Simulation-Based Solution of Stochastic Mathematical Programs with Complementarity Constraints

Ilker Birbil, S.; Gürkan, G.; Listes, O.L.

Publication date:
2004

Citation for published version (APA):
Ilker Birbil, S., Gürkan, G., & Listes, O. L. (2004). *Simulation-Based Solution of Stochastic Mathematical Programs with Complementarity Constraints: Sample-Path Analysis*. (CentER Discussion Paper; Vol. 2004-25). Operations research.

CentER

Discussion Paper

No. 2004–25

SIMULATION-BASED SOLUTION OF STOCHASTIC MATHEMATICAL PROGRAMS WITH COMPLEMENTARITY CONSTRAINTS: SAMPLE-PATH ANALYSIS

By Ş. İlker Birbil, G. Gürkan, O. Listeş

March 2004

ISSN 0924-7815

TILBURG UNIVERSITY

Simulation-based solution of stochastic mathematical programs with complementarity constraints: Sample-path analysis

Ş. İlker Birbil[†] Gül Gürkan[*] Ovidiu Listeş[*]

January 2004

Abstract

We consider a class of stochastic mathematical programs with complementarity constraints, in which both the objective and the constraints involve limit functions or expectations that need to be estimated or approximated. Such programs can be used for modeling "average" or steady-state behavior of complex stochastic systems. Recently, simulation-based methods have been successfully used for solving challenging stochastic optimization problems and equilibrium models. Here we broaden the applicability of so-called the sample-path method to include the solution of certain stochastic mathematical programs with equilibrium constraints. The convergence analysis of sample-path methods rely heavily on stability conditions. We first review necessary sensitivity results, then describe the method, and provide sufficient conditions for its almost-sure convergence. Alongside we provide a complementary sensitivity result for the corresponding deterministic problems. In addition, we also provide a unifying discussion on alternative set of sufficient conditions, derive a complementary result regarding the analysis of stochastic variational inequalities, and prove the equivalence of two different regularity conditions.

Keywords: stochastic mathematical programs with complementarity constraints, sample-path methods, simulation, stability, regularity conditions, mathematical programs with equilibrium constraints

JEL classification codes: C61, C62, C63

1 Introduction

In the last decade, many researchers have studied solution methods for mathematical programs with complementarity constraints (MPCC's). Primarily, these programs arise in reformulations of important optimization problems such as mathematical programs with equilibrium constraints (MPEC's) (Luo *et al.* (1996)), generalized semi-infinite programs

[†]Erasmus Research Institute of Management, Erasmus University Rotterdam, P.O. Box 1738, 3000 DR Rotterdam, The Netherlands.

[*]CentER for Economic Research, Tilburg University, P.O. Box 90153, 5000 LE Tilburg, The Netherlands. The research reported here was sponsored by the Netherlands Organization for Scientific Research (NWO), grant 016.005.005.

(Still (1999)), or bilevel optimization problems (Bard (1998) and Dempe (2002)). Naturally, the crucial role of the MPCC's also attracted the attention of the practitioners, since these mathematical programs are important modeling tools for numerous applications. Among such applications we mention here Stackelberg games of mathematical economics, congestion problems in transportation networks, or optimal design of mechanical systems; see Luo *et al.* (1996), Ferris and Pang (1997), Bard (1998), and Dempe (2002) for other applications.

When pursuing a realistic modeling approach, decision makers are usually faced with uncertainty. For instance, costs in a transportation network or measurements in a mechanical system are liable to intrinsic uncertainty by the nature of the problem. In certain cases, the uncertainty may be cast explicitly by means of random model parameters, an approach which leads to a stochastic mathematical programming formulation. In the present paper we focus on a particular formulation of stochastic mathematical programs that involve complementarity constraints. Alongside a detailed analysis of this problem, we also discuss a solution approach by a simulation-based method, so-called sample-path optimization.

Roughly speaking, sample-path methods are concerned with solving a problem of optimization and/or equilibrium, involving a limit function f_∞ which is not observable. However, one can observe functions f_n that almost surely converge pointwise to f_∞ as $n \to \infty$. In the kind of applications we have in mind, f_∞ could be a steady-state performance measure of a dynamic system or an expected value in a static system, and we use simulation to observe the f_n's. In systems that evolve over time, we simulate the operation of the system for, say, n time units and then compute an appropriate performance measure. In static systems we repeatedly observe instances of the system and compute an average. In both cases, to observe f_n at different parameter settings we use the method of common random numbers. Furthermore, in many cases derivatives or directional derivatives of the f_n can be obtained using well-established methods of gradient estimation such as infinitesimal perturbation analysis (IPA); see Ho and Cao (1991) and Glasserman (1991). The key point is the following: once we fix n and a sample point (using common random numbers), f_n becomes a deterministic function. Sample-path methods then solve the resulting deterministic problem (using f_n with the fixed sample path selected), and take the solution as an estimate of the true solution. Clearly, the availability of very powerful deterministic solvers (both for optimization and for equilibrium problems) makes this approach very attractive.

In general the results which provide theoretical support for sample-path methods are based on the sensitivity analysis of the corresponding deterministic problems. This aspect is illustrated in Robinson (1996) in the case of simulation optimization problems with deterministic constraints as well as in Gürkan *et al.* (1999a) in the case of stochastic variational inequalities. Following a similar argument, we build our sample-path analysis of stochastic MPCC's on the recent work of Scheel and Scholtes (2000), who set forth important sensitivity results for deterministic MPCC's.

In our view, the current paper makes the following contributions:

- It proposes a modeling structure for a class of stochastic MPCC's in which the expectations or limit functions must be approximated/estimated, for example by using simulation.

- It presents sufficient conditions for the convergence of the sample-path method applied to the class of stochastic MPCC's addressed.

- It provides complementary sensitivity results for deterministic MPCC's to those presented by Scheel and Scholtes (2000), in terms of weaker sufficient conditions for stability (by interpreting the approximating functions as non-parametric perturbations of the true functions).

- It provides a discussion on the assumptions involved in the stability analysis of both variational inequalities and MPCC's. Along the way, a complementary result to the main theorem of Gürkan *et al.* (1999a) dealing with stochastic variational inequalities is provided.

The remainder of this paper is organized as follows. In Section 2, we specify the problem under consideration and provide some related background material. In Section 3, we first restate some essential concepts, then use them in analyzing the sample-path solution of stochastic MPCC's. In Section 4, we provide a discussion on main results involved in our analysis and different assumptions required. Finally, in Section 5 we give a summary and conclude the paper. Our analysis uses some regularity conditions for generalized equations and nonlinear programs. Since these conditions are rather technical, we deal with them in Appendix 1. In addition, we often refer to the main theorem in Gürkan *et al.* (1999a); for convenience this theorem is restated in Appendix 2.

2 Background material and sample-path methods

In this section we review some background material related to MPCC's and sample-path methods. We work with an open set $\Theta \subseteq \mathbb{R}^{n_0}$ and twice differentiable functions $f : \Theta \to \mathbb{R}$, $g : \Theta \to \mathbb{R}^p$, $h : \Theta \to \mathbb{R}^q$, and $F : \Theta \to \mathbb{R}^{m \times l}$ with $m \geq 1$, $l \geq 2$, and

$$F(z) = \begin{bmatrix} F_{11}(z) & \dots & F_{1l}(z) \\ \vdots & \ddots & \vdots \\ F_{m1}(z) & \dots & F_{ml}(z) \end{bmatrix}.$$

Given these ingredients, the problem under consideration is the following mathematical program with complementarity constraints:

$$\begin{array}{lll} \text{MPCC} & \min & f(z) \\ & \text{s.t.} & \min\{F_{k1}(z), \dots, F_{kl}(z)\} = 0 \qquad k = 1, \dots, m \\ & & g(z) \leq 0 \\ & & h(z) = 0 \\ & & z \in \Theta\,, \end{array}$$

as discussed in Scheel and Scholtes (2000). If $z = (x, y) \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$, $l = 2$, $F_{k1}(x, y) = y_k$, and $G_k(x, y) := F_{k2}(x, y)$, then the constraints $\min\{F_{k1}(z), F_{k2}(z)\} = 0$, $k = 1, ..., m$, represent the parametric nonlinear complementarity problem

$$y \geq 0, \qquad G(x, y) \geq 0, \qquad y\, G(x, y) = 0, \tag{1}$$

with parameter x and variable y. It is well known (see e.g., Harker and Pang (1990)) that problem (1) is equivalent to solving a parametric variational inequality over the positive orthant $\mathbb{R}^{n_2}_+$. Given a closed convex set $C \subseteq \mathbb{R}^{n_2}$ and a function G from $\mathbb{R}^{n_2}$ to itself, the variational inequality problem, denoted by $\mathrm{VI}(G, C)$, is to find a point $y_0 \in C$ such that

$$G(y_0)(y - y_0) \geq 0 \qquad \text{for all } \ y \in C. \tag{2}$$

For any given parameter x, the problem (1) is then equivalent to $\mathrm{VI}(G(x, \cdot), \mathbb{R}^{n_2}_+)$. Moreover, problem (1) also includes, as a special case, the Karush-Kuhn-Tucker conditions of parametric variational inequalities defined over sets described by smooth systems of inequalities (see e.g., Luo *et al.* (1996)). Therefore, the MPCC's are indeed a very important subclass of the mathematical programs with equilibrium constraints (MPEC's), in which the essential constraints are explicitly formulated as parametric variational inequalities of the form "y solves $\mathrm{VI}(G(x, \cdot), \mathbb{R}^{n_2}_+)$". In MPEC's, x are usually referred to as the *upper level* variables, while y denote the variables at the *lower level* (or the inner problem). For details on the general MPEC problem as well as examples of interesting applications, we refer to the monographs by Luo *et al.* (1996) and Outrata *et al.* (1998).

In this paper we address MPCC's under uncertainty. More specifically, we are interested in a certain class of stochastic MPCC's, in which potentially, all of the defining functions f, F, g, or h (or some of their components) may represent a limit function or an expectation and therefore, can not be directly observed. Formally, we are interested in the following problem:

$$\begin{array}{lll} \text{SMPCC} & \min & f_\infty(z) \\ & \text{s.t.} & \min\{F^\infty_{k1}(z), \ldots, F^\infty_{kl}(z)\} = 0 \qquad k = 1, \ldots, m \\ & & g_\infty(z) \leq 0 \\ & & h_\infty(z) = 0 \\ & & z \in \Theta\,, \end{array}$$

where any of f_∞, F^∞, g_∞, or h_∞ (or some of their components) may be unobservable. A particular example in which expectations are involved in the objective function and in the complementarity constraints is the following:

$$\begin{array}{ll} \min & \mathbb{E}_\omega[f(x, y, \omega)] \\ \text{s.t.} & (x, y) \in Z \\ & y \geq 0, \quad \mathbb{E}_\omega[F(x, y, \omega)] \geq 0 \\ & y\, \mathbb{E}_\omega[F(x, y, \omega)] = 0, \end{array} \tag{3}$$

where $Z = \{\, (x, y) \mid g(x, y) \leq 0, \ h(x, y) = 0 \,\}$ and ω denotes the random element in the model. This formulation deals with problems in which all decisions, that is, at both upper and lower levels, must be made at once, *before* observing the random event ω. From this point of view, the stochastic MPCC (SMPCC) under consideration here differs from the stochastic programming extension of MPEC's as formulated in Patriksson and Wynter (1999), where the complementarity (or equilibrium) constraints are required to hold individually for every realization of ω and the lower level decisions are taken *after* the value of ω is observed. By analogy with some stochastic programming terminology, Lin *et al.* (2003) call the latter situation a lower level *wait-and-see* problem with upper level decisions x and lower-level decisions y. Lin *et al.* (2003) also discuss a different variant, somewhat closer to ours, in which both x and y must be determined a priori to the random event ω. However, their variant makes use of a recourse variable depending on ω and eventually, some "adjusted" complementarity constraints are yet required to hold individually, for each ω. This second variant in Lin *et al.* (2003) is called a *here-and-now* model. Our variant of SMPCC differs from either formulation above, in that it does not impose individual realization constraints, but rather complementarity constraints at an "average" level. For a more formal and detailed discussion on different SMPCC formulations as well as some concrete examples we refer to our accompanying paper Birbil *et al.* (2004).

We propose to solve problem SMPCC using the so-called sample-path method. The basic case of sample-path optimization, concerning the solution of simulation optimization problems with deterministic constraints, appeared in Plambeck *et al.* (1993, 1996) and was analyzed in Robinson (1996). Plambeck *et al.* (1993, 1996) used infinitesimal perturbation analysis (IPA) for gradient estimation. In the static case, a closely related technique centered around likelihood-ratio methods appeared in Rubinstein and Shapiro (1993) under the name of stochastic counterpart methods. The basic approach (and its variants) is also known as sample average approximation method in the stochastic programming literature; see for example Shapiro and Homem-De-Mello (1998), Kleywegt *et al.* (2001), and Linderoth *et al.* (2002).

Gürkan *et al.* (1996, 1999a) extended the basic idea of using sample-path information to solve stochastic equilibrium problems. They presented a framework to model such equilibrium problems as stochastic variational inequalities and provide conditions under which equilibrium points of approximating problems (computed via simulation and deterministic variational inequality solvers) converge almost surely to the solution of the limiting problem which can not be observed. Gürkan *et al.* (1999a) also considered a numerical application about finding the equilibrium prices and quantities of natural gas in the European market. This work was used further in Gürkan *et al.* (1999b) for establishing almost-sure convergence of sample-path methods when dealing with stochastic optimization problems with stochastic constraints.

In order to guarantee the closeness of solutions of the approximating variational inequalities to the solution of the real problem, a certain functional convergence of the data functions should be imposed. The specific property required is called *continuous convergence* and will be denoted by $\xrightarrow{\mathcal{C}}$; it is equivalent to uniform convergence on compact sets to a continuous limit. For an elementary treatment of the relationship between different types of functional convergence, see Kall (1986), and for a comprehensive treatment of continuous convergence and related issues, see Rockafellar and Wets (1998). In particular, Theorem 1 and Corollary 7 of Kall (1986), and Theorems 7.11, 7.17, and 7.18 of Rockafellar and Wets (1998) provide other equivalent notions of convergence or conditions that imply continuous convergence. In the sequel we are going to employ the term continuous convergence for the analysis of the problem SMPCC as well. For convenience we restate the following definition.

Definition 1 *A sequence f_n of real-valued functions defined on $\mathbb{R}^k$ converges continuously to a real-valued function f defined on $\mathbb{R}^k$ (written $f_n \xrightarrow{\mathcal{C}} f$) at x_0 if for any sequence $\{x_n\}$ converging to x_0, one has $f_n(x_n) \to f(x_0)$. If $f_n \xrightarrow{\mathcal{C}} f$ at x for any $x \in \mathbb{R}^k$, then we say that f_n converges to f on $\mathbb{R}^k$ and write $f_n \xrightarrow{\mathcal{C}} f$. A sequence of functions from $\mathbb{R}^k$ into $\mathbb{R}^m$ converges continuously if each of the m component functions does so.*

The discussion above about variational inequalities and continuous convergence plays an important role in our subsequent sections because the stationarity conditions for MPCC's may be expressed in terms of variational inequalities and moreover, these conditions can be solved approximately to find estimates of the true solutions. In this case, we can apply a similar argument as in Gürkan et. al. (1999a) and justify the validity of the approximating solutions. However, although useful, continuous convergence by itself guarantees neither the existence of such approximating solutions nor their convergence. To guarantee these, one needs to impose an additional regularity condition. We will elaborate this in the next section.

3 Solution of stochastic mathematical programs with complementarity constraints

We focus in this section on solving the problem SMPCC, in which no explicit description is available, in general, for any of the defining functions. For ease of notation, however, the ∞ scripts in SMPCC are omitted from now on. We start by reviewing some useful concepts pertaining to the framework of deterministic MPCC's.

It is well known that, from the viewpoint of nonlinear programming, the complementarity constraints involving the function F are problematic, irrespective of the properties of F, since no solution z can be a strictly feasible point (i.e. it is impossible for all the entries of $F(z)$ in one row to be positive). Consequently, the standard Mangasarian-Fromovitz constraint qualification is violated at every feasible point and one needs to deal with the complementarity constraints explicitly. In the next subsection we review some important stationarity concepts for MPCC's as discussed in Scheel and Scholtes (2000).

3.1 Constraint qualifications and stationarity concepts for MPCC's

One can associate with an MPCC the following Lagrangian function:

$$L(z, \Gamma, \lambda, \mu) = f(z) - F(z)\Gamma + g(z)\lambda + h(z)\mu \,, \tag{4}$$

where $\lambda \in \mathbb{R}^p$, $\mu \in \mathbb{R}^q$, and $\Gamma \in \mathbb{R}^{m\times l}$ are the corresponding Lagrange multipliers and $F(z)\Gamma = \sum_i \sum_j F_{ij}(z)\Gamma_{ij}$ is the inner product of the two $m \times l$-matrices. We will refer in the sequel to three constraint qualifications for MPCC's, restated below in the form used by Scheel and Scholtes (2000). At a feasible point z, the linear independence constraint qualification (LICQ) for MPCC is said to be satisfied if the gradients

$$\begin{aligned} &\nabla F_{ki}(z)\,, && (k,i): \; F_{ki}(z) = 0\,, \\ &\nabla g_r(z)\,, && r: \; g_r(z) = 0\,, \\ &\nabla h_s(z)\,, && s = 1, ..., q \end{aligned}$$

are linearly independent. The Mangasarian-Fromovitz constraint qualification (MFCQ) holds when the gradients

$$\begin{aligned} &\nabla F_{ki}(z)\,, && (k,i): \; F_{ki}(z) = 0\,, \\ &\nabla h_s(z)\,, && s = 1, ..., q \end{aligned}$$

are linearly independent and that there exists a vector v orthogonal to the these gradients such that

$$\nabla g_r(z)v < 0\,, \qquad r:\ g_r(z) = 0.$$

Finally, the strict Mangasarian-Fromovitz constraint qualification (SMFCQ) is said to be satisfied if there exist Lagrange multipliers Γ, λ, and μ such that the gradients

$$\begin{array}{ll} \nabla F_{ki}(z)\,, & (k,i):\ F_{ki}(z) = 0\,, \\ \nabla g_j(z)\,, & j:\ \lambda_j > 0\,, \\ \nabla h_s(z)\,, & s = 1, ..., q \end{array}$$

are linearly independent and there exist a vector v such that

$$\begin{array}{rl} \nabla F_{ki}(z)v = 0\,, & (k,i):\ F_{ki}(z) = 0\,, \\ \nabla h_s(z)v = 0\,, & s = 1, ..., q\,, \\ \nabla g_j(z)v = 0\,, & j:\ \lambda_j > 0\,, \\ \nabla g_r(z)v < 0\,, & r:\ g_r(z) = \lambda_r = 0. \end{array}$$

It is well-known that LICQ implies SMFCQ, while SMFCQ implies MFCQ.

Scheel (1995) and Luo *et al.* (1996) proposed the local decomposition approach to MPCC, in which the feasible region of the MPCC can be expressed locally as the union of the feasible regions of certain nonlinear programs constructed using the problem data. This approach is reviewed in Scheel and Scholtes (2000). Based on this argument, they provide a thorough discussion on different stationarity concepts for MPCC's and how they relate to the local minima. A point z is called a *weakly stationary point* for MPCC if there exist multipliers Γ, λ, and μ such that

$$\begin{array}{rcll} \nabla_z L(z, \Gamma, \lambda, \mu) & = & 0 & \\ \min\{F_{k1}(z), \ldots, F_{kl}(z)\} & = & 0 & k = 1, \ldots, m \\ h(z) & = & 0 & \\ g(z) & \leq & 0 & \\ \lambda & \geq & 0 & \\ g_r(z)\lambda_r & = & 0 & r = 1, \ldots, p \\ F_{ki}(z)\Gamma_{ki} & = & 0 & k = 1, \ldots, m, \quad i = 1, \ldots, l\,. \end{array} \tag{5}$$

In general, a weakly stationary point may preserve this property for arbitrary local shapes of the objective function. Nevertheless, under appropriate assumptions, the multipliers associated with a weakly stationary point can provide valuable information about the local geometry of the problem. Moreover, weak stationarity provides the basis for the definition of stronger stationarity concepts. A point z is called a *C-stationary point* for MPCC if, in addition to (5), it holds moreover that

$$\Gamma_{ki}\Gamma_{kj} \geq 0 \qquad \forall\, k\ \forall\, (i,j):\ \ F_{ki}(z) = F_{kj}(z) = 0\,. \tag{6}$$

C-stationarity is a necessary optimality condition under the MFCQ. If in addition to (5), it holds furthermore that

$$\Gamma_{ki} \geq 0 \qquad \forall\, k\ \forall\, i:\ \exists\, j \neq i\ \text{ such that }\ F_{ki}(z) = F_{kj}(z) = 0\,, \tag{7}$$

then z is called a *strongly stationary point* for MPCC. Strong stationarity becomes a necessary condition for optimality if SMFCQ is satisfied.

Scheel and Scholtes (2000) prove that under SMFCQ (or under the stronger condition LICQ), a point z is a strongly stationary point for MPCC if and only if z is a *B-stationary point*; that is, a feasible point of the MPCC for which $\nabla f(z)d \geq 0$ for every d satisfying

$$\begin{aligned} \min\{\,\nabla F_{ki}(z)d \mid i:\ F_{ki}(z)=0\,\} &= 0 \qquad && \forall\, k=1,\ldots,m \\ \nabla g_r(z)d &\leq 0 && \forall\, r:\ g_r(z)=0 \\ \nabla h(z)d &\leq 0\,. \end{aligned} \tag{8}$$

Thus, B-stationarity at z means that z is a local minimizer of the *linearized* MPCC which is obtained by linearizing locally all data functions.

In the sequel we need one more definition. A solution (z,Γ,λ,μ) of (5) is said to satisfy the *upper level strict complementarity condition* (ULSC) if $\Gamma_{ki}\neq 0$ for every (k,i) with the property that $F_{ki}(z)=F_{kj}(z)=0$ for some $j\neq i$. In the next subsection we set forth the analysis of the sample-path method applied to problem SMPCC.

3.2 Solution of SMPCC's using the sample-path method

As mentioned earlier, we consider the situation where we can not observe the functions f, g, h, and F. However, suppose we can observe some sequences of functions $\{f_n\}$, $\{g^n\}$, $\{h^n\}$, and $\{F^n\}$ for $n\in\mathbb{N}$, which approximate f, g, h, and F, respectively. Then we are concerned with sufficient conditions under which the solutions of SMPCC can be approximated by the solutions of the sequence of MPCC's of the following type:

$$\begin{array}{llll} \mathrm{MPCC}_n & \min & f_n(z) & \\ & \text{s.t.} & \min\{F^n_{k1}(z),\ldots,F^n_{kl}(z)\}=0 & \qquad k=1,\ldots,m \\ & & g^n(z)\leq 0 & \\ & & h^n(z)=0 & \\ & & z\in\Theta\,. & \end{array}$$

Assuming that the functions $\{f_n\}$, $\{g^n\}$, $\{h^n\}$, and $\{F^n\}$ are twice differentiable, the weak stationarity conditions for MPCC_n are of the form (5), where the true functions are replaced by the approximating functions (and their derivatives). As these conditions represent a system which approximates system (5), the strategy could be to solve such an approximating system for n sufficiently large.

However, an important way of envisioning the approximating setup is to regard the approximating functions as estimates of the true functions obtained from a simulation run of length n. In this sense, the above setting may not be always achievable in applications. In the context of option pricing, Gürkan *et al.* (1996) provide an example in which an unobservable function f_∞ is approximated by a sequence $\{f_n : n\in\mathbb{N}\}$ of step functions. Hence, each f_n has a finite (but large) number of discontinuity points and a zero derivative on the rest of the domain, which makes it extremely difficult to optimize. On the other hand, the authors show that the derivative ∇f_∞ of f_∞ may be approximated by a sequence g_n of nicely behaved (smooth) functions. Clearly, in this example g_n does not coincide with ∇f_n (at the points where the latter is defined). Such examples indicate that in practise it is important to work with assumptions as weak as possible. Therefore, we focus here on a more general context as explained below.

We assume that one can observe some sequences of functions $\{g^n\}$, $\{h^n\}$, $\{F^n\}$, $\{a^n\}$, $\{b^n\}$, $\{c^n\}$, and $\{d^n\}$ for $n \in \mathbb{N}$, which approximate g, h, F, ∇f, ∇g, ∇h, and ∇F respectively. In the elaboration of our main result we will use the following notation:

$$J(z) = (\nabla f(z), \nabla g(z), \nabla h(z), \nabla F(z)), \tag{9}$$

$$J^n(z) = (a^n(z), b^n(z), c^n(z), d^n(z)), \tag{10}$$

$$dL^n(z, \Gamma, \lambda, \mu) = a^n(z) - d^n(z)\Gamma + b^n(z)\lambda + c^n(z)\mu\,. \tag{11}$$

In this setting, we are actually concerned with sufficient conditions under which the solutions of (5) can be approximated by the solutions of the sequence of systems of the following type:

$$\begin{aligned} dL^n(z, \Gamma, \lambda, \mu) &= 0 \\ \min\{F^n_{k1}(z), \dots, F^n_{kl}(z)\} &= 0 \qquad k = 1, \dots, m \\ h^n(z) &= 0 \\ g^n(z) &\leq 0 \\ \lambda &\geq 0 \\ g^n_r(z)\lambda_r &= 0 \qquad r = 1, \dots, p \\ F^n_{ki}(z)\Gamma_{ki} &= 0 \qquad k = 1, \dots, m, \quad i = 1, \dots, l\,. \end{aligned} \tag{12}$$

Notice that if $b^n = \nabla g^n$, $c^n = \nabla h^n$ and $d^n = \nabla F^n$ for every $n \in \mathbb{N}$ and if moreover, f is approximated by a sequence of functions f_n such that $a^n = \nabla f_n$ for every $n \in \mathbb{N}$, then indeed, the approximating problem (12) represents the weak stationarity conditions for an approximating program MPCC$_n$ of the type above. In the sequel we work in the more general context as given by (12).

The following theorem contains our main result on existence and convergence of approximating solutions. Roughly speaking, it says that if the SMPCC defined by the limit functions has a weakly stationary point $\bar{z}$ satisfying a regularity condition, then for sufficiently good approximations of the limit functions the approximating problems must have weakly stationary points close to $\bar{z}$.

Theorem 1 *Let Θ be an open set in $\mathbb{R}^{n_0}$. Suppose that f, g, h, and F are functions from Θ to $\mathbb{R}$, $\mathbb{R}^p$, $\mathbb{R}^q$, and $\mathbb{R}^{m\times l}$ respectively, which are twice differentiable and that J is defined as in (9). Let $\bar{z} \in \Theta$, $\bar{\Gamma} \in \mathbb{R}^{m\times l}$, $\bar{\lambda} \in \mathbb{R}^p$, and $\bar{\mu} \in \mathbb{R}^q$. Suppose that $\{J^n \mid n = 1, 2, ...\}$ are random functions defined on Θ, as in (10), $\{dL^n \mid n = 1, 2, ...\}$ are random functions defined as in (11), $\{g^n \mid n = 1, 2, ...\}$ are random functions from Θ to $\mathbb{R}^p$, $\{h^n \mid n = 1, 2, ...\}$ are random functions from Θ to $\mathbb{R}^q$, and $\{F^n \mid n = 1, 2, ...\}$ are random functions from Θ to $\mathbb{R}^{m\times l}$ such that for all $z \in \Theta$ and all finite n the random variables $J^n(z)$, $g^n(z)$, $h^n(z)$, and $F^n(z)$ are defined on a common probability space $(\Omega, \mathcal{F}, P)$. Let $L(z, \Gamma, \lambda, \mu)$ be defined as in (4) and assume the following:*

1) *With probability one, each J^n for $n = 1, 2, ...$ is continuous and $J^n \xrightarrow{\mathcal{C}} J$.*
2) *With probability one, each g^n for $n = 1, 2, ...$ is continuous and $g^n \xrightarrow{\mathcal{C}} g$.*
3) *With probability one, each h^n for $n = 1, 2, ...$ is continuous and $h^n \xrightarrow{\mathcal{C}} h$.*

4) With probability one, each F^n for $n = 1, 2, ...$ is continuous and $F^n \xrightarrow{\mathcal{C}} F$.

5) $(\bar{z}, \bar{\Gamma}, \bar{\lambda}, \bar{\mu})$ is a solution of (5) (that is, a weakly stationary point of the SMPCC).

6) $\nabla_z L$ has a strong Fréchet derivative $\nabla^2_{zz} L(\bar{z}, \bar{\Gamma}, \bar{\lambda}, \bar{\mu})$ at the point $(\bar{z}, \bar{\Gamma}, \bar{\lambda}, \bar{\mu})$ and all the matrices

$$\begin{pmatrix} \nabla^2_{zz} L(\bar{z}, \bar{\Gamma}, \bar{\lambda}, \bar{\mu}) & -\nabla_z F_I(\bar{z})^\top & \nabla_z g_R(\bar{z})^\top & \nabla_z h(\bar{z})^\top \\ \nabla_z F_I(\bar{z}) & 0 & 0 & 0 \\ -\nabla_z g_R(\bar{z}) & 0 & 0 & 0 \\ \nabla_z h(\bar{z}) & 0 & 0 & 0 \end{pmatrix} \tag{13}$$

with

$$\{(k,i) \mid \bar{\Gamma}_{ki} \neq 0\} \subseteq I \subseteq \{(k,i) \mid F_{ki}(\bar{z}) = 0\} \text{ and } \forall\, k = 1, ..., m \;\exists\, i \in \{1, ..., l\} : (k,i) \in I\,,$$

$$\text{and } \; \{r \mid \bar{\lambda}_r > 0\} \subseteq R \subseteq \{r \mid g_r(\bar{z}) = 0\}\,,$$

have the same nonvanishing determinantal sign.

Then, there exist compact subsets $C_0 \subset \Theta$ containing $\bar{z}$, $U_0 \subset \mathbb{R}^{m\times l}$ containing $\bar{\Gamma}$, $V_0 \subset \mathbb{R}^p$ containing $\bar{\lambda}$, and $W_0 \subset \mathbb{R}^q$ containing $\bar{\mu}$, neighborhoods $Y_1 \subset \Theta$ of $\bar{z}$, $U_1 \subset \mathbb{R}^{m\times l}$ of $\bar{\Gamma}$, $V_1 \subset \mathbb{R}^p$ of $\bar{\lambda}$, and $W_1 \subset \mathbb{R}^q$ of $\bar{\mu}$, a constant $\alpha > 0$ and a set $\Delta \subset \Omega$ of measure zero, with the following properties: for $n = 1, 2, ...$ and $\omega \in \Omega$ let

$$\begin{aligned} \xi_n(\omega) := \sup\{\, &\|(dL^n(\omega, z, \Gamma, \lambda, \mu), (g^n, h^n, F^n)(\omega, z)) - (\nabla_z L(z, \Gamma, \lambda, \mu), (g, h, F)(z))\| : \\ &(z, \Gamma, \lambda, \mu) \in C_0 \times U_0 \times V_0 \times W_0 \,\}\,, \end{aligned}$$

$$Z_n(\omega) := \{\, (z, \Gamma, \lambda, \mu) \in Y_1 \times U_1 \times V_1 \times W_1 \mid (z, \Gamma, \lambda, \mu) \text{ solves (12) corresponding to } \omega \,\}\,.$$

For each $\omega \notin \Delta$ there is then a finite integer N_ω such that for every $n \geq N_\omega$ the set $Z_n(\omega)$ is a nonempty, compact subset of $B((\bar{z}, \bar{\Gamma}, \bar{\lambda}, \bar{\mu}), \alpha\, \xi_n(\omega))$.

Proof. Determine $\Delta \subset \Omega$ of measure zero such that off Δ the properties listed in assumptions 1) − 4) hold for all n. We fix an $\omega \notin \Delta$ and suppress ω from the rest of discussion.

The nonvanishing determinantal sign property in assumption 6) implies that LICQ holds at $\bar{z}$, so the multipliers $\bar{\Gamma}$, $\bar{\lambda}$, and $\bar{\mu}$ are unique. Define

$$I_{\neq} = \{(k,i) \mid \bar{\Gamma}_{ki} \neq 0\}\,, \quad I_F(\bar{z}) = \{(k,i) \mid F_{ki}(\bar{z}) = 0\}\,, \quad I_F^n(z) = \{(k,i) \mid F^n_{ki}(z) = 0\}\,,$$

$$\text{and } \; \mathcal{I} := \{\, I \mid I_{\neq} \subseteq I \subseteq I_F(\bar{z}) \text{ and } \forall\, k = 1, ..., m \;\exists\, i \in \{1, ..., l\} \colon (k,i) \in I \}\,.$$

Following a similar argument to Theorem 12 of Scheel and Scholtes (2000), we first show that there exist a neighborhood $Z^* \subset \Theta$ and an integer $N^* \in \mathbb{N}$ such that $z \in Z^*$ (together with some multipliers) is a solution of (12) for $n \geq N^*$ if and only if z (together with some multipliers) solves a problem of type

$$\begin{aligned} dL^n(z, \Gamma, \lambda, \mu) &= 0 \\ F^n_I(z) &= 0 \\ F^n_{I^c}(z) &\geq 0 \\ \Gamma_{I^c} &\geq 0 \\ F^n_{ki}(z)\Gamma_{ki} &= 0 \qquad \forall\, (k,i) \in I^c \end{aligned} \tag{14}$$

$$
\begin{aligned}
h^n(z) &= 0 \\
g^n(z) &\le 0 \\
\lambda &\ge 0 \\
g_r^n(z)\lambda_r &= 0 \qquad r = 1, \ldots, p\,,
\end{aligned}
$$

with $I \in \mathcal{I}$ and $I^c = \{1, ..., m\} \times \{1, ..., l\} \setminus I$ is the complement of I. Note that if problem (12) represented the weak stationarity conditions for an approximating program MPCC_n (see our remark above), then problem (14) would represent the stationarity conditions of one of the nonlinear programs associated to the point $\bar{z}$ in the local decomposition approach to that MPCC_n.

Take a neighborhood $Z^* \subset \Theta$ of $\bar{z}$ and an integer $N^* \in \mathbb{N}$ such that for every $z \in Z^*$ and every $n \ge N^*$ the inclusion $I_F^n(z) \subseteq I_F(\bar{z})$ holds. If z together with some multipliers Γ, λ, and μ satisfy (12) for a certain n, then $(z, \Gamma, \lambda, \mu)$ satisfy (14) for that n and for $I = I_F^n(z)$. If z is sufficiently close to $\bar{z}$ and n is large enough then, in the view of LICQ, the corresponding multipliers Γ_{ki} (which satisfy (12) for the considered n) are close to $\bar{\Gamma}_{ki}$. This can be seen by applying the classical implicit function theorem to the system of linear equations $\nabla_z L(z, \Gamma, \lambda, \mu) = 0$, in the variables (Γ, λ, μ). So, reducing Z^* and increasing N^* if necessary, we may assume that $I_{\neq} \subseteq I_F^n(z)$ for all $z \in Z^*$ and all $n \ge N^*$. Moreover, since z satisfies (12), it follows that for every $k \in \{1, ..., m\}$ there exists $i \in \{1, ..., l\}$ such that $F_{ki}^n(z) = 0$. Hence we conclude that $(z, \Gamma, \lambda, \mu)$ solves (14) for some index set I from the collection $\mathcal{I}$.

To see the converse, suppose that $(z^I, \Gamma^I, \lambda^I, \mu^I)$ satisfy (14) for some n and for some I from $\mathcal{I}$. Then z^I satisfies (12), since for every k there exists at least one index i with $F_{ki}^n(z) = 0$. Moreover, the relations in (14) imply the relations in (12).

Now we are going to treat closer a partial problem of the form (14). Let I be an arbitrary set from $\mathcal{I}$. Since $(\bar{z}, \bar{\Gamma}, \bar{\lambda}, \bar{\mu})$ satisfy (5), it follows that $(\bar{z}, \bar{\Gamma}, \bar{\lambda}, \bar{\mu})$ also satisfy the following conditions:

$$
\begin{aligned}
\nabla_z L(z, \Gamma, \lambda, \mu) &= 0 \\
F_I(z) &= 0 \\
F_{I^c}(z) &\ge 0 \\
\Gamma_{I^c} &\ge 0 \\
F_{ki}(z)\Gamma_{ki} &= 0 \qquad \forall\, (k, i) \in I^c \\
h(z) &= 0 \\
g(z) &\le 0 \\
\lambda &\ge 0 \\
g_r(z)\lambda_r &= 0 \qquad r = 1, \ldots, p\,.
\end{aligned}
\tag{15}
$$

Our next step is to examine the sufficient conditions under which a solution of (15) can be approximated by the solutions of problems of type (14). Notice that (15) is equivalent with the following generalized equation (see Appendix 1):

$$
0 \in \begin{bmatrix} \nabla_z L(z, \Gamma, \lambda, \mu) \\ F_I(z) \\ F_{I^c}(z) \\ -g(z) \\ h(z) \end{bmatrix} + N_{\mathbb{R}^{n_0} \times \mathbb{R}^{|I|} \times \mathbb{R}_+^{|I^c|} \times \mathbb{R}_+^p \times \mathbb{R}^q} \begin{bmatrix} z \\ \Gamma_I \\ \Gamma_{I^c} \\ \lambda \\ \mu \end{bmatrix}. \tag{16}
$$

For a general introduction to generalized equations and their use in reformulating complementarity we refer to Robinson (1979). Similarly, (14) is equivalent with the generalized equation:

$$0 \in \begin{bmatrix} dL^n(z,\Gamma,\lambda,\mu) \\ F_I^n(z) \\ F_{I^c}^n(z) \\ -g^n(z) \\ h^n(z) \end{bmatrix} + N_{\mathbb{R}^{n_0}\times\mathbb{R}^{|I|}\times\mathbb{R}_+^{|I^c|}\times\mathbb{R}_+^p\times\mathbb{R}^q} \begin{bmatrix} z \\ \Gamma_I \\ \Gamma_{I^c} \\ \lambda \\ \mu \end{bmatrix} . \tag{17}$$

Hence we regard (17) as an approximation to (16) and check in the sequel that for these generalized equations (which are in turn equivalent to solving appropriate variational inequalities) the assumptions of Theorem 2 in Gürkan *et al.* (1999a) are satisfied. For the ease of reference, the last theorem is included as Theorem 5 in our Appendix 2.

Clearly, by our assumptions 1) – 4), the functions in (17) are continuous and converge continuously to the corresponding functions in (16). Since $(\bar{z},\bar{\Gamma},\bar{\lambda},\bar{\mu})$ satisfy (15), it is straightforward that $(\bar{z},\bar{\Gamma},\bar{\lambda},\bar{\mu})$ is a solution of (16). From our assumption 6) and the twice differentiability of F, g, and h, we obtain that ∇L has a strong Fréchet derivative $\nabla^2 L(\bar{z},\bar{\Gamma},\bar{\lambda},\bar{\mu})$ at the point $(\bar{z},\bar{\Gamma},\bar{\lambda},\bar{\mu})$. Moreover, for all the index sets K and R with $I \subseteq K \subseteq I_F(\bar{z})$ and $\{r \mid \bar{\lambda}_r > 0\} \subseteq R \subseteq \{r \mid g_r(\bar{z}) = 0\}$, the matrices (13) have the same nonvanishing determinantal sign. In Appendix 1 we show that this nonvanishing determinantal sign condition is equivalent with the fact that the generalized equation (16) is strongly regular at $(\bar{z},\bar{\Gamma},\bar{\lambda},\bar{\mu})$ in the sense of Robinson (1980), which, in turn, is equivalent with the coherent orientation condition of Theorem 5 (for a detailed discussion of the second equivalence, see the appendix in Gürkan *et al.* (1999a)).

Thus we can apply Theorem 5 to (16) and (17) and obtain that there exist compact sets $C_0^I \subset \Theta$ containing $\bar{z}$, $U_0^I \subset \mathbb{R}^{m\times l}$ containing $\bar{\Gamma}$, $V_0^I \subset \mathbb{R}^p$ containing $\bar{\lambda}$, and $W_0^I \subset \mathbb{R}^q$ containing $\bar{\mu}$, neighborhoods $Y_1^I \subset Z^*$ of $\bar{z}$, $U_1^I \subset \mathbb{R}^{m\times l}$ of $\bar{\Gamma}$, $V_1^I \subset \mathbb{R}^p$ of $\bar{\lambda}$, and $W_1^I \subset \mathbb{R}^q$ of $\bar{\mu}$, a constant $\alpha^I > 0$ with the properties: for $n = 1, 2, ...$ let

$$\begin{aligned} \xi_n^I = \sup\{ &\|(dL^n(z,\Gamma,\lambda,\mu),(g^n,h^n,F^n)(z)) - (\nabla_z L(z,\Gamma,\lambda,\mu),(g,h,F)(z))\| : \\ &(z,\Gamma,\lambda,\mu) \in C_0^I \times U_0^I \times V_0^I \times W_0^I \} \end{aligned}$$

and

$$Z_n^I := \{ (z,\Gamma,\lambda,\mu) \in Y_1^I \times U_1^I \times V_1^I \times W_1^I \mid (z,\Gamma,\lambda,\mu) \text{ solves (17)} \} .$$

Then there is a finite integer $N_1^I \geq N^*$ such that for each $n \geq N_1^I$ the set Z_n^I is a nonempty, compact subset of $B((\bar{z},\bar{\Gamma},\bar{\lambda},\bar{\mu}),\alpha^I\,\xi_n^I)$.

Our last step is to construct the necessary sets mentioned in the conclusion of the theorem based on the collection of the I sets ($I \in \mathcal{I}$) we found above. For this, let

$$C_0 = \bigcup_{I\in\mathcal{I}} C_0^I\ , \quad U_0 = \bigcup_{I\in\mathcal{I}} U_0^I\ , \quad V_0 = \bigcup_{I\in\mathcal{I}} V_0^I\ , \quad W_0 = \bigcup_{I\in\mathcal{I}} W_0^I\ ,$$

$$Y_1 = \bigcap_{I\in\mathcal{I}} Y_1^I\ , \quad U_1 = \bigcap_{I\in\mathcal{I}} U_1^I\ , \quad V_1 = \bigcap_{I\in\mathcal{I}} V_1^I\ , \quad W_1 = \bigcap_{I\in\mathcal{I}} W_1^I\ ,$$

and $\alpha = \max\{ \alpha^I \mid I \in \mathcal{I} \}$. Let I be an index set from $\mathcal{I}$. As $\xi_n^I \to 0$ for $n \to \infty$, it follows that there is an integer $N_2^I \geq N_1^I$ such that $Z_n^I \subset Y_1 \times U_1 \times V_1 \times W_1$ for $n \geq N_2^I$. Now let $N = \max\{ N_2^I \mid I \in \mathcal{I} \}$. Then for $n \geq N$ we have that the set Z_n defined in the statement of the theorem can be expressed as $Z_n = \bigcup\{Z_n^I \mid I \in \mathcal{I}\}$. This relation together with the definition of ξ_n lead us to the conclusion that for $n \geq N$ the set Z_n is a nonempty, compact subset of $B((\bar{z}, \bar{\Gamma}, \bar{\lambda}, \bar{\mu}), \alpha\, \xi_n)$. ∎

If we are concerned with the approximation of C-stationary or strongly stationary points of the SMPCC, the question arises if one can strengthen the approximating problem (12) by imposing the additional constraints

$$\Gamma_{ki}\Gamma_{kj} \geq 0 \qquad \forall\, k\; \forall\, (i,j): \;\; F^n_{ki}(z) = F^n_{kj}(z) = 0 \tag{18}$$

or respectively,

$$\Gamma_{ki} \geq 0 \qquad \forall\, k\; \forall\, i: \;\exists\, j \neq i \;\text{ such that }\; F^n_{ki}(z) = F^n_{kj}(z) = 0\,. \tag{19}$$

This can be done if assumption 6) in Theorem 1 is strengthen with the ULSC as shown by the following result.

Theorem 2 *Assume the setting of Theorem 1 and assume that conditions 1)–5) hold. Moreover, assume the following:*

6′) The ULSC holds at $(\bar{z}, \bar{\Gamma}, \bar{\lambda}, \bar{\mu})$ and all the matrices

$$\begin{pmatrix} \nabla^2_{zz} L(\bar{z}, \bar{\Gamma}, \bar{\lambda}, \bar{\mu}) & -\nabla_z F_I(\bar{z})^\top & \nabla_z g_R(\bar{z})^\top & \nabla_z h(\bar{z})^\top \\ \nabla_z F_I(\bar{z}) & 0 & 0 & 0 \\ -\nabla_z g_R(\bar{z}) & 0 & 0 & 0 \\ \nabla_z h(\bar{z}) & 0 & 0 & 0 \end{pmatrix} \tag{20}$$

with $I = \{(k,i) \mid F_{ki}(\bar{z}) = 0\}$ and $\{r \mid \bar{\lambda}_r > 0\} \subseteq R \subseteq \{r \mid g_r(\bar{z}) = 0\}$, have the same nonvanishing determinantal sign.

1°. If $\bar{\Gamma}$ satisfies (6) (that is, $\bar{z}$ is a C-stationary point of the SMPCC), then the conclusions of Theorem 1 hold with

$$Z_n(\omega) := \{(z, \Gamma, \lambda, \mu) \in Y_1 \times U_1 \times V_1 \times W_1 \mid (z, \Gamma, \lambda, \mu) \textit{ solves } (12) \textit{ corresponding to } \omega, \textit{ together with } (18)\,\}.$$

2°. If $\bar{\Gamma}$ satisfies (7) (that is, $\bar{z}$ is a strongly stationary point of the SMPCC), then the conclusions of Theorem 1 hold with

$$Z_n(\omega) := \{(z, \Gamma, \lambda, \mu) \in Y_1 \times U_1 \times V_1 \times W_1 \mid (z, \Gamma, \lambda, \mu) \textit{ solves } (12) \textit{ corresponding to } \omega, \textit{ together with } (19)\,\}.$$

Proof. As above, we start by determining a subset $\Delta \subset \Omega$ of measure zero such that off Δ the properties listed in assumptions 1)– 4) hold for all n. We fix again an $\omega \notin \Delta$ and suppress ω from the rest of discussion.

Again, we initially follow an argument similar to Theorem 11 of Scheel and Scholtes (2000). The nonvanishing determinantal sign property in assumption 6′) implies that LICQ holds at $\bar{z}$, so the multipliers $\bar{\Gamma}$, $\bar{\lambda}$, and $\bar{\mu}$ are unique. Moreover, given the LICQ, if $(z, \Gamma, \lambda, \mu)$ satisfy (12) for n large enough, then the multipliers (Γ, λ, μ) are close to $(\bar{\Gamma}, \bar{\lambda}, \bar{\mu})$.

Hence, in the view of ULSC, $(z, \Gamma, \lambda, \mu)$ satisfy (12) for z close to $\bar{z}$ and n large enough if and only if

$$\begin{aligned} dL^n(z, \Gamma, \lambda, \mu) &= 0 \\ F^n_{ki}(z) &= 0 \qquad (k,i): \ F_{ki}(\bar{z}) = 0 \\ h^n(z) &= 0 \\ g^n(z) &\leq 0 \\ \lambda &\geq 0 \\ g^n_r(z)\lambda_r &= 0 \qquad r = 1, \ldots, p \\ \Gamma_{ki} &= 0 \qquad (k,i): \ F_{ki}(\bar{z}) > 0\,. \end{aligned} \tag{21}$$

Conditions (21) are in fact conditions (14) with $I = \{(k,i) \mid F_{ki}(\bar{z}) = 0\}$. Obviously, $(\bar{z}, \bar{\Gamma}, \bar{\lambda}, \bar{\mu})$ satisfy in this case the conditions

$$\begin{aligned} \nabla L_z(z, \Gamma, \lambda, \mu) &= 0 \\ F_{ki}(z) &= 0 \qquad (k,i) \in I \\ h(z) &= 0 \\ g(z) &\leq 0 \\ \lambda &\geq 0 \\ g_r(z)\lambda_r &= 0 \qquad r = 1, \ldots, p \\ \Gamma_{ki} &= 0 \qquad (k,i) \in I^c\,, \end{aligned} \tag{22}$$

that is, conditions (15) with I as defined above. Basically, due to the ULSC, the situation here corresponds to the situation from the proof of Theorem 1 where the collection of index sets $\mathcal{I}$ reduces to just one element $\mathcal{I} = \{\, I\,\}$. The conclusion follows then from a similar argument as in the proof of Theorem 1. Moreover, for z close to $\bar{z}$ and n large enough we have that

$$\{\,(k,i) \mid F^n_{ki}(z) = 0\} \subseteq \{\,(k,i) \mid F_{ki}(\bar{z}) = 0\}$$

and the following can be asserted:

1°. Suppose that $\bar{\Gamma}$ satisfies (6). Since ULSC holds at $(\bar{z}, \bar{\Gamma}, \bar{\lambda}, \bar{\mu})$ and Γ is close to $\bar{\Gamma}$, it follows that Γ satisfies (18).

2°. Similarly, suppose that $\bar{\Gamma}$ satisfies (7). Again Γ is close to $\bar{\Gamma}$ and this together with ULSC imply that Γ will satisfy (19) for such n. ▮

The ULSC condition has been used by Scheel and Scholtes (2000) in order to guarantee the local uniqueness of the (weakly) stationary point of a parametric MPCC as the parameter varies. Besides uniqueness, this condition also makes sure that strong stationarity and C-stationarity are preserved under small perturbations of the problem.

In general, without an additional condition such as ULSC, the conclusion of Theorem 2 fails. This can be seen by considering the SMPCC and the approximating MPCC$_n$'s of the same form as in Example 10 part (2) of Scheel and Scholtes (2000), corresponding to the values $t_\infty = 0$ and $t_n = 1/n$, respectively. Then the origin is a strongly stationary (B-stationary) point for the SMPCC. However, for any $n \in \mathbb{N}$, the origin is only a weakly stationary point for the approximating program MPCC_n and this program has no C-stationary points close to the origin. Clearly, in this case the limit program satisfies the weaker condition mentioned above at the origin, but not the stronger ULSC.

The example mentioned above also shows that the conclusion of Theorem 2 also fails in general if the ULSC is replaced by the weaker condition that at least one of the multipliers Γ_{ki} is positive for those (k,i) with the property that $F_{ki}(z) = F_{kj}(z) = 0$ for some $j \neq i$.

4 Discussion

We have established in Section 3 the convergence of the sample-path method for stochastic MPCC's based on sample-path convergence for stochastic variational inequalities. Here we provide a more detailed discussion on the sufficient conditions that guarantee almost-sure convergence.

We denote by $\mathbb{N}^\infty = \mathbb{N} \cup \{\infty\}$ the one point compactification of $\mathbb{N}$, organized as a topological space with the topology $\tau = \mathcal{P}(\mathbb{N}) \cup \{\mathbb{N}^\infty \setminus \{1,2,...,n\} : n \in \mathbb{N}\}$, where $\mathcal{P}(\mathbb{N})$ is the collection of all subsets of $\mathbb{N}$ (i.e., the discrete topology on $\mathbb{N}$); see for example page 183 of Munkres (1975). First we state the following auxiliary result.

Lemma 1 *Let $g : \mathbb{N}^\infty \times \mathbb{R}^k \to \mathbb{R}$ be a real function defined on the product topological space $\mathbb{N}^\infty \times \mathbb{R}^k$ and let $x_0 \in \mathbb{R}^k$ be an arbitrary point. Then the following statements are equivalent:*

a) g is continuous at the point (∞, x_0).

b) The partial function $g(\infty,\cdot)$ is continuous at x_0 and the partial functions $g(n,\cdot)$ converge continuously to $g(\infty,\cdot)$ at the point x_0 as $n \to \infty$.

The next result follows immediately through the application of Theorem 2.1 from Robinson (1980).

Theorem 3 *Let Θ be an open subset of $\mathbb{R}^k$, x_0 a point in Θ, $C \subseteq \mathbb{R}^k$ a closed convex set, and $f : \Theta \to \mathbb{R}^k$ a function which has a Fréchet derivative ∇f defined on Θ. Moreover, let $\{ f_n \mid n \in \mathbb{N} \}$ be a sequence of random functions from Θ to $\mathbb{R}^k$, such that for all finite $x \in \Theta$ and all finite n the random variables $f_n(x)$ are defined on a common probability space $(\Omega, \mathcal{F}, P)$. Suppose that the following conditions are satisfied:*

1) With probability one, each function f_n for $n = 1,2,...$ has a Fréchet derivative ∇f_n defined on Θ.

2) With probability one, $f_n \xrightarrow{\mathcal{C}} f$ at x_0 and $\nabla f_n \xrightarrow{\mathcal{C}} \nabla f$ at x_0.

3) x_0 is a solution of the generalized equation

$$0 \in f(x) + N_C(x). \tag{23}$$

4) The generalized equation (23) is strongly regular at x_0.

Then there exist a neighborhood W of x_0, a scalar $\alpha > 0$ and a set $\Delta \subset \Omega$ of measure zero, with the following properties: for $n \in \mathbb{N}$ and $\omega \in \Omega$ let

$$\xi_n(\omega) = \|f_n(\omega, x_0) - f(x_0)\|$$

and

$$X_n(\omega) := \{ x \in C \cap W \mid 0 \in f_n(\omega, x_0) + N_C(x) \}.$$

Then for each $\omega \notin \Delta$ there is a finite integer N_ω such that for every $n \geq N_\omega$ the set $X_n(\omega)$ is a singleton contained in $B(x_0, \alpha\, \xi_n(\omega))$.

Proof. First we determine Δ of measure zero such that off Δ the properties in hypotheses 1) and 2) hold for all n. We fix an $\omega \notin \Delta$ and suppress ω from here on.

We apply Theorem 2.1 from Robinson (1980) with the following setting: $P = \mathbb{N}^\infty$, $p_0 = \infty$, $f : \mathbb{N}^\infty \times \Theta \to \mathbb{R}^k$ with $f(\infty, x) = f(x)$, $f(n, x) = f_n(x)$ for any $n \in \mathbb{N}$ and $\nabla f(\infty, x) = \nabla f(x)$, $\nabla f(n, x) = \nabla f_n(x)$ is the partial Fréchet derivative of f with respect to x. Given the hypotheses 1) and 2), it follows from Lemma 1 that both $f(\cdot,\cdot)$ and $\nabla f(\cdot,\cdot)$ are continuous at the point (∞, x_0). Suppose that λ is the Lipschitz constant associated with the strong regularity property of assumption 4). Thus we obtain that for every $\epsilon > 0$ there exist an $N_\epsilon \in \mathbb{N}$ and a neighborhood W_ϵ of x_0, and a single-valued function $x : \mathbb{N}^\infty \setminus \{1, 2, ..., N_\epsilon - 1\} \to W_\epsilon$ such that for every $n \in \mathbb{N}^\infty \setminus \{1, 2, ..., N_\epsilon - 1\}$, x_n is the *unique* solution in W_ϵ of the generalized equation

$$0 \in f_n(x) + N_C(x)\,. \tag{24}$$

Further, for any $p, q \in \mathbb{N}^\infty \setminus \{1, 2, ..., N_\epsilon - 1\}$ one has

$$\|x_p - x_q\| \leq (\lambda + \epsilon)\|f_p(x_q) - f_q(x_q)\|\,. \tag{25}$$

Clearly, $x_\infty = x_0$ by our assumption 3). Hence, (25) implies in particular that for every $n \in \mathbb{N}$, $n \geq N_\epsilon$ one has

$$\|x_n - x_0\| \leq (\lambda + \epsilon)\|f_n(x_0) - f(x_0)\| = (\lambda + \epsilon)\xi_n\,, \tag{26}$$

that is, $x_n \in B(x_0, (\lambda + \epsilon)\xi_n)$.

Now, if we choose for example $\epsilon = 1$, we obtain that there exist a finite integer $N := N_1$ and a neighborhood $W := W_1$ of x_0 as well as a positive constant $\alpha := \lambda + 1$ such that for each $n \in \mathbb{N}$, $n \geq N$ the generalized equation (24) has a unique solution $x_n \in W$, that is, the set $X_n = \{x_n\}$ reduces to a singleton. Moreover, one has that $x_n \in B(x_0, \alpha\, \xi_n)$. ∎

Thus, such a convergence result for the sample-path method applied to variational inequalities defined on *arbitrary* closed convex sets can be obtained with little effort by employing a special case of the main result in Robinson (1980). In particular, under the assumptions in Theorem 3, not only are the existence and the convergence of approximate solutions guaranteed, but also the local uniqueness of these solutions can be established. However, as noted by the example of Gürkan *et al.* (1996), in some applications the assumptions on the existence and the continuous convergence of the derivatives $\nabla f_n \xrightarrow{\mathcal{C}} \nabla f$ at x_0 may prove to be quite strong.

In contrast with assumptions 1) and 2) of Theorem 3, Theorem 5 in Appendix 2 (a restatement of Theorem 2 in Gürkan *et al.* (1999a)) makes no explicit assumption on the derivatives of the functions f_n, but only on the continuity and the continuous convergence of these approximating functions. In Gürkan *et al.* (1999a), variational inequalities over polyhedral convex sets were considered. By employing degree theoretic concepts and normal maps techniques, the existence and convergence of the approximating solutions was established, but the local uniqueness of these solutions is no longer guaranteed. Consequently, in order to guarantee convergence of the approximate solutions, the continuous convergence $f_n \xrightarrow{\mathcal{C}} f$ is required to hold on the whole domain Θ (or at least on some neighborhood of the point x_0). Nevertheless, as already noted by Gürkan *et al.* (1999a), for computational purposes one may typically need more with respect to the derivatives of the approximating functions f_n. In this sense, Theorem 3 and Theorem 5 could be viewed as complementary results on the convergence of the sample-path method for stochastic variational inequalities.

Similar to Theorem 3, if one is willing to make the strong assumptions with respect to the existence and (local) continuous convergence of both first and second order derivatives of the data functions f, F, g, and h, then utilizing the one point compactification concept leads to discrete versions of the stability results of Scheel and Scholtes (2000) for the stochastic MPCC's, such as the one below.

Theorem 4 *Let Θ be an open set in $\mathbb{R}^{n_0}$. Suppose that f, F, g, and h are functions from Θ to $\mathbb{R}$, $\mathbb{R}^{m\times l}$, $\mathbb{R}^p$, and $\mathbb{R}^q$, respectively, which are twice differentiable. Let $\bar{z} \in \Theta$, $\bar{\Gamma} \in \mathbb{R}^{m\times l}$, $\bar{\lambda} \in \mathbb{R}^p$, and $\bar{\mu} \in \mathbb{R}^q$. Suppose that $\{f_n\}$, $\{F^n\}$, $\{g^n\}$, and $\{h^n\}$, $n \in \mathbb{N}$, are sequences of random functions defined from Θ to $\mathbb{R}$, $\mathbb{R}^{m\times l}$, $\mathbb{R}^p$, and $\mathbb{R}^q$, respectively, such that for all $z \in \Theta$ and all finite n the random variables $f_n(z)$, $F^n(z)$, $g^n(z)$, and $h^n(z)$ are defined on a common probability space $(\Omega, \mathcal{F}, P)$. Let $L(z, \Gamma, \lambda, \mu)$ be defined as in (4) and assume the following:*

1) With probability one, each of the functions f_n, F^n, g^n, and h^n, for $n \in \mathbb{N}$, is twice differentiable on Θ.

2) With probability one, the sequences $\{F^n\}$, $\{g^n\}$, and $\{h^n\}$ converge continuously to F, g, and h, respectively, at the point $\bar{z}$;

3) With probability one, the sequences $\{\nabla f_n\}$, $\{\nabla F^n\}$, $\{\nabla g^n\}$, and $\{\nabla h^n\}$ converge continuously to ∇f, ∇F, ∇g, and ∇h, respectively, at the point $\bar{z}$;

4) With probability one, the sequences $\{\nabla^2 f_n\}$, $\{\nabla^2 F^n\}$, $\{\nabla^2 g^n\}$, and $\{\nabla^2 h^n\}$ converge continuously to $\nabla^2 f$, $\nabla^2 F$, $\nabla^2 g$, and $\nabla^2 h$, respectively, at the point $\bar{z}$.

5) $(\bar{z}, \bar{\Gamma}, \bar{\lambda}, \bar{\mu})$ is a weakly stationary point of SMPCC.

6) All the matrices

$$\begin{pmatrix} \nabla^2_{zz} L(\bar{z}, \bar{\Gamma}, \bar{\lambda}, \bar{\mu}) & -\nabla_z F_I(\bar{z})^\top & \nabla_z g_R(\bar{z})^\top & \nabla_z h(\bar{z})^\top \\ \nabla_z F_I(\bar{z}) & 0 & 0 & 0 \\ -\nabla_z g_R(\bar{z}) & 0 & 0 & 0 \\ \nabla_z h(\bar{z}) & 0 & 0 & 0 \end{pmatrix}$$

with

$$\{(k,i) \mid \bar{\Gamma}_{ki} \neq 0\} \subseteq I \subseteq \{(k,i) \mid F_{ki}(\bar{z}) = 0\} \text{ and } \forall\, k = 1, ..., m\ \exists\, i \in \{1, ..., l\} : (k,i) \in I\,,$$

$$\text{and } \{r \mid \bar{\lambda}_r > 0\} \subseteq R \subseteq \{r \mid g_r(\bar{z}) = 0\}\,,$$

have the same nonvanishing determinantal sign.

Then, there exist neighborhoods Y of $\bar{z}$, U of $\bar{\Gamma}$, V of $\bar{\lambda}$, and W of $\bar{\mu}$, a constant $\alpha > 0$ and a set $\Delta \subset \Omega$ of measure zero, with the following properties: for $n \in \mathbb{N}$ and $\omega \in \Omega$ let

$$\xi_n(\omega) := \|(dL^n(\omega, \bar{z}, \bar{\Gamma}, \bar{\lambda}, \bar{\mu}), (g^n, h^n, F^n)(\omega, \bar{z})) - (\nabla_z L(\bar{z}, \bar{\Gamma}, \bar{\lambda}, \bar{\mu}), (g, h, F)(\bar{z}))\|\,,$$

$$Z_n(\omega) := \{\, (z, \Gamma, \lambda, \mu) \in Y \times U \times V \times W \mid (z, \Gamma, \lambda, \mu) \textit{ weakly stationary for } MPCC_n(\omega)\}\,.$$

For each $\omega \notin \Delta$ there is then a finite integer N_ω such that for every $n \geq N_\omega$ the set $Z_n(\omega)$ is a nonempty, finite subset of $B((\bar{z}, \bar{\Gamma}, \bar{\lambda}, \bar{\mu}), \alpha\, \xi_n(\omega))$.

Theorem 4 is a discrete correspondent to Theorem 12 in Scheel and Scholtes (2000), in the sense that the perturbations given by a continuous parameter in the latter are replaced in the former by sequences of approximations (or alternatively, discrete perturbations). An immediate proof for this theorem is almost identical to the proof of Theorem 1 apart from using Theorem 3 instead of Theorem 5. Alternatively, it may be proven as Theorem 12 in Scheel and Scholtes (2000), with using a discrete correspondent of the existence and uniqueness result for certain nonlinear programs, namely their Theorem 9. As stated in Scheel and Scholtes (2000), Theorem 9 in turn is a version of results of Robinson (1980) and Kojima (1980) (cf. Scholtes (1994)). In this case, the set of stationary points is guaranteed to be finite, which may not be true under the weaker assumptions of Theorem 1.

A discrete correspondent to Theorem 11 in Scheel and Scholtes (2000) can also be stated similarly and proven based on the additional assumption that ULSC holds at the solution point. Again, in this case the local uniqueness of stationary points can be also guaranteed, which may not be the case under the weaker assumptions of Theorem 2.

In summary, the (local) non-uniqueness of the aprroximating solutions in Theorem 1 results both from relaxing the ULSC and from the assumptions in Theorem 5. In Theorem 2, the non-uniqueness relates to the assumptions only on the data functions themselves instead of assumptions involving second order derivatives as suggested by Theorem 3.

Finally, we note that the stability results in Scheel and Scholtes (2000) focus on the piecewise smoothness of the stationary points as functions of continuous perturbations of the deterministic problem, whereas the discrete situations discussed here provide sequential estimates of the approximation errors, which are of main interest when solving the stochastic problem with a simulation-based method, such as sample-path optimization.

5 Summary and conclusions

In this paper we have provided theoretical support for extending the applicability of a simulation-based method, known as sample-path optimization, to an important class of stochastic mathematical programs with complementarity constraints. The type of mathematical programs studied here are suitable for modeling "average" or steady-state behavior of complex stochastic systems. By including unobservable limit functions or expectations in both the objective and the constraints, this class of programs differs from any of the existing stochastic formulations of MPCC's known to us. In particular, we deal with the situation in which the functions do not have an analytic form, but may be approximated/estimated by using simulation. For such problems, we have provided the convergence analysis as well as an extensive discussion on sufficient conditions and regularity concepts involved in using the sample-path method.

The analysis presented in this paper follows a similar strategy as in Robinson (1980) or Gürkan *et al.* (1999a,1999b), in the sense that the existence and convergence of solutions of the approximating problems is proven under a set of sufficient conditions. From this point of view, our work complements the results presented by Patriksson and Wynter (1999) and Lin *et al.* (2003), even though they both deal with different classes of stochastic mathematical programs with complementarity constraints than the one we considered. Patriksson and Wynter (1999) are concerned exclusively about existence issues, whereas

Lin *et al.* (2003) assume the existence and convergence of some approximating solutions and study their asymptotic behavior.

We have striven for a self-contained paper that uses several concepts from the recent work by Scheel and Scholtes (2000) concerning the sensitivity analysis for deterministic mathematical programs with complementarity constraints. Moreover, we worked in a setting involving some weaker conditions for stability in terms of the available higher order information for the problem functions. The approximating/estimating functions in this setting can be also interpreted as slight (non-parametric) perturbations of the true functions. In this respect, we believe that the kind of results we have presented here can be seen as complementary to the sensitivity results in Scheel and Scholtes (2000).

In order not to overburden, the present paper does not illustrate the discussed methodology on an application. In the accompanying paper Birbil *et al.* (2004), which targets mainly the practitioners who would be interested in using the on-going methodology for specific applications, we provide details on several practical issues. There we discuss in further detail our SMPCC formulation, thoroughly compare it with other stochastic MPEC models from the literature, and outline a generic algorithm to implement the sample-path method for SMPCC's. Moreover we illustrate how uncertainty may be modeled in two applications: taxation in natural gas market and toll pricing in traffic networks. We then present numerical results on the latter problem by using the current state-of-the-art software for deterministic MPCC's.

In addition, in Appendix 1 we have shown that the nonvanishing determinantal sign condition guarantees coherent orientation. We established this result by proving the equivalence between the nonvanishing determinantal sign condition and the strong regularity conditions given by Robinson (1980). However, our algebraic proof does not give a clear geometrical meaning to the condition itself. A direct link between the nonvanishing determinantal sign condition and the coherent orientation could provide some interesting additional insight.

Appendix 1: Regularity conditions for generalized equations and nonlinear programs

For a closed convex set $C \subseteq \mathbb{R}^k$ and a function f from an open set $\Omega \subseteq \mathbb{R}^k$ to $\mathbb{R}^k$, the variational inequality determined by f and C can be expressed in an equivalent way as a generalized equation; that is, an inclusion of the form

$$0 \in f(x) + N_C(x), \tag{27}$$

where $N_C(x)$ denotes the normal cone of C at x:

$$N_C(x) := \begin{cases} \{\, y \in \mathbb{R}^k \mid y(c-x) \leq 0 \text{ for all } c \in C \,\} & \text{if } x \in C, \\ \emptyset & \text{if } x \notin C. \end{cases}$$

Now suppose that x_0 is a solution of (27) and that f is Fréchet differentiable at x_0. Then the generalized equation (27) is called *strongly regular* at x_0 if there are neighborhoods X of x_0 and Y of the origin in $\mathbb{R}^k$ with the property that the *linearized* generalized equation

$$y \in f(x_0) + \nabla f(x_0)(x - x_0) + N_C(x) \tag{28}$$

defines a single-valued, Lipschitzian map $x(y)$ from Y to X, such that for each $y \in Y$, $x(y)$ is the unique solution in X of (28). The property of strong regularity was introduced by Robinson (1980) and its equivalence to another well-known property called *coherent orientation* was established in the appendix of Gürkan *et al.* (1999a).

Robinson (1980) showed how strong regularity can be applied to the standard nonlinear programming problem

$$\begin{array}{rl} \min & f(x) \\ \text{s.t.} & g(x) \leq 0 \\ & h(x) = 0\,, \end{array} \tag{29}$$

where f, g, and h are twice Fréchet differentiable functions from some open set $\Theta \subseteq \mathbb{R}^n$ into $\mathbb{R}$, $\mathbb{R}^p$, and $\mathbb{R}^q$, respectively. We summarize these findings as follows. Let $L(x,u,v) := f(x) + u^\top g(x) + v^\top h(x)$ denote the Lagrangian with multipliers $u \in \mathbb{R}^p$ and $v \in \mathbb{R}^q$ associated with the problem (29). The first order necessary optimality conditions for (29) are

$$\begin{aligned} \nabla_x L(x,u,v) &= 0 \\ g(x) &\leq 0 \\ h(x) &= 0 \\ u &\geq 0 \\ u^\top g(x) &= 0\,. \end{aligned} \tag{30}$$

These conditions can be written as the following generalized equation

$$0 \in \begin{bmatrix} \nabla_x L(z,u,v) \\ -g(x) \\ h(x) \end{bmatrix} + N_{\mathbb{R}^n \times \mathbb{R}^p_+ \times \mathbb{R}^q} \begin{bmatrix} x \\ u \\ v \end{bmatrix}. \tag{31}$$

Suppose that $(\bar{x}, \bar{u}, \bar{v})$ is a solution of (31). Then we can partition the vectors $g(\bar{x})$ and $\bar{u}$ into some smaller vectors corresponding to the following index sets:

$$\begin{aligned} I^+ &= \{\, i \mid g_i(\bar{x}) = 0, \quad \bar{u} > 0 \,\}, \\ I^0 &= \{\, i \mid g_i(\bar{x}) = 0, \quad \bar{u} = 0 \,\}, \\ I^- &= \{\, i \mid g_i(\bar{x}) < 0, \quad \bar{u} = 0 \,\}. \end{aligned}$$

Thus, we denote by g_{I^+}, g_{I^0}, and g_{I^-} the partial functions constructed with the components of g corresponding to I^+, I^0, and I^-, respectively. Given this notation, Robinson (1980) showed that the generalized equation (31) is strongly regular at the solution point $(\bar{x}, \bar{u}, \bar{v})$ if and only if the matrix

$$\begin{bmatrix} \nabla^2_{xx} L(\bar{x},\bar{u},\bar{v}) & \nabla_x h(\bar{x})^\top & \nabla_x g_{I^+}(\bar{x})^\top \\ \nabla_x h(\bar{x}) & 0 & 0 \\ -\nabla_x g_{I^+}(\bar{x}) & 0 & 0 \end{bmatrix} \tag{32}$$

is nonsingular and the matrix

$$\begin{bmatrix} \nabla_x g_{I^0}(\bar{x}) & 0 & 0 \end{bmatrix} \begin{bmatrix} \nabla^2_{xx} L(\bar{x},\bar{u},\bar{v}) & \nabla_x h(\bar{x})^\top & \nabla_x g_{I^+}(\bar{x})^\top \\ \nabla_x h(\bar{x}) & 0 & 0 \\ -\nabla_x g_{I^+}(\bar{x}) & 0 & 0 \end{bmatrix}^{-1} \begin{bmatrix} \nabla_x g_{I^0}(\bar{x})^\top \\ 0 \\ 0 \end{bmatrix} \tag{33}$$

has all its principal minors positive (i.e. it is a P-matrix). Clearly, when I^0 is vacuous, then the strong regularity condition reduces to the nonsingularity of the matrix (32).

In the sequel, we assume that I^0 is not vacuous and we show that the two conditions above are equivalent with the following condition used by Scheel and Scholtes (2000): for all index sets I with

$$\{\, i \mid \bar{u}_i > 0 \,\} = I^+ \subseteq I \subseteq I^+ \cup I^0 = \{\, i \mid g_i(\bar{x}) = 0 \,\},$$

the matrices

$$\begin{bmatrix} \nabla^2_{xx} L(\bar{x}, \bar{u}, \bar{v}) & \nabla_x h(\bar{x})^\top & \nabla_x g_I(\bar{x})^\top \\ \nabla_x h(\bar{x}) & 0 & 0 \\ -\nabla_x g_I(\bar{x}) & 0 & 0 \end{bmatrix} \tag{34}$$

have the same nonvanishing determinantal sign. This equivalence follows immediately from the following elementary results from linear algebra.

Lemma 2 *Let m and r be positive integers and A be an $(m+r) \times (m+r)$ matrix of the form*

$$A = \begin{bmatrix} X & Y^\top \\ -Y & 0 \end{bmatrix},$$

where X is a nonsingular $m \times m$ matrix. Then $\det(A) = \det(X) \det(Y X^{-1} Y^\top)$.

The formula in Lemma 2 is a special case of a more general algebraic result; see for example page 46 of Lancaster and Tismenetsky (1985).

Proposition 1 *Let X be an $m \times m$ matrix and Y be an $r \times m$ matrix. Then the following two statements are equivalent:*

1) X is nonsingular and $Y X^{-1} Y^\top$ has all principal minors positive.

2) All the matrices

$$A_R = \begin{bmatrix} X & Y_R^\top \\ -Y_R & 0 \end{bmatrix}$$

with $R \subseteq \{1, ..., r\}$ and Y_R the sub-matrix of Y formed with the rows in R, have the same nonvanishing determinantal sign.

Proof. Notice that if we take $R = \emptyset$ in 2), then $Y_\emptyset$ is vacuous and hence $X = A_\emptyset$. Thus, both 1) and 2) assume that X is nonsingular. Now let $R \subseteq \{1, ..., r\}$, $R \neq \emptyset$ be an arbitrary index set. Then, according to Lemma 2, we have that

$$\det(A_R) = \det(X) \det(Y_R X^{-1} Y_R^\top).$$

But $Y_R X^{-1} Y_R^\top$ is the principal minor of the matrix $Y X^{-1} Y^\top$ determined by rows i and columns i with $i \in R$. Given the determinantal relation in Lemma 2, it is straightforward that $\det(Y_R X^{-1} Y_R^\top) > 0$ if and only if $\det(A_R)$ and $\det(X)$ have the same nonvanishing sign. As this holds for an arbitrary R, the equivalence between 1) and 2) follows then immediately. ∎

The equivalence of the two regularity conditions can be seen immediately by applying Proposition 1 with X being the matrix (32) and $Y = \begin{bmatrix} \nabla_x g_{I^0}(\bar{x}) & 0 & 0 \end{bmatrix}$.

This algebraic proof provides a short way of bridging the two technical conditions for regularity. However, it does not provide any insight into the geometrical interpretation of strong regularity. To this end, we note that the nonvanishing determinantal sign condition is in fact reminiscent of the coherent orientation condition, which has by definition a clear geometrical meaning. For a thorough discussion on coherent orientation in the general setting as well as its equivalence with strong regularity we refer again to the appendix of Gürkan *et al.* (1999a).

Appendix 2: Sample-path solution of stochastic variational inequalities

In order to make this paper self-contained and to ease the reference in Section 3, we restate here Theorem 2 from Gürkan *et al.* (1999a).

Theorem 5 *Let Θ be an open subset of $\mathbb{R}^k$ and let C be a polyhedral convex set in $\mathbb{R}^k$. Let x_0 be a point of Θ, and suppose f is a function from Θ to $\mathbb{R}^k$. Let $\{f_n \mid n = 1, 2, \dots\}$ be random functions from Θ to $\mathbb{R}^k$ such that for all $x \in \Theta$ and all finite n the random variables $f_n(x)$ are defined on a common probability space $(\Omega, \mathcal{F}, P)$. Let $z_0 = x_0 - f(x_0)$ and assume the following:*

1) With probability one, each f_n for $n = 1, 2, \dots$ is continuous and $f_n \xrightarrow{\mathcal{C}} f$.

2) x_0 solves the variational inequality defined by f and C.

3) f has a strong Fréchet derivative $df(x_0)$ at x_0, and $df(x_0)_{K(z_0)}$ is coherently oriented.

Then the restriction of f_C to a neighborhood of z_0 has an inverse that is Lipschitzian with some modulus λ. Further, there exist a compact subset $C_0 \subset C \cap \Theta$ containing x_0, a neighborhood $\Theta_1 \subset \Theta$ of x_0, a scalar $\beta > 0$, and a set $\Delta \subset \Omega$ of measure zero, with the following properties: for $n = 1, 2, \dots$, $\omega \in \Omega$, and $y \in \mathbb{R}^k$ with $\|y\| \leq \beta$, let

$$\xi_n(\omega) = \sup_{x \in C_0} \|f_n(\omega, x) - f(x)\|,$$

and

$$X_n(\omega, y) := \{x \in C \cap \Theta_1 \mid \text{for each } c \in C,\ (f_n(\omega, x) - y)(c - x) \geq 0\}.$$

For each $\omega \notin \Delta$ there is then a finite integer N_ω such that for each $n \geq N_\omega$ the set $X_n(\omega, y)$ is a nonempty, compact subset of $B(x_0, \lambda(\xi_n(\omega) + \|y\|))$.

References

[1] J. F. Bard. Practical Bilevel Optimization. Kluwer Academic Publishers, Dordrecht, 1998.

[2] Ş. İ. Birbil, G. Gürkan, and O. Listeş. Simulation-based solution of stochastic mathematical programs with complementarity constraints: Applications. Working paper, Tilburg University and Erasmus University Rotterdam, The Netherlands, 2004.

[3] S. Dempe. Foundations of Bilevel Programming. Kluwer Academic Publishers, Dordrecht, 2002.

[4] M. C. Ferris and J. S. Pang. Engineering and economic applications of complementarity problems. *SIAM Review*, 39:669–713, 1997.

[5] P. Glasserman. Gradient Estimation via Perturbation Analysis. Kluwer, Norwell, MA., USA, 1991.

[6] G. Gürkan, A. Y. Özge, and S. M. Robinson. Sample-path solution of stochastic variational inequalities, with applications to option pricing. In J. M. Charnes, D. M. Morrice, D. M. Brunner, and J. J. Swain, editors, Proceedings of the 1996 Winter Simulation Conference, pages 337–344, 1996.

[7] G. Gürkan, A. Y. Özge, and S. M. Robinson. Sample-path solution of stochastic variational inequalities. *Mathematical Programming*, 84:313–333, 1999a.

[8] G. Gürkan, A. Y. Özge, and S. M. Robinson. Solving stochastic optimization problems with stochastic constraints: An application in network design. In D. T. Sturrock, P. A. Farrington, H. B. Nembhard, and G. W. Evans, editors, Proceedings of the 1999 Winter Simulation Conference, pages 471–478, 1999b.

[9] P. T. Harker and J. S. Pang. Finite-dimensional variational inequality and nonlinear complementarity problem: A survey of theory, algorithms and applications. *Mathematical Programming*, 48:161–220, 1990.

[10] Y. C. Ho and X. R. Cao. Perturbation Analysis of Discrete Event Dynamical Systems. Kluwer, Norwell, MA., USA, 1991.

[11] P. Kall. Approximation to optimization problems: An elementary review. *Mathematics of Operations Research*, 11:9–18, 1986.

[12] A. J. Kleywegt, A. Shapiro, and T. Homem-De-Mello. The sample average approximation method for stochastic discrete optimization. *SIAM Journal of Optimization*, 12:479–502, 2001.

[13] P. Lancaster and M. Tismenetsky. The Theory of Matrices with Applications. Computer Science and Applied Mathematics. Academic Press, San Diego, California, U.S., 1985.

[14] G. H. Lin, X. Chen, and M. Fukushima. Smoothing implicit programming approaches for stochastic mathematical programs with linear complementarity constraints. Technical Report, Department of Applied Mathematics and Physics, Kyoto University, Japan, 2003.

[15] J. T. Linderoth, A. Shapiro, and S. J. Wright. The empirical behavior of sampling methods for stochastic programming. Optimization Technical Report 02-01, Computer Sciences Department, University of Wisconsin-Madison, Madison, WI, USA, 2002.

[16] Z. Q. Luo, J. S. Pang, and D. Ralph. Mathematical Programs with Equilibrium Constraints. Cambridge University Press, Cambridge, U.K., 1996.

[17] J. R. Munkres. Topology: A First Course. Prentice Hall, Englewood Cliffs, New Jersey, 1975.

[18] J. V. Outrata, M. Kocvara, and J. Zowe. Nonsmooth Approach to Optimization Problems with Equilibrium Constraints: Theory, Applications and Numerical Results. Kluwer Academic Publishers, Boston, MA, 1998.

[19] M. Patriksson and L. Wynter. Stochastic mathematical programs with equilibrium constraints. *Operations Research Letters*, 25:159–167, 1999.

[20] E. L. Plambeck, B. R. Fu, S. M. Robinson, and R. Suri. Throughput optimization in tandem production lines via nonsmooth programming. In J. Schoen, editor, *Proceedings of 1993 Summer Computer Simulation Conference*, pages 70–75. San Diego, CA: Society for Computer Simulation, 1993.

[21] E. L. Plambeck, B. R. Fu, S. M. Robinson, and R. Suri. Sample-path optimization of convex stochastic performance functions. *Mathematical Programming*, 75:137–176, 1996.

[22] S. M. Robinson. Generalized equations and their solutions, Part I: Basic theory. *Mathematical Programming Study*, 10:128–141, 1979.

[23] S. M. Robinson. Strongly regular generalized equations. *Mathematics of Operations Research*, 5:43–62, 1980.

[24] S. M. Robinson. Analysis of sample-path optimization. *Mathematics of Operations Research*, 21(3):513–528, 1996.

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

[26] R. Y. Rubinstein and A. Shapiro. Discrete Event Systems: Sensitivity Analysis and Stochastic Optimization by the Score Function Method. Wiley, Chichester and New York, 1993.

[27] H. Scheel. Ein Straffunktionsansatz für Optimierungsprobleme mit Gleichgewichtsrestriktionen, Diploma Thesis. Institut für Statistik und Mathematische Wirtschaftstheorie, Universität Karlsruhe, 76128 Karlsruhe, Germany, 1995.

[28] H. Scheel and S. Scholtes. Mathematical programs with complementarity constraints: stationarity, optimality and sensitivity. *Mathematics of Operations Research*, 25(1):1–22, 2000.

[29] S. Scholtes. Introduction to piecewise differentiable equations, Habilitation Thesis. Institut für Statistik und Mathematische Wirtschaftstheorie, Universität Karlsruhe, 76128 Karlsruhe, Germany, 1994.

[30] A. Shapiro and T. Homem-De-Mello. A simulation-based approach to two-stage stochastic programming with recourse. *Mathematical Programming*, 81:301–325, 1998.

[31] G. Still. Generalized semi-infinite programming: theory and methods. *European Journal of Operational Research*, 119:301–303, 1999.