

# ABSORBING GAMES WITH A SIGNALLING STRUCTURE

JEAN-MICHEL COULOMB

*The University of York*  
*York, UK.*

**Abstract.** The perfect monitoring assumption is replaced by a signalling structure that models, for each player, the information about the opponent's previous action. The value of a stochastic game may not exist any more. For the class of absorbing games, we are able to show the existence of the max min and the min max. The value does not exist if they are different.

## 1. Introduction

### 1.1. KOHLBERG'S EXAMPLE

Let us consider the *Big Match* game (BM) (introduced by [10]):

	L	R
T	$1^*$	$0^*$
B	0	1

where the absorbing payoffs are indicated by a star. Player 1 is the maximizer and player 2 is the minimizer. Suppose that player 1 is never informed of the previous action of player 2. Kohlberg [11] observes that, for any  $\eta > 0$ , against any strategy  $\sigma$  of player 1, player 2 has a response  $\tau$  such that  $\limsup_{k \rightarrow \infty} \gamma_k(\sigma, \tau) \leq \eta$ . Remember that  $\gamma_k(\sigma, \tau)$  is the expected average of the first  $k$  payoffs.

The argument is as follows. Note that the probability  $\mathbb{P}_\sigma(\theta < k)$  of absorption before stage  $k$  is independent of the actions of player 2. The construction of an  $\eta$ -best response of player 2 is divided in two cases:

- (i)  $\mathbb{P}_\sigma(\theta < \infty) = 1$ .
- (ii)  $\mathbb{P}_\sigma(\theta < \infty) < 1$ .

In case (i), player 2 always selects R. Absorption occurs with probability 1 and the absorbing payoff is  $0^*$ .

In case (ii), let  $\bar{K} > 0$  be such that:

$$IP_\sigma(\infty > \theta > \bar{K}) < \eta/2. \quad (1)$$

Inequality (1) means that the probability of absorption after stage  $\bar{K}$  is very small. Also, the payoff received during the first  $\bar{K}$  stages has a small impact on the average payoff of the first  $k$  stages when  $k$  is large enough.

Player 2 selects R till stage  $\bar{K}$  and thereafter selects L. Thus, for  $k$  sufficiently large so that  $\bar{K}/k < \eta/2$ , we have

$$\gamma_k(\sigma, \tau) \leq \bar{K}/k + IP_\sigma(\infty > \theta > \bar{K}) < \eta.$$

## 1.2. ORGANIZATION OF THE CHAPTER

Section 2 presents the originality of the present chapter: a signalling structure (Definition 1). This is a rupture with the usual assumption of perfect monitoring and requires the elaboration of a new approach. The main result is the existence of the max min of any absorbing game with a signalling structure.<sup>1</sup> It is formulated twice, though in a different way: Theorem 1 is a general statement underlying the depth of the problem and Theorem 5.1 [8] says that the max min is expressed by (5). Surely, there are dual statements for the min max. Chronologically, Theorem 2 was established first [4]. It is the same statement but in the case where the players do not have any information at all about their opponent's past actions.

Theorems 3 and 4 [7] are used as a stepping stone for Theorem 5. Again, these are particular cases of Theorem 5: the underlying absorbing game is a *Generalized Big Match* (GBM) as defined by Definition 4.

Section 6 develops the functional approach of BM games [5]. It is an innovation that became quite convenient when dealing with signalling structures.

The definition of the max min  $\underline{v}$  involves two conditions (Section 3.1): (i) player 1 guarantees  $\underline{v}$  and (ii) player 2 defends  $\underline{v}$ . On the one hand, to prove (i), we will construct a strategy of player 1 that is effective (in a precise sense) against all the strategies of player 2. On the other hand, to prove (ii), we will construct a response of player 2 against any given strategy of player 1. Notice that for (i) it is possible to establish a strong property of uniformity with respect to the strategies of player 2. However,

<sup>1</sup>There is actually a highly non-trivial adaptation of Mertens and Neyman's argument [13] that generalizes this result to any finite stochastic game with a signalling structure [9], [15].

for (ii) such a thing is not exactly relevant since the construction of the response of player 2 depends on the strategy of player 1.

The proof of Theorem 5 is planned as follows.

- In order to show that player 1 guarantees (5), one needs to prove first that player 1 guarantees (3) in a GBM game. This is done in Sections 6 and 7.
- We demonstrate that player 2 defends (5) in Sections 5.2 and 5.3. For a GBM game, an adaptation of these arguments would prove that player 2 defends (3). The key ideas are illustrated by examples (Sections 1.1 and 4.1).

## 2. Framework

### 2.1. NOTATIONS

Recall that an absorbing game  $\Gamma$  consists of:

- A finite set  $A$  (resp.  $B$ ) of actions  $a$  (resp.  $b$ ) for player 1 (resp. 2).
- A probability  $p : A \times B \rightarrow [0, 1]$  of absorption. Note that it can be linearly extended to  $\Delta(A) \times \Delta(B)$ .
- An absorbing payoff  $r^* : A \times B \rightarrow \mathbb{R}$  and a non-absorbing payoff  $r : A \times B \rightarrow \mathbb{R}$ . Both are extended to  $\Delta(A) \times \Delta(B)$  in the following way: for  $(\alpha, \beta) \in \Delta(A) \times \Delta(B)$

$$p(\alpha, \beta)r^*(\alpha, \beta) = \sum_{(a,b) \in A \times B} \alpha(a)\beta(b)p(a, b)r^*(a, b)$$

and

$$(1 - p(\alpha, \beta))r(\alpha, \beta) = \sum_{(a,b) \in A \times B} \alpha(a)\beta(b)(1 - p(a, b))r(a, b).$$

Observe that these extensions are neither bilinear nor properly defined for every pair  $(\alpha, \beta)$  of mixed actions. This is not a problem, because if  $p(\alpha, \beta) = 0$  for instance, then the conditional absorbing payoff  $r^*(\alpha, \beta)$  does not play any role!

### 2.2. SIGNALLING STRUCTURE

In the standard information case (perfect monitoring), each player is informed of his opponent's previous action (and remembers it). The set of *signalling structures* explores a wider range of information received by the players about their opponent's action. Formally, let  $L_1$  be a finite set of signals for player 1.

**Definition 1** *A signalling structure for player 1 is a mapping*

$$\psi_1 : A \times B \rightarrow \Delta(L_1).$$

*A signalling structure is deterministic if  $\psi_1(A \times B) \subset L_1$  ( $L_1$  here stands for a subspace of  $\Delta(L_1)$ ).*

After each stage, assuming that the pair  $(a, b)$  was chosen by the players, a lottery of distribution  $\psi_1(a, b)$  is performed and player 1 is informed of its outcome instead of being informed of  $b$ . Observe that the players are supposed to remember their own action. We find it convenient to include his own previous action in player 1's signal.

**Definition 2** *A strategy of player 1 is a sequence  $\sigma = (\sigma_k)_{k \geq 1}$  of mappings*

$$\sigma_k : L_1^{k-1} \rightarrow \Delta(A).$$

A set of dual definitions for player 2 would introduce  $L_2, \psi_2$  and strategies  $\tau$ . A signalling structure for the game is a pair  $\psi = (\psi_1, \psi_2)$ .

**Example 1** The perfect monitoring corresponds to the (deterministic) signalling structure  $\delta = (\delta_1, \delta_2)$  with  $\delta_i : (a, b) \mapsto (a, b)$  ( $i = 1, 2$ ).

The case where no information is received (Kohlberg's example) corresponds to the (deterministic) signalling structure  $\pi = (\pi_1, \pi_2)$  with  $\pi_1 : (a, b) \mapsto a$  and  $\pi_2 : (a, b) \mapsto b$ .

**Definition 3** *An absorbing game with a signalling structure is a pair  $(\Gamma, \psi)$ , where  $\Gamma$  is an absorbing game and  $\psi$  is a signalling structure.*

The difficulty created by a signalling structure is twofold:

- (a) Even if the signalling structure is deterministic, player 1 ignores the actual action of player 2. A signal says only that the action of player 2 is in a given set of possible actions.
- (b) One signal is usually not sufficient because player 1 should assess average distributions of signals induced by the actions of player 2.

### 3. Results

#### 3.1. EXISTENCE OF THE MAX MIN

Two different signalling structures create two different games. To make this point clearer, consider  $(\Gamma, \psi)$  and let us denote by  $G^1(\psi)$  the set of  $\gamma \in \mathbb{R}$  having the following property: player 1 has a strategy  $\sigma$  and there exists  $K > 0$  such that for all  $k > K$  and any strategy  $\tau$  of player 2:

$$\gamma_k(\sigma, \tau) > \gamma,$$

where, as in Section 1.1,  $\gamma_k(\sigma, \tau)$  is the expected average payoff of the first  $k$  stages, provided that the players use  $(\sigma, \tau)$ . Similarly, there is a set  $G^2(\psi)$  for player 2.

Obviously, the closure of  $G^1(\psi)$  is an interval  $\bar{G}^1(\psi) = ]-\infty, \underline{v}(\psi)]$ . By duality, the closure of  $G^2(\psi)$  is an interval  $\bar{G}^2(\psi) = [\bar{v}(\psi), -\infty[$ . By definition, player 1 (resp. 2) *guarantees*  $\gamma$  if  $\gamma \in \bar{G}^1(\psi)$  (resp.  $\gamma \in \bar{G}^2(\psi)$ ).

The example of Section 1.1 shows that for  $(BM, \delta)$  and  $(BM, \pi)$ , the sets  $\bar{G}^1(\delta) = ]-\infty, 1/2]$  (by [3]) and  $\bar{G}^1(\pi) = ]-\infty, 0]$  are different.

In order to show that  $(\Gamma, \delta)$  (with perfect monitoring) has a value [11], one finds a quantity  $v$  such that: (i) player 1 guarantees  $v$  and (ii) player 2 guarantees  $v$ . Observe that  $\underline{v}(\delta) = \bar{v}(\delta) = v$  and that (i) and (ii) are dual properties.

Considering  $(\Gamma, \psi)$  (without perfect monitoring), a question is: what is  $\underline{v}(\psi)$  (resp.  $\bar{v}(\psi)$ )? To answer this question we shall proceed as follows. First, one finds a quantity  $\underline{v}$  such that (i') player 1 guarantees  $\underline{v}$ .

Second, one would like to show that if  $\gamma > \underline{v}$  then  $\gamma \notin \bar{G}^1(\psi)$ , that is, for any strategy  $\sigma$  of player 1, for any  $k > 0$  large enough, there exists a strategy  $\tau = \tau(\sigma, k)$  such that

$$\gamma_k(\sigma, \tau) \leq \gamma.$$

This implies that  $\underline{v} = \underline{v}(\psi)$ .

Actually, we are going to show a stronger property: (ii') player 2 *defends*  $\underline{v}$ . This means that for any strategy  $\sigma$  of player 1, there exists a strategy  $\tau = \tau(\sigma)$  such that

$$\limsup_{k \rightarrow \infty} \gamma_k(\sigma, \tau) \leq \underline{v}.$$

If (i') and (ii') are satisfied, then  $\underline{v}$  is called the max min of  $(\Gamma, \psi)$ . Note that it is unique.

To show that the max min exists is a difficult task in the sense that we have to characterize  $\underline{v}$  and explicitly construct strategies of player 1 and 2.

**Theorem 1** 1.  $(\Gamma, \psi)$  has a max min denoted by  $\underline{v}$ .

2. It depends only on  $\psi_1$  (the signalling structure of player 1).

By inversion of the role played by the players, Theorem 1 would imply that any game with a signalling structure has a min max denoted by  $\bar{v} = \bar{v}(\psi)$  which depends only on the signalling structure of player 2. Observe that:

$$\underline{v} \leq v \leq \bar{v}.$$

### 3.2. NO INFORMATION CASE

The characterization of the max min is not simple in general. However, when the signalling structure is  $\pi$  (no information case), the max min has a straightforward expression.

Let us define  $\Phi : \Delta(A) \times \Delta(B) \rightarrow \mathbb{R}$  by

$$\Phi(\alpha, \beta) = \begin{cases} r^*(\alpha, \beta) & \text{if } p(\alpha, \beta) > 0 \\ r(\alpha, \beta) & \text{otherwise.} \end{cases}$$

**Theorem 2** *The max min of  $(\Gamma, \pi)$  is*

$$\underline{v} = \sup_{\alpha} \min_b \Phi(\alpha, b).$$

Note that there is a dual formula for  $\bar{v}$ .

#### 4. Generalized Big Match (GBM) Games

##### 4.1. AN EXAMPLE

Let  $\Gamma$  be the following absorbing game:

	L	C	R
T	100	1*	0*
B	100	0	1

In the game  $(\Gamma, \delta)$  (perfect monitoring), player 2 has no interest in playing L. This is because player 1 guarantees 1/2 by ignoring stages at which player 2 plays L. It is as if player 1 played the BM game with a very high payoff (100) from time to time. Player 1 can ignore those stages because he knows when player 2 has selected L.

The situation is radically different when player 1 ignores the action chosen by player 2. Let us modify  $\delta$  by assuming that instead of  $\delta_1$ , the (deterministic) signalling structure of player 1 is

$$\psi_1 : \{T, B\} \times \{L, C, R\} \rightarrow \{\ell_1, \ell_2\},$$

defined by  $\psi(T, L) = \psi(B, L) = \psi(B, C) = \ell_1$  and  $\psi(B, R) = \ell_2$ . The signal we assign to  $(T, C)$  and  $(T, R)$  does not matter since there is absorption. The max min of  $(\Gamma, \psi)$  is zero because it turns out that L enlarges the set of strategies available to player 2. Since all payoffs are positive, it is enough to show that player 2 defends 0. Let us construct an  $\eta$ -best response against a given strategy  $\sigma$  of player 1.

First, observe that if player 1 selects B and receives the signal  $\ell_1$  then it is not possible for him to say whether player 2 has played L or C. In particular,  $\beta_1 = (1 - \eta/2, 0, \eta/2)$  and  $\beta_2 = (0, 1 - \eta/2, \eta/2)$  induce exactly the same distribution of signals on row B. Somehow  $\beta_1$  and  $\beta_2$  are “equivalent.” A general definition is given in Section 4.3 (Definition 5). Note that  $\beta_1$  insures the absorbing payoff  $0^*$  and that  $\beta_2$  insures the non-absorbing payoff  $\eta/2$ .

The argument of Section 1.1 is repeated but in a more sophisticated manner since the probability  $\mathbb{P}_{\sigma,\tau}(\theta < k)$  of absorption before stage  $k$  depends on  $\sigma$  and  $\tau$  and not only on  $\sigma$  (strategy of player 1).

Assume that  $\tau$  consists of playing identically and independently distributed (i.i.d.) the mixed action  $\beta_1$ . Distinguish between (i) and (ii) as in Section 1.1. If (i) holds then  $\tau$  is a best response. If (ii) holds then a large stage  $\bar{K}$  is fixed as in (1). Observe that the probability of player 1 selecting T after stage  $\bar{K}$  is negligible. Suppose that instead of playing (i.i.d.)  $\beta_1$ , player 2 plays (i.i.d.)  $\beta_2$  after stage  $\bar{K}$ , thus characterizing a modified strategy  $\tau'$  that is an  $\epsilon$ -best response.

#### 4.2. DEFINITION

**Definition 4** A Generalized BM game (GBM) is made of two rows. Up to some permutation of the columns, there is no loss of generality in assuming that it has the following form:

	$B_1$			$B_2$			$B_0$		
$T$	.	...	.	★	...	★	★	...	★
$B$	.	...	.	.	...	.	★	...	★

where a ★ indicates that the probability of absorption is strictly positive.  $B_1 = [1, n]$  is the set of non-absorbing columns,  $B_2 = [n + 1, m]$  is the set of absorbing columns and  $B_0 = [m + 1, s]$  is the set of completely absorbing columns.

Further restrictions are made:

**Assumption 1** For any  $b \in B_1 \cup B_2 \cup B_0$

$$r^*(T, b) = r^*(B, b) = r^*(b)$$

and

$$r(T, b) = r(B, b) = r(b).$$

In the sequel,  $r(\cdot)$  and  $r^*(\cdot)$  are extended to  $\Delta(B_1 \cup B_2)$  as follows:  $r(\beta) = r(B, \beta)$  and  $r^*(\beta) = r^*(T, \beta)$ .

GBM games are considered in [11] as an intermediate step when solving absorbing games (with perfect monitoring). Considering [13], this can be avoided.

#### 4.3. CHARACTERIZATION OF THE MAX MIN

Let  $(GBM, \psi)$  be such that  $B_0 = \emptyset$ , i.e., there is no completely absorbing column available to player 2.

Let us denote by  $\phi : B_1 \cup B_2 \rightarrow \Delta(L)$  (which is linearly extended to  $\Delta(B_1 \cup B_2)$ ) the signalling structure  $\psi_1(B, \cdot)$  of player 1 that is induced by  $\psi_1$  on  $B$ .

**Definition 5** *Two mixed actions  $\beta \in \Delta(B_1 \cup B_2)$  and  $\beta' \in \Delta(B_1 \cup B_2)$  are equivalent if  $\phi(\beta) = \phi(\beta')$ . In such a case let us write  $\beta \sim \beta'$ .*

Two mixed actions are equivalent if they induce the same distribution of signals for player 1 on  $B$ . Let  $\underline{r}^* : \Delta(B_1 \cup B_2) \rightarrow \mathbb{R}$  be the mapping:

$$\underline{r}^*(\beta) = \begin{cases} r^*(\beta) & \text{if } p(T, \beta) > 0 \\ r(\beta) & \text{otherwise.} \end{cases}$$

The signalling structure plays a role in the characterization of the max min only through the equivalence  $\sim$ .

**Theorem 3**

$$\underline{v} = \inf_{\beta \sim \beta'} \max[\underline{r}^*(\beta), r(\beta')]. \quad (2)$$

Let us denote by  $GBM'$  the game obtained by adding a set  $B_0$  of completely absorbing columns. Using  $\underline{v}$ , the max min  $\underline{v}'$  of  $(GBM', \psi)$  has a simple expression.

**Theorem 4**

$$\underline{v}' = \min \left[ \min_{b \in B_0} r^*(b), \underline{v} \right]. \quad (3)$$

Note that if  $T$  and  $B$  of  $GBM'$  are the same, then  $B_2 = \emptyset$  and the max min is given by

$$\min \left[ \min_{b \in B_0} r^*(b), \min_{b \in B_1} r(b) \right] \quad (4)$$

## 5. Absorbing Games

### 5.1. CHARACTERIZATION OF THE MAX MIN

In order to express the max min  $\underline{v}$  of any  $(\Gamma, \psi)$ :

1. We construct a set of auxiliary GBM games and we obtain a set  $M$  of max min.
2.  $\underline{v}$  is the supremum of  $M$ .

Let us denote by  $R(\alpha)$  the (virtual) row induced by playing i.i.d. the mixed action  $\alpha \in \Delta(A)$ . For a column  $b \in B$ , the absorbing (resp. non-absorbing) payoff is  $r^*(\alpha, b)$  (resp.  $r(\alpha, b)$ ) and the probability of absorption is  $p(\alpha, b)$ . Note that  $\psi_1$  can be linearly extended on  $R(\alpha)$ :

$$\psi_1(\alpha) = \psi_1(\alpha, \cdot) : B \rightarrow \Delta(L).$$

Let us slightly perturb  $\alpha$  with a pair  $(\eta, \alpha') \in ]0, 1[ \times \Delta(A)$ , thus obtaining  $\bar{\alpha} = \bar{\alpha}(\eta)$  defined by

$$\bar{\alpha}(\eta) = (1 - \eta)\alpha + \eta\alpha'.$$

By looking at the non-absorbing and completely absorbing columns, observe that the game obtained by playing either  $R(\alpha)$  or  $R(\bar{\alpha})$  is close (up to  $\eta$ ) to a GBM game (satisfying Assumption 1), denoted by  $\Gamma(\alpha, \alpha')$ , satisfying the conditions of Assumption 1. Let us denote by  $\underline{v}(\alpha, \alpha')$  the max min of  $(\Gamma(\alpha, \alpha'), \psi(\alpha))$ . We have the following result:

**Theorem 5** *The max min of  $(\Gamma, \psi)$  is given by:*

$$\underline{v} = \sup_{\alpha, \alpha'} \underline{v}(\alpha, \alpha'). \quad (5)$$

**Proof.** Here we prove only the first part: player 1 guarantees  $\underline{v}$  in  $\Gamma$ . The second part is dealt with in Sections 5.2 and 5.3.

In  $(\Gamma(\alpha, \alpha'), \psi(\alpha))$ , an  $\eta/2$ -optimal strategy  $\sigma$  of player 1 specifies a probability of playing T (resp. B) for any given history.

With  $\bar{\alpha} = \bar{\alpha}(\eta/2)$ ,  $\sigma$  induces an  $\eta$ -optimal strategy  $\tilde{\sigma}$  in  $\Gamma$ : if T (resp. B) of  $(\Gamma(\alpha, \alpha'), \psi(\alpha))$  is selected then player 1 chooses his action according to  $\bar{\alpha}$  (resp.  $\alpha$ ). ■

## 5.2. RESPONSE AGAINST A MIXED ACTION

What we do in this section and the next one is to make explicit a general formulation of the idea of Section 1.1 or Section 4.1. In both cases, the underlying game is a GBM. It is easy to find a pair of equivalent mixed actions of player 2 by looking at rows T and B. In particular, the equivalence between two mixed actions is clearly defined by row B. When considering an absorbing game, we construct a pair of mixed actions  $(\beta, \bar{\beta}) \in \Delta(B) \times \Delta(B)$  of player 2 as a response against a given mixed action  $\alpha$  of player 1. The mixed actions  $\beta$  and  $\bar{\beta}$  are equivalent on a subset of  $\alpha$ 's support.

Formally, fix  $\eta > 0$  and let us define two mappings  $\mu : \Delta(A) \rightarrow \Delta(B)$  (resp.  $\bar{\mu} : \Delta(A) \rightarrow \Delta(B)$ ) as follows.

If there exists  $b \in B$  such that  $p(\alpha, b) = 0$  and  $r(\alpha, b) \leq \underline{v} + \eta/2$  then  $\mu(\alpha) = \bar{\mu}(\alpha) = b$ . Otherwise,  $B^*(\alpha) = \{b \in B | p(\alpha, b) > 0, r^*(\alpha, b) \leq \underline{v} + \eta/2\}$  is a non-empty set because (notations of 5.1) the max min  $\underline{v}(\alpha, \alpha)$  of  $\Gamma(\alpha, \alpha)$  is given by (4) and satisfies  $\underline{v}(\alpha, \alpha) \leq \underline{v}$ . There exists  $\kappa > 0$  such that, if we choose  $\bar{\eta} \in (0, \eta/2)$  small enough, then the following proposition holds:

**Proposition 1** *Any  $\alpha \in \Delta(A)$  can be decomposed in such a way that*

$$\max_{b \in B^*(\alpha)} p(\alpha, b) < \bar{\eta} \quad (6)$$

into two mixed actions of disjoint support  $\alpha^{(1)}$  and  $\alpha^{(2)}$ , i.e.,

$$\alpha = \lambda\alpha^{(1)} + (1 - \lambda)\alpha^{(2)}$$

so that there exists  $(\beta, \bar{\beta}) \in \Delta(B) \times \Delta(B)$  satisfying:

1.  $p(\alpha^{(1)}, \beta) \geq \kappa$ .
2.  $r^*(\alpha^{(1)}, \beta) \leq \underline{v} + \eta/2$ .
3.  $p(\alpha^{(2)}, \beta) = p(\alpha^{(2)}, \bar{\beta}) = 0$ .
4.  $\psi(\alpha^{(2)}, \beta) = \psi(\alpha^{(2)}, \bar{\beta})$ .
5.  $r(\alpha^{(2)}, \bar{\beta}) \leq \underline{v} + \eta/2$ .

- If (6) holds, then let us define  $\mu(\alpha) = \beta$  and  $\bar{\mu}(\alpha) = \bar{\beta}$ .
- If (6) does not hold, then let us choose  $b \in B^*(\alpha)$  such that:

$$p(\alpha, b) \geq \bar{\eta} \quad (7)$$

and define  $\mu(\alpha) = \bar{\mu}(\alpha) = b$ .

Observe that if the signalling structure is  $\pi$  (no information case) then the conclusion of Proposition 1 translates as follows: the pair  $(\beta, \bar{\beta})$  of mixed actions satisfying items 1 - 5 can be chosen as a pair of actions  $(b, \bar{b}) \in B \times B$  and note that item 4 is always true.

**Sketch of Proposition 1's proof.** We rely on “ $Q$ -saturation” where  $Q > 0$  is a sufficiently large constant depending on  $\eta > 0$ . This means that starting from  $B^{(0)} = B^*(\alpha)$ , we construct a (finite) family of increasing subsets of  $B$

$$B^{(0)} \subset B^{(1)} \subset \dots$$

that remain unchanged and equal to a subset  $\bar{B}$  after a while. More precisely, suppose that  $B^{(k)}$  is constructed. Define the following subset of  $A$ :

$$A^{(k)} = \{a \in A \mid \exists b \in B_k, p(a, b) > 0\}.$$

The subset  $B^{(k+1)}$  is defined as:

$$B^{(k)} \cup \{b \in B \mid \sum_{a \in A \setminus A_k} \alpha(a)p(a, b) \leq Q \sum_{a \in A_k} \alpha(a)\}.$$

The subset

$$\bar{A} = \{a \in A \mid \exists b \in \bar{B}, p(a, b) > 0\}$$

characterizes  $\alpha^{(2)} = \alpha(\bar{A})$  ( $\alpha$  reduced to  $\bar{A}$  and renormalized) and  $\alpha^{(1)} = \alpha(A \setminus \bar{A})$ . The existence of a pair of mixed actions  $(\beta, \bar{\beta})$  is deduced from the study of the GBM game induced by  $\alpha^{(1)}$  and  $\alpha^{(2)}$ .

### 5.3. CONSTRUCTION OF AN $\eta$ -BEST RESPONSE

Given a strategy  $\sigma = (\sigma_k)_{k \geq 1}$  of player 1, let us describe an  $\eta$ -best response  $\tau$  of player 2. As we shall point out,  $\tau$  is a sequence of mixed actions, indicative of the fact that player 2 does not need any information about player 1's actions to construct an  $\eta$ -best response. The pair of mappings  $(\mu, \bar{\mu})$  is the basic building block of the strategy  $\tau$ . We proceed as follows.

1. A sequence of mixed actions  $\tilde{\alpha} = \alpha_1, \alpha_2, \dots$  (resp.  $\tilde{\beta} = \beta_1, \beta_2, \dots$ ) of player 1 (resp. of player 2) is constructed recursively with  $\alpha_1 = \sigma_1$ . Suppose that  $\alpha_1, \dots, \alpha_k$  and  $\beta_1, \dots, \beta_{k-1}$  are determined. Define

$$\beta_k = \mu(\alpha_k)$$

and then, introduce:

$$\alpha_{k+1} = \mathbb{E}[\sigma_{k+1} | \theta > k]$$

assuming that player 1 (resp. player 2) uses  $\sigma$  (resp.  $\beta_1, \dots, \beta_k$ ).

2. If  $\mathbb{P}_{\sigma, \tilde{\beta}}(\theta < \infty) = 1$  then  $\tau$  consists of playing  $\tilde{\beta}$ .
3. Otherwise, i.e.,  $\mathbb{P}_{\sigma, \tilde{\beta}}(\theta < \infty) < 1$ , let  $\bar{K} > 0$  be an integer such that

$$\frac{\mathbb{P}_{\sigma, \tilde{\beta}}(\infty > \theta > \bar{K})}{1 - \mathbb{P}_{\sigma, \tilde{\beta}}(\theta < \infty)} < \bar{\eta}/2. \quad (8)$$

If we denote by  $\bar{\beta}_k$  the mixed action  $\bar{\mu}(\alpha_k)$ , then  $\tau$  consists of playing

$$\beta_1, \dots, \beta_{\bar{K}}, \bar{\beta}_{\bar{K}+1}, \bar{\beta}_{\bar{K}+2}, \dots$$

By construction, when player 2 uses  $\tilde{\beta}$ , playing  $\tilde{\alpha}$  or  $\sigma$  is the same. Note that, by (8) for any  $k > \bar{K}$ , either  $B^*(\alpha_k) = \emptyset$  or (6) holds. In the latter case, the mixed action  $\alpha_k = \lambda_k \alpha_k^{(1)} + (1 - \lambda_k) \alpha_k^{(2)}$  amounts to the choice of the (1) (resp. the (2)) row with probability  $\lambda_k$  (resp.  $1 - \lambda_k$ ) and then the mixed action  $\alpha_k^{(1)}$  (resp.  $\alpha_k^{(2)}$ ).

By 1) of Proposition 1, the total probability of choosing the (1) row after stage  $\bar{K}$  is negligible. By 3) and 4)  $\beta_k$  and  $\bar{\beta}_k$  do not absorb against  $\alpha_k^{(2)}$  and induce the same distribution of signals. Therefore, playing  $\bar{\beta}_k$  instead of  $\beta_k$  cannot be detected by player 1 with a probability close to one. It implies that against  $\tau$ ,  $\sigma$  and  $\tilde{\alpha}$  induce a close expected payoff. Observe that  $\tau$  generates an absorbing payoff under  $\underline{v} + \eta/2$  by 2) when it matters and a non-absorbing payoff under  $\underline{v} + \eta/2$  after stage  $\bar{K}$  by 5).

## 6. The Functional Approach of BM Games

In the present section, we make the assumption of perfect monitoring in order to study a technique which takes care of difficulty (b) of Section 2.2 when employed in the framework of signalling structures.

## 6.1. PRINCIPLE OF THE STRATEGY

Let us consider the following BM game:

	L	R
T	$r^*(1)$	$r^*(2)$
B	$r(1)$	$r(2)$

with the perfect monitoring assumption (the signalling structure is  $\delta$ ). Note that its value  $v$  is

$$v = \min_{\beta_1 + \beta_2 = 1} \max [\beta_1 r^*(1) + \beta_2 r^*(2), \beta_1 r(1) + \beta_2 r(2)] \quad (9)$$

and that it is a particular case of (2). Without loss of generality, let us assume that  $v = 0$  and that the absolute value of the payoffs is less than  $1/8$ .

**Lemma 1** (9) leads to three different cases:

- (i)  $r^*(b) \geq 0$  for  $b = L, R$ .
- (ii)  $r(b) \geq 0$  for  $b = L, R$ .
- (iii) There exists  $\zeta > 0$  such that  $r^*(\beta) \geq -\zeta r(\beta)$  for any  $\beta \in \Delta(B)$ . As we will later on show, there is no loss of generality in assuming that  $\zeta < 1$  (see Section 6.4).

Cases (i) or (ii) are easy because player 1 should play either T or B all the time. In case (iii), real difficulties appear. If player 1 selects T then he takes a risk since he may have a strictly negative payoff forever. On the contrary, playing B all the time is not a good idea since player 2 may choose the column with a strictly negative payoff.

The idea consists of playing T with a small probability that depends on the history. If the actual choice of player 1 is T then the play is over. Otherwise (player 1 selects B), the probability of playing T and the action chosen by player 2 measure the risk that player 1 has taken. Player 1 should monitor the overall risk during a play. How shall we quantify it in order to obtain some “robustness” with respect to signal processing when the assumption of perfect monitoring is removed? We propose to introduce the functions  $f(\cdot)$  and  $\epsilon(\cdot)$  defined as follows ( $M > 0$  is a constant chosen large enough):

$$f(\varrho) = \begin{cases} \frac{1}{1+(M+\varrho)^2} & \text{if } M + \varrho \geq 4 \\ 1/17 & \text{otherwise.} \end{cases}$$

$$\epsilon(\varrho) = \begin{cases} -\frac{1}{M+\varrho} & \text{if } M + \varrho \geq 4 \\ -1/4 & \text{otherwise.} \end{cases}$$

The function  $f(\cdot)$  helps to calculate the probability of playing T at a given stage and  $\epsilon(\cdot)$  counts the overall risk that is allowed from some stage on. There is some similarity with [3] but the choice of the functions  $f(\cdot)$  and  $\epsilon(\cdot)$  is justified by a deep relationship between the two functions:

**Lemma 2** *For any  $|\eta| \leq 1/4$ :*

$$\eta f(\varrho) + (1 - f(\varrho))\epsilon(\varrho - \eta) \geq \epsilon(\varrho). \quad (10)$$

## 6.2. DEFINITION OF THE STRATEGY

As long as player 1 does not play T, he needs to evaluate a parameter  $\varrho_k$  updated as follows ( $\varrho_1 = 0$ ):

$$\varrho_k = \varrho_{k-1} - (r^*(b_{k-1}) + \max[\epsilon(\varrho_{k-1}), -1/16])$$

where  $b_{k-1}$  is the action chosen by player 2 at stage  $k - 1$ .

The strategy  $\sigma$  of player 1 consists of selecting T with probability  $f(\varrho_k)$  at stage  $k$ . By playing in such a way, the parameter  $\varrho$ , on average, will not become “too negative,” i.e., for instance lower than  $-3\eta$  for any  $\eta > 0$  provided that sufficiently many stages are taken into account. More precisely, let  $K > 0$  be an integer such that

$$\left(\frac{1}{17}\right)^{[\eta K]} + \left(\frac{1}{17}\right)^{[\eta(K+1)]} + \dots \leq \eta. \quad (11)$$

where  $[x]$  is the largest integer  $p \leq x$ . We have the following lemma.

**Lemma 3** *If  $k > (K + 2)/\eta$  then for any strategy  $\tau$  of player 2*

$$\frac{1}{k} \mathbb{E}_{\sigma, \tau}[\varrho_{\min(\theta, k+1)}] \geq -3\eta. \quad (12)$$

**Proof.** Note that the set of histories with no absorption before stage  $k > K$  such that for all  $k' \in [k(1 - \eta), k]$

$$\varrho_{k'} + M < 4$$

is of probability smaller than

$$\left(\frac{1}{17}\right)^{[\eta k]}$$

because player 1 has to select T with a probability  $1/17$  for at least  $[\eta k]$  stages.

Otherwise, there exists  $k' \in [k(1 - \eta), k]$  such that:

$$\varrho_{k'} + M \geq 4$$

that implies

$$\varrho_k + M \geq 4 - \eta k.$$

The time  $\theta$  of absorption is either  $< K + 2$  or  $\geq K + 2$ . In the former case, since we average by  $k > (K + 2)/\eta$ , the contribution to (12) is smaller than  $\eta$ . In the latter case, the previous observations apply to  $k = \theta - 1 > K$ . Therefore, for  $k > K/\eta$  (12) holds. ■

### 6.3. SUBMARTINGALE INEQUALITY

Let us define the (discrete) process  $Y = (Y_k)_{k \geq 1}$  by:

$$Y_k = \begin{cases} \epsilon(\varrho_k) & \text{if } k \leq \theta \\ r^*(b_\theta) + \max[\epsilon(\varrho_\theta), -1/16] & \text{otherwise} \end{cases}$$

(remember that  $\theta$  is the stage at which absorption occurs).

**Proposition 2** *Assuming that player 1 uses the strategy  $\sigma$  defined in Section 6.2,  $Y$  is a submartingale for any strategy of player 2.*

**Proof.** Use (10) with  $\eta = r^*(b_{k-1}) + \epsilon(\varrho_{k-1})$  or  $\eta = r^*(b_{n-1}) - 1/16$ . ■

We are able to prove the next proposition ( $K$  is the same as in Lemma 3):

**Proposition 3** *For any  $k > K/\eta$  and for any strategy  $\tau$  of player 2*

$$\gamma_k(\sigma, \tau) \geq -3\eta/\zeta - 1/M. \quad (13)$$

**Proof.** Recall that by assumption  $\zeta \leq 1$  (and is independent of  $\eta$ ).

– If  $k \leq \theta$  then

$$\begin{aligned} \zeta(r(b_k) - \epsilon(\varrho_k)) &\geq -(r^*(b_k) + \epsilon(\varrho_k)) \\ &\geq -(r^*(b_k) + \max[\epsilon(\varrho_k), -1/16]). \end{aligned}$$

– If  $k > \theta$  then

$$\zeta[r^*(b_\theta) - (r^*(b_\theta) + \epsilon(\varrho_\theta))] \geq 0.$$

By summing the previous inequalities, it turns out that the average payoff  $\gamma_k$  satisfies:

$$\zeta[k\gamma_k - (Y_1 + \dots + Y_k)] \geq \varrho_{\min(\theta, k+1)}.$$

By Proposition 2, observe that  $\mathbb{E}[Y_k] \geq Y_1 = -1/M$ . The proof of (13) is terminated by application of Lemma 3. ■

### 6.4. CASE OF A GBM GAME WITH PERFECT MONITORING

Note that if we consider a GBM game (satisfying Assumption 1), then almost the same strategy is applicable. The differences, by increasing order of complexity, are as follows.

1. The probability of absorption on T is anything between 0 and 1 (depending on the action chosen by player 2).
2. There are non-absorbing columns ( $B_1 \neq \emptyset$ ).
3. There are completely absorbing columns ( $B_0 \neq \emptyset$ ).

Let us sketch briefly how we should proceed. First, suppose that only item 1 applies. Case (iii) of Lemma 1 would take the following form: there exists  $\zeta > 0$  such that for any  $\beta \in \Delta(B)$

$$p(T, \beta)r^*(T, \beta) \geq -\zeta r(B, \beta).$$

We are now able to justify why we can assume  $\zeta < 1$ . Because, if it is not the case, then replace T by T' defined as the following mixed action:

$$T' = \frac{1}{\zeta}T + (1 - \frac{1}{\zeta})B.$$

The game made of B and T' has exactly the same value as the original game.

We have to modify the functions  $f(\cdot)$  and  $\epsilon(\cdot)$  to accommodate the fact that item 1 holds. In Section 7, we show (in the case of a signalling structure) how it looks. The parameter is updated in a slightly different manner:

$$\varrho_k = \varrho_{k-1} - p(T, b_{k-1})(r^*(b_{k-1}) + \max[\epsilon(\varrho_{k-1}), -1/16])$$

where  $b_{k-1}$  is the action chosen by player 2 at stage  $k - 1$ .

Second, suppose that items 1 and 2 apply. Any non-absorbing action of player 2 is ignored by player 1, that is, he does not update his parameter. This is possible because, unlike the example of Section 4.1, player 1 knows when player 2 chooses one of these non-absorbing columns. If the corresponding non-absorbing payoff is strictly higher than the value of the BM game (it is always higher or equal), the non-absorbing column should not be played by player 2.

Third, suppose that items 1, 2 and 3 apply. Each time player 2 selects a completely absorbing column, there is a large probability of absorption. However, the absorbing payoff does not constitute a risk for player 1 in the sense that the submartingale inequality is preserved.

## 7. Extension of the Proof to GBM Games with a Signalling Structure

### 7.1. METHODOLOGY

We assume that there is no completely absorbing column. The  $\epsilon$ -optimal strategy described in Section 6.2 will be modified so that player 1 guarantees

$\underline{v}$  given by (2). Observe that if there are completely absorbing columns, we would have to prove that player 1 guarantees  $\underline{v}'$  given by (3). The argument we are about to develop should be adapted in the same spirit as for item 3 of Section 6.4.

Without loss of generality, we assume that  $\underline{v} = 0$ . We suppose that the cases corresponding to (i) and (ii) of Lemma 1 are excluded (player 1 would either always play T or B). Let us define

$$\nu = \min_{b \in B_2} [p(T, b)r^*(b)] < 0 \quad (14)$$

and let  $\bar{b} \in B_2$  be one of the actions that achieve the minimum.

The case corresponding to (iii) of Lemma 1 leads to:

**Lemma 4** *There exists  $\zeta > 0$  (without limiting our generality  $\zeta < 1$ ) such that for any pair of equivalent mixed actions  $\beta_1 \sim \beta_2$  we have:*

$$p(T, \beta_1)r^*(T, \beta_1) \geq -\zeta r(B, \beta_2). \quad (15)$$

## 7.2. DESCRIPTION OF THE STRATEGY

A random signal received by player 1 after any stage is not enough information to update the probability of choosing T at the next stage. This is due to the fact that the  $\sim$  relation between mixed actions induces a piecewise linearity of the worst absorbing payoff associated with a theoretical distribution of signals.

Therefore, the sequence of stages is decomposed into blocks  $\mathcal{B}_1, \mathcal{B}_2, \dots$  of length  $N > 0$  where  $N$  is a large integer. More precisely,  $\mathcal{B}_k$  is the set of stages  $(k-1)N+1, \dots, kN$ . The parameter  $\varrho$  is updated at the beginning of each block, thus obtaining a sequence  $\varrho_1, \varrho_2, \dots$  used to play on the corresponding block.

The minimal non-zero probability of absorption  $\min_{b \in \{n+1, m\}} p(T, b)$  on row T is denoted by  $\omega$ . The new forms of the functions  $f(\cdot)$  and  $\epsilon(\cdot)$  are as follows.

$$f(\varrho) = \begin{cases} \frac{1}{\omega(1+(M+\varrho)^2)} & \text{if } M + \varrho \geq 4N/\omega \\ \frac{\omega}{\omega^2 + 4N^2} & \text{otherwise} \end{cases}$$

and

$$\epsilon(\varrho) = \begin{cases} -\frac{N}{\omega(M+\varrho)} & \text{if } M + \varrho \geq 4N/\omega \\ -1/4 & \text{otherwise.} \end{cases}$$

The original functions of Section 6.1 can be recovered if one takes  $\omega = 1$  and  $N = 1$ .

At the end of block  $\mathcal{B}$ , player 1 updates the parameter using a stream of signals  $\tilde{l} = l_1, \dots, l_N \in L^N$  received on  $\mathcal{B}$ ; let us denote by  $\bar{l} \in \Delta(L)$  its

empirical distribution. A fundamental observation is that player 1 updates a parameter using  $\tilde{l}$  because he does not know the previous actions  $b_1, \dots, b_N$  of player 2. Let us denote by  $\bar{l} \in \Delta(L)$  the empirical distribution of  $\tilde{l}$ .

**Definition 6** *Given  $\eta > 0$ , a block  $\mathcal{B}$  is reliable if the empirical distribution  $\bar{l}$  of signals and the theoretical distribution of signals  $\psi(\beta)$  induced by  $\beta$  (empirical distribution of  $b_1, \dots, b_N$ ) on  $B$  are such that*

$$\|\psi(\beta) - \bar{l}\| \leq \eta^2.$$

The integer  $N > 0$  (length of the blocks) is chosen so that a block is reliable with probability larger than  $1 - \eta^2$  and thus uniformly with respect to the actions of player 2 (law of large numbers).

Given  $\beta \in \Delta(B)$  and a value  $\varrho$  of the parameter, let us define

$$\ell(\beta, \varrho, \eta) = p(T, \beta)[r^*(T, \beta) + \max[\epsilon(\varrho), -1/16] - 2\eta] \quad (16)$$

and

$$\underline{\ell}(\tilde{l}, \varrho, \eta) = \min\{\ell(\beta, \varrho, \eta) \mid \|\psi(\beta) - \bar{l}\| \leq \eta\}. \quad (17)$$

Starting from  $\varrho_1 = 0$ , let  $\sigma$  be the following strategy of player 1:

- Play i.i.d. on block  $\mathcal{B}_k$  with probability  $f(\varrho_k)$  on  $T$ .
- Provided that no absorption occurred and that the stream of signals received on the block  $\mathcal{B}_k$  is  $\tilde{l}_k$ , the new value  $\varrho_{k+1}$  of the parameter is:

$$\varrho_{k+1} = \varrho_k - \underline{\ell}(\tilde{l}_k, \varrho, \eta).$$

### 7.3. SUBMARTINGALE INEQUALITY

In order to show that this strategy  $\sigma$  indeed provides the desirable result let us extend (10) to the new functions  $f(\cdot)$  and  $\epsilon(\cdot)$ . Let us denote by  $\Delta_N(B)$  the set of  $\beta \in \Delta(B)$  that can be obtained as the empirical distribution of  $N$  actions  $b_1, \dots, b_N$ .

**Lemma 5** *For any  $\beta \in \Delta_N(B)$*

$$Nf(\varrho)\ell(\beta, \varrho, 2\eta) + (1 - Np(T, \beta))\epsilon(\varrho - \ell(\beta, \varrho, \eta)) \geq \epsilon(\varrho). \quad (18)$$

Observe that (18) is not an exact formula but an approximation to the first order of the risks of absorption that have been taken. It is easy to see that the error on the left-hand side is at most  $4N\omega^2 f^2(\varrho)$ . Remember that in (16) we have introduced a small correcting factor  $2\eta > 0$  that is designed, in part, to play the role of a “shock absorber.” Remember that when perfect monitoring holds, this small correction is useless. Up to taking  $N$  large, we can assume that:

$$4N\omega^2 f^2(\varrho) - \eta\omega Nf(\varrho) \leq 0,$$

which allows us to keep the inequality without adding error terms.

Recall that player 1 updates the parameter using a stream of signals  $\tilde{l}$ . He may have no knowledge of the actions  $b_1, \dots, b_N$  of player 2 and therefore of its empirical distribution  $\beta$ . Therefore, (18) is not sufficient except when  $\epsilon(\varrho - \ell(\beta, \varrho, 2\eta)) = -1/4$  since  $-1/4$  is the lowest value that  $\epsilon(\cdot)$  may take. The “theoretical”  $\epsilon(\varrho - \ell(\beta, \varrho, 2\eta))$  should be related to the “empirical”  $\epsilon(\varrho - \ell(\tilde{l}, \varrho, 2\eta))$ . If a block is reliable then this is straightforward by construction (17). But we have to take into account the possibility that a block is unreliable. One should describe the effect of dispersion around the theoretical distribution. The property (19) satisfied by the functions  $f(\cdot)$  and  $\epsilon(\cdot)$  allows us to overcome the problem.

**Lemma 6** *Let two values of the parameter  $\varrho'' \geq \varrho'$  be such that  $M + \varrho' \geq 4N/\omega$ . We have:*

$$\epsilon(\varrho') - \epsilon(\varrho'') \geq -2Nf(\varrho')(\varrho'' - \varrho'). \quad (19)$$

From now on, we separate in two cases:

1. If  $p(T, \beta) \leq \omega/2$  then let us prove that:

$$\mathbb{E}_{\tilde{l}}[\epsilon(\varrho - \ell(\tilde{l}, \varrho, 2\eta))] \geq \epsilon(\varrho - \ell(\beta, \varrho, 2\eta)).$$

To do so, observe that our condition on  $p(T, \beta)$  implies

$$p(T, \beta)r^*(T, \beta) \geq \nu/2.$$

Moreover, observe that since  $p(T, \beta) \leq p(T, \bar{b})$  holds:

$$\ell(\beta, \varrho, 2\eta) \geq \ell(\bar{b}, \varrho, 2\eta)$$

where  $\bar{b}$  is defined after (14).

Note that  $\beta' = (1 - \eta/2 + \eta^2)\beta + (\eta/2 - \eta^2)\bar{b}$  satisfies

$$\|\beta - \beta'\| \leq 3\eta/4$$

(provided that  $\eta > 0$  is chosen small enough) and

$$\ell(\beta', \varrho, 2\eta) \leq \ell(\beta, \varrho, 2\eta) + \frac{\eta}{8}\ell(\bar{b}, \varrho, 2\eta). \quad (20)$$

Therefore, with a probability larger than  $1 - \eta^2$ , making sure that

$$\begin{aligned} \|\bar{l} - \beta'\| &\leq \|\bar{l} - \beta\| + \|\beta - \beta'\| \\ &\leq \eta, \end{aligned}$$

by (20) we have

$$\underline{\ell}(\tilde{l}, \varrho, 2\eta) \leq \ell(\beta, \varrho, 2\eta) + \frac{\eta}{8}\ell(\bar{b}, \varrho, 2\eta). \quad (21)$$

By (19), comparing  $\epsilon(\varrho - \underline{\ell}(\tilde{l}, \varrho, 2\eta))$  and  $\epsilon(\varrho - \ell(\beta, \varrho, 2\eta))$ , one obtains an excess which is at least:

$$-2Nf(\varrho) \times \frac{\eta}{8}\ell(\bar{b}, \varrho, 2\eta).$$

On the other side, with a probability lower than  $\eta^2$ , by (19) there is a drop of at most  $2Nf(\varrho - 1)$  and therefore of at most  $3Nf(\varrho)$ .

2. Suppose that  $p(T, \beta) > \omega/2$ . Note that by similar arguments we have:

$$\mathbb{E}_{\tilde{l}}[\epsilon(\varrho - \underline{\ell}(\tilde{l}, \varrho, 2\eta))] \geq \epsilon(\varrho - \ell(\beta, \varrho, 2\eta)) - 3N\eta^2 f(\varrho).$$

Let us use the last part of the small error margin  $2\eta$  in (16) by observing that the drop of at most  $3N\eta^2 f(\varrho)$  is less than the excess  $Nf(\varrho)\omega\eta/2$ .

#### 7.4. END OF THE PROOF

Since a block is unreliable with a probability smaller than  $\eta^2$ , it can be shown that if  $K_1 > 0$  is large enough, the proportion of unreliable blocks among  $\mathcal{B}_1, \mathcal{B}_2, \dots, \mathcal{B}_k$ ,  $k > K_1$ , is smaller than  $2\eta^2$  (law of large numbers) with a probability (conditional to the fact that no absorption occurs before block  $\mathcal{B}_{k+1}$ ) larger than  $1 - \eta^2$ . Such a history (with no absorption before block  $\mathcal{B}_{k+1}$ ) is called  $k$ -reliable. From this observation, we derive two consequences.

The first one consists of the extension of (12) of Lemma 3. Let  $K > K_1$  be a large integer such that:

$$\left(\frac{\omega^2}{\omega^2 + 4N^2}\right)^{[K\eta\sqrt{\eta}]} + \left(\frac{\omega^2}{\omega^2 + 4N^2}\right)^{[(K+1)\eta\sqrt{\eta}]} + \dots \leq \eta.$$

Keep in mind that each exponent is defined up to a constant factor depending on  $\bar{b}$ , as we will see later on.

Since we want to prove a result involving the expectation, we reduce our focus to reliable histories. Let us consider  $k$ -reliable histories ( $k > K$ ) such that the largest integer  $k' \leq k$  satisfying:

$$\varrho_{k'} + M \geq 4N/\omega$$

satisfies also

$$k - k' < k\sqrt{\eta}.$$

This implies that  $\varrho_k + M \geq 4N/\omega - k\sqrt{\eta}$ . Otherwise, and this is the difficult case, we have:

$$k - k' \geq k\sqrt{\eta}. \quad (22)$$

The set of  $k$ -reliable histories such that (22) holds is of probability smaller than

$$\left(\frac{\omega^2}{\omega^2 + 4N^2}\right)^{[k\eta\sqrt{\eta}]}.$$

To prove it, note that there should exist at least  $[k\eta\sqrt{\eta}]$  blocks (up to a multiplicative constant) in  $[k' + 1, k]$  such that at least one action is an absorbing column (on T). Otherwise, there would be at least  $k(1 - \eta - 2\eta^2)\sqrt{\eta}$  reliable blocks such that there is no action corresponding to an absorbing column. Each one of these blocks induces an increase of parameter  $\varrho$  that is of the order of  $\eta$  (by (21) with  $\ell(\beta, \varrho, 2\eta) = 0$ ) and therefore from  $k'$  to  $k$  the parameter should increase. Clearly this would be a contradiction to the construction of  $k'$ .

Let us now turn to the equivalent of Proposition 3. This is done by considering reliable histories and the application of Lemma 4 in a way very similar to the proof of Proposition 3.

## 8. Conclusion

We hope that we have been able to convey the message that when dealing with a signalling structure, additional techniques must be employed. We believe that the functions  $\epsilon()$  and  $f()$  are the right tools for that. This is because Lemma 2 and Lemma 6 hold at the same time.

Nevertheless, the construction of a submartingale is still the core of our argument. There is nothing surprising about that since it appears from the outset in [3].

**ACKNOWLEDGEMENTS.** The writing of this chapter has been supported by a grant from the Australian Research Council no. A69703141

## References

1. Bewley, T. and Kohlberg, E. (1976) The asymptotic theory of stochastic games, *Mathematics of Operations Research* **1**, 197–208.
2. Bewley, T. and Kohlberg, E. (1978) On stochastic games with stationary optimal strategies, *Mathematics of Operations Research* **2**, 104–125.
3. Blackwell, D. and Ferguson, T.S. (1968) The big match, *Annals of Mathematical Statistics* **39**, 159–163.
4. Coulomb, J.M. (1992) Repeated games with absorbing states and no signals, *International Journal of Game Theory* **21**, 161–174.

5. Coulomb, J.M. (1996) A note on big match, *ESAIM: Probability and Statistics* **1**, 89–93.
6. Coulomb, J.M. (1997) On the value of discounted stochastic games, Working Paper 275, Institute of Mathematical Economics, University of Bielefeld.
7. Coulomb, J.M. (1999) Generalized big-match, *Mathematics of Operations Research* **24**, 795–816.
8. Coulomb, J.M. (2001) Repeated games with absorbing states and signaling structure, *Mathematics of Operations Research* **26**, 286–303.
9. Coulomb, J.M. (2002) Stochastic games without perfect monitoring, submitted to *International Journal of Game Theory*.
10. Gillette, D. (1957) Stochastic games with zero-stop probabilities, in M. Dresher, A.W. Tucker and P. Wolfe (eds.), *Contributions to the Theory of Games, Vol. III*, Princeton University Press, Princeton, NJ, pp. 179–187.
11. Kohlberg, E. (1974) Repeated games with absorbing states, *Annals of Statistics* **2**, 724–738.
12. Kohlberg, E. (1975) Optimal strategies in repeated games with incomplete information, *International Journal of Game Theory* **4**, 7–24.
13. Mertens, J-F. and Neyman, A. (1981) Stochastic games, *International Journal of Game Theory* **10**, 53–56.
14. Mertens, J-F., Sorin, S. and Zamir, S. (1994) Repeated games, CORE Discussion Papers 9420, 9421, 9422, Université Catholique de Louvain, Louvain-la-Neuve, Belgium.
15. Rosenberg, D., Solan, E. and Vieille, N. (2003) The maxmin value of stochastic games with imperfect monitoring, submitted to *International Journal of Game Theory*.
16. Shapley, L.S. (1953) Stochastic games, *Proceedings of the National Academy of Sciences of the U.S.A.* **39**, 1095–1100 (Chapter 1 in this volume).