SIAM J. CONTROL OPTIM.
Vol. 40, No. 6, pp. 1821–1839

ε-EQUILIBRIA FOR STOCHASTIC GAMES WITH UNCOUNTABLE STATE SPACE AND UNBOUNDED COSTS*

ANDRZEJ S. NOWAK† AND EITAN ALTMAN‡

Abstract. We study a class of noncooperative stochastic games with unbounded cost functions and an uncountable state space. It is assumed that the transition law is absolutely continuous with respect to some probability measure on the state space. Undiscounted stochastic games with expected average costs are considered first. It is shown under a uniform geometric ergodicity assumption that there exists a stationary ε-equilibrium for each $\varepsilon > 0$. The proof is based on recent results on uniform bounds for convergence rates of Markov chains [S. P. Meyn and R. L. Tweedie, *Ann. Appl. Probab.*, 4 (1994), pp. 981–1011] and on an approximation method similar to that used in [A. S. Nowak, *J. Optim. Theory Appl.*, 45 (1985), pp. 591–602], where an ε-equilibrium in stationary policies was shown to exist for the bounded discounted costs. The stochastic game is approximated by one with a countable state space for which a stationary Nash equilibrium exists (see [E. Altman, A. Hordijk, and F. M. Spieksma, *Math. Oper. Res.*, 22 (1997), pp. 588–618]); this equilibrium determines an ϵ-equilibrium for the original game. Finally, new results for the existence of stationary ε-equilibrium for discounted stochastic games are given.

Key words. nonzero-sum stochastic games, approximate equilibria, general state space, long run average payoff criterion

AMS subject classifications. Primary, 90D10, 90D20; Secondary, 90D05, 93E05

PII. S0363012900378371

1. Introduction.

This paper treats nonzero-sum stochastic games with general state space and unbounded cost functions. Our motivation for studying unbounded costs comes from applications of stochastic games to queuing theory and telecommunication networks (see [2, 3, 4, 38]). We assume that the transition law is absolutely continuous with respect to some probability measure on the state space. For the expected average cost case, we impose some stochastic stability conditions, considered often in the theory of Markov chains in general state space [25, 26]. These assumptions imply the so-called ν-geometric ergodicity condition for Markov chains governed by stationary multipolicies of players. Using an approximation technique similar to that in [29], we prove the existence of stationary ε-equilibria in m-person average cost games satisfying the mentioned stability conditions and some standard regularity assumptions. A similar result is stated for discounted stochastic games, but then we do not impose any ergodicity assumptions. To obtain an ε-equilibrium, we apply a recent result by Altman, Hordijk, and Spieksma [4] given for nonzero-sum stochastic games with countably many states. Completely different approximation schemes for stochastic games with a separable metric state space were given by Rieder [39] and Whitt [48]. As in [29], they considered only (bounded) discounted stochastic games.

The passage from finite (or even countably infinite) state space with possibly unbounded cost turns out to be quite a tough problem. In fact, the question of the existence of stationary Nash equilibria in nonzero-sum stochastic games with uncountable state space remains open even in the discounted case. Only some special

*Received by the editors September 18, 2000; accepted for publication (in revised form) August 18, 2001; published electronically March 5, 2002.

http://www.siam.org/journals/sicon/40-6/37837.html

†Institute of Mathematics, Zielona Góra University, Podgórna 50, 65-246 Zielona Góra, Poland (nowak@im.pwr.wroc.pl).

‡INRIA, 2004 Route des Lucioles, B.P.93, 06902 Sophia-Antipolis Cedex, France (altman@sophia.inria.fr).

classes of games are known to possess a stationary Nash equilibrium. For example, Parthasarathy and Sinha [37] 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 [30] to a class of uniformly ergodic average cost games. There are papers on certain economic games in which a stationary equilibrium is shown to exist by exploiting a very special transition and payoff structure; see, for example, [5, 7]. Mertens and Parthasarathy [24] reported the existence of nonstationary subgame-perfect Nash equilibria in a class of discounted stochastic games with norm continuous transition probabilities. Some results for nonzero-sum stochastic games with additive reward and transition structure (and, in particular, games with complete information) are given by Küenle [19, 20]. Finally, Harris, Reny, and Robson [13] proved the existence of correlated subgame-perfect equilibria in a class of stochastic games with weakly continuous transition probabilities. We would like to point out that the only papers which deal with nonzero-sum average cost stochastic games with uncountable state space are [20] and [30]. In the zero-sum case, the theory of stochastic games with uncountable state spaces is much more complete. Mertens and Neyman [23] provided some conditions for the existence of value, and Maitra and Sudderth [21, 22] developed a general theory of zero-sum stochastic games with limsup payoffs. Stationary optimal strategies exist in the average cost zero-sum games only if some ergodicity conditions are imposed in the model; see, for example, [31, 34, 15, 17, 18].

In this paper, we make use of an extension of Federgruen's work [11] given by Altman, Hordijk, and Spieksma [4]. Other approaches (based on different assumptions) to nonzero-sum stochastic games with countably many states can be found in [6, 40] and [33]. Some results on sensitive equilibria in a class of ergodic stochastic games are discussed in [33, 16, 35]. To close this brief overview of the existing literature, we note that the theory of stochastic games is much more complete in the case of finite state and action spaces. On one hand, many deep existence theorems are available at the moment; see [23, 22, 44, 45] and some references therein. On the other hand, a theory of algorithms for solving special classes of stochastic games with finitely many states and actions is also well developed [12].

In order to study the uncountable state space, we make use of Lyapunov-type techniques [25] (which also allows us to treat unbounded costs) and of approximations based on discretization. Unfortunately, the discretization to a countable state space does not directly yield a setting for which we can apply the existing theory for stochastic games with a countable state [4]. For example, the Foster (or Lyapunov)-type conditions that have been used for countable Markov chains always involved the requirement of a negative drift outside a finite set, whereas our discretization provides a negative drift outside a countable set. Also, ensuring that the approximating game maintains the same type of ergodic structure as the initial game turned out to be a highly complex problem. The fact that our model allows us to handle unbounded costs is very useful in stochastic games occurring in queueing and in networking applications; see [2, 3, 4, 38], in which bounded costs turn out to be unnatural.

The involved process of discretization given in our paper, which requires assumptions that may be restrictive in some applications, may suggest that, when possible, other equilibrium concepts might be sought instead of the Nash equilibrium. Indeed, some results on the existence of stationary *correlated* equilibria are available at the moment [36, 30, 10]. This type of equilibrium allows for some coordination between players, and the proof of existence is considerably simpler.

This paper is organized as follows. In section 2, we describe our game model. Section 3 is devoted to studying the average cost games. In section 4, we examine discounted stochastic games. An appendix is given in section 5, which contains some auxiliary results on piecewise constant policies in controlled Markov chains.

2. The model and notation. Before presenting the model, we collect some basic definitions and notation. Let $(\Omega, \mathcal{F})$ be a measurable space, where $\mathcal{F}$ is the σ-field of subsets in Ω. By $\mathbb{P}(\Omega)$ we denote the space of all probability measures on $(\Omega, \mathcal{F})$. If Ω is a metric space, then $\mathcal{F}$ is assumed to be the Borel σ-field in Ω. Let $(S, \mathcal{G})$ be another measurable space. We write $P(\cdot|\omega)$ to denote a transition probability from Ω into S. Recall that $P(\cdot|\omega) \in \mathbb{P}(S)$ for each $\omega \in \Omega$, and $P(D|\cdot)$ is a measurable function for each $D \in \mathcal{G}$.

We now describe the game model:

(i) S—the *state space*, endowed with a *countably generated* σ-field $\mathcal{G}$.

(ii) X^i—a *compact metric action space* for player i, $i = 1, 2, \ldots, m$. Let $X = X^1 \times X^2 \times \cdots \times X^m$. We assume that X is given the Borel σ-field.

(iii) $c^i : S \times X \to \mathbb{R}$—a product measurable *cost (payoff) function* for player i.

(iv) $Q(\cdot|s, x)$—a (product measurable) transition probability from $S \times X$ into S, called the *law of motion among states*.

We assume that actions are chosen by the players at discrete times $k = 1, 2, \ldots$. At each time k, the players observe the current state s_k and choose their actions independently of one another. In other words, they select a vector $x_k = (x_k^1, \ldots, x_k^m)$ of actions, which results in a cost $c^i(s_k, x_k)$ at time k incurred by player i, and in a transition to a new state, whose distribution is given by $Q(\cdot|s_k, x_k)$. Let $H_1 = S$ and let $H_n = S \times X \times S \times X \times \cdots \times S$ ($2n - 1$ factors) be the space of all n-stage histories of the game, endowed with the product σ-field. A randomized *policy* γ^i for player i is a sequence $\gamma^i = (\gamma_1^i, \gamma_2^i, \ldots)$, where each γ_n^i is a (product measurable) transition probability $\gamma_n^i(\cdot|h_n)$ from H_n into X^i. The *class of all policies* for player i will be denoted by Γ^i. Let U^i be the set of all transition probabilities u^i from S into X^i. A *Markov policy* for player i is a sequence $\gamma^i = (u_1^i, u_2^i, \ldots)$, where $u_k^i \in U^i$ for every k. A Markov policy γ^i for player i is called *stationary* if it is of the form $\gamma^i = (u^i, u^i, \ldots)$ for some $u^i \in U^i$. Every stationary policy $(u^i, u^i, \ldots)$ for player i can thus be identified with $u^i \in U^i$. Denote by $\Gamma = \prod_{i=1}^m \Gamma^i$ the *set of all multipolicies*, and by U the *subset of stationary multipolicies*. Let $H = S \times X \times S \times X \times \cdots$ be the space of all infinite histories of the game, endowed with the product σ-field. For any $\gamma \in \Gamma$ and every initial state $s_1 = s \in S$, a probability measure P_s^γ and a stochastic process $\{S_k, X_k\}$ are defined on H in a canonical way, where the random variables S_k and X_k describe the state and the action, respectively, chosen by the players on the kth stage of the game (see Proposition V.1.1 in [28]). Thus, for each initial state $s \in S$, any multipolicy $\gamma \in \Gamma$, and any finite horizon n, the *expected n-stage cost* of player i is

$$J_n^i(s, \gamma) = E_s^\gamma \left(\sum_{k=1}^n c^i(S_k, X_k) \right),$$

where E_s^γ means the expectation operator with respect to the probability measure P_s^γ. (Later on we make an assumption on the functions c^i that assures that all the expectations considered in this paper are well defined.)

The *average cost per unit time* to player i is defined as

$$J^i(s, \gamma) = \limsup_{n \to \infty} J_n^i(s, \gamma)/n.$$

If β is a fixed real number in $(0,1)$, called the *discount factor*, then the *expected discounted cost* to player i is

$$D^i(s,\gamma) = E_s^\gamma\left(\sum_{k=1}^{\infty}\beta^{k-1}c^i(S_k,X_k)\right).$$

For any multipolicy $\gamma = (\gamma^1,\dots,\gamma^m) \in \Gamma$ and a policy σ^i for player i, we define (γ^{-i},σ^i) to be the multipolicy obtained from γ by replacing γ^i with σ^i.

Let $\varepsilon \geq 0$. A multipolicy γ is called an ε-equilibrium for the average cost stochastic game if for every player i and any policy $\sigma^i \in \Gamma^i$,

$$J^i(s,\gamma) \leq J^i(s,(\gamma^{-i},\sigma^i)) + \varepsilon.$$

We similarly define ε-equilibria for the expected discounted cost games. Of course, a 0-equilibrium will be called a Nash equilibrium.

To ensure the existence of ε-equilibrium strategies for the players in the stochastic game, we will accept some regularity conditions on the primitive data, and in the average expected cost case we will also impose some general Lyapunov stability assumptions on the transition structure.

In both the discounted and average cost cases, we make the following assumptions.

C1: For each player i and $s \in S$, $c^i(s,\cdot)$ is continuous on X. Moreover, there exists a measurable function $\nu : S \to [1,\infty)$ such that

$$L \overset{\text{def}}{=} \sup_{s\in S, x\in X, i=1,\dots,m} \frac{|c^i(s,x)|}{\nu(s)} < \infty. \tag{2.1}$$

C2: There exists a probability measure $\varphi \in \mathbb{P}(S)$ such that

$$Q(B|s,x) = \int_B z(s,t,x)\varphi(dt)$$

for each $B \in \mathcal{G}$ and $(s,x) \in S \times X$. Moreover, we assume that if $x_n \to x_0$ in X, then

$$\lim_{n\to\infty}\int_S |z(s,t,x_n) - z(s,t,x_0)|\nu(t)\varphi(dt) = 0,$$

where ν was defined above (2.1).

Remark 2.1. Let w be a measurable function such that $1 \leq w(s) \leq \nu(s) + \delta$ for all $s \in S$ and for some $\delta > 0$. If $x_n \to x_0$ in X as $n \to \infty$, then

$$\int_S |z(s,t,x_n) - z(s,t,x_0)|w(s)\varphi(dt) \to 0.$$

This follows from **C2**, since $\nu \geq 1$ implies that

$$\int_S |z(s,t,x_n) - z(s,t,x_0)|\varphi(dt) \to 0.$$

3. The undiscounted stochastic game. To formulate our further assumptions, we introduce some helpful notation. Let $s \in S$, $u = (u^1, \dots, u^m) \in U$. We set

$$c^i(s,u) = \int_{X^1} \cdots \int_{X^m} c^i(s, x^1, \dots, x^m) u^1(dx^1|s) \cdots u^m(dx^m|s),$$

and, for any set $D \in \mathcal{G}$, we set

$$Q(D|s,u) = \int_{X^1} \cdots \int_{X^m} Q(D|s, x^1, \dots, x^m) u^1(dx^1|s) \cdots u^m(dx^m|s).$$

By $Q^n(\cdot|s,u)$, we denote the n-step transition probability induced by Q and the multipolicy $u \in U$.

C3 (Drift inequality): Let $\nu : S \to [1, \infty)$ be some given measurable function. There exists a set $C \in \mathcal{G}$ such that ν is bounded on C and for some $\xi \in (0,1)$ and $\eta > 0$ we have

$$\int_S \nu(t) Q(dt|s,x) \le \xi \nu(s) + \eta 1_C(s)$$

for each $(s,x) \in S \times X$. Here 1_C is the characteristic function of the set C.

C4: There exist a $\lambda \in (0,1)$ and a probability measure ζ concentrated on the set C such that

$$Q(D|s,x) \ge \lambda \zeta(D)$$

for any $s \in C$, $x \in X$ and for each measurable set $D \subset C$.

For any measurable function $w : S \to R$, we define the ν-weighted norm as

$$\|w\|_\nu = \sup_{s \in S} \frac{|w(s)|}{\nu(s)}.$$

We write L_ν^∞ to denote the Banach space of all measurable functions w for which $\|w\|_\nu$ is finite.

Condition **C3** implies that, outside a set C, the function ν decreases under any stationary multipolicy u; i.e.,

$$E_s^u \left(\nu(S_{k+1}) - \nu(S_k)|S_k\right) \le -(1-\xi)\nu(S_k) \le -(1-\xi). \tag{3.1}$$

This is known as a drift condition. If (i) the state space is countable, (ii) the set C is finite, and (iii) the state space is communicating under a stationary policy u, then (3.1) implies that the Markov chain (when using u) is ergodic. (This is the well known Foster criterion for ergodicity; see, e.g., [27].)

In the uncountable infinite state space, the same drift condition should be used to obtain the ergodicity condition. However, the finiteness of the set C is replaced by a weaker assumption. Namely, C has to be a *small* set or a *petite set* [25]; condition **C4** is a simple sufficient condition for the set C to be small.

Beyond the ergodicity of the Markov chain $\{S_k\}$ under a stationary multipolicy, Foster-type criteria (i.e., conditions **C3**–**C4**) also ensure the finiteness of the expectation $E_s^u \nu(S_k)$ in steady state, as well the finiteness of the expected cost $E_s^u w(S_k)$

for every potential cost function $w \in L_\nu^\infty$; moreover, they provide a geometric rate of convergence of the expected costs at time k to the steady state cost for $w \in L_\nu^\infty$. These statements will be made precise below.

Note that **C3**–**C4** provide uniform conditions for ergodicity; i.e., ξ, ζ, C, and λ do not depend on the actions (or on the policies). This will be needed in order for approximating games (with countable state space) to have stationary Nash equilibria [4].

LEMMA 3.1. *Assume* **C3**–**C4**. *Then the following properties hold.*

C5: *For every $u \in U$, the corresponding Markov chain is aperiodic and ψ_u-irreducible for some σ-finite measure ψ_u on $\mathcal{G}$. (The latter condition means that if $\psi_u(D) > 0$ for some set $D \in \mathcal{G}$, then the chance that the Markov chain (starting at any $s \in S$ and induced by u) ever enters D is positive.) Thus the state process $\{S_n\}$ is a positive recurrent Markov chain with the unique invariant probability measure denoted by π_u.*

C6: *For every stationary multipolicy u,*

(a)

$$\int_S \nu(s)\pi_u(ds) < \infty.$$

(b) *$\{S_n\}$ is ν-uniformly ergodic; that is, there exist $\theta > 0$ and $\alpha \in (0,1)$ such that*

$$\left| \int_S w(t)Q^n(dt|s,u) - \int_S w(t)\pi_u(dt) \right| \le \nu(s)\|w\|_\nu \theta\alpha^n$$

for every $w \in L_\nu^\infty$ and $s \in S$, $n \ge 1$.

Proof. **C3**–**C4** imply that for any stationary u, the chain is positive Harris recurrent (see Theorem 11.3.4 in [25]). It is thus ψ_u-irreducible (see Chapter 9 of [25]). The aperiodicity (and, in fact, strong aperiodicity) follows from condition **C4** (see [25, p. 116]). This establishes **C5**. **C6** follows from Theorem 2.3 in [26]. ∎

Remark 3.1. From Lemma 3.1 it follows that for any player i and $u \in U$ we have

$$J^i(u) := \int_S c^i(s,u)\pi_u(ds) = J^i(s,u);$$

that is, the expected average cost of player i is independent of the initial state. Theorem 2.3 in [26] implies that the constants α and θ in Lemma 3.1 depend only on ξ, η, λ, and $\nu_C = \sup_{s\in C} \nu(s)$ (and, in particular, they do not depend on u). **C1**, **C3**, and **C4** imply that the expected costs considered in this section are well defined for any multipolicy $\gamma \in \Gamma$; see [34] or [14].

In what follows, whenever we assume **C1**–**C4**, we shall take the same function ν in **C1** as in **C3**. We are now ready to state our first main result.

THEOREM 3.1. *Consider an undiscounted stochastic game satisfying* **C1**–**C4**. *Then for any $\varepsilon > 0$ there exists a stationary ε-equilibrium.*

The proof of this result is based on an approximation technique and consists of several steps which will be described later on. Before proving the result, we briefly mention the approach and the steps we are using, the difficulties, and the way we overcome these difficulties.

Basic idea behind the proof. Our basic goal is to approximate our game by a sequence of m-person games with countable state spaces and compact action spaces and which have equilibria in stationary policies; based on such approximating games, we shall construct a stationary policy which is an ϵ-equilibrium for the original game. The basic idea here is similar to the one already used in [29] for the problem with discounted cost. However, the situation here is much more involved; indeed, in the discounted case one does not need to bother about the ergodic structure of the approximating games in order to show that they possess equilibrium in stationary policies. Here, in contrast, we need to carefully construct the approximating games so as to ensure that they not only have the required ergodic property but also are uniform ergodic and have some additional "good" properties for the cost. Our first step in the proof will be to construct such approximating games, which will also satisfy conditions **C1**–**C4**. The function ν, as well as the other objects that appear in these assumptions, will be approximated as well. (We will have to show, for example, that the approximation of ξ is indeed within $(0,1)$, etc.) The approximation of the game in a way that allows conditions similar to **C1**–**C4** to hold is done in the next two subsections.

Properties similar to **C2**–**C4** were used in [4] to establish the existence of equilibria in stationary policies for games with countable state space; the properties imply, for example, that the costs are continuous in the policies. Unfortunately, the counterpart of property **C4** that is used to establish ergodicity in the literature of countable state Markov chains (or for Markov decision processes, or for Markov games) requires that the set C that appears in conditions **C3**–**C4** be finite. Unfortunately, we *were not able to* come up with a direct approximation scheme for which C is finite. To overcome this problem, we first use some results from [26] to obtain uniform ergodicity results for the approximating chains. Using a key theorem from [41], this will be shown to imply that there exist some function (instead of the original approximation of ν) and constants for which properties **C3**–**C4** hold and for which C is a singleton. This is done in subsection 3.3.

3.1. Transition operators and their $\boldsymbol{\nu}$-weighted norms. If $f \in L_\nu^\infty$ and σ is a finite signed measure on $(S, \mathcal{G})$, then for convenience we set

$$\sigma(f) = \int_S f(s)\sigma(ds),$$

provided that this integral exists. Let P_1 and P_2 be transition subprobabilities from S into S. Define

$$\|P_1 - P_2\|_\nu = \sup_{s\in S} \sup_{|f|\le\nu} \frac{|P_1(f|s) - P_2(f|s)|}{\nu(s)}. \tag{3.2}$$

We will also use the definition (3.2) in the case in which P_1 and P_2 are probability measures on $(S, \mathcal{G})$, or when one of them is zero. Note that if $P_2 = 0$ and P_1 is a transition probability, then it follows from (3.2) that

$$\|P_1\|_\nu = \sup_{s\in S} \frac{P_1(\nu|s)}{\nu(s)}.$$

If P_1 and P_2 are transition probabilities and $\|P_1 - P_2\| < \infty$, then $P_1 - P_2$ induces a bounded linear operator from L_ν^∞ into itself, and $\|P_1 - P_2\|_\nu$ is its operator norm (see Lemma 16.1.1 in [25]).

We now come back to our game model and accept the following notation. For any $u \in U$, we use $Q(u)$ to denote the operator on L_ν^∞ defined by $Q(u)f(s) = Q(f|s, u)$,

$s \in S$, and $f \in L_\nu^\infty$. By **C3**, we have

$$\|Q(u)\|_\nu \le \xi + \eta. \tag{3.3}$$

Clearly, (3.3) implies that $Q(u)$ is (under condition **C3**) a bounded linear operator from L_ν^∞ into itself. By $\Pi(u)$ we denote the invariant probability measure operator given by

$$\Pi(u)f = \pi_u(f),$$

where π_u is the invariant probability measure for $Q(\cdot|s,u)$, $u \in U$, and $f \in L_\nu^\infty$.

3.2. Approximating games. We define Γ_A to be the class of stochastic games that "resemble" stochastic games with countably many states and can be used to approximate the original game. The games in Γ_A will depend on some parameter $\delta > 0$. The transition probability in a game belonging to Γ_A is denoted by Q_δ, and the cost function of player i is denoted by c_δ^i.

We introduce some notation:

- $\mathbb{N}$—the set of positive integers,
- $C(X)$—the Banach space of all continuous functions on X, endowed with the supremum norm $\|\cdot\|$,
- $L_\nu^1 = L_\nu^1(S, \mathcal{G}, \varphi)$—the Banach space of measurable functions $f : S \to \mathbb{R}$ such that $\int_S |f(s)|\nu(s)\varphi(ds) < \infty$.

We assume that each game $G_\delta \in \Gamma_A$ corresponds to some sequences $\{Y_n\}$, $\{c_n^i\}$, $\{z_n\}$, and $\{\nu_n\}$, where n belongs to some subset $\mathbb{N}_1 \subset \mathbb{N}$ and $\{Y_n\}$ is a measurable partition of the state space such that $Y_n \subset C$ or $Y_n \subset S \setminus C$ for each $n \in \mathbb{N}_1$ (the set C is introduced in assumption **C3**),

$$c_\delta^i(s,x) = c_n^i(x), \quad \text{and} \quad Q_\delta(B|s,x) = \int_B z_n(t,x)\varphi(dt)$$

for all $s \in Y_n$, $x \in X$, and $n \in \mathbb{N}_1$. Moreover, ν_n are rational numbers and $\nu_n \ge 1$ for all $n \in \mathbb{N}_1$. Define $\nu_\delta(s) \overset{\text{def}}{=} \nu_n$ if $s \in Y_n$.

We will show that for each $\delta > 0$ it is possible to construct a game G_δ such that $c_n^i \in C(X)$ and $z_n(\cdot, x) \in L_\nu^1$ while $z_n(s, \cdot) \in C(X)$ for all $n \in \mathbb{N}_1$, $x \in X$, and $s \in S$.

Because in our approximation we need to preserve (in some sense) condition **C4**, we consider the following subset $\Delta \subset L_\nu^1$: $\phi \in \Delta$ if and only if ϕ is a density function such that

$$\int_D \phi(s)\varphi(ds) \ge \lambda\zeta(D) \tag{3.4}$$

for each $D \in \mathcal{G}$ such that $D \subset C$. Our assumption **C4** implies that $\Delta \ne \emptyset$. It is obvious that Δ is *convex*. Suppose that $\phi_n \in \Delta$ and $\phi_n \to \phi \in L_\nu^1$. Since $\nu \ge 1$, then $\phi_n \to \phi$ in L^1. By Scheffe's theorem, ϕ is a density function. Moreover, ϕ satisfies (3.4). Thus, we have shown that Δ is a *closed* and *convex* subset of L_ν^1.

Let V be the space of all continuous mappings from X into Δ with the metric ρ defined by

$$\rho(\phi_1, \phi_2) = \max_{x \in X} \int_S |\phi_1(x)(s) - \phi_2(x)(s)|\nu(s)\varphi(ds). \tag{3.5}$$

Since $\mathcal{G}$ is countably generated, L_1 is separable. As in [47, Theorem I.5.1], we can prove the following.

LEMMA 3.2. *V is a complete separable metric space.*

Note that the proof of Theorem I.5.1 in [47] makes use of the convexity of the range space of the continuous mappings involved. In our case, the range space Δ is also convex.

For each $s \in \mathbf{S}$, the transition probability density z of the original game induces elements $\phi(s, \cdot)$ of V by

$$\phi(s, x) = z(s, \cdot, x).$$

From the product measurability of z on $S \times S \times X$, it follows that $s \to \phi(s, \cdot)$ is a measurable mapping from S into V.

We introduce the following notation:

- $\{\phi_k\}$—a countable dense subset of V (see Lemma 3.2),
- $\{c_k\}$—a countable dense set in $C(X)$,
- $\{r_k\}$, $r_k \geq 1$, where $\{r_k\}$ is the set of all rational numbers satisfying $r_k \geq 1$.

Let $0 < \delta < 1$ be fixed. Define for any $k, k_1, \ldots, k_m, l$

$$B(k, k_1, \ldots, k_m, l) = \left\{ s \in S : \rho(\phi(s, \cdot), \phi_k) + \sum_{i=1}^{m} \left\| c^i(s, \cdot) - c_{k_i} \right\| + |\nu(s) - r_l| < \delta \right\}.$$

Let τ be a (fixed) one-to-one correspondence between $\mathbb{N}$ and $\mathbb{N} \times \cdots \times \mathbb{N} = \mathbb{N}^{m+2}$. Define $T_n \stackrel{\text{def}}{=} B(\tau(n)),\ n \in \mathbb{N}$. Next, set $\bar{Y}_1 \stackrel{\text{def}}{=} T_1$ and $\bar{Y}_k \stackrel{\text{def}}{=} T_k - \cup_{j<k} \bar{Y}_j$ for $k \geq 2$.

Let $\{Y_n\}$ be the enumeration of all nonempty sets $\bar{Y}_k$. Clearly, $\{Y_n\}$ is a measurable countable partition of the state space, and n belongs to some $\mathbb{N}_1 \subset \mathbb{N}$.

If necessary, we can modify (trivially) this partition in such a way that $Y_n \subset C$ or $Y_n \subset S \setminus C$ for each n. Note that for each $n \in \mathbb{N}_1$ and each set Y_n there correspond some $z_n \in V$ and $c_n^i \in C(X)$, so that we obtain a game $G_\delta \in \Gamma_A$. Moreover, we have

$$\rho(\phi(s, \cdot), z_n) < \delta$$

($\phi(s, x) = z(s, \cdot, x)$, by definition) for each $n \in \mathbb{N}_1$ and $s \in Y_n$. This implies that

$$\|Q(u) - Q_\delta(u)\|_\nu < \delta \tag{3.6}$$

for every $u \in U$. Next, we have

$$\left\| c^i(s, \cdot) - c_n^i \right\| < \delta$$

for each $n \in \mathbb{N}_1$ and $s \in Y_n$. If we set $c_\delta^i(s, x) = c_n^i(x)$ for $s \in Y_n, x \in X$, we obtain

$$\sup_{s \in S} \sup_{x \in X} |c^i(s, x) - c_\delta^i(s, x)| \leq \delta. \tag{3.7}$$

We also have

$$|\nu(s) - \nu_\delta(s)| < \delta \tag{3.8}$$

for every $s \in S$.

3.3. Equivalence with a game with a countable state space. Next, we shall show that the G_δ game has an equilibrium in the class of stationary multipolicies. This will be done in the proof of the following lemma.

LEMMA 3.3. *Assume that the stochastic game satisfies* **C1**–**C4** *and* $\xi_\delta \stackrel{\text{def}}{=} 3\delta+\xi < 1$. *Then*

(i) *the game* G_δ *satisfies* **C3** *with* ξ *and* ν *replaced by* ξ_δ *and* ν_δ, *respectively, and it satisfies* **C4**; *moreover,*

(ii) *it has a Nash equilibrium in the class of stationary multipolicies.*

Proof. (i) From (3.6), it follows that

$$Q_\delta(\nu|s,x) \le Q(\nu|s,x) + \delta\nu(s)$$

for every $s \in S$ and $x \in X$. Since **C3** holds for the original game, this implies that

$$Q_\delta(\nu|s,x) \le (\delta+\xi)\nu(s) + \eta 1_C(s). \tag{3.9}$$

From (3.8) and (3.9), we conclude that

$$Q_\delta(\nu_\delta|s,x) \le \delta + (\delta+\xi)\delta + (\delta+\xi)\nu_\delta(s) + \eta 1_C(s).$$

Hence

$$Q_\delta(\nu_\delta|s,x) \le \xi_\delta\nu_\delta(s) + \eta 1_C(s) \tag{3.10}$$

for every $s \in S$ and $x \in X$; i.e., a condition of the type **C3** holds. Condition **C4** follows from the construction of Δ (above (3.4)).

(ii) Consider the approximating games under the further assumption that every player i restricts to the class U_0^i of policies that are piecewise constant: u^i belongs to U_0^i if and only if $s \to u^i(\cdot|s)$ is constant on each set Y_n of the partition $\{Y_n\}$ of S. Denote by U_0 the set of all stationary piecewise constant multipolicies. Every game G_δ with the above restriction is equivalent to a stochastic game denoted by $\bar{G}$ with the countable state space $\mathbb{N}_1$ (defined in our approximation procedure). Because every stationary multipolicy in $\bar{G}$ corresponds to a multipolicy in U_0, we will use U_0 also to denote the set of *all stationary multipolicies* in $\bar{G}$. The cost functions in $\bar{G}$ are $c_n^i \in C(X)$, where $n \in \mathbb{N}_1$. The transition probabilities in $\bar{G}$ are given by

$$P_{mn}(u) = Q_\delta(Y_n|s,u)$$

for all $s \in Y_m$, $u \in U_0$, and $m,n \in \mathbb{N}_1$. Let $P(u)$ denote the transition probability matrix corresponding to any $u \in U_0$. Finally, the piecewise constant function ν_δ induces a function $\mu : \mathbb{N}_1 \to [1,\infty)$ by $\mu(n) = \nu_\delta(s), s \in Y_n, n \in \mathbb{N}_1$. (Sometimes we will identify μ with the column vector, and $\mu(n)$ with its nth coordinate.)

Fix δ such that $\xi_\delta < 1$. Applying Lemma 3.1 to the game G_δ, we conclude that it satisfies **C5**. By part (i) and Theorem 2.3 in [26], this game also satisfies **C6** (with possibly different constants θ_1 and α_1). A simple translation of **C5** to the game $\bar{G}$ with countably many states says that for any $u \in U_0$ the Markov chain with the transition probability matrix $P(u)$ has a single ergodic class and is aperiodic. On the other hand, a translation of **C6** and the fact that $\|Q_\delta(u)\|_{\nu_\delta} \le \xi_\delta + \eta$ (which follows from (3.10)) mean that the Markov chain is μ-uniform geometric ergodic; see [9, 41]. By Key Theorem II and Lemma 5.3(ii) in [41], there exist a nonempty finite set $M \subset \mathbb{N}_1$, a function $\tilde{\mu} : \mathbb{N}_1 \to [1,\infty)$, and some $b \in (0,1)$ such that

$$\sum_{n \notin M} P_{kn}(u)\tilde{\mu}(n) \le b\tilde{\mu}(k) \tag{3.11}$$

for every $k \in \mathbb{N}_1$ and $u \in U_0$. This property is called the $\tilde{\mu}$-uniform geometric recurrence (see [8, 9, 41]) and is Assumption A2(1) in [4]. The function $\tilde{\mu}$ is given by

$$\tilde{\mu}(k) = \mu(k) + \sup_{u \in U_0} \left(\sum_{n=1}^{\infty} {}_M P^n(u)\mu \right)(k),$$

where $k \in \mathbb{N}_1$ (${}_M P(u)$ is the matrix $P(u)$ in which we replace the columns corresponding to states $m \in M$ by zeros); see [4, pp. 99–100]. Note that this new function $\tilde{\mu}$ is μ-bounded (i.e., $\sup_{n\in\mathbb{N}_1} \tilde{\mu}(n)/\mu(n) < \infty$) and vice versa. Indeed, $\mu(k) \le \tilde{\mu}(k)$ for each k. On the other hand, by (3.11), we have $({}_M P(u)\tilde{\mu})(k) \le b\tilde{\mu}(k)$ for any $k \in \mathbb{N}_1$ and $u \in U_0$. Hence

$$\tilde{\mu}(k) = \mu(k) + \sup_{u \in U_0} \left(\sum_{n=1}^{\infty} {}_M P^n(u)\mu \right)(k) \le \mu(k) + \sup_{u \in U_0} ({}_M P(u)\tilde{\mu})(k) \le \mu(k) + b\tilde{\mu}(k).$$

Thus, $\tilde{\mu}(k) \le \mu(k)/(1-b)$ for every $k \in \mathbb{N}_1$. Since $\tilde{\mu}$ is μ-bounded and vice versa, this implies the $\tilde{\mu}$-continuity of the immediate costs (recall **C1**) and of the transition probabilities (recall **C2** and Remark 2.1). This is Assumption 1* in [4]. Since both assumptions 1* as well as 2(1) in [4] hold, it follows that the game $\bar{G}$ has a stationary Nash equilibrium $u^* \in U_0$, and consequently G_δ has a stationary equilibrium (also denoted by u^*) in the class U_0 of all stationary piecewise constant multipolicies. It now follows from Lemma 5.1 in the appendix that u^* is a Nash equilibrium for the game G_δ in the class U of all stationary multipolicies. ∎

3.4. Uniform convergence of the steady state probabilities and costs, and proof of the main result. Let $J^i_\delta(u)$ be the expected average cost for player i in the game G_δ when a stationary multipolicy u is used (see Remark 3.1 and Lemma 3.3(i)).

Let $Q_\delta(u)$ and $\Pi_\delta(u)$ denote the transition probability and the invariant probability measure operators under any stationary multipolicy $u \in U$ in the approximating game.

LEMMA 3.4. *Under* **C1**–**C4***,*

(i)

$$\lim_{\delta\to 0} \|\Pi_\delta(u) - \Pi(u)\|_\nu = 0$$

uniformly in $u \in U$.

(ii)

$$\lim_{\delta\to 0} \left| J^i_\delta(u) - J^i(u) \right| = 0$$

uniformly in $u \in U$.

Proof. (i) If $\xi_\delta = 3\delta + \xi < 1$, then (by Lemma 3.3) the games G_δ satisfy **C4** and **C3**, with ξ replaced by ξ_δ. By (3.10), we have $\|Q_\delta(u)\|_{\nu_\delta} \le \xi_\delta + \eta$ for all $u \in U$. This and Theorem 2.3 in [26] (applied to the games G_δ) imply that there exists a δ_0 such that

$$\sup_{\delta\le\delta_0} \sup_{u\in U} \|\Pi_\delta(u)\|_{\nu_\delta} < \infty.$$

Hence

$$K_0 \overset{\text{def}}{=} \sup_{\delta\le\delta_0} \sup_{u\in U} \|\Pi_\delta(u)\|_\nu < \infty. \tag{3.12}$$

The rest of the proof is an adaptation of the proof of Proposition 1 in [42]. By Lemma 3.1 and Remark 3.1, there exist some $\theta > 0$ and $\alpha \in (0,1)$ such that

$$\sup_{u \in U} \|Q^n(u) - \Pi(u)\|_\nu \leq \theta\alpha^n$$

for every n. Hence, there exists an n_0 such that

$$\sup_{n \geq n_0} \sup_{u \in U} \left\| \frac{1}{n} \sum_{i=0}^{n-1} Q^i(u) - \frac{1}{n}(I - Q(u)) - \Pi(u) \right\|_\nu < 1;$$

$Q^0 = I$ is the identity operator. Therefore for each $n \geq n_0$ there exists a ν-bounded transition operator

$$\Phi_n(u) \stackrel{\text{def}}{=} \left(I + \Pi(u) - \frac{1}{n} \sum_{i=0}^{n-1} Q^i(u) + \frac{1}{n}(I - Q(u)) \right)^{-1}$$

and

$$K_1 \stackrel{\text{def}}{=} \sup_{n \geq n_0} \sup_{u \in U} \|\Phi_n(u)\|_\nu < \infty. \tag{3.13}$$

Define

$$Z_n(u) \stackrel{\text{def}}{=} I + \frac{1}{n} \sum_{i=1}^{n-1} \sum_{j=1}^{i-1} (Q^j(u) - \Pi(u)).$$

We have

$$K_2 \stackrel{\text{def}}{=} \sup_{n \geq n_0} \sup_{u \in U} \|Z_n(u)\|_\nu < \infty. \tag{3.14}$$

A direct calculation yields

$$(I - Q(u) + \Pi(u)) Z_n(u) \Phi_n(u) = I. \tag{3.15}$$

Clearly, (3.15) implies that

$$\Pi_\delta(u)(I - Q(u) + \Pi(u)) Z_n(u) \Phi_n(u) = \Pi_\delta(u),$$

so that

$$\Pi_\delta(u)(I - Q(u)) Z_n(u) \Phi_n(u) + \Pi(u) Z_n(u) \Phi_n(u) = \Pi_\delta(u). \tag{3.16}$$

From (3.15), we infer that

$$\Pi(u)(I - Q(u) + \Pi(u)) Z_n(u) \Phi_n(u) = \Pi(u).$$

Therefore

$$\Pi(u) Z_n(u) \Phi_n(u) = \Pi(u).$$

Substituting into (3.16), we obtain

$$\Pi_\delta(u)(I - Q(u)) Z_n(u) \Phi_n(u) = \Pi_\delta(u) - \Pi(u),$$

and consequently

$$\|\Pi_\delta(u) - \Pi(u)\|_\nu = \|\Pi_\delta(u)(Q_\delta(u) - Q(u))Z_n(u)\Phi_n(u)\|_\nu .$$

Combining this with (3.6) and (3.12)–(3.14), we obtain

$$\|\Pi_\delta(u) - \Pi(u)\|_\nu \le \|Q_\delta(u) - Q(u)\|_\nu K_0K_1K_2 < \delta K_0K_1K_2.$$

The proof of statement (i) is finished.

(ii) Using L defined in (2.1) and (3.7), we obtain

$$\begin{aligned} |J^i(u) - J^i_\delta(u)| &= |\Pi(u)c^i(\cdot,u) - \Pi_\delta(u)c^i_\delta(\cdot,u)| \\ &\le |\Pi(u)c^i(\cdot,u) - \Pi_\delta(u)c^i(\cdot,u)| + |\Pi_\delta(u)(c^i(\cdot,u) - c^i_\delta(\cdot,u))| \\ &\le L\nu(s_0) \sup_{|w|\le\nu} \frac{|\Pi(u)w - \Pi_\delta(u)w|}{\nu(s_0)} + \delta \\ &\le L\nu(s_0) \sup_{s\in S}\sup_{|w|\le\nu} \frac{|\Pi(u)w - \Pi_\delta(u)w|}{\nu(s)} + \delta \\ &= L\nu(s_0)\,\|\Pi(u) - \Pi_\delta(u)\|_\nu + \delta, \end{aligned}$$

where s_0 is an arbitrary state. Now (ii) follows from (i). □

A version of Lemma 3.4 corresponding with a *bounded* function ν was established by Stettner [42]. When ν is bounded, an elementary proof of Lemma 3.4 (stated as an extension of Ueno's lemma [46]) is possible [32].

Proof of Theorem 3.1. Choose some $\epsilon > 0$. According to Lemma 3.4 there exists some δ such that for all $u \in U$,

$$\left|J^i(u) - J^i_\delta(u)\right| \le \epsilon. \tag{3.17}$$

Let $u^* \in U$ be a Nash equilibrium for the game G_δ in the class U of multipolicies (its existence follows from Lemma 3.3). It then follows from (3.17) that u^* is an ϵ-equilibrium (in the class U) for the original game. The fact that u^* is an ϵ-equilibrium in the class Γ of all multipolicies follows from Theorem 3 and Remark 1 in [34] (or [18, 14] in the Borel state space framework). □

4. The discounted stochastic game. In this section, we drop conditions **C3** and **C4**. However, in the unbounded cost case, we make the following assumption.

C7: There exists $\alpha \in [\beta, 1)$ such that

$$\beta Q(\nu|s,x) \le \alpha\nu(s)$$

for each $s \in S$ and $x \in X$.

Using **C7**, we can easily prove that, for any $s \in S$, any multipolicy $\gamma \in \Gamma$, and any number of stages k, we have

$$|\beta^{k-1}E^\gamma_s(c^i(S_k,X_k))| \le \beta^{k-1}E^\gamma_s(|c^i(S_k,X_k)|) \le L\beta^{k-1}E^\gamma_s(\nu(S_k)) \le L\alpha^{k-1}\nu(s),$$

where L is the constant defined in **C1**. This gives us the following lemma.

Lemma 4.1. *Assume* **C1** *and* **C7**. *Then for every player i the expected discounted cost $D^i(s,\gamma)$ is well defined (absolutely convergent) for each $s \in S$ and $\gamma \in \Gamma$.*

We are ready to formulate our main result in this section.

Theorem 4.1. *Any discounted stochastic game satisfying conditions* **C1**, **C2**, *and* **C7** *has a stationary ε-equilibrium for any $\varepsilon > 0$.*

Before we give the proof of this theorem, we state some auxiliary results. Let Δ_1 be the set of *all density functions* in L^1_ν. Clearly, Lemma 3.2 holds true if Δ is replaced by Δ_1. Applying the approximation scheme from section 3 to the present situation, we construct a game G_δ for any $\delta > 0$ such that (3.7) holds and, moreover, we have

$$\sup_{|f|\le\nu} |Q(f|s,u) - Q_\delta(f|s,u)| \le \delta \tag{4.1}$$

for each $s \in S$ and any stationary multipolicy $u \in U$.

Fix player i and set

$$K^n(s,u) = E^u_s(c^i(S_n, X_n)) \quad \text{and} \quad K^n_\delta(s,u) = E^u_s(c^i_\delta(S_n, X_n)),$$

where $s \in S$ and $u \in U$. Clearly, $K^n(s,u)$ is the nth stage cost for player i under stationary multipolicy u when the game starts at an initial state $s \in S$.

LEMMA 4.2. *Assume* **C1** *and* **C7**. *Then, for each $s \in S$ and $u \in U$, we have*

$$|K^n(s,u) - K^n_\delta(s,u)| \le \delta(1 + (n-1)L)\left(\frac{\alpha}{\beta}\right)^{n-1}.$$

Proof. The proof proceeds by induction. For $n = 1$ the inequality follows immediately from (3.7). We now give the induction step. Note that

$$\begin{aligned} |K^{n+1}(s,u) - K^{n+1}_\delta(s,u)| &= |Q(K^n(\cdot,u)|s,u) - Q_\delta(K^n_\delta(\cdot,u)|s,u)| \\ &\le |Q(K^n(\cdot,u)|s,u) - Q_\delta(K^n(\cdot,u)|s,u)| \\ &\quad + |Q_\delta(K^n(\cdot,u)|s,u) - Q_\delta(K^n_\delta(\cdot,u)|s,u)|. \end{aligned}$$

Using (4.1), our induction hypothesis, and the obvious inequality

$$K^n(s,u) \le L\left(\frac{\alpha}{\beta}\right)^{n-1}\nu(s),$$

which holds for every $s \in S$ and $u \in U$, we obtain

$$\begin{aligned} |K^{n+1}(s,u) - K^{n+1}_\delta(s,u)| &\le \delta L\left(\frac{\alpha}{\beta}\right)^{n-1} + \delta(1 + (n-1)L)\left(\frac{\alpha}{\beta}\right)^{n-1} \\ &= \delta(1 + nL)\left(\frac{\alpha}{\beta}\right)^{n-1} \le \delta(1 + nL)\left(\frac{\alpha}{\beta}\right)^{n}, \end{aligned}$$

which ends the proof. ∎

From Lemmas 4.1 and 4.2, we infer the following result.

LEMMA 4.3. *Assume* **C1** *and* **C7**. *If $D^i_\delta(s,u)$ is the expected β-discounted cost for player i in the game G_δ, then*

$$|D^i(s,u) - D^i_\delta(s,u)| \le \delta(1 + \alpha(L-1))(1-\alpha)^{-2}$$

for each $s \in S$ and $u \in U$.

The game G_δ is characterized by the cost functions c^i_δ, transition probability Q_δ, and the function ν_δ. Note that if δ is sufficiently small, then the game G_δ satisfies condition **C1** with L replaced by $2L$. From our approximation scheme (the new definition of the space V) and Remark 2.1, it follows also that **C2** is satisfied in our game G_δ. Since $|\nu(s) - \nu_\delta(s)| < \delta$ for all $s \in S$, we have (by (4.1))

$$\beta \int_S \nu_\delta(t) Q_\delta(dt|s,x) \leq \alpha\nu_\delta(s) + \alpha\delta + 2\beta\delta < \alpha_0 \nu_\delta(s),$$

where $\alpha_0 = \alpha + \alpha\delta + 2\delta$ and $s \in S$, $x \in X$. Note that $\beta < \alpha_0$, and if δ is sufficiently small, then $\alpha_0 < 1$, and thus G_δ satisfies condition **C7** with α replaced by α_0. Let δ_0 be a positive number such that for every $\delta < \delta_0$ the game G_δ satisfies conditions of type **C1**, **C2**, and **C7**. In particular, we have $\beta < \alpha_0 < 1$.

LEMMA 4.4. *If $\delta < \delta_0$, then the game G_δ has a Nash equilibrium in the class U of all stationary multipolicies.*

Proof. We use a transformation to bounded cost games similar to that of [43, p. 101]. One may define the new discount factor $\tilde{\beta} \stackrel{\text{def}}{=} \alpha_0$ and the functions

$$\tilde{c}^i(s,x) = \frac{c_n^i(x)}{\nu_\delta(s)}, \qquad \tilde{z}(s,t,x) = \frac{\beta z_n(t,x)\nu_\delta(t)}{\alpha_0 \nu_\delta(s)},$$

where $s \in Y_n$, $t \in S$, and $x \in X$. This transformation ensures that the new costs $\tilde{c}^i$ are bounded and that

$$q(\cdot|s,x) \stackrel{\text{def}}{=} \int_S \tilde{z}(s,t,x)\varphi(dt)$$

is a transition subprobability such that $q(Y_n|s,x)$ is continuous in x for each n and $s \in S$. Moreover, it implies that

$$\tilde{D}^i(s,u) = \frac{D_\delta^i(s,u)}{\nu_\delta(s)}, \tag{4.2}$$

where $\tilde{D}^i(s,u)$ is the expected discounted cost for player i under any $u \in U$ in the transformed (bounded) game. Similarly, as in section 3 we can recognize the game G_δ as a game with countably many states. By [11], such a game has a stationary Nash equilibrium. In other words, our bounded game has an equilibrium $u^\star$ in the class U_0 of all piecewise constant multipolicies. It now follows from Lemma 5.2 in the appendix that $u^\star$ is an equilibrium for the bounded game in the class U. By (4.2), we infer that $u^\star$ is also an equilibrium (in the class U of all stationary multipolicies) for the game G_δ. ∎

Proof of Theorem 4.1. Fix $\varepsilon > 0$. By Lemma 4.3, there exists $\delta < \delta_0$ such that

$$|D^i(s,u) - D_\delta^i(s,u)| \leq \varepsilon/2$$

for each $s \in S$ and $u \in U$. It follows from Lemma 4.4 that the game G_δ has an equilibrium $u^\star$ in the class U. Clearly, $u^\star$ is an ε-equilibrium in the class U for the original game. The fact that $u^\star$ is also an ε-equilibrium in Γ follows from Theorem 2 and Remark 1 in [34] (or [14] in the case of Borel state space games). ∎

5. Appendix. In this section we restrict our attention to the approximating games and state some auxiliary results on sufficiency of piecewise constant policies in the sense that they can be used to dominate any other policy. Related statements are proven in [1] for countable state space models. Their extension to the present situation would require new notation and some additional measure theoretic work.

Therefore, in this section we restrict ourselves to stationary policies, and in such a case we can use different methods which are based on some standard arguments from the dynamic programming literature [14].

Let G_δ be an approximating game corresponding to a partition of the state space. Fix player i and a stationary *piecewise constant* multipolicy u^{-i} for the other players. For any $s \in S$ and $f \in U^i$ set

$$c(s, f) = c^i_\delta(s, (u^{-i}, f)) \quad \text{and} \quad q(\cdot|s, f) = Q_\delta(\cdot|s, (u^{-i}, f)).$$

Recall that U^i_0 denotes the set of all piecewise constant stationary policies for player i.

Consider the Markov decision process (MDP) with the state space S, the action space X^i, the cost function c, and the transition probability q.

The average cost case. We assume that δ is sufficiently small so that the MDP satisfies conditions **C1**–**C4** (restricted to the one-player case). Let $J_n(s, f)$ $(J(f))$ denote the expected n-stage (expected average) cost (in the MDP) under stationary policy f.

LEMMA 5.1. *Assume* **C1–C4***, and consider the average cost MDP described above. Then for each $f \in U^i$ there exists some $f_0 \in U^i_0$ such that*

$$J(f_0) \leq J(f).$$

Proof. Let $f \in U^i$ and $g = J(f)$. By Lemma 3.1, our MDP satisfies condition **C6** with ν replaced by ν_δ and possibly different constants. It is well known that in such a case the function

$$h(s) \stackrel{\text{def}}{=} E^f_s \left[\sum_{n=1}^{\infty} (c(S_n, X^i_n) - g) \right]$$

is well defined and $h \in L^\infty_{\nu_\delta}$. Moreover, we have

$$g + h(s) = c(s, f) + q(h|s, f) \quad \text{for each} \quad s \in S.$$

For the details, see [14] and [25]. Our approximating game (and thus the MDP) satisfies continuity conditions **C1**–**C2**. Because the cost function c and the transition probability correspond to a partition of the state space (and, in addition, the other players use stationary piecewise constant multipolicy u^{-i}), this implies that one can find some $f_0 \in U^i_0$ such that

$$c(s, f_0) + q(h|s, f_0) \leq c(s, f) + q(h|s, f) = g + h(s)$$

for all $s \in S$. Iterating this inequality, we obtain

$$J_n(s, f_0) + q^n(h|s, f_0) \leq ng + h(s)$$

for all $s \in S$. Hence

$$\frac{J_n(s, f_0)}{n} + \frac{q^n(h|s, f_0)}{n} \leq g + \frac{h(s)}{n}$$

for each n, and consequently

$$J(f_0) \leq g = J(f).$$

For a detailed discussion of the fact that **C3** implies that $q^n(h|s, f_0)/n \to 0$ as $n \to \infty$, consult [14] or [34]. $\square$

The discounted cost case. We now assume that the stochastic game satisfies **C1**, **C2**, and **C7**. If δ is sufficiently small, then both G_δ and the aforementioned MDP satisfy **C1**, **C2**, and **C7**, but with different constants (see section 4). Let $f \in U^i$. By $D_n(s, f)$ ($D(s, f)$) we denote the expected n-stage discounted (total discounted) cost in the MDP under policy f.

LEMMA 5.2. *Assume* **C1**, **C2**, *and* **C7**, *and consider the discounted MDP described above. Then for each $f \in U^i$ there exists some $f_0 \in U_0^i$ such that*

$$D(s, f_0) \le D(s, f) \quad \textit{for every} \quad s \in S.$$

Proof. Set $d(s) = D(s, f)$, $s \in S$. Under our assumptions, we have

$$d(s) = c(s, f) + \beta q(d|s, f)$$

for all $s \in S$ (see Lemma 4.1). From our compactness and continuity conditions, **C7**, and the construction of the approximating game, it follows that there exists some $f_0 \in U_0^i$ such that

$$c(s, f_0) + \beta q(d|s, f_0) \le c(s, f) + \beta q(d|s, f) = d(s)$$

for each $s \in S$. Hence

$$D_n(s, f_0) + \beta^n q^n(d|s, f_0) \le d(s)$$

for each n and $s \in S$, and consequently

$$D(s, f_0) \le D(s, f) \quad \text{for each} \quad s \in S.$$

The fact that $\beta^n q^n(d|s, f_0) \to 0$ as $n \to \infty$ follows easily from **C7**. $\square$

REFERENCES

[1] E. ALTMAN, *Constrained Markov Decision Processes*, Chapman and Hall, London, 1998.

[2] E. ALTMAN, *Non zero-sum stochastic games in admission, service and routing control in queueing systems*, QUESTA, 23 (1996), pp. 259–279.

[3] E. ALTMAN AND A. HORDIJK, *Zero-sum Markov games and worst-case optimal control of queueing systems*, QUESTA, 21, (1995), pp. 415–447.

[4] E. ALTMAN, A. HORDIJK, AND F. M. SPIEKSMA, *Contraction conditions for average and α-discount optimality in countable state Markov games with unbounded rewards*, Math. Oper. Res., 22 (1997), pp. 588–618.

[5] R. AMIR, *Continuous stochastic games of capital accumulation with convex transitions*, Games Econom. Behav., 15 (1996), pp. 111–131.

[6] V. S. BORKAR AND M. K. GHOSH, *Denumerable state stochastic games with limiting average payoff*, J. Optim. Theory Appl., 76 (1993), pp. 539–560.

[7] L. O. CURTAT, *Markov equilibria of stochastic games with complementarities*, Games Econom. Behav., 17 (1996), pp. 177–199.

[8] R. DEKKER AND A. HORDIJK, *Recurrence conditions for average and Blackwell optimality in denumerable state Markov decision chains*, Math. Oper. Res., 17 (1992), pp. 271–290.

[9] R. DEKKER, A. HORDIJK, AND F. M. SPIEKSMA, *On the relation between recurrence and ergodicity properties in denumerable Markov decision chains*, Math. Oper. Res., 19 (1994), pp. 539–559.

[10] D. DUFFIE, J. GEANAKOPLOS, A. MAS-COLELL, AND A. MCLENNAN, *Stationary Markov equilibria*, Econometrica, 62 (1994), pp. 745–782.

[11] A. FEDERGRUEN, *On N-person stochastic games with denumerable state space*, Adv. in Appl. Probab., 10 (1978), pp. 452–471.

[12] J. A. Filar and K. Vrieze, *Competitive Markov Decision Processes*, Springer-Verlag, New York, 1997.
[13] C. Harris, P. Reny, and A. Robson, *The existence of subgame-perfect equilibrium in continuous games with almost perfect information: A case for public randomization*, Econometrica, 63 (1995), pp. 507–544.
[14] O. Hernández-Lerma and J. B. Lasserre, *Further Topics in Discrete-Time Markov Control Processes*, Springer-Verlag, New York, 1999.
[15] O. Hernández-Lerma and J. B. Lasserre, *Zero-sum stochastic games in Borel spaces: Average payoff criteria*, SIAM J. Control Optim., 39 (2001), pp. 1520–1539.
[16] A. Jaśkiewicz, *On strong* 1*-optimal policies in Markov control processes with Borel state spaces*, Bull. Polish Acad. Sci., 48 (2000), pp. 439–450.
[17] A. Jaśkiewicz, *Zero-sum semi-Markov games*, SIAM J. Control Optim., to appear.
[18] A. Jaśkiewicz and A. S. Nowak, *On the optimality equation for zero-sum ergodic stochastic games*, Math. Methods Oper. Res., 54 (2001), pp. 291–301.
[19] H.-U. Küenle, *Equilibrium strategies in stochastic games with additive cost and transition structure*, Int. Game Theory Rev., 1 (1999), pp. 131–147.
[20] H.-U. Küenle, *Stochastic games with complete information and average cost criterion*, Ann. Internat. Soc. Dynam. Games, 5 (2000) pp. 325–338.
[21] A. Maitra and W. D. Sudderth, *Borel stochastic games with limsup payoff*, Ann. Probab., 21 (1993), pp. 861–885.
[22] A. Maitra and W. D. Sudderth, *Discrete Gambling and Stochastic Games*, Springer-Verlag, New York, 1996.
[23] J.-F. Mertens and A. Neyman, *Stochastic games*, Internat. J. Game Theory, 10 (1981), pp. 53–66.
[24] J.-F. Mertens and T. Parthasarathy, *Equilibria for discounted stochastic games*, Research paper 8750, CORE, Université Catholique de Louvain, Louvain-la-Neuver, Belgium, 1987.
[25] S. P. Meyn and R. L. Tweedie, *Markov Chains and Stochastic Stability*, Springer-Verlag, New York, 1993.
[26] S. P. Meyn and R. L. Tweedie, *Computable bounds for geometric convergence rates of Markov chains*, Ann. Appl. Probab., 4 (1994), pp. 981–1011.
[27] S. P. Meyn and R. L. Tweedie, *State-dependent criteria for convergence of Markov chains*, Ann. Appl. Probab., 4 (1994), pp. 149–168.
[28] J. Neveu, *Mathematical Foundations of the Calculus of Probability*, Holden–Day, San Francisco, 1965.
[29] A. S. Nowak, *Existence of equilibrium stationary strategies in discounted noncooperative stochastic games with uncountable state space*, J. Optim. Theory Appl., 45 (1985), pp. 591–602.
[30] A. S. Nowak, *Stationary equilibria for nonzero-sum average payoff ergodic stochastic games with general state space*, Ann. Internat. Soc. Dynam. Games, 1 (1994), pp. 231–246.
[31] A. S. Nowak, *Zero-sum average payoff stochastic games with general state space*, Games Econom. Behav., 7 (1994), pp. 221–232.
[32] A. S. Nowak, *A generalization of Ueno's inequality for n-step transition probabilities*, Applicationes Mathematicae, 25 (1998), pp. 295–299.
[33] A. S. Nowak, *Sensitive equilibria for ergodic stochastic games with countable state spaces*, Math. Methods Oper. Res., 50 (1999), pp. 65–76.
[34] A. S. Nowak, *Optimal strategies in a class of zero-sum ergodic stochastic games*, Math. Methods Oper. Res., 50 (1999), pp. 399–420.
[35] A. S. Nowak and A. Jaśkiewicz, *Remarks on Sensitive Equilibria in Stochastic Games with Additive Reward and Transition Structure*, Technical report, Institute of Mathematics, University of Zielona Góra, Zielona Góra, Poland, 2002.
[36] A. S. Nowak and T. E. S. Raghavan, *Existence of stationary correlated equilibria with symmetric information for discounted stochastic games*, Math. Oper. Res., 17 (1992), pp. 519–526.
[37] T. Parthasarathy and S. Sinha, *Existence of stationary equilibrium strategies in non-zero-sum discounted stochastic games with uncountable state space and state independent transitions*, Internat. J. Game Theory, 18 (1989), pp. 189–194.
[38] O. Passchier, *The Theory of Markov Games and Queueing Control*, Ph.D. thesis, Department of Mathematics and Computer Science, Leiden University, Leiden, The Netherlands, 1998.
[39] U. Rieder, *Equilibrium plans for nonzero-sum Markov games*, in Game Theory and Related Topics, O. Moeschlin and D. Pallaschke, eds., North–Holland, Amsterdam, 1979, pp. 91–102.
[40] L. I. Sennott, *Nonzero-sum stochastic games with unbounded costs: discounted and average*

cost cases, Z. Oper. Res., 40 (1994), pp. 145–162.

[41] F. M. Spieksma, *Geometrically Ergodic Markov Chains and the Optimal Control of Queues*, Ph.D. thesis, Department of Mathematics and Computer Science, Leiden University, Leiden, The Netherlands, 1990.

[42] L. Stettner, *On nearly self-optimizing strategies for a discrete-time uniformly ergodic adaptive model*, Appl. Math. Optim., 27 (1993), pp. 161–177.

[43] J. Van der Wal, *Stochastic Dynamic Programming*, Mathematical Center Tracts 139, Mathematisch Centrum, Amsterdam, 1981.

[44] N. Vieille, *Equilibrium in* 2*-player stochastic games* 1: *A reduction*, Israel J. Math., 119 (2000), pp. 55–91.

[45] N. Vieille, *Equilibrium in* 2*-player stochastic games* 2: *The case of recursive games*, Israel J. Math., 119 (2000), pp. 93–126.

[46] T. Ueno, *Some limit theorems for temporally discrete Markov processes*, J. Fac. Sci. Univ. Tokyo, 7 (1957), pp. 449–462.

[47] J. Warga, *Optimal Control of Differential and Functional Equations*, Academic Press, New York, 1977.

[48] W. Whitt, *Representation and approximation of noncooperative sequential games*, SIAM J. Control Optim., 18 (1980), pp. 33–48.