

## On the Existence of Stationary Nash Equilibria in Average Stochastic Games with Finite State and Action Spaces

Dmitrii Lozovanu<sup>1</sup> and Stefan Pickl<sup>2</sup>

<sup>1</sup> *Institute of Mathematics and Computer Science of Moldova Academy of Sciences,  
Academiei 5, Chisinau, MD-2028, Moldova,*

*E-mail: lozovanu@math.md*

<sup>2</sup> *Institute for Theoretical Computer Science, Mathematics and Operations Research,  
Universität der Bundeswehr München, 85577 Neubiberg-München, Germany,*

*E-mail: stefan.pickl@unibw.de*

**Abstract** We consider infinite  $n$ -person stochastic games with limiting average payoffs criteria for the players. The main results of the paper are concerned with the existence of stationary Nash equilibria and determining the optimal strategies of the players in the games with finite state and action spaces. We present conditions for the existence of stationary Nash equilibria in the considered games and propose an approach for determining the optimal stationary strategies of the players if such strategies exist.

**Keywords:** Markov decision processes, Average stochastic games, Stationary Nash equilibria, Optimal stationary strategies

### 1. Introduction

In this paper we study the problem of the existence and determining stationary Nash equilibria in average stochastic games with finite state and action spaces. Stochastic games, named sometimes Markov games, were introduced by Shapley, 1953. He considered two-person zero-sum stochastic games for which he proved the existence of the value and the optimal stationary strategies of the players with respect to a discounted payoff criterion. Later, this class of games has been extended to general  $n$ -person stochastic games with discounted and average payoffs criteria (Gillette, 1957; Fink, 1964; Takahashi, 1964; Vrieze, 1987; Filar et al., 1991; Neyman and Sorin, 2003). The most important results for  $n$ -person stochastic games with discounted payoffs have been obtained by Fink, 1964, Takahashi, 1964 and Sobel, 1971 who proved the existence of stationary Nash equilibria in such games. Schultz, 1986, Filar et al., 1991 showed that the problem of determining stationary Nash equilibria in a general  $n$ -person stochastic game with discounted payoffs can be represented as a nonlinear programming problem with linear constraints and the global minimum of the objective function equal to zero. Mertens and Neyman, 1981 studied two-person zero-sum games and proved the existence of uniform  $\varepsilon$ -optimal strategies for the players, i.e. they showed that for every  $\varepsilon > 0$  each of the two players has a strategy that guarantees the discounted value up to  $\varepsilon$  for every discount factor sufficiently close to 0. These results have been extended by Vieille, 2009 to an arbitrary noncooperative stochastic game of two players and afterwards they have been used for studying the problem of the existence of Nash equilibria in non-stationary strategies for two-player average stochastic games (Vieille, 2002, 2009, Solan, 2009; Solan and Vieille, 2010). Algorithmic approaches concerned with determining the optimal strategies of the players in some classes of stochastic games can be found in (Schultz, 1986; Filar et al., 1991; Neyman and Sorin, 2003; Solan, 2009).

The  $n$ -person stochastic games with limiting average payoffs have been studied by many authors (Neyman and Sorin, 2003; Rogers, 1969; Sobel, 1971; Solan, 2009; Solan and Vieille, 2010; Vieille, 2002; Vieille, 2009; Vrieze, 1987) however the existence of stationary Nash equilibria has been proved only for some classes of such games. Rogers, 1969 and Sobel, 1971 showed that stationary Nash equilibria exist for nonzero-sum stochastic games with average payoffs when the transition probability matrices induced by any stationary strategies of the players are unichain. An important class of average stochastic games for which stationary Nash equilibria exist represents stochastic positional games (Lozovanu, 2018, 2019). Furthermore Lozovanu, 2018, 2019 shown that for average stochastic positional games with unichain property and for two-player zero-sum stochastic positional games there exist stationary Nash equilibria in pure strategy. The main results concerned with the existence and determining Nash equilibria in two-player average stochastic games can be found in (Mertens and Neyman, 1981; Neyman and Sorin, 2003; Vieille, 2002, 2009, Vieille, 2009; Solan, 2009; Solan and Vieille, 2010). In the general case for an average stochastic game with given starting state a stationary Nash equilibrium may not exist. This fact has been shown by Flesch et al., 1997 that constructed an example of a 3-player average stochastic game with fixed starting state for which a stationary Nash equilibrium does not exist. Moreover, they shown that for an  $m$ -player ( $m \geq 3$ ) average stochastic game may not exist also stationary  $\varepsilon$ -equilibrium ( $\varepsilon > 0$ ). In general, for an average stochastic game there may exist a nonempty subset of states such that if the game starts in one of them then a stationary Nash equilibrium exists (Tijs and Vrieze, 1986). However, the problem of determining the initial states in an average stochastic game for which stationary equilibria exist is an open problem.

In this contribution we consider average stochastic games with finite state and action spaces. We show that an arbitrary average stochastic game in stationary strategies can be represented as a game in normal form where each payoff is quasi-monotonic (quasi-concave and quasi-convex) with respect to the strategy of the corresponding player. Furthermore we show that if the game in normal form has a pure Nash equilibrium then such an equilibrium corresponds to a stationary Nash equilibrium of the average stochastic game and vice versa. Based on this result and results of Debreu, 1952, Glicksberg, 1952, Dasgupta and Maskin, 1986, Reny, 1999 related to existence of Nash equilibria in the games with quasi-concave (quasi-convex) payoffs we formulate conditions for the existence and determining stationary Nash equilibria in average stochastic games.

## 2. Average Stochastic Games in Pure and Mixed Stationary Strategies

We first present the framework of a  $n$ -person stochastic game and then specify the formulation of stochastic games with average payoffs when the players use pure and mixed stationary strategies.

### 2.1. The Framework of a $n$ -person Average Stochastic Game

A stochastic game with  $n$  players consists of the following elements:

- a state space  $X$  (which we assume to be finite);
- a finite set  $A^i(x)$  of actions with respect to each player  $i \in \{1, 2, \dots, n\}$  for an arbitrary state  $x \in X$ ;

- a payoff  $f^i(x, a)$  with respect to each player  $i \in \{1, 2, \dots, n\}$  for each state  $x \in X$  and for an arbitrary action vector  $a \in \prod_i A^i(x)$ ;
- a transition probability function  $p : X \times \prod_{x \in X} \prod_{i=1}^n A^i(x) \times X \rightarrow [0, 1]$  that gives the probability transitions  $p_{x,y}^a$  from an arbitrary  $x \in X$  to an arbitrary  $y \in Y$  for a fixed action vector  $a \in \prod_i A^i(x)$ , where  $\sum_{y \in X} p_{x,y}^a = 1, \forall x \in X, a \in \prod_i A^i(x)$ ;
- a starting state  $x_0 \in X$ .

The game starts in the state  $x_0$  and the play proceeds in a *sequence of stages*. At *stage*  $t$  the players observe state  $x_t$  and simultaneously and independently choose actions  $a_t^i \in A^i(x_t)$ ,  $i = 1, 2, \dots, n$ . Then nature selects a state  $y = x_{t+1}$  according to probability transitions  $p_{x_t,y}^{a_t}$  for the given action vector  $a_t = (a_t^1, a_t^2, \dots, a_t^n)$ . Such a play of the game produces a sequence of states and actions  $x_0, a_0, x_1, a_1, \dots, x_t, a_t, \dots$  that defines a stream of stage payoffs  $f_t^1 = f^1(x_t, a_t)$ ,  $f_t^2 = f^2(x_t, a_t), \dots, f_t^n = f^n(x_t, a_t)$ ,  $t = 0, 1, 2, \dots$ . The *infinite average stochastic game* is the game with payoffs of the players

$$\omega_{x_0}^i = \liminf_{t \rightarrow \infty} \mathbb{E} \left( \frac{1}{t} \sum_{\tau=0}^{t-1} f_\tau^i \right), \quad i = 1, 2, \dots, n,$$

where  $\mathbb{E}$  is the expectation operator with respect to the probability measure in a Markov process with rewards induced by the initial state  $x_0$  and the corresponding vector actions  $a_t$  in the states  $x_t \in X$  (see Kallenberg, 2016; Puterman, 2005). Here  $\omega_{x_0}^i$  expresses the *average payoff per transition* of player  $i$  in the infinite game. Each player in this game has the aim to maximize his average payoff per transitions. In the case  $n = 1$  this game becomes the average Markov decision problem with a probability transition function  $p : X \times \prod_{x \in X} A(x) \times X \rightarrow [0, 1]$  and step rewards  $f(x, a) = f^1(x, a)$  in the states  $x \in X$  for given actions  $a \in A(x) = A^1(x)$ .

In the paper we will study the stochastic games when players use pure and mixed stationary strategies of selection the actions in the states.

## 2.2. Pure and Mixed Stationary Strategies of the Players

A *strategy (policy) of player*  $i \in \{1, 2, \dots, n\}$  in a stochastic game is a mapping  $s^i$  that provides for every state  $x_t \in X$  a probability distribution over the set of actions  $A^i(x_t)$ . If these probabilities take only values 0 and 1, then  $s^i$  is called a *pure strategy*, otherwise  $s^i$  is called a *mixed strategy*. If these probabilities depend only on the state  $x_t = x \in X$  (i. e.  $s^i$  do not depend on  $t$ ), then  $s^i$  is called a *stationary strategy*, otherwise  $s^i$  is called a non-stationary strategy.

Thus, a *pure stationary strategy* of player  $i \in \{1, 2, \dots, n\}$  can be regarded as a map  $s^i : x \rightarrow a^i \in A^i(x)$  for  $x \in X$  that determines for each state  $x$  an action  $a^i \in A^i(x)$ , i.e.  $s^i(x) = a^i$ . Obviously, the corresponding sets of pure stationary strategies  $S^1, S^2, \dots, S^n$  of the players in the game with finite state and action spaces are finite sets.

In the following we will identify a pure stationary strategy  $s^i(x)$  of player  $i$  with the set of boolean variables  $s_{x,a^i}^i \in \{0, 1\}$ , where for a given  $x \in X$   $s_{x,a^i}^i = 1$  if

and only if player  $i$  fixes the action  $a^i \in A^i(x)$ . So, we can represent the set of pure stationary strategies  $S^i$  of player  $i$  as the set of solutions of the following system:

$$\begin{cases} \sum_{a^i \in A^i(x)} s_{x,a^i}^i = 1, & \forall x \in X; \\ s_{x,a^i}^i \in \{0, 1\}, & \forall x \in X, \forall a^i \in A^i(x). \end{cases}$$

If in this system we change the restriction  $s_{x,a^i}^i \in \{0, 1\}$  for  $x \in X, a^i \in A^i(x)$  by the condition  $0 \leq s_{x,a^i}^i \leq 1$  then we obtain the set of stationary strategies in the sense of Shapley, 1953, where  $s_{x,a^i}^i$  is treated as the probability of the choices of the action  $a^i$  by player  $i$  every time when the state  $x$  is reached by any route in the dynamic stochastic game. Thus, we can identify the set of mixed stationary strategies of the players with the set of solutions of the system

$$\begin{cases} \sum_{a^i \in A^i(x)} s_{x,a^i}^i = 1, & \forall x \in X; \\ s_{x,a^i}^i \geq 0, & \forall x \in X, \forall a^i \in A^i(x) \end{cases} \quad (1)$$

and for a given profile  $s = (s^1, s^2, \dots, s^n)$  of mixed strategies  $s^1, s^2, \dots, s^n$  of the players the probability transition  $p_{x,y}^s$  from a state  $x$  to a state  $y$  can be calculated as follows

$$p_{x,y}^s = \sum_{(a^1, a^2, \dots, a^n) \in A(x)} \prod_{k=1}^n s_{x,a^k}^k p_{x,y}^{(a^1, a^2, \dots, a^n)}. \quad (2)$$

In the sequel we will distinguish stochastic games in pure and mixed stationary strategies.

### 2.3. Average Stochastic Games in Pure Stationary Strategies

Let  $s = (s^1, s^2, \dots, s^n)$  be a profile of pure stationary strategies of the players and denote by  $a(s) = (a^1(s), a^2(s), \dots, a^n(s)) \in \prod_{x \in X} \prod_{i=1}^n A^i(x)$  the action vector

that corresponds to  $s$  and determines the probability distributions  $p_{x,y}^s = p_{x,y}^{a(s)}$  in the states  $x \in X$ . Then the average payoffs per transition  $\omega_{x_0}^1(s), \omega_{x_0}^2(s), \dots, \omega_{x_0}^n(s)$  for the players are determined as follows

$$\omega_{x_0}^i(s) = \sum_{y \in X} q_{x_0,y}^s f^i(y, a(s)), \quad i = 1, 2, \dots, n,$$

where  $q_{x_0,y}^s$  represent the limiting probabilities in the states  $y \in X$  for the Markov process with a probability transition matrix  $P^s = (p_{x,y}^s)$  when the transitions start in  $x_0$ . So, if for the Markov process with probability matrix  $P^s$  the corresponding limiting probability matrix  $Q^s = (q_{x,y}^s)$  is known then  $\omega_x^1, \omega_x^2, \dots, \omega_x^n$  can be determined for an arbitrary starting state  $x \in X$  of the game. The functions  $\omega_{x_0}^1(s), \omega_{x_0}^2(s), \dots, \omega_{x_0}^n(s)$  on  $S = S^1 \times S^2 \times \dots \times S^n$  define a game in normal form that we denote by  $\langle \{S^i\}_{i=1,n}, \{\omega_{x_0}^i(s)\}_{i=1,n} \rangle$ . This game corresponds to an *average stochastic game in pure stationary strategies* that in extended form is determined by the tuple  $(X, \{A^i(x)\}_{i=1,n}, \{f^i(x, a)\}_{i=1,n}, p, x_0)$ .

If an arbitrary profile  $s = (s^1, s^2, \dots, s^n)$  of pure stationary strategies in a stochastic game induces a probability matrix  $P^s$  that corresponds to a Markov unichain then we say that the game possesses the unichain property and shortly we call it *unichain stochastic game*; otherwise we call it *multichain stochastic game*.

#### 2.4. Average Stochastic Games in Mixed Stationary Strategies

Let  $s = (s^1, s^2, \dots, s^n)$  be a profile of mixed stationary strategies of the players. Then elements of the probability transition matrix  $P^s = (p_{x,y}^s)$  in the Markov process induced by  $s$  can be calculated according to (2). Therefore if  $Q^s = (q_{x,y}^s)$  is the limiting probability matrix of  $P^s$  then the average payoffs per transition  $\omega_{x_0}^1(s), \omega_{x_0}^2(s), \dots, \omega_{x_0}^n(s)$  for the players are determined as follows

$$\omega_{x_0}^i(s) = \sum_{y \in X} q_{x_0,y}^s f^i(y, s), \quad i = 1, 2, \dots, n, \quad (3)$$

where

$$f^i(y, s) = \sum_{(a^1, a^2, \dots, a^n) \in A(y)} \prod_{k=1}^n s_{y, a^k}^k f^i(y, a^1, a^2, \dots, a^n) \quad (4)$$

expresses the average payoff (immediate reward) in the state  $y \in X$  of player  $i$  when the corresponding stationary strategies  $s^1, s^2, \dots, s^n$  have been applied by players  $1, 2, \dots, n$  in  $y$ .

Let  $\bar{S}^1, \bar{S}^2, \dots, \bar{S}^n$  be the corresponding sets of mixed stationary strategies for the players  $1, 2, \dots, n$ , i.e. each  $\bar{S}^i$  for  $i \in \{1, 2, \dots, n\}$  represents the set of solutions of system (2). The functions  $\omega_{x_0}^1(s), \omega_{x_0}^2(s), \dots, \omega_{x_0}^n(s)$  on  $\bar{S} = \bar{S}^1 \times \bar{S}^2 \times \dots \times \bar{S}^n$ , defined according to (3), (4), determine a game in normal form that we denote by  $\langle \{\bar{S}^i\}_{i=1, \dots, n}, \{\omega_{x_0}^i(s)\}_{i=1, \dots, n} \rangle$ . This game corresponds to an *average stochastic game in mixed stationary strategies* that in extended form is determined by the tuple  $(X, \{A^i(x)\}_{i=1, \dots, n}, \{f^i(x, a)\}_{i=1, \dots, n}, p, x_0)$ .

#### 2.5. Average Stochastic Games with Random Starting State

In the paper we will consider also average stochastic games in which the starting state is chosen randomly according to a given distribution  $\{\theta_x\}$  on  $X$ . So, for a given stochastic game we will assume that the play starts in the states  $x \in X$  with probabilities  $\theta_x > 0$  where  $\sum_{x \in X} \theta_x = 1$ . If the players use mixed stationary strategies of selection the actions in the states then the payoff functions

$$\psi_\theta^i(s^1, s^2, \dots, s^n) = \sum_{x \in X} \theta_x \omega_x^i(s^1, s^2, \dots, s^n), \quad i = 1, 2, \dots, n$$

on  $\bar{S} = \bar{S}^1 \times \bar{S}^2 \times \dots \times \bar{S}^n$  define a game in normal form  $\langle \{\bar{S}^i\}_{i=1, \dots, n}, \{\psi_\theta^i(s)\}_{i=1, \dots, n} \rangle$  that in extended form is determined by the following tuple  $(X, \{A^i(x)\}_{i=1, \dots, n}, \{f^i(x, a)\}_{i=1, \dots, n}, p, \{\theta_x\})$ . In the case  $\theta_x = 0, \forall x \in X \setminus \{x_0\}$ ,  $\theta_{x_0} = 1$  the considered game becomes a stochastic game with fixed starting state  $x_0$ . In analogues way we can specify the game in normal form  $\langle \{\bar{S}^i\}_{i=1, \dots, n}, \{\psi_\theta^i(s)\}_{i=1, \dots, n} \rangle$  for the average stochastic game with random starting state  $x_0$  when players use pure stationary strategies of selection the actions in the states.

#### 2.6. Definition of Stationary Nash Equilibria

Let  $s = (s^1, s^2, \dots, s^n) \in \bar{S}^*$ . Define  $s^{-i} = (s^1, s^2, \dots, s^{i-1}, s^{i+1}, \dots, s^n)$  as the vector of stationary strategies of all players other than  $i$  and denote  $s = (s^i, s^{-i})$ ,  $i = 1, 2, \dots, n$ . The profile  $s^* = (s^{1*}, s^{2*}, \dots, s^{n*})$  is called *stationary*

Nash equilibrium for an average stochastic game  $\langle \{\overline{S}^i\}_{i=1,n}, \{\omega_{x_0}^i(s)\}_{i=1,n} \rangle$  with given starting state  $x_0$  if

$$\omega_{x_0}^i(s^{i*}, s^{-i*}) \geq \omega_{x_0}^i(s^i, s^{-i*}), \quad \forall s^i \in \overline{S}^i, \quad i = 1, 2, \dots, n. \quad (5)$$

The profile  $s^* = (s^{1*}, s^{2*}, \dots, s^{n*})$  is called *stationary Nash equilibrium* for an average stochastic game  $\langle \{\overline{S}^i\}_{i=1,n}, \{\psi_\theta^i(s)\}_{i=1,n} \rangle$  when the starting state is chosen randomly according to a given distribution  $\{\theta_x\}$  on  $X$  if

$$\psi_\theta^i(s^{i*}, s^{-i*}) \geq \psi_\theta^i(s^i, s^{-i*}), \quad \forall s^i \in \overline{S}^i, \quad i = 1, 2, \dots, n. \quad (6)$$

### 3. An Approach for Determining Stationary Nash Equilibria in Average Stochastic Games with Unichain Property

In this section we show that an unichain average stochastic game in stationary strategies can be represented as a continuous game in normal form where the payoffs are quasi-monotonic with respect to the corresponding strategies of the players. Using such a model we propose an approach for determining stationary Nash equilibria for unichain average stochastic games.

#### 3.1. A Continuous Model for the Average Markov Decision Problem with Unichain Property

In (Lozovanu, 2011) has been shown that an average Markov decision problem with unichain property can be formulated as the following optimization problem:  
*Maximize*

$$\psi(s, q) = \sum_{x \in X} \sum_{a \in A(x)} f(x, a) s_{x,a} q_x, \quad (7)$$

*subject to*

$$\left\{ \begin{array}{l} q_y - \sum_{x \in X} \sum_{a \in A(x)} p_{x,y}^a s_{x,a} q_x = 0, \quad \forall y \in X; \\ \sum_{x \in X} q_x = 1; \\ \sum_{a \in A(x)} s_{x,a} = 1, \quad \forall x \in X; \\ s_{x,a} \geq 0, \quad \forall x \in X, a \in A(x). \end{array} \right. \quad (8)$$

Here  $f(x, a)$  represents the step reward in the state  $x \in X$  for a given action  $a \in A(x)$  in the unichain problem and  $p_{x,y}^a$  expresses the probability transition from  $x \in X$  to  $y \in X$  for  $a \in A(x)$ . The variables  $s_{x,a}$  correspond to strategies of selection of the actions  $a \in A(x)$  in the states  $x \in X$  and  $q_x$  for  $x \in X$  represent the corresponding limiting probabilities in the states  $x \in X$  for the probability transition matrix  $P^s = (p_{x,y}^s)$  induced by the stationary strategy  $s$ .

In this problem the average reward  $\psi(s, q)$  is maximized under the conditions (8) that determines the set of feasible stationary strategies in the unichain problem. An optimal solution  $(s^*, q^*)$  of problem (7), (8) with  $s_{x,a}^* \in \{0, 1\}$  corresponds to an optimal stationary strategy  $s^* : X \rightarrow A$  where  $a^* = s^*(x)$  for  $x \in X$  if  $s_{x,a}^* = 1$ . Using the notations  $\alpha_{x,a} = s_{x,a} q_x$ , for  $x \in X, a \in A(x)$ , problem (7), (8) can be

easily transformed into the following linear programming problem:

*Maximize*

$$\bar{\psi}(\alpha) = \sum_{x \in X} \sum_{a \in A(x)} f(x, a) \alpha_{x,a} \quad (9)$$

*subject to*

$$\left\{ \begin{array}{l} q_y - \sum_{x \in X} \sum_{a \in A(x)} p_{x,y}^a \alpha_{x,a} = 0, \quad \forall y \in X; \\ \sum_{x \in X} q_x = 1; \\ \sum_{a \in A(x)} \alpha_{x,a} - q_x = 0, \quad \forall x \in X; \\ \alpha_{x,a} \geq 0, \quad \forall x \in X, a \in A(x). \end{array} \right. \quad (10)$$

This problem can be simplified by eliminating  $q_x$  from (10) and finally we obtain the problem in which it is necessary to maximize the objective function (9) on the set of solutions of the following system:

$$\left\{ \begin{array}{l} \sum_{a \in A(y)} \alpha_{y,a} - \sum_{x \in X} \sum_{a \in A(x)} p_{x,y}^a \alpha_{x,a} = 0, \quad \forall y \in X; \\ \sum_{x \in X} \sum_{a \in A(x)} \alpha_{x,a} = 1; \\ \alpha_{x,a} \geq 0, \quad \forall x \in X, a \in A(x). \end{array} \right. \quad (11)$$

Based on the relationship mentioned above between problem (7), (8) and problem (9), (11) in (Lozovanu, 2011) the following result has been announced.

**Lemma 1.** *Let an average Markov decision problem be given, where an arbitrary stationary strategy  $s$  generates a Markov unichain, and consider the function*

$$\psi(s) = \sum_{x \in X} \sum_{a \in A(x)} f(x, a) s_{x,a} q_x$$

where  $q_x$  for  $x \in X$  satisfy the condition

$$\left\{ \begin{array}{l} q_y - \sum_{x \in X} \sum_{a \in A(x)} p_{x,y}^a s_{x,a} q_x = 0, \quad \forall y \in X; \\ \sum_{x \in X} q_x = 1. \end{array} \right.$$

Then the function  $\psi(s)$  on the set  $\bar{S}$  of solutions of the system

$$\left\{ \begin{array}{l} \sum_{a \in A(x)} s_{x,a} = 1, \quad \forall x \in X; \\ s_{x,a} \geq 0, \quad \forall x \in X, a \in A(x) \end{array} \right.$$

depends only on  $s_{x,a}$  for  $x \in X, a \in A(x)$ , and  $\psi(s)$  is quasi-monotonic on  $\bar{S}$  (i.e.  $\psi(s)$  is quasi-concave and quasi-convex on  $\bar{S}$  Boyd and Vandenberghe, 2004). Moreover,  $\psi(s) = \omega_x(s), \forall x \in X$ .

The full proof of this lemma in a more general form is presented in (Lozovanu, 2018).

### 3.2. Stationary Equilibria for Average Stochastic Games with Unichain Property

An average stochastic game with unichain property can be formulated in terms of stationary strategies as follows.

Let  $\bar{S} = \bar{S}^1 \times \bar{S}^2 \times \dots \times \bar{S}^n$ , where each  $\bar{S}^i$  for  $i \in \{1, 2, \dots, n\}$  represents the set of solutions of system (2), i.e.  $\bar{S}^i$  represents the set of mixed stationary strategies for player  $i$ . On  $\bar{S}$  we define the average payoffs for the players as follows:

$$\psi^i(s^1, s^2, \dots, s^n) = \sum_{x \in X} \sum_{(a^1, a^2, \dots, a^n) \in A(x)} \prod_{k=1}^n s_{x, a^k}^k f^i(x, a^1, a^2, \dots, a^n) q_x, \quad i = 1, 2, \dots, n,$$

where  $q_x$  for  $x \in X$  are determined uniquely from the following system of linear equations

$$\begin{cases} \sum_{x \in X} \sum_{(a^1, a^2, \dots, a^n) \in A(x)} \prod_{k=1}^n s_{x, a^k}^k p_{x, y}^{(a^1, a^2, \dots, a^n)} q_x = q_y, \quad \forall y \in X; \\ \sum_{x \in X} q_x = 1, \end{cases}$$

where  $s^i \in \bar{S}^i$ ,  $i = 1, 2, \dots, n$ . The functions  $\psi^i(s^1, s^2, \dots, s^n)$ ,  $i = 1, 2, \dots, n$  on  $\bar{S}$  define a game in normal form  $\langle \{\bar{S}^i\}_{i=1, n}, \{\psi^i(s)\}_{i=1, n} \rangle$  that corresponds to a stationary average stochastic game with unichain property, where  $\psi^i(s^1, s^2, \dots, s^n) = \omega_x^i(s^1, s^2, \dots, s^n)$ ,  $\forall x \in X$ ,  $i = 1, 2, \dots, n$ .

From Lemma 1 we obtain the following result.

**Lemma 2.** *For an arbitrary unichain stochastic game  $\langle \{\bar{S}^i\}_{i=1, n}, \{\psi^i(s)\}_{i=1, n} \rangle$  each payoff function  $\psi^i(s^i, \hat{s}^{-i})$ ,  $i \in \{1, 2, \dots, n\}$  is quasi-monotonic with respect to  $s^i \in \bar{S}^i$  for arbitrary fixed  $\hat{s}^{-i} \in \bar{S}^{-i}$ .*

Based on Lemma 2 and results from (Debreu, 1952; Glicksberg, 1952) we obtain the following theorem.

**Theorem 1.** *Let  $\langle \{\bar{S}^i\}_{i=1, n}, \{\psi^i(s)\}_{i=1, n} \rangle$  be an average stochastic game determined by  $(X, A, \{X_i\}_{i=1, n}, \{f^i(x, a)\}_{i=1, n}, p, x)$ . If for an arbitrary  $s = (s^1, s^2, \dots, s^n) \in S$  of the game the transition probability matrix  $P^s = (p_{x, y}^s)$  corresponds to a Markov unichain then for the game  $\langle \{\bar{S}^i\}_{i=1, n}, \{\psi^i(s)\}_{i=1, n} \rangle$  there exists a Nash equilibrium  $s^* = (s^{1*}, s^{2*}, \dots, s^{n*})$  which is a Nash equilibrium for an arbitrary starting state  $x \in X$  of the game.*

*Proof.* According to Lemma 2 each payoff  $\psi^i(s^i, \hat{s}^{-i})$ ,  $i \in \{1, 2, \dots, n\}$  is quasi-monotonic with respect to  $s^i \in \bar{S}^i$  for fixed  $\hat{s}^{-i} \in \bar{S}^{-i}$ . Additionally, each payoff  $\psi^i(s)$ ,  $i \in \{1, 2, \dots, n\}$  is continuous on  $\bar{S}$  because the stochastic game is unichain. Then according to (Debreu, 1952; Glicksberg, 1952) the game  $\langle \{\bar{S}^i\}_{i=1, n}, \{\psi^i(s)\}_{i=1, n} \rangle$  possesses a pure Nash equilibrium  $s^* \in \bar{S}$  which is a stationary Nash equilibrium for the unichain average stochastic game with an arbitrary starting state  $x \in X$ .  $\square$



Thus, if we find a pure Nash equilibrium  $s^*$  for the game in normal form  $\langle \{\bar{S}^i\}_{i=1,n}, \{\psi^i(s)\}_{i=1,n} \rangle$  then  $s^*$  is a stationary Nash equilibrium for the average stochastic game with unichain property.

#### 4. Some Results for a Multichain Average Markov Decision Problem

In this section we extend the results from Section 3.1. for the multichain average Markov decision problem, i.e. we show how this decision problem can be formulated in terms of stationary strategies. These results we shall use in the next section for the average stochastic games in general case.

##### 4.1. A Linear Programming Approach for a Multichain Decision Problem

The basic model that we shall use in the sequel for formulation and studying a Markov decision problem in terms of stationary strategies represents the following linear programming problem (Kallenberg, 2016; Lozovanu and Pickl, 2015; Puterman, 2005):

*Maximize*

$$\bar{\psi}(\alpha, \beta) = \sum_{x \in X} \sum_{a \in A(x)} f(x, a) \alpha_{x,a} \quad (12)$$

*subject to*

$$\begin{cases} \sum_{a \in A(y)} \alpha_{y,a} - \sum_{x \in X} \sum_{a \in A(x)} p_{x,y}^a \alpha_{x,a} = 0, & \forall y \in X; \\ \sum_{a \in A(y)} \alpha_{y,a} + \sum_{a \in A(y)} \beta_{y,a} - \sum_{x \in X} \sum_{a \in A(x)} p_{x,y}^a \beta_{x,a} = \theta_y, & \forall y \in X; \\ \alpha_{x,a} \geq 0, \quad \beta_{y,a} \geq 0, & \forall x \in X, \quad a \in A(x), \end{cases} \quad (13)$$

where  $\theta_y$  for  $y \in X$  represent arbitrary positive values that satisfy the condition  $\sum_{y \in X} \theta_y = 1$ . Recall that  $f(x, a)$  denotes the step reward in a state  $x \in X$  for a given action  $a \in A(x)$  in the decision problem and  $p_{x,y}^a$  represent the corresponding probability transitions from a state  $x \in X$  to the states  $y \in X$  for  $a \in A(x)$ , where  $\sum_{y \in X} p_{x,y}^a = 1$ .

This problem generalizes the unichain linear programming model (9), (11) from Section 3.1.. In (13) the restrictions

$$\sum_{a \in A(y)} \alpha_{y,a} + \sum_{a \in A(y)} \beta_{y,a} - \sum_{x \in X} \sum_{a \in A(x)} p_{x,y}^a \beta_{x,a} = \theta_y, \quad \forall y \in X \quad (14)$$

with the condition  $\sum_{y \in X} \theta_y = 1$  generalize the constraint  $\sum_{x \in X} \sum_{a \in A(y)} \alpha_{y,a} = 1$  in the unichain model. This constraint is obtained if we sum (14) over  $y$ . The relationship between feasible solutions of problem (12), (13) and stationary strategies in the average Markov decision problem is the following (see Puterman, 2005):

Let  $(\alpha, \beta)$  be a feasible solution of the linear programming problem (12), (13) and denote  $X_\alpha = \{x \in X \mid \sum_{a \in X} \alpha_{x,a} > 0\}$ . Then  $(\alpha, \beta)$  possesses the properties that

$\sum_{a \in A(x)} \beta_{x,a} > 0$  for  $x \in X \setminus X_\alpha$  and a stationary strategy that corresponds to  $(\alpha, \beta)$  is determined as follows

$$s_{x,a} = \begin{cases} \frac{\alpha_{x,a}}{\sum_{a \in A(x)} \alpha_{x,a}} & \text{if } x \in X_\alpha; \\ \frac{\beta_{x,a}}{\sum_{a \in A(x)} \beta_{x,a}} & \text{if } x \in X \setminus X_\alpha, \end{cases} \quad (15)$$

where  $s_{x,a}$  expresses the probability of choosing the actions  $a \in A(x)$  in the states  $x \in X$ . It is easy to see that the set of feasible solutions of problem (12),(13) generate through (15) the set of stationary strategies  $S$  that corresponds to the set of solution of the following system

$$\begin{cases} \sum_{a \in A(x)} s_{x,a} = 1, & \forall x \in X; \\ s_{x,a} \geq 0, & \forall x \in X, \forall a \in A(x) \end{cases}$$

In (Kallenberg, 2016; Lozovanu and Pickl, 2015; Puterman, 2005) the problem (12), (13) is regarded as the dual model of the following linear programming problem:

Minimize

$$\phi(\varepsilon, \omega) = \sum_{x \in X} \theta_x \omega_x \quad (16)$$

subject to

$$\begin{cases} \varepsilon_x + \omega_x \geq f(x, a) + \sum_{y \in X} p_{x,y}^a \varepsilon_y, & \forall x \in X, \forall a \in A(x); \\ \omega_x \geq \sum_{y \in X} p_{x,y}^a \omega_y, & \forall x \in X, \forall a \in A(x). \end{cases} \quad (17)$$

The optimal value of the objective function in this problem as well as the optimal values of the objective functions in problems (12), (13) and (16), (17) express the optimal average reward when the initial state is chosen according to distribution  $\{\theta_x\}$ . Solving problem (16), (17) we obtain the value  $\omega_x^*$  for each  $x \in X$  that represents the optimal average reward when a transition starts in  $x$  with probability equal to 1. This means that if  $(\alpha^*, \beta^*)$  is an optimal solution of problem (12), (13) then we can determine the optimal strategy  $s^*$  and the optimal values of object functions of problems (16), (17) and (12), (13), where  $\phi(\varepsilon^*, \omega^*) = \bar{\psi}(\alpha^*, \beta^*)$ . An arbitrary optimal solution of problem (12), (13) or of problem (16), (17) determine an optimal strategy  $s^*$  that is an optimal stationary strategy for the multichain decision problem with an arbitrary starting state  $x \in X$ .

**Remark 1.** Problems (12),(13) and (16),(17) can be considered also for the case when  $\theta_x = 0$  for some  $x \in X$ . In particular, if  $\theta_x = 0, \forall x \in X \setminus \{x_0\}$  and  $\theta_{x_0} = 1$  then these problems are transformed into the models with fixed starting state  $x_0$ . In this case for a feasible solution  $(\alpha, \beta)$  the subset  $X \setminus X_\alpha$  may contain states for which  $\sum_{a \in A(x)} \beta_{x,a} = 0$ . In such states it couldn't be used (15) for determining  $s_{x,a}$ . Formula (15) can be used for determining the strategies  $s_{x,a}$  in the states  $x \in X$

for which either  $\sum_{a \in A(x)} \alpha_{x,a} > 0$  or  $\sum_{a \in A(x)} \beta_{x,a} > 0$  and these strategies determine the value of the objective function in the decision problem. In the states  $x \in X_0$ , where

$$X_0 = \{x \in X \mid \sum_{a \in A(x)} \alpha_{x,a} = 0, \sum_{a \in A(x)} \beta_{x,a} = 0\},$$

the strategies of the selection of the actions may be arbitrary because they do not affect the value of the objective function.

#### 4.2. A Multichain Markov Decision Model in Terms of Stationary Strategies

The multichain average Markov decision model in terms of stationary strategies that generalizes the unichain model (7), (8) from Section 3.1. is the following:

*Maximize*

$$\psi(s, q, w) = \sum_{x \in X} \sum_{a \in A(x)} f(x, a) s_{x,a} q_x \quad (18)$$

*subject to*

$$\left\{ \begin{array}{l} q_y - \sum_{x \in X} \sum_{a \in A(x)} p_{x,y}^a s_{x,a} q_x = 0, \quad \forall y \in X; \\ q_y + w_y - \sum_{x \in X} \sum_{a \in A(x)} p_{x,y}^a s_{x,a} w_x = \theta_y, \quad \forall y \in X; \\ \sum_{a \in A(y)} s_{y,a} = 1, \quad \forall y \in X; \\ s_{x,a} \geq 0, \quad \forall x \in X, \forall a \in A(x); \quad w_x \geq 0, \quad \forall x \in X, \end{array} \right. \quad (19)$$

where  $\theta_y$  are the same values as in problem (12), (13) and  $s_{x,a}$ ,  $q_x$ ,  $w_x$  for  $x \in X$ ,  $a \in A(x)$  represent the variables that must be found.

**Theorem 2.** *Optimization problem (18), (19) determines the optimal stationary strategies of the multichain average Markov decision problem.*

*Proof.* Indeed, if we assume that each action set  $A(x), x \in X$  contains a single action  $a'$  then system (13) is transformed into the following system of equations

$$\left\{ \begin{array}{l} q_y - \sum_{x \in X} p_{x,y} q_x = 0, \quad \forall y \in X; \\ q_y + w_y - \sum_{x \in X} p_{x,y} w_x = \theta_y, \quad \forall y \in X \end{array} \right.$$

with conditions  $q_y, w_y \geq 0$  for  $y \in X$ , where  $q_y = \alpha_{y,a'}$ ,  $w_y = \beta_{y,a'}$ ,  $\forall y \in X$  and  $p_{x,y} = p_{x,y}^{a'}$ ,  $\forall x, y \in X$ . This system uniquely determines  $q_x$  for  $x \in X$  and determines  $w_x$  for  $x \in X$  up to an additive constant in each recurrent class of  $P = (p_{x,y})$  (see Puterman, 2005). Here  $q_x$  represents the limiting probability in the state  $x$  when transitions start in the states  $y \in X$  with probabilities  $\theta_y$  and therefore the condition  $q_x \geq 0$  for  $x \in X$  can be released. Note that  $w_x$  for some states may be negative, however always the additive constants in the corresponding recurrent classes can be chosen so that  $w_x$  became nonnegative. In general, we can observe that in (19) the condition  $w_x \geq 0$  for  $x \in X$  can be released and

this does not affect the value of the objective function of the problem. In the case  $|A(x)| = 1, \forall x \in X$  the average cost is determined as  $\psi = \sum_{x \in X} f(x)q_x$ , where  $f(x) = f(x, a), \forall x \in X$ .

If the action sets  $A(x), x \in X$  may contain more than one action then for a given stationary strategy  $s \in \bar{S}$  of a selection of the actions in the states we can find the average cost  $\psi(s)$  in a similar way as above by considering the probability matrix  $P^s = (p_{x,y}^s)$ , where

$$p_{x,y}^s = \sum_{a \in A(x)} p_{x,y}^a s_{x,a} \tag{20}$$

expresses the probability transition from a state  $x \in X$  to a state  $y \in X$  when the strategy  $s$  of selections of the actions in the states is applied. This means that we have to solve the following system of equations

$$\begin{cases} q_y - \sum_{x \in X} p_{x,y}^s q_x = 0, & \forall y \in X; \\ q_y + w_y - \sum_{x \in X} p_{x,y}^s w_x = \theta_y, & \forall y \in X. \end{cases}$$

If in this system we take into account (20) then this system can be written as follows

$$\begin{cases} q_y - \sum_{x \in X} \sum_{a \in A(x)} p_{x,y}^a s_{x,a} q_x = 0, & \forall y \in X; \\ q_y + w_y - \sum_{x \in X} \sum_{a \in A(x)} p_{x,y}^a s_{x,a} w_x = \theta_y, & \forall y \in X. \end{cases} \tag{21}$$

An arbitrary solution  $(q, w)$  of the system of equations (21) uniquely determines  $q_y$  for  $y \in X$  that allows us to determine the average cost per transition

$$\psi(s) = \sum_{x \in X} \sum_{a \in X} f(x, a) s_{x,a} q_x \tag{22}$$

when the stationary strategy  $s$  is applied. If we are seeking for an optimal stationary strategy then we should add to (21) the conditions

$$\sum_{a \in A(x)} s_{x,a} = 1, \forall x \in X; \quad s_{x,a} \geq 0, \forall x \in X, a \in A(x) \tag{23}$$

and to maximize (22) under the constraints (21), (23). In such a way we obtain problem (18), (19) without conditions  $w_x \geq 0$  for  $x \in X$ . As we have noted the conditions  $w_x \geq 0$  for  $x \in X$  do not affect the values of the objective function (18) and therefore we can preserve such conditions that show the relationship of the problem (18), (19) with problem (12), (13).  $\square$

**Corollary 1.** *If  $\theta_x > 0, \forall x \in X$  then an arbitrary optimal strategy  $s^*$  of problem (18),(19) is an optimal stationary strategy for the multichain average decision problem with an arbitrary starting state  $x \in X$ . If  $\theta_x = 0, \forall x \in X \setminus \{x_0\}$  and  $\theta_{x_0} = 1$  then an optimal strategy  $s^*$  of problem (18), (19) is an optimal stationary strategy for the multichain decision problem with starting state  $x_0$ .*

The relationship between feasible solutions of problem (12), (13) and feasible solutions of problem (18), (19) can be established on the basis of the following lemma.

**Lemma 3.** Let  $(s, q, w)$  be a feasible solution of problem (18), (19). Then

$$\alpha_{x,a} = s_{x,a}q_x, \quad \beta_{x,a} = s_{x,a}w_x, \quad \forall x \in X, a \in A(x) \quad (24)$$

represent a feasible solution  $(\alpha, \beta)$  of problem (12), (13) and  $\psi(s, q, w) = \bar{\psi}(\alpha, \beta)$ . If  $(\alpha, \beta)$  is a feasible solution of problem (12), (13) with  $\theta_x > 0, \forall x \in X$  then a feasible solution  $(s, q, w)$  of problem (18), (19) can be determined as follows:

$$s_{x,a} = \begin{cases} \frac{\alpha_{x,a}}{\sum_{a \in A(x)} \alpha_{x,a}} & \text{for } x \in X_\alpha, a \in A(x); \\ \frac{\beta_{x,a}}{\sum_{a \in A(x)} \beta_{x,a}} & \text{for } x \in X \setminus X_\alpha, a \in A(x); \end{cases} \quad (25)$$

$$q_x = \sum_{a \in A(x)} \alpha_{x,a}, \quad w_x = \sum_{a \in A(x)} \beta_{x,a} \quad \text{for } x \in X.$$

If  $(\alpha, \beta)$  is a feasible solution of problem (12), (13) for which  $\theta_x = 0, \forall x \in X \setminus \{x_0\}$  then a feasible solution  $(s, q, w)$  of problem (18), (19) can be determined as follows:

$$s_{x,a} = \begin{cases} \frac{\alpha_{x,a}}{\sum_{a \in A(x)} \alpha_{x,a}} & \text{for } x \in X_\alpha, a \in A(x); \\ \frac{\beta_{x,a}}{\sum_{a \in A(x)} \beta_{x,a}} & \text{for } x \in X \setminus (X_\alpha \cup X_0), a \in A(x); \\ \text{arbitrary} & \text{for } x \in X_0, a \in A(x), \end{cases}$$

$$q_x = \sum_{a \in A(x)} \alpha_{x,a}, \quad w_x = \sum_{a \in A(x)} \beta_{x,a} \quad \text{for } x \in X,$$

where  $X_0 = \{x \in X \mid \sum_{a \in A(x)} \alpha_{x,a} = 0, \sum_{a \in A(x)} \beta_{x,a} = 0\}$ .

*Proof.* If  $(s, q, w)$  is a feasible solution of problem (18), (19) and  $(\alpha, \beta)$  is determined according to (24) then by introducing (24) in (12), (13) we obtain that (13) is transformed in (19) and  $\psi(s, q, w) = \bar{\psi}(\alpha, \beta)$ , i.e.  $(\alpha, \beta)$  is a feasible solution of problem (12), (13). The second part of lemma follows directly from the properties of a feasible solutions of problems (12), (13) and (18), (19).  $\square$

### 4.3. The Main Properties of the Problem in Stationary Strategies

Using problem (18), (19) we can now extend the results from Section 3.1. for the general case of an average Markov decision problem.

**Theorem 3.** Let an average Markov decision problem be given and consider the function

$$\psi(s) = \sum_{x \in X} \sum_{a \in A(x)} f_{(x,a)} s_{x,a} q_x, \quad (26)$$

where  $q_x$  for  $x \in X$  satisfies the condition

$$\begin{cases} q_y - \sum_{x \in X} \sum_{a \in A(x)} p_{x,y}^a s_{x,a} q_x = 0, & \forall y \in X; \\ q_y + w_y - \sum_{x \in X} \sum_{a \in A(x)} p_{x,y}^a s_{x,a} w_x = \theta_y, & \forall y \in X. \end{cases} \quad (27)$$

Then on the set  $\bar{S}$  of solutions of the system

$$\begin{cases} \sum_{a \in A(x)} s_{x,a} = 1, & \forall x \in X; \\ s_{x,a} \geq 0, & \forall x \in X, a \in A(x) \end{cases} \quad (28)$$

the function  $\psi(s)$  depends only on  $s_{x,a}$  for  $x \in X, a \in A(x)$  and  $\psi(s)$  is quasi-monotonic on  $\bar{S}$  (i.e.  $\psi(s)$  is quasi-convex and quasi-concave on  $\bar{S}$ ).

The proof of this theorem can be found in (Lozovanu, 2018).

## 5. The Main Results for Average Stochastic Games

In this section we extend the results from Section 3.2. for the case of multichain average stochastic game in stationary strategies. We show that a multichain average stochastic game in normal form can be formulated as the game in which the payoffs possess of quasi-monotonic property with respect to the corresponding strategies of the players. Based on this property we present some conditions for the existence of stationary Nash equilibria in the multichain average stochastic game.

### 5.1. A Normal Form of Average Stochastic Game in Stationary Strategies

The multichain average stochastic game in stationary strategies that generalizes the unichain game model from Section 3.2. is the following:

Let  $\bar{S}^i, i \in \{1, 2, \dots, n\}$  be the set of solutions of the system

$$\begin{cases} \sum_{a^i \in A^i(x)} s_{x,a^i}^i = 1, & \forall x \in X; \\ s_{x,a^i}^i \geq 0, & \forall x \in X, a^i \in A^i(x). \end{cases} \quad (29)$$

that determines the set of stationary strategies of player  $i$ . Each  $\bar{S}^i$  is a convex compact set and an arbitrary extreme point corresponds to a basic solution  $s^i$  of system (29), where  $s_{x,a^i}^i \in \{0, 1\}, \forall x \in X, a^i \in A^i(x)$ , i.e each basic solution of this system corresponds to a pure stationary strategy of player  $i$ . On the set  $\bar{S} = \bar{S}^1 \times \bar{S}^2 \times \dots \times \bar{S}^n$  we define  $n$  payoff functions

$$\begin{cases} \psi_\theta^i(s^1, s^2, \dots, s^n) = \sum_{x \in X} \sum_{(a^1, a^2, \dots, a^n) \in A(x)} \prod_{k=1}^n s_{x,a^k}^k f^i(x, a^1, a^2 \dots a^n) q_x, \\ i = 1, 2, \dots, n, \end{cases} \quad (30)$$

where  $q_x$  for  $x \in X$  are determined uniquely from the following system of linear equations

$$\begin{cases} q_y - \sum_{x \in X} \sum_{(a^1, a^2, \dots, a^n) \in A(x)} \prod_{k=1}^n s_{x, a^k}^k p_{x, y}^{(a^1, a^2, \dots, a^n)} q_x = 0, & \forall y \in X; \\ q_y + w_y - \sum_{x \in X} \sum_{(a^1, a^2, \dots, a^n) \in A(x)} \prod_{k=1}^n s_{x, a^k}^k p_{x, y}^{(a^1, a^2, \dots, a^n)} w_x = \theta_y, & \forall y \in X, \end{cases} \quad (31)$$

for an arbitrary fixed  $s = (s^1, s^2, \dots, s^n) \in \bar{S}$ . The functions  $\psi_\theta^i(s^1, s^2, \dots, s^n)$ ,  $i = 1, 2, \dots, n$ , represent the payoff functions for the average stochastic game in normal form  $\langle \{\bar{S}^i\}_{i=1, \dots, n}, \{\psi_\theta^i(s)\}_{i=1, \dots, n} \rangle$ . This game is determined by the tuple  $(X, \{A^i(x)\}_{i=1, \dots, n}, \{f^i(x, a)\}_{i=1, \dots, n}, p, \{\theta_y\})$  where  $\theta_y$  for  $y \in X$  are given nonnegative values such that  $\sum_{y \in X} \theta_y = 1$ .

If  $\theta_y = 0, \forall y \in X \setminus \{x_0\}$  and  $\theta_{x_0} = 1$  then we obtain an average stochastic game in normal form  $\langle \{\bar{S}^i\}_{i=1, \dots, n}, \{\omega_{x_0}^i(s)\}_{i=1, \dots, n} \rangle$  when the starting state  $x_0$  is fixed, i.e.  $\psi_\theta^i(s^1, s^2, \dots, s^n) = \omega_{x_0}^i(s^1, s^2, \dots, s^n)$ ,  $i = 1, 2, \dots, n$ . So, in this case the game is determined by  $(X, \{A^i(x)\}_{i=1, \dots, n}, \{f^i(x, a)\}_{i=1, \dots, n}, p, x_0)$ .

If  $\theta_y > 0, \forall y \in X$  and  $\sum_{y \in X} \theta_y = 1$  then we obtain an average stochastic game when the play starts in the states  $y \in X$  with probabilities  $\theta_y$ . In this case for the payoffs of the players in the game in normal form we have

$$\psi_\theta^i(s^1, s^2, \dots, s^n) = \sum_{y \in X} \theta_y \omega_y^i(s^1, s^2, \dots, s^n), \quad i = 1, 2, \dots, n. \quad (32)$$

## 5.2. The Main Properties of Average Stochastic Games in Normal Form

Based on results from the previous section we can prove the following results.

**Theorem 4.** *Let  $\langle \{\bar{S}^i\}_{i=1, \dots, n}, \{\psi_\theta^i(s)\}_{i=1, \dots, n} \rangle$  be the game in normal form for the average stochastic game in stationary strategies determined by  $(X, \{A^i(x)\}_{i=1, \dots, n}, \{f^i(x, a)\}_{i=1, \dots, n}, p, \{\theta_y\})$  where  $\theta_y > 0, \forall y \in X, \sum_{y \in X} \theta_y = 1$ . If for this game there exists a Nash equilibrium  $s^* = (s^{1*}, s^{2*}, \dots, s^{n*})$  then it is a Nash equilibrium for the game in normal form  $\langle \{\bar{S}^i\}_{i=1, \dots, n}, \{\omega_y^i(s)\}_{i=1, \dots, n} \rangle$  with an arbitrary  $y \in X$ , i.e.  $s^* = (s^{1*}, s^{2*}, \dots, s^{n*})$  is a stationary Nash equilibrium of the average stochastic game with an arbitrary starting state  $y \in X$ . Conversely, if for an arbitrary starting state  $y \in X$  the corresponding game in normal form  $\langle \{\bar{S}^i\}_{i=1, \dots, n}, \{\omega_y^i(s)\}_{i=1, \dots, n} \rangle$  has a Nash equilibrium then for an arbitrary distribution function  $\{\theta_y\}$  on  $X$  with  $\theta_y > 0, \forall y \in X$  ( $\sum_{y \in X} \theta_y = 1$ ) the corresponding game in normal form  $\langle \{\bar{S}^i\}_{i=1, \dots, n}, \{\psi_\theta^i(s)\}_{i=1, \dots, n} \rangle$  of the average stochastic game determined by  $(X, \{A^i(x)\}_{i=1, \dots, n}, \{f^i(x, a)\}_{i=1, \dots, n}, p, \{\theta_y\})$  has a Nash equilibrium  $s^* = (s^{1*}, s^{2*}, \dots, s^{n*})$  which is a Nash equilibrium for each of the game in normal form  $\langle \{\bar{S}^i\}_{i=1, \dots, n}, \{\omega_y^i(s)\}_{i=1, \dots, n} \rangle$  with the corresponding starting states  $y \in Y$ .*

*Proof.* ( $\Rightarrow$ ) Let  $s^* = (s^{1*}, s^{2*}, \dots, s^{n*})$  be a Nash equilibrium for the game in normal form  $\langle \{\bar{S}^i\}_{i=1, \dots, n}, \{\psi_\theta^i(s)\}_{i=1, \dots, n} \rangle$  determined by  $(X, \{A^i(x)\}_{i=1, \dots, n}, \{f^i(x, a)\}_{i=1, \dots, n}, p, \{\theta_y\})$ , where  $\theta_y > 0, \forall y \in X, \sum_{y \in X} \theta_y = 1$ . Then  $(s^{1*}, s^{2*}, \dots, s^{n*})$  is a Nash

equilibrium for the average stochastic game  $\langle \{\bar{S}^i\}_{i=1,n}, \{\psi_{\theta'}^i(s)\}_{i=1,n} \rangle$  with an arbitrary distribution  $\{\theta'_y\}$  on  $X$ , where  $\theta'_y > 0, \forall y \in X, \sum_{y \in X} \theta'_y = 1$ , i.e

$$\psi_{\theta'}^i(s^{i*}, s^{-i*}) \geq \psi_{\theta'}^i(s^i, s^{-i*}), \quad \forall s^i \in \bar{S}^i, \quad i = 1, 2, \dots, n.$$

If here we express  $\psi_{\theta'}^i$  via  $\omega_y^i$  using (32) then we obtain

$$\sum_{y \in X} \theta'_y (\omega_y^i(s^{i*}, s^{-i*}) - \omega_y^i(s^i, s^{-i*})) \geq 0, \forall s^i \in \bar{S}^i, \quad i = 1, 2, \dots, n.$$

This property holds for arbitrary  $\theta'_y > 0, \forall y \in X$  such that  $\sum_{y \in Y} \theta'_y = 1$  and therefore for an arbitrary  $y \in X$  we have

$$\omega_y^i(s^{i*}, s^{-i*}) - \omega_y^i(s^i, s^{-i*}) \geq 0, \quad \forall s^i \in \bar{S}^i, \quad i = 1, 2, \dots, n.$$

So,  $(s^{1*}, s^{2*}, \dots, s^{n*})$  is a Nash equilibrium for each of the game in normal form  $\langle \{\bar{S}^i\}_{i=1,n}, \{\omega_y^i(s)\}_{i=1,n} \rangle$  with the corresponding starting states  $y \in X$ .

( $\Leftarrow$ ) Assume that for each starting state  $y \in X$  the average stochastic game  $\langle \{\bar{S}^i\}_{i=1,n}, \{\omega_y^i(s)\}_{i=1,n} \rangle$  has a Nash equilibrium. Let us show that for the game  $\langle \{\bar{S}^i\}_{i=1,n}, \{\psi_{\theta}^i(s)\}_{i=1,n} \rangle$ , determined by  $(X, \{A^i(x)\}_{i=1,n}, \{f^i(x, a)\}_{i=1,n}, p, \{\theta_y\})$ , where  $\theta_y > 0, \forall y \in X, \sum_{y \in X} \theta_y = 1$  there exists a Nash equilibrium. We prove this using an auxiliary average stochastic game with a new starting state  $z$  and the set of states  $X \cup \{z\}$ , where for an arbitrary state  $x \in X$  each player  $i \in \{1, 2, \dots, n\}$  has the same set of actions  $A^i(x)$ , the same payoffs  $f^i(x, a)$  for  $a \in A(x)$  and the same transition probability distributions  $p_{x,y}^a$  for  $a \in A(x)$  as in the game determined by  $(X, \{A^i(x)\}_{i=1,n}, \{f^i(x, a)\}_{i=1,n}, p, \{\theta_y\})$ ; in the state  $z$  of the auxiliary game each player  $i \in \{1, 2, \dots, n\}$  has a single action  $a_z^i$  and  $A(z)$  contains a unique profile  $a_z = (a_z^1, a_z^2, \dots, a_z^n)$  for which  $p_{z,z}^{a_z} = 0, p_{z,y}^{a_z} = \theta_y, \forall y \in X$  and  $f^i(z, a_z) = 0, i = 1, 2, \dots, n$ . Obviously, for the auxiliary average stochastic game with starting state  $z$ , determined by  $(X \cup \{z\}, \{A^i(x) \cup A^i(z)\}_{i=1,n}, \{f^i(x, a), f^i(z, a_z)\}_{i=1,n}, p \cup \{p_{x,y}^{a_z}\}, z)$  there exists a stationary Nash equilibrium because a Nash equilibrium exists for an arbitrary average stochastic game  $\langle \{\bar{S}^i\}_{i=1,n}, \{\omega_y^i(s)\}_{i=1,n} \rangle$  with starting state  $y \in Y$ . Taking into account that the auxiliary game is equivalent to the average stochastic game, determined by  $(X, \{A^i(x)\}_{i=1,n}, \{f^i(x, a)\}_{i=1,n}, p, \{\theta_y\})$ , where  $\theta_y > 0, \forall y \in X, \sum_{y \in X} \theta_y = 1$ , we obtain that the considered average stochastic game with a random starting state has a stationary Nash equilibrium  $s^* = (s^{1*}, s^{2*}, \dots, s^{n*})$  which is a stationary Nash equilibrium for the average stochastic game  $\langle \{\bar{S}^i\}_{i=1,n}, \{\omega_y^i(s)\}_{i=1,n} \rangle$  with an arbitrary starting state  $y \in Y$ .  $\square$

From Theorem 3 we can easily obtain the following result.

**Lemma 4.** For an arbitrary game in normal form  $\langle \{\bar{S}^i\}_{i=1,n}, \{\psi_{\theta}^i(s)\}_{i=1,n} \rangle$  with  $\theta_x \geq 0, \forall x \in X, \sum_{y \in X} \theta_y = 1$  each payoff function  $\psi_{\theta}^i(s^1, s^2, \dots, s^n), i \in \{1, 2, \dots, n\}$  possesses the property that  $\psi_{\theta}^i(s^i, \hat{s}^{-i})$  is quasi-monotonic with respect to  $s^i \in \bar{S}^i$  for arbitrary fixed  $\hat{s}^{-i} \in \bar{S}^{-i}$ .

*Proof.* Indeed, if players  $1, 2, \dots, i-1, i+1, \dots, n$  fix their stationary strategies  $\hat{s}^k \in \bar{S}^k, k = 1, 2, \dots, i-1, i+1, \dots, n$ , then we obtain an average decision



problem with respect to  $s^i \in \bar{S}^i$  and an average cost function  $\psi_\theta^i(s^i, \hat{s}^{-i})$ . According to Theorem 3  $\psi_\theta^i(s^i, \hat{s}^{-i})$  possesses the property that the value of this function is uniquely determined by  $s^i \in \bar{S}^i$  and it is quasi-monotone with respect to  $s^i$  on  $\bar{S}^i$ .  $\square$

Using this lemma we can prove the following result.

**Theorem 5.** *Let  $\langle \{\bar{S}^i\}_{i=1,n}, \{\psi_\theta^i(s)\}_{i=1,n} \rangle$  be the normal form game for the average stochastic game determined by  $(X, A, \{X_i\}_{i=1,n}, \{f^i(x, a)\}_{i=1,n}, p, \{\theta_x\})$  where  $\theta_x > 0$ ,  $\forall x \in X$ ,  $\sum_{y \in X} \theta_y = 1$ . If each function  $\psi_\theta^i$ ,  $i \in \{1, 2, \dots, n\}$  is continuous on  $\bar{S} = \bar{S}^1 \times \bar{S}^2 \times \dots \times \bar{S}^n$  then the game  $\langle \{\bar{S}^i\}_{i=1,n}, \{\psi_\theta^i(s)\}_{i=1,n} \rangle$  possesses a Nash equilibrium  $s^* = (s^{1*}, s^{2*}, \dots, s^{n*})$  which is a stationary Nash equilibrium for the average stochastic game with an arbitrary starting state  $y \in X$ .*

*Proof.* Indeed, according to Lemma 4 each function  $\psi^i(s^1, s^2, \dots, s^n)$ ,  $i \in \{1, 2, \dots, n\}$  satisfies the condition that  $\psi^i(s^i, \bar{s}^{-i})$ , is quasi-monotonic with respect to  $s^i \in \bar{S}^i$  for arbitrary fixed  $\bar{s}^{-i} \in \bar{S}^{-i}$ . In the considered game each subset  $\bar{S}^i$  is convex and compact and according to the condition of the theorem each payoff function  $\psi_\theta^i(s^1, s^2, \dots, s^n)$ ,  $i \in \{1, 2, \dots, n\}$  is continuous on  $\bar{S}$ . Based on results from (Dasgupta and Maskin, 1986; Debreu, 1952; Reny, 1999; Simon, 1987) these conditions provide the existence of a Nash equilibrium  $s^* = (s^{1*}, s^{2*}, \dots, s^{n*})$  for the game  $\langle \{\bar{S}^i\}_{i=1,n}, \{\psi_\theta^i(s)\}_{i=1,n} \rangle$ . According to Theorem 4 such an equilibrium is a Nash equilibrium for the game  $\langle \{\bar{S}^i\}_{i=1,n}, \{\omega_y^i(s)\}_{i=1,n} \rangle$  with an arbitrary starting state  $y \in X$ .  $\square$

**Remark 2.** Theorems 4 and 5 are valid also for the case of the game  $\langle \{\bar{S}^i\}_{i=1,n}, \{\psi_\theta^i(s)\}_{i=1,n} \rangle$  when  $\theta_y = 0$  for some  $y \in X$ , however in this case we obtain stationary Nash equilibria only for the games  $\langle \{\bar{S}^i\}_{i=1,n}, \{\omega_z^i(s)\}_{i=1,n} \rangle$  with starting states  $z \in X^+$ , where  $X^+ = \{z \in X | \theta_z > 0\}$ .

**Remark 3.** Theorem 5 holds also for the case when the payoffs are not continuous but satisfy so-called graph-continuous property from (Dasgupta and Maskin, 1986).

## 6. Stationary Equilibria for Average Stochastic Positional Games

Average stochastic positional games have been introduced in (Lozovanu, 2018) as a generalization of mean payoff games from (Ehrenfeucht and Mycielski, 1979). An average stochastic positional game represents an average stochastic game in which the set of states is divided into several disjoint subsets such that each subset represents the position set for one of the player and each player controls the Markov process only in his position set. In such a game each player chooses actions in his position set in order to maximize his average reward per transition.

An average stochastic positional game with  $n$  players is determined by the following elements:

- a state space  $X$  (which we assume to be finite);
- a partition  $X = X_1 \cup X_2 \cup \dots \cup X_n$  where  $X_i$  represents the position set of player  $i \in \{1, 2, \dots, n\}$ ;

- a finite set  $A(x)$  of actions in each state  $x \in X$ ;
- a step reward  $f^i(x, a)$  with respect to each player  $i \in \{1, 2, \dots, n\}$  in each state  $x \in X$  and for an arbitrary action  $a \in A(x)$ ;
- a transition probability function  $p : X \times \prod_{x \in X} A(x) \times X \rightarrow [0, 1]$  that gives the probability transitions  $p_{x,y}^a$  from an arbitrary  $x \in X$  to an arbitrary  $y \in X$  for a fixed action  $a \in A(x)$ , where  $\sum_{y \in X} p_{x,y}^a = 1, \forall x \in X, a \in A(x)$ ;
- a starting state  $x_0 \in X$ .

The game starts at the moment of time  $t = 0$  in the state  $x_0$  where the player  $i \in \{1, 2, \dots, m\}$  who is the owner of position  $x_0$  ( $x_0 \in X_i$ ) chooses an action  $a_0 \in A(x_0)$  and determines the rewards  $f^1(x_0, a_0), f^2(x_0, a_0), \dots, f^m(x_0, a_0)$  for the corresponding players  $1, 2, \dots, m$ . After that the game passes to a state  $y = x_1 \in X$  according to probability distribution  $\{p_{x_0,y}^{a_0}\}$ . At the moment of time  $t = 1$  the player  $k \in \{1, 2, \dots, n\}$  who is the owner of the state position  $x_1$  ( $x_1 \in X_k$ ) chooses an action  $a_1 \in A(x_1)$  and players  $1, 2, \dots, m$  receive the corresponding rewards  $f^1(x_1, a_1), f^2(x_1, a_1), \dots, f^n(x_1, a_1)$ . Then the game passes to a state  $y = x_2 \in X$  according to probability distribution  $\{p_{x_1,y}^{a_1}\}$  and so on indefinitely. Such a play of the game produces a sequence of states and actions  $x_0, a_0, x_1, a_1, \dots, x_t, a_t, \dots$  that defines a stream of stage rewards  $f^1(x_t, a_t), f^2(x_t, a_t), \dots, f^n(x_t, a_t), t = 0, 1, 2, \dots$ . The average stochastic positional game is the game with payoffs of the players

$$\omega_{x_0}^i = \liminf_{t \rightarrow \infty} \mathbb{E} \left( \frac{1}{t} \sum_{\tau=0}^{t-1} f^i(x_\tau, a_\tau) \right), \quad i = 1, 2, \dots, n.$$

If  $p_{x,a} \in \{0, 1\}$  then the average stochastic positional game becomes a mean payoff game. The problem of the existence of pure and mixed stationary equilibria in a stochastic positional games has been studied in (Lozovanu, 2018; Lozovanu, 2019). The pure and mixed stationary strategies in such a game can be defined in analogous way as for a stochastic game, taking into account that each player select actions only in his state positions and determines in these states the step rewards for all players. Thus, a stationary strategy of player  $i \in \{1, 2, \dots, n\}$  in a stochastic positional game is a mapping  $s^i$  that provides for every state  $x \in X_i$  a probability distribution over the set of actions  $A(x)$ . This means that the set of stationary strategies  $\bar{S}^i$  of player  $i \in \{1, 2, \dots, n\}$  can be identified with the set of solutions of the system

$$\begin{cases} \sum_{a^i \in A(x)} s_{x,a}^i = 1, & \forall x \in X_i; \\ s_{x,a}^i \geq 0, & \forall x \in X_i, \forall a \in A(x). \end{cases} \quad (33)$$

The payoffs  $\psi_\theta^i(s^1, s^2, \dots, s^m), i = 1, 2, \dots, n$  on  $\bar{S} = \bar{S}^1 \times \bar{S}^2 \times \dots \times \bar{S}^n$  for the game in normal form for the considered positional game we can obtain from (30), (31) if we take into account the particularity of the average stochastic positional game. So,

$$\psi_\theta^i(s^1, s^2, \dots, s^m) = \sum_{k=1}^m \sum_{x \in X_k} \sum_{a \in A(x)} s_{x,a}^k f^i(x, a) q_x, \quad i = 1, 2, \dots, m, \quad (34)$$

where  $q_x$  for  $x \in X$  are determined uniquely from the following system of linear equations

$$\begin{cases} q_y - \sum_{k=1}^m \sum_{x \in X_k} \sum_{a \in A(x)} s_{x,a}^k p_{x,y}^a q_x = 0, & \forall y \in X; \\ q_y + w_y - \sum_{k=1}^m \sum_{x \in X_k} \sum_{a \in A(x)} s_{x,a}^k p_{x,y}^a w_x = \theta_y, & \forall y \in X \end{cases} \quad (35)$$

In (Lozovanu, 2018) is shown that the game  $\langle \{\bar{S}^i\}_{i=1,m}, \{\psi_\theta^i(s)\}_{i=1,m} \rangle$  defined according to (2) - (35) possesses a Nash equilibrium which is a stationary Nash equilibrium for the average stochastic positional game with an arbitrary starting state  $y \in X$ . Moreover, in (Lozovanu, 2019) is shown that for two-player zero-sum stochastic positional games and for  $n$ -player average stochastic positional games with unichain property there exist stationary Nash equilibria in pure strategies.

## 7. Conclusion

An arbitrary average stochastic game with finite state and action spaces can be formulated in terms of stationary strategies as a game in normal form where each payoff is quasi-monotonic (quasi-concave and quasi-convex) with respect to the strategy of the corresponding player. Such a normal form game (the game model from Section 5) allows to determine all stationary Nash equilibria of the average stochastic game if stationary Nash equilibria exist. If the payoffs of the game in normal form are continuous or graph continuous then stationary Nash equilibria exist. For an average stochastic game with unichain property and for an average stochastic positional game stationary Nash equilibria always exist and all stationary equilibria can be found by using the corresponding game models in normal form from Section 4 and Section 6. For two-player zero-sum average stochastic positional games and for  $n$ -player average stochastic positional games with unichain property there exist stationary equilibria in pure strategies

## References

- Boyd, S. and Vandenberghe, L. (2004). *Convex optimization*. Cambridge university press.
- Dasgupta, P. and Maskin, E. (1986). *The existence of equilibrium in discontinuous economic games, I: Theory*. The Review of economic studies, **53**, 1–26.
- Debreu, G. (1952). *A social equilibrium existence theorem*. Proceedings of the National Academy of Sciences, **38**, 886–893.
- Ehrenfeucht, A., Mycielski, J. (1979). *Positional strategies for mean payoff games*. Int. J. of Game Theory, **8**, 109–113.
- Filar, J., Vrieze, K. (1997). *Competitive Markov Decision Processes*. New York, NY, Springer.
- Filar, J. A., Schultz, T. A., Thuijsman, F., Vrieze, O. (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 the Hiroshima University, ser. math., **28**, 89–93.
- Flesch, J., Thuijsman, F., Vrieze, K. (1997). *Cyclic Markov equilibria in stochastic games*. International Journal of Game Theory, **26**, 303–314.
- Gillette, D. (1957). *Stochastic games with zero stop probabilities*. Contributions to the Theory of Games, **3**, 179–187.

- Glicksberg, I. L. (1952). *A further generalization of the Kakutani fixed point theorem with application to Nash equilibrium points*. Proceedings of the American Mathematical Society, **38**, 170–174.
- Kallenberg, L. (2016). *Markov decision processes*. University of Leiden, Netherland 2016.
- Lozovanu, D. (2011). *The game-theoretical approach to Markov decision problems and determining Nash equilibria for stochastic positional games*. International Journal of Mathematical Modelling and Numerical Optimisation, **2**, 162–174.
- Lozovanu, D. (2018). *Stationary Nash equilibria for average stochastic positional games*. Chapter 9 in the book "Frontiers of Dynamic Games, Static and Dynamic Game Theory: Foundation and Application"(L.Petrosyan et al. eds), Springer, 139–163.
- Lozovanu, D. (2019). *Pure and Mixed Stationary Nash equilibria for average stochastic positional games*. Chapter 8 in the book "Frontiers of Dynamic Games, Static and Dynamic Game Theory: Foundation and Application"(L.Petrosyan et al. eds), Springer, 131–174.
- Lozovanu, D., Pickl, S. (2015). *Optimization of stochastic discrete systems and control on complex networks*. Springer.
- Mertens, J.-F., Neyman, A. (1981). *Stochastic games*. International Journal of Game Theory, **10**, 53–66.
- Neyman, A., Sorin, S. (2003). *Stochastic games and applications*. NATO science series, C, 569, Mathematical and physical sciences, Kluwer Academic Publishers.
- Puterman, M. L. (2005). *Markov decision processes: Discrete stochastic dynamic programming*. Wiley, New Jersey.
- Reny, P. J. (1999). *On the existence of pure and mixed strategy Nash equilibria in discontinuous games*. Econometrica, **67**, 1029–1056.
- Rogers, P. D. (1969). *Nonzero-sum stochastic games*. Technical Report, DTIC Document.
- 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, **39**, 1095–1100.
- Simon, L. K. (1987). *Games with discontinuous payoffs*. The Review of Economic Studies, **54**, 569–597.
- Sobel, M. J. (1971). *Noncooperative stochastic games*. The Annals of Mathematical Statistics, **42**, 1930–1935.
- Solan, E. (2009). *Stochastic games*. In: *Encyclopedia of Complexity and Systems Science*. 8698–8708, Springer.
- Solan, E., Vieille, N. (2010). *Computing uniformly optimal strategies in two-player stochastic games*. Economic Theory, **42**, 237–253.
- Takahashi, M. (1964). *Equilibrium points of stochastic non-cooperative n-person games*. Journal of Science of the Hiroshima University, Series AI, Math., **28**, 95–99.
- Tijs, S., Vrieze, O. (1986). *On the existence of easy initial states for undiscounted stochastic games*. Mathematics of Operations Research, **11**, 506–513.
- Vieille, N. (2002). *Stochastic games: Recent results*. Handbook of Game Theory with Economic Applications, **3**, 1833–1850.
- Vieille, N. (2009). *Equilibrium in 2-person stochastic games I,II*. Israel Journal of Mathematics, 8698–8708.
- Vrieze, O. J. (1987). *Stochastic games with finite state and action spaces*. CWI Tracts, **33**, 1–221.