N–PERSON STOCHASTIC GAMES: EXTENSIONS OF THE FINITE STATE SPACE CASE AND CORRELATION

ANDRZEJ S. NOWAK
Zielona Góra University
Zielona Góra, Poland

1. Introduction

In this chapter, we present a framework for m-person stochastic games with an infinite state space. Our main purpose is to present a correlated equilibrium theorem proved by Nowak and Raghavan [42] for discounted stochastic games with a measurable state space, where the correlation of the different players' strategies employs only "public signals" [16]. We will also provide a detailed survey of the literature containing related results, some approximation theorems for general state space stochastic games (the existence of ε-equilibria), and the existence of equilibria in some classes of countable state space stochastic games with applications to queueing models.

We consider *m-person non-zero-sum stochastic games* for which:
(i) S is a nonempty Borel state space.
(ii) X_i is a nonempty compact metric *space of actions* for player i. We put $X = X_1 \times X_2 \times \cdots \times X_m$.
(iii) $A_i(s)$ is a nonempty compact subset of X_i and represents the *set of actions available* to player i in state s. We assume that $\{(s,a) : s \in S \text{ and } a \in A_i(s)\}$ is a Borel subset of $S \times X_i$. Put

$$A(s) = A_1(s) \times A_2(s) \times \cdots \times A_m(s), \quad s \in S.$$

(iv) $r_i : S \times X \mapsto \mathbb{R}$ is a bounded Borel-measurable *payoff function* for player i. It is assumed that $r_i(s, \cdot)$ is continuous on X, for every $s \in S$.
(v) p is a Borel-measurable transition probability from $S \times X$ to S, called the *law of motion* among states. If s is a state at some stage of the game and the players select an $x \in A(s)$, then $p(\cdot|s,x)$ is the probability distribution of the next state of the game. We assume that the transition probability $p(\cdot|s,x)$ has a density function, say $z(\cdot|s,x)$, with respect to a fixed probability

measure μ on S, satisfying the following L_1 continuity condition:
For any sequence of joint action tuples $x^n \to x^0$,

$$\int_S \mid z(t|s, x^n) - z(t|s, x^0) \mid \mu(dt) \to 0 \quad \text{as } n \to \infty.$$

The L_1 continuity above is satisfied via Scheffe's theorem when $z(s|t, \cdot)$ is continuous on X. It implies the norm continuity of the transition probability $p(\cdot \mid s, x)$ with respect to $x \in X$ (see Theorem 16.11 in [7]).

The game is played in discrete time with past history as common knowledge for all players. A *strategy* for a player is a Borel-measurable mapping which associates with each given history a probability distribution on the set of actions available to him. A *stationary strategy* for player i is a Borel-measurable mapping which associates with each state $s \in S$ a probability distribution on the set $A_i(s)$ of actions available to him at s, independent of the history that led to the state s. A stationary strategy for player i can thus be identified with a Borel-measurable transition probability f from S to X_i such that $f(A_i(s) \mid s) = 1$, for every $s \in S$.

Let $H^\infty = S \times X \times S \times \cdots$ be the space of all infinite histories of the game, endowed with the product σ-algebra. For any profile of strategies $\pi = (\pi_1, \ldots, \pi_m)$ of the players and every initial state $s_1 = s \in S$, a probability measure P_s^π and a stochastic process $\{s_n, a_n\}$ are defined on H^∞ in a canonical way, where the random variables s_n and a_n describe the state and the actions chosen by the players, respectively, on the n-th stage of the game (see Chapter 7 in [5]). Thus, for each profile of strategies $\pi = (\pi_1, \ldots, \pi_m)$, any finite horizon T, and every initial state $s \in S$, the *expected T-stage payoff* to player i is

$$\gamma_i^T(\pi)(s) = E_s^\pi \left(\sum_{n=1}^{T} r_i(s_n, a_n) \right).$$

Here E_s^π means the expectation operator with respect to the probability measure P_s^π. If $\beta \in (0,1)$ is a fixed *discount factor*, then the *β-discounted expected payoff* to player i is

$$\gamma_i^\beta(\pi)(s) = E_s^\pi \left(\sum_{n=1}^{\infty} \beta^{n-1} r_i(s_n, a_n) \right).$$

The *expected average payoff (per unit time)* for player i is defined as

$$\gamma_i(\pi)(s) = \liminf_{T \to \infty} \frac{1}{T} \gamma_i^T(\pi)(s).$$

Let $\pi^* = (\pi_1^*, \ldots, \pi_m^*)$ be a fixed profile of strategies of the players. For any strategy π_i of player i, we write (π_{-i}^*, π_i) to denote the strategy profile obtained from π^* by replacing π_i^* with π_i.

A strategy profile $\pi^* = (\pi_1^*, \ldots, \pi_m^*)$ is called a *Nash equilibrium* for the β-discounted stochastic game iff no unilateral deviations from it are profitable, that is, for each $s \in S$,

$$\gamma_i^\beta(\pi^*)(s) \geq \gamma_i^\beta(\pi_{-i}^*, \pi_i)(s),$$

for every player i and any strategy π_i. Of course, Nash equilibria are analogously defined for the finite horizon and limiting average payoff stochastic games.

Remark 1 The question of the existence of stationary Nash equilibria in non-zero-sum stochastic games remains open. Only some special classes of games are known to possess a stationary Nash equilibrium. For example, Parthasarathy and Sinha [44] proved the existence of stationary Nash equilibria in discounted stochastic games with finitely many actions for the players and state-independent nonatomic transition probabilities. Their result was extended by Nowak [35] to a class of uniformly ergodic average payoff games. Also, some classes of discounted non-zero-sum stochastic games with perfect information or additive payoff and reward structure (ARAT games) are known to possess stationary equilibria [25], [23]. It is shown in [26] that ARAT games also have a nonstationary nonrandomized equilibrium. There are papers on certain economic (resource extraction, capital accumulation or consumption and investment) games where a stationary equilibrium is shown to exist by exploiting a very special transition and payoff structure (for example, [4], [11], [14]). The transition probabilities in [4] and [14] are assumed to be weakly continuous and the state space is a segment of real line. Very recently, stationary Nash equilibria were shown to exist in a pretty large class of discounted stochastic games (including some stochastic games of capital accumulation) in which the transition probability is a combination of finitely many measures with coefficients depending on states and actions of the players [40]. Mertens and Parthasarathy [28], [29], [30] showed the existence of nonstationary subgame-perfect Nash equilibria in a class of discounted stochastic games assuming that the transition probabilities are norm-continuous with respect to the actions. A simplification of their proof is given in [48]. For finite-horizon games a Markov equilibrium can be obtained by a simple algorithm based on backward induction [46].

2. Correlated Equilibria

In this section, we extend the sets of strategies available to the players in the sense that we allow them to correlate their choices in a natural way described below. The resulting solution is a kind of extensive-form correlated equilibrium [16].

Suppose that $\{\xi_n : n \geq 1\}$ is a sequence of so-called *signals*, drawn independently from $[0,1]$ according to the uniform distribution. Suppose that at the beginning of each period n of the game the players are informed not only of the outcome of the preceding period and the current state s_n, but also of ξ_n. Then the information available to them is a vector $h^n = (s_1, \xi_1, x_1, \ldots, s_{n-1}, \xi_{n-1}, x_{n-1}, s_n, \xi_n)$, where $s_i \in S$, $x_i \in A(s_i)$, $\xi_i \in [0,1]$. We denote the set of such vectors by H^n.

An *extended strategy* for player i is a sequence $\pi_i = (\pi_i^1, \pi_i^2, \ldots)$, where every π_i^n is a Borel-measurable transition probability from H^n to X_i such that $\pi_i^n(A_i(s_n) \mid h^n) = 1$ for each history $h^n \in H^n$. An *extended stationary strategy* for player i is a strategy $\pi_i = (\pi_i^1, \pi_i^2, \ldots)$ such that each π_i^n depends on the current state s_n and the last signal ξ_n only. In other words, a strategy π_i of player i is called stationary if there exists a transition probability f from $S \times [0,1]$ to X_i such that for every period n of the game and each history $h^n \in H^n$, we have $\pi_i^n(\cdot \mid h^n) = f(\cdot \mid s_n, \xi_n)$. Assuming that the players use extended strategies we actually assume that they play the stochastic game with the extended state space $S \times [0,1]$. The law of motion, say $\bar{p}$, in the extended state space model is obviously the product of the original law of motion p and the uniform distribution η on $[0,1]$. More precisely, for any $s \in S$, $\xi \in [0,1]$, $a \in A(s)$, any Borel sets $C \subset S$ and $D \subset [0,1]$, $\bar{p}(C \times D \mid s, \xi, a) = p(C \mid s, a)\eta(D)$.

For any profile of extended strategies $\pi = (\pi_1, \ldots, \pi_m)$ of the players, the β-discounted (undiscounted) payoff to player i is a function of the initial state s_1 and the first signal ξ_1 and is denoted by $\gamma_i^\beta(\pi)(s_1, \xi_1)$ $[\gamma_i(\pi)(s_1, \xi_1)]$.

We say that $f^* = (f_1^*, \ldots, f_m^*)$ is a *Nash equilibrium in the average payoff stochastic game in the class of extended strategies* if for each initial state $s_1 \in S$,

$$\int_0^1 \gamma_k^\beta(f^*)(s_1, \xi_1)\eta(d\xi_1) \geq \int_0^1 \gamma_k^\beta(f_{-k}^*, \pi_k)(s_1, \xi_1)\eta(d\xi_1),$$

for every player k and any extended strategy π_k.

A Nash equilibrium in extended strategies is also called a *correlated equilibrium with public signals*. The reason is that after the outcome of any period of the game, the players can coordinate their next choices by exploiting the next (known to all of them, i.e., public) signal and using some coordination mechanism telling which (pure or mixed) action is to be played by each of them. In many applications, we are particularly interested in stationary equilibria. In such a case the coordination mechanism can be represented by a family of $m+1$ measurable functions $\lambda^1, \ldots, \lambda^{m+1} : S \mapsto [0,1]$ such that $\sum_{i=1}^{m+1} \lambda^i(s) = 1$ for every $s \in S$. A stationary Nash equilibrium in the class of extended strategies can be constructed then by using a family of $m+1$ stationary strategies $f_i^1, \dots, f_i^{m+1}$, given for each player i, and the following coordination rule. If the game is at a state s on the n-th stage and a random number ξ_n is selected, then each player i is suggested to use $f_i^k(\cdot \mid s)$, where k is the least index for which $\sum_{j=1}^k \lambda^j(s) \geq \xi_n$. The λ^j and f_i^j fixed above induce an extended stationary strategy f_i^* for each player i as follows

$$f_i^*(\cdot \mid s, \xi) = f_i^1(\cdot \mid s) \quad \text{if} \quad \xi \leq \lambda^1(s), \quad s \in S,$$

and

$$f_i^*(\cdot \mid s, \xi) = f_i^k(\cdot \mid s) \quad \text{if} \quad \sum_{j=1}^{k-1} \lambda^j(s) < \xi \leq \sum_{j=1}^{k} \lambda^j(s),$$

for $s \in S$, $2 \leq k \leq m+1$. Because the signals are independent and uniformly distributed in $[0,1]$, it follows that at any period of the game and for any current state s, the number $\lambda^j(s)$ can be interpreted as the probability that player i is suggested to use $f_i^j(\cdot \mid s)$ as his mixed action. Now it is quite obvious that a strategy profile $(f_1^*, \dots, f_m^*)$ obtained by the above construction is a stationary Nash equilibrium in the class of extended strategies of the players in a game iff no player i can unilaterally improve upon his expected payoff by changing any of his strategies f_i^j, $j = 1, \dots, m+1$.

The following result was proved by Nowak and Raghavan [42] by a fixed-point argument.

Theorem 1 *Every non-zero-sum discounted stochastic game satisfying (i) through (v) has a stationary correlated equilibrium with public signals.*

Remark 2 A related result to Theorem 1 is reported in Duffie et al. [12]. They used some stronger assumptions about the primitive data of the game (for example the transition probability), but showed that there exists a stationary correlated equilibrium which induces an ergodic process. Nonstationary correlated subgame-perfect equilibria in a class of dynamic games with weakly continuous transition probabilities were studied by Harris et al. [18]. [34] studied an in some sense weaker correlated scheme.

3. Auxiliary Results and the Proof

Let S and Y be nonempty Borel spaces. We assume that B is a Borel subset of $S \times Y$ whose projection on the horizontal axis is S. For each $s \in S$, put $B(s) := \{y \in Y : (s, y) \in B\}$. The set-valued mapping $s \mapsto B(s)$ is said to be *lower measurable* if for every open set G in Y the set $\{s \in S : B(s) \cap G \neq \emptyset\}$ is Borel.

From Theorem 3 in [22], we obtain the following auxiliary result.

Lemma 1 *Assume that $B(s)$ is compact for each $s \in S$. Then B is a Borel set if and only if $s \mapsto B(s)$ is lower measurable.*

It is well known that every Borel set B for which the sets $B(s)$ are compact has a Borel-measurable uniformization (selection) [9]. However, to make use of many helpful results described in the literature in terms of set-valued mappings, we shall apply the characterization given in Lemma 1. From Kuratowski and Ryll-Nardzewski [27], we immediately obtain:

Lemma 2 *The lower measurable set-valued mapping $s \mapsto B(s)$ has a Borel-measurable selector, that is, there exists a Borel-measurable function $f : S \mapsto Y$ such that $f(s) \in B(s)$ for every $s \in S$.*

Castaing gave the following useful characterization of lower measurable set-valued mappings (see [10] or [20] for the proof).

Lemma 3 *The compact set-valued mapping $s \mapsto B(s)$ defined above is lower measurable if and only if there exists a countable family of Borel-measurable functions $f_n : S \mapsto Y$ such that B(s) = closure of $\{f_n(s)\}$ in Y for every $s \in S$.*

From Lemma 1, we infer the following corollary.

Lemma 4 *Let D be a lower measurable mapping from S into nonempty compact subsets of Y. Let $u : S \times Y \mapsto \mathbb{R}$ be a Borel-measurable real-valued function such that, for every $s \in S$, $u(s, \cdot)$ is continuous on D(s). Define F by*

$$F(s) = \{x \in D(s) : u(s, x) = 0\}$$

and assume that $F(s)$ is nonempty for every $s \in S$. Then F is a lower measurable set-valued mapping.

As a corollary to Theorem 7.1 in [20], we obtain:

Lemma 5 *Assume that F is a lower measurable set-valued mapping from S into nonempty compact subsets of Y, T is a metric space, $u : S \times Y \mapsto T$ is a mapping such that, $u(s, \cdot)$ is continuous on Y for every $s \in S$, $u(\cdot, y)$ is measurable for every $y \in Y$. Suppose there is a measurable mapping $g : S \mapsto T$ such that $g(s) \in \{u(s, y) : y \in F(s)\}$ for every $s \in S$. Then there exists a measurable selector f of F such that*

$$g(s) = u(s, f(s)) \quad \textit{for every} \quad s \in S.$$

Let V be the space of all μ-equivalence classes of Borel-measurable functions $v : S \mapsto \mathbb{R}$ such that $|v(s)| \leq C$ μ-a.e., where C is a fixed constant such that $|r_k(s, x)| \leq C$ for each $s \in S$, $x \in A(s)$ and for every player k. It is obvious that V is a compact and metrizable subset of $L_\infty = L_\infty(S, \mu)$, when endowed with the weak-star topology $\sigma(L_\infty, L_1)$

[13]. Let $V^m = V \times V \times \cdots \times V$ (m times). We endow V^m with the product topology, so V^m is a compact metrizable space too. Obviously, V^m is convex. We shall define a set-valued mapping on V^m into its compact convex subsets and show that this mapping has a fixed point corresponding to a correlated equilibrium payoff vector for the stochastic game. To do this, we associate with each $v = (v_1, \ldots, v_m) \in V^m$ and $s \in S$ the non-zero-sum game $\Gamma_v(s)$ where the payoff for player k corresponding to any strategy m-tuple $x = (x_1, \ldots, x_m) \in A(s)$ is

$$u_k(s,x)(v) = (1-\beta)r_k(s,x) + \beta \int_S v_k(t)p(dt|s,x).$$

Under our continuity assumptions, the set of Nash equilibria in $\Gamma_v(s)$, denoted by $N_v(s)$, is a nonempty compact subset of $\Delta(X)$, the space of probability measures on the product space $X = X_1 \times X_2 \times \cdots X_m$, equipped with the weak topology.

Lemma 6 *For any $v \in V^m$, $s \mapsto N_v(s)$ is a lower measurable compact set-valued mapping.*

Proof. For any probability measures μ_k on X_k $(k = 1, \ldots, m)$ put

$$u_k(s, \mu_1, \ldots, \mu_m)(v) = \int \cdots \int u_k(s, x_1, \ldots, x_m)(v)\mu_1(dx_1) \times \cdots \times \mu_m(dx_m),$$

and for any probability measures $p_k \in \Delta(A_k(s))$ and $v \in V^m$ set

$$\begin{aligned} &u(s, p_1, \ldots, p_m)(v) \\ &\quad = u_1(s, p_1, \ldots, p_m)(v) - \max_{\mu_1 \in \Delta(A_1(s))} u_1(s, \mu_1, p_2, \ldots, p_m)(v) + \cdots \\ &\quad\; + u_m(s, p_1, \ldots, p_m)(v) - \max_{\mu_m \in \Delta(A_m(s))} u_m(s, p_1, \ldots, p_{m-1}, \mu_m)(v). \end{aligned}$$

Note that, for any $s \in S$, $v \in V^m$, we have

$$N_s(v) = \{(p_1, \ldots, p_m) : \quad u(s, p_1, \ldots, p_m)(v) = 0\}.$$

The set-valued mapping $s \mapsto \Delta(A_k(s))$ is lower measurable [21], and so is $s \mapsto \Delta(A_1(s)) \times \cdots \times \Delta(A_m(s))$ [20]. By our continuity assumptions, the function $u(s, \ldots)(v)$ satisfies the conditions of Lemma 4, and thus the result follows. ∎

Let $P_v(s)$ be the set of payoff vectors in $\Gamma_v(s)$ corresponding to all Nash equilibria from the set $N_v(s)$. By $coP_v(s)$, we denote the convex hull of $P_v(s)$.

Lemma 7 *Both the set-valued mappings $s \mapsto P_v(s)$ and $s \mapsto coP_v(s)$ are lower measurable.*

Proof. To prove that $s \mapsto P_v(s)$ is lower measurable one can use Castaing's characterization of the lower measurable mapping $s \mapsto N_v(s)$ stated in Lemma 3. The lower measurability of $s \mapsto coP_v(s)$ follows from Theorem 9.1 in [20]. ∎

By Lemmas 2 and 7, the compact set-valued mapping $s \mapsto coP_v(s)$ has a measurable selector. Let M_v be the set of all μ-equivalence classes of Borel-measurable selectors of $s \mapsto coP_v(s)$, $v \in V^m$ is fixed.

Lemma 8 *The mapping $v \mapsto M_v$ is convex compact-valued and uppersemicontinuous.*

Proof. Obviously, M_v is convex for every $v \in V^m$. Assume that $v^n \to v$ in V^m as $n \to \infty$, $w^n \in M_{v^n}$ for each n, and $w^n \to w$. We have to prove that $w \in M_v$. First we note that, for any $s \in S$, the game $\Gamma_{v^n}(s)$ converges to the game $\Gamma_v(s)$ in the sense that, for every player k,

$$l_k^n(s) := \max_{x \in A(s)} |u_k(s,x)(v_k^n) - u_k(s,x)(v_k)| \to 0 \quad \text{as} \quad n \to \infty.$$

Here v_k^n (v_k) is the k-th component of v^n (v). Clearly,

$$l_k^n(s) = \max_{x \in A(s)} \left| \beta \int_S v_k^n(t) p(dt|s,x) - \beta \int_S v_k(t) p(dt|s,x) \right|.$$

We shall prove that $l_k^n(s) \to 0$ as $n \to \infty$. Suppose that the convergence to zero does not take place. Then there exist a positive number α and an infinite set J of positive integers such that $l_k^n(s) > \alpha$ for all $n \in J$. For each $n \in J$, let x^n be a point in $A(s)$ at which the above maximum is attained. We can assume without loss of generality that $x^n \to x^0 \in A(s)$ as $n \to \infty$. We obviously have

$$\begin{aligned} l_k^n(s) \le \beta & \left| \int_S (v_k^n(t) - v_k(t)) p(dt|s,x^0) \right| \\ & + \beta \left| \int_S v_k^n(t) p(dt|s,x^n) - \int_S v_k^n(t) p(dt|s,x^0) \right| \\ & + \beta \left| \int_S v_k(t) p(dt|s,x^n) - \int_S v_k(t) p(dt|s,x^0) \right|. \end{aligned}$$

The first term on the right-hand side of this inequality tends to zero as $n \to \infty$, because $p(\cdot|s,x^0) \ll \mu$ and $v_k^n \to v_k$ as $n \to \infty$ in the $\sigma(L_\infty, L_1)$ topology. The second and also the third term are each less than or equal to $C\|p(\cdot|s,x^n) - p(\cdot|s,x^0)\|$, ($\|\cdot\|$ is the total variation norm) which tends to zero as $n \to \infty$. This contradicts our assumption that $l_k^n(s) > \alpha > 0$ for all $n \in J$. Thus the convergence of $l_k^n(s)$ to zero is proved.

We assume that $w^n \to w$ in the product weak-star topology on V^m. Then some sequence $\{\hat{w}^n\}$ of convex combinations of these functions converges to w almost surely, say $\hat{w}^n(s) \to w(s)$ for every $s \in S - S_1$, where $\mu(S_1) = 0$ (see Mazur's theorem in [13]). For any $\epsilon > 0$, let $coP_v^\epsilon(s)$ be the ϵ-neighborhood of $coP_v(s)$. Using the fact that $l_k^n(s) \to 0$ as $n \to \infty$, for every player k and $s \in S$, it is easy to show that $w(s) \in coP_v^\epsilon(s)$ for every $s \in S - S_1$ and $\epsilon > 0$. Hence $w(s) \in coP_v(s)$, $s \in S - S_1$. This proves that $v \mapsto M_v$ is uppersemicontinuous. Similarly, we conclude that M_v is closed for any v and because $M_v \subset V^m$ and V^m is compact for every $v \in V^m$, M_v is compact as well. ∎

Proof of Equilibrium Theorem. From the Kakutani–Glicksberg fixed-point theorem [17] and Lemma 8 we infer that there exists a Borel-measurable mapping $v : S \mapsto \mathbb{R}^m$ such that $v(s) \in coP_v(s)$ for all $s \in S - S_1$, where $\mu(S_1) = 0$. By the random version of Caratheodory's theorem (see [10]) there exist Borel-measurable functions $\lambda^1, \dots, \lambda^{m+1} : S \mapsto [0,1]$ such that $\sum_{i=1}^{m+1} \lambda^i(s) = 1$ and there exist Borel-measurable mappings $u^1, \dots, u^{m+1} : S \mapsto \mathbb{R}^m$ such that for all states $s \in S - S_2$ where $\mu(S_2) = 0$, we have $u^i(s) \in P_v(s)$ and

$$v(s) = \sum_{i=1}^{m+1} \lambda^i(s) u^i(s) \quad \text{for all} \quad s \in S - (S_1 \cup S_2).$$

By Lemma 5 there exist Borel-measurable mappings $f^i : S \mapsto \Delta(X_1) \times \cdots \times \Delta(X_m)$ such that $f^i(s) \in N_v(s)$, and

$$u_k^i(s) = (1-\beta) r_k(s, f^i(s)) + \beta \int_S v_k(t) p(dt|s, f^i(s)), \quad i = 1, \dots, m+1,$$

for every player k and for each $s \in S - (S_1 \cup S_2)$. Here u_k^i is the k-th component of u^i. Hence, if we put $f_\lambda = \sum_{i=1}^{m+1} \lambda^i f^i$, we get

$$v_k(s) = (1-\beta) r_k(s, f_\lambda(s)) + \beta \int_S v_k(t) p(dt|s, f_\lambda(s)), \quad s \in S - (S_1 \cup S_2).$$

Let $w = v/(1-\beta)$. Then $f_\lambda(s) \in coN_w(s)$ and consequently

$$w_k(s) = r_k(s, f_\lambda(s)) + \beta \int_S w_k(t) p(dt|s, f_\lambda(s)), \quad s \in S - (S_1 \cup S_2).$$

Let f_0 be any Borel-measurable selector of $s \mapsto N_w(s)$. We put

$$w_k^*(s) = \begin{cases} w_k(s) & \text{if} \quad s \in S - (S_1 \cup S_2), \\ r_k(s, f_0(s)) + \beta \int_S w_k(t) p(dt|s, f_0(s)) & \text{if} \quad s \in S_1 \cup S_2. \end{cases}$$

Observe that if we set

$$f_\lambda^*(s) = \begin{cases} f_\lambda(s) & \text{for} \quad s \in S - (S_1 \cup S_2) \\ f_0(s) & \text{for} \quad s \in S_1 \cup S_2, \end{cases}$$

then we get (because $p \ll \mu$, $\mu(S_1 \cup S_2) = 0$)

$$w_k^*(s) = r_k(s, f_\lambda^*(s)) + \beta \int_S w_k^*(t) p(dt|s, f_\lambda^*(s))$$

and $f_\lambda^*(s)$ belongs to the convex hull of $N_{w^*}(s)$ for every $s \in S$. Using standard results on discounted dynamic programming [5], [22], we infer that

$$w_k^*(s) = \gamma_k^\beta(f_\lambda^*)(s) \quad \text{for every } s \in S,$$

and since $f_\lambda^*(s) \in coN_{w^*}(s)$, f_λ^* is a correlated equilibrium for our game. ∎

We now present an extension of Theorem 1 to undiscounted stochastic games obtained in Nowak [35]. We need some additional assumptions on the transition probability p. For any stationary strategy profile f and $n \geq 1$, let $p^n(\cdot \mid s, f)$ denote the n-step transition probability determined by p and f.

GE (*Uniform Geometric Ergodicity*): There exist scalars $\alpha \in (0, 1)$ and $\delta > 0$ for which the following holds: for any profile f of stationary strategies of the players, there exists a probability measure p_f on S such that

$$\|p^n(\cdot \mid s, f) - p_f(\cdot)\| \leq \delta\alpha^n \qquad \text{for each } n \geq 1.$$

Here $\|\cdot\|$ denotes the total variation norm in the space of finite signed measures on S.

Now the main result of Nowak [35] can be formulated.

Theorem 2 *Every non-zero-sum average payoff stochastic game satisfying (i) through (v) and* **GE** *has a stationary correlated equilibrium with public signals.*

The basic idea of the proof of Theorem 2 is rather simple. Let C be any positive number such that $|r_k| \leq C$ for every player k. Then, for every discount factor β, and any stationary correlated equilibrium f_λ^β obtained in Theorem 1, $(1-\beta)\gamma_k(f_\lambda^\beta)(\cdot)$ is in the compact ball $B(C)$ with radius C in the space $L^\infty(S, \mu)$, endowed with the weak-star topology $\sigma(L^\infty, L_1)$. Therefore, it is possible to find a sequence $\{\beta_n\}$ of discount factors which converges to one and $(1-\beta_n)\gamma_k(f_\lambda^{\beta_n})$ converges to some function $J_k \in B(C)$. Using **GE**, it is shown that J_k are constant equilibrium functions of the players, and $f_\lambda^{\beta_n}$ converges (in some sense) to a stationary correlated equilibrium for the undiscounted game.

Remark 3 Condition **GE** is rather difficult to check. However, some nice characterizations of this property can be given by so-called *drift inequality* [31], [32] with a bounded solution. A much weaker form of **GE** called V–geometric ergodicity (allowing for unbounded solutions to the drift inequality) was applied in [37], [41] for studying zero-sum stochastic games and some approximation problems. For a further discussion of generalized versions of the **GE** assumption with applications to decision processes see [19], [24]. In [39], one can find a version of Theorem 2 given for some class of semi-Markov games where the time between successive jumps from one state to another is random.

Remark 4 Nonzero-sum discounted stochastic games with *countable state spaces* (and bounded payoff functions) are known to possess a stationary equilibrium [15], [43]. Federgruen [15] extended this result to average payoff non-zero-sum stochastic games with countably many states, satisfying condition **GE**. His result [15] was considerably generalized by Altman, Hordijk and Spieksma [3] who assumed a μ-recurrence condition implying a weaker version of the geometric ergodicity property of Markov chains induced by stationary strategies of players. Other extensions of Federgruen's result to a variety of games satisfying different stochastic stability or recurrence assumptions are contained in [8], [36], [45], [47]. The daily payoff functions (often called cost functions) are assumed in these papers to be unbounded, which is motivated by interesting applications of countable state space stochastic games to *queueing theory*; see also [1], [2], [3], [45], [49].

Remark 5 The problem of approximating a stochastic game with a measurable state space by countable state space games is natural and was first studied by Whitt [50]. He considered separable state space discounted game models with uniformly continuous payoff and transition probability functions. By a suitable approximation he established the existence of ε-equilibria in the class of stationary strategies. Extensions of Whitt's result to some classes of uniformly continuous and ergodic undiscounted stochastic games are given in [6], [38]. The existence of stationary ϵ-equilibrium strategies in discounted stochastic games satisfying assumptions accepted in this chapter (where no continuity properties of the payoff and transition functions with respect to the state variable are imposed) was proved in [33]. The approximation technique used in [33] is completely different from Whitt's approach [50]. Extensions of Nowak's result [33] to some classes of non-zero-sum stochastic games with unbounded cost functions were recently given by Altman and Nowak [41]. The expected average cost criterion is also considered in [41] but under certain stochastic stability conditions widely discussed in the theory of Markov control processes and queueing networks [31], [32], [19].

Acknowledgements. The author thanks A. Neyman and S. Sorin for their help in the preparation of this chapter.

References

1. Altman, E. (1996) Non-zero-sum stochastic games in admission, service and routing control in queueing systems, *Queueing Systems, Theory and Applications* **23**, 259–279.
2. Altman, E. and Hordijk, A. (1995) Zero-sum Markov games and worst-case optimal control of queueing systems, *Queueing Systems, Theory and Applications* **21**, 415–447.
3. Altman, E., Hordijk, A. and Spieksma, F.M. (1997) Contraction conditions for average and α-discount optimality in countable state Markov games with unbounded rewards, *Mathematics of Operations Research* **22**, 588–618.
4. Amir, R. (1996) Continuous stochastic games of capital accumulation with convex transitions, *Games and Economic Behavior* **15**, 111–131.
5. Bertsekas, D.P. and Shreve, S.E. (1978) *Stochastic Optimal Control: The Discrete Time Case*, Academic Press, New York.
6. Bielecki, T.B. (1996) Approximations of dynamic Nash games with general state and action spaces and ergodic costs for the players, *Applicationes Mathematicae* **24**, 195–202.
7. Billingsley, P. (1979) *Probability and Measure*, John Wiley and Sons, New York.
8. Borkar, V.S. and Ghosh, M.K. (1993) Denumerable state stochastic games with limiting average payoff, *Journal of Optimization Theory and Applications* **76**, 539–560.
9. Brown, L.D. and Purves, R. (1973) Measurable selections of extrema, *Annals of Statistics* **1**, 902–912.
10. Castaing C. and Valadier, M. (1977) *Convex Analysis and Measurable Multifunctions*, Lecture Notes in Mathematics 580, Springer-Verlag, New York.
11. Curtat, L.O. (1996) Markov equilibria of stochastic games with complementarities, *Games and Economic Behavior* **17**, 177–199.
12. Duffie, D., Geanakoplos, J., Mas-Colell, A. and McLennan, A. (1994) Stationary Markov equilibria, *Econometrica* **62**, 745–782.
13. Dunford, N. and Schwartz, J.T. (1958) *Linear Operators I: General Theory*, Wiley Interscience Publishers, New York.
14. Dutta, P.K. and Sundaram, R.K. (1992) Markovian equilibrium in a class of stochastic games: Existence theorems for discounted and undiscounted models, *Economic Theory* **2**, 197–214.
15. Federgruen, A. (1978) On n-person stochastic games with denumerable state space, *Advances in Applied Probability* **10**, 452–471.
16. Forges, F. (1986) An approach to communication equilibria, *Econometrica* **54**, 1375–1385.
17. Glicksberg, I.E. (1952) A further generalization of the Kakutani fixed-point theorem with application to Nash equilibrium points, *Proceedings of the American Mathematical Society* **3**, 170–174.
18. Harris, C., Reny, P. and Robson, A. (1995) The existence of subgame-perfect equilibrium in continuous games with almost perfect information: A case for public randomization, *Econometrica* **63**, 507–544.
19. Hernández-Lerma, O. and Lasserre, J.B. (1999) *Further Topics on Discrete-Time Markov Control Processes*, Springer-Verlag, New York.

20. Himmelberg, C.J. (1975) Measurable relations, *Fundamenta Mathematicae* **87**, 53–72.
21. Himmelberg, C.J. and Van Vleck, F.S. (1975) Multifunctions with values in a space of probability measures, *Journal of Mathematical Analysis and Applications* **50**, 108–112.
22. Himmelberg, C.J., Parthasarathy, T. and Van Vleck, F.S. (1976) Optimal plans for dynamic programming problems, *Mathematics of Operations Research* **1**, 390–394.
23. Himmelberg, C.J., Parthasarathy, T., Raghavan, T.E.S. and Van Vleck, F.S. (1976) Existence of p-equilibrium and optimal stationary strategies in stochastic games, *Proceedings of the American Mathematical Society* **60**, 245–251.
24. Jaśkiewicz, A. (2001) An approximation approach to ergodic semi-Markov control processes, *Mathematical Methods of Operations Research* **54**, 1–19.
25. Küenle, H.-U. (1999) Stochastic games with complete information and average cost criterion, *Annals of the International Society of Dynamic Games* **5**, 325–338.
26. Küenle, H.-U. (1999) Equilibrium strategies in stochastic games with additive cost and transition structure, *International Game Theory Review* **1**, 131–147.
27. Kuratowski, K. and Ryll-Nardzewski, C. (1965) A general theorem on selectors, *Bulletin of the Polish Academy of Sciences* (*Ser. Mathematics*) **13**, 379–403.
28. Mertens, J.-F. and Parthasarathy, T. (1987) Equilibria for discounted stochastic games, CORE Disscusion Paper 8750, Université Catholique de Louvain, Louvain-la-Neuve, Belgium (Chapter 10 in this volume).
29. Mertens J.-F. and Parthasarathy, T. (1991) Nonzero-sum stochastic games, in T.E.S. Raghavan et al. (eds.), *Stochastic Games and Related Topics, Essays in Honor of L.S. Shapley*, Kluwer Academic Publishers, Dordrecht, pp. 145–148.
30. Mertens J.-F., Sorin, S. and Zamir, S. (1994) Repeated games, CORE Disscusion Papers 9420, 9421, 9422, Université Catholique de Louvain, Louvain-la-Neuve, Belgium.
31. Meyn, S.P. and Tweedie, R.L. (1993) *Markov Chains and Stochastic Stability*, Springer-Verlag, New York.
32. Meyn, S.P. and Tweedie, R.L. (1994) Computable bounds for geometric convergence rates of Markov chains, *Annals of Applied Probability* **4**, 981–1011.
33. Nowak, A.S. (1985) Existence of equilibrium stationary strategies in discounted noncooperative stochastic games with uncountable state space, *Journal of Optimization Theory and Applications* **45**, 591–602.
34. Nowak, A.S. (1991) Existence of correlated weak equilibria in discounted stochastic games with general state space, in T.E.S. Raghavan et al. (eds.), *Stochastic Games and Related Topics, Essays in Honor of L.S. Shapley*, Kluwer Academic Publishers, Dordrecht, pp. 135–143.
35. Nowak, A.S. (1994) Stationary equilibria for non-zero-sum average payoff ergodic stochastic games with general state space, *Annals of the International Society of Dynamic Games* **1**, 231–246.
36. Nowak, A.S. (1999) Sensitive equilibria for ergodic stochastic games with countable state spaces, *Mathematical Methods of Operations Research* **50**, 65–76.
37. Nowak, A.S. (1999) Optimal strategies in a class of zero-sum ergodic stochastic games, *Mathematical Methods of Operations Research* **50**, 399–419.
38. Nowak, A.S. (1999) On approximations of nonzero–sum uniformly continuous ergodic stochastic games, *Applicationes Mathematicae* **26**, 221–228.
39. Nowak, A.S. (2000) Some remarks on equilibria in semi-Markov games, *Applicationes Mathematicae* **27**, 385–394.
40. Nowak, A.S. (2002) On a new class of non-zero-sum discounted stochastic games having stationary Nash equilibrium points, *International Journal of Game Theory*, to appear.
41. Nowak, A.S. and Altman, E. (2001) On ε-equilibria for stochastic games with uncountable state space and unbounded costs, *SIAM Journal on Control and*

Optimization **40**, 1821–1839.
42. Nowak, A.S. and Raghavan, T.E.S. (1992) Existence of stationary correlated equilibria with symmetric information for discounted stochastic games, *Mathematics of Operations Research* **17**, 519–526.
43. Parthasarathy, T. (1973) Discounted, positive and noncooperative stochastic games, *International Journal of Game Theory* **2**, 25–37.
44. Parthasarathy, T. and Sinha, S. (1989) Existence of stationary equilibrium strategies in non-zero-sum discounted games with uncountable state space and state-independent transition, *International Journal of Game Theory* **18**, 189-194.
45. Passchier, O. (1996) The theory of Markov games and queueing control, Ph.D. thesis, Department of Mathematics and Computer Science, Leiden University, Leiden, The Netherlands.
46. Rieder, U. (1979) Equilibrium plans for non-zero-sum Markov games, in O. Moeschlin and D. Pallaschke (eds.), *Game Theory and Related Topics*, North-Holland, Amsterdam, pp. 91–102.
47. Sennott, L.I. (1994) Nonzero-sum stochastic games with unbounded costs: Discounted and average cost cases, *Zeitschrift für Operations Research* **40**, 145–162.
48. Solan, E. (1998) Discounted stochastic games, *Mathematics of Operations Research* **23**, 1010–1021.
49. Spieksma, F. M. (1990) Geometrically ergodic Markov chains and the optimal control of queues, Ph.D. thesis, Department of Mathematics and Computer Science, Leiden University, Leiden, The Netherlands.
50. Whitt, W. (1980) Representation and approximation of noncooperative sequential games, *SIAM Journal on Control and Optimization* **18**, 33–48.