

Stationary Nash Equilibria for Two-Player Average Stochastic Games with Finite State and Action Spaces

Dmitrii Lozovanu¹ and Stefan Pickl²

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

E-mail: lozovanu@math.md

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

E-mail: stefan.pickl@unibw.de

Abstract The problem of the existence and determining stationary Nash equilibria in two-player average stochastic games with finite state and action spaces is considered. We show that an arbitrary two-player average stochastic game can be formulated in the terms of stationary strategies where each payoff is graph-continuous and quasimonotonic with respect to player's strategies. Based on this result we ground an approach for determining the optimal stationary strategies of the players in the considered games. Moreover, based on the proposed approach a new proof of the existence of stationary Nash equilibria in two-player average stochastic games is derived and the known methods for determining the optimal strategies for the games with quasimonotonic payoffs can be applied.

Keywords: two-players stochastic games, average payoffs, stationary Nash equilibria, optimal stationary strategies

1. Introduction

The aim of this paper is to propose a new approach for determining stationary Nash equilibria in two-player average stochastic games with finite state and action spaces. We ground such an approach by using a new model in stationary strategies for the considered class of average stochastic games. We show that the payoffs of the players in the proposed model are quasimonotonic (i.e quasiconvex and quasiconcave) with respect to the corresponding strategies of the players and satisfy the graph-continuity property in the sense of Dasgupta and Maskin, 1986. Based on these results a new proof of the existence of stationary Nash equilibria in the considered two-player average stochastic games is obtained and a new approach for determining the optimal stationary strategies of the players is proposed.

Note that two-player stochastic games with average and discounted payoffs have been studied by Mertens and Neyman, 1981, Vielle, 2000, Solan and Vielle, 2010, who proved the existence of stationary Nash equilibria and proposed computing procedures for determining the optimal stationary strategies of the players in two-player stochastic games. The approach we propose for two-player average stochastic games differ from the mentioned ones and it can be extended for n -player average stochastic games if the mentioned graph-continuous property for the payoffs holds. However the graph-continuous property for the average stochastic games with $n \geq 3$ players may not take place. It is well known that for an n -player average stochastic game ($n \geq 3$) a stationary Nash equilibrium may not exist. This fact has been shown by Flesch, Thuijman and Vrieze, 1997, who constructed an example of 3-player average stochastic game with fixed starting state for which a stationary Nash

equilibrium does not exist. Tijds and Vrieze, 1986, have shown that for an arbitrary average stochastic game with a finite set of states always exists a non empty subset of starting states for which a stationary Nash equilibrium exists. In the general case the problem of determining the states in average stochastic games for which stationary Nash equilibria exist is an open problem.

2. A Two-Player Average Stochastic Game in Stationary Strategies

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

2.1. The Framework of a Two-Person Stochastic Game

A stochastic game with two players consists of the following elements:

- a state space X (which we assume to be finite);
- a finite set $A^1(x)$ of actions of player 1 for an arbitrary state $x \in X$;
- a finite set $A^2(x)$ of actions of player 2 for an arbitrary state $x \in X$;
- a payoff $f^1(x, a)$ with respect to player 1 for each state $x \in X$ and for an arbitrary action vector $a = (a^1, a^2) \in A^1(x) \times A^2(x)$;
- a payoff $f^2(x, a)$ with respect to player 2 for each state $x \in X$ and for an arbitrary action vector $a = (a^1, a^2) \in A^1(x) \times A^2(x)$;
- a transition probability function $p : X \times \prod_{x \in X} (A^1(x) \times A^2(x)) \times X \rightarrow [0, 1]$ that gives the probability transitions $p_{x,y}^a$ from an arbitrary $x \in X$ to an arbitrary $y \in Y$ for every action vector $a = (a^1, a^2) \in A^1(x) \times A^2(x)$, where $\sum_{y \in X} p_{x,y}^a = 1, \forall x \in X, a \in A^1(x) \times A^2(x)$;
- 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 players observe state x_t and simultaneously and independently choose actions $a_t^i \in A^i(x_t), i = 1, 2$. Then nature selects a state $y = x_{t+1}$ according to probability transitions $p_{x_t,y}^{a_t}$ for the given action vector $a_t = (a_t^1, a_t^2)$. 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

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

where $\omega_{x_0}^i$ expresses the average payoff per transition of player i in an infinite game.

Each player has the aim to maximize his average payoff per transition. In the case $i = 1$ this game becomes the average Markov decision problem with a transition probability function $p : X \times \prod_{x \in X} A(x) \times X \rightarrow [0, 1]$ and immediate 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 of the actions in the states.

2.2. Pure and mixed stationary strategies of the players

A strategy of player $i \in \{1, 2\}$ in a stochastic game is a mapping s^i that for every state $x_t \in X$ provides 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 does not depend on t), then s^i is called a *stationary strategy*, otherwise s^i is called a non-stationary strategy. This means that a *pure stationary strategy* of player $i \in \{1, 2\}$ can be regarded as a map

$$s^i : x \rightarrow a^i \in A^i(x) \text{ for } x \in X$$

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

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

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

2.3. Average stochastic games in pure stationary strategies

Let $s = (s^1, s^2)$ be a profile of pure stationary strategies of the players and denote by $a(s) = (a^1(s), a^2(s)) \in A^1(x) \times A^2(x)$ 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 $\omega_{x_0}^1(s), \omega_{x_0}^2(s)$ for the players are determined as follows

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

where $q_{x_0,y}^s$ represent the limiting probabilities in the states $y \in X$ for the Markov process with 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 ω_x^1, ω_x^2 can be determined for an arbitrary starting state $x \in X$ of the game. The functions $\omega_{x_0}^1(s), \omega_{x_0}^2(s)$ on $S = S^1 \times S^2$ define a game in normal form that we denote $\langle \{\bar{S}^i\}_{i=1,2}, \{\omega_{x_0}^i(s)\}_{i=1,2} \rangle$. This game corresponds to an *average stochastic game in pure stationary strategies* that in extended form is determined by the tuple $(X, \{A^i(x)\}_{i=1,2}, \{f^i(x, a)\}_{i=1,2}, p, x_0)$.

If an arbitrary profile $s = (s^1, s^2)$ 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 $s = (s^1, s^2)$ 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 (3). Therefore if $Q^s = (q_{x,y}^s)$ is the limiting probability matrix of P^s then the average payoffs per transition $\omega_{x_0}^1(s), \omega_{x_0}^2(s)$ for the players are determined as follows

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

where

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

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

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

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

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

on $\bar{S} = \bar{S}^1 \times \bar{S}^2$ define a game in normal form $\langle \{\bar{S}^i\}_{i=1,2}, \{\psi_\theta^i(s)\}_{i=1,2} \rangle$ that in extended form is determined by $(X, \{A^i(x)\}_{i=1,2}, \{f^i(x, a)\}_{i=1,2}, p, \{\theta_x\})$. 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 .

3. Some Auxiliary Results

In this section we present some auxiliary results for the average Markov decision problem in the terms of stationary strategies and some auxiliary results related to the existence of pure-strategy Nash equilibria in n -person games.

3.1. Optimal Stationary Policies in the Average Markov Decision Problem

It is well-known that an optimal stationary policy (strategy) for the average Markov decision problem can be found by using the following linear programming model (see Puterman, 2005):

Maximize

$$\psi(\alpha, \beta) = \sum_{x \in X} \sum_{a \in A(x)} f(x, a) \alpha_{x,a} \quad (5)$$

subject to

$$\begin{cases} \sum_{a \in A(y)} \alpha_{y,a} - \sum_{x \in X} \sum_{a \in A(x)} p_{x,y}^a \alpha_{x,a} = 0, \quad \forall y \in X; \\ \sum_{a \in A(y)} \alpha_{y,a} + \sum_{a \in A(y)} \beta_{y,a} - \sum_{x \in X} \sum_{a \in A(x)} p_{x,y}^a \beta_{x,a} = \theta_y, \quad \forall y \in X; \\ \alpha_{x,a} \geq 0, \quad \beta_{y,a} \geq 0, \quad \forall x \in X, a \in A(x), \end{cases} \quad (6)$$

where θ_y for $y \in X$ represent arbitrary positive values that satisfy the condition $\sum_{y \in X} \theta_y = 1$, where θ_y for $y \in Y$ are treated as the probabilities of choosing the starting state $y \in Y$. In the case $\theta_y = 1$ for $y = x_0$ and $\theta_y = 0$ for $y \in X \setminus \{x_0\}$ we obtain the linear programming model for an average Markov decision problem with fixed starting state x_0 .

This linear programming model corresponds to a multichain case of an average Markov decision problem. If each stationary strategy in the decision problem induces an ergodic Markov chain then the restrictions (6) can be replaced by the restrictions

$$\begin{cases} \sum_{a \in A(y)} \alpha_{y,a} - \sum_{x \in X} \sum_{a \in A(x)} p_{x,y}^a \alpha_{x,a} = 0, \quad \forall y \in X; \\ \sum_{y \in X} \sum_{a \in A(y)} \alpha_{y,a} = 1; \\ \alpha_{y,a} \geq 0, \quad \forall y \in X, a \in A(y). \end{cases} \quad (7)$$

In the linear programming model (5),(6) the restrictions

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

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 linear programming model (5),(7) for the ergodic case.

The relationship between feasible solutions of problem (5),(6) and stationary strategies in the average Markov decision problem is the following: Let (α, β) be a feasible solution of the linear programming problem (5), (6) and denote $X_\alpha = \{x \in X \mid \sum_{a \in X} \alpha_{x,a} > 0\}$. Then (α, β) possesses the properties that $\sum_{a \in A(x)} \beta_{x,a} > 0$ for $x \in X \setminus X_\alpha$ and a stationary strategy $s_{x,a}$ that correspond to (α, β) is determined as

$$s_{x,a} = \begin{cases} \frac{\alpha_{x,a}}{\sum_{a \in A(x)} \alpha_{x,a}} & \text{if } x \in X_\alpha; \\ \frac{\beta_{x,a}}{\sum_{a \in A(x)} \beta_{x,a}} & \text{if } x \in X \setminus X_\alpha, \end{cases} \quad (8)$$

where $s_{x,a}$ expresses the probability of choosing the actions $a \in A(x)$ in the states $x \in X$. Thus, s can be regarded as a mapping that for every state $x \in X$ provides a probability distribution over the set of actions $A(x)$; if these probabilities take only values 0 and 1, then s corresponds to a pure stationary strategy, otherwise it corresponds to a mixed stationary strategy.

Using the linear programming problem (5),(7) Lozovanu, 2016, showed that an average Markov decision problem in terms of stationary strategies can be formulates as follows:

Maximize

$$\psi(s, q, w) = \sum_{x \in X} \sum_{a \in A(x)} f(x, a) s_{x,a} q_x \quad (9)$$

subject to

$$\begin{cases} q_y - \sum_{x \in X} \sum_{a \in A(x)} p_{x,y}^a s_{x,a} q_x = 0, & \forall y \in X; \\ q_y + w_y - \sum_{x \in X} \sum_{a \in A(x)} p_{x,y}^a s_{x,a} w_x = \theta_y, & \forall y \in X; \\ \sum_{a \in A(y)} s_{y,a} = 1, & \forall y \in X; \\ s_{x,a} \geq 0, \forall x \in X, \forall a \in A(x); \quad w_x \geq 0, \forall x \in X, \end{cases} \quad (10)$$

where θ_y are the same values as in problem (5), (6) and $s_{x,a}$, q_x , w_x for $x \in X$, $a \in A(x)$ represent the variables that must be found, where q_x for $x \in X$ express the limiting probabilities in the states for the corresponding strategy s .

The main property that we shall use for the average stochastic game is represented by the theorem that has been proven by Lozovanu, 2016.

Theorem 1. *Let an average Markov decision problem be given and consider the function*

$$\psi(s) = \sum_{x \in X} \sum_{a \in A(x)} f_{(x,a)} s_{x,a} q_x, \quad (11)$$

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

$$\begin{cases} q_y - \sum_{x \in X} \sum_{a \in A(x)} p_{x,y}^a s_{x,a} q_x = 0, & \forall y \in X; \\ q_y + w_y - \sum_{x \in X} \sum_{a \in A(x)} p_{x,y}^a s_{x,a} w_x = \theta_y, & \forall y \in X. \end{cases} \quad (12)$$

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

$$\begin{cases} \sum_{a \in A(x)} s_{x,a} = 1, & \forall x \in X; \\ s_{x,a} \geq 0, & \forall x \in X, a \in A(x) \end{cases} \quad (13)$$

the function $\psi(s)$ depends only on $s_{x,a}$ for $x \in X, a \in A(x)$ and $\psi(s)$ is quasimonotone on S (i.e. $\psi(s)$ is quasiconvex and quasiconcave on S).

Remark 1. The function (1) on \bar{S} depends only on $s_{x,a}$ for $x \in X, a \in A(x)$ because system (12) uniquely determines $q_x, \forall x \in X$ for a given $s \in S$.

3.2. Existence of Pure Nash equilibria in n -Player Games with Quasimonotonic Payoffs

Let $\langle \bar{S}_{i=\overline{1,n}}, f^i(s)_{i=\overline{1,n}} \rangle$ be an n -player game in normal form, where $\bar{S}^i, i = \overline{1,n}$ represent the corresponding sets of strategies (pure strategies) of the players $1, 2, \dots, n$ and $f^i : \prod_{j=1}^n \bar{S}^j \rightarrow R^1, i = \overline{1,n}$ represent the corresponding payoffs of these players.

Let $s = (s^1, s^2, \dots, s^n)$ be a profile of strategies of the players, $s \in \bar{S} = \prod_{j=1}^n \bar{S}^j$, and define $s^{-i} = (s^1, s^2, \dots, s^{i-1}, s^{i+1}, \dots, s^n)$, $\bar{S}^{-i} = \prod_{j=1(j \neq i)}^n \bar{S}^j$ where $s^{-i} \in \bar{S}^{-i}$.

Thus, for an arbitrary $s \in \bar{S}$ we can write $s = (s^i, s^{-i})$.

Fan, 1966 extended the well-known equilibrium result of Nash, 1951 to the games with quasiconcave payoffs. He proved the following theorem:

Theorem 2. Let $\bar{S}^i, i = \overline{1,n}$ be non-empty, convex and compact sets. If each payoff $f^i : \bar{S} \rightarrow R^1, i \in \{1, 2, \dots, n\}$, is continuous on \bar{S} and quasiconcave with respect to s^i on \bar{S}^i , then the game $\langle \bar{S}_{i=\overline{1,n}}, f^i(s)_{i=\overline{1,n}} \rangle$ possesses a pure-strategy Nash equilibrium.

Dasgupta and Maskin, 1986 considered a class of games with discontinuous payoffs and proved a pure Nash equilibria existence result for the case when the payoffs are upper semi-continuous and graph-continuous.

The payoff $f^i : \prod_{j=1}^n \bar{S}^j \rightarrow R^1$ is upper semi-continuous if for any sequence $\{s_k\} \subseteq \bar{S}$ such that $\{s_k\} \rightarrow s$ holds $\limsup_{k \rightarrow \infty} f^i(s_k) \leq f^i(s)$.

The payoff $f^i : \prod_{j=1}^n \bar{S}^j \rightarrow R^1$ is graph-continuous if for all $\bar{s} \in \bar{S}$ there exists a function $F^i : \bar{S}^{-i} \rightarrow \bar{S}^i$ with $F^i(\bar{s}^{-i})$ such that $f^i(F^i(\bar{s}^{-i}), s^{-i})$ is continuous at $s^{-i} = \bar{s}^{-i}$.

Dasgupta and Maskin proved the following theorem.

Theorem 3. Let \bar{S}^i , $i = \overline{1, n}$ be non-empty, convex and compact sets. If each payoff $f^i : \bar{S} \rightarrow R^1$, $i \in \{1, 2, \dots, n\}$, is upper semi-continuous on \bar{S} , graph-continuous and quasiconcave with respect to s^i on \bar{S}^i , then the game $\langle \bar{S}_{i=\overline{1, n}}, f^i(s)_{i=\overline{1, n}} \rangle$ possesses a pure-strategy Nash equilibrium.

In the following we need to extend this theorem for the case when each payoff $f^i(s^i, s^{-i})$, $i = 1, 2, \dots, n$ is quasimonotonic with respect to s^i on \bar{S} , i.e. $f^i(s^i, s^{-i})$ is quasiconvex and quasiconcave with respect to s^i on \bar{S}^i . We can observe that in this case the reaction correspondences of the players

$$\phi^i(s^{-i}) = \{s^i \in S^i | f^i(s^i, s^{-i}) = \max_{\bar{s}^i \in \bar{S}^i} f^i(\bar{s}^i, s^{-i})\}, \quad i = 1, 2, \dots, n$$

are compact and convex valued and therefore the upper semi-continuous condition can be released. So, in this case the theorem can be formulated as follows.

Theorem 4. Let \bar{S}^i , $i = \overline{1, n}$ be non-empty, convex and compact sets. If each payoff $f^i : \bar{S} \rightarrow R^1$, $i \in \{1, 2, \dots, n\}$, is graph-continuous and quasimonotonic with respect to s^i on \bar{S}^i , then the game $\langle \bar{S}_{i=\overline{1, n}}, f^i(s)_{i=\overline{1, n}} \rangle$ possesses a pure-strategy Nash equilibrium.

4. The Main Results

In this section we present the results concerned with the existence and determining stationary Nash equilibria for a two-player average stochastic game in stationary strategies. For this case we formulate this game in normal form.

4.1. The Game Model in Normal Form

The game model in normal form for the considered two-player average stochastic game is the following:

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

$$\begin{cases} \sum_{a^i \in A^i(x)} s_{x, a^i}^i = 1, & \forall x \in X; \\ s_{x, a^i}^i \geq 0, & \forall x \in X, a^i \in A^i(x). \end{cases} \quad (14)$$

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 (14), where $s_{x, a^i}^i \in \{0, 1\}$, $\forall x \in X$, $a^i \in A(x)$, i.e. such a solution corresponds to a pure stationary strategy of player i . On the set $\bar{S} = \bar{S}^1 \times \bar{S}^2$ we define the payoff functions

$$\psi_{\theta}^i(s^1, s^2) = \sum_{x \in X} \sum_{(a^1, a^2) \in A^1(x) \times A^2(x)} s_{x, a^1}^1 s_{x, a^2}^2 f^i(x, (a^1, a^2)) q_x, \quad i = 1, 2 \quad (15)$$

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

$$\begin{cases} q_y - \sum_{x \in X} \sum_{(a^1, a^2) \in A^1(x) \times A^2(x)} s_{x, a^1}^1 s_{x, a^2}^2 p_{x, y}^{(a^1, a^2)} q_x = 0, & \forall y \in X; \\ q_y + w_y - \sum_{x \in X} \sum_{(a^1, a^2) \in A^1(x) \times A^2(x)} s_{x, a^1}^1 s_{x, a^2}^2 p_{x, y}^{(a^1, a^2)} w_x = \theta_y, & \forall y \in X, \end{cases} \quad (16)$$

for an arbitrary fixed profile $s = (s^1, s^2) \in \bar{S}$. The functions $\psi_\theta^i(s^1, s^2)$, $i = 1, 2$ represent the payoff functions for the average stochastic game in normal form that we denote by $\langle \{\bar{S}^i\}_{i=1,2}, \{\psi_\theta^i(s)\}_{i=1,2} \rangle$. This game is determined by the tuple $(X, \{A^i(x)\}_{i=1,2}, \{f^i(x, a)\}_{i=1,2}, p, \{\theta_y\})$ where θ_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 $\langle \{\bar{S}^i\}_{i=1,2}, \{\omega_{x_0}^i(s)\}_{i=1,2} \rangle$ when the starting state x_0 is fixed, i.e. $\psi_\theta^i(s^1, s^2) = \omega_{x_0}^i(s^1, s^2)$, $i = 1, 2$. So, in this case the game is determined by $(X, \{A^i(x)\}_{i=1,2}, \{f^i(x, a)\}_{i=1,2}, p, x_0)$.

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 θ_y . In this case for the payoffs of the players in the game in normal form we have

$$\psi_\theta^i(s^1, s^2, \dots, s^n) = \sum_{y \in X} \theta_y \omega_y^i(s^1, s^2, \dots, s^n), \quad i = 1, 2. \quad (17)$$

4.2. Stationary Nash Equilibria Existence Results

As we have noted the existence of stationary Nash equilibria in two-player average stochastic games has been shown by Vielle, 2000. Here we show that this result can be derived also from the following theorem.

Theorem 5. *The game $\langle \{\bar{S}^i\}_{i=1,2}, \{\psi_\theta^i(s)\}_{i=1,2} \rangle$ possesses a pure-strategy Nash equilibrium $s^* = (s^{1*}, s^{2*})$ which is a stationary Nash equilibrium for the two-player average stochastic game determined by $(X, \{A^i(x)\}_{i=1,2}, \{f^i(x, a)\}_{i=1,2}, p, \{\theta_y\})$. Moreover, if $s^* = (s^{1*}, s^{2*})$ is a pure-strategy Nash equilibrium for the game $\langle \{\bar{S}^i\}_{i=1,2}, \{\psi_\theta^i(s)\}_{i=1,2} \rangle$, where $\theta_y > 0, \forall y \in X$ then s^* is a stationary Nash equilibrium for the two-player average stochastic game $(X, \{A^i(x)\}_{i=1,2}, \{f^i(x, a)\}_{i=1,2}, p, y)$ with an arbitrary starting state $y \in Y$.*

Proof. The proof of the existence of a pure-strategy Nash equilibrium for the game $\langle \{\bar{S}^i\}_{i=1,2}, \{\psi_\theta^i(s)\}_{i=1,2} \rangle$ follows from Theorems 4. Indeed, according to Theorem 1 the payoff $\psi_\theta^1(s^1, s^2)$ is quasimonotonic with respect to s^1 on \bar{S}^1 for a fixed $s^2 \in \bar{S}^2$ and the payoff $\psi_\theta^2(s^1, s^2)$ is quasimonotonic with respect to s^2 on \bar{S}^2 for a fixed $s^1 \in \bar{S}^1$. The graph-continuous property of payoffs functions also follows from Theorem 1 (see the proof of theorem in Lozovanu, 2016). Note that the graph-continuous property for payoffs holds only for two-players games.

Now let us prove the second part of the theorem. Let $s^* = (s^{1*}, s^{2*})$ be a pure-strategy Nash equilibrium for the game $\langle \{\bar{S}^i\}_{i=1,2}, \{\psi_\theta^i(s)\}_{i=1,2} \rangle$ determined by $(X, \{A^i(x)\}_{i=1,2}, \{f^i(x, a)\}_{i=1,2}, p, \{\theta_y\})$, where $\theta_y > 0, \forall y \in X, \sum_{y \in X} \theta_y = 1$. Then $s^* = (s^{1*}, s^{2*})$ is a Nash equilibrium for the average stochastic game $\langle \{\bar{S}^i\}_{i=1,2}, \{\psi_{\theta'}^i(s)\}_{i=1,2} \rangle$ 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

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

If here we express $\psi_{\theta'}^i$ via ω_y^i using (17) then we obtain

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

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

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

So, $s^* = (s^{1*}, s^{2*})$ is a Nash equilibrium for an arbitrary average stochastic game $\langle \{\bar{S}^i\}_{i=1,n}, \{\omega_y^i(s)\}_{i=1,n} \rangle$ with an arbitrary starting state $y \in X$.

Remark 2. The graph-continuous property of the payoffs in the case $n > 2$ players may fail to hold and therefore Theorem 5 couldn't be extended to general n -player games.

So, the problem of determining the optimal stationary strategies in a two-player average stochastic game determined by $(X, \{A^i(x)\}_{i=1,2}, \{f^i(x, a)\}_{i=1,2}, p, \{\theta_y\})$ can be found if we find the optimal strategies of the game $\langle \{\bar{S}^i\}_{i=1,2}, \{\psi_\theta^i(s)\}_{i=1,2} \rangle$, where the strategy sets \bar{S}^1, \bar{S}^2 and the payoff functions $\psi_\theta^1(s), \psi_\theta^2(s)$ are determined according to (14)-(16).

5. Conclusion

For a two-player average stochastic game a stationary Nash equilibrium exists and the optimal stationary strategies of the players can be found by determining the the optimal pure strategies for the game in normal form presented in this paper.

References

- Dasgupta, P. and Maskin, E. (1986). *The existence of equilibrium in discontinuous economic games*. Review of Economic Studies, **53**, 1–26.
- Fan, K. (1966). *Application of a theorem concerned sets with convex sections*. Math. Ann. **163** 189-203.
- Flesch, J., Thuijsman, F. and Vrieze, K. (1997). *Cyclic Markov equilibria in stochastic games*. International Journal of Game Theory, **26**, 303–314 66.
- Lozovanu, D. (2016). *Stationary Nash equilibria for average stochastic games*. Buletinul A.S.R.M, ser. Math. **2(81)** 71–92.
- Mertens, J. F. and Neyman, A. (1981) *Stochastic games*. International Journal of Game Theory, **10**, 53–66 .
- Nash, J. *Non-cooperative games*. Ann. Math. **54**, 286-293.
- Puterman, M. (2005). *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. John Wiley, New Jersey.
- Shapley, L. (1953). *Stochastic games*. Proc. Natl. Acad. Sci., U.S.A., **39**, 1095–1100.
- Solan, E. and Vieille, N. (2010). *Computing uniform optimal strategies in two-player stochastic games*. Economic Theory (special issue on equilibrium computation), **42**, 237–253.
- Tijs, S. and Vrieze, O. (1986). *On the existence of easy initial states for undiscounted stochastic games*. Math. Oper. Res., **11**, 506–513.
- Vieille, N. (2000). *Equilibrium in 2-person stochastic games I ,II*. Israel J. Math., **119(1)**, 55–126.