

Stochastic Games with a Single Controller and Incomplete Information

Dinah Rosenberg*, Eilon Solan[†] and Nicolas Vieille[‡] §

May 6, 2002

Abstract

We study stochastic games with incomplete information on one side, where the transition is controlled by one of the players.

We prove that if the informed player also controls the transition, the game has a value, whereas if the uninformed player controls the transition, the max-min value, as well as the min-max value, exist, but they may differ.

We discuss extensions to the case of incomplete information on both sides.

*Laboratoire d'Analyse Géométrie et Applications, Institut Galilée, Université Paris Nord, avenue Jean-Baptiste Clément, 93430 Villetaneuse, France. e-mail: dinah@zeus.math.univ-paris13.fr

[†]MEDS Department, Kellogg School of Management, Northwestern University, *and* the School of Mathematical Sciences, Tel Aviv University, Tel Aviv 69978, Israel. e-mail: eilons@post.tau.ac.il

[‡]Département Finance et Economie, HEC, 1, rue de la Libération, 78 Jouy-en-Josas, France. e-mail: vieille@hec.fr

§We acknowledge the financial support of the Arc-en-Ciel/Keshet program for 2001/2002. The research of the second author was supported by the Israel Science Foundation (grant No. 03620191).

1 Introduction

In a seminal work, Aumann and Maschler (1968, 1995) introduced infinitely repeated two player zero-sum games with incomplete information on one side. Those are repeated games where the payoff matrix is known by one player, say player 1, but is not known by the other player: all player 2 knows is that the payoff matrix was drawn according to some known probability distribution from a finite set of possible matrices.

Aumann and Maschler proved that those games have a value.

The issue faced by player 1 is the optimal use of information. On the one hand, player 1 needs to reveal his information (at least partially) in order to make use of it. On the other hand, any piece of information that is revealed to player 2 can later be exploited against player 1.

In the strategies devised by Aumann and Maschler player 1 reveals part of his information at the first stage, but no further information is revealed along the game.

When the underlying game is a stochastic game rather than a repeated game, the difficulties the players face are more serious. Is it optimal for player 1 to reveal information only once in every state, or will he reveal information several times in each state? Player 1 has no incentive to reveal information little by little in repeated games since a reply of player 2 could always be to wait until player 1 reveals all the information he will ever reveal, and interim payoffs are irrelevant in the long-run. In stochastic games, on the contrary, the game can move meanwhile to a different state that can be more or less favorable to the informed player. This is why player 1 might be willing to reveal his information little by little in stochastic games: he will have more opportunities to get different transitions according to his information while player 2 is still ignorant of it.

Player 2, on the other hand, has to play optimally whatever be the actual payoff matrix. In Aumann and Maschler, he plays a Blackwell approachability strategy. The issue is here to define the analogue for stochastic games.

Sorin (1984, 1985) and Sorin and Zamir (1991) studied classes of stochastic games with incomplete information on one side that have a single non-absorbing state, and proved that these games have a min-max value, a max-min value, and that the values of the n -stage (resp. λ -discounted) games converge as n goes to infinity (resp. as λ goes to 0) to the max-min value. Rosenberg and Vieille (2000) studied recursive games with incomplete information on one side, and proved that the max-min value exists, and is equal to the limit of the values of n -stage games (resp. λ -discounted games) as n goes to infinity (resp. as λ goes to 0).

In the present paper we study stochastic games in which one player controls the transition; that is, the evolution of the stochastic state depends on the actions of one player, but is independent of the actions of his opponent.

We show that if player 1 (who is the informed player) controls the transition then the game admits a value, while if player 2 controls the transition then the game admits a min-max value and a max-min value, but the two may differ.

The techniques and the characterizations provided extend the ideas of Aumann and Maschler for incomplete information games to our framework.

In the last section of the paper we extend the existence result for the max-min value and the min-max value to the case of stochastic games with a single controller and incomplete information on both sides; that is, when each of the players has some partial private information about the payoff matrix of the game.

2 The Model and the Main Results

2.1 The Model

A *two-player zero-sum stochastic game* G is described by: (i) a finite set Ω of states, and an initial state $\omega \in \Omega$, (ii) finite action sets I and J for the two players, (iii) a transition rule $q : \Omega \times I \times J \rightarrow \Delta(\Omega)$, where $\Delta(\Omega)$ is the simplex of probability distributions over Ω , and (iv) a reward function $g : \Omega \times I \times J \rightarrow \mathbf{R}$.

A *two-player zero-sum stochastic game with incomplete information* is described by a finite collection $(G_k)_{k \in K}$ of stochastic games, together with a distribution $p \in \Delta(K)$ over K . We assume that the games G_k differ only through their reward functions g^k , but they all have the same sets of states and actions, and the same transition rule. We denote the common transition rule by q .

The game is played in stages. An element $k \in K$ is chosen according to p . Player 1 is informed of k , while player 2 is not. At every stage n , the two players choose simultaneously actions $i_n \in I$ and $j_n \in J$, and ω_{n+1} is drawn according to $q(\cdot \mid \omega_n, i_n, j_n)$. Both players are informed of (i_n, j_n, ω_{n+1}) .

We parametrize the game by the initial distribution p and by the initial state ω , and denote it by $\Gamma(p, \omega)$. We write Γ for $(\Gamma(p, \omega))_{(p, \omega) \in \Delta(K) \times \Omega}$.

Few remarks are in order. This model is an extension of the classical model of zero-sum stochastic games. It is also an extension of Aumann and Maschler's model of repeated games with incomplete information, where a zero-sum matrix game is first drawn using p , then played repeatedly over time. Here, Nature chooses a stochastic game, that is then played over time. Note that the reward function $g^k(\omega_n, i_n, j_n)$ is not told to player 2 (but is known to player 1).

We assume w.l.o.g. that $0 \leq g^k \leq 1$ for every $k \in K$, and we identify each $k \in K$ with the probability measure over K that gives weight 1 to k .

2.2 Strategies and values

Players may base their choices on the stochastic states the play has visited so far, as well as on past choices of actions (of the two players). Player 1 can base his choices also on the state of the world k .

The space of histories of length n is $H_n = (\Omega \times I \times J)^n \times \Omega$, the space of finite histories is $H = \cup_{n \in \mathbf{N}} H_n$, and the space of plays (infinite histories) is $H_\infty = (\Omega \times I \times J)^\infty$. H_n defines naturally a finite algebra \mathcal{H}_n over H_∞ . We equip H_∞ with the σ -algebra $\vee_{n \in \mathbf{N}} \mathcal{H}_n$ spanned by all finite cylinders. A (behavioral) *strategy* of player 1 is a function $\sigma : K \times H \rightarrow \Delta(I)$. A strategy for player 2 is a function $\tau : H \rightarrow \Delta(J)$. A strategy $\sigma = (\sigma_k)_{k \in K}$ of player 1 is *non revealing* if σ_k is independent of $k \in K$.¹

A strategy σ is stationary if the mixed action played at every stage depends only on the current state. We identify each vector $x = (x_\omega)_{\omega \in \Omega} \in (\Delta(I))^\Omega$ with the stationary strategy that plays the mixed action x_ω whenever the game visits ω . Stationary strategies of player 2 are defined analogously.

Every distribution p , initial stochastic state ω , and pair of strategies (σ, τ) induce a probability $\mathbf{P}_{p, \omega, \sigma, \tau}$ over $K \times H_\infty$ (equipped with the product σ -algebra). We denote by $\mathbf{E}_{p, \omega, \sigma, \tau}$ the corresponding expectation operator.

¹The strategy is non revealing in the sense that knowledge of the strategy σ and of past play does not enable player 2 to gain information on k . This property relies on the fact that transitions are independent of k .

We let k, ω_n, i_n, j_n denote respectively the actual game being played, the current state at stage n and the actions played at stage n . These are random variables.

Define

$$\gamma_N(p, \omega, \sigma, \tau) = \mathbf{E}_{p, \omega, \sigma, \tau} [\bar{g}_N],$$

where $\bar{g}_N = \frac{1}{N} \sum_{n=1}^N g^k(\omega_n, i_n, j_n)$ is the average payoff over the first N periods. For fixed strategies σ, τ , $\gamma_N(p, \omega, \sigma, \tau)$ is linear in p , and 1-Lipshitz.

We recall the definitions of max-min value, the min-max value and the (uniform) value. The notion of strong guaranteeing is non-standard.

Definition 1 *Player 1 can guarantee $w \in \mathbf{R}$ in the game $\Gamma(p, \omega)$ if for every $\epsilon > 0$ there exists a strategy σ of player 1 and $N \in \mathbf{N}$, such that*

$$\forall \tau, \forall n \geq N, \gamma_n(p, \omega, \sigma, \tau) \geq w - \epsilon.$$

We say that such a strategy σ guarantees $w - \epsilon$ in $\Gamma(p, \omega)$.

Player 1 can guarantee a function $w : \Delta(K) \times \Omega \rightarrow \mathbf{R}$ if player 1 can guarantee $w(p, \omega)$ in the game $\Gamma(p, \omega)$ for every $(p, \omega) \in \Delta(K) \times \Omega$.

Note that, due to the Lipshitz property on payoffs and the compactness of $\Delta(K)$, the integer N in Definition 1 can be chosen to be independent of (p, ω) . The definition of a function that is guaranteed by player 2 is similar, with the roles of the two players exchanged.

Definition 2 *Player 2 can defend $w \in \mathbf{R}$ in the game $\Gamma(p, \omega)$ if for every $\epsilon > 0$ and every strategy σ of player 1, there exists a strategy τ of player 2 and $N \in \mathbf{N}$ such that*

$$\forall n \geq N, \gamma_n(p, \omega, \sigma, \tau) \leq w + \epsilon. \quad (1)$$

We say that such a strategy τ defends $w + \epsilon$ against σ in $\Gamma(p, \omega)$.

Player 2 can defend a function $w : \Delta(K) \times \Omega \rightarrow \mathbf{R}$ if player 2 can defend $w(p, \omega)$ in the game $\Gamma(p, \omega)$ for every $(p, \omega) \in \Delta(K) \times \Omega$.

The definition of a function that is defended by player 1 is similar, with the roles of the two players exchanged. Note that player 1 can guarantee (resp. defend) $\max\{w, w'\}$ as soon as he can guarantee (resp. defend) both w and w' . Similarly, player 2 can guarantee (resp. defend) $\min\{w, w'\}$ as soon as he can guarantee (resp. defend) both w and w' .

Definition 3 *A function $w : \Delta(K) \times \Omega \rightarrow \mathbf{R}$ is:*

- *the (uniform) value of Γ if both players can guarantee w .*
- *the max-min value of Γ if player 1 can guarantee w , and player 2 can defend w .*
- *the min-max value of Γ if player 1 can defend w , and player 2 can guarantee w .*

Note that the value exists if, and only if, the max-min value and min-max value exist and coincide.

The value (*resp.* max-min value, min-max value) is denoted by v (*resp.* \underline{v}, \bar{v}) when it exists. Observe that $\underline{v} \leq \bar{v}$ whenever the two exist. Note that each of the functions \underline{v} and \bar{v} is 1-Lipshitz in p , as soon as it exists.

2.3 Related literature

Most of the literature deals with the polar cases where either Ω or K is a singleton. In the former case, the game is a repeated game with incomplete information. Such games have a value, see Aumann and Maschler (1995). Moreover, an explicit formula for the value exists. Letting $u^*(p)$ be the value of the matrix game with payoff function $\sum_k p_k g^k(\cdot, \cdot)$, the value of the repeated game with incomplete information is the concavification $\text{cav}(u^*)$ of u^* (see Section 3.1 for definitions).

When K is a singleton the game is a standard stochastic game. Such games have a value, see Mertens and Neyman (1981).

For general stochastic games with incomplete information, little is known, but some classes were studied in the literature. For “Big Match” games, Sorin (1984, 1985), Sorin and Zamir (1991) proved the existence of the max-min value and min-max value. These values may differ.

For recursive games, Rosenberg and Vieille (2000) proved that the max-min value exists, and provided an example where the value does not exist.

2.4 Statements of the results

In the present paper we consider games where a single player controls the transition.

Definition 4 *Player 1 controls the transition if for every $\omega \in \Omega$ and $i \in I$ the transition $q(\cdot | \omega, i, j)$ does not depend on j . Player 2 controls the transition if the symmetric property holds. We then simply write $q(\cdot | \omega, i)$ or $q(\cdot | \omega, j)$ depending on who controls transitions.*

We prove the following two results.

Theorem 5 *If player 1 controls the transition, the value exists.*

Theorem 6 *If player 2 controls the transition, both the min-max and max-min values exist.*

We provide an example of a game where player 2 controls the transition, and $\bar{v} \neq \underline{v}$. We also provide a characterization of \bar{v} and \underline{v} as a unique solution of a functional equation.

We prove no result on the existence of the limit of the values of the finitely repeated games. In the games analyzed so far (see Section 2.3), this limit is known to exist, and coincides with \underline{v} . This property is conjectured to hold in general by Mertens (1987).

3 Various tools

This section gathers a few results that we use in the sequel. The first three subsections introduce few extensions of tools used in the analysis of games with incomplete information.

For three vectors $a, b, c \in \mathbf{R}^K$, $c = a + b$ if and only if $c_k = a_k + b_k$ for every $k \in K$, $c = \max\{a, b\}$ if and only if $c_k = \max\{a_k, b_k\}$ for every $k = 1, \dots, K$, and $a \geq b$ if and only if $a_k \geq b_k$ for every $k = 1, \dots, K$. For a scalar $r \in \mathbf{R}$, $c = a + r$ if and only if $c_k = a_k + r$ for every $k = 1, \dots, K$, and $c = ra$ if and only if $c_k = ra_k$ for every $k = 1, \dots, K$. Finally, unless otherwise stated, the norm we use is the uniform norm.

3.1 Concavification

Given a continuous function $u : \Delta(K) \rightarrow \mathbf{R}$, we denote by $\text{cav } u$ its concavification, namely the least concave function v defined over $\Delta(K)$, such that $v \geq u$. It is the function whose hypograph is the convex hull of the hypograph of u . Similarly, we denote by $\text{vex } u$ its convexification, namely the largest convex function v such that $v \leq u$. Both $\text{cav } u$ and $\text{vex } u$ are well-defined. Thus, cav and vex are functional operators that act on real-valued functions defined on $\Delta(K)$.

Lemma 7 (see, e.g., Laraki (2001)). *The two operators cav and vex map continuous functions into continuous functions, and C -Lipshitz functions into C -Lipshitz functions.*

Lemma 8 *The two operators cav and vex are non-expansive.*

Proof. It is proven in De Meyer (1996, Lemma 2.1) that for any two real valued continuous functions over $\Delta(K)$, u and v ,

$$\|u^{**} - v^{**}\| \leq \|u^* - v^*\| \leq \|u - v\|,$$

where $u^*(x) = \inf\{\langle y, x \rangle - u(y), y \in \mathbf{R}^K\}$ is the dual of u . Since $u^{**} = \text{cav}(u)$, the result follows.

The argument for the operator vex is analogous. ■

The following lemma is classical (see, e.g., Mertens, Sorin and Zamir (1994, Corollary V.1.3), or the discussion in Zamir (1992, p.118)).

Lemma 9 *Assume that player 1 can guarantee u . Then player 1 can guarantee $\text{cav } u$.*

The following result will be useful later.

Lemma 10 *Let $(\mathcal{A}_i)_{i \in I}$ be a finite collection of convex closed upwards comprehensive sets and let \mathcal{A} be the set $\{a \in \mathbf{R}^K \mid a = \max_{i \in I} a_i, a_i \in \mathcal{A}_i\}$. Then*

$$f_{\mathcal{A}}(p) = (\text{cav } \max_{i \in I} f_{\mathcal{A}_i})(p),$$

where, for any convex upwards comprehensive set \mathcal{B} , $f_{\mathcal{B}}(p) = \inf_{a \in \mathcal{B}} \langle a, p \rangle$.

Proof. Since each \mathcal{A}_i is upwards comprehensive, \mathcal{A} coincides with $\cap_i \mathcal{A}_i$. Therefore $f_{\mathcal{A}} \geq f_{\mathcal{A}_i}$ for each i . In particular $f_{\mathcal{A}} \geq \max_{i \in I} f_{\mathcal{A}_i}$. Since $f_{\mathcal{A}}$ is concave, $f_{\mathcal{A}} \geq \text{cav } \max_{i \in I} f_{\mathcal{A}_i}$.

To prove the opposite inequality, we first observe that if \mathcal{B} is convex, closed and upwards comprehensive, one has

$$\mathcal{B} = \{a \in \mathbf{R}^K \mid \langle a, p \rangle \geq f_{\mathcal{B}}(p) \text{ for each } p \in \Delta(K)\}. \quad (2)$$

Set $g = \text{cav } \max_{i \in I} f_{\mathcal{A}_i}$, and

$$\mathcal{D} = \{a \in \mathbf{R}^K \mid \langle a, p \rangle \geq g(p) \text{ for each } p \in \Delta(K)\}.$$

Since $g \geq f_{\mathcal{A}_i}$ for each $i \in I$, and using (2) with $\mathcal{B} = \mathcal{A}_i$, one has $\mathcal{D} \subseteq \mathcal{A}_i$. Therefore, $\mathcal{D} \subseteq \mathcal{A}$ which readily implies $g \geq f_{\mathcal{A}}$. ■

3.2 Approachability

We present here the basic approachability result of Blackwell (1956), in the framework of stochastic games. Let G be a stochastic game with payoffs in \mathbf{R}^K . The description of such a game is the same as that of a zero-sum stochastic game given in Section 2.1, except that the reward function now takes values in \mathbf{R}^K . The definition of strategies in this framework is similar to that given in Section 2.2.

We denote $\bar{g}_N = \frac{1}{N} \sum_{n=1}^N g(\omega_n, i_n, j_n) \in \mathbf{R}^K$, the average vector payoff in the first N stages.

Definition 11 *A vector $a \in \mathbf{R}^K$ is approachable by player 2 at ω if for every $\varepsilon > 0$, there is a strategy τ of player 2 and $N \in \mathbf{N}$ such that:²*

$$\forall \sigma, \mathbf{E}_{\omega, \sigma, \tau} \left[\sup_{n \geq N} (\bar{g}_n - a)^+ \right] \leq \varepsilon.$$

We say that such a strategy τ approaches $a + \varepsilon$ at ω .

In words, for every ε player 2 has a strategy such that, the average payoff vector will eventually not exceed $a + \varepsilon$. Note that a is approachable if and only if $a + \varepsilon$ is approachable for every $\varepsilon > 0$, so that the set of approachable vectors is closed and upwards comprehensive.

Our definition slightly differs from that of Blackwell (1956), where the strategy τ is required to be independent of ε (*i.e.*, the original definition of Blackwell reads as: $\exists \tau, \forall \varepsilon > 0$, etc.). Any vector a that is approachable in Blackwell's sense is also approachable in our sense. The two definitions are not equivalent. However, it is easily checked that, if a is approachable (in our sense) at each state, it is also approachable in Blackwell's sense.

Every stochastic game with incomplete information $\Gamma(p, \omega)$ induces a stochastic game with vector payoffs $\Gamma^V(\omega)$, in which the payoff coordinates are given by the payoff functions of the component games (G_k) of $\Gamma(p, \omega)$.

The following two Lemmas relate approachable vectors in Γ^V to quantities in $\Gamma(p, \omega)$. The first one is immediate.

Lemma 12 *If $a \in \mathbf{R}^K$ is approachable at ω in the game Γ^V , then player 2 can guarantee $\langle a, p \rangle$ in $\Gamma(p, \omega)$ for each $p \in \Delta(K)$.*

We now state Blackwell's sufficient condition for approachability in this context. Denote by $u_\infty(p, \omega)$ the uniform value of the zero-sum stochastic game with reward function $\sum_{k \in K} p_k g^k(\omega, \cdot, \cdot)$. The existence of u_∞ follows by Mertens and Neyman (1981). We also denote by $u_n(p, \omega)$ the value of the n -stage version of that game (thus, $\lim_{n \rightarrow \infty} u_n = u_\infty$ and the limit is uniform in p).

Proposition 13 *If $\text{cav } u_\infty(p, \omega) \leq \langle a, p \rangle$ for every $(p, \omega) \in \Delta(K) \times \Omega$, then a is approachable in Γ^V by player 2 at ω , for each $\omega \in \Omega$.*

In this statement (and in later ones), $\text{cav } u_\infty$ is the concavification of u_∞ with respect to the first variable, p : $\text{cav } u_\infty(p, \omega) = (\text{cav } u_\infty(\cdot, \omega))(p)$.

²For every real $a \in \mathbf{R}$, $a^+ = \max\{a, 0\}$.

Sketch of the proof: let $\varepsilon > 0$ and choose N such that $\|u_N - u_\infty\| \leq \varepsilon$. We then view successive blocks of N stages as successive stages in the repetition of the N -stage game. We apply directly Blackwell's result (noting that Blackwell's proof still holds when the stage game changes from stage to stage, with payoffs remaining bounded).

A more general result was proved by Milman (2000, Theorem 2.1.1). For results with similar flavor, see Shimkin and Shwartz (1993).

3.3 Information revelation

Let σ be a given strategy of player 1. For $n \in \mathbf{N}$, we denote by p_n the conditional distribution over K given \mathcal{H}_n : it is the belief held by player 2 about the true game being played.³ The difference $\|p_n - p_{n+1}\|_1$ may be interpreted as the amount of information that is revealed at stage n .

It is well-known (see, e.g., Sorin (2002, Lemma 3.4)) that, for each τ ,

$$\mathbf{E}_{p,\omega,\sigma,\tau} \left[\sum_{n=1}^{\infty} \|p_n - p_{n+1}\|_1^2 \right] \leq |K|. \quad (3)$$

Given $p \in \Delta(K)$, we denote by σ^p the average non revealing strategy defined by $\sigma^p(h) = \sum_{k \in K} p(k)\sigma(k, h)$, for each finite history h . It is very convenient to relate the benefit derived by player 1 from using his information at a given stage to the amount of information revealed at that stage. Let $n \in \mathbf{N}$ be given. The expected payoff at stage n , conditional on past play is

$$\mathbf{E}_{p,\omega,\sigma,\tau} [g_n | \mathcal{H}_n] = \sum_{k \in K} p_n(k) g^k(\omega_n, \sigma(k, h_n), \tau(h_n)),$$

where $\sigma(k, h_n)$ and $\tau(h_n)$ are the mixed moves used by the two players at that stage.⁴ By Proposition 3.2 and Lemma 3.13 in Sorin (2002),

$$|\mathbf{E}_{p,\omega,\sigma,\tau} [g_n | \mathcal{H}_n] - \langle p_n, g(\omega_n, \sigma^{p_n}(h_n), \tau(h_n)) \rangle| \leq \mathbf{E} [\|p_n - p_{n+1}\|_1 | \mathcal{H}_n]. \quad (4)$$

Definition 14 Let $\tilde{\mathcal{T}}$ be a set of strategies of player 2. Let $\varepsilon > 0$ and σ be given. The strategy $\tilde{\tau} \in \tilde{\mathcal{T}}$ is ε -exhausting information given (p, ω) and σ if $\tilde{\tau}$ maximises $\mathbf{E}_{p,\omega,\sigma,\tilde{\tau}} \left[\sum_{n=1}^{\infty} \|p_n - p_{n+1}\|_1^2 \right]$ up to ε over $\tilde{\mathcal{T}}$.

This notion is relative to the class $\tilde{\mathcal{T}}$. Which class of strategies is meant will always be clear.

Lemma 15 Let $\tilde{\mathcal{T}}, \varepsilon, \sigma, (p, \omega)$ as in Definition 14. Let $\tilde{\tau} \in \tilde{\mathcal{T}}$ be an ε -exhausting strategy given (p, ω) and σ , and let $N \in \mathbf{N}$ be such that $\mathbf{E}_{p,\omega,\sigma,\tilde{\tau}} \left[\sum_{n=N}^{\infty} \|p_n - p_{n+1}\|_1^2 \right] \leq \varepsilon$. Then for each strategy $\tau \in \tilde{\mathcal{T}}$ that coincides with $\tilde{\tau}$ until stage N one has

$$\mathbf{E}_{p,\omega,\sigma,\tau} \left[\sum_{n=N}^{\infty} \|p_n - p_{n+1}\|_1^2 \right] \leq 2\varepsilon, \text{ and } \mathbf{E}_{p,\omega,\sigma,\tau} [\|p_l - p_N\|_1] \leq \sqrt{2\varepsilon} \text{ for each } l \geq N.$$

³The value of p_n at a specific atom of \mathcal{H}_n depends only on σ . Since the distribution on \mathcal{H}_n depends on τ , the law of p_n depends on both σ and τ .

⁴There is a notational inconsistency here, since the right-hand side is the value of the left-hand side on a typical atom of \mathcal{H}_n .

Proof. The first inequality needs no proof. Note that for each $l \geq N$,

$$(\mathbf{E}_{p,\omega,\sigma,\tau} [\|p_l - p_N\|_1])^2 \leq \mathbf{E}_{p,\omega,\sigma,\tau} [\|p_l - p_N\|_1^2] = \mathbf{E}_{p,\omega,\sigma,\tau} \left[\sum_{n=N}^{l-1} \|p_n - p_{n+1}\|_1^2 \right], \quad (5)$$

where the equality follows since (p_n) is a martingale. The second inequality follows. ■

The next lemma is specific to stochastic games with incomplete information.

Lemma 16 *Let (σ, τ) be given. For every $p \in \Delta(K)$, every $\omega \in \Omega$, and every $l \in \mathbf{N}$, one has*

$$|\mathbf{E}_{p,\omega,\sigma,\tau} [\bar{g}_l] - \mathbf{E}_{p,\omega,\sigma^p,\tau} [\bar{g}_l]| \leq 4\mathbf{E}_{p,\omega,\sigma,\tau} \left[\sum_{m=1}^l \|p_m - p_{m+1}\|_1 \right].$$

Proof. For notational convenience, we abbreviate $\mathbf{E}_{p,\omega,\sigma,\tau}$ and $\mathbf{E}_{p,\omega,\sigma^p,\tau}$ to \mathbf{E} and $\tilde{\mathbf{E}}$ respectively, and to \mathbf{P} and $\tilde{\mathbf{P}}$ the corresponding probability distributions. Let $n \leq l$ be given. Since σ^p is non revealing, and by the Lipschitz property,

$$\begin{aligned} & \left| \langle p_n, g(\omega_n, \sigma^{p_n}(h_n), \tau(h_n)) \rangle - \tilde{\mathbf{E}} [g_n | \mathcal{H}_n] \right| \\ &= \left| \langle p_n, g(\omega_n, \sigma^{p_n}(h_n), \tau(h_n)) \rangle - \langle p, g(\omega_n, \sigma^p(h_n), \tau(h_n)) \rangle \right| \\ &\leq 2 \|p_n - p\|_1. \end{aligned} \quad (6)$$

By (4), it follows that

$$\left| \mathbf{E} [g_n | \mathcal{H}_n] - \tilde{\mathbf{E}} [g_n | \mathcal{H}_n] \right| \leq 2 \|p_n - p\|_1 + \|p_n - p_{n+1}\|_1. \quad (7)$$

On the other hand, it is easily checked that the probabilities \mathbf{P}^n and $\tilde{\mathbf{P}}^n$ induced by \mathbf{P} and $\tilde{\mathbf{P}}$ on \mathcal{H}_n satisfy

$$\left\| \mathbf{P}^n - \tilde{\mathbf{P}}^n \right\|_1 \leq \mathbf{E} \left[\sum_{m=1}^n \|p_m - p_{m+1}\|_1 \right]. \quad (8)$$

By (7) and (8),

$$\left| \mathbf{E} [g_n] - \tilde{\mathbf{E}} [g_n] \right| \leq 4\mathbf{E} \left[\sum_{m=1}^n \|p_m - p_{m+1}\|_1 \right],$$

which implies the result. ■

3.4 A partition of states

In this section we define a partition of the set of states, that will be extensively used in the sequel. It hinges on the fact that a single player controls the transitions, but it does not matter who is the controller. The partition is similar to the one defined by Ross and Varadarajan (1991) for Markov decision processes, who also provide an algorithm to calculate it.

We assume that player 1 controls the transition. The partition when player 2 controls the transition is defined analogously. Since transitions are independent of player 2's actions, we here omit player 2's strategy from the notations, whenever convenient.

Given $\omega \in \Omega$, we denote by

$$r_\omega = \min \{n \in \mathbf{N}, \omega_n = \omega\}$$

the stage of the first visit to ω . By convention, the minimum over an empty set is $+\infty$.

Definition 17 Let $\omega_1, \omega_2 \in \Omega$. We say that ω_1 leads to ω_2 if $\omega_1 = \omega_2$, or if $\mathbf{P}_{\omega_1, \sigma}(r_{\omega_2} < +\infty) = 1$ for some strategy σ of player 1.

Note that the relation *leads to* is reflexive and transitive.

We define an equivalence relation over Ω by

$$\omega \leftrightarrow \omega' \text{ if and only if } \omega \text{ leads to } \omega' \text{ and } \omega' \text{ leads to } \omega.$$

The equivalence classes of this relation are called *communicating sets*. Given $\omega \in \Omega$, we let C_ω denote the communicating set that contains ω , and we define

$$I_\omega = \{i \in I \mid q(C_\omega \mid \omega, i) = 1\}.$$

The set I_ω may (but does not have to) be empty only if $|C_\omega| = 1$. Actions in I_ω are called *stay* actions, and any state ω such that $I_\omega = \emptyset$ is a *null* state. The set of non-null states is Ω_c . Note that $C_\omega \subseteq \Omega_c$ whenever $\omega \in \Omega_c$.

Lemma 18 $\omega \in \Omega_c$ if and only if there is a stationary strategy x_{C_ω} such that C_ω is a recurrent set for x .

Thus, $I_\omega = \emptyset$ if and only if ω is transient for every stationary strategy x .

Proof. We start with the direct implication. Let $\omega \in \Omega_c$. For $\omega' \in C_\omega$, define $x_{\omega'} \in \Delta(A)$ by

$$x_{\omega'}[i] = \begin{cases} 0 & i \notin I_{\omega'} \\ 1/|I_{\omega'}| & i \in I_{\omega'}. \end{cases}$$

and let x be any stationary strategy that coincides with $x_{\omega'}$ in each state $\omega' \in C_\omega$. It is easy to show that C_ω is recurrent under x .

The reverse implication is straightforward. ■

It is useful to distinguish the communicating sets that are recurrent sets for a fully mixed stationary strategy x . The corresponding set of states is denoted Ω_0 . Thus, $\omega \in \Omega_0$ if and only if $I_{\omega'} = I$ for every $\omega' \in C_\omega$.

Lemma 19 Assume player 1 controls transitions. Let $\omega \in \Omega$ and $\omega' \in C_\omega$. If one of the players can guarantee w in $\Gamma(p, \omega)$, he can also guarantee w in $\Gamma(p, \omega')$.

Proof. Assume first player 1 can guarantee w in $\Gamma(p, \omega)$. Let σ be a strategy that guarantees $w - \varepsilon$ in $\Gamma(p, \omega)$, and let σ^* be the strategy that plays x_{C_ω} until r_ω , then switches to σ . In the game $\Gamma(p, \omega')$, the strategy σ^* guarantees $w - \varepsilon'$, for each $\varepsilon' > \varepsilon$.

Assume now player 2 can guarantee w in $\Gamma(p, \omega)$, but assume to the contrary that he *cannot* guarantee w in $\Gamma(p, \omega')$, for some $\omega' \in C_\omega$. Then for every $\varepsilon > 0$, every strategy τ of player 2, and every $N \in \mathbf{N}$ there is a strategy $\sigma_{\tau, N}$ of player 1 and $n_{\tau, N} \geq N$ such that $\gamma_{n_{\tau, N}}(p, \omega', \sigma_{\tau, N}, \tau) > w + \varepsilon$. Let ε, τ and N be given. Let σ^* be the strategy of player 1 defined as follows. Play x_{C_ω} until stage $r_{\omega'}$, and then switch to $\sigma_{\tau_{r_{\omega'}}, M}$, where $\tau_{r_{\omega'}}$ is the strategy induced by τ after stage $r_{\omega'}$, and $M \geq N$ is sufficiently large. One can verify that if M is sufficiently large then there is $n' \geq N$ such that $\gamma_{n'}(p, \omega, \sigma^*, \tau) > w + \varepsilon/2$. ■

3.5 Auxiliary games

As for the analysis of repeated zero-sum games with lack of information on one side, it is convenient to introduce an average game in which no player is informed of the realization of k .

For notational ease, assume that player 1 is the controller. For every $p \in \Delta(K)$ and every non null state $\omega \in \Omega$, we denote by $\tilde{\Gamma}_R(p, \omega)$ the zero-sum stochastic game with: (i) initial state ω , (ii) state space C_ω , (iii) reward function $\sum_k p_k g^k$, (iv) action sets $I_{\omega'}$ and J at each state $\omega' \in C_\omega$, and (v) transition function induced by q .

In the case where player 2 is the controller, the game $\tilde{\Gamma}_R(p, \omega)$ is defined by restricting player 2's action set to $J_{\omega'}$ in each state $\omega' \in C_\omega$.

Thus, $\tilde{\Gamma}_R(p, \omega)$ is the stochastic game in which player 1 is not informed of the realization of k (or does not use his information), and the controller is restricted to stay actions. Since the controller can use only stay actions, the game remains in C_ω forever. The letter R is a reminder for *restricted*, while the symbol $\tilde{\cdot}$ stands for average.

Note that $\tilde{\Gamma}_R(p, \omega)$ is a single controller game. Denote by $\tilde{u}(p, \omega)$ its value. Note that $\tilde{u}(p, \omega) = u_\infty(p, \omega)$ for each $\omega \in \Omega_0$.⁵

By convention, if ω is a null state, we set $\tilde{u}(p, \omega) = -\infty$ if player 1 controls the transition and $\tilde{u}(p, \omega) = +\infty$ if player 2 controls the transition. By Lemma 19, for every communicating set C , $\tilde{u}(p, \omega)$ is independent of $\omega \in C$.

Proposition 20 *For every $\omega \in \Omega_0$ and every $p \in \Delta(K)$ the value $v(p, \omega)$ of $\Gamma(p, \omega)$ exists and is equal to $\text{cav } \tilde{u}(p, \omega)$ ($= \text{cav } u_\infty(p, \omega)$).*

Thus, restricted to Ω_0 , the game is similar to a standard repeated game with incomplete information.

Proof. The proof of this lemma is similar to the proof for repeated games with incomplete information on one side. Clearly player 1, by not using his information, can guarantee $\tilde{u}(p, \omega)$. By Lemma 9, player 1 can guarantee $\text{cav } \tilde{u}(p, \omega)$.

The proof that player 2 can guarantee $\text{cav } \tilde{u}$ is based on approachability results, and follows closely classical lines. Let $a \in \mathbf{R}^K$ be such that

$$\begin{aligned} \langle a, p \rangle &= \text{cav } \tilde{u}(p, \omega) \\ \langle a, q \rangle &\geq \text{cav } \tilde{u}(q, \omega) \text{ for each } q \in \Delta(K). \end{aligned}$$

If $\text{cav } \tilde{u}(\cdot, \omega)$ is differentiable at p , then a is defined by the hyperplane tangent to $\text{cav } \tilde{u}(\cdot, \omega)$ at p . By Proposition 13, a is approachable. By Lemma 12, player 2 can guarantee $\text{cav } \tilde{u}$. ■

Let $\Gamma_R(p, \omega)$ be a game similar to $\tilde{\Gamma}_R(p, \omega)$, but in which player 1 is informed of k . Thus, $\Gamma_R(p, \omega)$ differs from $\Gamma(p, \omega)$ only in that actions of the controller are restricted.

A similar argument as the one used in the proof of Proposition 20 proves the following:

Lemma 21 *Let ω be a non null state. Then $\Gamma_R(p, \omega)$ has a value, which is $\text{cav } \tilde{u}(p, \omega)$.*

We denote by Γ_R^V the stochastic game with vector payoffs in which the controller is restricted to stay actions.

⁵By Filar (1981), both players have optimal stationary strategies. We will not use this fact.

3.6 Functional equations

Let \mathcal{B} denote the set of functions $w : \Delta(K) \times \Omega \rightarrow [0, 1]$ that are 1-Lipshitz with respect to p . We here define three operators on \mathcal{B} that will be used to characterize the solutions of the game.

When transitions are controlled by player 1, we define T_1 by

$$T_1 w(p, \omega) = \text{cav max} \left\{ \text{cav } \tilde{u}, \max_{\omega' \in C_\omega, i \notin I_{\omega'}} \mathbf{E} [w \mid \omega', i] \right\} (p, \omega). \quad (9)$$

By convention, a maximum over an empty set is $-\infty$. In this expression, $\mathbf{E} [w \mid \omega', i]$ stands for the expectation of w under $q(\cdot \mid \omega', i)$.

When transitions are controlled by player 2, we define T_2 and T_3 by

$$T_2 w(p, \omega) = \text{cav min} \left\{ \tilde{u}, \min_{\omega' \in C_\omega, j \notin J_{\omega'}} \mathbf{E} [w \mid \omega', j] \right\} (p, \omega)$$

$$T_3 w(p, \omega) = \text{min} \left\{ \text{cav } \tilde{u}, \min_{\omega' \in C_\omega, j \notin J_{\omega'}} \mathbf{E} [w \mid \omega', j] \right\} (p, \omega).$$

Since the maximum (or minimum) of a finite number of elements of \mathcal{B} belongs to \mathcal{B} , and since concavification preserves Lipshitz properties, all three operators T_1, T_2 and T_3 map \mathcal{B} into \mathcal{B} . Note that T_i is monotonic: $w_1 \leq w_2$ implies that $T_i w_1 \leq T_i w_2$.

We now assume that player 1 controls transitions, and prove few results on T_1 . When transitions are controlled by player 2, identical results hold for both T_2 and T_3 , proofs being analogous hence omitted.

Proposition 22 1. T_1 has a unique fixed point w .

2. The sequences (w_n^0) and (w_n^1) defined by $w_0^j = j$, $w_{n+1}^j = T_1 w_n^j$ for $j = 0, 1$, are monotonic and converge uniformly to w .

3. w coincides with $\text{cav } \tilde{u}$ on Ω_0 .

4. If $f \in \mathcal{B}$ satisfies $f \leq T_1 f$ (resp. $f \geq T_1 f$), then $f \leq w$ (resp. $f \geq w$).

By induction on n , the sequences $(w_n^j)_n$ associated with T_3 , are sequences of concave functions. Thus, the fixed points of T_2 and T_3 are concave functions.

Proof. Plainly, **2** follows from **1**, by monotonicity of T_1 . Since $\text{cav } \tilde{u}(p, \omega)$ is constant on every communicating set, so is $T_1 w(p, \omega)$, for every $w \in \mathcal{B}$. Since $I_\omega = I$ for every $\omega \in \Omega_0$, $T_1 w(p, \omega) = \text{cav } \tilde{u}(p, \omega)$ for every $w \in \mathcal{B}$, every $\omega \in \Omega_0$, and every $p \in \Delta(K)$. Thus, **3** will follow from **1**. We now prove **1**. By Ascoli's characterization, \mathcal{B} is a compact metric space when endowed with the uniform norm. Since T_1 is non-expansive, it is continuous on \mathcal{B} , hence it has a fixed point.

We prove uniqueness by contradiction. Let w_1 and w_2 be two distinct fixed points of T_1 , and assume w.l.o.g. that $\delta := \max_{(p, \omega) \in \Delta(K) \times \Omega} (w_1(p, \omega) - w_2(p, \omega)) > 0$. Let

$$D = \{\omega \in \Omega, w_1(p, \omega) - w_2(p, \omega) = \delta \text{ for some } p \in \Delta(K)\}$$

contain those states where the difference is maximal. Since both $w_1(p, \cdot)$ and $w_2(p, \cdot)$ are constant on each communicating set, $C_\omega \subseteq D$ for each $\omega \in D$.

Since $w_1 = w_2$ on Ω_0 , $D \subseteq \Omega \setminus \Omega_0$. Let $\omega \in D$ be given, and let $p_0 \in \Delta(K)$ be an extreme point of the convex hull of the set $\{p \in \Delta(K) : w_1(p, \omega) - w_2(p, \omega) = \delta\}$. Thus, $w_1(p_0, \omega) - w_2(p_0, \omega) = \delta > 0$. Since $w_1(\cdot, \omega)$ is concave, it also follows that $(p_0, w_1(p_0, \omega))$ is an extreme point of the hypograph of the concave function $w_1(\cdot, \omega)$. This implies

$$w_1(p_0, \omega) = \max \left\{ \text{cav } \tilde{u}, \max_{\omega' \in C_\omega, i \notin I_\omega} \mathbf{E} [w_1 | \omega', i] \right\} (p_0, \omega).$$

Since $w_1(p_0, \omega) > w_2(p_0, \omega) \geq \text{cav } \tilde{u}(p_0, \omega)$, one has $w_1(p_0, \omega) = \mathbf{E} [w_1(p_0, \cdot) | \omega', i]$ for some $\omega' \in C_\omega$ and $i \notin I_{\omega'}$. Since $T_1 w_2 = w_2$, $w_2(p_0, \omega) \geq \mathbf{E} [w_2(p_0, \cdot) | \omega', i]$, and therefore

$$\delta = w_1(p_0, \omega) - w_2(p_0, \omega) \leq \mathbf{E} [w_1(p_0, \cdot) - w_2(p_0, \cdot) | \omega', i].$$

By the definition of D , this implies that $q(D | \omega', i) = 1$.

Thus, for every $\omega \in D$ there exists $\omega' \in C_\omega$ and $i \notin I_{\omega'}$ that satisfy $q(D | \omega', i) = 1$. This implies the existence of $\bar{\omega}, \tilde{\omega} \in D$ such that $C_{\bar{\omega}} \neq C_{\tilde{\omega}}$ and $\bar{\omega} \leftrightarrow \tilde{\omega}$ – a contradiction. This proves **1**.

To prove **4**, we assume that $\delta = \max_{(p, \omega) \in \Delta(K) \times \Omega} (f(p, \omega) - w(p, \omega)) > 0$, and repeat the second part of the proof of **1** to obtain a contradiction. ■

4 Lack of information on one side

4.1 Preliminaries

We here single out a useful lemma. The Lemma concerns a standard stochastic game G , and its version G_R in which player 1 is restricted to stay actions. Thus, K is a singleton.

Lemma 23 *Let G be a zero-sum stochastic game, with transitions controlled by player 1, and let $\omega \in \Omega$. If player 2 can guarantee $\alpha \in \mathbf{R}$ in $G_R(\omega)$, and he can guarantee $w : \Omega \rightarrow \mathbf{R}$ in G , then he can also guarantee $\max \{\alpha, \max_{\omega' \in C_\omega, i \notin I_{\omega'}} \mathbf{E} [w | \omega', i]\}$ in $G(\omega)$.*

Proof. By Lemma 19 player 2 can guarantee α in $G_R(\omega')$ for every $\omega' \in C_\omega$. Let τ_1 be a strategy that guarantees $\alpha + \varepsilon$ in $G_R(\omega')$ for every $\omega' \in C_\omega$, and let τ_2 be a strategy that guarantees $w + \varepsilon$ in G . Let $N \in \mathbf{N}$ be such that for every $n \geq N$, every $\omega' \in C_\omega$ and every σ in $G_R(\omega)$, $\gamma_n(\omega', \sigma, \tau_1) \leq \alpha + \varepsilon$, while for every σ in G , $\omega' \in \Omega$, $\gamma_n(\omega', \sigma, \tau_2) \leq w(\omega') + \varepsilon$.

Define $\nu = 1 + \inf \{n \geq 1, i_n \notin I_{\omega_n}\}$. Define a strategy τ for player 2 as follows.

- At stage ν , τ forgets past play and start following τ_2 .
- Before stage ν , τ plays in blocks of size N (the last block may be shorter). In block l , where $lN < \nu$, τ forgets past play and follows $\tau_1(\omega_{lN})$ for N stages.

Let σ be an arbitrary pure strategy. We will compute an upper bound on $\mathbf{E}_{\omega, \sigma, \tau} [\bar{g}_n]$, for n sufficiently large. Set $L_* = \lceil \frac{\ln \varepsilon}{\ln(1-\varepsilon)} \rceil$, and take $n \geq N_1 := \lceil L_* N / \varepsilon^2 \rceil$. Denote by $\bar{g}_{m_1 \rightarrow m_2}$ the average payoff from stage m_1 to stage m_2 . With $\theta^* := \lceil \frac{\nu-1}{N} \rceil$, and since payoffs are non negative, one has the inequality

$$\bar{g}_n \leq \frac{N\theta^*}{n} \bar{g}_{N\theta^*} + \frac{n+1-\nu}{n} \bar{g}_{\nu \rightarrow n}. \quad (10)$$

On the event $\nu \leq n - N$, one has

$$\mathbf{E}_{\omega, \sigma, \tau} [\bar{g}_{\nu \rightarrow n} | \mathcal{H}_\nu] = \mathbf{E}_{\omega, \sigma_\nu, \tau_2} [\bar{g}_{n-\nu+1}] \leq w(\omega_\nu) + \varepsilon, \quad (11)$$

where σ_ν is the strategy induced by σ after ν . Since σ is pure, $\nu - 1$ is a stopping time and, using (11),

$$\mathbf{E}_{\omega, \sigma, \tau} [\bar{g}_{\nu \rightarrow n} | \mathcal{H}_{\nu-1}] \leq \mathbf{E} [w | \omega_{\nu-1}, i_{\nu-1}] + \varepsilon. \quad (12)$$

On the other hand, on the event $\nu > n - N$, $\frac{n+1-\nu}{n} \leq \varepsilon$. Therefore, using (12),

$$\mathbf{E}_{\omega, \sigma, \tau} \left[\frac{n+1-\nu}{n} \bar{g}_{\nu \rightarrow n} \right] \leq \beta \mathbf{E}_{\omega, \sigma, \tau} \left[\frac{n+1-\nu}{n} \right] + \varepsilon. \quad (13)$$

We now proceed to the first term in the decomposition (10) of \bar{g}_n . For each l , we let $\pi_l = \mathbf{P}_{\omega, \sigma, \tau} [\nu \leq (l+1)N | \mathcal{H}_{lN+1}]$. By the choice of N ,

$$\begin{aligned} \mathbf{E}_{\omega, \sigma, \tau} [\bar{g}_{lN+1 \rightarrow (l+1)N} | \mathcal{H}_{lN+1}] &\leq \alpha + 2\varepsilon \text{ on the event } \pi_l < \varepsilon \\ \text{and } \mathbf{E}_{\omega, \sigma, \tau} [\bar{g}_{lN+1 \rightarrow (l+1)N} | \mathcal{H}_{lN+1}] &\leq 1 \text{ otherwise.} \end{aligned}$$

By taking expectations, this yields

$$\mathbf{E}_{\omega, \sigma, \tau} [\bar{g}_{lN+1 \rightarrow (l+1)N}] \leq (\alpha + 2\varepsilon) \mathbf{P}_{\omega, \sigma, \tau} (\theta^* > l) + \mathbf{P}_{\omega, \sigma, \tau} (\pi_l \geq \varepsilon, \theta^* > l).$$

By summation over l , one has

$$\mathbf{E}_{\omega, \sigma, \tau} \left[\sum_{l=0}^{\theta^*-1} \bar{g}_{lN+1 \rightarrow (l+1)N} \right] \leq (\alpha + 2\varepsilon) \mathbf{E}_{\omega, \sigma, \tau} [\theta^*] + \mathbf{E}_{\omega, \sigma, \tau} [\tilde{N}_{\theta^*}], \quad (14)$$

where $\tilde{N}_m = |\{l < m : \pi_l \geq \varepsilon\}|$. Plainly, $\mathbf{E}_{\omega, \sigma, \tau} [\tilde{N}_{\theta^*}] \leq \frac{1}{\varepsilon}$. Thus, (14) rewrites

$$\mathbf{E}_{\omega, \sigma, \tau} \left[\frac{\nu^*}{n} \bar{g}_{\nu^*} \right] \leq (\alpha + 2\varepsilon) \mathbf{E}_{\omega, \sigma, \tau} \left[\frac{\nu^*}{n} \right] + \frac{N}{n\varepsilon}. \quad (15)$$

The result follows by (10), (12), (13) and (15). ■

We shall need a variant of the previous result, whose proof is identical to the previous proof. Consider the stochastic game with incomplete information $\Gamma(p, \omega)$, where ω is a non-null state and assume that transitions are controlled by player 2. Assume that player 1 can guarantee a function w . Then player 1 can also guarantee $\min \{\tilde{u}, \min_{\omega' \in C_\omega, j \notin J_{\omega'}} \mathbf{E} [w | \omega', j]\}$ (p, ω) in $\Gamma(p, \omega)$.

4.2 Transitions Controlled by Player 1

In this section we assume that transitions are controlled by player 1.

Proposition 24 *The unique fixed point of T_1 is the value of Γ .*

Proof. Let w be the unique fixed point of T_1 , and fix $\epsilon > 0$ once and for all.

Step 1: Player 1 can guarantee w in Γ

By Lemma 21 player 1 can guarantee $\text{cav } \tilde{u}$. Assume that player 1 can guarantee w_n^0 for some $n \in \mathbf{N}$. Let $p \in \Delta(K)$ and $\omega \in \Omega$ be given. Plainly, for every $\omega' \in C_\omega$ and every $i \notin I_\omega$ player 1 can guarantee $\mathbf{E}[w_n^0 | \omega', i](p, \omega)$ in $\Gamma(p, \omega')$; first he plays the action i at ω' , and then a strategy that guarantees $w_n^0(p, \cdot)$ (up to ϵ). By Lemma 19, he can guarantee $\mathbf{E}[w_n^0 | \omega', i](p, \omega)$ in $\Gamma(p, \omega)$. Therefore, he can guarantee $T_1 w_n^0 = w_{n+1}^0$ in Γ . Since player 1 can guarantee $w_0^0 = 0$, the result follows by Lemma 9.

We now prove that player 2 can guarantee w .

Step 2: Definition of approachable sets.

For $\omega \in \Omega$, let \mathcal{B}_ω be the set of vectors approachable in Γ^V by player 2 at ω . We also define

$$\mathcal{A}_\omega = \{a \in \mathbf{R}^K, \langle a, p \rangle \geq \text{cav } \tilde{u}(p, \omega) \text{ for every } p\}.$$

By Proposition 13 and Lemma 21, \mathcal{A}_ω is the set of vectors approachable by player 2 at ω in the stochastic game with vector payoffs Γ_R^V . Both sets \mathcal{A}_ω and \mathcal{B}_ω are non-empty, closed, convex and (upwards) comprehensive.

For every $\omega \in \Omega$ define

$$\mathcal{C}_\omega = \left\{ c = \max \left\{ a, \max_{\omega' \in C_\omega, i \notin I_\omega'} \mathbf{E}[b(\cdot) | \omega', i] \right\}, \text{ where } a \in \mathcal{A}_\omega, b(\omega'') \in \mathcal{B}_{\omega''} \text{ for every } \omega'' \in \Omega \right\}.$$

Step 3: $\mathcal{C}_\omega \subseteq \mathcal{B}_\omega$.

Fix $c \in \mathcal{C}_\omega$. Let τ_1 be a strategy that approaches $a + \epsilon$ at ω , and let τ_2 be a strategy that approaches $b(\omega'') + \epsilon$ at each state ω'' . For each k the strategy τ_1 guarantees $a^k + \epsilon$ in the $\Gamma(k, \omega)$, and τ_2 has similar properties. By Lemma 23, applied independently to each G^k , the strategy obtained by concatenation of τ_1 and τ_2 guarantees $\max\{a^k, \max_{\omega' \in C_\omega, i \notin I_\omega'} \mathbf{E}[b^k(\cdot) | \omega', i]\} + 3\epsilon = c^k + 3\epsilon$ in G^k .

Step 4: Player 2 can guarantee w .

Let $f(p, \omega) = \inf_{a \in \mathcal{B}_\omega} \langle a, p \rangle$ and $h(p, \omega) = \inf_{a \in \mathcal{C}_\omega} \langle a, p \rangle$, so that by **Step 3** $f \leq h$. By Lemma 12 player 2 can guarantee $\langle a, p \rangle$ in $\Gamma(p, \omega)$ for every $a \in \mathcal{B}_\omega$. Therefore he can guarantee $f(p, \omega)$ as well. By Lemma 10, the definition of \mathcal{C}_ω may be rephrased as

$$h = \text{cav } \max \left\{ \text{cav } \tilde{u}, \max_{\omega' \in C_\omega, i \notin I_\omega'} \mathbf{E}[f | \omega', i] \right\} = T_1 f.$$

Thus, $f \leq T_1 f$. By Proposition 22(4), $f \leq w$. Therefore, player 2 can guarantee w . ■

4.3 Transitions Controlled by Player 2

In this section we assume that transitions are controlled by player 2.

4.3.1 The max-min Value

Lemma 25 *The unique fixed point of T_2 is the max-min value of Γ .*

Proof. Let w be the fixed point of T_2 , and fix $\varepsilon > 0$.

Step 1: Player 1 can guarantee w .

Assume player 1 can guarantee w_m^0 for some $m \in \mathbf{N}$. By the remark following Lemma 23, player 1 can guarantee $\min \{ \tilde{u}, \min_{\omega' \in C_\omega, j \notin J_{\omega'}} \mathbf{E} [w_m^0 \mid \omega', j] \}$. Hence player 1 can also guarantee $\text{cav} \min \{ \tilde{u}, \min_{\omega' \in C_\omega, j \notin J_{\omega'}} \mathbf{E} [w_m^0 \mid \omega', j] \} = w_{m+1}^0$. Since player 1 can guarantee $w_0^0 = 0$, the result follows.

We now prove that player 2 can defend w . Assume that player 2 can defend w_m^1 for some $m \in \mathbf{N}$, and let σ be an arbitrary strategy of player 1.

Step 2: Definition of a reply

Given (p, ω) , we let $\tau_1(p, \omega)$ be a strategy that guarantees $\tilde{u}(p, \omega) + \varepsilon$ in $\tilde{\Gamma}_R(p, \omega)$. Choose $N_1 \in \mathbf{N}$ such that $\gamma_n(p, \omega, \tilde{\sigma}, \tau_1(p, \omega)) \leq \tilde{u}(p, \omega) + \varepsilon$ for every $n \geq N_1$ and every non revealing strategy $\tilde{\sigma}$ of player 1.

By the remark that follows Definition 1, N_1 can be chosen independently of (p, ω) . Let $\tilde{\mathcal{T}}$ be the set of strategies of player 2 in $\tilde{\Gamma}_R(p, \omega)$, and let $\tilde{\tau} \in \tilde{\mathcal{T}}$ be an $\varepsilon^2/32N_1^2$ -exhausting information strategy given σ and (p, ω) . Choose $N \in \mathbf{N}$ such that

$$\mathbf{E}_{p, \omega, \sigma, \tilde{\tau}} \left[\sum_{n=N}^{+\infty} \|p_n - p_{n+1}\|_1^2 \right] \leq \frac{\varepsilon^2}{32N_1^2}.$$

We define τ by

- Play $\tilde{\tau}$ up to stage N .
- At stage N compute $\beta_N := \min \{ \tilde{u}, \min_{\omega' \in C_\omega, j \notin J_{\omega'}} \mathbf{E} [w_m^1 \mid \omega', j] \} (p_N, \omega_N)$.
 - If $\beta_N = \tilde{u}(p_N, \omega_N)$, play by successive blocks of length N_1 : in the $b + 1$ th block play the strategy $\tau_1(p_{N+bN_1}, \omega_{N+bN_1})$.
 - Otherwise, switch to a strategy that defends the quantity $\min_{\omega' \in C_\omega, j \notin J_{\omega'}} \mathbf{E} [w_m^1 \mid \omega', j] (p_N, \omega_N) + \varepsilon$ against σ_N , where σ_N is the strategy induced by σ after stage N .

Step 3: The computation

We here prove that τ defends $w_{m+1}^1(p, \omega) + 6\varepsilon$ in $\Gamma(p, \omega)$. We abbreviate $\mathbf{E}_{p, \omega, \sigma, \tau}$ to \mathbf{E} . First, we provide an upper bound on the average payoff $\mathbf{E} [\bar{g}_{N \rightarrow N+n} \mid \mathcal{H}_N]$ between stages N and $N + n$ on the event

$$A := \{ \beta_N = \tilde{u}(p_N, \omega_N) \}. \quad (16)$$

Take first $n = N_1$. By definition,

$$\mathbf{E} [\bar{g}_{N \rightarrow N+N_1-1} \mid \mathcal{H}_N] = \mathbf{E}_{p_N, \omega_N, \sigma_N, \tau_1(p_N, \omega_N)} [\bar{g}_{N_1}].$$

By the choice of N_1 ,

$$\mathbf{E}_{p_N, \omega_N, \sigma_N^{p_N}, \tau_1(p_N, \omega_N)} [\bar{g}_{N_1}] \leq \tilde{u}(p_N, \omega_N) + \varepsilon. \quad (17)$$

On the other hand, by Lemma 16,

$$\begin{aligned} & \left| \mathbf{E}_{p_N, \omega_N, \sigma_N^{p_N}, \tau_1(p_N, \omega_N)} [\bar{g}_{N_1}] - \mathbf{E}_{p_N, \omega_N, \sigma_N, \tau_1(p_N, \omega_N)} [\bar{g}_{N_1}] \right| \\ & \leq 4\mathbf{E}_{p_N, \omega_N, \sigma_N, \tau_1(p_N, \omega_N)} \left[\sum_{m=1}^{N_1} \|p_m - p_{m+1}\|_1 \right]. \end{aligned}$$

Thus, using (17),

$$\mathbf{E} [\bar{g}_{N \rightarrow N+N_1-1} | \mathcal{H}_N] \leq \tilde{u}(p_N, \omega_N) + \varepsilon + 4\mathbf{E} \left[\sum_{m=N}^{N+N_1-1} \|p_m - p_{m+1}\|_1 | \mathcal{H}_N \right].$$

The same computation applies to any block of N_1 stages. Specifically, for each $b \geq 0$,

$$\begin{aligned} \mathbf{E} [\bar{g}_{N+bN_1 \rightarrow N+(b+1)N_1-1} | \mathcal{H}_{N+bN_1}] & \leq \tilde{u}(p_{N+bN_1}, \omega_{N+bN_1}) + \varepsilon \\ & + 4\mathbf{E} \left[\sum_{m=N+bN_1}^{N+(b+1)N_1-1} \|p_m - p_{m+1}\|_1 | \mathcal{H}_{N+bN_1} \right]. \end{aligned}$$

Since $\tilde{u}(p, \cdot)$ is constant on every communicating set, and since $\tilde{u}(\cdot, \omega)$ is Lipschitz, $\tilde{u}(p_{N+bN_1}, \omega_{N+bN_1}) \leq \tilde{u}(p_N, \omega_N) + \|p_{N+bN_1} - p_N\|_1$. By taking expectations on the event A (defined by (16)), one gets, by Lemma 15,

$$\begin{aligned} \mathbf{E} [\mathbf{1}_A \bar{g}_{N+bN_1 \rightarrow N+(b+1)N_1-1}] & \leq \mathbf{E} [\mathbf{1}_A \tilde{u}(p_N, \omega_N)] + \varepsilon + \mathbf{E} [\|p_{N+bN_1} - p_N\|_1] \\ & + 4\mathbf{E} \left[\sum_{m=N+bN_1}^{N+(b+1)N_1-1} \|p_m - p_{m+1}\|_1 \right] \\ & \leq \mathbf{E} [\mathbf{1}_A \tilde{u}(p_N, \omega_N)] + 3\varepsilon. \end{aligned}$$

By averaging over blocks, one obtains for every $n \geq \frac{2}{\varepsilon}(N + N_1)$,

$$\mathbf{E} [\mathbf{1}_A \bar{g}_n] \leq \mathbf{E} [\mathbf{1}_A \tilde{u}(p_N, \omega_N)] + 4\varepsilon. \quad (18)$$

On the other hand, there is $N_2 \in \mathbf{N}$ such that for every $n \geq N_2$,

$$\mathbf{E} [\bar{g}_n | \mathcal{H}_N] \leq \min_{\omega' \in C_\omega, j \notin J_{\omega'}} \mathbf{E} [w_m^1 | \omega', j] (p_N, \omega_N) + 2\varepsilon \text{ on the event } \bar{A}. \quad (19)$$

By taking expectations, (18) and (19) yield

$$\begin{aligned} \mathbf{E} [\bar{g}_n] & \leq \mathbf{E} \left[\min \left\{ \tilde{u}, \min_{\omega' \in C_\omega, j \notin J_{\omega'}} \mathbf{E} [w_m^1 | \omega', j] \right\} (p_N, \omega_N) \right] + 6\varepsilon \\ & \leq \text{cav} \min \left\{ \tilde{u}, \min_{\omega' \in C_\omega, j \notin J_{\omega'}} \mathbf{E} [w_m^1 | \omega', j] \right\} (p, \omega) + 6\varepsilon. \end{aligned}$$

for every $n \geq \max\{N_2, \frac{2}{\varepsilon}(N + N_1)\}$. ■

4.4 The min-max Value

Lemma 26 *The unique fixed point of T_3 is the min-max value of Γ .*

Proof. Let w be the unique fixed point of T_3 , and fix $\varepsilon > 0$.

We first prove that player 2 can guarantee w . Assume that player 2 can guarantee w_m^1 for some $m \in \mathbf{N}$, and let (p, ω) be given. Plainly, for each $\omega' \in C_\omega, j \notin J_{\omega'}$, player 2 can guarantee $\mathbf{E}[w_m^1 | \omega', j]$ in $\Gamma(p, \omega')$ by first playing j at ω' , and then a strategy that guarantees w_m^1 (up to ε). By Lemma 19, he can guarantee $\mathbf{E}[w_m^1 | \omega', j]$ in $\Gamma(p, \omega)$ as well. By Lemma 21, player 2 can guarantee $\text{cav } \tilde{u}$. Thus, he can guarantee $T_3 w_m^1 = w_{m+1}^1$. Since he can guarantee w_0^1 , the result follows.

We now prove that player 1 can defend w_m^0 for each $m \in \mathbf{N}$. Clearly, player 1 can defend $w_0^0 = 0$. Assume that player 1 can defend w_m^0 for some $m \in \mathbf{N}$. Let a strategy τ of player 2 and $(p, \omega) \in \Delta(K) \times \Omega$ be given. Set $\nu = 1 + \inf\{n \geq 1, j_n \notin J_{\omega_n}\}$. The supremum of $\mathbf{P}_{p, \omega, \sigma, \tau}(\nu < \infty)$ over all strategies σ coincides with the supremum over all non revealing strategies σ .⁶ Denote by σ^* a *non revealing* strategy that achieves the supremum up to ε . We choose N such that $\mathbf{P}_{p, \omega, \sigma^*, \tau}(\nu > N) \leq \varepsilon$. The strategy σ^* thus exhausts the probability of leaving the initial communicating set. Denote by $\tau_{\min\{\nu, N\}}$ the strategy induced by τ after stage $\min\{\nu, N\}$.

On the event $\nu > N$, there is a strategy $\tilde{\tau}$ in $\Gamma_R(p, \omega)$ such that $\|\mathbf{P}_{p, \omega_N, \sigma, \tilde{\tau}} - \mathbf{P}_{p, \omega_N, \sigma, \tau_N}\| \leq \mathbf{P}_{p, \omega_N, \sigma, \tau_N}(\nu < +\infty)$ for every non revealing strategy σ in $\Gamma_R(p, \omega)$. This strategy depends on the history up to stage N .

We now define the reply σ of player 1 to τ as follows: play σ^* up to stage $\min\{\nu, N\}$.

- If $\nu > N$, switch to a strategy that defends $\text{cav } \tilde{u}(p, \omega) + \varepsilon$ in $\Gamma_R(p, \omega_N)$ against $\tilde{\tau}$;
- If $\nu \leq N$, switch to a strategy that defends $w(p, \omega_\nu) + \varepsilon$ against τ_ν .

Since there are finitely many histories of length N , the set of strategies $(\tau_{\min\{\nu, N\}})$ is finite. It is straightforward to check that σ defends

$$\min \left\{ \text{cav } \tilde{u}, \min_{\omega' \in C_\omega, j \notin J_{\omega'}} \mathbf{E}[w_m^1 | \omega', j] \right\} (p, \omega) + 2\varepsilon = w_{m+1}^1(p, \omega) + 2\varepsilon$$

against τ . ■

4.5 An example

Since $\min\{\text{cav } f, g\}$ may be strictly bigger than $\text{cav}(\min\{f, g\})$, the max-min value and the min-max value may differ.

Consider the following game, where player 2 controls the transitions, and $|\omega| = |K| = 2$, $|I| = 2$ and $|J| = 5$.

⁶This is true since, given σ that approximates the supremum up to ε , the non revealing strategy σ' that is defined by $\sigma'_k = \sigma_l$ for every k , where l maximizes $(p_j)_j$, approximates the supremum up to $|K|\varepsilon$.

		$k = 1$					$k = 2$						
		j_1	j_2	j_3	j_4	j_5	j_1	j_2	j_3	j_4	j_5		
T		4	0	0	0	0	0	0	0	0	0	ω_2	
B		0	0	0	0	0	0	4	4	4	4		
		j_1	j_2	j_3	j_4	j_5	j_1	j_2	j_3	j_4	j_5		
T		0	1	1	3	\uparrow	3	0	1	0	\uparrow	ω_1	
B		0	1	0	3	\uparrow	3	1	1	0	\uparrow		

Figure 1

The initial state is ω_1 (bottom two matrices). If in ω_1 player 2 chooses j_5 , the game moves to ω_2 , which is absorbing. If player 2 chooses another action in ω_1 , the game remains in ω_1 . Payoffs are as appears in Figure 1 (the definition of $g^k(\omega_1, \cdot, j_5)$ is irrelevant).

Note that $I_{\omega_1} = \{j_1, j_2, j_3, j_4\}$, $\Omega_0 = \{\omega_2\}$, and $C_{\omega_1} = \{j_5\}$.

The game $\Gamma_R(p, \omega_1)$ is similar to Example 3.3 in Aumann and Maschler (1995). As calculated in Aumann and Maschler,

$$f(p) = \tilde{u}(p, \omega_1) = \begin{cases} 3p_1 & 0 \leq p_1 \leq 2 - \sqrt{3} \\ 1 - p_1(1 - p_1) & 2 - \sqrt{3} \leq p_1 \leq \sqrt{3} - 1 \\ 3(1 - p_1) & \sqrt{3} - 1 \leq p_1 \leq 1 \end{cases} .$$

Note that $\text{cav } f \neq f$.

The game $\Gamma_R(p, \omega_2)$ is similar the game presented in Aumann and Maschler (1995, I.2), with all payoffs multiplied by 4.⁷ As calculated in Aumann and Maschler,

$$g(p) = \tilde{u}(p, \omega_2) = 4p_1(1 - p_1).$$

As proved above, the max-min value when the initial state is ω_1 is $(\text{cav } \min\{f, g\})(p)$, while the min-max value is $\min\{\text{cav } f, g\}(p)$.

The function f is linear on both intervals $[0, 2 - \sqrt{3}]$ and $[\sqrt{3} - 1, 1]$, and convex on $[2 - \sqrt{3}, \sqrt{3} - 1]$. Since $f(2 - \sqrt{3}) = f(\sqrt{3} - 1) = 3(2 - \sqrt{3})$, $\text{cav } f$ is piecewise linear and equal to $3(2 - \sqrt{3})$ on $[2 - \sqrt{3}, \sqrt{3} - 1]$. Thus, $\text{cav } f(1/2) = 3(2 - \sqrt{3})$, and $g(1/2) = 1$, therefore $\min(\text{cav } f, g)(1/2) = 3(2 - \sqrt{3})$.

On the other hand, a straightforward computation yields

$$\min(f, g)(p) = \begin{cases} f(p) = 3p_1 & \text{if } p \leq 1/4 \\ g(p) = 4p_1(1 - p_1) & \text{if } 1/4 \leq p \leq \frac{5 - \sqrt{5}}{10} \\ f(p) = 1 - p_1(1 - p_1) & \text{if } \frac{5 - \sqrt{5}}{10} \leq p \leq \frac{5 + \sqrt{5}}{10} \\ g(p) = 4p_1(1 - p_1) & \text{if } \frac{5 + \sqrt{5}}{10} \leq p \leq 3/4 \\ f(p) = 3p_1(1 - p_1) & \text{if } 3/4 \leq p \end{cases}$$

⁷We added the actions j_3, j_4, j_5 , which do not change the calculation of the value. For our purposes, we could have multiplied all payoffs by any α , $3 < \alpha < 3/(\sqrt{3} - 1)$.

Therefore $\text{cav min}(f, g)(p) = \min(f, g)(p)$ if $p \leq \frac{5-\sqrt{5}}{10}$ or $\frac{5+\sqrt{5}}{10} \leq p$ and is linear between $\frac{5-\sqrt{5}}{10}$ and $\frac{5+\sqrt{5}}{10}$. In particular $\text{cav min}(f, g)(1/2) = 4/5$.

So in this example $\min(\text{cav } f, g)(1/2) \neq \text{cav min}(f, g)(1/2)$.

5 Incomplete Information on Both Sides

5.1 The model

We now extend our model to the case of incomplete information on both sides; that is, each player has some private information on the game that is to be played. Formally the model is extended as follows. For more details we refer to Mertens, Sorin and Zamir (1994) or Sorin (2002).

A *two-player zero-sum stochastic game with incomplete information on both sides* is described by a finite collection $(G_{k,l})_{k \in K, l \in L}$ of stochastic games, together with a distribution $p \in \Delta(K)$ and a distribution $s \in \Delta(L)$. We assume that the games $G_{k,l}$ differ only through their reward functions $g^{k,l}$, but they all have the same sets of states Ω and actions I and J , and the same transition rule q .

The game is played in stages. A pair $(k, l) \in K \times L$ is chosen according to $p \otimes s$. Player 1 is informed of k , and player 2 of l . At every stage n , the two players choose simultaneously actions $i_n \in I$ and $j_n \in J$, and ω_{n+1} is drawn according to $q(\cdot \mid \omega_n, i_n, j_n)$. Both players are informed of (i_n, j_n, ω_{n+1}) .

We will assume throughout this section that transitions are controlled by player 1.

Since the ideas are similar to the case of incomplete information on one side, we only sketch the proofs.

5.2 Related literature

The main results in this framework are related to the case with no transition (repeated games with incomplete information) and are due to Aumann, Maschler and Stearns (1968, see also Aumann and Maschler (1995)), and Mertens and Zamir (1971, 1980). As in the case of incomplete information on one side we denote by $u(p, s)$ the value of the matrix game with action sets I and J and matrix payoff $\left(\sum_{k \in K, l \in L} p^k s^l g^{k,l}(i, j) \right)_{i,j}$. Given $f : \Delta(K) \times \Delta(L) \rightarrow \mathbf{R}$, we let $\text{cav}_p f$ denote the smallest function that is above f and concave in p , and $\text{vex}_s f$ denotes the largest function that is below f and convex in s .

The min-max value of a repeated game with incomplete information exists, and is equal to $\text{vex}_s \text{cav}_p u(p, s)$. The max-min value exists and is equal to $\text{cav}_p \text{vex}_s u(p, s)$.

5.3 Partitioning the states and the average restricted game

Since player 1 controls transitions, the partition defined in Section 4 extends to this case, as well as the definition of the average restricted game $\tilde{\Gamma}_R(p, s, \omega)$ in which none of the players has any information. Denote by $\tilde{u}(p, s, \omega)$ the value of $\tilde{\Gamma}_R(p, s, \omega)$. In addition, we define the average restricted game $\tilde{\Gamma}_R^1(p, s, \omega)$ (resp. $\tilde{\Gamma}_R^2(p, s, \omega)$) in which player 1 (resp. player 2) is informed of k (resp. l) while his opponent gets no information. Our first goal is to extend Proposition 20.

Proposition 27 *For every $(\omega, p, s) \in \Omega_0 \times \Delta(K) \times \Delta(L)$, the min-max value of $\Gamma(p, s, \omega)$ exists and is equal to $\text{vex}_s \text{cav}_p \tilde{u}(p, s, \omega)$. Similarly the max-min value exists and is equal to $\text{cav}_p \text{vex}_s \tilde{u}(p, s, \omega)$.*

Proof. The proof follows the proof for repeated games with incomplete information, using the tools developed in the previous sections. We shall only sketch the arguments, and refer for details to Zamir (1992).

First, we explain how player 2 can guarantee $\text{vex}_s \text{cav}_p \tilde{u}(p, s, \omega)$. When player 2 ignores his information, he faces a game with incomplete information on one side with parameter set K and payoffs $\sum_{l \in L} s^l g^{k,l}$. By Proposition 20, player 2 can guarantee $\text{cav}_p \tilde{u}(p, s, \omega)$ in this game. Therefore by Lemma 9 (with the roles of the two players exchanged), he can also guarantee $\text{vex}_s \text{cav}_p \tilde{u}(p, s, \omega)$.

To prove that player 1 can defend $\text{vex}_s \text{cav}_p \tilde{u}(p, s, \omega)$, we adapt Zamir (1992, Theorem 4.1). Let τ be a given strategy of player 2. As in Step 2 of the proof of Lemma 25, we let player 1 play first an ε -exhausting strategy $\tilde{\sigma}$ given τ . This strategy may be chosen to be non revealing (see, e.g., Sorin (2002), ch. IV, Lemma 4.1). Player 1 switches at some stage N to a strategy that defends $\text{cav}_p \tilde{u}(p, s_N, \omega_N)$ (up to ε) in $\Gamma(p, s_N, \omega_N)$ against the continuation strategy τ_N (see Step 3 of Lemma 25). Since $\tilde{u}(\cdot, \cdot, \omega) = \tilde{u}(\cdot, \cdot, \omega_N)$, $\text{cav}_p \tilde{u}(p, s_N, \omega_N) = \text{cav}_p \tilde{u}(p, s_N, \omega)$. Therefore player 1 defends $\mathbf{E}_{p,s,\omega,\tilde{\sigma},\tau} [\text{cav}_p \tilde{u}(p, s_N, \omega)] \geq \text{vex}_s \text{cav}_p \tilde{u}(p, s, \omega)$. ■

5.4 The max-min value and the min-max value

Let \mathcal{B} denote the set of all functions $w : \Delta(K) \times \Delta(L) \times \Omega \rightarrow [0, 1]$ that are 1-Lipshitz with respect to p and s . Denote by T_4 and T_5 the operators on \mathcal{B} defined by

$$T_4 w(p, s, \omega) = \text{cav}_p \max \left\{ \text{cav}_p \text{vex}_s \tilde{u}, \max_{\omega' \in C_\omega, i \notin I_{\omega'}} \mathbf{E}[w \mid \omega', i] \right\} (p, s, \omega), \quad (20)$$

and

$$T_5 w(p, s, \omega) = \text{vex}_s \text{cav}_p \max \left\{ \text{cav}_p \tilde{u}, \max_{\omega' \in C_\omega, i \notin I_{\omega'}} \mathbf{E}[w \mid \omega', i] \right\} (p, s, \omega). \quad (21)$$

Our main result is the following.

Theorem 28 1. The mappings T_4 and T_5 have unique fixed points, denoted respectively by \underline{v} and \bar{v} .

2. The function \underline{v} is the max-min value of the game.

3. The function \bar{v} is the min-max value of the game.

Note that if player 2 has no information, there is no vex operator in (20) and (21) and both T_4 and T_5 reduce to T_1 . If player 1 has no information, there is no cav operator in (20) and (21), T_4 and T_5 reduce respectively to T_3 and T_2 with the roles of the players reversed.

Proof. The first assertion follows the same lines as the proof of Proposition 22.

We now prove the second assertion. For $j = 0, 1$, we define the sequence $(w_n^j)_{n \geq 0}$ by $w_0^j = j$ and $w_{n+1}^j = T_4 w_n^j$. We follow the inductive proof of Proposition 24 Step 1, or the first part of Lemma 26.

The sequence (w_n^0) is increasing and converges uniformly to \underline{v} . It is clear that player 1 can guarantee w_0^0 . Assuming player 1 can guarantee w_n^0 , we prove that he can guarantee w_{n+1}^0 . By Lemma 9 it is sufficient to show that he can guarantee both $\text{cav}_p \text{vex}_s \tilde{u}(p, s, \omega)$ and $\max_{\omega' \in C_\omega, i \notin I_{\omega'}} \mathbf{E}[w_n^0 \mid \omega', i]$, which is true by Proposition 27 and by Step 1 of Proposition 24.

To prove that player 2 can defend \underline{v} , we combine several ideas from the preceding sections. Let σ be given, and let \tilde{T} be the set of non revealing strategies of player 2. We let τ_σ be a non revealing strategy that ε -exhausts the information contained in σ , and choose N as previously. Denote by $\nu = 1 + \min\{n \geq 1, i_n \notin I_{\omega_n}\}$. Player 2 plays according to τ_σ up to stage $\min\{\nu, N\}$.

- If $\nu \leq N$, from stage ν on he defends $w_n^1(p_\nu, s, \omega_\nu)$.
- If $\nu > N$, we first use the idea of Lemma 26, with the roles of the two players exchanged. Specifically, we define a non revealing strategy τ_N^σ that exhausts the probability of leaving the initial communicating set, given the strategy σ_N induced by σ after stage N . Choose N' such that $\mathbf{P}_{p_N, s, \omega_N, \sigma_N, \tau_N^\sigma}(\nu > N') \leq \varepsilon$. Player 2 plays τ_N^σ up to stage $\min\{\nu, N + N'\}$.
 - If $\nu \leq N + N'$ player 2 switches to a strategy that defends $w_n(p_\nu, s, \omega_\nu) + \varepsilon$.
 - If $\nu > N + N'$, following Steps 2 and 3 of Lemma 25, player 2 starts to play in blocks of length N_1 . In the b th block he forgets past play and follows a strategy that defends $\text{vex}_s \tilde{u}(p_{N+N'+bN_1}, s, \omega_{N+N'+bN_1})$ in the restricted game $\tilde{\Gamma}_R^2(p_{N+N'+bN_1}, s, \omega_{N+N'+bN_1})$ against the average continuation strategy $\sigma_{N+N'+bN_1}^{p_{N+N'+bN_1}}$ of player 1.

We now turn to the third assertion. We first prove that player 2 can guarantee \bar{v} . By following Steps 2, 3 and 4 of Lemma 24, one proves that player 2 guarantees $\text{cav}_p \max\{\text{cav}_p \tilde{u}, \max_{\omega' \in C_\omega, i \notin I_{\omega'}} \mathbf{E}[\bar{v} \mid \omega', i]\}(p, s, \omega)$. Hence, by Lemma 9 (with the roles of the two players exchanged), he can guarantee $\text{vex}_s \text{cav}_p \max\{\text{cav}_p \tilde{u}, \max_{\omega' \in C_\omega, i \notin I_{\omega'}} \mathbf{E}[\bar{v} \mid \omega', i]\} = \bar{v}$.

We now prove that player 1 can defend \bar{v} . We first follow Step 2 of Lemma 25. Given τ , we let σ^τ be a strategy in $\tilde{\Gamma}_R^2(p, s, \omega)$ that exhausts the information contained in τ , and we choose N such that $\mathbf{E}_{p, s, \omega, \sigma^\tau, \tilde{\tau}} \left[\sum_{n=N}^{\infty} \|p_n - p_{n+1}\|_1^2 \right] \leq \varepsilon$. Player 1 plays σ^τ up to stage N . He then switches to a strategy that guarantees $\text{cav}_p \max\{\text{cav}_p \tilde{u}(\cdot, s_N, \omega), \max_{\omega' \in C_\omega, i \notin I_{\omega'}} \mathbf{E}[\bar{v} \mid \omega', i]\}(p, s_N)$ in $\tilde{\Gamma}_R^1(p, s_N, \omega_N)$, as in the proof of Proposition 24. The result follows. ■

References

- [1] Aumann R.J. and Maschler M.B. (1968) Repeated Games of Incomplete Information: the Zero-Sum Extensive Case, in *Report of the U.S. Arms Control and Disarmament Agency ST-143*, Washington D.C., Chapter III, 37-116
- [2] Aumann R.J. and Maschler M.B. (1995) Repeated Games with Incomplete Information, The MIT Press
- [3] Aumann R.J., Maschler M.B. and Stearns R.E. (1968) Repeated Games of Incomplete Information: an Approach to the Non-Zero Sum Case, in *Report of the U.S. Arms Control and Disarmament Agency ST-143*, Washington D.C., Chapter IV, 117-216
- [4] Blackwell D. (1956) An Analog of the Minimax Theorem for Vector Payoffs, *Pacific J. Math.* **6**, 1-8
- [5] De Meyer B. (1996) Repeated Games, Duality and the Central Limit Theorem, *Math. Oper. Res.*, **21**, 237-251

- [6] Filar J.A. (1981) Ordered Field Property for Stochastic Games when the Player who Controls Transitions Changes from State to State, *J. Optim. Th. Appl.*, **34**, 503-515
- [7] Laraki R. (2001) On the Regularity of the Convexification Operator on a Compact Set, Cahiers du Laboratoire d'Econometrie de l'Ecole Polytechnique, 2001-005, Paris, France
- [8] Mertens J.F. (1987) Repeated Games, *Proceedings of the International Congress of Mathematicians, Berkeley, 1986*, Gleason A.M. eds, American Mathematical Society, 1528-1577
- [9] Mertens J.F. and Neyman A. (1981) Stochastic Games, *Internat. J. Game Th.*, **10**, 53-66
- [10] Mertens J.F., Sorin S. and Zamir S. (1994) Repeated Games, CORE Discussion Paper 9420-2
- [11] Mertens J.F. and Zamir S. (1971) The Value of Two-Person Zero-Sum Repeated Games with Lack of Information on Both Sides, *Internat. J. Game Theory*, **1**, 39-64
- [12] Mertens J.F. and Zamir S. (1980) Minmax and Maxmin of Repeated Games with Incomplete Information, *Internat. J. Game Theory*, **9**, 201-215
- [13] Milman E. (2000) Uniform Properties of Stochastic Games and Approachability, Master Thesis, Tel Aviv University
- [14] Rosenberg D. and Vieille N. (2000) The Maxmin of Recursive Games with Incomplete Information on One Side, *Math. Oper. Res.*, **25**, 23-35
- [15] Ross K.W. and Varadarajan R. (1991) Multichain Markov Decision Processes with a Sample Path Constraint: A Decomposition Approach, *Math. Oper. Res.*, **16**, 195-207
- [16] Shimkin N. and Shwartz A. (1993) Guaranteed Performance Regions in Markovian Systems with Competing Decision Makers, *IEEE Trans. Automat. Control*, **38**, 84-95
- [17] Sorin S. (1984) Big Match with Lack of Information on One Side (Part 1), *Internat. J. Game Theory*, **13**, 201-255
- [18] Sorin S. (1985) Big Match with Lack of Information on One Side (Part 2), *Internat. J. Game Theory*, **14**, 173-204
- [19] Sorin S. (2002) A First Course on Zero-Sum Repeated Games, *Mathematiques et applications*, **37**, Springer
- [20] Sorin S. and Zamir S. (1991) Big Match with lack of information on one side II, in *Stochastic Games and Related Topics*, T.E.S. Raghavan et al., eds, Kluwer, 101-112
- [21] Zamir S. (1992) Repeated Games of Incomplete Information: Zero-sum, in *Handbook of Game Theory with Economic Applications*, Volume 1 (eds. Aumann R.J. and Hart S.), Elsevier Science Publishers B.V.