EQUILIBRIUM PROBLEMS WITH EQUILIBRIUM CONSTRAINTS: STATIONARITIES, ALGORITHMS, AND APPLICATIONS

A DISSERTATION

SUBMITTED TO THE DEPARTMENT OF

MANAGEMENT SCIENCE AND ENGINEERING

AND THE COMMITTEE ON GRADUATE STUDIES

OF STANFORD UNIVERSITY

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

Che-Lin Su

September 2005

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

Richard W. Cottle
(Principal Advisor)

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

Hung-Po Chao

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

Kenneth L. Judd

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

Michael A. Saunders

Approved for the University Committee on Graduate Studies.

Abstract

Equilibrium problem with equilibrium constraints (EPECs) often arise in engineering and economics applications. One important application of EPECs is the multi-leader-follower game in economics, where each leader is solving a Stackelberg game formulated as a mathematical program with equilibrium constraints (MPEC). Motivated by applied EPEC models for studying the strategic behavior of generating firms in deregulated electricity markets, the aim of this thesis is to study theory, algorithms, and new applications for EPECs.

We begin by reviewing the stationarity conditions and algorithms for MPECs. Then, we generalize Scholtes's regularization scheme for solving MPECs. We define EPEC stationarity concepts in Chapter 3. We propose a sequential nonlinear complementarity (SNCP) method for solving EPECs and establish its convergence. We present the numerical results of the SNCP method and give a comparison with two best-reply iterations, nonlinear Jacobi and nonlinear Gauss-Seidel, on a set of randomly generated test problems. The computational experience to date shows that both the SNCP algorithm and the nonlinear Gauss-Seidel method outperform the nonlinear Jacobi method.

We investigate the issue of existence of an EPEC solution in Chapter 4. In general, an EPEC solution may not exist because of nonconvexity of the associated MPECs. However, we show that the existence result can be established for the spot-forward market model proposed by Allaz and Vila and the two-period Cournot game studied by Saloner. We observe that the mathematical structure of the spot-forward market model is similar to that of the multiple leader Stackelberg model analyzed by Sherali when new variables are introduced for spot market sales. Consequently, we are able to adapt Sherali's analysis to establish the existence of a forward market equilibrium for M asymmetric producers with nonidentical linear cost functions.

In Chapter 5, we present a novel MPEC approach for computing solutions of incentive problems in economics. Specifically, we consider deterministic contracts as well as contracts with action and/or compensation lotteries for the static moral-hazard problem, and formulate each case as an MPEC. We propose a hybrid procedure that combines the best features of the MPEC approach and the LP lottery approach. The hybrid procedure obtains a solution that is, if not global, at least as good as an LP solution. It also preserves the fast local convergence property by applying an SQP algorithm to MPECs. Numerical results on an example show that the hybrid procedure outperforms the LP approach in both computational time and solution quality in terms of the optimal objective value.

Acknowledgments

Many Ph.D. students said that their thesis would not be completed if not for their thesis advisors. What is even more true for me is that my thesis would not even be started if not for my thesis advisor, Professor Richard W. Cottle. When I first came to Stanford, all I knew about optimization was from an undergraduate linear programming course. I did not have much research or any teaching experience, either. Professor Cottle did not mind to work with a rookie like me. He greatly encouraged me to transfer from the Master to the Ph.D. program after I took two Ph.D.-level courses with him (one of which being the famous two-person class on linear complementarity problems). Over the years, he patiently taught me everything I know, not only on optimization theory, but also the right attitude on conducting research and toward life. He diligently and carefully read my weekly research reports for almost two years, and after that, read each of my three working papers uncountably many times. His critical thinking and invaluable guidance helped me improve the quality of my research. His support of my discussions and collaboration with other researchers has been tremendous. My work has been constantly motivated by one goal: to achieve the high standards he has set for me, and it will continue to be so in my future academic career. For everything that they have done for me during my years at Stanford, I am deeply indebted to him and his wife, Sue Cottle.

I would like to express my gratitude to Professor Michael A. Saunders, who always makes his time for me from his busy schedule. His exceptionally careful readings of my papers and this thesis and his detailed comments greatly helped to polish my work. I will always remember his cheerful spirit and will miss his companionship in the late night hours, even during the Christmas and New Year holidays!

I thank Professor Hung-Po Chao for serving on my oral and reading committee. His expertise on electric power industry and his insights on economic models in general provided a different perspective on my research. I am also grateful for the

financial support from Electric Power Research Institute during the early stages of my research. It allowed me to build the theoretical foundation of this thesis in Chapter 2 and Chapter 3.

Special thanks are due to Dr. Kenneth L. Judd, whose influence on this thesis was far greater than I could have ever expected. Starting in early 2003, I spent about a year and a half unsuccessfully looking for new applications of MPECs and EPECs in economics and operations management. At the suggestion of Professor Jong-Shi Pang in summer 2004, I went to Ken to discuss my research. Ken provided the last missing piece of the puzzle by suggesting the moral-hazard problem. Chapter 5 represents joint work with him. However, Ken's influence and generosity did not stop there. He has provided me numerous opportunities and much professional advice to start my academic career. I am thankful for my pleasant association with him in this past year, and I look forward to our continued collaboration in the future.

I thank Professor Robert B. Wilson for chairing and Professor Feryal Erhun for serving on my oral defense committee. The Jerome Kaseberg Doolan Fellowship from Stanford University in this past year allowed me to investigate new research topics in computational economics.

The Systems Optimization Laboratory has been a great working environment. I appreciate the friendship of Kien-Ming Ng, Michael Friedlander, Byunggyoo Kim, Maureen Doyle, Alexis Collomb, Uday Shanbag and Kaustuv. Fellow OR students, including Peter Salzman, Yen Lin Chia, Sanatan Rai, Hernan Awad, Yuqing Sun, Kahn Mason and Jiawei Zhang, also made my life at Stanford an enriching and enjoyable experience. Jocelyn was the sole source of emotional support in my early years at Stanford, and I am deeply indebted to her.

Last but definitely not the least, I thank my family for their unconditional support in all aspects of my life. This thesis is dedicated to them.

Contents

Abstract **vi**

Acknowledgments **viii**

List of Tables and Figures **xi**

Notation **xiii**

1 Introduction **1**

2 Mathematical Program with Equilibrium Constraints **5**

2.1 Preliminary on MPECs 5

2.2 A Generalization of Scholtes's Regularization 10

3 Equilibrium Problem with Equilibrium Constraints **19**

3.1 Formulation and stationarity conditions 19

3.2 Algorithms for Solving EPECs 22

3.2.1 Diagonalization methods 22

3.2.2 Sequential NCP method 25

3.3 Implementation and Numerical Comparison 27

3.3.1 Discussion of numerical results 32

4 Forward-Spot Market Model **35**

4.1 The Two-Period Forward Market Model 36

4.1.1 The production game 36

4.1.2 Forward market equilibrium 37

4.2 Existence of a Forward Market Equilibrium 38

4.3 An EPEC Approach for Computing a Forward Market Equilibrium 45

4.3.1 An example with three producers 47

4.4 Conclusions 49

5 Computation of Moral-Hazard Problems **51**
5.1 Introduction 51
5.2 The Basic Moral-Hazard Model 54
5.2.1 The deterministic contract 54
5.2.2 A mixed-integer NLP formulation 56
5.2.3 An MPEC formulation 57
5.3 Moral-Hazard Problems with Lotteries 60
5.3.1 The contract with action lotteries 60
5.3.2 An NLP formulation with star-shaped feasible region . . . 61
5.3.3 MPEC formulations 63
5.3.4 The contract with compensation lotteries 64
5.3.5 The contract with action and compensation lotteries . . . 67
5.3.6 Linear programming approximation 68
5.4 A Hybrid Approach toward Global Solution 71
5.5 An Example and Numerical Results 72
5.6 Conclusions and Future Work 76

Bibliography **77**

Tables and Figures

Tables

3.1 Input parameters 28
3.2 Notation for numerical results 29
3.3 Information on test problems 29
3.4 Results for SNCP method 30
3.5 Results for Gauss-Seidel with 10^{-6} convergence accuracy 30
3.6 Results for Jacobi with 10^{-6} convergence accuracy 31
3.7 Results for Gauss-Seidel with 10^{-4} convergence accuracy 31
3.8 Results for Jacobi with 10^{-4} convergence accuracy 32

4.1 Notation used for computational results on forward market model 48
4.2 Computational results on forward market model 48

5.1 Problem characteristics of various approaches 73
5.2 The value of parameters used in Example 1 73
5.3 MINLP solutions for Example 1 74
5.4 MPEC solutions for Example 1 74
5.5 LP solutions for Example 1 75
5.6 Hybrid solutions for Example 1 75

Figures

5.1 Feasible region of an MPEC 59
5.2 Star-shaped feasible region 63

Notation

R^n	real n-dimensional space
$x\,(\in R^n)$	an n-dimensional vector
x^{T}	the transpose of a vector x
$\{x^\nu\}$	a sequence of vectors $x^1, x^2, x^3, \ldots$
$\|x\|$	norm of x (2-norm unless otherwise stated)
e_m	an m-dimensional vectors of all ones
	(m is sometimes omitted)
$x^{\mathrm{T}}y$	the standard inner product of vectors x and $y \in R^n$
$x \perp y$	orthogonality of vectors x and $y \in R^n$
$x \circ y$	$(x_i y_i)$, the Hadamard product of vectors x and $y \in R^n$
$f(x)$	objective function (a scalar)
∇f	gradient of f (a column vector)
$\nabla^2 f$	Hessian of f (an $n \times n$ matrix)
$g : R^n \to R^m$	a mapping from domain R^n into range R^m ($m \geq 2$)
∇g	the $m \times n$ Jacobian of a mapping $g : R^n \to R^m$ ($m \geq 2$)
$\mathcal{I}_g(\bar{x})$	the set of active constraints at $\bar{x}$ for $g(x) \geq 0$
$\mathcal{I}_g^c(\bar{x})$	the set of nonactive constraints at $\bar{x}$ for $g(x) \geq 0$

Chapter 1

Introduction

An *equilibrium problem with equilibrium constraints* (EPEC) is a member of a new class of mathematical programs that often arise in engineering and economics applications. One important application of EPECs is the multi-leader-follower game [47] in economics. In noncooperative game theory, the well-known Stackelberg game (single-leader-multi-follower game) can be formulated as an optimization problem called a *mathematical program with equilibrium constraints* (MPEC) [33, 43], in which followers' optimal strategies are solutions of complementarity problems or variational inequality problems based on the leader's strategies. Analogously, the more general problem of finding equilibrium solutions of a multi-leader-follower game, where each leader is solving a Stackelberg game, is formulated as an EPEC. Consequently, one may treat an EPEC as a two-level hierarchical problem, which involves finding equilibria at both lower and upper levels. More generally, an EPEC is a mathematical program to find equilibria that simultaneously solve several MPECs, each of which is parameterized by decision variables of other MPECs.

Motivated by applied EPEC models for studying the strategic behavior of generating firms in deregulated electricity markets [5, 22, 49], the aim of this thesis is to study the stationarities, algorithms, and new applications for EPECs. The main contributions of this thesis are summarized below.

In Chapter 2, we review the stationarity conditions and algorithms for mathematical programs with equilibrium constraints. We generalize Scholtes's regularization scheme for solving MPECs and establish its convergence. Chapter 3 begins by defining EPEC stationarities, followed by a new class of algorithm, called a *sequential nonlinear complementarity* (SNCP) method, which we propose for solving EPECs. Numerical approaches used by researchers in engineering and economics fields to solve EPEC models fall into the category of *diagonalization methods* (or the *best-reply iteration* in the economics literature), a cyclic procedure that solves the MPECs one at a time. The main drawback with this procedure is that it may fail to find an EPEC solution even if one exists. To avoid this disadvantage, we propose a sequential nonlinear complementarity (SNCP) algorithm for solving EPECs and establish the convergence of this algorithm. We present the numerical results of the SNCP method and give a comparison with two best-reply iteration methods, nonlinear Jacobi and nonlinear Gauss-Seidel, on a set of EPEC test problems, randomly generated by *EPECgen*, a MATLAB program we develop to study the numerical performance of various algorithms. The computational experience to date shows that both the SNCP algorithm and the nonlinear Gauss-Seidel method outperform the nonlinear Jacobi method.

In Chapter 4, we investigate the issue of existence of an EPEC solution. In general, an EPEC solution may not exist because of nonconvexity of the associated MPECs. However, we show that the existence result can be established for specific models; for example, the forward-spot market model proposed by Allaz and Vila [1] or the two-period Cournot game studied by Saloner [56, 57] and Pal [44]. In [1], Allaz and Vila showed that one can solve for the forward market Nash equilibrium in closed-form when the inverse demand function is affine and the producers have identical linear cost functions. However, their approach is not valid for the case of producers with nonidentical linear cost functions. As a result, the existence of a forward market Nash equilibrium is in jeopardy. By introducing additional variables for the spot market sales, we observe that the two-period forward market model has structure similar to that of the multiple leader Stackelberg model analyzed by Sherali [59]. Consequently, we can adapt Sherali's analysis to establish the existence of a forward market equilibrium when the M producers have different linear cost functions. We further demonstrate the use of the SNCP method to solve the EPEC formulation of this model on an example with three asymmetric producers.

The aim of Chapter 5 is to extend the applicability of MPECs and EPECs to new research fields. One particular topic we study is incentive problems in economics. As one of the most active research topics in economics in the past three decades, incentive problems such as contract design, optimal taxation and regulation, and multiproduct pricing, can be naturally formulated as an MPEC or an EPEC. In Chapter 5, we present a novel MPEC approach for computing solutions of the static moral-hazard problem. We consider deterministic contracts as well as contracts with action and/or compensation lotteries, and formulate each case as an MPEC. We investigate and compare solution properties of the MPEC approach to those of the linear programming (LP) approach with lotteries. We propose a hybrid procedure that combines the best features of both. The hybrid procedure obtains a solution that is, if not global, at least as good as an LP solution. It also preserves the fast local convergence property by applying an SQP algorithm to MPECs. Numerical results on an example show that the hybrid procedure outperforms the LP approach in both computational time and solution quality in terms of the optimal objective value.

Chapter 2

Mathematical Program with Equilibrium Constraints

In this chapter, we first define mathematical programs with equilibrium constraints (MPECs). We propose a new regularization scheme for MPECs that includes Scholtes's approach [62] and the two-sided relaxation scheme proposed by DeMiguel, Friedlander, Nogales, and Scholtes [9] as special cases. We show that the convergence theorems studied in [62] can be carried over to our approach.

2.1 Preliminary on MPECs

We consider an MPEC formulated as a nonlinear program with complementarity constraints:

$$\begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & g(x) \leq 0, \quad h(x) = 0, \\ & 0 \leq G(x) \perp H(x) \geq 0, \end{array} \tag{2.1}$$

where $f : R^n \to R,\ g : R^n \to R^p,\ h : R^n \to R^q,\ G : R^n \to R^m$, and $H : R^n \to R^m$ are twice continuously differentiable functions. Given a feasible vector $\bar{x}$ of the MPEC (2.1), we define the following index sets of active and inactive constraints:

$$\begin{array}{rlrl} \mathcal{I}_G(\bar{x}) := & \{i \mid G_i(\bar{x}) = 0\}, & \mathcal{I}_G^c(\bar{x}) := & \{i \mid G_i(\bar{x}) > 0\}, \\ \mathcal{I}_H(\bar{x}) := & \{i \mid H_i(\bar{x}) = 0\}, & \mathcal{I}_H^c(\bar{x}) := & \{i \mid H_i(\bar{x}) > 0\}, \\ \mathcal{I}_{GH}(\bar{x}) := & \{i \mid G_i(\bar{x}) = H_i(\bar{x}) = 0\}, & \mathcal{I}_g(\bar{x}) := & \{i \mid g_i(\bar{x}) = 0\}, \end{array} \tag{2.2}$$

where $\mathcal{I}_{GH}(\bar{x})$ is known as the *degenerate* set. If $\mathcal{I}_{GH}(\bar{x}) = \emptyset$, then the feasible vector $\bar{x}$ is said to fulfill the *strict complementarity* conditions.

Associated with any given feasible vector $\bar{x}$ of MPEC (2.1), there is a nonlinear program, called the *tightened* NLP (TNLP($\bar{x}$)) [46, 58]:

$$\begin{aligned} \text{minimize} \quad & f(x) \\ \text{subject to} \quad & g(x) \leq 0, \quad h(x) = 0, \\ & G_i(x) = 0, \quad i \in \mathcal{I}_G(\bar{x}), \\ & G_i(x) \geq 0, \quad i \in \mathcal{I}_G^c(\bar{x}), \\ & H_i(x) = 0, \quad i \in \mathcal{I}_H(\bar{x}), \\ & H_i(x) \geq 0, \quad i \in \mathcal{I}_H^c(\bar{x}). \end{aligned} \tag{2.3}$$

Similarly, there is a *relaxed* NLP (RNLP($\bar{x}$)) [46, 58] defined as follows:

$$\begin{aligned} \text{minimize} \quad & f(x) \\ \text{subject to} \quad & g(x) \leq 0, \quad h(x) = 0, \\ & G_i(x) = 0, \quad i \in \mathcal{I}_H^c(\bar{x}), \\ & G_i(x) \geq 0, \quad i \in \mathcal{I}_H(\bar{x}), \\ & H_i(x) = 0, \quad i \in \mathcal{I}_G^c(\bar{x}), \\ & H_i(x) \geq 0, \quad i \in \mathcal{I}_G(\bar{x}). \end{aligned} \tag{2.4}$$

It is well known that an MPEC cannot satisfy the standard constraint qualifications, such as *linear independence constraint qualification* (LICQ) or *Mangasarian-Fromovitz constraint qualification* (MFCQ), at any feasible point [6, 58]. This implies that the classical KKT theorem on necessary optimality conditions (with the assumption that LICQ or MFCQ is satisfied at local minimizers) is not appropriate in the context of MPECs. One then needs to develop suitable variants of CQs and concepts of stationarity for MPECs. Specifically, the MPEC-CQs are closely related to those of the RNLP (2.4).

Definition 2.1. The MPEC (2.1) is said to satisfy the *MPEC-LICQ (MPEC-MFCQ)* at a feasible point $\bar{x}$ if the corresponding RNLP($\bar{x}$) (2.4) satisfies the

LICQ (MFCQ) at $\bar{x}$.

In what follows, we define B(ouligand)-stationarity for MPECs. We also summarize various stationarity concepts for MPECs introduced in Scheel and Scholtes [58].

Definition 2.2. Let $\bar{x}$ be a feasible point for the MPEC (2.1). We say that $\bar{x}$ is a *Bouligand-* or *B-stationary* point if $d = 0$ solves the following linear program with equilibrium constraints (LPEC) with the vector $d \in R^n$ being the decision variable:

$$\begin{array}{ll} \text{minimize} & \nabla f(\bar{x})^{\mathrm{T}} d \\ \text{subject to} & g(\bar{x}) + \nabla g(\bar{x})^{\mathrm{T}} d \leq 0, \quad h(\bar{x}) + \nabla h(\bar{x})^{\mathrm{T}} d = 0, \\ & 0 \leq G(\bar{x}) + \nabla G(\bar{x})^{\mathrm{T}} d \perp H(\bar{x}) + \nabla H(\bar{x})^{\mathrm{T}} d \geq 0. \end{array} \tag{2.5}$$

B-stationary points are good candidates for local minimizers of the MPEC (2.1). However, checking B-stationarity is difficult because it may require checking the optimality of $2^{|\mathcal{I}_{GH}(\bar{x})|}$ linear programs [33, 58].

Definition 2.3. We define the *MPEC Lagrangian* with the vector of *MPEC multipliers* $\lambda = (\lambda^g, \lambda^h, \lambda^G, \lambda^H)$ as in Scheel and Scholtes [58] and Scholtes [62]:

$$\mathcal{L}(x, \lambda) = f(x) + (\lambda^g)^{\mathrm{T}} g(x) + (\lambda^h)^{\mathrm{T}} h(x) - (\lambda^G)^{\mathrm{T}} G(x) - (\lambda^H)^{\mathrm{T}} H(x). \tag{2.6}$$

Notice that the complementarity constraint $G(x)^{\mathrm{T}} H(x) = 0$ does not appear in the MPEC Lagrangian function. This special feature distinguishes MPECs from standard nonlinear programming problems.

The following four concepts of MPEC stationarity, stated in increasing strength, are introduced in Scheel and Scholtes [58].

Definition 2.4. A feasible point $\bar{x}$ of the MPEC (2.1) is called *weakly stationary* if there exists a vector of MPEC multipliers $\bar{\lambda} = (\bar{\lambda}^g, \bar{\lambda}^h, \bar{\lambda}^G, \bar{\lambda}^H)$ such that $(\bar{x}, \bar{\lambda})$ is a KKT stationary point of the TNLP (2.3), i.e., $(\bar{x}, \bar{\lambda})$ satisfies the following conditions:

$$
\begin{aligned}
&\nabla_x \mathcal{L}(\bar{x}, \bar{\lambda}) = \\
&\nabla f(\bar{x}) + \nabla g(\bar{x})^{\mathrm{T}} \bar{\lambda}^g + \nabla h(\bar{x})^{\mathrm{T}} \bar{\lambda}^h - \nabla G(\bar{x})^{\mathrm{T}} \bar{\lambda}^G - \nabla H(\bar{x})^{\mathrm{T}} \bar{\lambda}^H = 0,
\end{aligned}
$$

$$
\begin{array}{lr}
 & h(\bar{x}) = 0; \quad g(\bar{x}) \le 0, \quad \bar{\lambda}^g \ge 0, \quad (\bar{\lambda}^g)^{\mathrm{T}} g(\bar{x}) = 0, \\
i \in \mathcal{I}_G(\bar{x}) : & G_i(\bar{x}) = 0, \\
i \in \mathcal{I}_G^c(\bar{x}) : & G_i(\bar{x}) \ge 0, \quad \bar{\lambda}_i^G \ge 0, \quad \bar{\lambda}_i^G G_i(\bar{x}) = 0, \\
i \in \mathcal{I}_H(\bar{x}) : & H_i(\bar{x}) = 0, \\
i \in \mathcal{I}_H^c(\bar{x}) : & H_i(\bar{x}) \ge 0, \quad \bar{\lambda}_i^H \ge 0, \quad \bar{\lambda}_i^H H_i(\bar{x}) = 0.
\end{array}
\tag{2.7}
$$

In addition, the feasible vector $\bar{x}$ is called

(a) a *C(larke)-stationary point* if $\bar{\lambda}_i^G \bar{\lambda}_i^H \ge 0 \quad \forall\, i \in \mathcal{I}_{GH}(\bar{x})$.

(b) a *M(ordukhovich)-stationary point* if either ($\bar{\lambda}_i^G > 0$, $\bar{\lambda}_i^H > 0$) or ($\bar{\lambda}_i^G \bar{\lambda}_i^H = 0$) $\forall\, i \in \mathcal{I}_{GH}(\bar{x})$.

(c) a *strongly stationary point* if $\bar{\lambda}_i^G \ge 0, \bar{\lambda}_i^H \ge 0 \quad \forall\, i \in \mathcal{I}_{GH}(\bar{x})$.

Notice that by Definition 2.4, a point $\bar{x}$ is a strongly stationary point of the MPEC (2.1) if $(\bar{x}, \bar{\lambda})$ is a KKT pair of the RNLP (2.4).

Definition 2.5 (Upper-level strict complementarity (ULSC)). A weakly stationary point $\bar{x}$ is said to satisfy ULSC if there exist MPEC multipliers $\bar{\lambda} = (\bar{\lambda}^g, \bar{\lambda}^h, \bar{\lambda}^G, \bar{\lambda}^H)$ satisfying (2.7) with $\bar{\lambda}_i^G \bar{\lambda}_i^H \neq 0$ for all $i \in \mathcal{I}_{GH}(\bar{x})$.

See [58] for a discussion of these various stationarity conditions and their relations to others in the literature such as Clarke's generalized stationarity.

The following two theorems relate the strongly stationary point to B-stationary point and local minimizers of MPECs.

Theorem 2.6 ([58]). *If a feasible point $\bar{x}$ is a strong stationary point for the MPEC (2.1), then it is a B-stationary point. Conversely, if $\bar{x}$ is a B-stationary point for the MPEC (2.1), and if the MPEC-LICQ holds at $\bar{x}$, then it is a strongly stationary point.*

Theorem 2.7 ([46, 58]). *If the MPEC-LICQ holds at a local minimizer $\bar{x}$ of the MPEC (2.1), then $\bar{x}$ is a strongly stationary point with a unique vector of MPEC multipliers $\bar{\lambda} = (\bar{\lambda}^g, \bar{\lambda}^h, \bar{\lambda}^G, \bar{\lambda}^H)$.*

Recently, researchers have shown that MPECs can be solved reliably and efficiently [2, 9, 13, 14, 32, 53] using standard nonlinear optimization solvers by reformulating the MPEC (2.1) as the following equivalent nonlinear program:

$$\begin{aligned} \text{minimize} \quad & f(x) \\ \text{subject to} \quad & g(x) \le 0, \quad h(x) = 0, \\ & G(x) \ge 0, \quad H(x) \ge 0, \\ & G(x) \circ H(x) \le 0. \end{aligned} \tag{2.8}$$

The key observation in proving convergence of such an approach is that strong stationarity is equivalent to the KKT conditions of the equivalent NLP (2.8). This is stated in the following theorem.

Theorem 2.8 ([2, 14, 30, 62])). *A vector $\bar{x}$ is a* strongly stationary point *of the MPEC (2.1) if and only if it is a KKT point of nonlinear program (2.8), i.e., there exists a vector of Lagrangian multipliers $\hat{\lambda} = (\hat{\lambda}^g, \hat{\lambda}^h, \hat{\lambda}^G, \hat{\lambda}^H, \hat{\lambda}^{GH})$, such that $(\bar{x}, \hat{\lambda})$ satisfies the following conditions:*

$$\begin{aligned} \nabla f(\bar{x}) + \nabla g(\bar{x})^{\mathrm{T}} \hat{\lambda}^g + \nabla h(\bar{x})^{\mathrm{T}} \hat{\lambda}^h & \\ -\nabla G(\bar{x})^{\mathrm{T}}[\hat{\lambda}^G - H(\bar{x}) \circ \hat{\lambda}^{GH}] - \nabla H(\bar{x})^{\mathrm{T}}[\hat{\lambda}^H - G(\bar{x}) \circ \hat{\lambda}^{GH}] & = 0, \\ h(\bar{x}) = 0; \quad g(\bar{x}) \le 0, \quad \hat{\lambda}^g \ge 0, \quad (\hat{\lambda}^g)^{\mathrm{T}} g(\bar{x}) & = 0, \\ G(\bar{x}) \ge 0, \quad \hat{\lambda}^G \ge 0, \quad (\hat{\lambda}^G)^{\mathrm{T}} G(\bar{x}) & = 0, \\ H(\bar{x}) \ge 0, \quad \hat{\lambda}^H \ge 0, \quad (\hat{\lambda}^H)^{\mathrm{T}} G(\bar{x}) & = 0, \\ G(\bar{x}) \circ H(\bar{x}) \le 0, \quad \hat{\lambda}^{GH} \ge 0, \quad (\hat{\lambda}^{GH})^{\mathrm{T}} [G(\bar{x}) \circ H(\bar{x})] & = 0. \end{aligned} \tag{2.9}$$

2.2 A Generalization of Scholtes's Regularization

In this section, we present a generalization of Scholtes's regularization scheme [62]. Our approach suggests relaxing the complementarity constraints and perturbing the coefficients in the objective function and constraints simultaneously. Hence, Scholtes's scheme is a special case of our approach if the objective function and constraints are not perturbed. We show that the convergence analysis studied in [62] can be extended to our method without any difficulty. The convergence results of our method will be applied to establish the convergence of the sequential nonlinear complementarity algorithm in the next section.

For any mapping $F : R^n \times \mathcal{A}^F \to R^m$, where R^n is the space of variables and $\mathcal{A}^F$ is the space of (fixed) parameters, we denote the mapping as $F(x; \bar{a}^F)$ with $x \in R^n$ and $\bar{a}^F \in \mathcal{A}^F$. The order of elements in $\bar{a}^F$ is mapping specific. For any positive sequence $\{t\}$ tending to 0, we perturb the parameters in F and denote the new parameter vector as a_t^F with $a_t^F \to \bar{a}^F$ as $t \to 0$, and $a_t^F = \bar{a}^F$ when $t = 0$. Note that the perturbation on $\bar{a}^F$ does not require the perturbed vector a_t^F to be parameterized by t.

To facilitate the presentation, we let $\Omega := \{f, g, h, G, H\}$ denote the collection of all the functions in the MPEC (2.1). With the notation defined above, the MPEC (2.1) is presented as

$$\begin{array}{ll} \text{minimize} & f(x; \bar{a}^f) \\ \text{subject to} & g(x; \bar{a}^g) \le 0, \quad h(x; \bar{a}^h) = 0, \\ & 0 \le G(x; \bar{a}^G) \perp H(x; \bar{a}^H) \ge 0, \end{array} \tag{2.10}$$

where $\bar{a}^\omega \in \mathcal{A}^\omega$, for all $\omega \in \Omega$.

For any positive sequence $\{t\}$ tending to 0, we perturb every parameter vector $\bar{a}^\omega$ and denote the perturbed parameter vector as a_t^ω for all $\omega \in \Omega$. The perturbed vector a_t^ω should satisfy the following two conditions for all $\omega \in \Omega$:

$$a_t^\omega \to \bar{a}^\omega, \quad \text{as } t \to 0^+. \tag{2.11}$$

$$a_t^\omega = \bar{a}^\omega, \quad \text{when } t = 0. \tag{2.12}$$

As $\{t\} \to 0^+$, we are solving a sequence of perturbed NLPs, denoted by Reg(t):

$$\begin{array}{ll} \text{minimize} & f(x; a_t^f) \\ \text{subject to} & g(x; a_t^g) \leq 0, \quad h(x; a_t^h) = 0, \\ & G(x; a_t^G) \geq 0, \quad H(x; a_t^H) \geq 0, \\ & G(x; a_t^G) \circ H(x; a_t^H) \leq t\, e. \end{array} \tag{2.13}$$

In what follows, we extend Theorem 3.1, Theorem 3.3, and Corollary 3.4 in [62] to our relaxation method. The proof closely follows the one given by Scholtes in [62] and is included here for the completeness. We first state two technical lemmas. Lemma 2.9 states that the NLP-LICQ at a feasible point carries over to all feasible points in a sufficiently small neighborhood. Lemma 2.10 extends similar results to MPECs.

Lemma 2.9. *Consider the nonlinear program*

$$\begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & h(x) = 0, \\ & g(x) \leq 0, \end{array} \tag{2.14}$$

where $f : R^n \to R^1$, $h : R^n \to R^l$ and $g : R^n \to R^m$ are twice continuously differentiable functions. If the NLP-LICQ holds at a feasible point $\bar{x}$ of (2.14), then there exists a neighborhood $\mathcal{N}(\bar{x})$ such that the NLP-LICQ holds at every feasible point $x \in \mathcal{N}(\bar{x})$.

Proof. Let $A(\bar{x})$ be the Jacobian matrix of active constraints at $\bar{x}$ of (2.14). Since NLP-LICQ holds at $\bar{x}$, the rows of $A(\bar{x})$ are linearly independent. Then, there exists a basis $A_{\bullet\beta}(\bar{x})$ in $A(\bar{x})$ and the determinant of $A_{\bullet\beta}(\bar{x})$, denoted as $\det(A_{\bullet\beta}(\bar{x}))$, is nonzero. Since $\det(A_{\bullet\beta}(x))$ is a continuous function of x, it follows that there exists a neighborhood $\mathcal{N}(\bar{x})$ such that for all feasible points x of (2.14) in $\mathcal{N}(\bar{x})$:

$$\begin{aligned} \det(A_{\bullet\beta}(x)) &\neq 0, \\ \mathcal{I}_g(x) &\subseteq \mathcal{I}_g(\bar{x}), \\ \mathcal{I}_h(x) &\subseteq \mathcal{I}_h(\bar{x}). \end{aligned} \tag{2.15}$$

This further implies that for every such x the gradient vectors of the active constraints in (2.14) are linearly independent. Hence, NLP-LICQ holds at every feasible point $x \in \mathcal{N}(\bar{x})$. ■

Given a feasible point $\bar{x}$ of Reg(t) (2.13), we define the following index sets of active constraints:

$$
\begin{aligned}
\mathcal{I}_g(\bar{x}, t) &:= \{i \mid g_i(\bar{x}; a_t^g) = 0\}, \\
\mathcal{I}_h(\bar{x}, t) &:= \{i \mid h_i(\bar{x}; a_t^h) = 0\}, \\
\mathcal{I}_G(\bar{x}, t) &:= \{i \mid G_i(\bar{x}; a_t^G) = 0\}, \\
\mathcal{I}_H(\bar{x}, t) &:= \{i \mid H_i(\bar{x}; a_t^H) = 0\}, \\
\mathcal{I}_{GH}(\bar{x}, t) &:= \{i \mid G_i(\bar{x}; a_t^H) H_i(\bar{x}; a_t^H) = t\}.
\end{aligned}
\tag{2.16}
$$

Lemma 2.10. *If the MPEC-LICQ holds at the feasible point $\bar{x}$ of the MPEC (2.10), then there exists a neighborhood $\mathcal{N}(\bar{x})$ and a scalar $\bar{t} > 0$ such that for every $t \in (0, \bar{t}\,)$, the NLP-LICQ holds at every feasible point $x \in \mathcal{N}(\bar{x})$ of Reg(t) (2.13).*

Proof. This follows from Lemma 2.9 and the following relations on the index sets of active constraints:

$$
\begin{aligned}
\mathcal{I}_g(x, t) &\subseteq \mathcal{I}_g(\bar{x}), \\
\mathcal{I}_h(x, t) &\subseteq \mathcal{I}_h(\bar{x}), \\
\mathcal{I}_G(x, t) \cup \mathcal{I}_H(x, t) \cup \mathcal{I}_{GH}(x, t) &\subseteq \mathcal{I}_G(\bar{x}) \cup \mathcal{I}_H(\bar{x}), \\
\mathcal{I}_G(x, t) \cap \mathcal{I}_{GH}(x, t) &= \emptyset, \\
\mathcal{I}_H(x, t) \cap \mathcal{I}_{GH}(x, t) &= \emptyset,
\end{aligned}
\tag{2.17}
$$

which hold for all x in a sufficiently small neighborhood $\mathcal{N}(\bar{x})$ and all $t \in (0, \bar{t})$ for sufficiently small $\bar{t} > 0$.

For every feasible x of (2.13) in $\mathcal{N}(\bar{x})$, by Lemma 2.9 and (2.17), the system

$$\begin{aligned} & \sum_{i \in \mathcal{I}_g(x,t)} \lambda_i^g \nabla g_i(x) + \sum_{i \in \mathcal{I}_h(x,t)} \lambda_i^h \nabla g_i(x) \\ + & \sum_{i \in \mathcal{I}_G(x,t)} \lambda_i^G \nabla G_i(x) + \sum_{i \in \mathcal{I}_H(x,t)} \lambda_i^H \nabla H_i(x) \\ + & \sum_{i \in \mathcal{I}_{GH}(x,t)} [(\lambda_i^{GH} H_i(x; a_t^H)) \nabla G_i(x) + (\lambda_i^{GH} G_i(x; a_t^G)) \nabla H_i(x)] = 0 \end{aligned} \tag{2.18}$$

implies that $\lambda_i^g = \lambda_i^h = \lambda_i^G = \lambda_i^H = \lambda_i^{GH} G_i(x; a_t^G) = \lambda_i^{GH} H_i(x; a_t^H) = 0$. Since

$$G_i(x; a_t^G) H_i(x; a_t^H) = t,$$

we have

$$G_i(x; a_t^G) > 0, \quad H_i(x; a_t^H) > 0,$$

and thus, $\lambda_i^{GH} = 0$. This proves that the NLP-LICQ holds at x. ■

Theorem 2.11. *Let $\{t_\nu\}$ be a sequence of positive scalars tending to zero as $\nu \to \infty$. Denote a stationary point of Reg(t_ν) by x_ν for each ν and let the sequence $\{x_\nu\}$ converge to $\bar{x}$. Suppose the MPEC-LICQ holds at $\bar{x}$. Then*

(i) The point $\bar{x}$ is a C-stationary point of the MPEC (2.1).

(ii) If, for each ν, the point x_ν also satisfies second-order necessary optimality conditions for Reg(t_ν), then $\bar{x}$ is an M-stationary point of the MPEC (2.1).

(iii) Moreover, if the ULSC assumption holds at $\bar{x}$, then $\bar{x}$ is a B-stationary point of the MPEC (2.1).

Proof. First, by (2.11) and (2.12), it is easy to see that $\bar{x}$ is a feasible point of the MPEC (2.1). Let $\lambda_\nu = (\lambda_\nu^g, \lambda_\nu^h, \lambda_\nu^G, \lambda_\nu^H, \lambda_\nu^{GH})$ be the Lagrangian multipliers of Reg(t_ν) (2.13) at the stationary point x_ν, and let

$$\mathcal{I}_0 = \{i \,|\, i \in \mathcal{I}_{GH}(x_\nu, t_\nu) \text{ for infinitely many } \nu\}.$$

Since MPEC-LICQ holds at $\bar{x}$, the multiplier vector λ_ν is unique in the following

KKT system of Reg(t_ν) (2.13) for sufficiently small t_ν:

$$
\begin{array}{rcl}
\nabla f(x_\nu; a^f_{t_\nu}) + \nabla g(x_\nu; a^g_{t_\nu})^{\mathrm{T}}\lambda^g_\nu + \nabla h(x_\nu; a^h_{t_\nu})^{\mathrm{T}}\lambda^h_\nu & & \\
- \nabla G(x_\nu; a^G_{t_\nu})^{\mathrm{T}}\lambda^G_\nu - \nabla H(x_\nu; a^H_{t_\nu})^{\mathrm{T}}\lambda^H_\nu & & \\
+ \nabla G(x_\nu; a^G_{t_\nu})^{\mathrm{T}}[H(x_\nu; a^H_{t_\nu}) \circ \lambda^{GH}_\nu] + \nabla H(x_\nu; a^H_{t_\nu})^{\mathrm{T}}[G(x_\nu; a^G_{t_\nu}) \circ \lambda^{GH}_\nu] & = & 0, \\
h(x_\nu; a^h_{t_\nu}) & = & 0, \\
g(x_\nu; a^g_{t_\nu}) \leq 0,\ \lambda^g_\nu \geq 0,\ g(x_\nu; a^g_{t_\nu})^{\mathrm{T}}\lambda^g_\nu & = & 0, \\
G(x_\nu; a^G_{t_\nu}) \leq 0,\ \lambda^G_\nu \geq 0,\ G(x_\nu; a^G_{t_\nu})^{\mathrm{T}}\lambda^G_\nu & = & 0, \\
H(x_\nu; a^H_{t_\nu}) \leq 0,\ \lambda^H_\nu \geq 0,\ H(x_\nu; a^H_{t_\nu})^{\mathrm{T}}\lambda^H_\nu & = & 0, \\
G(x_\nu; a^G_{t_\nu}) \circ H(x_\nu; a^H_{t_\nu}) - t_\nu\, e \leq 0,\ \lambda^{GH}_\nu & \geq & 0, \\
[G(x_\nu; a^G_{t_\nu}) \circ H(x_\nu; a^H_{t_\nu}) - t_\nu\, e]^{\mathrm{T}}\lambda^{GH}_\nu & = & 0.
\end{array}
\tag{2.19}
$$

(i) Define $\tilde{\lambda}^G_{i,\nu} = -\lambda^{GH}_{i,\nu} H_i(x_\nu; a^H_{t_\nu})$ and $\tilde{\lambda}^H_{i,\nu} = -\lambda^{GH}_{i,\nu} G_i(x_\nu; a^G_{t_\nu})$ for $i \in \mathcal{I}_{GH}(x_\nu, t_\nu)$ and rewrite the first equation in (2.19) as

$$
\begin{aligned}
& -\nabla f(x_\nu; a^f_{t_\nu}) \\
= \quad & \sum_{i\in\mathcal{I}_g(x_\nu)} \lambda^g_{i,\nu}\nabla g_i(x_\nu; a^g_{t_\nu}) + \sum_{i\in\mathcal{I}_h(x_\nu)} \lambda^h_{i,\nu}\nabla h_i(x_\nu; a^h_{t_\nu}) \\
& - \sum_{i\in\mathcal{I}_G(x_\nu,t_\nu)} \lambda^G_{i,\nu}\nabla G_i(x_\nu; a^G_{t_\nu}) - \sum_{i\in\mathcal{I}_H(x_\nu,t_\nu)} \lambda^H_{i,\nu}\nabla H_i(x_\nu; a^H_{t_\nu}) \\
& - \sum_{i\in\mathcal{I}_{GH}(x_\nu,t_\nu)\cap\mathcal{I}^c_G(\bar{x})} \tilde{\lambda}^H_{i,\nu}\left[\nabla H_i(x_\nu; a^H_{t_\nu}) + \frac{G_i(x_\nu; a^G_{t_\nu})}{H_i(x_\nu; a^H_{t_\nu})}\nabla G_i(x_\nu; a^G_{t_\nu})\right] \\
& - \sum_{i\in\mathcal{I}_{GH}(x_\nu,t_\nu)\cap\mathcal{I}^c_H(\bar{x})} \tilde{\lambda}^G_{i,\nu}\left[\nabla G_i(x_\nu; a^G_{t_\nu}) + \frac{H_i(x_\nu; a^H_{t_\nu})}{G_i(x_\nu; a^G_{t_\nu})}\nabla H_i(x_\nu; a^H_{t_\nu})\right] \\
& - \sum_{i\in\mathcal{I}_{GH}(x_\nu,t_\nu)\cap\mathcal{I}_G(\bar{x})\cap\mathcal{I}_H(\bar{x})} \left[\tilde{\lambda}^G_{i,\nu}\nabla G_i(x_\nu; a^G_{t_\nu}) + \tilde{\lambda}^H_{i,\nu}\nabla H_i(x_\nu; a^H_{t_\nu})\right].
\end{aligned}
\tag{2.20}
$$

For every sufficient large ν, we construct a matrix $A(x_\nu)$ with rows being the transpose of the following vectors:

$$
\begin{aligned}
\nabla g_i(x_\nu; a^g_{t_\nu}), &\quad i \in \mathcal{I}_g(\bar{x}),\\
\nabla h_i(x_\nu; a^h_{t_\nu}), &\quad i \in \mathcal{I}_h(\bar{x}),\\
-\nabla G_i(x_\nu; a^G_{t_\nu}), &\quad i \in \mathcal{I}_G(\bar{x}) \setminus (\mathcal{I}_{GH}(x_\nu, t_\nu) \cap \mathcal{I}^c_H(\bar{x})),\\
-\nabla G_i(x_\nu; a^G_{t_\nu}) - \frac{H_i(x_\nu; a^H_{t_\nu})}{G_i(x_\nu; a^G_{t_\nu})}\nabla H_i(x_\nu; a^H_{t_\nu}), &\quad i \in \mathcal{I}_{GH}(x_\nu, t_\nu) \cap \mathcal{I}^c_H(\bar{x}),\\
-\nabla H_i(x_\nu; a^H_{t_\nu}), &\quad i \in \mathcal{I}_H(\bar{x}) \setminus (\mathcal{I}_{GH}(x_\nu, t_\nu) \cap \mathcal{I}^c_G(\bar{x})),\\
-\nabla H_i(x_\nu; a^H_{t_\nu}) - \frac{G_i(x_\nu; a^G_{t_\nu})}{H_i(x_\nu; a^H_{t_\nu})}\nabla G_i(x_\nu; a^G_{t_\nu}), &\quad i \in \mathcal{I}_H(\bar{x}) \setminus (\mathcal{I}_{GH}(x_\nu, t_\nu) \cap \mathcal{I}^c_G(\bar{x})).
\end{aligned}
$$

Then (2.20) can be represented as an enlarged system of equations $A(x_\nu)^{\mathrm{T}} y_\nu = -\nabla f(x_\nu)$ with some components in y_ν set to 0. The sequence of matrices $\{A(x_\nu)\}$ converges to the matrix $A(\bar{x})$ with linearly independent rows

$$
\begin{aligned}
\nabla g_i(\bar{x}; \bar{a}^g), &\quad i \in \mathcal{I}_g(\bar{x}),\\
\nabla h_i(\bar{x}; \bar{a}^h), &\quad i \in \mathcal{I}_h(\bar{x}),\\
\nabla G_i(\bar{x}; \bar{a}^G), &\quad i \in \mathcal{I}_G(\bar{x}),\\
\nabla H_i(\bar{x}; \bar{a}^H), &\quad i \in \mathcal{I}_H(\bar{x}).
\end{aligned}
$$

It follows that y_ν converges to a unique vector, $\bar{\lambda} = (\bar{\lambda}^g, \bar{\lambda}^h, \bar{\lambda}^G, \bar{\lambda}^H)$, with

$$
\begin{aligned}
&\lim_{\nu\to\infty} \lambda^g_{i,\nu} = \bar{\lambda}^g_i \ge 0, \quad \lim_{\nu\to\infty} \lambda^h_{i,\nu} = \bar{\lambda}^h_i,\\
i \notin \mathcal{I}_0 :\ &\lim_{\nu\to\infty} \lambda^G_{i,\nu} = \bar{\lambda}^G_i \ge 0, \quad \lim_{\nu\to\infty} \lambda^H_{i,\nu} = \bar{\lambda}^H_i \ge 0,\\
i \in \mathcal{I}_0 :\ &\begin{cases} -\lim\limits_{\nu\to\infty} \lambda^{GH}_{i,\nu} H_i(x_\nu; a^H_{t_\nu}) = \bar{\lambda}^G_i \le 0,\\ -\lim\limits_{\nu\to\infty} \lambda^{GH}_{i,\nu} G_i(x_\nu; a^G_{t_\nu}) = \bar{\lambda}^H_i \le 0. \end{cases}
\end{aligned}
\tag{2.21}
$$

This completes the proof of (i).

(ii) Suppose $\bar{x}$ is not an M-stationary point of the MPEC (2.1). Then there exists an index $j \in \mathcal{I}_{GH}(\bar{x})$ such that

$$-\lim_{\nu\to\infty} \lambda_{j,\nu}^{GH} H_j(x_\nu; a_{t_\nu}^H) = \bar{\lambda}_j^G < 0, \quad -\lim_{\nu\to\infty} \lambda_{j,\nu}^{GH} G_j(x_\nu; a_{t_\nu}^G) = \bar{\lambda}_j^H \le 0.$$

From (i), this further implies that $j \in \mathcal{I}_0$ and $G_j(x_\nu; a_{t_\nu}^G)H_j(x_\nu; a_{t_\nu}^H) = t_\nu$ for every sufficiently large ν.

For every ν, we construct a matrix $B(x_\nu)$ with rows being the transpose of the vectors

$$\begin{aligned} &\nabla g_i(x_\nu; a_{t_\nu}^g), \quad i \in \mathcal{I}_g(\bar{x}), \\ &\nabla h_i(x_\nu; a_{t_\nu}^h), \quad i \in \mathcal{I}_h(\bar{x}), \\ &\nabla G_i(x_\nu; a_{t_\nu}^G), \quad i \in \mathcal{I}_G(\bar{x}), \\ &\nabla H_i(x_\nu; a_{t_\nu}^H), \quad i \in \mathcal{I}_H(\bar{x}). \end{aligned}$$

The sequence of matrices $\{B(x_\nu)\}$ converges to the matrix $B(\bar{x})$ with linearly independent rows

$$\begin{aligned} &\nabla g_i(\bar{x}; \bar{a}^g), \quad i \in \mathcal{I}_g(\bar{x}), \\ &\nabla h_i(\bar{x}; \bar{a}^h), \quad i \in \mathcal{I}_h(\bar{x}), \\ &\nabla G_i(\bar{x}; \bar{a}^G), \quad i \in \mathcal{I}_G(\bar{x}), \\ &\nabla H_i(\bar{x}; \bar{a}^H), \quad i \in \mathcal{I}_H(\bar{x}). \end{aligned}$$

Since the MPEC-LICQ holds at $\bar{x}$, it follows that the rows in the matrix $B(x_\nu)$ are linearly independent for every sufficiently large ν. Consequently, the following system has no solutions for ν large enough:

$$B(x_\nu)^{\mathrm{T}} z_\nu = 0, \quad z_\nu \neq 0. \tag{2.22}$$

By Gale's theorem of alternatives [34, p. 34], the following system has a solution

d_ν for every sufficiently large ν:

$$\begin{array}{rcll}
\nabla g_i(x_\nu; a^g_{t_\nu})^{\mathrm{T}} d_\nu &=& 0, & i \in \mathcal{I}_g(\bar{x}), \\
\nabla h_i(x_\nu; a^h_{t_\nu})^{\mathrm{T}} d_\nu &=& 0, & i \in \mathcal{I}_h(\bar{x}), \\
\nabla G_i(x_\nu; a^G_{t_\nu})^{\mathrm{T}} d_\nu &=& 0, & i \in \mathcal{I}_G(\bar{x}),\ i \neq j, \\
\nabla H_i(x_\nu; a^H_{t_\nu})^{\mathrm{T}} d_\nu &=& 0, & i \in \mathcal{I}_H(\bar{x}),\ i \neq j, \\
\nabla G_j(x_\nu; a^G_{t_\nu})^{\mathrm{T}} d_\nu &=& 1, & \\
\nabla H_j(x_\nu; a^H_{t_\nu})^{\mathrm{T}} d_\nu &=& -\dfrac{H_j(x_\nu; a^H_{t_\nu})}{G_j(x_\nu; a^G_{t_\nu})}, &
\end{array}$$

and we represent the system as

$$B(x_\nu)\, d_\nu = b_\nu. \tag{2.23}$$

Similarly, the following system has a solution $\bar{d}$:

$$\begin{array}{rcll}
\nabla g_i(\bar{x}; \bar{a}^g)^{\mathrm{T}} \bar{d} &=& 0, & i \in \mathcal{I}_g(\bar{x}), \\
\nabla h_i(\bar{x}; \bar{a}^h)^{\mathrm{T}} \bar{d} &=& 0, & i \in \mathcal{I}_h(\bar{x}), \\
\nabla G_i(\bar{x}; \bar{a}^G)^{\mathrm{T}} \bar{d} &=& 0, & i \in \mathcal{I}_G(\bar{x}),\ i \neq j, \\
\nabla H_i(\bar{x}; \bar{a}^H)^{\mathrm{T}} \bar{d} &=& 0, & i \in \mathcal{I}_H(\bar{x}),\ i \neq j, \\
\nabla G_j(\bar{x}; \bar{a}^G)^{\mathrm{T}} \bar{d} &=& 1, & \\
\nabla H_j(\bar{x}; \bar{a}^G)^{\mathrm{T}} \bar{d} &=& -\bar{\lambda}^G / \bar{\lambda}^H, &
\end{array}$$

and we represent the system as

$$B(\bar{x})\, \bar{d} = \bar{b}. \tag{2.24}$$

Below, we construct a bounded sequence $\{d_\nu\}$ converging to $\bar{d}$. Without loss of generality, we can assume that there exists an index set β such that $B_{\bullet\beta}(\bar{x})$ is a basis in $B(\bar{x})$ and $B_{\bullet\beta}(x_\nu)$ is a basis in $B(x_\nu)$ for every sufficient large ν. Furthermore, the vector $\bar{d}$ is a basic solution of (2.24) associated with the basis

$B_{\bullet\beta}(\bar{x})$ with $\bar{d}_\beta$ satisfying

$$B_{\bullet\beta}(\bar{x})\bar{d}_\beta = \bar{b}$$

and the rest of the components in $\bar{d}$ being 0.

Similarly, for every sufficiently large ν, the vector d_ν is a basic solution of (2.23) associated with the basis $B_{\bullet\beta}(x_\nu)$ with $(d_\nu)_\beta$ satisfying

$$B_{\bullet\beta}(x_\nu)(d_\nu)_\beta = b_\nu$$

and the rest of the components in d_ν being 0.

From (2.21), it is clear that $H_j(x_\nu; a^H_{t_\nu})/G_j(x_\nu; a^G_{t_\nu}) \to \bar{\lambda}^G/\bar{\lambda}^H$, and hence, $b_\nu \to \bar{b}$ as $\nu \to \infty$. With $B(x_\nu)$ converging to $B(\bar{x})$, it follows that the sequence $\{d_\nu\}$ is bounded and $d_\nu \to \bar{d}$ as $\nu \to \infty$.

It is easy to see that d_ν is a critical direction of Reg(t_ν) (2.13) at x_ν for ν large enough. If the constraint $G_j(x; a^G_{t_\nu})H_j(x; a^H_{t_\nu}) \le t_\nu$ is active at x_ν, we examine the term associated with this constraint in the Lagrangian function of Reg(t_ν) for the second-order necessary optimality conditions. In particular,

$$\begin{aligned}
&\lambda^{GH}_{j,\nu}\, d_\nu^{\mathrm{T}} \nabla^2 \left(G_j(x_\nu; a^G_{t_\nu})H_j(x_\nu; a^H_{t_\nu}) - t_\nu\right) d_\nu \\
&= \lambda^{GH}_{j,\nu} H_j(x_\nu; a^H_{t_\nu})\, d_\nu^{\mathrm{T}} \nabla^2 G_j(x_\nu; a^G_{t_\nu})\, d_\nu + \lambda^{GH}_{j,\nu} G_j(x_\nu; a^G_{t_\nu})\, d_\nu^{\mathrm{T}} \nabla^2 H_j(x_\nu; a^H_{t_\nu})\, d_\nu \\
&\quad - \lambda^{GH}_{j,\nu} H_j(x_\nu; a^H_{t_\nu}) \frac{2}{G_j(x_\nu; a^G_{t_\nu})}.
\end{aligned}$$

While the first two terms in the above equation are bounded, the third term

$$-\lambda^{GH}_{j,\nu} H_j(x_\nu; a^H_{t_\nu}) \frac{2}{G_j(x_\nu; a^G_{t_\nu})} \to -\infty, \quad \text{as } \nu \to \infty,$$

since $\lambda^{GH}_{j,\nu} H_j(x_\nu; a^H_{t_\nu}) \to -\bar{\lambda}_j > 0$ and $G_j(x_\nu; a^G_{t_\nu}) \to 0^+$. It is easy to check that all other terms in $d_\nu^{\mathrm{T}} \nabla^2 \mathcal{L}(x_\nu, \lambda_\nu) d_\nu$ are bounded, and hence, the second-order necessary optimality condition of Reg(t_ν) (2.13) fails at x_ν for sufficiently large ν.
(iii) Since, from (ii), $\bar{x}$ is an M-stationary point and the ULSC holds at $\bar{x}$, it follows that $\bar{x}$ is a strongly stationary point, and hence, a B-stationary point. ■

Chapter 3

Equilibrium Problem with Equilibrium Constraints

In this chapter, we define EPECs and their stationarity concepts. Diagonalization methods [4, 5, 22, 25, 49] have been widely used by researchers in engineering fields to solve EPECs. The absence of convergence results for diagonalization methods is one of their main drawbacks. We briefly discuss the convergence properties of the diagonalization methods, in which every MPEC subproblem is solved as an equivalent nonlinear program [13, 14]. We also propose a sequential nonlinear complementarity (NCP) approach for solving EPECs and establish the convergence of this algorithm. Finally, we present the numerical results of the SNCP method and give a comparison with two diagonalization methods, nonlinear Jacobi and nonlinear Gauss-Seidel, on a set of randomly generated EPEC test problems.

3.1 Formulation and stationarity conditions

An EPEC is a problem of finding an equilibrium point that solves several MPECs simultaneously. Since practical applications of EPEC models often arise from multi-leader-follower game settings, we consider the EPEC consisting of MPECs with shared decision variables and shared equilibrium constraints. In particular, we assume the EPEC consists of K MPECs, and for each $k = 1, \ldots, K$, the k-th MPEC has the following form with independent decision variables $x^k \in R^{n_k}$ and shared decision variables $y \in R^{n_0}$:

$$\begin{array}{ll} \text{minimize} & f^k(x^k, y; \bar{x}^{-k}) \\ \text{subject to} & g^k(x^k, y; \bar{x}^{-k}) \leq 0, \quad h^k(x^k, y; \bar{x}^{-k}) = 0, \\ & 0 \leq G(x^k, y; \bar{x}^{-k}) \perp H(x^k, y; \bar{x}^{-k}) \geq 0, \end{array} \tag{3.1}$$

where $f^k : R^n \to R$, $g^k : R^n \to R^{p_k}$, $h^k : R^n \to R^{q_k}$, $G : R^n \to R^m$ and $H : R^n \to R^m$ are twice continuously differentiable functions in both $x = (x^k)_{k=1}^K$ and y, with $n = \sum_{k=0}^K n_k$. The notation $\bar{x}^{-k}$ means that $x^{-k} = (x^j)_{j=1}^K \setminus x^k$ $(\in R^{n-n_k-n_0})$ is not a variable but a fixed vector. This implies that we can view (3.1), denoted by MPEC($\bar{x}^{-k}$), as being parameterized by $\bar{x}^{-k}$. Given $\bar{x}^{-k}$, we assume the solution set of the k-th MPEC is nonempty and denote it by SOL(MPEC($\bar{x}^{-k}$)). Notice that in the above formulation, each MPEC shares the same equilibrium constraints, represented by the complementarity system

$$0 \leq G(x, y) \perp H(x, y) \geq 0.$$

The EPEC, associated with K MPECs defined as above, is to find a Nash equilibrium $(x^*, y^*) \in R^n$ such that

$$(x^{k*}, y^*) \in \text{SOL}(\text{MPEC}(x^{-k*})) \quad \forall\, k = 1, \ldots, K. \tag{3.2}$$

Mordukhovich [39] studies the necessary optimality conditions of EPECs in the context of multiobjective optimization with constraints governed by parametric variational systems in infinite-dimensional space. His analysis is based on advanced tools of variational analysis and generalized differential calculus [37, 38]. Since we only consider finite-dimensional optimization problems, following Hu [25], we use the KKT approach and define stationary conditions for EPECs by applying those for MPECs.

Definition 3.1. We call a vector (x^*, y^*) a B-stationary (strongly stationary, M-stationary, C-stationary, weakly stationary) point of the EPEC (3.2) if for each $k = 1, \ldots, K$, (x^{k*}, y^*) is a B-stationary (strongly stationary, M-stationary, C-stationary, weakly stationary) point for the MPEC(x^{-k*}).

Theorem 3.2. *Let (x^*, y^*) be a (possibly local) equilibrium point of EPEC (3.2). If for each $k = 1, \dots, K$, the MPEC-LICQ holds at (x^{k*}, y^*) for MPEC(x^{-k*}) (3.1), then (x^*, y^*) is an EPEC strongly stationary point. In particular, there exist vectors $\lambda^* = (\lambda^{1*}, \dots, \lambda^{K*})$ with $\lambda^{k*} = (\lambda^{g,k*}, \lambda^{h,k*}, \lambda^{G,k*}, \lambda^{H,k*}, \lambda^{GH,k*})$ such that (x^*, y^*, λ^*) solves the system*

$$\left.\begin{aligned}
\nabla_{x^k} f^k(x^k, y; x^{-k}) \\
+\nabla_{x^k} g^k(x^k, y; x^{-k})^{\mathrm{T}}\lambda^{g,k} + \nabla_{x^k} h^k(x^k, y; x^{-k})^{\mathrm{T}}\lambda^{h,k} \\
-\nabla_{x^k} G(x^k, y; x^{-k})^{\mathrm{T}}\lambda^{G,k} - \nabla_{x^k} H(x^k, y; x^{-k})^{\mathrm{T}}\lambda^{H,k} \\
+\nabla_{x^k} G(x^k, y; x^{-k})^{\mathrm{T}}[H(x^k, y; x^{-k}) \circ \lambda^{GH,k}] \\
+\nabla_{x^k} H(x^k, y; x^{-k})^{\mathrm{T}}[G(x^k, y; x^{-k}) \circ \lambda^{GH,k}] &= 0 \\
\nabla_y f^k(x^k, y; x^{-k}) \\
+\nabla_y g^k(x^k, y; x^{-k})^{\mathrm{T}}\lambda^{g,k} + \nabla_y h^k(x^k, y; x^{-k})^{\mathrm{T}}\lambda^{h,k} \\
-\nabla_y G(x^k, y; x^{-k})^{\mathrm{T}}\lambda^{G,k} - \nabla_y H(x^k, y; x^{-k})^{\mathrm{T}}\lambda^{H,k} \\
+\nabla_y G(x^k, y; x^{-k})^{\mathrm{T}}[H(x^k, y; x^{-k}) \circ \lambda^{GH,k}] \\
+\nabla_y H(x^k, y; x^{-k})^{\mathrm{T}}[G(x^k, y; x^{-k}) \circ \lambda^{GH,k}] &= 0 \\
h^k(x^k, y; x^{-k}) = 0 \\
0 \geq g^k(x^k, y; x^{-k}) \quad \perp \quad \lambda^{g,k} \geq 0, \\
0 \leq G(x^k, y; x^{-k}) \quad \perp \quad \lambda^{G,k} \geq 0, \\
0 \leq H(x^k, y; x^{-k}) \quad \perp \quad \lambda^{H,k} \geq 0, \\
0 \leq -G(x^k, y; x^{-k}) \circ H(x^k, y; x^{-k}) \quad \perp \quad \lambda^{GH,k} \geq 0,
\end{aligned}\right\} \tag{3.3}$$

$$k = 1, \dots, K.$$

Conversely, if (x^*, y^*, λ^*) is a solution of the above system (3.3), then (x^*, y^*) is a B-stationary point of the EPEC (3.2).

Proof. Since (x^*, y^*) is a (possibly local) equilibrium point of the EPEC (3.2), it follows that for each $k = 1, \dots, K$, the point (x^{k*}, y^*) is a (local) minimizer of the MPEC(x^{-k*}) (3.1). By applying Theorem 2.7 and Theorem 2.8 to the MPEC(x^{-k*}) (3.1) for $k = 1, \dots, K$, we can show that there exists a vector $\lambda^{k*} = (\lambda^{g,k*}, \lambda^{h,k*}, \lambda^{G,k*}, \lambda^{H,k*}, \lambda^{GH,k*})$ such that $(x^{k*}, y^*, \lambda^{k*})$ satisfies the conditions in the system (3.3) for each $k = 1, \dots, K$. Let $\lambda^* = (\lambda^{1*}, \dots, \lambda^{K*})$. Then, the vector (x^*, y^*, λ^*) is a solution of the system (3.3). Conversely, by Theorem 2.8, it is easy to check that for each $k = 1, \dots, K$, the vector (x^{k*}, y^*) is a strongly stationary point, and hence, B-stationary point (by Theorem 2.6) for the MPEC(x^{-k*}) (3.1). As a result, the vector (x^*, y^*) is a B-stationary point of the EPEC (3.2). ∎

3.2 Algorithms for Solving EPECs

To date, algorithms specifically designed for solving EPECs have not been developed in the literature. The approaches used by researchers in engineering fields to solve EPECs fall into the category of *Diagonalization methods* [4, 5, 22, 25, 49], which mainly rely on NLP solvers, or more appropriately, MPEC algorithms to solve one MPEC at a time and cyclicly repeat the same procedure for every MPEC until an equilibrium point is found. In the remainder of this section, we first describe two types of diagonalization method: nonlinear Jacobi and nonlinear Gauss-Seidel, and briefly discuss their convergence. We then present a new method called the *sequential nonlinear complementarity (SNCP) algorithm* for solving EPECs. This new method is based on simultaneously relaxing the complementarity constraints in each MPEC, and solves EPECs by solving a sequence of nonlinear complementarity problems. We also establish the convergence of the SNCP algorithm.

3.2.1 Diagonalization methods

Diagonalization methods [8, 45] were originally proposed to solve variational inequality problems. In [20], Harker applied a diagonalization (or *nonlinear Jacobi*) algorithm, to find a solution to a variational inequality formulation of the *Nash equilibrium* problem in an oligopolistic market.

Because of their conceptual simplicity and ease of implementation, diagonalization methods using NLP solvers have been natural choices for engineers and applied economists to solve EPEC models [4, 5] arising in deregulated electricity markets. In [22] and [49], MPEC algorithms (a penalty interior point algorithm in the former reference and a smoothing algorithm in the latter) are used in diagonalization methods to solve EPEC models. Below, we describe two diagonalization methods: nonlinear Jacobi and nonlinear Gauss-Seidel. The framework of the diagonalization methods presented here follows the one given in Hu [25].

The nonlinear Jacobi method for the EPEC (3.2) is described as follows:

Step 0. *Initialization.*
Choose a starting point $(x^{(0)}, y^{(0)}) = (x^{1,(0)}, \dots, x^{K,(0)}, y^{(0)})$, the maximum number of outer iterations J, and an accuracy tolerance $\varepsilon > 0$.

Step 1. *Loop over every MPEC.*
Suppose the current iteration point of (x, y) is $(x^{(j)}, y^{(j)})$. For each $k = 1, \dots, K$, the MPEC$(\bar{x}^{-k,(j)})$ is solved (using NLP solvers or MPEC algorithms) while fixing $\bar{x}^{-k,(j)} = (x^{1,(j)}, \dots, x^{k-1,(j)}, x^{k+1,(j)}, \dots, x^{K,(j)})$. Denote the x-part of the optimal solution of MPEC$(\bar{x}^{-k})$ by $x^{k,(j+1)}$.

Step 2. *Check convergence.*
Let $(x^{(j+1)}) = (x^{1,(j+1)}, \dots, x^{K,(j+1)})$. If $j < J$, then increase j by one and repeat Step 1. Otherwise, *stop* and *check the accuracy tolerance*: if $\|x^{k,(j+1)} - x^{k,(j)}\| < \varepsilon$ for $k = 1, \dots, K$, then accept and report the solution (x^J, y^J) ; otherwise, output "*No equilibrium point found*".

Note that the nonlinear Jacobi method does not use the most recently available information when computing $x^{k,(j+1)}$. For example, $x^{1,(j)}$ is used in the calculation of $x^{2,(j+1)}$, even though the vector, $x^{1,(j+1)}$ is known. If we revise the nonlinear Jacobi method so that we always use the new information, then we have another diagonalization method, the Gauss–Seidel method. Hence, the framework of the nonlinear Gauss–Seidel method for the EPEC (3.2) is the same as nonlinear

Jacobi, except that in Step 1., we have

$$\bar{x}^{-k,(j)} = (x^{1,(j+1)}, \ldots, x^{k-1,(j+1)}, x^{k+1,(j)}, \ldots, x^{K,(j)}).$$

The multi-firm algorithm proposed in [22] belongs to this category.

To solve each MPEC in Step 1 of diagonalization methods, one can solve the equivalent NLP (2.8) suggested by Fletcher and Leyffer [13] using off-the-shelf NLP solvers. For each $k = 1, \ldots, K$, the equivalent NLP formulation of MPEC($\bar{x}^{-k,(j)}$) is

$$\begin{array}{rrcl} \text{minimize} & f^k(x^k, y; \bar{x}^{-k,(j)}) & & \\ \text{subject to} & g^k(x^k, y; \bar{x}^{-k,(j)}) & \leq & 0, \\ & h^k(x^k, y; \bar{x}^{-k,(j)}) & = & 0, \\ & G(x^k, y; \bar{x}^{-k,(j)}) & \geq & 0, \\ & H(x^k, y; \bar{x}^{-k,(j)}) & \geq & 0, \\ \multicolumn{2}{r}{G(x^k, y; \bar{x}^{-k,(j)}) \circ H(x^k, y; \bar{x}^{-k,(j)})} & \leq & 0. \end{array} \tag{3.4}$$

We denote the above equivalent NLP of the k-th MPEC by NLP$^k(\bar{x}^{-k,(j)})$.

The following theorem states the convergence of diagonalization methods based on solving equivalent NLPs.

Theorem 3.3. *Let $\{(x^{(j)}, y^{(j)})\}$ be a sequence of solutions generated by a diagonalization (nonlinear Jacobi or nonlinear Gauss-Seidel) method, in which each MPEC is reformulated and solved as an equivalent NLP (3.4). Suppose the sequence $\{(x^{(j)}, y^{(j)})\}$ converges to (x^*, y^*) as $j \to \infty$. If, for each $k = 1, \ldots, K$, the MPEC-LICQ holds at (x^{k*}, y^*) for MPEC(x^{-k*}), then (x^*, y^*) is B-stationary for the EPEC (3.2).*

Proof. From Theorem 2.6 applied to the MPEC(x^{-k*}) for each $k = 1, \ldots, K$, the point (x^{k*}, y^*) is a B-stationary point, and hence, the point (x^*, y^*) is a B-stationary point for the EPEC (3.2). ∎

3.2.2 Sequential NCP method

We propose a new method for solving EPECs. Instead of solving an EPEC by cyclicly using an MPEC-based approach, our approach simultaneously relaxes the complementarity system in each MPEC($\bar{x}^{-k}$) to

$$\begin{aligned} &G(x^k, y; \bar{x}^{-k}) \geq 0, \quad H(x^k, y; \bar{x}^{-k}) \geq 0, \\ &G(x^k, y; \bar{x}^{-k}) \circ H(x^k, y; \bar{x}^{-k}) \leq te, \end{aligned} \tag{3.5}$$

and finds an equilibrium solution $(x^*(t), y^*(t)) = (x^{1*}(t), \dots, x^{K*}(t), y^*(t))$ of the following regularized NLPs, denoted as $\mathrm{Reg}^k(\bar{x}^{-k}; t)$ for $t > 0$:

$$\left.\begin{array}{rrcll} \text{minimize} & f^k(x^k, y; \bar{x}^{-k}) & & & \\ \text{subject to} & g^k(x^k, y; \bar{x}^{-k}) & \leq & 0, & (\lambda^{g,k}) \\ & h^k(x^k, y; \bar{x}^{-k}) & = & 0, & (\lambda^{h,k}) \\ & G(x^k, y; \bar{x}^{-k}) & \geq & 0, & (\lambda^{G,k}) \\ & H(x^k, y; \bar{x}^{-k}) & \geq & 0, & (\lambda^{H,k}) \\ \multicolumn{2}{r}{G(x^k, y; \bar{x}^{-k}) \circ H(x^k, y; \bar{x}^{-k})} & \leq & te, & (\lambda^{GH,k}) \end{array}\right\} \quad k = 1, \dots, K, \tag{3.6}$$

where, given $x^{-k*}(t)$ as the input parameter, $(x^{k*}(t), y^*(t))$ is a stationary point of the k-th regularized NLP, $\mathrm{Reg}^k(x^{-k*}(t); t)$.

Let $\mathcal{L}^k(x^k, y, \lambda^k; t)$ denote the Lagrangian function for the $\mathrm{Reg}^k(\bar{x}^{-k}; t)$. If $(x^*(t), y^*(t))$ is an equilibrium solution of (3.6) and LICQ holds at $(x^*(t), y^*(t))$ for each $\mathrm{Reg}^k(x^{-k*}(t); t)$, $k = 1, \dots, K$, then $(x^*(t), y^*(t))$ is a solution of the following mixed nonlinear complementarity problem, obtained by combining the first-order KKT system of each $\mathrm{Reg}^k(x^{-k*}(t); t)$ in (3.6):

$$\left.\begin{array}{r}
\nabla_{x^k}\mathcal{L}^k(x^k, y, \lambda^k; t) = 0 \\
\nabla_{y}\mathcal{L}^k(x^k, y, \lambda^k; t) = 0 \\
h^k(x^k, y; x^{-k}) = 0, \\
0 \geq g^k(x^k, y; x^{-k}) \quad \perp \quad \lambda^{g,k} \geq 0, \\
0 \leq G(x^k, y; x^{-k}) \quad \perp \quad \lambda^{G,k} \geq 0, \\
0 \leq H(x^k, y; x^{-k}) \quad \perp \quad \lambda^{H,k} \geq 0, \\
0 \leq te - G(x^k, y; x^{-k}) \circ H(x^k, y; x^{-k}) \quad \perp \quad \lambda^{GH,k} \geq 0,
\end{array}\right\} \tag{3.7}$$

$$k = 1, \ldots, K.$$

For convenience, we denote the above system by $\mathrm{NCP}(t)$.

While Scholtes's regularized scheme for MPECs can be described as solving an MPEC by solving a sequence of nonlinear programs (NLPs), our method is to solve an EPEC by solving a sequence of nonlinear complementarity problems (NCPs).

The following theorem states the convergence of the sequential NCP algorithm.

Theorem 3.4. *Let $\{t_\nu\}$ be a sequence of positive scalars tending to zero. Denote (x_ν, y_ν) as a solution to NCP(t_ν) (3.7) and let $\{(x_\nu, y_\nu)\}$ converge to (x^*, y^*) as $t_\nu \to 0$. Furthermore, assume, for each $k = 1, \ldots, K$ and for every ν, the point (x^k_ν, y_ν) satisfies the second-order optimality conditions of $\mathrm{Reg}^k(x^{-k}_\nu; t_\nu)$. If, for each $k = 1, \ldots, K$, the MPEC-LICQ and the ULSC hold at (x^{k*}, y^*) for MPEC(x^{-k*}), then (x^*, y^*) is B-stationary for the EPEC (3.2).*

Proof. For each $k = 1, \ldots, K$, the point (x^k_ν, y_ν) satisfies the second-order optimality condition of $\mathrm{Reg}^k(x_\nu^{-k}; t_\nu)$:

$$\begin{array}{ll} \text{minimize} & f^k(x^k, y; x_\nu^{-k}) \\ \text{subject to} & g^k(x^k, y; x_\nu^{-k}) \le 0,\ h^k(x^k, y; x_\nu^{-k}) = 0, \\ & G(x^k, y; x_\nu^{-k}) \ge 0,\ H(x^k, y; x_\nu^{-k}) \ge 0, \\ & G(x^k, y; x_\nu^{-k}) \circ H(x^k, y; x_\nu^{-k}) \le t_\nu\, e, \end{array}$$

and as $t_\nu \to 0^+$, we have $x_\nu^{-k} \to x^{-k*}$, and $(x_\nu^k, y_\nu) \to (x^{k*}, y^*)$. Since the MPEC-LICQ and the ULSC hold at (x^{k*}, y^*) for $\mathrm{MPEC}(\bar{x}^{-k})$, by Theorem 2.11, the point (x^{k*}, y^*) is B-stationary for each $k = 1, \ldots, K$. It follows that (x^*, y^*) is a B-stationary point for the EPEC (3.2). ■

3.3 Implementation and Numerical Comparison

To compare the numerical performance of EPEC algorithms, we develop a MATLAB program, *EPECgen*, to randomly generate EPEC test problems with known solutions. The number of MPECs within a generated EPECs can be arbitrary and is determined by the user. Furthermore, each MPEC is generated by the *Qpecgen* [27] program and has the following form:

$$\left.\begin{array}{ll} \underset{(x^k, y)}{\text{maximize}} & \frac{1}{2}(x^k, y)^{\mathrm{T}} P^k\,(x^k, y) + (c^k)^{\mathrm{T}} x^k + (d^k)^{\mathrm{T}} y \\ \text{subject to} & G^k x^k + H^k y + a^k \le 0, \\ & x^k \ge 0, \\ & 0 \le y \;\perp\; N^k x^k + \displaystyle\sum_{i=1, i\ne k}^{K} N^i \bar{x}^i + My + q \ge 0, \end{array}\right\} \tag{3.8}$$

$$k = 1, \ldots, K,$$

where the data for the i-th MPEC is given by the following generated vectors and matrices: $P^k \in R^{(n_k+m)\times(n_k+m)}, c^k \in R^{n_k}, d^k \in R^m, G^k \in R^{l_k \times n_k}, H^k \in R^{l_k \times m}, a^k \in R^{l_k}, N^k \in R^{m \times n_k}, M \in R^{m\times m}$, and $q \in R^m$. As stated in [27], the user has the freedom to control different properties of these MPEC and their solutions; for example, dimension of the problem, condition number of the matrix P^k, convexity of the objective function (whether P^k is positive semidefinite or not), and so on.

We have implemented the diagonalization methods and the sequential NCP algorithm on randomly generated EPECs consisting of two MPECs. For diagonalization methods, each MPEC is reformulated as an equivalent nonlinear program (2.8) and solved with TOMLAB/SNOPT [23, 24]. For the SNCP method, one can solve the complementarity system NCP(t) (3.7) as a set of constraints in an optimization problem with a constant objective function such as 0. However, such a naive implementation will result in numerical instabilities when t is small, because the set of Lagrange multipliers is unbounded for each MPEC. To stabilize the SNCP method, we minimize the sum of components in λ^{GH} and use TOMLAB/SNOPT to solve the sequence of optimization problems with $t = 1, 10^{-1}, \ldots, 10^{-15}$:

$$\begin{array}{ll} \text{minimize} & e^{\mathrm{T}}\lambda^{GH} \\ \text{subject to} & \mathrm{NCP}(t)\ (3.7). \end{array} \tag{3.9}$$

Table 3.1 summarizes the parameters used to generate the EPEC test problems. For the definition of these input parameters for each MPEC, see [27].

Table 3.1
Input parameters.

(n_1, n_2)	(8, 10)	(first_deg1, first_deg2)	(1, 1)
m	15	second_deg	3
(l_1, l_2)	(8, 8)	(mix_deg1, mix_deg2)	(1, 1)

Table 3.3 gives the random seeds used to generate each test problem and the objective function values of MPECs at generated solutions for each test problem. Numerical results for the methods SNCP, nonlinear Gauss-Seidel (with $\varepsilon = 1.0\mathrm{e}{-6}$), and nonlinear Jacobi (with $\varepsilon = 1.0\mathrm{e}{-6}$) on generated test problems are shown in Tables 5.4–5.6. To investigate the impact of the accuracy tolerance ε on the performance of diagonalization methods, we ran the nonlinear Gauss-Seidel and nonlinear Jacobi methods on the same test problems with

lower tolerance $\varepsilon = 1.0\text{e}{-4}$, and give the numerical results in Table 3.7–3.8. The notation used in these tables is explained in Table 3.2.

TABLE 3.2
Notation used for numerical results.

P	Problem number.
random_seed	The random seed used to generate each EPEC problem.
Time	Total time (in seconds) needed by the termination of algorithms.
Out Iter	# of outer iterations required by diagonalization methods.
Maj Iter	# of major iterations.
f1*	The obj val of MPEC1 at the found solution (x^{1*}, y^*).
f2*	The obj val of MPEC2 at the found solution (x^{2*}, y^*).
f1_{gen}	The obj val of MPEC1 at the generated solution (x^1_{gen}, y_{gen}).
f2_{gen}	The obj val of MPEC2 at the generated solution (x^2_{gen}, y_{gen}).
Norm	The 2-norm of the difference vector between the found solution (x^{1*}, x^{2*}, y^*) and the generated solution $(x^1_{gen}, x^2_{gen}, y_{gen})$.
flag1	= 0 if the algorithm finds an equilibrium point; = 1 if the algorithm is terminated by reaching the iteration limit; = 2 if the SNCP is terminated by the infeasibility message.
flag2	= 0 if cycling behavior is not observed for diagonalization methods; = 1 if cycling behavior is observed for diagonalization methods.

TABLE 3.3
Information on test problems.

P	random_seed	f1_{gen}	f2_{gen}
1	2.0e+5	−23.4844	22.8737
2	3.0e+5	−9.7748	−10.7219
3	4.0e+5	−16.7669	−4.6330
4	5.0e+5	−9.6054	−0.8600
5	6.0e+5	−46.9213	−11.1220
6	7.0e+5	−1.8838	−6.1389
7	8.0e+5	−14.9793	−12.1478
8	9.0e+5	−5.7299	−19.3843
9	1.0e+6	7.0672	−19.1931
10	1.1e+6	−3.2355	−17.3311

TABLE 3.4
Numerical results for SNCP method (Major Iteration limit = 120).

P	Maj Iter	T	f1*	f2*	Norm	flag1
1	120	127.9	−23.4844	22.8739	2.70e−4	1
2	75	69.0	−9.7749	−10.7219	7.06e−5	0
3	45	48.4	−16.7669	−4.6330	6.48e−5	0
4	68	73.1	−9.6052	−0.8600	1.57e−3	0
5	80	72.9	−46.9213	−11.1220	4.70e−13	0
6	120	83.9	−1.8839	−6.1393	1.15e−3	1
7	75	71.8	−14.9790	−12.1477	4.03e−4	0
8	120	126.7	−5.7300	−19.3844	1.08e−4	1
9	54	52.6	7.0672	−19.1930	1.68e−4	0
10	72	68.0	−3.2363	−17.3301	2.04e−3	0

TABLE 3.5
Numerical results for nonlinear Gauss-Seidel method ($J = 30$, $\varepsilon = 1.0e{-}6$).

P	Out Iter	Maj Iter	Time	f1*	f2*	Norm	flag1	flag2
1	30	619	96.3	−23.4844	22.8737	1.23e−4	1	0
2	26	821	114.8	−9.7805	−10.7263	6.25e−3	0	0
3	30	903	133.7	−16.7672	−4.6327	6.06e−4	1	0
4	19	1340	232.3	−9.6044	−0.8601	4.67e−3	0	0
5	7	118	17.8	−46.9213	−11.1221	1.02e−4	0	0
6	30	508	73.8	−1.7661	−6.1783	1.18e−1	1	0
7	11	1076	191.0	−14.9807	−12.1489	1.80e−3	0	0
8	30	320	61.8	−5.7228	−19.3929	1.00e−2	1	0
9	9	189	29.4	7.0672	−19.1930	7.24e−5	0	0
10	15	170	30.6	−3.2265	−17.3179	1.50e−2	0	0

TABLE 3.6

Numerical results for nonlinear Jacobi method ($J = 30$, $\varepsilon = 1.0e{-}6$).

P	Out Iter	Maj Iter	Time	f1*	f2*	Norm	flag1	flag2
1	30	695	104.8	−23.4844	22.8738	1.25e−4	1	0
2	30	1036	138.2	−9.7756	−10.7262	4.34e−3	1	1
3	30	807	104.6	−16.7668	−4.6327	4.24e−4	1	1
4	30	703	94.6	−9.6031	−0.8601	6.21e−3	1	1
5	30	375	69.2	−46.9213	−11.1221	6.06e−5	1	1
6	30	819	103.6	−1.8837	−6.1672	3.91e−2	1	1
7	30	667	94.0	−14.9790	−12.1494	2.30e−3	1	1
8	30	847	108.2	−5.7314	−19.3929	6.85e−3	1	1
9	30	624	97.6	7.0672	−19.1930	5.56e−5	1	1
10	30	766	98.7	−3.2819	−17.3179	4.76e−2	1	1

TABLE 3.7

Numerical results for nonlinear Gauss-Seidel method ($J = 30$, $\varepsilon = 1.0e{-}4$).

P	Out Iter	Maj Iter	Time	f1*	f2*	Norm	flag1	flag2
1	5	134	17.9	−23.4844	22.8738	1.25e−4	0	0
2	4	152	24.5	−9.7805	−10.7263	6.24e−3	0	0
3	5	149	19.1	−16.7672	−4.6327	6.68e−4	0	0
4	6	149	20.2	−9.6044	−0.8601	4.71e−3	0	0
5	5	100	14.4	−46.9213	−11.1220	1.02e−4	0	0
6	30	508	73.8	−1.7661	−6.1783	1.15e−1	1	0
7	6	130	18.0	−14.9807	−12.1489	1.80e−3	0	0
8	17	299	47.0	−5.7228	−19.3929	1.00e−2	0	0
9	7	187	27.3	7.0672	−19.1930	7.42e−5	0	0
10	7	149	20.8	−3.2265	−17.3179	1.49e−2	0	0

TABLE 3.8
Numerical results for nonlinear Jacobi method ($J = 30$, $\varepsilon = 1.0e{-4}$).

P	Out Iter	Maj Iter	Time	f1*	f2*	Norm	flag1	flag2
1	10	257	36.8	−23.4844	22.8738	1.26e−4	0	0
2	30	1036	138.2	−9.7756	−10.7262	4.34e−3	1	1
3	30	807	104.6	−16.7668	−4.6327	4.24e−4	1	1
4	30	703	94.6	−9.6054	−0.8600	6.21e−3	1	1
5	8	155	23.6	−46.9213	−11.1220	6.14e−5	0	0
6	30	819	103.6	−1.8837	−6.1672	3.91e−2	1	1
7	30	667	94.0	−14.9790	−12.1494	2.30e−3	1	1
8	30	847	108.2	−5.7314	−19.3929	6.85e−3	1	1
9	30	624	97.6	7.0672	−19.1930	5.56e−5	1	1
10	30	766	98.7	−3.2819	−17.3179	4.76e−2	1	1

3.3.1 Discussion of numerical results

From the numerical results, we have the following observations.

- The SNCP algorithm solves 7 test problems and is terminated for reaching the major iteration limit for 3 test problems. However, the vectors returned by SNCP method for those three test problems are close to the generated solutions with “Norm” of order 1.0e−4. It takes around one second to perform one major iteration. The number of major iterations for each of the 7 solved problem is consistently between 40 and 80.

- With $\varepsilon = 1.0e{-6}$, the nonlinear Gauss-Seidel method solves 7 test problems and is terminated for reaching the outer iteration limit for 4 test problems. However, with $\varepsilon = 1.0e{-4}$, the nonlinear Gauss-Seidel method succeeds in solving 9 test problems within at most 7 outer iterations. In fact, we observe that the nonlinear Gauss-Seidel method only needs a few (4 or 5) outer iterations to reach an accuracy of 1.0e−3 or 1.0e−4, and then makes slow progress to achieve higher accuracy of 1.0e−5 or 1.0e−6. Note that the cycling behavior is not observed for the nonlinear Gauss-Seidel method on any test problem.

- The results in Table 3.6 show that for each test problem, the nonlinear Jacobi method is terminated for reaching the outer iteration limit, although the solution vector returned is close to the generated solution. Surprisingly, the cycling behavior is observed for 9 out of the 10 problems, and the diagonalization method fails when cycling occurs. Even with lower accuracy ($\varepsilon = 1.0\text{e}{-4}$), the nonlinear Jacobi method solves only two test problems. This observation suggests that the nonlinear Jacobi method has difficulty achieving high accuracy and is less reliable.

- The comparison of "Norm" for the SNCP algorithm and the nonlinear Gauss-Seidel method seems to suggest that the SNCP algorithm is able to reach the generated solution $(x^1_{gen}, x^2_{gen}, y_{gen})$ more accurately than the nonlinear Gauss-Seidel method. The fact that all these methods return a solution close to the generated solution $(x^1_{gen}, x^2_{gen}, y_{gen})$ seems to indicate that the generated solution is isolated or locally unique. Further investigation of the properties of the generated solutions is needed.

- With the accuracy tolerance $\varepsilon = 1.0\text{e}{-6}$, it is difficult to say which method, SNCP or nonlinear Gauss-Seidel, is more efficient. However, it is clear that both methods outperform the nonlinear Jacobi method. If a user is willing to accept lower accuracy, e.g., $\varepsilon = 1.0\text{e}{-2}$ or $\varepsilon = 1.0\text{e}{-4}$, the nonlinear Gauss-Seidel method can be very efficient.

Chapter 4

Forward-Spot Market Model

Allaz and Vila [1] presented a forward market model with identical Cournot duopolists. They showed that even with certainty and perfect foresight, forward trading can improve market efficiency. Each of the producers will sell forward so as to make them worse off and make consumers better off than would be the case if the forward market did not exist. This phenomenon is similar to that of the prisoners' dilemma.

In the forward market model mentioned above, the inverse demand function is affine and the producers have the same linear cost function. Hence, one can solve for a Nash equilibrium of the forward market in closed form; see [1]. However, it is not clear that a Nash equilibrium would exist if the producers had nonidentical cost functions. Indeed, one can construct a simple example of Cournot duopolists with nonidentical linear cost functions for which the Allaz-Vila approach is not valid. If fact, the two-period forward market model belongs to a new class of mathematical programs called *Equilibrium Problems with Equilibrium Constraints* (EPECs), where each player solves a nonconvex *mathematical program with equilibrium constraints* (MPEC) [33], and a Nash equilibrium for an EPEC may not exist because of the nonconvexity in each player's problem. Pang and Fukushima [47] give a simple numerical example of such a case.

We observe that the mathematical structure of the two-period forward market model is similar to that of the multiple leader Stackelberg model analyzed by Sherali [59]. The similarity becomes evident when new variables are introduced for spot market sales. An immediate result is that we are able to adapt the analysis in [59] to establish the existence of a forward market equilibrium for M producers with nonidentical linear cost functions.

The remainder of this chapter is organized as follows. In the next section, we give a general formulation for the two-period forward market model with M producers. In Section 4.2, we reformulate the forward market equilibrium model by introducing new variables on spot market sales. Assuming that the inverse demand function is affine and allowing the producers to have nonidentical linear cost functions, we establish the existence of a forward market equilibrium. In Section 4.3, we use the sequential nonlinear complementarity (SNCP) algorithm proposed in Chapter 3 to compute a forward market Nash equilibrium for a three-producer example.

4.1 The Two-Period Forward Market Model

We use the following notation throughout this chapter:

M : number of producers,

f_i : producer i's forward sales in the first period,

Q_f : the total forward sales in the first period,

x_i : the production of producer i in the second period,

s_i : producer i's spot sales in the second period,

Q_s : the total spot sales in the second period,

$c_i(\cdot)$: the cost function of producer i,

$u_i(\cdot)$: the payoff function of producer i from the spot market in the second period,

$\pi_i(\cdot)$: the overall profit function of producer i,

$p_f(\cdot)$: the forward price (or inverse demand function) in the first period,

$p(\cdot)$: the spot price (or inverse demand function) in the second period.

4.1.1 The production game

Given the producers' forward position vector $f = (f_1, \ldots, f_M)$, the producers are playing a Cournot game in production quantities in the second period. For each $i = 1, \ldots, M$, producer i, assuming the production quantities $x_{-i} = (x_j)_{j=1,\, j \neq i}^{M}$ of the other producers are fixed at $\bar{x}_{-i}$, chooses the nonnegative production quantity x_i to maximize the payoff function in the second period:

$$[u_i(f_1, \ldots, f_M)](x_i, \bar{x}_{-i}) = p\left(x_i + \sum_{j=1,\, j\neq i}^{M} \bar{x}_j\right)(x_i - f_i) - c_i(x_i). \tag{4.1}$$

Indeed, if the producer i has already sold f_i in the forward market, it can only sell quantity $(x_i - f_i)$ in the spot market. The vector of production quantities $x^* = (x_1^*, \ldots, x_M^*)$ is said to be a *Nash-Cournot equilibrium* for the production game, if for each $i = 1, \ldots, M$, x_i^* solves

$$\underset{x_i \geq f_i}{\text{maximize}} \left\{ p(x_i + \sum_{j=1,\, j\neq i}^{M} x_j^*)(x_i - f_i) - c_i(x_i) \right\}. \tag{4.2}$$

Accordingly, we use $x(f) = (x_1(f), \ldots, x_M(f))$ to denote a Nash-Cournot production equilibrium x^* corresponding to the forward position vector f.

4.1.2 Forward market equilibrium

In the first period, the producers are playing a Cournot game in forward quantities. Assuming the forward position of other producers are fixed, producer i chooses his forward position $f_i(\geq 0)$ to maximize the overall profit function:

$$\begin{aligned} \pi_i(f_i, \bar{f}_{-i}) &= p_f\left(f_i + \sum_{j=1,\, j\neq i}^{M} \bar{f}_j\right)(f_i) + [u_i(f_i, \bar{f}_{-i})](x(f_i, \bar{f}_{-i})) \\ &= p\left(\sum_{j=1}^{M} x_j(f_i, \bar{f}_{-i})\right)(x_i(f_i, \bar{f}_{-i})) - c_i(x_i(f_i, \bar{f}_{-i})), \end{aligned} \tag{4.3}$$

since $p_f\left(f_i + \sum_{j=1,\, j\neq i}^{M} \bar{f}_j\right) = p\left(\sum_{j=1}^{M} x_j(f_i, \bar{f}_{-i})\right)$ under perfect foresight.

A vector $f^* = (f_1^*, \ldots, f_M^*)$ is said to be a *forward market equilibrium* if for $i = 1, \ldots, M$, f_i^* solves

$$\underset{f_i \geq 0}{\text{maximize}} \left\{ p\left(\sum_{j=1}^{M} x_j(f_i, f_{-i}^*)\right)\left(x_i(f_i, f_{-i}^*)\right) - c_i\left(x_i(f_i, f_{-i}^*)\right) \right\}. \tag{4.4}$$

Moreover, $x(f^*) = (x_j(f_j^*, f_{-j}^*,))_{j=1}^M$ is a Nash-Cournot equilibrium for the production game corresponding to the forward equilibrium f^*, as defined in (4.2).

4.2 Existence of a Forward Market Equilibrium

In [1], Allaz and Vila showed that one can solve for the forward market Nash equilibrium in closed form when demand and cost functions are affine and the producers have the same cost function, i.e., $c_i(x_i) = cx_i$, for $i = 1, \ldots, M$. In particular, in the case of Cournot duopolists, and the demand function $p(q) = a - q$, with $0 < c < a$, the unique forward market equilibrium outcome is

$$x_1 = x_2 = \frac{2(a-c)}{5}; \quad f_1 = f_2 = \frac{a-c}{5}; \quad p = c + \frac{a-c}{5}.$$

The purpose of this chapter is to establish an existence theorem for the forward market equilibrium when the M producers have *nonidentical* linear cost functions. Since the producers would only produce what they can sell, the production quantity equals the sum of forward sales and spot market sales. We then introduce new variable s_i for producer i's spot market sales and replace x_i by $f_i + s_i$ in the model of the production game to obtain the following equivalent formulation:

$$[u_i(f_1, \ldots, f_M)](s_i, \bar{s}_{-i}) = p\left(s_i + \sum_{j=1,\, j\neq i}^{M} \bar{s}_j + \sum_{j=1}^{M} f_j\right)(s_i) - c_i(s_i + f_i). \quad (4.5)$$

With the new variables s_i, we can define the spot market equilibrium. In particular, given a vector of forward positions $f \in R^M$, a vector of spot market sales $s^* \in R^M$ is said to be a *spot market equilibrium* if for each $i = 1, \ldots, M$, s_i^* solves

$$\underset{s_i \geq 0}{\text{maximize}} \left\{ p\left(s_i + \sum_{j=1,\, j\neq i}^{M} s_j^* + \sum_{j=1}^{M} f_j\right)(s_i) - c_i(s_i + f_i) \right\}. \quad (4.6)$$

Similarly, we use $s(f) = (s_1(f), \ldots, s_M(f))$ to denote the spot market equilibrium given the forward position f.

Lemma 4.1. *Given the forward sales f, if a vector $s(f)$ is a spot market equilibrium, then $x(f) = s(f) + f$ is a Nash equilibrium for the production game and* *vice versa.*

Proof. This is clear. ■

Following Lemma 4.1, an equivalent formulation for the forward market equilibrium (4.4) is as follows. A vector $f^* = (f_1^*, \ldots, f_M^*)$ is said to be a *forward market equilibrium* if for $i = 1, \ldots, M$, f_i^* solves

$$\begin{aligned} \underset{f_i \geq 0}{\text{maximize}} \quad & p\left(f_i + \sum_{j=1,\, j\neq i}^{M} f_j^* + \sum_{j=1}^{M} s_j(f_i, f_{-i}^*) \right) (f_i + s_i(f_i, f_{-i}^*)) \\ & -c_i(f_i + s_i(f_i, f_{-i}^*)). \end{aligned} \tag{4.7}$$

We observe that the spot and forward market equilibrium models ((4.6) and (4.7)) are similar to the multiple-leader Stackelberg model analyzed by Sherali [59]. The intuition is that in the forward market, every producer is a leader, while in the spot market, every producer becomes a follower, implementing his best response, given every producer's forward position in the first period. The two differences between the forward-spot market model and the multiple-leader Stackelberg model are:

(i) the cost function of a producer in the forward-spot market model is a function of both forward and spot sales;

(ii) each producer's revenue in the first period includes the spot market sales $s_i(f_i, f_{-i}^*)$ from the second period.

It turns out that these two differences are problematic. For (i), in contrast to the case in [59], the spot market sales $s(f)$ can not be simplified to a function of total forward sales $Q_f = \sum_{j=1}^M f_j$ only; this is due to the f_j term in the cost function $c_i(\cdot)$. Hence, the aggregate spot market reaction curve $Q_s(f) = \sum_{i=1}^M s_i(f)$ is a function of the forward sales f, and in general, can not be reduced to a function of the total forward market sales, Q_f, a variable in R^1. The aggregate spot market sales, Q_s, being a function of total forward market sales only, is crucial in the analysis. Below, we show that the aggregate spot market sales Q_s is indeed a function of the aggregate forward market sales Q_f when the inverse demand function is affine and producers' cost functions are linear.

Assumption 4.2. *We assume that the inverse demand function is* $p(z) := a - bz$ *for* $z \geq 0$ *with* $a, b > 0$*, and for each* $i = 1, \ldots, M$*, producer* i*'s cost function is* $c_i(z) := c_i z$ *with* $c_i > 0$.

Remark. Since the producers are maximizing their profits, and $p(z) < 0$ for all $z > a/b$, Assumption 4.2 also implies that no producer will produce more than a/b units.

Proposition 4.3. *Let Assumption 4.2 hold. Given producers' forward sales* f*, the spot market sales* $s_i(f)$*, for* $i = 1, \ldots, M$ *and the aggregate spot market sales* $Q_s(f)$ *can be simplified to a function of* Q_f*, and denoted as* $s_i(Q_f)$ *and* $Q_s(Q_f)$ *respectively.*

Proof. Substitute $p(z) = a - bz$, $c_i(z) = c_i z$, and $Q_f = \sum_{j=1}^{M} f_j$ into (4.6); then producer i's profit maximization problem in the spot market equilibrium model becomes

$$\underset{s_i \geq 0}{\text{maximize}} \left\{ -c_i f_i - \left[b \left(\sum_{j=1,\, j\neq i}^{M} s_j^* + Q_f \right) + c_i - a \right] s_i - b s_i^2 \right\}. \tag{4.8}$$

Since the objective function in (4.8) is strictly concave, the spot equilibrium s^* exists and is unique. Furthermore, it must satisfy the (necessary and sufficient) KKT conditions:

$$0 \leq s_i \perp b \left(\sum_{j=1}^{M} s_j + Q_f \right) + c_i - a + b s_i \geq 0, \quad i = 1, \ldots, M. \tag{4.9}$$

Since f_i does not appear explicitly in (4.9) and s_i^* is unique, we can denote s_i^* as $s_i(Q_f)$ for each $i = 1, \ldots, M$. Consequently, the aggregate spot market sales $Q_s = \sum_{i=1}^{M} s_i(Q_f)$ is a function of the total forward sales Q_f and is denoted as $Q_s(Q_f)$. ■

Define $\mathcal{I}(Q_f) := \{i : s_i(Q_f) > 0\}$ and let $|\mathcal{I}(Q_f)|$ be the cardinality of index set $\mathcal{I}(Q_f)$. Let $s_i^+(Q_f)$ and $Q_s^+(Q_f)$ denote the right-hand derivatives of $s_i(Q_f)$ and $Q_s(Q_f)$, respectively. The following theorem states the properties of $s_i(Q_f)$ and $Q_s(Q_f)$ corresponding to the forward sales Q_f.

Theorem 4.4. *Suppose Assumption 4.2 holds and let Q_f be the aggregate forward sales. For each $i = 1, \ldots, M$, the spot market sales $s_i(Q_f)$ and the aggregate spot market sales $Q_s(Q_f)$ satisfy the following properties for each $Q_f > 0$:*

(i) for $i = 1, \ldots, M$, each $s_i(Q_f)$ is a continuous, nonnegative, decreasing, piecewise linear concave function in $\{Q_f \geq 0 : s_i(Q_f) > 0\}$ with

$$s_i^+(Q_f) = \begin{cases} \dfrac{-1}{|\mathcal{I}(Q_f)| + 1} < 0 & \text{if } i \in \mathcal{I}(Q_f), \\ 0 & \text{otherwise;} \end{cases} \tag{4.10}$$

(ii) the aggregate spot sales $Q_s(Q_f)$ is a continuous, nonnegative, decreasing, piecewise linear convex function in Q_f for $\{Q_f \geq 0 : Q_s(Q_f) > 0\}$ with

$$Q_s^+(Q_f) = \begin{cases} \dfrac{-|\mathcal{I}(Q_f)|}{|\mathcal{I}(Q_f)| + 1} < 0 & \text{if } Q_s(Q_f) > 0, \\ 0 & \text{if } Q_s(Q_f) = 0; \end{cases} \tag{4.11}$$

(iii) the function $T(Q_f) = Q_f + Q_s(Q_f)$ is an increasing, piecewise linear convex function in $Q_f \geq 0$.

Proof. Notice that with the inverse demand function $p(\cdot)$ and cost function $c_i(\cdot)$ stated in Assumption 4.2, the objective function in (4.8) for the spot market equilibrium model is identical to the oligopolistic market equilibrium model studied in [40], except for an extra constant term $c_i f_i$ in (4.8). However, adding a constant to the objective function will not change the optimal solution. Indeed, we solve the same KKT system (4.9) for the equilibrium solutions of these two models.

For (i) and (ii), Sherali et al. [61] proved that for $i = 1, \ldots, M$, $s_i(Q_f)$ and $Q_s(Q_f)$ are continuous and nonnegative. For a fixed Q_f, if $\mathcal{I}(Q_f) \neq \emptyset$, then $Q_s(Q_f) > 0$ and from the KKT conditions (4.9), we have

$$b\left(Q_s(Q_f) + Q_f\right) + c_i - a + b s_i(Q_f) = 0, \quad \forall\, i \in \mathcal{I}(Q_f).$$

Summing over $i \in \mathcal{I}(Q_f)$, we obtain

$$b|\mathcal{I}(Q_f)|\left(Q_s(Q_f) + Q_f\right) + \sum_{i \in \mathcal{I}(Q_f)} c_i - |\mathcal{I}(Q_f)|a + bQ_s(Q_f) = 0,$$

which gives

$$Q_s(Q_f) = \frac{|\mathcal{I}(Q_f)|a - b|\mathcal{I}(Q_f)|Q_f - \sum_{i \in \mathcal{I}(Q_f)} c_i}{b(|\mathcal{I}(Q_f)| + 1)}, \tag{4.12}$$

and

$$s_i(Q_f) = \frac{a - bQ_f + \sum_{j \in \mathcal{I}(Q_f)} c_j}{b(|\mathcal{I}(Q_f)| + 1)} - \frac{c_i}{b}, \quad \forall\, i \in \mathcal{I}(Q_f). \tag{4.13}$$

Taking the right-hand derivatives of $Q_s(Q_f)$ and $s_i(Q_f)$ with respect to Q_f, we obtain

$$Q_s^+(Q_f) = \frac{-|\mathcal{I}(Q_f)|}{|\mathcal{I}(Q_f)| + 1} < 0, \quad \text{if } Q_s(Q_f) > 0, \tag{4.14}$$

and

$$s_i^+(Q_f) = -\frac{1}{|\mathcal{I}(Q_f)| + 1} < 0, \quad \text{if } i \in \mathcal{I}(Q_f). \tag{4.15}$$

Now, suppose that $s_i(Q_f) = 0$ and $s_i^+(Q_f) > 0$ for some $Q_f \geq 0$. Then there must exist a point $\hat{Q}_f$ near Q_f such that $s_i(\hat{Q}_f) > 0$ and $s_i^+(\hat{Q}_f) > 0$; this leads to a contradiction. Hence, if $s_i(Q_f) = 0$, then $s_i^+(Q_f) = 0$, and $s_i(\bar{Q}_f) = 0$ for all $\bar{Q}_f \geq Q_f$. In other words, if it is not profitable for the producer i to sell in the spot market for a given aggregate forward market sales Q_f, then he will not be active in the spot market for any aggregate forward sales $\bar{Q}_f$ greater than Q_f.

The same implication also holds for the aggregate spot market sales $Q_s(Q_f)$, i.e., if $Q_s(Q_f) = 0$, then $Q_s^+(Q_f) = 0$ and $Q_s(\bar{Q}_f) = 0$ for all $\bar{Q}_f \geq Q_f$.

Observe that $Q_f^1 > Q_f^2$ implies $\mathcal{I}(Q_f^2) \supseteq \mathcal{I}(Q_f^1)$. Furthermore, for $i = 1, \ldots, M$, we have

$$s_i^+(Q_f^2) > s_i^+(Q_f^1), \quad \text{if } \mathcal{I}(Q_f^2) \supset \mathcal{I}(Q_f^1);$$

$$s_i^+(Q_f^2) = s_i^+(Q_f^1), \quad \text{if } \mathcal{I}(Q_f^2) = \mathcal{I}(Q_f^1).$$

This proves that for each $i = 1, \ldots, M$, $s_i(Q_f)$ is nonincreasing, piecewise linear and concave in $Q_f \geq 0$. Similarly, we can establish that $Q_s^+(Q_f) = 0$ if $Q_s(Q_f) = 0$ and $Q_s(Q_f)$ is nonincreasing, piecewise linear and convex in $Q_f \geq 0$. This completes the proof for (ii).

For (iii), since $T^+(Q_f) = 1 + Q_s^+(Q_f) = \frac{1}{|\mathcal{I}(Q_f)|+1} > 0$, the function $T(Q_f)$ is increasing in Q_f for $Q_f \geq 0$. ■

Lemma 4.5. *Let Assumption 4.2 hold. For all $i = 1, \dots, M$ and any fixed vector $\bar{f}_{-i} = (\bar{f}_j)_{j \neq i}$, the objective function in (4.7) is strictly concave in f_i. Furthermore, there exists a unique optimal solution $f_i^*(\bar{f}_{-i})$ for the problem (4.7), with $0 \leq f_i^*(\bar{f}_{-i}) \leq a/b$.*

Proof. The proof closely follows the analysis of Lemma 1 in [59]. To ease the notation, we define $\bar{y}_i = \sum_j (\bar{f}_{-i})_j$ and

$$\begin{aligned} g_i(f_i, \bar{y}_i) = \quad & (a - f_i - \bar{y}_i - Q_s(f_i + \bar{y}_i)) \, (f_i + s_i(f_i + \bar{y}_i)) \\ & - (c_i f_i + c_i s_i(f_i + \bar{y}_i)) \, . \end{aligned} \tag{4.16}$$

Then for a fixed vector $\bar{y}_i$, producer i's profit maximization problem (4.7) can be written as

$$\underset{f_i \geq 0}{\text{maximize}} \quad \{g(f_i, \bar{y}_i)\} \, . \tag{4.17}$$

Let $g_i^+(f_i, \bar{y})$ denote the right-hand derivative of $g_i(f_i, \bar{y})$ with respect to f_i. To show that $g_i(f_i, \bar{y})$ is strictly concave in f_i, it suffices to show that for any $f_i \geq 0$, there exists a $\bar{\delta} > 0$ such that $g_i^+(f_i + \delta, \bar{y}) < g_i^+(f_i, \bar{y})$ for all $0 \leq \delta < \bar{\delta}$.

Without loss of generality, we assume $\mathcal{I}(f_i + \bar{y}_i) = \{i : s_i(f_i + \bar{y}_i) > 0\} \neq \emptyset$ at a given point $(f_i, \bar{y}_i)$; the assertion of the lemma on strict concavity of $g_i(f_i, \bar{y}_i)$ (4.16) in f_i is obvious when $\mathcal{I}(f_i + \bar{y}_i) = \emptyset$. We then choose

$$\bar{\delta} := \text{argmin}\{\delta : \mathcal{I}(f_i + \delta + \bar{y}_i) \subset \mathcal{I}(f_i + \bar{y}_i)\}.$$

It is easy to verify that for all $0 \leq \delta < \bar{\delta}$,

$$\begin{aligned} & s_i^+(f_i + \bar{y}_i) = s_i^+(f_i + \delta + \bar{y}_i), \qquad \text{for } i = 1, \dots, M, \\ \text{and} \quad & Q_s^+(f_i + \bar{y}_i) = Q_s^+(f_i + \delta + \bar{y}_i). \end{aligned} \tag{4.18}$$

Using $T(Q_f) = Q_f + Q_s(Q_f)$ and $p(z) = a - bz$ to further simplify the notation, we obtain

$$\begin{aligned}
&g_i^+(f_i+\delta,\bar{y}) - g_i^+(f_i,\bar{y}) = \{c_i(s_i^+(f_i+\bar{y}_i) - s_i^+(f_i+\delta+\bar{y}_i))\} \\
&+\{p(T(f_i+\delta+\bar{y}_i))(1+s_i^+(f_i+\delta+\bar{y}_i)) - p(T(f_i+\bar{y}_i))(1+s_i^+(f_i+\bar{y}_i))\} \\
&+\{p'(T(f_i+\delta+\bar{y}_i))T^+(f_i+\delta+\bar{y}_i)\,(f_i+\delta+s_i(f_i+\delta+\bar{y}_i)) \\
&-p'(T(f_i+\bar{y}_i))T^+(f_i+\bar{y}_i)\,(f_i+s_i(f_i+\bar{y}_i))\}.
\end{aligned} \tag{4.19}$$

Consider each term $\{\cdot\}$ in (4.19) separately. Since $s_i^+(f_i+\bar{y}_i) = s_i^+(f_i+\delta+\bar{y}_i)$, the first term $\{\cdot\}$ in (4.19) equals 0. Since $p(z)$ is strictly decreasing and $T(Q_f)$ is strictly increasing, the second term $\{\cdot\}$ in (4.19) is negative. Finally, since $p'(z) = -b < 0$, $T^+(f_i+\bar{y}_i) = T^+(f_i+\delta\bar{y}_i) > 0$, and

$$\delta + s_i(f_i+\delta+\bar{y}_i) - s_i(f_i+\delta+\bar{y}_i) = \delta(1 - \frac{1}{|\mathcal{I}(f_i+\bar{y}_i)|}) > 0,$$

the third term $\{\cdot\}$ in (4.19) is negative. This completes the proof that the objective function in (4.7) is strictly concave in f_i. Moreover, since no firm will produce beyond a/b units (see the remark on Assumption 4.2), the optimal solution $f_i^*(\bar{f}_{-i})$ is unique with $0 \le f_i^*(\bar{f}_{-i}) \le a/b$. ■

We are now ready to state the existence of the forward market equilibrium.

Theorem 4.6. *If Assumption 4.2 holds, then there exists a forward market equilibrium.*

Proof. For a given vector $f = (f_1, \ldots, f_M)$, we define a point-to-point map

$$F(f) = (f_1^*(f_{-1}), \ldots, f_M^*(f_{-M})).$$

In light of Lemma 4.5, $f_i^*(f_{-i})$ is the unique optimal solution of (4.7) for the corresponding vector f_{-i} for all $i = 1, \ldots, M$. If we can show that the map F is continuous, then by Brouwer's fixed point theorem, it will have a fixed point on the compact convex set $C = \{f : 0 \le f_i \le \frac{a}{b}, \quad \forall i = 1, \ldots, M\}$. Furthermore, this fixed point is a forward market equilibrium.

To show F is continuous, it suffices to show that $f_i^*(f_{-i})$ is continuous for $i = 1, \ldots, M$. Consider a sequence $\{f_{-i}^k\}$ converging to $\bar{f}_{-i}$, and let f_i^k denote

$f_i^*(f_{-i}^k)$ of (4.7) for the corresponding fixed vector f_{-i}^k. Since the sequence f_i^k is contained in the interval $[0, \frac{a}{b}]$, there exists a convergent subsequence. Without loss of generality, we assume the sequence $\{f_i^k\}$ converges to $\bar{f}_i$. It remains to show that $f_i^*(\bar{f}_{-i}) = \bar{f}_i$.

Choose any $\hat{f}_i \geq 0$. Since f_i^k is the optimal solution of (4.17), for a fixed vector f_{-i}^k, we have

$$g_i(f_i^k, \sum_j (f_{-i}^k)_j) \geq g_i(\hat{f}_i^k, \sum_j (f_{-i}^k)_j), \quad \forall k,$$

where $g_i(\cdot, \cdot)$ is defined in (4.16). As $k \to \infty$, by the continuity of $g_i(\cdot, \cdot)$, we obtain

$$g_i(\bar{f}_i, \sum_j (\bar{f}_{-i})_j) \geq g_i(\hat{f}_i, \sum_j (\bar{f}_{-i})_j).$$

Since $\hat{f}_i$ is chosen arbitrarily, the above inequality holds for any $\hat{f}_i \geq 0$. This implies that $\bar{f}_i$ is an optimal solution of (4.7) for the corresponding $\bar{f}_{-i}$. It follows that $\bar{f}_i = f_i^*(\bar{f}_{-i})$, since from Lemma 4.5, the optimal solution is unique. This completes the proof. ■

4.3 An EPEC Approach for Computing a Forward Market Equilibrium

The computation of a forward market equilibrium involves solving a family of concave maximization problems, as defined in (4.7). However, the objective functions are nonsmooth in these problems because $s_i(Q_f)$ is piecewise linear concave in Q_f on $\{Q_f : s_i(Q_f) > 0\}$; see Theorem 4.4. One might encounter difficulties in solving these nonsmooth problems with nonlinear programming solvers. An alternative to avoid the nonsmooth objective functions is to formulate the forward market equilibrium problem as an *equilibrium problem with equilibrium constraints* (EPEC). Since we do not know the explicit representation of s_i as a function of Q_f, we return to the original formulation of the forward market equilibrium model (4.4) and embed the necessary and sufficient KKT conditions for spot market equilibrium s^* as a set of constraints in producer i's profit maximization problem in the forward market:

$$\begin{aligned}
&\underset{(f_i, s, \theta_i)}{\text{maximize}} && (\theta_i - c_i)(f_i + s_i) \\
&\text{subject to} && \theta_i = a - b\left(f_i + e^{\mathrm{T}} s + \sum_{j \neq i} \bar{f}_j\right), \\
& && 0 \leq s \perp c - \theta_i\, e + bs \geq 0, \\
& && f_i \geq 0,
\end{aligned} \tag{4.20}$$

where $c = (c_1, \ldots, c_M)$, $f = (f_1, \ldots, f_M)$, and e is a vector of all ones of the proper dimension.

Observe that producer i's profit maximization problem (4.20) is an MPEC because it includes complementarity constraints

$$0 \leq s \perp c - \theta_i\, e + bs \geq 0. \tag{4.21}$$

Furthermore, each MPEC is parameterized by other producers' forward sales $\bar{f}_{-i}$. Following the notation in Chapter 3, we denote producer i's maximization problem by MPEC($\bar{f}_{-i}$).

The problem of finding a forward market equilibrium solution is formulated as an EPEC:

$$\begin{aligned}
&\text{Find } (f^*, s^*, \theta^*) \text{ such that for all } i = 1, \ldots, M, \\
&(f_i^*, s^*, \theta_i^*) \text{ solves MPEC}(f_{-i}^*)\,(4.20).
\end{aligned} \tag{4.22}$$

The following theorem expresses the strong stationarity conditions for a solution to the forward market equilibrium model (4.22) in which the i-th MPEC takes the form of (4.20).

Theorem 4.7. *Suppose (f^*, s^*, θ^*) is a solution for the forward market equilibrium model (4.22). If for each $i = 1, \ldots, M$, the MPEC-LICQ holds at a feasible point (f_i^*, s^*) for the i-th MPEC (4.20), then (f^*, s^*) is an EPEC* strongly stationary *point. In particular, there exist vectors $\lambda^* = (\lambda^{1*}, \ldots, \lambda^{M*})$ with*

$\lambda^{i*} = (\lambda^{f,i*}, \lambda^{s,i*}, \lambda^{c,i*}, \lambda^{sc,i*})$ *such that* (f^*, s^*, λ^*) *solves the system*

$$
\left.\begin{aligned}
-a + be^{\mathrm{T}}(f+s) + bf_i + c_i - \lambda^{f,i} - be^{\mathrm{T}}\lambda^{c,i} + bs^{\mathrm{T}}\lambda^{sc,i} \quad &= 0, \\
bf_i e - 2b[e_i \circ s] - \lambda^{s,i} - b\lambda^{c,i} + b\,[s \circ \lambda^{c,i}] \quad & \\
+ \left[\left(be^{\mathrm{T}}(f+s)e + c - ae + bs\right) \circ \lambda^{sc,i}\right] + b(ee^{\mathrm{T}} + I)[s \circ \lambda^{sc\,i}] \quad &= 0, \\
0 \le f_i \quad \perp \quad \lambda^{f,i} \ge 0, & \\
0 \le s \quad \perp \quad \lambda^{s,i} \ge 0, & \\
0 \le c - \left(a - b(e^{\mathrm{T}}f + e^{\mathrm{T}}s)\right)\, e + bs \quad \perp \quad \lambda^{c,i} \ge 0, & \\
0 \le -s \circ c - \left(a - b(e^{\mathrm{T}}f + e^{\mathrm{T}}s)\right)\, e + bs, \qquad \lambda^{sc,i} \ge 0, &
\end{aligned}\right\} \tag{4.23}
$$

$$
i = 1, \ldots, M.
$$

Conversely, if (f^*, x^*, λ^*) *is a solution to the system (4.23), then* (f^*, x^*) *is an EPEC* B(ouligand)-stationary *point of the forward market equilibrium model (4.22).*

Proof. This is Theorem 3.2. ■

In what follows, we apply the SNCP algorithm proposed in Chapter 3 to solve the EPEC formulation for the forward market equilibrium model (4.22) for each of the four scenarios. The SNCP algorithm finds a solution of the system (4.23), if one exists.

4.3.1 An example with three producers

Consider the case with three producers in the market. The cost function of producer i is

$$
c_i(z) = c_i z, \text{ with } (c_1, c_2, c_3) = (2, 3, 4).
$$

The inverse demand function is $p(z) = 10 - z$. We analyze the following four scenarios:

(1) no producer contracts forward sales;

(2) only one producer is allowed to contract forward sales;

(3) only two producers are allowed to contract forward sales;

(4) all three producers can participate in the forward market.

The computational results for each scenario are summarized in Table 4.2. The notation used in that table is explained in Table 4.1.

TABLE 4.1
Notation used for computational results.

Cases	List of producers allowed to sell in the forward market.
$f = (f_1, f_2, f_3)$	The vector of producers' forward sales.
$x = (x_1, x_2, x_3)$	The vector of producers' production quantities.
$\pi = (\pi_1, \pi_2, \pi_3)$	The vector of producers' overall profits.
p	The spot (and forward) price.

TABLE 4.2
Computational results on four scenarios.

Cases	f	x	π	p
No forward sales	(0, 0, 0)	(2.75, 1.75, 0.75)	$(7.563, 3.063, 0.563)$	4.75
Only producer 1	(3, 0, 0)	(5, 1, 0)	$(10, 1, 0)$	4
Only producer 2	(0, 3, 0)	(2, 4, 0)	$(4, 4, 0)$	4
Only producer 3	(0, 0, 1)	(2.5, 1.5, 1.5)	$(6.25, 2.25, 0.75)$	4.5
Only producer 1 & 2	(2, 1, 0)	(4, 2, 0)	$(8, 2, 0)$	4
Only producer 1 & 3	(7/3, 0, 1/3)	(40/9, 10/9, 1/3)	$(9.383, 1.235, 0.037)$	4.111
Only producer 2 & 3	(0, 2.25, 0.25)	(2.125, 3.375, 0.375)	$(4.516, 3.797, 0.047)$	4.125
All producers	(2, 1, 0)	(4, 2, 0)	$(8, 2, 0)$	4

We summarize some observations based on the computational results for the four scenarios. First, we notice that allowing more producers to participate in the forward market will not necessarily decrease the market clearing price. For example, the price is 4 when only producer 2 is allowed to produce in the forward market and it increases to 4.125 when producer 3 joins producer 2 to produce in the forward market. In contrast to the outcome of the Cournot game, each producer can increase production by being the only player in the forward market.

If the producers have the chance to sell forward, they can do so profitably. For example, if only producer 1 is allowed to sell forward, he can increase his profit from 7.5625 (the profit in the Cournot game) to 10; if only producer 2 is allowed to contract forward sales, producer 1 can increase his profit from 4 to 8 if he starts selling in the forward market, and similarly for producer 3. This is similar to the conclusion on the emergence of a forward market in [1] for the case of identical producers. However, if all producers participate in the forward market, producer 1 is better off and producers 2 and 3 are worse off than would be the case if the forward market did not exist; this phenomenon is in contrast to that of the prisoners' dilemma observed for the case of identical producers in [1].

Finally, the results suggest that the market is the most efficient (in terms of the clearing price) when producers 2 and 3 both participate in the forward market, in which case, producer 3 will not produce in the forward market even if he is allowed to do so.

4.4 Conclusions

We have established the existence of a forward market equilibrium for the case of M producers with nonidentical linear cost functions. We also proposed an EPEC formulation for the forward market equilibrium model and applied a sequential nonlinear complementarity algorithm to compute a forward market equilibrium. The computational results suggest that if a producer has the chance to sell forward, he can do so profitably. However, if all producers start selling forward contracts, they are worse off and the consumers are better off than would be the case if the forward market did not exist. This supplements Allaz and Vila's results on producers with identical linear cost functions.

Chapter 5

Computation of Moral-Hazard Problems

In this chapter, we study computational aspects of moral-hazard problems. In particular, we consider deterministic contracts as well as contracts with action and/or compensation lotteries, and formulate each case as a mathematical program with equilibrium constraints. We investigate and compare solution properties of the MPEC approach to that of the linear programming (LP) approach with lotteries. We propose a hybrid procedure that combines the best features of both. The hybrid procedure obtains a solution that is, if not global, at least as good as an LP solution. It also preserves the fast local convergence property by applying an SQP algorithm to MPECs. The numerical results on an example show that the hybrid procedure outperforms the LP approach in both computational time and solution quality in terms of the optimal objective value.

5.1 Introduction

We study mathematical programming approaches to solve moral-hazard problems. More specifically, we formulate moral-hazard problems with finitely many action choices, including the basic deterministic models and models with lotteries, as mathematical programs with equilibrium constraints. One advantage of using an MPEC formulation is that the size of resulting program is often orders of magnitude smaller than the linear programs derived from the LP lotteries approach [50, 51]. This feature makes the MPEC approach an appealing alternative when solving a large-scale linear program is computationally infeasible because of limitations on computer memory or computing time.

The moral-hazard model studies the relationship between a principal (leader) and an agent (follower) in situations in which the principal can neither observe nor verify an agent's action. The model is formulated as a bilevel program, in which the principal's upper-level decision takes the agent's best response to the principal's decision into account. Bilevel programs are generally difficult mathematical problems, and much research in the economics literature has been devoted to analyzing and characterizing solutions of the moral-hazard model (see Grossman and Hart [18] and the references therein). When the agent's set of actions is a continuum, an intuitive approach to simplifying the model is to assume the agent's optimal action lies in the interior of the action set. One then can treat the agent's problem as an unconstrained maximization problem and replace it by the first-order optimality conditions. This is called the *first-order approach* in the economics literature. However, Mirrlees [35, 36] showed that the first-order approach may be invalid because the lower-level agent's problem is not necessarily a concave maximization program, and that the optimal solution may fail to be unique and interior. Consequently, a sequence of papers [54, 19, 26] has developed conditions under which the first-order approach is valid. Unfortunately, these conditions are often more restrictive than is desirable.

In general, if the lower-level problem in a bilevel program is a convex minimization (or concave maximization) problem, one can replace the lower-level problem by the first-order optimality conditions, which are both necessary and sufficient, and reformulate the original bilevel problem as an MPEC. This idea is similar to the first-order approach to the moral-hazard problem with one notable difference: MPEC formulations include complementarity constraints. The first-order approach assumes that the solution to the agent's problem lies in the interior of the action set, and hence, one can treat the agent's problem as an unconstrained maximization problem. This assumption may also avoid issues associated with the failure of the constraint qualification at a solution. General bilevel programs do not make an interior solution assumption. As a result, the complementarity conditions associated with the Karush-Kuhn-Tucker multipliers for inequality constraints would appear in the first-order optimality conditions for the lower-level program. MPECs also arise in many applications in engineering (e.g., transportation, contact problems, mechanical structure design) and economics (Stackelberg games, optimal taxation problems). One well known theoretical difficulty with MPECs is that the standard constraint qualifications, such as the linear independence constraint qualification and the Mangasarian-Fromovitz constraint qualification, fail at every feasible point. A considerable amount of literature is devoted to refining constraint qualifications and stationarity conditions for MPECs; see Scheel and Scholtes [58] and the references therein. We also refer to the two-volume monograph by Facchinei and Pang [11] for theory and applications of equilibrium problems and to the monographs by Luo et al. [33] and Outrata et al. [43] for more details on MPEC theory and applications.

The failure of the constraint qualification conditions means that the set of Lagrange multipliers is unbounded and that conventional numerical optimization software may fail to converge to a solution. Economists have avoided these numerical problems by reformulating the moral-hazard problem as a linear program involving lotteries over a finite set of outcomes. See Townsend [64, 65] and Prescott [50, 51]. While this approach avoids the constraint qualification problems, it does so by restricting aspects of the contract, such as consumption, to a finite set of possible choices even though a continuous choice formulation would be economically more natural.

The purpose of this chapter is twofold: (1) to introduce to the economics community the MPEC approach, or more generally, advanced equilibrium programming approaches, to the moral-hazard problem; (2) to present an interesting and important class of incentive problems in economics to the mathematical programming community. Many incentive problems, such as contract design, optimal taxation and regulation, and multiproduct pricing, can be naturally formulated as an MPEC or an *equilibrium problem with equilibrium constraints* (EPEC). This greatly extends the applications of equilibrium programming to one of the most active research topics in economics in past three decades. The need for a global solution for these economical problems provides a motivation for the optimization community to develop efficient global optimization algorithms for MPECs and EPECs.

The remainder of this chapter is organized as follows. In the next section, we describe the basic moral-hazard model and formulate it as a mixed-integer nonlinear program and as an MPEC. In Section 5.3, we consider moral-hazard problems with action lotteries, with compensation lotteries, and with a combination of both. We derive MPEC formulations for each of these cases. We also compare the properties of the MPEC approach and the LP lottery approach. In Section 5.4, we develop a hybrid approach that preserves the desired global solution property from the LP lottery approach and the fast local convergence of the MPEC approach. The numerical efficiency of the hybrid approach in both computational speed and robustness of the solution is illustrated in an example in Section 5.5.

5.2 The Basic Moral-Hazard Model

5.2.1 The deterministic contract

We consider a moral-hazard model in which the agent chooses an action from a finite set $\mathcal{A} = \{a_1, \ldots, a_M\}$. The outcome can be one of N alternatives. Let $\mathcal{Q} = \{q_1, \ldots, q_N\}$ denote the outcome space, where the outcomes are dollar returns to the principal ordered from smallest to largest.

The principal can only observe the outcome, not the agent's action. However, the stochastic relationship between actions and outcomes, which is often called a *production technology*, is common knowledge to both the principal and the agent. Usually, the production technology is exogenously described by the probability distribution function, $p(q \,|\, a)$, which presents the probability of outcome $q \in \mathcal{Q}$ occurring given that action a is taken. We assume $p(q \,|\, a) > 0$ for all $q \in \mathcal{Q}$ and $a \in \mathcal{A}$; this is called the *full-support assumption*.

Since the agent's action is not observable to the principal, the payment to the agent is only based on the outcome observed by the principal. Let $\mathcal{C} \subset R$ be the set of all possible compensations.

Definition 5.1. A *compensation schedule* $c = (c(q_1), \ldots, c(q_N)) \in R^N$ is an agreement between the principal and the agent such that $c(q) \in \mathcal{C}$ is the payoff to the agent from the principal if outcome $q \in \mathcal{Q}$ is observed.

The agent's utility $u(x, a)$ is a function of the payment $x \in R$ received from the principal and of his action $a \in \mathcal{A}$. The principal's utility $w(q - x)$ is a function over net income $q - x$ for $q \in \mathcal{Q}$. We let $W(c,a)$ and $U(c,a)$ denote the *expected utility* to the principal and agent, respectively, of a compensation schedule $c \in R^N$ when the agent chooses action $a \in \mathcal{A}$, i.e.,

$$\begin{aligned} W(c,a) &= \sum_{q \in \mathcal{Q}} p(q \,|\, a)\, w\,(q - c(q))\,, \\ U(c,a) &= \sum_{q \in \mathcal{Q}} p(q \,|\, a)\, u\,(c(q), a)\,. \end{aligned} \tag{5.1}$$

Definition 5.2. A *deterministic contract* (proposed by the principal) consists of a recommended action $a \in \mathcal{A}$ to the agent and a compensation schedule $c \in R^N$.

The contract has to satisfy two conditions to be accepted by the agent. The first condition is the *participation constraint.* It states that the contract must give the agent an expected utility no less than a required utility level U^*:

$$U(c,a) \geq U^*. \tag{5.2}$$

The value U^* represents the highest utility the agent can receive from other activities if he does not sign the contract.

Second, the contract must be *incentive compatible* to the agent; it has to provide incentives for the agent not to deviate from the recommended action. In particular, given the compensation schedule c, the recommended action a must be optimal from the agent's perspective by maximizing the agent's expected utility function. The *incentive compatibility constraint* is given as follows:

$$a \in \operatorname{argmax}\{U(c,a) :\ a \in \mathcal{A}\}. \tag{5.3}$$

For a given U^*, a *feasible* contract satisfies the participation constraint (5.2) and the incentive compatibility constraint (5.3). The objective of the principal is to find an *optimal deterministic contract*, a feasible contract that maximizes his expected utility. A mathematical program for finding an optimal deterministic contract (c^*, a^*) is

$$\begin{aligned}
\underset{(c,a)}{\text{maximize}} \quad & W(c,a) \\
\text{subject to} \quad & U(c,a) \geq U^*, \\
& a \in \operatorname{argmax}\{U(c,a) : a \in \mathcal{A}\}.
\end{aligned} \tag{5.4}$$

Since there are only finitely many actions in $\mathcal{A}$, the incentive compatibility constraint (5.3) can be presented as the following set of inequalities:

$$U(c,a) \geq U(c,a_i), \quad \text{for } i = 1, \ldots, M. \tag{5.5}$$

These constraints ensure that the agent's expected utility obtained from choosing the recommendation action is no worse than that from choosing other actions. Replacing (5.3) by the set of inequalities (5.5), we have an equivalent formulation of the optimal contract problem:

$$\begin{aligned}
\underset{(c,a)}{\text{maximize}} \quad & W(c,a) \\
\text{subject to} \quad & U(c,a) \geq U^*, \\
& U(c,a) \geq U(c,a_i), \quad \text{for } i = 1, \ldots, M, \\
& a \in \mathcal{A} = \{a_1, \ldots, a_M\}.
\end{aligned} \tag{5.6}$$

5.2.2 A mixed-integer NLP formulation

The optimal contract problem (5.6) can be formulated as a mixed-integer nonlinear program. Associated with each action $a_i \in \mathcal{A}$, we introduce a binary variable $y_i \in \{0, 1\}$. Let $y = (y_1, \ldots, y_M) \in R^M$ and let e_M denote the vector of all ones in R^M. The mixed-integer nonlinear programming formulation for the optimal contract problem (5.6) is

$$\begin{aligned}
\underset{(c,y)}{\text{maximize}} \quad & W(c, \sum_{i=1}^{M} a_i y_i) \\
\text{subject to} \quad & U(c, \sum_{i=1}^{M} a_i y_i) \geq U^*, \\
& U(c, \sum_{i=1}^{M} a_i y_i) \geq U(c, a_j), \quad \forall\, j = 1, \ldots, M, \\
& e_M^{\mathrm{T}} y = 1, \\
& y_i \in \{0, 1\} \quad \forall\, i = 1, \ldots, M.
\end{aligned} \tag{5.7}$$

The above problem has N nonlinear variables, M binary variables, one linear constraint and $(M+1)$ nonlinear constraints. To solve a mixed-integer nonlinear program, one can use MINLP [12], BARON [55] or other solvers developed for this class of programs. For (5.7), since the agent will choose one and only one action, the number of possible combinations on the binary vector y is only M. One then can solve (5.7) by solving M nonlinear programs with $y_i = 1$ and the other $y_j = 0$ in the i-th nonlinear program, as Grossman and Hart suggested in [18] for the case where the principal is risk averse. They further point out that each nonlinear program is a convex program if the agent's utility function $u(x, a)$ can be written as $G(a) + K(a)V(x)$, where (1) V is a real-valued, strictly increasing, concave function defined on some open interval $\mathcal{I} = (\underline{I}, \bar{I}) \subset R$; (2) $\lim_{x \to \underline{I}} V(x) = -\infty$; (3) G, K are real-valued functions defined on $\mathcal{A}$ and K is strictly positive; (4) $u(x, a) \geq u(x, \hat{a}) \Rightarrow u(\hat{x}, a) \geq u(\hat{x}, \hat{a})$, for all a, $\hat{a} \in \mathcal{A}$, and x, $\hat{x} \in \mathcal{I}$. The above assumption implies that the agent's preferences over income lotteries are independent of his action.

5.2.3 An MPEC formulation

In general, a mixed-integer nonlinear program is a difficult optimization problem. Below, by considering a mixed-strategy reformulation of the incentive compatibility constraints for the agent, we can reformulate the optimal contract problem (5.6) as a mathematical program with equilibrium constraints (MPEC); see [33].

For $i = 1, \ldots, M$, let δ_i denote the probability that the agent will choose action a_i. Then, given the compensation schedule c, the agent chooses a mixed strategy profile $\delta^* = (\delta_1^*, \ldots, \delta_M^*) \in R^M$ such that

$$\delta^* \in \operatorname{argmax} \left\{ \sum_{k=1}^{M} \delta_k U(c, a_k) \;:\; e_M^{\mathrm{T}} \delta = 1,\, \delta \geq 0 \right\}. \tag{5.8}$$

Observe that the agent's mixed-strategy problem (5.8) is a linear program, and hence, its optimality conditions are necessary and sufficient.

The following lemma states the relationship between the optimal pure strategy a_i and the optimal mixed strategy δ^*. To ease the notation, we define

$$\begin{aligned} U(c) &= (U(c, a_1), \ldots, U(c, a_M)) \in R^M, \\ W(c) &= (W(c, a_1), \ldots, W(c, a_M)) \in R^M. \end{aligned} \tag{5.9}$$

Lemma 5.3. *Given a compensation schedule $\bar{c} \in R^N$, the agent's action $a_i \in \mathcal{A}$ is optimal for problem (5.3) iff there exists an optimal mixed strategy profile δ^* for problem (5.8) such that*

$$\begin{gathered} \delta_i^* > 0, \\ \sum_{k=1}^{M} \delta_k^* U(\bar{c}, a_k) = U(\bar{c}, a_i), \\ e_M^{\mathrm{T}} \delta^* = 1, \quad \delta^* \geq 0. \end{gathered}$$

Proof. If a_i is an optimal action of (5.3), then let $\delta^* = e_i$, the i-th column of the identity matrix of order M. It is easy to verify that all the conditions for δ^* are satisfied. Conversely, if a_i is not an optimal solution of (5.3), then there exists an action a_j such that $U(\bar{c}, a_j) > U(\bar{c}, a_i)$. Let $\tilde{\delta} = e_j$. Then $\tilde{\delta}^{\mathrm{T}} U(c) = U(\bar{c}, a_j) > U(\bar{c}, a_i) = \delta^{*\mathrm{T}} U(c)$. We have a contradiction. ∎

An observation following from Lemma 5.3 is stated below.

Lemma 5.4. *Given a compensation schedule $c \in R^N$, a mixed strategy profile δ*

is optimal for the linear program (5.8) iff

$$0 \le \delta \perp \left(\delta^{\mathrm{T}} U(c)\right) e_M - U(c) \ge 0, \qquad e_M^{\mathrm{T}} \delta = 1. \tag{5.10}$$

Proof. This follows from the optimality conditions and the strong duality theorem for the LP (5.8). ∎

Substituting the incentive compatibility constraint (5.5) by the system (5.10) and replacing $W(c,a)$ and $U(c,a)$ by $\delta^{\mathrm{T}} W(c)$ and $\delta^{\mathrm{T}} U(c)$, respectively, we derive an MPEC formulation of the principal's problem (5.6):

$$\begin{aligned} \underset{(c,\delta)}{\text{maximize}} \quad & \delta^{\mathrm{T}} W(c) \\ \text{subject to} \quad & \delta^{\mathrm{T}} U(c) \ge U^*, \\ & e_M^{\mathrm{T}} \delta = 1, \\ & 0 \le \delta \perp \left(\delta^{\mathrm{T}} U(c)\right) e_M - U(c) \ge 0. \end{aligned} \tag{5.11}$$

To illustrate the failure of constraint qualification at any feasible point of an MPEC, we consider the feasible region $\mathcal{F}_1 = \{(x,y) \in R^2 \mid x \ge 0,\ y \ge 0,\ xy = 0\}$. At the point $(\bar{x},\bar{y}) = (0,2)$, the first constraint $x \ge 0$ and the third constraint $xy = 0$ are binding. The gradients of the binding constraints at $(\bar{x},\bar{y})$ are $(1,0)$ and $(2,0)$, which are dependent. It is easy to verify that the gradient vectors of the binding constraints are indeed dependent at other feasible points.

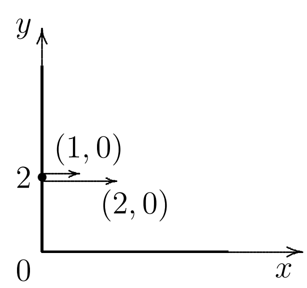

FIGURE 5.1 *The feasible region* $\mathcal{F}_1 = \{(x,y) \mid x \ge 0,\ y \ge 0,\ xy = 0\}$.

The following theorem states the relationship between the optimal solutions for the principal-agent problems (5.6) and the corresponding MPEC formulation (5.11).

Theorem 5.5. *If (c^*, δ^*) is an optimal solution for the MPEC (5.11), then (c^*, a_i^*), where $i \in \{j : \delta_j^* > 0\}$, is an optimal solution for the problem (5.6). Conversely, if (c^*, a_i^*) is an optimal solution for the problem (5.6), then (c^*, e_i) is an optimal solution for the MPEC (5.11).*

Proof. The statement follows directly from Lemma 5.4. ■

The MPEC (5.11) has $(N + M)$ variables, 1 linear constraint, 1 nonlinear constraint, and M complementarity constraints. Hence, the size of the problem grows linearly in the number of the outcomes and actions. As we will see in Section 5.4, this feature is the main advantage of using the MPEC approach rather than the LP lotteries approach.

5.3 Moral-Hazard Problems with Lotteries

In this section, we study moral-hazard problems with lotteries. In particular, we consider action lotteries, compensation lotteries, and a combination of both. For each case, we first give definitions for the associated lotteries and then derive the nonlinear programming or MPEC formulation.

5.3.1 The contract with action lotteries

Definition 5.6. A *contract with action lotteries* is a probability distribution over actions, $\pi(a)$, and a compensation schedule $c(a) = (c(q_1, a), \dots, c(q_N, a)) \in R^N$ for all $a \in \mathcal{A}$. The compensation schedule $c(a)$ is an agreement between the principal and the agent such that $c(q, a) \in \mathcal{C}$ is the payoff to the agent from the principal if outcome $q \in \mathcal{Q}$ is observed and the action $a \in \mathcal{A}$ is recommended by the principal.

In the definition of a contract with action lotteries, the compensation schedule $c(a)$ is contingent on both the outcome and the agent's action. Given this definition, one might raise the following question: if the principal can only observe the outcome, not the agent's action, is it reasonable to have the compensation schedule $c(a)$ contingent on the action chosen by the agent? After all, the principal does not know which action is implemented by the agent. One economic justification is as follows. Suppose that the principal and the agent sign a total of M contracts, each with different recommended action $a \in \mathcal{A}$ and compensation schedule $c(a)$ as a function of the recommended action, a. Then, the principal and the agent would go to an authority or a third party to conduct a lottery with probability distribution function $\pi(a)$ on which contract would be implemented on that day. If the i-th contract is drawn from the lottery, then the third party would inform both the principal and the agent that the recommended action for that day is a_i with the compensation schedule $c(a_i)$.

Arnott and Stiglitz [3] use *ex ante randomization* for action lotteries. This terminology refers to the situation that a random contract occurs before the recommended action is chosen. They demonstrate that the action lotteries will result in a welfare improvement if the principal's expected utility is nonconcave in the agent's expected utility. However, it is not clear what sufficient conditions would be needed for the statement in the assumption to be true.

5.3.2 An NLP formulation with star-shaped feasible region

When the principal proposes a contract with action lotteries, the contract has to satisfy the participation constraint and the incentive compatibility constraints. In particular, for a given contract $(\pi(a), c(a))_{a\in\mathcal{A}}$, the participation constraint requires the agent's expected utility to be at least U^*:

$$\sum_{a\in\mathcal{A}} \pi(a)U(c(a), a) \geq U^*. \tag{5.12}$$

For any recommended action a with $\pi(a) > 0$, it has to be incentive compatible with respect to the corresponding compensation schedule $c(a) \in R^N$. Hence, the incentive compatibility constraints are

$$\forall\, a \in \{\hat{a} : \pi(\hat{a}) > 0\} : \quad a = \operatorname{argmax}\{U(c(a), \tilde{a}) : \tilde{a} \in \mathcal{A}\}, \tag{5.13}$$

or equivalently,

$$\text{if } \pi(a) > 0, \text{ then } U(c(a), a) \geq U(c(a), a_i), \quad \text{for } i = 1, \ldots, M. \tag{5.14}$$

However, we do not know in advance whether $\pi(a)$ will be strictly positive at an optimal solution. One way to overcome this difficulty is to reformulate the solution-dependent constraints (5.14) as:

$$\forall\, a \in \mathcal{A}: \quad \pi(a)U(c(a), a) \geq \pi(a)U(c(a), a_i), \quad \text{for } i = 1, \ldots, M, \tag{5.15}$$

or in a compact presentation,

$$\pi(a)\left(U(c(a), a) - U(c(a), \tilde{a})\right) \geq 0, \quad \forall\, (a, \tilde{a}(\neq a)) \in \mathcal{A} \times \mathcal{A}. \tag{5.16}$$

Finally, since $\pi(\cdot)$ is a probability distribution function, we need

$$\begin{aligned} &\sum_{a \in \mathcal{A}} \pi(a) = 1, \\ &\pi(a) \geq 0, \quad \forall\, a \in \mathcal{A}. \end{aligned} \tag{5.17}$$

The principal chooses a contract with action lotteries that satisfies participation constraint (5.12), incentive compatibility constraints (5.16), and the probability measure constraint (5.17) to maximize his expected utility. An optimal contract with action lotteries $(\pi^*(a), c^*(a))_{a \in \mathcal{A}}$ is then a solution to the following nonlinear program:

$$\begin{aligned} \text{maximize} \quad & \sum_{a \in \mathcal{A}} \pi(a)W(c(a), a) \\ \text{subject to} \quad & \sum_{a \in \mathcal{A}} \pi(a)U(c(a), a) \geq U^*, \\ & \sum_{a \in \mathcal{A}} \pi(a) = 1, \\ \forall\, (a, \tilde{a}(\neq a)) \in \mathcal{A} \times \mathcal{A}: \quad & \pi(a)\left(U(c(a), a) - U(c(a), \tilde{a})\right) \geq 0, \\ & \pi(a) \geq 0, \quad \forall\, a \in \mathcal{A}. \end{aligned} \tag{5.18}$$

The nonlinear program (5.18) has $(N * M + M)$ variables and $(M * (M-1) + 2)$ constraints. In addition, its feasible region is highly nonconvex because of the last two sets of constraints in (5.18). As shown in Figure 5.2, the feasible region $\mathcal{F}_2 = \{(x,y) \,|\, xy \geq 0,\, x \geq 0\}$ is the union of the first quadrant and the y-axis. Furthermore, the standard nonlinear programming constraint qualification fails to hold at every point on the y-axis.

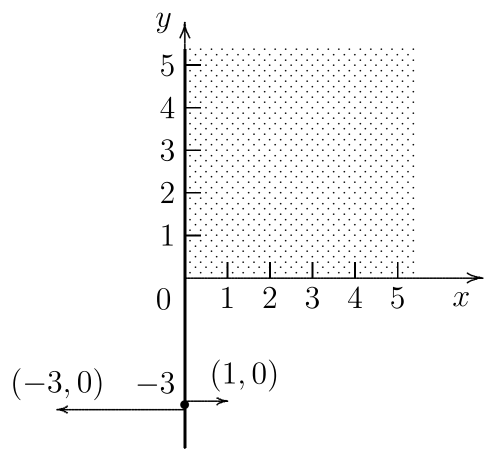

FIGURE 5.2 *The feasible region* $\mathcal{F}_2 = \{(x,y) \,|\, xy \geq 0,\, x \geq 0\}$.

5.3.3 MPEC formulations

Below, we introduce an MPEC formulation for the star-shaped problem (5.18). We first show that constraints of a star-shaped set $\mathcal{Z}_1 = \{z \in R^n \,|\, g_i(z) \geq 0,\, g_i(z)\, h_i(z) \geq 0,\, i = 1, \ldots, m\}$ can be rewritten as complementarity constraints if we introduce additional variables.

Proposition 5.7. *A point z is in $\mathcal{Z}_1 = \{z \in R^n \,|\, g_i(z) \geq 0,\, g_i(z)\, h_i(z) \geq 0,\, i = 1, \ldots, m\}$ iff there exists an s such that (z, s) is in $\mathcal{Z}_2 = \{(z,s) \in R^{n+m} \,|\, 0 \leq g(z) \perp s \geq 0,\, h(z) \geq -s\}$.*

Proof. Suppose that z is in $\mathcal{Z}_1$. If $g_i(z) > 0$, choose $s_i = 0$; if $g_i(z) = 0$, choose $s_i = -h_i(z)$. Then (z, s) is in $\mathcal{Z}_2$. Conversely, if (z, s) is in $\mathcal{Z}_2$, then $g_i(z)h_i(z) \geq g_i(z)(-s_i) = 0$ for all $i = 1, \ldots m$. Hence, the point z is in $\mathcal{Z}_1$. ■

Following Proposition 5.7, we introduce a variable $s(a,\tilde{a})$ for each pair $(a,\tilde{a}) \in \mathcal{A} \times \mathcal{A}$ for the incentive compatibility constraints in (5.18). We then obtain the following MPEC formulation with variables $(\pi(a), c(a), s(a,\tilde{a}))_{(a,\tilde{a})\in\mathcal{A}\times\mathcal{A}}$ for the optimal contract with action lottery problem:

$$\begin{aligned}
\text{maximize} \quad & \sum_{a\in\mathcal{A}} \pi(a) W(c(a),a) \\
\text{subject to} \quad & \sum_{a\in\mathcal{A}} \pi(a) U(c(a),a) \geq U^*, \\
& \sum_{a\in\mathcal{A}} \pi(a) = 1, \\
\forall\, (a,\tilde{a}(\neq a)) \in \mathcal{A}\times\mathcal{A}: \quad & U(c(a),a) - U(c(a),\tilde{a}) + s(a,\tilde{a}) \geq 0, \\
\forall\, (a,\tilde{a}(\neq a)) \in \mathcal{A}\times\mathcal{A}: \quad & 0 \leq \pi(a) \perp s(a,\tilde{a}) \geq 0.
\end{aligned} \tag{5.19}$$

Allowing the compensation schedules to be dependent on the agent's action will increase the principal's expected utility; see Theorem 5.8 below. The difference between the optimal objective value of the NLP (5.18) (or the MPEC(5.19)) and that of the MPEC (5.11) characterizes the principal's improved welfare from using an optimal contract with action lotteries.

Theorem 5.8. *The principal prefers an optimal contract with action lotteries to an optimal deterministic contract. His expected utility from choosing an optimal contract with action lotteries will be at least as good as that from choosing an optimal deterministic contract.*

Proof. This is clear. ∎

5.3.4 The contract with compensation lotteries

Definition 5.9. For any outcome $q \in \mathcal{Q}$, a *randomized compensation* $\tilde{c}(q)$ is a random variable on the set of compensations $\mathcal{C}$ with a probability measure $F(\cdot)$.

Remark If the set of compensations $\mathcal{C}$ is a closed interval $[\underline{c}, \bar{c}] \in R$, then the measure of $\tilde{c}(q)$ is a cumulative density function (cdf) $F : [\underline{c}, \bar{c}] \to [0, 1]$ with $F(\underline{c}) = 0$ and $F(\bar{c}) = 1$. In addition, $F(\cdot)$ is nondecreasing and right-continuous.

To simplify the analysis, we assume that every randomized compensation $\tilde{c}(q)$ has finite support.

Assumption 5.10 (Finite support for randomized compensation.). *For all $q \in \mathcal{Q}$, the randomized compensation $\tilde{c}(q)$ has finite support over an unknown set $\{c_1(q), c_2(q), \ldots, c_L(q)\}$ with a known L.*

An immediate consequence of Assumption 5.10 is that we can write $\tilde{c}(q) = c_i(q)$ with probability $p_i(q) > 0$ for all $i = 1, \ldots, L$ and $q \in \mathcal{Q}$. In addition, we have $\sum_{i=1}^{L} p_i(q) = 1$ for all $q \in \mathcal{Q}$. Notice that both $(c_i(q))_{i=1}^{L} \in R^L$ and $(p_i(q))_{i=1}^{L} \in R^L$ are endogenous variables and will be chosen by the principal.

Definition 5.11. A *compensation lottery* is a randomized compensation schedule $\tilde{c} = (\tilde{c}(q_1), \ldots, \tilde{c}(q_N)) \in R^N$, in which $\tilde{c}(q)$ is a randomized compensation satisfying Assumption 5.10 for all $q \in \mathcal{Q}$.

Definition 5.12. A *contract with compensation lotteries* consists of a recommended action a to the agent and a randomized compensation schedule $\tilde{c} = (\tilde{c}(q_1), \ldots, \tilde{c}(q_N)) \in R^N$.

Let $c^q = (c_i(q))_{i=1}^{L} \in R^L$ and $p^q = (p_i(q))_{i=1}^{L} \in R^L$. Given that the outcome q is observed by the principal, we let $\mathbf{w}(c^q, p^q)$ denote the principal's expected utility with respect to a randomized compensation $\tilde{c}(q)$, i.e.,

$$\mathbf{w}(c^q, p^q) = \mathbb{E}\, w(q - \tilde{c}(q)) = \sum_{i=1}^{L} p_i(q) w(q - c_i(q)).$$

With a randomized compensation schedule $\tilde{c}$ and a recommended action a, the principal's expected utility then becomes

$$\mathbb{E}\, W(\tilde{c}, a) = \sum_{q \in \mathcal{Q}} p(q \,|\, a) \left(\sum_{i=1}^{L} p_i(q) w(q - c_i(q)) \right) = \sum_{q \in \mathcal{Q}} p(q \,|\, a) \mathbf{w}(c^q, p^q). \quad (5.20)$$

Similarly, given a recommended action a, we let $\mathbf{u}(c^q, p^q, a)$ denote the agent's expected utility with respect to $\tilde{c}(q)$ for the observed outcome q:

$$\mathbf{u}(c^q, p^q, a) = \mathbb{E}\, u(\tilde{c}(q), a) = \sum_{i=1}^{L} p_i(q) u(c_i(q), a).$$

The agent's expected utility with a randomized compensation schedule $\tilde{c}$ and a recommended action a is

$$\mathbb{E}\, U(\tilde{c}, a) = \sum_{q \in \mathcal{Q}} p(q \mid a) \left(\sum_{i=1}^{L} p_i(q) u(c_i(q), a) \right) = \sum_{q \in \mathcal{Q}} p(q \mid a) \mathbf{u}(c^q, p^q, a). \quad (5.21)$$

To further simply to notation, we use $c_{\mathcal{Q}} = (c^q)_{q \in \mathcal{Q}}$ and $p_{\mathcal{Q}} = (p^q)_{q \in \mathcal{Q}}$ to denote the collection of variables c^q and p^q, respectively. We also let $\mathbf{W}(c_{\mathcal{Q}}, p_{\mathcal{Q}}, a)$ denote the principal's expected utility $\mathbb{E}\, W(\tilde{c}, a)$ as defined in (5.20), and similarly, $\mathbf{U}(c_{\mathcal{Q}}, p_{\mathcal{Q}}, a)$ for $\mathbb{E}\, U(\tilde{c}, a)$ as in (5.21).

An optimal contract with compensation lotteries $(c^*_{\mathcal{Q}}, p^*_{\mathcal{Q}}, a^*)$ is a solution to the following problem:

$$\begin{array}{ll} \text{maximize} & \mathbf{W}(c_{\mathcal{Q}}, p_{\mathcal{Q}}, a) \\ \text{subject to} & \mathbf{U}(c_{\mathcal{Q}}, p_{\mathcal{Q}}, a) \geq U^*, \\ & \mathbf{U}(c_{\mathcal{Q}}, p_{\mathcal{Q}}, a) \geq \mathbf{U}(c_{\mathcal{Q}}, p_{\mathcal{Q}}, a_i), \quad \forall\, i = 1, \ldots, M, \\ & a \in \mathcal{A} = \{a_1, \ldots, a_M\}. \end{array} \quad (5.22)$$

Define

$$\begin{aligned} \mathbf{W}(c_{\mathcal{Q}}, p_{\mathcal{Q}}) &= (\mathbf{W}(c_{\mathcal{Q}}, p_{\mathcal{Q}}, a_1), \ldots, \mathbf{W}(c_{\mathcal{Q}}, p_{\mathcal{Q}}, a_M)) \in R^M, \\ \mathbf{U}(c_{\mathcal{Q}}, p_{\mathcal{Q}}) &= (\mathbf{U}(c_{\mathcal{Q}}, p_{\mathcal{Q}}, a_1), \ldots, \mathbf{U}(c_{\mathcal{Q}}, p_{\mathcal{Q}}, a_M)) \in R^M. \end{aligned}$$

Following the derivation as in Section 2, we can reformulate the program for an optimal contract with compensation lotteries (5.22) as a mixed-integer nonlinear program with decision variables $(c_{\mathcal{Q}}, p_{\mathcal{Q}})$ and $y = (y_i)_{i=1}^{M}$:

$$\begin{aligned}
\text{maximize} \quad & \mathbf{W}(c_{\mathcal{Q}}, p_{\mathcal{Q}}, \sum_{i=1}^{M} a_i y_i) \\
\text{subject to} \quad & \mathbf{U}(c_{\mathcal{Q}}, p_{\mathcal{Q}}, \sum_{i=1}^{M} a_i y_i) \geq U^*, \\
& \mathbf{U}(c_{\mathcal{Q}}, p_{\mathcal{Q}}, \sum_{i=1}^{M} a_i y_i) \geq \mathbf{U}(c_{\mathcal{Q}}, p_{\mathcal{Q}}, a_j), \quad \forall\, j = 1, \ldots, M, \\
& e_M^{\mathrm{T}} y = 1, \\
& y_i \in \{0, 1\} \quad \forall\, i = 1, \ldots, M,
\end{aligned} \tag{5.23}$$

Similarly, the MPEC formulation with decision variables $(c_{\mathcal{Q}}, p_{\mathcal{Q}})$ and $\delta \in R^M$ is

$$\begin{aligned}
\text{maximize} \quad & \delta^{\mathrm{T}} \mathbf{W}(c_{\mathcal{Q}}, p_{\mathcal{Q}}) \\
\text{subject to} \quad & \delta^{\mathrm{T}} \mathbf{U}(c_{\mathcal{Q}}, p_{\mathcal{Q}}) \geq U^*, \\
& e_M^{\mathrm{T}} \delta = 1, \\
& 0 \leq \delta \perp \left(\delta^{\mathrm{T}} \mathbf{U}(c_{\mathcal{Q}}, p_{\mathcal{Q}})\right) e_M - \mathbf{U}(c_{\mathcal{Q}}, p_{\mathcal{Q}}) \geq 0.
\end{aligned} \tag{5.24}$$

Arnott and Stiglitz [3] call the compensation lotteries *ex post randomization*; this refers to the situation where the random compensation occurs after the recommended action is chosen or implemented. They show that if the agent is risk averse and his utility function is separable, and if the principal is risk neutral, then the compensation lotteries are not desirable.

5.3.5 The contract with action and compensation lotteries

Definition 5.13. A *contract with action and compensation lotteries* is a probability distribution over actions, $\pi(a)$, and a randomized compensation schedule $\tilde{c}(a) = (\tilde{c}(q_1, a), \ldots, \tilde{c}(q_N, a)) \in R^N$ for every $a \in \mathcal{A}$. The randomized compensation schedule $c(a)$ is an agreement between the principal and the agent such that $\tilde{c}(q, a) \in \mathcal{C}$ is a randomized compensation to the agent from the principal if outcome $q \in \mathcal{Q}$ is observed and the action $a \in \mathcal{A}$ is recommended by the principal.

Assumption 5.14. *For every action $a \in \mathcal{A}$, the randomized compensation schedule $\tilde{c}(q,a)$ satisfies the finite support assumption (Assumption 5.10) for all $q \in \mathcal{Q}$.*

With Assumption 5.14, the notation $c^q(a), p^q(a), c_{\mathcal{Q}}(a), p_{\mathcal{Q}}(a)$ is analogous to what we have defined in Section 5.3.1 and 5.3.2. Without repeating the same derivation process described earlier, we give the star-shaped formulation with variables $(\pi(a), c_{\mathcal{Q}}(a), p_{\mathcal{Q}}(a))_{a\in\mathcal{A}}$ for the optimal contract with action and compensation lotteries problem:

$$\begin{array}{ll}
\text{maximize} & \displaystyle\sum_{a\in\mathcal{A}} \pi(a)\mathbf{W}(c_{\mathcal{Q}}(a), p_{\mathcal{Q}}(a), a) \\
\text{subject to} & \displaystyle\sum_{a\in\mathcal{A}} \pi(a)\mathbf{U}(c_{\mathcal{Q}}(a), p_{\mathcal{Q}}(a), a) \geq U^*, \\
& \displaystyle\sum_{a\in\mathcal{A}} \pi(a) = 1, \\
\forall\, (a,\tilde{a}) \in \mathcal{A}\times\mathcal{A}: & \pi(a)\left(\mathbf{U}(c_{\mathcal{Q}}(a), p_{\mathcal{Q}}(a), a) - \mathbf{U}(c_{\mathcal{Q}}(a), p_{\mathcal{Q}}(a), \tilde{a})\right) \geq 0, \\
& \pi(a) \geq 0.
\end{array} \tag{5.25}$$

Following the derivation in Section 5.3.3, an equivalent MPEC formulation is with variables $(\pi(a), c_{\mathcal{Q}}(a), p_{\mathcal{Q}}(a), s(a,\tilde{a}))_{(a,\tilde{a})\in\mathcal{A}\times\mathcal{A}}$:

$$\begin{array}{ll}
\text{maximize} & \displaystyle\sum_{a\in\mathcal{A}} \pi(a)\mathbf{W}(c_{\mathcal{Q}}(a), p_{\mathcal{Q}}(a), a) \\
\text{subject to} & \displaystyle\sum_{a\in\mathcal{A}} \pi(a)\mathbf{U}(c_{\mathcal{Q}}(a), p_{\mathcal{Q}}(a), a) \geq U^*, \\
& \displaystyle\sum_{a\in\mathcal{A}} \pi(a) = 1, \\
\forall\, (a,\tilde{a}) \in \mathcal{A}\times\mathcal{A}: & \mathbf{U}(c_{\mathcal{Q}}(a), p_{\mathcal{Q}}(a), a) - \mathbf{U}(c_{\mathcal{Q}}(a), p_{\mathcal{Q}}(a), \tilde{a}) \geq -s(a,\tilde{a}), \\
\forall\, (a,\tilde{a}) \in \mathcal{A}\times\mathcal{A}: & 0 \leq \pi(a) \perp s(a,\tilde{a}) \geq 0.
\end{array} \tag{5.26}$$

5.3.6 Linear programming approximation

Townsend [64, 65] was among the first to use linear programming techniques to solve static incentive constrained problems. Prescott [50, 51] further apply linear programming specifically to solve moral-hazard problems. A solution obtained by the linear programming approach is an approximation to a solution to the MPEC (5.26). Instead of treating $c_{\mathcal{Q}}(a)$ as unknown variables, one can construct a grid Ξ with elements ξ to approximate the set $\mathcal{C}$ of compensations. By introducing probability measures associated with the action lotteries on $\mathcal{A}$ and compensation lotteries on Ξ, one can then approximate a solution to the moral-hazard problem with lotteries (5.26) by solving a linear program. More specifically, the principal chooses probability distributions $\pi(a)$, and $\pi(\xi|\, q, a)$ over the set of actions $\mathcal{A}$, the set of outcomes $\mathcal{Q}$, and the compensation grid Ξ. One then can reformulate the resulting nonlinear program as a linear program with decision variables $\pi = (\pi(\xi, q, a))_{\xi\in\Xi, q\in\mathcal{Q}, a\in\mathcal{A}}$:

$$\begin{aligned}
\text{maximize}_{(\pi)} \quad & \sum_{\xi,q,a} w(q-\xi)\pi(\xi,q,a) \\
\text{subject to} \quad & \sum_{\xi,q,a} u(\xi,a)\pi(\xi,q,a) \geq U^*, \\
\forall\,(a,\tilde{a}) \in \mathcal{A}\times\mathcal{A}: \quad & \sum_{\xi,q} u(\xi,a)\pi(\xi,q,a) \geq \sum_{\xi,q} u(\xi,\tilde{a})\frac{p(q|\tilde{a})}{p(q|a)}\pi(\xi,q,a) \\
\forall\,(\tilde{q},\tilde{a}) \in \mathcal{Q}\times\mathcal{A}: \quad & \sum_{\xi} \pi(\xi,\tilde{q},\tilde{a}) = p(\tilde{q}|\tilde{a})\sum_{\xi,q}\pi(\xi,q,\tilde{a}), \\
& \sum_{\xi,q,a}\pi(\xi,q,a) = 1, \\
& \pi(\xi,q,a) \geq 0 \quad \forall\,(\xi,q,a)\in\Xi\times\mathcal{Q}\times\mathcal{A}.
\end{aligned} \tag{5.27}$$

Note that the above linear program has $(|\Xi| * N * M)$ variables and $(M * (N + M - 1) + 2)$ constraints. The size of the linear program will grow enormously when one chooses a fine grid. For example, if there are 50 actions, 40 outputs, and 500 compensations, then the linear program has one million variables and 4452 constraints. It will become computationally intractable because of the limitation on computer memory, if not the time required. On the other hand, a solution of the LP obtained from a coarse grid will not be satisfactory if an accurate solution is needed. Prescott [51] points out that the constraint matrix of the linear program (5.27) has block angular structure. As a consequence, one can apply Dantzig-Wolfe decomposition to the linear program (5.27) to reduce the computer memory and computational time. Recall that the MPEC (5.11) for the optimal contract problem has only $(N+M)$ variables and M complementarity constraints with one linear constraint and one nonlinear constraint. Even with the use of the Dantzig-Wolfe decomposition algorithm to solve LP (5.27), choosing the "right" grid is still an issue. With the advances in both theory and numerical methods for solving MPECs in the last decade, we believe that the MPEC approach has greater advantages in solving a much smaller problem and in obtaining a more accurate solution.

The error from discretizing set of compensations $\mathcal{C}$ is characterized by the difference between the optimal objective value of LP (5.27) and that of MPEC (5.26).

Theorem 5.15. *The optimal objective value of MPEC (5.26) is at least as good as that of LP (5.27).*

Proof. It is sufficient to show that given a feasible point of LP (5.27), one can construct a feasible point for MPEC (5.26) with objective value equal to that of the LP (5.27).

Let $\pi = (\pi(\xi, q, a))_{\xi \in \Xi, q \in \mathcal{Q}, a \in \mathcal{A}}$ be a given feasible point of LP (5.27). Let

$$\pi(a) = \sum_{\xi \in \Xi, q \in \mathcal{Q}} \pi(\xi, q, a).$$

For every $q \in \mathcal{Q}$ and $a \in \mathcal{A}$, we define

$$\begin{aligned}
\mathcal{S}(q,a) &:= \{\xi \in \Xi \,|\, \pi(\xi, q, a) > 0\}, \\
L^q(a) &:= |\mathcal{S}(q,a)|, \\
c^q(a) &:= (\xi)_{\xi \in \mathcal{S}(q,a)} \\
p^q(a) &:= (\pi(\xi, q, a))_{\xi \in \mathcal{S}(q,a)}
\end{aligned}$$

It is easy to check that $\pi(a)$, $c^q(a)$ and $p^a(a)$ is a feasible for MPEC (5.26). Furthermore, its objective value is the same as that of π for the LP (5.27). ■

5.4 A Hybrid Approach toward Global Solution

One reason that nonconvex programs are not popular among economists is the issue of the need for global solutions. While local search algorithms for solving nonconvex programs have fast convergence properties near a solution, they are designed to find a local solution. Algorithms for solving MPECs are no exception. One heuristic in practice is to solve the same problem with several different starting points. It then becomes a trade-off between the computation time and the quality of the "best" solution found.

Linear programming does not suffer from the global solution issue. However, to obtain an accurate solution to a moral-hazard problem via the linear programming approach, one needs to use a very fine compensation grid. This often leads to large-scale linear programs with millions of variables and tens or hundreds of thousands of constraints, which might require excessive computer memory or time.

Certainly, there is a need to develop a global optimization method with fast local convergence for MPECs. Below, we propose a hybrid approach combining both MPECs and linear programming approaches to find a global solution (or at least better than the LP solution) of an optimal contract problem. The motivation for this hybrid method comes from the observation that the optimal objective value of the LP approach from a coarse grid could provide a lower bound on the optimal objective value of the MPEC as well as a good guess on the final recommended action a^*. We can then use this information to exclude some undesired local minimizers and to provide a good starting point when we solve the MPEC (5.11). This heuristic procedure toward a global solution of the MPEC (5.11) leads to the following algorithm.

A hybrid method for the optimal contract problem as MPEC (5.11)

Step 0: Construct a coarse grid Ξ over the compensation interval.

Step 1: Solve the LP (5.27) for the given grid Ξ.

Step 2:
$$\begin{cases} (2.1): & \text{Compute } p(a) = \sum_{\xi \in \Xi} \sum_{q \in \mathcal{Q}} \pi(\xi, q, a), \quad \forall\, a \in \mathcal{A}; \\ (2.2): & \text{Compute } \mathbb{E}[\xi(q)] = \sum_{\xi \in \Xi} \xi \pi(\xi, q, a), \quad \forall\, q \in \mathcal{Q}; \\ (2.3): & \text{Set initial point } c^0 = (\mathbb{E}[\xi(q)])_{q \in \mathcal{Q}} \text{ and } \delta^0 = (p(a))_{a \in \mathcal{A}}; \\ (2.4): & \text{Solve the MPEC (5.11) with starting point } (c^0, \delta^0). \end{cases}$$

Step 3: Refine the grid and repeat Step 1 and Step 2.

Remark If the starting point from an LP solution is close to the optimal solution of the MPEC (5.11), then the sequence of iterates generated by an SQP algorithm converges Q-quadratically to the optimal solution. See Proposition 5.2 in Fletcher et al. [14].

One can also develop similar procedures to find global solutions for optimal contract problems with action and/or compensation lotteries. However, the MPECs for contracts with lotteries are much more numerically challenging problems than the MPEC (5.11) for deterministic contracts.

5.5 An Example and Numerical Results

To illustrate the use of the mixed-integer nonlinear program (5.7), the MPEC (5.11) and the hybrid approaches, and to understand the effect of discretizing the set of compensations $\mathcal{C}$, we only consider problems of deterministic contracts without lotteries. We consider a two-outcome example in Karaivanov [29]. Before starting the computational work, we summarize in Table 5.1 the problem characteristics of various approaches to computing the optimal deterministic contracts.

TABLE 5.1
Problem characteristics of various approaches.

	MINLP (5.7)	MPEC (5.11)	LP (5.27)
Regular Variables	N	$N+M$	$\lvert\Xi\rvert * N * M$
Binary Variables	M	–	–
Constraints	$M+2$	2	$M*(N+M-1)+2$
Complementarity Const.	–	M	–

Example 1: No Action and Compensation Lotteries

Assume the principal is risk neutral with utility $w(q-c(q)) = q - c(q)$, and the agent is risk averse with utility

$$u(c(q), a) = \frac{c^{1-\gamma}}{1-\gamma} + \kappa \frac{(1-a)^{1-\delta}}{1-\delta}.$$

Suppose there are only two possible outcomes, e.g., a coin-flip. If the desirable outcome (high sale quantities or high production quantities) happens, then the principal receives $q_H = \$3$; otherwise, he receives $q_L = \$1$. For simplicity, we assume that the set of actions $\mathcal{A}$ consists of M equally-spaced effort levels within the closed interval $[0.01, 0.99]$. The production technology for the high outcome is described by $p(q = q_H|a) = a^\alpha$ with $0 < \alpha < 1$. Note that since 0 and 1 are excluded from the action set $\mathcal{A}$, the full-support assumption on production technology is satisfied.

The parameter values for the particular instance we solve are given in Table 5.2.

TABLE 5.2
The value of parameters used in Example 1.

γ	κ	δ	α	U^*	M
0.5	1	0.5	0.7	1	10

We solve this problem first as a mixed-integer nonlinear program (5.7) and then as an MPEC (5.11). For the LP lotteries approach, we start with 20 grid points in the compensation grid (we evenly discretize the compensation set $\mathcal{C}$ into 19 segments) and then increase the size of the compensation grid to $50, 100, 200, \ldots, 5000$.

We submitted the corresponding AMPL programs to the NEOS server [42]. The mixed-integer nonlinear programs were solved using the MINLP solver [12] on the computer host `newton.mcs.anl.gov`. To obtain fair comparisons between the LP, MPEC, and hybrid approaches, we chose SNOPT [17] to solve the associated mathematical programs. The AMPL programs were solved on the computer host `tate.iems.northwestern.edu`.

Table 5.3 gives the solutions returned by the MINLP solver to the mixed-integer nonlinear program (5.7). We use $y = 0$ and $y = e_M$ as starting points. In both cases, the MINLP solver returns a solution very quickly. However, it is not guaranteed to find a global solution.

TABLE 5.3
Solutions of the MINLP approach.

Starting Point	Regular Variables	Binary Variables	Constraints	Solve Time (in sec.)	Objective Value
$y = 0$	2	10	12	0.01	1.864854251
$y = e_M$	2	10	12	0.00	1.877265189

For solving the MPEC (5.11), we try two different starting points to illustrate the possibility of finding only a local solution. The MPEC solutions are given in Table 5.4 below.

TABLE 5.4
Solutions of the MPEC approach with two different starting points.

(22 variables and 10 complementarity constraints)

Starting Point	Read Time (in sec.)	Solve Time (in sec.)	# of Major Iterations	Objective Value
$\delta = 0$	0	0.07	45	1.079621424
$\delta = e_M$	0	0.18	126	1.421561553

The solutions for the LP lottery approach with different compensation grids are given in Table 5.5. Notice that the solve time increases faster than the size of the grid when $|\Xi|$ is of order 10^5 and higher, while the number of major iterations only increases about 3 times when we increase the grid size 250 times (from $|\Xi| = 20$ to $|\Xi| = 5000$).

Table 5.5
Solutions of the LP approach with 8 different compensation grids.

$\lvert\Xi\rvert$	# of Variables	Read Time (in sec.)	Solve Time (in sec.)	# of Iterations	Objective Value
20	400	0.01	0.03	31	1.876085819
50	1000	0.02	0.06	46	1.877252488
100	2000	0.04	0.15	53	1.877252488
200	4000	0.08	0.31	62	1.877254211
500	10000	0.21	0.73	68	1.877263962
1000	20000	0.40	2.14	81	1.877262184
2000	40000	0.83	3.53	71	1.877260460
5000	100000	2.19	11.87	101	1.877262793

Finally, for the hybrid approach, we first use the LP solution from a compensation grid with $|\Xi| = 20$ to construct a starting point for the MPEC (5.11). As one can see in Table 5.6, with a good starting point, it takes SNOPT only 0.01 seconds to find a solution to the example formulated as the MPEC (5.11). Furthermore, the optimal objective value is higher than that of the LP solution from a fine compensation grid with $|\Xi| = 5000$.

Table 5.6
Solutions of the hybrid approach for Example 1.

LP $\lvert\Xi\rvert$	Read Time (in sec.)	Solve Time (in sec.)	# of Iterations	Objective Value
20	0.01	0.03	31	1.876085819
MPEC Starting Point	Read Time (in sec.)	Solve Time (in sec.)	# of Major Iterations	Objective Value
$\delta_6 = 1, \delta_{i(\neq 6)} = 0$	0.02	0.01	13	1.877265298

5.6 Conclusions and Future Work

The purpose of this chapter is to introduce the MPEC approach and apply it to moral-hazard problems. We have presented MPEC formulations for optimal deterministic contract problems and optimal contract problems with action and/or compensation lotteries. We also formulated the former problem as a mixed-integer nonlinear program. To obtain a global solution, we have proposed a hybrid procedure that combines the LP lottery and the MPEC approaches. In this procedure, the LP solution from a coarse compensation grid provides a good starting point for the MPEC. We can then apply specialized MPEC algorithms with fast local convergence rate to obtain a solution. In a numerical example, we have demonstrated that the hybrid method is more efficient than using only the LP lottery approach, which requires the solution of a sequence of large-scale linear programs. Although we cannot prove that the hybrid approach will guarantee to find a global solution, it always finds one better than the solution from the LP lottery approach. We plan to test the numerical performance of the hybrid procedure on other examples such as the bank regulation example in [50] and the two-dimensional action choice example in [51].

One can extend the MPEC approach to single-principal multiple-agent problems without any difficulty. For multiple-principal multiple-agent models [41], it can be formulated as an equilibrium problem with equilibrium constraint. We will investigate these two topics in our future research.

Another important topic we plan to explore is the dynamic moral-hazard problem; see Phelan and Townsend [48]. In the literature, dynamic programming is applied to solve this model. We believe that there is an equivalent nonlinear programming formulation. Analogous to the hybrid procedure proposed in Section 4, an efficient method to solve this dynamic model is to combine dynamic programming and nonlinear programming.

Bibliography

[1] Allaz, B., Vila, J.-L. (1993): Cournot competition, forward markets and efficiency, *Journal of Economic Theory*, **59**, 1–16.

[2] Anitescu, M. (2000): On solving mathematical programs with complementarity constraints as nonlinear programs, Preprint ANL/MCS-P864-1200, Mathematics and Computer Science Division, Argonne National Laboratory, Argonne, IL.

[3] Arnott, A., Stiglitz, J.E. (1998): Randomization with asymmetric information, *The Rand Journal of Economics*, **19**, 344-362.

[4] Berry, C.A., Hobbs, B.F., Meroney, W.A., O'Neill, R.P., Stewart Jr., W.R. (1999): Understanding how market power can arise in network competition: A game theoretic approach, *Utility Policy*, **8**, 139–158.

[5] Cardell, J.B., Hitt, C.C., Hogan, W.W. (1997): Market power and strategic interaction in electricity networks, *Resource and Energy Economics*, **19**, 109–137.

[6] Chen, Y., Florian, M. (1995): The nonlinear bilevel programming problem: Formulations, regularity and optimality conditions, *Optimization*, **32**, 193–209.

[7] Cottle, R.W., Pang, J.-S., Stone, R.E. (1992): *The Linear Complementarity Problem*, Academic Press, Boston.

[8] Dafemos, S. (1983): An iterative scheme for variational inequalities, *Mathematical Programming*, **26**, 40–47.

[9] DeMiguel, A.V., Friedlander, M.P., Nogales, F.J., Scholtes, S. (2005): A two-sided relaxation scheme for mathematical programs with equilibrium constraints, *SIAM Journal on Optimization* (forthcoming).

[10] Ehrenmann, A. (2004): Manifolds of multi-leader Cournot equilibria, *Operations Research Letters*, **32**, 121–125.

[11] Facchinei, F., Pang, J.-S. (2003): *Finite-Dimensional Variational Inequalities and Complementarity Problems*, Springer Verlag, New York.

[12] Fletcher, R., Leyffer, S. (1999): User manual for MINLP_BB, Department of Mathematics, University of Dundee, UK.

[13] Fletcher, R., Leyffer, S. (2004): Solving mathematical programs with complementarity constraints as nonlinear programs, *Optimization Methods and Software*, **19**, 15–40.

[14] Fletcher, R., Leyffer, S., Ralph, D., Scholtes, S. (2002): Local convergence of SQP methods for mathematical programs with equilibrium constraints. Numerical Analysis Report NA/209, Department of Mathematics, University of Dundee, UK.

[15] Flegel, M.L., Kanzow, C. (2002): Optimality conditions for mathematical programs with equilibrium constraints: Fritz John and Abadie-type approaches, Working Paper, Institute of Applied Mathematics and Statistics, University of Würzburg, Germany.

[16] Fukushima, M., Pang, J.-S. (1999): Convergence of a smoothing continuation method for mathematical programs with equilibrium constraints, in (Théra, M., Tichatschke, R., ed.) *Ill-posed Variational Problems and Regularization Techniques*, 99–110, Springer-Verlag, New York.

[17] Gill, P.E., Murray, W., Saunders, M.A. (2002): SNOPT: An SQP algorithm for large-scale constrained optimization, *SIAM Journal on Optimization*, **12**, 979–1006.

[18] Grossman, S.J., Hart, O.D (1983): An analysis of the principal-agent problem, *Econometrica*, **51**, 7–46.

[19] Hart, O.D, Holmström, B. (1987): The theory of contracts, in (Bewley, T., ed.) *Advances in Economic Theory: Fifth World Congress*, 7–46, Cambridge University Press, Cambridge, UK.

[20] Harker, P.T. (1984): A variational inequality approach for the determination of oligopolistic market equilibrium, *Mathematical Programming*, **30**, 105–111.

[21] Harker, P.T. (1991): Generalized Nash games and quasi-variational inequalities, *European Journal of Operations Research*, **54**, 81–94.

[22] Hobbs, B.F., Metzler, C.B., Pang, J.-S. (2000): Strategic gaming analysis for electric power networks: An MPEC approach, *IEEE Transactions on Power Systems*, **15**, 638–645.

[23] Holmström, K. (1999): TOMLAB Optimization, `www.tomlab.biz`.

[24] Holmström, K. (1999): The Tomlab Optimization Environment in Matlab, *Advanced Modeling and Optimization*, **1**, 47–69.

[25] Hu, X. (2002): *Mathematical Programs with Complementarity Constraints and Game Theory Models in Electricity Markets,* Ph.D. Thesis, Department of Mathematics and Statistics, University of Melbourne.

[26] Jewitt, I. (1988): Justifying the first-order approach to principal-agent problems, *Econometrica*, **56**, 1177–1190.

[27] Jiang, H., Ralph, D. (1999): QPECgen, a MATLAB generator for mathematical programs with quadratic objectives and affine variational inequality constraints, *Computational Optimization and Applications*, **13**, 25–49.

[28] Judd, K.L. (1998): *Numerical Methods in Economics*, The MIT Press, Cambridge, MA.

[29] Karaivanov, A.K. (2001): Computing moral-hazard programs with lotteries using Matlab, Working paper, Department of Economics, University of Chicago.

[30] Leyffer, S. (2003): Mathematical programs with equilibrium constraints, Preprint ANL/MCS-P1026-0203, Mathematics and Computer Science Division, Argonne National Laboratory, Argonne, IL.

[31] Leyffer, S. (2003): Complementarity constraints as nonlinear equations: theory and numerical experiences, Preprint ANL/MCS-P1054-0603, Mathematics and Computer Science Division, Argonne National Laboratory, Argonne, IL.

[32] Leyffer, S., Lopez-Calva, G., Nocedal, J. (2005): Interior methods for mathematical programs with complementarity constraints, *SIAM Journal on Optimization* (forthcoming).

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

[34] Mangasarian, O.L. (1969): *Nonlinear Programming*, McGraw-Hill, New York.

[35] Mirrlees, J.A. (1975): The theory of moral-hazard and unobservable behavior: part I, Mimeo, Nuffield College, Oxford University, Oxford, UK.

[36] Mirrlees, J.A. (1999): The theory of moral-hazard and unobservable behavior: part I, *Review of Economic Studies*, **66**, 3–22.

[37] Mordukhovich, B.S. (2003): Subdifferential and superdifferential optimality conditions in nonsmooth minimization, Research Report, Department of Mathematics, Wayne State University, Detroit, MI.

[38] Mordukhovich, B.S. (2003): Optimization and equilibrium problems with equilibrium constraints, Research Report, Department of Mathematics, Wayne State University, Detroit, MI.

[39] Mordukhovich, B.S. (2004): Equilibrium problems with equilibrium constraints via multiobjective optimization, *Optimization Methods and Software*, **19**, 479–492.

[40] Murphy, F.H., Sherali, H.D., Soyster, A.L. (1982): A mathematical programming approah for determining oligopolistic market equilibrium, *Mathematical Programming*, **24**, 92–106.

[41] Myerson, R.B. (1982): Optimal coordination mechanisms in generalized principal-agent problems, *Journal of Mathematical Economics*, **10**, 67–81.

[42] The NEOS server for optimization, `www-neos.mcs.anl.gov/neos/`.

[43] Outrata, J., Kocvara, M., Zowe, J. (1998): *Nonsmooth Approach to Optimization Problems with Equilibrium Constraints: Theory, Applications, and Numerical Results*, Kluwer Academic Publishers, Dordrecht, The Netherlands.

[44] Pal, D. (1991): Cournot duopoly with two production periods and cost differentials, *Journal of Economic Theory*, **55**, 441–448.

[45] Pang, J.-S., Chan, D. (1982): Iterative methods for variational and complementarity problems, *Mathematical Programming*, **24**, 284–313.

[46] Pang, J.-S., Fukushima, M. (1999): Complementarity constraint qualifications and simplified B-stationarity conditions for mathematical programs with equilibrium constraints, *Computational Optimization and Applications*, **13**, 111–136.

[47] Pang, J.-S., Fukushima, M. (2005): Quasi-variational inequalities, generalized Nash equilibrium, and multi-leader-follower games, *Computational Management Science*, **2**, 21–56.

[48] Phelan, C., Townsend, R.M. (1991): Computing multi-period, information-constrained optima, *The Review of Economic Studies*, **58**, 853–881.

[49] Pieper, H. (2001): *Algorithms for Mathematical Programs with Equilibrium Constraints with Applications to Deregulated Electricity Markets,* Ph.D. Thesis, Department of Management Science and Engineering, Stanford University, Stanford, CA.

[50] Prescott, E.S. (1999): A primer on moral-hazard models, *Federal Reserve Bank of Richmond Quarterly Review*, **85**, 47–77.

[51] Prescott, E.S. (2004): Computing solutions to moral-hazard programs using the Dantzig-Wolfe decomposition algorithm, *Journal of Economic Dynamics and Control*, **28**, 777–800.

[52] Ralph, D., Wright, S.J. (2004): Some properties of regularization and penalization schemes for MPECs, *Optimization Methods and Software*, **19**, 527–556.

[53] Raghunathan, A., Biegler, L.T. (2002): Barrier methods for mathematical programs with complementarity constraints (MPCCs), Technical Report, Department of Chemical Engineering, Carnegie Mellon University, Pittsburgh, PA.

[54] Rogerson, W.P. (1985): The first-order approach to principal-agent problems, *Econometrica*, **53**, 1357–1368.

[55] Sahinidis, N.V., Tawarmalani, M. (2004): BARON 7.2: Global Optimization of Mixed-Integer Nonlinear Programs, *User's Manual*, Available at `http://www.gams.com/dd/docs/solvers/baron.pdf`.

[56] Saloner, G. (1986): The role of obsolescence and inventory costs in providing commitment, *International Journal of Industrial Organization*, **4**, 333–345.

[57] Saloner, G. (1987): Cournot duopoly with two production periods, *Journal of Economic Theory*, **42**, 183–187.

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

[59] Sherali, H.D. (1984): A multiple leader Stackelberg model and analysis, *Operations Research*, **32**, 390-404.

[60] Sherali, H.D., Leleno, J.M. (1988): A mathematical programming approach to a Nash-Cournot equilibrium analysis for a two-stage network of oligopolies, *Operations Research*, **36**, 682–702.

[61] Sherali, H.D., Soyster, A.L., Murphy, F.H. (1983): Stackelberg-Nash-Cournot equilibria: characterization and computations, *Operations Research*, **31**, 253–276.

[62] Scholtes, S. (2001): Convergence properties of a regularization scheme for mathematical programs with complementarity constraints, *SIAM Journal on Optimization*, **11**, 918–936.

[63] Scholtes, S., Stöhr, M. (2001): How stringent is the linear independence assumption for mathematical programs with complementarity constraints?, *Mathematics of Operations Research*, **26**, 851–863.

[64] Townsend, R.M. (1987): Economic organization with limited communication, *American Economic Review*, **77**, 954–971.

[65] Townsend, R.M. (1988): Information constrained insurance: the revelation principle extended, *Journal of Monetary Economics*, **21**, 411–450.

[66] Wolf, D.D., Smeers, Y. (1997): A stochastic version of a Stackelberg-Nash-Cournot Equilibrium model, *Management Science*, **43**, 190–197.