COMBINED SMOOTHING IMPLICIT PROGRAMMING AND PENALTY METHOD FOR STOCHASTIC MATHEMATICAL PROGRAMS WITH EQUILIBRIUM CONSTRAINTS [*]

Gui-Hua Lin[†], Xiaojun Chen[‡] and Masao Fukushima[§]

March 15, 2004

Abstract In this paper, we consider the stochastic mathematical program with equilibrium constraints (SMPEC), which can be thought as a generalization of the mathematical program with equilibrium constraints. Many decision problems can be formulated as SMPECs in practice. We discuss both here-and-now and lower-level wait-and-see decision problems. In particular, with the help of a penalty technique, we propose a combined smoothing implicit programming and penalty method for the here-and-now decision problem and a comprehensive convergence theory is also included. Furthermore, we remark that similar discussion applies to the lower-level wait-and-see model as well.

Key words. stochastic mathematical program with equilibrium constraints, implicit programming, complementarity, wait-and-see, here-and-now, quasi-Monte Carlo method.

1 Introduction

Mathematical program with equilibrium constraints (MPEC) is a constrained optimization problem in which the essential constraints are defined by a parametric variational inequality:

$$\begin{array}{ll} \text{minimize} & f(x,y) \\ \text{subject to} & x \in X, \\ & y \text{ solves } \mathrm{VI}(F(x,\cdot), C(x)). \end{array} \tag{1.1}$$

Here, X is a subset of $\Re^n$, $f : \Re^{n+m} \to \Re, F : \Re^{n+m} \to \Re^m$, $C : \Re^n \to 2^{\Re^m}$ are mappings, and $\mathrm{VI}(F(x,\cdot), C(x))$ denotes the variational inequality problem defined by the pair $(F(x,\cdot), C(x))$, i.e., y solves $\mathrm{VI}(F(x,\cdot), C(x))$ if and only if $y \in C(x)$ and

$$(v-y)^T F(x,y) \ge 0 \quad \forall v \in C(x).$$

Problem (1.1) can be regarded as a generalization of a bilevel programming problem and it therefore plays an important role in many fields such as engineering design, economic equilibrium, multilevel game, and mathematical programming itself. In the recent optimization study,

[*]This work was supported in part by the Scientific Research Grant-in-Aid from the Ministry of Education, Science, Sports, and Culture of Japan.

[†]Department of Applied Mathematics, Dalian University of Technology, Dalian 116024, China. Current address: Department of Applied Mathematics and Physics, Graduate School of Informatics, Kyoto University, Kyoto 606-8501, Japan. E-mail: ghlin@amp.i.kyoto-u.ac.jp.

[‡]Department of Mathematical System Science, Hirosaki University, Hirosaki 036-8560, Japan. E-mail: chen@cc.hirosaki-u.ac.jp.

[§]Department of Applied Mathematics and Physics, Graduate School of Informatics, Kyoto University, Kyoto 606-8501, Japan. E-mail: fuku@amp.i.kyoto-u.ac.jp.

MPECs have been receiving much attention, see the monograph of Luo et al. [11] and the attached references.

Stochastic programming is another important branch of mathematical programming that deals with problems in which optimal decisions are sought under uncertainty. Modelling the uncertainty by random objects may lead to diverse stochastic programming problems, a special case of which is the so-called two-stage stochastic program with recourse:

$$\begin{array}{ll} \text{minimize} & p(x) + E_\omega[Q(x,\omega)] \\ \text{subject to} & x \in X, \end{array} \tag{1.2}$$

where $p : \Re^n \to \Re$, E_ω means expectation with respect to the random variable $\omega \in \Omega$, and

$$Q(x,\omega) := \inf_{y \in Y(x,\omega)} g(y,\omega)$$

with $Y : \Re^n \times \Omega \to 2^{\Re^m}$ and $g : \Re^m \times \Omega \to \Re$. Problem (1.2) minimizes the sum of the cost of the master decision and the expected cost of the recourse decision, where "recourse" means the ability to take corrective action after random events have taken place. Many applications of such problems can be found in practice, especially in financial planning. For further details, see [2].

This paper deals with stochastic mathematical programs with equilibrium constraints (SMPECs). Our purpose is two-fold. The first is to give two formulations of SMPECs formally. The second is to present an approach for solving SMPECs and study its convergence properties.

The bilevel nature of MPECs allows the uncertainty to enter at different levels. In our first formulation, only the upper-level decision is made under an uncertain circumstance, and the lower-level decision is made after the random event ω is observed. This results in the following problem, which we call the *lower-level wait-and-see* model:

$$\begin{array}{ll} \text{minimize} & E_\omega[f(x,y(\omega),\omega)] \\ \text{subject to} & x \in X, \\ & y(\omega) \text{ solves } \mathrm{VI}(F(x,\cdot,\omega), C(x,\omega)) \quad \forall \omega \in \Omega, \end{array} \tag{1.3}$$

where $X \subseteq \Re^n, f : \Re^{n+m} \times \Omega \to \Re, F : \Re^{n+m} \times \Omega \to \Re^m$, and $C : \Re^n \times \Omega \to 2^{\Re^m}$. This type of SMPEC was studied by Patriksson and Wynter [15], in which the existence of solutions, the convexity and directional differentiability of an implicit objective function, and links between (1.3) and bilevel models have been investigated. Note that the wait-and-see model [17] in the classical stochstic programming study is not an optimization problem. However, the lower-level wait-and-see model (1.3) of SMPEC is an optimization problem in which essential variables consist of the upper-level decision x.

When $C(x,\omega) \equiv \Re^m_+$ for any $x \in X$ and any $\omega \in \Omega$ in problem (1.3), the variational inequality constraints reduce to the complementarity constraints and problem (1.3) is equivalent to the following stochastic mathematical program with complementarity constraints (SMPCC):

$$\begin{array}{ll} \text{minimize} & E_\omega[f(x,y(\omega),\omega)] \\ \text{subject to} & x \in X, \\ & y(\omega) \ge 0,\ F(x,y(\omega),\omega) \ge 0, \\ & y(\omega)^T F(x,y(\omega),\omega) = 0 \quad \forall \omega \in \Omega. \end{array} \tag{1.4}$$

On the other hand, if the set-valued function C in problem (1.3) is defined by

$$C(x,\omega) := \{y \in \Re^m | \;\; c(x,y,\omega) \le 0\},$$

where $c(\cdot,\cdot,\omega)$ is continuously differentiable, then, under some suitable conditions, the variational inequality problem VI$(F(x,\cdot,\omega), C(x,\omega))$ has an equivalent Karush-Kuhn-Tucker representation

$$\begin{aligned} &F(x,y(\omega),\omega) + \nabla_y c(x,y(\omega),\omega)\lambda(x,\omega) = 0, \\ &\lambda(x,\omega) \ge 0, \quad c(x,y(\omega),\omega) \le 0, \quad \lambda(x,\omega)^T c(x,y(\omega),\omega) = 0, \end{aligned}$$

where $\lambda(x,\omega)$ is the Lagrange multiplier vector [14]. As a result, problem (1.3) can be reformulated as a program like (1.4) under some conditions, see [11] for more details. Hence, problem (1.4) constitutes an important subclass of SMPECs.

Another formulation that we are particularly interested in is the following problem that requires us to make all decisions at once, before ω is observed:

$$\begin{array}{ll} \text{minimize} & E_\omega[f(x,y,\omega) + d^T z(\omega)] \\ \text{subject to} & x \in X, \\ & y \ge 0, \;\; F(x,y,\omega) + z(\omega) \ge 0, \\ & y^T(F(x,y,\omega) + z(\omega)) = 0, \\ & z(\omega) \ge 0 \quad \forall \omega \in \Omega. \end{array} \tag{1.5}$$

Here, both the decisions x and y are independent of the random variable ω, $z(\omega)$ is called a recourse variable, and $d \in \Re^m$ is a vector with positive elements. We call (1.5) a *here-and-now* model. Compared with the lower-level wait-and-see model (1.4), the here-and-now model (1.5) involves more variables and hence seems more difficult to deal with. Moreover, a feasible vector y in (1.5) is required to satisfy the complementarity condition for all $\omega \in \Omega$, which is different from the ordinary complementarity condition if Ω has more than one realization. Because of this restriction, some results for MPECs cannot be applied to (1.5) directly. Special new treatment has to be developed. In this paper, we will mainly be concerned with the here-and-now model (1.5), as the obtained results may be applied to the model (1.4) by suitable modification.

The following example illustrates the two models.

Example 1.1 There are a food company who makes picnic lunches and a vendor who sells lunches to hikers on every Sunday. The company and the vendor have the following contract:

C1: The vendor buys lunches from the company at the price $x \in [a,b]$ determined by the company, where a and b are two positive constants.

C2: The vendor decides the amount y of lunches that he buys from the company, where y must be no less than the minimum amount $c > 0$.

C3: The vendor pays the company for the whole lunches he buys, i.e., the vendor pays xy to the company.

C4: The vendor sells lunches to hikers at the price $2x$ and get the proceeds for the total number of lunches actually sold.

C5: Even if there are any unsold lunches, the vendor cannot return them to the company but he can dispose of the unsold lunches with no cost.

We suppose that the demand of lunches depends on the price and the weather on that day. Since the weather is uncertain, we may treat it as a random variable. More specifically, we suppose that the demand is given by the function

$$\phi(x,\omega) := D(\omega) - d(\omega)x, \quad \omega \in \Omega,$$

where $D(\omega) \geq 0$ and $d(\omega) \geq 0$ are random variables. Therefore, the actual amount of lunches sold is given by $\min(y, \phi(x,\omega))$, which also depends on the weather on that day.

The decisions by the company and the vendor are x and y, respectively. The company's objective is to maximize its total earnings xy, while the vendor's objective is to maximize its total earnings $2x\min(y,\phi(x,\omega)) - xy$. The latter problem may be written as

$$\begin{array}{ll} \text{maximize}_{y,t} & x(2t-y) \\ \text{subject to} & y \geq c, \; y - t \geq 0, \\ & D(\omega) - d(\omega)x - t \geq 0, \end{array}$$

whose optimality conditions are stated as

$$\begin{pmatrix} x \\ -2x \end{pmatrix} - u\begin{pmatrix} 1 \\ 0 \end{pmatrix} - v\begin{pmatrix} 1 \\ -1 \end{pmatrix} - w\begin{pmatrix} 0 \\ -1 \end{pmatrix} = 0, \tag{1.6}$$

$$0 \leq u \perp (y-c) \geq 0,$$

$$0 \leq v \perp (y-t) \geq 0,$$

$$0 \leq w \perp (D(\omega) - d(\omega)x - t) \geq 0. \tag{1.7}$$

Here, $\lambda \perp \mu$ means $\lambda\mu = 0$. It follows from (1.6) that

$$u = x - v, \quad w = 2x - v.$$

This implies that $w = x + u \geq a > 0$, which together with (1.7) yields $t = D(\omega) - d(\omega)x$. Thus the above optimality conditions may further be rewritten as

$$\begin{array}{l} 0 \leq (x-v) \perp (y-c) \geq 0, \\ 0 \leq v \perp (y - D(\omega) + d(\omega)x) \geq 0. \end{array} \tag{1.8}$$

Then the company's problem may be written as the following stochastic MPEC:

$$\begin{array}{ll} \text{minimize} & -xy \\ \text{subject to} & a \leq x \leq b, \quad \omega \in \Omega, \\ & 0 \leq (x-v) \perp (y-c) \geq 0, \\ & 0 \leq v \perp (y - D(\omega) + d(\omega)x) \geq 0. \end{array}$$

Now there are two cases.

Here-and-now model: Suppose that both the company and the vendor have to make decision on Saturday, without knowing the weather of Sunday. In this case, there is no (x,v) satisfying (1.8) for all $\omega \in \Omega$ in general. So, by introducing the recourse variables, the company's problem is represented as the following model:

$$\begin{array}{ll} \text{minimize} & -xy + \beta E_\omega[\, z(\omega)\,] \\ \text{subject to} & a \leq x \leq b, \\ & 0 \leq (x-v) \perp (y-c) \geq 0, \\ & 0 \leq v \perp (y - D(\omega) + d(\omega)x + z(\omega)) \geq 0, \\ & z(\omega) \geq 0 \quad \forall \omega \in \Omega, \end{array}$$

where $\beta > 0$ is a constant.

Lower-level wait-and-see model: Suppose that the company makes a decision on Saturday, but the vendor can make a decision on Sunday morning after knowing the weather of that day. In this case, the vendor's decision may depend on the observation of ω, which is given by $(y(\omega), v(\omega))$ that satisfies

$$\begin{aligned} &0 \leq (x - v(\omega)) \perp (y(\omega) - c) \geq 0, \\ &0 \leq v(\omega) \perp (y(\omega) - D(\omega) + d(\omega)x) \geq 0 \end{aligned}$$

for each $\omega \in \Omega$. Therefore the company's problem is represented as the following model:

$$\begin{array}{ll} \text{minimize} & E_\Omega[\,-y(\omega)x\,] \\ \text{subject to} & a \leq x \leq b, \\ & 0 \leq (x - v(\omega)) \perp (y(\omega) - c) \geq 0, \\ & 0 \leq v(\omega) \perp (y(\omega) - D(\omega) + d(\omega)x) \geq 0 \quad \forall \omega \in \Omega. \end{array}$$

Organization of the paper: In Section 3, we apply a penalty technique and present a smoothing implicit programming method for the discrete here-and-now problem and, in Section 4, we employ a quasi-Monte Carlo method for numerical integration to discretize the here-and-now problem with a continuous random variable. Comprehensive convergence theory is also included. In Section 5, we make some remarks to conclude the paper. Especially, we mention that the proposed approach for here-and-now models can be extended to the lower-level wait-and-see problems.

Notation used in the paper: Throughout, all vectors are thought as column vectors and $x[i]$ stands for the ith coordinate of $x \in \Re^n$, whereas for a matrix M, we denote by $M[i]$ the vector whose elements consist of the ith row of M. If $\mathcal{K}$ is an index set, we let $M[\mathcal{K}]$ be the principal submatrix of M whose elements consist of those of M indexed by $\mathcal{K}$. For any vectors u and v of the same dimension, we denote $u \perp v$ to mean $u^T v = 0$. For a given function $F : \Re^n \to \Re^m$ and a vector $x \in \Re^n$, $\nabla F(x)$ is the transposed Jacobian of F at x and $\mathcal{I}_F(x) := \{i \mid F_i(x) = 0\}$ stands for the active index set of F at x. In addition, e_i denotes the unit vector with $e_i[i] = 1$; I and O denote the identity matrix and the zero matrix with suitable dimension, respectively.

2 Preliminaries

In this section, we recall some basic concepts and properties that will be used later on. First we consider the standard smooth nonlinear programming problem:

$$\begin{array}{ll} \text{minimize} & f(z) \\ \text{subject to} & c_i(z) \leq 0, \quad i = 1, \cdots, t, \\ & c_i(z) = 0, \quad i = t+1, \cdots, \nu. \end{array} \tag{2.1}$$

We will use the standard definition of stationarity, i.e., a feasible point z is said to be *stationary* to (2.1) if there exists a Lagrange multiplier vector $\lambda \in \Re^\nu$ satisfying the Karush-Kuhn-Tucker conditions

$$\begin{aligned} &\nabla f(z) + \nabla c(z)\lambda = 0, \\ &\lambda[i] \geq 0, \ \lambda[i]c_i(z) = 0, \qquad i = 1, \cdots, t. \end{aligned}$$

We next consider the mathematical program with complementarity constraints:

$$\begin{array}{ll} \text{minimize} & f(z) \\ \text{subject to} & g(z) \le 0,\ h(z) = 0, \\ & G(z) \ge 0,\ H(z) \ge 0, \\ & G(z)^T H(z) = 0, \end{array} \tag{2.2}$$

where $f : \Re^s \to \Re, g : \Re^s \to \Re^p, h : \Re^s \to \Re^q$, and $G, H : \Re^s \to \Re^t$ are all continuously differentiable functions. Let $\mathcal{Z}$ denote the feasible region of the MPEC (2.2).

It is well-known that the MPEC (2.2) fails to satisfy a standard constraint qualification (CQ) at any feasible point [4], which causes a difficulty in dealing with MPECs by a conventional nonlinear programming approach. The following special CQ turns out to be useful in the study of MPECs.

Definition 2.1 The MPEC-*linear independence constraint qualification* (MPEC-LICQ) is said to hold at $\bar{z} \in \mathcal{Z}$ if the set of vectors

$$\left\{ \nabla g_l(\bar{z}), \nabla h_r(\bar{z}), \nabla G_i(\bar{z}), \nabla H_j(\bar{z}) \;\middle|\; l \in \mathcal{I}_g(\bar{z}),\ r = 1, \cdots, q,\ i \in \mathcal{I}_G(\bar{z}),\ j \in \mathcal{I}_H(\bar{z}) \right\}$$

is linearly independent.

Definition 2.2 [16] (1) $\bar{z} \in \mathcal{Z}$ is called a *Clarke or C-stationary* point of problem (2.2) if there exist multiplier vectors $\bar{\lambda} \in \Re^p, \bar{\mu} \in \Re^q$, and $\bar{u}, \bar{v} \in \Re^t$ such that $\bar{\lambda} \ge 0$ and

$$\nabla f(\bar{z}) + \sum_{i \in \mathcal{I}_g(\bar{z})} \bar{\lambda}[i] \nabla g_i(\bar{z}) + \sum_{i=1}^{q} \bar{\mu}[i] \nabla h_i(\bar{z}) - \sum_{i \in \mathcal{I}_G(\bar{z})} \bar{u}[i] \nabla G_i(\bar{z}) - \sum_{i \in \mathcal{I}_H(\bar{z})} \bar{v}[i] \nabla H_i(\bar{z}) = 0, \tag{2.3}$$

$$\bar{u}[i]\bar{v}[i] \ge 0, \quad i \in \mathcal{I}_G(\bar{z}) \cap \mathcal{I}_H(\bar{z}). \tag{2.4}$$

(2) $\bar{z} \in \mathcal{Z}$ is called a *strongly or S-stationary* point of problem (2.2) if there exist multiplier vectors $\bar{\lambda}, \bar{\mu}, \bar{u}$, and $\bar{v}$ such that (2.3) holds with

$$\bar{u}_i \ge 0, \quad \bar{v}_i \ge 0, \quad i \in \mathcal{I}_G(\bar{z}) \cap \mathcal{I}_H(\bar{z}).$$

It is easy to see that S-stationarity implies C-stationarity. Moreover, under the *strict complementarity condition* (namely, $\mathcal{I}_G(\bar{z}) \cap \mathcal{I}_H(\bar{z}) = \emptyset$), they are equivalent.

Definition 2.3 [5] Suppose that M is an $m \times m$ matrix. We call M a *P-matrix* if all the principal minors of M are positive, or equivalently,

$$\max_{1 \le i \le m} y[i](My)[i] > 0, \qquad 0 \ne \forall y \in \Re^m,$$

and we call M a P_0-*matrix* if all the principal minors of M are nonnegative, or equivalently,

$$\max_{1 \le i \le m} y[i](My)[i] \ge 0, \qquad \forall y \in \Re^m.$$

It is obvious that a P-matrix must be a P_0-matrix. However, the converse does not hold. In addition, if M is a P_0-matrix and μ is a positive number, then the matrix $M + \mu I$ is a P-matrix.

Definition 2.4 [5] A square matrix is said to be *nondegenerate* if all of its principal submatrices are nonsingular.

It is easy to see that a P-matrix is nondegenerate.

For given $N \in \Re^{m\times n}$, $M \in \Re^{m\times m}$, $q \in \Re^m$, and two positive numbers ϵ and μ, we define the function

$$\Phi_{\epsilon,\mu}(x, y, w; N, M, q) := \begin{pmatrix} Nx + (M + \epsilon I)y + q - w \\ \phi_\mu(y[1], w[1]) \\ \vdots \\ \phi_\mu(y[m], w[m]) \end{pmatrix}, \tag{2.5}$$

where $\phi_\mu : \Re^2 \to \Re$ is the perturbed Fischer-Burmeister function

$$\phi_\mu(a, b) := a + b - \sqrt{a^2 + b^2 + 2\mu^2}\ .$$

Then we have the following well-known result [3, 9].

Theorem 2.1 *Suppose that M is a P_0-matrix. Then, for given $x \in \Re^n$, $\epsilon > 0$, and $\mu > 0$, we have the following statements:*

(i) *The function $\Phi_{\epsilon,\mu}$ defined by (2.5) is continuously differentiable with respect to (y, w) and the Jacobian matrix $\nabla_{(y,w)}\Phi_{\epsilon,\mu}(x, y, w; N, M, q)$ is nonsingular everywhere;*

(ii) *The equation $\Phi_{\epsilon,\mu}(x, y, w; N, M, q) = 0$ has a unique solution $(y(x, \epsilon, \mu), w(x, \epsilon, \mu))$, which is continuously differentiable with respect to x and satisfies*

$$\begin{aligned} &y(x, \epsilon, \mu) > 0, \quad w(x, \epsilon, \mu) > 0, \\ &y(x, \epsilon, \mu)[i]w(x, \epsilon, \mu)[i] = \mu^2, \quad i = 1, \cdots, m. \end{aligned}$$

In the rest of the paper, to mitigate the notational complication, we assume $\epsilon = \mu$ and denote $\Phi_{\epsilon,\mu}$, $y(x, \epsilon, \mu)$, and $w(x, \epsilon, \mu)$ by Φ_μ, $y(x, \mu)$, and $w(x, \mu)$, respectively. Our analysis will remain valid, however, even though the two parameters are treated independently.

Suppose that M is a P_0-matrix and $\mu > 0$. Theorem 2.1 indicates that the smooth equation

$$\Phi_\mu(x, y, w; N, M, q) = 0 \tag{2.6}$$

gives two smooth functions $y(\cdot, \mu)$ and $w(\cdot, \mu)$. Note that

$$\phi_\mu(a, b) = 0 \quad \Longleftrightarrow \quad a \geq 0, b \geq 0, ab = \mu^2.$$

As a result, the equation (2.6) is equivalent to the system

$$\begin{aligned} &y \geq 0,\ Nx + (M + \mu I)y + q \geq 0, \\ &y[i]\Big(Nx + (M + \mu I)y + q\Big)[i] = \mu^2, \quad i = 1, \cdots, m \end{aligned} \tag{2.7}$$

in the sense that $y(x, \mu)$ solves (2.7) if and only if

$$\Phi_\mu(x, y(x, \mu), w(x, \mu); N, M, q) = 0$$

with

$$w(x, \mu) := Nx + (M + \mu I)y(x, \mu) + q\ .$$

Since (2.7) with $\mu = 0$ reduces to the linear complementarity problem

$$y \geq 0, \quad Nx + My + q \geq 0, \quad y^T(Nx + My + q) = 0, \tag{2.8}$$

we see that $y(x,\mu)$ tends to a solution of (2.8) as $\mu \to 0$, provided that it is convergent.

In our analysis, we will assume that $y(x,\mu)$ is bounded as $\mu \to 0$. In particular, if M is a P-matrix, then (2.8) has a unique solution for any x and it can be shown that $y(x,\mu)$ actually converges to it as $\mu \to 0$, even without using the regularization term μI in (2.7), see [3].

3 Combined Smoothing Implicit Programming and Penalty Method for Discrete Here-and-Now Problems

In this section, we consider the following here-and-now problem:

$$\begin{array}{ll} \text{minimize} & \displaystyle\sum_{\ell=1}^{L} p_\ell\Big(f(x,y,\omega_\ell) + d^T z_\ell\Big) \\ \text{subject to} & g(x) \leq 0,\ h(x) = 0, \\ & y \geq 0,\ N_\ell x + M_\ell y + q_\ell + z_\ell \geq 0, \\ & y^T(N_\ell x + M_\ell y + q_\ell + z_\ell) = 0, \\ & z_\ell \geq 0,\ \ell = 1,\cdots,L, \end{array} \tag{3.1}$$

which corresponds to the discrete case where $\Omega := \{\omega_1, \omega_2, \cdots, \omega_L\}$. Here, p_ℓ denotes the probability of the random event $\omega_\ell \in \Omega$, the functions $f : \Re^{n+m} \to \Re$, $g : \Re^n \to \Re^{s_1}$, $h : \Re^n \to \Re^{s_2}$ are all continuously differentiable, $N_\ell \in \Re^{m\times n}$, $M_\ell \in \Re^{m\times m}$, $q_\ell \in \Re^m$ are realizations of the random coefficients, d is a constant vector with positive elements, and z_ℓ is the recourse variable corresponding to ω_ℓ. Throughout we assume $p_\ell > 0$ for all $\ell = 1,\cdots,L$.

It is easy to see that problem (3.1) can be rewritten as

$$\begin{array}{ll} \text{minimize} & \displaystyle\sum_{\ell=1}^{L} p_\ell\Big(f(x,y,\omega_\ell) + d^T z_\ell\Big) \\ \text{subject to} & g(x) \leq 0,\ h(x) = 0,\ z_\ell \geq 0, \\ & N_\ell x + M_\ell y + q_\ell + z_\ell \geq 0,\ \ell = 1,\cdots,L, \\ & y \geq 0,\ Nx + My + q + \sum_{l=1}^{L} z_l \geq 0, \\ & y^T(Nx + My + q + \sum_{l=1}^{L} z_l) = 0 \end{array} \tag{3.2}$$

with $N := \sum_{l=1}^{L} N_l$, $M := \sum_{l=1}^{L} M_l$, and $q := \sum_{l=1}^{L} q_l$, or equivalently,

$$\begin{array}{ll} \text{minimize} & \displaystyle\sum_{\ell=1}^{L} p_\ell f(x,y,\omega_\ell) + \mathbf{d}^T\mathbf{z} \\ \text{subject to} & g(x) \leq 0,\ h(x) = 0, \\ & \mathbf{y} - \mathbf{D}y = 0,\ \mathbf{z} \geq 0, \\ & \mathbf{y} \geq 0,\ \mathbf{N}x + \mathbf{M}\mathbf{y} + \mathbf{q} + \mathbf{z} \geq 0, \\ & \mathbf{y}^T(\mathbf{N}x + \mathbf{M}\mathbf{y} + \mathbf{q} + \mathbf{z}) = 0, \end{array} \tag{3.3}$$

where

$$\mathbf{y} := \begin{pmatrix} y_1 \\ \vdots \\ y_L \end{pmatrix}, \quad \mathbf{z} := \begin{pmatrix} z_1 \\ \vdots \\ z_L \end{pmatrix}, \quad \mathbf{d} := \begin{pmatrix} p_1 d \\ \vdots \\ p_L d \end{pmatrix}, \quad \mathbf{D} := \begin{pmatrix} I \\ \vdots \\ I \end{pmatrix}, \tag{3.4}$$

and

$$\mathbf{N} := \begin{pmatrix} N_1 \\ \vdots \\ N_L \end{pmatrix}, \quad \mathbf{M} := \begin{pmatrix} M_1 & & O \\ & \ddots & \\ O & & M_L \end{pmatrix}, \quad \mathbf{q} := \begin{pmatrix} q_1 \\ \vdots \\ q_L \end{pmatrix}.$$

Note that both problems (3.2) and (3.3) are different from ordinary MPECs, because they require $y_1 = y_2 = \cdots = y_L$. This restriction makes the problems harder to deal with than ordinary MPECs. In particular, for any feasible point $(x, y, z_1, \cdots, z_L)$ of problem (3.2), $(Nx + My + q + \sum_{l=1}^{L} z_l)[i] = 0$ implies that $(N_\ell x + M_\ell y + q_\ell + z_\ell)[i] = 0$ holds for every ℓ. This indicates that the MPEC-LICQ does not hold for problem (3.2) in general. On the other hand, since L is usually very large in practice, problem (3.3) is a large-scale program with variables $(x, y, \mathbf{y}, \mathbf{z}) \in \Re^{n+(1+2L)m}$ so that some methods for MPECs may cause more computational difficulties. In this section, we will develop a combined smoothing implicit programming and penalty method for solving the ill-posed MPEC (3.2) directly.

A similar smoothing method for ordinary MPECs with linear complementarity constraints has been considered in [3]. However, several differences should be emphasized here: (a) In [3], the matrix $\mathbf{M}$ is assumed to be a P-matrix, whereas in this paper, we assume it to be a P_0-matrix only; (b) In order to make the new method applicable, in addition to smoothing, we employ a regularization technique and a penalty technique. We will investigate the limiting behavior of local optimal solutions and stationary points.

In addition, as mentioned above, the MPEC-LICQ does not hold for problem (3.2) in general. From now on, the MPEC-LICQ means the one for problem (3.3). On the other hand, because the complementarity constraints in problem (3.2) are lower dimensional, we use them to generate the subproblems.

3.1 SIPP method

Suppose that the matrix M in problem (3.2) is a P_0-matrix. We denote by Λ the matrix $(I, \cdots, I) \in \Re^{m \times mL}$. For each $(x, \mathbf{z})$ and $\mu_k > 0$, let $y(x, \Lambda\mathbf{z}, \mu_k)$ and $w(x, \Lambda\mathbf{z}, \mu_k)$ solve

$$\Phi_{\mu_k}\Big(x, y(x, \Lambda\mathbf{z}, \mu_k), w(x, \Lambda\mathbf{z}, \mu_k); N, M, q + \Lambda\mathbf{z}\Big) = 0. \tag{3.5}$$

The existence and differentiability of the above implicit functions follow from Theorem 2.1. Note that the implicit functions are denoted by $y(x, \Lambda\mathbf{z}, \mu_k)$ and $w(x, \Lambda\mathbf{z}, \mu_k)$, rather than $y(x, \mathbf{z}, \mu_k)$ and $w(x, \mathbf{z}, \mu_k)$, respectively. We then obtain an approximation of problem (3.2)

$$\begin{array}{ll} \text{minimize} & \displaystyle\sum_{\ell=1}^{L} p_\ell\Big(f(x, y(x, \Lambda\mathbf{z}, \mu_k), \omega_\ell) + d^T z_\ell\Big) \\ \text{subject to} & g(x) \le 0, \ h(x) = 0, \\ & N_\ell x + M_\ell y(x, \Lambda\mathbf{z}, \mu_k) + q_\ell + z_\ell \ge 0, \\ & z_\ell \ge 0, \ \ell = 1, \cdots, L. \end{array} \tag{3.6}$$

Since the feasible region of problem (3.6) is dependent on μ_k, (3.6) may not be easy to solve. Therefore, we apply a penalty technique to this problem and have the following approximation:

$$\begin{array}{ll} \text{minimize} & \theta_k(x, \mathbf{z}) \\ \text{subject to} & g(x) \le 0, \ h(x) = 0, \ \mathbf{z} \ge 0, \end{array} \tag{3.7}$$

where

$$\theta_k(x,\mathbf{z}) \;:=\; \sum_{\ell=1}^{L} p_\ell f(x, y(x,\Lambda\mathbf{z},\mu_k),\omega_\ell) + \mathbf{d}^T\mathbf{z}$$
$$+\rho_k \sum_{\ell=1}^{L} \psi\Big(-(N_\ell x + M_\ell y(x,\Lambda\mathbf{z},\mu_k) + q_\ell + z_\ell)\Big), \tag{3.8}$$

ρ_k is a positive parameter, $\psi : \Re^m \to [0,+\infty)$ is a smooth penalty function, and $z_\ell := (\mathbf{z}[(\ell-1)m+1],\cdots,\mathbf{z}[\ell m])^T$ for each ℓ. Some specific penalty functions will be given later. Note that, unlike problem (3.6), the feasible region of problem (3.7) is common for all k.

Now we present our method, called *combined smoothing implicit programming and penalty method* (SIPP), for problem (3.2): Choose two sequences $\{\mu_k\}$ and $\{\rho_k\}$ of positive numbers satisfying

$$\lim_{k\to\infty}\mu_k = 0, \quad \lim_{k\to\infty}\rho_k = +\infty, \quad \lim_{k\to\infty}\mu_k\rho_k = 0. \tag{3.9}$$

We then solve the problems (3.7) to get a sequence $\{(x^{(k)},\mathbf{z}^{(k)})\}$ and let

$$y^{(k)} := y(x^{(k)},\Lambda\mathbf{z}^{(k)},\mu_k).$$

Note that, by Theorem 2.1, problem (3.7) is a smooth mathematical program. Moreover, under some suitable conditions, (3.7) is a convex program, see Chen and Fukushima (2003) for details. Therefore, we may expect that problem (3.7) may be relatively easy to deal with, provided the evaluation of the function $y_\ell(x,\mu_k)$ is not very expensive.

In what follows, we denote by $\mathcal{F}$ and $\mathcal{X}$ the feasible regions of problems (3.2) and (3.7), respectively. Moreover, particular sequences generated by the method will be denoted by $\{x^{(k)}\}$, $\{y^{(k)}\}$, etc., while general sequences will be denoted by $\{x^k\}$, $\{y^k\}$, etc. Also, we use (3.4) to generate some related vectors such as $\mathbf{y}^{(k)},\mathbf{y}^*,\mathbf{z}^{(k)},\mathbf{z}^*$, and so on.

3.2 Convergence results

We investigate the limiting behavior of a sequence generated by SIPP in this subsection. The following lemma will be used later.

Lemma 3.1 *Suppose the matrix M in (3.2) is a P_0-matrix and, for any bounded sequence $\{(x^k,\mathbf{z}^k)\}$ in $\mathcal{X}$, $\{y(x^k,\Lambda\mathbf{z}^k,\mu_k)\}$ is bounded. If $(x^*,y^*,\mathbf{z}^*)\in\mathcal{F}$ and the submatrix $M[\mathcal{K}^*]$ is nondegenerate, where $\mathcal{K}^* := \{\, i \mid (Nx^*+My^*+q+\Lambda\mathbf{z}^*)[i]=0\}$, then there exist a neighborhood U^* of $(x^*,y^*,\mathbf{z}^*)$ and a positive constant π^* such that*

$$\|y(x,\Lambda\mathbf{z},\mu_k)-y\| \le \mu_k\pi^*(\|y\|+\sqrt{m}) \tag{3.10}$$

holds for any $(x,y,\mathbf{z})\in U^*\cap\mathcal{F}$ and any k.

A proof of this lemma is given in Appendix A. Now, we discuss the limiting behavior of the sequence of local optimal solutions of problems (3.7).

Theorem 3.1 *Let the matrix M in (3.2) be a P_0-matrix, $\psi:\Re^m\to[0,+\infty)$ be a continuously differentiable function satisfying*

$$\psi(0)=0, \quad \psi(y)\le\psi(y') \;\; \textit{for any } y'\ge y \textit{ in } \Re^m, \tag{3.11}$$

and, for each bounded sequence $\{(x^k, \mathbf{z}^k)\}$ in $\mathcal{X}$, $\{y(x^k, \Lambda\mathbf{z}^k, \mu_k)\}$ be bounded. Suppose that the sequence $\{(x^{(k)}, y^{(k)}, \mathbf{z}^{(k)})\}$ generated by SIPP with $(x^{(k)}, \mathbf{z}^{(k)})$ being a local optimal solution of problem (3.7) is convergent to $(x^*, y^*, \mathbf{z}^*) \in \mathcal{F}$. If there exists a neighborhood V^* of $(x^*, y^*, \mathbf{z}^*)$ such that $(x^{(k)}, \mathbf{z}^{(k)})$ minimizes θ_k over $V^*|_{\mathcal{X}} := \{(x, \mathbf{z}) \in \mathcal{X} \mid \exists\, y \text{ s.t. } (x, y, \mathbf{z}) \in V^*\}$ for all k large enough and the submatrix $M[\mathcal{K}^*]$ is nondegenerate with $\mathcal{K}^*$ being the same as in Lemma 3.1, then $(x^*, y^*, \mathbf{z}^*)$ is a local optimal solution of problem (3.2).

Proof. By Lemma 3.1, there exist a closed sphere $\mathcal{B} \subseteq V^*$ centered at the point $(x^*, y^*, \mathbf{z}^*)$ with positive radius and a positive number π^* such that (3.10) holds for any $(x, y, \mathbf{z}) \in \mathcal{F} \cap \mathcal{B}$ and every k. Since $\mathcal{F} \cap \mathcal{B}$ is a nonempty compact set, the problem

$$\begin{array}{ll} \text{minimize} & \displaystyle\sum_{\ell=1}^{L} p_\ell f(x, y, \omega_\ell) + \mathbf{d}^T\mathbf{z} \\ \text{subject to} & (x, y, \mathbf{z}) \in \mathcal{F} \cap \mathcal{B} \end{array} \tag{3.12}$$

has an optimal solution, say $(\bar{x}, \bar{y}, \bar{\mathbf{z}})$.

Suppose $(x, y, \mathbf{z}) \in \mathcal{F} \cap \mathcal{B}$. We then have from (3.8) and the mean-value theorem that

$$\begin{aligned} \theta_k(x, \mathbf{z}) \;=\; & \sum_{\ell=1}^{L} p_\ell \Big(f(x, y, \omega_\ell) + (y(x, \Lambda\mathbf{z}, \mu_k) - y)^T \nabla_y f(x, (1-t)y(x, \Lambda\mathbf{z}, \mu_k) + ty, \omega_\ell) \Big) \\ & + \mathbf{d}^T\mathbf{z} + \rho_k \sum_{\ell=1}^{L} \psi\Big(-(N_\ell x + M_\ell y(x, \Lambda\mathbf{z}, \mu_k) + q_\ell + z_\ell) \Big), \end{aligned} \tag{3.13}$$

where $t \in [0, 1]$. Note that, by (3.10),

$$\begin{aligned} \|(1-t)y(x, \Lambda\mathbf{z}, \mu_k) + ty\| &= \|(1-t)(y(x, \Lambda\mathbf{z}, \mu_k) - y) + y\| \\ &\leq \|y(x, \Lambda\mathbf{z}, \mu_k) - y\| + \|y\| \\ &\leq \mu_k \pi^* (\|y\| + \sqrt{m}) + \|y\|. \end{aligned}$$

This indicates that the set

$$\Big\{(x, (1-t)y(x, \Lambda\mathbf{z}, \mu_k) + ty) \;\Big|\; (x, y, \mathbf{z}) \in \mathcal{F} \cap \mathcal{B}, t \in [0, 1], k = 1, 2, \cdots \Big\}$$

is bounded. Similarly, we see that

$$\Big\{(x, tM_\ell(y - y(x, \Lambda\mathbf{z}, \mu_k))) \;\Big|\; (x, y, \mathbf{z}) \in \mathcal{F} \cap \mathcal{B}, \ell = 1, \cdots, L, t \in [0, 1], k = 1, 2, \cdots \Big\}$$

is also bounded. Then, by the continuous differentiability of both f and ψ, there exists a constant $\tau > 0$ such that, for $\ell = 1, \cdots, L$,

$$\|\nabla_y f(x, (1-t)y(x, \Lambda\mathbf{z}, \mu_k) + ty, \omega_\ell)\| \leq \tau, \tag{3.14}$$

$$\|\nabla\psi\Big(tM_\ell(y - y(x, \Lambda\mathbf{z}, \mu_k))\Big)\| \leq \tau \tag{3.15}$$

hold for any $(x, y, \mathbf{z}) \in \mathcal{F} \cap \mathcal{B}, t \in [0, 1]$, and every k. Noticing that $(x, y, \mathbf{z}) \in \mathcal{F} \cap \mathcal{B}$ implies $N_\ell x + M_\ell y + q_\ell + z_\ell \geq 0$ for each ℓ, we have from (3.11) and (3.15) that

$$\begin{aligned} \psi\Big(-(N_\ell x + M_\ell y(x, \Lambda\mathbf{z}, \mu_k) + q_\ell + z_\ell) \Big) &\leq \psi\Big(M_\ell(y - y(x, \Lambda\mathbf{z}, \mu_k)) \Big) \\ &= \psi\Big(M_\ell(y - y(x, \Lambda\mathbf{z}, \mu_k)) \Big) - \psi(0) \\ &= \nabla\psi\Big(t' M_\ell(y - y(x, \Lambda\mathbf{z}, \mu_k)) \Big)^T M_\ell \Big(y - y(x, \Lambda\mathbf{z}, \mu_k) \Big) \\ &\leq \tau \|M_\ell\| \, \|y - y(x, \Lambda\mathbf{z}, \mu_k)\|, \end{aligned}$$

where $t' \in [0,1]$ and the second equality follows from the mean-value theorem. This, together with (3.13)–(3.14) and (3.10), yields

$$\begin{aligned}\Big|\theta_k(x,\mathbf{z}) - \sum_{\ell=1}^{L} p_\ell f(x,y,\omega_\ell) - \mathbf{d}^T\mathbf{z}\Big| &\le \tau\|y(x,\Lambda\mathbf{z},\mu_k) - y\| + \Big(\tau\rho_k \sum_{\ell=1}^{L}\|M_\ell\|\Big)\|y - y(x,\Lambda\mathbf{z},\mu_k)\| \\ &\le \pi^*\tau\Big(\mu_k + \mu_k\rho_k\sum_{\ell=1}^{L}\|M_\ell\|\Big)(\|y\| + \sqrt{m})\end{aligned}$$

for any $(x,y,\mathbf{z}) \in \mathcal{F}\cap\mathcal{B}$ and k. In particular,

$$\Big|\theta_k(\bar{x},\bar{\mathbf{z}}) - \sum_{\ell=1}^{L} p_\ell f(\bar{x},\bar{y},\omega_\ell) - \mathbf{d}^T\bar{\mathbf{z}}\Big| \le \pi^*\tau\Big(\mu_k + \mu_k\rho_k\sum_{\ell=1}^{L}\|M_\ell\|\Big)(\|\bar{y}\| + \sqrt{m}). \tag{3.16}$$

Moreover, since ψ is always nonnegative, we have from the continuity of f that

$$\begin{aligned}\lim_{k\to\infty}\theta_k(x^{(k)},\mathbf{z}^{(k)}) &\ge \lim_{k\to\infty}\Big(\sum_{\ell=1}^{L} p_\ell f(x^{(k)},y^{(k)},\omega_\ell) + \mathbf{d}^T\mathbf{z}^{(k)}\Big) \\ &= \sum_{\ell=1}^{L} p_\ell f(x^*,y^*,\omega_\ell) + \mathbf{d}^T\mathbf{z}^*. \end{aligned}\tag{3.17}$$

Note that, by the fact that $\mathcal{F}\cap\mathcal{B} \subseteq V^*$, $(x^{(k)},\mathbf{z}^{(k)})$ is an optimal solution of the problem

$$\begin{array}{ll}\text{minimize} & \theta_k(x,\mathbf{z}) \\ \text{subject to} & (x,\mathbf{z}) \in \mathcal{X}_1 := \{\, (x,\mathbf{z}) \in \mathcal{X} \mid \exists\, y \text{ s.t. } (x,y,\mathbf{z}) \in \mathcal{F}\cap\mathcal{B}\,\},\end{array}$$

provided k is large enough, and $(\bar{x},\bar{\mathbf{z}})$ is a feasible point of this problem. We then have from (3.16) that, for every k sufficiently large,

$$\begin{aligned}\theta_k(x^{(k)},\mathbf{z}^{(k)}) &\le \theta_k(\bar{x},\bar{\mathbf{z}}) \\ &\le \sum_{\ell=1}^{L} p_\ell f(\bar{x},\bar{y},\omega_\ell) + \mathbf{d}^T\bar{\mathbf{z}} + \pi^*\tau\Big(\mu_k + \mu_k\rho_k\sum_{\ell=1}^{L}\|M_\ell\|\Big)(\|\bar{y}\| + \sqrt{m}). \end{aligned}\tag{3.18}$$

Therefore, taking into account the equality (3.17) and the assumption (3.9), we have by letting $k \to \infty$ in (3.18) that

$$\sum_{\ell=1}^{L} p_\ell f(x^*,y^*,\omega_\ell) + \mathbf{d}^T\mathbf{z}^* \le \sum_{\ell=1}^{L} p_\ell f(\bar{x},\bar{y},\omega_\ell) + \mathbf{d}^T\bar{\mathbf{z}},$$

while the converse inequality immediately follows from the fact that $(\bar{x},\bar{y},\bar{\mathbf{z}})$ is an optimal solution of problem (3.12). As a result, we have

$$\sum_{\ell=1}^{L} p_\ell f(x^*,y^*,\omega_\ell) + \mathbf{d}^T\mathbf{z}^* = \sum_{\ell=1}^{L} p_\ell f(\bar{x},\bar{y},\omega_\ell) + \mathbf{d}^T\bar{\mathbf{z}},$$

namely, $(x^*,y^*,\mathbf{z}^*)$ is an optimal solution of problem (3.12) and hence it is a local optimal solution of problem (3.2). This completes the proof. $\square$

It is not difficult to see that the function

$$\psi(y) := \sum_{i=1}^{m}\Big(\max(y[i],0)\Big)^{\sigma}, \tag{3.19}$$

where $\sigma \geq 2$ is a positive integer, satisfies the conditions assumed in Theorem 3.1. This function is often employed for solving constrained optimization problems. For more details, see [1].

We have discussed the convergence of local optimal solutions of problems (3.7). In practice, it may not be easy to obtain an optimal solution, whereas computation of stationary points may be relatively easy. Therefore, it is necessary to study the limiting behavior of stationary points of subproblems (3.7).

Theorem 3.2 *Suppose the matrix M in (3.2) is a P_0-matrix, the function $\psi : \Re^m \to [0, +\infty)$ is given by (3.19) with $\sigma = 2$, and $(x^{(k)}, \mathbf{z}^{(k)})$ is a stationary point of (3.7) for each k. Let $(x^*, y^*, \mathbf{z}^*) \in \mathcal{F}$ be an accumulation point of the sequence $\{(x^{(k)}, y^{(k)}, \mathbf{z}^{(k)})\}$ generated by SIPP. If the MPEC-LICQ is satisfied at $(x^*, y^*, \mathbf{y}^*, \mathbf{z}^*)$ in the MPEC (3.3), then $(x^*, y^*, \mathbf{z}^*)$ is a C-stationary point of problem (3.2). Furthermore, if y^* satisfies the strict complementarity condition, then $(x^*, y^*, \mathbf{z}^*)$ is S-stationary to (3.2).*

Although the results established in this theorem are interesting and important, its proof is somewhat lengthy and technical. To avoid disturbing the readability, we therefore give a detailed proof of the theorem in Appendix B.

4 Here-and-Now Problems with Continuous Random Variable

In this section, we consider the here-and-now problem

$$\begin{array}{ll} \text{minimize} & E_\omega[f(x, y, \omega) + d^T z(\omega)] \\ \text{subject to} & g(x) \leq 0,\ h(x) = 0, \\ & 0 \leq y \perp (N(\omega)x + M(\omega)y + q(\omega) + z(\omega)) \geq 0, \\ & z(\omega) \geq 0 \quad \forall \omega \in \Omega, \\ & x \in \Re^n,\ y \in \Re^m,\ z(\cdot) \in \mathcal{C}(\Omega), \end{array} \tag{4.1}$$

where ω is a continuous random variable, $\Omega := [a_1, b_1] \times \cdots \times [a_\nu, b_\nu] \subset \Re^\nu$, and $\mathcal{C}(\Omega)$ is the family of continuous functions from Ω into $\Re^m$. Here, g, h, d are the same as in Section 3, $f : \Re^{n+m} \times \Omega \to \Re$ is uniformly continuous with respect to (x, y) and continuous with respect to ω, $N : \Omega \to \Re^{m \times n}$, $M : \Omega \to \Re^{m \times m}$, and $q : \Omega \to \Re^m$ are all continuous. Without loss of generality, we assume that $\Omega := [0, 1]^\nu$. Let $\zeta : \Omega \to [0, +\infty)$ be the continuous probability density function of ω. Then we have

$$E_\omega[f(x, y, \omega) + d^T z(\omega)] = \int_\Omega \Big(f(x, y, \omega) + d^T z(\omega)\Big)\zeta(\omega) d\omega.$$

We next employ a quasi-Monte Carlo method [12] for numerical integration to discretize problem (4.1). This method uses an *uniformly distributed* infinite sequence $\Omega_\infty := \{\omega_1, \omega_2, \cdots\} \subseteq \Omega$, i.e., for any subinterval S of Ω, there holds

$$\lim_{L \to \infty} \frac{1}{L} \sum_{\ell=1}^{L} \delta_S(\omega_\ell) = \mathcal{L}(S), \tag{4.2}$$

where δ_S denotes the characteristic function of the set S and $\mathcal{L}(S)$ means the Lebesgue measure of S. Since the Lebesgue measure of any nonempty open set is positive, it is easy to see from (4.2) that Ω_∞ is dense in Ω. Therefore, the following problem is an appropriate discrete approximation of problem (4.1): For a given integer $L > 0$ and a given subset $\Omega_L := \{\omega_1, \cdots, \omega_L\} \subseteq \Omega_\infty$,

$$\begin{array}{ll} \text{minimize} & \displaystyle\frac{1}{L}\sum_{\ell=1}^{L} \zeta(\omega_\ell)\Big(f(x, y, \omega_\ell) + d^T z(\omega_\ell)\Big) \\ \text{subject to} & g(x) \le 0,\ h(x) = 0, \\ & 0 \le y \perp (N(\omega_\ell)x + M(\omega_\ell)y + q(\omega_\ell) + z(\omega_\ell)) \ge 0, \\ & z(\omega_\ell) \ge 0,\ \ell = 1, \cdots, L. \end{array} \tag{4.3}$$

This problem has been discussed in the last section.

Suppose that $(x^L, y^L, z^L(\omega_1), \cdots, z^L(\omega_L))$ is an optimal solution of problem (4.3) for each L and the sequence $\{(x^L, y^L)\}$ converges to a point (x^*, y^*) as $L \to +\infty$. Let us define

$$\tilde{z}^L(\omega_\ell) := \max\Big\{-(N(\omega_\ell)x^L + M(\omega_\ell)y^L + q(\omega_\ell)), 0\Big\}, \qquad \ell = 1, \cdots, L, \tag{4.4}$$

$$z^*(\omega) := \max\Big\{-(N(\omega)x^* + M(\omega)y^* + q(\omega)), 0\Big\}, \qquad \omega \in \Omega. \tag{4.5}$$

Since $\Omega_L \subseteq \Omega_{L'}$ for all $L \le L'$, we have

$$\lim_{L\to\infty} \tilde{z}^L(\omega_\ell) = z^*(\omega_\ell) \tag{4.6}$$

for any fixed ω_ℓ. Moreover, there holds

$$\begin{aligned} & \frac{1}{L}\sum_{\ell=1}^{L} \zeta(\omega_\ell)\Big(f(x^L, y^L, \omega_\ell) + d^T z^L(\omega_\ell)\Big) - \frac{1}{L}\sum_{\ell=1}^{L} \zeta(\omega_\ell)\Big(f(x^L, y^L, \omega_\ell) + d^T \tilde{z}^L(\omega_\ell)\Big) \\ = \ & \frac{1}{L}\sum_{\ell=1}^{L} \zeta(\omega_\ell) d^T\Big(z^L(\omega_\ell) - \tilde{z}^L(\omega_\ell)\Big) \\ = \ & \frac{1}{L}\sum_{\ell=1}^{L} \zeta(\omega_\ell) d^T \min\Big\{N(\omega_\ell)x^L + M(\omega_\ell)y^L + q(\omega_\ell) + z^L(\omega_\ell),\ z^L(\omega_\ell)\Big\} \ge 0, \end{aligned}$$

where the second equality follows from (4.4) and the inequality follows from the feasibility of $(x^L, y^L, z^L(\omega_1), \cdots, z^L(\omega_L))$ in (4.3). Thus, $(x^L, y^L, \tilde{z}^L(\omega_1), \cdots, \tilde{z}^L(\omega_L))$ is also an optimal solution of problem (4.3). We next show that (x^*, y^*) together with $z^*(\cdot)$ is an optimal solution of problem (4.1). Note that $z^*(\cdot) \in \mathcal{C}(\Omega)$ by the definition (4.5). Moreover, since $\Omega = [0, 1]^\nu$, any continuous function must be integrable on Ω.

Lemma 4.1 *Suppose the function $\xi : \Omega \to \Re$ is continuous. Then we have*

$$\lim_{L\to\infty} \frac{1}{L}\sum_{\ell=1}^{L} \xi(\omega_\ell)\zeta(\omega_\ell) = \int_\Omega \xi(\omega)\zeta(\omega)d\omega.$$

It is not difficult to prove this lemma by the results given in Chapter 2 of [12]. We then have from Lemma 4.1 immediately that, for any $z(\cdot) \in \mathcal{C}(\Omega)$,

$$\lim_{L\to\infty} \frac{1}{L}\sum_{\ell=1}^{L} \zeta(\omega_\ell)\Big(f(x, y, \omega_\ell) + d^T z(\omega_\ell)\Big) = \int_\Omega \Big(f(x, y, \omega) + d^T z(\omega)\Big)\zeta(\omega)d\omega \tag{4.7}$$

and particularly,

$$\lim_{L\to\infty} \frac{1}{L}\sum_{\ell=1}^{L} \zeta(\omega_\ell) = \int_\Omega \zeta(\omega)d\omega = 1. \tag{4.8}$$

Theorem 4.1 *The point (x^*, y^*) together with $z^*(\cdot)$ is an optimal solution of problem (4.1).*

Proof. We first prove that $(x^*, y^*, z^*(\cdot))$ is feasible to problem (4.1). To this end, since $N(\omega)x^* + M(\omega)y^* + q(\omega) + z^*(\omega) \geq 0$ holds by the definition (4.5), it is sufficient to show that

$$(y^*)^T\Big(N(\omega)x^* + M(\omega)y^* + q(\omega) + z^*(\omega)\Big) = 0 \qquad \forall \omega \in \Omega. \tag{4.9}$$

In fact, by the feasibility of $(x^L, y^L, \tilde{z}^L(\omega_1), \cdots, \tilde{z}^L(\omega_L))$ to problem (4.3) for each L, there holds

$$(y^L)^T\Big(N(\omega_\ell)x^L + M(\omega_\ell)y^L + q(\omega_\ell) + \tilde{z}^L(\omega_\ell)\Big) = 0, \qquad \ell \leq L.$$

In consequence, letting $L \to +\infty$ and noting that (4.6) holds for each fixed ω_ℓ, we have

$$(y^*)^T\Big(N(\omega_\ell)x^* + M(\omega_\ell)y^* + q(\omega_\ell) + z^*(\omega_\ell)\Big) = 0. \tag{4.10}$$

Since the sequence $\{\omega_\ell\}$ is dense in Ω and $N(\cdot), M(\cdot), q(\cdot), z^*(\cdot)$ are all continuous, we obtain (4.9) from (4.10) immediately.

Let $(x, y, z(\cdot))$ be an arbitrary feasible solution of (4.1). It is obvious that $(x, y, z(\omega_1), \cdots, z(\omega_L))$ is feasible to problem (4.3) for any L. Since $(x^L, y^L, \tilde{z}^L(\omega_1), \cdots, \tilde{z}^L(\omega_L))$ is an optimal solution of (4.3) as shown earlier, we have

$$\begin{aligned}
& \frac{1}{L}\sum_{\ell=1}^{L}\zeta(\omega_\ell)\Big(f(x^*, y^*, \omega_\ell) + d^T z^*(\omega_\ell)\Big) - \frac{1}{L}\sum_{\ell=1}^{L}\zeta(\omega_\ell)\Big(f(x, y, \omega_\ell) + d^T z(\omega_\ell)\Big) \\
\leq \; & \frac{1}{L}\sum_{\ell=1}^{L}\zeta(\omega_\ell)\Big(f(x^*, y^*, \omega_\ell) + d^T z^*(\omega_\ell)\Big) - \frac{1}{L}\sum_{\ell=1}^{L}\zeta(\omega_\ell)\Big(f(x^L, y^L, \omega_\ell) + d^T \tilde{z}^L(\omega_\ell)\Big) \\
= \; & \frac{1}{L}\sum_{\ell=1}^{L}\zeta(\omega_\ell)\Big[\Big(f(x^*, y^*, \omega_\ell) - f(x^L, y^L, \omega_\ell)\Big) + d^T\Big(z^*(\omega_\ell) - \tilde{z}^L(\omega_\ell)\Big)\Big] \\
\leq \; & \frac{1}{L}\sum_{\ell=1}^{L}\zeta(\omega_\ell)\Big(\Big|f(x^*, y^*, \omega_\ell) - f(x^L, y^L, \omega_\ell)\Big| + \Big|d^T\Big(z^*(\omega_\ell) - \tilde{z}^L(\omega_\ell)\Big)\Big|\Big).
\end{aligned} \tag{4.11}$$

Note that f is uniformly continuous with respect to (x, y) and, by (4.8), the sequence $\{\frac{1}{L}\sum_{\ell=1}^{L}\zeta(\omega_\ell)\}$ is bounded. This yields

$$\lim_{L\to\infty}\frac{1}{L}\sum_{\ell=1}^{L}\zeta(\omega_\ell)\Big|f(x^*, y^*, \omega_\ell) - f(x^L, y^L, \omega_\ell)\Big| = 0. \tag{4.12}$$

On the other hand, it is easy to see from the definitions (4.4) and (4.5) that

$$\Big|d^T\Big(z^*(\omega_\ell) - \tilde{z}^L(\omega_\ell)\Big)\Big| \leq \Big|d^T\Big(N(\omega_\ell)(x^* - x^L) + M(\omega_\ell)(y^* - y^L)\Big)\Big|, \qquad \ell = 1, \cdots, L.$$

By the boundedness of the functions $N(\cdot)$ and $M(\cdot)$ and the sequence $\{\frac{1}{L}\sum_{\ell=1}^{L}\zeta(\omega_\ell)\}$, we have

$$\lim_{L\to\infty}\frac{1}{L}\sum_{\ell=1}^{L}\zeta(\omega_\ell)\Big|d^T\Big(z^*(\omega_\ell) - \tilde{z}^L(\omega_\ell)\Big)\Big| = 0. \tag{4.13}$$

Thus, by letting $L \to +\infty$ in (4.11) and taking (4.7) and (4.12)–(4.13) into account, we obtain

$$\int_\Omega \Big(f(x^*, y^*, \omega) + d^T z^*(\omega)\Big)\zeta(\omega)d\omega \leq \int_\Omega \Big(f(x, y, \omega) + d^T z(\omega)\Big)\zeta(\omega)d\omega, \tag{4.14}$$

which means

$$E_\omega[f(x^*, y^*, \omega) + d^T z^*(\omega)] \leq E_\omega[f(x, y, \omega) + d^T z(\omega)].$$

This implies that the point (x^*, y^*) together with $z^*(\cdot)$ constitutes an optimal solution of problem (4.1). $\square$

5 Concluding Remarks

We have presented a combined smoothing implicit programming and penalty method for the here-and-now problems with linear complementarity constraints. For the lower-level wait-and-see problems, we may consider a similar but somewhat simpler approach. In particular, for the discrete model

$$
\begin{array}{ll}
\text{minimize} & \displaystyle\sum_{\ell=1}^{L} p_\ell f(x, y_\ell) \\
\text{subject to} & g(x) \le 0,\ h(x) = 0, \\
& y_\ell \ge 0,\ N_\ell x + M_\ell y_\ell + q_\ell \ge 0, \\
& y_\ell^T (N_\ell x + M_\ell y_\ell + q_\ell) = 0,\ \ell = 1, \cdots, L
\end{array}
\tag{5.1}
$$

with p_ℓ, N_ℓ, M_ℓ, and q_ℓ being the same as in Section 3, the subproblem corresponding to (3.7) becomes

$$
\begin{array}{ll}
\text{minimize} & \displaystyle\sum_{\ell=1}^{L} p_\ell f(x, y_\ell(x, \mu_k)) \\
\text{subject to} & g(x) \le 0,\ h(x) = 0,
\end{array}
$$

where $y_\ell(x, \mu)$ satisfies the equation $\Phi_\mu(x, y_\ell(x, \mu), w_\ell(x, \mu); N_\ell, M_\ell, q_\ell) = 0$ with $w_\ell(x, \mu) = N_\ell x + (M_\ell + \mu I) y_\ell(x, \mu) + q_\ell$ for each ℓ. Therefore, we do not need the penalty steps for (5.1).

On the other hand, recall that SMPECs contain the ordinary MPECs as a special subclass. As a result, all the conclusions remain true for standard MPECs. Comparing with the results given in the literature, the assumptions employed in this paper are relatively weak.

References

[1] D.P. Bertsekas, Constrained Optimization and Lagrange Multiplier Methods, Academic Press, New York, 1982.

[2] J.R. Birge and F. Louveaux, Introduction to Stochastic Programming, Springer, New York, 1997.

[3] X. Chen and M. Fukushima, *A smoothing method for a mathematical program with P-matrix linear complementarity constraints*, Computational Optimization and Applications, 27 (2004), 223–246.

[4] Y. Chen and M. Florian, *The Nonlinear Bilevel Programming Problem: Formulations, Regularity and Optimality Conditions*, Optimization, 32 (1995), 193–209.

[5] R.W. Cottle, J.S. Pang, and R.E. Stone, The Linear Complementarity Problem, Academic Press, New York, NY, 1992.

[6] A. Fischer, *A special Newton-type optimization method*, Optimization, 24 (1992), 269–284.

[7] H. Jiang and D. Ralph, *Smooth SQP methods for mathematical programs with nonlinear complementarity constraints*, SIAM Journal on Optimization, 10 (2000), 779–808.

[8] P. Kall and S.W. Wallace, Stochastic Programming, John Wiley & Sons, Chichester, 1994.

[9] C. Kanzow, *Some noninterior continuation methods for linear complementarity problems*, SIAM Journal on Matrix Analysis and Applications, 17 (1996), 851–868.

[10] G.H. Lin and M. Fukushima, *New relaxation method for mathematical programs with complementarity constraints*, Journal of Optimization Theory and Applications, 118 (2003), 81–116.

[11] Z.Q. Luo, J.S. Pang, and D. Ralph, Mathematical Programs with Equilibrium Constraints, Cambridge University Press, Cambridge, United Kingdom, 1996.

[12] H. Niederreiter, Random Number Generation and Quasi-Monte Carlo Methods, SIAM, Philadelphia, 1992.

[13] J.M. Ortega and W.C. Rheinboldt, Iterative Solution of Nonlinear Equations in Several Variables, Academic Press, 1970.

[14] J.S. Pang, *Complementarity problems*, Handbook on Global Optimization, R. Horst and P. Pardalos (eds.), Kluwer Academic Publishers, B.V. Dordrecht, 1994, 271–338.

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

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

[17] S. Vajda, Probabilistic Programming, Academic Press, New York, 1972.

[18] H. Xu, *An MPCC approach for stochastic Stackelberg-Nash-Cournot equilibrium*, manuscript, University of Southampton, Highfield Southampton, 2003.

Appendix A: Proof of Lemma 3.1. For any $(x, y, \mathbf{z}) \in \Re^{n+m+mL}$, we denote $w := Nx + My + q + \Lambda\mathbf{z}$. Since $(x, y, \mathbf{z}) \in \mathcal{F}$ implies

$$\Phi_0(x, y, w; N, M, q + \Lambda\mathbf{z}) = 0, \tag{A.1}$$

we have from (3.5) and (A.1) that

$$\begin{aligned} 0 &= \Phi_{\mu_k}(x, y(x, \Lambda\mathbf{z}, \mu_k), w(x, \Lambda\mathbf{z}, \mu_k); N, M, q + \Lambda\mathbf{z}) - \Phi_0(x, y, w; N, M, q + \Lambda\mathbf{z}) \\ &= \begin{pmatrix} M + \mu_k I & -I \\ I - D(x, \mathbf{z}, \mu_k) & D(x, \mathbf{z}, \mu_k) \end{pmatrix} \begin{pmatrix} y(x, \Lambda\mathbf{z}, \mu_k) - y \\ w(x, \Lambda\mathbf{z}, \mu_k) - w \end{pmatrix} - \mu_k \begin{pmatrix} -y \\ 2\mu_k a^k \end{pmatrix}. \end{aligned} \tag{A.2}$$

Here, $D(x, \mathbf{z}, \mu_k) := \operatorname{diag}\Big(a^k[i](y(x, \Lambda\mathbf{z}, \mu_k)[i] + y[i])\Big)$ and

$$\begin{aligned} a^k[i] &:= \frac{1}{\sqrt{(y(x, \Lambda\mathbf{z}, \mu_k)[i])^2 + (w(x, \Lambda\mathbf{z}, \mu_k)[i])^2 + 2\mu_k^2} + \sqrt{(y[i])^2 + (w[i])^2}} \\ &= \frac{1}{y(x, \Lambda\mathbf{z}, \mu_k)[i] + w(x, \Lambda\mathbf{z}, \mu_k)[i] + y[i] + w[i]}, \end{aligned} \tag{A.3}$$

where the last equality follows from (3.5) and (A.1). Moreover, it is easy to see that, for any i and k,

$$0 < a^k[i](y(x, \Lambda\mathbf{z}, \mu_k)[i] + y[i]) < 1.$$

We next prove that there exist a neighborhood U^* of $(x^*, y^*, \mathbf{z}^*)$ and a positive constant π^* such that

$$\left\| \begin{pmatrix} M + \mu_k I & -I \\ I - D(x, \mathbf{z}, \mu_k) & D(x, \mathbf{z}, \mu_k) \end{pmatrix}^{-1} \right\| \leq \pi^* \tag{A.4}$$

holds for any $(x, y, \mathbf{z}) \in U^* \cap \mathcal{F}$ and any k. Otherwise, there must be a subsequence $\{k_j\}$ of $\{k\}$ and a sequence $\{(x^j, y^j, \mathbf{z}^j)\} \subset \mathcal{F}$ such that

$$\begin{aligned} &\lim_{j\to\infty} (x^j, y^j, \mathbf{z}^j) = (x^*, y^*, \mathbf{z}^*), \\ &\lim_{j\to\infty} \begin{pmatrix} M + \mu_{k_j} I & -I \\ I - D(x^j, \mathbf{z}^j, \mu_{k_j}) & D(x^j, \mathbf{z}^j, \mu_{k_j}) \end{pmatrix} = \begin{pmatrix} M & -I \\ I - \hat{D} & \hat{D} \end{pmatrix} := H, \end{aligned}$$

where $\hat{D} := \operatorname{diag}(\hat{d}[1], \cdots, \hat{d}[m])$ satisfies $0 \leq \hat{d}[i] \leq 1$ for each i and the matrix H is singular. Note that, by (A.3),

$$\hat{d}[i] = \lim_{j\to\infty} \frac{y(x^j, \Lambda\mathbf{z}^j, \mu_{k_j})[i] + y^j[i]}{y(x^j, \Lambda\mathbf{z}^j, \mu_{k_j})[i] + w(x^j, \Lambda\mathbf{z}^j, \mu_{k_j})[i] + y^j[i] + w^j[i]}.$$

By the assumptions of the lemma, the sequence $\{y(x^j, \Lambda\mathbf{z}^j, \mu_{k_j})\}$ is bounded and then so is the sequence $\{w(x^j, \Lambda\mathbf{z}^j, \mu_{k_j})\}$. It is not difficult to see that

$$\hat{d}[i] = 1 \quad \Rightarrow \quad \lim_{j\to\infty} \Big(w(x^j, \Lambda\mathbf{z}^j, \mu_{k_j})[i] + w^j[i]\Big) = 0 \quad \Rightarrow \quad \lim_{j\to\infty} w^j[i] = 0.$$

This indicates that $\mathcal{K}^* \supseteq \hat{\mathcal{K}} := \{\, i \mid \hat{d}[i] = 1\}$. Taking into account the assumption that the submatrix $M[\mathcal{K}^*]$ is nondegenerate, we deduce that the submatrix $M[\hat{\mathcal{K}}]$ is nonsingular and then it is not difficult to show that H is nonsingular. This is a contradiction and hence we obtain (A.4).

Moreover, we have from Theorem 2.1(ii) that, for each k and i,

$$\begin{aligned} 2\mu_k a^k[i] &\leq \frac{2\mu_k}{\sqrt{(y(x, \Lambda\mathbf{z}, \mu_k)[i])^2 + (w(x, \Lambda\mathbf{z}, \mu_k)[i])^2 + 2\mu_k^2}} \\ &\leq \frac{2\mu_k}{\sqrt{2y(x, \Lambda\mathbf{z}, \mu_k)[i]w(x, \Lambda\mathbf{z}, \mu_k)[i] + 2\mu_k^2}} \\ &= 1. \end{aligned} \tag{A.5}$$

Thus, it follows from (A.2), (A.4), and (A.5) that

$$\begin{aligned} \|y(x, \Lambda\mathbf{z}, \mu_k) - y\| &\leq \left\| \begin{pmatrix} y(x, \Lambda\mathbf{z}, \mu_k) - y \\ w(x, \Lambda\mathbf{z}, \mu_k) - w \end{pmatrix} \right\| \\ &= \mu_k \left\| \begin{pmatrix} M + \mu_k I & -I \\ I - D(x, \mathbf{z}, \mu_k) & D(x, \mathbf{z}, \mu_k) \end{pmatrix}^{-1} \begin{pmatrix} -y \\ 2\mu_k a^k \end{pmatrix} \right\| \\ &\leq \mu_k \pi^* (\|y\| + \sqrt{m}). \end{aligned}$$

This completes the proof of Lemma 3.1. □

Appendix B: Proof of Theorem 3.2. Assume without loss of generality that the sequence $\{(x^{(k)}, y^{(k)}, \mathbf{z}^{(k)})\}$ converges to $(x^*, y^*, \mathbf{z}^*)$. By the MPEC-LICQ assumption, problem (3.7) satisfies the standard LICQ at $(x^{(k)}, \mathbf{z}^{(k)})$ for all k sufficiently large and so, by the stationarity of $(x^{(k)}, \mathbf{z}^{(k)})$, there exist unique Lagrange multiplier vectors λ^k, γ^k, and $\alpha^k := \Big((\alpha_1^k)^T, \cdots, (\alpha_L^k)^T\Big)^T$ such that

$$\nabla\theta_k(x^{(k)}, \mathbf{z}^{(k)}) + \begin{pmatrix} \nabla g(x^{(k)}) \\ O \\ \vdots \\ O \end{pmatrix} \lambda^k + \begin{pmatrix} \nabla h(x^{(k)}) \\ O \\ \vdots \\ O \end{pmatrix} \gamma^k - \begin{pmatrix} O & \cdots & O \\ I & \cdots & O \\ \vdots & \ddots & \vdots \\ O & \cdots & I \end{pmatrix} \alpha^k = 0, \tag{B.1}$$

$$h(x^{(k)}) = 0, \qquad 0 \le \lambda^k \perp \Big(- g(x^{(k)})\Big) \ge 0, \qquad 0 \le \alpha^k \perp \mathbf{z}^{(k)} \ge 0. \tag{B.2}$$

In the rest of the proof, we suppose k is large enough so that (B.1)–(B.2) and

$$\mathcal{I}_g(x^{(k)}) \subseteq \mathcal{I}_g(x^*) \tag{B.3}$$

are satisfied and furthermore, for each $\ell = 1, \cdots, L$, there hold

$$\begin{aligned} \mathcal{I}_{Z_\ell}^k &:= \Big\{ i \;\Big|\; z_\ell^{(k)}[i] = 0 \Big\} \subseteq \mathcal{I}_{Z_\ell}^* := \Big\{ i \;\Big|\; z_\ell^*[i] = 0 \Big\}, && \text{(B.4)} \\ \mathcal{I}_{W_\ell}^k &:= \Big\{ i \;\Big|\; \Big(N_\ell x^{(k)} + M_\ell y^{(k)} + q_\ell + z_\ell^{(k)}\Big)[i] = 0 \Big\} \\ &\subseteq \mathcal{I}_{W_\ell}^* := \Big\{ i \;\Big|\; (N_\ell x^* + M_\ell y^* + q_\ell + z_\ell^*)[i] = 0 \Big\}, && \text{(B.5)} \\ \mathcal{I}_Y^k &:= \Big\{ i \;\Big|\; y^{(k)}[i] = 0 \Big\} \subseteq \mathcal{I}_Y^* := \Big\{ i \;\Big|\; y^*[i] = 0 \Big\}, && \text{(B.6)} \\ \mathcal{I}_W^k &:= \Big\{ i \;\Big|\; \Big(N x^{(k)} + M y^{(k)} + q + \Lambda \mathbf{z}^{(k)}\Big)[i] = 0 \Big\} \\ &\subseteq \mathcal{I}_W^* := \Big\{ i \;\Big|\; (N x^* + M y^* + q + \Lambda \mathbf{z}^*)[i] = 0 \Big\}. && \text{(B.7)} \end{aligned}$$

It is clear that

$$\mathcal{I}_W^* = \cap_{\ell=1}^L \mathcal{I}_{W_\ell}^*. \tag{B.8}$$

Note that $\Phi_{\mu_k}(x, y(x, \Lambda\mathbf{z}, \mu_k), w(x, \Lambda\mathbf{z}, \mu_k); N, M, q + \Lambda\mathbf{z}) = 0$ is satisfied. By the implicit function theorem [13], we have

$$\begin{aligned} &\begin{pmatrix} \nabla_{(x,z)} y(x^{(k)}, \Lambda\mathbf{z}^{(k)}, \mu_k)^T \\ \nabla_{(x,z)} w(x^{(k)}, \Lambda\mathbf{z}^{(k)}, \mu_k)^T \end{pmatrix} \\ = & -\begin{pmatrix} \nabla_y \Phi_{\mu_k}(x^{(k)}, y^{(k)}, w^{(k)}; N, M, q + \Lambda\mathbf{z}^{(k)}) \\ \nabla_w \Phi_{\mu_k}(x^{(k)}, y^{(k)}, w^{(k)}; N, M, q + \Lambda\mathbf{z}^{(k)}) \end{pmatrix}^{-T} \nabla_{(x,z)} \Phi_{\mu_k}(x^{(k)}, y^{(k)}, w^{(k)}; N, M, q + \Lambda\mathbf{z}^{(k)})^T \\ = & -\begin{pmatrix} M + \mu_k I & -I \\ I - D^k & D^k \end{pmatrix}^{-1} \begin{pmatrix} N & I \\ O & O \end{pmatrix}, \end{aligned} \tag{B.9}$$

where

$$\begin{aligned} w^{(k)} &:= N x^{(k)} + M y^{(k)} + q + \Lambda \mathbf{z}^{(k)}, \\ D^k &:= \operatorname{diag}\left(\frac{y^{(k)}[1]}{y^{(k)}[1] + w^{(k)}[1]}, \cdots, \frac{y^{(k)}[m]}{y^{(k)}[m] + w^{(k)}[m]} \right), \end{aligned}$$

and the existence of the inverse matrix follows from Theorem 2.1. Furthermore, since

$$\begin{pmatrix} M + \mu_k I & -I \\ I - D^k & D^k \end{pmatrix}^{-1} = \begin{pmatrix} E^k D^k & E^k \\ -I + (M + \mu_k I) E^k D^k & (M + \mu_k I) E^k \end{pmatrix}$$

with $E^k := \Big(D^k M + I - (1-\mu_k)D^k\Big)^{-1}$, it follows from (B.9) that

$$\nabla_{(x,z)} y(x^{(k)}, \Lambda \mathbf{z}^{(k)}, \mu_k) = -\begin{pmatrix} N^T D^k (E^k)^T \\ D^k (E^k)^T \end{pmatrix}.$$

As a result, we have

$$\begin{aligned} \nabla_x y(x^{(k)}, \Lambda \mathbf{z}^{(k)}, \mu_k) &= -N^T D^k (E^k)^T, \qquad \text{(B.10)} \\ \nabla_z y(x^{(k)}, \Lambda \mathbf{z}^{(k)}, \mu_k) &= -D^k (E^k)^T. \qquad \text{(B.11)} \end{aligned}$$

Thus, from the definition of θ_k, (B.10)–(B.11), and by a straightforward calculus, (B.1) becomes

$$\begin{aligned} 0 \;=\; & \begin{pmatrix} \sum_{\ell=1}^L p_\ell(\nabla_x f(x^{(k)}, y^{(k)}, \omega_\ell) - N^T D^k (E^k)^T \nabla_y f(x^{(k)}, y^{(k)}, \omega_\ell)) \\ p_1 d - \sum_{\ell=1}^L p_\ell D^k (E^k)^T \nabla_y f(x^{(k)}, y^{(k)}, \omega_\ell) \\ \vdots \\ p_L d - \sum_{\ell=1}^L p_\ell D^k (E^k)^T \nabla_y f(x^{(k)}, y^{(k)}, \omega_\ell) \end{pmatrix} \\ & - \begin{pmatrix} N_1^T - N^T D^k (E^k)^T M_1^T & \cdots & N_L^T - N^T D^k (E^k)^T M_L^T \\ I - D^k (E^k)^T M_1^T & \cdots & -D^k (E^k)^T M_L^T \\ -D^k (E^k)^T M_1^T & \cdots & -D^k (E^k)^T M_L^T \\ \vdots & \vdots & \vdots \\ -D^k (E^k)^T M_1^T & \cdots & I - D^k (E^k)^T M_L^T \end{pmatrix} \beta^k \\ & + \begin{pmatrix} \nabla g(x^{(k)}) \\ O \\ \vdots \\ O \end{pmatrix} \lambda^k + \begin{pmatrix} \nabla h(x^{(k)}) \\ O \\ \vdots \\ O \end{pmatrix} \gamma^k - \begin{pmatrix} O & \cdots & O \\ I & \cdots & O \\ \vdots & \ddots & \vdots \\ O & \cdots & I \end{pmatrix} \alpha^k. \end{aligned}$$

Here, $\beta^k \in \Re^{mL}$ is given by $\beta^k := \Big((\beta_1^k)^T, \cdots, (\beta_L^k)^T\Big)^T$ with

$$\beta_\ell^k := 2\rho_k \max\Big(-(N_\ell x^{(k)} + M_\ell y^{(k)} + q_\ell + z_\ell^{(k)}),\ 0\Big)$$

for each ℓ. We then have

$$\begin{aligned} 0 \;=\; & \begin{pmatrix} \sum_{\ell=1}^L p_\ell \nabla_x f(x^{(k)}, y^{(k)}, \omega_\ell) \\ p_1 d \\ \vdots \\ p_L d \end{pmatrix} + \begin{pmatrix} \nabla g(x^{(k)}) \\ O \\ \vdots \\ O \end{pmatrix} \lambda^k + \begin{pmatrix} \nabla h(x^{(k)}) \\ O \\ \vdots \\ O \end{pmatrix} \gamma^k \\ & - \begin{pmatrix} O & \cdots & O \\ I & \cdots & O \\ \vdots & \ddots & \vdots \\ O & \cdots & I \end{pmatrix} \alpha^k - \begin{pmatrix} N_1^T & \cdots & N_L^T \\ I & \cdots & O \\ \vdots & \ddots & \vdots \\ O & \cdots & I \end{pmatrix} \beta^k - \begin{pmatrix} N^T \\ I \\ \vdots \\ I \end{pmatrix} v^k, \end{aligned} \qquad \text{(B.12)}$$

where v^k is defined by

$$v^k := D^k (E^k)^T \sum_{\ell=1}^L \Big(p_\ell \nabla_y f(x^{(k)}, y^{(k)}, \omega_\ell) - M_\ell^T \beta_\ell^k\Big). \qquad \text{(B.13)}$$

Furthermore, by letting

$$\sum_{\ell=1}^L p_\ell \nabla_y f(x^{(k)}, y^{(k)}, \omega_\ell) - \sum_{\ell=1}^L M_\ell^T \beta_\ell^k - u^k - M^T v^k = 0, \qquad \text{(B.14)}$$

we have

$$
\begin{aligned}
u^k &= \sum_{\ell=1}^{L}\Big(p_\ell\nabla_y f(x^{(k)},y^{(k)},\omega_\ell)-M_\ell^T\beta_\ell^k\Big)-M^Tv^k \\
&= \Big((E^k)^{-T}-M^TD^k\Big)(E^k)^T\sum_{\ell=1}^{L}\Big(p_\ell\nabla_y f(x^{(k)},y^{(k)},\omega_\ell)-M_\ell^T\beta_\ell^k\Big) \\
&= \Big(I-(1-\mu_k)D^k\Big)(E^k)^T\sum_{\ell=1}^{L}\Big(p_\ell\nabla_y f(x^{(k)},y^{(k)},\omega_\ell)-M_\ell^T\beta_\ell^k\Big), \qquad \text{(B.15)}
\end{aligned}
$$

where the second equality follows from (B.13). Combining (B.12) and (B.14) yields

$$
\begin{aligned}
0 \;=\; & \begin{pmatrix} \sum_{\ell=1}^L p_\ell\nabla_x f(x^{(k)},y^{(k)},\omega_\ell) \\ \sum_{\ell=1}^L p_\ell\nabla_y f(x^{(k)},y^{(k)},\omega_\ell) \\ p_1 d \\ \vdots \\ p_L d \end{pmatrix} + \begin{pmatrix} \nabla g(x^{(k)}) \\ O \\ O \\ \vdots \\ O \end{pmatrix}\lambda^k + \begin{pmatrix} \nabla h(x^{(k)}) \\ O \\ O \\ \vdots \\ O \end{pmatrix}\gamma^k \\
& - \begin{pmatrix} O & \cdots & O \\ O & \cdots & O \\ I & \cdots & O \\ \vdots & \ddots & \vdots \\ O & \cdots & I \end{pmatrix}\alpha^k - \begin{pmatrix} N_1^T & \cdots & N_L^T \\ M_1^T & \cdots & M_L^T \\ I & \cdots & O \\ \vdots & \ddots & \vdots \\ O & \cdots & I \end{pmatrix}\beta^k - \begin{pmatrix} O \\ I \\ O \\ \vdots \\ O \end{pmatrix}u^k - \begin{pmatrix} N^T \\ M^T \\ I \\ \vdots \\ I \end{pmatrix}v^k. \qquad \text{(B.16)}
\end{aligned}
$$

We can show that, when k is large sufficiently,

$$\beta_\ell^k[i]=0 \quad \text{as long as} \quad i\notin\mathcal{I}_{W_\ell}^*. \qquad \text{(B.17)}$$

In fact, if $i\notin\mathcal{I}_{W_\ell}^*$, namely, $(N_\ell x^*+M_\ell y^*+q_\ell+z_\ell^*)[i]>0$, then, when k is large enough, there must hold $(N_\ell x^{(k)}+M_\ell y^{(k)}+q_\ell+z_\ell^{(k)})[i]>0$ and hence

$$\beta_\ell^k[i]=2\rho_k\max\Big\{-(N_\ell x^{(k)}+M_\ell y^{(k)}+q_\ell+z_\ell^{(k)})[i],\ 0\Big\}=0.$$

Taking into account (B.2)–(B.7), we can rewrite (B.16) as

$$
\begin{aligned}
& \begin{pmatrix} \sum_{\ell=1}^L p_\ell\nabla_x f(x^{(k)},y^{(k)},\omega_\ell) \\ \sum_{\ell=1}^L p_\ell\nabla_y f(x^{(k)},y^{(k)},\omega_\ell) \\ p_1 d \\ \vdots \\ p_L d \end{pmatrix} - \sum_{i\notin\mathcal{I}_Y^*} u^k[i]\begin{pmatrix} 0 \\ e_i \\ 0 \\ \vdots \\ 0 \end{pmatrix} - \sum_{i\notin\mathcal{I}_W^*} v^k[i]\begin{pmatrix} N[i] \\ M[i] \\ e_i \\ \vdots \\ e_i \end{pmatrix} \\
=\; & -\sum_{i\in\mathcal{I}_g(x^*)}\lambda^k[i]\begin{pmatrix} \nabla g_i(x^{(k)}) \\ 0 \\ \vdots \\ 0 \end{pmatrix} - \sum_{i=1}^{s_2}\gamma^k[i]\begin{pmatrix} \nabla h_i(x^{(k)}) \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \sum_{\ell=1}^{L}\sum_{i\in\mathcal{I}_{Z_\ell}^*}\alpha_\ell^k[i]\begin{pmatrix} 0 \\ \vdots \\ e_i \\ \vdots \\ 0 \end{pmatrix} \\
& + \sum_{\ell=1}^{L}\sum_{i\in\mathcal{I}_{W_\ell}^*}\beta_\ell^k[i]\begin{pmatrix} N_\ell[i] \\ M_\ell[i] \\ \vdots \\ e_i \\ \vdots \\ 0 \end{pmatrix} + \sum_{i\in\mathcal{I}_Y^*} u^k[i]\begin{pmatrix} 0 \\ e_i \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \sum_{i\in\mathcal{I}_W^*} v^k[i]\begin{pmatrix} N[i] \\ M[i] \\ e_i \\ \vdots \\ e_i \end{pmatrix}. \qquad \text{(B.18)}
\end{aligned}
$$

We next prove

$$i \notin \mathcal{I}_Y^* \quad \Rightarrow \quad \lim_{k\to\infty} u^k[i] = 0, \tag{B.19}$$

$$i \notin \mathcal{I}_W^* \quad \Rightarrow \quad \lim_{k\to\infty} v^k[i] = 0. \tag{B.20}$$

To this end, since (B.15) and (B.13) imply that

$$u^k[i] \;=\; \frac{\mu_k y^{(k)}[i] + w^{(k)}[i]}{y^{(k)}[i] + w^{(k)}[i]} e_i^T (E^k)^T \sum_{\ell=1}^{L} \Big(p_\ell \nabla_y f(x^{(k)}, y^{(k)}, \omega_\ell) - M_\ell^T \beta_\ell^k\Big), \tag{B.21}$$

$$v^k[i] \;=\; \frac{y^{(k)}[i]}{y^{(k)}[i] + w^{(k)}[i]} e_i^T (E^k)^T \sum_{\ell=1}^{L} \Big(p_\ell \nabla_y f(x^{(k)}, y^{(k)}, \omega_\ell) - M_\ell^T \beta_\ell^k\Big), \tag{B.22}$$

it is enough to show that the sequence

$$\Big\{\Big\|(E^k)^T \sum_{\ell=1}^{L} \Big(p_\ell \nabla_y f(x^{(k)}, y^{(k)}, \omega_\ell) - M_\ell^T \beta_\ell^k\Big)\Big\|\Big\}$$

is bounded. For the purpose of contradiction, taking a further subsequence if necessary, we assume

$$\lim_{k\to\infty} \Big\|(E^k)^T \sum_{\ell=1}^{L} \Big(p_\ell \nabla_y f(x^{(k)}, y^{(k)}, \omega_\ell) - M_\ell^T \beta_\ell^k\Big)\Big\| = +\infty. \tag{B.23}$$

Note that (B.21) implies

$$|u^k[i]| \le \frac{\mu_k y^{(k)}[i] + w^{(k)}[i]}{y^{(k)}[i] + w^{(k)}[i]} \Big\|(E^k)^T \sum_{\ell=1}^{L} \Big(p_\ell \nabla_y f(x^{(k)}, y^{(k)}, \omega_\ell) - M_\ell^T \beta_\ell^k\Big)\Big\|$$

for each i and k. In consequence, we have

$$\begin{aligned} i \notin \mathcal{I}_Y^* \quad &\Rightarrow \quad \lim_{k\to\infty} \frac{\mu_k y^{(k)}[i] + w^{(k)}[i]}{y^{(k)}[i] + w^{(k)}[i]} = 0 \\ &\Rightarrow \quad \lim_{k\to\infty} \frac{|u^k[i]|}{\|(E^k)^T \sum_{\ell=1}^{L} \Big(p_\ell \nabla_y f(x^{(k)}, y^{(k)}, \omega_\ell) - M_\ell^T \beta_\ell^k\Big)\|} = 0. \end{aligned} \tag{B.24}$$

Similarly, we can show that

$$i \notin \mathcal{I}_W^* \quad \Rightarrow \quad \lim_{k\to\infty} \frac{|v^k[i]|}{\|(E^k)^T \sum_{\ell=1}^{L} \Big(p_\ell \nabla_y f(x^{(k)}, y^{(k)}, \omega_\ell) - M_\ell^T \beta_\ell^k\Big)\|} = 0. \tag{B.25}$$

Let d^k denote the vector on the right-hand side of equality (B.18). It then follows from (B.18) and (B.23)–(B.25) that

$$\lim_{k\to\infty} \frac{d^k}{\|(E^k)^T \sum_{\ell=1}^{L} \Big(p_\ell \nabla_y f(x^{(k)}, y^{(k)}, \omega_\ell) - M_\ell^T \beta_\ell^k\Big)\|} = 0. \tag{B.26}$$

Since the MPEC-LICQ holds at $(x^*, y^*, \mathbf{y}^*, \mathbf{z}^*)$ in problem (3.3), the vectors on the right-hand side of (B.18) are linearly independent when k is sufficiently large and so, by (B.26), all the sequences generated by dividing the multipliers that appear on the right-hand side of (B.18) by the number

$$\left\|(E^k)^T \sum_{\ell=1}^{L} \left(p_\ell \nabla_y f(x^{(k)}, y^{(k)}, \omega_\ell) - M_\ell^T \beta_\ell^k\right)\right\|$$

are convergent to 0 as $k \to \infty$. This fact, together with (B.24) and (B.25), implies that

$$\lim_{k\to\infty} \frac{u^k[i]}{\|(E^k)^T \sum_{\ell=1}^{L} \left(p_\ell \nabla_y f(x^{(k)}, y^{(k)}, \omega_\ell) - M_\ell^T \beta_\ell^k\right)\|} = 0, \tag{B.27}$$

$$\lim_{k\to\infty} \frac{v^k[i]}{\|(E^k)^T \sum_{\ell=1}^{L} \left(p_\ell \nabla_y f(x^{(k)}, y^{(k)}, \omega_\ell) - M_\ell^T \beta_\ell^k\right)\|} = 0 \tag{B.28}$$

hold for any i. However, noticing that

$$\begin{aligned} & \frac{u^k[i] + v^k[i]}{\|(E^k)^T \sum_{\ell=1}^{L} \left(p_\ell \nabla_y f(x^{(k)}, y^{(k)}, \omega_\ell) - M_\ell^T \beta_\ell^k\right)\|} \\ = \quad & \left(1 + \frac{\mu_k y^{(k)}[i]}{y^{(k)}[i] + w^{(k)}[i]}\right) \frac{e_i^T (E^k)^T \sum_{\ell=1}^{L} \left(p_\ell \nabla_y f(x^{(k)}, y^{(k)}, \omega_\ell) - M_\ell^T \beta_\ell^k\right)}{\|(E^k)^T \sum_{\ell=1}^{L} \left(p_\ell \nabla_y f(x^{(k)}, y^{(k)}, \omega_\ell) - M_\ell^T \beta_\ell^k\right)\|} \end{aligned}$$

holds for any i and k, there exists an index $\hat{i}$ such that

$$\lim_{k\to\infty} \frac{|u^k[\hat{i}] + v^k[\hat{i}]|}{\|(E^k)^T \sum_{\ell=1}^{L} \left(p_\ell \nabla_y f(x^{(k)}, y^{(k)}, \omega_\ell) - M_\ell^T \beta_\ell^k\right)\|} \geq \frac{1}{\sqrt{m}} \lim_{k\to\infty} \left(1 + \frac{\mu_k y^{(k)}[\hat{i}]}{y^{(k)}[\hat{i}] + w^{(k)}[\hat{i}]}\right) = \frac{1}{\sqrt{m}} > 0.$$

This contradicts (B.27) and (B.28). As a result, the implications (B.19) and (B.20) are true.

Consider equality (B.18) again. By (B.19) and (B.20), the left-hand side of equality (B.18) is convergent. Moreover, from the assumption that the MPEC-LICQ holds at $(x^*, y^*, \mathbf{y}^*, \mathbf{z}^*)$, we can prove that all the sequences of the multipliers that appear on the right-hand side of (B.18) are bounded. In fact, by letting $u_\ell^k := \left(I - (1-\mu_k)D^k\right)(E^k)^T \left(p_\ell \nabla_y f(x^{(k)}, y^{(k)}, \omega_\ell) - M_\ell^T \beta_\ell^k\right)$ for $\ell = 1, \cdots, L$,

$$\mathbf{v}^k := \begin{pmatrix} v^k \\ \vdots \\ v^k \end{pmatrix} \in \Re^{mL}, \quad \mathbf{u}^k := \begin{pmatrix} u_1^k \\ \vdots \\ u_L^k \end{pmatrix},$$

and

$$\begin{aligned} \mathbf{a}^k &:= \mathbf{u}^k + \mathbf{M}^T(\beta^k + \mathbf{v}^k), \\ \mathbf{b}^k &:= \mathbf{u}^k, \\ \mathbf{c}^k &:= \beta^k + \mathbf{v}^k, \end{aligned}$$

we obtain from (B.16) that

$$\begin{aligned} 0 \;=\; & \begin{pmatrix} \sum_{\ell=1}^{L} p_\ell \nabla_x f(x^{(k)}, y^{(k)}, \omega_\ell) \\ \sum_{\ell=1}^{L} p_\ell \nabla_y f(x^{(k)}, y^{(k)}, \omega_\ell) \\ 0 \\ \mathbf{d} \end{pmatrix} + \begin{pmatrix} \nabla g(x^{(k)}) \\ O \\ O \\ O \end{pmatrix} \lambda^k + \begin{pmatrix} \nabla h(x^{(k)}) \\ O \\ O \\ O \end{pmatrix} \gamma^k \\ & - \begin{pmatrix} O \\ O \\ O \\ I \end{pmatrix} \alpha^k + \begin{pmatrix} O \\ -\mathbf{D}^T \\ I \\ O \end{pmatrix} \mathbf{a}^k - \begin{pmatrix} O \\ O \\ I \\ O \end{pmatrix} \mathbf{b}^k - \begin{pmatrix} \mathbf{N}^T \\ O \\ \mathbf{M}^T \\ I \end{pmatrix} \mathbf{c}^k. \end{aligned} \tag{B.29}$$

Applying (B.19) and (B.20), it is not difficult to show that

$$\begin{aligned} \mathbf{y}^*[i] > 0 \quad &\Rightarrow \quad \lim_{k\to\infty} \mathbf{u}^k[i] = 0, \\ (\mathbf{N}x^* + \mathbf{M}\mathbf{y}^* + \mathbf{q} + \mathbf{z}^*)[i] > 0 \quad &\Rightarrow \quad \lim_{k\to\infty} \mathbf{v}^k[i] = 0. \end{aligned}$$

From (B.8), (B.17), and the MPEC-LICQ assumption, we see that all the sequences of the multiplier vectors in (B.29) are bounded, which implies the boundedness of the multiplier sequences that appear on the right-hand side of (B.18). In consequence, assuming these vector sequences are all convergent without loss of generality and letting $k \to \infty$ in (B.18), we obtain the equality corresponding to (2.3).

In addition, since both $y^{(k)}[i]$ and $w^{(k)}[i]$ are positive, we have from (B.21) and (B.22) that

$$u^k[i]v^k[i] \geq 0, \quad i = 1, \cdots, m.$$

This yields the results corresponding to (2.4). Therefore, $(x^*, y^*, \mathbf{z}^*)$ is a C-stationary point of problem (3.1). If, in addition, y^* satisfies the strict complementarity condition, then $(x^*, y^*, \mathbf{z}^*)$ is a S-stationary point by the definitions of C-stationarity and S-stationarity. This completes the proof of Theorem 3.2. $\square$