

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

Citation for published version (APA):

Dang, C., Herings, P. J.-J., & Li, P. (2020). *An Interior-Point Path-Following Method to Compute Stationary Equilibria in Stochastic Games*. Maastricht University, Graduate School of Business and Economics. GSBE Research Memoranda No. 001 <https://doi.org/10.26481/umagsb.20001>

## Document status and date:

Published: 17/02/2020

## DOI:

[10.26481/umagsb.20001](https://doi.org/10.26481/umagsb.20001)

## Document Version:

Publisher's PDF, also known as Version of record

## Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

## General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

[www.umlib.nl/taverne-license](http://www.umlib.nl/taverne-license)

## Take down policy

If you believe that this document breaches copyright please contact us at:

[repository@maastrichtuniversity.nl](mailto:repository@maastrichtuniversity.nl)

providing details and we will investigate your claim.

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

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

RM/20/001

**GSBE**

Maastricht University School of Business and Economics  
Graduate School of Business and Economics

PO Box 616  
NL- 6200 MD Maastricht  
The Netherlands

# An Interior-Point Path-Following Method to Compute Stationary Equilibria in Stochastic Games\*

Chuangyin Dang<sup>†</sup> P. Jean-Jacques Herings<sup>‡</sup> Peixuan Li<sup>§</sup>

December 16, 2019

## Abstract

Subgame perfect equilibrium in stationary strategies (SSPE) is the most important solution concept used in applications of stochastic games, which makes it imperative to develop efficient numerical methods to compute an SSPE. For this purpose, this paper develops an interior-point path-following method (IPM), which remedies a number of issues with the existing method called stochastic linear tracing procedure (SLTP). The homotopy system of IPM is derived from the optimality conditions of an artificial barrier game, whose objective function is a combination of the original payoff function and a logarithmic term. Unlike SLTP, the starting stationary strategy profile can be arbitrarily chosen and IPM does not need switching between different systems of equations. The use of a perturbation term makes IPM applicable to all stochastic games, whereas SLTP only works for a generic stochastic game. A transformation of variables reduces the number of equations and variables of by roughly one half. Numerical results show that our method is more than three times as efficient as SLTP.

**Keywords:** Stochastic Games, Subgame Perfect Equilibria, Stationary Strategies, Interior-point Method, Path-following Algorithm

---

\*This work was partially supported by GRF: CityU 11302715 of Hong Kong SAR Government.

<sup>†</sup>Department of Systems Engineering and Engineering Management, City University of Hong Kong. Email: mecdang@cityu.edu.hk

<sup>‡</sup>Department of Economics, Maastricht University. Email: p.herings@maastrichtuniversity.nl

<sup>§</sup>Department of Systems Engineering and Engineering Management, City University of Hong Kong. Email: peixuanli3-c@my.cityu.edu.hk

# 1 Introduction

The concept of stochastic game was introduced by Shapley in 1953 [27]. As a cornerstone in the area of game theory and economics, a stochastic game enriches the model of repeated games and has been applied in many economic situations of interest [2, 3, 10]. A stochastic game 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 get their instantaneous payoffs immediately. Subsequently, each player is 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.

As mentioned in [12], the concept of a subgame perfect equilibrium in stationary strategies (SSPE) is one of the most important concepts in stochastic games. A stationary strategy only depends on the current state rather than the entire history of states and action profiles, which is consistent with the principle that ‘bygones are bygones’ [21]. The existence of stationary equilibria in stochastic games has been discussed extensively in earlier papers, see [14] and [19] for wonderful reviews. [27] proved the existence of stationary equilibria for zero-sum games with finite action and state spaces. [11], [29], and [28] extended Shapley’s model to general  $n$ -person stochastic games. For the model with a finite state and action space they showed the existence of an SSPE.

The computation of SSPEs is very important in applications of stochastic games [16, 22, 23]. However, computation remains a challenging problem since the structure of stochastic games is very complicated. Homotopy methods as proposed by [8] and [25] are a class of powerful methods for solving problems that can be formulated as a fixed point problem. Examples of such problems are the computation of competitive equilibria in general equilibrium models [6, 33] and the computation of equilibria in non-cooperative game theory [5, 17].

In [16], the stochastic linear tracing procedure (SLTP) was developed to construct a differentiable path converging to an SSPE for a finite discounted stochastic game. Their method is the first globally convergent algorithm to solve for an SSPE in a stochastic game, indicating that the homotopy method can be applied to stochastic games as well. The purpose of SLTP is to extend the tracing procedure of Harsanyi and Selten [13], a reasoning process

to select one particular equilibrium among the set of all equilibria, to the class of stochastic games. Even though SLTP is an effective method to compute SSPEs, it 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 path induced by 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 [16], 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 work as well [15, 18], but also this approach leads to an increase in computation time. Finally, SLTP only works for generic stochastic games.

The idea of an ‘interior-point’ method was first proposed and applied in convex optimization, where points on the ‘path’ are restricted to the interior of the feasible set, thereby by-passing many boundary points. Interior-point methods are very efficient for large-scale linear and convex quadratic programming problems, see for instance [30, 31, 32]. In the past decade, this interior-point idea has been applied in market equilibrium problems and normal-form games [4, 7, 34], which also confirms its excellent numerical performance.

It is therefore natural to ask whether one can extend the ‘interior-point’ idea to stochastic games and if the performance of ‘interior-point’ methods is preserved in this complicated class of games. In this paper, we develop an interior-point 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 zero. IPM solves the earlier mentioned obstacles related to SLTP. The starting point of IPM can be arbitrarily chosen and there is no need to solve optimization problems to obtain it. The barrier term 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 zero, the equilibria of the artificial game are in totally mixed strategies. The switching between different systems of equations in SLTP is therefore avoided by IPM, see Figure 1 for an illustration. IPM makes use of a well-chosen perturbation term, which enables us to find an SSPE for every stochastic game. Finally, when implementing IPM to test its numerical performance, we show how a transformation of variables can be

used to reduce the number of equations and unknowns by about one half.

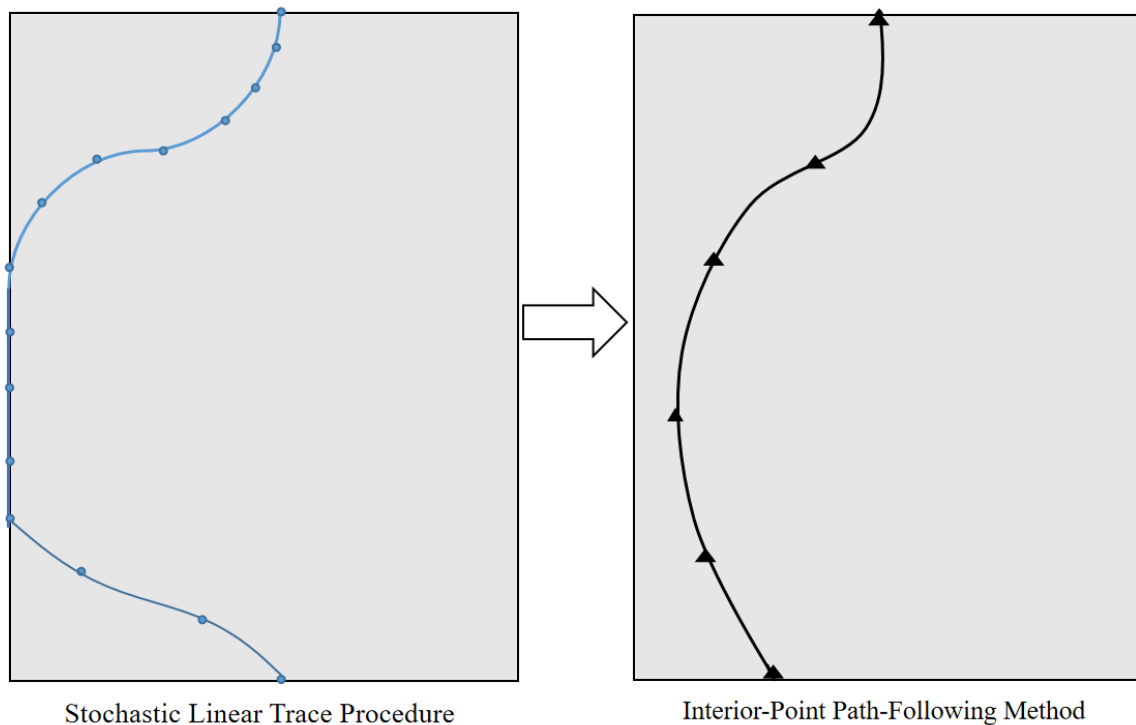


Figure 1: Different trajectories of the two methods

The remainder of this paper is organized as follows. We introduce some preliminaries about finite discounted stochastic games and formulate their equilibrium system in Section 2. In Section 3, we construct an artificial barrier stochastic game and propose our interior-point 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 to SLTP, which further illustrates the efficiency 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  $\Gamma = \langle N, \Omega, \{S_\omega^i\}_{(i,\omega) \in N \times \Omega}, \{u^i\}_{i \in N}, \pi, \delta \rangle$ , where

- $N = \{1, 2, \dots, n\}$  is the set of players.

- $\Omega = \{\omega_1, \omega_2, \dots, \omega_d\}$  is the state space.
- $S_\omega^i = \{s_{\omega_j}^i \mid j \in M_\omega^i\}$  is the set of actions of player  $i \in N$  in state  $\omega \in \Omega$ , where  $M_\omega^i = \{1, 2, \dots, m_\omega^i\}$  is the index set of actions of player  $i$  in state  $\omega$ .
- $S_\omega = \prod_{i=1}^n S_\omega^i$  is the set of action profiles in state  $\omega$ .
- Let  $S_\omega^{-i} = \prod_{k \in N \setminus \{i\}} S_\omega^k$ . Then,  $s_\omega = (s_{\omega_{j_1}}^1, s_{\omega_{j_2}}^2, \dots, s_{\omega_{j_n}}^n) \in S_\omega$  can be written as  $s_\omega = (s_{\omega_{j_i}}^i, s_\omega^{-i})$  with  $s_\omega^{-i} \in S_\omega^{-i}$ .
- $X_\omega^i = \{x_\omega^i \in \mathbb{R}_+^{m_\omega^i} \mid \sum_{j \in M_\omega^i} x_{\omega_j}^i = 1\}$  is the set of mixed actions of player  $i$  in state  $\omega$ . For  $x_\omega^i \in X_\omega^i$ , the probability assigned to  $s_{\omega_j}^i \in S_\omega^i$  equals  $x_{\omega_j}^i$ .
- $X_\omega = \prod_{i=1}^n X_\omega^i$  is the set of mixed action profiles in state  $\omega$ . If  $x_\omega \in X_\omega$  is played, then the probability that an action profile  $s_\omega = (s_{\omega_{j_1}}^1, s_{\omega_{j_2}}^2, \dots, s_{\omega_{j_n}}^n) \in S_\omega$  occurs is equal to  $\prod_{i=1}^n x_{\omega_{j_i}}^i$ .
- Let  $X_\omega^{-i} = \prod_{k \in N \setminus \{i\}} X_\omega^k$ . Then,  $x_\omega = (x_{\omega_{j_1}}^1, x_{\omega_{j_2}}^2, \dots, x_{\omega_{j_n}}^n) \in X_\omega$  can be written as  $x_\omega = (x_{\omega_{j_i}}^i, x_\omega^{-i})$  with  $x_\omega^{-i} \in X_\omega^{-i}$ .
- 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  $\kappa$  equals  $H^\kappa = \prod_{q=0}^{\kappa-1} (D \times \Omega)$ .
- $u^i : D \rightarrow R$  is the instantaneous payoff function of player  $i$ . We have

$$u^i(\omega, x_\omega) = \sum_{s_\omega \in S_\omega} u^i(\omega, s_\omega) \prod_{i=1}^n x_{\omega_{j_i}}^i = \sum_{j \in M_\omega^i} x_{\omega_j}^i u^i(\omega, s_{\omega_j}^i, x_\omega^{-i}).$$

- $\delta$  is the discount factor with  $0 < \delta < 1$ .
- $\pi(\bar{\omega} \mid \omega, s_\omega)$  is the probability that the system jumps from state  $\omega$  to state  $\bar{\omega}$  when the action profile  $s_\omega$  is chosen, where  $\pi(\bar{\omega} \mid \omega, s_\omega) \geq 0$  and  $\sum_{\bar{\omega} \in \Omega} \pi(\bar{\omega} \mid \omega, s_\omega) = 1$ .
- $\pi(\omega, s_\omega) = (\pi(\omega_1 \mid \omega, s_\omega), \pi(\omega_2 \mid \omega, s_\omega), \dots, \pi(\omega_d \mid \omega, s_\omega))$  is the state transition probability map.
- Let  $X = \prod_{i \in N} X^i$  with  $X^i = \prod_{\omega \in \Omega} X_\omega^i$ . An element of  $X$  has  $\bar{m} = \sum_{i \in N} \sum_{\omega \in \Omega} m_\omega^i$  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  $h \in \cup_{\kappa=0}^{\infty} H^{\kappa}$  induces a subgame of  $\Gamma$ . A strategy profile is a subgame perfect equilibrium if it induces a Nash equilibrium in every subgame of  $\Gamma$ .

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  $x \in X$ . The most important solution concept that has been used in applications of stochastic games is subgame perfect equilibrium in stationary strategies (SSPE). An SSPE is a stationary strategy profile that induces a Nash equilibrium in every subgame of  $\Gamma$ . 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  $x \in X$ , 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  $\mu_{\omega}^i(k)$ . The value of  $\mu_{\omega}^i(k)$  follows from the following system of recursive equations,

$$\mu_{\omega}^i(k+1) = u^i(\omega, x_{\omega}) + \delta \sum_{\bar{\omega} \in \Omega} \pi(\bar{\omega} | \omega, x_{\omega}) \mu_{\bar{\omega}}^i(k).$$

For any initial state  $\omega \in \Omega$ , let  $\mu_{\omega}^i = \lim_{k \rightarrow \infty} \mu_{\omega}^i(k)$  be the total expected payoff for player  $i$ . Since  $\mu_{\omega}^i(k)$  is an increasing, uniformly bounded, function of  $k$ ,  $\mu_{\omega}^i$  always exists. It is the unique solution to the linear system of equations

$$\mu_{\omega}^i = u^i(\omega, x_{\omega}) + \delta \sum_{\bar{\omega} \in \Omega} \pi(\bar{\omega} | \omega, x_{\omega}) \mu_{\bar{\omega}}^i. \quad (1)$$

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

$$\varphi^i(\omega, s_{\omega j}^i, x_{\omega}^{-i}, \mu^i(x)) = u^i(\omega, s_{\omega j}^i, x_{\omega}^{-i}) + \delta \sum_{\bar{\omega} \in \Omega} \pi(\bar{\omega} | \omega, s_{\omega j}^i, x_{\omega}^{-i}) \mu_{\bar{\omega}}^i(x),$$

where  $\mu^i(x) = (\mu_{\omega}^i(x))_{\omega \in \Omega}$  is the unique solution to the linear system (1). Then, for any stationary strategy profile  $x \in X$ , 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  $\omega$  can be found as the solution to the



optimization problem

$$\begin{aligned}
& \max_{\hat{x}_\omega^i \in X_\omega^i} \quad \sum_{j \in M_\omega^i} \hat{x}_{\omega j}^i \varphi^i(\omega, s_{\omega j}^i, x_\omega^{-i}, \mu^i(x)) \\
& \text{s.t.} \quad \hat{x}_{\omega j}^i \geq 0, \quad j \in M_\omega^i, \\
& \quad \quad \sum_{j \in M_\omega^i} \hat{x}_{\omega j}^i = 1.
\end{aligned} \tag{2}$$

A direct application of the optimality conditions yields

$$\begin{aligned}
& \varphi^i(\omega, s_{\omega j}^i, x_\omega^{-i}, \mu^i(x)) + \hat{\lambda}_{\omega j}^i - \hat{\beta}_\omega^i = 0, \quad \omega \in \Omega, \quad i \in N, \quad j \in M_\omega^i, \\
& \hat{\lambda}_{\omega j}^i \hat{x}_{\omega j}^i = 0, \quad \hat{\lambda}_{\omega j}^i \geq 0, \quad \hat{x}_{\omega j}^i \geq 0, \quad \omega \in \Omega, \quad i \in N, \quad j \in M_\omega^i, \\
& \sum_{j \in M_\omega^i} \hat{x}_{\omega j}^i = 1, \quad \omega \in \Omega, \quad i \in N.
\end{aligned} \tag{3}$$

We multiply both sides of the first group of equations in (3) with  $\hat{x}_{\omega j}^i$ , sum each side over  $j \in M_\omega^i$ , and obtain that

$$\hat{\beta}_\omega^i = \sum_{j \in M_\omega^i} \hat{x}_{\omega j}^i \varphi^i(\omega, s_{\omega j}^i, x_\omega^{-i}, \mu^i(x)) = u^i(\omega, \hat{x}_\omega^i, x_\omega^{-i}) + \sum_{\bar{\omega} \in \Omega} \pi(\bar{\omega} | \omega, \hat{x}_\omega^i, x_\omega^{-i}) \mu_{\bar{\omega}}^i(x).$$

The set of solutions to (3), which is denoted by  $B(x)$ , consists of best responses to the given stationary strategy profile  $x$ . It follows from the one-stage deviation principle that  $\hat{x}$  is a subgame perfect equilibrium in stationary strategies (SSPE) if and only if it is a fixed point of the best response correspondence, i.e.  $\hat{x} \in B(\hat{x})$ , see [12]. Then, letting  $x = \hat{x}$  and  $\hat{\mu}^i = \mu^i(\hat{x})$ , we get

$$\begin{aligned}
& \varphi^i(\omega, s_{\omega j}^i, \hat{x}_\omega^{-i}, \hat{\mu}^i) + \hat{\lambda}_{\omega j}^i - \hat{\beta}_\omega^i = 0, \quad \omega \in \Omega, \quad i \in N, \quad j \in M_\omega^i, \\
& \hat{\lambda}_{\omega j}^i \hat{x}_{\omega j}^i = 0, \quad \hat{\lambda}_{\omega j}^i \geq 0, \quad \hat{x}_{\omega j}^i \geq 0, \quad \omega \in \Omega, \quad i \in N, \quad j \in M_\omega^i, \\
& \sum_{j \in M_\omega^i} \hat{x}_{\omega j}^i = 1, \quad \omega \in \Omega, \quad i \in N.
\end{aligned}$$

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

$$\begin{aligned}
& u^i(\omega, s_{\omega j}^i, \hat{x}_\omega^{-i}) + \delta \sum_{\bar{\omega} \in \Omega} \pi(\bar{\omega} | \omega, s_{\omega j}^i, \hat{x}_\omega^{-i}) \hat{\mu}_{\bar{\omega}}^i + \hat{\lambda}_{\omega j}^i - \hat{\mu}_\omega^i = 0, \quad \omega \in \Omega, \quad i \in N, \quad j \in M_\omega^i, \\
& \hat{\lambda}_{\omega j}^i \hat{x}_{\omega j}^i = 0, \quad \hat{\lambda}_{\omega j}^i \geq 0, \quad \hat{x}_{\omega j}^i \geq 0, \quad \omega \in \Omega, \quad i \in N, \quad j \in M_\omega^i, \\
& \sum_{j \in M_\omega^i} \hat{x}_{\omega j}^i = 1, \quad \omega \in \Omega, \quad i \in N.
\end{aligned} \tag{4}$$

Thus far, we have reformulated the problem to find an SSPE as the equivalent problem of solving the nonlinear system of equations (4). In the remainder of this paper, we aim to explore an effective and efficient algorithm to compute a solution to the system (4).

### 3 An Interior-Point Path-Following Method

In general, it is difficult to solve the system (4) directly. [16] extended the linear tracing procedure of [13] 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 and the game of interest. It is shown in [16] 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 path-following method to find an SSPE. By introducing a homotopy variable  $t$  ranging from one to zero, 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 one to zero. 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 homotopy variable  $t$ . As  $t$  decreases to zero, an SSPE for the stochastic game of interest is obtained. Additionally, due to 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  $x \in X$ ,  $\mu^i = \mu^i(x) = (\mu_\omega^i(x))_{\omega \in \Omega}$  is defined to be the unique solution to the following linear system,

$$\mu_\omega^i = (1 - t)(u^i(\omega, x_\omega) + \delta \sum_{\bar{\omega} \in \Omega} \pi(\bar{\omega}|\omega, x_\omega) \mu_{\bar{\omega}}^i) + t^2, \quad \omega \in \Omega. \quad (5)$$

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

$$\begin{aligned} \max_{\hat{x}_\omega^i \in X_\omega^i} \quad & (1 - t) \sum_{j \in M_\omega^i} \hat{x}_{\omega j}^i \varphi^i(\omega, s_{\omega j}^i, x_\omega^{-i}, \mu^i(x)) + t^2 \sum_{j \in M_\omega^i} x_{\omega j}^{0,i} \ln \hat{x}_{\omega j}^i \\ \text{s.t.} \quad & \sum_{j \in M_\omega^i} \hat{x}_{\omega j}^i = 1, \end{aligned} \quad (6)$$

where  $x_\omega^{0,i}$  is an arbitrarily chosen totally mixed stationary strategy profile with  $\sum_{j \in M_\omega^i} x_{\omega j}^{0,i} = 1$ .

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 zero, 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

$$\begin{aligned} (1-t)\varphi^i(\omega, s_{\omega j}^i, x_{\omega}^{-i}, \mu^i(x)) + \hat{\lambda}_{\omega j}^i - \hat{\beta}_{\omega}^i &= 0, & \omega \in \Omega, i \in N, j \in M_{\omega}^i, \\ \hat{\lambda}_{\omega j}^i \hat{x}_{\omega j}^i - t^2 x_{\omega j}^{0,i} &= 0, & \omega \in \Omega, i \in N, j \in M_{\omega}^i, \\ \sum_{j \in M_{\omega}^i} \hat{x}_{\omega j}^i &= 1, & \omega \in \Omega, i \in N. \end{aligned} \quad (7)$$

Multiplying both sides of the first group of equations in the system (7) by  $\hat{x}_{\omega j}^i$  and summing over  $j$ , we obtain that

$$\hat{\beta}_{\omega}^i = (1-t)(u^i(\omega, \hat{x}_{\omega}^i, x_{\omega}^{-i}) + \delta \sum_{\bar{\omega} \in \Omega} \pi(\bar{\omega}|\omega, \hat{x}_{\omega}^i, x_{\omega}^{-i}) \mu_{\bar{\omega}}^i(x)) + t^2.$$

From a fixed point argument and (5), letting  $x = \hat{x}$ , we obtain that  $\hat{\beta}_{\omega}^i = \hat{\mu}_{\omega}^i$ . The equilibrium system of the artificial stochastic barrier game is therefore given by

$$\begin{aligned} (1-t)(u^i(\omega, s_{\omega j}^i, \hat{x}_{\omega}^{-i}) + \delta \sum_{\bar{\omega} \in \Omega} \pi(\bar{\omega}|\omega, s_{\omega j}^i, \hat{x}_{\omega}^{-i}) \hat{\mu}_{\bar{\omega}}^i) \\ + \hat{\lambda}_{\omega j}^i - \hat{\mu}_{\omega}^i &= 0, & \omega \in \Omega, i \in N, j \in M_{\omega}^i, \\ \hat{\lambda}_{\omega j}^i \hat{x}_{\omega j}^i - t^2 x_{\omega j}^{0,i} &= 0, & \omega \in \Omega, i \in N, j \in M_{\omega}^i, \\ \sum_{j \in M_{\omega}^i} \hat{x}_{\omega j}^i &= 1, & \omega \in \Omega, i \in N. \end{aligned} \quad (8)$$

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

$$\begin{aligned} \hat{\lambda}_{\omega j}^i - \hat{\mu}_{\omega}^i &= 0, & \omega \in \Omega, i \in N, j \in M_{\omega}^i, \\ \hat{\lambda}_{\omega j}^i \hat{x}_{\omega j}^i - x_{\omega j}^{0,i} &= 0, & \omega \in \Omega, i \in N, j \in M_{\omega}^i, \\ \sum_{j \in M_{\omega}^i} \hat{x}_{\omega j}^i &= 1, & \omega \in \Omega, i \in N. \end{aligned} \quad (9)$$

**Theorem 1.** *As  $t = 1$ , the system (8) has a unique solution  $(\hat{x}(1), \hat{\lambda}(1), \hat{\mu}(1))$  with*

$$\hat{x}_{\omega j}^i(1) = x_{\omega j}^{0,i}, \quad \hat{\lambda}_{\omega j}^i(1) = 1, \quad \text{and} \quad \hat{\mu}_{\omega}^i(1) = 1,$$

where  $\omega \in \Omega$ ,  $i \in N$ , and  $j \in M_{\omega}^i$ .

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

$$\begin{aligned} \max_{\hat{x}_{\omega}^i \in X_{\omega}^i} & \sum_{j \in M_{\omega}^i} x_{\omega j}^{0,i} \ln(\hat{x}_{\omega j}^i) \\ \text{s.t.} & \sum_{j \in M_{\omega}^i} \hat{x}_{\omega j}^i = 1, \end{aligned} \quad (10)$$

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

At  $t = 0$ , the definition of  $\mu$  in (5) is the same as that in (1), and the system (8) resumes to system (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 [20].

**Theorem 2 (Mas-Colell's fixed point theorem).** *Let  $S$  be a non-empty, compact and convex subset of  $\mathbb{R}^{\ell}$  and let  $f : S \times [0, 1] \rightarrow S$  be an upper hemi-continuous correspondence. 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  $x \in X$ , let  $\sigma_{\omega}^i(x, t)$  be all  $\hat{x}_{\omega}^i \in X_{\omega}^i$  that solve

$$\begin{aligned} \max_{\hat{x}_{\omega}^i \in X_{\omega}^i} \quad & (1-t) \sum_{j \in M_{\omega}^i} \hat{x}_{\omega j}^i (u^i(\omega, s_{\omega j}^i, x_{\omega}^{-i}) + \delta \sum_{\bar{\omega} \in \Omega} \pi(\bar{\omega} | \omega, s_{\omega j}^i, x_{\omega}^{-i}) \mu_{\bar{\omega}}^i(x)) \\ & + t^2 \sum_{j \in M_{\omega}^i} x_{\omega j}^{0,i} \ln(\hat{x}_{\omega j}^i) - t(1-t) \sum_{j \in M_{\omega}^i} \alpha_{\omega j}^i \hat{x}_{\omega j}^i \\ \text{s.t.} \quad & \sum_{j \in M_{\omega}^i} \hat{x}_{\omega j}^i = 1, \end{aligned} \quad (11)$$

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

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

$$\begin{aligned} (1-t)(u^i(\omega, s_{\omega j}^i, x_{\omega}^{-i}) + \delta \sum_{\bar{\omega} \in \Omega} \pi(\bar{\omega} | \omega, s_{\omega j}^i, x_{\omega}^{-i}) \mu_{\bar{\omega}}^i(x)) \\ + \hat{\lambda}_{\omega j}^i - \hat{\beta}_{\omega}^i - t(1-t)\alpha_{\omega j}^i = 0, \quad \omega \in \Omega, \quad i \in N, \quad j \in M_{\omega}^i, \\ \hat{\lambda}_{\omega j}^i \hat{x}_{\omega j}^i - t^2 x_{\omega j}^{0,i} = 0, \quad \omega \in \Omega, \quad i \in N, \quad j \in M_{\omega}^i, \\ \sum_{j \in M_{\omega}^i} \hat{x}_{\omega j}^i = 1, \quad \omega \in \Omega, \quad i \in N. \end{aligned} \quad (12)$$

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

We argue next that  $\sigma_\omega^i : X \times [0, 1] \rightarrow X_\omega^i$  is an upper hemi-continuous 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  $\sigma_\omega^i$  is closed, which is equivalent to  $\sigma_\omega^i$  being an upper hemi-continuous correspondence. Note that  $\sigma_\omega^i$  is a continuous function on  $X \times (0, 1]$  since 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,  $\sigma_\omega^i$  is single-valued.

From the above discussion,  $H(x, t)$  is obtained as a product of  $\sigma_\omega^i$  for  $\omega \in \Omega$  and  $i \in N$ . Then,  $H(x, t)$  is also an upper hemi-continuous correspondence.

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

$$\begin{aligned}
(1-t)(u^i(\omega, s_{\omega j}^i, \hat{x}_\omega^{-i}) + \delta \sum_{\bar{\omega} \in \Omega} \pi(\bar{\omega} | \omega, s_{\omega j}^i, \hat{x}_\omega^{-i}) \hat{\mu}_{\bar{\omega}}^i) \\
+ \hat{\lambda}_{\omega j}^i - \hat{\mu}_\omega^i - t(1-t)\alpha_{\omega j}^i = 0, \quad \omega \in \Omega, \quad i \in N, \quad j \in M_\omega^i, \\
\hat{\lambda}_{\omega j}^i \hat{x}_{\omega j}^i - t^2 x_{\omega j}^{0,i} = 0, \quad \omega \in \Omega, \quad i \in N, \quad j \in M_\omega^i, \\
\sum_{j \in M_\omega^i} \hat{x}_{\omega j}^i - 1 = 0, \quad \omega \in \Omega, \quad i \in N,
\end{aligned} \tag{13}$$

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

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

*Proof.* Comparing the systems (12) and (13), we find that letting  $x = \hat{x}$ , the system (12) totally resumes to the system (13). Then,  $\Phi$  can be rewritten as

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

The upper hemi-continuity of  $H$ , together with a direct application of Mas-Colell's fixed point theorem, leads to the conclusion of our corollary.  $\square$

All equations in (13) are polynomial. The set  $\Phi$  is therefore a semi-algebraic set, so all its components are also path-connected, i.e., any two points in a component can be joined by a path, see [26]. We obtain the following corollary.

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

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

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

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

**Theorem 3.** *For a generic choice of  $\alpha \in \mathbb{R}^{\bar{m}}$ , there exists a smooth path in  $\Delta$ , 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,  $\Phi$  contains a path-connected subset  $\Phi^c$  that intersects both the sets  $\mathbb{R}^{\bar{m}} \times \{1\}$  and  $\mathbb{R}^{\bar{m}} \times \{0\}$ . Moreover, we can choose  $\Phi^c$  such that it contains a single point in  $\mathbb{R}^{\bar{m}} \times \{0\}$ . The second group of equations in (13) determines a unique value for  $\lambda \in \mathbb{R}^{\bar{m}}$  for each  $(x, t) \in \mathbb{R}^{\bar{m}} \times \mathbb{R}$ . Next, the first group of linear equations in (13) pins down a unique value for  $\mu \in \mathbb{R}^{nd}$ . Thus  $\Delta$  has a path-connected subset that intersects both the sets  $\mathbb{R}^{\bar{m}} \times \mathbb{R}^{\bar{m}} \times \mathbb{R}^{nd} \times \{1\}$  and  $\mathbb{R}^{\bar{m}} \times \mathbb{R}^{\bar{m}} \times \mathbb{R}^{nd} \times \{0\}$ . By Theorem 1 there is a unique starting point at  $t = 1$ . We prove in Appendix I that the Jacobian matrix of  $p_0(x, \lambda, \mu, 1)$  is of full rank. For every  $\alpha \in \mathbb{R}^{\bar{m}}$ ,  $p_\alpha(x, \lambda, \mu, 1) = p_0(x, \lambda, \mu, 1)$ , so it follows that zero is a regular value of  $p_\alpha(x, \lambda, \mu, 1)$  on  $\mathbb{R}^{\bar{m}} \times \mathbb{R}^{\bar{m}} \times \mathbb{R}^{nd} \times \{1\}$ . Similarly, zero is also a regular value of  $p(x, \lambda, \mu, t; \alpha)$  since the Jacobian matrix of  $p(x, \lambda, \mu, t; \alpha)$  is of full-row rank, for all  $(x, \lambda, \mu, t; \alpha) \in \mathbb{R}^{\bar{m}} \times \mathbb{R}^{\bar{m}} \times \mathbb{R}^{nd} \times (0, 1) \times \mathbb{R}^{\bar{m}}$ , see Appendix I. By a direct application of the well-known transversality theorem, see [9], zero is a regular value of  $p_\alpha(x, \lambda, \mu, t)$  for almost all  $\alpha \in \mathbb{R}^{\bar{m}}$ . Then it follows that  $\Phi^c$  is a smooth one-dimensional manifold with boundary, which starts from the unique solution at  $t = 1$  and ends at an SSPE of the stochastic game of interest at  $t = 0$ . This completes the proof.  $\square$

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  $\|\alpha\|$  is small, the solutions to the system of equations (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  $t = 0$ , the 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  $\alpha$  into the equilibrium system to avoid degeneracies. In fact, for a generic choice of  $\alpha$ , our method finds an SSPE for every stochastic game. SLTP does not use an  $\alpha$  to perturb the system and only computes an SSPE for almost every stochastic game. In numerical experiments, we have always obtained convergence for  $\alpha = 0$ .

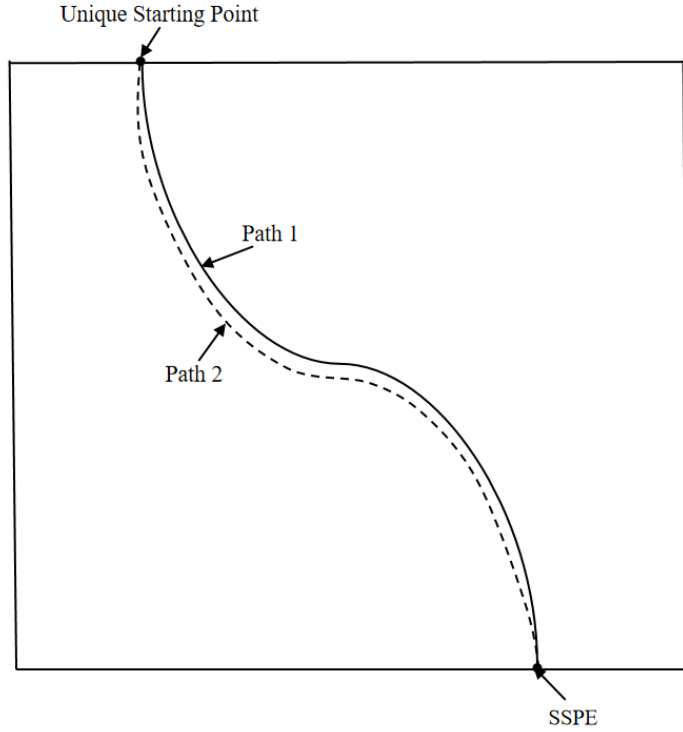


Figure 2: Path 1 and Path 2 represent the solution sets for the systems (13) and (8), respectively.

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  $j \in M_\omega^i$ , we write  $x_{\omega j}^i$  and  $\lambda_{\omega j}^i$  as a function of a variable  $y_{\omega j}^i$  and the homotopy parameter  $t$ ,

$$x_{\omega j}^i(y, t) = \left( \frac{\sqrt{(y_{\omega j}^i)^2 + 4t\sqrt{x_{\omega j}^{0,i} + y_{\omega j}^i}}}{2} \right)^2 \quad \text{and} \quad \lambda_{\omega j}^i(y, t) = \left( \frac{\sqrt{(y_{\omega j}^i)^2 + 4t\sqrt{x_{\omega j}^{0,i} - y_{\omega j}^i}}}{2} \right)^2.$$

Clearly,  $x(y, t)$  and  $\lambda(y, t)$  are continuously differentiable functions for all  $y \in \mathbb{R}^{\bar{m}}$  and  $t \in (0, 1]$ , and it holds that, for every  $\omega \in \Omega$ ,  $i \in N$ , and  $j \in M_\omega^i$ ,  $\lambda_{\omega j}^i(y, t)x_{\omega j}^i(y, t) = t^2 x_{\omega j}^{0,i}$ . We substitute these functions in (13) and obtain the following homotopy system,

$$\begin{aligned} (1-t)(u^i(\omega, s_{\omega j}^i, x_{\omega}^{-i}(\hat{y}, t)) + \delta \sum_{\bar{\omega} \in \Omega} \pi(\bar{\omega}|\omega, s_{\omega j}^i, x_{\omega}^{-i}(\hat{y}, t))\hat{\mu}_{\bar{\omega}}^i) \\ + \lambda_{\omega j}^i(\hat{y}, t) - \hat{\mu}_{\omega}^i - t(1-t)\alpha_{\omega j}^i = 0, \quad \omega \in \Omega, \quad i \in N, \quad j \in M_\omega^i, \quad (14) \\ \sum_{j \in M_\omega^i} x_{\omega j}^i(\hat{y}, t) - 1 = 0, \quad \omega \in \Omega, \quad i \in N. \end{aligned}$$

The system of equations (14) corresponds to our proposed interior path-following method (IPM). 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  $\alpha \in \mathbb{R}^{\bar{m}}$ .

## 4 Numerical Performance

We use a predictor-corrector algorithm, see [1] and [9] for details, to trace the smooth path generated by IPM. Moreover, we compare the numerical performance of IPM with SLTP based on [16], see Appendix II, to illustrate the efficiency of IPM. All experiments are run in MatLab software on a 2.00 GHz Windows PC with CORE i7. The stopping criterion is taken equal to  $t < 10^{-6}$ . Each experiment is implemented 10 times and the average number of iterations (AITER) and average computation time (ATIME) are recorded.

### 4.1 Fundamental Cases

In this subsection, we implement IPM to solve several basic stochastic games with different numbers of players, states, and actions, and compare its performance with SLTP. The discount factor is always taken equal to  $\delta = 0.95$ .

**Example 1.** Assume  $N = \{1, 2\}$ ,  $\Omega = \{\omega_1, \omega_2\}$ , and, for  $i = 1, 2$ ,  $S_{\omega_1}^i = \{s_{\omega_1 1}^i, s_{\omega_1 2}^i\}$  and  $S_{\omega_2}^i = \{s_{\omega_2 1}^i\}$ . The payoff matrices in states  $\omega_1$  and  $\omega_2$  are given by

$$\begin{array}{c|cc} \omega_1 & s_{\omega_1 1}^2 & s_{\omega_1 2}^2 \\ \hline s_{\omega_1 1}^1 & (1, -1) & (0, 0) \\ s_{\omega_1 2}^1 & (0, 0) & (3, -3) \end{array} \quad \text{and} \quad \begin{array}{c|c} \omega_2 & s_{\omega_2 1}^2 \\ \hline s_{\omega_2 1}^1 & (0, 0) \end{array}.$$

The transition probability matrices in states  $\omega_1$  and  $\omega_2$  are given by

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

IPM finds the SSPE

$$(((x_{\omega_1 1}^1, x_{\omega_1 2}^1), (x_{\omega_1 1}^2, x_{\omega_1 2}^2)), (x_{\omega_2 1}^1, x_{\omega_2 1}^2)) = (((0.67, 0.33), (0.67, 0.33)), (1, 1)).$$

Figure 3 shows the development of the variables  $y$  and  $t$  in the various iterations of IPM. The downward sloping (yellow) curve corresponds to  $t$ , the non-monotonic (red) curve to  $y_{\omega_1 1}^1$ , and the upward sloping (blue) curve to  $y_{\omega_1 2}^1$ .  $\square$

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

$$\begin{array}{c|cc} (\omega_k)_{k=1,2} & s_{\omega_k 1}^2 & s_{\omega_k 2}^2 \\ \hline s_{\omega_k 1}^1 & (0, 0) & (0, 0) \\ s_{\omega_k 2}^1 & (0, 0) & (0, 0) \end{array}, \quad \begin{array}{c|c} \omega_3 & s_{\omega_3 1}^2 \\ \hline s_{\omega_3 1}^1 & (1, -1) \end{array}, \quad \text{and} \quad \begin{array}{c|c} \omega_4 & s_{\omega_4 1}^2 \\ \hline s_{\omega_4 1}^1 & (-1, 1) \end{array}.$$



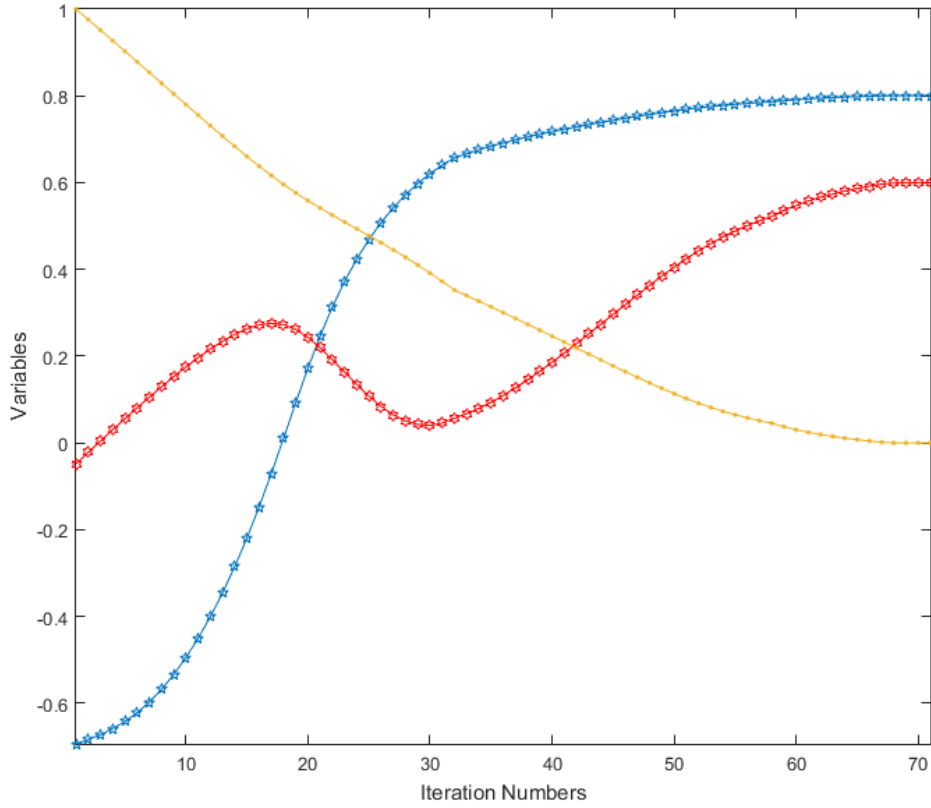


Figure 3: Development of the variables  $t$ ,  $y_{\omega_1}^1$ , and  $y_{\omega_12}^1$  along the homotopy path.

The transition probability matrix in state  $\omega_1$  is given by

$$\frac{\pi((\omega_1, \omega_2, \omega_3, \omega_4) | \omega_1)}{\begin{matrix} s_{\omega_11}^1 \\ s_{\omega_12}^1 \end{matrix}} \left| \begin{array}{cc} s_{\omega_11}^2 & s_{\omega_12}^2 \\ (1, 0, 0, 0) & (0, 0, 1, 0) \\ (0, 0, 1, 0) & (0, 1, 0, 0) \end{array} \right. ,$$

and the transition probability matrix in state  $\omega_2$  is given by

$$\frac{\pi((\omega_1, \omega_2, \omega_3, \omega_4) | \omega_2)}{\begin{matrix} s_{\omega_21}^1 \\ s_{\omega_22}^1 \end{matrix}} \left| \begin{array}{cc} s_{\omega_21}^2 & s_{\omega_22}^2 \\ (1, 0, 0, 0) & (0, 0, 0, 1) \\ (0, 0, 0, 1) & (0, 1, 0, 0) \end{array} \right. .$$

The states  $\omega_3$  and  $\omega_4$  are absorbing. IPM obtains the SSPE

$$((x_{\omega_11}^1, x_{\omega_11}^2), (x_{\omega_21}^1, x_{\omega_21}^2), (x_{\omega_31}^1, x_{\omega_31}^2), (x_{\omega_41}^1, x_{\omega_41}^2)) = ((0.86, 0.86), (0.14, 0.14), (1, 1), (1, 1)).$$

Figure 4 shows the development of the variables  $y$  and  $t$  in the various iterations of IPM. The downward sloping (yellow) curve corresponds to  $t$ , the non-monotonic (red) curve to  $y_{\omega_1}^1$ , and the upward sloping (blue) curve to  $y_{\omega_12}^1$ .  $\square$

**Example 3.** Assume  $N = \{1, 2\}$ ,  $\Omega = \{\omega_1, \omega_2, \omega_3\}$ , and, for  $i = 1, 2$ ,  $S_{\omega_1}^i = \{s_{\omega_11}^i, s_{\omega_12}^i\}$ ,

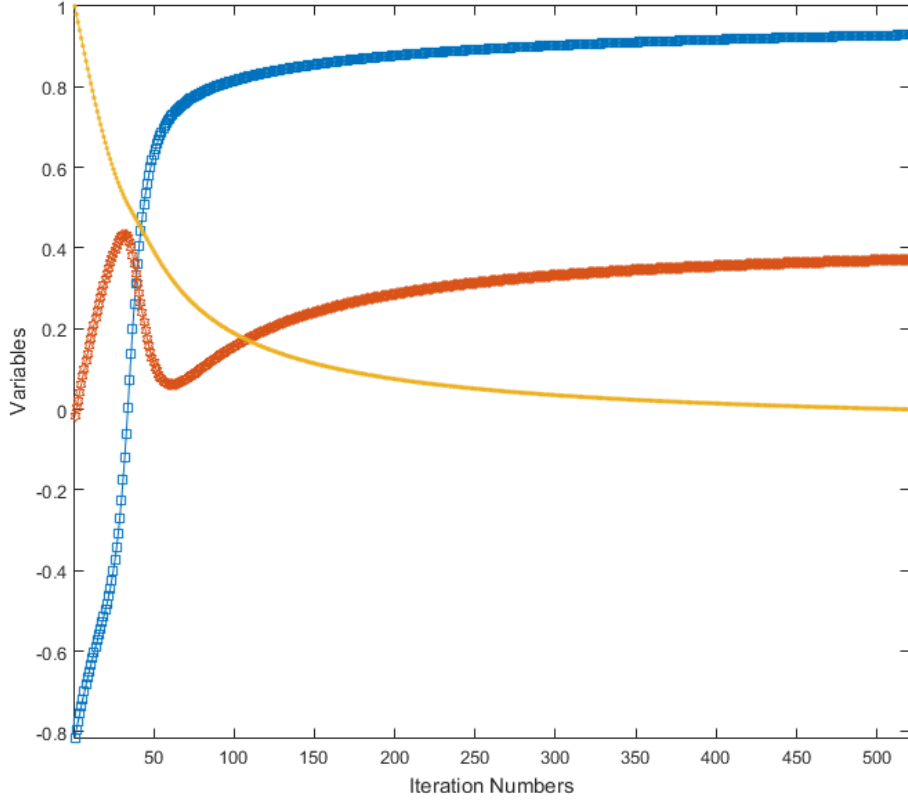


Figure 4: Development of the variables  $t$ ,  $y_{\omega_1}^1$ , and  $y_{\omega_12}^1$  along the homotopy path.

$S_{\omega_2}^i = \{s_{\omega_21}^i\}$ , and  $S_{\omega_3}^i = \{s_{\omega_31}^i\}$ . The payoff matrices are given by

$$\begin{array}{c|cc} \omega_1 & s_{\omega_11}^2 & s_{\omega_12}^2 \\ \hline s_{\omega_11}^1 & (1, -1) & (0, 0) \\ s_{\omega_12}^1 & (0, 0) & (1, -1) \end{array}, \quad \begin{array}{c|c} \omega_2 & s_{\omega_21}^2 \\ \hline s_{\omega_21}^1 & (0, 0) \end{array}, \quad \text{and} \quad \begin{array}{c|c} \omega_3 & s_{\omega_31}^2 \\ \hline s_{\omega_31}^1 & (1, -1) \end{array}.$$

The transition probability matrix in state  $\omega_1$  is given by

$$\begin{array}{c|cc} \pi((\omega_1, \omega_2, \omega_3) | \omega_1) & s_{\omega_11}^2 & s_{\omega_12}^2 \\ \hline s_{\omega_11}^1 & (1, 0, 0) & (0, 1, 0) \\ s_{\omega_12}^1 & (1, 0, 0) & (0, 0, 1) \end{array}.$$

The states  $\omega_2$  and  $\omega_3$  are absorbing. IPM generates the SSPE

$$((x_{\omega_11}^1, x_{\omega_11}^2), (x_{\omega_21}^1, x_{\omega_21}^2), (x_{\omega_31}^1, x_{\omega_31}^2)) = ((1, 0.5), (1, 1), (1, 1)).$$

□

**Example 4.** Assume  $N = \{1, 2\}$ ,  $\Omega = \{\omega_1, \omega_2, \omega_3\}$ , and, for  $i = 1, 2$ ,  $S_{\omega_1}^i = \{s_{\omega_11}^i, s_{\omega_12}^i\}$ ,  $S_{\omega_2}^i = \{s_{\omega_21}^i\}$ , and  $S_{\omega_3}^i = \{s_{\omega_31}^i\}$ . The payoff matrices are given by

$$\begin{array}{c|cc} \omega_1 & s_{\omega_11}^2 & s_{\omega_12}^2 \\ \hline s_{\omega_11}^1 & (1, 0) & (0, 2) \\ s_{\omega_12}^1 & (0, 1) & (1, 0) \end{array}, \quad \begin{array}{c|c} \omega_2 & s_{\omega_21}^2 \\ \hline s_{\omega_21}^1 & (0, 2) \end{array}, \quad \text{and} \quad \begin{array}{c|c} \omega_3 & s_{\omega_31}^2 \\ \hline s_{\omega_31}^1 & (1, 0) \end{array}.$$

The transition probability matrix in state  $\omega_1$  is given by

$$\frac{\pi((\omega_1, \omega_2, \omega_3) \mid \omega_1)}{\begin{matrix} s_{\omega_1 1}^1 \\ s_{\omega_1 2}^1 \end{matrix}} \left| \begin{array}{cc} s_{\omega_1 1}^2 & s_{\omega_1 2}^2 \\ (1, 0, 0) & (1, 0, 0) \\ (0, 1, 0) & (0, 0, 1) \end{array} \right.$$

The states  $\omega_2$  and  $\omega_3$  are absorbing. IPM finds that SSPE

$$((x_{\omega_1 1}^1, x_{\omega_1 1}^2), (x_{\omega_2 1}^1, x_{\omega_2 1}^2), (x_{\omega_3 1}^1, x_{\omega_3 1}^2)) = ((1, 0.5), (1, 1), (1, 1)).$$

□

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

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

Irrespective of the actions chosen, the transition probability between any two states is equal to 1/2. IPM find the SSPE

$$(((x_{\omega_1 1}^1, x_{\omega_1 2}^1, x_{\omega_1 3}^1), (x_{\omega_1 1}^2, x_{\omega_1 2}^2, x_{\omega_1 3}^2)), (x_{\omega_2 1}^1, x_{\omega_2 1}^2)) = (((0, 1, 0), (0, 1, 0)), (1, 1)).$$

□

We also implement SLTP to solve the five examples and present the average number of iterations (AITER) and computation time in seconds (ATIME) of IPM and SLTP in Tables 1 and 2, respectively. Moreover, we present I-Ratio and T-Ratio, where

$$\text{I-Ratio} = \frac{\text{AITER of IPM}}{\text{AITER of SLTP}},$$

$$\text{T-Ratio} = \frac{\text{ATIME of IPM}}{\text{ATIME of SLTP}}.$$

The five examples illustrate that IPM is an effective and efficient method to compute SSPEs. IPM generates an everywhere smooth path that is in the interior of the space of stationary strategy profiles, see Figures 3 and 4. Moreover, Tables 1 and 2 show that the average number of iterations and computation time for IPM is less than one-third of SLTP, which indicates that IPM is much more efficient than SLTP.

Table 1: Average number of iterations

AITER	IPM	SLTP	I-Ratio
Example 1	60	297	20.20%
Example 2	515	1771	29.07%
Example 3	355	844	42.06%
Example 4	497	1389	35.78%
Example 5	258	1227	21.02%

Table 2: Average computation time

ATIME	IPM	SLTP	T-Ratio
Example 1	1.03	5.10	20.20%
Example 2	25.42	99.33	25.51%
Example 3	8.69	24.71	35.16%
Example 4	13.01	40.20	32.36%
Example 5	4.57	25.67	17.80%

## 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 zero with probability ‘pd0,’ which indicates the probability density of the payoff matrix. Obviously, the larger pd0, the sparser the payoff matrix becomes. Here, we let pd0 be equal 0.00, 0.25, 0.50 and 0.75 to induce several groups of different games and test the performance of the two methods. We take the discount factor equal to  $\delta = 0.95$ .

From Table 3, 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. These results from numerous randomly generated stochastic games illustrate that IPM performs much better both in number of iterations and computation time than SLTP. This advantage becomes more pronounced as the scale of the problem gets larger.

## 4.3 More Complicated Cases

In [16], SSPEs are computed for stochastic games with a scale up to  $n = 5$ ,  $d = 5$ , and  $m = 5$ . The dimension of the homotopy is very sensitive to  $n$ ,  $d$ , and  $m$ . For example, if

Table 3: Numerical Performance and Comparisons

$(n, d, m)$	IPM			SLTP		I-Ratio	T-Ratio
	pd0	AITER	ATIME	AITER	ATIME		
(2,2,5)	0.00	2013	24.28	4220	55.58	47.70%	43.68%
	0.25	1869	24.68	4317	78.70	43.29%	31.35%
	0.50	1716	23.52	4228	76.54	40.58%	20.72%
	0.75	1461	20.18	2965	51.27	49.27%	39.36%
(2,5,3)	0.00	3685	100.24	6083	213.30	60.57%	46.99%
	0.25	2290	78.99	5877	283.45	38.96%	27.86%
	0.50	1762	63.07	4563	216.54	38.61%	29.12%
	0.75	1275	45.80	4522	219.86	28.19%	20.83%
(2,5,4)	0.00	3806	125.95	6592	284.56	57.73%	44.26%
	0.25	3259	137.33	7447	438.66	43.76%	31.30%
	0.50	2812	121.28	5914	341.06	47.54%	35.55%
	0.75	2050	87.92	6069	337.98	33.77%	26.01%
(2,5,5)	0.00	3810	179.20	8002	473.60	47.61%	37.83%
	0.25	3051	113.90	7202	352.21	42.36%	32.33%
	0.50	2503	93.47	6704	320.48	37.33%	29.16%
	0.75	1847	70.51	4702	237.95	39.28%	29.63%
(3,3,3)	0.00	2091	58.00	6596	234.03	31.70%	24.78%
	0.25	2560	90.09	6466	316.41	39.59%	28.47%
	0.50	1843	67.50	5596	271.46	32.93%	24.86%
	0.75	1066	36.56	3893	178.10	27.38%	20.52%
(3,3,5)	0.00	2864	130.97	8136	430.92	35.20%	30.39%
	0.25	2125	112.93	7842	546.24	27.09%	20.67%
	0.50	2119	116.46	6919	484.54	30.62%	24.03%
	0.75	1376	79.06	5356	371.25	25.69%	21.29%
(4,2,5)	0.00	1903	78.36	6945	324.50	27.40%	24.14%
	0.25	1725	77.82	6389	391.71	26.99%	19.86%
	0.50	959	47.66	6221	379.53	15.41%	12.55%
	0.75	1012	50.51	4348	255.39	23.27%	19.77%
(5,2,5)	0.00	1287	92.48	6930	571.21	18.57%	16.19%
	0.25	1039	94.98	8273	735.33	12.55%	12.91%
	0.50	787	74.66	10000	869.32	7.87%	8.58%
	0.75	757	73.75	6338	530.32	11.94%	13.90%

$(d, m) = (5, 5)$  and  $n$  increases from 5 to 6, then the number of variables in the homotopy system increases from 150 to 180.

In this subsection, we use IPM to solve problems with a scale up to  $n = 5$ ,  $d = 8$ , and  $m = 8$ . In each experiment, we let  $\text{pd0}=0.95$  and generate the payoff matrices and transition probabilities randomly. The average computation time for all experiments is recorded in Table 4. It is left blank if the computation time exceeds  $3 \times 10^4$  seconds.

Table 4: Numerical Results IPM

$n = 3$				
$d \backslash m$	5	6	7	8
5	230.90	358.78	747.82	1108.71
6	602.41	936.39	2162.49	3291.86
7	828.61	1065.73	2303.06	4137.06
8	1430.17	2146.85	3021.24	3117.04
$n = 4$				
$d \backslash m$	5	6	7	8
5	539.80	1021.91	1203.28	2255.70
6	1328.68	1554.25	2536.31	4761.73
7	1440.40	2230.62	2413.82	6202.49
8	1875.36	3681.46	4200.69	8576.12
$n = 5$				
$d \backslash m$	5	6	7	8
5	1737.74	3377.21	10403.29	27626.37
6	2897.06	10249.51	24284.54	
7	4378.30	17669.26		
8	10247.18			

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 number of iterations and computation time, which is consistent with the observations in [16].

## 5 Conclusion

In this paper, we extend the idea of interior-point methods, which has been proven to be very efficient for large-scale convex programming problems, to computing a subgame perfect

equilibrium in stationary strategies (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 zero, our path converges to an SSPE for the stochastic game of interest.

Our method (IPM) has several advantages over the alternative method (SLTP) presented in the literature to compute an SSPE. First, 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. SLTP has the same properties only for generic stochastic games and requires a transformation of variables to obtain smoothness everywhere. The effectiveness and efficiency of IPM is confirmed by numerous numerical experiments.

## Appendix I

This appendix shows that for any  $(x, \lambda, \mu, t; \alpha) \in \mathbb{R}^{\bar{m}} \times \mathbb{R}^{\bar{m}} \times \mathbb{R}^{nd} \times (0, 1] \times \mathbb{R}^{\bar{m}}$ , the Jacobian matrix of  $p(x, \lambda, \mu, t; \alpha)$  is of full-row rank. This result is utilized in the proof of Theorem 3. First, we consider the case where  $t \in (0, 1)$ . For simplicity, for  $k = 1, 2, 3$ , we denote the  $k$ -th group of equations in system of equations (13) by  $p^k(x, \lambda, \mu, t; \alpha)$ . The Jacobian matrix  $Dp(x, \lambda, \mu, t; \alpha)$  of  $p(x, \lambda, \mu, t; \alpha)$  is given by

$$\begin{pmatrix} \left(\frac{\partial p^1}{\partial x}\right)_{\bar{m} \times \bar{m}} & I_{\bar{m} \times \bar{m}} & \left(\frac{\partial p^1}{\partial \mu}\right)_{\bar{m} \times nd} & \left(\frac{\partial p^1}{\partial t}\right)_{\bar{m} \times 1} & (-t(1-t)I)_{\bar{m} \times \bar{m}} \\ \text{diag}(\lambda_{\omega_j}^i)_{\bar{m} \times \bar{m}} & \text{diag}(x_{\omega_j}^i)_{\bar{m} \times \bar{m}} & 0 & (-2x_{\omega_j}^{0,i}te)_{\bar{m} \times 1} & 0 \\ \text{diag}(e_{\omega_j}^i)_{nd \times \bar{m}} & 0 & 0 & 0 & 0 \end{pmatrix},$$

where  $I$  is the identity matrix,  $\text{diag}(\lambda_{\omega_j}^i)$  and  $\text{diag}(x_{\omega_j}^i)$  are two diagonal matrices, and  $e_{\omega_j}^i = (1, 1, \dots, 1)$  is an  $m_{\omega_j}^i$ -dimensional vector, so

$$\text{diag}(e_{\omega_j}^i) = \begin{pmatrix} 1 & \cdots & 1 & & \\ & & & \ddots & \\ & & & & 1 & \cdots & 1 \end{pmatrix}$$

Because  $0 < x_{\omega_j}^i < 1$  as  $0 < t < 1$ ,  $\text{diag}(x_{\omega_j}^i) \in \mathbb{R}^{\bar{m} \times \bar{m}}$  has full rank. Clearly,  $-t(1-t)I \in \mathbb{R}^{\bar{m} \times \bar{m}}$  is also of full rank and  $\text{diag}(e_{\omega_j}^i) \in \mathbb{R}^{nd \times \bar{m}}$  is of full-row rank. Then, the Jacobian

matrix  $Dp(x, \lambda, \mu, t; \alpha)$  is of full-row rank.

Next, we prove that as  $t = 1$ ,  $p_0(x, \lambda, \mu, 1)$  has full rank at  $(x(1), \lambda(1), \mu(1))$ . We have that

$$p_0(x, \lambda, \mu, 1) = \begin{cases} \lambda_{\omega j}^i - \mu_{\omega}^i, & \omega \in \Omega, i \in N, j \in M_{\omega}^i, \\ \lambda_{\omega j}^i x_{\omega j}^i - x_{\omega j}^{0,i}, & \omega \in \Omega, i \in N, j \in M_{\omega}^i, \\ \sum_{j \in M_{\omega}^i} x_{\omega j}^i - 1, & \omega \in \Omega, i \in N. \end{cases}$$

The Jacobian matrix is

$$Dp_0(x, \lambda, \mu, 1) = \begin{pmatrix} 0 & I_{\bar{m} \times \bar{m}} & \text{diag}(e_{\omega j}^{i\top})_{\bar{m} \times nd} \\ (\text{diag}(\lambda_{\omega j}^i))_{\bar{m} \times \bar{m}} & (\text{diag}(x_{\omega j}^i))_{\bar{m} \times \bar{m}} & 0 \\ (\text{diag}(e_{\omega j}^i))_{nd \times \bar{m}} & 0 & 0 \end{pmatrix}$$

It holds by Theorem 1 that, for every  $\omega \in \Omega$ ,  $i \in N$ , and  $j \in M_{\omega}^i$ ,  $x_{\omega j}^i(1) = x_{\omega j}^{0,i}$  and  $\lambda_{\omega j}^i(1) = 1$ . We denote  $M = Dp_0(x(1), \lambda(1), \mu(1), 1)$ ,  $E = \text{diag}(e_{\omega j}^i)$ , and  $X^0 = \text{diag}(x_{\omega j}^{0,i})$ .

Let  $z = (z_1, z_2, z_3) \in \mathbb{R}^{\bar{m}} \times \mathbb{R}^{\bar{m}} \times \mathbb{R}^{nd}$  be such that  $Mz = 0$ . We show that  $z = 0$ , which proves that  $M$  has full rank. From  $Mz = 0$ , it follows that

$$z_2 + E^{\top} z_3 = 0, \tag{15}$$

$$z_1 + X^0 z_2 = 0, \tag{16}$$

$$E z_1 = 0. \tag{17}$$

We multiply (15) by  $z_1^{\top}$  and obtain

$$0 = z_1^{\top} z_2 + z_1^{\top} E^{\top} z_3 = z_1^{\top} z_2, \tag{18}$$

where the second equality follows from (17). We multiply (16) by  $z_2^{\top}$  and obtain

$$0 = z_2^{\top} z_1 + z_2^{\top} X^0 z_2 = z_2^{\top} X^0 z_2, \tag{19}$$

where the last equality follows from (18). Since  $X^0$  is a diagonal matrix with strictly positive elements, (19) implies  $z_2 = 0$ . Next, (15) implies  $z_3 = 0$  and (16) implies  $z_1 = 0$ .

This completes the proof.

## Appendix II

As mentioned in Section 1, [16] proposed SLTP to compute a stationary equilibrium for a stochastic game. This appendix describes our implementation of SLTP. The starting point of SLTP consists of the vector of best responses against a given common prior belief  $p^0 \in X$ .



This vector of best responses can be computed by various methods like the value iteration algorithm or the policy iteration algorithm, see [24]. Here we follow another approach, where the computation of the vector of best responses is part of the homotopy. For  $t \in [0, 2]$ , we define

$$\zeta(t) = \begin{cases} 0 & \text{if } t \leq 1, \\ (t-1)^2 & \text{if } t > 1, \end{cases}$$

and

$$\nu(t) = \begin{cases} t & \text{if } t \leq 1 + \beta_0, \\ t - \frac{1}{4} \left( \frac{(t-1-2\beta_0)^2}{\beta_0} + 2(t-1-2\beta_0) + \beta_0 \right) & \text{if } 1 + \beta_0 < t < 1 + 3\beta_0, \\ 1 + 2\beta_0 & \text{if } t > 1 + 3\beta_0, \end{cases}$$

where  $\beta_0$  is a sufficiently small positive number to guarantee the differentiability of  $\nu(t)$ . The homotopy system of SLTP reads as

$$\begin{aligned} (1 - \zeta(t)) & \left( (1 - \nu(t)) (u^i(\omega, x_{\omega}^{-i}(y), s_{\omega_j}^i) + \delta \sum_{\bar{\omega} \in \Omega} \pi(\bar{\omega} | \omega, x_{\omega}^{-i}(y), s_{\omega_j}^i) \mu_{\bar{\omega}}^i) \right. \\ & \left. + \nu(t) (u^i(\omega, p_{\omega}^{0,-i}, s_{\omega_j}^i) + \delta \sum_{\bar{\omega} \in \Omega} \pi(\bar{\omega} | \omega, p_{\omega}^{0,-i}, s_{\omega_j}^i) \mu_{\bar{\omega}}^i) \right) - \zeta(t) (x_{\omega_j}^i(y) - x_{\omega_j}^{0,i}) \\ & + \lambda_{\omega_j}^i(y) - \mu_{\omega}^i - t(2-t)\alpha_{\omega_j}^i = 0, & \omega \in \Omega, i \in N, j \in M_{\omega}^i, \\ \sum_{j \in M_{\omega}^i} x_{\omega_j}^i(y) - 1 = 0, & \omega \in \Omega, i \in N. \end{aligned}$$

As the homotopy variable  $t$  descends from 2 to 0, the path approximates an SSPE for the stochastic game of interest.

## References

- [1] Allgower, E. L., & Georg, K. (2012). Numerical Continuation Methods: An Introduction (Vol. 13). Springer Science & Business Media.
- [2] Banks, J., & Duggan, J. (2000). A bargaining model of collective choice. *American Political Science Review*, 94(1), 73-88.
- [3] Chatterjee, K., Dutta B., Ray D., & Sengupta, K. (1993). A noncooperative theory of coalitional bargaining. *The Review of Economic Studies*, 60(2), 463-477.
- [4] Chen, Y., & Dang, C. (2016). A reformulation-based smooth path-following method for computing Nash equilibria. *Economic Theory Bulletin*, 4(2), 231-246.
- [5] Chen, Y., & Dang, C. (2018). A differentiable homotopy method to compute perfect equilibria. Forthcoming in *Mathematical Programming*, 1-33.

- [6] Dang, C. (2013). Simplicial methods for approximating fixed point with applications in combinatorial optimizations. *Handbook of Combinatorial Optimization*, 3015-3056.
- [7] Dang, C., Ye, Y., & Zhu, Z. (2011). An interior-point path-following algorithm for computing a Leontief economy equilibrium. *Computational Optimization and Applications*, 50(2), 223-236.
- [8] Eaves, B. C. (1972). Homotopies for computation of fixed points. *Mathematical Programming*, 3(1), 1-22.
- [9] Eaves, B. C., & Schmedders, K. (1999). General equilibrium models and homotopy methods. *Journal of Economic Dynamics and Control*, 23(9-10), 1249-1279.
- [10] Ericson, R., & Pakes, A. (1995). Markov-perfect industry dynamics: A framework for empirical work. *The Review of Economic Studies*, 62(1), 53-82.
- [11] Fink, A. M. (1964). Equilibrium in a stochastic  $n$ -person game. *Journal of Science of Hiroshima University Series A-I*, 28(1), 89-93.
- [12] Fudenberg, D., & Tirole, J. (1991). *Game Theory*. Cambridge, Massachusetts.
- [13] Harsanyi, J. C., & Selten, R. (1988). *A General Theory of Equilibrium Selection in Games*. MIT Press, Cambridge.
- [14] He, W., & Sun, Y. (2017). Stationary Markov perfect equilibria in discounted stochastic games. *Journal of Economic Theory*, 169, 35-61.
- [15] Herings, P. J. J., & Peeters, R. J. (2001). A differentiable homotopy to compute Nash equilibria of  $n$ -person games. *Economic Theory*, 18(1), 159-185.
- [16] Herings, P. J. J., & Peeters, R. J. (2004). Stationary equilibria in stochastic games: Structure, selection, and computation. *Journal of Economic Theory*, 118(1), 32-60.
- [17] Herings, P. J. J., & Peeters, R. J. (2010). Homotopy methods to compute equilibria in game theory. *Economic Theory*, 42(1), 119-156.
- [18] Herings, P. J. J., & Schmedders, K. (2006). Computing equilibria in finance economies with incomplete markets and transaction costs. *Economic Theory*, 27(3), 493-512.
- [19] Jaskiewicz, A., & Nowak, A. S. (2016). Non-zero-sum stochastic games. *Handbook of Dynamic Game Theory*, 1-64.

- [20] Mas-Colell, A. . (1974). A note on a theorem of f. browder. *Mathematical Programming*, 6(1), 229-233.
- [21] Maskin, E., & Tirole, J. (2001). Markov perfect equilibrium: I. Observable actions. *Journal of Economic Theory*, 100(2), 191-219.
- [22] McKelvey, R. D., & McLennan, A. (1996). Computation of equilibria in finite games. *Handbook of Computational Economics*, 1, 87-142.
- [23] Pakes, A., & McGuire, P. (2001). Stochastic algorithms, symmetric Markov perfect equilibrium, and the curse of dimensionality. *Econometrica*, 69(5), 1261-1281.
- [24] Puterman, M. L. (1994). *Markov Decision Processes, Discrete Stochastic Dynamic Programming*. John Wiley and Sons, Hoboken, New Jersey.
- [25] Scarf, H. (1967). The approximation of fixed points of a continuous mapping. *SIAM Journal on Applied Mathematics*, 15(5), 1328-1343.
- [26] Schanuel, S. H., Simon, L.K., & Zame, W.R. (1991). The algebraic geometry of games and the tracing procedure. In *Game Equilibrium Models II: Methods, Morals and Markets* (pp. 9-43). Springer-Verlag, Berlin.
- [27] Shapley, L. S. (1953). Stochastic games. *Proceedings of the National Academy of Sciences*, 39(10), 1095-1100.
- [28] Sobel, M. J. (1971). Noncooperative stochastic games. *The Annals of Mathematical Statistics*, 42(6), 1930-1935.
- [29] Takahashi, M. (1964). Equilibrium points of stochastic non-cooperative  $n$ -person games. *Journal of Science of the Hiroshima University Series A-I*, 28(1), 95-99.
- [30] Ye, Y. (1991). An  $O(n^3L)$  potential reduction algorithm for linear programming. *Mathematical programming*, 50(1-3), 239-258.
- [31] Ye, Y. (1992). A potential reduction algorithm allowing column generation. *SIAM Journal on Optimization*, 2(1), 7-20.
- [32] Ye, Y. (1997). *Interior Point Algorithms: Theory and Analysis*. Springer, New York.

- [33] Zhan, Y., & Dang, C. (2018). A smooth path-following algorithm for market equilibrium under a class of piecewise-smooth concave utilities. *Computational Optimization and Applications*, 71(2), 381-402.
- [34] Zhu, Z., Dang, C., & Ye, Y. (2012). A FPTAS for computing a symmetric Leontief competitive economy equilibrium. *Mathematical Programming*, 131(1-2), 113-129.