

Nonstationary Zero Sum Stochastic Games with Incomplete Observation ¹

D. Leão, J.B.R. do Val
UNICAMP, Universidade de Campinas
Fac. of Electrical and Computer Eng.
Depto. de Telemática, C.P. 6101
13081-970, Campinas, SP, Brazil
jbosco@dt.fee.unicamp.br

M.D. Fragoso
LNCC/CNPq, Lab. Nac. de Comp. Científica
Department of Systems and Control
Av. Getúlio Vargas, 333 - C.P. 95113
25650-070 Petrópolis, RJ, Brazil
frag@alpha.lncc.br

Abstract

We study in this paper a discrete-time, two-players, zero-sum stochastic dynamic game similar to that introduced by Schäl (1981). The novelty here, regarding previous works, is that besides the fact that the players do not have access to the opponent's earlier choices, they do not have all information about the state of the game. The main results are a minimax theorem and the existence of optimal strategy for one player.

1 Introduction

We study a two-players, zero-sum discrete-time nonstationary stochastic game with Borel model data, where one player has nonempty compact metric action space. The information structure is a generalization of the stochastic game model introduced by Schäl (1981). Here, besides the fact that at any stage of the game neither player has any information about his opponent's earlier choices, they also do not necessarily have all information about the state of the game. The main contribution in this paper is to establish the value of the game, optimal strategy for one player and ϵ -optimal strategy for the other player.

First results for the game problem with uncountable state space were given in Maitra and Parthasarathy (1971), where it was solved the problem formulated in Shapley (1953) for compact state and actions spaces with discounted ($0 < \alpha < 1$) or positive ($\alpha = 1$) payoff. For some extensions of the results of Maitra and

Parthasarathy we refer to Kumar and Shiau (1981) where these results were extended to deal with Polish state spaces and compact action spaces.

Although the stochastic game literature forms a large body of theory nowadays, there are few works dealing with the incomplete observation case. We mention Haurie and Tolwinski (1990) in the study of cooperative games, for which the players have access to the system state but not to the move of the others players. We also mention Schäl (1981) and Nowak (1989) where the players do not have access to opponent's earlier choices and Bernhard (1992) where the observations are restricted to an output variable.

In the present approach, we use a suitable topology on the space of probabilities induced by the strategies introduced by Schäl (1975), which provides some regularity to the stochastic game. This regularity in conjunction with convexity of the class of probabilities induced by the strategies provide us with the elements to apply the minimax theorem of Sion (1958) and establish the main result.

2 Set up and main result

A Borel space Y is a nonempty Borel subset of a separable complete metric space endowed with its Borel σ -algebra $\beta(Y)$. We denote $P(Y)$ the set of all probability defined on $\beta(Y)$. Given two Borel spaces X and Y , we define a transition probability from X to Y as a mapping $f(\cdot | \cdot) : \beta(Y) \times X \rightarrow [0, 1]$, which satisfies the following conditions

a) $f(\cdot | x)$ is a probability on $\beta(Y)$, for each $x \in X$.

¹Research supported in part by CNPq, Grant No. 300721/86.2, Grant No. 520169/97.2 and PRONEX Grant No. 015/98 'Control of Dynamical Systems'.

b) $f(A | \cdot)$ is a measurable mapping on X , for each $A \in \beta(Y)$.

Consider Y_1, Y_2, \dots a sequence of Borel spaces. Throughout this paper we assume that the product space $Y_1 \times \dots \times Y_n$ and $Y_1 \times Y_2 \times \dots$ are given with their product topologies and product σ -algebras. It is well known that the product σ -algebra $\beta(Y_1 \times \dots \times Y_n)$ is identical to $\beta(Y_1) \times \dots \times \beta(Y_n)$. A similar result is also valid for the product space $Y_1 \times Y_2 \times \dots$, cf. Bertsekas and Shreve (1978, prop 7.13). For each $n \geq 2$, we denote a point of $Y_1 \times \dots \times Y_n$ by \mathbf{y}_n . We shall use the same notation for a random sequence. The set of real numbers will be denoted by \mathbb{R} and the set of real numbers augmented by the point $+\infty$ by $\overline{\mathbb{R}}$.

A zero-sum nonstationary stochastic game with incomplete observation is defined by a sequence of objects $\{\mathbb{S}_n, \mathbb{O}_n, \mathbb{A}_n, \mathbb{B}_n, P_n^1, P_n^2, u_n : n \in \mathbb{N}\}$, where:

i) \mathbb{S}_n is the state space at stage n . It is assumed to be a Borel space, for each $n \geq 1$.

ii) \mathbb{O}_n is a Borel space standing for the observable space, for each $n \geq 1$.

iii) \mathbb{A}_n and \mathbb{B}_n are the spaces of actions available at stage n for players I and II, respectively. We assume that \mathbb{A}_n is a Borel space and \mathbb{B}_n is a nonempty compact metric space endowed with the Borel σ -algebra.

Consider,

$$H_1 = \mathbb{S}_1 \times \mathbb{O}_1 \times \mathbb{A}_1 \times \mathbb{B}_1,$$

$$H_n = H_{n-1} \times \mathbb{S}_n \times \mathbb{O}_n \times \mathbb{A}_n \times \mathbb{B}_n; \quad n \geq 2,$$

$$H_\infty = \mathbb{S}_1 \times \mathbb{O}_1 \times \mathbb{A}_1 \times \mathbb{B}_1 \times \mathbb{S}_2 \times \mathbb{O}_2 \times \mathbb{A}_2 \times \mathbb{B}_2 \times \dots$$

Hence, H_n is the set of histories of the game for horizon $n \geq 1$, whereas H_∞ is the set of all infinite histories of the game.

iv) $\{P_n^1 : n \geq 1\}$ and $\{P_n^2 : n \geq 1\}$ are the laws of motion. At the initial stage, P_1^1 is a probability on $\beta(\mathbb{S}_1)$ and P_1^2 is a transition probability from \mathbb{S}_1 to \mathbb{O}_1 . Then, we can define a joint probability P_1 on $\beta(\mathbb{S}_1 \times \mathbb{O}_1)$ which satisfies

$$P_1(\mathbb{S}_1 \times \mathbb{O}_1) = \int_{\mathbb{S}_1} P_1^2(\mathbb{O}_1 | x_1) P_1^1(dx_1),$$

for any $\mathbb{S}_1 \in \beta(\mathbb{S}_1)$ and $\mathbb{O}_1 \in \beta(\mathbb{O}_1)$. In the sequence, for each $n \geq 2$,

$$\begin{cases} P_n^1(dx_n | h_{n-1}) \\ P_n^2(dz_n | h_{n-1}, x_n) \end{cases}$$

denote transition probabilities from H_{n-1} to \mathbb{S}_n and $H_{n-1} \times \mathbb{S}_n$ to \mathbb{O}_n , respectively. With these elements, we obtain

$$\begin{aligned} P_n(\mathbb{S}_n \times \mathbb{O}_n | h_{n-1}) \\ = \int_{\mathbb{S}_n} P_n^2(\mathbb{O}_n | h_{n-1}, x_n) P_n^1(dx_n | h_{n-1}), \end{aligned}$$

for all $\mathbb{S}_n \in \beta(\mathbb{S}_n)$, $\mathbb{O}_n \in \beta(\mathbb{O}_n)$ and $h_{n-1} \in H_{n-1}$. Then P_n is a transition probability from H_{n-1} to $\mathbb{S}_n \times \mathbb{O}_n$ for each $n \geq 2$.

We assume that $P_n(\cdot | h_{n-1})$ is dominated by a probability $\lambda_n \in \mathcal{P}(\mathbb{S}_n \times \mathbb{O}_n)$ for all $h_{n-1} \in H_{n-1}$ and $n \geq 2$. Consider the density $c_n : H_{n-1} \times \mathbb{S}_n \times \mathbb{O}_n \rightarrow \mathbb{R}$, e.g., the measurable function that satisfies

$$\begin{aligned} P_n(\mathbb{S}_n \times \mathbb{O}_n | h_{n-1}) \\ = \int_{\mathbb{S}_n} \int_{\mathbb{O}_n} c_n(h_{n-1}, x_n, z_n) \lambda_n(dx_n, dz_n), \end{aligned}$$

for all $\mathbb{S}_n \in \beta(\mathbb{S}_1)$, $\mathbb{O}_n \in \beta(\mathbb{O}_n)$ and $h_{n-1} \in H_{n-1}$ ($n \geq 2$). We also assume that the density c_n has the following property: for each $\mathbf{x}_n \in \mathbb{S}_1 \times \dots \times \mathbb{S}_n$, $\mathbf{z}_n \in \mathbb{O}_1 \times \dots \times \mathbb{O}_n$ and $\mathbf{a}_{n-1} \in \mathbb{A}_1 \times \dots \times \mathbb{A}_{n-1}$, the section

$$\mathbf{b}_{n-1} \longrightarrow c_n(\mathbf{x}_{n-1}, \mathbf{z}_{n-1}, \mathbf{a}_{n-1}, \mathbf{b}_{n-1}; x_n, z_n)$$

is lower semicontinuous on $\mathbb{B}_1 \times \dots \times \mathbb{B}_{n-1}$.

v) $u_n : H_n \rightarrow \overline{\mathbb{R}}$ is a measurable function bounded below, standing for the payoff function at stage $n \geq 1$. We assume that for each $\mathbf{x}_n, \mathbf{z}_n$ and \mathbf{a}_n , the section

$$\mathbf{b}_n \longrightarrow u_n(\mathbf{x}_n, \mathbf{z}_n, \mathbf{a}_n, \mathbf{b}_n)$$

is lower semicontinuous on $\mathbb{B}_1 \times \dots \times \mathbb{B}_n$. Each u_n may be recognized as a function on H_∞ , and we also assume that the sequence $\{u_n\}$ is nondecreasing and

$$u_n \uparrow u$$

is the infinite horizon payoff function.

With these objects, the game is played as follows: Given an initial probability P_1^1 on $\beta(\mathbb{S}_1)$, the game starts at state $x_1 \in \mathbb{S}_1$ and the observable component moves to $z_1 \in \mathbb{O}_1$ according to the probability $P_1^2(\cdot | x_1)$. The players observe z_1 and independently choose actions $a_1 \in \mathbb{A}_1$ and $b_1 \in \mathbb{B}_1$, respectively. Then the game moves to a new state $x_2 \in \mathbb{S}_2$ via the probability $P_2^1(\cdot | x_1, z_1, a_1, b_1)$, the players observe $z_2 \in \mathbb{O}_2$ according to the probability $P_2^2(\cdot | x_1, z_1, a_1, b_1; x_2)$ and,

independently, choose actions $a_2 \in \mathbb{A}_2$ and $b_2 \in \mathbb{B}_2$, respectively. Once more the game moves to a new state $x_3 \in \mathbb{S}_3$ according to the probability $P_3^1(\cdot | \mathbf{x}_2, \mathbf{z}_2, \mathbf{a}_2, \mathbf{b}_2)$ and so on. The result of such an infinite sequence of moves is a point $h \in H_\infty$ and the player II pays player I the amount $u(h)$. Note that at the n -th stage of the game, each player knows the sequence of observable components \mathbf{z}_n and his own previous choices. We denote:

$$\begin{aligned} H_1^1 &= H_1^2 = \mathbb{O}_1, \\ H_n^1 &= \mathbb{O}_1 \times \mathbb{A}_1 \times \mathbb{O}_2 \times \cdots \times \mathbb{A}_{n-1} \times \mathbb{O}_n; \quad n \geq 2 \\ H_n^2 &= \mathbb{O}_1 \times \mathbb{B}_1 \times \mathbb{O}_2 \times \cdots \times \mathbb{B}_{n-1} \times \mathbb{O}_n; \quad n \geq 2, \\ H_\infty^1 &= \mathbb{O}_1 \times \mathbb{A}_1 \times \mathbb{O}_2 \times \mathbb{A}_2 \times \cdots, \\ H_\infty^2 &= \mathbb{O}_1 \times \mathbb{B}_1 \times \mathbb{O}_2 \times \mathbb{B}_2 \times \cdots. \end{aligned}$$

Hence, H_n^1 and H_n^2 are the sets of histories of the game for players I and II, respectively.

Let $\Pi_n [\Delta_n]$ be the set of all transition probabilities π_n [δ_n] from H_n^1 to \mathbb{A}_n [H_n^2 to \mathbb{B}_n]. A strategy for player I [II] is a sequence $\pi = (\pi_n)$ [$\delta = (\delta_n)$], where $\pi_n \in \Pi_n$ [$\delta_n \in \Delta_n$] for all $n \geq 1$. Denote by $\Pi [\Delta]$ the set of all strategies for player I [II].

By the Ionescu-Tulcea theorem, cf. Neveu (1965, prop V.1.1), each pair of strategies $\pi = (\pi_n)$ and $\delta = (\delta_n)$, together with the law of motion $\{P_n : n \geq 1\}$, uniquely defines a probability $P_{\pi, \delta}$ on $\beta(H_\infty)$ such that, for every measurable function $g : H_n \rightarrow \overline{\mathbb{R}}$ bounded below, we have that

$$\begin{aligned} E_{\pi, \delta}[g] &= \int g(h) P_{\pi, \delta}(dh) \\ &= \int g(\mathbf{x}_n, \mathbf{z}_n, \mathbf{a}_n, \mathbf{b}_n) \pi_n(da_n | \mathbf{a}_{n-1}; \mathbf{z}_n) \cdot \\ &\delta_n(db_n | \mathbf{b}_{n-1}; \mathbf{z}_n) P_n(dx_n, dz_n | \mathbf{x}_{n-1}, \mathbf{z}_{n-1}, \mathbf{a}_{n-1}, \mathbf{b}_{n-1}) \cdots \\ &\cdots \pi_1(da_1 | z_1) \delta_1(db_1 | z_1) P_1(dx_1, dz_1), \end{aligned}$$

where g is regarded as a function on H_∞ . Thus, each pair of strategies defines an expected payoff

$$E_{\pi, \delta}[u_n] = \int u_n(h_n) P_{\pi, \delta}(dh_n); \quad n \geq 1.$$

Considering the assumption (v), it follows from monotone convergence theorem that for each $\pi \in \Pi$ and $\delta \in \Delta$

$$E_{\pi, \delta}[u] = \lim_{n \rightarrow \infty} E_{\pi, \delta}[u_n].$$

vi) We assume that

$$E_{\pi, \delta}[u] < \infty; \quad \forall \delta \in \Delta \text{ and } \pi \in \Pi.$$

Define,

$$v_* = \sup_{\pi \in \Pi} \inf_{\delta \in \Delta} E_{\pi, \delta}[u]$$

and

$$v^* = \inf_{\delta \in \Delta} \sup_{\pi \in \Pi} E_{\pi, \delta}[u].$$

If $v_* = v^* = v$ this common value is called *value* of the game. Suppose the value of the game exists. A strategy $\pi^* \in \Pi$ is called ε -optimal ($\varepsilon > 0$) for player I, if

$$\begin{aligned} \inf_{\delta \in \Delta} E_{\pi^*, \delta}[u] &> v - \varepsilon; \quad v < \infty; \\ \inf_{\delta \in \Delta} E_{\pi^*, \delta}[u] &> \frac{1}{\varepsilon}; \quad v = \infty. \end{aligned}$$

A strategy $\delta^* \in \Delta$ is called *optimal* for player II, if

$$\sup_{\pi \in \Pi} E_{\pi, \delta^*}[u] = v.$$

With these elements we can announce the main result of this paper.

Theorem 1 Assume (i)-(vi). The stochastic game has a value v . Player II has an optimal strategy whereas player I, for any $\varepsilon > 0$, has an ε -optimal strategy.

Remark 1: The information structure considered in this paper is a generalization of Schäl's stochastic game model, see Schäl (1981) and Nowak (1989). In these papers, the players choose actions independently, but they have full information on the state of the game. A further aspect is the optimality concept. Here, the initial distribution $P_1^1 \in \mathcal{P}(\mathbb{S}_1)$ is fixed, thus, the optimal (ε -optimal) strategies depend on P_1^1 . Such strategies are often called in the literature \bar{p} -optimal (or, \bar{p} - ε -optimal), cf. Hinderer (1970) or Schäl (1981).

Remark 2: In (iv), we give the transition probabilities P_n^1 and P_n^2 and then, we define the transition probabilities P_n . Conversely, if the transition probabilities P_n are given, we can determine P_n^1 as the marginal distribution of P_n :

$$P_n^1(S_n | h_{n-1}) = P_n(S_n \times \mathbb{O}_n | h_{n-1}),$$

for $S_n \in \beta(\mathbb{S}_n)$ and $h_{n-1} \in H_{n-1}$, and P_n^2 as the regular conditional probability, cf. Bertsekas and Shreve (1978, coro 7.27.1), which satisfies

$$P_n(S_n \times \mathbb{O}_n | h_{n-1}) = \int_{S_n} P_n^2(\mathbb{O}_n | h_{n-1}, x_n) P_n^1(dx_n | h_{n-1}),$$

for $h_{n-1} \in H_{n-1}$, $S_n \in \beta(\mathbb{S}_n)$ and $O_n \in \beta(\mathbb{O}_n)$.

Remark 3: The convergence assumption (v) includes the discounted and positive stochastic games, see Kumar and Shiau (1981). Optimal strategies for player I need not to exist even with finite action spaces, see Everett (1957) and Kumar and Shiau (1981).

Remark 4: In order to proof the minimax theorem, it is convenient to use that the set of probabilities

$$M_n = \{P_n(\cdot | h_{n-1}) : h_{n-1} \in H_{n-1}\} \subset P(\mathbb{S}_n \times \mathbb{O}_n),$$

for $n \geq 2$, are dominated by a probability $\lambda_n \in P(\mathbb{S}_n \times \mathbb{O}_n)$.

3 Proof of the main result

We shall prepare the proof of Lemmas 1 and 2 in the sequel, and introduce the extra elements in expressions (4), (5) and (7). For each pair of strategies $\pi \in \Pi$ and $\delta \in \Delta$, it follows from Fubini's theorem and (iv), that

$$\begin{aligned} & \int u_n(h_n) P_{\pi, \delta}(dh_n) = \\ & = \int u_n(h_n) \pi_n(da_n | \mathbf{a}_{n-1}; \mathbf{z}_n) \delta_n(db_n | \mathbf{b}_{n-1}; \mathbf{z}_n) \\ & \quad P_n(dx_n, dz_n | \mathbf{x}_{n-1}, \mathbf{z}_{n-1}, \mathbf{a}_{n-1}, \mathbf{b}_{n-1}) \cdots \\ & \quad \cdots \pi_1(da_1 | z_1) \delta_1(db_1 | z_1) P_1(dx_1, dz_1) \\ & = \int u_n(h_n) c_n(h_{n-1}; x_n, z_n) \\ & \quad \cdots c_2(h_1; x_2, z_2) \pi_n(da_n | \mathbf{a}_{n-1}; \mathbf{z}_n) \cdots \pi_1(da_1 | z_1) \\ & \quad \delta_n(db_n | \mathbf{b}_{n-1}; \mathbf{z}_n) \cdots \delta_1(db_1 | z_1) \lambda_n(dx_n, dz_n) \\ & \quad \cdots \lambda_2(dx_2, dz_2) P_1(dx_1, dz_1). \quad (1) \end{aligned}$$

Since for each $n \geq 1$, $(\mathbb{S}_n, \beta(\mathbb{S}_n))$ and $(\mathbb{O}_n, \beta(\mathbb{O}_n))$ are Borel spaces, there exist transition probabilities $\bar{\lambda}_n$ from \mathbb{O}_n to \mathbb{S}_n , cf. Bertsekas and Shreve (1978, corollary 7.27.1), which satisfies

$$P_1(S_1 \times O_1) = \int_{O_1} \bar{\lambda}_1(S_1 | z_1) \mu_1(dz_1), \quad (2)$$

where μ_1 is the marginal of P_1 on $(\mathbb{O}_1, \beta(\mathbb{O}_1))$, such that

$$\lambda_n(S_n \times O_n) = \int_{O_n} \bar{\lambda}_n(S_n | z_n) \mu_n(dz_n); \quad n \geq 2, \quad (3)$$

holds for each $S_n \in \beta(\mathbb{S}_n)$ and $O_n \in \beta(\mathbb{O}_n)$, where μ_n is the marginal of λ_n on $(\mathbb{O}_n, \beta(\mathbb{O}_n))$. Using equations

(1), (2) and (3), and Fubini's theorem, we have that

$$\begin{aligned} & \int u_n(h_n) P_{\pi, \delta}(dh_n) = \\ & = \int u_n(h_n) c_n(h_{n-1}; x_n, z_n) \cdots c_2(h_1; x_2, z_2) \\ & \quad \cdot \bar{\lambda}_n(dx_n | z_n) \cdots \bar{\lambda}_1(dx_1 | z_1) \\ & \quad \cdot \pi_n(da_n | \mathbf{a}_{n-1}; \mathbf{z}_n) \cdots \pi_1(da_1 | z_1) \\ & \quad \delta_n(db_n | \mathbf{b}_{n-1}; \mathbf{z}_n) \cdots \delta_1(db_1 | z_1) \mu_n(dz_n) \cdots \mu_1(dz_1). \end{aligned}$$

Hence, for each pair of strategies $\pi \in \Pi$ and $\delta \in \Delta$, we define

$$\begin{aligned} & P_\delta(O_1 \times B_1 \times \cdots \times O_n \times B_n) = \\ & \int \delta_n(B_n | \mathbf{b}_{n-1}; \mathbf{z}_n) \mu_n(dz_n) \cdots \delta_1(B_1 | z_1) \mu_1(dz_1), \end{aligned} \quad (4)$$

where $O_k \in \beta(\mathbb{O}_k)$ and $B_k \in \beta(\mathbb{B}_k)$ for each $1 \leq k \leq n$, and

$$\begin{aligned} & P_\pi(O_1 \times A_1 \times \cdots \times O_n \times A_n) = \\ & \int \pi_n(A_n | \mathbf{b}_{n-1}; \mathbf{z}_n) \mu_n(dz_n) \cdots \pi_1(A_1 | z_1) \mu_1(dz_1), \end{aligned} \quad (5)$$

where $O_k \in \beta(\mathbb{O}_k)$ and $A_k \in \beta(\mathbb{A}_k)$ for each $1 \leq k \leq n$. From Ionescu-Tulcea theorem, cf. Neveu (1965, prop V.1.1), there exists a unique probability P_δ (P_π) defined on $\beta(H_\infty^2)$ ($\beta(H_\infty^1)$) satisfying (4) (respectively, (5)). We denote

$$P_\Pi = \{P_\pi : \pi \in \Pi\} \quad \text{and} \quad P_\Delta = \{P_\delta : \delta \in \Delta\}.$$

For each $n \geq 1$, we denote

$$l_n(h_n) = u_n(h_n) c_n(h_{n-1}; x_n, z_n) \cdots c_2(h_1; x_2, z_2).$$

Hence, from (iv) and (v), for any \mathbf{x}_n , \mathbf{z}_n and \mathbf{a}_n , the section

$$\mathbf{b}_n \longrightarrow l_n(\mathbf{x}_n, \mathbf{z}_n, \mathbf{a}_n, \mathbf{b}_n)$$

is lower semicontinuous on $\mathbb{B}_1 \times \cdots \times \mathbb{B}_n$. Moreover, for any pair of strategies $\pi \in \Pi$ and $\delta \in \Delta$, we define

$$\begin{aligned} U_n[P_\pi, P_\delta] & = \int l_n(h_n) \bar{\lambda}_n(dx_n | z_n) \cdots \bar{\lambda}_1(dx_1 | z_1) \\ & \quad \delta_n(db_n | \mathbf{b}_{n-1}; \mathbf{z}_n) \cdots \delta_1(db_1 | z_1) P_\pi(d\mathbf{z}_n, d\mathbf{a}_n) \\ & = \int l_n(h_n) \bar{\lambda}_n(dx_n | z_n) \cdots \bar{\lambda}_1(dx_1 | z_1) \\ & \quad \pi_n(da_n | \mathbf{a}_{n-1}; \mathbf{z}_n) \cdots \pi_1(da_1 | z_1) P_\delta(d\mathbf{z}_n, d\mathbf{b}_n) \\ & = \int u_n(h_n) P_{\pi, \delta}(dh_n). \end{aligned}$$

Bearing in mind the monotone convergence theorem, we define

$$\begin{aligned} U[P_\pi, P_\delta] &= \lim_{n \rightarrow \infty} U_n[P_\pi, P_\delta] \\ &= \lim_{n \rightarrow \infty} \int u_n(h_n) P_{\pi, \delta}(dh_n) = \int u(h) P_{\pi, \delta}(dh). \end{aligned} \quad (6)$$

Thus, as a consequence of expressions (4), (5) and (6), we consider the following lemma, as established by Nowak (1989, lem 5).

Lemma 1 *We have that*

$$v^* = \inf_{\delta \in \Delta} \sup_{\pi \in \Pi} E_{\pi, \delta}[u] = \inf_{P_\delta \in P_\Delta} \sup_{P_\pi \in P_\Pi} U[P_\pi, P_\delta]$$

and

$$v_* = \sup_{\pi \in \Pi} \inf_{\delta \in \Delta} E_{\pi, \delta}[u] = \sup_{P_\pi \in P_\Pi} \inf_{P_\delta \in P_\Delta} U[P_\pi, P_\delta].$$

For each pair of strategies $\pi \in \Pi$ and $\delta \in \Delta$, we denote

$$\begin{aligned} g_n(\mathbf{z}_n, \mathbf{b}_n) &= \int l_n(h_n) \bar{\lambda}_n(dx_n | z_n) \cdots \bar{\lambda}_1(dx_1 | z_1) \\ &\quad \pi_n(da_n | \mathbf{a}_{n-1}; \mathbf{z}_n) \cdots \pi_1(da_1 | z_1). \end{aligned}$$

Then, it follows from (6) that for each pair of strategies $\pi \in \Pi$ and $\delta \in \Delta$

$$\begin{aligned} U[P_\pi, P_\delta] &= \lim_{n \rightarrow \infty} U_n[P_\pi, P_\delta] \\ &= \lim_{n \rightarrow \infty} \int u_n(h_n) P_{\pi, \delta}(dh_n) = \lim_{n \rightarrow \infty} \int g_n dP_\delta, \end{aligned} \quad (7)$$

for each $n \geq 1$.

Lemma 2 *The function $g_n : H_n^2 \times \mathbb{B}_n \rightarrow \bar{\mathbb{R}}$ above is measurable, bounded below and for each $\mathbf{z}_n \in \mathbb{O}_1 \times \cdots \times \mathbb{O}_n$, the section $\mathbf{b}_n \rightarrow g_n(\mathbf{z}_n, \mathbf{b}_n)$ is lower semicontinuous on $\mathbb{B}_1 \times \cdots \times \mathbb{B}_n$, for all $n \geq 1$.*

Proof: For each $n \geq 1$, it is clear that g_n is measurable and bounded below. If we fix $\mathbf{z}_n \in \mathbb{O}_1 \times \cdots \times \mathbb{O}_n$, it follows from Bertsekas and Shreve (1978, prop 7.31) that the section $\mathbf{b}_n \rightarrow g_n(\mathbf{z}_n, \mathbf{b}_n)$ is lower semicontinuous. \square

With elements (4), (5) and (7) we have a structure suitable for the minimax theorem of Sion (1958), also similar to that in Schäl (1981). In order to apply the minimax theorem, we need a topology on $P(H_\infty^2)$ such that P_Δ is compact and for any $P_\pi \in P_\Pi$ the section

$$P_\delta \longrightarrow U[P_\pi, P_\delta]; \quad P_\delta \in P_\Delta$$

is lower semicontinuous on P_Δ . Consider (Ω, F) a Borel space and (E, E) a compact metric space endowed with the Borel σ -algebra. We denote $B(\Omega)$ the set of all bounded measurable functions $f : \Omega \rightarrow \mathbb{R}$ and $\bar{B}(\Omega)$ for the set of all measurable functions $g : \Omega \rightarrow \bar{\mathbb{R}}$ which are bounded below. Moreover, $\mathcal{C}(E)$ will denote the subset of all continuous functions in $B(E)$ and $\bar{\mathcal{C}}(E)$ is the subset of all lower semicontinuous function in $\bar{B}(E)$. Finally, we write

$$\begin{aligned} Q(\Omega \times E) &= \{f \in B(\Omega \times E) : f(w, \cdot) \in \mathcal{C}(E), w \in \Omega\}, \\ \bar{Q}(\Omega \times E) &= \{f \in \bar{B}(\Omega \times E) : \\ &\quad \text{there exists } (f_n) \subset Q(\Omega \times E) \text{ such that } f_n \uparrow f\}. \end{aligned}$$

A result for the proof of the main result states a characterization of the elements of $\bar{Q}(\Omega \times E)$. The following statements are equivalent, cf. Schäl (1979, eq 4.1):

- a) $g \in \bar{B}(\Omega \times E)$ and the section $g(w, \cdot) \in \bar{\mathcal{C}}(E)$, for all $w \in \Omega$.
- b) $g \in \bar{Q}(\Omega \times E)$.

For any $n \geq 1$, we consider

$$\Omega = \mathbb{O}_1 \times \mathbb{O}_2 \times \cdots \times \mathbb{O}_n; \quad E = \mathbb{B}_1 \times \mathbb{B}_2 \times \cdots \times \mathbb{B}_n.$$

Thus, $f \in Q(H_n^2 \times \mathbb{B}_n)$ if and only if $f \in B(H_n^2 \times \mathbb{B}_n)$ and the sections

$$f(\mathbf{z}_n, \cdot) \in C(\mathbb{B}_1 \times \cdots \times \mathbb{B}_n),$$

for all $\mathbf{z}_n \in \mathbb{O}_1 \times \cdots \times \mathbb{O}_n$. Similarly, $f \in \bar{Q}(H_n^2 \times \mathbb{B}_n)$ if and only if there exists a sequence of functions $(f_n) \subset Q(H_n^2 \times \mathbb{B}_n)$ such that

$$f_n \uparrow f; \quad (n \rightarrow \infty).$$

The ws^∞ -topology on $P(H_\infty^2)$ is defined as the coarsest topology for which the mappings

$$\mu \longrightarrow \int f d\mu; \quad f \in \bigcup_{n \geq 1} \bar{Q}(H_n^2 \times \mathbb{B}_n)$$

are lower semicontinuous on $P(H_\infty^2)$.

Lemma 3 (Schäl 1979, theo 7.11) *The subsets $P_\Pi \subset P(H_\infty^1)$ and $P_\Delta \subset P(H_\infty^2)$ are convex. In other words, for each $0 < \lambda < 1$ and $\pi^1, \pi^2 \in P_\Pi$,*

$$\lambda P_{\pi^1} + (1 - \lambda) P_{\pi^2} \in P_\Pi.$$

Lemma 4 Let $P(H_\infty^2)$ be endowed with the ws^∞ -topology. Then, for each $P_\pi \in \Pi$, the section

$$P_\delta \longrightarrow U[P_\pi, P_\delta]; \quad P_\delta \in P_\Delta$$

is lower semicontinuous on P_Δ . Moreover, the set P_Δ is compact.

Proof: Since $P(H_\infty^2)$ is endowed with the ws^∞ -topology, and for all $n \in \mathbb{N}$:

- a. \mathbb{B}_n is compact;
- b. μ_n is independent of \mathbf{b}_n , then s -continuous on $\mathbb{B}_1 \times \dots \times \mathbb{B}_n$;
- c. $g_n \in \overline{Q}(H_n^2 \times \mathbb{B}_n)$ (Lemma 2).

We have only to apply Schäl (1975, theo 6.6) and equation (7). \square

Lemma 5 Consider $P_\pi \in P_\Pi$ and $P_\delta \in P_\Delta$. Then, the sections $U[\cdot, P_\delta]$ and $U[P_\pi, \cdot]$ are affine on P_Π and P_Δ , respectively.

Proof: It follows from Lemma 3 and equation (7).

Proof of the Theorem: It follows from Lemmas 1, 4 and 5, that we can apply the minimax theorem of Sion (1958, theo 4.2') to obtain that the zero sum game has a value, namely,

$$\sup_{P_\pi \in P_\Pi} \inf_{P_\delta \in P_\Delta} U[P_\pi, P_\delta] = \inf_{P_\delta \in P_\Delta} \sup_{P_\pi \in P_\Pi} U[P_\pi, P_\delta] = v.$$

Moreover, using semicontinuity and compactness arguments, there exists $\psi \in P_\Delta$ such that

$$\sup_{P_\pi \in P_\Pi} U[P_\pi, \psi] = \inf_{P_\delta \in P_\Delta} \sup_{P_\pi \in P_\Pi} U[P_\pi, P_\delta] = v.$$

Recalling to the definition of supremum, for all $\varepsilon > 0$ there exists an strategy $\theta_\varepsilon \in P_\Pi$ which satisfies

$$\begin{cases} \inf_{P_\delta \in P_\Delta} U[\theta_\varepsilon, P_\delta] > v - \varepsilon, & \text{if } v < \infty; \\ \inf_{P_\delta \in P_\Delta} U[\theta_\varepsilon, P_\delta] > \frac{1}{\varepsilon}, & \text{if } v = \infty. \end{cases}$$

Then, each $\delta^* \in \{\delta \in \Delta : P_\delta = \psi\}$ is an optimal strategy for player II and each $\pi^* \in \{\pi \in \Pi : P_\pi = \theta_\varepsilon\}$ is an ε -optimal strategy for player I. \square

With the ws^∞ -topology on $P(H_\infty^2)$ we obtained the regularity properties (semicontinuity and compactness) necessary to apply the minimax theorem of Sion.

References

- [1] Bertsekas D. P., Shreve S. E., (1978) Stochastic Optimal Control: The discrete-Time Case, Academic Press, New York.
- [2] Bernhard P. (1992) Information and Strategies in Dynamic Games, *SIAM J. Control and Op.* **32** 212–228.
- [3] Everett (1957) Recursive Games, *Ann. of Math. Stud.* **39** 47-78
- [4] Haurie A. and Tolwinski B. (1990) Cooperative Equilibria in Discounted Sequential Games, *J. of Optimization Theory and Appl.* **64** 511–535.
- [5] Hinderer K. (1970) Foundations of Nonstationary Dynamic Programming with Discrete-Time Parameter, Springer-Verlag, New York
- [6] Kumar P. R., Shiau T. H. (1981) Existence of Value and Randomized Strategies in Zero-Sum Discrete-time Stochastic Dynamic Games, *SIAM J. on Control and Opt.* **19** 617–634.
- [7] Maitra A., Parthasarathy T. (1971) On Stochastic Games II, *J. of Optimization Theory and Appl.* **8** 154–160.
- [8] Neveu J. (1965) Mathematical Foundations of the Calculus of Probability, Holden-Day, San Francisco.
- [9] Nowak A. S. (1989) Existence of optimal strategies in zero-sum nonstationary stochastic games with lack of information on both sides, *SIAM J. Control and Opt.* **27**
- [10] Schäl M. (1975) On dynamic programming: compactness of the space of policies, *Stochastic Process and Applications* **3** 345-364
- [11] Schäl M. (1979) On dynamic programming and statistical decision theory, *Annals Statistic* **7** 432-445
- [12] Schäl M. (1981) Stochastic nonstationary two-person zero-sum games, *Z. Angew. Math. Mech.* **61** 352-353
- [13] Sion M.(1958) On general minimax theorem, *Pacific Journal Math.* **VIII** 171-176
- [14] Shapley L. S. (1953) Stochastic Games, Proc. Nat. Acad. Sci. USA **39** 1095–1100.