Zero-Sum Two-Person Repeated Games with Public Uncertain Duration Process

Abraham Neyman and Sylvain Sorin*

June 2001

Abstract

We consider repeated two-person zero-sum games where the number of repetitions θ is unknown. The information about the uncertain duration is identical to both players and can change during the play of the game. This is described by an uncertain duration process Θ. To each repeated game Γ and uncertain duration process Θ is associated the Θ repeated game Γ_Θ with value V_Θ. We establish a recursive formula for the value V_Θ. We study asymptotic properties of the value $v_\Theta = V_\Theta/E(\theta)$ as the expected duration $E(\theta)$ goes to infinity. We extend and unify several asymptotic results on the existence of $\lim v_n$ and $\lim v_\lambda$ and their equality to $\lim v_\Theta$. This analysis applies in particular to stochastic games and repeated games of incomplete information.

*Neyman: Institute of Mathematics; and Center for Rationality and Interactive Decision Theory; The Hebrew University of Jerusalem, Givat Ram, Jerusalem 91904, Israel ⟨*e-mail: aneyman@math.huji.ac.il*⟩;
This research was in part supported by The Israel Science Foundation grant 382/98.
Sorin: Laboratoire d'Econométrie, Ecole Polytechnique, 1 rue Descartes, 75005 Paris and Equipe Combinatoire, UFR 921, Université P. et M. Curie - Paris 6, 175 rue du Chevaleret, 75013 Paris, France ⟨*e-mail: sorin@poly.polytechnique.fr*⟩.

1 Introduction

We consider two-person zero-sum repeated games with an uncertain number of stages. The model thus consists of two basic components:

a) First a repeated game Γ is given and described as follows. M is a state space on which a family of normal form two-person zero-sum games is defined by move spaces I and J for Player 1 and Player 2 respectively, and payoff g from $I \times J \times M$ to $\mathbb{R}$ (all sets considered are finite). The initial state m_1 is chosen at random and the players receive some information about it, say a_1 (resp. b_1) for Player 1 (resp. Player 2). The choice of the triple (m_1, a_1, b_1) is performed according to some probability π on $M \times A \times B$, where A and B are signal sets. In addition, after each stage the players obtain some further information about the previous choice of moves and both the previous and the new state. This is represented by a map Q from $M \times I \times J$ to probabilities on $M \times A \times B$. At stage t, given the state m_t and the moves (i_t, j_t), a triple $(m_{t+1}, a_{t+1}, b_{t+1})$ is chosen at random according to the distribution $Q(m_t, i_t, j_t)$. The new state is m_{t+1}, and the signal a_{t+1} (resp. b_{t+1}) is transmitted to Player 1 (resp. Player 2). A play of the game is thus a sequence $m_1, a_1, b_1, i_1, j_1, m_2, a_2, b_2, i_2, j_2, \cdots$, while the information of Player 1 before his play at stage t is a private history of the form $(a_1, i_1, a_2, i_2, \cdots, a_t)$ and similarly for Player 2. The associated sequence of stage payoffs is $g_1, g_2, \cdots$ with $g_t = g(i_t, j_t, m_t)$. Note that a play of the game consists of a sequence of states, signals and moves and so is independent of the sequence of payoffs. The repeated game Γ is thus represented by the tuple $\langle M, I, J, g, \pi, Q, A, B \rangle$ and this description is public knowledge.

Two special cases of repeated games that have been extensively considered are stochastic games and repeated games with incomplete information. We describe a stochastic game in the simplest framework. The initial signal to the players is the state: $a_1 = b_1 = m_1$ and at each subsequent stage the signal to both players is the previous pair of moves and the new state: $a_{t+1} = b_{t+1} = \{i_t, j_t, m_{t+1}\}$. Hence, formally $A = B = M \cup I \times J \times M$. It follows that a play can be identified with a sequence $m_1, i_1, j_1, m_2, i_2, j_2, \cdots$ and the information of each player before his play at stage t is the sequence of states and moves $m_1, i_1, j_1, \cdots, i_{t-1}, j_{t-1}, m_t$. In addition, since the initial state is publicly known, the analysis is usually done conditionally to m_1; hence the initial probability π is replaced by a Dirac mass at some point in M.

As for the game with lack of information on both sides Γ (dependent case) the traditional description is as follows. For each m in M we have an $I \times J$ real-valued payoff matrix G^m. Nature chooses $m \in M$ according to a publicly known probability distribution p on M. Each player gets partial information regarding the actual state $m \in M$; Player 1 (resp. Player 2) observes the realization of a signal $\ell^1(m)$ (resp. $\ell^2(m)$). Equivalently, $M^1 = \{M_1^1, \ldots, M_C^1\}$ and $M^2 = \{M_1^2, \ldots, M_D^2\}$ are two partitions of the set M, and following the choice of $m \in M$, Player 1 is informed of c and Player 2 is informed of d where $m \in M_c^1 \bigcap M_d^2$. The state m is chosen once and for all according to p, and the game G^m is played repeatedly, with perfect monitoring. Using the general presentation above, π is the probability induced by p on $M \times C \times D$ and $g(i,j,m) = G_{ij}^m$. Finally, $Q(m_t, i_t, j_t)$ is the unit mass on (m_t, i_t, j_t) and $a_{t+1} = b_{t+1} = (i_t, j_t)$; hence formally $A = C \cup (I \times J)$ and $B = D \cup (I \times J)$. One basic distinction between the two classes is that in stochastic games the information of the players on the state space is at each stage identical while it is asymmetric in repeated games with incomplete information. More generally, we refer to any repeated game with identical information about the state as a stochastic game.

b) The second component is the number of repetitions θ, unknown to the players. θ is an integer-valued random variable defined on a probability space $(\Omega, \mathcal{B}, \mu)$ with finite expectation $E(\theta)$. The players receive partial information about the value of θ via a sequence of public signals $s_0, s_1, \cdots, s_t, \cdots$. Each signal s_t is a measurable function defined on the probability space $(\Omega, \mathcal{B}, \mu)$ and with finite range S. All these random variables are independent of those occurring in Γ. This defines an uncertain duration process $\Theta = \langle(\Omega, \mathcal{B}, \mu), (s_t)_{t \geq 0}, \theta\rangle$.

To each repeated game Γ and uncertain duration process Θ is associated an extended repeated game Γ_Θ, called the Θ-repeated game. The Θ-repeated game is played essentially like the original game Γ but, in addition, following the play at stage t, namely (i_t, j_t), and before the next play at stage $t+1$, the players receive the public signal $s_t(\omega)$. Formally, a play is now an infinite sequence $s_0, m_1, a_1, b_1, i_1, j_1, s_1, m_2, a_2, b_2, i_2, j_2, s_2, \cdots$, while the information of Player 1 corresponds to finite private histories like $(s_0, a_1, i_1, s_1, a_2, i_2, s_2, \cdots, a_t)$. Note that the sequence of payoffs $g_1, g_2, \cdots$ is not affected by the process Θ, but the total payoff in Γ_Θ is $\sum_{t=1}^{\infty} g_t I(\theta \geq t)(\omega) = \sum_{t=1}^{\theta(\omega)} g_t$. Since the play after stage t on the event $\theta \leq t$ is irrelevant, one can enlarge the signal space so that the signal s_t conveys also the information whether $\theta \leq t$ or not. Thus, we can assume that θ is a stopping time w.r.t. the filtration generated by the signals, namely the increasing sequence of fields $\mathcal{F}_t = \sigma(s_0, \cdots, s_t)$, $t = 0, 1, \cdots$.

A strategy σ for Player 1 in Γ is a map from private histories in Γ to $\Delta(I)$: probabilities on the set I of moves and τ is defined similarly for Player 2. The corresponding sets of strategies are denoted Σ and $\mathcal{T}$. A similar definition holds for Γ_Θ: a strategy σ for Player 1 in Γ_Θ is a map from private histories in Γ_Θ to $\Delta(I)$. The associated sets of strategies are Σ_Θ and $\mathcal{T}_\Theta$. A strategy $\sigma \in \Sigma$ is identified with the strategy $\widetilde{\sigma} \in \Sigma_\Theta$ that ignores the public signals $s_0, s_1, s_2, \ldots$, i.e., $\widetilde{\sigma}(s_0, a_1, i_1, s_1, \ldots, a_t) = \sigma(a_1, i_1, \ldots, a_t)$. Thus Σ and $\mathcal{T}$ are identified with subsets of Σ_Θ and $\mathcal{T}_\Theta$ respectively.
Given a repeated game $\Gamma = \langle M, I, J, g, \pi, Q, A, B \rangle$ and an uncertain duration process Θ, a pair of strategies (σ, τ) in $\Sigma_\Theta \times \mathcal{T}_\Theta$ induces a distribution on plays; hence on the sequence of payoffs.
The value $v_\Theta(\Gamma)$ of the normalized Θ-repeated game is

$$v_\Theta(\Gamma) = \frac{V_\Theta(\Gamma)}{E(\theta)}$$

with

$$\begin{aligned} V_\Theta(\Gamma) &= \max_{\sigma \in \Sigma_\Theta} \min_{\tau \in \mathcal{T}_\Theta} E_{\sigma,\tau,\mu} \left(\textstyle\sum_{t=1}^{\infty} g_t I(\theta \geq t) \right) \\ &= \min_{\tau \in \mathcal{T}_\Theta} \max_{\sigma \in \Sigma_\Theta} E_{\sigma,\tau,\mu} \left(\textstyle\sum_{t=1}^{\infty} g_t I(\theta \geq t) \right). \end{aligned}$$

The existence of $V_\Theta(\Gamma)$ follows from the usual minmax theorem: considering mixed strategies, the payoff is bilinear and (even jointly) continuous and both strategy spaces are convex and compact.

We are interested in the asymptotic behavior of $v_\Theta(\Gamma)$ as the expected duration $E(\theta)$ goes to ∞.

Remarks. 1) The above model of public uncertain duration extends naturally to a model of asymmetric uncertain duration with n players. The asymmetric uncertainty is modeled by private signals, i.e., the signal s_t is a profile $s_t = (s_t^0, s_t^1, \cdots, s_t^n)$ of signals for each player and the information to, say, Player 1 before the play at stage t is thus the private history $(s_0^1, a_1, i_1, s_1^1, a_2, i_2, s_2^1, \cdots, a_t)$. Then a strategy σ of Player 1 is a map from such histories to the set of mixed moves $\Delta(I)$.
2) The previous extension of a game via a random duration process applies as well to any multistage game with an associated sequence of stage payoffs.
3) The case of uncertain duration rather than of uncertain duration process corresponds to the signaling structure where there are no signals upon

the duration: $s_t(\omega)$ is a constant which is independent of t and ω. In this case Θ reduces to θ and the strategy sets Σ_Θ and $\mathcal{T}_\Theta$ are equal to Σ and $\mathcal{T}$. In particular the evaluation of the payoffs can be written as $\sum_t \rho_t g_t$ with $\rho_t = \mu(\theta \geq t)/E_\mu(\theta)$.
The case where θ is deterministic ($\theta = n$) is the classical n-stage repeated game. The payoff is $E_{\sigma,\tau,\mu}(\frac{1}{n}\sum_{t=1}^n g_t)$ and we write v_n for v_Θ.
The λ-discounted game corresponds to $\mu(\theta \geq t) = (1-\lambda)^{t-1}$. Since $E(\theta) = 1/\lambda$ the payoff is $E_{\sigma,\tau,\mu}(\sum_{t=1}^\infty \lambda(1-\lambda)^{t-1} g_t)$ and we use the notation v_λ for v_Θ.

2 Initial results

The next property confirms the robustness of the uniform value.
We first recall the definition. Following [5] (Chapter IV, Section 1) we say that Player 1 can guarantee w in Γ if: for any $\varepsilon > 0$, there exists a strategy σ of Player 1 in Γ and a number of stages T such that, for any strategy τ of Player 2 in Γ and any $t \geq T$,

$$E_{\sigma,\tau}(\sum_{\ell=1}^t g_\ell) \geq t\ (w - \varepsilon).$$

Similarly, Player 2 can guarantee w in Γ if: for any $\varepsilon > 0$, there exists a strategy τ of Player 2 in Γ and a number of stages T such that for any strategy σ of Player 1 in Γ and any $t \geq T$,

$$E_{\sigma,\tau}(\sum_{\ell=1}^t g_\ell) \leq t\ (w + \varepsilon).$$

The uniform value v_∞ exists if both players can guarantee it.

Theorem 1 *If Player 1 can guarantee w in Γ then* $\liminf_{E(\theta)\to\infty} v_\Theta(\Gamma) \geq w$.

In particular, if the infinite game Γ has a uniform value v_∞, then

$$\lim_{E(\theta)\to\infty} v_\Theta(\Gamma) = v_\infty.$$

Proof. Assume that Player 1 can guarantee w in Γ. Given $\varepsilon > 0$, let σ be a strategy of Player 1 in Γ and let T be a positive integer so that for every strategy τ of Player 2 in Γ and every $t \geq T$,

$$E_{\sigma,\tau}(\sum_{\ell=1}^{t} g_\ell) \geq t(w - \varepsilon).$$

Given a pure strategy τ of Player 2 in Γ_Θ and ω in Ω we denote by τ_ω the strategy of Player 2 in Γ given by

$$\tau_\omega(b_1, j_1, b_2, \cdots, b_t) = \tau(s_0(\omega), b_1, j_1, s_1(\omega), b_2, j_2, s_2(\omega), \cdots, b_t).$$

It follows that if τ is a pure strategy of Player 2 in Γ_Θ and ω in Ω satisfies $t = \theta(\omega) \geq T$, then

$$E_{\sigma,\tau_\omega}(\sum_{\ell=1}^{t} g_\ell) \geq t(w - \varepsilon)$$

and therefore for every $\omega \in \Omega$

$$E_{\sigma,\tau_\omega}(\sum_{\ell=1}^{t} g_\ell) \geq t(w - \varepsilon) - \|g\|T$$

where $\|g\| = \sup_{I\times J\times M} |g(i,j,m)|$. As $E_{\sigma,\tau,\mu}(\sum_{\ell=1}^{\theta} g_\ell) = E_\mu(E_{\sigma,\tau_\omega}(\sum_{\ell=1}^{\theta(\omega)} g_\ell))$ we deduce that for any strategy τ of Player 2 in Γ_Θ

$$E_{\sigma,\tau,\mu}(\sum_{t=1}^{\theta} g_t) \geq E(\theta)(w - \varepsilon) - \|g\|T.$$

■

Remark. The same proof holds even if the signals of the random duration process are private.

3 The extended recursive structure

3.1 Recursive structure

Shapley [14] associates to a two-person zero-sum stochastic game the Shapley operator $\mathbf{\Psi}$ that maps real-valued functions defined on the state space M to themselves:

$$\mathbf{\Psi}(f)[m] = \sup_{x\in\Delta(I)} \inf_{y\in\Delta(J)} \{g(x,y,m) + E_{x,y,m}\left[f(m')\right]\} \tag{1}$$

where $g(x,y,m)$ is the bilinear extension of $g(\cdot,\cdot,m)$ to $\Delta(I)\times\Delta(J)$ and the expectation $E_{x,y,m}$ is with respect to the law of m' given by $Q(x,y,m)$: the bilinear extension to $\Delta(I)\times\Delta(J)$ of the transition $Q(.,.,m)$. $\mathbf{\Psi}(f)[m]$ is interpreted as the value of a one-stage game played as the one-shot stochastic game with the payoff function being the sum of the stage payoff of the stochastic game and the value of the function f at the new state. The iterates of the operator $\mathbf{\Psi}$ evaluated at 0 express the values of the finitely repeated stochastic game:

$$\mathbf{\Psi}^n(0) = V_n. \tag{2}$$

A similar operator can be introduced for any repeated game considered in Section 1 ([5], Chapter IV, Section 3). In contrast to the case of stochastic games where the operator acts on functions defined on the state space of the game, the general case involves operators on functions defined on an auxiliary enlarged state space M'. In fact, to each repeated game $\Gamma = \langle M, I, J, g, \pi, Q, A, B\rangle$, one associates an auxiliary stochastic game $\Gamma' = \langle M', I', J', g', \pi', Q'\rangle$ having the same n-stage (and λ-discounted) value. A formula like (1) defines the corresponding Shapley operator and the n-stage value is expressed by the n-th iterate of the operator evaluated at 0.

The purpose of this section is to extend formulas (1) and (2) for repeated games (hence in particular also for stochastic games) with an uncertain duration process. Given an uncertain duration process Θ, equality of the type $V_\Theta(\Gamma) = \mathbf{\Psi}^\Theta(0)$ will be obtained where $\mathbf{\Psi}$ is the auxiliary Shapley operator. The generalized iterate $\mathbf{\Psi}^\Theta$ will in fact be defined for any nonexpansive mapping $\mathbf{\Psi}$.

Next we illustrate auxiliary Shapley operators $\mathbf{\Psi}$ associated with several specific classes of repeated games:

1) Repeated games with a publicly known state.

This is the class where, at every stage t, each private signal a_t or b_t contains at least the current state m_t. Since in this framework the state is public knowledge, the Shapley operator introduced in the perfect monitoring case (1) and acting on functions defined on M itself, is the auxiliary Shapley operator.

2) Repeated games with publicly known probability distribution on the state

space.
This is the class with public signals and perfect monitoring, i.e., where at every stage t, a_t equals b_t and contains at least (i_t, j_t), namely the pair of moves. The auxiliary stochastic game Γ' can be chosen as follows: the state space $M'= \Delta(M)$ is the space of probability distributions on M; $I' = I$, $J' = J$, $g'(i,j,.)$ is the linear extension of $g(i,j,.)$ to M'. Finally, for each $\mu \in M'$, $Q'(i,j,\mu)$ is the probability on M' defined as follows: let $\rho(i,j,\mu)(a) = E_\mu(Q(i,j,m)(a))$ be the probability of signal a induced in Γ by (i,j,μ). Each signal a determines, given (i,j,μ), a conditional probability distribution $\mu(a)$ on the state space M, hence a point in M'. Then $Q'(i,j,\mu)\,[\nu] = \sum_{a;\mu(a)=\nu} \rho(i,j,\mu)(a)$.
The resulting auxiliary Shapley operator is:

$$\mathbf{\Psi}(f)[\mu] = \sup_{x\in\Delta(I)} \inf_{y\in\Delta(J)} \{g'(x,y,\mu) + E_{x,y,\mu}\,[f(\mu')]\}.$$

3) Repeated games with incomplete information: the independent case with perfect monitoring.
M is a product space $K\times L$, π is a product probability $p\otimes q$ with $p\in\Delta(K)$, $q\in\Delta(L)$ and in addition $a_1 = k$ and $b_1 = \ell$. The state $m = (k,\ell)$ corresponds to the type of the players and each player knows his own type and holds a prior on the other player's type. From stage 1 on, the state is fixed and the information of the players is $a_{t+1} = b_{t+1} = \{i_t, j_t\}$.
The auxiliary stochastic game Γ' can be chosen as follows: the state space M' is $\Delta(K)\times\Delta(L)$ and is interpreted as the space of beliefs on the true state; $I' = \Delta(I)^K$ and $J' = \Delta(J)^L$ correspond to type-dependent mixed moves of the players; g' is defined on $I'\times J'\times M'$ by $g'(x,y,p,q) = \sum_{k,\ell} p^k q^\ell g(x^k, y^\ell, k, \ell)$. The transition Q' is introduced next: given (x,y,p,q), let $x(i) = \sum_k x^k_i p^k_1$ and $p(i)$ be the conditional probability on K given the move i, explicitly $p^k(i) = \frac{p^k x^k_i}{x(i)}$ (and similarly for y and q). Then $Q'(x,y,p,q)(p',q') = \sum_{i,j;(p(i),q(j))=(p',q')} x(i)y(j)$.
The resulting auxiliary Shapley operator is:

$$\mathbf{\Psi}(f)[p,q] = \sup_{x\in\Delta(I)^K} \inf_{y\in\Delta(J)^L} \{\textstyle\sum_{k,\ell} p^k q^\ell g(x^k, y^\ell, k, \ell) + \displaystyle\sum_{i,j} x(i)y(j) f(p(i), q(j))\}.$$

4) Repeated games where Player 1 knows more than Player 2.
This is the case where a_t determines b_t and j_t and w.l.o.g. determines also i_t. The auxiliary stochastic game Γ' can be chosen as follows: the state

space M' is $\Delta(\Delta(M))$. The set $S = \Delta(M)$ describes the possible beliefs of Player 1 on the state space M and the set $M' = \Delta(S)$ stands for the beliefs of Player 2 about Player 1's beliefs. $I' = \Delta(I)^S$, $J' = \Delta(J)$, $g'(x, y, \mu) = \int_S [\sum_j y_j \int_M \sum_i g(i, j, m) x_i^s s(dm)] \mu(ds)$. It remains to specify $Q'(x, y, \mu)$. A signal a of Player 1 defines (through Q) a posterior probability on M, namely a point in S. Given a signal b and a move j of Player 2, Player 2 can compute the conditional distribution on the signals a of Player 1, hence a point in M'. Finally (x, y, μ) and Q determine the law of the signals b. The resulting auxiliary Shapley operator is:

$$\mathbf{\Psi}(f)[\mu] = \sup_{x \in \Delta(I)^S} \inf_{y \in \Delta(J)} \{g'(x, y, \mu) + E_{x,y,\mu}[f(\mu')]\}.$$

Note that the above procedure aims at building from a repeated game Γ, another repeated game Γ', which is in fact a stochastic game, such that the n-stage and λ-discounted values of both games coincide. However, there is in general no direct relation between strategies in both games; in particular optimal strategies in Γ' need not necessarily have counterpart optimal strategies in the original repeated game Γ. In examples 1 and 2 above, the state space in the auxiliary stochastic game Γ' is public knowledge in the original game Γ, enabling a natural map from strategies in Γ' to strategies in Γ, preserving optimality. In example 3, the auxiliary state space corresponds to beliefs of the players, which can be computed on the basis of their type-dependent mixed moves, but which are not observable during the play of the original game Γ. Finally, in example 4, the auxiliary state space that models the probabilistic information of Player 2, relies again on the knowledge of Player 1's strategy that Player 2 does not observe in the play of Γ. However, the better-informed Player 1 can compute it, hence optimal strategies for Player 1 in Γ' translate also to optimal strategies in Γ.

We turn to the construction of the auxiliary Shapley operator for an arbitrary repeated game Γ as defined in Section 1. It relies on the recursive structure based on the representation of an information scheme, namely a probability distribution on $M \times A \times B$ (where A and B are arbitrary signal sets for Players 1 and 2 respectively) ([5], Chapter IV, Section 3).
Given M, there exists a space W with the following properties [7]:
1) There exists a homeomorphism φ from W to $\Delta(M \times W)$.
2) If W^i denotes a copy of W, any information scheme induces a probability P on $\mathcal{U} = M \times W^1 \times W^2$ such that, for $\{i, j\} = \{1, 2\}$, the conditional probability

$P(\cdot|w^i)$ on $M \times W^j$ coincides with $\varphi(w^i)$, P a.s.
The set of such probabilities (called consistent) is denoted $\mathcal{P}$. The set W^i is called the type set of player i and $\mathcal{U}$ is called the universal belief space. Given a consistent probability P, a type w^i in W^i, which corresponds to the signal to Player i, is thus identified with its image $\varphi(w^i)$ which is a distribution on $M \times W^j$ ($j \neq i$), namely on states and types of the opponent. It coincides with the beliefs he computes, given his type.

The game $\Gamma = \langle M, I, J, g, \pi, Q, A, B\rangle$ is value-equivalent to $\Gamma' = \langle M', I', J', g', \pi', Q'\rangle$ where $M' = \mathcal{P}$, the signal space to player i would be W^i and the corresponding distribution on $\mathcal{U}$ would be $P_1 = P$. Similarly, given σ and τ, strategies in Γ, the distribution on plays up to moves at stage t defines a consistent probability P_t on $\mathcal{U}$ which corresponds to the state of knowledge on M at stage t. The game Γ is value-equivalent to a new game where σ and τ are played for $t-1$ stages and where the game restarts at stage t with P_t (as a function of the play and σ and τ) as initial distribution on state and signals. The (behavioral) strategies used at stage t in Γ allow us to construct P_{t+1}. They can be replaced, for each player i, by a function of his type w^i_t at that stage that induces the same payoff. Hence the play at time t is specified by P_t and maps α_t and β_t from type sets to mixed actions. This triple determines the stage payoff via the marginal distribution on $I \times J \times M$ and P_{t+1} as a function $H(\alpha_t, \beta_t, P_t)$ in $\Delta(M \times W^1 \times W^2)$.
The extension of the Shapley operator (1) is the following operator $\mathbf{\Psi}$ defined on bounded functions on $\mathcal{P}$:

$$\mathbf{\Psi}(f)[P] = \sup_{\alpha} \inf_{\beta} \{g'(\alpha, \beta, P) + f[H(\alpha, \beta, P)]\} \tag{3}$$

where α (resp. β) is a map from W^1 to $\Delta(I)$ (resp. W^2 to $\Delta(J)$) and $g'(\alpha, \beta, P)$ denotes the expectation of $g(i, j, m)$ with respect to the probability induced by α, β and P on $I \times J \times M$. Hence the auxiliary repeated game is actually a stochastic game with a deterministic transition defined by the function H.

The previous explicit examples used this property with a reduction of both spaces $\mathcal{U}$ and $\mathcal{P}$ to subspaces $\mathcal{U}_0$ and $\mathcal{P}_0$ where the support of each probability P in $\mathcal{P}_0$ is included in $\mathcal{U}_0$.
In example 1 as well as in standard stochastic games satisfying (1), rather than dealing with probability distributions on states, the public knowledge of the state in Γ allows us to choose $M' = M$ in Γ', but the transition is no longer deterministic. A similar remark applies to example 2 where one

works with $\Delta(M)$ rather than $\Delta(\Delta(M))$. Also in the last two examples public knowledge of the moves or comparison of the information enables us to reduce the level of iterations needed when working with H.

3.2 Uncertain duration process and extended orbits of a nonexpansive operator

Let $\mathbf{\Psi}$ be a nonexpansive map from a Banach space X to itself. Following [9], we generalize here the iterates $\mathbf{\Psi}^n$ to operators $\mathbf{\Psi}^\Theta$ that act on X and are defined for an arbitrary uncertain duration process $\Theta = \langle(\Omega, \mathcal{B}, \mu), (s_t)_{t\geq 0}, \theta\rangle$. $\mathbf{\Psi}^\Theta$ captures the idea of a "generalized" random number of iterations of the nonexpansive map $\mathbf{\Psi}$. Moreover, when θ is a stopping time w.r.t. the increasing sequence of algebras $\mathcal{F}_t = \sigma(_0, \ldots, s_t)$, the domain of the operator $\mathbf{\Psi}^\Theta$ extends from X to all $\mathcal{F}_\theta$-measurable functions $x : \Omega \to X$.

To an uncertain duration process Θ, where θ is a stopping time, we associate a probability tree as follows. The terminal nodes $T = T_\Theta$ are all finite sequences of signals $\nu = (s_0, \ldots, s_t)$ with positive μ probability and $t = \theta(s_0, \ldots, s_t)$. The set of nodes $N = N_\Theta$ is the set of all initial segments $(s_0, \ldots, s_r)$, $r \leq t$, of a terminal node $(s_0, \ldots, s_t)$. The root of the tree is the node of the empty string of signals $\nu = \emptyset$. The probability measure μ on $(\Omega, \mathcal{B})$ induces a probability measure $\bar{\mu}$ on the countable set of terminal nodes T. Given a node $\nu = (s_0, \ldots, s_r)$ and an integrable function $f : \Omega \to I\!\!R$ we denote by $E(f \mid \nu)$ the value of the conditional expectation $E(f \mid \mathcal{F}_r)$ at ν.

The next theorem asserts that give an integrable function x on the terminal nodes and a nonexpansive map Ψ there is a uniquely determined Ψ iterate $\bar{x}$ defined on all non-terminal nodes.

Theorem 2 *Given a function $x : T \to X$ (which is identified with a function $x : \Omega \to X$, measurable w.r.t. $\mathcal{F}_\theta$) with finite expectation $E_{\bar{\mu}}(\|x\|) = E_\mu(\|x\|) < \infty$, and a nonexpansive map $\mathbf{\Psi} : X \to X$, there is a unique* *extension of the function x to a function $\bar{x}$ defined on all nodes N such that*

$$\bar{x}(\emptyset) = E(\bar{x}(s_0)), \tag{4}$$

and for every non-terminal node $\nu = (s_0, \ldots, s_r) \neq \emptyset$,

$$\bar{x}(\nu) = \mathbf{\Psi}(E(\bar{x}(\nu, s_{r+1}) \mid \nu)) \tag{5}$$

$$\|\bar{x}(\nu)\| \leq E(\theta - r \mid \nu)\|\mathbf{\Psi}(0)\| + E(\|x\| \mid \nu). \tag{6}$$

In addition, given two functions $x, y : T \to X$ with finite expectation, the following inequality holds:

$$\|\bar{x}(\nu) - \bar{y}(\nu)\| \leq E(\|x - y\| \mid \nu). \tag{7}$$

Proof. If $\theta = 0$, equation (4) defines uniquely $\bar{x}(\emptyset)$ and the inequalities (6) and (7) hold. Assume that $E(\theta) > 0$. We first prove the lemma for the case where θ is bounded, by induction on the number of nodes in N_Θ. For every node $\nu = (s_0, \ldots, s_r)$ let $k(\nu) = \max\{\theta - r \mid s_0, \ldots, s_r\}$. Let $\nu \neq \emptyset$ be a node with $k(\nu) = 1$; equivalently, ν is a maximal non-terminal node (i.e., all successor nodes are terminal nodes). Given two functions x, y from T_Θ to X, equation (5) defines $\bar{x}(\nu)$ and $\bar{y}(\nu)$ and $\|\bar{x}(\nu) - \bar{y}(\nu)\| \leq E(\|y - x\| \mid \nu)$ by nonexpansiveness. Therefore, by the induction hypothesis, $\bar{x}$ and $\bar{y}$, which are uniquely defined on all nodes by backward induction using (5), will satisfy for any other non-terminal node ν', $\|\bar{x}(\nu') - \bar{y}(\nu')\| \leq E(\|x - y\| \mid \nu')$, i.e., (7) holds. (6) holds as well by backwards induction. In fact, fix a node ν with $k(\nu) = 1$; note that $\bar{x}(\nu)$ is defined by (5), and replace it by a terminal node. This defines a stopping time θ' with associated process Θ'. Set $y : T_{\Theta'} \to X$ by $y(\nu) = \bar{x}(\nu)$ and $y = x$ at all other terminal nodes in $T_{\Theta'}$. Note that $\|y(\nu)\| = \|\bar{x}(\nu)\| \leq \|\mathbf{\Psi}(0)\| + E(\|x\| \mid \nu)$. At any node $\nu' = (s_0, \cdots, s_r)$ in $N_{\Theta'} \setminus T_{\Theta'}$, one has $\bar{x}(\nu') = \bar{y}(\nu')$ so that $\|\bar{x}(\nu')\| = \|\bar{y}(\nu')\| \leq E(\theta' - r \mid \nu')\|\mathbf{\Psi}(0)\| + E(\|y\| \mid \nu')$ hence $\|\bar{x}(\nu')\| \leq E(\theta' - r \mid \nu')\|\mathbf{\Psi}(0)\| + E(\|x\| \mid \nu') + \text{Prob}(\nu \mid \nu')\|\mathbf{\Psi}(0)\| = E(\theta - r \mid \nu')\|\mathbf{\Psi}(0)\| + E(\|x\| \mid \nu')$.

Assume finally that θ is unbounded with finite expectation. Fix a node $\nu = (s_0, \ldots, s_r)$ and a sufficiently large $n \geq r$. Define the stopping time $\theta \wedge n$ and let $\Theta \wedge n = \langle (\Omega, \mathcal{B}, \mu), (s_t)_{t\geq 0}, \theta \wedge n \rangle$ be the associate uncertain duration process. Consider y and z, two functions on $T_{\Theta\wedge n}$ that coincide with x on $T_{\Theta\wedge n} \cap T_\Theta$ and such that for every $\nu' \in T_{\Theta\wedge n}$,

$$\|\bar{y}(\nu')\| + \|\bar{z}(\nu')\| \leq 2E((\theta - n)^+ \mid \nu')\|\mathbf{\Psi}(0)\| + 2E(\|x\| \mid \nu').$$

It follows that

$$\|\bar{y}(\nu) - \bar{z}(\nu)\| \leq E(\|y - z\| \mid \nu) \leq 2E((\theta - n)^+ \mid \nu)\|\mathbf{\Psi}(0)\| + 2E(\|x\| I(\theta > n) \mid \nu)$$

and the upper bound goes to 0 as $n \to \infty$. In particular, if y_n coincides with x on $T_{\Theta\wedge n} \cap T_\Theta$ and equals 0 on $T_{\Theta\wedge n} \setminus T_\Theta$, $\bar{y}_n(\nu)$ converges to a limit denoted by $\bar{x}(\nu)$. The last argument provesexistence of the extension and the previous one shows uniqueness.

■

Definition. The Θ-iterate of $\mathbf{\Psi}$ is defined on the set of $\mathcal{F}_\theta$-measurable function $x : \Omega \to X$ by:

$$\mathbf{\Psi}^\Theta(x) = \bar{x}(\emptyset).$$

Comments. Equations (5), (6)and (7) are in fact true in a more general setting. Let θ' be another $\mathcal{F}_t$-stopping time and write Θ' for the associated uncertain duration process. Assume $\theta' \leq \theta$ so that $T_{\Theta'} \subset N_\Theta$. Define a $\mathcal{F}_{\theta'}$-measurable function y by $y(\nu) = \bar{x}(\nu)$ for ν in $T_{\Theta'}$. Then

$$\mathbf{\Psi}^\Theta(x) = \mathbf{\Psi}^{\Theta'}(y). \tag{8}$$

If θ and θ' are two stopping times (w.r.t. $(\mathcal{F}_t)_{t\geq 0}$) with finite expectations, $\mathbf{\Psi}^\Theta(0) = \mathbf{\Psi}^{\Theta\wedge\Theta'}(y)$ where $E(\|y\|) \leq E(\theta - (\theta \wedge \theta'))\|\mathbf{\Psi}(0)\|$ and thus $\|\mathbf{\Psi}^\Theta(0) - \mathbf{\Psi}^{\Theta\wedge\Theta'}(0)\| \leq E(\theta - (\theta \wedge \theta'))\|\mathbf{\Psi}(0)\|$. Similarly, $\|\mathbf{\Psi}^{\Theta'}(0) - \mathbf{\Psi}^{\Theta\wedge\Theta'}(0)\| \leq E(\theta' - (\theta \wedge \theta'))\|\mathbf{\Psi}(0)\|$. As $|\theta - \theta'| = \theta - (\theta \wedge \theta') + \theta' - (\theta \wedge \theta')$, we have,

$$\|\mathbf{\Psi}^{\Theta'}(0) - \mathbf{\Psi}^\Theta(0)\| \leq E(|\theta' - \theta|)\|\mathbf{\Psi}(0)\|.$$

If $\nu = (s_0, \ldots, s_r)$ is a non-terminal node of the uncertain duration process Θ, we denote by $\Theta(\nu)$ the remaining uncertain duration process after ν (in particular $\theta(\nu) = \theta - r$). The associated probability tree is thus the sub-tree with root ν, endowed with the corresponding conditional probability. If ν

is a terminal node we identify the identity operator with $\mathbf{\Psi}^{\Theta(\nu)}$. With this notation one has for every $\nu \in N_\Theta$,

$$\bar{x}(\nu) = \mathbf{\Psi}^{\Theta(\nu)}(x),$$

and thus in particular we obtain, for any non-terminal node $\nu = (s_0, \ldots, s_r) \neq \emptyset$:

$$\mathbf{\Psi}^{\Theta(\nu)}(\cdot) = \mathbf{\Psi}(E(\mathbf{\Psi}^{\Theta(\nu, s_{r+1})}(\cdot)) \mid \nu). \tag{9}$$

Note that equation (6) is needed for uniqueness only in the case where θ is unbounded. However uniqueness follows also with a weaker requirement: $\|\bar{x}(\nu)\| \leq KE(\theta - r \mid \nu)\|\mathbf{\Psi}(0)\| + KE(\|x\| \mid \nu)$ for some constant $K \geq 1$. Nevertheless, some bound is needed. Indeed, if θ is unbounded and $(s_0, s_1, \ldots)$ is an infinite sequence of signals so that $\mu(s_0, \ldots, s_t) > 0$ for every t, for any $z \in X$ we can find a function $\bar{y} : N_\Theta \to X$ that coincides with our defined $\bar{x}$ on all nodes ν which are not initial segments of the infinite sequence $(s_0, s_1, \ldots)$ and so that $\bar{y}(\emptyset) = z$ and $\bar{y}$ obeys (5). Indeed, define inductively $\bar{y}(s_0, \ldots, s_t)$ so that (5) holds also for $\bar{y}(s_0, \ldots, s_{t-1})$.

3.3 Uncertain duration process and extended recursive formula

Here we establish an extended recursive formula for the value of the Θ-repeated game Γ_Θ.
Since the law of θ is independent of the moves and states, one obtains that the value of the game Γ extended by the uncertain duration process Θ satisfies the following extension of (2):

Theorem 3

$$V_\Theta(\Gamma) = \mathbf{\Psi}^\Theta(0) \tag{10}$$

where $\mathbf{\Psi}$ *is given by (3).*

Proof. As the duration signals are public, following, e.g., ([5], Chapter IV, Section 3), the equality holds for any bounded uncertain duration process

Θ. Moreover, as $V_{\Theta\wedge n}$ and $\mathbf{\Psi}^{\Theta\wedge n}(0)$ converge to V_Θ and $\mathbf{\Psi}^{\Theta}(0)$ respectively, the equalities $V_{\Theta\wedge n} = \mathbf{\Psi}^{\Theta\wedge n}(0)$ imply that $V_\Theta = \mathbf{\Psi}^{\Theta}(0)$. ∎

Equation (9) implies that for any non-terminal node $\nu = (s_0, \cdots, s_r) \neq \emptyset$,

$$V_{\Theta(\nu)}(\Gamma) = \mathbf{\Psi}(E(V_{\Theta(\nu,s_{r+1})}(\Gamma) \mid \nu)) \tag{11}$$

and from (4) one has:

$$V_\Theta(\Gamma) = E(V_{\Theta(s_0)}(\Gamma)).$$

In particular, this recursive formula implies the following property on optimal strategies.

Theorem 4 *Assume that the state variable $P_t \in \mathcal{P}$ is public knowledge. Then each player has an optimal strategy which at each stage t is only a function of P_t and $\Theta(\nu)$, $\nu \in \mathcal{F}_t$.*

To be more specific, consider a stochastic game with a publicly known state, as previously defined in Section 3.1, example 1. The above result implies that both players have optimal strategies that depend only upon the remaining uncertain duration process $\Theta(\nu)$ and the current state m_t. Hence the value v_Θ is the same whatever the additional information on moves may be. However, in the case of full monitoring or at least when the signals a_t and b_t allow us to compute g_{t-1}, Mertens and Neyman [4] proved the existence of a uniform value. Theorem 1 thus implies:

Corollary 1 *In a stochastic game with a publicly known state, $\lim_{E(\theta)\to\infty} v_\Theta$ exists and is independent of the set of uncertain duration processes Θ.*

4 Operator approach

In this section we extend previously known inequalities of the values of the n-stage game v_n and the λ-discounted game v_λ to corresponding inequalities of the values v_Θ of the repeated games with uncertain duration processes.

4.1 Variational bounds

The nonexpansive operators arising in repeated games act on spaces of real-valued functions endowed with the uniform norm and in addition these operators are monotonic. Let $\mathbf{\Psi}$ be a monotonic nonexpansive mapping on a vector space $\mathcal{F}$ of bounded real functions that contain the constants. We first recall a definition from [16] extending [12].

Definition 1 *$\mathcal{L}^+$ is the set of functions f in $\mathcal{F}$ for which there exists L such that:*

$$\mathbf{\Psi}(Kf) \leq (K+1)f, \qquad \forall K \geq L.$$

Such a function yields an upper bound for the iterates: $\mathbf{\Psi}^n(0) \leq nf + 2L\|f\|$ which implies that $\limsup_{n\to\infty} \frac{\mathbf{\Psi}^n(0)}{n} \leq f$ for any $f\in\mathcal{L}^+$, see [12] and [16] . The next result generalizes this inequality to any uncertain duration process.

Theorem 5 *Assume $f\in\mathcal{L}^+$. For any uncertain duration process Θ,*

$$\mathbf{\Psi}^{\Theta}(0) \leq E(\theta)f + 2L\|f\|.$$

Proof. Assume $f\in\mathcal{L}^+$ and let L be the constant associated to f. We prove the stronger property that for any uncertain duration process Θ and every function $K : T_\Theta \to I\!R$ with $K \geq L$ and $E(K) < \infty$,

$$\mathbf{\Psi}^{\Theta}(Kf) \leq (E(K) + E(\theta))f. \tag{12}$$

Obviously, (12) holds when $E(\theta) = 0$. The proof of (12) for bounded uncertain duration processes (with $E(\theta) > 0$) is by induction on N_Θ, the number of nodes of the probability tree associated with Θ. Let Θ' be an uncertain duration process (defined on the same duration signal space $(\Omega, , \mu, (s_t)_{t\geq 0})$) with $\theta - 1 \leq \theta' < \theta$. Thus $|N_{\Theta'}| < |N_\Theta|$. Define the function $K' : N_{\Theta'} \to I\!R$ by $K'(\nu) = E(\theta - r) + E(K \mid \nu)$ for a terminal node $\nu = (s_0, \ldots, s_r)$ of Θ'. As $f \in \mathcal{L}^+$, we have $\mathbf{\Psi}^{\Theta(\nu)}(Kf) \leq (E(\theta - r \mid \nu) + E(K \mid \nu))f = K'(\nu)f$. As $K'(\nu) = E(\theta - r \mid \nu) + E(K \mid \nu) \geq L$, the induction hypothesis implies that $\mathbf{\Psi}^{\Theta'}(K'f) \leq (E(\theta') + E(K'))f$. Note that $E(\theta') + E(K') = E(\theta) + E(K)$ and thus $\mathbf{\Psi}^{\Theta'}(K'f) \leq (E(\theta)+E(K))f$. As $\mathbf{\Psi}^{\Theta}(Kf) = \mathbf{\Psi}^{\Theta'}(K'f)$ we conclude that $\mathbf{\Psi}^{\Theta}(Kf) \leq (E(\theta)+E(K))f$. We prove (12) for an unbounded duration process by truncation: define the function $K \wedge n$ on the terminal nodes of $\Theta \wedge n$ by $(K \wedge n)(\nu) = E(K \mid \nu)$. As $\mathbf{\Psi}^{\Theta \wedge n}((K \wedge n)f) \to_{n\to\infty} \mathbf{\Psi}^{\Theta}(Kf)$ and $E(\theta \wedge n) + E(K \wedge n) \to_{n\to\infty} E(\theta)+E(K)$, (12) holds for any uncertain duration process. ∎

A function f belongs to $\mathcal{C}^+$ if it satisfies: For all $\delta > 0$, there exists L_δ such that:

$$\mathbf{\Psi}(Kf) \leq (K+1)f + \delta, \qquad \forall K \geq L_\delta.$$

Similarly a function f belongs to $\mathcal{C}^-$ if it satisfies: For all $\delta > 0$, there exists L_δ such that:

$$\mathbf{\Psi}(Kf) \geq (K+1)f - \delta, \qquad \forall K \geq L_\delta.$$

If a function f belongs to $\mathcal{C}^+$ then

$$\mathbf{\Psi}(K(f+\delta)) \leq (K+1)(f+\delta), \qquad \forall K \geq L_\delta,$$

and thus $f+\delta$ belongs to $\mathcal{L}^+$ for all $\delta > 0$; one obtains then an upper bound:

Corollary 2 *Let $f \in \mathcal{C}^+$, then*

$$\limsup_{E(\theta)\to\infty} \frac{\mathbf{\Psi}^{\Theta}(0)}{E(\theta)} \leq f.$$

We now apply this property to continuous absorbing games: these are stochastic games where only one state, say m, is nonabsorbing. The action spaces are compact sets X and Y. Given (x, y) in $X \times Y$ the game starting from m remains in stage m with probability $q(x, y)$ and the payoff in this event is $g(x, y)$. Otherwise there is an absorbing payoff $\rho(x, y)$. These three functions are continuous on $X \times Y$. Proposition 7 and Corollary 8 in [12] prove that in this case the intersection of the closure of $\mathcal{C}^+$ and of $\mathcal{C}^-$ is nonempty; hence reduced to one point.

Corollary 3 *If Γ is a continuous absorbing game, there exists a real w such that*

$$\lim_{E(\theta)\to\infty} v_\Theta(\Gamma) = w.$$

In the same spirit let us consider continuous recursive games. These are stochastic games where the set M_0 of nonabsorbing states is finite and the payoff is 0 at each of these states. The action spaces are compact sets X and Y. The absorbing payoffs as well as the transitions are continuous on $X \times Y$. From Proposition 16 in [15], which proves that in this case also the intersection of the closure of $\mathcal{C}^+$ and of $\mathcal{C}^-$ is nonempty, one deduces the following result:

Corollary 4 *If Γ is a continuous recursive game, there exists w in $\mathbb{R}^{M_0}$ such that:*

$$\lim_{E(\theta)\to\infty} v_\Theta(\Gamma) = w.$$

4.2 Bounded variation of v_λ

We first recall a result from [9] dealing with a nonexpansive mapping $\mathbf{\Psi}$ on a Banach space X. Define $v_n = \frac{\mathbf{\Psi}^n(0)}{n}$ and for $0 < \lambda \leq 1$ let $V_\lambda = \frac{v_\lambda}{\lambda}$ be the (unique) fixed point of the mapping $x \mapsto \mathbf{\Psi}((1-\lambda)x)$.

Definition 2 *The function $\lambda \mapsto v_\lambda$ is of bounded variation (over $(0,1]$) if there exists a constant C such that for any increasing sequence λ_i with $0 < \lambda_i \leq \lambda_{i+1} \leq 1$,*

$$\sum_i \|v_{\lambda_{i+1}} - v_{\lambda_i}\| \leq C.$$

If the function $\lambda \mapsto v_\lambda$ has bounded variation then v_λ converges to a limit w as $\lambda \to 0^+$ and it is shown further in [9] that it implies the convergence of v_n to the same limit.
We will establish here a similar property under an additional monotonicity hypothesis on the uncertain duration process.

Definition 3 *The uncertain duration process is monotonic if for every terminal node $\nu = (s_0, \ldots, s_r)$, the conditional expectations $E(\theta - t \mid s_0, \ldots, s_t)$ decrease in t, $0 \leq t \leq r$.*

The interpretation is that the expected remaining duration decreases over time, implying that the relative weight of the present increases as the process evolves. Typical examples include finite length (where the ratio is $1/n$ if the remaining duration is n) and discounted factor uncertain duration (where the ratio is the constant λ).
We follow the proof in [9] to obtain:

Theorem 6 *Assume v_λ is of bounded variation. Set $w = \lim_{\lambda \to 0+} v_\lambda$. For every $\varepsilon > 0$, there exists N such that for any monotonic duration process satisfying $E(\theta) > N$:*

$$\|v_\Theta - w\| \leq \varepsilon.$$

Proof. On the event $\{t < \theta\}$ we set $\rho_t = E(\theta - t \mid s_0, \ldots, s_t)$, $\lambda_t = 1/\rho_t$, $w_t = v_{\lambda_t}$. On the event $\{\theta \leq t\}$ we set $\rho_t = 0$ and $w_t = 0$. We will prove that

$$\|\mathbf{\Psi}^\Theta(0) - V_{1/E(\theta)}\| \leq E(\sum_{t \geq 0} \rho_{t+1} \|w_{t+1} - w_t\|). \tag{13}$$

For every $t \geq 0$ we define the random variables U_t and W_t as follows: on the event $\{t < \theta\}$ we define $U_t = \mathbf{\Psi}^{\Theta(s_o, \ldots, s_t)}(0)$ and $W_t = V_{\lambda_t} = v_{\lambda_t}/\lambda_t$; on the event $\{t \geq \theta\}$ we set $U_t = W_t = 0$. As θ is a stopping time the X-valued random variables U_t and W_t are measurable w.r.t. $\mathcal{F}_t$.
On the event $\{t < \theta\}$, $U_t = \mathbf{\Psi}(E(U_{t+1} \mid \mathcal{F}_t))$, $W_t = \mathbf{\Psi}((1 - \lambda_t)W_t)$ and $E(\rho_{t+1} \mid \mathcal{F}_t) = \rho_t - 1$. Therefore using the nonexpansiveness of $\mathbf{\Psi}$ followed by the triangle inequality and thereafter the equality $E(\rho_{t+1} \mid \mathcal{F}_t) = \rho_t - 1$, we have on $\theta > t$:

$$\begin{aligned}
\|U_t - W_t\| &\leq \|E(U_{t+1} \mid \mathcal{F}_t) - ((1 - \lambda_t)W_t\| \\
&\leq \|E(U_{t+1} - W_{t+1} \mid \mathcal{F}_t)\| + \|E(W_{t+1} \mid \mathcal{F}_t) - (1 - \lambda_t)W_t\| \\
&\leq E(\|U_{t+1} - W_{t+1}\| \mid \mathcal{F}_t) + \|E(\rho_{t+1} w_{t+1} - (\rho_t - 1)w_t \mid \mathcal{F}_t)\| \\
&\leq E(\|U_{t+1} - W_{t+1}\| \mid \mathcal{F}_t) + E(\rho_{t+1}\|w_{t+1} - w_t\| \mid \mathcal{F}_t).
\end{aligned}$$

On the event $\theta \leq t$, $\|U_t - W_t\| = 0$. Therefore we have everywhere:

$$\|U_t - W_t\| \leq E(\|U_{t+1} - W_{t+1}\| \mid \mathcal{F}_t) + E(\rho_{t+1}\|w_{t+1} - w_t\| \mid \mathcal{F}_t).$$

Summing the expectations (conditional on $\mathcal{F}_0$) of the above inequalities over $t \geq 0$ we deduce that:

$$\|U_0 - W_0\| \leq E(\|U_M - W_M\| \mid \mathcal{F}_0)) + E(\sum_{0 \leq t < M} \rho_{t+1}\|w_{t+1} - w_t\| \mid \mathcal{F}_0).$$

As $\|U_M\| + \|W_M\| \leq 2\rho_M\|\mathbf{\Psi}(0)\|$, $E(\|U_M - W_M\| \mid \mathcal{F}_0) \leq 2E(\rho_M \mid \mathcal{F}_0)\|\mathbf{\Psi}(0)\|$ which converges to 0 as $M \to \infty$ and therefore:

$$\|U_0 - W_0\| \leq E(\sum_{t \geq 0} \rho_{t+1}\|w_{t+1} - w_t\| \mid \mathcal{F}_0).$$

Observe that $\rho_0 = E(\theta)$ and therefore $\mathbf{\Psi}^{\Theta}(0) = E(U_0)$ and $V_{1/E(\theta)} = E(W_0)$, which proves (13).
Fix $\varepsilon > 0$ and let K be sufficiently large so that the variation of v_λ over the interval $(0, 1/K)$ is less than ε. Assume that $E(\theta) > K$ and let θ' be the smallest r so that $E(\theta - r \mid (s_0, \ldots, s_r)) < K$. We have

$$\|\mathbf{\Psi}^{\Theta}(0) - V_{1/E(\theta)}\| \leq E(\sum_{0 \leq t < \theta'} \rho_{t+1}\|v_{\lambda_{t+1}} - v_{\lambda_t}\|) + E(\sum_{t \geq \theta'} \rho_{t+1}\|v_{\lambda_{t+1}} - v_{\lambda_t}\|).$$

The monotonicity of the uncertain duration process Θ implies that the sequence λ_t is monotonic. Therefore, for any $t \geq 0$, $\rho_t \leq \rho_0$, and for every $t \geq \theta'$, $\rho_{t+1} \leq K$. Hence:

$$\begin{aligned}\|\mathbf{\Psi}^{\Theta}(0) - V_{1/E(\theta)}\| &\leq E(\rho_1 \sum_{0 \leq t < \theta'} \|v_{\lambda_{t+1}} - v_{\lambda_t}\|) + E(K \sum_{t \geq \theta'} \|v_{\lambda_{t+1}} - v_{\lambda_t}\|) \\ &\leq \rho_1 \varepsilon + KC,\end{aligned}$$

where C bounds the variation of the function v_λ.
Thus if $E(\theta) > KC/\varepsilon$ we deduce that $\|v_\Theta - v_{1/E(\theta)}\| < 2\varepsilon$ implying that $\|v_\Theta - w\| < 3\varepsilon$. ∎

The inequality (13) has an alternative formulation using the probability tree associated with the uncertain duration process Θ. For every terminal node $\nu = (s_0, \ldots, s_r)$ define

$$f(\nu) = \sum_{t=1}^{r} (r - t)\|w(s_0, \ldots, s_t) - w(s_0, \ldots, s_{t-1})\|$$

where $\rho(s_0,\ldots,s_t) = E(\theta - t \mid s_0,\ldots,s_t)$ and $w(s_0,\ldots,s_t) = v_{1/\rho(s_0,\ldots,s_t)}$ if $(s_0,\ldots,s_t)$ is a non-terminal node and $= 0$ if $(s_0,\ldots,s_t)$ is a terminal node. Recall that $\bar{\mu}$ is the probability induced on the terminal nodes. The alternative formulation is:

$$\|\mathbf{\Psi}^{\Theta}(0) - V_{1/E(\theta)}\| \leq E_{\bar{\mu}}(f(\nu))$$

5 Game with lack of information on both sides

We consider here games with incomplete information as defined in Section 1. In the case of finitely repeated games, it is proved in [6] that $v(p) = \lim_{n\to\infty} v_n(p)$ exists. Moreover, the error term $\|v_n - v\|$ is bounded by a constant times $\frac{1}{\sqrt{n}}$. For the λ-discounted game it is also proved in [6] that $\|v_\lambda - v\|$ is bounded by a term of the order of $\sqrt{\lambda}$.

The purpose of this section is to extend this result of [6] to general public uncertain duration process.

Theorem 7 *The limit of $v_\Theta(p)$ as $E(\theta) \to \infty$ exists, equals $v(p)$ and*

$$\|v_\Theta - v\| \leq O\left(\frac{1}{\sqrt{E(\theta)}}\right).$$

Proof. We use the minmax theorem. It is thus sufficient to prove that there is a constant R such that for every uncertain duration process Θ and every strategy τ of Player 2, there is a strategy σ of Player 1 such that

$$E_{p,\sigma,\tau,\mu}\left(\sum_{t\geq 1} g_t I(\theta \geq t)\right) \geq v(p)E(\theta) - R\sqrt{E(\theta)}. \tag{14}$$

We denote by $\mathcal{H}_t$ the σ-algebra generated by the sequence of moves $i_1, j_1, \ldots, i_{t-1}, j_{t-1}$ and by the sequence $s_1,\ldots,s_t$ of public signals prior to the play at stage t. W.l.o.g. we assume that the event $\theta \geq t$ is measurable w.r.t. $\mathcal{H}_t$. Let $\|G\| = \max_{k,i,j} |G^k_{ij}|$.

Let Θ and τ be given. From Theorem 4.4 and the proof of Proposition 4.3 in [6], there exists a strategy σ of Player 1 and a martingale

$p_1, \tilde{p}_1, \ldots, p_t, \tilde{p}_t, \ldots$, with values in $\Delta(K)$ and $p_1 = p$, and where p_t is measurable w.r.t. $\mathcal{H}_t$, such that

$$v(p_t) \text{ is a submartingale,} \tag{15}$$

$$E_{p,\sigma,\tau,\mu}(g_t \mid \mathcal{H}_t) \geq v(p_t) - \|G\| E(\|p_{m+1} - \tilde{p}_t\|_1 \mid \mathcal{H}_t), \tag{16}$$

and for any $\ell \geq t$, conditionally on $\mathcal{H}_t$, the following pairs of random variables are independent:

$$I(\theta \geq \ell) \text{ and } g_t \tag{17}$$

$$I(\theta \geq \ell) \text{ and } p_{t+1} - \tilde{p}_t, \tag{18}$$

and

$$I(\theta \geq \ell) \text{ and } p_{t+1}. \tag{19}$$

Therefore, by (16), (17), (18) and (19),

$$E_{p,\sigma,\tau,\mu}(\sum_{t\geq 1} I(\theta \geq t) g_t) = E_{p,\sigma,\tau,\mu}(\sum_{t\geq 1} E(I(\theta \geq t) g_t \mid \mathcal{H}_t)) \geq$$

$$\sum_{t\geq 1} E_{p,\sigma,\tau,\mu}(I(\theta \geq t) v(p_t)) - \|G\| \sum_{t\geq 1} E_{p,\sigma,\tau,\mu}(E(I(\theta \geq t)\|p_{t+1} - \tilde{p}_t\|_1 \mid \mathcal{H}_t)).$$

In what follows we write E for short instead of $E_{p,\sigma,\tau,\theta}$. Note that

$$E(\sum_{t\geq 1} E(I(\theta \geq t)\|p_{t+1} - \tilde{p}_t\|_1 \mid \mathcal{H}_t)) = E(\sum_{k\in K} \sum_{t\geq 1} I(\theta \geq t)|p_{t+1}(k) - \tilde{p}_t(k)|).$$

But we have:

Lemma 1 *For any martingale $\{q_t\}$ with values in $[0,1]$ and expectation q:*

$$\sum_{m\geq 1} E(|q_{m+1} - q_m|^2) \leq q(1-q)$$

$$E(\sum_{m\geq 1} I(\theta{\geq}m)|q_{m+1} - q_m|) \leq \sqrt{E(\theta)}\sqrt{q(1-q)}.$$

Proof. The first inequality follows from the fact that the differences $(q_{m+1} - q_m)$ of the martingale $\{q_t\}$ are uncorrelated. Hence:

$$\sum_{m\geq 1}^{M} E(|q_{m+1} - q_m|^2) = E((\sum_{m\geq 1}^{M} q_{m+1} - q_m)^2) \leq E((q_{M+1} - q)^2)$$

and the variance of q_{M+1} is at most $q(1-q)$.
For the second property one has:

$$E(\sum_{m\geq 1} I(\theta\geq m)|q_{m+1} - q_m|) = \sum_{m\geq 1} E(I(\theta\geq m)|q_{m+1} - q_m|),$$

hence, by using twice the Cauchy Schwarz inequality:

$$\leq \sum_{m\geq 1} \sqrt{E(I(\theta \geq m))E(|q_{m+1} - q_m|^2)}$$

$$\leq \sqrt{\sum_{m\geq 1} E(I(\theta \geq m)) \sum_{m\geq 1} E(|q_{m+1} - q_m|^2)}$$

$$\leq \sqrt{E(\theta)}\sqrt{q(1-q)}$$

by using the previous bound. ∎
Hence:

$$E(\sum_{t\geq 1} E(I(\theta \geq t)\|p_{t+1} - \tilde{p}_t\|_1 \mid \mathcal{H}_t)) \leq \sum\nolimits_k (E(\theta))^{1/2} \sqrt{p^k(1-p^k)}.$$

Next we show that:

$$E(\sum_{t\geq 1} I(\theta \geq t)v(p_t)) \geq E(\theta)v(p_1).$$

Fix $\ell \geq t \geq 1$. As $I(\theta \geq \ell)$ and p_t are independent conditionally on $\mathcal{H}_{t-1}$, and as $v(p_t)$ is a submartingale,

$$E(I(\theta \geq \ell)v(p_t) \mid \mathcal{H}_{t-1}) \geq E(I(\theta \geq \ell)v(p_{t-1}) \mid \mathcal{H}_{t-1}).$$

Therefore,

$$E(I(\theta \geq t)v(p_t)) \geq E(I(\theta \geq t))v(p_1),$$

so that formula (14) holds with $R = \#K\|G\|$. ∎

Comments
In the framework of games with state-independent signalling matrices [3], the proof of Theorem 4 implies, using Lemma 4.6 in ([5], p.355), that:

$$\|v_\Theta - v\| \leq O(E(\theta)^{-\frac{1}{3}}).$$

6 Open Problems

Among other classes where the existence of the limit of the value of the discounted game has been established are recursive games with incomplete information on both sides [13] and absorbing games with lack of information on one side [10]. In both cases one has in addition $\lim_{\lambda\to 0} v_\lambda = \lim_{n\to\infty} v_n$. The asymptotic behavior of v_Θ in such games deserves further research.

In the framework of general dynamic programming, [2] proves the equivalence between the uniform convergence of the functions v_n and the uniform convergence of the functions v_λ. Whether this property has an extension to some family v_Θ is unknown. However, in the special class of finite action and state spaces with signals, the result of [11] on the existence of a uniform value proves in particular together with our Theorem 1 that v_Θ converges to a limit as $E(\Theta) \to \infty$. Partial results extending this equivalence property to models with uncertain duration are in [8].

References

[1] Aumann R.J. and M. Maschler (1995), *Repeated Games with Incomplete Information*, with the collaboration of R. Stearns, MIT Press.

[2] Lehrer E. and S. Sorin (1992), A Uniform Tauberian Theorem in Dynamic Programming, *Mathematics of Operations Research*, **17**, 303-307.

[3] Mertens J.-F. (1971), The Value of Two-Person Zero-Sum Repeated Games: The Extensive Case, *International Journal of Game Theory*, **1**, 217-227.

[4] Mertens J.-F. and A. Neyman (1981), Stochastic Games, *International Journal of Game Theory*, **10**, 53-66.

[5] Mertens J.-F., S. Sorin and S. Zamir (1994), *Repeated Games*, C.O.R.E. D.P. 9420, 9421, 9422.

[6] Mertens J.-F. and S. Zamir (1971), The Value of Two-Person Zero-Sum Repeated Games with Lack of Information on Both Sides, *International Journal of Game Theory*, **1**, 39-64.

[7] Mertens J.-F. and S. Zamir (1985), Formulation of Bayesian Analysis for Games with Incomplete Information, *International Journal of Game Theory*, **14**, 1-29.

[8] Monderer D. and S. Sorin (1993), Asymptotic Properties in Dynamic Programming, *International Journal of Game Theory*, **22**, 1-11.

[9] Neyman A. (1998), Nonexpansive Mappings and Stochastic Games, preprint.

[10] Rosenberg D. (1999), Zero-Sum Absorbing Games with Incomplete Information on One Side: Asymptotic Analysis, *SIAM Journal on Control and Optimization*, **39**, 208-225.

[11] Rosenberg D., E. Solan and N. Vieille (2001), Blackwell Optimality in Markov Decision Processes with Partial Observation, preprint.

[12] Rosenberg D. and S. Sorin (2001), An Operator Approach to Zero-Sum Repeated Games, *Israel Journal of Mathematics*, **121**, 221-246.

[13] Rosenberg D. and N. Vieille (2000), The Maxmin of Recursive Games with Lack of Information on One Side, *Mathematics of Operations Research*, **25**, 23-35.

[14] Shapley L.S. (1953), Stochastic Games, *Proceedings of the National Academy of Sciences of the U.S.A*, **39**, 1095-1100.

[15] Sorin S. (2000), Operator Approach to Stochastic Games, in *Stochastic Games and Applications*, A. Neyman and S. Sorin (eds.), Kluwer Academic Publishers, to appear.

[16] Sorin S. (2001), Asymptotic Properties of Monotonic Nonexpansive Mappings, preprint.