Existence of the uniform value in repeated games with a more informed controller

November 29th 2012

Fabien Gensbittel[*], Miquel Oliu-Barton[†], Xavier Venel[‡]

Abstract

We prove that in a general zero-sum repeated game where the first player is more informed than the second player and controls the evolution of information on the state, the uniform value exists. This result extends previous results on Markov decision processes with partial observation (Rosenberg, Solan, Vieille [11]), and repeated games with an informed controller (Renault [10]). Our formal definition of a more informed player is more general than the inclusion of signals, allowing therefore for imperfect monitoring of actions. We construct an auxiliary stochastic game whose state space is the set of second order beliefs of player 2 (beliefs about beliefs of player 1 on the true state variable of the initial game) with perfect monitoring and we prove it has a value by using a result of Renault [10]. A key element in this work is to prove that player 1 can use strategies of the auxiliary game in the initial game in our general framework, which allows to deduce that the value of the auxiliary game is also the value of our initial repeated game by using classical arguments.

---

[*]TSE (GREMAQ, Université Toulouse 1 Capitole), France. fabien.gensbittel@tse-fr.eu
[†]Equipe Combinatoire et Optimisation, Université Paris VI, France. miquel.oliu.barton@normalesup.org
[‡]School of Mathematical Sciences, Tel Aviv University, Tel Aviv 69978, Israel. xavier.venel@sip.univ-tlse1.fr

Contents

1 Introduction

Zero-sum repeated games with incomplete information were introduced by Aumann and Maschler in 1966 [1] in order to study repeated interactions between two players having a different information. The authors also introduced a notion of value for these games usually called uniform value and proved its existence for games with incomplete information on one side. Mertens and Neyman [4] proved that the uniform value exists for finite stochastic games and several works were devoted since then to prove the existence of the uniform value for some subclasses of the general model of repeated games. Recently, Renault proved in [10] that the uniform value exists in repeated games with an informed controller using an approach based on an existence result for dynamic programming problems (Renault, [9]). The existence theorem in [10] requires that the first player observes the state variable at each stage and controls and observes the evolution of the beliefs of the second player on the state variable.

In the present work, we prove that the uniform value exists in the class of repeated games with a more informed controller. Our existence result requires that the first player is more informed about the state variable than the second player and also that he controls the evolution of beliefs of the second player. A weaker version of our result was conjectured in the conclusion of [10], and it was suggested that the proof may be based on an auxiliary game whose state space would be the pair of beliefs of both players about the original state variable. We show that the analysis requires actually to introduce an auxiliary game whose state space is the set of

second order beliefs of the less informed player and provide a set of weaker assumptions than those suggested in [10], allowing to deal with imperfect monitoring of actions.

The paper is organized as follows: In section 2, we describe the general model of repeated games and introduce three assumptions that formalize the notion of a more informed controller. In section 3, we check that several models previously studied in the literature satisfy these three assumptions. Section 4 is dedicated to a discussion of the assumptions and a precise study of their implications. In addition, we provide a second version of the theorem with stronger, but easier to check, assumptions. The last section 5 is dedicated to the proof of existence of the uniform value. We introduce there an auxiliary stochastic game with perfect monitoring on an auxiliary state variable which represents the beliefs of player 2 about the beliefs of player 1 about the state variable of the original game. We prove that this auxiliary game has a uniform value using the main theorem of Renault [10] and that player 1 can use optimal strategies in this auxiliary stochastic game in order to play optimally in the original repeated game. Finally, we prove that player 2 can also guarantee this value by playing by blocks, so that both games have a uniform value and these values are equal.

2 Model

2.1 General definitions and notation

For any metric space X, let $\Delta(X)$ denote the set of Borel probability distributions on X. If X is a finite set (endowed with the discrete metric) of cardinal $|X|$, then $\Delta(X)$ is precisely the $|X|$-dimensional simplex. $\Delta_f(X) \subset \Delta(X)$ denotes the probability distributions supported on a finite subset of X and δ_x denotes the Dirac measure on $x \in X$.

A zero-sum repeated game is described by a 8-tuple $(K, I, J, g, C, D, \pi, q)$, where K is the state space, I and J are the action sets for player 1 and 2 respectively, g is a payoff function $g : K \times I \times J \to [0, 1]$, and C and D are the signal sets for player 1 and 2 respectively. $\pi \in \Delta_f(K \times \mathbb{N} \times \mathbb{N})$ denotes the initial probability and $q : K \times I \times J \to \Delta(K \times C \times D)$ denotes the transition function.

The game is played as follows: At the beginning of the game, the triple (k_1, c_1, d_1) is chosen according to the initial probability distribution $\pi \in \Delta_f(K \times \mathbb{N} \times \mathbb{N})$. For each stage $m \geq 1$, player 1 observes the signal c_m and player 2 observes the signal d_m. Then both players choose actions $(i_m, j_m) \in I \times J$ based on their own past actions and on the sequence of signals they observed (i.e. we assume perfect recall). Given the state k_m and the actions (i_m, j_m), a new triple $(k_{m+1}, c_{m+1}, d_{m+1}) \in K \times C \times D$ is chosen according to the probability distribution $q(k_m, i_m, j_m)$. The payoff for stage m of player 1 is $g(k_m, i_m, j_m)$ and the game proceeds to stage $m + 1$. The stage payoffs are not directly observed by the players and cannot be deduced, in general, from their observations. The sets of initial signals can be any finite subset of $\mathbb{N}$. This generalization is for technical reasons only. Indeed, it will very convenient in the sequel to consider this possibly larger set of initial signals in order to have a simple way to deal with the recursive structure of the game.

The information held by player 1 before his play at stage m, called player 1's private history, is given by

$$h_m^I \triangleq (c_1, i_1, \ldots, c_{m-1}, i_{m-1}, c_m) \in \mathbb{N} \times (I \times C)^{m-1}.$$

Similarly, the information held by player 2 is represented by

$$h_m^{II} \triangleq (d_1, j_1, \ldots, d_{m-1}, j_{m-1}, d_m) \in \mathbb{N} \times (J \times D)^{m-1}.$$

Let $\mathcal{H}^I$ (resp. $\mathcal{H}^{II}$) denote the set of all finite private histories for player 1 (resp. of all private histories for player 2). We assume the sets K, I, J, C and D are all finite and that the description of the model is common knowledge.

Instead of $\mathbb{N}$, the sets of initial signals will often be denoted by C' and D' where C' and D' are finite subsets of $\mathbb{N}$. We will also write abusively that $\pi \in \Delta(K \times C' \times D')$. The initial signals will still be denoted by (c_1, d_1). Reciprocally, given finite sets C' and D', any $\pi \in \Delta(K \times C' \times D')$ can be seen as an element of $\Delta_f(K \times \mathbb{N} \times \mathbb{N})$ using some enumerations of C' and D'. The main advantage is that any couple of finite private histories can be embedded in $\mathbb{N} \times \mathbb{N}$ via some enumerations. This advantage will become clear in the proof.

Strategies   A behavior strategy for player 1 is a map from private histories $\mathcal{H}^I$ to probabilities over I . The set of behavior strategies of player 1 is denoted by Σ. Every strategy $\sigma \in \Sigma$ corresponds to a sequence $\{\sigma_m\}_{m \geq 1}$, where σ_m is defined on the set of histories up to stage m. That is,

$$\sigma_m : \mathbb{N} \times (I \times C)^{m-1} \to \Delta(I).$$

Similarly, a behavior strategy τ for player 2 is a map from private histories $\mathcal{H}^{II}$ to probability distributions over J. The set of behavior strategies of player 2 is denoted by $\mathcal{T}$. Any τ corresponds to a sequence $\{\tau_m\}_{m \geq 1}$, with

$$\tau_m : \mathbb{N} \times (J \times D)^{m-1} \to \Delta(J).$$

The initial distribution π, the transition function q and a behavior strategy profile $(\sigma, \tau) \in \Sigma \times \mathcal{T}$ induce a unique probability distribution over the set of plays $K \times \mathbb{N} \times \mathbb{N} \times (K \times C \times D \times I \times J)^\infty$, denoted by $\mathbb{P}_{\sigma\tau}^\pi$. Let $\mathbb{E}_{\sigma\tau}^\pi = \mathbb{E}_{\mathbb{P}_{\sigma\tau}^\pi}$ denote the expectation with respect to the probability $\mathbb{P}_{\sigma\tau}^\pi$.

Evaluations of the payoff   A second component of the model is the way in which the total payoff of player 1 is evaluated, in terms of the sequence of stage payoffs $\{g(k_m, i_m, j_m)\}_{m \geq 1}$. The two classical evaluations correspond to the n-stage game and the λ-discounted game. In the former, the payoff function is the expected Cesàro mean of the stage payoffs of the n first stages, i.e.

$$\gamma_n(\pi, \sigma, \tau) = \mathbb{E}_{\sigma\tau}^\pi \Big[\frac{1}{n} \sum_{m=1}^n g(k_m, i_m, j_m) \Big].$$

In the latter, the payoff is taken as the expected Abel sum, with respect to the discount factor $0 < \lambda \leq 1$, i.e.

$$\gamma_\lambda(\pi, \sigma, \tau) = \mathbb{E}_{\sigma\tau}^\pi \Big[\lambda \sum_{m \geq 1} (1-\lambda)^{m-1} g(k_m, i_m, j_m) \Big].$$

More generally, one may consider any *compact evaluation*. That is, for any $\theta \in \Delta(\mathbb{N}^*)$, let

$$\gamma_\theta(\pi, \sigma, \tau) \quad = \quad \mathbb{E}_{\sigma\tau}^\pi \Big[\sum_{m \geq 1} \theta_m g(k_m, i_m, j_m) \Big]. \tag{2.1}$$

Denote by $\Gamma_\theta(\pi)$ the 8-tuple defined above together with the θ-evaluation.

The value function For any $\pi \in \Delta_f(K \times \mathbb{N} \times \mathbb{N})$ and any $\theta \in \Delta(\mathbb{N}^*)$, $\Gamma_\theta(\pi)$ is known to have a value, denoted by $v_\theta(\pi)$. It satisfies

$$v_\theta(\pi) = \sup_{\sigma \in \Sigma} \inf_{\tau \in \mathcal{T}} \gamma_\theta(\pi, \sigma, \tau) = \inf_{\tau \in \mathcal{T}} \sup_{\sigma \in \Sigma} \gamma_\theta(\pi, \sigma, \tau).$$

Remark 2.1. *These general evaluations will be used in section [5](), and we will only need to consider probabilities with finite support (i.e. $\theta \in \Delta_f(\mathbb{N}^*)$).*

Let $\sigma \in \Sigma$ be a behavior strategy and let $h \in \mathcal{H}^I$ be some finite private history of player 1. We denote by $\sigma(h)$ the behavior strategy of player 1 after the history h. Equivalently, $\sigma(h)$ is the restriction of the map σ to the subset of histories beginning with h. In particular, given some strategy profile $(\sigma, \tau) \in \Sigma \times \mathcal{T}$ and two signals $(c, d) \in C' \times D'$, consider the profile $(\sigma(c), \tau(d))$. It may be interpreted as a strategy profile in a game in which the players have no initial signals. More formally, for any $p \in \Delta(K)$, we will use the following notation:

$$\gamma_\theta(k, \sigma(c), \tau(d)) \triangleq \gamma_\theta(\delta_{(k,c,d)}, \sigma, \tau), \text{ and } \quad \gamma_\theta(p, \sigma(c), \tau(d)) \triangleq \gamma_\theta(p \otimes \delta_{(c,d)}, \sigma, \tau).$$

With this notation, the payoff can be written as

$$\gamma_\theta(\pi, \sigma, \tau) = \mathbb{E}_\pi[\gamma_\theta(k, \sigma(c), \tau(d))] = \sum_{(k,c,d) \in K \times C' \times D'} \gamma_\theta(k, \sigma(c), \tau(d)) \pi(k, c, d).$$

Alternatively, one can consider the game as having *per se* infinitely many stages.

Uniform value The infinitely repeated game is denoted by $\Gamma_\infty(\pi)$. Let us present here some important definitions relative to the game $\Gamma_\infty(\pi)$. Its value will be called the *uniform value* and denoted by $v_\infty(\pi)$.

Definition 2.2. *Let v be a real number,*

- *Player 1 can guarantee v in $\Gamma_\infty(\pi)$ if for any $\epsilon > 0$ there exists a strategy $\sigma \in \Sigma$ of player 1 and an integer $N \in \mathbb{N}$, such that*

$$\forall n \geq N, \ \forall \tau \in \mathcal{T}, \ \gamma_n(\pi, \sigma, \tau) \geq v - \epsilon.$$

We say that such a strategy σ guarantees $v - \epsilon$ in $\Gamma_\infty(\pi)$ and define

$$\underline{v}_\infty(\pi) \triangleq \sup\{v \in \mathbb{R} \mid player\ 1\ can\ garantee\ v\}.$$

- *Player 2 can guarantee v in $\Gamma_\infty(\pi)$ if for any $\epsilon > 0$ there exists a strategy $\tau \in \mathcal{T}$ of player 2 and an integer $N \in \mathbb{N}$, such that*

$$\forall n \geq N, \ \forall \sigma \in \Sigma, \ \gamma_n(\pi, \sigma, \tau) \leq v + \epsilon.$$

We say that such a strategy τ guarantees $v + \epsilon$ in $\Gamma_\infty(\pi)$ and define

$$\overline{v}_\infty(\pi) \triangleq \inf\{v \in \mathbb{R} \mid player\ 2\ can\ garantee\ v\}.$$

- *If $\underline{v}_\infty(\pi) = \overline{v}_\infty(\pi)$ the uniform value exists and we denote by $v_\infty(\pi)$ the common value.*

The existence of a uniform value v_∞ is stronger than the existence of a limit value (or asymptotic value), in the sense that it implies (see e.g. Neyman and Sorin [6, Theorem 1] for more general evaluations)

$$\lim_{\lambda \to 0} v_\lambda = \lim_{n \to \infty} v_n = v_\infty.$$

2.2 Model with a more informed controller

We will consider a particular class of the general model presented above, which generalizes both the class of repeated games considered by Renault [10] and the model of Partially Observable Markov Decision Processes (see section 3.1). As usual in games with incomplete information, we call belief of player 1 at stage m about some random variable ξ the conditional law of ξ given the information held by player 1 at stage m. In the sequel, the first order beliefs of a player denote beliefs about the state variable k, and second order beliefs of a player denote beliefs about the first order beliefs of his opponent.

We assume the following three hypotheses at every stage m of the game:

$(a1)$ Player 1's first order belief is more accurate than player 2's first order belief.

$(a2)$ Player 1 can compute the second order beliefs of player 2.

$(a3)$ Player 1 controls the evolution of second order beliefs of player 2.

The main result of this paper is to establish the existence of the uniform value under these assumptions. Let the formal transcription of $(a1) - (a3)$ defined below be denoted by $(A1) - (A3)$.

Theorem 2.3. *Let Γ be a repeated game with a more informed controller, i.e. such that assumptions $(A1)$, $(A2)$, $(A3)$ hold. Then the uniform value exists.*

Remark 2.4. *It was already pointed out in the literature (see e.g. Mertens [3]) that in games with a more informed player, the analysis of beliefs can be restricted to second order beliefs of the less informed player. In this work, the definition of more informed is slightly more general than the inclusion of signals and a similar reduction is made formally in Lemma 4.4.*

2.3 Formal assumptions

Let us present here a rigorous transcription of the informal assumptions $(a1) - (a3)$. In the next section, we will present some of the models which satisfy our three assumptions and to which, consequently, Theorem 2.3 applies.

Let us start with some notations:

Given some probability distribution $\mu \in \Delta_f(X \times Y)$ over a product, we denote by

$$\mu(x) \triangleq \sum_{y \in Y} \mu(x, y).$$

For any random variable ξ defined on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ and $\mathcal{F}$ a sub σ-algebra of $\mathcal{A}$, let $\mathcal{L}_{\mathbb{P}}(\xi \mid \mathcal{F})$ denote the conditional distribution of ξ given $\mathcal{F}$, which is seen as a $\mathcal{F}$-measurable random variable[1] and let $\mathcal{L}_{\mathbb{P}}(\xi)$ denote the distribution of ξ.

In the sequel, both functions g and q are linearly extended to $\Delta(K) \times I \times J$.

Assumption $(a1)$ can now be formalized as follows:

---

[1] All random variables appearing here take only finitely many values so that the definition of conditional laws does not require any additional care about measurability.

$(A1)$ $\forall n \geq 1$, $\forall (\sigma, \tau) \in \Sigma \times \mathcal{T}$, $\quad \mathcal{L}_{\mathbb{P}^{\pi}_{\sigma\tau}}(k_n \mid h_n^I, h_n^{II}) = \mathcal{L}_{\mathbb{P}^{\pi}_{\sigma\tau}}(k_n \mid h_n^I)$.

In words, at every stage and given any strategy profile, player 2's information does not contain any information about the state variable that is not already contained in player 1's information. Assumption $(A1)$ is equivalent to the conditional independence of k_n and h_n^{II}, given h_n^I, under the probability $\mathbb{P}^{\pi}_{\sigma\tau}$.

For $n = 1$, this equation does not depend on σ and τ and it can be reformulated as

$(A1a)$ $\pi(c)\pi(k, c, d) = \pi(k, c)\pi(c, d)$, $\forall (k, c, d) \in K \times C' \times D'$.

In order to model the players' information about the state variable at stage n, we need to define three variables x_n, y_n and η_n. Before choosing their first action, the players receive signals $(c_1, d_1) \in C' \times D'$. The (random) variable

$$x_1 \triangleq \mathcal{L}_\pi(k_1 \mid c_1) \in \Delta(K)$$

represents the *first order beliefs of player* 1 about the initial state. Let $x_1(c_1) \in \Delta(K)$ denote its realization, i.e. the beliefs of player 1 once he has received the signal $c_1 \in C'$. Thus, $x_1(c_1) = \mathcal{L}_\pi(k_1|c_1)$ and each signal $c_1 \in C'$ occurs with probability $\pi(c_1)$, so that $\mathcal{L}_\pi(x_1) = \sum_{c_1 \in C'} \pi(c_1)\delta_{x_1(c_1)}$. Similarly, define the *second order beliefs of player* 2, i.e. beliefs about player 1's beliefs about the initial state

$$y_1 \triangleq \mathcal{L}_\pi(x_1|d_1) \in \Delta_f(\Delta(K)).$$

With probability $\pi(d_1)$, player 2's beliefs about player 1's beliefs (about the state variable) are distributed as follows:

$$y_1(d_1) = \sum_{c_1 \in C'} \pi(c_1|d_1)\delta_{x_1(c_1)} \in \Delta_f(\Delta(K)),$$

with a slight abuse of notations since we write $\pi(c_1|d_1)$ instead of $\pi(c_1 = c|d_1)$ with a sum over $c \in C'$. Finally, let η_1 be the distribution of the second order beliefs of player 2

$$\eta_1 \triangleq \mathcal{L}_\pi(y_1) = \sum_{d_1 \in D'} \pi(d_1)\delta_{y(d_1)} \in \Delta_f(\Delta_f(\Delta(K))).$$

Notice that the Dirac measures involved in the definition of $\mathcal{L}_\pi(x_1)$ or of $\mathcal{L}_\pi(y_1)$ refer to different spaces: the former refers to $\Delta(K)$, the latter to $\Delta_f(\Delta(K))$.

More generally, for some fixed strategy profile $(\sigma, \tau) \in \Sigma \times \mathcal{T}$, let us denote the first order beliefs of player 1 at stage n by $x_n \in \Delta(K)$, the second order beliefs of player 2 at stage n by $y_n \in \Delta_f(\Delta(K))$, and the distribution of y_n by η_n.

Definition 2.5. *Put* $x_n \triangleq \mathcal{L}_{\mathbb{P}^{\pi}_{\sigma\tau}}(k_n \mid h_n^I)$, $y_n \triangleq \mathcal{L}_{\mathbb{P}^{\pi}_{\sigma\tau}}(x_n|h_n^{II})$, *and* $\eta_n \triangleq \mathcal{L}_{\mathbb{P}^{\pi}_{\sigma\tau}}(y_n)$.

Let us illustrate these definitions through the following example.

Example 2.6. *Let* $K = \{k_1, k_2\}$ *be set of states space,* $U = \{u_1, u_2\}$ *a set of public signals and* $S = \{s_1, s_2, s_3\}$ *a set of private signals for player* 1. *Using the notations above, let* $C = U \times S$

(resp. $D = U$) be the set of signals for player 2 (resp. 2). We consider $\pi \in \Delta(K \times S \times U)$ defined by

$$
\begin{array}{cc}
\begin{array}{c} \end{array} & \begin{array}{cc} u_1 & u_2 \end{array} \\
\begin{array}{c} s_1 \\ s_2 \\ s_3 \end{array} & \begin{pmatrix} \frac{8}{24} & 0 \\ \frac{1}{24} & \frac{3}{24} \\ 0 & 0 \end{pmatrix}
\end{array}
\qquad
\begin{array}{cc}
\begin{array}{c} \end{array} & \begin{array}{cc} u_1 & u_2 \end{array} \\
\begin{array}{c} s_1 \\ s_2 \\ s_3 \end{array} & \begin{pmatrix} 0 & 0 \\ \frac{1}{24} & \frac{3}{24} \\ \frac{2}{24} & \frac{6}{24} \end{pmatrix}
\end{array}.
$$
$$
\qquad\quad k_1 \qquad\qquad\qquad\qquad k_2
$$

It is more convenient here to use $K \times S \times U$ but π can be understood as a probability on $K \times C \times D$. To simplify notations, let us identify $\Delta(K)$ with $[0,1]$ with the convention that $p \in \Delta(K)$ is identified with $p(k_1)$. If player 1 receives signal (s_1, u_1), then $x_1 = 1$. If he receives (s_2, u_1) or (s_2, u_2), then $x_1 = \frac{1}{2}$. Finally, if he receives (s_3, u_1) or (s_3, u_2), then $x_1 = 0$. The value of x_1 depends only on his private signal.

We now compute the second order beliefs of player 2. If player 2 receives u_1 then his beliefs about the private signal of player 1 are $\frac{8}{24}\delta_{s_1} + \frac{2}{24}\delta_{s_2} + \frac{2}{24}\delta_{s_3}$, so that

$$
y_1(u_1) = \frac{8}{24}\delta_1 + \frac{2}{24}\delta_{\frac{1}{2}} + \frac{2}{24}\delta_0.
$$

If player 2 receives u_2, then we obtain

$$
y_1(u_2) = \frac{6}{24}\delta_{\frac{1}{2}} + \frac{6}{24}\delta_0.
$$

To conclude, player 2 receives each signal with probability $\frac{1}{2}$, so that η_1 is equal to

$$
\eta_1 = \frac{1}{2}\delta_{\frac{8}{24}\delta_1 + \frac{2}{24}\delta_{\frac{1}{2}} + \frac{2}{24}\delta_0} + \frac{1}{2}\delta_{\frac{6}{24}\delta_{\frac{1}{2}} + \frac{6}{24}\delta_0}.
$$

Assumption $(a2)$ will be split in two parts $(A2a)$ and $(A2b)$. At first, we assume that player 1 is able to compute the variable y_1, which is a constraint on the initial probability π only.

$(A2a)$ There exists a map $f_1 = f_1^\pi : C' \to \Delta(\Delta(K))$ such that $y_1 = f_1(c_1)$, π-almost surely.

Assuming $(A2a)$, we can introduce a special class of strategies for player 1 which will be needed for the second part of the formal assumption.

Definition 2.7. If $\pi \in \Delta_f(K \times \mathbb{N} \times \mathbb{N})$ fulfills $(A2a)$, a strategy $\sigma \in \Sigma$ is called a reduced strategy if it depends on the initial signal c_1 in C' only through (x_1, y_1). Let $\Sigma'(\pi) \subset \Sigma$ denote the subset of reduced strategies.

The second part of $(A2)$ requires that when player 1 is using a reduced strategy, the variable y_2 has to be h_2^I-measurable. Formally:

$(A2b)$ $\forall \pi \in \Delta_f(K \times \mathbb{N} \times \mathbb{N})$ satisfying $(A2a)$, $\forall \sigma \in \Sigma'(\pi), \forall \tau \in \mathcal{T}$, $\exists f_2 = f_2^{\pi,\sigma,\tau} : \mathbb{N} \times I \times C \to \Delta(\Delta(K))$ such that $y_2 = f_2(c_1, i_1, c_2)$, $\mathbb{P}_{\sigma,\tau}^\pi$-almost surely.

The introduction of reduced strategies for player 1 is necessary in order to exclude non relevant correlations between players (see example 4.1 in section 4). It will be shown in Lemma 4.4 and in the proof of the main Theorem that there is no loss in restricting player 1 to reduced strategies.

In order to state the last assumption, we reduce the set of initial probabilities.

Definition 2.8. *Let $\Delta_f^*(K \times \mathbb{N} \times \mathbb{N})$ be the set of probability distributions satisfying $(A1a)$ and $(A2a)$.*

Assumption $(a3)$ can now be formalized as

$(A3)$ $\forall \pi \in \Delta_f^*(K \times \mathbb{N} \times \mathbb{N})$, $\forall \sigma \in \Sigma'(\pi)$, η_2 is independent of $\tau \in \mathcal{T}$.

Remark 2.9. *Assumptions $(A1, A2)$ imply that the properties $(A1a)$ and $(A2a)$ of the initial probability π are preserved by the transition when player 1 plays reduced strategies. Precisely, for all (σ, τ) with σ reduced, the law of (k_2, h_2^I, h_2^{II}) under $\mathbb{P}_{\sigma\tau}^\pi$, seen as an element of $\Delta_f(K \times \mathbb{N} \times \mathbb{N})$, belongs to the set $\Delta_f^*(K \times \mathbb{N} \times \mathbb{N})$. We will prove in section 4 that even if the two last assumptions $(A2b)$ and $(A3)$ are stated in terms of y_2 and η_2, it is possible to extend these properties by induction for y_n and η_n for appropriate strategies. Thus the formal assumptions are coherent with the informal assumptions. In particular, player 1 can compute the auxiliary variables (x_2, y_2, η_2) without knowing the strategy of player 2 and therefore play again a reduced strategy at the second stage (i.e. which depends only on (x_2, y_2)).*

3 Applications

We present in this section several models which satisfy our assumptions.

3.1 Partially Observable Markov Decision Processes.

A POMDP is a one-player game, given by a tuple (K, I, C, g, q, π), where K is the state space, I is the action set, C is the signals set, $g : K \times I \to [0, 1]$ is the payoff function, $q : K \times I \to \Delta(K \times C)$ is the transition function and π is an initial distribution on $K \times \mathbb{N}$. In the finite framework, the existence of the uniform value has been proven by Rosenberg, Solan and Vieille [11] and it was extended by Renault [9] to arbitrary set of actions and signals with the additional assumption that all the probabilities appearing in the transition or in the definition of strategies have finite support. We will only consider here the finite case.

Formally, a POMDP can be seen as a repeated game in which player 2 is dummy (i.e. his action set J is a singleton). Since player 2 has only one action, his information plays no role here. The assumptions $(A1) - (A3)$ hold obviously.

3.2 Repeated game with a perfectly informed controller.

The model of a repeated game with an informed controller introduced by Renault [10] fulfils our assumptions. In this model, player 1 is perfectly informed of the state and of the signal of player 2, in the sense that he can deduce the true state variable and the signal of player 2 from his signals. Moreover, the transition q is such that player 2 has no influence on the joint distribution of the pair made by the state variable and his signal.

In [10], the sets of initial signals are $C' = C$ and $D' = D$. Formally, the first assumption (i.e. that player 1 is perfectly informed of the state and of the signal of player 2) is given by

(HA') There exists two mappings $\hat{k} : C \to K$ and $\hat{d} : C \to D$ such that, if E denotes $\{(k, c, d) \in K \times C \times D, \hat{k}(c) = k, \hat{d}(c) = d\}$, then

$$\pi(E) = 1, \text{ and } q(k, i, j)(E) = 1, \ \forall (k, i, j) \in K \times I \times J.$$

The second assumption is formalized by

(HB') Player 1 controls the transition in the sense that the marginal of the transition q on $K \times D$ does not depend on player 2's action. For $k \in K$, $i \in I$, $j \in J$, we denote by $\overline{q}(k, i)$ the marginal of $q(k, i, j)$ on $K \times D$.

Let us check that this model satisfies our assumptions. Assuming that the initial distribution $\pi \in \Delta(K \times C \times D)$ fulfils assumption (HA'), we have

$$y_1(d_1) = \sum_{c_1 \in C} \pi(c_1 | d_1) \delta_{\mathcal{L}_\pi(k_1 | c_1, d_1)} = \sum_{c_1 \in C} \pi(c_1 | d_1) \delta_{\delta_{\hat{k}(c_1)}} = \sum_{k_1 \in K} \pi(k_1 | d_1) \delta_{\delta_{k_1}}.$$

We deduce that π can be seen as an element of $\Delta_f^*(K \times \mathbb{N} \times \mathbb{N})$. Formally, we have to verify our assumptions, starting from any initial distribution in $\Delta_f^*(K \times \mathbb{N} \times \mathbb{N})$. From now on, initial signals (c_1, d_1) belong to arbitrary finite subsets of $\mathbb{N}$ denoted by C', D' as in the previous section[2].

First, note that any stage $m \geq 2$, player 1's first order belief at each stage is a Dirac mass on the current state (i.e. $x_n = \delta_{k_n}$). Thus, adding the signal of player 2 to the signal of player 1 does not change the beliefs of the latter, which proves $(A1)$. It also implies that the second order beliefs of player 2 can be identified with the first order beliefs of player 2. Let $\pi \in \Delta_f^*(K \times \mathbb{N} \times \mathbb{N})$, τ be a strategy for player 2 and σ a reduced strategy for player 1. Recall that σ is function only of (x_1, y_1) which are by assumption c_1 measurable and that that y_1 is d_1-measurable, so that there exist two functions h_1 and f_1 such that, with probability 1, $\mathcal{L}_\pi(x_1 | d_1) = y_1 = h_1(d_1) = f_1(c_1)$. It follows that

$$\mathbb{P}_{\sigma\tau}^\pi(k_1, x_1, d_1, i_1, j_1, k_2, d_2) = \pi(d_1) h_1(d_1)(x_1) x_1(k_1) \sigma(x_1, h_1(d_1))(i_1) \tau(d_1)(j_1) \overline{q}(k_1, i_1)(k_2, d_2).$$

We deduce that

$$\begin{aligned}
y_2(d_1, j_1, d_2) &= \sum_{c_1, i_1, c_2} \mathbb{P}_{\sigma\tau}^\pi(c_1, i_1, c_2 | d_1, j_1, d_2) \delta_{\mathcal{L}_{\mathbb{P}_{\sigma\tau}^\pi}(k_2 | c_1, i_1, c_2)} \\
&= \sum_{c_1, i_1, c_2} \mathbb{P}_{\sigma\tau}^\pi(c_1, i_1, c_2 | d_1, j_1, d_2) \delta_{\delta_{\hat{k}(c_2)}} \\
&= \sum_{k_2} \mathbb{P}_{\sigma\tau}^\pi(k_2 | d_1, j_1, d_2) \delta_{\delta_{k_2}} \\
&= \frac{\sum_{k_1, x_1, k_2, i_1} \pi(d_1) h_1(d_1)(x_1) \sigma(x_1, h_1(d_1))(i_1) \overline{q}(k_1, i_1)(k_2, d_2) \delta_{\delta_{k_2}}}{\sum_{k_1', x_1', k_2', i_1'} \pi(d_1) y_1(d_1)(x_1') \sigma(x_1', y_1(d_1))(i_1') \overline{q}(k_1', i_1')(k_2', d_2)} \\
&= \frac{\sum_{k_1, x_1, k_2, i_1} f_1(c_1)(x_1) \sigma(x_1, f_1(c_1))(i_1) \overline{q}(k_1, i_1)(k_2, d_2) \delta_{\delta_{k_2}}}{\sum_{k_1', x_1', k_2', i_1'} f_1(c_1)(x_1') \sigma(x_1', f_1(c_1))(i_1') \overline{q}(k_1', i_1')(k_2', d_2)}
\end{aligned}$$

---

[2]One may easily reduce the analysis to a smaller set of initial probabilities, but we chose to keep this general formulation since the reduction does not really simplify the proofs.

where the above equalities hold almost surely whenever the conditional probabilities are well-defined. We deduce that y_2 is only a function of c_1 and d_2 which does not depend on τ, so that the function $f_2(c_1, c_2) = y_2(c_1, \hat{d}(c_2))$ proves that assumption $(A2)$ is satisfied. Finally the distribution of the random variable y_2 is equal to

$$\eta_2 = \sum_{c_1, d_2} \mathbb{P}^\pi_{\sigma\tau}(c_1, d_2) \delta_{y_2(c_1, d_2)}.$$

Using the preceding result and that the marginal of q on D does not depend on the action played by player 2, η_2 does not depend on τ and $(A3)$ is satisfied.

Remark 3.1. *In a previous work, Renault [8] studied the particular case where the state follows a Markov chain $f : K \to \Delta(K)$, player 1 observes the state and both players observe the actions. This is easily seen as a particular case of the above model. In a more recent work, Neyman [7] proved the existence of the uniform value when allowing for any signalling structure on the actions. This last result is not covered by our main theorem since in this case, player 1 cannot control player 2's information about the state variable.*

3.3 Player 1 is more informed about the state.

In this last paragraph, we assume that actions are observed by both players after each stage. Moreover, both players receive a public signal in a set U, player 1 receives a signal in a set S and player 2 has influence on the joint distribution of the state and signals in $K \times U \times S$.

Formally, it is a repeated game where $C = I \times J \times S \times U$, $D = I \times J \times U$ and the transition function satisfies the following two conditions. At first, the signal u is public and the actions are observed:

$$\forall (k, i, j) \in K \times I \times J, \quad \sum_{k', s, u \in K \times S \times U} q(k, i, j)(k', (i, j, u), (i, j, s, u)) = 1.$$

Secondly, there exists a function $\hat{q}$ from $K \times I$ to $\Delta(K \times S \times U)$ such that

$$\forall (k, k', i, j, s, u) \in K \times K \times I \times J \times S \times U, \quad q(k, i, j)(k', (i, j, s, u), (i, j, u)) = \hat{q}(k, i)(k', s, u).$$

Let us stress out that the transition q in itself depends on player 2 since it has to reveal his actions but as we will see our assumptions are still satisfied. It was already noticed in Renault [10] that it is too restrictive to assume that the transition is fully controlled by player 1. This model is a natural generalization of Renault's model, dropping the (important) condition of Player 1 to know the state at every stage. However, it does not allow for imperfect monitoring of actions as in the previous examples.

Let us check that this model satisfies our assumptions. According to the description of the model, an initial distribution π of (k_1, s_1, u_1) can be seen as an element of $\Delta^*_f(K \times \mathbb{N} \times \mathbb{N})$ since the signal of player 2 is contained in the signal of player 1. As for the previous example, we will start with a general initial probability $\pi \in \Delta^*_f(K \times \mathbb{N} \times \mathbb{N})$ and initial signals (c_1, d_1). At first, note that apart from the initial signal d_1, histories of player 2 are contained in histories of player 1, so that assumption $(A1)$ reduces to

$$\forall n \geq 1, \; \forall (\sigma, \tau) \in \Sigma \times \mathcal{T}, \quad \mathcal{L}_{\mathbb{P}^\pi_{\sigma\tau}}(k_n \mid h_n^I, d_1) = \mathcal{L}_{\mathbb{P}^\pi_{\sigma\tau}}(k_n \mid h_n^I).$$

This property is true for $n = 1$ by assumption. Let us proceed by induction on n. Assume that $n \geq 2$ and that the property is proved for $n - 1$, i.e. that

$$x_{n-1} = \mathcal{L}_{\mathbb{P}^\pi_{\sigma\tau}}(k_{n-1} \mid h^I_{n-1}, d_1) = \mathcal{L}_{\mathbb{P}^\pi_{\sigma\tau}}(k_{n-1} \mid h^I_{n-1}).$$

Again, let $\hat{q}$ be linearly extended to $\Delta(K) \times I$. It follows that,

$$\mathcal{L}_{\mathbb{P}^\pi_{\sigma\tau}}(k_n, s_n, u_n | h^I_{n-1}, d_1, i_{n-1}, j_{n-1}) = \hat{q}(x_{n-1}, i_{n-1}) = \mathcal{L}_{\mathbb{P}^\pi_{\sigma\tau}}(k_n, s_n, u_n | h^I_{n-1}, i_{n-1}, j_{n-1}),$$

$(A1)$ follows then directly by disintegration. Moreover, we deduce that

$$\mathbb{P}^\pi_{\sigma,\tau}(k_n = k | h^I_n) = \frac{\hat{q}(x_{n-1}, i_1)(k, s_n, u_n)}{\sum_{k' \in K} \hat{q}(x_{n-1}, i_1)(k', s_n, u_n)}. \tag{3.1}$$

The latter proves that x_2 can be expressed as a function of (x_1, i_1, s_2, u_2) which does not depend on τ. Recall then that by assumption there exist functions h_1 and f_1 such that with probability 1, we have

$$y_1 = \mathcal{L}_\pi(x_1 | d_1) = h_1(d_1) = f_1(c_1).$$

If player 1 uses a reduced strategy σ and player 2 uses a strategy τ, we have

$$\mathbb{P}^\pi_{\sigma,\tau}(k_1, x_1, d_1, i_1, j_1, k_2, s_2, u_2) = \pi(d_1)h_1(d_1)(x_1)x_1(k_1)\sigma(x_1, h_1(d_1))(i_1)\tau(d_1)(j_1)\hat{q}(k_1, i_1)(k_2, s_2, u_2).$$

We deduce that

$$y_2(d_1, i_1, j_1, u_2) = \sum_{x_1, i_1, s_2} \mathbb{P}^\pi_{\sigma,\tau}(x_1, i_1, s_2 | d_1, j_1, u_2)\delta_{\mathcal{L}_{\mathbb{P}^\pi_{\sigma,\tau}}(k_2 | x_1, d_1, i_1, j_1, s_2, u_2)}$$

$$= \sum_{x_1, i_1, s_2} \mathbb{P}^\pi_{\sigma,\tau}(x_1, i_1, s_2 | d_1, j_1, u_2)\delta_{x_2(x_1, i_1, s_2, u_2)}.$$

From the previous formula, we deduce

$$\mathbb{P}^\pi_{\sigma,\tau}(x_1, i_1, s_2 | d_1, j_1, u_2) = \frac{\sum_{k_1} \pi(d_1)h_1(d_1)(x_1)x_1(k_1)\sigma(x_1, h_1(d_1))(i_1)\tau(d_1)(j_1)\hat{q}(k_1, i_1)(s_2, u_2)}{\sum_{k'_1, x'_1, i'_1, s'_2} \pi(d_1)h_1(d_1)(x'_1)x'_1(k'_1)\sigma(x'_1, h_1(d_1))(i'_1)\tau(d_1)(j_1)\hat{q}(k'_1, i'_1)(s'_2, u_2)}$$

$$= \frac{\sum_{k_1} f_1(c_1)(x_1)x_1(k_1)\sigma(x_1, f_1(c_1))(i_1)\hat{q}(k_1, i_1)(s_2, u_2)}{\sum_{k'_1, x'_1, i'_1, s'_2} f_1(c_1)(x'_1)x'_1(k'_1)\sigma(x'_1, f_1(c_1))(i'_1)\hat{q}(k'_1, i'_1)(s'_2, u_2)}.$$

Thus, y_2 does not depend on τ nor on (j_1, d_1). Player 1, knowing i_1, c_1 and u_2, can compute y_2, which proves that assumption $(A2)$ is satisfied. Finally, the distribution of the random variable y_2 is equal to

$$\eta_2 = \sum_{c_1, i_1, u_2} \mathbb{P}^\pi_{\sigma,\tau}(c_1, i_1, u_2)\delta_{y_2(c_1, i_1, u_2)}.$$

Since the function $\hat{q}$ does not depend on j_1, we deduce as above that η_2 does not depend on τ and therefore that $(A3)$ is satisfied.

4 Discussion of the assumptions

In this section, we discuss several implications of the formal assumptions $(A1)$ $,(A2)$, $(A3)$. At first, we show that it is necessary to introduce the notion of reduced strategy in order to exclude non relevant correlations between the players. Then, in order to answer to a suggestion made in [10], we show that the analysis cannot be made in terms of first order beliefs only, and that it is necessary to introduce second order beliefs. Lemma 4.4 shows that the value can be expressed as a function of second order beliefs. Then, we prove that player 1 can compute his first order beliefs, the second order beliefs of player 2 and the distribution of these beliefs without knowing the strategy of player 2 as soon as he plays a Markovian strategy with respect to the beliefs at each stage. Finally, we give a weaker version of the theorem where the assumptions are formulated more directly in terms of the data of the game.

4.1 Necessity of reduced strategies.

The introduction of reduced strategies for player 1 is necessary in order to exclude non relevant correlations between players as shown in example 4.1 below. It will be shown in Lemma 4.4 and in the main Theorem that in our model, there is no loss in restricting player 1 to reduced strategies.

Example 4.1. *Let* $K = \{k_1, k_2\}$, $I = \{T, B\}$, $C = \{a, b\}$, $D = \{\alpha, \beta\}$ *and* J *any finite set. The transition* q *depends only on the action of player 1 and is described by the matrices*

$$
\begin{array}{cc}
T \\
B
\end{array}
\begin{pmatrix}
\delta_{k_1} \\
\frac{1}{2}\delta_{k_1} + \frac{1}{2}\delta_{k_2}
\end{pmatrix}
\begin{pmatrix}
\frac{1}{2}\delta_{k_1} + \frac{1}{2}\delta_{k_2} \\
\delta_{k_2}
\end{pmatrix}
\cdot
\quad .
$$
$$
\quad k_1 \qquad\qquad k_2
$$

At each stage (including at the initial stage), the signals of the players are randomly chosen independently of the state variable with distribution

$$
\begin{array}{cc}
& \alpha \quad\ \beta \\
\begin{array}{c} a \\ b \end{array} &
\begin{pmatrix}
1/6 & 2/6 \\
2/6 & 1/6
\end{pmatrix}.
\end{array}
$$

It is clear that signals do not contain any information on the state variable. However, assume that the initial state is k_1 *and that player 1 plays at the first stage action* T *if he receives the signal* a *and action* B *if he receives the signal* b. *The second order beliefs* y_2 *of player 2 will differ if his initial private signal is equal to* α *or* β. *Since player 1 is not able to compute the initial signal of player 2, he is not able to compute the variable* y_2 *at the second stage. Nevertheless, when considering reduced strategies, signals can be omitted and player 1 is able to compute the beliefs of player 2 which implies that* $(A2)$ *is satisfied.*

4.2 Second-order beliefs

Renault [10] conjectured that the pair of distributions of first order beliefs of both players could be sufficient auxiliary variables. We present here an example showing the necessity to take into account second order beliefs in the sense that there exist a game and two initial probabilities π and π' such that the law of first-order beliefs are the same under π and π' while the values differ.

Example 4.2. *We consider again the situation of example 2.6. Recall that $K = \{k_1, k_2\}$, that there are two public signals $U = \{u_1, u_2\}$ available to both players and three private signals $S = \{s_1, s_2, s_3\}$ for player 1. The set of signals of player 1 is $C = U \times S$ and the set of signals of player 2 is $D = U$. Let $\pi, \pi' \in \Delta(K \times S \times U)$ be defined by*

$$
\begin{array}{cc}
& \begin{array}{cc} u_1 & u_2 \end{array} \\
\begin{array}{c} s_1 \\ s_2 \\ s_3 \end{array} &
\begin{pmatrix} \frac{8}{24} & 0 \\ \frac{1}{24} & \frac{3}{24} \\ 0 & 0 \end{pmatrix}
\end{array}
\qquad
\begin{array}{cc}
& \begin{array}{cc} u_1 & u_2 \end{array} \\
\begin{array}{c} s_1 \\ s_2 \\ s_3 \end{array} &
\begin{pmatrix} 0 & 0 \\ \frac{1}{24} & \frac{3}{24} \\ \frac{2}{24} & \frac{6}{24} \end{pmatrix}
\end{array}.
$$
$$
\qquad\qquad k_1 \qquad\qquad\qquad\qquad k_2
$$

and

$$
\begin{array}{cc}
& \begin{array}{cc} u_1 & u_2 \end{array} \\
\begin{array}{c} s_1 \\ s_2 \\ s_3 \end{array} &
\begin{pmatrix} \frac{6}{24} & \frac{2}{24} \\ \frac{3}{24} & \frac{1}{24} \\ 0 & 0 \end{pmatrix}
\end{array}
\qquad
\begin{array}{cc}
& \begin{array}{cc} u_1 & u_2 \end{array} \\
\begin{array}{c} s_1 \\ s_2 \\ s_3 \end{array} &
\begin{pmatrix} 0 & 0 \\ \frac{3}{24} & \frac{1}{24} \\ 0 & \frac{8}{24} \end{pmatrix}
\end{array}.
$$
$$
\qquad\qquad k_1 \qquad\qquad\qquad\qquad k_2
$$

We will identify $\Delta(K)$ and $[0,1]$ as in example 2.6. The beliefs of Player 2 about the state are the same in both cases and are equal to $\frac{1}{2}\delta_{\frac{3}{4}} + \frac{1}{2}\delta_{\frac{1}{4}}$. Similarly the beliefs of Player 1 are $\frac{1}{3}\delta_1 + \frac{1}{3}\delta_{\frac{1}{2}} + \frac{1}{3}\delta_0$ in both games. Moreover player 1 observes the signal of player 2, so that assumption $(A2)$ is satisfied. Thus, the laws of first-order beliefs are not sufficient to discriminate between π and π'. Let Γ be the repeated game where $K = \{k_1, k_2\}$, $I = \{T, B\}$, $J = \{L, R\}$ and payoff g given by

$$
\begin{pmatrix} 0 & 1 \\ 1 & 2 \end{pmatrix} \quad \begin{pmatrix} 1 & \frac{1}{2} \\ 1 & 0 \end{pmatrix}.
$$
$$
k_1 \qquad\quad k_2
$$

The average payoff matrix with coefficients $(\frac{1}{2}, \frac{1}{2})$ is $\begin{pmatrix} \frac{1}{2} & \frac{3}{4} \\ 1 & 1 \end{pmatrix}$.

Let us prove that $v_1(\pi)$ and $v_1(\pi')$ are different. If player 1 receives s_3, then his beliefs on the state is 0 and thus Top is a weakly dominant action. If he receives s_1 or s_2, his beliefs is 1 or $\frac{1}{2}$ and Bottom is a strictly dominant action. From the point of view of player 2, playing Left if receiving u_1 and playing Right if receiving u_2 is a best reply to this strategy. Using these strategies, we find that $v_1(\pi) = \frac{7}{8}$ and $v_1(\pi') = \frac{11}{12}$.

Let us prove that if assumptions $(A1a)$ and $(A2a)$ hold, then $v_\theta(\pi)$ depends only on the law of second order beliefs of player 2.

Definition 4.3. *For all $\pi \in \Delta_f^*(K \times \mathbb{N} \times \mathbb{N})$, define*

$$
\Phi(\pi) \triangleq \mathcal{L}_\pi \left(\mathcal{L}_\pi \left(\mathcal{L}_\pi(k_1 | c_1) | d_1 \right) \right).
$$

Lemma 4.4. *Let $\pi, \pi' \in \Delta_f^*(K \times \mathbb{N} \times \mathbb{N})$. If $\Phi(\pi) = \Phi(\pi')$, then $v_\theta(\pi) = v_\theta(\pi')$.*

Proof. Let (σ, τ) be a pair of behavior strategies in $\Gamma_\theta(\pi)$. It is enough to show that $v_\theta(\pi)$ depends on π only through $\eta_1 = \Phi(\pi)$. Recall that $x_1 := \mathcal{L}_\pi(k_1|c_1)$ and $y_1 := \mathcal{L}_\pi(\mathcal{L}_\pi(k_1|c_1)|d_1)$. Note that y_1 is a function of d_1, and that x_1 is a function of c_1. Moreover, by assumption $(A2a)$, there exists a map $f_1^\pi : C' \to \Delta(\Delta(K))$ such that

$$f_1^\pi(c_1) = y_1 \quad \pi\text{-almost surely.}$$

Let us construct a reduced version of the game $\Gamma_\theta(\pi)$ in which player 1 and player 2 are constrained to choose strategies that depend only on c_1 and d_1 through the variables (x_1, y_1) and y_1 respectively, and keeping the same payoff function. This game has a value since the sets of possible values of (x_1, y_1) is finite and this value is exactly the value of $\Gamma_\theta(\tilde{\pi})$ where $\tilde{\pi}$ is the joint distribution of $(k_1, (x_1, y_1), y_1)$ seen as an element of $\Delta_f^*(K \times \mathbb{N} \times \mathbb{N})$.

The sets of strategies in $\Gamma_\theta(\tilde{\pi})$ (denoted by $\Sigma'(\pi)$ and $\mathcal{T}'(\pi)$) can be seen as subsets of Σ and $\mathcal{T}$ via the previous identification and we will prove that both games have the same value and that $v_\theta(\tilde{\pi})$ depends only on η_1.

Assume at first that $\tau \in \mathcal{T}'(\pi)$ and $\sigma \in \Sigma$ and let μ denote the joint law of $(k_1, c_1, d_1, x_1, y_1)$ induced by π. By disintegration, we have

$$\gamma_\theta(\pi, \sigma, \tau) = \int_{K \times \mathbb{N} \times \Delta(K) \times \Delta(\Delta(K))} \gamma_\theta(k_1, \sigma(c_1), \tau(y_1)) d\mu(k_1, c_1, x_1, y_1),$$

$$= \int_{\mathbb{N} \times \Delta(K) \times \Delta(\Delta(K))} \left(\int_K \gamma_\theta(k_1, \sigma(c_1), \tau(y_1)) d\mathcal{L}_\mu(k_1|c_1, x_1, y_1) \right) d\mu(c_1, x_1, y_1),$$

$$= \int_{\mathbb{N} \times \Delta(K) \times \Delta(\Delta(K))} \int_K \gamma_\theta(k_1, \sigma(c_1), \tau(y_1)) d\mathcal{L}_\mu(k_1|c_1) d\mu(c_1, x_1, y_1),$$

$$= \int_{\mathbb{N} \times \Delta(K) \times \Delta(\Delta(K))} \langle \gamma_\theta(\cdot, \sigma(c_1), \tau(y_1)), x_1 \rangle_{\mathbb{R}^K} d\mu(c_1, x_1, y_1),$$

where we used that $\mathcal{L}_\mu(k_1|c_1, x_1, y_1) = \mathcal{L}_\mu(k_1|c_1)$ since (x_1, y_1) are c_1-measurable and the notations $\gamma_\theta(., \sigma(c_1), \tau(y_1))$ for $(\gamma_\theta(k, \sigma(c_1), \tau(y_1)))_{k \in K} \in \mathbb{R}^K$ and $\langle \cdot, \cdot \rangle_{\mathbb{R}^K}$ for the scalar product in $\mathbb{R}^K$. Taking the supremum over all strategies of player 1, we obtain

$$\sup_{\sigma \in \Sigma} \gamma_\theta(\pi, \sigma, \tau) = \int_{\mathbb{N} \times \Delta(K) \times \Delta(\Delta(K))} \sup_{\sigma(c_1)} \langle \gamma_\theta(\cdot, \sigma(c_1), \tau(y_1)), x_1 \rangle_{\mathbb{R}^K} d\mu(c_1, x_1, y_1).$$

The supremum inside the integral is achieved by strategies depending only on (x_1, y_1) since these variables are c_1 measurable. It means that there exists an optimal strategy in $\Sigma'(\pi)$, which proves

$$\inf_{\tau \in \mathcal{T}'(\pi)} \sup_{\sigma \in \Sigma} \gamma_\theta(\pi, \sigma, \tau) = \inf_{\tau \in \mathcal{T}'(\pi)} \sup_{\sigma \in \Sigma'(\pi)} \gamma_\theta(\pi, \sigma, \tau).$$

Moreover the value of the reduced game depends only on η_1 since taking the infimum over $\tau \in \mathcal{T}'(\pi)$,

$$\inf_{\tau \in \mathcal{T}'(\pi)} \sup_{\sigma \in \Sigma'(\pi)} \gamma(\pi, \sigma, \tau) \tag{4.1}$$

$$= \int_{\Delta(\Delta(K))} \left[\inf_{\tau(y_1)} \int_{\Delta(K)} \left(\sup_{\sigma(x_1, y_1)} \langle \gamma_\theta(\cdot, \sigma(x_1, y_1), \tau(y_1)), x_1 \rangle_{\mathbb{R}^K} \right) d\mathcal{L}_\mu(x_1 \mid y_1) \right] d\mu(y_1), \tag{4.2}$$

15

which depends only on the law of y_1, since $y_1 = \mathcal{L}_\mu(x_1 \mid d_1) = \mathcal{L}_\mu(x_1 \mid y_1)$.

Let us prove a dual equality starting with $\sigma \in \Sigma'(\pi)$ and $\tau \in \mathcal{T}$:

$$\gamma_\theta(\pi, \sigma, \tau) = \int_{K \times \mathbb{N} \times \mathbb{N} \times \Delta(K) \times \Delta(\Delta(K))} \gamma_\theta(k_1, \sigma(x_1, y_1), \tau(d_1)) d\mu(k_1, c_1, d_1, x_1, y_1),$$

$$= \int_{\mathbb{N} \times \mathbb{N} \times \Delta(K) \times \Delta(\Delta(K))} \left(\int_K \gamma_\theta(k_1, \sigma(x_1, y_1), \tau(d_1)) d\mathcal{L}_\mu(k_1 | c_1, d_1, x_1, y_1) \right) d\mu(c_1, d_1, x_1, y_1)$$

$$= \int_{\mathbb{N} \times \mathbb{N} \times \Delta(K) \times \Delta(\Delta(K))} \langle \gamma_\theta(\cdot, \sigma(x_1, y_1), \tau(d_1)), x_1 \rangle_{\mathbb{R}^K} d\mu(c_1, d_1, x_1, y_1)$$

$$= \int_{\mathbb{N} \times \Delta(K) \times \Delta(\Delta(K))} \langle \gamma_\theta(\cdot, \sigma(x_1, y_1), \tau(d_1)), x_1 \rangle_{\mathbb{R}^K} d\mu(d_1, x_1, y_1)$$

$$= \int_{\mathbb{N} \times \Delta(\Delta(K))} \left(\int_{\Delta(K)} \langle \gamma_\theta(\cdot, \sigma(x_1, y_1), \tau(d_1)), x_1 \rangle_{\mathbb{R}^K} d\mathcal{L}_\mu(x_1 \mid d_1, y_1) \right) d\mu(d_1, y_1).$$

For the second equality, we used that $\mathcal{L}_\mu(k_1 \mid c_1, d_1, x_1, y_1) = \mathcal{L}_\mu(k_1 \mid c_1, d_1) = \mathcal{L}_\mu(k_1 \mid c_1) = x_1$ which follows from the fact that (x_1, y_1) is c_1-measurable and assumption $(A1)$. Taking the infimum over all $\tau \in \mathcal{T}$, it follows that

$$\inf_{\tau \in \mathcal{T}} \gamma(\pi, \sigma, \tau) = \int_{\mathbb{N} \times \Delta(\Delta(K))} \inf_{\tau(d_1)} \left(\int_{\Delta(K)} \langle \gamma_\theta(\cdot \sigma(x_1, y_1), \tau(d_1)), x_1 \rangle_{\mathbb{R}^K} d\mathcal{L}_\mu(x_1 \mid d_1) \right) d\mu(d_1, y_1).$$

$$(4.3)$$

The infimum inside the integral is achieved for strategies depending only on $y_1 = \mathcal{L}_\mu(x_1 \mid d_1)$ since y_1 is d_1-measurable. We proved that

$$\sup_{\sigma \in \Sigma'(\pi)} \inf_{\tau \in \mathcal{T}} \gamma_\theta(\pi, \sigma, \tau) = \sup_{\sigma \in \Sigma'(\pi)} \inf_{\tau \in \mathcal{T}'(\pi)} \gamma_\theta(\pi, \sigma, \tau).$$

Finally, using that $\mathcal{T}'(\pi) \subset \mathcal{T}$ and $\Sigma'(\pi) \subset \Sigma$, it follows that

$$v_\theta(\pi) = \sup_{\sigma \in \Sigma} \inf_{\tau \in \mathcal{T}} \gamma_\theta(\pi, \sigma, \tau) \geq \sup_{\sigma \in \Sigma'(\pi)} \inf_{\tau \in \mathcal{T}} \gamma_\theta(\pi, \sigma, \tau) = v_\theta(\tilde{\pi}),$$

$$v_\theta(\pi) = \inf_{\tau \in \mathcal{T}} \sup_{\sigma \in \Sigma} \gamma_\theta(\pi, \sigma, \tau) \leq \inf_{\tau \in \mathcal{T}'(\pi)} \sup_{\sigma \in \Sigma} \gamma_\theta(\pi, \sigma, \tau) = v_\theta(\tilde{\pi}),$$

which proves the equality. Since $v_\theta(\tilde{\pi})$ depends only on η_1, the proof is complete. □

4.3 Player 1 can compute his beliefs without knowing player 2's strategy.

Assumption $(A1)$ can be reformulated as a couple of assumptions $(A1a)$ and $(A1b)$ which are expressed in terms of π, i.e. the initial information, and q, i.e. the evolution of the information structure for stages $m \geq 2$ respectively.

(A1a) The probability $\pi \in \Delta(K \times C' \times D')$ is such that

$$\forall (k, c', d') \in K \times C' \times D', \qquad \pi(c')\pi(k, c', d') = \pi(k, c')\pi(c', d')$$

(A1b) There exists a map F from $\Delta(K) \times I \times C$ to $\Delta(K)$ such that

$$\forall (p,i,j,c,d,k) \in \Delta(K) \times I \times J \times C \times D \times K, \ q(p,i,j)[k,c,d] = F(p,i,c)[k] \sum_{k' \in K} q(p,i,j)[k',c,d].$$

Note that $(A1a)$ is equivalent to $\pi(k|c',d') = \pi(k|c')$ for any (k,c',d') such that $\pi(c',d') > 0$. Similarly $(A1b)$ could be written in terms of conditional probabilities, though we shall distinguish events with probability 0. In addition, it highlights the first important consequence of assumption $(A1)$: player 1 can compute his beliefs about the state variable (i.e. the conditional distribution in the right-hand-side of $(A1)$) without knowing the strategy, nor the signals, of his opponent.

Proposition 4.5. *Assuming $(A1)$, player 1 can compute x_n for each $n \geq 1$ without knowing the strategy of player 2.*

The proof of the Proposition follows directly from the following Lemma.

Lemma 4.6. *Assumptions $(A1)$ and $(A1a + A1b)$ are equivalent. Furthermore, the map F from $\Delta(K) \times I \times C$ to $\Delta(K)$ defined in $(A1b)$ is such that for all $n \geq 2$ and for all strategy profile (σ, τ)*

$$x_n = F(x_{n-1}, i_{n-1}, c_{n-1}), \quad \mathbb{P}^\pi_{\sigma\tau}\text{-almost surely.}$$

Proof. Using the definition of conditional independence, assumption $(A1)$ at stage 1 is equivalent to $(A1a)$. It remains to prove that $(A1)$ for $n \geq 2$ implies $(A1b)$ and the converse. Assume that π fulfils $(A1a)$ and let $(\sigma_1, \tau_1) \in \Delta(I)^{C'} \times \Delta(J)^{D'}$ be strategies with full support. By construction, we have

$$\mathcal{L}_{\mathbb{P}^\pi_{\sigma\tau}}(k_2, c_2, d_2 \mid k_1, c_1, i_1, d_1, j_1) = q(k_1, i_1, j_1) \in \Delta(K \times C \times D).$$

It follows, using the tower property of conditional expectation and $(A1a)$ that

$$\mathcal{L}_{\mathbb{P}^\pi_{\sigma\tau}}(k_2, c_2, d_2 \mid c_1, i_1, d_1, j_1) = q(x_1, i_1, j_1),$$

where, by definition, x_1 can be written as a function of c_1. On one hand, one obtains by disintegration

$$\mathbb{P}^\pi_{\sigma\tau}(k_2 = k \mid c_2, d_2, c_1, i_1, d_1, j_1)(\sum_{\widetilde{k} \in K} q(x_1, i_1, j_1)[\widetilde{k}, c_2, d_2]) = q(x_1, i_1, j_1)[k, c_2, d_2].$$

On the other hand, the conditional law $\mathcal{L}_{\mathbb{P}^\pi_{\sigma\tau}}(k_2 \mid c_1, i_1, c_2)$ is characterized by the following expression

$$\mathbb{P}^\pi_{\sigma\tau}(k_2 = k \mid c_1, i_1, c_2)(\sum_{\widetilde{k}, \widetilde{d}_1, \widetilde{d}_2, \widetilde{j}_1} \pi(c_1, \widetilde{d}_1)\tau_1(\widetilde{d}_1)[\widetilde{j}_1]q(x_1(c_1), i_1, \widetilde{j}_1)[\widetilde{k}, c_2, \widetilde{d}_2])$$

$$= \sum_{\widetilde{d}_1, \widetilde{d}_2, \widetilde{j}_1} \pi(c_1, \widetilde{d}_1)\tau_1(\widetilde{d}_1)[\widetilde{j}_1]q(x_1(c_1), i_1, \widetilde{j}_1)[k, c_2, \widetilde{d}_2].$$

Assumption $(A1)$ for $n = 2$ implies that these two conditional probabilities are equal, which in turn implies

$$\frac{q(x_1, i_1, j_1)[k, c_2, d_2]}{\sum_{\widetilde{k} \in K} q(x_1, i_1, j_1)[\widetilde{k}, c_2, d_2]} = \frac{\sum_{\widetilde{d}_1, \widetilde{d}_2, \widetilde{j}_1} \pi(c_1, \widetilde{d}_1)\tau_1(\widetilde{d}_1)[\widetilde{j}_1]q(x_1(c_1), i_1, \widetilde{j}_1)[k, c_2, \widetilde{d}_2]}{\sum_{\widetilde{k}, \widetilde{d}_1, \widetilde{d}_2, \widetilde{j}_1} \pi(c_1, \widetilde{d}_1)\tau_1(\widetilde{d}_1)[\widetilde{j}_1]q(x_1(c_1), i_1, \widetilde{j}_1)[\widetilde{k}, c_2, \widetilde{d}_2]} \quad (4.4)$$

whenever the left-hand side is well-defined. Since τ_1 has full support, this implies that the right-hand side is also well-defined in this case and does not depend on d_1, j_1, d_2. Moreover, for all $p \in \Delta(K)$, we can choose an initial distribution π such that $\pi(x_1 = p) > 0$. It follows that there exists a function F such that

$$F(p, i, c)[k] = \frac{q(p, i, j)[k, c, d]}{\sum_{k' \in K} q(p, i, j)[k', c, d]},$$

whenever the right hand side is well-defined for some (j, d) and extended by $1/|K|$ (say) otherwise.

For the converse assertion, we already mentioned that $(A1a)$ implies $(A1)$ for $n = 1$. We are therefore allowed to write the following formula for the conditional laws,

$$\mathbb{P}(k_2 = k \mid c_2, d_2, c_1, i_1, d_1, j_1) = \frac{q(x_1, i_1, j_1)[k, c_2, d_2]}{\sum_{\widetilde{k} \in K} q(x_1, i_1, j_1)[\widetilde{k}, c_2, d_2]}. \tag{4.5}$$

It follows therefore that

$$\mathbb{P}(k_2 = k \mid c_2, d_2, c_1, i_1, d_1, j_1) = F(x_1, i_1, c_2), \quad \mathbb{P}^\pi_{\sigma\tau}\text{-almost surely,}$$

and since the right-hand-side is measurable with respect to the history of player 1, we have the equality

$$\begin{aligned}
\mathbb{P}(k_2 = k \mid c_2, c_1, i_1) &= \mathbb{E}\big[\mathbb{P}(k_2 = k \mid c_2, d_2, c_1, i_1, d_1, j_1)|c_1, i_1, c_2\big], \\
&= F(x_1, i_1, c_2), \\
&= \mathbb{P}(k_2 = k \mid c_2, d_2, c_1, i_1, d_1, j_1).
\end{aligned}$$

which proves $(A1)$ and our last assertion for $n = 2$. Finally the distribution of $(k_2, (c_1, i_1, c_2), (d_1, j_1, d_2))$, seen as an element of $\Delta_f(K \times \mathbb{N} \times \mathbb{N})$, fulfils $(A1a)$. Applying exactly the same argument with these new initial signals allows us therefore to conclude by induction on n. □

4.4   Player 1 can compute the beliefs of player 2.

The assumptions $(A1)$ and $(A2)$ are independent, as shown in example 4.7 below. However, $(A2)$ really makes sense only when player 1 is better informed.

Example 4.7. *Let $\Gamma = (K, I, J, C, D, q, g)$ be such that player 1 is in the dark and player 2 is perfectly informed: $K = \{\alpha, \beta\}$, I and J are finite, C is a singleton $\{c\}$ and $D = K$. The payoff mapping is anything and the state is randomly chosen at each stage with probability $(1/2, 1/2)$. Player 1 observes nothing and player 2 learns the state. It is clear that player 1's signal is less accurate than player 2's, so that assumption $(A1)$ is not satisfied. On the other hand, $(A2)$ is satisfied since player 1 knows the beliefs of player 2 about himself which is $(\frac{1}{2}, \frac{1}{2})$ whatever are the signals.*

Under the assumptions $(A1)$ and $(A2)$, if player 1 plays a reduced strategy, he can compute y_2, the belief of player 2 about his own belief on the state, without knowing the strategy of player 2.

Lemma 4.8. *Assume (A1b) and (A2b), and let $\pi \in \Delta_f^*(K \times \mathbb{N} \times \mathbb{N})$. Then, for all $\sigma \in \Sigma'(\pi)$, there exists a map $f_2 = f_2^{\pi,\sigma}$ such that for all $\tau \in \mathcal{T}$*

$$y_2 = f_2(h_2^I), \ \mathbb{P}_{\sigma\tau}^\pi - almost\ surely.$$

Proof. It is sufficient to prove that the map f_2 appearing in $(A2b)$ does not depend on τ. Note that since we assumed $(A1)$, we have $x_2 = F(x_1(c_1), i_1, c_2)$ almost surely, where F is defined in $(A1b)$. Moreover, the conditional probability

$$\mathbb{P}_{\sigma\tau}^\pi(c_1 = \tilde{c}_1, i_1 = \tilde{i}_1, c_2 = \tilde{c}_2 | d_1, j_1, d_2)$$

$$= \frac{\pi(\tilde{c}_1, d_1)\sigma(x_1(\tilde{c}_1), y_1(d_1))(\tilde{i}_1)\tau(d_1)(j_1)q(x_1(\tilde{c}_1, \tilde{i}_1, j_1)(\tilde{c}_2, d_2)}{\sum_{c_1', i_1', c_2'} \pi(x', d_1)\sigma(x_1(c_1'), y_1(d_1))(i_1')\tau(d_1)(j_1)q(x_1(c_1'), i_1', j_1)(c_2', d_2)}$$

$$= \frac{\pi(\tilde{c}_1, d_1)\sigma(x_1(\tilde{c}_1), y_1(d_1))(\tilde{i}_1)q(x_1(\tilde{c}_1, \tilde{i}_1, j_1)(\tilde{c}_2, d_2)}{\sum_{c_1', i_1', c_2'} \pi(x', d_1)\sigma(x_1(c_1'), y_1(d_1))(i_1')q(x_1(c_1'), i_1', j_1)(c_2', d_2)}$$

does not depend on τ. There exists therefore a map $y_2(d_1, j_1, d_2)$ which does not depend on τ, defined by the above expression everywhere it makes sense and arbitrarily elsewhere. Let τ^* be a strategy with full support. Using $(A2b)$, there exists a map f_2^{π,σ,τ^*} such that

$$y_2(d_1, j_1, d_2) = f_2^{\pi,\sigma,\tau^*}(c_1, i_1, c_2), \ \mathbb{P}_{\sigma\tau^*}^\pi - almost\ surely.$$

The previous computation shows that the conditional law of (c_1, i_1, c_2) given (d_1, j_1, d_2) does not depend on τ. Therefore, if the event

$$\{d_1 = \tilde{d}_1, j_1 = \tilde{j}_1, d_2 = \tilde{d}_2, c_1 = \tilde{c}_1, i_1 = \tilde{i}_1, c_2 = \tilde{c}_2\}$$

has positive probability under $\mathbb{P}_{\sigma\tau}^\pi$, it also has positive probability under $\mathbb{P}_{\sigma\tau^*}^\pi$. We deduce that $f_2^{\pi,\sigma,\tau} = f_2^{\pi,\sigma,\tau^*}$, $\mathbb{P}_{\sigma\tau}^\pi$-almost surely for all τ, which concludes the proof. $\square$

Let us now prove that player 1 is able to play a strategy which is Markovian with respect to the beliefs. The idea is to prove by induction that if player 1 plays a strategy which depends at stage $n-1$ only on (x_{n-1}, y_{n-1}), then he can compute the variables (x_n, y_n) at stage n and play at stage n a strategy which depends only on (x_n, y_n), etc... Formally, we have the following.

Lemma 4.9. *For all $\pi \in \Delta_f^*(K \times \mathbb{N} \times \mathbb{N})$, and for any sequence of $\Delta(I)$-valued measurable functions $\psi_1, \psi_2, ...$ defined on $\Delta(K) \times \Delta_f(\Delta(K))$, there exists a strategy σ such that for all τ and for all n*

$$\sigma(h_n^I) = \psi_n(x_n, y_n), \quad \mathbb{P}_{\sigma\tau}^\pi - almost\ surely.$$

Proof. We will prove the result by induction. It is obviously true for $n = 1$ due to the definition of $\Delta_f^*(K \times \mathbb{N} \times \mathbb{N})$. For $n = 2$, due to the Lemmas 4.6 and 4.8, player 1 can compute x_2 and y_2 as a function of h_2^I independently of the chosen strategy of player 2. However, to prove the property for $n \geq 3$, we cannot rely on the same argument. It would be tempting to say that the distribution of (k_2, h_2^I, h_2^{II}) belongs to $\Delta_f^*(K \times \mathbb{N} \times \mathbb{N})$ and to apply the preceding argument when starting from this new initial distribution. But this would be wrong since this distribution may depend on τ. To overcome this problem, it is sufficient to prove that

the map f_2 appearing in $(A2b)$ and the distribution η_2 appearing in $(A3)$ depend on π only through $\Phi(\pi)$. Indeed, in this case, reasoning by induction, player 1 can compute η_n as a function of η_{n-1} and his new signals, x_n using Lemma 4.6, and y_n using the map given by $(A2b)$ which will depend only on η_{n-1} and his own strategy.

Let us prove this assertion. Let $\pi \in \Delta_f^*(K \times \mathbb{N} \times \mathbb{N})$, and $\sigma \in \Sigma'(\pi)$ be a reduced strategy, which implies that there exists a map $\psi : K \times \Delta(K) \to \Delta(I)$ such that π-almost surely $\sigma_1(c_1) = \psi(x_1, y_1)$. Assumption $(A3)$ implies that η_2 is a function of the initial distribution π and σ_1 only. We denote it by $\eta_2(\pi, \sigma)$. We now prove that $\eta_2(\pi, \sigma)$ and the map $f_2^{\pi, \sigma}$ appearing in $(A2b)$ depend only on the projection of π, $\eta_1 = \Phi(\pi)$, and on the map ψ.

At first, given $\eta_1 \in \Delta_f(\Delta_f(\Delta(K)))$, we can construct a canonical probability $\overline{\pi}$ with finite support on $K \times \Delta(K) \times \Delta_f(\Delta(K))$ defined by $\overline{\pi}(k, p, z) = p^k z(p) \eta(z)$. Applying $(A3)$ and $(A2)$ in the game $\Gamma(\overline{\pi})$ if player 1 plays $\overline{\sigma}_1 = \psi$, there exists a distribution $\eta_2(\overline{\pi}, \psi)$ and a map $f_2^{\overline{\pi}, \psi} : K \times \Delta(K)$ such that $y_2 = f_2^{\overline{\pi}, \psi}(x_1, y_1)$ almost surely and y_2 has law $\eta_2(\overline{\pi}, \psi)$ for all τ. Recall that π is such that $\Phi(\pi) = \eta_1$ and that σ_1 is such that $\sigma_1(c_1) = \psi(x_1, y_1)$. Note also that d_1 and x_1 are conditionally independent given y_1 under π. Therefore, for any τ_1, the joint law of $(x_1, y_1, i_1, j_1, c_2, d_2)$ is the same under $\mathbb{P}^{\pi}_{\sigma_1, \tau_1}$ and under the probability $\mathbb{P}^{\overline{\pi}}_{\psi, \tau_1'}$ where τ_1' is defined as follows: choose d_1 using some exogenous lottery such that the conditional law of d_1 given y_1 is the same as under π and then play $\tau_1(d_1)$. We deduce that $\eta_2(\pi, \sigma_1) = \eta_2(\overline{\pi}, \psi)$ and $y_2 = f_2^{\overline{\pi}, \psi}(x_1, y_1)$ under the probability $\mathbb{P}^{\pi}_{\sigma_1, \tau_1}$ which concludes the proof. $\qquad \square$

4.5 A stronger version of the theorem

To conclude this section, let us state a couple of stronger assumptions, which are expressed in terms of the data of the game more directly: Player 1 can deduce exactly the signal received by player 2 and player 2 can not influence the joint law of (x_2, d_2).

Definition 4.10. *For all, $x, i, j \in \Delta(K) \times I \times J$, let $q_{C \times D}(x, i, j)$ denote the marginal distribution on $C \times D$ induced by $q(x, i, j)$, i.e. $q_{C \times D}(x, i, j)(c, d) = \sum_{k, \tilde{k}} x(k) q(k, i, j)(\tilde{k}, c, d)$.*

Let also $H_{x,i}$ the map defined on $C \times D$ by

$$H_{x,i}(c, d) = (F(x, i, c), d) \in \Delta(K) \times D.$$

With these notations, we can define a set of assumptions on the marginal of q. The assumptions $(A1)$, $(A2a)$ are unchanged and we define $(A'2b)$ and $(A'3)$.

$(A'2b)$ Player 1 knows the signal of player 2 i.e. there exists a map $h : C \to D$ such that for all $(k, i, j) \in K \times I \times J$, $\sum_{c \in C} q(k, i, j)[c, h(c)] = 1$.

$(A'3)$ The image probability $\phi(x, i)$ of $q_{C \times D}(x, i, j)$ by the map $H_{x,i}$ does not depend on j.

Corollary 4.11. *Let Γ be such that assumptions $(A1), (A2a), (A'2b)$ and $(A'3)$ are true. Then:*

$$\textit{For all } \pi \in \Delta_f^*(K \times \mathbb{N} \times \mathbb{N}), \quad \Gamma(\pi) \textit{ has a uniform value.}$$

The proof of this corollary follows directly from the next Lemma.

Lemma 4.12. *If $A1$ and $A2a$ hold, then $A'2b$ and $A'3$ imply $A2b$ and $A3$.*

Proof. It follows from the definitions and from Lemma 4.6 that

$$\mathcal{L}_{\mathbb{P}^\pi_{\sigma\tau}}(x_2, d_2 \mid c_1, d_1, i_1, j_1) = \mathcal{L}_{\mathbb{P}^\pi_{\sigma\tau}}(F(x_1, i_1, c_2), d_2 \mid c_1, d_1, i_1, j_1)$$
$$= \mathcal{L}_{\mathbb{P}^\pi_{\sigma\tau}}(H_{x_1, i_1}(c_2, d_2) \mid c_1, d_1, i_1, j_1) = \phi(x_1, i_1),$$

since (x_1, i_1) is measurable with respect to (c_1, d_1, i_1, j_1) and $\phi(x_1, i_1)$ is the image probability of $q_{C \times D}(x_1, i_1, j_1)$ by the map H_{x_1, i_1}. Therefore, the conditional law of the pair (x_2, d_2) does not depend on the strategy of player 2. Precisely, we have

$$\mathcal{L}_{\mathbb{P}^\pi_{\sigma\tau}}(x_2, d_2 \mid d_1, j_1) = \mathbb{E}^\pi_{\sigma\tau}[\phi(x_1, i_1) \mid d_1, j_1].$$

Since j_1 and (x_1, i_1) are conditionally independent given d_1 it follows that

$$\mathcal{L}_{\mathbb{P}^\pi_{\sigma\tau}}(x_2, d_2 \mid d_1, j_1) = \mathbb{E}^\pi_{\sigma\tau}[\phi(x_1, i_1) \mid d_1].$$

The right hand side does not depend on τ, so $\mathcal{L}_{\mathbb{P}^\pi_{\sigma\tau}}(x_2, d_2 \mid d_1, j_1)$ does not depend on τ and j_1, and the same is true for the (unconditional) law of (x_2, y_2). As a consequence, the law of y_2 (denoted η_2) does not depend on τ which proves $(A3)$. It remains to prove that player 1 can compute the auxiliary random variable y_2. Using $(A2a)$ and that σ is reduced, i_1 can be written as a measurable function of (x_1, y_1) and of an independent random variable u uniformly distributed on $[0, 1]$. Recall that the conditional law of x_1 given d_1 is y_1, so that

$$\mathcal{L}_{\mathbb{P}^\pi_{\sigma\tau}}(x_2, d_2 \mid d_1, j_1) = \mathbb{E}^\pi_{\sigma\tau}[\phi(x_1, i_1(x_1, y_1, u)) \mid d_1]$$
$$= \int_{\Delta(K) \times [0,1]} \phi(x, i_1(x, y_1, u)) dy_1[x] du.$$

Player 1 can compute the conditional law of (x_2, d_2) given (d_1, j_1) since it depends only on (y_1, σ). Moreover by assumption $(A'2b)$, he can deduce d_1 from his initial signal c_1, so he is able to compute y_2 which proves $(A2b)$. $\square$

5 Proof of Theorem 2.3.

The proof is divided into three steps. First, using Lemma 4.4, we define a value function $\hat{v}$ on $\Delta_f(\Delta_f(\Delta(K)))$ and prove that it is concave and Lipschitz. Secondly, we introduce an auxiliary game $\mathcal{G}$ on $\Delta_f(\Delta(K))$ and check it satisfies some (slightly) weakened assumptions needed to apply a Theorem of Renault [10]. This implies the existence of a uniform value in the auxiliary game. Finally we show that both players can guarantee this value in the original game: player 2 by playing by blocks and player 1 by using optimal Markovian strategies in the auxiliary game.

5.1 The canonical value function $\hat{v}_\theta$

In view of Lemma 4.4, it is appropriate to work directly on the set $\Delta_f(\Delta_f(\Delta(K)))$, i.e. for any π, π' such that $\Phi(\pi) = \Phi(\pi')$ the value of the game is the same. At first, given $\eta \in \Delta_f(\Delta_f(\Delta(K)))$, there is a canonical way to build a distribution π such that $\phi(\pi) = \eta$.

Definition 5.1. *Let* $\Gamma = (K, I, J, C, D)$ *be a repeated game. For any* $\eta \in \Delta_f(\Delta_f(\Delta(K)))$, *we define* $D' = \text{supp}(\eta) \subset \Delta_f(\Delta(K))$ *and* $C' = D' \times \left(\cup_{z \in \text{supp}(\eta)} \text{supp}(z) \right)$. *By definition of* η, *these sets are finite and we can define* $\pi \in \Delta_f^*(K \times C' \times D')$ *by*

$$\forall (k, p, z) \in K \times \Delta(K) \times \Delta_f(\Delta(K)), \pi(k, (p, z), z) = \eta(z) z(p) p(k).$$

To canonical game $\Gamma(\pi)$ *will be denoted* $\widehat{\Gamma}(\eta)$, *and its value* $\hat{v}_\theta(\eta)$. *If* $\eta = \delta_z$ *for some* $z \in \Delta_f(\Delta(K))$, *we will use the shorter notations* $\widehat{\Gamma}(z) = \widehat{\Gamma}(\delta_z)$ *and* $\hat{v}_\theta(z)$ *for the value.*

Informally, the game $\widehat{\Gamma}(\eta)$ proceeds as follows: η is common knowledge, player 2 is informed about the realization z of a random variable of law η (player 2 learns his beliefs). Then player 1 is informed about z (his opponent's beliefs) and about the realization p of a random variable of law z (his own beliefs). The state variable is finally selected according to p, but none of the players observe it. If $\eta = \delta_z$, for some $z \in \Delta_f(\Delta(K))$, then the set of initial signals for player 2 is reduced to a singleton. In this case, player 1 receives a partial information about the state, whereas player 2 only knows the joint distribution over the state and player 1's signal. Using these notations, Lemma 4.4 implies that if $\pi, \pi' \in \Delta_f^*(K \times \mathbb{N} \times \mathbb{N})$ are such that $\Phi(\pi) = \Phi(\pi')$, we have that $v_\theta(\pi) = v_\theta(\pi') = \hat{v}_\theta(\Phi(\pi))$.

In order to study the regularity of the canonical value function, let us recall some properties of the Wasserstein distance

Let (Z, d) be a compact metric space and $Lip_1(Z)$ the set of 1-Lipschitz functions on Z. The function

$$\mathbf{d} : \Delta(Z) \times \Delta(Z) : (\mu, \nu) \to \sup_{f \in Lip_1(Z)} \int_Z f d\mu - \int_Z f d\nu$$

is a distance on $\Delta(Z)$ which makes $\Delta(Z)$ compact. Moreover, for all $\mu, \nu \in \Delta(Z)$

$$\mathbf{d}(\mu, \nu) = \min_{\pi \in \mathcal{P}(\mu, \nu)} \int_{Z \times Z} |y - x| d\pi(x, y),$$

where $\mathcal{P}(\mu, \nu)$ is the set of probabilities on $Z \times Z$ having for marginals μ and ν (see e.g. [14]).

If f is a bounded measurable function on Z, define $\tilde{f} : \Delta(Z) \to \mathbb{R}$ by $\tilde{f}(\mu) = \int_Z f d\mu$. Then

$$\tilde{f} \in Lip_1(\Delta(Z), \mathbf{d}) \Leftrightarrow f \in Lip_1(Z).$$

In the following, $\Delta(K)$ is endowed with the ℓ_1-norm induced by $\mathbb{R}^K$ and $\Delta(\Delta(K))$ is endowed with the Wasserstein metric $\mathbf{d}$ induced by the metric space $(\Delta(K), \ell_1)$.

Lemma 5.2. *Let* $\eta \in \Delta_f(\Delta_f(\Delta(K)))$ *and* $z \in \Delta_f(\Delta(K))$. *Then* $\hat{v}_\theta(\eta)$ *is linear on* $\Delta_f(\Delta_f(\Delta(K)))$ *and the mapping on* $\Delta(\Delta(K))$, $\hat{v}_\theta(z)$ *is 1-Lipschitz for the Wasserstein metric* $\mathbf{d}$.

Proof. The first assertion is immediate since by definition both players learn the realization of η. Let $z, z' \in \Delta_f(\Delta(K))$. By definition of the Wasserstein distance, there exists $\mu \in \Delta(\Delta(K) \times \Delta(K))$ such that the first marginal is z, the second is z' and

$$\mathbf{d}(z, z') = \int_{\Delta(K) \times \Delta(K)} \|p - p'\|_1 d\mu(p, p').$$

We denote by $\mathcal{L}_\mu(p | p')$ the conditional law of p given p'.

Let $\sigma \in \Sigma$ be a behavior strategy for player 1 in the game $\widehat{\Gamma}(z)$. As in Section 2.2, $\sigma(p)$ denotes the strategy of player 1 conditionally on the signal p. Let us construct a general strategy for P1 as follows. Let $(\Omega, \mathbb{P}) = ([0,1], dx)$ be the auxiliary probability space[3] that will be used as a "tossing coin". The classical representation result of Blackwell-Dubins (see [2]) asserts that there exists a jointly Borel-measurable map $\phi : \Omega \times \Delta(\Delta(K)) \mapsto \Delta(K)$ such that for all $\nu \in \Delta(\Delta(K))$, $\phi(\cdot, \nu)$ is a ν-distributed random variable. Therefore, the map $\sigma'(\omega, p') = \sigma(\phi(\omega, \mu(p|p')))$ defines a general strategy which is equivalent to a behavior strategy by Kuhn's theorem. It follows that

$$
\begin{aligned}
\gamma_\theta(z', \sigma', \tau) &= \int_{\Delta(K) \times \Omega} \gamma_\theta(p', \sigma'(\omega, p'), \tau) dz'(p') \otimes d\mathbb{P}(\omega), \\
&= \int_{\Delta(K)} \left(\int_\Omega \gamma_\theta(p', \sigma(\phi(\omega, \mu(p|p'))), \tau) d\mathbb{P}(\omega) \right) dz'(p'), \\
&= \int_{\Delta(K)} \left(\int_{\Delta(K)} \gamma_\theta(p', \sigma(p), \tau) d\mathcal{L}_\mu(p|p') \right) d\mu(p'), \\
&= \int_{\Delta(K) \times \Delta(K)} \gamma_\theta(p', \sigma(p), \tau) d\mu(p, p'),
\end{aligned}
$$

where the last equality follows from $dz'(p') = d\mu(p')$. Recall that by assumption g takes values in $[0,1]$. Consequently, $\gamma(p, \sigma(p), \tau) \in [0,1], \forall p, \sigma, \tau$. Hence

$$
\begin{aligned}
|\gamma_\theta(z', \sigma', \tau) - \gamma_\theta(z, \sigma, \tau)| &\leq \int_{\Delta(K) \times \Delta(K)} |\gamma_\theta(p, \sigma(p), \tau) - \gamma_\theta(p', \sigma(p), \tau)| d\mu(p, p'), \\
&\leq \int_{\Delta(K) \times \Delta(K)} \|p - p'\|_1 d\mu(p, p'), \\
&= \mathbf{d}(z, z').
\end{aligned}
$$

It follows that $|\hat{v}_\theta(z) - \hat{v}_\theta(z')| \leq \mathbf{d}(z, z')$, for any $z, z' \in \Delta_f(\Delta(K))$. $\qquad \square$

Note that usually, the underlying space is $\Delta(K)$ with discrete metric on K and, in order to prove that the value is 1-Lipschitz, we can use the same strategy in $\Gamma(z)$ in $\Gamma(z')$. Here, we cannot use directly σ. The state space is $\Delta_f(\Delta(K))$ with the norm 1 on $\Delta(K)$, and two states may be close while having disjoint supports. Therefore an optimal strategy σ in $\Gamma(z)$ may have no sense in z'. The idea behind the above proof is to construct, given σ in $\Gamma(z)$, a strategy σ' in $\Gamma(z')$ which behaves in z' like σ in z.

Example 5.3. *Assume that $K = \{k_1, k_2\}$ and let $z = \delta_{\frac{1}{2}}$ and $z' = \frac{1}{2}\delta_{\frac{1}{2}-\epsilon} + \frac{1}{2}\delta_{\frac{1}{2}+\epsilon}$ be two initial distributions in $\Delta_f(\Delta(K)$ (where we identified $\Delta(K)$ and $[0,1]$). A strategy in $\Gamma(z)$ is defined only at $\frac{1}{2}$ since it can be modified elsewhere without altering the payoff. Therefore an optimal strategy σ in in $\Gamma(z)$ can play anything in $\frac{1}{2} - \epsilon$ and in $\frac{1}{2} + \epsilon$ since no regularity for σ is required. The good way to use the proximity between z and z' is to always play as if the initial distribution was $\frac{1}{2}$. Here we have to define σ' such that for all $z \in \Delta(K)$, $\sigma'(z) = \sigma(\frac{1}{2})$.*

---

[3]Using a continuum of alternatives is clearly unnecessary but allows to simplify the proof.

Lemma 5.4 (Splitting procedure). *The mapping $\hat{v}_\theta(z)$ is concave on $\Delta_f(\Delta(K))$.*

Proof. We follow the same scheme as for games with incomplete information on one side (see e.g. [5] Corollary 1.3 p.184). Let $\lambda \in [0,1]$ and let Y be a random variable with values in $\{0,1\}$, such that $\mathbb{P}(Y = 0) = \lambda$. Let $z, z' \in \Delta_f(\Delta(K))$. The random variable P is selected according to the distribution z if $Y = 0$ and z' if $Y = 1$, the state variable k_1 is finally selected according to p if $P = p$. Compare now the two following situations: on one hand, the game with initial signals (Y, P) for player 1 and nothing for player 2 and on the other hand the game with initial signals (Y, P) for player 1 and Y for player 2. These two distributions of initial signals and states fulfill our assumptions and it's clear that the value of the second is less or equal than the value of the first for any evaluation $\theta \in \Delta_f(\mathbb{N}^*)$ since the set of behavior strategies of player 2 in the second game is larger than in the first game. Translating this inequality using $\hat{v}$, we deduce directly

$$\hat{v}_\theta(\delta_{\lambda z + (1-\lambda)z'}) \geq \hat{v}_\theta(\lambda \delta_z + (1-\lambda)\delta_{z'}) = \lambda \hat{v}_\theta(z) + 1 - \lambda \hat{v}_\theta(z'),$$

which proves the Lemma. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

5.2 Auxiliary game $\mathcal{G}$

Let $X = \Delta_f(\Delta(K))$ be the state space, which corresponds to player 2's belief about player 1's belief about the current state. It is a convex relatively compact subset of a normed vector space and we are going to express the auxiliary game and the recursive formula on this state space.

Let $\mathcal{G}$ be the stochastic game defined by

- the state space $X = \Delta_f(\Delta(K))$,

- the action space $A = \{f : \Delta(K) \to \Delta(\mathcal{I}), \text{measurable}\}$ for player 1,

- the action space $B = \Delta(J)$ for player 2,

- the payoff function $G : X \times A \times B \to [0,1]$ defined, for any $z \in X$ by

$$G(z, a, b) = \sum_{p \in supp(z)} \sum_{(i,j) \in I \times J} b(j)a(p,i)g(p,i,j))z(p),$$

  where supp(z) stands for the support of z,

- the transition function $\ell : X \times A \times B \to \Delta_f(X)$ is defined as $\ell(z, a, b) = \Phi(Q(z, a, b))$, where $Q(z, a, b) \in \Delta_f((K) \times (\Delta(K) \times C) \times (D))$ is the induced joint distribution of $(k_2, (p, i_1, c_2), (j_1, d_2))$ in the canonical game $\widehat{\Gamma}(\delta_z)$ where players play at the first stage $\sigma_1 = a$ and $\tau_1 = b$. The sets C, D, K and supp(z) being finite and using assumptions $(A1)$ and $(A2)$, we may consider Q as an element in $\Delta_f^*(K \times \mathbb{N} \times \mathbb{N})$.

Let us recall the definition of Choquet order on $\Delta_f(X)$.

Definition 5.5. *The order $\leq$ on $\Delta_f(X)$ called (reversed) Choquet order is defined by the relation*

$$\mu \leq \nu \Leftrightarrow \text{For all continuous concave function on } X, \ \tilde{f}(\mu) \leq \tilde{f}(\nu).$$

24

We aim to apply a weakened version of Renault [10] to the game $\mathcal{G}$, thus let us first recall the hypotheses of the Theorem as they appear in the original article.

Hypotheses 5.1.

H1) The map ℓ does not depend on b.

H2) X is a compact convex subset of a normed vector space,

H3) A and B are convex compact subsets of some topological vector spaces,

H4) $(a \mapsto G(z, a, b))$ is concave upper semi-continuous $\forall (z, b) \in X \times B$ and $(b \mapsto G(z, a, b))$ is convex and lower semi-continuous $\forall (z, a) \in X \times A$.

H5) There exists a subset $\mathcal{C}$ of 1-Lipschitz functions containing $\phi(1, 0)$ such that for all f in $\mathcal{C}$, $\alpha \in [0, 1]$, the function $\phi(\alpha, f)$ is in $\mathcal{C}$, where $\phi(\alpha, f)$ is defined by

$$\forall z \in \Delta_f(X) \ \phi(\alpha, f)(z) = \sup_{a \in A} \min_{b \in B} \left\{ \alpha G(z, a, b) + (1 - \alpha)\tilde{f}(\ell(z, a)) \right\}.$$

H6) The mapping $a \mapsto \ell(z, a)$ is concave for the Choquet order and continuous.

H7) (Splitting assumption) Let z be a convex combination in $\Delta_f(\Delta(K))$, $z = \sum_{s=1}^{S} \lambda_s z_s$ and $(a_s)_{s \in S}$ be a family of actions in A^S. Then there exists $a \in A$ such that

$$\ell(z, a) \geq \sum_{s \in S} \lambda_s \ell(z_s, a_s) \ and \ \min_{b \in B} G(z, a, b) \geq \sum_{s \in S} \lambda_s \min_{b \in B} G(z_s, a_s, b).$$

The main consequence of assumption ($A3$) is that player 2 cannot influence the transition in the auxiliary game so the map ℓ does not depend on b, i.e.

$$\forall (z, a) \in X \times A, \forall b, b' \in B, \ \ell(z, a, b) = \ell(z, a, b').$$

Thus ($H1$) is satisfied and from now on, we will work under the shorter notation $l(z, a)$ for $l(z, a, b)$.

The hypotheses ($H2, H3, H4, H6, H7$) ensure the application of Sion's theorem in several steps of Renault's proof. Here they are not all satisfied since, for example, the set A is not compact. However, it is well known that adding some geometrical hypotheses allows to weaken the topological assumptions in Sion's theorem (see, for instance, Proposition A.8 in Sorin's monography [13]). For instance, if A is a convex set, B is a compact convex subset of a topological vector space, $(a \mapsto G(z, a, b))$ is concave $\forall (z, b) \in X \times B$ and $(b \mapsto G(z, a, b))$ is convex and lower semi-continuous $\forall (z, a) \in X \times A$, Sion's result applies to the one-stage game: the game $\mathcal{G}_1(z)$ has a value. They can be replaced without altering the proof by the following hypotheses.

Hypotheses 5.2.

H2') X is a relatively compact convex subset of a normed vector space.

H3') B is a convex compact subset of a topological vector space, A is a convex set.

H4') $(a \mapsto G(z, a, b))$ is concave $\forall (z, b) \in X \times B$ and $(b \mapsto G(z, a, b))$ is convex and lower semi-continuous $\forall (z, a) \in X \times A$.

H6') The mapping $a \mapsto \ell(z, a)$ is concave for the Choquet order.

Assumption $(H2')$ is satisfied since the Wasserstein distance can be extended to a norm on the space of finite signed measures. Moreover assumptions $(H3')$ and $(H4')$ are clearly satisfied. Therefore, we need to prove $(H6')$ and $(H7)$.

Lemma 5.6. *The game $\mathcal{G}$ fulfills $H6'$ and $H7$.*

Proof. Let z be a convex combination in X, $z = \sum_{s=1}^{S} \lambda_s z_s$ and $(a_s)_{s \in S}$ be a family of actions in A^S. Denote $\mu(z_s, a_s) \in \Delta_f(\Delta(K) \times I)$ the joint law induced on $\Delta(K) \times I$ by (z_s, a_s). By disintegration, there exists $a \in A$ such that $\mu(z,a) = \sum_{s \in S} \lambda_s \mu(z_s, a_s)$. A first, note that $Q(z,a,b) = \sum_{s \in S} \lambda_s Q(z_s, a_s, b)$.

Given (z, a, b), we consider the canonical game $\hat{\Gamma}(z)$. In this game, a pair (k, p) is chosen according to the probability $\pi \in \Delta_f(K \times \Delta(K))$ defined by $\pi(k,p) = p^k z(p)$ for all $(k, p) \in K \times \Delta(K)$. Then, player 1 receives the signal $c_1 = p$ and player 2 receives no initial signal. We associate to (a, b) a pair strategies for the first stage (σ_1, τ_1) by $\sigma_1(p) = a(p)$ and $\tau_1 = b$. Then, $Q(z,a,b)$ denotes the joint distribution of $(k_2, (p, i_1, c_2), (j_1, d_2))$. Since the conditional law of (k_2, c_2, d_2) given (p, i_1, j_1) is $q(p, i_1, j_1)$, it follows that $Q(z,a,b)$ is bilinear with respect to $(\mu(z,a), b)$ (with abusive notations).

Let $\rho = Q(z, a, b)$ (resp. $\rho_s = Q(z_s, a_s, b)$) and $c' = (p, i_1, c_2)$ (resp. $d' = (j_1, d_2)$). Let $C' = (\cup_{s \in S} supp(z_s)) \times I \times C$ and $D' = J \times D$. By construction,

$$\ell(z,a) = \Phi(\rho) = \mathcal{L}_\rho(\mathcal{L}_\rho(\mathcal{L}_\rho(k_2 \mid c') \mid d')) = \sum_{d' \in D'} \rho(d') \delta_{\mathcal{L}_\rho(\mathcal{L}_\rho(k_2 \mid c') \mid d')}.$$

Using Lemma 4.6, we have the following equality ρ-almost surely

$$\mathcal{L}_\rho(k_2 \mid p, i_1, c_1) = F(p, i_1, c_2).$$

This implies that

$$\mathcal{L}_\rho(\mathcal{L}_\rho(k_2 \mid c'), d') = \mathcal{L}_\rho(F(c'), d') \in \Delta_f(C'' \times D')$$

where $C'' = F(C')$. A similar equality holds with ρ_s instead of ρ for all $s \in S$. By definition of $l(z,a)$, we deduce that

$$\ell(z,a) = \mathcal{L}_\rho(\mathcal{L}_\rho(F(c') \mid d')) = \Psi(\mathcal{L}_\rho(F(c'), d')),$$

where Ψ is the disintegration map defined by

$$\Psi : \Delta(C'' \times D') \to \Delta_f(\Delta(C'')) : m \to \sum_{d' \in D'} m(d') \delta_{\mathcal{L}_m(c'' \mid d')},$$

where $\mathcal{L}_m(c'' \mid d')$ denoted the conditional law of c'' given d'. It was proved in Renault [10] (Lemma 4.16) that Ψ is concave for the Choquet order on $\Delta_f(\Delta(C''))$. However, C'' being a finite subset of $\Delta(K)$, $\Delta(C'')$ is identified as a compact convex subset of X. It follows easily that the convex order on $\Delta_f(\Delta(C''))$ coincides with the order induced by the convex order on $\Delta_f(X)$.

We conclude that the first part of H7 holds since

$$l(z,a) = \Psi(\sum_{s \in S} \lambda_s \mathcal{L}_{\rho_s}(F(c'), d')) \geq \sum_{s \in S} \lambda_s \Psi(\mathcal{L}_{\rho_s}(F(c'), d')) = \sum_{s \in S} \lambda_s l(z_s, a_s).$$

For the second part of H7, it is sufficient to note that (again with abusive notations) $\mu(z,a) \mapsto G(z,a,b)$ is linear so that for all $b \in B$

$$G(z,a,b) = \sum_{s \in S} \lambda_s G(z_s, a_s, b),$$

which implies the result. Finally, in case $z_s = z$ for all s, the same arguments also imply $(H6')$ since in this case one can choose $a = \sum_{s \in S} \lambda_s a_s$ in the above proof. $\square$

The proof of the following proposition follows from Proposition 3.21 in Renault [10].

Proposition 5.7. *Assuming* $(H1, H2', H3', H4', H6', H7)$*, then for any* $\theta \in \Delta_f(\mathbb{N}^*)$ *and any* $\eta \in \Delta_f(X)$*, the game* $\mathcal{G}_\theta(\eta)$ *has a value* $w_\theta(\eta)$ *such that*

$$\forall z \in X, w_\theta(z) = \sup_{a \in A} \min_{b \in B} \{\theta_1 G(z,a,b) + (1-\theta_1) w_{\theta^+}(\ell(z,a))\}, \tag{5.1}$$

$$= \min_{b \in B} \sup_{a \in A} \{\theta_1 G(z,a,b) + (1-\theta_1) w_{\theta^+}(\ell(z,a))\}, \tag{5.2}$$

where θ^+ *is defined by* $\theta_t^+ = \frac{\theta_{t+1}}{\sum_{m \geq 2} \theta_m}$ *for* $t \geq 1$ *whenever* $\sum_{m \geq 2} \theta_m > 0$ *and is defined arbitrarily otherwise. Moreover, in* $\mathcal{G}_\theta(\eta)$*, player 1 has* ε*-optimal Markov strategies for all* $\varepsilon > 0$ *and player 2 has optimal Markov strategies.*

In order to prove the last assumption $(H5)$, we first prove that the value of the game $\mathcal{G}_\theta$ is equal to the canonical value function $\hat{v}_\theta$. Since we proved that the canonical value is 1-Lipschitz, it will imply using the previous Proposition that the set of functions $\mathcal{C} = \{v_\theta, \theta \in \Delta_f(\mathbb{N}^*)\}$ satisfies $(H5)$.

We now prove that the value functions of both games are the same. The proof is classic and consists to show that both families of functions are linked by the same recursive formula.

Proposition 5.8. *For all* $\theta \in \Delta_f(\mathbb{N}^*)$ *and for any* $z \in X$*,* $w_\theta(z) = \hat{v}_\theta(z)$*.*

Corollary 5.9. *The game* $\mathcal{G}$ *fulfills* $(H5)$*.*

Proof of Proposition 5.8. Notice first that $\hat{v}_1(z) = w_1(z)$ for all z. This comes indeed almost from the definition

$$\hat{v}_1(z) = \sup_{\sigma_1 : \Delta(K) \to \Delta(I)} \inf_{b \in \Delta(J)} \int_{\Delta(K)} g(p, \sigma_1(p), b) dz(p)$$

$$= \sup_{a \in A} \min_{b \in \Delta(J)} G(z,a,b)$$

$$= \min_{b \in \Delta(J)} \sup_{a \in A} G(z,a,b)$$

$$= w_1(z).$$

It is enough to prove that w and $\hat{v}$ satisfy the same recurrence formula. We will prove that $\hat{v}$ satisfies the recurrence formula in $\mathcal{G}$, i.e.

$$\hat{v}_\theta(z) = \sup_{a \in A} \min_{b \in B} \theta_1 G(z,a,b) + (1-\theta_1) \hat{v}_{\theta^+}(\ell(z,a))$$

$$= \min_{b \in B} \sup_{a \in A} \theta_1 G(z,a,b) + (1-\theta_1) \hat{v}_{\theta^+}(\ell(z,a)).$$

27

We prove the recursive formula by induction on the greatest element in the support of θ. If $\theta = \delta_1$, it follows from the preceding equality. Fix now $n \geq 2$, and assume that the proposition is true for every θ supported by $\{1, ..., n-1\}$. Let $z \in \Delta_f(\Delta(K))$. We first prove that player 1 can defend in $\hat{\Gamma}_\theta(z)$ the quantity

$$\min_{b \in B} \sup_{a \in A} \left(\theta_1 G(z, a, b) + (1 - \theta_1) \hat{v}_{\theta+}(\ell(z, a)) \right).$$

Using the canonical representation $\hat{\Gamma}$, $\hat{v}_\theta(z) = v_\theta(\pi)$ where $\pi \in \Delta_f(K \times \Delta(K) \times \Delta_f(\Delta(K)))$ is defined by

$$\forall (k, p, x) \in K \times \Delta(K) \times \Delta_f(\Delta(K)), \ \pi(k, p, x) = p(k) z(p) \mathbb{1}_{x=z}.$$

Consider the game $\Gamma_\theta(\pi)$. Let $\varepsilon > 0$ and τ be a strategy of player 2. Denoting by b the law induced by τ_1, let $a^* \in A$ an action which realizes the supremum up to ε in the expression

$$\theta_1 G(z, a, b) + (1 - \theta_1) \hat{v}_{\theta+}(\ell(z, a)).$$

Let σ^* be an ε-optimal strategy in the game $\Gamma_{\theta+}(Q(z, a^*, b))$. Define then σ by $\sigma_1 = a^*$ and for all $n \in \mathbb{N}^*$, $h_n^I = (p, i_1, c_2, ..., i_{n-1}, c_n)$, $\sigma_n(h_n^I) = \sigma_{n-1}^*(c', h_{n-1}^{1,+})$ where $c' = (p, i_1, c_2)$ and $h_{n-1}^{1,+} = (i_2, c_3, .. i_{n-1}, c_n)$. We have

$$\gamma_\theta(\mu, \sigma, \tau) = \theta_1 G(z, a^*, b) + (1 - \theta_1) \gamma_{\theta+}(Q(z, a^*, b), \sigma^*, \tau^+),$$

where τ^+ is a continuation strategy. Precisely, for all $n \in \mathbb{N}^*$, $\tau_{n-1}^+(d', h_{n-1}^{2,+}) = \tau_n(h_n^{II})$ with $h_{n-1}^{2,+} = (j_2, d_3, .., j_{n-1}, d_n)$, $h_n^{II} = (d', h_{n-1}^{2,+})$ and $d' = (j_1, d_2)$ is the "signal" for player 2 given by $Q(z, a^*, b)$.

Therefore, σ^* and τ^+ can be seen as behavior strategies in a new game with initial signals corresponding to the past history in the original game and since σ^* is ε-optimal in $\Gamma(Q(z, a^*, b))$, we have

$$\begin{aligned}
\gamma_\theta(\mu, \sigma, \tau) &\geq \theta_1 G(z, a^*, b) + (1 - \theta_1) v_{\theta+}(Q(z, a^*, b)) - \varepsilon \\
&\geq \sup_{a \in A} \theta_1 G(z, a, b) + (1 - \theta_1) v_{\theta+}(Q(z, a, b)) - 2\varepsilon \\
&= \sup_{a \in A} \theta_1 G(z, a, b) + (1 - \theta_1) \hat{v}_{\theta+}(\ell(z, a)) - 2\varepsilon \\
&\geq \min_{b \in B} \sup_{a \in A} \theta_1 G(z, a, b) + (1 - \theta_1) \hat{v}_{\theta+}(\ell(z, a)) - 2\varepsilon.
\end{aligned}$$

It follows that $\hat{v}_\theta(z) \geq \min_{b \in B} \sup_{a \in A} \theta_1 G(z, a^*, b) + (1 - \theta_1) \hat{v}_{\theta+}(\ell(z, a))$ by sending ε to zero.

Let us show that player 2 can defend $\sup_{a \in A} \min_{b \in B} (\theta_1 G(z, a, b) + (1 - \theta_1) \hat{v}_{\theta+}(\ell(z, a))$ in $\Gamma(\mu)$. Fix a strategy σ of player 1 and let $a = \sigma_1$, there exists $b^* \in B$ achieving $\min_b G(z, a, b)$. We also choose τ^* an optimal strategy for player 2 in the game $\Gamma_{\theta+}(Q(z, a, b^*))$. This defines a strategy τ such that

$$\begin{aligned}
\gamma_\theta^\mu(\sigma, \tau) &= \theta_1 G(z, a, b^*) + (1 - \theta_1) \gamma_{\theta+}^{Q(z,a,b^*)}(\sigma^+, \tau^*) \\
&\leq \theta_1 G(z, a, b^*) + (1 - \theta_1) v_{\theta+}(Q(z, a, b^*)) \\
&= \theta_1 G(z, a, b^*) + (1 - \theta_1) \hat{v}_{\theta+}(\ell(z, a)) \\
&= \min_{b \in B} \theta_1 G(z, a, b) + (1 - \theta_1) \hat{v}_{\theta+}(\ell(z, a)).
\end{aligned}$$

Thus $\hat{v}_\theta(z) \leq \sup_{a \in A} \min_{b \in B} (\theta_1 G(z, a, b) + (1 - \theta_1) \hat{v}_{\theta+}(\ell(z, a))$. Finally, since the maxmin is always smaller than the minmax, all the intermediate inequalities are equalities. $\square$

5.3 Existence of the uniform value

Let us at first recall the first main result proved in [10] which holds under our set of weakened assumptions.

Theorem 5.10 (Renault(2012))**.** *Assume that* $H1, H'2, H'3, H'4, H5, H'6, H7$ *hold. Then for every initial distribution* $\eta \in \Delta_f(X)$*, the game has a uniform value* $w^*(\eta)$*. Moreover player 1 can guarantee* $w^*(\eta)$ *with a Markov strategy:*

$$\forall \epsilon > 0, \ \exists \sigma \in \Sigma^M, \exists N_0 \in \mathbb{N}, \ \forall N \geq N_0 \ \forall \tau' \in \mathcal{T}, \ \gamma_N(\eta, \sigma, \tau') \geq w^*(\eta) - \epsilon.$$

and we have $w^*(\eta) = \inf_{n \geq 1} \sup_{m \geq 0} w_{m,n}(\eta)$*.*

In order to conclude the proof, we show that both players can guarantee

$$v^*(\pi) = \inf_{n \geq 1} \sup_{m \geq 0} v_{m,n}(\pi),$$

where $v_{m,n}(\pi) = v_{\theta_{m,n}}(\pi)$ and $\theta_{m,n}$ is the uniform law between stage m and $m + n$.
The game $\mathcal{G}(z)$ satisfies assumptions $H1, ..., H7'$ so it has a uniform value given by

$$w^*(z) = \inf_{n \geq 1} \sup_{m \geq 0} w_{m,n}(z).$$

And by proposition 5.8, the value in $\mathcal{G}$ and in the reduced game are equal, so if $\pi \in \Delta_f^*(K \times \mathbb{N} \times \mathbb{N})$ we have

$$v^*(\pi) = \inf_{n \geq 1} \sup_{m \geq 0} v_{m,n}(\pi) = \inf_{n \geq 1} \sup_{m \geq 0} \hat{v}_{m,n}(\Phi(\pi)) = \inf_{n \geq 1} \sup_{m \geq 0} w_{m,n}(\Phi(\pi)) = w^*(\Phi(\pi)),$$

Thus player 1 can guarantee $v^*(\pi)$ in $\mathcal{G}(\Phi(\pi))$ with a Markov strategy. Let us check that he can guarantee $v^*(\pi)$ in the game $\widehat{\Gamma}(\Phi(\pi))$ or equivalently in $\Gamma(\pi)$.

Proposition 5.11. *Any Markovian strategy* $\widehat{\sigma}$ *of player 1 in* $\mathcal{G}_\infty(z)$ *induces a strategy* σ *in* $\hat{\Gamma}_\infty(z)$ *guaranteeing the same amount.*

Proof. Let $\widehat{\sigma}$ be a behavior strategy in $\mathcal{G}_\infty(z)$. Let us describe the strategy σ. Player 1 plays at the first round in $\Gamma_\infty(z)$ the mixed action $\widehat{\sigma}_1(z)(p)$ where p is his initial signal. Then, at round n, he plays the mixed action $\widehat{\sigma}_n(y_n)(x_n)$. That this strategy is a well-defined strategy follows from Lemma 4.9.

It remains to prove that this strategy guarantees the same quantity as $\widehat{\sigma}$. Let us fix $n \in \mathbb{N}^*$, we will prove that there exists a best reply $\widetilde{\tau}$ to σ in $\widehat{\Gamma}_n(z)$ which can be seen as a strategy $\hat{\tau}$ in $\mathcal{G}_\theta(z)$ and such that

$$\gamma_\theta(z, \sigma, \widetilde{\tau}) = \widehat{\gamma}_\theta(z, \widehat{\sigma}, \widehat{\tau}).$$

We will proceed by backward induction. Let us fix a best reply τ to σ in $\widehat{\Gamma}_\theta(z)$. We will construct a strategy $\widetilde{\tau}$ which depends at stage m on h_m^{II} only through y_m. Recall that σ is fixed so that $y_m(h_m^{II})$ can be computed by player 2. At first let us replace τ_n by

$$\widetilde{\tau}_n(y_n) = \mathbb{E}_{\mathbb{P}_{\sigma\tau}^z}[\tau(h_n^{II}) \mid y_n].$$

Note that this conditional expectation depends on the strategies σ, τ up to stage $n-1$. Let us prove that the payoff at the last stage n is not modified.

$$
\begin{aligned}
\mathbb{E}_{\mathbb{P}^z_{\sigma\tau}}[g(k_n, i_n, j_n)] &= \mathbb{E}_{\mathbb{P}^z_{\sigma\tau}}[\mathbb{E}_{\mathbb{P}^z_{\sigma\tau}}[g(k_n, i_n, j_n) \mid h^I_n, h^{II}_n]] \\
&= \mathbb{E}_{\mathbb{P}^z_{\sigma\tau}}[\gamma_1(x_n, \sigma_n(y_n, x_n), \tau_n(h^{II}_n))] \\
&= \mathbb{E}_{\mathbb{P}^z_{\sigma\tau}}[\mathbb{E}_{\mathbb{P}^z_{\sigma\tau}}[\gamma_1(x_n, \sigma_n(y_n, x_n), \tau_n(h^{II}_n)) \mid h^{II}_n]] \\
&= \mathbb{E}_{\mathbb{P}^z_{\sigma\tau}}[\int \gamma_1(x, \sigma_n(y_n, x), \tau_n(h^{II}_n))dy_n[x]] \\
&= \mathbb{E}_{\mathbb{P}^z_{\sigma\tau}}[\mathbb{E}_{\mathbb{P}^z_{\sigma\tau}}[\int \gamma_1(x, \sigma_n(y_n, x), \tau_n(h^{II}_n))dy_n[x] \mid y_n]] \\
&= \mathbb{E}_{\mathbb{P}^z_{\sigma\tau}}[\int \gamma_1(x, \sigma_n(y_n, x), \mathbb{E}_{\mathbb{P}^z_{\sigma\tau}}[\tau_n(h^{II}_n) \mid y_n])dy_n[x]] \\
&= \mathbb{E}_{\mathbb{P}^z_{\sigma\tau}}[\int \gamma_1(x, \sigma_n(y_n, x), \widetilde{\tau}_n(y_n))dy_n[x]] \\
&= \mathbb{E}_{\mathbb{P}^z_{\sigma,(\tau_1,...,\tau_{n-1},\widetilde{\tau}_n)}}[g(k_n, i_n, j_n)].
\end{aligned}
$$

The above equations show that the expected payoff at stage n when player 2 is playing the best reply $(\tau_1, ..., \tau_{n-1}, \widetilde{\tau}_n)$ against σ is a function of σ and of the law of y_n. Assume now that at step m, we have proved that there exists a best reply to σ of player 2 such that the sum of expected payoffs for the stages $m+1, ..., n$ is a function of σ and of the law of $(y_{m+1}, .., y_n)$ only. We can replace $\tau_m(h^{II}_m)$ by $\widetilde{\tau_m}(y_m) = \mathbb{E}_{\mathbb{P}^z_{\sigma\tau}}[\tau_m(h^{II}_m) \mid y_m]$ without modifying the expected payoff of stage m with the same argument as above. Using assumption $(A3)$, Lemma 4.9 and the definition of σ, the law of $(y_{m+1}, ...y_n)$ is not modified by this operation which proves that this modified strategy is still a best reply to σ. $\qquad\square$

Secondly, we prove that Player 2 can guarantee $v^*(\pi)$ by splitting the stage in blocks and playing on each block separately since he has no influence on the transition. The following results are quite similar to the corresponding ones proved in Renault [10] and are reproduced here since their proofs are very short.

Lemma 5.12. *For every $\pi \in \Delta^*_f(K \times C' \times D')$, $n \geq 1$ and $m \geq 1$, $\forall \tau_1, ..., \tau_m$, $\exists \tau_{m+1}, ..., \tau_{m+n}$ such that the strategy $\tau_1, ..., \tau_m, ..., \tau_{m+n}$ of player 2 is optimal in the game $\Gamma_{m,n}$.*

Proof. Let $\pi \in \Delta(K \times C' \times D')$, $n \geq 1$ and $m \geq 0$, and $\tau_1, ..., \tau_m$ such that $\tau_i : D' \times (J \times D)^{i-1} \to \Delta(J)$. We define $\mathcal{T}^*$ the subset strategies of player 2 which start with $\tau_1, ..., \tau_m$ and we consider the game with the evaluation $\theta_{m,n}$ and the set of strategies Σ and $\mathcal{T}^*$. It can be seen as the mixed extension of a finite game, thus the value exists and will be denoted $v^*_{m,n}(\pi)$. Since the set of strategies of player 2 is smaller than $\mathcal{T}$, we have $v_{m,n}(\pi) \leq v^*_{m,n}(\pi)$. But using the same method as for proving the recursive formula of Proposition 5.8, for any σ, we can build a strategy which defends $v_{m,n}(\pi)$. Both values are therefore equal and any optimal strategy in the restricted game satisfies the conclusion of the lemma. $\qquad\square$

Proposition 5.13. *For every $\pi \in \Delta^*_f(K \times C' \times D')$, player 2 can guarantee $v^*(\pi)$ in the game $\Gamma_\infty(\pi)$.*

Proof. We prove that for all $n \in \mathbb{N}$, Player 2 can guarantee the payoff $\sup_{m \geq 0} v_{m,n}(p)$. Let $n \in \mathbb{N}$ be a number of stages, then for each $L \in \mathbb{N}$ we split the game of length nL in L blocks of length n: $B_1, ..., B_L$. We define the strategy τ^* by induction on the block.

Let τ be an optimal strategy in $\Gamma_{1,n}(\pi)$ then we set $\tau_i^* = \tau_i$ for all $i \in \{1, .., n\}$. Once we have constructed $\tau_1^*, ..., \tau_{nl}^*$ for some $1 \leq l \leq L - 1$, we define the game $\Gamma_{nL+1,n}^{\#}(\pi)$ where the player 2 has to play τ_i^* for all $i \leq nl$. We have $v_{nL+1,n}^{\#}(\pi) = v_{nL+1,n}(\pi)$ using the preceding Lemma. Let τ be an optimal strategy in $\Gamma_{nL+1,n}^{\#}(\pi)$ and set $\tau_i^* = \tau_i$ for all $i \in \{nL + 1, ..., (n + 1)L\}$. We have

$$\gamma_{Ln}(\sigma, \tau^*) = \frac{1}{nL} \mathbb{E}_{\sigma\tau^*}^{\pi} \left(\sum_{m=0}^{Ln} g(k_m, i_m, j_m) \right) = \frac{1}{nL} \sum_{d=0}^{L-1} \mathbb{E}_{\sigma\tau^*}^{\pi} \left(\sum_{m=dn+1}^{(d+1)n} g(k_m, i_m, j_m) \right)$$

$$\leq \frac{1}{L} \sum_{d=0}^{L-1} v_{dn+1,n}(\pi) \leq \frac{1}{L} \sum_{d=0}^{L-1} \sup_{m \geq 0} v_{m,n}(\pi)$$

$$\leq \sup_{m \geq 0} v_{m,n}(\pi).$$

The payoff being bounded, we deduce that this strategy guarantees $v_{m,n}(\pi)$. Finally, Player 2 can guarantee the minimum on $n \in \mathbb{N}$, $\inf_{n \in \mathbb{N}} \sup_{m \geq 0} v_{m,n}(\pi) = v^*(\pi)$. $\square$

Since each player can guarantee $v^*(\pi)$, the game has a uniform value given by $v^*(\pi)$ which concludes the proof of Theorem 2.3.

References

[1] R. J. Aumann and M. Maschler, *Repeated games with incomplete information*, with the collaboration of R. Stearns, Cambridge, MA: MIT Press, 1995.

[2] D. Blackwell and L.E. Dubins, *An extension of Skorohod's almost sure representation theorem*, Proceedings of the American Mathematical Society, 1983, **89**, 691-692.

[3] J. F. Mertens, *Repeated Games*, Proceedings of the ICM. Berkeley, California, 1986.

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

[5] J. F. Mertens, S. Sorin and S. Zamir, *Repeated Games*, CORE Discussion Papers 9420, 9421 and 9422, Universite Catholique De Louvain, Belgium, 1994.

[6] A. Neyman and S. Sorin, *Repeated games with public uncertain duration process*, International Journal of Game Theory, 2010, **39**, 29-52.

[7] A. Neyman, *Existence of optimal strategies in Markov games with incomplete information*, International Journal of Game Theory, 2008, **37**, 581-596.

[8] J. Renault, *The value of Markov chain games with lack of information on one side*, Mathematics of Operations Research, 2006, **31**, 490-512.

[9] J. Renault, *Uniform value in Dynamic Programming*, Journal of the European Mathematical Society, 2011, **13**, 309-330.

[10] J. Renault, *The value of repeated games with an informed controller*, Mathematics of Operations Research, 2012, **37**, 154-179.

[11] D. Rosenberg, E. Solan and N. Vieille, *Blackwell optimality in Markov decision processes with partial observation*, Annals of statistics, 2002, **30**, 1178-1193.

[12] D. Rosenberg, E. Solan and N. Vieille, *Stochastic games with a single controller and incomplete information*, SIAM journal on control and optimization, 2004, **43**, 86-110.

[13] S. Sorin, *A first course on zero-sum repeated games*, Mathématiques & Applications, Springer, 2002.

[14] C. Villani, *Topics in Optimal Transportation.* Graduate Studies in Mathematics, Vol. 58. American Mathematical Society, 2003.