On the Existence and Determining Stationary Nash Equilibria for Switching Controller Stochastic Games

Dmitrii Lozovanu[1] and Stefan Pickl[2]

[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 In this paper we consider the problem of the existence and determining stationary Nash equilibria for switching controller stochastic games with discounted and average payoffs. The set of states and the set of actions in the considered games are assumed to be finite. For a switching controller stochastic game with discounted payoffs we show that all stationary equilibria can be found by using an auxiliary continuous noncooperative static game in normal form in which the payoffs are quasi-monotonic (quasi-convex and quasi-concave) with respect to the corresponding strategies of the players. Based on this we propose an approach for determining the optimal stationary strategies of the players. In the case of average payoffs for a switching controller stochastic game we also formulate an auxiliary noncooperative static game in normal form with quasi-monotonic payoffs and show that such a game possesses a Nash equilibrium if the corresponding switching controller stochastic game has a stationary Nash equilibrium.

Keywords: Stochastic game, Switching control, Stationary strategies, Stationary Nash equilibrium

1. Introduction

A switching controller stochastic game is a stochastic game in which the transition probabilities in a state are controlled only by one of the players. So, the set of states of an m-player switching controller stochastic game can be divided into m disjoint subsets where each subset represents the set of states for one of the player that governs the transition probabilities. The problem of determining stationary Nash equilibria in switching controller stochastic games recently has been studied by Bayraktar et al., 2016; Dubey et al., 2017; Krishnamurthy, N., Neogy, S.K., 2020. For a switching controller stochastic game with discounted payoffs stationary equilibria exist because stationary Nash equilibria exist for a discounted stochastic game in general. Schultz, 1992 showed that the optimal stationary strategies of the players in a discounted switching controller stochastic game can be found by solving an auxiliary linear complimentary problem. In the case of switching control stochastic games with average payoffs stationary Nash equilibria are known to exist only for some special classes of games, however, for such an arbitrary game Thuijman and Raghavan, 1997 showed the existence of ϵ- Nash equilibria. In general, the question of the existence of stationary Nash equilibria for an average stochastic game is an open problem.

In this paper we propose some new results concerned with determining stationary Nash equilibria in switching controller stochastic games with discounted and average

payoffs. We show that all stationary equilibria for a discounted switching controller stochastic game can be obtained as Nash equilibria for an auxiliary noncooperative continuous static game in normal form where the payoffs are quasi-monotonic (quasi-convex and quasi-concave) with respect to the corresponding strategies of the players. Based on this we propose an approach for determining the optimal stationary strategies of the players for a discounted switching controller stochastic game. For an average switching controller stochastic game we also formulate an auxiliary noncooperative static game with quasi-monotonic payoffs and show that such a game has a Nash equilibrium if the corresponding average switching controller stochastic game has a stationary Nash equilibrium.

2. Formulation of the Basic Game Models

An m-player switching controller stochastic game consists of the following elements:

- a finite set of states X;

- a finite set of actions $A^i(x)$ in each state $x \in X$ for an arbitrary player $i \in \{1, 2, \ldots, m\}$;

- a reward $r^i_{x,a}$ for each player $i \in \{1, 2, \ldots, m\}$ in each state $x \in X$ for an arbitrary action vector $a = (a^1, a^2 \ldots, a^m)$, where $a^i \in A^i(x), i = 1, 2 \ldots, m$;

- a partition $X = X_1 \cup X_2 \cup \cdots \cup X_m$, of the set of states X, where X_i represents the set of controllable states of player $i \in \{1, 2, \ldots, m\}$;

- a transition probability function $p^i : X_i \times \prod_{x \in X_i} A^i(x) \times X \to [0, 1]$ for each $i \in \{1, 2, \ldots, m\}$ that gives the transition probabilities $p^{a^i}_{x,y}$ of player i from an arbitrary $x \in X_i$ and each $a^i \in A^i(x)$ to an arbitrary $y \in X$ where $\sum_{y \in X} p^{a^i}_{x,y} = 1, \forall x \in X_i, \forall a^i \in A^i(x)$;

- a starting state $x_0 \in X$.

The game starts in the state x_0 at the moment of time $t = 0$ where the players simultaneously and independently fix their actions $a^i_0 \in A^i(x_0)$, $i = 1, 2, \ldots, m$. After that the players receive the corresponding rewards $r^i_{x_0,a_0}$, $i = 1, 2, \ldots, m$ in x_0 for the given action vector $a_0 = (a^1_0, a^2_0, \ldots, a^m_0)$. Then the game passes randomly to a state $x_1 \in X$ according to the probability distribution $\{p^{a^1_0}_{x_0,y}\}_{y \in X}$. At the moment of time $t = 1$ players observe the state $x_1 \in X$ and again simultaneously and independently select their actions $a^i_1 \in A^i(x_1)$, $i = 1, 2, \ldots, m$ in the state x_1 and receive the corresponding rewards $r^i_{x_1,a_1}, i = 1, 2, \ldots, m$ for the given action vector $a_1 = (a^1_1, a^2_1, \ldots, a^m_1)$. Then the game passes randomly to a state $x_2 \in X$ according to a probability distribution $\{p^{a^1_1}_{x_1,y}\}_{y \in X}$. In general, at the moment of time t the players observe the state $x_t \in X$, fix their actions $a^i_t \in A^i(x_t)$, $i = 1, 2, \ldots, m$ in x_t and receive the corresponding rewards $r^i_{x_1,a_t}$, $i = 1, 2, \ldots, m$ in x_t for the given action vector $a_t = (a^1_t, a^2_t, \ldots, a^m_t)$. Such a play of the game produces a sequence of states and actions $x_0, a_0, x_1, a_1, \ldots, x_t, a_t, \ldots$ that defines a stream of stage rewards $r^1_{x_t,a_t}, r^2_{x_t,a_t}, \ldots, r^m_{x_t,a_t}, \quad t = 0, 1, 2, \ldots$.

The average switching-controller stochastic game is the game with payoffs of the players

$$\omega_{x_0}^i = \lim_{t \to \infty} \inf \, \mathsf{E}\left(\frac{1}{t}\sum_{\tau=0}^{t-1} r_{x_\tau,a_\tau}^i\right), \quad i = 1, 2, \ldots, m$$

where E is the expectation operator with respect to the probability measure induced by the Markov process with actions chosen by the first player and given starting state x_0. Each player in this game has the aim to maximize his average reward per transition.

The discounted switching-controller stochastic game with a given discount factor γ is the game with payoffs of the players

$$\sigma_{x_0}^i = \mathsf{E}\left(\sum_{\tau=0}^{\infty} \gamma^\tau r_{x_\tau,a_\tau}^i\right).$$

Each player in this game has the aim to maximize his discounted sum of stage rewards.

We will study these games in the case when the players use stationary strategies of selection the actions in the states.

3. Switching Controller Stochastic Games in Stationary Strategies

A *strategy* $s^i(x_t)$ of player $i \in \{1, 2, \ldots, m\}$ in a switching controller 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 of 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, we can identify the set of stationary strategies S^i of player i with the set of solutions of the following system

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

in which the basic solutions corresponds to the set of pure stationary strategies.

Let $s = (s^1, s^2, \ldots, s^m) \in S = S^1 \times S^2 \times \cdots \times S^m$ be a profile of stationary strategies (pure or mixed strategies) of the players. Taking into account that the transition probabilities in each state are controlled only by one of the players, we have that the dynamics of the game is determined by the stochastic matrix $P^s = (p_{x,y}^s)$, where the elements $p_{x,y}^s$ are calculated as follows

$$p_{x,y}^s = \sum_{a^i \in A^i(x)} s_{x,a^i}^i p_{x,y}^{a^i}, \;\; \text{for } x \in X_i \text{ and } y \in X, \;\; i = 1, 2, \ldots, m. \tag{1}$$

If $Q^s = (q_{x,y}^s)$ is the limiting probability matrix of P^s then the average payoffs per transition $\omega_{x_0}^1(s), \; \omega_{x_0}^2(s), \ldots, \omega_{x_0}^m(s)$ for the players are determined as follows

$$\omega_{x_0}^i(s) = \sum_{y \in X} q_{x_0,y}^s r_{y,s}^i, \quad i = 1, 2, \ldots, m, \tag{2}$$

where

$$r^i_{y,s} = \sum_{(a^1,a^2,\ldots,a^m)\in A(y)} \prod_{k=1}^n s^k_{y,a^k} r^i_{y,(a^1,a^2,\ldots,a^m)} \tag{3}$$

expresses the average payoff (immediate reward) in the state $y \in X$ of player i when the corresponding stationary strategies $s^1, s^2, \ldots, s^m$ have been applied by players $1, 2, \ldots, m$ in y. Here $A(y) = \prod_{i=1}^m A^i(y)$.

The functions $\omega^1_{x_0}(s)$, $\omega^2_{x_0}(s)$, ..., $\omega^m_{x_0}(s)$ on $S = S^1 \times S^2 \times \cdots \times S^m$, defined according to (2), (3), determine a game in normal form that we denote by $\langle \{S^i\}_{i=\overline{1,m}}, \{\omega^i_{x_0}(s)\}_{i=\overline{1,m}} \rangle$. This game corresponds to the *average switching controller stochastic game in stationary strategies* that in extended form is determined by the tuple $(\{X_i\}_{i=\overline{1,m}}, \{A^i(x)\}_{x\in X, i=\overline{1,m}}, \{r^i_{x,a}\}_{i=\overline{1,m}}, \{p^i\}_{i=\overline{1,m}}, x_0)$.

A discounted switching controller stochastic game in stationary strategies with given discount factor γ we define as follows. Let $s = (s^1, s^2, \ldots, s^m)$ be a profile of stationary strategies (pure or mixed strategies) of the players. Then the elements of the probability transition matrix $P^s = (p^s_{x,y})$ induced by s can be calculated according to (1) and we can find the matrix $W^s = (w^s_{x,y})$, where $W^s = (I - \gamma P^s)^{-1}$. After that we can determine the payoffs for the players as follows

$$\sigma^i_{x_0}(s) = \sum_{y\in X} w_{x_0,y} r^i_{y,s}, \quad i = 1, 2, \ldots, m,$$

where $r^i_{y,s}$ is calculated according to (3). These payoffs on S define a normal form game $\langle \{S^i\}_{i=\overline{1,m}}, \{\sigma^i_{x_0}(s)\}_{i=\overline{1,m}} \rangle$ that corresponds to the *discounted switching controller stochastic game in stationary strategies*. In the extended form this game is determined by tuple $(\{X_i\}_{i=\overline{1,m}}, \{A^i(x)\}_{x\in X, i=\overline{1,m}}, \{r^i_{x,a}\}_{i=\overline{1,m}}, \{p^i\}_{i=\overline{1,m}}, \gamma, x_0)$.

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

$$\omega^i_\theta(s) = \sum_{x\in X} \theta_x \omega^i_x(s), \quad i = 1, 2, \ldots, m$$

on S define a game in normal form $\langle \{S^i\}_{i=\overline{1,m}}, \{\omega^i_\theta(s)\}_{i=\overline{1,m}} \rangle$ that in the extended form is determined by the tuple $(\{X_i\}_{i=\overline{1,m}}, \{A^i(x)\}_{x\in X, i=\overline{1,m}}, \{r^i_{x,a}\}_{i=\overline{1,m}}, p, \theta)$. In the case $\theta_x = 0, \forall x \in X \setminus \{x_0\}$, $\theta_{x_o} = 1$ the considered game becomes an average switching controller stochastic game with a fixed starting state x_0.

Similarly, for a discounted switching controller stochastic game we can consider the case when the starting state is chosen randomly according to a given distribution $\theta = \{\theta_x\}$ on X. In this case we can define the payoffs

$$\sigma^i_\theta(\mathbf{s}) = \sum_{x\in X} \theta_x \sigma^i_x(\mathbf{s}), \quad i = 1, 2, \ldots, m$$

on S and we obtain the game in normal form $\langle \{S^i\}_{i=\overline{1,m}}, \{\sigma^i_\theta(s)\}_{i=\overline{1,m}} \rangle$ that is determined by the tuple $(\{X_i\}_{i=\overline{1,m}}, \{A^i(x)\}_{x\in X, i=\overline{1,m}}, \{f^i(x,a)\}_{i=\overline{1,m}}, p, \gamma, \theta)$.

Note that an average switching controller stochastic game in the case $m = 1$ becomes an average Markov decision problem that is determined by the tuple $(X,\ \{A(x)\}_{x\in X},\ \{r_{x,a}\},\ p,\ \theta)$, where $X = X_1, A(x) = A^1(x), r_{x,a} = r_{x,a}^1$ and $a = a^1 \in A(x)$ for $x \in X$. Similarly, a discounted switching controller stochastic game in the case $m = 1$ becomes a discounted Markov decision problem that is determined by the tuple $(X,\ \{A(x)\}_{x\in X},\ \{r_{x,a}\},\ p,\gamma,\theta)$.

4. Preliminaries

In this section we present some properties of Markov decision problems with average and expected total discounted reward optimization criteria that we shall use in the following two sections for studying the switching controller stochastic games.

Let us consider a discounted Markov decision problem that is determined by the tuple $(X,\ \{A(x)\}_{x\in X},\ \{r_{x,a}\},\ p,\gamma,\theta)$. Lozovanu and Pickl, 2018 showed that this decision problem can be formulated and studied in terms of stationary strategies because the expected total discounted reward $\sigma_\theta(s)$ for a given stationary strategy s when the starting state $y \in X$ is chosen randomly according to probability distribution $\theta = \{\theta_y\}_{y\in X}$ can be represented as

$$\sigma_\theta(s) = \sum_{x\in X} \sum_{a\in A(x)} r_{x,a} s_{x,a} q_x$$

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

$$q_y - \gamma \sum_{x\in X} \sum_{a\in A(x)} s_{x,a} q_x = \theta_y, \forall y \in X.$$

Furthermore, the following theorem has been proven.

Theorem 1. *Let a discounted Markov decision problem be given and consider the function*

$$\sigma_\theta(s) = \sum_{x\in X} \sum_{a\in A(x)} r_{x,a} s_{x,a} q_x$$

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

$$q_y - \gamma \sum_{x\in X} \sum_{a\in A(x)} s_{x,a} q_x = \theta_y, \forall y \in X.$$

Then on the set S of the solutions of the system

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

the function $\sigma_\theta(s)$ depends only on $s_{x,a}$ for $x \in X$, $a \in A(x)$ and $\sigma_\theta(s)$ is quasi-monotonic on S, i.e. $\sigma_\theta(s)$ is quasi-convex and quasi-concave on S (see Boyd and Vandenberghe, 2004). Moreover $\sigma_\theta(s)$ is continuous on S.

Thus, a discounted Markov decision problem in stationary strategies can be represented as a problem of maximization of quasi-monotonic function $\sigma_\theta(s)$ on a polyhedron set S.

We can present some similar results for an average Markov decision problem.

Let us consider an average Markov decision problem that is determined by the tuple $(X, \{A(x)\}_{x \in X}, \{r_{x,a}\}, p, \theta)$. Lozovanu, D., 2018 showed that this decision problem can be formulated and studied in terms of stationary strategies. In this case the expected average reward $\omega_\theta(s)$ for a given stationary strategy s when the starting state $y \in X$ is chosen randomly according to probability distribution $\{\theta_{y \in X}\}$ can be represented as

$$\omega_\theta(s) = \sum_{x \in X} \sum_{a \in A(x)} r_{x,a} s_{x,a} q_x$$

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

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

For the considered problem the following theorem has been proven.

Theorem 2. *Let an average Markov decision problem be given and consider the function*

$$\omega_\theta(s) = \sum_{x \in X} \sum_{a \in A(x)} r_{x,a} s_{x,a} q_x$$

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

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

Then on the set S of the solutions of the system

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

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

So, an average Markov decision problem in stationary strategies can be represented as a problem of maximization of a quasi-monotonic function $\omega_\theta(s)$ on a polyhedron set S. However, here, the function $\omega_\theta(s)$ on S may be discontinuous, unless the fact that it is well defined on S.

The switching controller stochastic game models in stationary strategies with discounted and average payoffs that we present in the next two sections contain the game variants of the corresponding Markov decision problems in stationary strategies formulated above.

5.   Determining Stationary Nash Equilibria for a Discounted Switching Controller Stochastic Game

The existence of stationary Nash equilibria for discounted stochastic games with finite state and action spaces has been proven by Fink, 1964 and Takahashi, 1964. Based on a constructive proof of this result suitable algorithms for determining stationary Nash equilibria for a discounted stochastic game has been elaborated (see Kallenberg, 2016). Lozovanu and Pickl, 2018 proposed the following approach for determining stationary equilibria in a discounted stochastic game with finite state and action spaces.

Let $(X, \{A^i(x)\}_{i=\overline{1,m}}, \{r^i_{x,a}\}_{i=\overline{1,m}}, p, \gamma, \theta)$ be the tuple that determines a discounted stochastic game where X is the finite set of states; $A^i(x)$ is the set of actions of player $i \in \{1, 2, \dots, m\}$ in the state $x \in X$; $r^i_{x,a}$ is the reward of player $i \in \{1, 2, \dots, m\}$ in the state x for an action vector $a = (a^1, a^2, \dots, a^m)$ in x; $p^i : X \times \prod_{x \in X} \prod_{i=1}^m A^i(x) \times X \to [0, 1]$ is the probability function that gives transition probabilities $p^a_{x,y}$ from an arbitrary $x \in X$ to an arbitrary $y \in X$ for each action vector $a = (a^1, a^2, \dots, a^m)$ in x; γ is a given discount factor and $\theta = \{\theta_y\}_{y \in X}$ is a distribution function on X where θ_y expresses the probability that the game starts in y. Lozovanu and Pickl, 2018 showed that all stationary equilibria for such a game represents Nash equilibria of the following noncooperative static game in normal form $\langle \{S^i\}_{i=\overline{1,m}}, \{\sigma^i_\theta(s)\}_{i=\overline{1,m}} \rangle$, where each set of strategies S^i, $i \in \{1, 2, \dots, m\}$ represents the set of solutions of the system

$$
\begin{cases}
\sum\limits_{a^i \in A^i(x)} s^i_{x,a^i} = 1, & \forall x \in X; \\
\qquad s^i_{x,a^i} \geq 0, & \forall x \in X, \ \ \forall a \in A^i(x)
\end{cases}
\tag{4}
$$

and each payoff $\sigma^i_{x_0}(s^1, s^2, \dots, s^m))$ is defined as follows

$$
\begin{cases}
\sigma^i_\theta(s^1, s^2, \ \dots, \ s^m) = \sum\limits_{x \in X} \ \sum\limits_{(a^1, a^2, \dots, a^m) \in A(x)} \prod\limits_{k=1}^m s^k_{x,a^k} r^i_{x,a^1,a^2 \dots a^m} q_x, \\
\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad i = 1, 2, \dots, m,
\end{cases}
\tag{5}
$$

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

$$
q_y - \gamma \sum_{x \in X} \ \sum_{(a^1, a^2, \dots, a^m) \in A(x)} \prod_{k=1}^m s^k_{x,a^k} p^{(a^1, a^2, \dots, a^m)}_{x,y} q_x = \theta_y, \qquad \forall y \in X;
\tag{6}
$$

for an arbitrary $s = (s^1, s^2, \dots, s^m) \in S = S^1 \times S^2 \times \cdots \times S^m$, where S^i represents the set of solutions of system (4).

Each payoff $\sigma^i_{x_0}(s^1, s^2, \dots, s^m))$ in the considered auxiliary noncooperative game possesses the properties that it is continuous on S and quasi-monotonic with respect to the strategy s^i on S^i. Therefore, based on results of Dasgupta and Maskin, 1986 the game $\langle \{S^i\}_{i=\overline{1,m}}, \{\sigma^i_\theta(s)\}_{i=\overline{1,m}} \rangle$ has a Nash equilibrium and this equilibrium is a stationary Nash equilibrium of the discounted stochastic game.

Taking into account that a discounted switching controller stochastic game represent a special case of a discounted stochastic game then we can specify the auxiliary

game $\langle \{S^i\}_{i=\overline{1,m}}, \{\sigma_\theta^i(s)\}_{i=\overline{1,m}} \rangle$ for a special case. So, if in (6) we take into account that for each set X_i the transition probabilities in the states $x \in X_i$ are controlled only by player i then the payoff $\sigma_\theta^i(s)$ of player $i \in \{1, 2, \ldots, m\}$ on S is defined as follows

$$\begin{cases} \sigma_\theta^i(s^1, s^2, \ldots, s^m) = \sum_{x \in X} \sum_{(a^1, a^2, \ldots, a^m) \in A(x)} \prod_{k=1}^m s_{x,a^k}^k \cdot r_{x,(a^1,a^2\ldots a^m)}^i \cdot q_x, \\ \\ \hspace{6cm} i = 1, 2, \ldots, m, \end{cases} \quad (7)$$

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

$$q_y - \gamma \sum_{k=1}^m \sum_{x \in X_k} \sum_{a^k \in A^k(x)} s_{x,a^k}^k \cdot p_{x,y}^{a^k} \cdot q_x = \theta_y, \qquad \forall y \in X; \qquad (8)$$

for an arbitrary fixed $s = (s^1, s^2, \ldots, s^m) \in S = S^1 \times S^2 \times \cdots \times S^m$.

Each payoff function $\sigma_\theta^i(s^1, s^2, \ldots, s^m)$ on S is continuous. Moreover, Lozovanu and Pickl, 2018 showed that each payoff $\sigma_\theta^i(s^1, s^2, \ldots, s^m)$ on S is quasi-monotonic with respect to strategy s^i on S^i. Therefore, based on the results of Fan, 1966 and Dasgupta and Maskin, 1986 the following theorem holds.

Theorem 3. *The game $\langle \{S^i\}_{i=\overline{1,m}}, \{\sigma_\theta^i(s)\}_{i=\overline{1,m}} \rangle$ has a Nash equilibrium $s^* = (s^{1*}, s^{2*}, \ldots, s^{m*}) \in S = S^1 \times S^2 \times \cdots \times S^m$ that is a stationary Nash equilibrium of the discounted switching controlled stochastic game determined by the tuple $(\{X_i\}_{i=\overline{1,m}}, \{A^i(x)\}_{i=\overline{1,m}}, \{r_{x,a}^i\}_{i=\overline{1,m}}, p, \gamma, x_0)$. Moreover this equilibrium is a stationary Nash equilibrium of the discounted switching controller stochastic game with an arbitrary starting state $x \in X$.*

So, a stationary Nash equilibrium for a discounted switching controller stochastic game determined by the tuple $(\{X_i\}_{i=\overline{1,m}}, \{A^i(x)\}_{i=\overline{1,m}}, \{r_{x,a}^i\}i = \overline{1,m}, p, \gamma, \theta)$ can be found by computing a Nash equilibrium of the noncooperative static game $\langle \{S^i\}_{i=\overline{1,m}}, \{\sigma_\theta^i(s)\}_{i=\overline{1,m}} \rangle$, where the payoffs are determined according to (7),(8).

6. On Determining Stationary Nash Equilibria for an Average Switching Controller Stochastic Game

In this section we show that if an average switching controller stochastic game has a stationary Nash equilibrium then all such equilibria can be found from an auxiliary noncooperative static game in normal form $\langle \{S^i\}_{i=\overline{1,m}}, \{\omega_\theta^i(s)\}_{i=\overline{1,m}} \rangle$ where each payoffs $\omega_\theta^i(s^1, s^2, \ldots, s^m)$ are quasi-monotonic with respect to the corresponding strategies of the players. To prove this we shall use the conditions obtained by Lozovanu and Pickl, 2020 concerned with the existence of stationary Nash equilibria for an average stochastic game with finite state and action spaces.

Let $(X, \{A^i(x)\}_{i=\overline{1,m}}, \{r_{x,a}^i\}_{i=\overline{1,m}}, p, \theta)$ be the tuple that determines an average stochastic game where X is the finite set of states; $A^i(x)$ is the set of actions of player $i \in \{1, 2, \ldots, m\}$ in the state x; $r_{x,a}^i$ is the reward of player $i \in \{1, 2, \ldots, m\}$ in the state x for an action vector $a = (a^1, a^2, \ldots, a^m)$ in x; $p : X \times \prod_{x \in X} \prod_{i=1}^m A^i(x) \times X \to [0, 1]$ is the probability function that gives transition probabilities $p_{x,y}^a$ from an arbitrary $x \in X$ to an arbitrary $y \in X$ for each

action vector $a = (a^1, a^2, \ldots, a^m)$ in x and $\theta = \{\theta_y\}_{y \in X}$ is a distribution function on X where θ_y expresses the probability that the average stochastic game starts in y.

Lozovanu and Pickl, 2020 showed that if the average stochastic game determined by the tuple $(X, \{A^i(x)\}_{i=\overline{1,m}}, \{r^i_{x,a}\}_{i=\overline{1,m}}, p, \theta)$ has stationary Nash equilibria then for such a game an auxiliary noncooperative game $\langle \{S^i\}_{i=\overline{1,m}}, \{\omega^i_\theta(s)\}_{i=\overline{1,m}} \rangle$ with quasi-monotonic payoffs $\omega^i_\theta(s)$, $i = 1, 2, \ldots, m$ on $S = S^1 \times S^2 \times, \ldots, \times S^m$ for which Nash equilibria represents stationary Nash equilibria of the average stochastic game can be constructed. The set of strategies S^i of player $i \in \{1, 2, \ldots, m\}$ represents the set of solutions of the system

$$
\begin{cases}
\sum_{a^i \in A^i(x)} s^i_{x,a^i} = 1, & \forall x \in X; \\
s^i_{x,a^i} \geq 0, & \forall x \in X, \ \forall a^i \in A^i(x)
\end{cases}
\tag{9}
$$

and each payoff $\omega^i_{x_0}(s^1, s^2, \ldots, s^m))$ is defined as follows

$$
\begin{cases}
\omega^i_\theta(s^1, s^2, \ldots, s^m) = \sum_{x \in X} \sum_{(a^1, a^2, \ldots, a^m) \in A(x)} \prod_{k=1}^{m} s^k_{x,a^k} \cdot r^i_{x,(a^1, a^2 \ldots a^m)} \cdot q_x, \\
\hspace{8cm} i = 1, 2, \ldots, m,
\end{cases}
\tag{10}
$$

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

$$
\begin{cases}
q_y - \sum_{x \in X} \sum_{(a^1, a^2, \ldots, a^m) \in A(x)} \prod_{k=1}^{m} s^k_{x,a^k} \cdot p^{(a^1, a^2, \ldots, a^m)}_{x,y} \cdot q_x = 0, & \forall y \in X; \\
q_y + w_y - \sum_{x \in X} \sum_{(a^1, a^2, \ldots, a^m) \in A(x)} \prod_{k=1}^{m} s^k_{x,a^k} \cdot p^{(a^1, a^2, \ldots, a^m)}_{x,y} \cdot w_x = \theta_y, & \forall y \in X,
\end{cases}
\tag{11}
$$

for a fixed strategy $s = (s^1, s^2, \ldots, s^m) \in S$. In general, an average stochastic game may have no stationary Nash equilibria (see Flesch et al., 1997) and then the game $\langle \{S^i\}_{i=\overline{1,m}}, \{\omega^i_\theta(s)\}_{i=\overline{1,m}} \rangle$ has no Nash equilibrium. However if the game $\langle \{S^i\}_{i=\overline{1,m}}, \{\omega^i_\theta(s)\}_{i=\overline{1,m}} \rangle$ has a Nash equilibrium $s^* = (s^{1*}, s^{2*}, \ldots, s^{m*})$ then this equilibrium is a stationary Nash equilibrium of the average switching stochastic game for an arbitrary starting state $y \in X$ (see Lozovanu and Pickl, 2020). In the unichain case of the average stochastic game, when the matrix P^s is unichain for an arbitrary $s \in S$, system (11) can be replaced by the following system

$$
\begin{cases}
q_y - \sum_{x \in X} \sum_{(a^1, a^2, \ldots, a^m) \in A(x)} \prod_{k=1}^{m} s^k_{x,a^k} \cdot p^{(a^1, a^2, \ldots, a^m)}_{x,y} \cdot q_x = 0, & \forall y \in X; \\
\hspace{6cm} \sum_{x \in X} q_x = 1
\end{cases}
\tag{12}
$$

In this case the game $\langle \{S^i\}_{i=\overline{1,m}}, \{\omega^i_\theta(s)\}_{i=\overline{1,m}} \rangle$ has a Nash equilibrium $s^* = (s^{1*}, s^{2*}, \ldots, s^{m*})$ that is a stationary Nash equilibrium of the average stochastic game for an arbitrary starting state $y \in X$.

We can specify for an average switching controller game the auxiliary game model $\langle \{S^i\}_{i=\overline{1,m}}, \{\omega^i_\theta(s)\}_{i=\overline{1,m}} \rangle$ for the average stochastic game . If we specify the

auxiliary noncooperative game $\langle \{S^i\}_{i=\overline{1,m}}, \{\omega_\theta^i(s)\}_{i=\overline{1,m}} \rangle$ for an average switching controller stochastic game determined by the tuple $(\{X_i\}_{i=\overline{1,m}}, \{A^i(x)\}_{x\in X, i=\overline{1,m}}, \{r_{x,a}^i\}_{i=\overline{1,m}}, p, \theta)$ then we obtain the game where the sets of strategies S^i, $i = 1, 2, \ldots, m$ represent the corresponding sets of solutions of the systems

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

and the payoffs

$$\begin{cases} \omega_\theta^i(s^1, s^2, \ldots, s^m) = \displaystyle\sum_{x \in X} \sum_{(a^1, a^2, \ldots, a^m) \in A(x)} \prod_{k=1}^m s_{x,a^k}^k \cdot r_{x,(a^1, a^2 \ldots a^m)}^i \cdot q_x, \\ \hspace{8cm} i = 1, 2, \ldots, m, \end{cases} \tag{14}$$

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

$$\begin{cases} q_y - \displaystyle\sum_{k=1}^m \sum_{x \in X_k} \sum_{a^k \in A^k(x)} s_{x,a^k}^k \cdot p_{x,y}^{a^k} \cdot q_x = 0, & \forall y \in X; \\ q_y + w_y - \displaystyle\sum_{k=1}^m \sum_{x \in X_k} \sum_{a^k \in A^k(x)} s_{x,a^k}^k \cdot p_{x,y}^{a^k} \cdot w_x = \theta_y, & \forall y \in X, \end{cases} \tag{15}$$

for a fixed $s = (s^1, s^2, \ldots, s^m) \in S$.

Based on the results bove we can formulate the following theorem.

Theorem 4. *Let be given an average switching controller stochastic game determined by the tuple $(\{X_i\}_{i=\overline{1,m}}, \{A^i(x)\}_{x\in X, i=\overline{1,m}}, \{r_{x,a}^i\}_{i=\overline{1,m}}, \{p^i\}_{i=\overline{1,m}}, x_0)$ and let $\langle \{S^i\}_{i=\overline{1,m}}, \{\omega_\theta^i(s)\}_{i=\overline{1,m}} \rangle$ be the auxiliary noncooperative static game with quasi-monotonic payoffs $\omega_\theta^i(s)$, $i = 1, 2, \ldots, m$ for the average switching controller stochastic game. Then the game $\langle \{S^i\}_{i=\overline{1,m}}, \{\omega_\theta^i(s)\}_{i=\overline{1,m}} \rangle$ has a Nash equilibrium if and only if the average switching controller game has a stationary Nash equilibrium.*

So, if an average switching controller stochastic game has stationary Nash equilibria then such equilibria can be found as Nash equilibria of the auxiliary noncooperative game $\langle \{S^i\}_{i=\overline{1,m}}, \{\omega_\theta^i(s)\}_{i=\overline{1,m}} \rangle$, where S^i and $\omega_\theta^i(s)$, $i = 1, 2, \ldots, m$ are defined according to (13)-(15). In general, the problem of determining a Nash equilibrium for the game $\langle \{S^i\}_{i=\overline{1,m}}, \{\omega_\theta^i(s)\}_{i=\overline{1,m}} \rangle$ may be a difficult problem because the payoffs in the considered auxiliary game may be discontinuous with respect to all strategies of the players. If the paypffs of the auxiliary game are continuous or graph-continuous then according to Dasgupta and Maskin, 1986 we obtain that the average switching controller stochastic game possesses a stationary Nash equilibrium. An important class of average switching controller stochastic games for which stationary Nash equilibria exist represents the average stochastic positional games (see Lozovanu, D., 2018) and single controller stochastic games (see Rosenberg et al.,2004). For the unichain games the payoffs of the auxiliary game are continuous and therefore in this case the average switching controller stochastic

game possesses a stationary Nash equilibrtum. In this case the system (15) can be replaced by the following system

$$
\begin{cases}
q_y - \sum\limits_{k=1}^{m} \sum\limits_{x \in X_k} \sum\limits_{a^k \in A^k(x)} s_{x,a^k}^k \cdot p_{x,y}^{a^k} \cdot q_x = 0, & \forall y \in X; \\
\sum\limits_{y \in X} q_y = 1
\end{cases}
$$

and we obtain a more simple game model.

7. Conclusion

A discounted switching controller stochastic game always possesses stationary Nash equilibria and all such equilibria can be found as Nash equilibria of the auxiliary noncooperative game with quasi-monotonic payoffs from Section 5. This auxiliary noncooperative game always has a Nash equilibrium because the payoffs are quasi-monotonic with respect to the corresponding strategies of the players and continuous with respect to all strategies of the players. For an average switching controller stochastic game also can be considered an auxiliary noncooperative static game with quasi-monotonic payoffs as it is shown in Section 6. However in this case the payoffs may be not continuous with respect to all strategies of the players and therefore the problem of determining a Nash equilibrium for the auxiliary static game may be a difficult problem. If the payoffs in the auxiliary game are continuous or graph-continuous in the sense of Dasgupta and Maskin, 1986 then the average switching controller stochastic game has a stationary Nash equilibrium. In the unichain case an average switching controller stochastic game always has a stationary Nash equilibrium because the payoffs in the auxiliary static game are continuous with respect to all strategies of the players. In general, the existence and determining stationary Nash equilibria for an average switching controller stochastic game is a difficult open problem (see Thuijman and Raghavan, 1997, Flesch et al., 1997). Nevertheless for an average switching controller stochastic game we can state that the proposed auxiliary noncooperative static game from Section 6 has a Nash equilibrium if and only if the average switching controller stochastic game has a stationary Nash equilibrium. An important class of average switching controller stochastic games for which stationary Nash equilibria exist represents the average stochastic positional games and single controller stochastic games. So, if we determine a Nash equilibrium for the auxiliary static game then we determine a stationary Nash equilibrium for the average switching controller stochastic game.

References

Boyd, S., Vandenberghe, L. (2004). *Convex optimization.* Cambridge University Press.

Bayraktar, E., Cosso, A., Pham, H. (2016). *Robost feedback switching control: Dynamic programming and viscosity solutions.* SIAM Journal on Control and Optimization, **54**, 5, 2594–2628.

Dasgupta, P. and Maskin, E. (1986). *The existence of an equilibrium in discontinuous economic games.* The Review of Economic Studies, **53**, 1–26.

Dubey, D., Neogy, S.,G.K., Ghorui, D. (2017) *Completely mixed strategies for generalzed bimatrix and switching controller stochastic games.* Dynamic games and Applications, **7**, 4, 535–554.

Fan, K. (1966). *Application of a theorem concerned sets with convex sections.* Math. Ann.,**1963**, 189–203.

Fink, A. (1964). *Equilibrium in a Stochastic n-Person Game.* Journal of Science of the Hiroshima University, **28**, 89–93.

Flesch, J., Thuijsman, F., Vrieze, K. (1997). Cyclic Markov equilibria in stochastic games. International Journal of Game Theory, **26**, 303–314.

Kallenberg L. (2016). *Markov Decision Processes.* University of Leiden.

Krishnamurthy, N., Neogy, S. K. (2020) *On Lemke processibility of LCP formulations for solving discounted swiching control stochastic games.* Annals of Operations Research, **295**, 2, 633–644.

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:Fondation and Aplication"(L.Petrossyan et al.eds.), Springer, 139–163.

Lozovanu, D., Pickl, S. (2018). *Pure Stationary Nash Equilibria for Discounted Stochastic Positional Games.* Contributions to game theory and management, **12**, 246–260.

Lozovanu, D., Pickl, S. (2020). *On the existence of stationary Nash equilibrium in average stochastic games with finite state and action spaces.* Contributions to game theory and management, **13**, 304–323.

Rosenberg, D., Solan, E., Vieille, N. (2004). *Stochastic games with a single controller and incomplete information.* SIAM Journal on Control and Optimization, **43**, 1, 86–110.

Schultz, T. (1992). *Linear complimentary and discounted switching-controller stochastic games.* Journal of Optimization Theory and Applications , **73**, 89–99.

Takahashi, M. (1964). *Equilibrium Points of Stochastic Non-Cooperative n-Person Games.* Journal of Science of the Hiroshima University, **28**, 95–99.

Thuijsman, F., Raghavan, T.E.S. (1997). *Perfect information stochastic games and related classes.* International Journal of Game Theory, **26**, 403–408.