Stochastic mathematical programs with equilibrium constraints

Michael Patriksson* and Laura Wynter†

June 16, 1999

Abstract

We introduce Stochastic Mathematical Programs with Equilibrium Constraints (SMPEC), which generalize MPEC models by explicitly incorporating possible uncertainties in the problem data to obtain robust solutions to hierarchical problems. For this problem, we establish results on the existence of solutions, and on the convexity and directional differentiability of the implicit upper-level objective function, both for continuously and discretely distributed probability distributions. In so doing, we establish links between SMPEC models and two-stage stochastic programs with recourse. We also discuss basic parallel iterative algorithms for discretely distributed SMPEC problems.

Key words: Bilevel programming, variational inequality problems, stochastic programming, existence of solutions

1 Introduction

The present paper serves to introduce a framework for hierarchical decision-making under uncertainty. Hierarchical decision-making problems are encountered in a wide variety of domains in the engineering and experimental natural sciences, and in regional planning, management, and economics. These problems are all defined by the presence of two or more objectives with a prescribed order of priority or information. We consider in this paper a sub-class of these problems having two levels, or objectives. We refer to the upper level as the objective having the highest priority and/or information level; it is defined in terms of an optimization with respect to one set of variables. The lower-level problem, which

*Department of Mathematics, Chalmers University of Technology, Gothenburg, Sweden, email: mipat@math.chalmers.se

†PRISM, Computer Science Department, Université de Versailles, Versailles, France, email: Laura.Wynter@prism.uvsq.fr

in the most general case is described by a variational inequality, is then a supplementary problem parameterized by the upper-level variables. These models are known as generalized bilevel programming problems, or mathematical programs with equilibrium constraints (MPEC); see, for example, Luo et al. [13].

In applications relating to Stackelberg game theory, economics, and decision analysis, a number of the problem inputs will often be subject to uncertainty. This is true in particular with respect to costs, demands, and system capacities, which are subject to fluctuations and/or are difficult to measure. In hierarchical models of engineering design and physical phenomena, external conditions and measurement or manufacturing errors introduce uncertainty into the problems. In both of these cases, the uncertainty can be included explicitly by generalizing some of the problem parameters to random variables. However, this generalization complexifies the model significantly; resolution strategies will in many cases require some approximation methods to solve the resulting stochastic programs.

In the simplest case, the expected values of the random variables could be substituted for their distributions and a deterministic model then solved. However, in a nonlinear problem subject to constraints, the effect of this simplification can be quite costly. Indeed, not only will the optimal cost of the expected value solution not necessarily represent the average of the possible optimal costs, but the solution may not even be feasible with respect to the realized values of the random variables. In the case of bilevel pricing and other applications of the Stackelberg game model (consider, for example, the network design problem of Marcotte [14]), user demands are often known only probabilistically. A robust model should then take into account explicitly the range of possible demand values in order to minimize the costs of insufficient capacity when the true demands are realized. Similarly, in topology optimization in structural mechanics (see, for example, Marti [16], and Christiansen et al. [3]), uncertainty enters naturally through random properties of the materials used and possible obstacles, and through conditions that will affect the structure once it has been built, such as varying weather conditions and external forces (or, loads).

To take into account explicitly the variability of the random inputs, as well as the possible infeasibility, we consider a stochastic programming extension of the mathematical programming problem with equilibrium constraints.

We thus define the following stochastic MPEC:

$$[\text{SMPEC-}\Omega] \qquad \begin{aligned} &\text{minimize} \quad E_\omega\left[f(x, y(\omega))\right] := \int_\Omega f(x, y(\omega))\, dF(\omega), && \text{(1a)} \\ &\text{subject to} \quad x \in \mathcal{X}, && \text{(1b)} \end{aligned}$$

where, for every $\omega \in \Omega \subset \Re^l$, $y(\omega) \in \mathcal{S}(x, \omega) := \{\, y \in \Re^m \mid -T(x, \omega, y) \in N_{\mathcal{Y}(x,\omega)}(y) \,\}$ denotes the set of solutions to the lower-level variational inequality problem, parameterized by the upper-level variable, x. The random variable ω is defined on a probability space $(\Omega, \mathcal{A}, P)$. Further, E_ω denotes the expected value with respect to $\omega \in \Omega$, $f : \Re^{n+m} \mapsto \Re \cup \{+\infty\}$ denotes the upper-level objective function, F denotes the cumulative distribution function, assumed to be continuously differentiable, with $\int_\Omega dF(\omega) = 1$, by construction, and $\mathcal{X} \subseteq \Re^n$ denotes the upper-level feasible set.

The lower-level problem is defined by the parameterized mapping $T(x, \omega, \cdot)$ and feasible set $\mathcal{Y}(x, \omega)$ [presumed convex], and $N_{\mathcal{Y}(x,\omega)}(y)$ denotes the normal cone to the set $\mathcal{Y}(x, \omega)$ at $y \in \mathcal{Y}(x, \omega)$. Note that the lower-level problem may admit multiple solutions or no solution for certain values of x and ω; in the former case, we assume that the value of y chosen is that in the set $\mathcal{S}(x, \omega)$ which yields the lowest value of the integral in (1a). In this way, we may define an objective function for [SMPEC-Ω] which depends on x only.

In practice, it is often impossible to evaluate exactly the expected value in (1a) for problems of realistic size due to the difficulty in calculating the multiple integrals. Therefore, in many applications, it is convenient to make the assumption/approximation that the random variables $\omega \in \Omega$ are discrete or that the joint distribution function can be adequately discretized. In this case the expected value reduces to a sum over a discrete set $\mathcal{L}$ of random events, which we number by $\ell = 1, \ldots, |\mathcal{L}|$. The probability of each such event is then given by $\rho_\ell > 0$, with $\sum_{\ell \in \mathcal{L}} \rho_\ell = 1$, by construction. We can then reformulate (1) as the following discretely distributed stochastic MPEC:

[SMPEC-$\mathcal{L}$]
$$\text{minimize} \quad E_\ell\left[f(x, y_\ell)\right] := \sum_{\ell \in \mathcal{L}} \rho_\ell f(x, y_\ell), \tag{2a}$$
$$\text{subject to} \quad x \in \mathcal{X}, \tag{2b}$$

where, for each $\ell \in \mathcal{L}$, $y_\ell \in \mathcal{S}_\ell(x) := \{\, y \in \Re^m \mid -T_\ell(x, y) \in N_{\mathcal{Y}_\ell(x)}(y) \,\}$.

We note that the deterministic problem, which we will refer to as [MPEC], is obtained when $|\mathcal{L}| = 1$. (This observation will also aid in determining the properties of [SMPEC-$\mathcal{L}$].)

A special case of [MPEC] is bilevel programming, which is obtained when the lower-level variational inequality problem reduces to the optimality conditions for an optimization problem, that is, when, for all $x \in \mathcal{X}$, $T(x, \cdot) = \nabla_y t(x, \cdot)$ for some function $t : \mathcal{X} \times \Re^m \mapsto \Re$. Usually, bilevel programming is formulated in terms of the corresponding optimization problem, thus leading to the formulation

[BP]
$$\text{minimize} \quad f(x, y), \tag{3a}$$
$$\text{subject to} \quad x \in \mathcal{X}, \tag{3b}$$

where $y \in \arg\min_{z \in \mathcal{Y}(x)} t(x, z)$.

We will consider as well a special case of [BP], in which the upper-level objective function f depends on the lower-level solution only in the sense of its optimal value. This problem has been analyzed in Shimizu et al. [21], and is defined as follows:

[BPOV]
$$\text{minimize} \quad f(x, p(x)), \tag{4a}$$
$$\text{subject to} \quad x \in \mathcal{X}, \tag{4b}$$

where

$$p(x) := \inf_{y \in \mathcal{Y}(x)} t(x, y). \tag{4c}$$

The linear case of [BPOV] is denoted [BLPOV].

The formulations, as well as the notation, of the continuously and discretely distributed stochastic extensions of these special cases, [SBP], [SBPOV], and [SBLPOV], are the obvious.

We also introduce the classical canonical model of two-stage stochastic linear programming with recourse (e.g., Birge and Louveaux [2]):

[2S-SLP]

$$\text{minimize} \quad c^T x + Q(x), \tag{5a}$$

$$\text{subject to} \quad x \in \mathcal{X}, \tag{5b}$$

where $Q(x) := E_\omega[Q(x, \omega)]$, and $Q(x, \omega) := \inf_y \{\, q^T y \mid Wy = \omega - Ax; \quad y \geq 0 \,\}$, $\omega \in \Omega$, denotes the optimal value of the second-stage linear program.

Similarly, a non-linear form of the two-stage stochastic program with recourse is given as follows:

[2S-SNLP]

$$\text{minimize} \quad E_\omega[f(x, Q(x, \omega))], \tag{6a}$$

$$\text{subject to} \quad x \in \mathcal{X}, \tag{6b}$$

where $Q(x, \omega) := \inf_{y \in \mathcal{Y}(x,\omega)} t(x, y, \omega)$, $\omega \in \Omega$. We note that nonlinear forms of two-stage stochastic programs with recourse have been studied to a large extent only for the quadratic case (e.g., Rockafellar and Wets [20] and Birge et al. [1]).

In the rest of the paper, we shall investigate some relationships among the continuously and discretely distributed stochastic extensions of the models [MPEC], [BP], [BPOV], and [2S-SLP]. In particular, we analyze the links between these problem classes, previously not considered together, and present new results on the existence of solutions, and directional differentiability and convexity properties of the implicit upper-level objective function. A final section then briefly mentions some possibilities for constructing parallel iterative resolution methods for the case of discretely distributed probability distributions. We note that a number of related properties of expectation functionals can be found in Wets [22].

First, let us establish the complexity relations between the problem classes. These are illustrated in Figure 1. (The notation [A] $\subset$ [B] denotes that the class of problems represented by [A] is included in that of [B].)

All the relations described in Figure 1 follow directly from the problem formulations. Most interesting here is the equality, which equates two-stage stochastic programming with recourse to the optimal value form of stochastic bilevel programming problems.

The figure also illustrates that a particular linear form of [SBPOV] reduces to the classic two-stage stochastic linear program with recourse, [2S-SLP], namely a stochastic version of the *right-hand side perturbation* model, as discussed in [21, page 189], in which the upper-level variable is located only on the right-hand side

Figure 1: Relationships among problem classes

$$\begin{array}{ccccccc}
 & & & & [\text{SBLPOV}] & \supset & [\text{2S-SLP}] \\
 & & & & \cap & & \cap \\
[\text{SMPEC}] & \supset & [\text{SBP}] & \supset & [\text{SBPOV}] & = & [\text{2S-SNLP}] \\
\cup & & \cup & & \cup & & \\
[\text{MPEC}] & \supset & [\text{BP}] & \supset & [\text{BPOV}] & &
\end{array}$$

of the lower-level constraints. (We will see later that this simplification enjoys, in many cases, certain convexity properties.)

Since [BP] is $\mathcal{NP}$-hard, even in the linear case (e.g., Hansen et al. [8]), it is clear that the problem [SMPEC] introduced in this paper is $\mathcal{NP}$-hard, as are the other problems in Figure 1, with the exception of the discrete version of the problem [2S-SLP], which is equivalent to a linear program, and is thus in $\mathcal{P}$.

2 Existence of solutions

Consider the following general hierarchical programming model, which encompasses all those presented in Section 1.

[HP]

$$\text{minimize} \quad f(x, v), \tag{7a}$$

$$\text{subject to} \quad x \in \mathcal{X}, \tag{7b}$$

where $v \in \mathcal{S}(x)$, and $\mathcal{S} : \Re^n \mapsto 2^{\Re^s}$ denotes a point–to–set mapping on $\Re^n$. We let $\operatorname{gr} \mathcal{S} := \{ (x, v) \in \Re^{n+s} \mid v \in \mathcal{S}(x) \}$ denote the graph of the mapping $\mathcal{S}$.

The main result of this section is a general existence result for [HP], which in the forthcoming will be specialized to allow existence to be established for all of the problem classes introduced in Section 1.

First, we show that [HP] encompasses the problems of Figure 1.

Proposition 2.1 *The problems* [SMPEC-$\mathcal{L}$] *and* [SMPEC-Ω] *reduce to* [HP].

Proof: In the former case, let $v = (v_\ell)_{\ell \in \mathcal{L}}$, where $v_\ell := y_\ell$, $\mathcal{S} := \prod_{\ell \in \mathcal{L}} \mathcal{S}_\ell$, and $f(x, v) := \sum_{\ell \in \mathcal{L}} \rho_\ell f(x, y_\ell)$. In the latter case, let $v := y(\omega)$, and $f(x, v) := E_\omega[f(x, y(\omega))]$. $\square$

Proposition 2.2 *The problem* [BPOV] *reduces to* [HP].

Proof: Let $v := p(x)$ and $\mathcal{S}(x) := \{p(x)\}$. $\square$

Consider the following assumptions, which will be necessary for demonstrating the existence of optimal solutions to the problem classes of interest:

(i) The upper-level constraint set, $\mathcal{X} \subseteq \Re^n$, is nonempty and closed.

(ii) The lower-level constraint set, $\mathcal{Y}(x)$, is nonempty for at least one $x \in \mathcal{X}$, and bounded for all $x \in \mathcal{X}$.

(iii) (Slater CQ) The lower-level constraint set, $\mathcal{Y}(x)$, is of the form $\mathcal{Y}(x) := \{\, y \in \Re^m \mid g_i(x,y) \leq 0, \quad i = 1,\dots,k \,\}$, where each function $g_i : \Re^{n+m} \mapsto \Re$ is continuous on $\Re^{n+m}$ and convex in y for each $x \in \mathcal{X}$. Further, either $g_i(x,\cdot) = g_i(\cdot)$, $i = 1,\dots,k$, that is, $\mathcal{Y}(x) = \mathcal{Y}$, or for each $x \in \mathcal{X}$ there is a $y \in \Re^m$ such that $g_i(x,y) < 0$, $i = 1,\dots,k$.

(iv) There exists an $(x,v) \in \mathcal{Z} := \{\, (x,v) \in \operatorname{gr} \mathcal{S} \mid x \in \mathcal{X} \,\}$ with $f(x,v) < \infty$.

(v) (Inf-compactness) f is lower semi-continuous (l.s.c.), proper, and has bounded lower level sets on $\mathcal{Z}$.

Theorem 2.3 (Existence of optimal solutions to [HP]). *Let the assumptions* (i), (iv), *and* (v) *be satisfied. Further, assume that* $\operatorname{gr} \mathcal{S}$ *is closed. Then,* [HP] *has at least one optimal solution.*

Proof: The assumptions imply the inf-compactness of the extended function $f + \delta_{\mathcal{Z}}$, where $\delta_{\mathcal{Z}}$ is the indicator function for the closed set $\mathcal{Z}$. The result then follows from Weierstrass' Theorem. □

A similar result is found in Zhang [23].

Next, we specialize this result to the problems defined in Section 1, and give sufficient conditions for the hypotheses to be satisfied.

Note that the conditions required in the preceding existence result are weaker than those of many previously considered requirements on bilevel model formulations, and as such, may be particularly interesting for a number of important applications.

For example, the following two corollaries allow one to establish the existence of an optimal solution to stochastic (and indeed also deterministic) structural optimization problems in the presence of zero lower bounds on the truss bar volumes. (See [3] for further details.)

In the following corollaries, the assumptions above are expressed using the notation of Propositions 2.1 and 2.2, for each problem class.

Corollary 2.4 (Existence of optimal solutions to [SMPEC-$\mathcal{L}$]). *Let the assumption* (iii) *hold for each* $\ell \in \mathcal{L}$*, and the mapping* $T_\ell : \Re^{n+m} \mapsto \Re^m$ *be continuous on* $\mathcal{X} \times \Re^m$*. Then,* $\operatorname{gr} \mathcal{S}_\ell$ *is closed for each* $\ell \in \mathcal{L}$*. Suppose further that the assumption* (i) *holds, that the assumption* (iv) *holds for some* x *over all the sets* $\mathcal{Z}_\ell$*,* $\ell \in \mathcal{L}$*, and that the assumption* (v) *holds with* $\mathcal{Z}$ *replaced by* $\cup_{\ell \in \mathcal{L}} \mathcal{Z}_\ell$*. Then, there exists at least one optimal solution to* [SMPEC-$\mathcal{L}$].

Proof: Each set $\mathcal{S}_\ell(x)$ can be equivalently described by

$$\mathcal{S}_\ell(x) = \{ y \in \mathcal{Y}_\ell(x) \mid T_\ell(x,y)^T(z-y) \geq 0, \quad \forall z \in \mathcal{Y}_\ell(x) \}, \qquad x \in \mathcal{X}.$$

By the assumption (iii), either $\mathcal{Y}_\ell(x) = \mathcal{Y}_\ell$ holds, in which case the closedness of $\operatorname{gr} \mathcal{S}_\ell$ follows from the continuity of the mappings T_ℓ and g_ℓ, or the Slater condition holds, in which case the closedness of $\mathcal{S}_\ell$ follows from Hogan [11, Lemma 1]. □

COROLLARY 2.5 (Existence of optimal solutions to [SMPEC-Ω]). *Let the assumption* (iii) *hold for each $\omega \in \Omega$. Further, let $T : \Re^{n+l+m} \mapsto \Re^m$ and $g : \Re^{n+l+m} \mapsto \Re^k$ be continuous on $\mathcal{X} \times \Omega \times \Re^m$. Then, $\operatorname{gr} \mathcal{S}$ is closed. Suppose further that the assumption* (i) *holds, that the assumption* (iv) *holds for some x over all the sets $\mathcal{Z}(\omega)$, $\omega \in \Omega$, almost surely, and that the assumption* (v) *holds with $\mathcal{Z}$ replaced by $\cup_{\omega\in\Omega}\mathcal{Z}(\omega)$. Then, there exists at least one optimal solution to* [SMPEC-Ω].

Proof: The continuity assumptions on T and g, together with the assumption (iii) for each $\omega \in \Omega$, imply that the mapping $(x,\omega) \mapsto \mathcal{S}(x,\omega)$ is closed on $\mathcal{X} \times \Omega$, and hence also the set $\operatorname{gr} \mathcal{S}$ is closed, following the proof of Corollary 2.4. It follows from the lower semi-continuity of f that the integrand in (1a) is l.s.c. in ω (and hence integrable), and that, by the linearity of the integration operation and the fact that f is proper, the integral in (1a) also is l.s.c. Since there exists an $x \in \mathcal{X}$ such that this integral has a finite value, by the assumption (v), it is inf-compact. The desired result then follows as in Theorem 2.3. □

COROLLARY 2.6 (Existence of optimal solutions to [SBP-$\mathcal{L}$]). *Let the assumption* (iii) *hold for each $\ell \in \mathcal{L}$, and the function $t_\ell : \Re^{n+m} \mapsto \Re$ be continuous on $\mathcal{X} \times \Re^m$. Then, $\operatorname{gr} \mathcal{S}_\ell$ is closed for each $\ell \in \mathcal{L}$. Suppose further that the assumption* (i) *holds, that the assumption* (iv) *holds for some x over all the sets $\mathcal{Z}_\ell$, $\ell \in \mathcal{L}$, and that the assumption* (v) *holds with $\mathcal{Z}$ replaced by $\cup_{\ell\in\mathcal{L}}\mathcal{Z}_\ell$. Then, there exists at least one optimal solution to* [SBP-$\mathcal{L}$].

Proof: The proof is similar to that of Corollary 2.4, the closedness result following from [11, Lemma 1]. □

COROLLARY 2.7 (Existence of optimal solutions to [SBP-Ω]). *Let the assumption* (iii) *hold for each $\omega \in \Omega$. Further, let $t : \Re^{n+l+m} \mapsto \Re$ and $g : \Re^{n+l+m} \mapsto \Re^k$ be continuous on $\mathcal{X} \times \Omega \times \Re^m$. Then, $\operatorname{gr} \mathcal{S}$ is closed. Suppose further that the assumption* (i) *holds, the assumption* (iv) *holds for some x over all the sets $\mathcal{Z}(\omega)$, $\omega \in \Omega$, almost surely, and that the assumption* (v) *holds with $\mathcal{Z}$ replaced by $\cup_{\omega\in\Omega}\mathcal{Z}(\omega)$. Then, there exists at least one optimal solution to* [SBP-Ω].

Proof: The proof follows from that of Corollary 2.5. □

COROLLARY 2.8 (Existence of optimal solutions to [SBPOV-$\mathcal{L}$]). *Under the assumptions of Corollary* 2.6*, there exists at least one optimal solution to* [SBPOV-$\mathcal{L}$].

Proof: The result is a special case of Corollary 2.6. ☐

COROLLARY 2.9 (Existence of optimal solutions to [SBPOV-Ω]). *Under the assumptions of Corollary* 2.7*, there exists at least one optimal solution to* [SBPOV-Ω].

Proof: The result is a special case of Corollary 2.7. ☐

The following is an example of a result that can be established for [2S-SLP] by using the links to bilevel programming. (Existence results for two-stage stochastic linear programming based on general linear programming properties can be found in [2, Section 3.1].

COROLLARY 2.10 (Existence of an optimal solution to [2S-SLP]). *Let assumption* (ii) *hold for some* $x \in \mathcal{X}$ *and for each* $\omega \in \Omega$*. Suppose further that assumption* (i) *holds, and that the upper-level feasible set,* $\mathcal{X}$*, is bounded. Then, there exists at least one optimal solution to* [2S-SLP].

Proof: By the linearity of all of the problem functions, (i) and (ii) along with the boundedness of $\mathcal{X}$ imply (iv) and (v). Then, the result follows from that of Corollary 2.7. ☐

3 Convexity

We next establish the convexity of some interesting special cases of [SMPEC].

THEOREM 3.1 (Convexity of [SBPOV-$\mathcal{L}$]). *In addition to the assumptions of Corollary* 2.8*, suppose, for each* $\ell \in \mathcal{L}$*, that* t_ℓ *and* $g_{i,\ell}$*,* $i = 1, \ldots, k$*, are convex on* $\Re^{n+m}$*. Then, each function* $p_\ell : \mathcal{X} \mapsto \Re \cup \{+\infty\}$*, defined by* (4c)*, is convex on* $\mathcal{X}$*. Further, assume that* $\mathcal{X}$ *is convex, and that the function* $f : \mathcal{X} \times \Re \mapsto \Re$ *is convex on* $\mathcal{X} \times \Re$ *and increasing in its second argument. Then, the implicit upper-level objective function* $x \mapsto \sum_{\ell \in \mathcal{L}} \rho_\ell f(x, p_\ell(x))$ *is convex on* $\mathcal{X}$*, so that* [SBPOV-$\mathcal{L}$] *is a convex problem.*

Proof: We need only to establish the convexity of p_ℓ, $\ell \in \mathcal{L}$, on $\mathcal{X}$, but this result follows from the assumptions and Geoffrion [7, Theorem 5]. ☐

THEOREM 3.2 (Convexity of [SBPOV-Ω]). *In addition to the assumptions of Corollary* 2.9*, suppose, for each* $\omega \in \Omega$*, that* $t(\cdot, \omega, \cdot)$ *and* $g_i(\cdot, \omega, \cdot)$*,* $i = 1, \ldots, k$*, are convex on* $\Re^{n+m}$*. Then, each function* $p(\cdot, \omega) : \mathcal{X} \mapsto \Re \cup \{+\infty\}$*, defined by*

(4c), *is convex on* $\mathcal{X}$. *Further, assume that* $\mathcal{X}$ *is convex, and that the function* $f : \mathcal{X} \times \Re \mapsto \Re$ *is convex on* $\mathcal{X} \times \Re$ *and increasing in its second argument. Then, the implicit upper-level objective function* $x \mapsto E_\omega[f(x, p(x, \omega))]$ *is convex on* $\mathcal{X}$, *so that* [SBPOV-Ω] *is a convex problem.*

Proof: The convexity of $p(\cdot, \omega)$, $\omega \in \Omega$, on $\mathcal{X}$ follows as in the proof of Theorem 3.1. The linearity of the integration then yields the desired result. $\square$

It is clear that the convexity of [2S-SLP] is a special case of that of [SBPOV-Ω]. (Compare, for example, with [2, Corollary 3.33]).

4 Differentiability

In this section, we will examine the directional differentiability of the implicit upper-level objective function of [HP], and, as before, specialize that result to the different problem classes. Let us first consider a number of assumptions.

(a) The function f is continuously differentiable on $\mathcal{X} \times \Re^m$.

(b) The lower-level constraint set, $\mathcal{Y}(x)$, is of the form $\mathcal{Y}(x) := \{\, y \in \Re^m \mid g_i(x, y) \leq 0, \quad i = 1, \dots, k \,\}$, where each function $g_i : \Re^{n+m} \mapsto \Re$ is twice continuously differentiable on $\Re^{n+m}$ and convex in y for each $x \in \mathcal{X}$. Further, for each $x \in \mathcal{X}$, $\mathcal{Y}(x) \neq \emptyset$ and $\mathcal{Y}(x) \subset B$, for some open and bounded set $B \in \Re^m$.

(c) (Linear independence CQ) Let $\mathcal{I}(x, y) := \{\, i = 1, \dots, k \mid g_i(x, y) = 0 \,\}$. Then, for each $x \in \mathcal{X}$ and $y \in \mathcal{S}(x)$, the partial gradients $\nabla_y g_i(x, y)$, $i \in \mathcal{I}(x, y)$, are linearly independent.

(d) The mapping T is continuously differentiable on $\mathcal{X} \times \Re^m$ and strongly monotone in y for each $x \in \mathcal{X}$.

(d′) The mapping T is continuously differentiable on $\mathcal{X} \times \Re^m$ and monotone in y for each $x \in \mathcal{X}$.

The following Theorem, presented without proof, follows directly from Robinson [18, 19].

THEOREM 4.1 (Directional differentiability for [HP]). *Let the assumption* (a) *be satisfied and the implicit mapping* $x \mapsto \mathcal{S}(x)$ *be locally Lipschitz continuous on* $\mathcal{X}$. *Then, the implicit upper-level objective function* $x \mapsto f(x, \mathcal{S}(x))$ *of* [HP] *is locally Lipschitz continuous and directionally differentiable on* $\mathcal{X}$.

Next, we provide sufficient conditions for the directional differentiability of the implicit upper-level objective functions for the problem classes presented in Section 1.

Corollary 4.2 (Directional differentiability for [SMPEC-$\mathcal{L}$]). *Let the assumption* (a) *hold, and the assumptions* (b)–(d) *be satisfied for each* $\ell \in \mathcal{L}$. *Then, the implicit upper-level objective function* $x \mapsto \sum_{\ell \in \mathcal{L}} \rho_\ell f(x, y_\ell(x))$ *of* [SMPEC-$\mathcal{L}$] *is locally Lipschitz continuous and directionally differentiable on* $\mathcal{X}$.

Proof: By Robinson [18, Theorem 2.1], the assumptions (a)–(d) imply that for each $\ell \in \mathcal{L}$, the implicit mapping $x \mapsto \mathcal{S}_\ell(x)$ is locally Lipschitz continuous. Then, by Theorem 4.1, the result follows. □

Corollary 4.3 (Directional differentiability for [SMPEC-Ω]). *Let the assumptions of Corollary* 2.5 *hold. Further, let the assumption* (a) *hold, and the assumptions* (b)–(d) *be satisfied, almost surely. Then, the implicit upper-level objective function* $x \mapsto E_\omega[f(x, y(x, \omega))]$ *of* [SMPEC-Ω] *is locally Lipschitz continuous and directionally differentiable on* $\mathcal{X}$.

Proof: Note that under the assumptions of Corollary 2.5, the integral in (1a) exists. As in Corollary 4.2, the implicit function $x \mapsto \mathcal{S}(x, \omega)$ is locally Lipschitz continuous. The integration preserves the locally Lipschitz continuity, since the Lipschitz constant of f remains bounded on any bounded subset of $\mathcal{X}$, and is therefore integrable. The result then follows by Theorem 4.1. □

The following result, adapted from Fiacco [6], illustrates that the requirements needed to ensure the directional differentiability of the implicit upper-level objective function of [SBPOV] are weaker than those of [SMPEC].

Corollary 4.4 (Directional differentiability for [SBPOV-$\mathcal{L}$]). *Let the assumption* (a) *hold, and the assumptions* (b), (c), *and* (d$'$) *be satisfied for each* $\ell \in \mathcal{L}$. *Then, the implicit upper-level objective function* $x \mapsto \sum_{\ell \in \mathcal{L}} \rho_\ell f(x, p_\ell(x))$ *of* [SBPOV-$\mathcal{L}$] *is locally Lipschitz continuous and directionally differentiable on* $\mathcal{X}$.

Proof: By the result of [6], under the assumptions (b), (c) and (d$'$), the optimal value function $x \mapsto p_\ell(x)$ is directionally differentiable on $\mathcal{X}$, for all $\ell \in \mathcal{L}$. Then, by Corollary 4.2, with assumption (a), the result follows. □

Corollary 4.5 (Directional differentiability for [SBPOV-Ω]). *Let the assumptions of Corollary* 4.3 *hold, except that the assumption* (d) *is replaced by the weaker assumption* (d$'$). *Then, the implicit upper-level objective function* $x \mapsto E_\omega[f(x, p(x, \omega))]$ *of* [SBPOV-Ω] *is locally Lipschitz continuous and directionally differentiable on* $\mathcal{X}$.

Proof: Follows directly from using Corollary 4.3 and arguments similar to those in the proof of Corollary 4.4. □

5 Algorithms for SMPEC

Deterministic MPEC models are notoriously difficult and time-consuming to solve because of their non-convexity coupled with the nondifferentiability of the implicit upper-level objective function. (Differentiability can only be asserted under the additional condition that strict complementarity holds.) The introduction of a random distribution on some or all of its parameters causes an even greater increase in the problem size and computational complexity. For this reason, the development of efficient methods is primordial, as are the use of parallel strategies, whenever possible.

In this paper, we have presented a taxonomy which highlights the principal features of each problem type, and which therefore facilitates the adaptation of algorithms.

For example, descent algorithms developed for deterministic MPEC can be adapted to SMPEC as long as the local Lipschitz continuity and directional differentiability of the stochastic objective function can be maintained. Corollaries 4.2 and 4.3 present the conditions needed. Similarly, algorithms for solving BPOV can be adapted to SBPOV as long as the conditions of Corollary 4.4 or 4.5 are satisfied.

Further, in the case of [SMPEC-$\mathcal{L}$] or [SBPOV-$\mathcal{L}$], subgradients can be calculated efficiently if one makes use of the decomposability of the problem.

For example, based on the notation of [SMPEC-$\mathcal{L}$], and following the approach of Outrata and Zowe [17] for [MPEC], a subgradient of the implicit upper-level function f at x is given by the formula

$$\xi_f(x) := \sum_{\ell \in \mathcal{L}} \rho_\ell [\nabla_x f(x, y_\ell) + \nabla_x L_\ell(x, y_\ell, \lambda_\ell)^T d_{y_\ell} - \nabla_x g_{\mathcal{J}(x,\ell)}(x, y_\ell)^T d_{\lambda_{\mathcal{J}(x,\ell)}}], \quad (8)$$

where, for each $\ell \in \mathcal{L}$, $L_\ell(x, y_\ell, \lambda_\ell) := T_\ell(x, y_\ell) + \nabla_y g(x, y_\ell)^T \lambda_\ell$ is the Lagrangean of the lower-level problem, with $y_\ell := y_\ell(x)$ for brevity, and λ_ℓ being the (unique, under the conditions of Corollary 4.2) vector of multipliers for the parameterized lower-level constraints. Furthermore, for each $\ell \in \mathcal{L}$, the set $\mathcal{J}(x, \ell)$ is chosen such that $\mathcal{I}(x, \ell) \supseteq \mathcal{J}(x, \ell) \supseteq \mathcal{I}_+(x, \ell)$ holds, where $\mathcal{I}(x, \ell)$ denotes the set of active lower-level constraints at (x, y_ℓ), and $\mathcal{I}_+(x, \ell)$ is the subset for which the multiplier values are positive, and $(d_{y_\ell}, d_{\lambda_{\mathcal{J}(x,\ell)}})$ solves the linear system

$$\begin{bmatrix} \nabla_y L_\ell(x, y_\ell, \lambda_\ell) & -\nabla_y g_{\mathcal{J}(x,\ell)}(x, y_\ell)^T \\ \nabla_y g_{\mathcal{J}(x,\ell)}(x, y_\ell) & 0^{m \times |\mathcal{J}(x,\ell)|} \end{bmatrix} \begin{pmatrix} d_{y_\ell} \\ d_{\lambda_{\mathcal{J}(x,\ell)}} \end{pmatrix} = \begin{pmatrix} -\nabla_y f(x, y_\ell) \\ 0^m \end{pmatrix}. \quad (9)$$

Each component of the subgradient can be calculated independently and in parallel. The subgradient can then be used in an algorithm for the heuristic solution of the problem or be embedded within a more sophisticated algorithm. A simple scheme is to take, at some iteration t,

$$x^{t+1} := P_{\mathcal{X}}[x^t - \gamma_t \xi_f(x^t)],$$

for some $\gamma_t > 0$ determined through an inexact line search. Note that, at points of nondifferentiability, the method may break down because the negative of the subgradient may then not be a descent direction; in order to obtain a well-defined iteration at such points, one can envisage utilizing a step length which is the maximum of the one supplied by a backtracking line search and the result of a predetermined step length formula used in traditional subgradient optimization techniques. Outrata and Zowe [17] utilize the deterministic analog of the systems (8) and (9) in the development of a bundle method for the solution of [MPEC], which immediately extends to the solution of [SMPEC-$\mathcal{L}$].

Consider the following parallel resolution strategy for this model. In some cases, one may identify a cluster of similar scenarios. Allocating these to the same processor, one may solve the corresponding lower-level problems by utilizing efficient reoptimization procedures given that any of them have been solved to optimality, since the optimal solution to any one of them is feasible as well as near-optimal to all the others. Further, for scenarios with slightly differing sets, $\mathcal{J}(x, \ell)$, consider sorting subsets of the scenarios so that $\mathcal{J}(x, \ell_1) \subset \mathcal{J}(x, \ell_2)$, and so on. Then, one may solve the preceding linear systems in sequence, expanding the matrix with the necessary rows and columns and utilizing the solution to the former system as a starting point in the search for the next. The fact that the choice of $\mathcal{J}(x, \ell)$ is arbitrary in the range of active constraints may also assist in the construction of the afore-mentioned clusters (and the subsequent solutions of the linear systems) through a minimization of the number of scenarios with distinct values of $\mathcal{J}(x, \ell)$.

Another algorithmic method that can be transfered to [SMPEC-$\mathcal{L}$] is the class of penalty methods for [MPEC]. Assume that for each $\ell \in \mathcal{L}$ there is a (continuous) function, $\psi_\ell : \Re^{n+m} \mapsto \Re \cup \{+\infty\}$, which satisfies $\psi_\ell(x, y) \geq 0$ for all $(x, y) \in \mathcal{X} \times \Re^m$, and

$$\psi_\ell(x, y) = 0 \qquad \Longleftrightarrow \qquad y \in \mathcal{S}_\ell(x).$$

Then, the inclusion (2b) can be represented by a single equation in y. The penalized problem is then defined for $\eta_\ell > 0$ by

$$\begin{aligned} \text{minimize} \quad & \sum_{\ell \in \mathcal{L}} \rho_\ell [f(x, y_\ell) + \eta_\ell \psi_\ell(x, y_\ell)], \\ \text{subject to} \quad & x \in \mathcal{X}. \end{aligned}$$

In the deterministic case, such approaches have been devised by Harker and Choi [9], and Marcotte and Zhu [15]. See also Larsson and Patriksson [12] for other examples of differentiable, and typically non-convex, merit functions for variational inequality problems and Facchinei and Soares [5] for NCP-functions for the VIP primal–dual optimality conditions.

Again, this penalty approach can be easily parallelized, since the evaluation of the penalty function and its derivatives can be done independently over the different scenarios.

Several other methods for solving BP and MPEC problems have appeared in the literature in recent years (see, e.g., [13, 21] for overviews, and the recent smoothing method of Facchinei et al. [4]). These methods can generally be adapted to solving the discretely-distributed stochastic models presented here.

In the case of large-scale models, and when the number of scenarios is very high, however, simple parallelization strategies over the scenarios may however not make the model tractable. In these cases it may be necessary to develop further approximation schemes, such as approximating the lower-level problem, or constructing aggregation/disaggregation strategies with respect to the scenarios.

In the case of continuously distributed SMPEC problems like [SMPEC-Ω] or [SBPOV-Ω], no such decomposition over scenarios is available, and the presence of the multiple integral in the objective function generally necessitates numerical integration at each upper-level iteration. Birge et al. [1] propose an approximate Newton method as well as a modification of the stochastic decomposition method of Higle and Sen [10] to solve quadratic stochastic programs with recourse. These problems are in fact quadratic instances of SBPOV with a separable objective function and polyhedral constraints. Other strategies for solving [SMPEC-Ω] and [SBPOV-Ω] may make use of random sampling of the continuous distribution, as is done within the class of stochastic quasi-gradient methods.

References

[1] J. R. Birge, X. Chen, L. Qi, and Z. Wei, "A stochastic Newton method for stochastic quadratic programs with recourse," technical report, Department of Industrial and Operations Engineering, University of Michigan, Ann Arbor, MI, 1995.

[2] J. R. Birge, and F. Louveaux, *Stochastic Programming*, Springer-Verlag, New York, NY, 1997.

[3] S. Christiansen, M. Patriksson, and L. Wynter, "Stochastic bilevel programming in structural optimization," technical report, PRISM, Computer Science Department, Université de Versailles, Versailles, France, 1999. Submitted.

[4] F. Facchinei, H. Jiang, and L. Qi, "A smoothing method for mathematical programs with equilibrium constraints," *Mathematical Programming* **85**, 107–134 (1999).

[5] F. Facchinei and J. Soares, "A new merit function for nonlinear complementarity problems and a related algorithm," *SIAM Journal on Optimization* **7**, 225–247 (1997).

[6] A. V. Fiacco, "Optimal value continuity and differential stability bounds under the Mangasarian–Fromovitz constraint qualification," in: *Mathematical Programming with Data Perturbations II*, A. V. Fiacco, ed., Marcel Dekker, New York, NY, 1983, pp. 65–90.

[7] A. M. Geoffrion, "Elements of large-scale mathematical programming," *Management Science* **16**, 652–691 (1970).

[8] P. Hansen, B. Jaumard, and G. Savard, "New branch-and-bound rules for linear bilevel programming," *SIAM Journal on Scientific and Statistical Computing* **13**, 1194–1217 (1992).

[9] P. T. Harker and S.-C. Choi, "A penalty function approach for mathematical programs with variational inequality constraints," *Information and Decision Technologies* **17**, 41–50 (1991).

[10] J. L. Higle and S. Sen, "Stochastic decomposition: An algorithm for two stage linear programs with recourse," *Mathematics of Operations Research* **16**, 650–669 (1991).

[11] W. W. Hogan, "Directional derivatives for extremal-value functions with applications to the completely convex case," *Operations Research* **21**, 188–209 (1973).

[12] T. Larsson and M. Patriksson, "A class of gap functions for variational inequalities," *Mathematical Programming* **64**, 53–79 (1994).

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

[14] P. Marcotte, "Network design problem with congestion effects: A case of bilevel programming," *Mathematical Programming* **34**, 142–162 (1986).

[15] P. Marcotte and D. Zhu, "Exact and inexact penalty methods for the generalized bilevel programming problem," *Mathematical Programming* **74**, 141–157 (1996).

[16] K. Marti, ed., "Structural Reliability and Stochastic Structural Optimization," special issue of *Mathematical Methods of Operations Research* **46**, no. 3 (1997).

[17] J. Outrata, and J. Zowe, "A numerical approach to optimization problems with variational inequality constraints," *Mathematical Programming* **68**, 105–130 (1995).

[18] S. M. Robinson, "Strongly regular generalized equations," *Mathematics of Operations Research* **5**, 43–62 (1980).

[19] S. M. Robinson, "An implicit-function theorem for a class of nonsmooth functions," *Mathematics of Operations Research* **16**, 282–309 (1991).

[20] R. T. Rockafellar, and R. J.-B. Wets, "A Lagrangian finite generation technique for solving linear–quadratic problems in stochastic programming," *Mathematical Programming Study* **28**, 63–93 (1986).

[21] K. Shimizu, Y. Ishizuka, and J. Bard, *Nondifferentiable and Two-Level Mathematical Programming*, Kluwer Academic Publishers, Boston, MA, 1997.

[22] R. J.-B. Wets, "Stochastic programming", Chapter VIII in *Optimization*, volume 1 in *Handbooks in Operations Research and Management Science*, G. L. Nemhauser, A. H. G. Rinnooy Kan and M. J. Todd, eds., North-Holland, Amsterdam, 1989, pp. 573–629.

[23] R. Zhang, "Problems of hierarchical optimization in finite dimension," *SIAM Journal on Optimization* **4**, 521–536 (1994).