Stationary Equilibria in Stochastic Games: Structure, Selection, and Computation[*]

P. Jean-Jacques Herings[†]   Ronald J.A.P. Peeters[‡]

17th November 2000

Abstract

This paper is the first to introduce an algorithm to compute stationary equilibria in stochastic games, and shows convergence of the algorithm for almost all such games. Moreover, since in general the number of stationary equilibria is overwhelming, we pay attention to the issue of equilibrium selection. We do this by extending the linear tracing procedure to the class of stochastic games, called the stochastic tracing procedure.

From a computational point of view, the class of stochastic games possesses substantial difficulties compared to normal form games. Apart from technical difficulties, there are also conceptual difficulties, for instance the question how to extend the linear tracing procedure to the environment of stochastic games.

We prove that there is a generic subclass of the class of stochastic games for which the stochastic tracing procedure is a compact one-dimensional piecewise differentiable manifold with boundary. Furthermore, we prove that the stochastic tracing procedure generates a unique path leading from any exogenously specified prior belief, to a stationary equilibrium.

A well-chosen transformation of variables is used to formulate an everywhere differentiable homotopy function, whose zeros describe the (unique) path generated by the stochastic tracing procedure. Because of differentiability we are able to follow this path using standard path-following techniques. This yields a globally convergent algorithm that is easily and robustly implemented on a computer using existing software routines.

As a by-product of our results, we extend a recent result on the generic finiteness of stationary equilibria in stochastic games to oddness of equilibria.

JEL classification: C63, C72, C73

Keywords: Game theory; Stochastic games; Computation of equilibria; Linear tracing procedure.

1 Introduction

Many economic situations of interest can be modeled as a stochastic game. Recent work of for instance Olley and Pakes (1996), Pakes and Ericson (1998), Pakes and McGuire (1996), and Bergemann and Välimäki (1996) is devoted to the application of stochastic games to problems emerging in the industrial organization literature. In our opinion, further progress in this research program can be made by developing methods to solve stochastic games numerically. Numerical solution methods allow researchers to go back and forth between the implications of economic theory and the characteristics of alternative datasets. See also Judd (1997) and McKelvey and McLennan (1996) for an expression of the important role of computational methods in the further development of economic theory.

The aim of this paper is to present an algorithm to compute a stationary equilibrium for an arbitrary finite n-person discounted stochastic game. Even though such a game may possess many non-stationary equilibria, there are good reasons to focus on stationary equilibria. Several motivations for analyzing stationary equilibria can be found in Maskin and Tirole (1997). Stationary strategies prescribe the simplest form of behavior that is consistent with rationality, stationarity captures the notion that 'bygones are bygones' more completely than does the concept of subgame-perfect equilibrium, and it embodies the principle that 'minor causes should have minor effects', that is, only those aspects of the past that are 'significant' should have an appreciable influence on behavior. The pragmatic motivations they give are that in applied theory, the focus on stationary strategies allows for clean, unobstructed analysis of the influence of the state variables, that stationary strategies substantially reduce the number of parameters to be estimated in dynamic (econometric) models, and that stationary models can be simulated.

For many normal form games there is a vast multiplicity of equilibria, see McLennan (1999). There is no reason to expect that the situation is different for the multiplicity of stationary equilibria in stochastic games. For this reason, we also find it essential to provide an algorithm that selects a particular stationary equilibrium. We introduce a variant of the tracing procedure which allows selection within the class of stationary equilibria of stochastic games. Finally, the algorithm should be fast in that it allows for the computation of a stationary equilibrium for non-trivial stochastic games within reasonable time limits.

Stochastic games were introduced by Shapley (1953). He considered both finite and infinite horizon two-person zero-sum stochastic games with finite state space and finite action spaces. Shapley proved that such games have a value and that both players possess optimal stationary strategies with respect to the discounted payoff criterion.

Fink (1964), Takahashi (1964), and Sobel (1971) extended Shapley's model to general

n-person stochastic games. For the model with finite state space and finite action spaces they showed the existence of a stationary equilibrium. In Breton, Filar, Haurie, and Schultz (1986), Schultz (1986), Filar, Schultz, Thuijsman, and Vrieze (1991), and Breton (1991) the problem of finding discounted stationary equilibria in the general n-person stochastic game is reduced to that of finding a global minimum in a nonlinear program with linear constraints. Solving this nonlinear program is equivalent to solving a certain nonlinear system for which it is known that the objective value in the global minimum is zero. But, as is noted in Breton (1991), the convergence of an optimization algorithm to the global optimum is not guaranteed.

In this paper we propose an algorithm that is shown to converge to a stationary equilibrium for a generic n-person stochastic game. This algorithm is the first globally convergent algorithm to solve for an equilibrium in an arbitrary n-person stochastic game. The algorithm also deals with the equilibrium selection problem in that it computes the stationary equilibrium selected by a natural extension of the linear tracing procedure of Harsanyi and Selten (1988), which we call the stochastic tracing procedure.

In Harsanyi and Selten (1988) the tracing procedure is defined for normal form games and for extensive form games with a perfect recall information structure. Algorithms to compute the Nash equilibrium selected by the tracing procedure in normal form games are proposed in van den Elzen and Talman (1999), Herings and van den Elzen (1998), and Herings and Peeters (1999). For extensive form games, Harsanyi and Selten (1988) first transform the game into one in standard form and subsequently define the tracing procedure for that class of games. Computation of the Nash equilibrium selected by the tracing procedure in extensive form games is the topic of von Stengel, van den Elzen, and Talman (1996), who invoke the sequence form to calculate such equilibria efficiently. Since expected utility in stationary strategies does not hold in stochastic games, it is not possible to transform a stochastic game into one in standard form. The way to extend the tracing procedure to the class of stochastic games is neither straightforward nor unique.

The algorithm belongs to the class of homotopy methods. The formulation as a differentiable homotopy makes it possible to apply standard path-following techniques that are available in professionally programmed software. This makes implementation on a computer an easy exercise. As a by-product of the proof that the algorithm converges for a generic stochastic game, we obtain the result that for a generic stochastic game the stochastic tracing procedure yields a path leading to a unique stationary equilibrium.

As another by-product of our convergence proof, we obtain an extension of a recent result of Haller and Lagunoff (2000). Their main result is that the set of stationary equilibria in a stochastic game is generically finite. A corollary to our main result is that this

finite number of equilibria is odd.

The paper has been organized as follows. In Section 2 we describe the model of the finite discounted stochastic game. In Section 3 the restriction to stationary strategies is made and the set of stationary equilibrium strategies is characterized. The definition of the stochastic tracing procedure is given in Section 4. In Section 5, the properties of the stochastic tracing procedure are studied. It is shown that for almost every stochastic game, the stochastic tracing procedure is formed by a finite union of arcs and loops. Using a well-chosen transformation of variables, the stochastic tracing procedure is described by the zeros of an everywhere differentiable homotopy function in Section 6. Section 7 discusses the implementation of the homotopy algorithm and provides some numerical results.

2 Finite Discounted Stochastic Games

In this paper we study finite discounted stochastic games. Such a game is defined as follows.

Definition 2.1 A *finite discounted stochastic game* is an ordered sextuple

$$\Gamma = \left\langle N, \Omega, \{S^i_\omega\}_{(i,\omega)\in N\times\Omega}, \{u^i\}_{i\in N}, \pi, \delta \right\rangle,$$

where N, Ω and S^i_ω are finite non-empty sets, $S_\omega = \bigtimes_{i\in N} S^i_\omega$, u^i is a real-valued function on the set $H = \{(\omega, s_\omega) \,|\, \omega \in \Omega,\ s_\omega \in S_\omega\}$, where π is a map $\pi : H \to \Delta(\Omega)$ with $\Delta(\Omega)$ the family of probability distributions on the space Ω and δ is a discount factor.

The game parameters have the following meaning.

- $N = \{1, \ldots, n\}$ is the player set.

- $\Omega = \{\omega_1, \ldots, \omega_z\}$ is the state space.

- $S^i_\omega = \{s^i_{\omega 1}, \ldots, s^i_{\omega m^i_\omega}\}$ is the action set of player $i \in N$ in state $\omega \in \Omega$.

- $u^i : H \to \mathbb{R}$ is the payoff function of player $i \in N$; if in state $\omega \in \Omega$ the players action choices are $s_\omega = (s^1_\omega, \ldots, s^n_\omega) \in S_\omega = \bigtimes_{i\in N} S^i_\omega$, then player i gets an instantaneous payoff of $u^i(\omega, s_\omega)$.

- $\pi : H \to \Delta(\Omega)$ is the transition map. For each $(\omega, s_\omega) \in H$, we can identify $\pi(\omega, s_\omega)$ with the vector $(\pi(\omega_1 \,|\, \omega, s_\omega), \ldots, \pi(\omega_z \,|\, \omega, s_\omega))$. Here $\pi(\bar\omega \,|\, \omega, s_\omega)$ represents the probability that the system jumps to state $\bar\omega$ if in state ω the strategy-tuple s_ω is played. Hence, $\pi(\bar\omega \,|\, \omega, s_\omega) \geq 0$ and $\sum_{\bar\omega \in \Omega} \pi(\bar\omega \,|\, \omega, s_\omega) = 1$.

- $\delta \in (0, 1)$ is the discount factor and is used to discount future payoffs.

Such a stochastic game corresponds to a dynamic system which can be in different states and where at certain stages the players can influence the course of the play. We consider the infinite horizon model and the set of stages is assumed to be identical with the set $\mathbb{N} = \{0, 1, \ldots\}$. Players know the game itself and that this knowledge is common knowledge among all the players. Moreover, the initial state ω^0 at stage $k = 0$ is common knowledge to the players.

For further analysis the following notations are introduced: $S^* = \bigcup_{(i,\omega) \in N \times \Omega} S_\omega^i$, $S = \bigtimes_{(i,\omega) \in N \times \Omega} S_\omega^i$, $m^* = |S^*| = \sum_{(i,\omega) \in N \times \Omega} m_\omega^i$, and $m = |S| = \prod_{(i,\omega) \in N \times \Omega} m_\omega^i$. The game proceeds as follows. All players i select at the initial state, simultaneously and independently of each other (possibly by a chance experiment) an action $s_{\omega^0}^i \in S_{\omega^0}^i$. Now two things happen, both depending on the current state ω^0 and the action choices $s_{\omega^0}^1, \ldots, s_{\omega^0}^n$,

(a) player i earns $u^i(\omega^0, s_{\omega^0})$,

(b) the system jumps to the next state ω^1 according to the outcome of a chance experiment. The probability that the next state will be $\bar{\omega}$ equals $\pi(\bar{\omega} \mid \omega^0, s_{\omega^0})$.

Subsequently, prior to the next stage $k = 1$, all players are informed about the previous actions chosen by the players, and of the new state ω^1. At stage $k = 1$, the above procedure is repeated, starting from the state ω^1.

We assume that the game is of perfect recall, i.e., at each stage each player remembers all past actions chosen by all players and all past states that have occurred. Note that for finite stochastic games, each stage game resembles a normal form game. However, contrary to the situation with normal form games, the game does not consist of a single play, but jumps according to the probability measure $\pi(\cdot \mid \omega, s_\omega)$ to the next state and continues dynamically. In choosing an action in a certain state, a player not only takes into account the immediate payoff, but also his opportunities in future states.

Like in normal form games, the players are allowed to randomize their pure actions. A mixed strategy of player i in state ω is a probability distribution on S_ω^i. We identify the set of all probability distributions on S_ω^i with $\Sigma_\omega^i = \{\sigma_\omega^i \in \mathbb{R}_+^{m_\omega^i} \mid \sum_{s_{\omega j}^i \in S_\omega^i} \sigma_{\omega j}^i = 1\}$. For $\sigma_\omega^i \in \Sigma_\omega^i$, the probability assigned to pure strategy $s_{\omega j}^i$ is given by $\sigma_{\omega j}^i$. The strategy space of the normal form game in state ω is therefore equal to $\Sigma_\omega = \bigtimes_{i \in N} \Sigma_\omega^i$. Given a mixed strategy combination $\sigma_\omega \in \Sigma_\omega$ and a strategy $\bar{\sigma}_\omega^i \in \Sigma_\omega^i$, we denote by $(\sigma_\omega^{-i}, \bar{\sigma}_\omega^i)$ the mixed strategy that results from replacing σ_ω^i by $\bar{\sigma}_\omega^i$. If a mixed strategy combination $\sigma_\omega \in \Sigma_\omega$ is played, then the instantaneous expected payoff of player i is denoted by $u^i(\omega, \sigma_\omega)$ and the expected transition to state $\bar{\omega}$ by $\pi(\bar{\omega} \mid \omega, \sigma_\omega)$.

5

At the different stages this randomization may depend on the history of the game up to that stage.

Definition 2.2 A *history* up to a stage k is a sequence $h^k = (\omega^0, s_{\omega^0}, \omega^1, s_{\omega^1}, \ldots, \omega^{k-1}, s_{\omega^{k-1}})$ that could have occurred up to time k, $k \geq 1$. Here ω^κ represents the state and s_{ω^κ} the actions of the players at stage κ, $\kappa = 0, \ldots, k-1$.

Obviously, the set of histories up to time k equals $H^k = \mathsf{X}_{\kappa=0}^{k-1} H$, i.e., the k-fold Cartesian product of H.

A *behavior strategy* σ^{ik} of player i specifies for each stage k, each state ω^k at time k, and each history h^k a probability distribution $\sigma^{ik}(h^k, \omega^k)$ on the action space $S_{\omega^k}^i$ of player i in state ω^k. Then $\sigma^{ik}(s_{\omega^k}^i \mid h^k, \omega^k)$ is the probability with which player i chooses action $s_{\omega^k}^i \in S_{\omega^k}^i$ at time k if state ω^k and history h^k have occurred.

A *stationary strategy* for player i is a behavior strategy for which $\sigma^{ik}(h^k, \omega^k)$ is of the form $\sigma^i(\omega^k)$, i.e., a stationary strategy is a history and time independent strategy which depends on the state only. In the sequel, a stationary strategy for player i will be denoted by the symbol ρ^i. We define $\rho^i = (\rho_{\omega_1}^i, \ldots, \rho_{\omega_z}^i)$, where ρ_ω^i is a probability measure on the action space S_ω^i for each $\omega \in \Omega$. So $\rho_\omega^i \in \Sigma_\omega^i$. If player i decides to play the stationary strategy ρ^i, then every time that the system is in state ω, player i selects his pure action according to ρ_ω^i.

The above concepts lead to the following definition.

Definition 2.3 A *behavior strategy* σ^i for player i is a sequence $\sigma^{i0}, \sigma^{i1}, \ldots$ where $\sigma^{i0} \in \Sigma^i := \mathsf{X}_{\omega \in \Omega} \Sigma_\omega^i$ and $\sigma^{ik} : H^k \to \Sigma^i$ for all $k \geq 1$. A *stationary strategy* ρ^i for player i is an element of Σ^i.

Given initial state ω and strategy σ, the stream of expected payoffs is evaluated by

$$U^i(\omega, \sigma) := \sum_{k=0}^{\infty} \delta^k \cdot U^{ik}(\omega, \sigma),$$

where $U^{ik}(\omega, \sigma)$ denotes the expected instantaneous utility at stage k. Here, $U^i(\omega, \sigma)$ equals the total discounted expected payoff of player i when the discount factor equals δ, the starting state is ω and the strategy-tuple σ is played. Since the state and action spaces are assumed to be finite, $U^i(\omega, \sigma)$ exists.

Definition 2.4 A strategy-tuple σ is an *equilibrium* if and only if σ^i is a best response to σ^{-i} for all $i \in N$.

If $n = 2$ and the payoffs are zero-sum, then the definition characterizes a minimax solution (see Shapley (1953)).

3 Stationary Equilibria

In the sequel of this paper we will restrict ourselves to stationary strategies.[1] Suppose that ω is the initial state and the players decide to play a stationary strategy-tuple ρ. The total discounted expected payoff of player i is denoted by $U^i(\omega, \rho)$. The instantaneous payoff player i obtains in stage $k = 0$ equals $u^i(\omega, \rho_\omega)$. The probability that at the next stage the state will be $\bar{\omega}$ equals $\pi(\bar{\omega} \,|\, \omega, \rho_\omega)$. This gives rise to the following theorem (see e.g. Fink (1964)).

Theorem 3.1 *When ρ is a stationary strategy-tuple and ω is the initial state, the expected payoffs are given by the following recursive formula*

$$U^i(\omega, \rho) = u^i(\omega, \rho) + \delta \cdot \sum_{\bar{\omega} \in \Omega} \pi(\bar{\omega} \,|\, \omega, \rho_\omega) U^i(\bar{\omega}, \rho).$$

Definition 3.2 A stationary strategy-combination $\rho \in \Sigma$ is a *stationary equilibrium* if it is a Nash equilibrium in stationary strategies.

It is a very well-known result that there exists a stationary best response to stationary strategies ρ^{-i}. Therefore, the following theorem holds.

Theorem 3.3 *A stationary strategy-combination $\bar{\rho} \in \Sigma$ is a stationary equilibrium if for all i and all ω*

$$U^i(\omega, \bar{\rho}) \geq U^i(\omega, \bar{\rho}^{-i}, \rho^i), \qquad \rho^i \in \Sigma^i.$$

Another famous result, see Fink (1964), Takahashi (1964), or Sobel (1971), is the existence of a stationary equilibrium.

Theorem 3.4 *Every finite discounted stochastic game has a stationary equilibrium.*

The remainder of this section is devoted to the characterization of the set of stationary equilibria that is useful for numerical computations. Given that the other players play ρ^{-i} and the initial state is $\bar{\omega}$, player i faces the Markov decision problem of maximizing

$$U^i(\bar{\omega}, \rho^{-i}, \rho^i)$$

subject to

$$\rho^i_{\omega j} \geq 0 \qquad (s^i_{\omega j} \in S^i_\omega, \omega \in \Omega),$$
$$\sum_{s^i_{\omega j} \in S^i_\omega} \rho^i_{\omega j} - 1 = 0 \qquad (\omega \in \Omega),$$

---

[1]In fact, we restrict ourselves to a class which is even tighter. Namely, the one of perfect stationary strategies, where the strategy chosen is as if each state is reached or might be the initial state.

for all possible initial states $\bar{\omega} \in \Omega$. If we define by $\mu_\omega^i(k)$ the present value of the total expected payoff for a system in state ω with k transitions remaining, we obtain the basic recurrence relation

$$\mu_\omega^i(k+1) = u^i(\omega, \rho_\omega) + \delta \cdot \sum_{\bar{\omega} \in \Omega} \pi(\bar{\omega} \mid \omega, \rho_\omega) \mu_{\bar{\omega}}^i(k)$$

or in vector form

$$\mu^i(k+1) = u^i(\rho) + \delta \Pi(\rho) \mu^i(k).$$

Note that[2]

$$
\begin{aligned}
\mu^i &= \lim_{k \to \infty} \mu^i(k) = \lim_{k \to \infty} \left(\sum_{\ell=0}^{k-1} (\delta \Pi(\rho))^\ell u^i(\rho) + (\delta \Pi(\rho))^k \mu^i(0) \right) \\
&= \sum_{\ell=0}^{\infty} (\delta \Pi(\rho))^\ell u^i(\rho) = [I - \delta \Pi(\rho)]^{-1} u^i(\rho).
\end{aligned}
$$

The vector μ^i may be called the vector of present values of player i, because each of its elements μ_ω^i is the present value of an infinite number of future expected payoffs discounted by the discount factor δ with ω the initial state.

Because we are interested in the sequential decision process for large k (in fact for k equal to infinity), we substitute the present values $\mu_\omega^i = \lim_{k \to \infty} \mu_\omega^i(k)$ for the quantities $\mu_\omega^i(k)$ to obtain the recursive relation

$$\mu_\omega^i = u^i(\omega, \rho_\omega) + \delta \cdot \sum_{\bar{\omega} \in \Omega} \pi(\bar{\omega} \mid \omega, \rho_\omega) \mu_{\bar{\omega}}^i.$$

Given that the other players play ρ^{-i}, player i maximizes μ_ω^i subject to $\rho_\omega^i \in \Sigma_\omega^i$ for all $\omega \in \Omega$.

Because player i maximizes each μ_ω^i for $\omega \in \Omega$, he also maximizes $\sum_{\omega \in \Omega} \mu_\omega^i$. Given that the other players play ρ^{-i}, player i maximizes

$$\sum_{\omega \in \Omega} \mu_\omega^i$$

subject to

$$
\begin{aligned}
&u^i(\omega, \rho_\omega^{-i}, \rho_\omega^i) + \delta \cdot \sum_{\bar{\omega} \in \Omega} \pi(\bar{\omega} \mid \omega, \rho_\omega^{-i}, \rho_\omega^i) \mu_{\bar{\omega}}^i - \mu_\omega^i = 0 \qquad (\omega \in \Omega), \\
&\sum_{s_{\omega j}^i \in S_\omega^i} \rho_{\omega j}^i - 1 = 0 \qquad (\omega \in \Omega), \\
&\rho_{\omega j}^i \geq 0 \qquad (s_{\omega j}^i \in S_\omega^i, \omega \in \Omega).
\end{aligned}
$$

---

[2]By Hadamard's theorem all eigenvalues of the matrix between the squared brackets have absolute value larger than 0. Thus zero is not an eigenvalue of that matrix and the inverse exists. Furthermore, the spectral radius of $\delta \Pi(\rho)$ is less than one. Therefore it holds that $[I - \delta \Pi(\rho)]^{-1} = \sum_{k=0}^{\infty} \delta^k \Pi(\rho)^k$. From this it is easily seen that the value of element (i, j) of the matrix $[I - \delta \Pi(\rho)]^{-1}$ gives the discounted expected number of times that the state is j when i is the initial state and ρ is played.

The necessary and sufficient first-order conditions for this maximization problem are

$$\alpha_\omega^i \left\{ u^i(\omega, \rho_\omega^{-i}, s_{\omega j}^i) + \delta \cdot \sum_{\bar{\omega} \in \Omega} \pi(\bar{\omega} \,|\, \omega, \rho_\omega^{-i}, s_{\omega j}^i) \mu_{\bar{\omega}}^i \right\} + \gamma_{\omega j}^i - \beta_\omega^i = 0$$
$$(s_{\omega j}^i \in S_\omega^i, \omega \in \Omega),$$

$$\gamma_{\omega j}^i \geq 0, \quad \rho_{\omega j}^i \geq 0, \quad \gamma_{\omega j}^i \rho_{\omega j}^i = 0 \qquad (s_{\omega j}^i \in S_\omega^i, \omega \in \Omega),$$

$$\sum_{s_{\omega j}^i \in S_\omega^i} \rho_{\omega j}^i - 1 = 0 \qquad (\omega \in \Omega),$$

$$u^i(\omega, \rho_\omega^{-i}, \rho_\omega^i) + \delta \cdot \sum_{\bar{\omega} \in \Omega} \pi(\bar{\omega} \,|\, \omega, \rho_\omega^{-i}, \rho_\omega^i) \mu_{\bar{\omega}}^i - \mu_\omega^i = 0 \qquad (\omega \in \Omega),$$

$$\alpha^i [I - \delta \Pi(\rho^{-i}, \rho^i)] = \mathbb{1}.$$

Here, α, β and γ are the Lagrange multipliers of the first, second and third set of constraints, and $\mathbb{1}$ denotes the vector containing ones only. Then, for ρ^i a best response to ρ^{-i},

$$
\begin{aligned}
0 &= \alpha_\omega^i \left\{ u^i(\omega, \rho_\omega^{-i}, \rho_\omega^i) + \delta \cdot \sum_{\bar{\omega} \in \Omega} \pi(\bar{\omega} \,|\, \omega, \rho_\omega^{-i}, \rho_\omega^i) \mu_{\bar{\omega}}^i \right\} \\
&\qquad\qquad + \sum_{s_{\omega j}^i \in S_\omega^i} \gamma_{\omega j}^i \rho_{\omega j}^i - \sum_{s_{\omega j}^i \in S_\omega^i} \beta_\omega^i \rho_{\omega j}^i \\
&= \alpha_\omega^i \mu_\omega^i - \beta_\omega^i \qquad (\omega \in \Omega).
\end{aligned}
$$

Therefore, since $\alpha^i = [I - \delta \Pi(\rho^{-i}, \rho^i)]^{-1} \mathbb{1} = \frac{1}{1-\delta} \mathbb{1} \gg 0$,[3]

$$\frac{\beta_\omega^i}{\alpha_\omega^i} = \mu_\omega^i.$$

After division by α_ω^i, the following necessary and sufficient conditions remain, where $\lambda_{\omega j}^i$ is defined as the ratio of $\gamma_{\omega j}^i$ and α_ω^i

$$u^i(\omega, \rho_\omega^{-i}, s_{\omega j}^i) + \delta \cdot \sum_{\bar{\omega} \in \Omega} \pi(\bar{\omega} \,|\, \omega, \rho_\omega^{-i}, s_{\omega j}^i) \mu_{\bar{\omega}}^i + \lambda_{\omega j}^i - \mu_\omega^i = 0 \qquad (s_{\omega j}^i \in S_\omega^i, \omega \in \Omega),$$

$$\lambda_{\omega j}^i \geq 0, \quad \rho_{\omega j}^i \geq 0, \quad \lambda_{\omega j}^i \rho_{\omega j}^i = 0 \qquad (s_{\omega j}^i \in S_\omega^i, \omega \in \Omega),$$

$$\sum_{s_{\omega j}^i \in S_\omega^i} \rho_{\omega j}^i - 1 = 0 \qquad (\omega \in \Omega).$$

Here, $\lambda_{\omega j}^i$ is the shadowprice of playing strategy $s_{\omega j}^i$, i.e., the disutility from a one-shot deviation at $t = 0$ of a marginal increase in the probability $\rho_{\omega j}^i$ by which pure strategy $s_{\omega j}^i$ is played, and μ_ω^i is the expected payoff of player i when the initial state is ω, ρ^{-i} is played by his opponents, and player i chooses a best response. The last equality, $\sum_{s_{\omega j}^i \in S_\omega^i} \rho_{\omega j}^i - 1 = 0$, makes sure that ρ_ω^i is a member of Σ_ω^i for all $\omega \in \Omega$.

Since for a stationary equilibrium it holds that a strategy-tuple constitutes mutually best responses, we have found that the set of stationary equilibria can be fully characterized by the system of equalities and inequalities in Theorem 3.5.

---

[3]Since the value of element (i, j) of the matrix $[I - \delta \Pi(\rho)]^{-1}$ gives the discounted expected number of times that the state is j when i is the initial state and ρ is played, the rows sum up to $\frac{1}{1-\delta}$, which is larger than 0 (even large than 1).

Theorem 3.5 *A strategy $\rho \in \Sigma$ is a stationary equilibrium if and only if it is part of a solution to*

$$u^i(\omega, \rho_\omega^{-i}, s_{\omega j}^i) + \delta \cdot \sum_{\bar{\omega} \in \Omega} \pi(\bar{\omega} \mid \omega, \rho_\omega^{-i}, s_{\omega j}^i) \mu_{\bar{\omega}}^i + \lambda_{\omega j}^i - \mu_\omega^i = 0$$
$$(s_{\omega j}^i \in S_\omega^i, \omega \in \Omega, i \in N),$$

$$\lambda_{\omega j}^i \geq 0 \quad \rho_{\omega j}^i \geq 0, \quad \lambda_{\omega j}^i \rho_{\omega j}^i = 0 \qquad (s_{\omega j}^i \in S_\omega^i, \omega \in \Omega, i \in N),$$

$$\sum_{s_{\omega j}^i \in S_\omega^i} \rho_{\omega j}^i - 1 = 0 \qquad (\omega \in \Omega, i \in N).$$

The system above suggests that only one-shot deviations have to be considered. We explain intuitively the reason of this surprising phenomenon.

Suppose the other players play stationary strategies ρ^{-i} and suppose that ρ^i is the best stationary response of player i. If player i is not able to improve his utility by a deviation to his strategy ρ^i in one stage only, then it follows by a backward induction argument, that neither finitely many deviations to his strategy will make player i better off. Suppose now that player i can improve his utility by infinitely many changes. Then, by a profit-to-go argument, player i can also increase his payoff by finitely many changes, which is not possible.[4]

We end this section with an example in which it is shown that expected utility in stationary strategies does not hold for the class of stochastic games. This causes a number of technical difficulties for the convergence proof of our algorithm.

Example 3.6 The stochastic game of Figure 1 is a game in which there are two states p_H and p_L and one player. In each state the player can choose between actions q_H and q_L. In the upper-left corner of each square the intermediate payoff of the player is displayed. In the lower-right corner the transition probabilities are given. So, if the player chooses q_H when the state is p_H, then he earns 1 today and the next state will be p_L with probability 1. If the player chooses q_L when the state is p_L, then he earns 1 today and the next state will be p_H for sure. Future payoffs are discounted by a factor $1/2$.

Consider the strategies $\rho^- = (q_L, q_H)$ and $\rho^+ = (q_H, q_L)$.[5] Then $U(p_L, \rho^-) = U(p_H, \rho^-) = 1$, $U(p_L, \rho^+) = 2/3$ and $U(p_H, \rho^+) = 4/3$. However,

$$U(p_H, \tfrac{1}{2}\rho^- + \tfrac{1}{2}\rho^+) = \tfrac{6}{5} \quad \neq \quad \tfrac{7}{6} = \tfrac{1}{2} \cdot 1 + \tfrac{1}{2} \cdot \tfrac{4}{3} = \tfrac{1}{2} U(p_H, \rho^-) + \tfrac{1}{2} U(p_H, \rho^+),$$
$$U(p_L, \tfrac{1}{2}\rho^- + \tfrac{1}{2}\rho^+) = \tfrac{4}{5} \quad \neq \quad \tfrac{5}{6} = \tfrac{1}{2} \cdot 1 + \tfrac{1}{2} \cdot \tfrac{2}{3} = \tfrac{1}{2} U(p_L, \rho^-) + \tfrac{1}{2} U(p_L, \rho^+).$$

---

[4]Define $M = \max_{(i,\omega,s_\omega)} |u^i(\omega, s_\omega)|$. Then the maximum payoff a player can earn from time k on is bounded from above by $\delta^k(1 + \delta + \delta^2 + \ldots)M = \frac{\delta^k}{1-\delta}M$, the so-called maximum 'profit-to-go' value. Suppose player i is able to improve his utility by ε by means of infinitely many changes. When k grows large the profit-to-go value is at a certain point less then ε (this is when $k > \log(\frac{\varepsilon(1-\delta)}{M})/\log(\delta)$). This means that the utility improvement by changes until time k (finitely many changes) was positive.

[5]The first argument is the strategy when the state is p_H and the second when the state is p_L.
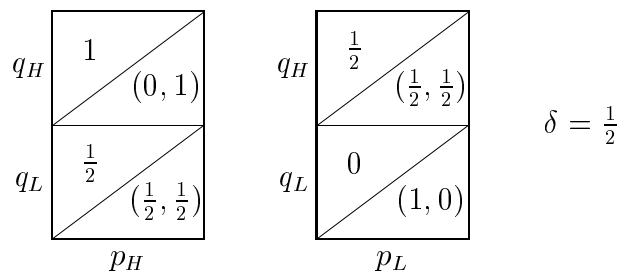
q_H 1 $(0,1)$ q_H $\frac{1}{2}$ $(\frac{1}{2}, \frac{1}{2})$ $\delta = \frac{1}{2}$

q_L $\frac{1}{2}$ $(\frac{1}{2}, \frac{1}{2})$ q_L 0 $(1,0)$

p_H p_L

Figure 1: Example.

So, expected utility does not hold. $\triangleleft$

In the literature on the computation of Nash equilibria in normal form games, a distinction is made between 2-player games on the one hand and 3 or more players on the other hand. For the class of 2-player games exact algorithms are possible, because of the bilinear structure of such games (see, for instance the algorithm of Lemke and Howson (1964)). For stochastic games this distinction disappears. The system of equations of Theorem 3.5 is not bilinear, even for 2-player games. The paper of Parthasarathy and Raghavan (1981) presents an example of a 2-player stochastic game with only rational numbers of payoffs and transition probabilities. The unique Nash equilibrium involves strategies with irrational probabilities. This means that the ordered field property does not hold. A straightforward application of a Lemke-Howson type algorithm to 2-player stochastic games is therefore not possible. This shows one more time that the class of stochastic games is considerably more difficult than the class of normal form games.

4 The Stochastic Tracing Procedure

The linear tracing procedure as presented in Harsanyi and Selten (1988) models a process of convergent expectations by which rational players will come to adopt, and expect each other to adopt, a particular equilibrium as a solution for a given game. Before applying the tracing procedure, every player is assumed to have a subjective probability distribution expressing his expectation about the strategic choices of the other players. Each player is assumed to use the same theory to determine his subjective probability distributions, which makes that all players have the same expectations about the other players. This common subjective probability distribution is called the prior. In the naive Bayesian approach, all players choose best responses to their prior beliefs and would in this way reach a strategy-combination that does not constitute an equilibrium in general. In the linear

11

tracing procedure, the information on the best responses is only gradually fed back into the expectations of the players. As the linear tracing procedure proceeds, both the priors and their best responses will gradually change until both converge to some equilibrium of the game.

In Harsanyi and Selten (1988) the linear tracing procedure is defined for normal form games and for extensive form games with a perfect recall information structure. For a normal form game $\Gamma = \langle N, \{S^i\}_{i \in N}, \{u^i\}_{i \in N} \rangle$ and a prior $p \in \Sigma$ the linear tracing procedure is defined by tracing a curve in the set of Nash equilibria of the games $\Gamma^t = \langle N, \{S^i\}_{i \in N}, \{v^i(t)\}_{i \in N} \rangle$ for $t \in [0, 1]$, where $v^i(t, s) = (1 - t)u^i(p^{-i}, s^i) + tu^i(s)$. For extensive form games, they first transform the game into one in standard form and subsequently define the tracing procedure for that class of games. They did not define the tracing procedure for stochastic games, which are games with instantaneous payoffs and infinite time horizon. Since expected utility in stationary strategies does not hold in stochastic games (see Example 3.6) it is not possible to transform this game into one in standard form. The extension of the tracing procedure to the class of stochastic games is far from obvious.

There are at least four ways to extend the tracing procedure of Harsanyi and Selten to the setting of stochastic games. Choices have to be made whether a player holds correlated beliefs within a state or not, and whether a player holds correlated beliefs across time or not. For the extension of the linear tracing procedure to stochastic games that we study in this paper, we assume that beliefs are correlated within states and that they are not correlated across time.

Correlation within states means that when a player knows that some opponent plays according to the prior (which he expects with probability $1 - t$), he expects all opponents to play according to the prior. This is equivalent to the way Harsanyi and Selten define the tracing procedure for normal form games.

Absence of correlation across time means that even when a player knows that his opponents are playing according to the prior today, these opponents might not play according to the prior in future stages. In all future events he faces independent lotteries which assigns probability $1 - t$ to play against the prior strategies of his opponents. Assuming that beliefs are not correlated across time captures the assumption of stationarity. The beliefs of a player depend only on the state reached and not on the time at which it is reached.

Consider some stochastic game Γ and some prior $p \in \Sigma$. For every $t \in [0, 1]$, the stochastic tracing procedure generates a stationary equilibrium of the stochastic game $\Gamma^t = \langle N, \Omega, \{S^i_\omega\}_{\omega \in \Omega, i \in N}, \{v^i(t)\}_{i \in N}, \tilde{\pi}(t), \delta \rangle$, where the instantaneous payoff function $v^i(t)$:

$\Omega \times S \to \mathbb{R}$ of player i is defined by

$$v^i(t; \omega, s_\omega) = (1 - t)u^i(\omega, p_\omega^{-i}, s_\omega^i) + tu^i(\omega, s_\omega)$$

and where the transition mapping $\tilde{\pi}(t)$ is defined by

$$\tilde{\pi}(t; \bar{\omega} \mid \omega, s_\omega) = (1 - t)\pi(\bar{\omega} \mid \omega, p_\omega^{-i}, s_\omega^i) + t\pi(\bar{\omega} \mid \omega, s_\omega).$$

Note that $\tilde{\pi}(t)$ may be different for different players when $t < 1$. This is consistent with the fact that the tracing procedure should be thought of as a reasoning process. The mapping $\tilde{\pi}(t)$ should be thought of as what the players think that the transition probabilities are in the stochastic game Γ^t. The expected payoff of player i is easily shown to satisfy the recursive relation

$$V^i(t; \omega, \rho) = v^i(t; \omega, \rho_\omega) + \delta \cdot \sum_{\bar{\omega} \in \Omega} \tilde{\pi}(t; \bar{\omega} \mid \omega, \rho_\omega) V^i(t; \bar{\omega}, \rho).$$

The stochastic game Γ^0 corresponds to a trivial stochastic game, where all players believe that all their opponents play with probability 1 according to the prior belief. The stochastic game Γ^1 coincides with the original stochastic game Γ. A best response against a strategy combination $\rho^{-i} \in \Sigma^{-i}$ in the stochastic game Γ^t corresponds to a best response against the stationary probability distribution $(1 - t)[p^{-i}] + t[\rho^{-i}]$ on S^{-i} in the stochastic game Γ.

The stochastic tracing procedure $\mathcal{S}(\Gamma, p)$ is defined as the set of pairs (t, ρ) for which it holds that ρ is a stationary equilibrium of the stochastic game Γ^t, i.e.,

$$\mathcal{S}(\Gamma, p) = \Big\{ (t, \rho) \in [0, 1] \times \Sigma \mid \rho \text{ is a stationary equilibrium of } \Gamma^t \Big\}.$$

The stochastic tracing procedure is said to be feasible if there exists a path in $\mathcal{S}(\Gamma, p)$ connecting a best response against the prior to a stationary equilibrium of the stochastic game Γ, i.e., there exists a continuous function $\gamma : [0, 1] \to \mathcal{S}(\Gamma, p)$ such that $\gamma(0) \in \mathcal{S}(\Gamma, p) \cap (\{0\} \times \Sigma)$ and $\gamma(1) \in \mathcal{S}(\Gamma, p) \cap (\{1\} \times \Sigma)$. In general there may be many trajectories $\gamma([0, 1])$ that link a stationary equilibrium of Γ^0 to a stationary equilibrium of Γ^1. If this trajectory is unique, then the stochastic tracing procedure is said to be well-defined. If the stochastic tracing procedure is well-defined, then it selects a unique stationary equilibrium of the stochastic game Γ.

For a simple proof of the feasibility of the linear tracing procedure for normal form games see Herings (2000) and for the well-definedness see Herings and Peeters (1999). The derivation of such properties for the stochastic tracing procedure is part of present paper.

5   Structure of the Stochastic Tracing Procedure

The size of any stochastic game Γ can be characterized by a vector $\eta = (n, z, \{m_\omega^i\}_{\omega=1,\ldots,z}^{i=1,\ldots,n})$, specifying the number of players, the number of states, and the number of pure strategies available to a player in a state. We will call η the size vector of a stochastic game Γ. For any possible size vector η, transition map π consistent with η and discount factor δ, the set $\mathcal{G}(\eta, \pi, \delta)$ of all stochastic games Γ possessing η as their size vector, having π as their transition map, and δ as discount factor, is called the size class generated by (η, π, δ).

Every stochastic game Γ in the size class $\mathcal{G}(\eta, \pi, \delta)$ is characterized by a vector $u(\Gamma)$ that contains exactly nm payoffs. We identify a stochastic game Γ with the vector $u(\Gamma)$, and we identify the size class $\mathcal{G}(\eta, \pi, \delta)$ with the set of all possible real vectors of size nm, that is, with an nm-dimensional Euclidean space. It is now possible to define the distance $\varrho(\Gamma, \Gamma')$ between two stochastic games and the Lebesgue measure of a set of stochastic games.

A given mathematical statement $\mathcal{Z}$ is said to be true for an open set of (almost all, set of generic) stochastic games if, for every possible size class $\mathcal{G}(\eta, \pi, \delta)$, the set $\tilde{\mathcal{G}}(\mathcal{Z})$ of all stochastic games Γ in $\mathcal{G}(\eta, \pi, \delta)$ for which the statement $\mathcal{Z}$ is true is open (has full Lebesgue measure, is open and of full Lebesgue measure) with respect to $\mathcal{G}(\eta, \pi, \delta)$.

Note that we fix η, π and δ. It is possible to fix only η, and to define open sets of (almost all, generic sets of) games if these properties hold for every possible size class $\mathcal{G}(\eta)$. All our proofs can be readily adapted to get results corresponding to these definitions.

Let a subset B^* of S^* be given with the property that for every player i and for every state ω there is at least one pure strategy $s_{\omega j}^i$ in B^*, so $B_\omega^i = B^* \cap S_\omega^i \neq \emptyset$, for every player i, for every state ω. Such a set B^* is called admissible. The sets B^* are used to decompose $\mathcal{S}(\Gamma, p)$ in subsets $\mathcal{S}(\Gamma, p, B^*)$, each having a differentiable manifold structure. The set $\mathcal{S}(\Gamma, p, B^*)$ contains those elements of $\mathcal{S}(\Gamma, p)$ where only strategies in B^* are played with positive probability. It is defined by

$$\mathcal{S}(\Gamma, p, B^*) = \left\{ (t, \rho) \in \mathcal{S}(\Gamma, p) \, \middle| \right.$$
$$s_{\omega j}^i \notin B^* \Rightarrow \rho_{\omega j}^i = 0$$
$$\left. s_{\omega j}^i \in B^* \Rightarrow s_{\omega j}^i \in \operatorname{argmax}_{s_{\omega \ell}^i \in S_\omega^i} V^i(t; \omega, \rho^{-i}, \rho_{-\omega}^i, s_{\omega \ell}^i) \right\}.$$

It follows that

$$\mathcal{S}(\Gamma, p) = \bigcup_{B^*} \mathcal{S}(\Gamma, p, B^*).$$

Two sets $\mathcal{S}(\Gamma, p, B^*)$ and $\mathcal{S}(\Gamma, p, \bar{B}^*)$ can only have a point (t, ρ) in common if there is a

player i and a strategy $s_{\omega j}^i$ such that $\rho_{\omega j}^i = 0$ and $s_{\omega j}^i \in \text{argmax}_{s_{\omega \ell}^i \in S_\omega^i} V^i(t; \omega, \rho^{-i}, \rho_{-\omega}^i, s_{\omega \ell}^i)$, so $s_{\omega j}^i$ is a best response to $(\rho^{-i}, \rho_{-\omega}^i)$ that is played with probability zero.

To analyze the structure of $\mathcal{S}(\Gamma, p)$ and the sets $\mathcal{S}(\Gamma, p, B^*)$, we design systems of equalities and inequalities whose solutions characterize these sets. By Theorem 3.5, an element (t, ρ) belongs to $\mathcal{S}(\Gamma, p)$ if and only if it is part of a solution to

$$v^i(t; \omega, \rho_\omega^{-i}, s_{\omega j}^i) + \delta \cdot \sum_{\bar{\omega} \in \Omega} \tilde{\pi}(t; \bar{\omega} \mid \omega, \rho_\omega^{-i}, s_{\omega j}^i) \mu_{\bar{\omega}}^i + \lambda_{\omega j}^i - \mu_\omega^i = 0$$
$$(s_{\omega j}^i \in S_\omega^i, \omega \in \Omega, i \in N),$$

$$\lambda_{\omega j}^i \geq 0, \quad \rho_{\omega j}^i \geq 0, \quad \lambda_{\omega j}^i \rho_{\omega j}^i = 0 \qquad (s_{\omega j}^i \in S_\omega^i, \omega \in \Omega, i \in N),$$
$$\sum_{s_{\omega j}^i \in S_\omega^i} \rho_{\omega j}^i - 1 = 0 \qquad (\omega \in \Omega, i \in N).$$

Given an admissible subset B^*, we define $\mathcal{O}(\Gamma, p, B^*)$ as the set of solutions (t, ρ, λ, μ) to the following system of equalities and inequalities:

$$(1) \qquad v^i(t; \omega, \rho_\omega^{-i}, s_{\omega j}^i) + \delta \cdot \sum_{\bar{\omega} \in \Omega} \tilde{\pi}(t; \bar{\omega} \mid \omega, \rho_\omega^{-i}, s_{\omega j}^i) \mu_{\bar{\omega}}^i + \lambda_{\omega j}^i - \mu_\omega^i = 0$$
$$(s_{\omega j}^i \in S_\omega^i, \omega \in \Omega, i \in N),$$

$$(2) \qquad \rho_{\omega j}^i = 0 \qquad (s_{\omega j}^i \notin B_\omega^i, \omega \in \Omega, i \in N),$$

$$(3) \qquad \lambda_{\omega j}^i = 0 \qquad (s_{\omega j}^i \in B_\omega^i, \omega \in \Omega, i \in N),$$

$$(4) \qquad \sum_{s_{\omega j}^i \in S_\omega^i} \rho_{\omega j}^i - 1 = 0 \qquad (\omega \in \Omega, i \in N),$$

$$(5) \qquad \rho_{\omega j}^i \geq 0 \qquad (s_{\omega j}^i \in B_\omega^i, \omega \in \Omega, i \in N),$$

$$(6) \qquad \lambda_{\omega j}^i \geq 0 \qquad (s_{\omega j}^i \notin B_\omega^i, \omega \in \Omega, i \in N),$$

$$(7) \qquad t \geq 0,$$

$$(8) \qquad 1 - t \geq 0.$$

The fact that for stochastic games the system of equalities and inequalities differs from the case of normal form games, is not the only difficulty. Since expected utility does not hold, we can not use any property that is derived from it. In particular, it is not even obvious that there is a unique best response to the prior in pure stationary strategies.

Theorem 5.1 implies that $(t, \rho) \in \mathcal{S}(\Gamma, p, B^*)$ if and only if there exists $\lambda \in \mathbb{R}^{m^*}$ and $\mu \in \mathbb{R}^{nz}$ such that the equalities (1)-(4) and the inequalities (5)-(8) are satisfied.

Theorem 5.1 *Let a stochastic game $\Gamma \in \mathcal{G}(\eta)$ and a prior $p \in \Sigma$ be given. For all admissible subsets B^* of S^*, the sets $\mathcal{S}(\Gamma, p, B^*)$ and $\mathcal{O}(\Gamma, p, B^*)$ are C^∞ diffeomorphic.*

Proof Let B^* be an admissible subset of S^*. For every $i \in N$ and $\omega \in \Omega$ we take an element $\bar{s}_\omega^i \in B^*$. We define a function $f : [0, 1] \times \Sigma \to \mathbb{R} \times \mathbb{R}^{m^*} \times \mathbb{R}^{m^*} \times \mathbb{R}^{nz}$ by $f(t, \rho) = (t, \rho, \lambda, \mu)$, where (in vector form with a slight abuse of notation) $\mu^i = [I - \delta \cdot \tilde{\Pi}(t; \rho^{-i}, \bar{s}^i)]^{-1} v^i(t; \rho^{-i}, \bar{s}^i)$

15

and $\lambda_{\omega\ell}^i = \mu_\omega^i - v^i(t; \omega, \rho_\omega^{-i}, s_{\omega\ell}^i) - \delta \cdot \sum_{\bar\omega \in \Omega} \tilde\pi(t; \bar\omega \mid \omega, \rho_\omega^{-i}, s_{\omega\ell}^i)\mu_{\bar\omega}^i$. Then $f(t, \rho) \in \mathcal{O}(\Gamma, p, B^*)$ if and only if $(t, \rho) \in \mathcal{S}(\Gamma, p, B^*)$. Note that f defined in this way is a C^∞ diffeomorphism.
$\square$

From Theorem 5.1 it follows that for all $(t, \rho) \in \mathcal{S}(\Gamma, p, B^*)$ there is a unique λ and a unique μ such that $(t, \rho, \lambda, \mu) \in \mathcal{O}(\Gamma, p, B^*)$. Vice versa, for all $(t, \rho, \lambda, \mu) \in \mathcal{O}(\Gamma, p, B^*)$ it holds that $(t, \rho) \in \mathcal{S}(\Gamma, p, B^*)$.

The analysis of the system of equalities and inequalities (1)-(8) provides the following result.

Theorem 5.2 *For an open set of stochastic games and priors $(\Gamma, p) \in \mathcal{G}(\eta) \times \Sigma$ with full Lebesgue measure, for all admissible subsets B^* of S^*, $\mathcal{S}(\Gamma, p, B^*)$ is a compact 1-dimensional C^∞ manifold with boundary. Moreover, a boundary point (t, ρ) of $\mathcal{S}(\Gamma, p, B^*)$ is either*

(i) *not a boundary point of $\mathcal{S}(\Gamma, p, \bar{B}^*)$ for all $\bar{B}^* \neq B^*$ and lies in $\{0, 1\} \times \Sigma$, or*

(ii) *is a boundary point of exactly one $\mathcal{S}(\Gamma, p, \bar{B}^*)$ with $\bar{B}^* \neq B^*$ and belongs to $(0, 1) \times \Sigma$. Moreover, B^* and $\bar{B}^*$ differ in exactly one element, say $s_{\omega j}^i$, for which $\rho_{\omega j}^i = 0$ and $s_{\omega j}^i$ is a best response to ρ in Γ^t.*

Proof The proof of this theorem is analogous to the proofs of Theorems 3.2 and 3.3 of Herings and Peeters (1999), and involves three lemmas. The only thing left to do is to prove those lemmas for the stochastic game situation, which is done in the Appendix. The inference is then analogously to the proofs of Herings and Peeters (1999), since it only uses the structural manifold properties of the lemmas.

In Lemma A.1, we prove that the set of solutions to (1)-(8) is one-dimensional if the inequalities are strict. Then in Lemma A.2, we prove that the set of solutions to (1)-(8) is zero-dimensional if exactly one of the inequalities is binding. Finally, it is proved in Lemma A.3 that there is no solution to (1)-(8) with more than one binding inequality.

It follows that $\mathcal{O}(\Gamma, p, B^*)$ is a 1-dimensional manifold with boundary, and a point in $\mathcal{O}(\Gamma, p, B^*)$ is a boundary point if and only if exactly one of the inequalities in (5)-(8) holds with equality. By Theorem 5.1 these properties carry over to $\mathcal{S}(\Gamma, p, B^*)$.

It is easily seen that for a boundary point alternative (i) holds when the binding inequality comes from (7) or (8) and that alternative (ii) holds when the binding inequality comes from (5) or (6).
$\square$

This theorem implies that, for almost every Γ and p, for all admissible sets B^* of S^* the

set $\mathcal{S}(\Gamma, p, B^*)$ consists of a finite number of smooth arcs and loops.[6] Each arc has two boundary points. The structure of $\mathcal{S}(\Gamma, p, B^*)$ is therefore a simple one; all kinds of complications like bifurcations, spirals, higher dimensional solution sets, diverging behavior, etc. are excluded.

If $\mathcal{S}(\Gamma, p, B^*)$ has a boundary point in $(0, 1) \times \Sigma$, then there is a unique admissible subset $\bar{B}^*$ such that $\mathcal{S}(\Gamma, p, \bar{B}^*)$ has this boundary point in common with $\mathcal{S}(\Gamma, p, B^*)$. The cardinality of $\bar{B}^*$ is one less or one greater than the cardinality of B^*, depending on whether in the common boundary point a strategy in B^* is played with probability zero, or whether a strategy not in B^* is a best response. The uniqueness of the set $\bar{B}^*$ implies that the sets $\mathcal{S}(\Gamma, p, B^*)$ and $\mathcal{S}(\Gamma, p, \bar{B}^*)$ are nicely linked to each other.

If $\mathcal{S}(\Gamma, p, B^*)$ has a boundary point in $\{0, 1\} \times \Sigma$, then this point does not belong to any other set $\mathcal{S}(\Gamma, p, B^*)$. This implies that such a boundary point is also a boundary point of $\mathcal{S}(\Gamma, p)$.

Formally, the structure of $\mathcal{S}(\Gamma, p)$ is as follows.

Theorem 5.3 *For an open set of stochastic games $\Gamma \in \mathcal{G}(\eta)$ and priors $p \in \Sigma$ with full Lebesgue measure, $\mathcal{S}(\Gamma, p)$ is a compact 1-dimensional piecewise C^∞ manifold with boundary. The boundary of $\mathcal{S}(\Gamma, p)$ is given by the intersection of $\mathcal{S}(\Gamma, p)$ and $\{0, 1\} \times \Sigma$. There is a unique boundary point in $\{0\} \times \Sigma$ which corresponds to a strategy-combination in pure strategies.*

Proof The intuition of the proof of the first part of this theorem is given in the text above the theorem. A rigorous proof can be given using the Lemke-Howson argument. For a rigorous development of this argument, see for instance Herings and Peeters (1999).

The second part of the theorem, that there is generically a unique boundary point in $\{0\} \times \Sigma$ and that this boundary point is in pure strategies, remains to be shown.

Suppose there is a best response $\bar{\rho}^i$ to the prior in mixed strategies. Then for some state ω player i plays under $\bar{\rho}^i$ at least two pure strategies with positive probability mass, say $s^i_{\omega 1}$ and $s^i_{\omega 2}$. Since $\bar{\rho}^i$ is an optimal strategy it is part of a solution to the necessary and sufficient conditions above Theorem 3.5. So,

$$\lambda^i_{\omega 1} = \lambda^i_{\omega 2} = 0 \quad \text{and}$$
$$u^i(\omega, p^{-i}_\omega, s^i_{\omega 1}) + \delta \cdot \sum_{\bar{\omega} \in \Omega} \pi(\bar{\omega} \mid \omega, p^{-i}_\omega, s^i_{\omega 1}) \mu^i_{\bar{\omega}} = \mu^i_\omega =$$
$$= u^i(\omega, p^{-i}_\omega, s^i_{\omega 2}) + \delta \cdot \sum_{\bar{\omega} \in \Omega} \pi(\bar{\omega} \mid \omega, p^{-i}_\omega, s^i_{\omega 2}) \mu^i_{\bar{\omega}}.$$

---

[6]Let X be a topological space. It is called an arc if it is homeomorphic to the closed unit interval $[0, 1]$ and it has two boundary points in this case; it is called a loop if it is homeomorphic to the unit circle in $\mathbb{R}^2$ and it has no boundary points in that case.

It follows that every strategy ρ^i with $\rho^i_{\omega 1}, \rho^i_{\omega 2} \geq 0$ and $\rho^i_{-\{\omega 1, \omega 2\}} = \bar{\rho}^i_{-\{\omega 1, \omega 2\}}$ is a best response to the prior. This contradicts the local finiteness of the number of solutions at $t = 0$. Therefore, generically, all best responses are in pure strategies.

Suppose there are two pure best responses $\bar{\rho}^i$ and $\hat{\rho}^i$ against the prior. Then $\bar{\mu}^i_\omega = \hat{\mu}^i_\omega$ for all $\omega \in \Omega$. Take $\rho^i = \frac{1}{2}\bar{\rho}^i + \frac{1}{2}\hat{\rho}^i$, $\mu^i = \bar{\mu}^i$, and $\lambda^i_{\omega j} = 0$ if $\rho^i_{\omega j} > 0$ and $\lambda^i_{\omega j} = \bar{\lambda}^i_{\omega j}$ otherwise. Now indeed the complementarity conditions are satisfied and ρ^i is a member of Σ^i.

Suppose $\rho^i_{\omega j} > 0$. Then it holds that $\bar{\rho}^i_{\omega j} > 0$ or $\hat{\rho}^i_{\omega j} > 0$; suppose without loss of generality that $\bar{\rho}^i_{\omega j} > 0$. Then

$$
\begin{aligned}
0 &= u^i(\omega, p^{-i}_\omega, s^i_{\omega j}) + \delta \cdot \sum_{\bar{\omega} \in \Omega} \pi(\bar{\omega} \mid \omega, p^{-i}_\omega, s^i_{\omega j}) \bar{\mu}^i_{\bar{\omega}} - \bar{\mu}^i_\omega = \\
&= u^i(\omega, p^{-i}_\omega, s^i_{\omega j}) + \delta \cdot \sum_{\bar{\omega} \in \Omega} \pi(\bar{\omega} \mid \omega, p^{-i}_\omega, s^i_{\omega j}) \mu^i_{\bar{\omega}} + \lambda^i_{\omega j} - \mu^i_\omega.
\end{aligned}
$$

Suppose $\rho^i_{\omega j} = 0$. Then it holds that $\bar{\rho}^i_{\omega j} = 0$ and therefore $\lambda^i_{\omega j} = \bar{\lambda}^i_{\omega j}$. The equalities above hold again.

It follows that ρ^i is a best response to the prior. This contradicts that, generically, best responses are pure. $\qquad \square$

For almost every Γ and p, the set $\mathcal{S}(\Gamma, p)$ consists of a finite number of arcs and loops. Although it is not necessarily the case that these arcs and loops are smooth, the number of non-differentiabilities is finite at most. Theorem 5.2 implies that all arcs in $\mathcal{S}(\Gamma, p)$ start and end in $\{0, 1\} \times \Sigma$. Each such path consists of a finite sequence of smooth arcs of the sets $\mathcal{S}(\Gamma, p, B^*)$. A loop in $\mathcal{S}(\Gamma, p)$ consists either of a finite sequence (at least two) of differentiable arcs of the sets $\mathcal{S}(\Gamma, p, B^*)$ or is a loop of a set $\mathcal{S}(\Gamma, p, B^*)$. See Figure 2 for an impression of the structure of $\mathcal{S}(\Gamma, p)$.
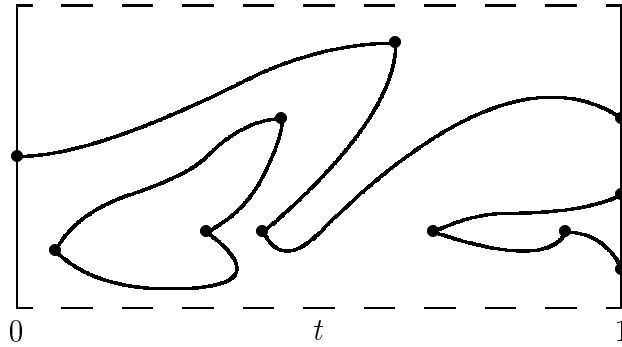

Figure 2: The structure of $\mathcal{S}(\Gamma, p)$.

Generically, each player i has a unique best response to the prior for all possible initial states, so generically there is exactly one point of $\mathcal{S}(\Gamma, p)$ that belongs to $\{0\} \times \Sigma$. This point

18

is both a corner point of $\{0\} \times \Sigma$ and a boundary point of $\mathcal{S}(\Gamma, p, B^*)$, where B^* consists of the following nz strategies: for each possible initial state and for each player the best response to the prior. Given some initial state ω, the determination of the best response against the prior in state ω, involves the determination of the behavior in all other states $\bar{\omega}$ as well. So the nz independent strategies are found by solving n optimization problems, as opposed to nz independent optimization problems. By Theorem 5.3 this point is the starting point of a uniquely defined arc of $\mathcal{S}(\Gamma, p)$. This arc is the unique feasible path of $\mathcal{S}(\Gamma, p)$ that transforms prior beliefs into (stationary) equilibrium beliefs.

Corollary 5.4 *For an open set of stochastic games $\Gamma \in \mathcal{G}(\eta)$ and priors $p \in \Sigma$ with full Lebesgue measure, the stochastic tracing procedure is well-defined.*

By following the feasible path starting in the unique point $\mathcal{S}(\Gamma, p) \cap (\{0\} \cap \Sigma)$ we find a stationary equilibrium of the stochastic game Γ. The set $\mathcal{S}(\Gamma, p) \cap (\{1\} \times \Sigma)$ consists of all stationary equilibria of the stochastic game Γ. Precisely one of these stationary equilibria is an element of the feasible path of $\mathcal{S}(\Gamma, p)$. Any other stationary equilibrium is a boundary point of $\mathcal{S}(\Gamma, p)$ and is therefore part of some arc of $\mathcal{S}(\Gamma, p)$. A moment of reflection makes clear that the remaining stationary equilibria are pairwise connected by arcs from $\mathcal{S}(\Gamma, p)$, and so the number of stationary equilibria is odd.

Corollary 5.5 *For an open set of stochastic games $\Gamma \in \mathcal{G}(\eta)$ with full Lebesgue measure, the number of stationary equilibria is odd.*

The generic oddness of the number of Nash equilibria in normal form games is a well-known result of Rosenmüller (1971), Wilson (1971), and Harsanyi (1973). The generic finiteness of the number of stationary equilibria in stochastic games is a recent result of Haller and Lagunoff (2000). The corollary shows that Haller and Lagunoff's result can be sharpened to oddness.

The observations made so far suggest the following algorithm for the computation of the stationary equilibrium selected by the stochastic tracing procedure in n-person stochastic games. Define the admissible set B^* that contains for each player i the best response to the prior for all possible initial states, and start with a point $(0, \rho)$ in $\mathcal{S}(\Gamma, p, B^*)$ such that ρ_ω^i is a best response of player i to the prior when ω is the initial state. Theorem 5.3 implies that B^* and ρ are uniquely determined. The equalities (1)-(4) belonging to this B^* determine at least a part of the feasible path. As long as the inequalities (5)-(8) hold with strict inequality we do not change our B^*. As soon as one of the inequalities from (5) or (6) gets binding, we change B^*. When the binding inequality belongs to (5), say $\rho_{\omega j}^i = 0$ while $s_{\omega j}^i \in B^*$, we delete $s_{\omega j}^i$ out of B^*. Obviously, this cannot happen for the

starting B^*. If the binding inequality belongs to (6), say $\lambda^i_{\omega j} = 0$ while $s^i_{\omega j} \notin B^*$, we have to add strategy $s^i_{\omega j}$ to B^*. In both situations there is a strategy $s^i_{\omega j}$ for which $\rho^i_{\omega j} = 0$ and $\lambda^i_{\omega j} = 0$. In general this leads to a kink in the feasible path of the stochastic tracing procedure. This procedure is repeated over and over again, until the inequality (8) is binding and a stationary equilibrium is found. Note that inequality (7) is only binding in the starting point.

6 Smoothing the Stochastic Tracing Procedure

The previous section presents a method that can be used for the computation of a stationary equilibrium. However, switching homotopies can be a serious problem in terms of computing time. There are $\prod_{(i,\omega) \in N \times \Omega} (2^{m^i_\omega} - 1)$ different sets B^*, whereas each one of them may be generated several times in the course of the algorithm.

We follow a suggestion in Garcia and Zangwill (1981), also used in Herings and Peeters (1999), and formulate one, everywhere differentiable, homotopy by using a well-chosen transformation of variables. Define, for $\alpha \in \mathbb{R}^{m^*}$,

$$\rho^i_{\omega j}(\alpha) = [\max\{0, \alpha^i_{\omega j}\}]^2 \quad \text{and} \quad \lambda^i_{\omega j}(\alpha) = [\max\{0, -\alpha^i_{\omega j}\}]^2.$$

After this transformation of variables, the system of equalities and inequalities becomes

$$v^i(t; \omega, \rho^{-i}_\omega(\alpha), s^i_{\omega j}) + \delta \cdot \sum_{\bar{\omega} \in \Omega} \tilde{\pi}(t; \bar{\omega} \mid \omega, \rho^{-i}_\omega(\alpha), s^i_{\omega j}) \mu^i_{\bar{\omega}} + \lambda^i_{\omega j}(\alpha) - \mu^i_\omega = 0$$
$$(s^i_{\omega j} \in S^i_\omega, \omega \in \Omega, i \in N),$$
$$\lambda^i_{\omega j}(\alpha) \geq 0, \quad \rho^i_{\omega j}(\alpha) \geq 0, \quad \lambda^i_{\omega j}(\alpha)\rho^i_{\omega j}(\alpha) = 0 \qquad (s^i_{\omega j} \in S^i_\omega, \omega \in \Omega, i \in N),$$
$$\sum_{s^i_{\omega j} \in S^i_\omega} \rho^i_{\omega j}(\alpha) - 1 = 0 \qquad (\omega \in \Omega, i \in N).$$

Owing to the proposed transformation of variables, the conditions $\lambda^i_{\omega j}(\alpha) \geq 0$, $\rho^i_{\omega j}(\alpha) \geq 0$ and $\lambda^i_{\omega j}(\alpha)\rho^i_{\omega j}(\alpha) = 0$ are trivially satisfied. We can reformulate the (in)equalities in (1)-(8) that characterize the set $\mathcal{O}(\Gamma, p, B^*)$ by considering solutions $(t, \alpha, \mu) \in \mathbb{R} \times \mathbb{R}^{m^*} \times \mathbb{R}^{nz}$, with $\alpha^i_{\omega j} \geq 0$ if $s^i_{\omega j} \in B^*$ and $\alpha^i_{\omega j} \leq 0$ if $s^i_{\omega j} \notin B^*$, to

(a) $\quad v^i(t; \omega, \rho^{-i}_\omega(\alpha), s^i_{\omega j}) + \delta \cdot \sum_{\bar{\omega} \in \Omega} \tilde{\pi}(t; \bar{\omega} \mid \omega, \rho^{-i}_\omega(\alpha), s^i_{\omega j}) \mu^i_{\bar{\omega}} + \lambda^i_{\omega j}(\alpha) - \mu^i_\omega = 0$
$$(s^i_{\omega j} \in S^i_\omega, \omega \in \Omega, i \in N),$$

(b) $\quad \sum_{s^i_{\omega j} \in S^i_\omega} \rho^i_{\omega j}(\alpha) - 1 = 0 \qquad (\omega \in \Omega, i \in N),$

(c) $\quad t \geq 0,$

(d) $\quad -t + 1 \geq 0.$

Notice that the same system of equalities and inequalities can be used, irrespective of the set B^*. The role of B^* is taken over by the sign-combinations of the components of the vector α.

Counting equations and unknowns in the system (a)-(d) shows that there is one degree of freedom, and therefore one expects a 1-dimensional solution set. Consider any solution (t, α, μ) to (a)-(d). When $\alpha_{\omega j}^i = 0$, then both $\rho_{\omega j}^i(\alpha)$ and $\lambda_{\omega j}^i(\alpha)$ are zero. This implies that we have exactly two admissible subsets of S^* for which the set of (in)equalities (1)-(8) are satisfied. If along a solution curve $\alpha_{\omega j}^i$ is increasing while passing zero, then $\rho_{\omega j}^i(\alpha)$ gets positive and $B_{\text{new}}^* = B_{\text{old}}^* \cup \{s_{\omega j}^i\}$. If $\alpha_{\omega j}^i$ is decreasing while passing zero, then $\lambda_{\omega j}^i(\alpha)$ gets negative and $B_{\text{new}}^* = B_{\text{old}}^* \setminus \{s_{\omega j}^i\}$. When $\alpha_{\omega j}^i$ passes zero, a kink appears in the method proposed in Section 5.

The left-hand sides of the equalities (a)-(b) specify the homotopy $\mathcal{H} : [0,1] \times \mathbb{R}^{m^*} \times \mathbb{R}^{nz} \to \mathbb{R}^{m^*} \times \mathbb{R}^{nz}$,

$$
\mathcal{H}(t, \alpha, \mu) =
\begin{pmatrix}
v^i(t; \omega, \rho_\omega^{-i}(\alpha), s_{\omega j}^i) + \delta \cdot \sum_{\bar{\omega} \in \Omega} \tilde{\pi}(t; \bar{\omega} \,|\, \omega, \rho_\omega^{-i}(\alpha), s_{\omega j}^i) \mu_{\bar{\omega}}^i + \lambda_{\omega j}^i(\alpha) - \mu_\omega^i = 0 \\
(s_{\omega j}^i \in S_\omega^i, \omega \in \Omega, i \in N) \\
\sum_{s_{\omega j}^i \in S_\omega^i} \rho_{\omega j}^i(\alpha) - 1 = 0 \qquad (\omega \in \Omega, i \in N)
\end{pmatrix}.
$$

The homotopy function $\mathcal{H}$ is continuously differentiable. The inequalities (c) and (d) are satisfied as the homotopy takes $[0,1]$ as the domain for the variable t. Further, $\mathcal{H}$ has the salient feature that its zeros describe the linear tracing procedure, $(t, \alpha, \mu) \in \mathcal{H}^{-1}(\{0\})$ if and only if $(t, \rho(\alpha)) \in \mathcal{S}(\Gamma, p)$.

The set $\mathcal{H}^{-1}(\{0\})$ consists of finitely many differentiable arcs and loops. All arcs start and end in $\{0, 1\} \times \mathbb{R}^{m^*} \times \mathbb{R}^{nz}$. Loops have no points in common with $\{0, 1\} \times \mathbb{R}^{m^*} \times \mathbb{R}^{nz}$. There is exactly one arc that starts in $\{0\} \times \mathbb{R}^{m^*} \times \mathbb{R}^{nz}$ and that ends in $\{1\} \times \mathbb{R}^{m^*} \times \mathbb{R}^{nz}$. All other arcs start and end in $\{1\} \times \mathbb{R}^{m^*} \times \mathbb{R}^{nz}$ and connect two points inducing stationary equilibria of the stochastic game Γ. Starting at the unique point $(0, \alpha^0, \mu^0) \in \mathcal{H}^{-1}(\{0\})$ at $t = 0$ and following the path described by the zeros of $\mathcal{H}$, we end up in a point $(1, \tilde{\alpha}, \tilde{\mu}) \in \mathcal{H}^{-1}(\{0\})$. This point generates the stationary equilibrium $\rho(\tilde{\alpha})$ of Γ selected by the stochastic tracing procedure. See Figure 3 for an impression of the structure of $\mathcal{H}^{-1}(\{0\})$.

The structure of $\mathcal{H}^{-1}(\{0\})$ is even simpler than the one of $\mathcal{S}(\Gamma, p)$. Not only are complications like bifurcations, spirals, higher dimensional solutions sets, diverging behavior, etc., excluded. The arcs and loops in $\mathcal{H}^{-1}(\{0\})$ are differentiable everywhere. It is the transformation of variables that smoothes out the kinks. As a direct consequence, it is possible to calculate the derivative at each point of the feasible path, which makes it possible to follow
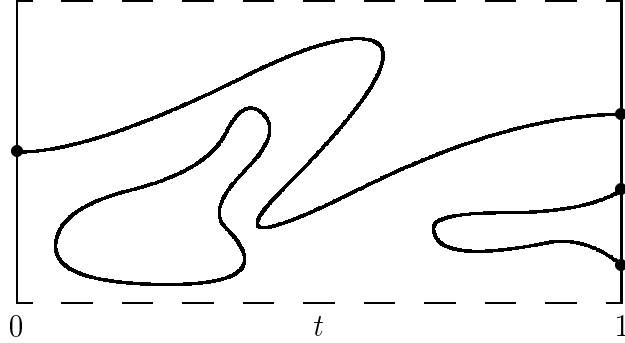
Figure 3: The structure of $\mathcal{H}^{-1}(\{0\})$.

the path by means of many easily implementable numerical methods, including methods to solve ordinary differential equations.

7 Implementation

The stochastic game Γ^0 naturally decomposes into n mutually independent and separate Markov decision problems, one for each player. We have shown that generically a Markov decision problem yields a unique optimal pure stationary strategy. The combination of all optimal strategies (for each player one) induces the starting point of our algorithm. This point can be determined analytically since there are finitely many pure stationary strategies in each decision problem. We only have to compute the total discounted payoffs for each pure stationary strategy of a player and observe which one generates the highest payoff.

Once we have determined the starting point, the numerical process starts by following the homotopy-path from that point on. We have programmed a number of FORTRAN-subroutines belonging to the software-package HOMPACK[7], a FORTRAN77 program (see Watson, Billups and Morgan (1987)). HOMPACK provides three qualitative different algorithms for tracking the zero curve of the homotopy: ordinary differential equation-based, normal flow, and augmented Jacobian matrix. In general, the first algorithm is the most robust of the three algorithms, but is also the slowest, sometimes by a wide margin. Being risk-averse we used this algorithm, called FIXPDF, for implementation. FIXPDF is an ordinary differential equation-based algorithm working with dense Jacobian matrices.

We parameterize the homotopy path by pathlength τ. Thus $t = t(\tau)$, $\alpha = \alpha(\tau)$ and $\mu = \mu(\tau)$ along the homotopy path, and $\mathcal{H}(t(\tau), \alpha(\tau), \mu(\tau)) = 0$ identically in τ. The

---

[7]http://www.netlib.org/hompack/index.html

differential equation is characterized by

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\mathcal{H}(t(\tau),\alpha(\tau),\mu(\tau)) = \partial\mathcal{H}(t(\tau),\alpha(\tau),\mu(\tau)) \cdot \begin{pmatrix} \mathrm{d}t/\mathrm{d}\tau \\ \mathrm{d}\alpha/\mathrm{d}\tau \\ \mathrm{d}\mu/\mathrm{d}\tau \end{pmatrix} = 0,$$

$$\left\| \left(\frac{\mathrm{d}t}{\mathrm{d}\tau}, \frac{\mathrm{d}\alpha}{\mathrm{d}\tau}, \frac{\mathrm{d}\mu}{\mathrm{d}\tau} \right) \right\|_2 = 1.$$

and initial conditions given by

$$(t(0),\alpha(0),\mu(0)) = \mathcal{H}^{-1}(\{0\}) \cap (\{0\} \times \mathbb{R}^{m^*} \times \mathbb{R}^{nz}),$$

so $t(0) = 0$ and $(\alpha(0),\mu(0))$ corresponds to the best response to the prior. The homotopy path corresponds to the trajectory of this initial value problem. When $t(\bar{\tau}) = 1$, $(\alpha(\bar{\tau}),\mu(\bar{\tau}))$ is a zeropoint of $\mathcal{H}$ and $\rho(\alpha(\bar{\tau}))$ is a stationary equilibrium of stochastic game Γ with equilibrium payoff $\mu(\bar{\tau})$. Complete details for solving the initial value problem are given in Watson (1979) and Watson and Fenner (1980).

We have implemented the homotopy function as described in Section 6 for stochastic games with 2 states, 3 players, with in each state 2 strategies for each player. It may be possible to improve on computing times by rescaling utilities and probabilities, or by using an α-transformation with a power different from 2 (but larger than 1, to keep differentiability) or multiplied by a positive scalar different from 1. An extensive digress on the optimal numerical implementation is beyond the aim of the current paper.

Within the class of stochastic games with 2 states, 3 players and for each player 2 strategies in each state, we have generated five randomly chosen stochastic games and five randomly chosen priors. All payoffs, transition probabilities, and priors are chosen independently from one another out of the uniform distribution on $[0, 1]$. The transition probabilities and the priors are of course normalized to sum up to 1. The discount factor is fixed at 0.95.

The maximal inaccuracy of our calculation amounts to 10^{-8}, which means that the 2-norm of the value of the homotopy function is less than 10^{-8} in the computed equilibrium.

In Table 1 the computing times to compute a stationary equilibrium for each game-prior pair are displayed. The mean time to compute a stationary equilibrium is 1.26 seconds with a standard deviation of 0.23 seconds. For the games we found on average 1.60 different equilibria with a standard deviation of 0.89 when 5 different priors where used. So, the algorithm is quite fast and multiple equilibria can be found.

	prior 1	prior 2	prior 3	prior 4	prior 5
game 1	1.31	1.09	1.21	1.38	1.32
game 2	0.82	1.37	1.27	1.43	1.10
game 3	1.59	1.21	1.43	1.31	1.42
game 4	1.27	1.15	1.27	1.65	0.93
game 5	1.54	0.77	0.88	1.26	1.48

Table 1: Computation times in seconds

A   Proofs

To make the proofs as transparent as possible, we need some notations and definitions from the theory of regular constraint sets as presented in Jongen, Jonker and Twilt (1983), see also Herings (1997) for a first application of this theory in economics.

For some $r \geq 1$ a subset M of $\mathbb{R}^m$ is called a k-dimensional C^r manifold with generalized boundary (MGB), if for every $\bar{x} \in M$ there exists a C^r diffeomorphism $\phi : U \to V$, where U is an open subset of $\mathbb{R}^m$ containing $\bar{x}$ and V is open in $\mathbb{R}^m$, and some integer $\ell(\bar{x}) \geq 0$, such that $\phi(\bar{x}) = 0$ and $\phi(U \cap M)$ equals $\{y \in V \mid y_i = 0, \ i = 1, \ldots, m - k, \ \text{and} \ y_i \geq 0, \ i = m - k + 1, \ldots, m - k + \ell(\bar{x})\}$. If for every element $\bar{x}$ of an MGB M it holds that $\ell(\bar{x}) \leq 1$, then M is called a manifold with boundary and the set of elements $\bar{x}$ for which $\ell(\bar{x}) = 1$ is an $(k-1)$-dimensional manifold, called the boundary of M.

One way to show that a set is an MGB is by showing that it is a regular constraint set. Let J^1 and J^2 be two finite index sets and let g_j for all $j \in J^1$ and h_j for all $j \in J^2$, be C^r functions defined on some open subset X of $\mathbb{R}^m$. We define

$$M[g, h] = \left\{ x \in X \ \middle| \ g_j(x) = 0, \ \forall j \in J^1, \ \text{and} \ h_j(x) \geq 0, \ \forall j \in J^2 \right\}.$$

For $x \in X$ we define the set $J^0(x) = \{j \in J^2 \mid h_j(x) = 0\}$. If for every $\bar{x} \in M[g, h]$ it holds that

$$\left\{ \partial_x g_j(\bar{x}) \ \middle| \ j \in J^1 \right\} \cup \left\{ \partial_x h_j(\bar{x}) \ \middle| \ j \in J^0(\bar{x}) \right\}$$

is a set of independent vectors, then $M[g, h]$ is called a C^r regular constraint set (RCS). In Jongen, Jonker and Twilt (1983) it is shown that every C^r RCS is an $(m - |J^1|)$-dimensional C^r MGB with $\ell(\bar{x}) = |J^0(\bar{x})|$ for every $\bar{x} \in M[g, h]$.

In this entire appendix, we fix a size vector η, a transition mapping π, a discount factor δ, and a prior $p \in \Sigma$. For any tuple of utility functions u and any admissible subset B^*, we define the functions $g^{B^*, u} : \mathbb{R}^{1 + 2m^* + nz} \to \mathbb{R}^{2m^* + nz}$ and $h^{B^*, u} : \mathbb{R}^{1 + 2m^* + nz} \to \mathbb{R}^{m^* + 2}$ in such

a way that $g^{B^*,u}$ equals the left-hand side of the equalities (1)-(4) and $h^{B^*,u}$ the left-hand side of the inequalities (5)-(8),

$$
g^{B^*,u}(t,\rho,\lambda,\mu) = \begin{cases}
v^i(t;\omega,\rho_\omega^{-i},s_{\omega j}^i) + \lambda_{\omega j}^i - \mu_\omega^i \\
\quad + \delta \cdot \sum_{\bar\omega\in\Omega}\tilde\pi(t;\bar\omega\,|\,\omega,\rho_\omega^{-i},s_{\omega j}^i)\mu_{\bar\omega}^i & (s_{\omega j}^i \in S_\omega^i, \omega\in\Omega, i\in N) \\
\rho_{\omega j}^i & (s_{\omega j}^i \notin B_\omega^i, \omega\in\Omega, i\in N) \\
\lambda_{\omega j}^i & (s_{\omega j}^i \in B_\omega^i, \omega\in\Omega, i\in N) \\
\sum_{s_{\omega j}^i\in S_\omega^i}\rho_{\omega j}^i - 1 & (\omega\in\Omega, i\in N)
\end{cases}
$$

and

$$
h^{B^*,u}(t,\rho,\lambda,\mu) = \begin{cases}
\rho_{\omega j}^i & (s_{\omega j}^i \in B_\omega^i, \omega\in\Omega, i\in N) \\
\lambda_{\omega j}^i & (s_{\omega j}^i \notin B_\omega^i, \omega\in\Omega, i\in N) \\
t \\
-t+1
\end{cases}.
$$

A point $(t,\rho,\lambda,\mu) \in \mathcal{O}(\Gamma,p,B^*)$ is a solution of (1)-(8) if and only if $g^{B^*,u}(t,\rho,\lambda,\mu) = 0$ and $h^{B^*,u}(t,\rho,\lambda,\mu) \geq 0$.

Let the functions $g^{B^*} : \mathbb{R}^{1+2m^*+nz}\times \mathbb{R}^{mn} \to \mathbb{R}^{2m^*+nz}$ and $h^{B^*} : \mathbb{R}^{1+2m^*+nz}\times \mathbb{R}^{mn} \to \mathbb{R}^{m^*+2}$ be defined such that $g^{B^*}(t,\rho,\lambda,\mu,u) = g^{B^*,u}(t,\rho,\lambda,\mu)$ and $h^{B^*}(t,\rho,\lambda,\mu,u) = h^{B^*,u}(t,\rho,\lambda,\mu)$ for all $u \in \mathbb{R}^{mn}$. Figure 4 presents the matrix of derivatives of the functions g^{B^*} and h^{B^*} with respect to all variables. The stars $(\star)$ in the matrix need not to be specified for our analysis.

Each row in Figure 4 corresponds to one of the equalities and inequalities in (1)-(8). To make the figure more clear, derivatives with respect to ρ are first taken for $s_{\omega j}^i \in B^*$. The same applies to the derivatives with respect to λ and the ordering of the equalities in (1). From the properties listed below the matrix, it follows immediately that the matrices $\partial_u v$ and E have full row rank.

The structure of the proof of Theorem 5.2 is as follows. First, it is shown that for almost all games u, $M[g^{B^*,u}, h^{B^*,u}]$ is a regular constraint set, from which the manifold structure of $\mathcal{L}(\Gamma,p,B^*)$ follows immediately. Next it is shown that there is an open set of full measure for which the manifold structure holds. For the first part of the proof, we show that for almost all games u, for every $\bar x \in M[g^{B^*,u}, h^{B^*,u}]$,

$$
\left\{ \partial_x g_j^{B^*,u}(\bar x) \,\Big|\, j \in J^1 \right\} \cup \left\{ \partial_x h_j^{B^*,u}(\bar x) \,\Big|\, j \in J^0(\bar x) \right\}
$$

is a set of independent vectors. To show this we need three lemmas, Lemmas A.1, A.2 and A.3. Lemma A.1 handles points $\bar x$ for which $J^0(\bar x) = \emptyset$, Lemma A.2 deals with points $\bar x$ such that the cardinality of $J^0(\bar x)$ is one, and Lemma A.3 implies that $J^0(\bar x)$ contains one element at most.

$$
\begin{array}{c|c|cc|cc|c|c}
 & \partial_t & \multicolumn{2}{c|}{\partial_\rho} & \multicolumn{2}{c|}{\partial_\lambda} & \partial_\mu & \partial_u \\
 & & B^* & S^*\backslash B^* & B^* & S^*\backslash B^* & & \\
\hline
(1)\ m^* & \star & \multicolumn{2}{c|}{\star} & \begin{matrix}I & 0\\ 0 & I\end{matrix} & & \star & \partial_u v(\cdot) \\
\hline
(2)\ m^* & 0 & 0 & I & \multicolumn{2}{c|}{0} & 0 & 0 \\
(3)\ & 0 & \multicolumn{2}{c|}{0} & I & 0 & 0 & 0 \\
\hline
(4)\ nz & 0 & \multicolumn{2}{c|}{E} & \multicolumn{2}{c|}{0} & 0 & 0 \\
\hline
(5)\ m^* & 0 & I & 0 & \multicolumn{2}{c|}{0} & 0 & 0 \\
(6)\ & 0 & \multicolumn{2}{c|}{0} & 0 & I & 0 & 0 \\
\hline
(7)\ 1 & 1 & \multicolumn{2}{c|}{0} & \multicolumn{2}{c|}{0} & 0 & 0 \\
(8)\ 1 & -1 & \multicolumn{2}{c|}{0} & \multicolumn{2}{c|}{0} & 0 & 0 \\
\hline
 & 1 & \multicolumn{2}{c|}{m^*} & \multicolumn{2}{c|}{m^*} & nz & mn
\end{array}
$$

(Row (1) right labels: B^* and $S^*\backslash B^*$.)

where

$$
E((i',\omega'),s^i_{\omega j}) = \begin{cases} 1 & \text{if } (i',\omega') = (i,\omega) \\ 0 & \text{otherwise} \end{cases}
$$

$$
\partial_{u^{i'}(\omega',s^{-i'}_{\omega'},s^{i'}_{\omega'j'})}v^i(t;\omega,\rho^{-i}_\omega,s^i_{\omega j}) = 0 \quad \text{if } s^i_{\omega j} \neq s^{i'}_{\omega'j'} \quad \text{for all } s^{-i'}_{\omega'} \in S^{-i'}_{\omega'}
$$

$$
\sum_{s^{-i}_\omega \in S^{-i}_\omega} \partial_{u^i(\omega,s^{-i}_\omega,s^i_{\omega j})}v^i(t;\omega,\rho^{-i}_\omega,s^i_{\omega j}) = 1.
$$

Figure 4: The matrix of partial derivatives of g^{B^*} and h^{B^*}.

Lemma A.1 *Let a prior $p \in \Sigma$ and an admissible subset B^* of S^* be given. Then, for almost all u, $g^{B^*,u} \pitchfork \{0\}$.*

Proof Consider a point $(\bar{t},\bar{\rho},\bar{\lambda},\bar{\mu},\bar{u})$ such that $g^{B^*}(\bar{t},\bar{\rho},\bar{\lambda},\bar{\mu},\bar{u}) = 0$. The matrix of partial derivatives of g^{B^*} at $(\bar{t},\bar{\rho},\bar{\lambda},\bar{\mu},\bar{u})$ is given by the rows (1)-(4) in Figure 4. We show first that this matrix has full row rank, from which it follows that $g^{B^*} \pitchfork \{0\}$.

Since $\partial_u v(\cdot)$ has full row rank and the derivative with respect to u in (2)-(4) is zero, it is sufficient to show that the part of the matrix given by (2)-(4) has full row rank. Since the derivative with respect to λ in (3) has full row rank, whereas the derivative with respect to λ in (2) and (4) is zero, it is sufficient to show that the rows in (2) and (4) are independent. The admissibility of B^* implies that the derivative with respect to ρ_{B^*} in (4) has full row rank. Since, the derivative with respect to ρ_{B^*} in (2) is zero, the only thing left to prove is that (2) has independent rows, which is obvious from the derivative with respect to $\rho_{S^*\backslash B^*}$. Consequently, $g^{B^*} \pitchfork \{0\}$.

By the transversality theorem (see, for example Mas-Colell (1985), theorem I.2.2) and since g^{B^*} is a C^∞ function, it follows that the complement of the set $\left\{ u \in \mathbb{R}^{mn} \mid g^{B^*,u} \pitchfork \{0\} \right\}$

has Lebesgue measure zero. □

Lemma A.2 *Let a prior $p \in \Sigma$ and an admissible subset B^* of S^* be given. Moreover, let an inequality $j' \in \{1, \ldots, m^* + 2\}$ be given. Then, for almost all u, $(g^{B^*,u}, h_{j'}^{B^*,u}) \pitchfork \{0\}$.*

Proof Consider a point $(\bar{t}, \bar{\rho}, \bar{\lambda}, \bar{\mu}, \bar{u})$ such that $g^{B^*}(\bar{t}, \bar{\rho}, \bar{\lambda}, \bar{\mu}, \bar{u}) = 0$ and $h_{j'}^{B^*}(\bar{t}, \bar{\rho}, \bar{\lambda}, \bar{\mu}, \bar{u}) = 0$. The matrix of partial derivatives of $(g^{B^*}, h_{j'}^{B^*})$ at $(\bar{t}, \bar{\rho}, \bar{\lambda}, \bar{\mu}, \bar{u})$ is given in Figure 4 by the rows (1)-(4) and a single row related to $h_{j'}^{B^*}$ in (5)-(8). We show first that this matrix has full row rank, from which it follows that $(g^{B^*}, h_{j'}^{B^*}) \pitchfork \{0\}$.

If row j' belongs to (6), (7) or (8), then it follows from the derivative with respect to $\lambda_{S^* \setminus B^*}$ or from the derivative with respect to t that row j' has rank 1. Since all other derivatives in row j' are zero, it follows as in the proof of Lemma A.1 that the rows of (1)-(4) together with row j' are independent. Consider the case where row j' belongs to (5). Following the first part of the proof of Lemma A.1, it suffices to prove that (2) and (4) together with row j' are independent. Inequality $h_{j'}^{B^*}$ states that $\rho_{\omega j'}^i \geq 0$, $s_{\omega j'}^i \in B^*$, and this inequality is now required to hold with equality. Since $\sum_{s_{\omega j}^i \in S_\omega^i} \rho_{\omega j}^i - 1 = 0$, it follows that $|B_\omega^i| = |B^* \cap S_\omega^i| > 1$. Then the row related to $s_{\omega j'}^i$ together with (2) and (4) trivially form an independent set. Consequently, $(g^{B^*}, h_{j'}^{B^*}) \pitchfork \{0\}$. By the transversality theorem it follows that the complement of $\{u \in \mathbb{R}^{mn} \mid (g^{B^*,u}, h_{j'}^{B^*,u}) \pitchfork \{0\}\}$ has Lebesgue measure zero. □

Lemma A.3 *Let a prior $p \in \Sigma$ and an admissible subset B^* of S^* be given. Moreover, let inequalities $j', j'' \in \{1, \ldots, m^* + 2\}$ with $j' \neq j''$ be given. Then, for almost all u, $(g^{B^*,u}, h_{j'}^{B^*,u}, h_{j''}^{B^*,u}) \pitchfork \{0\}$.*

Proof Consider a point $(\bar{t}, \bar{\rho}, \bar{\lambda}, \bar{\mu}, \bar{u})$ such that $g^{B^*}(\bar{t}, \bar{\rho}, \bar{\lambda}, \bar{\mu}, \bar{u}) = 0$, $h_{j'}^{B^*}(\bar{t}, \bar{\rho}, \bar{\lambda}, \bar{\mu}, \bar{u}) = 0$, and $h_{j''}^{B^*}(\bar{t}, \bar{\rho}, \bar{\lambda}, \bar{\mu}, \bar{u}) = 0$. The matrix of partial derivatives of $(g^{B^*}, h_{j'}^{B^*}, h_{j''}^{B^*})$ at $(\bar{t}, \bar{\rho}, \bar{\lambda}, \bar{\mu}, \bar{u})$ is given in Figure 4 by the rows (1)-(4) and two rows related to $h_{j'}^{B^*}$ and $h_{j''}^{B^*}$ in (5)-(8). We show first that this matrix has full row rank, from which it follows that $(g^{B^*}, h_{j'}^{B^*}, h_{j''}^{B^*}) \pitchfork \{0\}$.

The case where the two rows are not equal to (7) and (8) is similar to the proof of Lemma A.2. Rows (7) and (8) are not independent. However, they cannot be binding simultaneously, because then it holds that $t = 0$ and $t = 1$. Consequently, $(g^{B^*}, h_{j'}^{B^*}, h_{j''}^{B^*}) \pitchfork \{0\}$. It follows that the complement of the set $\{u \in \mathbb{R}^{mn} \mid (g^{B^*,u}, h_{j'}^{B^*,u}, h_{j''}^{B^*,u}) \pitchfork \{0\}\}$ has Lebesgue measure zero. □

References

Bergemann D., and J. Välimäki (1996), "Learning and Strategic Pricing", *Econometrica*, *64*, 1125-1149.

Breton M. (1991), "Algorithms for Stochastic Games", in T.E.S. Raghavan, T.S. Ferguson, T. Parthasarathy, and O.J. Vrieze (eds.) *Stochastic Games and Related Topics*, Series C: Game Theory, Mathematical Programming and Operations Research, 45-57.

Breton M., J.A. Filar, A. Haurie, and T.A. Schultz (1986), "On the computation of Equilibria in Discounted Games", in T. Basar (ed.) *Dynamic Games and Applications in Economics*, Lecture Notes on Economics and Mathematical Systems no.205, Springer-Verlag, Berlin.

Elzen A.H. van den, and Talman, A.J.J. (1999), "An Algorithmic Approach towards the Tracing Procedure for Bi-matrix Games", *Games and Economic Behavior, 28*, 130-145.

Filar J.A., T.A. Schultz, F. Thuijsman, and O.J. Vrieze (1991), "Nonlinear programming and stationary equilibria in stochastic games", *Mathematical Programming, 50*, 227-237.

Fink A.M. (1964), "Equilibrium in a stochastic n-person game", *Journal of Science of Hiroshima University Series A-I, 28*, 89-93.

Garcia C.B., and W.I. Zangwill (1981), *Pathways to Solutions, Fixed Points, and Equilibria*, Prentice Hall, Englewood Cliffs.

Haller H., and R. Lagunoff (2000), "Genericity and Markovian Behavior in Stochastic Games", *Econometrica, 68*, 1231-1248.

Harsanyi J.C. (1973), "Oddness of the Number of Equilibrium Points", *International Journal of Game Theory, 2*, 235-250.

Harsanyi J.C., and R. Selten (1988), *A General Theory of Equilibrium Selection in Games*, MIT Press, Cambridge, Massachusetts.

Herings P.J.J. (1997), "A Globally and Universally Stable Price Adjustment Process", *Journal of Mathematical Economics, 27*, 163-193.

Herings P.J.J. (2000), "Two Simple Proofs of the Feasibility of the Linear Tracing Procedure", *Economic Theory, 15*, 485-490.

Herings P.J.J., and A.H. van den Elzen (1998), "Computation of the Nash Equilibrium Selected by the Tracing Procedure in N-Person Games", CentER Discussion Paper, No. 9804, Tilburg University. *Accepted for Games and Economic Behavior subject to minor suggestion for modification.*

Herings P.J.J., and R.J.A.P. Peeters (1999), "A Differentiable Homotopy to Compute Nash

Equilibria of n-Person Games", Meteor Research Memorandum 99/33, University of Maastricht, Maastricht. *Forthcoming Economic Theory.*

Jongen H.T., P. Jonker, and F. Twilt (1983), *Nonlinear Optimization in $\mathbb{R}^n$, I. Morse Theory, Chebyshev Approximation*, Methoden und Verfahren der mathematische Physik, 29, Peter Lang, Frankfurt.

Judd K.L. (1997), "Computational Economics and Economic Theory: Substitutes and Complements?", *Journal of Economic Dynamics and Control, 21*, 907-942.

Lemke C.E., and J.T. Howson Jr. (1964), "Equilibrium Points of Bimatrix Games", *SIAM Journal on Applied Mathematics, 12*, 413-423.

Mas-Colell A. (1985), *The Theory of General Economic Equilibrium: A Differentiable Approach*, Cambridge University Press, Cambridge.

Maskin E., and J. Tirole (1997), "Markov Perfect Equilibrium, I: Observable Actions", mimeo.

McKelvey R., and A. McLennan (1996), "Computation of Equilibria in Finite Games", in H.M. Amman, D.A. Kendrick, and J. Rust (eds.), *Handbook of Computational Economics, Volume I*, Elsevier Science.

McLennan A. (1999), "The Expected Number of Nash Equilibria of a Normal Form Game", mimeo.

Olley G.S., and A. Pakes (1996) "The Dynamics of Productivity in the Telecommunications Equipment Industry", *Econometrica, 64*, 1263-1298.

Pakes A., and R. Ericson (1998), "Empirical Implications of Alternative Models of Firm Dynamics", *Journal of Economic Theory, 79*, 1-45.

Pakes A., and P. McGuire (1996), "Stochastic Algorithms, Symmetric Markov Perfect Equilibrium, and the 'Curse' of Dimensionality", mimeo. *Forthcoming Econometrica.*

Parthasarathy, T., and T.E.S. Raghavan (1981), "An Orderfield Property for Stochastic Games when One Player Controls Transition Probabilities", *Journal of Optimization Theory and Applications, 33*, 375-392.

Rosenmüller J. (1971), "On a Generalization of the Lemke-Howson Algorithm to Noncooperative N-Person Games", *SIAM Journal on Applied Mathematics, 21*, 73-79.

Schultz T.A. (1986), *Mathematical Programming and Stochastic Games*, Ph.D. Thesis, The John Hopkins University, Baltimore, Maryland.

Shapley L.S. (1953), "Stochastic Games", *Proceedings of the National Academy of Sciences of the United States of America, 39*, 1095-1100.

Sobel M.J. (1971), "Non-cooperative Stochastic Games", *The Annals of Mathematical Statistics, 42*, 1930-1935.

Stengel B. von, A. van den Elzen, and A.J.J. Talman (1996), "Tracing Equilibria in Extensive Games by Complementary Pivoting", CentER Discussion Paper, No. 9686, Tilburg University.

Takahashi M. (1964), "Equilibrium points of stochastic, noncooperative n-person games", *Journal of Science of Hiroshima University Series A-I, 28*, 95-99.

Watson L.T. (1979), "A Globally Convergent Algorithm for Computing Fixed Points of C^2 Maps", *Applied Mathematic and Computation, 5*, 297-311.

Watson L.T., S.C. Billups, and A.P. Morgan (1987), "HOMPACK: A Suite of Codes for Globally Convergent Homotopy Algorithms", *ACM Transactions on Mathematical Software, 13*, 281-310.

Watson L.T., and D. Fenner (1980), "Chow-Yorke Algorithm for Fixed Points or Zeros of C^2 Maps", *ACM Transactions on Mathematical Software, 6*, 252-260.

Wilson, R. (1971), "Computing Equilibria of N-Person Games", *SIAM Journal on Applied Mathematics, 21*, 80-87.