



**HAL**  
open science

**Zero-sum repeated games: counterexamples to the  
existence of the asymptotic value and the conjecture  
 $\max\min=\lim v(n)$**

Bruno Ziliotto

► **To cite this version:**

Bruno Ziliotto. Zero-sum repeated games: counterexamples to the existence of the asymptotic value and the conjecture  $\max\min=\lim v(n)$ . *Annals of Probability*, 2016, 10.1214/14-AOP997. hal-00824039v2

**HAL Id: hal-00824039**

**<https://hal.science/hal-00824039v2>**

Submitted on 23 May 2013

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Zero-sum repeated games: counterexamples to the existence of the asymptotic value and the conjecture $\max\min=\lim v_n$

Bruno Ziliotto \*

May 23, 2013

## Abstract

We provide an example of a two-player zero-sum repeated game with public signals and perfect observation of the actions, where neither the value of the lambda-discounted game nor the value of the n-stage game converges, when respectively lambda goes to 0 and n goes to infinity. It is a counterexample to two long-standing conjectures, formulated by Mertens [6]: first, in any zero-sum repeated game, the asymptotic value exists, and secondly, when Player 1 is more informed than Player 2, Player 1 is able to guarantee the limit value of the n-stage game in the long run. The aforementioned example involves seven states, two actions and two signals for each player. Remarkably, players observe the payoffs, and play in turn (at each step the action of one player only has an effect on the payoff and the transition). Moreover, it can be adapted to fit in the class of standard stochastic games where the state is not observed.

## Introduction

In a general zero-sum repeated game, at each step, the two players simultaneously choose an action and receive a payoff which depends on the actions and the current state of nature (Player 1's payoff is the opposite of Player 2's payoff). Following that, a new state is drawn from a distribution depending on the actions and the state of the current step. Last, each player receives a private signal which provides him with information on the action of the other player and on the past and new state. The state space, action sets and signal sets are assumed to be finite.

There are several ways to evaluate the global payoff in a repeated game. In the *discounted game with parameter*  $\lambda$ , the payoff is the discounted mean  $\lambda \sum_{m \geq 1} \lambda(1-\lambda)^{m-1} g_m$ , where  $g_m$  is the payoff at stage  $m$ . In the *n-stage repeated game*, the payoff is the Cesaro mean  $\frac{1}{n} \sum_{m=1}^n g_m$ .

Properties of repeated games with long horizon (that is to say, when  $n$  is big or  $\lambda$  is small) have been widely studied in the literature. Following Zamir [25], we distinguish two main approaches.

The asymptotic approach is the study of the asymptotic properties of the value  $v_n$  of the  $n$ -stage repeated game and the value  $v_\lambda$  of the  $\lambda$ -discounted game, when respectively  $n$  goes to  $+\infty$  and  $\lambda$  goes to 0. The main question is whether these two quantities converge and have the same limit. When this is the case, the game is said to have an *asymptotic value*.

For standard stochastic games (that is to say, repeated games where the state and the actions are perfectly observed), Bewley and Kohlberg [2] have shown the existence of an asymptotic value (see also Olli Barton [10] for a recent alternative proof). For repeated games with incomplete information on both sides (actions are perfectly observed, the state does not evolve and players receive only one private signal about the initial state at the beginning), Mertens and Zamir [9] have proved that the asymptotic value does exist and have characterized it.

---

\*TSE (GREMAQ, Université Toulouse 1 Capitole), 21 allée de Brienne, 31000 Toulouse, France.  
E-mail: bruno.ziliotto@univ-tlse1.fr  
This research was supported by the ANR Jeudy (ANR-10-BLAN 0112) and the GDR 2932

The asymptotic value represents a target payoff in the long-horizon game. A natural question is whether players have strategies which guarantee this quantity in the long run, that is to say, strategies which are approximately optimal in any game  $\Gamma_n$  and  $\Gamma_\lambda$  with  $n$  big enough and  $\lambda$  small enough. When this is the case, the game is said to have a *uniform value*. Note that the existence of the uniform value yields the existence of the asymptotic value. Standard stochastic games (see Mertens and Neyman [7]) and repeated games with incomplete information on one side (see Aumann Maschler [1]) have a uniform value. In the case of imperfect observation of the actions, or incomplete information on both sides, a repeated game may fail to have a uniform value (see respectively Coulomb [4] and Rosenberg, Solan and Vieille [16], and Aumann and Maschler [1]).

The results stated above, and the study of the Big Match with lack of information on one side by Sorin (see [20] and [21]) have led to the following two conjectures (see Mertens [6] and Mertens, Sorin and Zamir [8]):

**1** In a repeated game,  $(v_n)$  and  $(v_\lambda)$  converge and have the same limit.

**2** In a repeated game where Player 1 is more informed than Player 2,  $(v_n)$  converges to the maxmin of the game.

By definition, Player 1 is more informed than Player 2 if he observes what Player 2 observes (for more details, see Coulomb [3]). The maxmin is the greatest number Player 1 can guarantee in the long-horizon game.

More recent results have confirmed these conjectures in the following classes of repeated games: absorbing games with incomplete information on one side (Rosenberg [15]), recursive games with incomplete information on one side (Rosenberg and Vieille [17]), repeated games with an informed controller (Renault [13]), Markov chain games with lack of information on one side (Renault [11]), and repeated games with a more informed controller (Gensbittel, Oliu Barton and Venel [5]). When there is only one player, the two conjectures boil down to the existence of a uniform value. In this framework, they hold true, even in a more general setting such as compact action sets, continuous transition function and payoff, and more general evaluations of the payoff than Cesaro means (see respectively Renault [12] and Renault and Venel [14]).

One of the simplest category of repeated games where **1** and **2** were still open are state-blind repeated games: actions are public but players observe nothing about the state. In this model, Venel [23] has shown the existence of a uniform value, under the additional hypothesis that transitions are commutative. Note that if **2** holds true, then state-blind repeated games should have a uniform value. Indeed, in this case Player 1 is more informed than Player 2, and symmetrically Player 2 is more informed than Player 1. Therefore the maxmin should be equal to the minmax. More generally, if **2** holds true, then repeated games with public signals and perfect observation of the actions (repeated games where signals are the same for both players and include past actions) should have a uniform value.

We provide an example of a repeated game with public signals and perfect observation of the actions, where neither  $(v_n)$  nor  $(v_\lambda)$  converge, which contradicts **1**. In particular, this game has no uniform value, which also contradicts **2**. One remarkable feature of this game is that at each step, only the action of one player influences the payoff and the transition. Moreover, players observe their payoff.

Note that if  $f$  is a continuous function  $f : [0, 1] \rightarrow [0, 1]$ , the Cesaro mean of its iterates  $\frac{1}{n} \sum_{m=1}^n f^m$  may fail to converge. Hence, as far as the asymptotic value is concerned, either the state space has to be finite, or an additional assumption on the transition function has to be made, like nonexpansiveness (see Renault and Venel [14]).

In the case of compact stochastic games (finite state space, compact action sets, continuous transition and payoff functions and perfect observation of state and actions), Vigerat [24] has recently provided an example with no asymptotic value.

The paper is organized as follows. In section 1, we recall the model of repeated game and some basic concepts. In section 2, we present our main counterexample and show that  $(v_\lambda)$  does not converge. In

section 3, we construct a similar game where  $(v_n)$  does not converge. In section 4, we show how our counterexample adapts to other classes of games. We notably exhibit an alternative example to Vigerál [24] of a compact stochastic game with no asymptotic value.

## 1 Generalities

$\mathbb{N}$  denotes the set of non-negative integers, and  $\mathbb{N}^* := \mathbb{N} \setminus \{0\}$ .

If  $C$  is any countable set, we denote by  $\Delta(C)$  the set of probability measures on  $C$ :  $\Delta(C) := \{p \in \mathbb{R}_+^C \mid \sum_{c \in C} p_c = 1\}$ . To define some  $p \in \Delta(C)$ , we will often write  $p := \sum_{c \in C} p_c \cdot c$ .

### 1.1 General model of repeated game

A repeated game  $\Gamma$  is defined by:

- A state space  $K$
- An action set  $I$  (resp.  $J$ ) for Player 1 (resp. 2)
- A signal set  $A$  (resp.  $B$ ) for Player 1 (resp. 2)
- An initial probability  $p$  on  $K \times A \times B$
- A transition function  $q : K \times I \times J \rightarrow \Delta(K \times A \times B)$
- A bounded payoff function  $g : K \times I \times J \rightarrow \mathbb{R}$ .

We assume  $K, I, J, A$  and  $B$  to be finite.

The game proceeds as below:

- Before game starts, a triplet  $(k_1, a_0, b_0)$  is drawn according to  $p$ .  $k_1$  is the initial state, and Player 1 (resp. 2) gets the signal  $a_0$  (resp.  $b_0$ ).
- At step  $m \geq 1$ , both players choose an action simultaneously and independently,  $i_m \in I$  (resp.  $j_m \in J$ ) for Player 1 (resp. 2). The payoff at stage  $m$  is  $g(k_m, i_m, j_m)$ . A triplet  $(k_{m+1}, a_m, b_m)$  is drawn from  $q(k_m, i_m, j_m)$ . The signal  $a_m$  (resp.  $b_m$ ) is announced to Player 1 (resp. 2). The game switches to state  $k_{m+1}$ , and goes to the next step.

We call *history of the game before step  $m$*  the random sequence

$$\mathcal{H}_m := (a_0, b_0, k_1, i_1, j_1, a_1, b_1, k_2, \dots, i_{m-1}, j_{m-1}, a_{m-1}, b_{m-1}, k_m).$$

The set of all possible histories before step  $m$  is

$$H_m = A \times B \times K \times (I \times J \times A \times B \times K)^{m-1}.$$

The set of all possible histories is  $H_\infty = A \times B \times K \times (I \times J \times A \times B \times K)^{\mathbb{N}^*}$ .

The private history of Player 1 (resp. 2) before step  $m$  is the sequence  $(a_0, i_1, a_1, \dots, i_{m-1}, a_{m-1})$  (resp.  $(b_0, j_1, b_1, \dots, j_{m-1}, b_{m-1})$ ), and the set of all private histories for Player 1 (resp. 2) is  $H_m^1 = A \times (I \times A)^{m-1}$  (resp.  $H_m^2 = B \times (J \times B)^{m-1}$ ).

We denote  $H^1 = \bigcup_{m \geq 1} H_m^1$  and  $H^2 = \bigcup_{m \geq 1} H_m^2$ .

A *pure strategy* for Player 1 (resp. 2) is a map  $s : H^1 \rightarrow I$  (resp.  $t : H^2 \rightarrow J$ ).

A *behavioral strategy* for Player 1 (resp. 2) is a map  $\sigma : H^1 \rightarrow \Delta(I)$  (resp.  $\tau : H^2 \rightarrow \Delta(J)$ ). The set of all behavioral strategies for Player 1 (resp. 2) is denoted  $\Sigma$  (resp.  $\mathcal{T}$ ).

An initial probability  $p \in \Delta(K \times A \times B)$  and a couple of (pure or behavioral) strategies  $(\sigma, \tau) \in \Sigma \times \mathcal{T}$  naturally induce a unique probability measure  $\mathbb{P}_{\sigma, \tau}^p$  on the set of all possible histories of the game  $H_\infty$ .

We denote  $g_m$  the random payoff at stage  $m \geq 1$ :  $g_m := g(k_m, i_m, j_m)$ . For  $\lambda \in (0, 1]$ , we call  *$\lambda$ -discounted game* the game  $\Gamma_\lambda^p$  with normal form  $(\Sigma, \mathcal{T}, \gamma_\lambda^p)$ , where the payoff  $\gamma_\lambda^p : \Sigma \times \mathcal{T} \rightarrow \mathbb{R}$  is defined by

$$\gamma_\lambda^p(\sigma, \tau) = \mathbb{E}_{\sigma, \tau}^p \left( \sum_{m \geq 1} \lambda(1 - \lambda)^{m-1} g_m \right)$$

For  $n \in \mathbb{N}^*$ , we call *n-stage repeated game* the game  $\Gamma_n^p$  with normal form  $(\Sigma, \mathcal{T}, \gamma_n^p)$ , where the payoff  $\gamma_n^p : \Sigma \times \mathcal{T} \rightarrow \mathbb{R}$  is defined by

$$\gamma_n^p(\sigma, \tau) := \mathbb{E}_{\sigma, \tau}^p \left( \frac{1}{n} \sum_{m=1}^n g_m \right)$$

It is easy to show, using Sion's theorem (see [19]), that  $\Gamma_\lambda^p$  (resp.  $\Gamma_n^p$ ) has a value  $v_\lambda : \Delta(K \times A \times B) \rightarrow \mathbb{R}$  (resp.  $v_n : \Delta(K \times A \times B) \rightarrow \mathbb{R}$ ):

$$\begin{aligned} v_\lambda(p) &= \max_{\sigma \in \Sigma} \min_{\tau \in \mathcal{T}} \gamma_\lambda^p(\sigma, \tau) = \min_{\tau \in \mathcal{T}} \max_{\sigma \in \Sigma} \gamma_\lambda^p(\sigma, \tau) \\ v_n(p) &= \max_{\sigma \in \Sigma} \min_{\tau \in \mathcal{T}} \gamma_n^p(\sigma, \tau) = \min_{\tau \in \mathcal{T}} \max_{\sigma \in \Sigma} \gamma_n^p(\sigma, \tau) \end{aligned}$$

### 1.1.1 Asymptotic approach

**Definition**  $\Gamma$  has an *asymptotic value* if  $(v_n)$  and  $(v_\lambda)$  converge pointwise to the same limit (when  $n \rightarrow +\infty$  and  $\lambda \rightarrow 0$ ).

**Remark** It is easy to see that for all  $(n, \lambda) \in \mathbb{N}^* \times (0, 1]$ ,  $v_n$  and  $v_\lambda$  are  $\|g\|_\infty$  Lipschitz. Thus as far as these sequences are concerned, pointwise and uniform convergence are equivalent.

### 1.1.2 Uniform approach

Let  $p \in \Delta(K \times A \times B)$ .

**Definition** Player 1 (resp. 2) can *guarantee*  $\alpha \in \mathbb{R}$  in  $\Gamma_\infty(p)$  if for all  $\epsilon > 0$ , there exists  $\sigma^* \in \Sigma$  (resp.  $\tau^* \in \mathcal{T}$ ) and  $n_0 \in \mathbb{N}^*$  such that for all  $\tau \in \mathcal{T}$  (resp.  $\sigma \in \Sigma$ ) and  $n \geq n_0$

$$\gamma_n^p(\sigma^*, \tau) \geq \alpha - \epsilon \quad (\text{resp. } \gamma_n^p(\sigma, \tau^*) \leq \alpha + \epsilon)$$

The *maxmin* of  $\Gamma_\infty(p)$  is  $\sup \{ \alpha \mid \text{Player 1 can guarantee } \alpha \text{ in } \Gamma_\infty(p) \}$ .

The *minmax* of  $\Gamma_\infty(p)$  is  $\inf \{ \alpha \mid \text{Player 2 can guarantee } \alpha \text{ in } \Gamma_\infty(p) \}$ .

When the minmax is equal to the maxmin, this quantity is called the *uniform value*, and is denoted by  $v_\infty(p)$ .

**Definition** Player 1 (resp. 2) can *defend uniformly*  $\alpha \in \mathbb{R}$  in  $\Gamma_\infty(p)$  if for all  $\epsilon > 0$ , for all  $\tau \in \mathcal{T}$  (resp.  $\sigma \in \Sigma$ ), there exists  $\sigma \in \Sigma$  (resp.  $\tau \in \mathcal{T}$ ) and  $n_0 \in \mathbb{N}^*$  such that for all  $n \geq n_0$

$$\gamma_n^p(\sigma, \tau) \geq \alpha - \epsilon \quad (\text{resp. } \gamma_n^p(\sigma, \tau) \leq \alpha + \epsilon)$$

When Player 1 can defend uniformly the minmax, the minmax is called *uniform minmax*. When Player 2 can defend uniformly the maxmin, the maxmin is called *uniform maxmin*.

### Remarks

- The existence of the uniform value yields the existence of the asymptotic value.
- When the uniform value exists, the maxmin and minmax are uniform.

## 1.2 Repeated games with public signals and perfect observation of the actions

A repeated game  $\Gamma$  with public signals and perfect monitoring of the actions is a repeated game where the signals received by both players are public and include the actions. Explicitly, with the notations of subsection 1.1,  $A = B$ , for all  $m \in \mathbb{N}$   $a_m = b_m$ , and  $(i_m, j_m)$  is measurable with respect to  $a_m$ . We fix such a game  $\Gamma$ .

For  $m \geq 1$ , we denote by  $p_m$  the conditional probability on the state  $k_m$  at stage  $m$ , given the past history  $\mathcal{H}_m$ : for all  $k \in K$ ,  $p_m(k) = \mathbb{P}(k_m = k | \mathcal{H}_m)$ .  $p_m$  represents the common belief at stage  $m$  about the current state  $k_m$ . It turns out that the triplet  $(p_m, i_m, j_m)$  is the only relevant information conveyed by the signal  $a_m$ . Formally, let  $\tilde{\Gamma}$  be the auxiliary repeated game with state space  $\Delta(K)$ , action sets  $I$  and  $J$ , transition function  $\tilde{q} : \Delta(K) \times I \times J \rightarrow \Delta_f(\Delta(K))$  defined by  $\tilde{q}(p, i, j) := \mathbb{P}(p_2 = p | p_1 = p, i_1 = i, j_1 = j)$ , and payoff function  $\tilde{g} : \Delta(K) \times I \times J$  defined by  $\tilde{g}(p, i, j) := \sum_{k \in K} g(k, i, j)$ .

As for the signaling structure, both players observe perfectly the state and the actions.

For all  $\lambda \in (0, 1]$  (resp.  $n \in \mathbb{N}^*$ ),  $\tilde{\Gamma}_\lambda$  (resp.  $\tilde{\Gamma}_n$ ) and  $\Gamma_\lambda$  (resp.  $\Gamma_n$ ) have the same value, and optimal strategies in the first game induce optimal strategies in the second one, and reciprocally. Thus in what follows we identify  $\Gamma$  with  $\tilde{\Gamma}$ , and set  $q := \tilde{q}$ , and  $g := \tilde{g}$ .

By an easy generalization of Shapley [18],  $v_\lambda$  is the unique solution of the following functional equation

$$f(p) = \max_{x \in \Delta(I)} \min_{y \in \Delta(J)} \{ \lambda g(p, x, y) + (1 - \lambda) \mathbb{E}_{x,y}^p(f) \} \quad (1)$$

$$= \min_{y \in \Delta(J)} \max_{x \in \Delta(I)} \{ \lambda g(p, x, y) + (1 - \lambda) \mathbb{E}_{x,y}^p(f) \} \quad (2)$$

where the unknown is a continuous function  $f : \Delta(K) \rightarrow \mathbb{R}$ ,

$$\mathbb{E}_{x,y}^p(f) := \sum_{(p', i, j) \in \Delta(K) \times I \times J} x(i)y(j)q(p, i, j)(p')f(p') \text{ and } g(p, x, y) := \sum_{(i, j) \in I \times J} x(i)y(j)g(p, i, j).$$

**Remark** Note that in this framework the belief hierarchy is trivial: Players have the same information. In the general model of repeated game, it may not be the case, and a recursive equation similar to (1) is more difficult to write (see Coulomb [3]).

One can deduce for any  $\lambda \in (0, 1]$  the existence of *stationnary* optimal strategies in  $\Gamma_\lambda$ : a stationary strategy is a strategy which depends only on the state  $p_m$ . We have the following refinement:

**Definition** A player is said to *control*  $p \in \Delta(K)$  if in this state the transition  $q(p, \cdot)$  and the payoff  $g(p, \cdot)$  do not depend on the action of the other player.

**Lemma 1.1** *Assume that each state  $p \in \Delta(K)$  is controlled by one player. Let  $\Gamma'$  be the restriction of  $\Gamma$  to pure stationary strategies. Let  $\lambda \in (0, 1]$ . Then  $\Gamma'_\lambda$  has a value  $v'_\lambda$ , and  $v'_\lambda = v_\lambda$ . Moreover, optimal strategies in  $\Gamma'_\lambda$  are also optimal in  $\Gamma_\lambda$ .*

**Proof** Assume Player 1 (resp. 2) controls some  $p \in \Delta(K)$ . Let  $\lambda \in (0, 1]$ . Then (1) boils down to

$$v_\lambda(p) = \max_{x \in \Delta(I)} \{ \lambda g(p, x) + (1 - \lambda) \mathbb{E}_x^p(v_\lambda) \}$$

and respectively (2) boils down to

$$v_\lambda(p) = \min_{y \in \Delta(J)} \{ \lambda g(p, y) + (1 - \lambda) \mathbb{E}_y^p(v_\lambda) \}$$

The right term is linear in  $x$  (resp.  $y$ ), hence the maximum (resp. minimum) is reached for some  $i \in I$  (resp.  $j \in J$ ). It yields the existence of a pure optimal stationary strategy for Player 1 (resp. 2) in  $\Gamma_\lambda$ . In particular,  $\Gamma'_\lambda$  has a value and  $v'_\lambda = v_\lambda$ . If  $x, y$  are optimal strategies in  $\Gamma'_\lambda$ , they are also optimal in (1) and (2), and therefore in  $\Gamma_\lambda$ .

## 2 A repeated game with public signals and perfect observation of the actions where $(v_\lambda)$ does not converge

In subsection 2.1 we present the counterexample. Then we describe the equivalent game with perfect observation of state and actions (see subsection 1.2). Though this game seems a bit intricate at first

sight, it turns out that for each discount factor  $\lambda$  in  $(0, 1]$ , the discounted game is equivalent to a one-shot game played on  $\mathbb{N} \times 2\mathbb{N}$ . This last game is easy to analyze, and we show in subsection 2.4 that its discounted value does not converge.

## 2.1 Description of the game

Let us consider the following repeated game with public signals and perfect observation of the actions  $\Gamma$ , with state space  $K = \{1^*, 1^{++}, 1^T, 1^+, 0^*, 0^{++}, 0^+\}$ , action sets  $I = J = \{C, Q\}$  and signal sets  $A = B = \{D, D'\}$ . The initial state will usually be taken as  $1^{++}$ .

Payoffs are independent of actions, and are 1 in states  $1^*, 1^{++}, 1^T$  and  $1^+$ , and 0 in states  $0^*, 0^{++}$  and  $0^+$ .

$1^{++}, 1^T, 1^+$  are controlled by Player 2, in the sense that the transition on these states  $q(1^{++}, \cdot)$ ,  $q(1^T, \cdot)$  and  $q(1^+, \cdot)$  do not depend on the actions of Player 1. Similarly, Player 1 controls the states  $0^{++}$  and  $0^+$ . Hence  $q$  can be seen as a map from  $K \times \{C, Q\}$  to  $\Delta(K \times \{D, D'\})$ . Last,  $1^*$  and  $0^*$  are *absorbing* states: it means that once  $1^*$  or  $0^*$  is reached, the game remains forever in this state, and the payoff does not depend on the actions (*absorbing* payoff).

The following table describes the transitions of the game in the states controlled by Player 2, that is to say  $1^{++}, 1^T$  and  $1^+$ :

$$\begin{aligned} q(1^{++}, C) &:= \frac{1}{2} \cdot (1^T, D) + \frac{1}{2} \cdot (1^{++}, D') \\ q(1^T, C) &:= \frac{1}{8} \cdot (1^{++}, D) + \frac{3}{8} \cdot (1^+, D) + \frac{1}{2} \cdot (1^{++}, D') \\ q(1^+, C) &:= \frac{1}{2} \cdot (1^+, D) + \frac{1}{2} \cdot (1^{++}, D') \\ q(1^{++}, Q) &:= q(1^T, Q) := (1^*, D') \\ q(1^+, Q) &:= (0^{++}, D) \end{aligned}$$

We now describe the transitions in the states controlled by Player 1, that is to say  $0^{++}$  and  $0^+$ :

$$\begin{aligned} q(0^{++}, C) &:= \frac{1}{4} \cdot (0^{++}, D) + \frac{1}{4} \cdot (0^+, D) + \frac{1}{2} \cdot (0^{++}, D') \\ q(0^+, C) &:= \frac{1}{2} \cdot (0^+, D) + \frac{1}{2} \cdot (0^{++}, D') \\ q(0^{++}, Q) &:= (0^*, D') \\ q(0^+, Q) &:= (1^{++}, D) \end{aligned}$$

## 2.2 Equivalent repeated game with perfect observation

Recall that from subsection 1.2,  $\Gamma$  is equivalent in terms of value to a repeated game with perfect observation of the state and actions  $\tilde{\Gamma}$ , with state variable  $p_m = k_m | \mathcal{H}_m$ . In this subsection, we give the exact expression of the transition  $\tilde{q}$  and payoff  $\tilde{g}$  of such a game.

We define the following elements of  $\Delta(K)$ :

$$\begin{aligned} 1_n &:= 2^{-n} \cdot 1^{++} + (1 - 2^{-n}) \cdot 1^+ \\ 1_n^T &:= 2^{-n} \cdot 1^T + (1 - 2^{-n}) \cdot 1^+ \\ 0_n &:= 2^{-n} \cdot 0^{++} + (1 - 2^{-n}) \cdot 0^+ \end{aligned}$$

Let  $P_1 := \bigcup_{n \in \mathbb{N}} 1_{2n}$ ,  $P_1^T := \bigcup_{n \in \mathbb{N}} 1_{2n}^T$ ,  $P_2 := \bigcup_{n \in \mathbb{N}} 0_n$ , and  $P = P_1 \cup P_1^T \cup P_2 \cup \{1^*, 0^*\}$ . Note that in  $\tilde{\Gamma}$ , Player 1 controls all the states in  $P_2$ , and Player 2 controls all the states in  $P_1$  and  $P_1^T$ . In  $P$ , the transition  $\tilde{q}$

depends only on the action of one player, and is the following:

$$\begin{aligned}
\tilde{q}(1_{2n}, C) &:= \frac{1}{2} \cdot 1_{2n}^T + \frac{1}{2} \cdot 1^{++} \\
\tilde{q}(1_{2n}^T, C) &:= \frac{1}{2} \cdot 1_{2n+2} + \frac{1}{2} \cdot 1^{++} \\
\tilde{q}(1_{2n}, Q) &:= \tilde{q}(1_{2n}^T, Q) := (1 - 2^{-2n}) \cdot 0^{++} + 2^{-2n} \cdot 1^* \\
\tilde{q}(0_n, C) &:= \frac{1}{2} \cdot 0_{n+1} + \frac{1}{2} \cdot 0^{++} \\
\tilde{q}(0_n, Q) &:= (1 - 2^{-n}) \cdot 1^{++} + 2^{-n} \cdot 0^*
\end{aligned} \tag{3}$$

Conditionnally to the game leaving from any  $p \in P$ , the common belief about the state  $p_m$  remains in  $P$ . Hence the preceding equations describe completely the transitions of the game leaving from state  $p$ . The following proposition summarizes what has been shown above:

**Proposition 2.1** *Let  $p \in P$ .  $\Gamma^p$  is equivalent to a repeated game with perfect observation of state and actions  $\tilde{\Gamma}^p$ , with state space  $P$ , action sets  $I = J = \{C, Q\}$ , transition  $\tilde{q}$  described by (3), and a payoff function  $\tilde{g}$  which depends only on the state, such that for all  $p$  in  $P_1 \cup P_1^T$  (resp.  $p$  in  $P_2$ ),  $\tilde{g}(p) = 1$  (resp.  $\tilde{g}(p) = 0$ ).*

From now on we identify  $\Gamma$  with  $\tilde{\Gamma}$ , and set  $g := \tilde{g}$  and  $q := \tilde{q}$ .

Let us explain informally the dynamics of the game. Assume that the game starts in  $p_1 = 0^{++}$ . Player 1 wants to go to state  $1^{++}$ . If he plays  $Q$  immediately, the game is absorbed in  $0^*$ , which is very bad for him. If he never plays  $Q$ , the payoff is 0 forever, which is also bad. If he plays  $C$  until the state is  $0_n$ , and then  $Q$ , with probability  $2^{-n}$  the state is absorbed in  $0^*$  (we will often call  $2^{-n}$  the "absorbing risk"), and with probability  $(1 - 2^{-n})$  the games goes to  $1^{++}$ .

To reach  $0_n$  from  $0^{++}$ , Player 1 needs on average  $2^n$  steps. Hence there is a trade-off between staying not too long in states of type 0, and reducing the risk of absorbing in  $0^*$ . Basically Player 1 needs to wait  $2^n$  steps to take a risk of absorbing in  $0^*$  equal to  $2^{-n}$ .

For Player 2, it is the same principle. Assume that the game starts in  $p_1 = 1^{++}$ . Player 2 plays  $C$  until reaching  $1_{2n}$  or  $1_{2n}^T$ , and then  $Q$ . With probability  $2^{-2n}$  the game is absorbed in  $1^*$ , and with probability  $(1 - 2^{-2n})$  the game goes to state  $0^{++}$ .

To reach  $1_{2n}$ , Player 2 needs on average  $2^{2n}$  steps. Player 2 can also play  $Q$  in  $1_{2n}^T$ , but it is not a good strategy, since such a state is harder to reach than  $1_{2n}$  ( $2^{2n+1}$  steps on average) but leads to the same absorbing risk  $2^{-2n}$ . Note that the time needed by Player 2 to go from  $1^{++}$  to  $1_{2n}$  is on average the same as the time needed by Player 1 to go from  $0^{++}$  to  $0_{2n}$ .

Therefore, the only asymmetry of the game is that Player 1 can take any absorbing risks of the form  $2^{-n}$ , whereas Player 2 can only take absorbing risks of the form  $2^{-2n}$ .

### 2.3 Equivalent one-shot game on $\mathbb{N} \times 2\mathbb{N}$

We fix some  $\lambda \in (0, 1)$ .

The aim of this section is to prove the following proposition:

**Proposition 2.2**  $\Gamma_\lambda^{1^{++}}$  has the same value as the one-shot game  $G_\lambda$  with action set  $\mathbb{N}$  for Player 1,  $2\mathbb{N}$  for Player 2, and payoff

$$g_\lambda(a, b) := \frac{1 - f_\lambda(b)}{1 - f_\lambda(a)f_\lambda(b)}$$

where

$$f_\lambda(n) := \frac{(1 - 2^{-n})(1 - \lambda^2)}{1 + 2^{n+1}\lambda(1 - \lambda)^{-n} - \lambda}$$

Moreover, optimal strategies in  $G_\lambda$  induce optimal strategies in  $\Gamma_\lambda^{1^{++}}$ .



Let  $a, b \in \mathbb{N}^2$ . We consider strategies  $\sigma(a)$  and  $\tau(b)$  for Player 1 and 2 of the following form:  
if  $p_m = 0_n$  for some  $n \geq a$ , play  $Q$ , otherwise play  $C$ .

if  $p_m = 1_n$  for some  $n \geq b$ , play  $Q$ , otherwise play  $C$ .

When there is no ambiguity, we will simply denote  $a$  (resp.  $b$ ) for  $\sigma(a)$  (resp.  $\tau(b)$ ).

**Remark** If Player 1 (resp. 2) plays  $\sigma(a)$  (resp.  $\tau(b)$ ), the state  $0_n$  (resp.  $1_{2n}, 1_{2n}^T$ ) with  $n > a$  (resp.  $2n > b$ ) is never reached in  $\Gamma^{1^{++}}$ .

**Proposition 2.3** *Let  $\Gamma'$  be the restriction of  $\Gamma$  to pure stationary strategies. In  $\Gamma'_\lambda^{1^{++}}$ , any strategy for Player 1 (resp. 2) is dominated by some  $\sigma(a)$  (resp.  $\tau(b)$ ).*

**Proof** First note that a pure stationary strategy for one player in  $\Gamma$  can be seen as a map from  $\mathbb{N}$  to  $\{C, Q\}$ . Indeed Player 1 only influences the game when the state is in  $P_2$ , and Player 2 only influences the game when the state is in  $P_1$ .

Let  $(\sigma, \tau) \in \Sigma \times \mathcal{T}$  be pure stationary strategies.  $\sigma(n)$  represents the action of Player 1 when the state is  $0_n$ .  $\tau(2n)$  (resp.  $\tau(2n+1)$ ) represents the action of Player 2 when the state is  $1_{2n}$  (resp.  $1_{2n}^T$ ). Let  $a = \inf \{n \geq 1 | \sigma(n) = Q\}$ . If  $a = +\infty$ , then  $\sigma$  is dominated by  $\sigma(0)$ . If  $a < +\infty$ , by the preceding remark,  $\sigma$  and  $\sigma(a)$  are equivalent in  $\Gamma^{1^{++}}$ : for all  $\tau \in \mathcal{T}$ ,  $\mathbb{P}_{\sigma, \tau}^{1^{++}} = \mathbb{P}_{\sigma(a), \tau}^{1^{++}}$ .

Let  $b = \inf \{n \geq 1 | \tau(n) = Q\}$ . If  $b = +\infty$ , then  $\tau$  is dominated by  $\tau(0)$ . Assume  $b$  is even. Still by the preceding remark,  $\tau$  and  $\tau(b)$  are the same strategies in  $\Gamma^{1^{++}}$ : for all  $\sigma \in \Sigma$ ,  $\mathbb{P}_{\sigma, \tau}^{1^{++}} = \mathbb{P}_{\sigma, \tau(b)}^{1^{++}}$ . If  $b$  is odd, then  $\tau(b-1)$  dominates  $\tau(b)$ . Indeed, as mentioned above, it is always better for Player 2 to quit in  $1_{2n}$  rather than in  $1_{2n}^T$ .

**Proposition 2.4** *Let  $\Gamma''$  be the restriction of  $\Gamma$  to strategies of the form  $\sigma(a)$  and  $\sigma(b)$ ,  $(a, b) \in \mathbb{N} \times 2\mathbb{N}$ . Then  $\Gamma_\lambda^{1^{++}}$  and  $\Gamma''_\lambda^{1^{++}}$  have the same value. Moreover, any optimal strategy in  $\Gamma''_\lambda^{1^{++}}$  is also optimal in  $\Gamma_\lambda^{1^{++}}$ .*

**Proof** It stems directly from Proposition 2.3 and Lemma 1.1.

To complete the proof of Proposition 2.2, it remains to compute  $\gamma_\lambda^{1^{++}}(a, b)$  (the payoff in the game  $\Gamma_\lambda^{1^{++}}$ , when Player 1 plays  $\sigma(a)$  and Player 2 plays  $\tau(b)$ ), for some  $(a, b) \in \mathbb{N} \times 2\mathbb{N}$ , and show that  $\gamma_\lambda^{1^{++}}(a, b) = g_\lambda(a, b)$ .

Our aim is first to compute the average time spent by Player 2 (resp. 1) in states of type 1 (resp. 0) before he plays  $Q$ , when he plays the strategy  $\sigma(b)$  (resp.  $\sigma(a)$ ), going from  $1^{++}$  (resp.  $0^{++}$ ).

Let  $T_a^1 = \inf \{m \geq 1 | i_m = Q\}$  and  $T_b^2 = \inf \{m \geq T_a^1 + 1 | j_m = Q\} - T_a^1$ .

Let  $(X_n)_{n \geq 0}$  be the Markov chain on  $\mathbb{N}$  with transition  $\pi : \mathbb{N} \rightarrow \Delta(\mathbb{N})$  defined by  $\pi(n) = \frac{1}{2} \cdot (n+1) + \frac{1}{2} \cdot 0$ , and take  $X_0 := 0$ . Let  $T_n = \inf \{n' \geq 0 | X_{n'} = n\}$ .  $T_n$  represents the random time needed by the Markov chain to go from 0 to  $n$ . By definition of the transition of  $\Gamma$  (see (3)) and the Markov property, under  $\mathbb{P}_{\sigma(a), \sigma(b)}^{1^{++}}$ ,  $T_a^1$  (resp.  $T_b^2$ ) and  $T_a + 1$  (resp.  $T_b + 1$ ) have the same law. We have the following lemma:

**Lemma 2.5** *For all  $n \in \mathbb{N}$*

$$\mathbb{E}((1-\lambda)^{T_n}) = \frac{1+\lambda}{1+2^{n+1}\lambda(1-\lambda)^{-n}-\lambda}$$

**Proof** For  $n = 0$  the result is clear. Let  $n \geq 1$ .

Going from 0, to get to state  $n$  the Markov chain  $(X_n)$  has to reach  $n-1$ . Then it goes to  $n$  (resp. 0), if the realization of a certain Bernoulli random variable  $B$  is 1 (resp. 0). Therefore, by the Markov property,  $T_n$  satisfies

$$T_n = T_{n-1} + 1_{B=1} + 1_{B=0}(1 + T'_n)$$

where  $T'_n$  is an independent copy of  $T_n$ . Note that  $T_{n-1}$ ,  $B$  and  $T'_n$  are independent. Therefore we have

$$\mathbb{E}((1-\lambda)^{T_n}) = (1-\lambda)\mathbb{E}((1-\lambda)^{T_{n-1}})\mathbb{E}((1-\lambda)^{1_{B=0}T'_n})$$

and

$$\begin{aligned}\mathbb{E}\left((1-\lambda)^{1_{B=0}T'_n}\right) &= \mathbb{E}\left(1_{B=0}(1-\lambda)^{T'_n}\right) + \mathbb{E}(1_{B=1}) \\ &= \frac{1}{2}\left(1 + \mathbb{E}\left((1-\lambda)^{T'_n}\right)\right)\end{aligned}$$

Let  $x_n := \mathbb{E}\left((1-\lambda)^{T_n}\right) = \mathbb{E}\left((1-\lambda)^{T'_n}\right)$ . We have

$$x_n = \frac{1-\lambda}{2}x_{n-1}(1+x_n)$$

Dividing by  $x_n x_{n-1}$  (clearly all  $x_n$  are strictly positive) and taking  $u_n := \frac{1}{x_n}$  yields

$$u_{n-1} = \frac{1-\lambda}{2}(1+u_n)$$

and

$$u_n = \frac{2}{1-\lambda}u_{n-1} - 1$$

Let  $v_n := u_n - \frac{1-\lambda}{1+\lambda}$ . Then

$$v_n = \frac{2}{1-\lambda}v_{n-1}$$

Since  $T_0 = 0$ ,  $u_0 = 1$ ,  $v_0 = \frac{2\lambda}{1+\lambda}$  and we have for all  $n \in \mathbb{N}^*$

$$v_n = 2^n(1-\lambda)^{-n} \frac{2\lambda}{1+\lambda}$$

Thus

$$u_n = \frac{2^{n+1}\lambda(1-\lambda)^{-n} + 1 - \lambda}{1 + \lambda}$$

Replacing  $u_n$  by  $\frac{1}{x_n}$  gives the lemma.

We now give the expression of  $\gamma_\lambda^{1^{++}}(a, b)$ :

**Proposition 2.6**

$$\gamma_\lambda^{1^{++}}(a, b) = \frac{1 - f_\lambda(b)}{1 - f_\lambda(a)f_\lambda(b)}$$

where

$$f_\lambda(n) = \frac{(1-2^{-n})(1-\lambda^2)}{1 + 2^{n+1}\lambda(1-\lambda)^{-n} - \lambda}$$

**Proof**  $\gamma_\lambda^{1^{++}}(a, b)$  satisfies the following recursive equation:

$$\begin{aligned}\gamma_\lambda^{1^{++}}(a, b) &= 2^{-b} + (1-2^{-b})\mathbb{E}\left(\sum_{m=1}^{T_b^2} \lambda(1-\lambda)^{m-1}1 + \sum_{m=T_b^2+1}^{T_a^1+T_b^2} \lambda(1-\lambda)^{m-1}0\right) \\ &\quad + (1-2^{-b})(1-2^{-a})\mathbb{E}\left((1-\lambda)^{T_a^1+T_b^2}\right)\gamma_\lambda^{1^{++}}(a, b)\end{aligned}$$

The  $2^{-b}$  corresponds to the probability that the game be absorbed in  $1^*$  when Player 2 plays  $Q$ : in this case the payoff at any step is 1. If the game is not absorbed at that point, then the payoff from step 1

until the step when Player 1 plays  $Q$  is the second term of the equation. When Player 1 plays  $Q$ , with probability  $2^{-a}$  the game is absorbed in  $0^*$ , and with probability  $(1 - 2^{-a})$  the game is back in state  $1^{++}$ : it is the third term of the equation.

We deduce that

$$\gamma_\lambda^{1^{++}}(a, b) = \frac{1 - (1 - 2^{-b})\mathbb{E}\left((1 - \lambda)^{T_b^2}\right)}{1 - (1 - 2^{-a})(1 - 2^{-b})\mathbb{E}\left((1 - \lambda)^{T_a^1 + T_b^2}\right)}$$

Since  $T_a^1$  and  $T_b^2$  are independent, we have  $\mathbb{E}\left((1 - \lambda)^{T_a^1 + T_b^2}\right) = \mathbb{E}\left((1 - \lambda)^{T_a^1}\right)\mathbb{E}\left((1 - \lambda)^{T_b^2}\right)$ . Recall now that  $T_a^1$  (resp.  $T_b^2$ ) and  $T_a + 1$  (resp.  $T_b + 1$ ) have the same law. It yields

$$\gamma_\lambda^{1^{++}}(a, b) = \frac{1 - (1 - 2^{-b})(1 - \lambda)\mathbb{E}\left((1 - \lambda)^{T_b}\right)}{1 - (1 - \lambda)^2(1 - 2^{-a})(1 - 2^{-b})\mathbb{E}\left((1 - \lambda)^{T_a}\right)\mathbb{E}\left((1 - \lambda)^{T_b}\right)}$$

We get the proposition by applying Lemma 2.5.

## 2.4 Asymptotic study of $G_\lambda$ and proof of the main theorem

We first determine optimal strategies in  $G_\lambda$ :

**Proposition 2.7** *Let  $(a^*, b^*) \in \operatorname{argmax}_{n \in \mathbb{N}} f_\lambda \times \operatorname{argmax}_{n \in 2\mathbb{N}} f_\lambda$ . Then  $a^*$  (resp.  $b^*$ ) is a dominant strategy for Player 1 (resp. 2) in  $G_\lambda$ . In particular, they are optimal strategies in  $\Gamma_\lambda^{1^{++}}$ .*

**Proof** We have  $\lim_{n \rightarrow +\infty} f_\lambda(n) = 0$ , therefore  $a^*$  and  $b^*$  are well defined. Observe that the function  $(x, y) \rightarrow \frac{1 - y}{1 - xy}$ , defined on  $[0, 1]^2$ , is increasing in  $x$  and decreasing in  $y$ , and that  $f_\lambda(\mathbb{N}) \subset [0, 1]$ . Hence  $a^*$  and  $b^*$  are dominant strategies in  $G_\lambda$ , and by Proposition 2.2 they are optimal strategies in  $\Gamma_\lambda^{1^{++}}$ .

To study  $f_\lambda$ , it is convenient to make the change of variables  $r = 2^{-n}$ , and define

$$\widehat{f}_\lambda : \begin{cases} [0, 1] & \longrightarrow \mathbb{R} \\ r & \longmapsto \frac{1 - r}{1 + 2\lambda r^{-s} - \lambda} \quad \text{where } s := 1 - \frac{\ln(1 - \lambda)}{\ln(2)} > 1. \end{cases}$$

Note that for all  $n \in \mathbb{N}$ ,  $f_\lambda(n) = (1 - \lambda^2)\widehat{f}_\lambda(2^{-n})$ .

**Lemma 2.8**  *$\widehat{f}_\lambda$  reaches its maximum at one unique point  $r^*(\lambda)$ , is strictly increasing on  $[0, r^*(\lambda)]$ , and strictly decreasing on  $[r^*(\lambda), 1]$ . Moreover, for all  $c > 0$ ,  $\widehat{f}_\lambda(c\sqrt{2\lambda}) \underset{\lambda \rightarrow 0}{=} 1 - (c + c^{-1})\sqrt{2\lambda} + o(\sqrt{\lambda})$ , and  $r^*(\lambda) \underset{\lambda \rightarrow 0}{\sim} \sqrt{2\lambda}$ .*

**Proof**

$$\widehat{f}'_\lambda(r) = \frac{-(1 + 2\lambda r^{-s} - \lambda) - (1 - r)(-2\lambda s r^{-s-1})}{(1 + 2\lambda r^{-s} - \lambda)^2}$$

The numerator of this expression is equal to  $h_\lambda(r) := \lambda - 1 + 2\lambda(-(1 + s)r + s)r^{-s-1}$ . We have

$$h'_\lambda(r) = -2\lambda(s(1 + s)(1 - r))r^{-s-2}$$

We have  $h'_\lambda < 0$  on  $(0, 1)$ ,  $\lim_{r \rightarrow 0} h_\lambda(r) = +\infty$  and  $h_\lambda(1) = -(1 + \lambda)$ . Hence there exists  $r^*(\lambda) \in (0, 1)$  such that  $h_\lambda$  is strictly positive on  $(0, r^*(\lambda)]$ , and strictly negative on  $[r^*(\lambda), 1]$ . Thus  $\widehat{f}_\lambda$  is strictly increasing on  $[0, r^*(\lambda)]$ , and strictly decreasing on  $[r^*(\lambda), 1]$ .

If  $c > 0$ , we have

$$\widehat{f}_\lambda(c\sqrt{2\lambda}) \underset{\lambda \rightarrow 0}{=} 1 - (c + c^{-1})\sqrt{2\lambda} + o(\sqrt{\lambda})$$

Let  $\epsilon > 0$ . Applying the last relation to  $c = 1 - \epsilon$ ,  $c = 1$  and  $c = 1 + \epsilon$  shows that for  $\lambda$  small enough,  $\widehat{f}_\lambda(\sqrt{2\lambda}) > f_\lambda((1 - \epsilon)\sqrt{2\lambda})$  and  $\widehat{f}_\lambda(\sqrt{2\lambda}) > f_\lambda((1 + \epsilon)\sqrt{2\lambda})$ . Thus, for  $\lambda$  small enough,  $r^*(\lambda) \in \left[ (1 - \epsilon)\sqrt{2\lambda}, (1 + \epsilon)\sqrt{2\lambda} \right]$ . We deduce that  $r^*(\lambda) \underset{\lambda \rightarrow 0}{\sim} \sqrt{2\lambda}$ .

We can now prove our main result:

**Theorem 2.9**  $(v_\lambda)$  does not converge when  $\lambda \rightarrow 0$ .

**Proof** Set  $\lambda_m = 2^{-4m-1}$  and  $\mu_m = 2^{-4m-3}$ . Hence  $\sqrt{2\lambda_m} = 2^{-2m}$  and  $\sqrt{2\mu_m} = 2^{-2m-1}$ . By Lemma 2.8, for  $m$  big enough,

$$\operatorname{argmax}_{n \in \mathbb{N}} f_{\lambda_m} = \operatorname{argmax}_{n \in 2\mathbb{N}} f_{\lambda_m} = \{2m\}$$

Hence by Proposition 2.7, we have

$$v_{\lambda_m}(1^{++}) = \frac{1 - f_{\lambda_m}(2m)}{1 - f_{\lambda_m}(2m)^2} = \frac{1}{1 + f_{\lambda_m}(2m)}$$

By Lemma 2.8,  $f_{\lambda_m}(2m)$  converges to 1, thus  $(v_{\lambda_m}(1^{++}))$  converges to  $\frac{1}{2}$ .

Still by Lemma 2.8, for  $m$  big enough, we have

$$\operatorname{argmax}_{n \in \mathbb{N}} f_{\mu_m} = \{2m + 1\} \quad \text{and} \quad \operatorname{argmax}_{n \in 2\mathbb{N}} f_{\mu_m} \subset \{2m, 2m + 2\}$$

Therefore

$$v_{\mu_m}(1^{++}) = \min \left( \frac{1 - f_{\mu_m}(2m)}{1 - f_{\mu_m}(2m)f_{\mu_m}(2m + 1)}, \frac{1 - f_{\mu_m}(2m + 2)}{1 - f_{\mu_m}(2m + 2)f_{\mu_m}(2m + 1)} \right)$$

By Lemma 2.8, we have

$$\begin{aligned} f_{\mu_m}(2m + 1) &\underset{m \rightarrow +\infty}{=} 1 - 2\sqrt{2\mu_m} + o(\sqrt{\mu_m}) \\ f_{\mu_m}(2m) &\underset{m \rightarrow +\infty}{=} 1 - \frac{5}{2}\sqrt{2\mu_m} + o(\sqrt{\mu_m}) \\ f_{\mu_m}(2m + 2) &\underset{m \rightarrow +\infty}{=} 1 - \frac{5}{2}\sqrt{2\mu_m} + o(\sqrt{\mu_m}) \end{aligned}$$

Hence

$$\frac{1 - f_{\mu_m}(2m)}{1 - f_{\mu_m}(2m)f_{\mu_m}(2m + 1)} \underset{m \rightarrow +\infty}{\sim} \frac{\frac{5}{2}\sqrt{2\mu_m}}{\left(2 + \frac{5}{2}\right)\sqrt{2\mu_m}} = \frac{5}{9}$$

And similarly

$$\frac{1 - f_{\mu_m}(2m + 2)}{1 - f_{\mu_m}(2m + 2)f_{\mu_m}(2m + 1)} \underset{m \rightarrow +\infty}{\sim} \frac{\frac{5}{2}\sqrt{2\mu_m}}{\left(2 + \frac{5}{2}\right)\sqrt{2\mu_m}} = \frac{5}{9}$$

$(v_{\lambda_m}(1^{++}))$  and  $(v_{\mu_m}(1^{++}))$  converge to a different limit, hence  $(v_\lambda)$  does not converge.

**Remark** More generally, if  $p \in P \setminus \{1^*, 0^*\}$ ,  $(v_\lambda(p))$  does not converge. Indeed, let  $n \in \mathbb{N}$  and  $N \geq n$ , and consider the following strategy  $\sigma$  for Player 1 in  $\Gamma^{0n}$ : play  $C$  until  $p_m = 0_N$ , then play  $Q$ , and play optimal in  $\Gamma^{1^{++}}$ .  $\sigma$  guarantees asymptotically  $v_\lambda(1^{++}) - \frac{1}{N}$  in  $\Gamma^{0n}$ :  $\liminf_{\lambda \rightarrow 0} v_\lambda(0_n) \geq \liminf_{\lambda \rightarrow 0} v_\lambda(1^{++}) - \frac{1}{N}$ , and with  $N \rightarrow +\infty$ ,  $\liminf_{\lambda \rightarrow 0} v_\lambda(0_n) \geq \liminf_{\lambda \rightarrow 0} v_\lambda(1^{++})$ . With the same kind of argument, one can show that for all  $(p, p') \in P^2 \setminus \{1^*, 0^*\}$ ,  $\lim_{\lambda \rightarrow 0} |v_\lambda(p) - v_\lambda(p')| = 0$ , which gives the result.

### 3 From $(v_\lambda)$ to $(v_n)$

Let  $l \geq 2$ . We construct a repeated game with public signals and perfect observation of the actions  $\Gamma(l)$ , which  $\lambda$ -discounted value  $(v_\lambda^l)$  and  $n$ -stage value  $(v_n^l)$  fail to converge, for  $l$  big enough.

The state space is  $K = \{1^{++}, 1^{T_1}, 1^{T_2}, \dots, 1^{T_{2l-1}}, 1^+, 1^*, 0^{++}, 0^{T_1}, 0^{T_2}, \dots, 0^{T_{l-1}}, 0^+, 0^*\}$ , actions sets are  $I = J = \{C, Q\}$ , and signal sets are  $A = B = \{D, D'\}$ . The initial state will usually be taken as  $1^{++}$ . Payoffs are independent of actions, and are 1 in states belonging to  $1^*, 1^{++}, 1^{T_1}, \dots, 1^{T_{2l-1}}, 1^+$ , and 0 in states belonging to  $0^*, 0^{++}, 0^{T_1}, \dots, 0^{T_{l-1}}, 0^+$ .

$1^{++}, 1^{T_1}, \dots, 1^{T_{2l-1}}, 1^+$  are controlled by Player 2, in the sense that the transition on these states  $q(1^{++}, \cdot), q(1^{T_1}, \cdot), \dots, q(1^{T_{2l-1}}, \cdot)$  and  $q(1^+, \cdot)$  do not depend on the actions of Player 1.

Similarly, Player 1 controls the states  $0^{++}, 0^{T_1}, \dots, 0^{T_{l-1}}, 0^+$ . Hence  $q$  can be seen as a map from  $K \times \{C, Q\}$  to  $\Delta(K \times \{D, D'\})$ . Last,  $1^*$  and  $0^*$  are absorbing states.

The following table describes the transitions of the game in the states controlled by Player 2, that is to say  $1^{++}, 1^{T_1}, \dots, 1^{T_{2l-1}}, 1^+$ .  $m$  is any integer in  $[[0, 2l - 2]]$ , and  $1^{T_0} := 1^{++}$ .

$$\begin{aligned} q(1^{T_m}, C) &:= \frac{1}{2} \cdot (1^{T_{m+1}}, D) + \frac{1}{2} \cdot (1^{++}, D') \\ q(1^{T_{2l-1}}, C) &:= 2^{-2l-1} \cdot (1^{++}, D) + \frac{1}{2}(1 - 2^{-2l}) \cdot (1^+, D) + \frac{1}{2} \cdot (1^{++}, D') \\ q(1^+, C) &:= \frac{1}{2} \cdot (1^+, D) + \frac{1}{2} \cdot (1^{++}, D') \\ q(1^{T_m}, Q) &:= q(1^{T_{2l-1}}, Q) := (1^*, D') \\ q(1^+, Q) &:= (0^{++}, D) \end{aligned}$$

We now describe the transitions in the states controlled by Player 1, that is to say  $0^{++}, 0^{T_1}, \dots, 0^{T_{l-1}}, 0^+$ .  $m$  is any integer in  $[[0, l - 2]]$ , and  $0^{T_0} := 0^{++}$ .

$$\begin{aligned} q(0^{T_m}, C) &:= \frac{1}{2} \cdot (0^{T_{m+1}}, D) + \frac{1}{2} \cdot (0^{++}, D') \\ q(0^{T_{l-1}}, C) &:= 2^{-l-1} \cdot (0^{++}, D) + \frac{1}{2}(1 - 2^{-l}) \cdot (0^+, D) + \frac{1}{2} \cdot (0^{++}, D') \\ q(0^+, C) &:= \frac{1}{2} \cdot (0^+, D) + \frac{1}{2} \cdot (0^{++}, D') \\ q(0^{T_m}, Q) &:= q(0^{T_{l-1}}, Q) := (0^*, D') \\ q(0^+, Q) &:= (1^{++}, D) \end{aligned}$$

Recall that we have shown that for  $\lambda \in (0, 1]$  fixed, the discounted game  $\Gamma_\lambda$  of section 2 was equivalent in terms of value and optimal strategies to a one-shot game with actions sets  $\mathbb{N}$  for Player 1 (resp.  $2\mathbb{N}$  for Player 2) and with payoff function  $g_\lambda$  which expression is given in Proposition 2.2.

With exactly the same analysis, we can show that  $\Gamma(l)_\lambda^{1^{++}}$  is equivalent to a game played on  $l\mathbb{N} \times 2l\mathbb{N}$ :

**Proposition 3.1**  $\Gamma(l)_\lambda^{1^{++}}$  has the same value as the one-shot game  $G_\lambda(l)$ , with action set  $l\mathbb{N}$  for Player 1,  $2l\mathbb{N}$  for Player 2, and payoff

$$g_\lambda^l(a, b) := g_\lambda(a, b)$$

Moreover, optimal strategies in  $G_\lambda(l)$  induce optimal strategies in  $\Gamma(l)_\lambda^{1^{++}}$ .

For  $m \geq 1$ , let  $\lambda_m := 2^{-4lm-1}$  and  $\mu_m := 2^{-4lm-2l-1}$ . Proceeding exactly the same way as in 2.4, we get

**Proposition 3.2**

$$\lim_{m \rightarrow +\infty} v_{\lambda_m}^l(1^{++}) = \frac{1}{2} \quad \text{and} \quad \lim_{m \rightarrow +\infty} v_{\mu_m}^l(1^{++}) = \frac{2^l + 2^{-l}}{2^l + 2^{-l} + 2}$$

We will show that for  $l$  big enough, the value  $(v_n^l)$  of the game  $\Gamma(l)_n$  does not converge by comparing it to the value  $(v_\lambda^l)$  of the game  $\Gamma(l)_\lambda$ , using a similar technique as Vigerat [24].

We fix some  $l \geq 2$ .

We start with a lemma, which can be immediately deduced from the proof of Theorem C.8 in Sorin [22]:

**Lemma 3.3** *Let  $\Gamma$  be any repeated game with public signals and perfect observation of the actions. Let  $n_0, n \in \mathbb{N}^*$ , and for  $m \in \mathbb{N}^*$  set  $w_m := v_{\frac{\cdot}{m}}$ . Then the following inequality holds:*

$$\|v_n - w_n\|_\infty \leq \frac{n_0}{n} \|v_{n_0} - w_{n_0}\|_\infty + \sum_{m=n_0}^{n-1} \|w_m - w_{m+1}\|_\infty$$

**Proof** Let  $m \geq 1$  and  $p \in \Delta(K)$ . We have the following dynamic programming principle (see Sorin [22]):

$$v_m(p) = \max_{x \in \Delta(I)} \min_{y \in \Delta(J)} \left\{ \frac{1}{m} g(p, x, y) + \frac{m-1}{m} \mathbb{E}_{x,y}^p(v_{m-1}) \right\} \quad (4)$$

and

$$w_m(p) = \max_{x \in \Delta(I)} \min_{y \in \Delta(J)} \left\{ \frac{1}{m} g(p, x, y) + \frac{m-1}{m} \mathbb{E}_{x,y}^p(w_m) \right\} \quad (5)$$

Let  $x \in \Delta(I)$  optimal in (4) and  $y \in \Delta(J)$  optimal in (5). We have

$$v_m(p) \leq \frac{1}{m} g(p, x, y) + \frac{m-1}{m} \mathbb{E}_{x,y}^p(v_{m-1})$$

$$w_m(p) \geq \frac{1}{m} g(p, x, y) + \frac{m-1}{m} \mathbb{E}_{x,y}^p(w_m)$$

The combination of these two inequalities gives

$$v_m(p) - w_m(p) \leq \frac{m-1}{m} \|v_{m-1} - w_m\|_\infty$$

Taking  $x' \in \Delta(I)$  optimal in (4) and  $y' \in \Delta(J)$  optimal in (5) gives the symmetric inequality:

$$w_m(p) - v_m(p) \leq \frac{m-1}{m} \|v_{m-1} - w_m\|_\infty$$

Hence

$$\|v_m - w_m\|_\infty \leq \frac{m-1}{m} \|v_{m-1} - w_m\|_\infty$$

and

$$m \|v_m - w_m\|_\infty \leq (m-1) \|v_{m-1} - w_{m-1}\|_\infty + (m-1) \|w_{m-1} - w_m\|_\infty$$

Let  $n, n_0 \geq 1$ . Summing the last inequality from  $n_0 + 1$  to  $n$  yields

$$n \|v_n - w_n\|_\infty \leq n_0 \|v_{n_0} - w_{n_0}\|_\infty + \sum_{m=n_0}^{n-1} m \|w_m - w_{m+1}\|_\infty$$

Dividing by  $n$  gives the lemma.

To have more simple notations, if  $(\lambda, p) \in [0, 1) \times \Delta(K)$ ,  $v'_\lambda(p)$  designates the derivative of  $v_\lambda(p)$  with respect to  $\lambda$ , evaluated in  $\lambda$ , and it is the same for  $f'_\lambda(p)$ . Moreover, when there is no ambiguity, we will write  $v_n$  for  $v_n^l$  and  $v_\lambda$  for  $v_\lambda^l$ .

**Remark** When  $\|v'_\lambda\|_\infty \underset{\lambda \rightarrow 0}{=} o(\lambda^{-1})$ , one can easily deduce from the above inequality that  $(v_n)$  and  $(v_\lambda)$  have the same accumulation points (see Vigerat [24]). In our example we can only say that  $\|v'_\lambda\|_\infty \underset{\lambda \rightarrow 0}{=} O(\lambda^{-1})$ , which is not sufficient. We need a sharper majoration of the derivative, given by the following lemma.

For  $m \geq 1$ , we define  $\lambda_m := 2^{-4lm-1}$  and  $\mu_m := 2^{-4lm-2l-1}$ .

**Lemma 3.4** *There exists  $m_0 \geq 1$  such that for all  $m \geq m_0$  and  $\mu_m \leq \mu \leq 2^{\frac{1}{2}-1}\mu_m$*

$$\|v'_\mu\|_\infty \leq 2^{-\frac{1}{2}}\mu^{-1}$$

and for all  $\lambda_m \leq \lambda \leq 2^{l-1}\lambda_m$

$$\|v'_\lambda\|_\infty \leq 2^{\frac{l}{2}+2}\lambda^{-\frac{1}{2}}$$

**Proof** Let  $\mu \in \bigcup_{m \geq 1} [\mu_m, 2^{l-1}\mu_m]$ . We denote by  $m(\mu)$  the unique  $m$  such that  $\mu \in [\mu_m, 2^{l-1}\mu_m]$ . By a similar analysis as in Lemma 2.8, one can easily show that for  $\mu$  small enough,  $a(\mu) = 2lm(\mu) + l$  is an optimal strategy for Player 1 in  $G_\mu(l)$ . Moreover, for  $\mu$  small enough, there exists  $\mu_{m(\mu)}^0 \leq \mu_{m(\mu)} \leq 2^{l-1}\mu_{m(\mu)}$  such that for  $\mu_{m(\mu)} \leq \mu \leq \mu_{m(\mu)}^0$  (resp.  $\mu_{m(\mu)}^0 \leq \mu \leq 2^{l-1}\mu_{m(\mu)}$ ),  $b'(\mu) = 2lm(\mu) + 2l$  (resp.  $b(\mu) = 2lm(\mu)$ ) is an optimal strategy for Player 2 in  $G_\mu(l)$ . The two cases  $b(\mu) = 2lm(\mu)$  and  $b'(\mu) = 2lm(\mu) + 2l$  being similar, we only treat the first one.

Let  $C_1(\mu) := 2^{-2lm(\mu)-l}(2\mu)^{-\frac{1}{2}} = \sqrt{\frac{\mu_{m(\mu)}}{\mu}}$  and  $C_2(\mu) := 2^{-2lm(\mu)}(2\mu)^{-\frac{1}{2}} = 2^l \sqrt{\frac{\mu_{m(\mu)}}{\mu}}$ . Note that  $2^{-a(\mu)} = C_1(\mu)\sqrt{2\mu}$  and  $2^{-b(\mu)} = C_2(\mu)\sqrt{2\mu}$ . Moreover,  $C_1$  and  $C_2$  are bounded and bounded away from 0.

The fact that optimal strategies in  $\Gamma_\mu^{1++}$  are locally constant with respect to  $\mu$  allows us to compute easily the derivative of  $v_\mu$ . First

$$f'_\mu(a(\mu)) = (1-2^{-a(\mu)}) \frac{-2\mu(1+2^{a(\mu)+1}\mu(1-\mu)^{-a(\mu)} - \mu) - (1-\mu^2)(2^{a(\mu)+1}((1-\mu)^{-a(\mu)} - a(\mu)\mu(1-\mu)^{a(\mu)-1}) - 1)}{(1+2^{a(\mu)+1}\mu(1-\mu)^{-a(\mu)} - \mu)^2}$$

We deduce that

$$f'_\mu(a(\mu)) \underset{\mu \rightarrow 0}{=} -2C_1(\mu)^{-1}(2\mu)^{-\frac{1}{2}} + o\left(\mu^{-\frac{1}{2}}\right) \quad (6)$$

The same equality holds replacing  $a(\mu)$  by  $b(\mu)$  and  $C_1(\mu)$  by  $C_2(\mu)$ .

The same computation as in Lemma 2.8 gives

$$\begin{aligned} f_\mu(a(\mu)) \underset{\mu \rightarrow 0}{=} & 1 - (C_1(\mu) + C_1(\mu)^{-1})(2\mu)^{\frac{1}{2}} + o\left(\mu^{\frac{1}{2}}\right) \\ f_\mu(b(\mu)) \underset{\mu \rightarrow 0}{=} & 1 - (C_2(\mu) + C_2(\mu)^{-1})(2\mu)^{\frac{1}{2}} + o\left(\mu^{\frac{1}{2}}\right) \end{aligned}$$

We can now differentiate  $v_\mu(1^{++})$  (we omit the dependence of  $a$  and  $b$  in  $\mu$ ):

$$\begin{aligned} v'_\mu(1^{++}) = g'_\mu(a, b) &= \frac{-f'_\mu(b)(1-f_\mu(a)f_\mu(b)) + (1-f_\mu(b))(f'_\mu(a)f_\mu(b) + f_\mu(a)f'_\mu(b))}{(1-f_\mu(a)f_\mu(b))^2} \\ &= \frac{f'_\mu(a)f_\mu(b)(1-f_\mu(b)) - f'_\mu(b)(1-f_\mu(a))}{(1-f_\mu(a)f_\mu(b))^2} \end{aligned}$$

When  $\mu$  goes to 0, the numerator of this expression is (we omit the dependence in  $\mu$ )

$$-2C_1^{-1}(C_2 + C_2^{-1}) + 2C_2^{-1}(C_1 + C_1^{-1}) + o(1) = 2(C_2^{-1}C_1 - C_1^{-1}C_2) + o(1)$$

Hence

$$v'_\mu(1^{++}) \underset{\mu \rightarrow 0}{=} \frac{(C_2^{-1}C_1 - C_1^{-1}C_2)}{(C_1 + C_1^{-1} + C_2 + C_2^{-1})^2} \mu^{-1} + o(\mu^{-1})$$

If  $\mu_m \leq \mu \leq 2^{\frac{1}{2}-1}\mu_m$ , we have

$$\left| \frac{(C_2^{-1}C_1 - C_1^{-1}C_2)}{(C_1 + C_1^{-1} + C_2 + C_2^{-1})^2} \right| \leq C_1^{-1}C_2^{-1} = 2^{-l} \frac{\mu}{\mu_{m(\mu)}} \leq 2^{-\frac{1}{2}-1}$$

The last two relations show that for  $m$  big enough and  $\mu_m \leq \mu \leq 2^{\frac{l}{2}-1}\mu_m$

$$|v'_\mu(1^{++})| \leq 2^{-\frac{l}{2}-1}\mu^{-1}$$

We can easily show that  $\mu \rightarrow \mu^{\frac{l}{2}} |v'_\mu(p) - v'_\mu(1^{++})|$  is uniformly bounded in  $p \in \Delta(K)$  such that  $p(0^*) = p(1^*) = 0$ . Indeed for example

$$v_\mu(0^{++}) = \mathbb{E} \left( 2^{-a(\mu)} + (1 - 2^{-a(\mu)}) \left( \sum_{m=1}^{T_{a(\mu)}} \mu(1 - \mu)^{m-1} + (1 - \mu)^{T_{a(\mu)}} v_\mu(1^{++}) \right) \right)$$

Thus

$$v_\mu(0^{++}) = 1 - f_\mu(a(\mu))v_\mu(1^{++})$$

and

$$v'_\mu(0^{++}) = -f'_\mu(a(\mu))v_\mu(1^{++}) - f_\mu(a(\mu))v'_\mu(1^{++})$$

and (6) gives the result. Hence for  $m$  big enough and  $\mu_m \leq \mu \leq 2^{\frac{l}{2}-1}\mu$

$$\|v'_\mu\|_\infty \leq 2^{-\frac{l}{2}}\mu^{-1}$$

The proof of the second part is similar. Let  $\lambda \in \bigcup_{m \geq 1} [\lambda_m, 2^{l-1}\lambda_m]$ . We denote by  $m(\lambda)$  the unique  $m$  such that  $\lambda \in [\lambda_m, 2^{l-1}\lambda_m]$ . For  $\lambda$  small enough,  $a(\lambda) = 2lm(\lambda)$  is an optimal strategy for both players in  $G_\lambda(l)$ , hence for such a  $\lambda$

$$v'_\lambda(1^{++}) = -\frac{f_\lambda(a(\lambda))'}{(1 + f_\lambda(a(\lambda)))^2}$$

Let  $C(\lambda) := \sqrt{\frac{\lambda_{m(\lambda)}}{\lambda}}$ . Then as in (6)

$$f'_\lambda(a(\lambda)) \underset{\lambda \rightarrow 0}{=} -2C(\lambda)^{-1}(2\lambda)^{-\frac{l}{2}} + o\left(\lambda^{-\frac{l}{2}}\right)$$

Since  $C(\lambda)^{-1} \leq 2^{\frac{l-1}{2}}$  and  $f_\lambda(a(\lambda))$  goes to 1 when  $\lambda$  goes to 0, we get the desired inequality for  $v'_\lambda(1^{++})$ , and it extends to  $v'_\lambda(p)$  in the same way.

**Theorem 3.5** *There exists  $l_0 \in \mathbb{N}^*$  such that for all  $l \geq l_0$ ,  $(v_n^l)$  and  $(v_\lambda^l)$  do not converge.*

**Proof** Let  $l \geq 2$ . Recall that from Proposition 3.2,  $v_{\lambda_m}^l \rightarrow \frac{1}{2}$  and  $v_{\mu_m}^l \rightarrow \frac{2^l + 2^{-l}}{2^l + 2^{-l} + 2} := w(l)$ . In particular,  $(v_\lambda^l)$  does not converge.

Let  $m \geq m_0$ . Let  $n(m) := \mu_m^{-1} = 2^{4l+2l+1}$  and  $n_0(m) := 2^{-\lfloor \frac{l}{2} \rfloor + 1} n(m)$ . We compare  $v_{n(m)}^l$  and  $v_{\mu_m}^l$ , using Lemma 3.3:

$$\|v_{n(m)}^l - v_{\mu_m}^l\|_\infty \leq \frac{n_0(m)}{n(m)} \left\| v_{n_0(m)}^l - v_{2^{\lfloor \frac{l}{2} \rfloor - 1} \mu_m}^l \right\|_\infty + \sum_{m'=n_0(m)}^{n(m)-1} \left\| v_{\frac{1}{m'}}^l - v_{\frac{1}{m'+1}}^l \right\|_\infty$$

The term on the left is smaller than  $2^{-\lfloor \frac{l}{2} \rfloor + 1}$ , and by the Mean Value theorem and Lemma 3.4 the term on the right is smaller than  $2^{-\frac{l}{2}} \int_{(n(m))^{-1}}^{(n_0(m))^{-1}} \frac{1}{x} dx = 2^{-\frac{l}{2}} \left( \left\lfloor \frac{l}{2} \right\rfloor - 1 \right)$ . Making  $m$  going to infinity, we deduce that

$$\left| \liminf_{m \rightarrow +\infty} v_{n(m)}^l(1^{++}) - w(l) \right| \leq 2^{-\lfloor \frac{l}{2} \rfloor + 1} + 2^{-\frac{l}{2}} \left( \left\lfloor \frac{l}{2} \right\rfloor - 1 \right)$$



Note that  $\lim_{l \rightarrow +\infty} w(l) = 1$  and that the term on the right goes to 0 when  $l$  goes to infinity.

Applying again lemma 3.3 for  $n(m) := \lambda_m^{-1}$  and  $n_0(m) := 2^{-\lfloor \frac{l}{2} \rfloor + 1} n(m)$  gives also an inequality of the form

$$\left| \liminf_{m \rightarrow +\infty} v_{n(m)}^l(1^{++}) - \frac{1}{2} \right| \leq t(l)$$

where  $\lim_{l \rightarrow +\infty} t(l) = 0$ . Hence for  $l$  big enough,  $(v_n^l(1^{++}))$  does not converge.

## 4 Extension to other classes of games

In this section, we show that the counterexample can be adapted to fit in other classes of games.

### 4.1 State-blind repeated games

Let us consider the following state-blind repeated game  $\Gamma$ , with state space  $K = \{1^*, 1^{++}, 1^T, 1^+, 0^*, 0^{++}, 0^+\}$ , action sets  $I = \{T, B, Q\}$  for Player 1 and  $J = \{L, R, Q\}$  for Player 2.  $0^*$  and  $1^*$  are absorbing states. The payoff is 1 in states  $1^{++}$ ,  $1^T$  and  $1^+$ , and 0 in states  $0^{++}$  and  $0^+$ . The transitions are described below:

1 <sup>++</sup>			
	L	R	Q
T	1 <sup>++</sup>	1 <sup>T</sup>	1 <sup>*</sup>
B	1 <sup>T</sup>	1 <sup>++</sup>	1 <sup>*</sup>
Q	0 <sup>*</sup>	0 <sup>*</sup>	0 <sup>*</sup>

1 <sup>T</sup>			
	L	R	Q
T	1 <sup>++</sup>	$\frac{3}{4}1^+ + \frac{1}{4}1^{++}$	1 <sup>*</sup>
B	$\frac{3}{4}1^+ + \frac{1}{4}1^{++}$	1 <sup>++</sup>	1 <sup>*</sup>
Q	0 <sup>*</sup>	0 <sup>*</sup>	0 <sup>*</sup>

1 <sup>+</sup>			
	L	R	Q
T	1 <sup>++</sup>	1 <sup>+</sup>	0 <sup>++</sup>
B	1 <sup>+</sup>	1 <sup>++</sup>	0 <sup>++</sup>
Q	0 <sup>*</sup>	0 <sup>*</sup>	0 <sup>*</sup>

0 <sup>++</sup>			
	L	R	Q
T	$\frac{1}{2}0^{++} + \frac{1}{2}0^+$	0 <sup>++</sup>	1 <sup>*</sup>
B	0 <sup>++</sup>	$\frac{1}{2}0^{++} + \frac{1}{2}0^+$	1 <sup>*</sup>
Q	0 <sup>*</sup>	0 <sup>*</sup>	1 <sup>*</sup>

0 <sup>+</sup>			
	L	R	Q
T	0 <sup>+</sup>	0 <sup>++</sup>	1 <sup>*</sup>
B	0 <sup>++</sup>	0 <sup>+</sup>	1 <sup>*</sup>
Q	1 <sup>++</sup>	1 <sup>++</sup>	1 <sup>*</sup>

Recall that in this model, both players observe nothing about the state, but observe past actions.

As in section 2, we consider the equivalent game with full information played on  $\Delta(K)$ . As in the example with public signals,  $P$  (see section 2.2 for the notations) is stable under the dynamics of the game: leaving from some  $p \in P$ , the state remains in  $P$ . It is clearly optimal for Player 1 (resp. 2) to play  $\left(\frac{1}{2}, \frac{1}{2}\right)$  when the state lies in  $P_2$  (resp.  $P_1 \cup P_1^T$ ). Under this type of strategy, the dynamics of the game coincide with the counterexample of section 2. Therefore the values  $(v_\lambda)$  and  $(v_n)$  are the same, and thus do not converge.

### 4.2 Repeated games with one informed player\*

We now investigate a repeated game with perfect observation of the actions where Player 2 is fully informed about the state, but Player 1 has no information about it. As usual, both players observe past actions.

The state space is  $K = \{1^*, 1, 0^*, 0^{++}, 0^+\}$ , action sets are  $I = \{T, B, Q\}$  for Player 1 and  $J = \{L, R\}$  for Player 2.  $0^*$  and  $1^*$  are absorbing states. The payoff is 1 in state 1, and 0 in states  $0^{++}$  and  $0^+$ . The transitions are described below:

\*We thank Guillaume Vigeral for his help to design this example.

1		
	L	R
T	1	$0^{++}$
B	$0^{++}$	$1^*$
Q	$0^*$	$0^*$

$0^{++}$		
	L	R
T	$\frac{1}{2}0^{++} + \frac{1}{2}0^+$	$0^{++}$
B	$0^{++}$	$\frac{1}{2}0^{++} + \frac{1}{2}0^+$
Q	$0^*$	$0^*$

$0^+$		
	L	R
T	$0^+$	$0^{++}$
B	$0^{++}$	$0^+$
Q	1	1

Compared to the game of subsection 4.1, states  $1^{++}$ ,  $1^T$  and  $1^+$  have been replaced by one single state 1, which is similar to the state  $\omega^+$  of Vigerl [24]. The other states have not been changed.

In the equivalent game played on  $\Delta(K)$ , the game going from 1 remains in  $\{1\} \cup P_2$ . It is easy to see that optimal stationary strategies for Player 1 in  $\Gamma_\lambda$  when the state is in  $P_2$  are the same as in the game of subsection 4.1 (or as in the counterexample of section 2), and that  $(1 - \sqrt{\lambda}, \sqrt{\lambda})$  is an asymptotically optimal strategy for both players when the state is 1 (see Vigerl [24]). Thus when the state is 1, the game stays in 1 a number of steps of order  $\lambda^{-\frac{1}{2}}$ , and the probability of absorbing in  $1^*$  before going to  $0^{++}$  is of order  $\sqrt{\lambda}$ . Hence the dynamics of the game is similar to the example of section 2, and we can prove in the same way that  $(v_\lambda)$  oscillates. Note that in this example, Player 2's situation is even better than Player 1's situation in the previous example. Indeed one can prove that

$$\liminf_{\lambda \rightarrow 0} v_\lambda < \limsup_{\lambda \rightarrow 0} v_\lambda < \frac{1}{2}$$

### 4.3 Stochastic games with compact action sets

We now study a repeated game with perfect observation (states and actions are known by both players) but where  $I$  and  $J$  are compact. It yields an alternative counterexample to Vigerl [24], which is equivalent in terms of dynamics to the example of section 2. The state space is  $K = \{1^*, 1, 0^*, 0\}$ , and actions sets are  $I = [0, 1]$  and  $J = \{0\} \cup \cup_{m \in \mathbb{N}} 4^{-m}$ . The transition  $q$  is defined by  $q(1, x, y) := (1 - y) \cdot 1 + (y - y^2) \cdot 0 + y^2 \cdot 1^*$  and  $q(0, x, y) := (1 - x) \cdot 0 + (x - x^2) \cdot 1 + x^2 \cdot 0^*$ . Hence Player 1 controls 0 and Player 2 controls 1.

Let  $\lambda \in (0, 1]$ . A pure stationary strategy in  $\Gamma_\lambda$  for Player 1 (resp. 2) can be seen as an element of  $I$  (resp.  $J$ ).

**Remark**  $x \in I$  corresponds to the absorbing risk  $2^{-a}$  in the example of section 2. Indeed, when Player 1 plays  $x$  in state 0, on average he waits approximately  $x^{-1}$  steps before switching to state 1, and the probability of absorbing in  $0^*$  before reaching 1 is approximately  $x$ . Recall that in the example of section 2, when Player 1 plays  $a \in \mathbb{N}$ , he waits on average  $2^a$  steps before quitting, and when he quits the game is absorbed in  $0^*$  with probability  $2^{-a}$ . It is the same for Player 2. As in our first example, Player 2 can not take any absorbing risk: only  $y = 4^{-m}$  for some  $m \in \mathbb{N}$ , or  $y = 0$ . But Player 1 can take any absorbing risk in  $[0, 1]$ . That is why we expect  $(v_\lambda)$  to oscillate, just as in the first example.

We now compute the payoff in  $\Gamma_\lambda$  given by a couple of strategies  $(x, y) \in I \times J$ :

$$\gamma_\lambda(x, y) = \frac{(1 - (1 - \lambda)(1 - y^2))(1 - (1 - \lambda)(1 - x))}{(1 - (1 - \lambda)(1 - xy))(1 - (1 - \lambda)(1 - x)(1 - y))}$$

For any  $x \in [0, 1]$  (resp.  $y \in [0, 1]$ )  $\gamma_\lambda(x, \cdot)$  (resp.  $\gamma_\lambda(\cdot, y)$ ) is convex (resp. concave) and reaches its minimum (resp. its maximum) at  $y^*$  (resp.  $x^*$ ) such that

$$x^* = y^* = \frac{\sqrt{\lambda} - \lambda}{1 - \lambda}$$

For  $m \geq 1$ , we define  $\lambda_m := 2^{-2m}$  and  $\mu_m := 2^{-2m-1}$ . Then for  $m$  big enough,  $x_m = y_m = \sqrt{\lambda_m}$  are optimal strategies in  $\Gamma_{\lambda_m}$ . We have

$$\begin{aligned}\gamma_{\lambda_m}(x_m, y_m) &= \frac{(1 - (1 - \lambda_m)(1 - x_m))}{(1 - (1 - \lambda_m)(1 - x_m)(1 - y_m))} \\ &\sim \frac{\sqrt{\lambda_m}}{2\sqrt{\lambda_m}} = \frac{1}{2}\end{aligned}$$

Hence  $\lim_{m \rightarrow +\infty} v_{\lambda_m}(1) = \frac{1}{2}$ .

For  $m$  big enough,  $x_m = \sqrt{\mu_m}$  is an optimal strategy for Player 1 in  $\Gamma_{\mu_m}$ , and either  $y_m = 2\sqrt{\mu_m}$  or  $y'_m = \frac{1}{2}\sqrt{\mu_m}$  is an optimal strategy for Player 2 in  $\Gamma_{\mu_m}$ . We have

$$\gamma_{\mu_m}(x_m, y_m) \sim \frac{(\mu_m + 4\mu_m)\sqrt{\mu_m}}{(\mu_m + 2\mu_m)3\sqrt{\mu_m}} = \frac{5}{9}$$

And similarly  $\gamma_{\mu_m}(x_m, y'_m) \sim \frac{5}{9}$ , hence  $\lim_{m \rightarrow +\infty} v_{\mu_m}(1) = \frac{5}{9}$

Therefore  $(v_\lambda)$  does not converge.

## 5 Conclusion and open problems

We have shown that in the general model of repeated games, the asymptotic value may fail to exist, even when both players have the same information. In addition to the very simple signaling structure (two public signals, perfect observation of the actions), the example of section 2 presents some remarkable properties: players observe the payoffs, and players play in turn. A natural question that arises is the following: can we characterize the class of repeated games which have an asymptotic value? As recalled in the introduction, this class contains standard stochastic games, absorbing games with lack of information on one side, and repeated games with one informed controller. We leave for further research the study of the common denominator between these classes of games.

Concerning the link between asymptotic approach and uniform approach, our example contradicts the conjecture “ $\maxmin = \lim_{n \rightarrow \infty} v_n$  when Player 1 is more informed than Player 2”. Nonetheless, we conjecture that in our example,  $\liminf_{n \rightarrow +\infty} v_n(1^{++}) = \frac{1}{2} = \maxmin(1^{++})$  and  $\limsup_{n \rightarrow +\infty} v_n(1^{++}) = \frac{5}{9} = \minmax(1^{++})$ . It suggests that the conjecture of Mertens could be replaced by “When Player 1 is more informed than Player 2, Player 1 can guarantee  $\maxmin = \liminf v_n$ ”.

We used in section 3 a technique similar to Vigerl [24] to prove the divergence of  $(v_n)$ , exploiting the regularity of  $(v_\lambda)$ . But this method does not apply to the original example of section 2. Is there another way to prove the divergence of  $(v_n)$  in the original example, and to compute  $\liminf_{n \rightarrow +\infty} v_n$  and  $\limsup_{n \rightarrow +\infty} v_n$ ? More generally, the link between the asymptotic properties of  $(v_n)$  and  $(v_\lambda)$  requires further investigation.

Last, our example shows that, in the framework of non-zero sum stochastic games with public signals and perfect observation of the actions, the set of  $\lambda$ -discounted Nash equilibrium payoff  $E_\lambda$  may fail to converge in the sense of the Hausdorff distance and when  $\lambda$  goes to zero. This question of convergence is still open when the public signals include the state and the actions.

## Acknowledgments

I would like to thank Jérôme Renault for his help to clarify the proof of the main result of this paper, and Fabien Gensbittel for his assistance to prove Lemma 2.5. I also thank Guillaume Vigerl, Xavier Venel and Sylvain Sorin for interesting suggestions.

## References

- [1] R.J. Aumann and M. Maschler. Repeated games with incomplete information. *MIT Press*, 1995.
- [2] T. Bewley and E. Kohlberg. The asymptotic theory of stochastic games. *Mathematics of Operations Research*, 1(3):197–208, 1976.
- [3] J-M. Coulomb. Games with a recursive structure. In *Stochastic games and applications*, volume 570, pages 427–442. Springer, 2003.
- [4] J-M Coulomb. Stochastic games without perfect monitoring. *International journal of game theory*, 32(1):73–96, 2003.
- [5] F. Gensbittel, M. Oliu-Barton, and X. Venel. Existence of the uniform value in repeated games with a more informed controller. *arXiv preprint arXiv:1301.1935*, 2013.
- [6] J.F. Mertens. Repeated games. *Proceedings of the International Congress of Mathematicians Berkeley, California, USA*, pages 1528–1577, 1986.
- [7] J.F. Mertens and A. Neyman. Stochastic games. *International Journal of Game Theory*, 10(2):53–66, 1981.
- [8] J.F. Mertens, S. Sorin, and S. Zamir. *Repeated games*. Univ. Catholique de Louvain, Center for Operations Research & Econometrics, 1994.
- [9] J.F. Mertens and S. Zamir. The value of two-person zero-sum repeated games with lack of information on both sides. *International Journal of Game Theory*, 1(1):39–64, 1971.
- [10] M. Oliu-Barton. The asymptotic value in finite stochastic games. *arXiv preprint arXiv:1301.2450*, 2013.
- [11] J. Renault. The value of markov chain games with lack of information on one side. *Mathematics of Operations Research*, 31(3):490–512, 2006.
- [12] J. Renault. Uniform value in dynamic programming. *Journal of the European Mathematical Society*, 13(2):309–330, 2011.
- [13] J. Renault. The value of repeated games with an informed controller. *Mathematics of operations Research*, 37(1):154–179, 2012.
- [14] J. Renault and X. Venel. A distance for probability spaces, and long-term values in markov decision processes and repeated games. *Arxiv preprint arXiv:1202.6259*, 2012.
- [15] D. Rosenberg. Zero sum absorbing games with incomplete information on one side: Asymptotic analysis. *SIAM Journal on Control and Optimization*, 39(1):208–225, 2000.
- [16] D. Rosenberg, E. Solan, and N. Vieille. The maxmin value of stochastic games with imperfect monitoring. *International journal of game theory*, 32(1):133–150, 2003.
- [17] D. Rosenberg and N. Vieille. The maxmin of recursive games with incomplete information on one side. *Mathematics of Operations Research*, pages 23–35, 2000.
- [18] L.S. Shapley. Stochastic games. *Proceedings of the National Academy of Sciences of the United States of America*, 39(10):1095, 1953.
- [19] M. Sion. *General Minimax Theorems*. United States Air Force, Office of Scientific Research, 1957.
- [20] S. Sorin. “big match” with lack of information on one side (part i). *International Journal of Game Theory*, 13(4):201–255, 1984.

- [21] S. Sorin. “big match” with lack of information on one side (part ii). *International Journal of Game Theory*, 14(3):173–204, 1985.
- [22] S. Sorin. Operators and repeated games. In *A first course on zero-sum repeated games*, volume 37, chapter C, pages 173–190. Springer, 2002.
- [23] X. Venel. Commutative stochastic games. 2010.
- [24] G. Vigeral. A zero-sum stochastic game with compact action sets and no asymptotic value. *Dynamic Games and Applications*, pages 1–15, 2013.
- [25] Shmuel Zamir. On the notion of value for games with infinitely many stages. *The Annals of Statistics*, 1(4):791–796, 1973.