



**HAL**  
open science

## Markov games with frequent actions and incomplete information

Pierre Cardaliaguet, Catherine Rainer, Dinah Rosenberg, Nicolas Vieille

► **To cite this version:**

Pierre Cardaliaguet, Catherine Rainer, Dinah Rosenberg, Nicolas Vieille. Markov games with frequent actions and incomplete information. 2013. <hal-00843475>

**HAL Id: hal-00843475**

**<https://hal.science/hal-00843475v1>**

Preprint submitted on 11 Jul 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.



HAL Authorization

# Markov games with frequent actions and incomplete information

P. Cardaliaguet\*      Catherine Rainer†      Dinah Rosenberg‡  
Nicolas Vieille§

July 11, 2013

## Abstract

We study a two-player, zero-sum, stochastic game with incomplete information on one side in which the players are allowed to play more and more frequently. The informed player observes the realization of a Markov chain on which the payoffs depend, while the non-informed player only observes his opponent's actions. We show the existence of a limit value as the time span between two consecutive stages vanishes; this value is characterized through an auxiliary optimization problem and as the solution of an Hamilton-Jacobi equation.

**Key-words:** Markov games, incomplete information, zero-sum games, Hamilton-Jacobi equations, repeated games.

**A.M.S. classification :** 91A05, 91A15, 60J10

## 1 Introduction

This paper contributes to the expanding literature on dynamic games with asymmetric information, in which information parameters change with time, see e.g. Athey and Bagwell

---

\*Ceremade, Université Paris-Dauphine, Place du Maréchal de Lattre de Tassigny, 75775 Paris cedex 16 (France). cardaliaguet@ceremade.dauphine.fr.

†Université de Bretagne Occidentale, 6, avenue Victor-le-Gorgeu, B.P. 809, 29285 Brest cedex, France.  
e-mail: Catherine.Rainer@univ-brest.fr

‡Département ESD, HEC Paris. rosenberg@hec.fr

§Département ESD, HEC Paris. vieille@hec.fr

(2008), Mailath and Samuelson (2001), Phelan (2006), Renault (2006), Wiseman (2008), Neyman (2008) and more recently Escobar and Toikka (2013) and Renault, Solan and Vieille (2012). In these papers, payoff-relevant types are private information, and follow Markov processes.

The mathematical analysis of such games for a fixed discount factor  $\delta$  remains beyond reach, and all of the above papers (as well as most of the literature on repeated games, see Mailath and Samuelson (2006)) focus on the limiting case where  $\delta \rightarrow 1$ , with the interpretation that players are "very patient". Yet, another, equally interesting interpretation which is consistent with  $\delta \rightarrow 1$  is that the players get the opportunity to play very frequently. That the two interpretations may lead to sharply contrasted results was first pointed in Abreu, Milgrom and Pearce (1991) for repeated games with imperfect monitoring, see also Fudenberg and Levine (2007) for a recent elaboration on this issue. Recently, this point was also convincingly made in Peski and Wiseman (2012), which analyzes stochastic games with frequent actions, and which contrasts results with those proven in Hörner, Sugaya, Takahashi and Vieille (2011) for stochastic games with patient players.

We here adhere to this alternative interpretation. Our goal is to clear the mathematical problems in the analysis of the benchmark case of two-player, zero-sum games, and we revisit the model of Renault (2006) as follows. The interaction between two players is affected by a payoff-relevant type, which evolves in continuous time according to a Markov process  $(s_t)_{t \geq 0}$ . The two players choose actions at discrete time instants (called stages). Together with the current type, these actions determine the (unobserved) payoff of player 2 to player 1. The realizations of  $(s_t)$  are observed by player 1, but not by player 2, who only observes past action choices of player 1. Players discount future payoffs at a fixed and common discount rate  $r > 0$ , and the time span between two consecutive stages is  $\frac{1}{n}$ .

We prove the existence and provide a characterization of the limit value, as the time span  $\frac{1}{n}$  goes to zero. While our setup is directly inspired from Renault (2006), our analysis is significantly different. In Renault (2006), the transition rates between any two consecutive stages remain constant when players get more patient. (At least when the process of types is irreducible,) the initial private information of a player has a finite lifetime, and the limit value does not depend on the initial distribution. Here instead, transitions rates are of the order of  $\frac{1}{n}$ : as players play more often, the probability that the state changes from one stage to the next vanishes. As a result, the limit value does depend on the initial distribution.

We first analyze the case of exogenous transitions (transition rates do not depend on action choices). Adapting techniques from the literature on repeated games with incomplete information, see Aumann and Maschler (1995), we give a semi-explicit formula of the limit value as the value of an auxiliary optimization problem, which we use to get explicit formulas in a number of cases. Using PDE techniques, we provide an alternative characterization of the limit value as the unique solution (in a weak sense) to a non-standard Hamilton-Jacobi (HJ)

equation. This equation can be understood as the infinitesimal counterpart of a dynamic programming principle.

We next expand significantly this framework to allow first for endogenous transitions (transition rates do depend on actions) and next, for incomplete information on both sides (each player observes and controls his own Markov chain). In both settings we show that the limit value exists and is characterized as the unique (weak) solution of a HJ equation. Our techniques for this analysis (viscosity solutions of Hamilton-Jacobi equations, passage to the limit in these equations) are reminiscent of the ones developed for differential games, as in Evans and Souganidis (1984). However, because of the information asymmetry, the Hamilton-Jacobi equation satisfied by the limit value takes the form of an obstacle problem, much as in the case of differential games with incomplete information (Cardaliaguet and Rainer (2009a), Cardaliaguet (2009)), yet with a significant difference. Indeed, information is here disclosed to the informed player(s) through time (and not only at the initial instant); this leads to a new HJ equation and to a slightly different definition of weak solution (cf. the discussion after Definition 1 and at the beginning of the proof of the comparison principle). The passage from discrete games to continuous equations partially relies on methods developed for repeated games in Vieille (1992), Laraki (2002), Cardaliaguet, Laraki and Sorin (2012), and the starting point of our analysis for incomplete information on both sides is inspired by Gensbittel and Renault (2012).

The paper is organized as follows. We first present the model and state the main results (Section 2), which we illustrate through several examples in Section 3 and which we prove in Section 4. Games with endogenous transitions are analyzed in Section 5, while Section 6 is devoted to games with incomplete information on both sides. In the appendix we collect the proofs of several technical facts, including a new comparison principle adapted to our framework.

## 2 Model and main result

### 2.1 Model

We start with the simpler version of the model. There is a finite set of states  $S$  of cardinal  $|S|$ . With each state is associated a zero-sum game with finite action sets  $A$  and  $B$  and payoff function  $g(s, \cdot, \cdot)$ , where  $g : S \times A \times B \rightarrow \mathbf{R}$ . Time is continuous, and the state  $s_t$  at time  $t \geq 0$  follows a Markov chain with law  $\mathbf{P}$ , initial distribution  $p \in \Delta(S)$  and generator  $R = (\rho_{ss'})_{s,s' \in S}$ . For  $s \neq s'$ ,  $\rho_{ss'}$  is thus the rate of transitions from  $s$  to  $s'$ , while  $-\rho_{ss} = \sum_{s' \neq s} \rho_{ss'}$  is the rate of transitions out of state  $s$ . We denote by  $P(\cdot)$  the transition semi-group of  $(s_t)_{t \geq 0}$ , so that  $P_h(s, s') = \mathbf{P}(s_{t+h} = s' \mid s_t = s)$  for all  $t, h \geq 0$  and  $s, s' \in S$ .

The map  $t \mapsto P_t$  is a solution to the Kolmogorov equation  $P'_t = RP_t$ , and is given by  $P_t = \exp(tR)$ .

Given  $n \in \mathbf{N}^*$ , we let  $G_n(p)$  denote the following, two-player game with infinitely many stages. In each stage  $k \in \mathbf{N}$ , players choose actions  $a_k$  and  $b_k$  in  $A$  and  $B$ , and the payoff is given by  $g(s_k^{(n)}, a_k, b_k)$  where  $s_k^{(n)} := s_{k/n}$  is the state at time  $\frac{k}{n}$ . Along the play, player 1 observes past and current realizations of the states  $s_k^{(n)}$  and both players observe past actions of player 1, but payoffs are not observed.<sup>1</sup>

We view  $G_n(p)$  as the discretized version of a continuous-time game, with stage  $k$  of  $G_n(p)$  taking place at physical time  $\frac{k}{n}$ . As  $n$  increases, the time span between two consecutive stages shrinks and the players get the option to play more and more frequently. In physical time, players discount future payoffs at the fixed, positive rate  $r > 0$ . Hence, the weight of stage  $k$  in  $G_n(p)$  is  $\int_{k/n}^{(k+1)/n} re^{-rt} dt = \lambda_n(1 - \lambda_n)^k$ , where  $\lambda_n := 1 - e^{-r/n}$ . Note that  $\lambda_n \rightarrow 0$  as  $n \rightarrow +\infty$  (and  $1 - \lambda_n$  may be interpreted as the discount factor between two consecutive stages in  $G_n(p)$ ).

We denote by  $\tilde{v}_n(p)$  the value of the game  $G_n(p)$ .<sup>2</sup> From the perspective of the literature on repeated games, the game  $G_n(p)$  is thus a discounted game, with discount factor  $1 - \lambda_n$ .

## 2.2 Results

Our main result is the existence of  $\lim_{n \rightarrow +\infty} \tilde{v}_n(p)$ , together with different characterizations of the limit. We need a few definitions.

Define  $\mathcal{S}(p)$  to be the set of adapted, càdlàg processes  $(p_t)_{t \geq 0}$ , defined on some filtered probability space  $(\Omega, \mathcal{F}, \mathbf{P}, (\mathcal{F}_t)_{t \geq 0})$ , with values in  $\Delta(S)$ , and such that, for each  $t, h \geq 0$ , one has

$$\mathbf{E}[p_{t+h} \mid \mathcal{F}_t] = {}^T P(h)p_t, \quad \mathbf{P} - \text{a.s.} \quad (1)$$

Given  $\tilde{p} \in \Delta(S)$ , we denote by  $u(\tilde{p})$  the value of the one-shot, zero-sum game  $\Gamma(\tilde{p})$  with action sets  $A$  and  $B$  and payoff function

$$g(\tilde{p}, a, b) := \sum_{s \in S} \tilde{p}(s)g(s, a, b).$$

That is,  $u(\tilde{p}) = \max_{x \in \Delta(A)} \min_{y \in \Delta(B)} g(\tilde{p}, x, y) = \min_{y \in \Delta(B)} \max_{x \in \Delta(A)} g(\tilde{p}, x, y)$ .

**Theorem 1.** *The sequence  $(\tilde{v}_n(\cdot))_{n \in \mathbf{N}}$  converges uniformly, and **P1** and **P2** hold, with  $v(p) = \lim_{n \rightarrow \infty} \tilde{v}_n(p)$ .*

---

<sup>1</sup>Whether or not actions of player 2 are observed is irrelevant.

<sup>2</sup>We abstain from using the notation  $v_n(p)$ , which is associated with games with  $n$  stages.

$$\mathbf{P1} \quad v(p) = \max_{(p_t) \in \mathcal{S}(p)} \mathbf{E} \left[ \int_0^{+\infty} r e^{-rt} u(p_t) dt \right].$$

$\mathbf{P2}$   $v(\cdot)$  is the unique viscosity solution of the equation

$$\min \{ rv(p) + H(p, Dv(p)) ; -\lambda_{\max}(p, D^2v(p)) \} = 0 \quad \text{in } \Delta(S), \quad (2)$$

where  $H(p, \xi) = -\langle^T R p, \xi \rangle - ru(p)$ .

Few comments are in order. We first comment on  $\mathbf{P2}$ , and on the intuitive content of equation (2). Assuming  $v(\cdot)$  (extended to a neighborhood of  $\Delta(S)$ ) is a smooth function,  $Dv(p)$  and  $D^2v(p)$  stand respectively for the gradient and Hessian matrix of  $v(\cdot)$  at  $p$ , while, loosely speaking,  $\lambda_{\max}(p, D^2v(p))$  is the maximal eigenvalue of the restriction of  $D^2v(p)$  to the tangent space of  $\Delta(S)$  (all formal definitions will be provided later). According to (2),

- (i)  $-\lambda_{\max}(p, D^2v(p)) \geq 0$ , so that the limit value  $v$  is concave. This concavity property—which actually holds for each  $\tilde{v}_n$  (thanks to the so-called splitting results, such as Propositions 2.2 and 2.3 in Sorin (2002))—can also be established using  $\mathbf{P1}$ .
- (ii) the inequality  $rv(p) + H(p, Dv(p)) \geq 0$  always holds on  $\Delta(S)$ , with equality at any point  $p$  where  $v$  is strictly concave (or, more precisely, at which  $-\lambda_{\max}(p, D^2v(p)) > 0$ ).

It turns out that the Hamilton-Jacobi equation

$$rw(p) + H(p, Dw(p)) = 0, \quad p \in \Delta(S)$$

characterizes the limit value of the auxiliary game in which no player observes  $(s_t)$  – the PDE actually being the infinitesimal version of the dynamic programming equation. In our game, the equality  $rv(p) + H(p, Dv(p)) = 0$  must intuitively therefore hold wherever it is optimal for player 1 not to disclose information. For this reason, the set

$$\mathcal{H} := \{p \in \Delta(S) \mid rv(p) + H(p, Dv(p)) = 0\} \quad (3)$$

is called the *non-revealing set*.

In general however, one cannot hope the limit value  $v(\cdot)$  to be smooth. For this reason, the interpretation of the equation (2) is in the viscosity sense, see Definition 1 in Section 5.

To illustrate  $\mathbf{P1}$ , let us specialize Theorem 1 to the case where  $R$  is identically 0. In such a case, the state  $s_0$  is drawn at time 0 and remains fixed throughout time. The game  $G_n(p)$  thus reduces to a truly repeated game with incomplete information à la Aumann and Maschler (1995). Note that  $\mathcal{S}(p)$  is then equal to the set of càdlàg martingales with values in  $\Delta(S)$  and initial value  $p$ .

Consider two  $\Delta(S)$ -valued processes  $(p_t)_{t \geq 0}$  and  $(\tilde{p}_\tau)_{\tau \in [0,1]}$ , such that  $\tilde{p}_\tau = p_{\frac{\ln(1-\tau)}{r}}$  or equivalently  $p_t = \tilde{p}_{1-e^{-rt}}$  a.s., for each  $t \geq 0$  and  $\tau \in [0, 1]$ . Observe that  $(p_t)$  is a martingale iff  $(\tilde{p}_\tau)_{\tau \in [0,1]}$  is a martingale, and

$$\mathbf{E} \left[ \int_0^\infty r e^{-rt} u(p_t) dt \right] = \mathbf{E} \left[ \int_0^1 u(\tilde{p}_\tau) d\tau \right].$$

Therefore, denoting by  $\mathcal{M}_{[0,1]}(p)$  the set of càdlàg martingales defined over  $[0, 1]$ , with values in  $\Delta(S)$  and starting from  $p$ , one has

$$v(p) = \max_{(\tilde{p}_\tau) \in \mathcal{M}_{[0,1]}(p)} \mathbf{E} \left[ \int_0^1 u(\tilde{p}_\tau) d\tau \right], \quad (4)$$

a well-known formula for repeated games (see Section 3.7.2 in Sorin (2002)). In a sense, the assertion **P1** thus provides the appropriate generalization of (4) to the case of an arbitrary transition rate matrix  $R$ .

**P1** and **P2** provide two alternative (and independent) characterizations of the limit value, as the value of an auxiliary optimization problem, and as a solution to a Hamilton-Jacobi PDE. We next state a verification theorem, which relies on **P2** to give a sufficient condition under which a process in  $\mathcal{S}(p)$  is optimal in **P1**.

**Theorem 2.** *Assume that  $v$  is of class  $C^2$  in a neighborhood of  $\Delta(S)$ . Let  $p \in \Delta(S)$  and  $(p_t) \in \mathcal{S}(p)$  be given, and assume that (i) and (ii) below hold.*

(i) **P**-a.s., one has  $p_s \in \mathcal{H}$  and  $v(p_s) - v(p_{s-}) = \langle Dv(p_{s-}), p_s - p_{s-} \rangle$  for all  $s \geq 0$ ,

(ii)  $(p_t)$  has no continuous martingale part.

Then  $(p_t)$  achieves the maximum in **P1**.

**Remark 3.** As  $R$  is the generator of the transition semi-group  $P$ , (1) implies that each process  $(p_t) \in \mathcal{S}(p)$  can be decomposed, **P**-a.s. for all  $t \geq 0$ , as

$$p_t = p + \int_0^t {}^T R p_s ds + m_t, \quad (5)$$

where  $(m_t)$  is a martingale in the filtration generated by  $(p_t)$ . This martingale itself can be decomposed into a continuous and a purely discontinuous part (see Protter (2005)).

The most important condition in Theorem 2 is (i), which states that the “information process”  $(p_t)$  must live in the non-revealing set  $\mathcal{H}$  and can jump only on the flat parts of the graph of the limit value  $v(\cdot)$ ; this condition is known to be sufficient in a class of simpler games, such as in Cardaliaguet and Rainer (2009b). Condition (ii) is often satisfied in practice, as the examples in the next section show.

*Proof.* We write the Itô formula for  $e^{-rt}v(p_t)$ , using the decomposition (5):

$$e^{-rt}v(p_t) = v(p) - \int_0^t r e^{-rs} v(p_s) ds + \int_0^t e^{-rs} Dv(p_{s-}) dp_s + \frac{1}{2} \int_0^t e^{-rs} D^2v(p_s) d\langle m^c \rangle_s + \sum_{0 < s \leq t} e^{-rs} (v(p_s) - v(p_{s-}) - \langle Dv(p_{s-}), p_s - p_{s-} \rangle),$$

where  $(m_t^c)$  is the continuous part of the martingale  $(m_t)$  and  $\langle m^c \rangle$  its quadratic variation. Under the assumptions (i) and (ii), the two last terms in this equation vanish. Then, replacing  $(p_t)$  by its martingale decomposition and taking expectations on both sides, we get

$$e^{-rt}\mathbf{E}[v(p_t)] = v(p) - \mathbf{E} \left[ \int_0^t e^{-rt} (rv(p_s) - \langle {}^T R p_s, Dv(p_s) \rangle) ds \right].$$

It is time now to apply the assumption  $p_s \in \mathcal{H}$ , which leads to

$$e^{-rt}\mathbf{E}[v(p_t)] = v(p) - \mathbf{E} \left[ \int_0^t r e^{-rt} u(p_s) ds \right].$$

The result follows when letting  $t \rightarrow +\infty$ . □

### 3 Examples and Applications

We here illustrate how **P1** can be used to provide explicit formulas for  $v(p)$  in various cases. This section is organized as follows. We first provide in Lemmas 1 and 2 respectively upper and lower bounds on  $v(\cdot)$  which always hold. We next identify several cases where these bounds coincide, thereby pinning down  $v(\cdot)$ . We finally discuss two examples in more detail.

#### 3.1 Upper and lower bounds for the limit value

Let  $(p_t^*)$  be defined by  $p_t^* = {}^T P_t p$ . The process  $(p_t^*)$  is the unique deterministic process in  $\mathcal{S}(p)$ . It is the process of beliefs held by player 2 when player 1 plays in a non-revealing manner (that is, ignores his own private information) or equivalently, the beliefs of an outside observer who would not observe the informed player's actions. Observe that  $\mathbf{E}[p_t] = p_t^*$  for every  $t \geq 0$  and  $(p_t) \in \mathcal{S}(p)$ .

We denote by  $\text{cav } u := \Delta(S) \rightarrow \mathbf{R}$  the concavification of  $u(\cdot)$ .

**Lemma 1.** *One has*

$$v(p) \leq \int_0^\infty r e^{-rt} \text{cav } u(p_t^*) dt.$$

*Proof.* For any  $(p_t) \in \mathcal{S}(p)$ , one has

$$\mathbf{E} \int_0^\infty r e^{-rt} u(p_t) dt = \int_0^\infty r e^{-rt} \mathbf{E}[u(p_t)] dt \leq \int_0^\infty r e^{-rt} \text{cav } u(\mathbf{E}[p_t]) dt = \int_0^\infty r e^{-rt} \text{cav } u(p_t^*) dt,$$

where the first equality follows from Fubini Theorem, and the first inequality follows from the inequality  $u \leq \text{cav } u$  and from Jensen inequality.  $\square$

For  $s \in S$ , we denote by  $\delta_s \in \Delta(S)$  the probability measure which assigns probability one to  $s$ .

**Lemma 2.** *One has*

$$v(p) \geq \int_0^\infty r e^{-rt} u(p_t^*) dt,$$

and

$$v(p) \geq \sum_{s \in S} u(\delta_s) \int_0^\infty r e^{-rt} p_t^*(s) dt.$$

*Proof.* These lower bounds for  $v(p)$  are obtained when computing  $\mathbf{E} \int_0^\infty r e^{-rt} u(p_t) dt$  for specific processes  $(p_t) \in \mathcal{S}(p)$ .

The first lower bound is obtained when setting  $p_t := p_t^*$ . Intuitively, the right-hand side is then the amount which is secured by the strategy which plays at each  $t$  an optimal (non-revealing) strategy in the average game associated with the current belief of player 2.

The second lower bound obtains when setting  $p_t := \delta_{s_t}$ . Indeed, one then has

$$\begin{aligned} \mathbf{E} \int_0^\infty r e^{-rt} u(p_t) dt &= \int_0^\infty r e^{-rt} \mathbf{E}[u(\delta_{s_t})] dt \\ &= \sum_{s \in S} u(\delta_s) \int_0^\infty r e^{-rt} \mathbf{P}(s_t = s) dt \\ &= \sum_{s \in S} u(\delta_s) \int_0^\infty r e^{-rt} p_t^*(s) dt. \end{aligned}$$

Intuitively, the right-hand side is then the amount which is secured by a strategy which would announce at each  $t$  the current state, and then play optimally in the corresponding game.  $\square$

**Corollary 4.** *If  $u$  is concave, then  $v(p) = \int_0^\infty r e^{-rt} u(p_t^*) dt$ .*

*If  $u$  is convex, then  $v(p) = \sum_{s \in S} u(\delta_s) \int_0^\infty r e^{-rt} p_t^*(s) dt$ .*

*Proof.* If  $u$  is concave, then  $u = \text{cav } u$ , and the result follows from Lemmas 1 and 2 (first lower bound).

If  $u$  is convex, then  $\text{cav } u(p) = \sum_{s \in S} p(s)u(\delta_s)$  for each  $p \in \Delta(S)$ , and the result again follows from Lemmas 1 and 2 (second lower bound).  $\square$

We now illustrate in these two simple cases the alternative characterization **P2** in Theorem 1. If  $u$  is smooth and concave, then the map

$$w(p) := \int_0^\infty r e^{-rt} u(p_t^*) dt = \int_0^\infty r e^{-rt} u({}^T P_t p) dt$$

is concave and satisfies

$$r w(p) + H(p, Dw(p)) = r w(p) - \langle {}^T R p, Dw(p) \rangle - r u(p) = 0.$$

Therefore  $w$  is a solution to the Hamilton-Jacobi equation

$$\min \{ r v(p) + H(p, Dv(p)) ; -\lambda_{\max}(p, D^2 v(p)) \} = 0 \quad \text{in } \Delta(S). \quad (6)$$

By **P2**, this shows (again) that  $v = w$ . Recalling the definition of the non-revealing set  $\mathcal{H}$  in (3), we here have  $\mathcal{H} = \Delta(S)$ . As the deterministic process  $(p_t^*)$  satisfies conditions (i) and (ii) of Theorem 2, it is optimal in **P1**: in other words, player 1 does not reveal anything.

Assume instead that  $u$  is smooth and convex. Then the map

$$w(p) := \sum_{s \in S} u(\delta_s) \int_0^\infty r e^{-rt} p_t^*(s) dt$$

satisfies

$$r w(p) + H(p, Dw(p)) = r \sum_s p_s u(\delta_s) - r u(p) \geq 0,$$

because  $u$  is convex. As  $D^2 w(p) = 0$ ,  $w$  solves (6) (actually one has to be more cautious here and to use the notion of viscosity solution of Definition 1). From this and **P2**, it follows that  $v = w$ . Moreover, the non-revealing set is given by  $\mathcal{H} = \{p : u(p) = \text{cav}(u)(p)\}$  and thus contains  $S$ . Then the process  $(p_t := \delta_{s_t})$  satisfies conditions (i) and (ii) of Theorem 2, so that it is optimal in **P1**: player 1 reveals all his information.

In both the concave and the convex cases,  $v(p)$  is given by  $\int_0^\infty r e^{-rt} \text{cav } u(p_t^*) dt$ . We show that this latter formula is valid in many cases beyond the concave and convex case, but not always.

### 3.2 Two-state games

In the following, we focus on the case where  $S := \{s_1, s_2\}$  contains only two states, and we identify a probability measure over  $S$  with the probability assigned to state  $s_1$ . In particular,  $u$  will be viewed as a function defined over  $[0, 1]$ . We denote by  $p_\infty^* := \lim_{t \rightarrow \infty} p_t^* \in [0, 1]$  the unique invariant measure of  $(s_t)$ , and let  $\underline{p}, \bar{p} \in [0, 1]$  be such that  $\underline{p} \leq p_\infty^* \leq \bar{p}$ , and  $(p_\infty^*, \text{cav } u(p_\infty^*)) = \alpha(\underline{p}, u(\underline{p})) + (1 - \alpha)(\bar{p}, u(\bar{p}))$ , for some  $\alpha \in [0, 1]$ .

Such distributions  $\underline{p}$  and  $\bar{p}$  always exist, but need not to be uniquely defined.

**Lemma 3.** *Assume  $\underline{p} < p_\infty^* < \bar{p}$ . One has*

$$v(p) = \int_0^\infty r e^{-rt} \text{cav } u(p_t^*) dt \text{ for each } p \in [\underline{p}, \bar{p}]. \quad (7)$$

*If moreover the equality  $u = \text{cav } u$  holds on  $[0, \underline{p}]$  (respectively on  $[\bar{p}, 1]$ ), then (7) holds on the interval  $[0, \underline{p}]$  (respectively on  $[\bar{p}, 1]$ ).*

If we further assume that  $\text{cav } u$  is of class  $\mathcal{C}^1$  with  $\text{cav } u(p) > u(p)$  in  $(\underline{p}, \bar{p})$ , then one can easily check that the non-revealing set defined in (3) satisfies  $\{\underline{p}, \bar{p}\} \subset \mathcal{H} \subset [0, \underline{p}] \cup [\bar{p}, 1]$ . In particular, (7) follows from the construction—in the proof below—of a process  $(p_t) \in \mathcal{S}(p)$  such that  $p_t \in \{\underline{p}, \bar{p}\}$  a.s.: this process satisfies (i) and (ii) of Theorem 2 (because  $v$  is affine on  $[\underline{p}, \bar{p}]$ ) and therefore is optimal in **P1**. If, moreover, the equality  $u = \text{cav } u$  holds in  $[0, \underline{p}]$ , then one has  $[0, \underline{p}] \cup \{\bar{p}\} \subset \mathcal{H}$ . We show in the proof below that there is a process  $(p_t) \in \mathcal{S}(p)$  with  $p_t \in [0, \underline{p}] \cup \{\bar{p}\}$  a.s.: the same arguments as above show that this process is optimal.

*Proof.* Define  $\theta := \inf\{t : p_t^* \in [\underline{p}, \bar{p}]\}$ , the first time at which the “average” belief enters the interval  $[\underline{p}, \bar{p}]$ . Note that  $\theta < +\infty$  and that  $p_t^* \in [\underline{p}, \bar{p}]$  for every  $t \geq \theta$ .

The result follows from the fact, proven below, that there is a process  $(p_t) \in \mathcal{S}(p)$  such that  $p_t = p_t^*$  for  $t \leq \theta$ , and  $p_t \in \{\underline{p}, \bar{p}\}$  for  $t \geq \theta$ , **P**-a.s. Indeed, for any such process, one has

$$\begin{aligned} \mathbf{E} \int_0^\infty r e^{-rt} u(p_t) dt &= \int_0^\theta r e^{-rt} u(p_t^*) dt + \int_\theta^\infty r e^{-rt} \mathbf{E}[u(p_t)] dt \\ &= \int_0^\theta r e^{-rt} \text{cav } u(p_t^*) dt + \int_\theta^\infty r e^{-rt} \text{cav } u(p_t^*) dt, \end{aligned}$$

which will conclude the proof of the lemma.

We now construct the process  $(p_t)$ . For  $t \geq \theta$ , define  $Q_t := \begin{pmatrix} q_{11}(t) & 1 - q_{11}(t) \\ 1 - q_{22}(t) & q_{22}(t) \end{pmatrix}$  by

$$q_{11}(t) = \frac{1}{\bar{p} - \underline{p}} (\bar{p} - (\underline{p} \times p_{11}(t) + (1 - \underline{p})p_{21}(t)))$$

and

$$q_{22}(t) = \frac{1}{\bar{p} - \underline{p}} (1 - \bar{p} - ((1 - \bar{p})p_{22}(t) + \bar{p} \times p_{12}(t))).$$

Intuitively,  $Q_t$  is the transition matrix between the two "states"  $\underline{p}$  and  $\bar{p}$  induced by  $P_t$ . To see why, observe that, when starting from  $\underline{p}$ , the probability of being in state  $s_1$  at time  $t$  is  $\underline{p}p_{11}(t) + (1 - \underline{p})p_{21}(t)$ , which is equal to  $q_{11}(t)\underline{p} + (1 - q_{11}(t))\bar{p}$ . Similarly, when starting from  $\bar{p}$ , the probability of being in  $s_1$  at time  $t$  is  $\bar{p}p_{11}(t) + (1 - \bar{p})p_{21}(t) = (1 - q_{22}(t))\underline{p} + q_{22}(t)\bar{p}$ . An elementary computation using the Kolmogorov equation  $P'_t = RP_t$  yields  $Q'_t = \tilde{R}Q_t$ , where the rate matrix  $\tilde{R} = \begin{pmatrix} -\tilde{\rho}_{12} & \tilde{\rho}_{12} \\ \tilde{\rho}_{21} & -\tilde{\rho}_{21} \end{pmatrix}$  is given by

$$\tilde{\rho}_{12} = \frac{(1 - \underline{p})\rho_{21} - \underline{p}\rho_{12}}{\bar{p} - \underline{p}} \text{ and } \tilde{\rho}_{21} = \frac{\bar{p}\rho_{12} - (1 - \bar{p})\rho_{21}}{\bar{p} - \underline{p}}.$$

Both  $\rho_{12}$  and  $\rho_{21}$  are positive. Therefore, there is a Markov process  $(q_t)_{t \geq \theta}$  with values in  $\Delta(\{\underline{p}, \bar{p}\})$ , with rate matrix  $\tilde{R}$  and initial distribution  $q_\theta$  defined by  $p_\theta^* = q_\theta(\underline{p})\underline{p} + q_\theta(\bar{p})\bar{p}$ .

By construction, the process  $(p_t)_{t \geq \theta}$  defined by  $p_t := q_t(\underline{p})\underline{p} + q_t(\bar{p})\bar{p}$  is a Markov process and  $\mathbf{E}[p_{t+h} | \mathcal{F}_t^p] =^T P_h p_t$  for all  $t \geq \theta, h \geq 0$ . Set now  $p_t = p_t^*$  for  $t < \theta$ . Then the process  $(p_t)_{t \geq 0}$  satisfies the desired properties.  $\square$

In all previous cases, the equality  $v(p) = \int_0^\infty re^{-rt} \text{cav } u(p_t^*) dt$  holds. This is however not always the case, as we now show.

**Example:** Let a game  $A, B$  and  $g : \{s_1, s_2\} \times A \times B$  be such that (i)  $u(0) = u(1) = 0$ , (ii)  $u(p) = 1$  for  $p \in [\frac{1}{3}, \frac{2}{3}]$ , and (iii)  $u$  is strictly convex on each of the intervals  $[0, \frac{1}{3}]$  and  $[\frac{2}{3}, 1]$ . Assume that transitions are such that  $p_\infty^* \in (\frac{1}{3}, \frac{2}{3})$ .<sup>3</sup>

**Proposition 1.** *For every  $p \notin [\frac{1}{3}, \frac{2}{3}]$ , one has  $v(p) < \int_0^\infty re^{-rt} \text{cav } u(p_t^*) dt$ .*

*Proof.* Fix  $p \in [0, \frac{1}{3})$  for concreteness. We argue by contradiction, and assume that

$$\int_0^\infty re^{-rt} \mathbf{E}[u(p_t)] dt = \int_0^\infty re^{-rt} \text{cav } u(p_t^*) dt,$$

for some process  $(p_t) \in \mathcal{S}(p)$ . Since  $\mathbf{E}[u(p_t)] \leq \text{cav } u(p_t^*)$  for all  $t$ , one has  $\mathbf{E}[u(p_t)] = \text{cav } u(p_t^*)$  for Leb-a.e.  $t \in \mathbf{R}^*$ .

---

<sup>3</sup>The existence of such a game follows from Proposition 6 in Lehrer and Rosenberg (2003), which does not appear in the published version of the paper, Lehrer and Rosenberg (2010).

Let  $\theta := \inf\{t : p_t^* \geq \frac{1}{3}\}$ . Observe that  $\text{cav } u(p_t^*) = 1$  for  $t \geq \theta$ , and  $\text{cav } u(p_t^*) = 3p_t^*$  for  $t \leq \theta$ .

For  $t < \theta$ , the equality  $\mathbf{E}[u(p_t)] = \text{cav } u(p_t^*)$  implies that the law of  $p_t$  is concentrated on  $\{0, \frac{1}{3}\}$ , with  $\mathbf{P}(p_t = \frac{1}{3}) = 3p_t^*$ . As  $t \rightarrow \theta^-$  we get  $\mathbf{P}(p_{\theta^-} = \frac{1}{3}) = 1$ , so that  $p_{\theta^-} = p_\theta^*$  a.s.. Then (1) implies that  $p_t = (e^{-tR(\theta-t)})p_\theta^* = p_t^*$  a.s. for  $t \in [0, \theta]$ , which is impossible since  $p_t \in \{0, \frac{1}{3}\}$  a.s..

□

Intuitively, maximizing  $\mathbf{E}[u(p_t)]$  leads player 1 to disclose information at time  $t$  which he later wishes he hadn't disclosed.

### 3.3 An explicit example

We conclude this section by providing an explicit formula for the limit value in an example due to Renault (2006) (see also Hörner, Rosenberg, Solan and Vieille (2010)). In that example, both players have two actions, and the payoffs in the two states are given by

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \text{ and } \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

Transitions occur at the rate  $\pi > 0$ , so that  $R = \begin{pmatrix} -\pi & \pi \\ \pi & -\pi \end{pmatrix}$ . Observe that  $R = M \begin{pmatrix} -2\pi & 0 \\ 0 & 0 \end{pmatrix} M^{-1}$ , where  $M = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$ , so that

$$P_t = e^{tR} = M \begin{pmatrix} e^{-2\pi t} & 0 \\ 0 & 1 \end{pmatrix} M^{-1} = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} + \frac{e^{-2t\pi}}{2} \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}.$$

Note that  $u(p) = p(1-p)$  is concave, hence

$$v(p) = \int_0^\infty r e^{-rt} u(p_t^*) dt = \int_0^\infty r e^{-rt} p_t^* (1-p_t^*) dt. \quad (8)$$

On the other hand,  $p_t^*$  is given by  $\begin{pmatrix} 1-p_t^* \\ p_t^* \end{pmatrix} = P(t) \begin{pmatrix} 1-p \\ p \end{pmatrix}$ . Integration in (8) leads to

$$v(p) = \frac{1}{4} - \frac{(2p-1)^2}{4} \times \frac{r}{r+4\pi}.$$

## 4 Proof of Theorem 1

In this section, we prove **P1**. Statement **P2** is a particular case of Theorem 7 below, and we postpone the proof to section 5. The proof of **P1** is divided in three parts. We first prove that

$$\liminf_n \tilde{v}_n(p) \geq \sup_{\mathcal{S}(p)} \mathbf{E} \left[ \int_0^\infty r e^{-rt} u(p_t) dt \right],$$

and next that

$$\limsup_n \tilde{v}_n(p) \leq \sup_{\mathcal{S}(p)} \mathbf{E} \left[ \int_0^\infty r e^{-rt} u(p_t) dt \right].$$

We finally show that the supremum is reached.

### 4.1 Step 1

Let  $(p_t) \in \mathcal{S}(p)$  be arbitrary. We will prove that  $\liminf_n \tilde{v}_n(p) \geq \mathbf{E} \left[ \int_0^\infty r e^{-rt} u(p_t) dt \right]$ . The proof will make use of Lemma 5 below. This lemma is conceptually similar to (but technically more involved than) the elementary, so-called splitting lemma (Aumann and Maschler (1995)) which we quote here.

**Lemma 4.** *Let a finite set  $L$ , and a probability  $p \in \Delta(S)$  be given, such that  $p = \sum_{l \in L} \alpha_l p_l$ , for some  $\alpha \in \Delta(L)$  and  $p_l \in \Delta(S)$  ( $l \in L$ ). Then there is a probability distribution  $\mathbf{P}$  over  $L \times S$  with marginals given by  $\alpha$  and  $p$ , and such that the conditional law of  $s$  given  $l$  is  $p_l$ .*

The usual interpretation of Lemma 4 is as follows. Assume some player, informed of the realization of  $s$ , draws  $l$  according to  $\mathbf{P}(l | s)$  and announces  $l$ . Then, the posterior belief of an uninformed player with prior belief  $p$  is equal to  $p_l$ . Lemma 4 formalizes the extent to which an informed player can “manipulate” the belief of an uninformed player by means of a public announcement.

Lemma 5 below is the appropriate generalization of Lemma 4 to a dynamic world with changing states. Some notation is required. We fix a Markov chain  $(\omega_m)_{m \in \mathbf{N}}$  over  $S$ , with initial law  $p \in \Delta(S)$ , transition matrix  $\Pi = (\pi(s' | s))_{s, s' \in S}$ , and law  $\mathbf{P}$ . Given a sequence  $\mu = (\mu_m)_{m \in \mathbf{N}}$ , where  $\mu_m$  is a transition function from  $(\Delta(S))^m \times S$  to  $\Delta(S)$ ,<sup>4</sup> we denote by  $\mu \circ \mathbf{P}$  the probability measure over  $(\Delta(S) \times S)^{\mathbf{N}}$  which is obtained as follows. Together with  $\Pi$ ,  $\mu_m$  induces a transition function  $\nu_m$  from  $\Delta(S)^m \times S$  to  $\Delta(S) \times S$  defined by

$$\nu_m(q^m, \omega_m; q_{m+1}, \omega_{m+1}) = \pi(\omega_{m+1} | \omega_m) \mu_m(q^m, \omega_{m+1}; q_{m+1}). \quad (9)$$

---

<sup>4</sup>We write  $\mu_m(q^m, s; \cdot)$ . Thus,  $\mu_m(q^m, s; \cdot)$  is a probability distribution over  $\Delta(S)$  for each given  $q^m = (q_0, \dots, q_m) \in \Delta(S)^m$ , and  $s \in S$ , and the probability  $\mu_m(q^m, s; A)$  assigned to a fixed (measurable) set  $A \subset \Delta(S)$  is measurable in  $(q^m, s)$ .

The distribution  $\mu \circ \mathbf{P}$  is the probability measure over  $(\Delta(S) \times S)^{\mathbf{N}}$  induced by the sequence  $(\nu_m)_{m \in \mathbf{N}}$  (by means of the Ionescu-Tulcea Theorem) and the initial distribution of  $(q_0, \omega_0)$  which assigns probability  $p(\omega_0)$  to  $(\{p\}, \omega_0) \in \Delta(S) \times S$ .

To follow-up on the above interpretation, we think of an uninformed player with belief  $p$  over  $\omega_0$ , and of an informed player who observes the successive realizations of  $(\omega_m)$ , and picks a new belief  $q_{m+1} \in \Delta(S)$  for the uninformed player, as a (random) function  $\mu_m$  of the earlier beliefs  $q^m = (q_0, \dots, q_m)$  and of the realized state  $\omega_{m+1}$  in stage  $m + 1$ . The distribution  $\mu \circ \mathbf{P}$  is the induced distribution over sequences of beliefs and states.

**Lemma 5.** *Let  $Q$  be a probability distribution over  $\Delta(S)^{\mathbf{N}}$  such that  $Q$ -a.s.,  $q_0 = q$  and that  $\mathbf{E}[q_{m+1} \mid q^m] =^T \Pi q_m$  for each  $m$ .*

*Then there exists a sequence  $\mu = (\mu_m)$  such that the probability measure  $\mu \circ \mathbf{P}$  satisfies **C1** and **C2** below.*

**C1** *The marginal of  $\mu \circ \mathbf{P}$  over  $\Delta(S)^{\mathbf{N}}$  is  $Q$ .*

**C2** *For each  $m \geq 0$ ,  $q_m$  is (a version of) the conditional law of  $\omega_m$  given  $q^m$ .*

The proof of Lemma 5 is in the Appendix. We now construct a behavior strategy  $\sigma_1$  of player 1 in  $G_n(p)$ . We let  $\alpha : \Delta(S) \rightarrow \Delta(A)$  be a (measurable) function such that  $\alpha(\tilde{p})$  is an optimal strategy of player 1 in the one-shot, average game  $\Gamma(\tilde{p})$ , for each  $\tilde{p} \in \Delta(S)$ .

For  $k \geq 0$ , we set  $s_k^{(n)} := s_{k/n}$  and  $p_k^{(n)} := p_{k/n}$ . The sequence  $(s_k^{(n)})_{k \in \mathbf{N}}$  is a Markov chain with transition function  $P_{1/n}$  and initial distribution  $p$ . We let  $(\mu_k)_{k \in \mathbf{N}}$  be the transition functions obtained by applying Lemma 5 with  $\omega_k := s_k^{(n)}$  and  $q_k := p_k^{(n)}$ .

According to  $\sigma_1$ , player 1 picks  $p_{k+1}^{(n)} \in \Delta(S)$  according to  $\mu_k(p_0^{(n)}, \dots, p_k^{(n)}, s_{k+1}^{(n)}; \cdot)$  then plays the mixed action  $\alpha(p_{k+1}^{(n)}) \in \Delta(A)$ .

Let  $\sigma_2$  be an arbitrary strategy of player 2 in  $G_n(p)$ . For any given stage  $k$ , one has

$$\begin{aligned} \mathbf{E}[g(s_k^{(n)}, a_k, b_k)] &= \mathbf{E} \left[ \mathbf{E} \left[ g(s_k^{(n)}, a_k, b_k) \mid p_0^{(n)}, \dots, p_k^{(n)} \right] \right] \\ &= \mathbf{E} \left[ \mathbf{E} \left[ g(s_k^{(n)}, \alpha(p_k^{(n)}), b_k) \mid p_0^{(n)}, \dots, p_k^{(n)} \right] \right] \\ &= \mathbf{E} \left[ \mathbf{E} \left[ g(p_k^{(n)}, \alpha(p_k^{(n)}), b_k) \mid p_0^{(n)}, \dots, p_k^{(n)} \right] \right] \\ &\geq \mathbf{E} \left[ \mathbf{E} \left[ u(p_k^{(n)}) \mid p_0^{(n)}, \dots, p_k^{(n)} \right] \right] = \mathbf{E}[u(p_k^{(n)})]. \end{aligned}$$

Summing over  $k$ , and denoting by  $t \mapsto p_t^{(n)}$  the step process equal to  $p_{\frac{k}{n}}$  over the interval  $[\frac{k}{n}, \frac{k+1}{n})$ , one has therefore

$$\tilde{v}_n(p) \geq \mathbf{E} \left[ \int_0^\infty r e^{-rt} u(p_t^{(n)}) dt \right]. \quad (10)$$

Since  $(p_t)$  is càdlàg, the map  $t \mapsto p_t$  has  $\mathbf{P}$ -a.s. at most countably many discontinuity points. Note also that  $\lim_n p_t^{(n)} = p_t$  at every continuity point. Thus, one has  $\lim_n p_t^{(n)} = p_t$   $\mathbf{P} \otimes \text{Leb}$ -a.s. . By dominated convergence, this implies

$$\lim_{n \rightarrow +\infty} \mathbf{E} \int_0^\infty r e^{-rt} u(p_t^{(n)}) dt = \mathbf{E} \int_0^\infty r e^{-rt} u(p_t) dt.$$

By (10), one thus has

$$\liminf_{n \rightarrow +\infty} \tilde{v}_n(p) \geq \mathbf{E} \int_0^\infty e^{-rt} u(p_t) dt.$$

## 4.2 Step 2

Let  $n \in \mathbf{N}^*$ , and  $\sigma_1$  be an arbitrary strategy of player 1. We adapt Aumann and Maschler (1995), and construct a reply  $\sigma_2$  of player 2 recursively. Together with  $\sigma_1$ , the strategy  $\sigma_2$  induces a probability distribution over plays of the game, denoted  $\mathbf{P}_\sigma$ . Given a stage  $k$ , we denote by  $p_k := \mathbf{P}_\sigma(s_{k/n} = \cdot \mid \mathcal{H}_k^{II})$  the belief of player 2 at the beginning of stage  $k$ , where  $\mathcal{H}_k^{II}$  is the information available to player 2, that is, the  $\sigma$ -algebra generated by  $(a_i, b_i)_{i=1, \dots, k-1}$ .<sup>5</sup> We let  $\sigma_2$  play in stage  $k$  a best reply in the average, one-shot game  $\Gamma(p_k)$  to the conditional distribution of  $a_k$  given  $\mathcal{H}_k^{II}$ .

We introduce the belief  $\tilde{p}_k := \mathbf{P}(s_{k/n} = \cdot \mid \mathcal{H}_{k+1}^{II})$  held by player 2 at the end of stage  $k$  (that is, after observing  $a_k$ ), so that  $p_{k+1} = {}^T P_{1/n} \tilde{p}_k$ .

By Lemmas 2.5 and 2.6 in Mertens, Sorin and Zamir (1994), one has<sup>6</sup>

$$\mathbf{E}[g(s_{k/n}, a_k, b_k) \mid \mathcal{H}_k^{II}] \leq u(p_k) + \mathbf{E} [|p_k - \tilde{p}_k|_1 \mid \mathcal{H}_k^{II}].$$

Taking expectations and summing over stages, one obtains

$$\mathbf{E} \left[ \lambda_n \sum_{k=0}^{\infty} (1 - \lambda_n)^k g_k \right] \leq \mathbf{E} \left[ \lambda_n \sum_{k=0}^{\infty} (1 - \lambda_n)^k u(p_k) \right] + \mathbf{E} \left[ \lambda_n \sum_{k=0}^{\infty} (1 - \lambda_n)^k |p_k - \tilde{p}_k|_1 \right] \quad (11)$$

We now introduce a process  $(\bar{p}_t)$  in  $\mathcal{S}(p)$  defined by  $\bar{p}_{k/n} = p_k$  and  $\bar{p}_t = {}^T P_{t-\frac{k}{n}} \bar{p}_{k/n}$  for each  $k$  and  $t \in [\frac{k}{n}, \frac{k+1}{n})$ .

Choose a constant  $c > 0$  such that  $|\bar{p}_t - \bar{p}_{k/n}|_1 \leq \frac{c}{n}$  for each  $k, n \in \mathbf{N}^*$  and  $t \in [\frac{k}{n}, \frac{k+1}{n})$ .

We first bound the first term on the right-hand side of (11):

$$\mathbf{E} \left[ \lambda_n \sum_{k=0}^{\infty} (1 - \lambda_n)^{k-1} u(p_k) \right] \leq \mathbf{E} \left[ \int_0^\infty r e^{-rt} u(\bar{p}_t) dt \right] + \sup_{t \geq 0} |u(\bar{p}_t) - u(\bar{p}_{\frac{1}{n} \lfloor nt \rfloor})|.$$

<sup>5</sup>The belief  $p_k$  is used to define  $\sigma_2$  in stage  $k$ , and the computation of  $p_k$  uses the definition of  $\sigma_2$  in the first  $k-1$  stages only. Hence, there is no circularity.

<sup>6</sup>denoting by  $|\cdot|_1$  the  $L^1$ -norm on  $\Delta(S)$ .

Since  $u$  is Lipschitz for the  $L^1$ -norm, one has, for some  $C$ ,

$$\mathbf{E} \left[ \lambda_n \sum_{k=0}^{\infty} (1 - \lambda_n)^k u(p_k) \right] \leq \sup_{(p_t) \in \mathcal{S}(p)} \mathbf{E} \left[ \int_0^{\infty} r e^{-rt} u(p_t) dt \right] + \frac{C}{n}.$$

Next, adapting Mertens Sorin and Zamir (1994), one has

$$\begin{aligned} \mathbf{E} \left[ \lambda_n \sum_{k=0}^{\infty} (1 - \lambda_n)^k |p_k - \tilde{p}_k|_1 \right] &= \sum_{s \in S} \lambda_n \sum_{k=0}^{\infty} (1 - \lambda_n)^k \mathbf{E} [|p_k(s) - \tilde{p}_k(s)|] \\ &\leq \sum_{s \in S} \left( \lambda_n \sum_{k=0}^{\infty} (1 - \lambda_n)^k \mathbf{E} [|p_k(s) - \tilde{p}_k(s)|^2] \right)^{\frac{1}{2}} \\ &= \sum_{s \in S} \left( \lambda_n \sum_{k=0}^{\infty} (1 - \lambda_n)^k (\mathbf{E}(p_k^2(s)) - \mathbf{E}(\tilde{p}_k^2(s))) \right)^{\frac{1}{2}}, \end{aligned}$$

which is also equal to

$$\sum_{s \in S} \left( \lambda_n \sum_{k=0}^{\infty} (1 - \lambda_n)^k (\mathbf{E}(p_k^2(s)) - \mathbf{E}(p_{k+1}^2(s)) + \mathbf{E}(p_{k+1}^2(s)) - \mathbf{E}(\tilde{p}_k^2(s))) \right)^{\frac{1}{2}}.$$

Therefore

$$\begin{aligned} \mathbf{E} \left[ \lambda_n \sum_{k=0}^{\infty} (1 - \lambda_n)^k |p_k - \tilde{p}_k|_1 \right] &\leq \sum_{s \in S} \left( \lambda_n + \lambda_n \sum_{k=0}^{\infty} (1 - \lambda_n)^k \mathbf{E} (p_{k+1}^2(s) - \tilde{p}_k^2(s)) \right)^{\frac{1}{2}} \\ &\leq |S| \left( \lambda_n + \frac{2c}{n} \right)^{\frac{1}{2}}. \end{aligned}$$

Plugging into (11), and since  $\sigma_1$  is arbitrary, this yields

$$\tilde{v}_n(p) \leq \sup_{(p_t) \in \mathcal{S}(p)} \mathbf{E} \left[ \int_0^{\infty} r e^{-rt} u(p_t) dt \right] + \frac{C}{n} + |S| \left( \lambda_n + \frac{2c}{n} \right)^{\frac{1}{2}},$$

and the inequality  $\limsup_{n \rightarrow \infty} \tilde{v}_n(p) \leq \sup_{(p_t) \in \mathcal{S}(p)} \mathbf{E} \left[ \int_0^{\infty} r e^{-rt} u(p_t) dt \right]$  follows.

### 4.3 Step 3

We conclude by proving that the supremum in **P1** is reached. First we remark that the claim **P1** can alternatively be written as

$$v(p) = \max_{\mathbf{P} \in \Sigma(p)} \mathbf{E}_{\mathbf{P}} \left[ \int_0^{+\infty} r e^{-rt} u(p_t) dt \right],$$

where, if  $\Omega$  denotes the set of càdlàg functions from  $\mathbf{R}^+$  to  $\Delta(S)$  and  $(p_t)$  the canonical process on  $\Omega$ ,  $\Sigma(p)$  is the set of probability measures  $\mathbf{P}$  on  $\Omega$  under which  $(p_t)$  belongs to  $\mathcal{S}(p)$ . This reformulation permits us to use classical arguments: We apply the tightness criterion of Meyer and Zheng (1984). Let  $(\mathbf{P}_n)_n$  be a maximizing sequence for **P1**. Although strictly speaking the coordinate process  $(p_t)$  need not be a quasimartingale, Theorem 4 in Meyer and Zheng (1984) still applies.<sup>7</sup>

Denote by  $\bar{\mathbf{P}}$  the weak limit of some subsequence of  $(\mathbf{P}_n)_n$ . It is routine to show that  $\bar{\mathbf{P}} \in \mathcal{S}(p)$ . Finally, since the functional  $\mathbf{E} \left[ \int_0^\infty e^{-rt} u(p_t) dt \right]$  is weak continuous,  $\bar{\mathbf{P}}$  is a maximizer in **P1**.

## 5 Games with endogenous transitions

In this section we extend Theorem 1 to games with endogenous transitions. We show that the limit value exists and is characterized as the unique viscosity solution of a Hamilton-Jacobi equation.

### 5.1 Model

We now introduce a more general model in which players control transitions. As before,  $S$  is a finite set of states,  $A$  and  $B$  are finite action sets and  $g : S \times A \times B \rightarrow \mathbf{R}$  denotes the payoff function. In contrast with the previous sections, we here assume that the generator depends on actions :  $R := (R(s, s'; a, b), s, s' \in S, a \in A, b \in B)$ , where for all  $(s, a, b) \in S \times A \times B$ ,

- for all  $s' \neq s$ ,  $R(s, s'; a, b) \geq 0$ ,
- $\sum_{s' \in S} R(s, s'; a, b) = 0$ .

For fixed  $(a, b) \in A \times B$ , we denote by  $t \mapsto P_t(a, b)$  the transition semi-group of the Markov chain with transition rates  $R(\cdot; a, b)$ . Given  $t \geq 0$ , and  $x \in \Delta(A)$ ,  $y \in \Delta(B)$ , we set  $P_t(x, y) := \sum_{a \in A, b \in B} x(a)y(b)P_t(a, b)$ .

For  $n \in \mathbf{N}^*$ ,  $G_n(p)$  now denotes the two-player game with infinitely many stages, where, at each stage  $k \in \mathbf{N}$ , players first choose actions  $a_k \in A$  and  $b_k \in B$ , the payoff is  $g(s_k^{(n)}, a_k, b_k)$ , and next  $s_{k+1}^{(n)}$  is drawn using  $P_{1/n}(s_k^{(n)}, \cdot; a_k, b_k)$ .

---

<sup>7</sup>One may e.g. consider the laws  $\mathbf{P}_n^T$  of the coordinate process stopped at  $T$  and then use a diagonal argument. Alternatively, we may apply Theorem 4 directly to the "damped" process  $q_t := e^{-at} p_t$  where  $a \in (0, r)$ , with values in the cone spanned by  $\Delta(S)$ .

The information structure of the game is as before: player 1 observes past and current realizations of the states  $s_k^{(n)}$  and both players observe past actions of their opponent, while payoffs are not observed. As before, the weight of stage  $k$  in  $G_n(p)$  is  $\lambda_n(1 - \lambda_n)^k$ , with  $\lambda_n := 1 - e^{-r/n}$ .

The value of the game  $G_n(p)$  is still denoted  $\tilde{v}_n(p)$ .

## 5.2 Viscosity Solutions

In this section, we introduce the Hamilton-Jacobi equation which characterizes the limit value and we define the notion of weak solution (in the viscosity sense) used in Theorem 1.

We first need to fix some notations. As the partial differential equations encountered below take place in the simplex  $\Delta(S)$ , we have to define a tangent vector space  $T_{\Delta(S)}(p)$  to the set  $\Delta(S)$  at each point  $p$ :

$$T_{\Delta(S)}(p) := \{z = (z_s)_{s \in S} \in \mathbf{R}^{|S|}, \exists \varepsilon > 0, p + \varepsilon z \in \Delta(S), p - \varepsilon z \in \Delta(S)\} .$$

For instance, if  $p$  belongs to the relative interior of  $\Delta(S)$ , one has  $T_{\Delta(S)}(p) = \{z \in \mathbf{R}^{|S|}, \sum_{s \in S} z_s = 0\}$ , while  $T_{\Delta(S)}(p) = \{0\}$  if  $p$  is an extreme point of  $\Delta(S)$ . We also define, for any symmetric matrix  $X \in \mathbf{R}^{|S| \times |S|}$ ,

$$\lambda_{\max}(p, X) := \sup \left\{ \frac{\langle Xz, z \rangle}{|z|^2}, z \in T_{\Delta(S)}(p) \setminus \{0\} \right\} . \quad (12)$$

By convention  $\lambda_{\max}(p, X) = -\infty$  if  $T_{\Delta(S)}(p) = \{0\}$ . Note that  $\lambda_{\max}(p, X)$  is a kind of maximal eigenvalue of  $X$  on the tangent space  $T_{\Delta(S)}(p)$ .

Given a continuous map  $H : \Delta(S) \times \mathbf{R}^{|S|} \rightarrow \mathbf{R}$ , we consider the Hamilton-Jacobi equation

$$\min \{rw(p) + H(p, Dw(p)) ; -\lambda_{\max}(p, D^2w(p))\} = 0 \quad \text{in } \Delta(S). \quad (13)$$

**Definition 1.** A map  $w : \Delta(S) \rightarrow \mathbf{R}$  is a viscosity subsolution of (13) if it is upper semicontinuous and if, for any smooth test function  $\phi : \mathbf{R}^{|S|} \rightarrow \mathbf{R}$  such that  $w - \phi$  has a local maximum on  $\Delta(S)$  at a point  $p \in \Delta(S)$ , one has

$$\min \{rw(p) + H(p, D\phi(p)) ; -\lambda_{\max}(p, D^2\phi(p))\} \leq 0 .$$

A map  $w : \Delta(S) \rightarrow \mathbf{R}$  is a viscosity supersolution of (13) if it is lower semicontinuous and if, for any smooth test function  $\phi : \mathbf{R}^{|S|} \rightarrow \mathbf{R}$  such that  $w - \phi$  has a local minimum on  $\Delta(S)$  at a point  $p \in \Delta(S)$ , one has

$$\min \{rw(p) + H(p, D\phi(p)) ; -\lambda_{\max}(p, D^2\phi(p))\} \geq 0 .$$

Finally,  $w : \Delta(S) \rightarrow \mathbf{R}$  is a viscosity solution of (13) if it is a subsolution and a supersolution of (13).

**Remark 5.** 1) This definition does not exactly match the standard notion of viscosity solution given, e.g., in Crandall, Ishii and Lions (1992): the reason is that we work with functions defined on the simplex  $\Delta(S)$ , instead of the entire space. It is not even quite the same as in recent papers dealing with differential games with incomplete information, see e.g. Cardaliaguet and Rainer (2009a). In these papers, no private information is ever disclosed after the initial time and the “dynamics” on the parameter  $p$  is simply the evolution of the belief of the non-informed player. As a consequence, a key property of these games is that the faces of the simplex are invariant under this dynamics: in terms of PDE, this is expressed by the fact that the conditions for supersolution only need to be tested in the relative interior of  $\Delta(S)$ . In the present framework, the variable  $p$  has a dynamics (the controlled Markov chain), which leaves the entire set  $\Delta(S)$  invariant, but not the faces. As a consequence, the equations have to hold *up to the boundary*, as in the so-called state-constraint problems.

2) In the above definitions, one can always replace the assumption that  $w - \phi$  has a local maximum or minimum by the condition  $w - \phi$  has a strict local maximum or minimum (see, e.g., Crandall et al. (1992)).

Uniqueness for the solution of (25) holds thanks to the following comparison principle, proved in the appendix. We assume that the Hamiltonian  $H : \Delta(S) \times \mathbf{R}^{|S|} \rightarrow \mathbf{R}$  satisfies the condition:

$$|H(p, q) - H(p, q')| \leq C|q - q'| \quad \forall p \in \Delta(S), q, q' \in \mathbf{R}^{|S|} \quad (14)$$

as well as

$$|H(p, q) - H(p', q)| \leq C|p - p'|(|q| + 1) \quad \forall p, p' \in \Delta(S), q \in \mathbf{R}^{|S|} \quad (15)$$

**Theorem 6.** *Under assumptions (14) and (15), if  $w_1$  is a continuous viscosity subsolution of (13) while  $w_2$  is a continuous viscosity supersolution of (13), then  $w_1 \leq w_2$  in  $\Delta(S)$ .*

In particular, equation (13) has at most one continuous viscosity solution.

**Examples:** We have already encountered several examples of subsolution and supersolution for the Hamilton-Jacobi equation (2): for instance, it can be checked that the maps  $w(p) := \int_0^\infty r e^{-rt} \text{cav } u(p_t^*) dt$  in Lemma 1 is a supersolution to (2), while the maps  $w_1(p) := \int_0^\infty r e^{-rt} u(p_t^*) dt$  and  $w_2(p) := \sum_{s \in S} u(\delta_s) \int_0^\infty r e^{-rt} p_t^*(s) dt$  in Lemma 2 are subsolutions of (2). Hence, by **P2** and Theorem 6, one has  $v \leq w$  and  $v \geq \max(w_1, w_2)$ . This provides an alternative proof of Lemmas 1 and 2.

### 5.3 The convergence result

In the endogenous case, it seems difficult to provide a characterization of  $\lim \tilde{v}_n$  of the type of **P1** in Theorem 1. However, characterization **P2** still holds:

**Theorem 7.** *The uniform limit  $\lim_{n \rightarrow +\infty} \tilde{v}_n(p)$  exists and is the unique viscosity solution of the Hamilton-Jacobi equation (13), where  $H$  is now given by*

$$H(p, \xi) = \min_{x \in \Delta(A)} \max_{y \in \Delta(B)} \left\{ -\langle {}^T R(x, y)p, \xi \rangle - rg(p, x, y) \right\}. \quad (16)$$

Note that, when the transition are independent of actions, one recovers statement **P2** of Theorem 1 as a particular case. As the map  $H$  defined by (16) satisfies conditions (14) and (15) above, Theorem 6 applies, and equation (13) has at most a unique viscosity solution.

### 5.4 Proof of Theorem 7

#### 5.4.1 Step 1: dynamic programming principle and regularity

As is well-known, the maps  $\tilde{v}_n$  are (uniformly) Lipschitz on  $\Delta(S)$ , concave and bounded, and satisfy the following dynamic programming principle:

$$\begin{aligned} \tilde{v}_n(p) &= \max_{x \in (\Delta(A))^S} \min_{y \in \Delta(B)} \left( \lambda_n g(p, x, y) + (1 - \lambda_n) \sum_{a \in A, b \in B} x(p)(a)y(b) \tilde{v}_n({}^T P_{\frac{1}{n}}(a, b) \hat{p}(x, a)) \right) \\ &= \min_{y \in \Delta(B)} \max_{x \in (\Delta(A))^S} \left( \lambda_n g(p, x, y) + (1 - \lambda_n) \sum_{a \in A, b \in B} x(p)(a)y(b) \tilde{v}_n({}^T P_{\frac{1}{n}}(a, b) \hat{p}(x, a)) \right) \end{aligned} \quad (17)$$

where  $x(p)(a) = \sum_{s \in S} p_s x_s(a)$  and, for  $a \in A$ ,  $\hat{p}(x, a) := \left( \frac{p_s x_s(a)}{x(p)(a)} \right)_{s \in S}$  is the conditional law of the state given  $a$ :

We will prove that any accumulation point of  $(\tilde{v}_n)$  for the uniform convergence is a viscosity solution of (13). Since (13) has a unique viscosity solution, this will imply the uniform convergence of the sequence  $(\tilde{v}_n)$ .

We thus consider a uniformly convergent subsequence of  $(\tilde{v}_n)$ . We denote by  $w$  the continuous limit, and relabel the subsequence as  $(\tilde{v}_n)$ .

#### 5.4.2 Step 2: $w$ is a viscosity supersolution

Let  $\phi$  be a smooth test function such that  $w - \phi$  has a strict local minimum on  $\Delta(S)$  at some point  $\bar{p} \in \Delta(S)$ . This implies the existence of a sequence  $(p_n)$  which converges to  $\bar{p}$  and such that  $\tilde{v}_n - \phi$  has a local minimum at  $p_n$  for any  $n$ : namely,

$$\tilde{v}_n(p) \geq \phi(p) - \phi(p_n) + \tilde{v}_n(p_n) \quad \text{for any } p \in \Delta(S). \quad (18)$$

As  $\tilde{v}_n$  is concave, the inequality  $\lambda_{\max}(p_n, D^2\phi(p_n)) \leq 0$  holds by definition of  $T_{\Delta(S)}(\bar{p})$ . Since  $T_{\Delta(S)}(p_n) \supset T_{\Delta(S)}(\bar{p})$  for  $n$  large enough, letting  $n \rightarrow +\infty$  yields  $-\lambda_{\max}(\bar{p}, D^2\phi(\bar{p})) \geq 0$ .

Let  $y_n \in \Delta(B)$  achieve the minimum in (17) for  $\tilde{v}_n(p_n)$ , so that

$$\tilde{v}_n(p_n) = \max_{x \in (\Delta(A))^S} \left( \lambda_n g(p_n, x, y_n) + (1 - \lambda_n) \sum_{a \in A, b \in B} x(p_n)(a) y_n(b) \tilde{v}_n({}^T P_{\frac{1}{n}}(a, b) \hat{p}_n(x, a)) \right)$$

Up to a subsequence, still denoted  $(y_n)$ , we may assume that  $(y_n)$  converges to some  $y \in \Delta(B)$ .

Let now  $x \in \Delta(S)$  be arbitrary. Applying the latter equation with the non-revealing strategy  $(x, \dots, x) \in \Delta(A)^S$ , one has (with some abuse of notation)  $x(p_n) = x$  and  $\hat{p}_n(x, a) = p_n$ , hence

$$\tilde{v}_n(p_n) \geq \lambda_n g(p_n, x, y_n) + (1 - \lambda_n) \sum_{a \in A, b \in B} x(a) y_n(b) \tilde{v}_n({}^T P_{\frac{1}{n}}(a, b) p_n).$$

Using (18), this yields

$$\begin{aligned} \tilde{v}_n(p_n) \geq & (1 - \lambda_n) (-\phi(p_n) + \tilde{v}_n(p_n)) \\ & + \lambda_n g(p_n, x, y_n) + (1 - \lambda_n) \sum_{a \in A, b \in B} x(a) y_n(b) \phi({}^T P_{\frac{1}{n}}(a, b) p_n) \end{aligned} \quad (19)$$

Since  $P_{\frac{1}{n}}(a, b) = I + \frac{1}{n} {}^T R(a, b) + o(\frac{1}{n})$ ,

$$\phi({}^T P_{\frac{1}{n}}(a, b) p_n) = \phi(p_n) + \frac{1}{n} \langle {}^T R(a, b) p_n, D\phi(p_n) \rangle + o(\frac{1}{n}).$$

Since  $\lambda_n = 1 - e^{-\frac{r}{n}} = \frac{r}{n} + o(\frac{1}{n})$ , the inequality (19) can then be rewritten

$$0 \geq -r \tilde{v}_n(p_n) + r g(p_n, x, y_n) + \sum_{a \in A, b \in B} x(a) y_n(b) \langle {}^T R(a, b) p_n, D\phi(p_n) \rangle + o(1).$$

Letting  $n \rightarrow +\infty$  then yields  $0 \geq -r w(\bar{p}) + r g(\bar{p}, x, y) + \langle {}^T R(x, y) \bar{p}, D\phi(\bar{p}) \rangle$ . Taking the infimum over  $x \in \Delta(A)$  gives:

$$r w(\bar{p}) + \max_{y \in \Delta(B)} \min_{x \in \Delta(A)} \{ -r g(\bar{p}, x, y) - \langle {}^T R(x, y) \bar{p}, D\phi(\bar{p}) \rangle \} \geq 0.$$

In conclusion we have proved that

$$\min \{ r w(\bar{p}) + H(\bar{p}, D\phi(\bar{p})) ; -\lambda_{\max}(\bar{p}, D^2\phi(\bar{p})) \} \geq 0,$$

as desired.

### 5.4.3 Step 3: $w$ is a viscosity subsolution

We will use the following technical remark, which follows from Cardaliaguet and Rainer (2009a) or from Step 1 in the proof of Proposition 4.4 in Grün (2012):

**Lemma 6.** *Let  $w : \Delta(S) \rightarrow \mathbf{R}$  be a concave function and  $\phi$  be a smooth test function such that  $w - \phi$  has a local maximum on  $\Delta(S)$  at a point  $\bar{p} \in \Delta(S)$ . If  $\lambda_{\max}(\bar{p}, D^2\phi(\bar{p})) < 0$ , then there is  $\delta > 0$  such that, for any  $p \in \Delta(S)$  such that  $p - \bar{p} \in T_{\Delta(S)}(\bar{p})$ ,*

$$w(p) \leq w(\bar{p}) + \langle D\phi(\bar{p}), p - \bar{p} \rangle - \delta|p - \bar{p}|^2. \quad (20)$$

Let  $\phi$  be a smooth test function such that  $w - \phi$  has a strict local maximum on  $\Delta(S)$  at some point  $\bar{p} \in \Delta(S)$ . If  $\lambda_{\max}(\bar{p}, D^2\phi(\bar{p})) \geq 0$ , then the desired inequality holds. So we may assume that  $\lambda_{\max}(\bar{p}, D^2\phi(\bar{p})) < 0$  and it remains to check that, in this case,  $rw(\bar{p}) + H(\bar{p}, D\phi(\bar{p})) \leq 0$ .

As before there are  $p_n \in \Delta(S)$  which converge to  $\bar{p}$  and at which  $\tilde{v}_n - \phi$  has a local maximum. Let now  $x_n = (x_{n,s})$  achieve the maximum in (17) for  $\tilde{v}_n(p_n)$ . Given an arbitrary  $y \in \Delta(B)$ , one thus has

$$\tilde{v}_n(p_n) \leq \lambda_n g(p_n, x_n, y) + (1 - \lambda_n) \sum_{a \in A, b \in B} x_n(p_n)(a)y(b)\tilde{v}_n({}^T P_{\frac{1}{n}}(a, b)\hat{p}_n(x_n, a)). \quad (21)$$

Since  $\lambda_n = o(1)$ , since  ${}^T P_{\frac{1}{n}}(a, b)\hat{p}_n(x_n, a) = \hat{p}_n(x_n, a) + o(1)$ , and using the uniform continuity of  $\tilde{v}_n$ , this implies

$$o(1) \leq \sum_{a \in A, b \in B} x_n(p_n)(a)y(b) (\tilde{v}_n(\hat{p}_n(x_n, a)) - \tilde{v}_n(p_n))$$

Let  $x = (x_s)_{s \in S}$  be the limit of (a subsequence of)  $(x_n)_n$ . Letting  $n \rightarrow +\infty$  in the above inequality we get

$$\begin{aligned} 0 &\leq \sum_{a \in A, b \in B} x(\bar{p})(a)y(b) (w(\hat{p}(x, a)) - w(\bar{p})) \\ &= \sum_{a \in A} x(\bar{p})(a) (w(\hat{p}(x, a)) - w(\bar{p})) \end{aligned} \quad (22)$$

As  $\sum_{a \in A} x(\bar{p})(a)\hat{p}(x, a) = \bar{p}$ , all the points<sup>8</sup>  $\hat{p}(x, a)$  belong to the same face of  $\Delta(S)$  as  $\bar{p}$ . Hence  $\hat{p}(x, a) - \bar{p} \in T_{\Delta(S)}(\bar{p})$  for any  $a \in A$ . Since  $w - \phi$  has a local maximum on  $\Delta(S)$  at  $\bar{p} \in \Delta(S)$  with  $\lambda_{\max}(\bar{p}, D^2\phi(\bar{p})) < 0$ , Lemma 6 states that there is  $\delta > 0$  such that, for all  $p \in \Delta(S)$  with  $p - \bar{p} \in T_{\Delta(S)}(\bar{p})$ ,

$$w(p) \leq w(\bar{p}) + \langle D\phi(\bar{p}), p - \bar{p} \rangle - \delta|p - \bar{p}|^2. \quad (23)$$

---

<sup>8</sup>such that  $x(\bar{p})(a) > 0$

Plugging (23) into (22) gives

$$\begin{aligned} 0 &\leq \sum_{a \in A} x(\bar{p})(a) (\langle D\phi(\bar{p}), \hat{p}_n(x_n, a) - \bar{p} \rangle - \delta |\hat{p}_n(x_n, a) - \bar{p}|^2) \\ &= -\delta \sum_{a \in A} x(\bar{p})(a) |\hat{p}_n(x_n, a) - \bar{p}|^2, \end{aligned}$$

because  $\sum_{a \in A} x(\bar{p})(a) \hat{p}_n(x_n, a) = \bar{p}$ . In particular,  $\hat{p}(x, a) = \bar{p}$  if  $x(\bar{p})(a) > 0$ . By definition of  $x(\bar{p})(a)$ , we have therefore  $x_s = x_{s'}$  for any  $s \neq s'$  such that  $\bar{p}_s > 0$  and  $\bar{p}_{s'} > 0$  (which means that  $x$  is non revealing). We denote by  $x \in \Delta(A)$  this common value and note that  $\hat{p}(x, a) = \bar{p}$  whenever  $x(a) > 0$ .

We now come back to (21) and use the concavity of  $\tilde{v}_n$  to deduce that

$$\tilde{v}_n(p_n) \leq \lambda_n g(p_n, x_n, y) + (1 - \lambda_n) \tilde{v}_n \left( \sum_{a \in A, b \in B} x_n(p_n)(a) y(b)^T P_{\frac{1}{n}}(a, b) \hat{p}_n(x_n, a) \right)$$

Since  $\tilde{v}_n - \phi$  has a strict local maximum at  $p_n \in \Delta(S)$ , we get

$$\begin{aligned} 0 &\leq \lambda_n (g(p_n, x_n, y) - \tilde{v}_n(p_n)) \\ &\quad + (1 - \lambda_n) \left( \phi \left( \sum_{a \in A, b \in B} x_n(p_n)(a) y(b)^T P_{\frac{1}{n}}(a, b) \hat{p}_n(x_n, a) \right) - \phi(p_n) \right) \end{aligned} \quad (24)$$

Observe next that

$$\begin{aligned} \sum_{a \in A, b \in B} x_n(p_n)(a) y(b)^T P_{\frac{1}{n}}(a, b) \hat{p}_n(x_n, a) &= p_n + \frac{1}{n} \sum_{a \in A, b \in B} x_n(p_n)(a) y(b)^T R(a, b) \hat{p}_n(x_n, a) + o\left(\frac{1}{n}\right) \\ &= p_n + \frac{1}{n} \sum_{a \in A, b \in B} x_n(p_n)(a) y(b)^T R(a, b) \bar{p} + o\left(\frac{1}{n}\right) \end{aligned}$$

where the second equality holds because  $\hat{p}_n(x_n, a) \rightarrow \bar{p}$ . Plugging into in (24) we obtain

$$0 \leq \lambda_n (g(p_n, x_n, y) - \tilde{v}_n(p_n)) + \frac{(1 - \lambda_n)}{n} \langle D\phi(p_n), \sum_{a \in A, b \in B} x_n(p_n)(a) y(b)^T R(a, b) \bar{p} \rangle + o\left(\frac{1}{n}\right)$$

Since  $x_n(p_n)(a) \rightarrow x(a)$ , multiplying by  $n$  and letting  $n \rightarrow +\infty$  yields

$$0 \leq r (g(\bar{p}, x, y) - w(\bar{p})) + \langle D\phi(\bar{p}), \sum_{a \in A, b \in B} x(a) y(b)^T R(a, b) \bar{p} \rangle$$

When rearranging, we find that

$$r w(\bar{p}) + \min_{x \in \Delta(A)} \max_{y \in \Delta(B)} (-r g(\bar{p}, x, y) - \langle D\phi(\bar{p}), {}^T R(x, y) \bar{p} \rangle) \leq 0.$$

Therefore  $w$  is a subsolution.

## 6 Incomplete information on both sides

### 6.1 Model

The approach developed in the previous section can also be adapted to games with lack of information on both sides, in which each player observes and controls a Markov chain. The framework is close to the one of Gensbittel and Renault (2012). In particular, we also assume that each player observes only one Markov chain. However, the fact that players play more and more often completely changes the nature of the results.

We assume that there are two controlled Markov chains  $(s_t^1)$  and  $(s_t^2)$  with values in the finite sets  $S^1$  and  $S^2$  respectively. The process  $(s_t^i)$  is observed and controlled by Player  $i = 1, 2$ . That is, the generator of  $(s_t^1)_{t \geq 0}$  is of the form  $(R^1(s, s'; a), s, s' \in S^1, a \in A)$ , and that of  $(s_t^2)$  is  $(R^2(s, s'; b), s, s' \in S^1, b \in B)$ . The assumptions on  $R^1$  and  $R^2$  are the same as in the previous section.

Much as before, for given  $a$ , we denote by  $P_t^1(a)$  the transition function of a Markov chain with transition rates  $R^1(\cdot; a)$ , and set  $P_t^1(x) := \sum_{a \in A} x(a)P_t^1(a)$  whenever  $x \in \Delta(A)$ . The transition function  $P_t^2(y)$  is defined similarly for  $y \in \Delta(B)$ .

In this new game, the payoff function depends on both states and actions:  $g : S^1 \times S^2 \times A \times B \rightarrow \mathbf{R}$ . The initial positions  $s_0^1$  and  $s_0^2$  of the chains are chosen independently with laws  $p^1 \in \Delta(S)$  and  $p^2 \in \Delta(S^2)$ . As before, the weight of stage  $k$  in  $G_n(p)$  is  $\lambda_n(1 - \lambda_n)^k$ , with  $\lambda_n := 1 - e^{-r/n}$ . The value of the game with initial distribution  $(p^1, p^2)$  is denoted by  $\tilde{v}_n(p^1, p^2)$ .

### 6.2 The Hamilton-Jacobi equation

In this setting, we have to introduce a slightly new type of Hamilton-Jacobi equation. The Hamiltonian is now a map  $H : \Delta(S^1) \times \Delta(S^2) \times \mathbf{R}^{|S^1|+|S^2|} \rightarrow \mathbf{R}$  and the Hamilton-Jacobi equation is given by the pair of inequalities

$$\begin{aligned} \max \{ \min \{ rw + H(p^1, p^2, Dw); -\lambda_{\max}(p^1, D_{11}^2 w) \}; -\lambda_{\min}(p^2, D_{22}^2 w) \} &\leq 0 \\ &\text{in } \Delta(S^1) \times \Delta(S^2), \\ \min \{ \max \{ rw + H(p^1, p^2, Dw); -\lambda_{\min}(p^2, D_{22}^2 w) \}; -\lambda_{\max}(p^1, D_{11}^2 w) \} &\geq 0 \\ &\text{in } \Delta(S^1) \times \Delta(S^2). \end{aligned} \quad (25)$$

In the above expressions,  $r > 0$  is the discount rate,  $w : \Delta(S^1) \times \Delta(S^2) \rightarrow \mathbf{R}$  is the unknown (formally extended to a neighborhood of  $\Delta(S^1) \times \Delta(S^2)$ ),  $Dw = (D_1 w, D_2 w)$  is the full gradient of  $w$  with respect to  $(p^1, p^2)$ ,  $D_1 w$  (resp.  $D_2 w$ ) being the derivative with respect to  $p_1$  (resp.  $p_2$ ),  $D_{11}^2 w$  (resp.  $D_{22}^2 w$ ) is the second order derivative of  $w$  with respect to  $p^1$  (resp.

$p^2$ ),  $\lambda_{\max}(p^1, X)$  is defined by (12) while

$$\lambda_{\min}(p^2, X) = \inf \left\{ \frac{\langle Xz, z \rangle}{|z|^2}, z \in T_{\Delta(S^2)}(p^2) \setminus \{0\} \right\} \quad (26)$$

where  $T_{\Delta(S^2)}(p^2)$  is the tangent space of  $\Delta(S^2)$  at  $p^2$ . By convention we set  $\lambda_{\min}(p^2, X) = +\infty$  if  $T_{\Delta(S^2)}(p^2) = \{0\}$ .

As before, one cannot expect equation (25) to have a smooth solution in general, and we use instead the following notion of viscosity solution:

**Definition 2.** A map  $w : \Delta(S^1) \times \Delta(S^2) \rightarrow \mathbf{R}$  is a viscosity subsolution of (25) if it is upper semicontinuous and if, for any smooth test function  $\phi : \mathbf{R}^{|S^1|+|S^2|} \rightarrow \mathbf{R}$  such that  $w - \phi$  has a local maximum on  $\Delta(S^1) \times \Delta(S^2)$  at  $(p^1, p^2) \in \Delta(S^1) \times \Delta(S^2)$ , one has

$$\max \left\{ \min \left\{ rw(p^1, p^2) + H(p^1, p^2, D\phi(p^1, p^2)) ; -\lambda_{\max}(p^1, D_{11}^2\phi(p^1, p^2)) \right\} ; -\lambda_{\min}(p^2, D_{22}^2\phi(p^1, p^2)) \right\} \leq 0 .$$

A map  $w : \Delta(S^1) \times \Delta(S^2) \rightarrow \mathbf{R}$  is a viscosity supersolution of (25) if it is lower semicontinuous and if, for any smooth test function  $\phi : \mathbf{R}^{|S^1|+|S^2|} \rightarrow \mathbf{R}$  such that  $w - \phi$  has a local minimum on  $\Delta(S^1) \times \Delta(S^2)$  at  $(p^1, p^2) \in \Delta(S^1) \times \Delta(S^2)$ , one has

$$\min \left\{ \max \left\{ rw(p^1, p^2) + H(p^1, p^2, D\phi(p^1, p^2)) ; -\lambda_{\min}(p^2, D_{22}^2\phi(p^1, p^2)) \right\} ; -\lambda_{\max}(p^1, D_{11}^2\phi(p^1, p^2)) \right\} \geq 0 .$$

Finally,  $w : \Delta(S^1) \times \Delta(S^2) \rightarrow \mathbf{R}$  is a viscosity solution of (25) if it is a sub- and a supersolution of (25).

Uniqueness of a viscosity solution for (25) holds thanks to a comparison principle, which generalizes Theorem 6. We will assume that  $H : \Delta(S^1) \times \Delta(S^2) \times \mathbf{R}^{|S^1|+|S^2|} \rightarrow \mathbf{R}$  satisfies the condition:

$$|H(p, q) - H(p, q')| \leq C|q - q'| \quad \forall p \in \Delta(S^1) \times \Delta(S^2), q, q' \in \mathbf{R}^{|S^1|+|S^2|} \quad (27)$$

as well as

$$|H(p, q) - H(p', q)| \leq C|p - p'|(|q| + 1) \quad \forall p, p' \in \Delta(S^1) \times \Delta(S^2), q \in \mathbf{R}^{|S^1|+|S^2|} \quad (28)$$

**Theorem 8.** Assume that (27) and (28) hold. Let  $w_1$  be a continuous viscosity subsolution of (25) and  $w_2$  be a continuous viscosity supersolution of (25). Then  $w_1 \leq w_2$  in  $\Delta(S^1) \times \Delta(S^2)$ .

### 6.3 The limit theorem

Here is our main result in the framework of lack of information on both sides.

**Theorem 9.** *The uniform limit  $v = \lim_{n \rightarrow +\infty} \tilde{v}_n$  exists and is the unique viscosity solution of Hamilton-Jacobi equation (25) where  $H$  is given by*

$$\begin{aligned} H(p^1, p^2, \xi^1, \xi^2) &= \min_{x \in \Delta(A)} \max_{y \in \Delta(B)} \{ -\langle {}^T R^1(x) p^1, \xi^1 \rangle - \langle {}^T R^2(y) p^2, \xi^2 \rangle - rg(p^1, p^2, x, y) \} \\ &= \max_{y \in \Delta(B)} \min_{x \in \Delta(A)} \{ -\langle {}^T R^1(x) p^1, \xi^1 \rangle - \langle {}^T R^2(y) p^2, \xi^2 \rangle - rg(p^1, p^2, x, y) \} \end{aligned} \quad (29)$$

for any  $(p^1, p^2, \xi^1, \xi^2) \in \Delta(S^1) \times \Delta(S^2) \times \mathbf{R}^{|S^1|} \times \mathbf{R}^{|S^2|}$ .

Note that the Hamiltonian defined in (29) satisfies conditions (27) and (28). So equation (25) has at most one viscosity solution.

### 6.4 Proof of Theorem 9

The proof is close to the one for Theorem 7. The main difference is that we have to deal with the fact that both players now have private information, which complicates the proof of the viscosity solution property. On the other hand, the problem is now symmetrical, so that it is enough to show the supersolution property, the argument for the subsolution being identical.

#### 6.4.1 Step 1: Dynamic programming principle and regularity

For  $(p^1, p^2) \in \Delta(S^1) \times \Delta(S^2)$ ,  $(x, y) \in (\Delta(A))^{|S^1|} \times (\Delta(B))^{|S^2|}$  and  $(a, b) \in A \times B$ , we set

$$x(p^1)(a) = \sum_{s \in S^1} p_s^1 x_s(a) \quad \text{and} \quad y(p^2)(b) = \sum_{s \in S^2} p_s^2 y_s(b)$$

and  $\hat{p}^1(x, a) := \left( \frac{p_s^1 x_s(a)}{x(p^1)(a)} \right)_{s \in S^1}$  and  $\hat{p}^2(y, b) := \left( \frac{p_s^2 y_s(b)}{y(p^2)(b)} \right)_{s \in S^2}$  denote the conditional distributions of the states given  $a$  and  $b$  respectively.

The dynamic programming principle for  $\tilde{v}_n$  reads

$$\begin{aligned}
& \tilde{v}_n(p^1, p^2) \\
&= \max_{x \in (\Delta(A))^{|S^1|}} \min_{y \in (\Delta(B))^{|S^2|}} \left( \lambda_n g(p^1, p^2, x, y) \right. \\
&\quad \left. + (1 - \lambda_n) \sum_{a \in A, b \in B} x(p^1)(a) y(p^2)(b) \tilde{v}_n({}^T P_{\frac{1}{n}}^1(a) \hat{p}^1(x, a), {}^T P_{\frac{1}{n}}^2(b) \hat{p}^2(y, b)) \right) \\
&= \min_{y \in (\Delta(B))^{|S^2|}} \max_{x \in (\Delta(A))^{|S^1|}} \left( \lambda_n g(p^1, p^2, x, y) \right. \\
&\quad \left. + (1 - \lambda_n) \sum_{a \in A, b \in B} x(p^1)(a) y(p^2)(b) \tilde{v}_n({}^T P_{\frac{1}{n}}^1(a) \hat{p}^1(x, a), {}^T P_{\frac{1}{n}}^2(b) \hat{p}^2(y, b)) \right)
\end{aligned} \tag{30}$$

As before, the maps  $\tilde{v}_n$  are uniformly Lipschitz and bounded, and we will prove that any (uniform) accumulation point of the sequence  $(\tilde{v}_n)$  is a viscosity solution of (25). Again up to a subsequence, we may assume that  $(\tilde{v}_n)$  converges to some continuous map  $w$ .

#### 6.4.2 Step 2: $w$ is a viscosity supersolution

Let  $\phi$  be a smooth test function such that  $w - \phi$  has a strict local minimum on  $\Delta(S^1) \times \Delta(S^2)$  at some point  $(\bar{p}^1, \bar{p}^2) \in \Delta(S^1) \times \Delta(S^2)$ . From standard arguments, this implies the existence of a sequence  $(p_n^1, p_n^2)$  which converges to  $(\bar{p}^1, \bar{p}^2)$  and such that  $\tilde{v}_n - \phi$  has a local minimum at  $(p_n^1, p_n^2)$  for any  $n$ : namely,

$$\tilde{v}_n(p^1, p^2) \geq \phi(p^1, p^2) - \phi(p_n^1, p_n^2) + \tilde{v}_n(p_n^1, p_n^2) \quad \text{for any } (p^1, p^2) \in \Delta(S^1) \times \Delta(S^2). \tag{31}$$

As  $\tilde{v}_n$  is concave in  $p^1$ , we must have  $\lambda_{\max}(p_n^1, D_{11}^2 \phi(p_n^1, p_n^2)) \leq 0$  by definition of  $T_{\Delta(S^1)}(p_n^1)$ . Since  $T_{\Delta(S^1)}(p_n^1) \supset T_{\Delta(S^1)}(\bar{p}^1)$  for  $n$  large enough, we get, as  $n \rightarrow +\infty$ :

$$-\lambda_{\max}(\bar{p}^1, D_{11}^2 \phi(\bar{p}^1, \bar{p}^2)) \geq 0.$$

It remains to check that

$$\max \left\{ r w(\bar{p}^1, \bar{p}^2) + H(\bar{p}^1, \bar{p}^2, D\phi(\bar{p}^1, \bar{p}^2)) ; -\lambda_{\min}(\bar{p}^2, D_{22}^2 \phi(\bar{p}^1, \bar{p}^2)) \right\} \geq 0.$$

For this we assume that  $\lambda_{\min}(\bar{p}^2, D_{22}^2 \phi(\bar{p}^1, \bar{p}^2)) > 0$  and we are left to prove that

$$r w(\bar{p}^1, \bar{p}^2) + H(\bar{p}^1, \bar{p}^2, D\phi(\bar{p}^1, \bar{p}^2)) \geq 0.$$

Let  $y_n \in (\Delta(B))^{|S^2|}$  be optimal in the dynamic programming equation (30) for  $\tilde{v}_n(p_n)$ :

$$\begin{aligned}
\tilde{v}_n(p_n^1, p_n^2) &= \max_{x \in (\Delta(A))^{|S^1|}} \left( \lambda_n g(p_n^1, p_n^2, x, y_n) \right. \\
&\quad \left. + (1 - \lambda_n) \sum_{a \in A, b \in B} x(p_n^1)(a) y_n(p_n^2)(b) \tilde{v}_n({}^T P_{\frac{1}{n}}^1(a) \hat{p}_n^1(x, a), {}^T P_{\frac{1}{n}}^2(b) \hat{p}_n^2(y, b)) \right).
\end{aligned}$$

Let  $y = (y_s)_{s \in S^2}$  be the limit of (a subsequence of)  $(y_n)$ . Fix  $x \in \Delta(A)$ . (With a slight abuse of notation), if Player 1 plays the non-revealing strategy  $(x, \dots, x) \in (\Delta(A))^{S^1}$ , we get  $x(p_n^1) = x$  and  $\hat{p}_n^1(x, a) = p_n^1$  and therefore

$$\begin{aligned} \tilde{v}_n(p_n^1, p_n^2) \geq & \lambda_n g(p_n^1, p_n^2, x, y_n) \\ & + (1 - \lambda_n) \sum_{a \in A, b \in B} x(a) y_n(p_n^2)(b) \tilde{v}_n({}^T P_{\frac{1}{n}}^1(a) p_n^1, {}^T P_{\frac{1}{n}}^2(b) \hat{p}_n^2(y, b)). \end{aligned} \quad (32)$$

Recalling that  $\lambda_n = o(1)$ , that  ${}^T P_{\frac{1}{n}}^1(a, b) p_n^1 = p_n^1 + o(1)$  and that  ${}^T P_{\frac{1}{n}}^1(a, b) \hat{p}_n^2(y, b) = \hat{p}_n^2(y, b) + o(1)$ , we get, letting  $n \rightarrow +\infty$  in (32),

$$w(\bar{p}^1, \bar{p}^2) \geq \sum_{b \in B} y(\bar{p}^2)(b) w(\bar{p}^1, \hat{p}^2(y, b)). \quad (33)$$

From (33) we conclude as in the proof of Theorem 7 that  $\hat{p}^2(y, b) = \bar{p}^2$  if  $y(\bar{p}^2)(b) > 0$ . Coming back to the definition of  $y(\bar{p}^2)(b) > 0$ , we have therefore that  $y_s = y_{s'}$  for any  $s \neq s'$  such that  $\bar{p}_s^2 > 0$  and  $\bar{p}_{s'}^2 > 0$ : this means that  $y$  is non revealing. We denote by  $y \in \Delta(B)$  this common value and note that  $\hat{p}^2(y, b) = \bar{p}$  whenever  $y(b) > 0$ .

With this in mind, we come back to (32), which becomes, since  $\tilde{v}_n$  is convex in  $p^2$ , and since the dynamics of  $(s_t^1)$  is independent of Player 2:

$$\begin{aligned} \tilde{v}_n(p_n^1, p_n^2) \geq & \lambda_n g(p_n^1, p_n^2, x, y_n) \\ & + (1 - \lambda_n) \sum_{a \in A} x(a) \tilde{v}_n \left( {}^T P_{\frac{1}{n}}^1(a) p_n^1, \sum_{b \in B} y_n(p_n^2)(b) {}^T P_{\frac{1}{n}}^2(b) \hat{p}_n^2(y, b) \right). \end{aligned}$$

We next use the fact that  $\tilde{v}_n - \phi$  has a local minimum at  $(p_n^1, p_n^2)$ :

$$\begin{aligned} 0 \geq & \lambda_n (g(p_n^1, p_n^2, x, y_n) - \tilde{v}_n(p_n^1, p_n^2)) \\ & + (1 - \lambda_n) \sum_{a \in A} x(a) \left( \phi({}^T P_{\frac{1}{n}}^1(a) p_n^1, \sum_{b \in B} y_n(p_n^2)(b) {}^T P_{\frac{1}{n}}^2(b) \hat{p}_n^2(y, b)) - \phi(p_n^1, p_n^2) \right) \end{aligned} \quad (34)$$

where

$${}^T P_{\frac{1}{n}}^1(a) p_n^1 = p_n^1 + \frac{1}{n} {}^T R^1(a) p_n^1 + o\left(\frac{1}{n}\right)$$

while, as  $\sum_{b \in B} y_n(p_n^2)(b) \hat{p}_n^2(y, b) = p_n^2$ ,

$$\sum_{b \in B} y_n(p_n^2)(b) {}^T P_{\frac{1}{n}}^2(b) \hat{p}_n^2(y, b) = p_n^2 + \frac{1}{n} \sum_{b \in B} y_n(p_n^2)(b) {}^T R^2(b) \hat{p}_n^2(y, b) + o\left(\frac{1}{n}\right).$$

Multiplying (34) by  $n$  and letting  $n \rightarrow +\infty$  gives therefore

$$0 \geq r (g(\bar{p}^1, \bar{p}^2, x, y) - w(\bar{p}^1, \bar{p}^2)) + \sum_{a \in A} x(a) \langle D\phi(\bar{p}^1, \bar{p}^2), \left( {}^T R^1(a) \bar{p}^1, \sum_{b \in B} y(b) {}^T R^2(b) \bar{p}^2 \right) \rangle.$$

Rearranging we find that

$$rw(\bar{p}) + \min_{x \in \Delta(A)} \max_{y \in \Delta(B)} \left( -g(\bar{p}^1, \bar{p}^2, x, y) - \langle D\phi(\bar{p}), ({}^T R^1(x)\bar{p}^1, {}^T R^2(y)\bar{p}^2) \rangle \right) \geq 0.$$

Therefore  $w$  is a supersolution.

## A Technical results

### A.1 Proof of Lemma 5

Let a probability measure  $Q$  over  $\Delta(S)^d N$  be as stated. Generic elements of  $(\Delta(S) \times S)^{\mathbf{N}}$  are denoted  $(q_m, s_m)_{m \in \mathbf{N}}$ . To avoid multiplying notations, and at the cost of a notational abuse, we will write  $Q(q^m; q_{m+1})$  for the conditional law of  $q_{m+1}$  given  $q^m$ . Given a probability measure  $\tilde{\mathbf{P}}$  over  $(\Delta(S) \times S)^{\mathbf{N}}$ , we similarly write  $\tilde{\mathbf{P}}(q^m, s_m; q_{m+1}, s_{m+1})$  for the law of  $(q_{m+1}, s_{m+1})$  given  $(q^m, s_m)$ ,  $\tilde{\mathbf{P}}(q^m, s_m, s_{m+1}; q_{m+1})$  for the law of  $q_{m+1}$  given  $(q^m, s_m, s_{m+1})$ , etc., with semi-colons separating conditioning variables from the others.

For  $m \geq 1$ , denote by  $\theta_m$  the transition function from  $\Delta(S)^m \times S$  to  $\Delta(S) \times S$  defined by

$$\theta_m(q^m; F, s_{m+1}) = \int_F q_{m+1}(s_{m+1}) Q(q^m; dq_{m+1}). \quad (35)$$

Intuitively,  $\theta_m(q^m; q_{m+1}, s_{m+1})$  is the probability obtained when first choosing  $q_{m+1}$  according to its (conditional) law  $Q(q^m; q_{m+1})$ , then picking  $s_{m+1}$  according to  $q_{m+1}$ , and we define the sequence  $\mu = (\mu_m)$  by

$$\mu_m(q^m, s_{m+1}; F) := \frac{\theta_m(q^m; F, s_{m+1})}{\theta_m(q^m; s_{m+1})},$$

so that  $\mu_m(q^m, s_{m+1}; q_{m+1})$  is “the conditional law of  $q_{m+1}$  given  $q^m$  and  $s_{m+1}$ ”.

We now prove by induction that the induced distribution  $\mu \circ \mathbf{P}$  over  $(\Delta(S) \times S)^{\mathbf{N}}$  satisfies **C1** and **C2**. We thus assume that for some  $m$ , (i) the conditional law of  $q_m$  given  $q^{m-1}$  (under  $\mu \circ \mathbf{P}$ ) is equal to  $Q(q^{m-1}; q_m)$ , (ii) the conditional law of  $s_m$  given  $q^m$  is equal to  $q_m$ , and prove that (i) and (ii) also hold for  $m+1$ .

For  $F \subseteq \Delta(S)$ , note first that by (ii), one has

$$\mu \circ \mathbf{P}(q^m; q_{m+1} \in F) = \sum_{s_m \in S} \mu \circ \mathbf{P}(q^m, s_m; F) q_m(s_m) = \sum_{s_m, s_{m+1} \in S} \mu \circ \mathbf{P}(q^m, s_m; F, s_{m+1}) q_m(s_m).$$

Since  $\mu \circ \mathbf{P}(q^m, s_m; F, s_{m+1}) = \mu_m(q^m, s_{m+1}; F) \pi(s_{m+1} | s_m)$  by definition, one also has

$$\mu \circ \mathbf{P}(q^m; q_{m+1} \in F) = \sum_{s_m, s_{m+1}} \frac{\theta_m(q^m; F, s_{m+1})}{\theta_m(q^m; s_{m+1})} q_m(s_m) \pi(s_{m+1} | s_m).$$

Observe next that, by the induction assumption, since  $\mathbf{E}[q_{m+1} \mid q^m] =^T \Pi q_m$  and since  $s_{m+1}$  and  $q^m$  are conditionally independent given  $s_m$ , one has

$$\theta_m(q^m; s_{m+1}) = \sum_{s_m} q_m(s_m) \pi(s_{m+1} \mid s_m).$$

Hence

$$\mu \circ \mathbf{P}(q^m; F) = \sum_{s_{m+1}} \theta_m(q^m; s_{m+1}, F) = Q(q^m; F), \quad (36)$$

as desired.

We now prove that  $\mu \circ \mathbf{P}(q^m, q_{m+1}; s_{m+1}) = q_{m+1}(s_{m+1})$ . One has (beware of the semicolons)

$$\mu \circ \mathbf{P}(q^m, q_{m+1}; s_{m+1}) = \frac{\mu \circ \mathbf{P}(q^m; q_{m+1}, s_{m+1})}{\mu \circ \mathbf{P}(q^m; q_{m+1})} = \frac{\mu \circ \mathbf{P}(q^m, s_{m+1}; q_{m+1}) \mu \circ \mathbf{P}(q^m; s_{m+1})}{Q(q^m; q_{m+1})} \quad (37)$$

using (conditional) Bayes laws and (36). Observe next that

$$\begin{aligned} \mu \circ \mathbf{P}(q^m, s_{m+1}; q_{m+1}) &= \mu_m(q^m, s_{m+1}; q_{m+1}) \\ &= \frac{\theta_m(q^m; q_{m+1}, s_{m+1})}{\theta_m(q^m; s_{m+1})} \end{aligned}$$

and

$$\begin{aligned} \mu \circ \mathbf{P}(q^m; s_{m+1}) &= \sum_{s_m} \mu \circ \mathbf{P}(q^m; s_m, s_{m+1}) \\ &= \sum_{s_m} \mu \circ \mathbf{P}(q^m; s_m), s_{m+1}) \times \mu \circ \mathbf{P}(q^m, s_m; s_{m+1}) \\ &= \sum_{s_m} q_m(s_m) \pi(s_{m+1} \mid s_m) = \mathbf{E}[q_{m+1}(s_{m+1}) \mid q^m], \end{aligned}$$

where the second equality holds by (9). Plugging into (37), this yields

$$\mu \circ \mathbf{P}(q^m, q_{m+1}; s_{m+1}) = \frac{\theta_m(q^m; q_{m+1}, s_{m+1})}{\theta_m(q^m; s_{m+1})} \times \frac{\mathbf{E}[q_{m+1}(s_{m+1}) \mid q^m]}{\theta_m(q^m; q_{m+1})}.$$

To conclude, recall that (see (35))

$$\theta_m(q^m; q_{m+1}, s_{m+1}) = \theta_m(q^m; q_{m+1}) q_{m+1}(s_{m+1}),$$

while

$$\theta_m(q_m; s_{m+1}) = \int_{\Delta(S)} q_{m+1}(s_{m+1}) Q(q^m; dq_{m+1}) = \mathbf{E}[q_{m+1}(s_{m+1}) \mid q^m],$$

so that  $\mu \circ \mathbf{P}(q^m, q_{m+1}; s_{m+1}) = q_{m+1}(s_{m+1})$ , as desired.

## A.2 Proof of the comparison principle

In this section we prove Theorem 8 (which implies Theorem 6). We follow here the proof of Crandall, Ishii and Lions (1992) for second order Hamilton-Jacobi equations. However their results do not apply directly, because the terms  $\lambda_{\min}$  and  $\lambda_{\max}$  introduce a strong degeneracy in the equation. This issue is also present in Cardaliaguet and Rainer (2009a), where it is dealt with by an induction argument over the dimension of the faces of the simplices, which relies on the fact that the restriction of solutions to faces are still solutions. This is no longer the case here. This forces us to revisit the proof, and to come back to the basic technique consisting in regularizing the solutions by inf- or sup convolution, and then in using Jensen Lemma.

Let  $w_1$  be a subsolution and  $w_2$  be a supersolution of (25). Our aim is to show that  $w_1 \leq w_2$ . We argue by contradiction, and assume that

$$M := \sup_{p \in \Delta(S^1) \times \Delta(S^2)} \{w_1(p) - w_2(p)\} > 0. \quad (38)$$

In order to use the special structure of the problem, we have to regularize the maps  $w_1$  and  $w_2$  by sup and inf-convolution respectively. This technique is standard and we refer to Crandall, Ishii and Lions (1992) for details. For  $\delta > 0$  and  $p \in \mathbf{R}^{|S^1|+|S^2|}$  we set

$$w_1^\delta(p) = \max_{p' \in \Delta(S^1) \times \Delta(S^2)} \left\{ w_1(p') - \frac{1}{2\delta} |p - p'|^2 \right\}$$

and

$$w_{2,\delta}(p) = \min_{p' \in \Delta(S^1) \times \Delta(S^2)} \left\{ w_2(p') + \frac{1}{2\delta} |p - p'|^2 \right\}$$

We note for later use that  $w_1^\delta$  and  $w_{2,\delta}$  are now defined over the entire space  $\mathbf{R}^{|S^1|+|S^2|}$ , that  $w_1^\delta$  is semiconvex while  $w_{2,\delta}$  is semiconcave (see Crandall et al (1992)). Moreover,

$$\lim_{|p| \rightarrow +\infty} |p|^{-1} w_1^\delta(p) = -\infty, \quad \lim_{|p| \rightarrow +\infty} |p|^{-1} w_{2,\delta}(p) = +\infty. \quad (39)$$

Setting

$$M_\delta = \sup_{p \in \mathbf{R}^{|S^1|+|S^2|}} \{w_1^\delta(p) - w_{2,\delta}(p)\} \quad (40)$$

we have:

**Lemma 7.** *For any  $\delta > 0$ , the problem (40) has at least one maximum point. If  $p_\delta$  is such a maximum point and if  $p'_\delta \in \Delta(S^1) \times \Delta(S^2)$  and  $p''_\delta \in \Delta(S^1) \times \Delta(S^2)$  are such that*

$$(i) \ w_1^\delta(p_\delta) = w_1(p'_\delta) - \frac{1}{2\delta} |p_\delta - p'_\delta|^2 \quad \text{and} \quad (ii) \ w_{2,\delta}(p_\delta) = w_2(p''_\delta) + \frac{1}{2\delta} |p_\delta - p''_\delta|^2 \quad (41)$$

then, as  $\delta \rightarrow 0$ ,  $M_\delta \rightarrow M$  while

$$\frac{|p_\delta - p'_\delta|^2}{2\delta} + \frac{|p_\delta - p''_\delta|^2}{2\delta} \rightarrow 0.$$

*Proof.* The existence of a maximum point is a straightforward consequence of (39). The rest of the statement is classical.  $\square$

Next we note that  $w_1^\delta$  and  $w_{2,\delta}$  are still respectively a subsolution and a supersolution of slightly modified equations:

**Lemma 8.** *Assume that  $v_1^\delta$  has a second order Taylor expansion at a point  $p$ . Then*

$$\min \{rw_1^\delta(p) + H(p', Dw_1^\delta(p)) ; -\lambda_{\max}((p')^1, D_{11}^2 w_1^\delta(p))\} \leq 0, \quad (42)$$

where  $p' = ((p')^1, (p')^2) \in \Delta(S^1) \times \Delta(S^2)$  is such that

$$w_1^\delta(p) = w_1(p') - \frac{1}{2\delta}|p - p'|^2.$$

Similarly, if  $w_{2,\delta}$  has a second order Taylor expansion at a point  $p$ , then

$$\max \{rw_{2,\delta}(p) + H(p'', Dw_{2,\delta}(p)) ; -\lambda_{\min}((p'')^2, D_{22}^2 w_{2,\delta}(p))\} \geq 0, \quad (43)$$

where  $p'' = ((p'')^1, (p'')^2) \in \Delta(S^1) \times \Delta(S^2)$  is such that

$$w_{2,\delta}(p) = w_2(p'') + \frac{1}{2\delta}|p - p''|^2.$$

*Proof.* We do the proof for  $w_1^\delta$ , the argument for  $w_{2,\delta}$  being symmetrical. Assume that  $w_1^\delta$  has a second order Taylor expansion at a point  $\bar{p}$  and set, for  $\gamma > 0$  small,

$$\phi_\gamma(p) = \langle Dw_1^\delta(\bar{p}), p - \bar{p} \rangle + \frac{1}{2} \langle D^2 w_1^\delta(\bar{p})(p - \bar{p}), p - \bar{p} \rangle + \frac{\gamma}{2}|p - \bar{p}|^2.$$

We also denote by  $\bar{p}'$  a point in  $\Delta(S^1) \times \Delta(S^2)$  such that

$$w_1^\delta(\bar{p}) = w_1(\bar{p}') - \frac{1}{2\delta}|\bar{p} - \bar{p}'|^2. \quad (44)$$

Then  $w_1^\delta - \phi_\gamma$  has a maximum at  $\bar{p}$ , which implies, by definition of  $w_1^\delta$ , that

$$w_1(p') - \frac{1}{2\delta}|p' - p|^2 \leq \phi_\gamma(p) - \phi_\gamma(\bar{p}) + w_1^\delta(\bar{p}) \quad \forall p \in \mathbf{R}^{|\mathcal{S}^1| \times |\mathcal{S}^2|}, p' \in \Delta(S^1) \times \Delta(S^2),$$

with an equality for  $(p, p') = (\bar{p}, \bar{p}')$ . If we choose  $p = p' - \bar{p}' + \bar{p}$  in the above formula, we get:

$$w_1(p') \leq \phi_\gamma(p' - \bar{p}' + \bar{p}) + \frac{1}{2\delta}|p' - \bar{p}'|^2 - \phi_\gamma(\bar{p}) + w_1^\delta(\bar{p}) \quad \forall p' \in \Delta(S^1) \times \Delta(S^2),$$

with an equality at  $p' = \bar{p}'$ . As  $w_1$  is a subsolution, we obtain therefore, using the right-hand side of the above inequality as a test function,

$$\min \{rw_1(\bar{p}') + H(\bar{p}', D\phi_\gamma(\bar{p})) ; -\lambda_{\max}((\bar{p}')^1, D_{11}^2\phi_\gamma(\bar{p}))\} \leq 0.$$

We note that  $D\phi_\gamma(\bar{p}) = Dw_1^\delta(\bar{p})$ ,  $D_{11}^2\phi_\gamma(\bar{p}) = D^2w_1^\delta(\bar{p}) + \gamma I$  and  $w_1(\bar{p}') \geq w_1^\delta(\bar{p})$  (by (44)). So letting  $\gamma \rightarrow 0$  we obtain the desired result.  $\square$

In order to exploit inequalities (42) and (43), we have to produce points at which  $w_1^\delta$  is strictly concave with respect to the first variable while  $w_2^\delta$  is strictly convex with respect to the second one. For this reason, we introduce a new penalization. For  $\sigma > 0$  and  $p = (p^1, p^2) \in \mathbf{R}^{|S^1|+|S^2|}$ , let us set  $\bar{\xi}_1(p^1) = (1 + |p^1|^2)^{\frac{1}{2}}$ ,  $\bar{\xi}_2(p^2) = (1 + |p^2|^2)^{\frac{1}{2}}$  and

$$M_{\delta,\sigma} = \sup_{p \in \mathbf{R}^{|S^1|+|S^2|}} \{w_1^\delta(p) - w_{2,\delta}(p) + \sigma\bar{\xi}_1(p^1) + \sigma\bar{\xi}_2(p^2)\}$$

Using (39), one easily checks that there exists a maximizer  $(\hat{p}^1, \hat{p}^2)$  to the above problem. In order to use Jensen's Lemma (Lemma A.3. in Crandall, Ishii and Lions [8]), we need this maximum to be strict. For this we modify slightly  $\bar{\xi}_1$  and  $\bar{\xi}_2$ : we set, for  $i = 1, 2$ ,  $\xi_i(p^i) = \bar{\xi}_i(p^i) - \sigma(1 + |p^i - \hat{p}^i|^2)^{\frac{1}{2}}$ . We will choose  $\sigma > 0$  so small that  $\xi_1$  and  $\xi_2$  still have a positive second order derivative. By definition,

$$M_{\delta,\sigma} = \sup_{p \in \mathbf{R}^{|S^1|+|S^2|}} \{w_1^\delta(p) - w_{2,\delta}(p) + \sigma\xi_1(p^1) + \sigma\xi_2(p^2)\}$$

and the above problem has a strict maximum at  $(\hat{p}^1, \hat{p}^2)$ . As the map  $p \rightarrow w_1^\delta(p) - w_{2,\delta}(p) + \sigma\xi_1(p^1) + \sigma\xi_2(p^2)$  is semiconcave, Jensen's Lemma states that, for any  $\varepsilon > 0$ , there is vector  $a_\varepsilon \in \mathbf{R}^{|S^1|+|S^2|}$  with  $|a_\varepsilon| \leq \varepsilon$ , such that the problem

$$M_{\delta,\sigma,\varepsilon} := \sup_{p \in \mathbf{R}^{|S^1|+|S^2|}} \{w_1^\delta(p) - w_{2,\delta}(p) + \sigma\xi_1(p^1) + \sigma\xi_2(p^2) + \langle a_\varepsilon, p \rangle\}$$

has a maximum point  $p_{\delta,\sigma,\varepsilon} \in \mathbf{R}^{|S^1|+|S^2|}$  at which the maps  $w_1^\delta$  and  $w_{2,\delta}$  have a second order Taylor expansion. From Lemma 8, we have

$$\min \{rw_1^\delta(p_{\delta,\sigma,\varepsilon}) + H(p'_{\delta,\sigma,\varepsilon}, Dw_1^\delta(p_{\delta,\sigma,\varepsilon})) ; -\lambda_{\max}((p'_{\delta,\sigma,\varepsilon})^1, D_{11}^2w_1^\delta(p_{\delta,\sigma,\varepsilon}))\} \leq 0, \quad (45)$$

and

$$\max \{rw_{2,\delta}(p_{\delta,\sigma,\varepsilon}) + H(p''_{\delta,\sigma,\varepsilon}, Dw_{2,\delta}(p_{\delta,\sigma,\varepsilon})) ; -\lambda_{\min}((p''_{\delta,\sigma,\varepsilon})^2, D_{22}^2w_{2,\delta}(p_{\delta,\sigma,\varepsilon}))\} \geq 0, \quad (46)$$

where  $p'_{\delta,\sigma,\varepsilon}$  and  $p''_{\delta,\sigma,\varepsilon}$  are points in  $\Delta(S^1) \times \Delta(S^2)$  at which one has

$$w_1^\delta(p_{\delta,\sigma,\varepsilon}) = w_1(p'_{\delta,\sigma,\varepsilon}) - \frac{1}{2\delta}|p_{\delta,\sigma,\varepsilon} - p'_{\delta,\sigma,\varepsilon}|^2 \text{ and } w_{2,\delta}(p_{\delta,\sigma,\varepsilon}) = w_2(p''_{\delta,\sigma,\varepsilon}) + \frac{1}{2\delta}|p_{\delta,\sigma,\varepsilon} - p''_{\delta,\sigma,\varepsilon}|^2.$$

Note for later use that

$$Dw_1^\delta(p_{\delta,\sigma,\varepsilon}) = -\frac{1}{\delta} (p_{\delta,\sigma,\varepsilon} - p'_{\delta,\sigma,\varepsilon}) \quad \text{and} \quad Dw_{2,\delta}(p_{\delta,\sigma,\varepsilon}) = \frac{1}{\delta} (p_{\delta,\sigma,\varepsilon} - p''_{\delta,\sigma,\varepsilon}). \quad (47)$$

By definition of  $M_{\delta,\sigma,\varepsilon}$  we have

$$w_1^\delta(p) \leq M_{\delta,\sigma,\varepsilon} + w_{2,\delta}(p) - \sigma(\xi_1(p^1) + \xi_2(p^2)) \langle a_\varepsilon, p \rangle \quad \forall p \in \mathbf{R}^{|S^1|+|S^2|},$$

with an equality at  $p_{\delta,\sigma,\varepsilon}$ . Hence

$$Dw_1^\delta(p_{\delta,\sigma,\varepsilon}) = Dw_{2,\delta}(p_{\delta,\sigma,\varepsilon}) - \sigma \begin{pmatrix} D_1 \xi_1(p_{\delta,\sigma,\varepsilon}^1) \\ D_2 \xi_2(p_{\delta,\sigma,\varepsilon}^2) \end{pmatrix} - a_\varepsilon \quad (48)$$

while

$$D^2 w_1^\delta(p_{\delta,\sigma,\varepsilon}) \leq D^2 w_{2,\delta}(p_{\delta,\sigma,\varepsilon}) - \sigma \begin{pmatrix} D_{1,1}^2 \xi_1(p_{\delta,\sigma,\varepsilon}^1) & 0 \\ 0 & D_{2,2}^2 \xi_2(p_{\delta,\sigma,\varepsilon}^2) \end{pmatrix}. \quad (49)$$

We now check that  $\lambda_{\max}((p'_{\delta,\sigma,\varepsilon})^1, D_{11}^2 w_1^\delta(p_{\delta,\sigma,\varepsilon})) < 0$ . For this, we come back to the definition of  $w_{2,\delta}$  and note that, for any  $p^1 \in \mathbf{R}^{|S^1|}$  and  $(p')^1 \in \Delta(S^1)$ ,

$$w_{2,\delta}(p^1, p_{\delta,\sigma,\varepsilon}^2) \leq w_2((p')^1, (p'_{\delta,\sigma,\varepsilon})^2) + \frac{1}{2\delta} (|p^1 - (p')^1|^2 + |(p_{\delta,\sigma,\varepsilon})^2 - (p'_{\delta,\sigma,\varepsilon})^2|^2),$$

with an equality at  $(p^1, (p')^1) = ((p_{\delta,\sigma,\varepsilon})^1, (p'_{\delta,\sigma,\varepsilon})^1)$ . If  $z \in T_{\Delta(S^2)}(p'_{\delta,\sigma,\varepsilon})^1$  with  $|z|$  small enough, taking  $p^1 := (p_{\delta,\sigma,\varepsilon})^1 + z$  and  $(p')^1 = (p'_{\delta,\sigma,\varepsilon})^1 + z$  gives

$$\begin{aligned} w_{2,\delta}((p_{\delta,\sigma,\varepsilon})^1 + z, p_{\delta,\sigma,\varepsilon}^2) &\leq \\ w_2((p'_{\delta,\sigma,\varepsilon})^1 + z, (p'_{\delta,\sigma,\varepsilon})^2) &+ \frac{1}{2\delta} (|(p_{\delta,\sigma,\varepsilon})^1 - (p'_{\delta,\sigma,\varepsilon})^1|^2 + |(p_{\delta,\sigma,\varepsilon})^2 - (p'_{\delta,\sigma,\varepsilon})^2|^2), \end{aligned}$$

with equality for  $z = 0$ . As  $w_2$  is concave with respect to the first variable, the above inequality implies that  $\lambda_{\max}((p'_{\delta,\sigma,\varepsilon})^1, D_{11}^2 w_{2,\delta}(p_{\delta,\sigma,\varepsilon})) \leq 0$ . In view of (49) we get therefore

$$\lambda_{\max}((p'_{\delta,\sigma,\varepsilon})^1, D_{11}^2 w_1^\delta(p_{\delta,\sigma,\varepsilon})) \leq -\sigma \lambda_{\min}((p'_{\delta,\sigma,\varepsilon})^1, D_{1,1}^2 \xi_1(p_{\delta,\sigma,\varepsilon}^1)) < 0$$

because  $D_{1,1}^2 \xi_1 > 0$  by construction. One can check in the same way that

$$\lambda_{\min}((p''_{\delta,\sigma,\varepsilon})^2, D_{22}^2 w_{2,\delta}(p_{\delta,\sigma,\varepsilon})) > 0.$$

So (45) and (46) become

$$r w_1^\delta(p_{\delta,\sigma,\varepsilon}) + H(p'_{\delta,\sigma,\varepsilon}, Dw_1^\delta(p_{\delta,\sigma,\varepsilon})) \leq 0 \quad (50)$$

and

$$r w_{2,\delta}(p_{\delta,\sigma,\varepsilon}) + H(p''_{\delta,\sigma,\varepsilon}, Dw_{2,\delta}(p_{\delta,\sigma,\varepsilon})) \geq 0. \quad (51)$$

We compute the difference of the two inequalities:

$$r(w_1^\delta(p_{\delta,\sigma,\varepsilon}) - w_{2,\delta}(p_{\delta,\sigma,\varepsilon})) + H(p'_{\delta,\sigma,\varepsilon}, Dw_1^\delta(p_{\delta,\sigma,\varepsilon})) - H(p''_{\delta,\sigma,\varepsilon}, Dw_{2,\delta}(p_{\delta,\sigma,\varepsilon})) \leq 0,$$

where, in view of assumption (28) and (47),

$$H(p'_{\delta,\sigma,\varepsilon}, Dw_1^\delta(p_{\delta,\sigma,\varepsilon})) \geq H(p_{\delta,\sigma,\varepsilon}, Dw_1^\delta(p_{\delta,\sigma,\varepsilon})) - \frac{C}{\delta} |p_{\delta,\sigma,\varepsilon} - p'_{\delta,\sigma,\varepsilon}|^2$$

while

$$H(p''_{\delta,\sigma,\varepsilon}, Dw_{2,\delta}(p_{\delta,\sigma,\varepsilon})) \leq H(p_{\delta,\sigma,\varepsilon}, Dw_{2,\delta}(p_{\delta,\sigma,\varepsilon})) + \frac{C}{\delta} |p_{\delta,\sigma,\varepsilon} - p''_{\delta,\sigma,\varepsilon}|^2.$$

So

$$\begin{aligned} r(w_1^\delta(p_{\delta,\sigma,\varepsilon}) - w_{2,\delta}(p_{\delta,\sigma,\varepsilon})) + H(p_{\delta,\sigma,\varepsilon}, Dw_1^\delta(p_{\delta,\sigma,\varepsilon})) - H(p_{\delta,\sigma,\varepsilon}, Dw_{2,\delta}(p_{\delta,\sigma,\varepsilon})) \\ \leq \frac{C}{\delta} \left( |p_{\delta,\sigma,\varepsilon} - p'_{\delta,\sigma,\varepsilon}|^2 + |p_{\delta,\sigma,\varepsilon} - p''_{\delta,\sigma,\varepsilon}|^2 \right). \end{aligned}$$

We now use assumption (27) on  $H$  combined with (48) to deduce:

$$r(w_1^\delta(p_{\delta,\sigma,\varepsilon}) - w_{2,\delta}(p_{\delta,\sigma,\varepsilon})) \leq \frac{C}{\delta} \left( |p_{\delta,\sigma,\varepsilon} - p'_{\delta,\sigma,\varepsilon}|^2 + |p_{\delta,\sigma,\varepsilon} - p''_{\delta,\sigma,\varepsilon}|^2 \right) + C(\varepsilon + \sigma), \quad (52)$$

since  $D\xi_1$  and  $D\xi_2$  are bounded. As  $\sigma$  and  $\varepsilon$  tend to 0, the  $p_{\delta,\sigma,\varepsilon}$ ,  $p'_{\delta,\sigma,\varepsilon}$  and  $p''_{\delta,\sigma,\varepsilon}$  converges (up to a subsequence) to  $p_\delta$ ,  $p'_\delta$  and  $p''_\delta$ , where  $p_\delta$  is a maximum in (40) and where  $p'_\delta$  and  $p''_\delta$  satisfy (41). Moreover (52) implies that

$$rM_\delta = r(w_1^\delta(p_\delta) - w_{2,\delta}(p_\delta)) \leq \frac{C}{\delta} \left( |p_\delta - p'_\delta|^2 + |p_\delta - p''_\delta|^2 \right).$$

We finally let  $\delta \rightarrow 0$ : in view of Lemma 7 the above inequality yields to  $M = \lim_{\delta \rightarrow 0} M_\delta \leq 0$ , which contradicts our initial assumption. Therefore  $w_1 \leq w_2$  and the proof is complete.

**Acknowledgments.** We thank Fabien Gensbittel for fruitful discussions.

This work has been partially supported by the French National Research Agency ANR-10-BLAN 0112.

## References

- [1] Abreu, D., P. Milgrom and D. Pearce (1991), Information and timing in repeated partnerships, *Econometrica*, **59**, 1713–1733.
- [2] Athey, S. and K. Bagwell (2008), Collusion with Persistent Cost Shocks, *Econometrica*, **76**(3), 493–540.

- [3] Aumann, R.J. and M.B. Maschler (1995), *Repeated Games with Incomplete Information*, The MIT Press.
- [4] Cardaliaguet P., Laraki R. and Sorin S. (2012) A Continuous Time Approach for the Asymptotic Value in Two-Person Zero-Sum Repeated Games. *SIAM Cont. Optimization*, **50**, 3, 1573–1596.
- [5] Cardaliaguet P. (2009), A double obstacle problem arising in differential game theory. *J. Math. Anal. Appl.* **360**, no. 1, 95–107.
- [6] Cardaliaguet P. and Rainer C. (2009) Stochastic differential games with asymmetric information. *Appl. Math. Optim.* **59** 1-36.
- [7] Cardaliaguet P.; Rainer C. (2009) On a continuous time game with incomplete information. *Math. Oper. Res.* **34**, 4, 769–794.
- [8] M. G. Crandall, H. Ishii, and P.-L. Lions (1992) User’s guide to viscosity solutions of second order partial differential equations. *Bull. Amer. Math. Soc. (N.S.)*, **27**(1):1–67.
- [9] Evans, L.C.; Souganidis, P.E. (1984) Differential games and representation formulas for solutions of Hamilton-Jacobi Equations. *Indiana Univ. Math. J.*, **282**, 487–502.
- [10] Fudenberg, D. and D.K. Levine (2007), Continuous time limits of repeated games with imperfect public monitoring, *Review of Economic Dynamics*, **10**, 173–192.
- [11] Gensbittel F. and Renault J. (2012) The value of Markov Chain Games with incomplete information on both sides. hal-00745986.
- [12] Grün, C. (2012) A probabilistic numerical scheme for stochastic differential games with incomplete information, *AMO*, **66**, 363–385.
- [13] Jacod, J. and A. Shiryaev (2003) Limit theorems for stochastic processes. Springer.
- [14] Hörner, J., D. Rosenberg, E. Solan and N. Vieille (2010). “On a Markov Game with One-Sided Incomplete Information,” *Operations Research*, **58**, 1107–1115.
- [15] Hörner, J., T. Sugaya, S. Takahashi and N. Vieille (2011). “Recursive Methods in Discounted Stochastic Games: An Algorithm for  $\delta \rightarrow 1$  and a Folk Theorem,” *Econometrica*, **79**, 1277–1318.
- [16] Laraki R. (2002), Repeated games with lack of information on one side: The dual differential approach, *Math. Oper. Res.*, **27**, 419440.

- [17] Lehrer E. and D. Rosenberg (2003), Information and its value in zero-sum repeated games, mimeo.
- [18] Lehrer E. and D. Rosenberg (2010), A note on the evaluation of information in zero-sum repeated games, *Journal of Mathematical Economics*, **46**, 393–399.
- [19] Mailath, G. and L. Samuelson (2001), Who Wants a Good Reputation?, *Review of Economic Studies*, **68**, 415–441.
- [20] Mailath, G. and L. Samuelson (2006), *Repeated Games and Reputations: Long-Run Relationships*, Oxford University Press.
- [21] Neyman, A. (2008), Existence of Optimal Strategies in Markov Games with Incomplete Information, *International Journal of Game Theory*, **37**, 518–596.
- [22] Peski, M. and T. Wiseman (2012). A Folk Theorem for Stochastic Games with Infrequent State Changes, working paper, University of Toronto.
- [23] Phelan, C. (2006), Public Trust and Government Betrayal, *Journal of Economic Theory*, **127**(1), 27–43.
- [24] Protter, P. (2005). *Stochastic Integration and Differential Equations*. Springer-Verlag
- [25] Renault, J. (2006), The Value of Markov Chain Games with Lack of Information on One Side, *Mathematics of Operations Research*, **31**, 490–512.
- [26] Renault, J., E. Solan and N. Vieille (2013). “Dynamic Sender-Receiver Games,” *Journal of Economic Theory*, forthcoming.
- [27] Sorin, S (2002). *A first course on zero-sum repeated games*, Springer.
- [28] Vieille N. (1992). Weak approachability, *Math. Oper. Res.*, **17**, 781–791.
- [29] Wiseman, T. (2008), Reputation and Impermanent Types, *Games and Economic Behavior*, **62**, 190–210.