)$$

as a complementarity constraint

$$(4.5) \qquad\qquad 0 \leq r - F(x, y, z, \xi) \perp z \geq 0.$$

THEOREM 4.6. *Let $\bar{x}$ be a local optimal solution of the true problem defined by (1.1) and (4.1). Assume (a) Assumptions 2.14 and 3.5 hold at $\bar{x}$ for every $\xi \in \Xi$; (b) for every $(y, z) \in \Gamma(\bar{x}, \xi)$ and every $\xi \in \Xi$; (b1) the NNAMCQ for complementarity constraint (equivalently MPEC-GMFCQ) holds or (4.1) has no inequality constraints and one of the following constraint qualifications holds:*

(b2) *(SRCQ) the matrix $\nabla_z F_{L,L}(\bar{x}, y, z, \xi)$ is nonsingular, and the Schur complement of the above matrix in the matrix*

$$
\left[
\begin{array}{cc}
\nabla_z F_{L,L}(\bar{x}, y, z, \xi) & \nabla_z F_{L,I_0}(\bar{x}, y, z, \xi) \\
\nabla_z F_{I_0,L}(\bar{x}, y, z, \xi) & \nabla_z F_{I_0,I_0}(\bar{x}, y, z, \xi)
\end{array}
\right]
$$

has positive principle minors;

(b3) *$-F$ is locally strongly monotone in z uniformly with respect to y; i.e., there exists positive constant δ independent of y, neighborhoods U_1 of y and U_2 of z such that*

$$
\langle -F(\bar{x}, y', z', \xi) + F(\bar{x}, y', z, \xi), z' - z \rangle \geq \delta\|z' - z\|^2,
$$
$$
\forall z' \in U_2 \cap R_+^m \; \forall z' \in R_+^m, y' \in U_1;
$$

(b4) *the rank of the matrix $\nabla_y F(\bar{x}, y, z, \xi)$ is m;*

(c) *the problem (1.1) is calm at $\bar{x}$.*

Then there exists an M-stationary point $(y(\omega), z(\omega))$ and corresponding multipliers $\gamma(\omega), \eta(\omega)$, together with first stage multipliers η^G, η^H, such that

$$
\left\{
\begin{array}{l}
0 \in \partial f_1(\bar{x}) + \mathbb{E}[\nabla_x f_2(\bar{x}, y(\omega), z(\omega), \xi(\omega)) + \nabla_x\psi(\bar{x}, y(\omega), z(\omega), \xi(\omega))^\top\gamma(\omega) \\
\quad + \nabla_x F(\bar{x}, y(\omega), z(\omega), \xi(\omega))^\top\eta(\omega)] + \partial\langle G, \eta^G\rangle(\bar{x}) + \partial\langle H, \eta^H\rangle(\bar{x}) + \mathcal{N}_Q(\bar{x}), \\
0 \leq \eta^G \perp -G(\bar{x}) \geq 0, \\
0 = \nabla_{y,z} f_2(\bar{x}, y(\omega), z(\omega), \xi(\omega)) + \nabla_{y,z}\psi(\bar{x}, y(\omega), z(\omega), \xi(\omega))^T\gamma(\omega) \\
\quad + \nabla_{y,z}F(\bar{x}, y(\omega), z(\omega), \xi)^T\eta(\omega) + \{(0, \zeta(\omega))\}, \\
0 = \psi(\bar{x}, y(\omega), z(\omega), \xi)^T\gamma(\omega), \quad \gamma(\omega) \geq 0, \\
\zeta_L(\omega) = 0, \quad \eta_{I_+}(\omega) = 0, \\
\forall i \in I_0, \; \text{either } \zeta_i(\omega) < 0, \eta_i(\omega) < 0, \; \text{or } \zeta_i(\omega)\eta_i(\omega) = 0.
\end{array}
\right.
$$

If, in addition, Assumption 3.9 *holds, then there exist measurable (random) multipliers* $\gamma(\omega), \eta(\omega)$ *such that the above optimality conditions hold.*

Proof. By Robinson [38, Theorem 3.1], condition (b2) is equivalent to the strong regularity condition of the generalized equation

$$0 \in F(\bar{x}, y, z, \xi) + \mathcal{N}_{\mathbb{R}_+^m}(z)$$

for each fixed $(\bar{x}, \xi)$. (b3) and (b4) are restatements of Proposition 2.13 parts (ii) and (iii), respectively. By Theorem 3.7, there exist selections

$$(y(\omega), z(\omega)) \in \Gamma(\bar{x}, \xi(\omega)), \quad (\gamma(\omega), \eta(\omega)) \in M(\bar{x}, y(\omega), z(\omega), \xi(\omega))$$

such that

$$\begin{cases} 0 \in \partial f_1(\bar{x}) + \mathbb{E}[\nabla_x f_2(\bar{x}, y(\omega), z(\omega), \xi(\omega)) + \nabla_x \psi(\bar{x}, y(\omega), z(\omega), \xi(\omega))^\top \gamma(\omega) \\ \quad + \nabla_x F(\bar{x}, y(\omega), z(\omega), \xi(\omega))^\top \eta(\omega)] + \partial \langle G, \eta^G \rangle(\bar{x}) + \partial \langle H, \eta^H \rangle(\bar{x}) + \mathcal{N}_Q(\bar{x}), \\ \eta^G \geq 0, \quad \langle \eta^G, G(\bar{x}) \rangle = 0. \end{cases}$$

By the definition of $M(\bar{x}, y(\omega), z(\omega), \xi(\omega))$, one has

$$0 \in \nabla_{y,z} f_2(\bar{x}, y(\omega), z(\omega), \xi(\omega)) + \nabla_{y,z} \psi(\bar{x}, y(\omega), z(\omega), \xi(\omega))^\top \gamma(\omega)$$
$$+ \nabla_{y,z} F(\bar{x}, y(\omega), z(\omega), \xi(\omega))^\top \eta(\omega)$$
$$+ \{0\} \times D^* \mathcal{N}_{\mathbb{R}_+^m}(z(\xi(\omega)), -F(\bar{x}, y(\omega), z(\omega), \xi(\omega)))(\eta(\omega)),$$
$$0 = \psi(\bar{x}, y(\omega), z(\omega), \xi(\omega))^T \gamma(\xi(\omega)), \quad \gamma(\xi(\omega)) \geq 0.$$

Therefore there exists $\zeta(\omega) \in D^* \mathcal{N}_{\mathbb{R}_+^m}(z(\xi(\omega)), -F(\bar{x}, y(\omega), z(\omega), \xi(\omega)))(\eta(\omega))$ such that

$$0 = \nabla_{y,z} f_2(\bar{x}, y(\omega), z(\omega), \xi) + \nabla_{y,z} \psi(\bar{x}, y(\omega), z(\omega), \xi(\omega))^T \gamma(\omega)$$
$$+ \nabla_{y,z} F(\bar{x}, y(\omega), z(\omega), \xi)^T \eta(\omega) + \{(0, \zeta(\omega))\},$$
$$0 = \psi(\bar{x}, y(\omega), z(\omega), \xi(\omega))^T \gamma(\xi(\omega)), \quad \gamma(\xi(\omega)) \geq 0.$$

By the definition of coderivative, $\zeta(\omega) \in D^* \mathcal{N}_{R_+^m}(z(\omega), -F(\bar{x}, y(\omega), z(\omega), \xi(\omega)))(\eta(\omega))$ if and only if $(\zeta(\omega), -\eta(\omega)) \in \mathcal{N}_{gph \mathcal{N}_{R_+^m}}(z(\omega), -F(\bar{x}, y(\omega), z(\omega), \xi(\omega)))$. Consequently, by Proposition 2.3, one has

$$0 = \psi(\bar{x}, y(\omega), z(\omega), \xi(\omega))^T \gamma(\omega), \quad \gamma(\omega) \geq 0,$$
$$\zeta_L(\omega) = 0, \quad \eta_{I_+}(\omega) = 0,$$
$$\forall i \in I_0, \text{ either } \zeta_i(\omega) < 0, \eta_i(\omega) < 0, \text{ or } \zeta_i(\omega)\eta_i(\omega) = 0.$$

Since an optimal solution must be an M-stationary point, the conclusion follows. The existence of measurable multipliers under Assumption 3.9 follows from Theorem 3.11. $\square$

Recall that Xu and Meng [52] investigated a class of SMPCCs where the underlying function in the complementarity constraint is assumed to be uniformly strongly monotone in z. They considered an optimality condition which is derived by reformulating the complementarity constraints as a system of nonsmooth equations and then characterized the optimality condition in terms of Clarke subdifferentials of the reformulated nonsmooth functions together with the corresponding Lagrange multipliers. Our result here has extended their optimality condition [52, Proposition 5.1] in the

following aspects: (a) the element under expectation operator is a singleton rather than a set as in [52, Proposition 5.1] which could be potentially large at a nonsmooth point; (b) we have included an inequality constraint $\psi \leq 0$; (c) the second stage problem here may have multiple solutions; (d) we use the set of M-multipliers for the second stage problem which may be strictly contained in the set of C-multipliers, and hence the resulting necessary condition is sharper.

We can establish the following sharper necessary optimality condition which utilizes S-multipliers instead of M-multipliers of the second stage problem.

THEOREM 4.7 (necessary optimality condition with S-multipliers). *Let $\bar{x}$ be a local solution of the true problem defined by* (1.1) *and* (4.1). *Suppose* (a) *Assumptions 2.14 and 3.5 hold at $\bar{x}$ for every $\xi \in \Xi$;* (b) *for every $(y, z) \in \Gamma(\bar{x}, \xi)$, MPEC-LICQ holds;* (c) *the SMPCC problem defined by* (1.1) *and* (4.1) *is calm at $\bar{x}$. Then there exist a measurable S-stationary point $(y(\omega), z(\omega))$ and corresponding measurable multipliers $\gamma(\omega), \eta(\omega)$, together with the multipliers η^G, η^H, such that*

$$
\begin{cases}
0 \in \partial f_1(\bar{x}) + \mathbb{E}[\nabla_x f_2(\bar{x}, y(\omega), z(\omega), \xi(\omega)) + \nabla_x \psi(\bar{x}, y(\omega), z(\omega), \xi(\omega))^\top \gamma(\omega) \\
\quad + \nabla_x F(\bar{x}, y(\omega), z(\omega), \xi(\omega))^\top \eta(\omega)] + \partial\langle G, \eta^G \rangle(\bar{x}) + \partial\langle H, \eta^H \rangle(\bar{x}) + \mathcal{N}_Q(\bar{x}), \\
0 \leq \eta^G \perp -G(\bar{x}) \geq 0, \\
0 = \nabla_{y,z} f_2(\bar{x}, y(\omega), z(\omega), \xi(\omega)) + \nabla_{y,z} \psi(\bar{x}, y(\omega), z(\omega), \xi(\omega))^T \gamma(\omega) \\
\quad + \nabla_{y,z} F(\bar{x}, y(\omega), z(\omega), \xi(\omega))^T \eta(\omega) + \{(0, \zeta(\omega))\}, \\
0 = \psi(\bar{x}, y(\omega), z(\omega), \xi(\omega))^T \gamma(\omega), \quad \gamma(\omega) \geq 0, \\
\zeta_L(\omega) = 0, \quad \eta_{I_+}(\omega) = 0, \\
\forall i \in I_0, \zeta_i(\omega) \leq 0, \eta_i(\omega) \leq 0.
\end{cases}
$$

Proof. By Theorem 3.7 there exists a selection $(y(\omega), z(\omega)) \in \Gamma(\bar{x}, \xi(\omega))$ and corresponding M-multiplier $\{(\gamma(\omega), \eta(\omega))\} \in M(\bar{x}, y(\omega), z(\omega), \xi(\omega))$ such that (3.8) holds. Since under MPEC-LICQ, any local optimal solution is an S-stationary point with a unique S-multiplier, the set of S-multipliers and the set of M-multipliers coincide (see [23, 53]). The precise expression in the theorem follows by applying the definition of an S-stationary point to (3.8). We now prove the measurability result under extra assumption (a). Recall from Theorem 3.6 that $q(\bar{x}, \xi(\omega))$ is a measurable selection of $\partial_x v(\bar{x}, \xi(\omega))$ and

$$
q(\bar{x}, \xi(\omega)) = \nabla_x f_2(\bar{x}, y(\omega), z(\omega), \xi(\omega)) + \nabla_x \psi(\bar{x}, y(\omega), z(\omega), \xi(\omega))^\top \gamma(\omega) \\
+ \nabla_x F(\bar{x}, y(\omega), z(\omega), \xi(\omega))^\top \eta(\omega).
$$

Hence the measurability of $(\gamma(\omega), \eta(\omega))$ follows from the inverse image theorem for the calculus of measurable maps (see [3, Theorem 8.2.9]). ⬚

It is important to note that the optimality conditions established in Theorem 4.7 do not require uniform inf-compactness as in Theorem 4.6. This is because the set of S-multipliers at an S-stationary point is a singleton, and consequently we may use the inverse image measurability instead of Filippov's theorem to obtain the measurability of S-multipliers. In fact if the set of M-multipliers $M(\bar{x}, y(\omega), z(\omega), \xi(\omega))$ in Theorem 4.6 is a singleton, then we can also conclude the measurability of the selection $(\gamma(\omega), \eta(\omega))$ without uniform inf-compactness in the same way.

To conclude this section, let us make a few more comments. The first order necessary conditions established in Theorems 4.6 and 4.7 are in terms of M- and S-stationarity. In deterministic MPEC, there are a number of other stationarities being considered such as B-stationarity and C-stationarity. It is therefore natural to ask whether we can derive the optimality conditions for SMPCC defined by (1.1) and

(4.1) in terms of B- and C-stationarity. The answer is yes. Indeed, one can easily use the sensitivity analysis in terms of C-multipliers by Lucet and Ye [21, Theorem 4.8] to derive the necessary optimality condition with C-multipliers under the weaker NNAMCQ with C-multipliers. A similar result can be derived for the B-multipliers under the piecewise MPEC MFCQ using [21, Theorem 4.11].

5. The classical two-stage stochastic program. In this section, we consider the more specific case of the second stage problem (1.2) when $\mathcal{C} = \mathbb{R}^m$, and consequently the variational inequality constraint reduces to an equality constraint and the SMPEC problem (1.1)–(1.2) becomes an ordinary two-stage stochastic program

$$
(5.1) \quad
\begin{aligned}
\min \quad & f_1(x) + \mathbb{E}\left[v(x, \xi(\omega))\right] \\
\text{s.t.} \quad & G(x) \leq 0, \\
& H(x) = 0, \\
& x \in Q,
\end{aligned}
$$

where $v(x, \xi)$ is the optimal value of the second stage problem

$$
(5.2) \quad
\begin{aligned}
\min_{y \in \mathbb{R}^l} \quad & f_2(x, y, \xi) \\
\text{s.t.} \quad & \psi(x, y, \xi) \leq 0, \\
& F(x, y, \xi) = 0.
\end{aligned}
$$

The problem has been well-studied in the literature of stochastic programming. For instance, Rockafellar and Wets [41] investigated first order necessary conditions of a similar class of two-stage stochastic programming problems where the underlying functions are convex but not necessarily continuously differentiable, and Hiriart-Urruty [18] took them further to nonconvex cases. Outrata and Römisch [32] derived first order necessary optimality conditions of the problem in terms of limiting subgradients. Their approach is similar to ours, that is, through the limiting subgradients of the value function of the second stage problem. However, since they used Mordukhovich's exchange rule [28, Lemma 6.18], their results require the probability space of ξ to be nonatomic. More recently, inspired by the need for the convergence analysis of the Monte Carlo sampling method applied to the two-stage stochastic program, Ralph and Xu [35] derived a couple of optimality conditions for the *first stage* problem by replacing the limiting subdifferential with the convex hull of the gradients of the Lagrange function of the second stage problems at local optimal solutions and stationary points. We will come back to this after our main results Theorem 5.2.

To proceed with the discussion, we need the standard boundedness condition Assumption 3.5 and the inf-compactness condition Assumption 2.14. The boundedness condition remains the same, while the inf-compactness condition may be more specific by replacing the variational inequality by an equality as follows.

Assumption 5.1 (inf-compactness). Let $(x, \xi) \in X \times \Xi$ be fixed. There exists a constant $\delta > 0$ such that the set

$$
\{y : \psi(x, y, \xi) \leq q, F(x, y, \xi) = r, f_2(x, y, \xi) \leq \alpha, (q, r) \in B(0, \delta)\}
$$

is bounded for every constant α.

THEOREM 5.2 (necessary optimality condition for the classical case). *Let $\bar{x}$ be a local optimal solution of the classical two-stage stochastic program and Assumptions 3.5 and 5.1 hold at $\bar{x}$ for every $\xi \in \Xi$. Assume that MFCQ holds for problem (5.2) at every $y \in \Gamma(\bar{x}, \xi)$, and the problem (5.1) is calm at $\bar{x}$. Then*

(i) *there exist η^G, η^H such that*

$$(5.3)\begin{cases} 0 \in \partial f_1(\bar{x}) + \mathbb{E}[\Psi(\bar{x}, \xi(\omega))] + \partial \langle G, \eta^G \rangle(\bar{x}) + \partial \langle H, \eta^H \rangle(\bar{x}) + \mathcal{N}_Q(\bar{x}), \\ 0 \leq \eta^G \perp -G(\bar{x}) \geq 0, \end{cases}$$

where $\Psi(x, \xi)$ is defined as

$$\Psi(x, \xi) := \bigcup_{y \in \Gamma(x, \xi)} \bigcup_{(\gamma, \eta) \in M(x, y, \xi)} \{\nabla_x f_2(x, y, \xi) + \nabla_x \psi(x, y, \xi)^T \gamma + \nabla_x F(x, y, \xi)^\top \eta\}$$

and $M(x, y, \xi)$ is the set of Lagrange multipliers of the second stage problem (5.2);

(ii) *there exist $y(\omega) \in \Gamma(\bar{x}, \xi(\omega))$ and $\gamma(\omega) \in \mathbb{R}^p$, $\eta(\omega) \in \mathbb{R}^m$, $\eta^G \in \mathbb{R}^s, \eta^H \in \mathbb{R}^r$ such that*

$$(5.4)\begin{cases} 0 \in \partial f_1(\bar{x}) + \mathbb{E}[\nabla_x f_2(\bar{x}, y(\omega), \xi(\omega)) + \nabla_x \psi(\bar{x}, y(\omega), \xi(\omega))^\top \gamma(\omega) \\ \qquad + \nabla_x F(\bar{x}, y(\omega), \xi(\omega))^\top \eta(\omega)] \\ \qquad + \partial \langle G, \eta^G \rangle(\bar{x}) + \partial \langle H, \eta^H \rangle(\bar{x}) + \mathcal{N}_Q(\bar{x}), \\ 0 \leq \eta^G \perp -G(\bar{x}) \geq 0, \\ 0 = \nabla_y f_2(\bar{x}, y(\omega), \xi(\omega)) + \nabla_y \psi(\bar{x}, y(\omega), \xi(\omega))^\top \gamma(\omega) \\ \qquad + \nabla_y F(\bar{x}, y(\omega), \xi(\omega))^\top \eta(\omega), \\ \langle \psi(\bar{x}, y(\omega), \xi(\omega)), \gamma(\omega) \rangle = 0, \quad \gamma(\omega) \geq 0; \end{cases}$$

(iii) *there exist a stationary point $y(\omega)$ and corresponding Lagrange multipliers $\gamma(\omega) \in \mathbb{R}^p$, $\eta(\omega) \in \mathbb{R}^m$, together with the first stage multipliers $\eta^G \in \mathbb{R}^s, \eta^H \in \mathbb{R}^r$, such that* (5.4) *holds.*

If Assumption 5.1 is strengthened to be uniform with respect to ξ, then $\Psi(\bar{x}, \xi(\omega))$ is measurable in statement (i) *and in statements* (ii)–(iii), *existence of measurable multipliers $\gamma(\omega) \in \mathbb{R}^p$, $\eta(\omega) \in \mathbb{R}^m$ are guaranteed.*

Observe first that statement (ii) is indeed Outrata and Römisch's Theorem 3.5 in [32]. Our statement is more general in the sense that the probability space here does not have to be nonatomic. See the theorem and its proof for details.

Let us drop $f_1(x)$. Then the strengthened version of Theorem 5.2 (i) under the uniform inf-compactness coincides with one of the optimality conditions derived by Ralph and Xu [35] when the probability measure of ξ is nonatomic. However, it might be interesting to point out that the conditions are derived in a different way: In [35], Ψ is considered as a relaxation of the Clarke subdifferential of the value function of (5.2), while here it is a relaxation of the limiting subdifferential of the value function. The results here are sharper when the probability measure is atomic.

Let us now discuss parts (ii) and (iii) of the theorem. The conditions are a combination of the classical KKT conditions of the second stage problem and the *new* optimality conditions of the first stage with the following characteristics: (a) the expected value of a gradient of the Lagrange function of the second stage problem with respect to x at a stationary point is used to reflect the derivative information from the second stage problem; (b) the limiting subdifferential instead of Clarke subdifferential of the first stage constraint functions is used; (c) the optimality condition is established under Clarke's calmness condition.

6. Final comments. The first order necessary optimality conditions we derived in this paper have potential implications in the study of numerical methods for solving the two-stage SMPEC problem (1.1)–(1.2). To explain this, let us consider the well-known Monte Carlo sampling method for the SMPEC. In [45], Shapiro and Xu

sketched an NLP relaxation approach for a two-stage SMPEC discretized through Monte Carlo sampling. The same approach can be applied to our problem albeit our second stage problem may have multiple local and/or global solutions. However, the convergence results might be significantly different: When we solve our two-stage discretized SMPEC, we are more likely to obtain a stationary point or a local optimal solution than a global optimal solution because our second stage problem is nonconvex and the variational inequality constraint has multiple solutions. Consequently, the approximate stationary solution of the first stage problem might converge to a stationary point characterized by the optimality condition (3.7) or an M-stationary point under some specific circumstances. This kind of asymptotic analysis has been recently carried out by Ralph and Xu in [35] for a classical two-stage stochastic programming problem where the second stage generally has multiple local and/or global optimal solutions. The optimality conditions we derived here lay down a foundation for the Monte Carlo sampling MPEC-NLP approach to be applied to the SMPEC problem (1.1)–(1.2). They might also be used for the convergence analysis of a stochastic approximation method proposed by Gaivoronski and Werner for solving a class of two-stage stochastic bilevel programming problems [11] (where the equilibrium conditions reformulated from KKT conditions of the lower level program typically have multiple solutions).

In summary, the second stage problem in a two-stage SMPEC usually has multiple local and/or global optimal solutions. The Monte Carlo sampling method coupled with the NLP-MPEC relaxation or the stochastic approximation method may be applied to solve it, and the statistical estimators obtained from the discretized SMPEC often converge to a stationary point characterized by one of our optimality conditions.

REFERENCES

[1] Z. Artstein and R. A. Vitale, *A strong law of large numbers for random compact sets*, Ann. Probab., 3 (1975), pp. 879–882.

[2] J.-P. Aubin, *Lipschitz behavior of solutions to convex minimization problems*, Math. Oper. Res., 9 (1994), pp. 87–111.

[3] J.-P. Aubin and H. Frankowska, *Set-Valued Analysis*, Birkhäuser, Boston, 1990.

[4] R. J. Aumann, *Integrals of set-valued functions*, J. Math. Anal. Appl., 12 (1965), pp. 1–12.

[5] S. Christiansen, M. Patriksson, and L. Wynter, *Stochastic bilevel programming in structural optimization*, Struct. Multidiscip. Optim., 21 (2001), pp. 361–371.

[6] F. H. Clarke, *Optimization and Nonsmooth Analysis*, Wiley, New York, 1983.

[7] F. H. Clarke, Y. S. Ledyaev, R. J. Stern, and P. R. Wolenski, *Nonsmooth Analysis and Control Theory*, Springer, New York, 1998.

[8] A. L. Dontchev and R. T. Rockafellar, *Characterizations of strong regularity for variational inequalities over polyhedral convex sets*, SIAM J. Optim., 6 (1996), pp. 1087–1105.

[9] J. Gauvin, *A necessary and sufficient regularity condition to have bounded multipliers in nonconvex programming*, Math. Program., 12 (1977), pp. 136–138.

[10] J. Gauvin and F. Dubeau, *Differential properties of the marginal functions in mathematical programming*, Math. Program. Stud., 19 (1982), pp. 101–119.

[11] A. Gaivoronski and A. Werner, *Modeling of competition and collaboration networks under uncertainty: Stochastic programs with recourse and bilevel structure*, IR-07-041, International Institute for Applied Systems Analysis, Laxenburg, Austria, 2007.

[12] R. Henrion, A. Jourani, and J. Outrata, *On the calmness of a class of multifunctions*, SIAM J. Optim., 13 (2002), pp. 603–618.

[13] R. Henrion and J. Outrata, *Calmness of constraint systems with applications*, Math. Program. Ser. B, 104 (2005), pp. 437–464.

[14] R. Henrion and J. Outrata, *On calculating the normal to a finite union of convex polyhedra*, Optimization, 57 (2008), pp. 57–78.

[15] R. Henrion, J. Outrata, and T. Surowiec, *On the co-derivative of normal cone mappings to inequality systems*, Nonlinear Anal., 71 (2009), pp. 1213–1226.

[16] R. Henrion and W. Römisch, *On M-stationary point for a stochastic equilibrium problem under equilibrium constraints in electricity spot market modeling*, Appl. Math. (N.Y.), 52 (2007), pp. 473–494.

[17] C. Hess, *Set-valued integration and set-valued probability theory: An overview*, in Handbook of Measure Theory, Vols. I and II, North–Holland, Amsterdam, 2002, pp. 617–673.

[18] J. B. Hiriart-Urruty, *Conditions necessaires d'optimalité pour un programme stochastique avec recours*, SIAM J. Control Optim., 16 (1978), pp. 317–329.

[19] W. W. Hogan, *Point-to-set maps in mathematical programming*, SIAM Rev., 15 (1973), pp. 591–603.

[20] G.-H. Lin, X. Chen, and M. Fukushima, *Solving stochastic mathematical programs with equilibrium constraints via approximation and smoothing implicit programming with penalization*, Math. Program., 116 (2009), pp. 343–368.

[21] Y. Lucet and J. J. Ye, *Sensitivity analysis of the value function for optimization problems with variational inequality constraints*, SIAM J. Control Optim., 40 (2001), pp. 699–723.

[22] Y. Lucet and J. J. Ye, *Erratum: Sensitivity analysis of the value function for optimization problems with variational inequality constraints*, SIAM J. Control Optim., 41 (2002), pp. 1315–1319.

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

[24] F. Meng and H. Xu, *A regularized sample average approximation method for stochastic mathematical programs with nonsmooth equality constraints*, SIAM J. Optim., 17 (2006), pp. 891–919.

[25] B. S. Mordukhovich, *Maximum principle in problems of time optimal control with nonsmooth constraints*, J. Appl. Math. Mech., 40 (1976), pp. 960–969.

[26] B. S. Mordukhovich, *Metric approximation and necessary optimality conditions for general classes of nonsmooth extremal problems*, Soviet Math. Dokl., 22 (1980), pp. 526–530.

[27] B. S. Mordukhovich, *Variational Analysis and Generalized Differentiation,* I: *Basic Theory*, Grundlehren Math. Wiss. 330, Springer, New York, 2006.

[28] B. S. Mordukhovich, *Variational Analysis and Generalized Differentiation,* II: *Applications*, Grundlehren Math. Wiss. 331, Springer, New York, 2006.

[29] J. Neveu, *Discrete-Parameter Martingales*, North–Holland, New York, 1975.

[30] J. Outrata, *Optimality conditions for a class of mathematical programs with equilibrium constraints*, Math. Oper. Res., 25 (1999), pp. 627–644.

[31] J. V. Outrata, *A generalized mathematical program with equilibrium constraints*, SIAM J. Control Optim., 38 (2000), pp. 1623–1638.

[32] J. Outrata and W. Römisch, *On optimization conditions for some nonsmooth optimization problems over L^p spaces*, J. Optim. Theory Appl., 126 (2005), pp. 411–438.

[33] M. Patriksson and L. Wynter, *Stochastic mathematical programs with equilibrium constraints*, Oper. Res. Lett., 25 (1999), pp. 159–167.

[34] R. A. Poliquin and R. T. Rockafellar, *Tilt stability of a local minimum*, SIAM J. Optim., 8 (1998), pp. 287–299.

[35] D. Ralph and H. Xu, *Asymptotic Analysis of Stationary Points of Sample Average Two-stage Stochastic Programs: A Generalized Equation Approach*, manuscript, 2008.

[36] S. M. Robinson, *Stability theory for systems of inequalities. Part I: Linear systems*, SIAM J. Numer. Anal., 12 (1975), pp. 754–769.

[37] S. M. Robinson, *Stability theory for systems of inequalities. Part II: Differentiable Nonlinear systems*, SIAM J. Numer. Anal., 13 (1976), pp. 497–513.

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

[39] S. M. Robinson, *Some continuity properties of polyhedral multifunctions*, Math. Program. Stud., 14 (1981), pp. 206–214.

[40] R. T. Rockafellar, *Convex Analysis*, Princeton University Press, Princeton, 1970.

[41] R. T. Rockafellar and R. J.-B. Wets, *Stochastic convex programming: Kuhn–Tucker conditions*, J. Math. Econom., 2 (1975), pp. 349–370.

[42] R. T. Rockafellar and R. J.-B. Wets, *Variational Analysis*, Springer, Berlin, 1998.

[43] A. Ruszczyński and A. Shapiro, eds., *Stochastic Programming*, Handbooks Oper. Res. Management Sci. 10, North–Holland, Amsterdam, 2003.

[44] A. Shapiro, *Stochastic mathematical programs with equilibrium constraints*, J. Optim. Theory Appl., 128 (2006), pp. 223–243.

[45] A. SHAPIRO AND H. XU, *Stochastic mathematical programs with equilibrium constraints, modeling and sample average approximation*, Optimization, 57 (2008), pp. 395–418.

[46] W. SONG, *Calmness and error bounds for convex constraint systems*, SIAM J. Optim., 17 (2006), pp. 353–371.

[47] A. TOMASGARD, Y. SMEERS, AND K. MIDTHUN, *Capacity booking in a transportation network with stochastic demand*, in Proceedings of the 20th International Symposium on Mathematical Programming, Chicago, GAMS, 2009.

[48] A. S. WERNER, *Bilevel Stochastic Programming Problems: Analysis and Application to Telecommunications*, Ph.D. dissertation, Norwegian University of Science and Technology, Torndheim, Norway, 2004.

[49] A. S. WERNER AND Q. WANG, *Resale in vertically separated markets: Profit and consumer surplus implications*, in Proceedings of the 20th International Symposium on Mathematical Programming, Chicago, GAMS, 2009.

[50] Z. WU AND J. J. YE, *First- and second-order conditions for error bounds*, SIAM J. Optim., 14 (2004), pp. 621–645.

[51] H. XU, *An implicit programming approach for a class of stochastic mathematical programs with complementarity constraints*, SIAM J. Optim., 16 (2006), pp. 670–696.

[52] H. XU AND F. MENG, *Convergence analysis of sample average approximation methods for a class of stochastic mathematical programs with equality constraints*, Math. Oper. Res., 32 (2007), pp. 648–668.

[53] J. J. YE, *Optimality conditions for optimization problems with complementarity constraints*, SIAM J. Optim., 9 (1999), pp. 374–387.

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

[55] J. J. YE, *Nondifferentiable multiplier rules for optimization and bilevel optimization problems*, SIAM J. Optim., 15 (2004), pp. 252–274.

[56] J. J. YE, *Necessary and sufficient optimality conditions for mathematical programs with equilibrium constraints*, J. Math. Anal. Appl., 307 (2005), pp. 305–369.

[57] J. J. YE AND X. Y. YE, *Necessary optimality conditions for optimization problems with variational inequality constraints*, Math. Oper. Res., 22 (1997), pp. 977–997.

[58] J. J. YE AND D. L. ZHU, *Optimality conditions for bilevel programming problems*, Optimization, 33 (1995), pp. 9–27.

[59] J. J. YE, D. L. ZHU, AND Q. J. ZHU, *Exact penalization and necessary optimality conditions for generalized bilevel programming problems*, SIAM J. Optim., 7 (1997), pp. 481–507.

[60] D. ZHANG, H. XU, AND Y. WU, *A two stage stochastic equilibrium model for electricity markets with two way contracts*, Math. Methods Oper. Res., to appear.