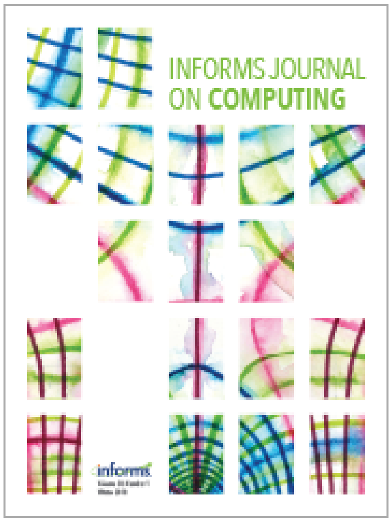
An Interior-Point Differentiable Path-Following Method to Compute Stationary Equilibria in Stochastic Games



INFORMS Journal on Computing

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

An Interior-Point Differentiable Path-Following Method to Compute Stationary Equilibria in Stochastic Games

Chuangyin Dang, P. Jean-Jacques Herings, Peixuan Li

informs
http://pubsonline.informs.org/journal/ijoc
INFORMS JOURNAL ON COMPUTING
Vol. 34, No. 3, May–June 2022, pp. 1403–1418
ISSN 1091-9856 (print), ISSN 1526-5528 (online)

An Interior-Point Differentiable Path-Following Method to Compute Stationary Equilibria in Stochastic Games

Chuangyin Dang,[a] P. Jean-Jacques Herings,[b] Peixuan Li[a,*]

[a] Department of Systems Engineering and Engineering Management, City University of Hong Kong, Kowloon, Hong Kong; [b] Department of Economics, Maastricht University, 6200 MD Maastricht, Netherlands
*Corresponding author

Contact: mecdang@cityu.edu.hk, https://orcid.org/0000-0003-4731-4616 (CD); p.herings@maastrichtuniversity.nl, https://orcid.org/0000-0002-1100-8601 (PJ-JH); peixuanli3-c@my.cityu.edu.hk, https://orcid.org/0000-0002-3792-7059 (PL)

Received: October 27, 2020
Revised: April 2, 2021; August 2, 2021; August 24, 2021; September 2, 2021; September 17, 2021
Accepted: September 28, 2021
Published Online in Articles in Advance: January 25, 2022

https://doi.org/10.1287/ijoc.2021.1139

Abstract. The subgame perfect equilibrium in stationary strategies (SSPE) is the most important solution concept in applications of stochastic games, making it imperative to develop efficient methods to compute an SSPE. For this purpose, this paper develops an interior-point differentiable path-following method (IPM), which establishes a connection between an artificial logarithmic barrier game and the stochastic game of interest by adding a homotopy variable. IPM brings several advantages over the existing methods for stochastic games. On the one hand, IPM provides a bridge between differentiable path-following methods and interior-point methods and remedies several issues of an existing homotopy method called the stochastic linear tracing procedure (SLTP). First, the starting stationary strategy profile can be arbitrarily chosen. Second, IPM does not need switching between different systems of equations. Third, the use of a perturbation term makes IPM applicable to all stochastic games rather than generic games only. Moreover, a well-chosen transformation of variables reduces the number of equations and variables by roughly one half. Numerical results show that the proposed method is more than three times as efficient as SLTP. On the other hand, the stochastic game can be reformulated as a mixed complementarity problem and solved by the PATH solver. We employ the proposed IPM and the PATH solver to compute SSPEs. Numerical results evince that for some stochastic games the PATH solver may fail to find an SSPE, whereas IPM is successful in doing so for all stochastic games, which confirms the reliability and stability of the proposed method.
Summary of Contribution: This paper incorporates the interior-point methods into a differentiable path-following method for computing stationary equilibria for stochastic games. This novel method brings excellent computational advantages and remedies several issues with the existing methods for stochastic games. We prove the global convergence of the proposed method and employ this method to solve numerous randomly generated stochastic games with different scales. Numerical results further confirm the high efficiency, stability, and universality of this method for stochastic games.

History: Accepted by Antonio Frangioni, Area Editor for Design & Analysis of Algorithms—Continuous.
Funding: This work was partially supported by a grant from Research Grants Council of the Hong Kong Special Administrative Region Government [GRF: CityU 11304620].
Supplemental Material: The software that supports the findings of this study is available within the paper and its Supplementary Information [https://pubsonline.informs.org/doi/suppl/10.1287/ijoc.2021.1139] or is available from the IJOC GitHub software repository (https://github.com/INFORMSJoC) at [http://dx.doi.org/10.5281/zenodo.5381196].

Keywords: stochastic games • subgame perfect equilibria • stationary strategies • interior-point method • path-following algorithm

1. Introduction

The concept of stochastic game was coined by Shapley (1953). As a cornerstone in the area of game theory, a stochastic game enriches the model of repeated games and has been extensively applied in a large body of situations of interest such as economics and inventory and supply chain management; see, for example, Chatterjee et al. (1993), Karatzas and Sudderth (1994), Ericson and Pakes (1995), Banks and Duggan (2000), Cachon and Netessine (2004), and Olsen and Parker (2014). A stochastic game of interest models a dynamic process played by a finite number of players in a sequence of stages, which vary with some observable states. Specifically, at the beginning of the first stage, all players are in the same initial state. They select their own actions independently and simultaneously, and they get their instantaneous payoffs immediately. Each player is subsequently informed about the actions of the other players at this stage, and the game moves to the second stage. The new state is selected with a probability that is based on the outcome of a chance experiment, which is determined by the previous state and action profile. The procedure is repeated over an infinite number of stages, and a series of such repetitive stage games gives rise to a stochastic game.

Subgame perfection is broadly considered a desirable property of an equilibrium. Therefore, as mentioned in Fudenberg and Tirole (1991), a subgame perfect equilibrium in stationary strategies (SSPE) is one of the most important concepts in stochastic games. A stationary strategy depends only on the current state rather than the entire history of states and action profiles, which is consistent with the principle of "letting bygones be bygones". The existence of stationary equilibria in stochastic games has been studied extensively in earlier papers; see He and Sun (2017) and Jaśkiewicz and Nowak (2018) for wonderful reviews. Shapley (1953) proved the existence of stationary equilibria for zero-sum games with finite action and state spaces. Fink (1964) and Sobel (1971) extended Shapley's model to general n-person cases. For the model with a finite state and action space, they showed the existence of an SSPE. Solan (1998) gave proof to the existence of an SSPE for an n-player discounted stochastic game with measurable action and state spaces. Maitra and Sudderth (1998) discussed several sufficient conditions for the existence of an SSPE under a Borel state space and compact metric action spaces. Moreover, another related work proposed by Nowak and Raghavan (1992) demonstrated that every non-zero-sum discounted stochastic game with measurable state space and compact metric action spaces admits a stationary correlated equilibrium with symmetric information.

The computation of SSPEs is very significant in applications of stochastic games, as stated in McKelvey and McLennan (1996), Pakes and McGuire (2001), and Herings and Peeters (2004). However, this computation remains a challenging problem because the structure of stochastic games is very complicated. Homotopy methods as proposed by Scarf (1967) Eaves (1972) and are a class of powerful methods for solving problems that can be formulated as a fixed point problem; see Saigal (1983). Examples of such problems are computing competitive equilibria in general equilibrium models (see, for instance, Den Elzen et al. (1994) and Zhan and Dang (2018)), computing equilibria in noncooperative game theory (see Herings and Peeters 2010 and Chen and Dang 2021), finding a solution to the variational inequality problem (see Zhao and Li 2001 and Zhan et al. 2020), and solving the piecewise linear approximation of mixed complementarity problems (MCPs) (see Ferris et al. 2000).

So far, there have been several homotopy-based path-following methods to compute SSPEs for stochastic games. For instance, a stochastic linear tracing procedure (SLTP) was developed in Herings and Peeters (2004), which extended a reasoning process in Harsanyi and Selten (1988) for equilibrium selection in normal-form games to the class of stochastic games and constructed a differentiable path converging to an SSPE for a generic finite discounted stochastic game. SLTP is the first globally convergent algorithm to solve for an SSPE in a stochastic game, indicating that homotopy methods can be applied to stochastic games as well. Govindan and Wilson (2009) extended the global Newton method to stochastic games, which induced a piecewise smooth homotopy path for finding SSPEs for any generic stochastic game. Eibelshäuser and Poensgen (2019) defined a Markov quantal response equilibrium with regard to a precision parameter and developed a logit homotopy path-following method to approach an SSPE for any stochastic game as the precision parameter goes to infinity. More recently, Li and Dang (2020) developed a modified version of SLTP with arbitrary starting point to compute an SSPE for any finite discounted stochastic game.

SLTP as developed by Herings and Peeters (2004) is based on game-theoretic arguments to select a particular SSPE. However, SLTP has not been designed to achieve the highest numerical efficiency. The starting stationary strategy profile of SLTP cannot be arbitrarily chosen but is a combination of solutions to several Markov decision problems, which have to be computed explicitly. Besides, the homotopy path of the classic SLTP is only piecewise differentiable, and one has to switch between several different systems of equations to follow it, which leads to additional computational burden. In Herings and Peeters (2004), it is shown that the switching between different systems can be avoided by a suitably chosen transformation of variables, a smoothing technique used in other works as well (see Herings and Peeters (2001) and Herings and Schmedders (2006)), but this approach also leads to an increase in computation time.

Through Karush-Kuhn-Tucker conditions, one can reformulate the problem of computing an SSPE as an equivalent problem of finding a solution to an MCP, which is a commonplace for describing applications in operations research, engineering, economics, and their related fields. Excellent reviews on these applications can be found in Murphy et al. (2016), Gutierrez et al. (2017), and the references therein. A number of approaches have been proposed for providing numerical solutions to MCPs such as a nonmonotone stabilized Newton method developed in Dirkse and Ferris (1995) and a semismooth algorithm introduced in Munson et al. (2001). Interior-point methods were first proposed for nonlinear programming in Fiacco and McCormick (1968, 1990) and have been substantially developed in the literature to solve large-scale convex optimization problems; see, for instance, Ye (1991, 1992, 2011). The key idea of the interior-point technique is to restrict the points on the "path" to the interior of the feasible set, thereby bypassing many boundary points. The application of interior-point methods to complementarity problems has attracted much attention in the literature. Simantiraki and Shanno (1997) introduced an infeasible interior-point method to compute a solution to the linear complementarity problem. Huang and Mehrotra (2017) developed a modified potential reduction interior-point method to solve monotone complementarity problems. An interior-point algorithm based on a modification of Newton's method was employed by Gutierrez et al. (2017) to solve mixed nonlinear complementarity problems. Although some of the mentioned methods have an excellent numerical performance for MCPs, the convergence or stability of these methods very much depends on the structure of the original problems. Hence, the existing methods to solve MCPs may fail to find SSPEs for stochastic games.

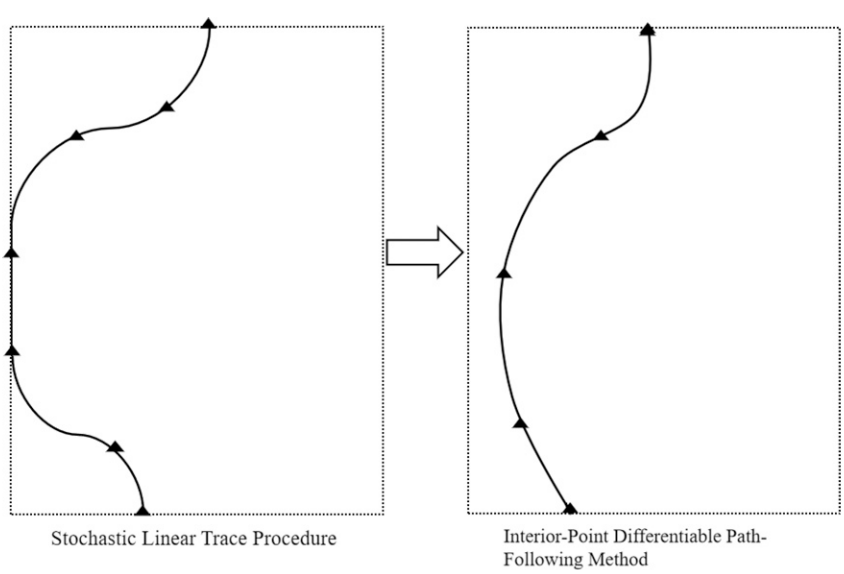
In the past decade, the idea of interior-point methods has been incorporated with homotopy methods and applied in market equilibrium problems and norm-form games, which presents satisfactory convergence properties and numerical performance. Some wonderful reviews about applications of homotopy-based interior-point methods can be found in Dang et al. (2011), Zhu et al. (2012), and Chen and Dang (2016). It is therefore natural to ask whether one can extend the "homotopy-based interior-point" idea to stochastic games and if the nice performance of the interior-point technique and good convergence properties of the homotopy methods are simultaneously preserved in this complicated class of games.

In this paper, we develop an interior-point differentiable path-following method (IPM) to compute SSPEs for finite discounted stochastic games. We achieve this by the incorporation of a logarithmic barrier term into the original payoff function and formulate an artificial barrier game, which deforms continuously from a trivial game to the stochastic game of interest. With this barrier game, a homotopy system is developed whose solutions induce an everywhere smooth path. Following the path, an SSPE for the stochastic game of interest is approximated as t descends to 0.

An advantage of IPM is that the starting point can be arbitrarily chosen, without the need to solve any optimization problem. Different starting points may lead to different SSPEs, and therefore repeated applications of IPM with randomly generated starting points may find various SSPEs. IPM introduces several advantages over SLTP and the PATH solver, a widely used method for solving general MCPs. First, IPM retains the nice property of global convergence the homotopy method possesses, and therefore the stability of IPM can be guaranteed. Second, the barrier term in IPM forces the points on the homotopy path to stay in the interior of its domain and never touch any boundary before t vanishes—that is, for any given t larger than 0, the equilibria of the artificial game are in totally mixed strategies. The switching between different systems of equations in SLTP can be therefore avoided by IPM; see Figure 1 for an illustration. Third, IPM makes use of a well-chosen perturbation term, which enables us to find an SSPE for every stochastic game. Finally, a well-chosen transformation of variables is used in IPM, which reduces the number of equations and unknowns by about one half and ensures a high numerical efficiency of IPM.

The remainder of this paper is organized as follows. In Section 2, we introduce some preliminaries about

Figure 1. Different Trajectories of the Two Methods



finite stochastic games and reformulate the concept of SSPE as the solution to a suitably chosen system of equations. In Section 3, we construct an artificial barrier stochastic game and propose our interior-point differentiable path-following method. We prove that our method is effective for computing an SSPE for any stochastic game. Extensive numerical results are reported in Section 4, where we develop a well-chosen transformation of variables to reduce the number of equations and variables in IPM by about one half. We compute SSPEs for stochastic games with different numbers of actions and players. We also compare the performance of IPM with SLTP and a widely used tool for MCPs called the PATH solver to further illustrate the efficiency and stability of IPM. Finally, this paper is concluded in Section 5.

2. Stationary Equilibria in Stochastic Games

2.1. Some Preliminaries

In this subsection, we introduce some basic notations and describe a finite discounted stochastic game as $\Gamma = \langle N, \Omega, \{S^i_\omega\}_{(i,\omega)\in N\times\Omega}, \{u^i\}_{i\in N}, \pi, \delta\rangle$, where

- Denote by $N = \{1, 2, \dots, n\}$ the set of players;
- Denote by $\Omega = \{\omega_1, \omega_2, \dots, \omega_d\}$ the state space;
- Denote by $S^i_\omega = \{s^i_{\omega j} \mid j \in M^i_\omega\}$ the set of actions of player $i \in N$ in state $\omega \in \Omega$, where $M^i_\omega = \{1, 2, \dots, m^i_\omega\}$ is the index set of actions of player i in state ω; and
- Denote by $S_\omega = \prod_{i=1}^n S^i_\omega$ the set of action profiles in state ω.
- Let $S^{-i}_\omega = \prod_{k\in N\setminus\{i\}} S^k_\omega$. Then, $s_\omega = (s^1_{\omega j_1}, s^2_{\omega j_2}, \dots, s^n_{\omega j_n}) \in S_\omega$ can be written as $s_\omega = (s^i_{\omega j_i}, s^{-i}_\omega)$ with $s^{-i}_\omega \in S^{-i}_\omega$.
- Denote by $X^i_\omega = \{x^i_\omega \in \mathbb{R}^{m^i_\omega}_+ \mid \sum_{j\in M^i_\omega} x^i_{\omega j} = 1\}$ the set of mixed actions of player i in state ω. For $x^i_\omega \in X^i_\omega$, the probability assigned to $s^i_{\omega j} \in S^i_\omega$ equals $x^i_{\omega j}$.
- Denote by $X_\omega = \prod_{i=1}^n X^i_\omega$ the set of mixed action profiles in state ω. If $x_\omega \in X_\omega$ is played, then the probability that an action profile $s_\omega = (s^1_{\omega j_1}, s^2_{\omega j_2}, \dots, s^n_{\omega j_n}) \in S_\omega$ occurs is equal to $\prod_{i=1}^n x^i_{\omega j_i}$.
- Let $X^{-i}_\omega = \prod_{k\in N\setminus\{i\}} X^k_\omega$. Then, $x_\omega = (x^1_{\omega j_1}, x^2_{\omega j_2}, \dots, x^n_{\omega j_n}) \in X_\omega$ can be written as $x_\omega = (x^i_{\omega j_i}, x^{-i}_\omega)$ with $x^{-i}_\omega \in X^{-i}_\omega$.
- Let $D = \{(\omega, s_\omega) \mid \omega \in \Omega, s_\omega \in S_\omega\}$. A history up to stage $\kappa \geq 0$ is a sequence $h_\kappa = ((\omega^0, s_{\omega^0}), (\omega^1, s_{\omega^1}), \dots, (\omega^{\kappa-1}, s_{\omega^{\kappa-1}}), \omega^\kappa)$. Then, the set of possible histories up to stage κ equals $H_\kappa = D^\kappa \times \Omega$.
- Let $u^i : D \to R$ be the instantaneous payoff function of player i. We have

$$u^i(\omega, x_\omega) = \sum_{s_\omega\in S_\omega} u^i(\omega, s_\omega)\prod_{i=1}^n x^i_{\omega j_i} = \sum_{j\in M^i_\omega} x^i_{\omega j} u^i(\omega, s^i_{\omega j}, x^{-i}_\omega).$$

- The term δ is the discount factor with $0 < \delta < 1$.
- Let $\pi(\varpi \mid \omega, s_\omega)$ be the probability that the system jumps from state ω to state ϖ when the action profile s_ω is chosen, where $\pi(\varpi \mid \omega, s_\omega) \geq 0$ and $\sum_{\varpi\in\Omega} \pi(\varpi \mid \omega, s_\omega) = 1$.
- Denote by $\pi(\omega, s_\omega) = (\pi(\omega_1 \mid \omega, s_\omega), \pi(\omega_2 \mid \omega, s_\omega), \dots, \pi(\omega_d \mid \omega, s_\omega))$ the state transition probability map.
- Let $X = \prod_{i\in N} X^i$ with $X^i = \prod_{\omega\in\Omega} X^i_\omega$. An element of X has $\overline{m} = \sum_{i\in N}\sum_{\omega\in\Omega} m^i_\omega$ components.

2.2. Equilibrium System

A strategy of player $i \in N$ is a function that assigns a feasible mixed action after each possible history. A strategy profile is a Nash equilibrium if no player has a profitable deviation from it (i.e., can choose a strategy that gives strictly higher payoffs given the strategies of the other players). Each history $h \in \bigcup_{\kappa=0}^\infty H_\kappa$ induces a subgame of Γ. A strategy profile is a subgame perfect equilibrium if it induces a Nash equilibrium in every subgame of Γ.

The strategy of player i is stationary if it depends only on the current state, so the player chooses the same mixed action at histories with the same current state. A stationary strategy of player $i \in N$ can therefore be represented by an element $x^i \in X^i$ and a stationary strategy profile by an element $\boldsymbol{x} \in X$. The most important solution concept that has been used in applications of stochastic games is the SSPE. An SSPE is a stationary strategy profile that induces a Nash equilibrium in every subgame of Γ. No player has a profitable deviation from an SSPE, even when deviations are not required to be stationary themselves.

We now reformulate the concept of SSPE as the solution to a suitably chosen system of equations. Given a stationary strategy profile $\boldsymbol{x} \in X$, we denote the present value of player $i \in N$ at state $\omega \in \Omega$ of the expected payoff of the next k stages by $\mu^i_\omega(k)$. The value of $\mu^i_\omega(k)$ follows from the following system of recursive equations, $\mu^i_\omega(k+1) = u^i(\omega, x_\omega) + \delta\sum_{\varpi\in\Omega} \pi(\varpi \mid \omega, x_\omega)\mu^i_\varpi(k)$. For any initial state $\omega \in \Omega$, let $\mu^i_\omega = \lim_{k\to\infty}\mu^i_\omega(k)$ be the total expected payoff for player i. Without loss of generality, we assume $\mu^i_\omega(k)$ is increasing in k.[1] Because $\mu^i_\omega(k)$ is a uniformly bounded function of k, μ^i_ω always exists. It is the unique solution to the linear system of equations:

$$\mu^i_\omega = u^i(\omega, x_\omega) + \delta\sum_{\varpi\in\Omega} \pi(\varpi \mid \omega, x_\omega)\mu^i_\varpi. \tag{1}$$

For simplicity, for any given stationary strategy profile $\boldsymbol{x} \in X$, we define

$$\varphi^i(\omega, s^i_{\omega j}, x^{-i}_\omega, \mu^i(\boldsymbol{x})) = u^i(\omega, s^i_{\omega j}, x^{-i}_\omega) + \delta\sum_{\varpi\in\Omega} \pi(\varpi \mid \omega, s^i_{\omega j}, x^{-i}_\omega)\mu^i_\varpi(\boldsymbol{x}),$$

where $\mu^i(\boldsymbol{x}) = (\mu^i_\omega(\boldsymbol{x}))_{\omega\in\Omega}$ is the unique solution to the linear system (1). Then, for any stationary strategy profile $\boldsymbol{x} \in X$, given a state $\omega \in \Omega$, the optimal mixed action of a player $i \in N$ who can only deviate once from x^i in state ω can be found as the solution to the optimization problem:

$$\begin{aligned}\max_{\hat{x}^i_\omega \in X^i_\omega} \quad & \sum_{j\in M^i_\omega} \hat{x}^i_{\omega j}\varphi^i(\omega, s^i_{\omega j}, x^{-i}_\omega, \mu^i(\boldsymbol{x}))\\ \text{s.t.} \quad & \hat{x}^i_{\omega j} \ge 0, j \in M^i_\omega,\\ & \sum_{j\in M^i_\omega} \hat{x}^i_{\omega j} = 1. \end{aligned} \tag{2}$$

A direct application of the optimality conditions yields

$$\begin{aligned}&\varphi^i(\omega, s^i_{\omega j}, x^{-i}_\omega, \mu^i(\boldsymbol{x})) + \hat{\lambda}^i_{\omega j} - \hat{\beta}^i_\omega = 0, \omega\in\Omega,\ i\in N,\ j\in M^i_\omega,\\ &\hat{\lambda}^i_{\omega j}\hat{x}^i_{\omega j} = 0,\ \hat{\lambda}^i_{\omega j} \ge 0,\ \hat{x}^i_{\omega j} \ge 0, \qquad \omega\in\Omega,\ i\in N,\ j\in M^i_\omega,\\ &\sum_{j\in M^i_\omega}\hat{x}^i_{\omega j} = 1, \qquad \omega\in\Omega,\ i\in N. \end{aligned} \tag{3}$$

We multiply both sides of the first group of equations in (3) with $\hat{x}^i_{\omega j}$, sum each side over $j \in M^i_\omega$, and obtain $\hat{\beta}^i_\omega = \sum_{j\in M^i_\omega} \hat{x}^i_{\omega j}\varphi^i(\omega, s^i_{\omega j}, x^{-i}_\omega, \mu^i(\boldsymbol{x})) = u^i(\omega, \hat{x}^i_\omega, x^{-i}_\omega) + \sum_{\varpi\in\Omega}\pi(\varpi \mid \omega, \hat{x}^i_\omega, x^{-i}_\omega)\mu^i_\varpi(\boldsymbol{x})$. The set of solutions to (3), which is denoted by $B(\boldsymbol{x})$, consists of best responses to the given stationary strategy profile $\boldsymbol{x}$. It follows from the one-stage deviation principle that $\hat{\boldsymbol{x}}$ is an SSPE if and only if it is a fixed point of the best response correspondence (i.e., $\hat{\boldsymbol{x}} \in B(\hat{\boldsymbol{x}})$); see Fudenberg and Tirole (1991). Then, letting $\boldsymbol{x}, \hat{\boldsymbol{x}}$ and $\hat{\mu}^i = \mu^i(\hat{\boldsymbol{x}})$, we get

$$\begin{aligned}&\varphi^i(\omega, s^i_{\omega j}, \hat{x}^{-i}_\omega, \hat{\mu}^i) + \hat{\lambda}^i_{\omega j} - \hat{\beta}^i_\omega = 0, \omega\in\Omega,\ i\in N,\ j\in M^i_\omega,\\ &\hat{\lambda}^i_{\omega j}\hat{x}^i_{\omega j} = 0,\ \hat{\lambda}^i_{\omega j} \ge 0,\ \hat{x}^i_{\omega j} \ge 0, \qquad \omega\in\Omega,\ i\in N,\ j\in M^i_\omega,\\ &\sum_{j\in M^i_\omega}\hat{x}^i_{\omega j} = 1, \qquad \omega\in\Omega,\ i\in N. \end{aligned}$$

It follows from (1) that $\hat{\beta}^i_\omega = \hat{\mu}^i_\omega$. From the preceding discussion, a stationary strategy profile $\hat{\boldsymbol{x}} \in X$ is an SSPE if and only if $\hat{\boldsymbol{x}}$ together with some $(\hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\mu}}) \in \mathbb{R}^{\overline{m}} \times \mathbb{R}^{nd}$ satisfies

$$\begin{aligned}&u^i(\omega, s^i_{\omega j}, \hat{x}^{-i}_\omega) + \delta\sum_{\varpi\in\Omega}\pi(\varpi \mid \omega, s^i_{\omega j}, \hat{x}^{-i}_\omega)\hat{\mu}^i_\varpi + \hat{\lambda}^i_{\omega j} - \hat{\mu}^i_\omega = 0,\\ &\qquad\qquad \omega\in\Omega,\ i\in N,\ j\in M^i_\omega,\\ &\hat{\lambda}^i_{\omega j}\hat{x}^i_{\omega j} = 0,\ \hat{\lambda}^i_{\omega j} \ge 0,\ \hat{x}^i_{\omega j} \ge 0, \qquad \omega\in\Omega,\ i\in N,\ j\in M^i_\omega,\\ &\sum_{j\in M^i_\omega}\hat{x}^i_{\omega j} = 1, \qquad \omega\in\Omega,\ i\in N. \end{aligned} \tag{4}$$

Thus far, we have reformulated the problem to find an SSPE as the equivalent problem of solving the nonlinear system of equations (4). To be more precise, (4) is actually an MCP, which is a generation of the nonlinear complementarity problem and provides a broad unifying setting for the study of equilibrium problems; see Ferris et al. (2013). In the remainder of this paper, we aim to explore an effective and efficient algorithm to compute a solution to System (4).

3. An Interior-Point Differentiable Path-Following Method

In general, it is difficult to solve System (4) directly. Herings and Peeters (2004) extended the linear tracing procedure of Harsanyi and Selten (1988) to the class of stochastic games. We refer to the resulting method as the stochastic linear tracing procedure (SLTP). SLTP is a homotopy method, which starts from an artificial game where all players optimize against given prior beliefs. The homotopy variable t corresponds to the weight that is put on the artificial game. It is shown in Herings and Peeters (2004) that SLTP converges to an SSPE for almost every stochastic game. The homotopy path of SLTP is piecewise differentiable. To follow it, one either has to switch between different systems of equations or use a transformation of variables, which makes the homotopy path smooth. The latter operation comes at the expense of computational speed.

To avoid switching between different systems of equations, we propose an interior-point differentiable path-following method to find an SSPE. By introducing a homotopy variable t ranging from 1 to 0, we incorporate a logarithmic barrier term into the payoff functions and formulate an artificial barrier stochastic game, which continuously deforms from a trivial stochastic game with a unique solution to the stochastic game of interest as t descends from 1 to 0. With this artificial game, we develop a smooth path, which is constructed as the collection of equilibria for the artificial game at different levels of the homotopy variable t. As t decreases to 0, an SSPE for the stochastic game of interest is obtained. Additionally, because of the existence of the barrier term, each point on the path is restricted to the interior before t vanishes.

Consider a player $i \in N$. For any given $t \in [0,1]$ and any given stationary strategy profile $\boldsymbol{x} \in X$, $\mu^i = \mu^i(\boldsymbol{x}) = (\mu^i_\omega(\boldsymbol{x}))_{\omega\in\Omega}$ is defined to be the unique solution to the following linear system:

$$\mu^i_\omega = (1-t)(u^i(\omega, x_\omega) + \delta\sum_{\varpi\in\Omega}\pi(\varpi \mid \omega, x_\omega)\mu^i_\varpi) + t^2, \quad \omega\in\Omega. \tag{5}$$

Then, for any stationary strategy profile $\boldsymbol{x} \in X$, for every $\omega \in \Omega$, player $i \in N$ solves the optimization problem:

$$\max_{\hat{x}^i_\omega \in X^i_\omega} \quad (1-t)\sum_{j\in M^i_\omega} \hat{x}^i_{\omega j}\varphi^i(\omega, s^i_{\omega j}, x^{-i}_\omega, \mu^i(x)) + t^2 \sum_{j\in M^i_\omega} x^{0,i}_{\omega j}\ln \hat{x}^i_{\omega j}$$
$$\text{s.t.} \quad \sum_{j\in M^i_\omega}\hat{x}^i_{\omega j} = 1, \tag{6}$$

where $x^{0,i}_\omega$ is an arbitrary totally mixed stationary strategy profile with $\sum_{j\in M^i_\omega} x^{0,i}_{\omega j} = 1$.

The term in front of the logarithmic part is equal to t^2 rather than t. The reason will become clear at the end of this section, where a transformation of variables is introduced to reduce the number of equations and unknowns. To guarantee differentiability of the system of equations after the transformation of variables, the term t^2 is needed rather than t. When t is equal to 0, the logarithmic part is not taken into account. For t positive, we only maximize over mixed actions without zero components.

The optimality conditions of problem (6) read as

$$(1-t)\varphi^i(\omega, s^i_{\omega j}, x^{-i}_\omega, \mu^i(x)) + \hat{\lambda}^i_{\omega j} - \hat{\beta}^i_\omega = 0, \omega\in\Omega,\ i\in N,\ j\in M^i_\omega,$$
$$\hat{\lambda}^i_{\omega j}\hat{x}^i_{\omega j} - t^2 x^{0,i}_{\omega j} = 0, \qquad \omega\in\Omega,\ i\in N,\ j\in M^i_\omega,$$
$$\sum_{j\in M^i_\omega}\hat{x}^i_{\omega j} = 1, \qquad \omega\in\Omega,\ i\in N. \tag{7}$$

Multiplying both sides of the first group of equations in System (7) by $\hat{x}^i_{\omega j}$ and summing over j, we obtain that $\hat{\beta}^i_\omega = (1-t)[u^i(\omega, \hat{x}^i_\omega, x^{-i}_\omega) + \delta\sum_{\varpi\in\Omega}\pi(\varpi \mid \omega, \hat{x}^i_\omega, x^{-i}_\omega)\mu^i_\varpi(x)] + t^2$. From a fixed point argument and (5), letting $x = \hat{x}$, we obtain that $\hat{\beta}^i_\omega = \hat{\mu}^i_\omega$. The equilibrium system of the artificial stochastic barrier game is therefore given by

$$(1-t)(u^i(\omega, s^i_{\omega j}, \hat{x}^{-i}_\omega) + \delta\sum_{\varpi\in\Omega}\pi(\varpi \mid \omega, s^i_{\omega j}, \hat{x}^{-i}_\omega)\hat{\mu}^i_\varpi)$$
$$+\hat{\lambda}^i_{\omega j} - \hat{\mu}^i_\omega = 0, \qquad \omega\in\Omega,\ i\in N,\ j\in M^i_\omega,$$
$$\hat{\lambda}^i_{\omega j}\hat{x}^i_{\omega j} - t^2 x^{0,i}_{\omega j} = 0, \qquad \omega\in\Omega,\ i\in N,\ j\in M^i_\omega,$$
$$\sum_{j\in M^i_\omega}\hat{x}^i_{\omega j} = 1, \qquad \omega\in\Omega,\ i\in N. \tag{8}$$

As $t = 1$, System (8) becomes a system that is very easy to solve:

$$\hat{\lambda}^i_{\omega j} - \hat{\mu}^i_\omega = 0, \qquad \omega\in\Omega,\ i\in N,\ j\in M^i_\omega,$$
$$\hat{\lambda}^i_{\omega j}\hat{x}^i_{\omega j} - x^{0,i}_{\omega j} = 0, \omega\in\Omega,\ i\in N,\ j\in M^i_\omega,$$
$$\sum_{j\in M^i_\omega}\hat{x}^i_{\omega j} = 1, \qquad \omega\in\Omega,\ i\in N. \tag{9}$$

Theorem 1. *As $t = 1$, System (8) has a unique solution, $(\hat{\boldsymbol{x}}(1), \hat{\boldsymbol{\lambda}}(1), \hat{\boldsymbol{\mu}}(1))$, with $\hat{x}^i_{\omega j}(1) = x^{0,i}_{\omega j}$, $\hat{\lambda}^i_{\omega j}(1) = 1$, and $\hat{\mu}^i_\omega(1) = 1$, where $\omega\in\Omega$, $i\in N$, and $j\in M^i_\omega$.*

Proof. By Problem (6), we know that as $t = 1$, for every $i\in N$, for every $\omega\in\Omega$, System (8) corresponds to the necessary and sufficient conditions of the following optimization problem:

$$\max_{\hat{x}^i_\omega\in X^i_\omega} \quad \sum_{j\in M^i_\omega} x^{0,i}_{\omega j}\ln(\hat{x}^i_{\omega j})$$
$$\text{s.t.} \quad \sum_{j\in M^i_\omega}\hat{x}^i_{\omega j} = 1, \tag{10}$$

which is a strictly convex optimization problem with a unique solution. The solution of Problem (10) is given by $\hat{x}^i_{\omega j} = x^{0,i}_{\omega j}$. From the system of equations (9), we obtain that for all $\omega\in\Omega$, $i\in N$, and $j\in M^i_\omega$, $(\hat{\lambda}^i_{\omega j}, \hat{\mu}^i_\omega) = (1, 1)$. Q.E.D.

At $t = 0$, the definition of $\boldsymbol{\mu}$ in (5) is the same as that in (1), and System (8) reduces to (4), the equilibrium system of the stochastic game of interest.

Next, we prove that the set of solutions to the system of equations (8) generates an everywhere smooth path from the arbitrarily chosen starting point x^0 to an SSPE of the stochastic game of interest.

For the analysis that follows next, we need Mas-Colell's fixed point theorem, proposed by Mas-Colell (1974).

Theorem 2 (Mas-Colell's Fixed Point Theorem). *Let S be a nonempty, compact, and convex subset of $\mathbb{R}^\ell$, and let $f: S\times[0,1]\to S$ be an upper hemicontinuous correspondence such that for every $(s,t)\in S\times[0,1]$, $f(s,t)$ is contractible. Then the set $F = \{(s,t)\in S\times[0,1] \mid s\in f(s,t)\}$ contains a connected subset F^c such that $(S\times\{1\})\cap F^c \neq \emptyset$ and $(S\times\{0\})\cap F^c\neq\emptyset$.*

For $i\in N$, $\omega\in\Omega$, and any strategy profile $\boldsymbol{x}\in X$, let $\sigma^i_\omega(\boldsymbol{x}, t)$ be all $\hat{x}^i_\omega\in X^i_\omega$ that solve

$$\max_{\hat{x}^i_\omega\in X^i_\omega} (1-t)\sum_{j\in M^i_\omega}\hat{x}^i_{\omega j}(u^i(\omega, s^i_{\omega j}, x^{-i}_\omega) + \delta\sum_{\varpi\in\Omega}\pi(\varpi|\omega, s^i_{\omega j}, x^{-i}_\omega)\mu^i_\varpi(\boldsymbol{x}))$$
$$+t^2\sum_{j\in M^i_\omega}x^{0,i}_{\omega j}\ln(\hat{x}^i_{\omega j}) - t(1-t)\sum_{j\in M^i_\omega}\alpha^i_{\omega j}\hat{x}^i_{\omega j}$$
$$\text{s.t.} \quad \sum_{j\in M^i_\omega}\hat{x}^i_{\omega j} = 1, \tag{11}$$

where $\alpha\in\mathbb{R}^{\overline{m}}$.

For any given $(\boldsymbol{x}, t)\in X\times[0,1]$, $H(\boldsymbol{x}, t)$ is defined as the set of all $\hat{\boldsymbol{x}}\in X$ satisfying the system of equations (12), which corresponds to the optimality conditions of Problem (11):

$$(1-t)(u^i(\omega, s^i_{\omega j}, x^{-i}_\omega) + \delta\sum_{\varpi\in\Omega}\pi(\varpi \mid \omega, s^i_{\omega j}, x^{-i}_\omega)\mu^i_\varpi(x))$$
$$+\hat{\lambda}^i_{\omega j} - \hat{\beta}^i_\omega - t(1-t)\alpha^i_{\omega j} = 0, \omega\in\Omega,\ i\in N,\ j\in M^i_\omega,$$
$$\hat{\lambda}^i_{\omega j}\hat{x}^i_{\omega j} - t^2 x^{0,i}_{\omega j} = 0, \qquad \omega\in\Omega,\ i\in N,\ j\in M^i_\omega,$$
$$\sum_{j\in M^i_\omega}\hat{x}^i_{\omega j} = 1, \qquad \omega\in\Omega,\ i\in N. \tag{12}$$

Compared with Problem (6), Problem (11) contains an additional term, $-t(1-t)\sum_{j\in M^i_\omega}\alpha^i_{\omega j}\hat{x}^i_{\omega j}$. When $t = 0$ or $t = 1$, this term disappears, and the two systems are completely the same. We use α as a perturbation term to avoid degeneracies. In the numerical implementation of our algorithm, perturbations were not needed, and we could always choose α equal to the zero vector.

We argue next that $\sigma^i_\omega : X \times [0,1] \to X^i_\omega$ is an upper hemicontinuous correspondence. This follows from the fact that the limit of any convergent sequence of solutions to (12) is a solution to (12), so the graph of σ^i_ω is closed, which is equivalent to σ^i_ω being an upper hemicontinuous correspondence. Note that σ^i_ω is a continuous function on $X \times (0,1]$ because the logarithmic term in the objective function of Problem (11) is strictly concave. Therefore, for any $t \in (0,1]$, Problem (11) is a strictly convex optimization model with a unique solution; that is, σ^i_ω is single valued.

From the preceding discussion, $H(x, t)$ is obtained as a product of σ^i_ω for $\omega \in \Omega$ and $i \in N$. Then, $H(x, t)$ is also an upper hemicontinuous correspondence.

Lemma 1. *For any $(x,t) \in X \times [0,1]$, $H(x, t)$ is a convex subset of the convex strategy space X.*

Proof. Recall that σ^i_ω is the set of solutions to Problem (11). Obviously, for any $(x,t) \in X^i_\omega \times (0,1]$, Problem (11) is a strictly concave optimization problem and σ^i_ω is single valued. That is, σ^i_ω is a convex subset of X^i_ω for any $(x,t) \in X^i_\omega \times (0,1]$. As $t = 0$, Problem (11) reduces to a linear optimization problem in $\hat{x}^i_\omega \in X^i_\omega$. Suppose $\hat{x}^{1,i}_\omega \in \sigma^i_\omega(x,0)$ and $\hat{x}^{2,i}_\omega \in \sigma^i_\omega(x,0)$. Then, for any constant $a \in [0,1]$, $a\hat{x}^{1,i}_\omega + (1-a)\hat{x}^{2,i}_\omega$ also optimizes the objective function of Problem (11), and $\sum_{j\in M^i_\omega}(a\hat{x}^{1,i}_{\omega j} + (1-a)\hat{x}^{2,i}_{\omega j}) = a + 1 - a = 1$. Then, $a\hat{x}^{1,i}_\omega + (1-a)\hat{x}^{2,i}_\omega \in \sigma^i_\omega(x,0)$. That is, $\sigma^i_\omega(x,t)$ is a convex-valued mapping on $X^i_\omega \times \{0\}$. Recall that $H(x, t)$ is obtained as a product of σ^i_ω for $\omega \in \Omega$ and $i \in N$. Therefore, $H(x, t)$ is a convex subset of X for any $(x,t) \in X \times [0,1]$. Q.E.D.

Let Φ be all $(\hat{x},t) \in X \times [0,1]$ satisfying the following system of equations:

$$
\begin{aligned}
&(1-t)(u^i(\omega, s^i_{\omega j}, \hat{x}^{-i}_\omega) + \delta \sum_{\varpi\in\Omega} \pi(\varpi \mid \omega, s^i_{\omega j}, \hat{x}^{-i}_\omega)\hat{\mu}^i_\varpi) \\
&\quad + \hat{\lambda}^i_{\omega j} - \hat{\mu}^i_\omega - t(1-t)\alpha^i_{\omega j} = 0, \omega \in \Omega,\ i \in N,\ j \in M^i_\omega, \\
&\hat{\lambda}^i_{\omega j}\hat{x}^i_{\omega j} - t^2 x^{0,i}_{\omega j} = 0, \qquad \omega \in \Omega,\ i \in N,\ j \in M^i_\omega, \\
&\sum_{j\in M^i_\omega} \hat{x}^i_{\omega j} - 1 = 0, \qquad \omega \in \Omega,\ i \in N,
\end{aligned} \tag{13}
$$

which is essentially the same as System (8) regardless of the perturbation term. Then the following corollary is established.

Corollary 1. *The set Φ has a connected subset Φ^c such that $(\mathbb{R}^{\overline{m}} \times \{1\}) \cap \Phi^c \neq \emptyset$ and $(\mathbb{R}^{\overline{m}} \times \{0\}) \cap \Phi^c \neq \emptyset$.*

Proof. Comparing Systems (12) and (13), we find that when $x = \hat{x}$, System (12) is identical to (13). Then, Φ can be rewritten as

$$\Phi = \{(\hat{x}, t) \in \mathbb{R}^{\overline{m}} \times [0,1] \mid \hat{x} = H(\hat{x}, t)\}.$$

The upper hemicontinuity of H, together with the fact that a convex set is contractible and a direct application of Mas-Colell's fixed point theorem, leads to the conclusion of our corollary. Q.E.D.

All equations in (13) are polynomial. The set Φ is therefore a semialgebraic set, so all its components are also path connected—that is, any two points in a component can be joined by a path; see Schanuel et al. (1991). We obtain the following corollary.

Corollary 2. *The set Φ has a path-connected subset Φ^c such that $(\mathbb{R}^{\overline{m}} \times \{1\}) \cap \Phi^c \neq \emptyset$ and $(\mathbb{R}^{\overline{m}} \times \{0\}) \cap \Phi^c \neq \emptyset$.*

Denote the left side of System (13) as $p(\boldsymbol{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}, t; \boldsymbol{\alpha})$, and for any given $\boldsymbol{\alpha} \in \mathbb{R}^{\overline{m}}$, let

$$p_\alpha(\boldsymbol{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}, t) = p(\boldsymbol{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}, t; \boldsymbol{\alpha}).$$

Fix some $\boldsymbol{\alpha} \in \mathbb{R}^{\overline{m}}$. The set of all $(\boldsymbol{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}, t) \in \mathbb{R}^{\overline{m}} \times \mathbb{R}^{\overline{m}} \times \mathbb{R}^{nd} \times [0,1]$ satisfying the system of Equations (13) is denoted by Δ. The following theorem states that our method is globally convergent.

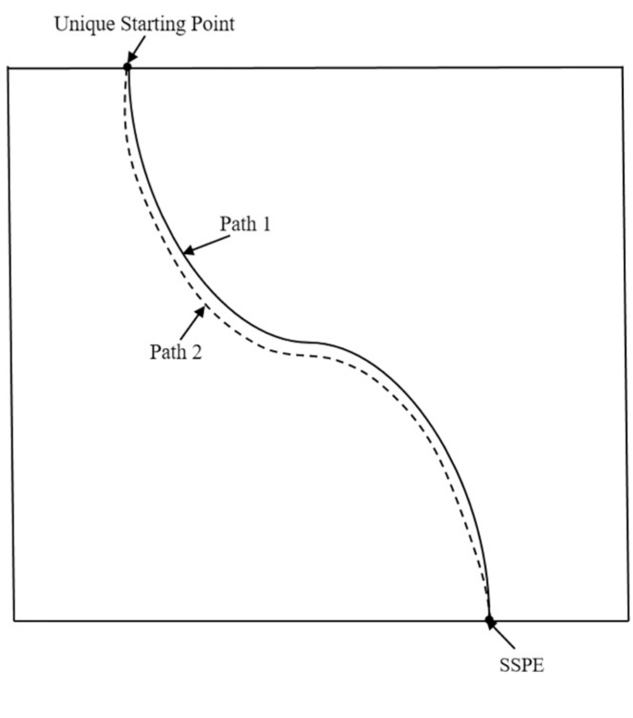
Theorem 3. *For a generic choice of $\alpha \in \mathbb{R}^{\overline{m}}$, there exists a smooth path in Δ, which starts from the totally mixed stationary strategy profile $x^0 \in X$ at $t = 1$ and ends at an SSPE for the stochastic game of interest at $t = 0$.*

Proof. As proved in Corollary 2, Φ contains a path-connected subset Φ^c that intersects both the sets $\mathbb{R}^{\overline{m}} \times \{1\}$ and $\mathbb{R}^{\overline{m}} \times \{0\}$. Now, we prove that Φ^c forms an everywhere smooth path. The second group of equations in (13) determines a unique value for $\boldsymbol{\lambda} \in \mathbb{R}^{\overline{m}}$ for each $(\boldsymbol{x}, t) \in \mathbb{R}^{\overline{m}} \times \mathbb{R}$. Next, the first group of linear equations in (13) pins down a unique value for $\boldsymbol{\mu} \in \mathbb{R}^{nd}$. Thus Δ has a path-connected subset that intersects both the sets $\mathbb{R}^{\overline{m}} \times \mathbb{R}^{\overline{m}} \times \mathbb{R}^{nd} \times \{1\}$ and $\mathbb{R}^{\overline{m}} \times \mathbb{R}^{\overline{m}} \times \mathbb{R}^{nd} \times \{0\}$. By Theorem 1, there is a unique starting point at $t = 1$. We prove in Appendix A of the online supplement that the Jacobian matrix of $p_0(\boldsymbol{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}, 1)$ is of full rank. For every $\boldsymbol{\alpha} \in \mathbb{R}^{\overline{m}}$, $p_\alpha(\boldsymbol{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}, 1) = p_0(\boldsymbol{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}, 1)$, so it follows that 0 is a regular value of $p_\alpha(\boldsymbol{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}, 1)$ on $\mathbb{R}^{\overline{m}} \times \mathbb{R}^{\overline{m}} \times \mathbb{R}^{nd} \times \{1\}$. Similarly, 0 is also a regular value of $p(\boldsymbol{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}, t; \boldsymbol{\alpha})$ because the Jacobian matrix of $p(\boldsymbol{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}, t; \boldsymbol{\alpha})$ is of full-row rank for all $(\boldsymbol{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}, t; \boldsymbol{\alpha}) \in \mathbb{R}^{\overline{m}} \times \mathbb{R}^{\overline{m}} \times \mathbb{R}^{nd} \times (0,1) \times \mathbb{R}^{\overline{m}}$; see Appendix A of the online supplement. By a direct application of the well-known transversality theorem (see Eaves and Schmedders 1999), 0 is a regular value of $p_\alpha(\boldsymbol{x},\boldsymbol{\lambda},\boldsymbol{\mu},t)$ for almost all $\boldsymbol{\alpha}\in\mathbb{R}^{\overline{m}}$. Then it follows that Φ^c is a smooth one-dimensional manifold with a boundary, which starts from the unique solution at $t=1$ and ends at an SSPE of the stochastic game of interest at $t=0$. Q.E.D.

Theorem 3 illustrates that the solutions to the system of equations (13) together with the associated Lagrangian multipliers form a smooth path. Clearly, the systems of equations (13) and (8) coincide at $t=1$ and $t=0$. When $\|\boldsymbol{\alpha}\|$ is sufficiently small, the solutions to Systems (13) and (8) are nearly the same, as is illustrated in Figure 2. Points on the smooth path 1 are approximate equilibria for the artificial stochastic game (6) on path 2. As $t=0$, System (13) corresponds to the stochastic game of interest. Thus, the end point of path 1 is an exact SSPE of the stochastic game of interest. We introduce $\boldsymbol{\alpha}$ into the equilibrium system to avoid degeneracies. In fact, for a generic choice of $\boldsymbol{\alpha}$, our method finds an SSPE for every stochastic game. SLTP does not use an $\boldsymbol{\alpha}$ to perturb the system and only computes an SSPE for almost every stochastic game. In numerical experiments, we have always obtained convergence for $\boldsymbol{\alpha}=\mathbf{0}$.

Finally, we use a suitably chosen transformation of variables to reduce the number of equations and unknowns by roughly one half, which improves the efficiency of the method. For every $\omega\in\Omega$, $i\in N$, and $j\in M^i_\omega$, we write $x^i_{\omega j}$ and $\lambda^i_{\omega j}$ as a function of a variable

Figure 2. Paths 1 and 2 Represent the Solution Sets for Systems (13) and (8), Respectively



$y^i_{\omega j}$ and the homotopy variable t,

$$x^i_{\omega j}(\boldsymbol{y},t)=\left(\frac{\sqrt{(y^i_{\omega j})^2+4t\sqrt{x^{0,i}_{\omega j}}}+y^i_{\omega j}}{2}\right)^2,$$

$$\lambda^i_{\omega j}(\boldsymbol{y},t)=\left(\frac{\sqrt{(y^i_{\omega j})^2+4t\sqrt{x^{0,i}_{\omega j}}}-y^i_{\omega j}}{2}\right)^2.$$

Clearly, $\boldsymbol{x}(\mathbf{y},t)$ and $\boldsymbol{\lambda}(\boldsymbol{y},t)$ are continuously differentiable functions for all $y\in\mathbb{R}^{\overline{m}}$ and $t\in(0,1]$, and it holds that for every $\omega\in\Omega$, $i\in N$, and $j\in M^i_\omega$, $\lambda^i_{\omega j}(\boldsymbol{y},t)x^i_{\omega j}(\boldsymbol{y},t)=t^2x^{0,i}_{\omega j}$. We substitute these functions into System (13) and obtain the following homotopy system:

$$\begin{aligned}&(1-t)(u^i(\omega,s^i_{\omega j},x^{-i}_\omega(\hat{\boldsymbol{y}},t))+\delta\sum_{\overline{\omega}\in\Omega}\pi(\overline{\omega}\mid\omega,s^i_{\omega j},x^{-i}_\omega(\hat{\boldsymbol{y}},t))\hat{\mu}^i_{\overline{\omega}})\\&\quad+\lambda^i_{\omega j}(\hat{\boldsymbol{y}},t)-\hat{\mu}^i_\omega-t(1-t)\alpha^i_{\omega j}=0,\omega\in\Omega,\ i\in N,\ j\in M^i_\omega,\\&\sum_{j\in M^i_\omega}x^i_{\omega j}(\hat{\boldsymbol{y}},t)-1=0,\qquad\qquad\omega\in\Omega,\ i\in N.\end{aligned}\tag{14}$$

The system of equations (14) corresponds to the proposed IPM, the interior-point differentiable path-following method. The transformation of variables only leads to a different parametrization of the homotopy path defined by (13), so it holds by Theorem 3 that IPM is globally convergent for a generic choice of $\boldsymbol{\alpha}\in\mathbb{R}^{\overline{m}}$.[2]

4. Numerical Performance

We use a predictor-corrector algorithm (see Eaves and Schmedders (1999) and Allgower and Georg (2012) for details) to trace the smooth path generated by IPM. To further illustrate the efficiency of IPM, we compare the numerical performance of IPM with an arbitrary starting SLTP developed by Li and Dang (2020), which is also an effective differentiable homotopy method for computing SSPEs. The range of the homotopy variable t in the arbitrary starting SLTP is $[0,2]$, where the extra part for $t\in[1,2]$ is to solve the Markov decision problems and compute the starting point for the standard SLTP. As t descends from 2 to 1, the starting point of the standard SLTP is immediately obtained. When t varies between 0 and 1, the homotopy system of the arbitrary starting SLTP is completely the same as that of the standard SLTP. As t approaches 0, the solution to the homotopy system leads to an SSPE for the stochastic game. More details about the arbitrary starting SLTP can be found in Appendix B of the online supplement.[3] All experiments are run using MatLab software on a 2.00 GHz Windows PC with CORE i7. The stopping criterion is taken equal to $t<10^{-6}$. In each experiment, the starting stationary strategy of IPM is randomly generated. When comparing IPM with SLTP, we set the experimental environment for both the methods to be exactly the same. Each experiment is run for 10 times, and we report the average number of iterations (AITER) and average computation time (ATIME) in this section. Furthermore, the computational results for all numerical instances and the source code that supports these results are available in the IJOC GitHub repository (Dang et al. 2021).

4.1. Fundamental Cases

In this subsection, we implement IPM to solve several fundamental stochastic games with different numbers of players, states, and actions. The five examples preliminarily illustrate that IPM is an effective and efficient method to compute SSPEs. The discount factor is always taken equal to $\delta = 0.95$.

Example 1. Assume $N = \{1,2\}$, $\Omega = \{\omega_1, \omega_2\}$, and, for $i = 1,2$, $S^i_{\omega_1} = \{s^i_{\omega_1 1}, s^i_{\omega_1 2}\}$ and $S^i_{\omega_2} = \{s^i_{\omega_2 1}\}$. The payoff matrices in states ω_1 and ω_2 are given by

ω_1	$s^2_{\omega_1 1}$	$s^2_{\omega_1 2}$
$s^1_{\omega_1 1}$	$(1,-1)$	$(0,0)$
$s^1_{\omega_1 2}$	$(0,0)$	$(3,-3)$

and

ω_2	$s^2_{\omega_2 1}$
$s^1_{\omega_2 1}$	$(0,0)$

.

The transition probability matrices in states ω_1 and ω_2 are given by

$\pi((\omega_1,\omega_2) \mid \omega_1)$	$s^2_{\omega_1 1}$	$s^2_{\omega_1 2}$
$s^1_{\omega_1 1}$	$(1,0)$	$(0,1)$
$s^1_{\omega_1 2}$	$(0,1)$	$(1,0)$

and

$\pi((\omega_1,\omega_2) \mid \omega_2)$	$s^2_{\omega_2 1}$
$s^1_{\omega_2 1}$	$(0,1)$

.

From the randomly generated starting point $(y^1_{\omega_1 1}, y^1_{\omega_1 2}, y^2_{\omega_1 1}, y^2_{\omega_1 2}, y^1_{\omega_2 1}, y^2_{\omega_2 1}, \mu^1_{\omega_1}, \mu^2_{\omega_1}, \mu^1_{\omega_2}, \mu^2_{\omega_2}, t) = (-0.2136, -0.3823, -0.4157, -0.1885, 0, 0, 0, 0, 0, 0, 1.0000)$, IPM finds the SSPE $(((x^1_{\omega_1 1}, x^1_{\omega_1 2}), (x^2_{\omega_1 1}, x^2_{\omega_1 2})), (x^1_{\omega_2 1}, x^2_{\omega_2 1})) = (((0.67, 0.33), (0.67, 0.33)), (1,1))$.

Figure 3 shows the development of the variables y and t in the various iterations of IPM. The downward sloping curve corresponds to t, the nonmonotonic curve to $y^1_{\omega_1 1}$, and the upward sloping curve to $y^1_{\omega_1 2}$.

Example 2. Assume $N = \{1,2\}$, $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\}$, and, for $i = 1,2$, $S^i_{\omega_1} = \{s^i_{\omega_1 1}, s^i_{\omega_1 2}\}$, $S^i_{\omega_2} = \{s^i_{\omega_2 1}, s^i_{\omega_2 2}\}$, $S^i_{\omega_3} = \{s^i_{\omega_3 1}\}$, and $S^i_{\omega_4} = \{s^i_{\omega_4 1}\}$. The payoff matrices are given by

$(\omega_k)_{k=1,2}$	$s^2_{\omega_k 1}$	$s^2_{\omega_k 2}$
$s^1_{\omega_k 1}$	$(0,0)$	$(0,0)$
$s^1_{\omega_k 2}$	$(0,0)$	$(0,0)$

,

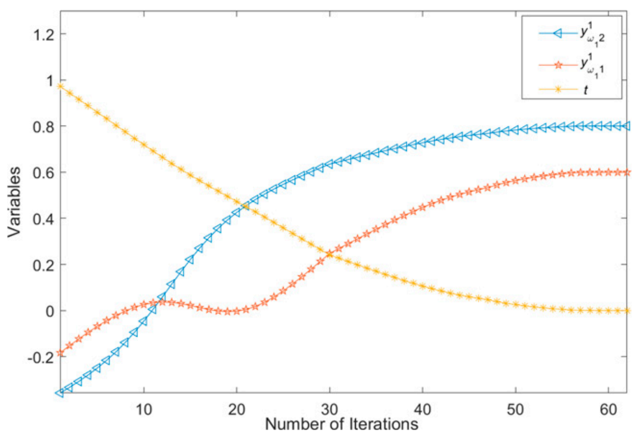
ω_3	$s^2_{\omega_3 1}$
$s^1_{\omega_3 1}$	$(1,-1)$

, and

ω_4	$s^2_{\omega_4 1}$
$s^1_{\omega_4 1}$	$(-1,1)$

.

Figure 3. (Color online) Development of the Variables t, $y^1_{\omega_1 1}$, and $y^1_{\omega_1 2}$ Along the Homotopy Path



The transition probability matrix in state ω_1 is given by

$\pi((\omega_1,\omega_2,\omega_3,\omega_4) \mid \omega_1)$	$s^2_{\omega_1 1}$	$s^2_{\omega_1 2}$
$s^1_{\omega_1 1}$	$(1,0,0,0)$	$(0,0,1,0)$
$s^1_{\omega_1 2}$	$(0,0,1,0)$	$(0,1,0,0)$

,

and the transition probability matrix in state ω_2 is given by

$\pi((\omega_1,\omega_2,\omega_3,\omega_4) \mid \omega_2)$	$s^2_{\omega_2 1}$	$s^2_{\omega_2 2}$
$s^1_{\omega_2 1}$	$(1,0,0,0)$	$(0,0,0,1)$
$s^1_{\omega_2 2}$	$(0,0,0,1)$	$(0,1,0,0)$

.

The states ω_3 and ω_4 are absorbing. Starting from the point $(-0.2187, -0.3759, -0.0116, -0.8479, -0.2900, -0.2958, -0.4846, -0.1430, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1.0000)$, IPM obtains the SSPE $((x^1_{\omega_1 1}, x^2_{\omega_1 1}), (x^1_{\omega_2 1}, x^2_{\omega_2 1}), (x^1_{\omega_3 1}, x^2_{\omega_3 1}), (x^1_{\omega_4 1}, x^2_{\omega_4 1})) = ((0.86, 0.86), (0.14, 0.14), (1,1), (1,1))$.

Figure 4 shows the development of the variables y and t in the various iterations of IPM. The downward sloping curve corresponds to t, the nonmonotonic curve to $y^1_{\omega_1 1}$, and the upward sloping curve to $y^1_{\omega_1 2}$.

Example 3. Assume $N = \{1,2\}$, $\Omega = \{\omega_1, \omega_2, \omega_3\}$, and, for $i = 1,2$, $S^i_{\omega_1} = \{s^i_{\omega_1 1}, s^i_{\omega_1 2}\}$, $S^i_{\omega_2} = \{s^i_{\omega_2 1}\}$, and $S^i_{\omega_3} = \{s^i_{\omega_3 1}\}$. The payoff matrices are given by

ω_1	$s^2_{\omega_1 1}$	$s^2_{\omega_1 2}$
$s^1_{\omega_1 1}$	$(1,-1)$	$(0,0)$
$s^1_{\omega_1 2}$	$(0,0)$	$(1,-1)$

,

ω_2	$s^2_{\omega_2 1}$
$s^1_{\omega_2 1}$	$(0,0)$

, and

ω_3	$s^2_{\omega_3 1}$
$s^1_{\omega_3 1}$	$(1,-1)$

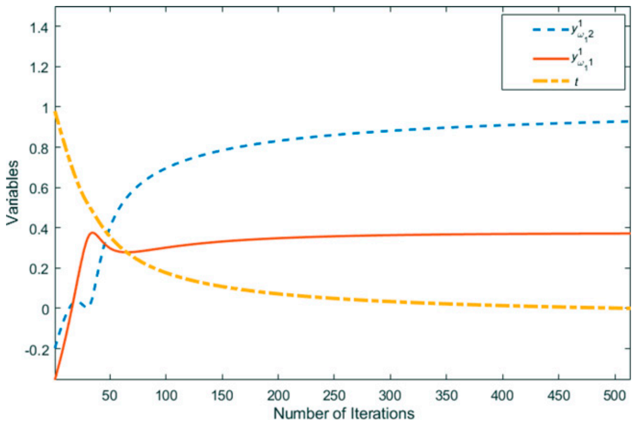
.

The transition probability matrix in state ω_1 is given by

$\pi((\omega_1,\omega_2,\omega_3) \mid \omega_1)$	$s^2_{\omega_1 1}$	$s^2_{\omega_1 2}$
$s^1_{\omega_1 1}$	$(1,0,0)$	$(0,1,0)$
$s^1_{\omega_1 2}$	$(1,0,0)$	$(0,0,1)$

.

Figure 4. (Color online) Development of the Variables t, $y^1_{\omega_1 1}$, and $y^1_{\omega_1 2}$ Along the Homotopy Path



The states ω_2 and ω_3 are absorbing. Starting from the point (−0.2429, −0.3467, −0.3788, −0.2164, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1.0000), IPM generates the SSPE $((x^1_{\omega_1 1}, x^2_{\omega_1 1}), (x^1_{\omega_2 1}, x^2_{\omega_2 1}), (x^1_{\omega_3 1}, x^2_{\omega_3 1})) = ((1, 0.5), (1, 1), (1, 1))$.

Example 4. Assume $N = \{1,2\}$, $\Omega = \{\omega_1, \omega_2, \omega_3\}$, and, for $i = 1,2$, $S^i_{\omega_1} = \{s^i_{\omega_1 1}, s^i_{\omega_1 2}\}$, $S^i_{\omega_2} = \{s^i_{\omega_2 1}\}$, and $S^i_{\omega_3} = \{s^i_{\omega_3 1}\}$. The payoff matrices are given by

$$\begin{array}{c|cc} \omega_1 & s^2_{\omega_1 1} & s^2_{\omega_1 2} \\ \hline s^1_{\omega_1 1} & (1,0) & (0,1) \\ s^1_{\omega_1 2} & (0,2) & (1,0) \end{array}, \quad \begin{array}{c|c} \omega_2 & s^2_{\omega_2 1} \\ \hline s^1_{\omega_2 1} & (0,2) \end{array}, \quad \text{and} \quad \begin{array}{c|c} \omega_3 & s^2_{\omega_3 1} \\ \hline s^1_{\omega_3 1} & (1,0) \end{array}.$$

The transition probability matrix in state ω_1 is given by

$$\begin{array}{c|cc} \pi((\omega_1, \omega_2, \omega_3) \mid \omega_1) & s^2_{\omega_1 1} & s^2_{\omega_1 2} \\ \hline s^1_{\omega_1 1} & (1,0,0) & (1,0,0) \\ s^1_{\omega_1 2} & (0,1,0) & (0,0,1) \end{array}.$$

The states ω_2 and ω_3 are absorbing. By choosing the starting point (−0.2094, −0.3876, −0.3135, −0.2728, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1.0000), IPM finds the SSPE $((x^1_{\omega_1 1}, x^2_{\omega_1 1}), (x^1_{\omega_2 1}, x^2_{\omega_2 1}), (x^1_{\omega_3 1}, x^2_{\omega_3 1})) = ((1, 0.5), (1, 1), (1, 1))$.

Example 5. Assume $N = \{1,2\}$, $\Omega = \{\omega_1, \omega_2\}$, and, for $i = 1,2$, $S^i_{\omega_1} = \{s^i_{\omega_1 1}, s^i_{\omega_1 2}, s^i_{\omega_1 3}\}$ and $S^i_{\omega_2} = \{s^i_{\omega_2 1}\}$. The payoff matrices are given by

$$\begin{array}{c|ccc} \omega_1 & s^2_{\omega_1 1} & s^2_{\omega_1 2} & s^2_{\omega_1 3} \\ \hline s^1_{\omega_1 1} & (1,1) & (0,0) & (-9,-9) \\ s^1_{\omega_1 2} & (0,0) & (0,0) & (-7,-7) \\ s^1_{\omega_1 3} & (-9,-9) & (-7,-7) & (-7,-7) \end{array} \quad \text{and} \quad \begin{array}{c|c} \omega_2 & s^2_{\omega_2 1} \\ \hline s^1_{\omega_2 1} & (0,0) \end{array}.$$

Irrespective of the actions chosen, the transition probability between any two states is equal to 1/2. From the starting point (−0.5645, −0.6302, −0.1793, −0.6422, −0.3859, −0.2966, 0, 0, 0, 0, 0, 0, 1.0000), IPM finds the SSPE $(((x^1_{\omega_1 1}, x^1_{\omega_1 2}, x^1_{\omega_1 3}), (x^2_{\omega_1 1}, x^2_{\omega_1 2}, x^2_{\omega_1 3})), (x^1_{\omega_2 1}, x^2_{\omega_2 1})) = (((0,1,0), (0,1,0)), (1,1))$.

4.2. Randomly Generated Cases

In this subsection, we randomly generate stochastic games for varying n, d, and m, where m denotes the number of actions for each player in each state. Payoffs are randomly drawn from the interval [−10, 10], but are set equal to 0 with probability "pd0," which indicates the probability density of the payoff matrix. Obviously, the larger the pd0, the sparser the payoff matrix becomes. To test the efficiency and stability of the proposed IPM, we will compare the numerical performance of IPM with that of SLTP and the PATH solver, which was established by Dirkse and Ferris (1995).

4.2.1. Comparisons with SLTP. We let pd0 be equal to 0.00, 0.25, 0.50, and 0.75 to induce several groups of different games and test the performance of IPM and SLTP. Note that the range of the homotopy variable t is [0,1] in IPM and [0,2] in SLTP. To make a fair comparison between the two methods, we also report the computational cost for $t \in [0,1]$ in SLTP. The discount factor is taken equal to $\delta = 0.95$. The average computation time (in seconds) (ATIME) and average number of iterations (AITER) for the two methods are reported in Tables 1 and 2, respectively, where T-ratio $= \frac{\text{ATIME of IPM}}{\text{ATIME of SLTP}}$, I-ratio $= \frac{\text{AITER of IPM}}{\text{AITER of SLTP}}$, SLTP-tratio $= \frac{\text{ATIME } (t \le 1) \text{ of SLTP}}{\text{ATIME of SLTP}}$, and SLTP-iratio $= \frac{\text{AITER}(t \le 1) \text{ of SLTP}}{\text{AITER of SLTP}}$.

From the results in Tables 1 and 2, for any given n, d, and m, it follows by inspecting each group of rows that a smaller pd0 leads to a more difficult problem. The results for numerous randomly generated stochastic games illustrate that IPM performs much better than SLTP in both the number of iterations and computation time. This advantage becomes more pronounced as the scale of the problem gets larger. Besides this, the columns of "SLTP-tratio" and "SLTP-iratio" illustrate that the computational cost for $t \in [1,2]$ in SLTP only takes up a very small proportion in the total computational cost. That is, in our experiments, the computational cost for finding the starting point of the standard SLTP is almost ignorable. Therefore, although the range of t in the two methods is not the same, this difference causes no problem for a fair comparison of the two methods.

4.2.2. Comparisons with the PATH Solver. Recall that stochastic games can be reformulated as MCPs. The PATH solver is an extensively used software package to solve MCPs in industry and academy settings; see Dirkse and Ferris (1995). Nevertheless, as shown in Ferris et al. (2000), the PATH solver is very sensitive

Table 1. Numerical Performance and Comparisons

(n, d, m)	pd0	IPM ATIME	SLTP ATIME	SLTP ATIME($t \leq 1$)	SLTP SLTP-tratio (%)	T-ratio (%)
(2,2,5)	0.00	25.37	84.87	74.00	87.19	29.89
	0.25	17.28	62.98	51.83	82.30	27.44
	0.50	13.05	44.20	37.08	83.89	29.51
	0.75	12.94	47.66	40.18	84.31	27.15
(2,5,3)	0.00	64.59	216.41	181.65	83.97	29.84
	0.25	63.59	221.15	178.57	80.74	28.75
	0.50	43.28	166.28	136.94	82.35	26.01
	0.75	26.72	106.41	86.47	81.26	25.11
(2,5,4)	0.00	91.27	320.56	270.23	84.30	28.47
	0.25	92.06	351.45	304.49	86.63	26.19
	0.50	38.62	155.60	124.76	80.18	24.81
	0.75	31.60	125.02	103.53	82.81	25.28
(2,5,5)	0.00	122.41	715.40	654.54	91.49	17.11
	0.25	85.50	412.11	361.38	87.69	20.74
	0.50	70.23	362.89	321.18	88.51	19.35
	0.75	68.73	389.65	358.15	91.92	17.64
(3,3,5)	0.00	75.80	580.83	548.76	94.48	13.05
	0.25	62.46	496.86	471.02	94.80	12.57
	0.50	48.24	286.69	266.45	92.94	16.82
	0.75	35.31	237.67	221.82	93.33	14.86
(4,2,5)	0.00	64.72	471.89	449.32	95.22	13.72
	0.25	31.28	330.65	310.25	93.83	9.46
	0.50	19.38	201.66	189.99	94.21	9.61
	0.75	15.87	170.88	159.21	93.30	9.29
(5,2,5)	0.00	85.23	887.29	848.85	95.69	9.61
	0.25	59.05	686.32	650.41	94.77	8.06
	0.50	58.35	551.41	522.77	94.80	10.58
	0.75	34.46	438.17	412.46	94.13	7.41

to the structure of the original problems and may lose effectiveness for game-theoretic problems. In the rest of this subsection, we choose $(n,m,d) = (2,2,5)$, $(n,m,d) = (3,2,5)$, $(n,m,d) = (4,2,5)$, and $(n,m,d) = (5,2,5)$ as 4 parameter constellations and randomly generate 10 stochastic games. Then we use the proposed IPM and the PATH solver to solve these stochastic games and record the feasibility and computation time of the two methods in Table 3, where "time1" is the computation time of IPM, "time2" is the computation time of the PATH solver, and "fail" indicates that the method fails to find an SSPE for the stochastic game. Moreover, pd0 is set to be 0.5.

From Table 3, we find that the PATH solver is really very efficient in some cases and even performs much better than the proposed IPM in computation time. However, as shown in the table, the PATH solver sometimes may fail to find a solution, and the success rate of the PATH solver seems to decrease when the scale of the stochastic game increases. For instance, among the entire 10 stochastic games with $(n,d,m) = (2,2,5)$, the PATH solver achieves 9 successes, whereas the number of successes becomes only 4 for the case in which $(n,d,m) = (5,2,5)$. To further study the success rates of the two methods, we randomly generated 100 stochastic games for different triples of (n, d, m) and record the success rates of the two methods for each triple. The developments of the success rates in n, d, and m are plotted in Figure 5.

Figure 5 shows that the success rate of the PATH solver decreases when n, d, and m increase. From the three subfigures in Figure 5, it is easy to see that the success rate is more sensitive to the change of d than to the changes of the other two factors. When the scale of the stochastic game becomes larger, the effectiveness of the PATH solver to find an SSPE becomes even lower, whereas IPM always achieves a 100% success rate for any stochastic game. As a result, IPM is more stable and robust to compute SSPEs in stochastic games.

4.3. More Complicated Random Cases

It is well known that the structure of stochastic games is very complicated, and the scale of the problem is very sensitive to the number of players n, the number of states d, and the number of actions m. For example, for the case of $(d,m) = (5,5)$, when n increases from 5 to 6, the number of variables in the homotopy system increases from 150 to 180. Additionally, an increase of n, d, and m results in a data explosion in stochastic games. For instance, in any state ω, when there are n players and m actions, the number of utilities for each

Table 2. Numerical Performance and Comparisons

(n, d, m)	pd0	IPM AITER	SLTP AITER	SLTP AITER($t \leq 1$)	SLTP SLTP-iratio (%)	I-ratio (%)
(2,2,5)	0.00	1,987	4,188	3,722	88.87	47.44
	0.25	1,461	4,287	3,545	82.70	34.08
	0.50	1,363	4,147	3,498	84.35	32.86
	0.75	1,317	3,700	3,190	86.21	35.59
(2,5,3)	0.00	3,741	6,541	5,528	84.51	57.19
	0.25	2,632	6,258	5,116	81.75	41.91
	0.50	1,691	4,803	4,026	83.82	35.21
	0.75	1,400	4,839	4,001	82.69	28.93
(2,5,4)	0.00	3,759	8,599	7,343	85.39	43.71
	0.25	3,016	8,470	7,433	87.76	35.61
	0.50	2,430	8,324	6,782	81.47	29.19
	0.75	1,752	6,513	5,523	84.81	26.90
(2,5,5)	0.00	3,840	9,776	9,019	92.26	39.28
	0.25	2,751	7,465	6,594	88.34	32.22
	0.50	2,544	7,873	7,065	89.74	32.31
	0.75	1,403	5,285	4,900	92.73	26.54
(3,3,5)	0.00	2,566	8,309	7,902	95.10	30.88
	0.25	2,096	7,723	7,377	95.53	27.14
	0.50	1,512	5,567	5,235	94.04	27.16
	0.75	1,399	5,627	5,308	94.33	24.86
(4,2,5)	0.00	1,256	6,548	6,316	96.45	19.18
	0.25	1,075	6,210	5,983	96.34	17.31
	0.50	941	4,981	4,734	95.04	18.89
	0.75	893	4,890	4,625	94.58	18.26
(5,2,5)	0.00	574	4,369	4,225	96.70	13.12
	0.25	538	3,543	3,535	99.77	15.18
	0.50	556	3,797	3,660	96.39	14.64
	0.75	416	3,018	2,881	95.66	13.82

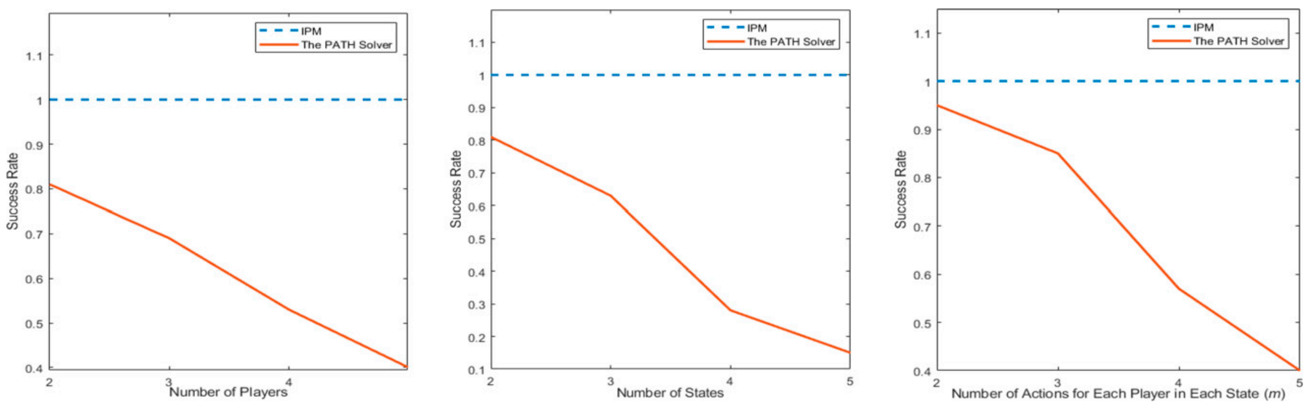
player is m^n in this state, which becomes huge when n is large. That is why we consider stochastic games with a limited number of players, states, and actions. With the existing methods such as SLTP proposed in Herings and Peeters (2004) and Li and Dang (2020), SSPEs can only be computed for stochastic games with a scale no more than $n = 5$, $d = 5$, and $m = 5$—that is, the number of utilities for each player in each state is less than 5^5. In this subsection, we use the proposed IPM to solve problems with a scale up to (5, 8, 8), where the number of utilities for each player in each state reaches 8^5. In each experiment, we let pd0 = 0.95 and generate the payoff matrices and transition probabilities randomly. The average computation time (in seconds) for all experiments is recorded in Table 4. It is left blank if the computation time exceeds 3×10^4 seconds.

It follows from Table 4 that the computation of an SSPE becomes much more difficult as the number of players n, states d, or actions for each player in each state m increases by a single unit. The variable n is the most influential factor for the number of iterations

Table 3. Numerical Comparisons

$(n,d,m)=(2,2,5)$ Problem	time1	time2	$(n,d,m)=(3,2,5)$ Problem	time1	time2	$(n,d,m)=(4,2,5)$ Problem	time1	time2	$(n,d,m)=(5,2,5)$ Problem	time1	time2
1	6.71	0.82	1	10.42	1.04	1	42.04	7.01	1	61.38	fail
2	3.87	0.78	2	18.84	fail	2	30.78	3.26	2	67.92	16.10
3	5.37	0.55	3	14.64	1.5	3	24.25	3.23	3	39.63	fail
4	11.11	2.38	4	26.43	2.96	4	12.70	fail	4	50.00	fail
5	10.01	0.79	5	12.71	1.29	5	16.18	5.02	5	42.89	10.55
6	22.76	fail	6	30.11	fail	6	35.48	fail	6	39.63	fail
7	5.94	1.14	7	11.76	1.42	7	25.78	3.96	7	68.06	40.72
8	3.77	0.29	8	8.04	2.16	8	11.57	fail	8	43.06	14.84
9	11.6	1.33	9	10.96	1.95	9	23.81	3.39	9	23.74	fail
10	4.06	0.34	10	11.94	2.81	10	11.01	fail	10	44.84	fail

Figure 5. (Color online) The Developments of the Success Rates in Three Factors



and computation time, which is consistent with the observations in Herings and Peeters (2004).

As previously mentioned, stochastic games with (n, d, m) greater than (5, 8, 8) result in a data explosion and are not considered in this paper. Several decomposition strategies were developed in the literature to cope with these issues in general equilibrium models and MCPs. Cabero et al. (2005) applied Bender's decomposition to address an electricity market equilibrium problem. A Gauss–Seidel decomposition scheme was presented in Ban et al. (2006) to convert a continuous network design problem with equilibrium constraints into multiple smaller-dimensional subproblems. Gabriel and Fuller (2010) designed a new Bender's decomposition approach to solve more general stochastic complementarity problems. A matrix-splitting decomposition method was proposed in Shanbhag et al. (2011) to deal with stochastic MCPs. Egging (2013) forms a decomposition-type algorithm to solve large-scale MCPs derived from multistage natural gas markets with market power exertion. Recall that a stochastic game can also be solved as an MCP. These papers suggest that the application of a well-designed decomposition strategy could reduce the computational cost for finding an SSPE and address larger-scale stochastic games. We leave the exploration of such an approach as an interesting topic for future research.

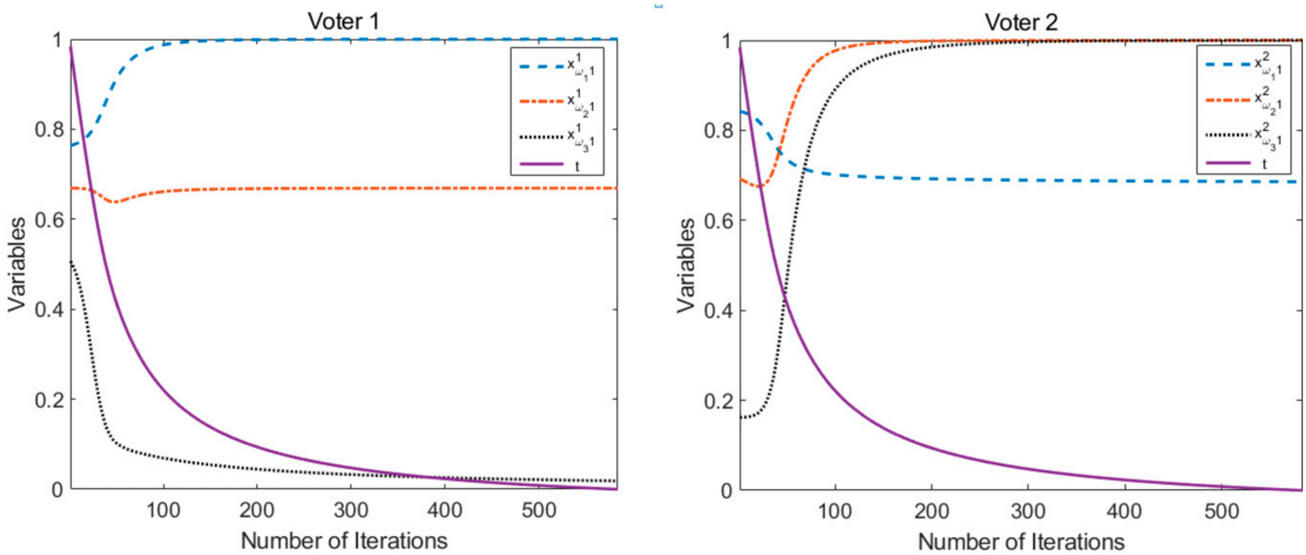
4.4. An Application in Bargaining Models

Stochastic games have many applications to various fields, where bargaining is one of the most popular topics and has been extensively studied in the past two decades; see Banks and Duggan (2006), Britz et al. (2010), Britz (2018), and the references therein. In this subsection, we study a legislative bargaining model based on a stochastic game paradigm. Suppose there are two voters who play an infinite-horizon bargaining game over three policies, p_1, p_2, and p_3. In any stage t, one of the policies, p_j, $j \in \{1,2,3\}$, will be

Table 4. Average Computational Time (in Seconds) of IPM

	m			
	5	6	7	8
d	$n = 3$			
5	230.90	358.78	747.82	1,108.71
6	602.41	936.39	2,162.49	3,291.86
7	828.61	1,065.73	2,303.06	4,137.06
8	1,430.17	2,146.85	3,021.24	3,117.04
	$n = 4$			
5	539.80	1,021.91	1,203.28	2,255.70
6	1,328.68	1,554.25	2,536.31	4,761.73
7	1,440.40	2,230.62	2,413.82	6,202.49
8	1,875.36	3,681.46	4,200.69	8,576.12
	$n = 5$			
5	1,737.74	3,377.21	10,403.29	27,626.37
6	2,897.06	10,249.51	24,284.54	
7	4,378.30	17,669.26		
8	10,247.18			

Figure 6. (Color online) Development of the Variables t, $x^i_{\omega_j 1}$ $(i = 1,2; j = 1,2,3)$ Along the Homotopy Path



proposed, and the voters simultaneously and independently make an accept/reject decision. The voting rule is that if at least one of the two voters chooses to accept, then p_j is accepted, and the two voters get their payoffs based on the outcome in the current stage. Otherwise, the preceding mentioned procedure will be repeated in the next stage $t + 1$. We say that the bargaining ends as soon as a policy is accepted, and in all the subsequent stages, the payoffs of the voters will be the same as those in the last stage. This bargaining model is essentially a stochastic game with two players, six states, and two actions for each player in each state. More specifically, $N = \{1,2\}$ and $\Omega = \{\omega_1, \omega_2, \ldots, \omega_6\}$, where $\omega_1 = \{p_1\,\textit{is proposed}\}$, $\omega_2 = \{p_2\,\textit{is proposed}\}$, and $\omega_3 = \{p_3\,\textit{is proposed}\}$. The states ω_4, ω_5, and ω_6 correspond to states in which the bargaining has ended, where $\omega_4 = \{p_1\ \textit{has been accepted}\}$, $\omega_5 = \{p_2\ \textit{has been accepted}\}$, and $\omega_6 = \{p_3\ \textit{has been accepted}\}$. In ω_1, ω_2, and ω_3, the voters have two actions: that is, $s^i_{\omega_j 1} = \textit{accept}$ or $s^i_{\omega_j 2} = \textit{reject}$ with $i \in N$ and $j \in \{1,2,3\}$. The utility matrices are given by

$$
\begin{array}{c|cc}
\omega_1 & s^2_{\omega_1 1} & s^2_{\omega_1 2} \\ \hline
s^1_{\omega_1 1} & (1,-1) & (1,-1) \\
s^1_{\omega_1 2} & (1,-1) & (0,0)
\end{array},\quad
\begin{array}{c|cc}
\omega_2 & s^2_{\omega_2 1} & s^2_{\omega_2 2} \\ \hline
s^1_{\omega_2 1} & (0,1) & (0,1) \\
s^1_{\omega_2 2} & (0,1) & (0,0)
\end{array},
$$

$$
\begin{array}{c|cc}
\omega_3 & s^2_{\omega_3 1} & s^2_{\omega_3 2} \\ \hline
s^1_{\omega_3 1} & (-1,0) & (-1,0) \\
s^1_{\omega_3 2} & (-1,0) & (0,0)
\end{array},\quad
\begin{array}{c|c}
\omega_4 & s^2_{\omega_4 1} \\ \hline
s^1_{\omega_4 1} & (1,-1)
\end{array},
$$

$$
\begin{array}{c|c}
\omega_5 & s^2_{\omega_5 1} \\ \hline
s^1_{\omega_5 1} & (0,1)
\end{array},\quad
\begin{array}{c|c}
\omega_6 & s^2_{\omega_6 1} \\ \hline
s^1_{\omega_6 1} & (-1,0)
\end{array}.
$$

If the current state is ω_1 and both voters reject the proposed policy, then the whole system will jump to ω_1, ω_2, and ω_3 with probability 1/3. Otherwise, the system will jump to state ω_4 with probability 1. If the current state is ω_2 and both voters reject the proposed policy, then the system will jump to ω_1, ω_2, and ω_3 with probability 1/3. Otherwise, the system will jump to state ω_5 with probability 1. If the current state is ω_3 and both voters reject the proposed policy, then the whole system will jump to ω_1, ω_2, and ω_3 with probability 1/3. Otherwise, the system will jump to state ω_6 with probability 1. The states ω_4, ω_5, and ω_6 are absorbing; that is, once the system reaches one of these three states, it will never leave it.

Let $\mathbf{0}$ denote a zero vector with dimension 12. From the randomly generated starting point, $(y^0, \mathbf{0}^\top, 1)^\top$ with $y^0 = (-0.1263,\ -0.5135,\ -0.0818,\ -0.6038,\ -0.1821,\ -0.4246,\ -0.1671,\ -0.4466,\ -0.2857,\ -0.3002,\ -0.5975,\ -0.0846, 0, 0, 0, 0, 0, 0)$, IPM eventually ends at a solution to this problem after 584 iterations and 21.48 s. The development of t and the mixed strategies in the first three states for the two voters in the various iterations are described in Figure 6.

5. Conclusions

In this paper, we extend the idea of interior-point methods, which have been proven to be very efficient for large-scale convex programming problems, to computing an SSPE in a finite discounted stochastic game. The basic idea of our method is to incorporate a logarithmic barrier term into the objective function of the stochastic game of interest and formulate an artificial barrier game. The solutions to the artificial game at different levels of the homotopy variable generate an everywhere smooth path. As the homotopy variable descends to 0, our path converges to an SSPE for the stochastic game of interest.

The proposed method (IPM) has several advantages over the existing alternative methods in the literature to compute an SSPE. The starting point of IPM can be arbitrarily chosen, and there is no need to solve an optimization problem to obtain it. The barrier function forces the homotopy path to stay in the interior of the strategy space, which avoids switching between different systems of equations or, alternatively, a computationally expensive transformation of variables. IPM fully exploits the differentiability of the problem, and for every stochastic game, the induced homotopy path is everywhere smooth. Moreover, IPM is able to find SSPEs for all stochastic games. The effectiveness and efficiency of IPM has been confirmed by numerous numerical experiments.

Endnotes

[1] For each player $i \in N$, one can add a constant c to the stage payoffs u^i such that they are all nonnegative. Then one obtains a new sequence $\{(\mu^i_\omega(k))'\}$ satisfying $(\mu^i_\omega(k+1))' \geq (\mu^i_\omega(k))'$. Subtracting $\frac{c}{1-\delta}$ from the limit of the new sequence, one gets the total expected payoffs μ^i_ω.

[2] It seems impossible to analyze the rate of convergence for the proposed IPM, because the problem of computing stationary equilibria for stochastic games is actually a PPAD-complete problem. As $t \in (0,1)$, the development of the homotopy path derived from our method very much relies on the structure of the stochastic game. As a result, for the proposed IPM, the computational complexity can only be obtained in each iteration, but the total number of iterations could not be theoretically attained.

[3] For simplicity, we refer to the arbitrary starting SLTP as SLTP in the remainder of this section.

References

Allgower EL, Georg K (2012) *Numerical Continuation Methods: An Introduction*, Springer Series in Computational Mathematics, vol. 13 (Springer-Verlag, Berlin).

Ban X, Liu HX, Lu J, Ferris MC (2006) Decomposition scheme for continuous network design problem with asymmetric user equilibria. *Transportation Res. Record* 1964(1):185–192.

Banks JS, Duggan J (2000) A bargaining model of collective choice. *Amer. Political Sci. Rev.* 94(1):73–88.

Banks JS, Duggan J (2006) A bargaining model of legislative policymaking. *J. Political Sci.* 1(1):49–85.

Britz V (2018) Rent-seeking and surplus destruction in unanimity bargaining. *Games Econom. Behav.* 109(C):1–20.

Britz V, Herings PJJ, Predtetchinski A (2010) Non-cooperative support for the asymmetric Nash bargaining solution. *J. Econom. Theory* 145(5):1951–1967.

Cabero J, Baillo Á, Cerisola S, Ventosa M (2005) Application of Benders decomposition to an equilibrium problem. *15th Power Systems Comput. Conf.*, Liege, August-2005, vol. 1 (Curran Associates, Red Hook, NY), 107–113.

Cachon GP, Netessine S (2004) Game theory in supply chain analysis. Simchi-Levi D, Wu SD, Shen Z-JM, eds. *Handbook of Quantitative Supply Chain Analysis: Modeling in the E-Business Era* (Springer, Boston), 13–65.

Chatterjee K, Dutta B, Ray D, Sengupta K (1993) A noncooperative theory of coalitional bargaining. *Rev. Econom. Stud.* 60(2): 463–477.

Chen Y, Dang C (2016) A reformulation-based smooth path-following method for computing Nash equilibria. *Econom. Theory Bull.* 4(2): 231–246.

Chen Y, Dang C (2021) A differentiable homotopy method to compute perfect equilibria. *Math. Programming* 185(1):77–109.

Dang C, Herings PJJ, Li P (2021) Data for an interior-point differentiable path-following method to compute stationary equilibria in stochastic games. http://dx.doi.org/10.5281/zenodo.5381196, https://github.com/INFORMSJoC/2020.0259.

Dang C, Ye Y, Zhu Z (2011) An interior-point path-following algorithm for computing a Leontief economy equilibrium. *Comput. Optim. Appl.* 50(2):223–236.

Den Elzen AV, Der Laan GV, Talman D (1994) An adjustment process for an economy with linear production technologies. *Math. Oper. Res.* 19(2):341–351.

Dirkse SP, Ferris MC (1995) The PATH solver: A nommonotone stabilization scheme for mixed complementarity problems. *Optim. Methods Software* 5(2):123–156.

Eaves BC (1972) Homotopies for computation of fixed points. *Math. Programming* 3(1):1–22.

Eaves BC, Schmedders K (1999) General equilibrium models and homotopy methods. *J. Econom. Dynam. Control* 23(9–10):1249–1279.

Egging R (2013) Benders decomposition for multi-stage stochastic mixed complementarity problems—Applied to a global natural gas market model. *Eur. J. Oper. Res.* 226(2):341–353.

Eibelshäuser S, Poensgen D (2019) Markov quantal response equilibrium and a homotopy method for computing and selecting Markov perfect equilibria of dynamic stochastic games. Preprint, submitted January 21, http://dx.doi.org/10.2139/ssrn.3314404.

Ericson R, Pakes A (1995) Markov-perfect industry dynamics: A framework for empirical work. *Rev. Econom. Stud.* 62(1):53–82.

Ferris MC, Mangasarian OL, Pang J-S, eds. (2013) *Complementarity: Applications, Algorithms and Extensions*, Applied Optimization, vol. 50 (Springer Science & Business Media, Dordrecht, Netherlands).

Ferris MC, Munson TS, Ralph D (2000) A homotopy method for mixed complementarity problems based on the PATH solver. Griffiths DF, Watson GA, eds. *Numerical Analysis 1999*, Chapman & Hall/CRC Research Notes in Mathematics, vol. 420 (CRC Press, Boca Raton, FL), 143–168.

Fiacco AV, McCormick GP (1968) Nonlinear programming: Sequential unconstrained minimization techniques. Technical report, Research Analysis Corporation, McLean, VA.

Fiacco AV, McCormick GP (1990) *Nonlinear Programming: Sequential Unconstrained Minimization Techniques*, Classics in Applied Mathematics (Society for Industrial and Applied Mathematics, Philadelphia).

Fink AM (1964) Equilibrium in a stochastic *n*-person game. *J. Sci. Hiroshima Univ. Ser. A-I* 28(1):89–93.

Fudenberg D, Tirole J (1991) *Game Theory* (MIT Press, Cambridge, MA).

Gabriel SA, Fuller J (2010) A benders decomposition method for solving stochastic complementarity problems with an application in energy. *Comput. Econom.* 35(4):301–329.

Govindan S, Wilson R (2009) Global Newton method for stochastic games. *J. Econom. Theory* 144(1):414–421.

Gutierrez AE, Mazorche SR, Herskovits J, Chapiro G (2017) An interior point algorithm for mixed complementarity nonlinear problems. *J. Optim. Theory Appl.* 175(2):432–449.

Harsanyi JC, Selten R (1988) *A General Theory of Equilibrium Selection in Games* (MIT Press, Cambridge, MA).

He W, Sun Y (2017) Stationary Markov perfect equilibria in discounted stochastic games. *J. Econom. Theory* 169(May):35–61.

Herings PJJ, Peeters RJ (2001) A differentiable homotopy to compute Nash equilibria of *n*-person games. *Econom. Theory* 18(1):159–185.

Herings PJJ, Peeters RJ (2004) Stationary equilibria in stochastic games: Structure, selection, and computation. *J. Econom. Theory* 118(1):32–60.

Herings PJJ, Peeters RJ (2010) Homotopy methods to compute equilibria in game theory. *Econom. Theory* 42(1):119–156.

Herings PJJ, Schmedders K (2006) Computing equilibria in finance economies with incomplete markets and transaction costs. *Econom. Theory* 27(3):493–512.

Huang KL, Mehrotra S (2017) Solution of monotone complementarity and general convex programming problems using a modified potential reduction interior point method. *INFORMS J. Comput.* 29(1):36–53.

Jaśkiewicz A, Nowak AS (2018) Non-zero-sum stochastic games. Başar T, Zaccour G, eds. *Handbook of Dynamic Game Theory* (Springer, Cham, Switzerland), 281–344.

Karatzas I, Sudderth SWD (1994) Construction of stationary Markov equilibria in a strategic market game. *Math. Oper. Res.* 19(4): 975–1006.

Li P, Dang C (2020) An arbitrary starting tracing procedure for computing subgame perfect equilibria. *J. Optim. Theory Appl.* 186(2): 667–687.

Maitra AP, Sudderth WD (1998) Subgame-perfect equilibria for stochastic games. *Math. Oper. Res.* 32(3):711–722.

Mas-Colell A (1974) A note on a theorem of F. Browder. *Math. Programming* 6(1):229–233.

McKelvey RD, McLennan A (1996) Computation of equilibria in finite games. Anman AH, Kendrick DA, Rust J, eds. *Handbook of Computational Economics*, vol. 1 (North-Holland, Amsterdam), 87–142.

Munson TS, Facchinei F, Ferris MC, Fischer A, Kanzow C (2001) The semismooth algorithm for large scale complementarity problems. *INFORMS J. Comput.* 13(4):294–311.

Murphy F, Pierru A, Smeers Y (2016) A tutorial on building policy models as mixed-complementarity problems. *Interfaces* 46(6):465–481.

Nowak AS, Raghavan TES (1992) Existence of stationary correlated equilibria with symmetric information for discounted stochastic games. *Math. Oper. Res.* 17(3):519–526.

Olsen TL, Parker RP (2014) On Markov equilibria in dynamic inventory competition. *Oper. Res.* 62(2):332–344.

Pakes A, McGuire P (2001) Stochastic algorithms, symmetric Markov perfect equilibrium, and the "curse" of dimensionality. *Econometrica* 69(5):1261–1281.

Saigal R (1983) A homotopy for solving large, sparse and structured fixed point problems. *Math. Oper. Res.* 8(4):557–578.

Scarf H (1967) The approximation of fixed points of a continuous mapping. *SIAM J. Appl. Math.* 15(5):1328–1343.

Schanuel SH, Simon LK, Zame WR (1991) The algebraic geometry of games and the tracing procedure. Selten R, ed. *Game Equilibrium Models* II (Springer, Berlin), 9–43.

Shanbhag UV, Infanger G, Glynn PW (2011) A complementarity framework for forward contracting under uncertainty. *Oper. Res.* 59(4): 810–834.

Shapley L (1953) Stochastic games. *Proc. Natl. Acad. Sci. USA* 39(10): 1095–1100.

Simantiraki EM, Shanno DF (1997) An infeasible-interior-point method for linear complementarity problems. *SIAM J. Optim.* 7(3):620–640.

Sobel MJ (1971) Noncooperative stochastic games. *Ann. Math. Statist.* 42(6):1930–1935.

Solan E (1998) Discounted stochastic games. *Math. Oper. Res.* 23(4): 1010–1021.

Ye Y (1991) An $o(n^3 l)$ potential reduction algorithm for linear programming. *Math. Programming* 50(1–3):239–258.

Ye Y (1992) A potential reduction algorithm allowing column generation. *SIAM J. Optim.* 2(1):7–20.

Ye Y (2011) *Interior Point Algorithms: Theory and Analysis* (John Wiley & Sons, New York).

Zhan Y, Dang C (2018) A smooth path-following algorithm for market equilibrium under a class of piecewise-smooth concave utilities. *Comput. Optim. Appl.* 71(2):381–402.

Zhan Y, Li P, Dang C (2020) A differentiable path-following algorithm for computing perfect stationary points. *Comput. Optim. Appl.* 76(2):571–588.

Zhao YB, Li D (2001) On a new homotopy continuation trajectory for nonlinear complementarity problems. *Math. Oper. Res.* 26(1): 119–146.

Zhu Z, Dang C, Ye Y (2012) A FPTAS for computing a symmetric Leontief competitive economy equilibrium. *Math. Programming* 131(1–2):113–129.