SECRET CORRELATION IN REPEATED GAMES WITH IMPERFECT MONITORING

OLIVIER GOSSNER AND TRISTAN TOMALA

Abstract. We characterize the maximum payoff that a team can guarantee against another in a class of repeated games with imperfect monitoring. Our result relies on the optimal trade-off for the team between optimization of stage-payoffs and generation of signals for future correlation.

JEL classification: C72, D82

Date: 11th May 2004.

1. Introduction

It many strategic situations, a group of players may find it beneficial to coordinate their action plans in a way that is hidden from other players. The manager of a sport team devices coordinated plans for the team members. Generals of allied armies need to keep their coordinated plans secret from ennemies. On the internet, coordinated attacks of systems (e.g. by viruses) are known to be much more dangerous than uncoordinated ones. The management of a firm coordinates the actions of the units of production in a way that is hidden from the competitors.

Coordination of a group of players needs to rely on the observation of a common signal by its members. This signal can arise from an external correlation device (Aumann, [Aum74]), or be the result of communication between the players (Forges, [For86]). In a repeated game with imperfect monitoring, players observe random and correlated signals (deterministic signals is a particular case) that depend on chosen actions. These games therefore feature both correlated signals and communication possibilities.

This article explores the possibilities of secret correlation between team members in a repeated game with imperfect monitoring. Our model opposes a team of players called team I to another team called team II viewed as a single player. Team I's member's action sets are denoted $A^i, i \in I$, and team II's action set is B. At each stage, team II observes a (possibly random) signal s about I's action profile a, drawn according to some probability distribution $q(s|a)$. Team I's members are informed of a, s, and possibly of II's actions (our result covers the cases in which team I has perfect, imperfect, or no observation of II's choice). The payoff to team I is a function of both team's action choices. In order to stress the value of secret correlation between team members, we assume that team II's goal is to minimize team I's payoff. Since team I has more information than team II about action choices, this extra information can be used as a correlation device for future actions. Our model allows to study the optimal trade-offs for team I between generation of signals for future correlation and use of correlation for present payoffs.

Our main result is a characterization of the best payoff that the team can guarantee against outside players as either the horizon of the game grows to infinity of the discount factor goes to one. We emphasize three reasons why characterizing the max min value is important.

First, the max min of the repeated game measures how successful team I is in correlating secretly its actions from outside players. Indeed, when no correlation is possible, the max min of the repeated game coincides with the max min in mixed strategies of the stage game. When full correlation is achievable, this max min equals the generally higher max min in correlated strategies of the stage game. In general, partial correlation only may be achievable, and the max min of the repeated game may lie between these two values.

Second, von Stengel and Koller [vSK97] proved that, in finite games opposing a team of players to one outside player, the max min payoff is a Nash payoff. Furthermore, it is the most natural Nash payoff to select since team members can guarantee this value. Combined with our result, we know that the maximal Nash payoff to the team in the repreated game with imperfect monitoring is the max min we characterize.

Finally, characterizations of max min payoffs in repeated games, are important for the general study of repeated games with imperfect monitoring. Indeed, the message of the "Folk Theorem" (see e.g. Fudenberg and Maskin [FM86]) is that in repeated games with perfect monitoring and sufficiently little discounting, the set of equilibrium payoffs is given by the set of feasible and individually rational payoffs of the one-shot game. Such payoffs can be enforced by a plan in which players follow a path generating the desired payoff, and any player that deviates from this plan gets punished by the others to his individually rational level. Generalizations of the "Folk Theorem" to games with imperfect monitoring yield two types of questions. First, the signalling structure may render deviations undetectable (e.g. Lehrer [Leh90]), so one needs to characterize detectable deviations. And second, assuming it is commonly known that a player has deviated, how harsh can this player be punished? This last question amounts to characterizing the $\min\max$ of player i in the repeated game, or the $\max\min$ of the team of players trying to minimize i's payoff.

The problem faced by the team consists in finding the optimal trade-off between using previous signals that are unknown to team II as correlation devices, and generating such signals for future use. We measure the amount of secret information contained in past signals by their entropy. Our main result characterizes the team's $\max\min$ payoff as the best payoff that can be obtained by a convex combination of correlated strategies under the constraint that the average entropy spent by the correlation devices does not exceed the average entropy of secret signals generated.

We motivate the problem by discussing examples in section 2, present the model and tools in section 3, and the main result in section 4. We illustrate our result in the simple cases of perfect and trivial observations in section 5, discuss computational applications in section 6, show an example of a signalling structure for which a folk theorem obtains in section 7, and conclude with extensions in section 8.

2. Examples

We consider a 3-player game where the team is $I = \{1, 2\}$ opposing player $II = \{3\}$. Player 1 chooses rows, player 2 chooses columns and player 3 chooses matrices. Consider the following one-shot game:

$$\begin{array}{ccc} & a & b \\ a & 1 & 0 \\ b & 0 & 0 \end{array} \qquad \begin{array}{cc} a & b \\ 0 & 0 \\ 0 & 1 \end{array}$$
$$L \qquad\qquad R$$

In the repeated game with perfect monitoring, the team guarantees the max min of the one-shot game, where the max runs over the independent probability distributions on $A^1 \times A^2$ that is, the team guarantees $\frac{1}{4}$.

Assume now that player 3 receives blank signals, i.e. has no information on the action profile of I, whereas players 1 and 2 observe each other's actions. Player 1 can then choose his first action uniformly, and the team can correlate their moves from stage 2 on according to player 1's first action. This way, the team guarantees $\frac{1}{2}$ from stage 2 on, where $\frac{1}{2}$ is computed as the max min of the one-shot game, where the max runs over the set of all probability distributions on $A^1 \times A^2$.

Consider now the case where team members observe each other's actions and the signal of player 3 is given by the following matrix:

$$\begin{array}{cc} a & b \end{array}$$

$$\begin{array}{cc} a & \begin{pmatrix} s & s' \\ s' & s \end{pmatrix} \\ b & \end{array}$$

Player 3 thus learns at each stage whether player 1 and 2 played the same action. Consider the following strategy of the team: at stage 1 each player randomizes between his two actions with equal probabilities. Let $\mathbf{a}_1^1$ be the random move of player 1 at stage 1. At each stage $n > 1$, play (a, a) if $\mathbf{a}_1^1 = a$ and play (b, b) if $\mathbf{a}_1^1 = a$. The signal of player 3 at stage 1 is uniformly distributed and conditional on this signal, $\mathbf{a}_1^1$ is also uniformly distributed. Since after stage 1 the signals will be constant, player 3 never learns anything about the value of $\mathbf{a}_1^1$. Players 1 and 2 are thus correlated from stage 2 on and I guarantees $\frac{1}{2}$.

Finally consider the case where team members observe each other's actions and the signal of player 3 is given by player 2's action, i.e. by the following matrix:

$$\begin{array}{ccc} & a & b \\ a & s & s' \\ b & s & s' \end{array}$$

As is the previous case, the move $\mathbf{a}_1^1$ of player 1 at stage 1 is unobserved by player 3 and may serve as a correlation device. Again let players 1 and 2 both randomize uniformly at the first stage and at stage 2, play (a, a) if $\mathbf{a}_1^1 = a$ and (b, b) if $\mathbf{a}_1^1 = a$. However, the move of player 2 at stage 2 reveals $\mathbf{a}_1^1$ and thus the correlation gained at stage 1 is lost after stage 2. The trade-off between generating signals for correlation and using this correlation appears here, the first stage generates a correlation device and the second uses it. Playing this two-stage strategy cyclically, the team guarantees $\frac{3}{8}$. We shall see in section 4.3 that the team can improve on this payoff.

3. Model and definitions

3.1. **The repeated game.** Let $I = \{1, \ldots, |I|\}$ be a finite set of players called *team* and II be another player. For each player $i \in I$, let A^i be player i's finite set of actions and let B be player II's finite set of actions. We denote $A = \prod_{i \in I} A^i$. At each stage $t = 1, 2, \ldots$, each player chooses an action in his own set of actions and if $(a, b) = ((a^i)_{i \in I}, b) \in A \times B$ is the action profile played, the payoff for each team player $i \in I$ is $g(a, b)$ where $g \colon A \times B \to \mathbb{R}$. The payoff for player II is $-g(a, b)$.

After each stage, if a is the action profile played by players $i \in I$, a signal s is drawn in a finite set S with probability $q(s|a)$, where q maps A to the set of probabilities on S. Player II observes (s, b), whereas team players observe (a, s, b). Thus, in our model all team members observe the same random signal that reveals the signal observed by player II.

We will use the following notations: for each finite set E, we let $\Delta(E)$ be the set of probabilities on E. We shall write an element $x \in \Delta(E)$ as the vector $x = (x(e))_{e \in E}$ with $x(e) \geq 0$ and $\sum_e x(e) \geq 0$. We denote by $\otimes$ the direct product of probabilities i.e. $(p \otimes q)(x, y) = p(x)q(y)$.

A history of length n for the team is an element h_n of $H_n = (A \times B \times S)^n$, and a history of length n player II is h_n^{II} element of $H_n^{II} = (B \times S)^n$, by convention $H_0 = H_0^{II} = \{\emptyset\}$. A behavioral strategy σ^i for a team player i is a mapping $\sigma^i \colon \cup_{n \geq 0} H_n \to \Delta(A^i)$ and a behavioral strategy τ for player II is a mapping $\tau \colon \cup_{n \geq 0} H_n^{II} \to \Delta(B)$. A profile of behavioral strategies $(\sigma, \tau) = ((\sigma^i)_{i \in I}, \tau)$ induces a probability distribution $\mathbf{P}_{\sigma,\tau}$ on the set of plays $(A \times B \times S)^\infty$ endowed with the product σ-algebra.

Given a discount factor $0 < \lambda < 1$, the discounted payoff for team I

induced by (σ,τ) is: $\gamma_\lambda(\sigma,\tau)=\mathbf{E}_{\sigma,\tau}[\sum_{n\geq 1}(1-\lambda)\lambda^{n-1}g(\mathbf{a}_n,\mathbf{b}_n)]$ where $(\mathbf{a}_n,\mathbf{b}_n)$ denotes the random action profile at stage n. The λ-discounted max min payoff of team I denoted v_λ is:

$$v_\lambda=\max_\sigma\min_\tau\gamma_\lambda(\sigma,\tau)$$

The average payoff for team I up to stage n is: $\gamma_n(\sigma,\tau)=\mathbf{E}_{\sigma,\tau}[\frac{1}{n}\sum_{m=1}^{n}g(\mathbf{a}_n,\mathbf{b}_n)]$. The n-stage max min payoff of team I denoted v_n is:

$$v_n=\max_\sigma\min_\tau\gamma_n(\sigma,\tau)$$

3.2. Best replies and Autonomous strategies.

We define here strategies for player II that play myopic best replies.

Definition 1. *Let σ be a strategy for the team, define inductively τ_σ as the strategy of player II that plays stage-best replies to σ:*

- *At stage 1, let $\tau_\sigma(\emptyset)\in\operatorname{argmin}_b g(\sigma(\emptyset),b)$;*
- *Assume that τ_σ is defined on histories of length less that $n+1$. For every history h_n^{II} of player II, let $x_{n+1}(h_n^{II})\in\Delta(A)$ be the distribution of the action profile of the team at stage $n+1$ given h_n^{II} and let $\tau_\sigma(h_n^{II})$ be in $\operatorname{argmin}_b g(x_{n+1}(h_n^{II}),b)$.*

We introduce now a class of strategies for the team against which the myopic best reply is a best reply in the repeated game. Call a strategy of a team player *autonomous* if it does not depend on player II's past moves that is for $i\in I$, $\sigma^i\colon\bigcup_n(A\times S)^n\to\Delta(A^i)$. Against a profile of autonomous strategies, the myopic best reply is a true best reply.

Lemma 2. *Let σ be a profile of autonomous strategies, for each stage n and strategy τ for player II, $\mathbf{E}_{\sigma,\tau_\sigma}g(\mathbf{a}_n,\mathbf{b}_n)\leq\mathbf{E}_{\sigma,\tau}g(\mathbf{a}_n,\mathbf{b}_n)$ and thus τ_σ is player II's best reply in any version of the repeated game.*

Proof. Consider the optimization problem of player II,

$$\min_{\tau} \mathbf{E}_{\sigma,\tau} \sum\nolimits_{n \geq 1} (1-\lambda)\lambda^{n-1} g(\mathbf{a}_n, \mathbf{b}_n)$$

Since player II's moves do not influence the play of the team, this amounts to solve for each n and history h_n^{II}, $\min_b \mathbf{E}_\sigma[g(\mathbf{a}_n, b) | h_n^{II}]$, and the solution is given by $\tau_s igma(h_n^{II})$. The same argument applies in the n-stage game. □

3.3. **Information theory tools.** The entropy of a finite random variable $\mathbf{x}$ with law P is by definition:

$$H(\mathbf{x}) = -\mathbf{E}[\log P(\mathbf{x})] = -\sum\nolimits_x P(x) \log P(x)$$

where log denotes the logarithm with base 2. Note that $H(\mathbf{x}) \geq 0$ and that $H(\mathbf{x})$ depends only on the law P of $\mathbf{x}$. The entropy of $\mathbf{x}$ is thus the entropy $H(P)$ of its distribution P, with $H(P) = -\sum_x P(x) \log P(x)$.

Let $(\mathbf{x}, \mathbf{y})$ be a couple of random variables with joint law P such that $\mathbf{x}$ is finite. The conditional entropy of $\mathbf{x}$ given $\{\mathbf{y} = y\}$ is the entropy of the conditional distribution $P(\mathbf{x}|y)$:

$$H(\mathbf{x} \mid y) = -\mathbf{E}[\log P(\mathbf{x} \mid y)]$$

The conditional entropy of $\mathbf{x}$ given $\mathbf{y}$ is the expected value of the previous:

$$H(\mathbf{x} \mid \mathbf{y}) = \int H(\mathbf{x} \mid y) dP(y)$$

If $\mathbf{y}$ is also finite, one has the following relation of additivity of entropies:

$$H(\mathbf{x}, \mathbf{y}) = H(\mathbf{y}) + H(\mathbf{x} \mid \mathbf{y})$$

4. The main result

The max min values v_λ, v_n are defined in terms of the data of the repeated game. Our main result is a characterization of their asymptotic values as the discount factor goes to 1 or the length of the game goes to infinity.

4.1. **Correlations systems.** Let σ be a strategy. Suppose that at stage n, the history for player II is $h_n^{II} = (b_1, s_1, \ldots, b_n, s_n)$. Let $h_n = (a_1, b_1, s_1, \ldots, a_n, b_n, s_n)$ be the history for the team. The mixed action played by the team at stage $n+1$ is $\sigma(h_n) = (\sigma^i(h_n))_{i \in I}$. Player II holds a belief on this mixed action, namely he believes that player II plays $\sigma(h_n)$ with probability $\mathbf{P}_\sigma(h_n | h_n^{II})$. The distribution of the action profile $\mathbf{a}_{n+1}$ given the information h_n^{II} of player II is $\sum_{h_n} \mathbf{P}_\sigma(h_n|h_n^{II})\sigma(h_n)$, element of the set $\Delta(A)$ of correlated distributions on A.

Definition 3. *Let $X = \otimes_{i \in I} \Delta(A^i)$ be the set of independent probability distributions on A. A* correlation system *is a probability distribution on X and we let $C = \Delta(X)$ be the set of all correlation systems.*

X is a closed subset of $\Delta(A)$ and thus C is compact with respect to the weak-$*$ topology.

Assume that at some stage n, after some history h_n^{II}, the distribution of $\sigma(h_n)$ conditional on h_n^{II} is c. The play of the game at this stage is as if: h_n were drawn according to the probability distribution c and announced to each player of the team but not to player II and given h_n, each team player chooses a mixed action. This generates a random action profile for the team and a random signal. We study the variation of uncertainty of player II regarding the total history, measuring uncertainty by entropy.

Definition 4. *Let c be a correlation system and $(\mathbf{x}, \mathbf{a}, \mathbf{s})$ be a random variable in $X \times A \times S$ such that the law of $\mathbf{x}$ is c, the law of $\mathbf{a}$ given $\{\mathbf{x} = x\}$ is x and the law of $\mathbf{s}$ given $\{\mathbf{a} = a\}$ is $q(\cdot|a)$. The entropy variation of c is:*

$$\Delta H(c) = H(\mathbf{a}, \mathbf{s} \mid \mathbf{x}) - H(\mathbf{s})$$

The entropy variation is the difference between the entropy gained by the team and the entropy lost. The entropy gain is the additional uncertainty contained in $(\mathbf{a}, \mathbf{s})$; the entropy loss is the entropy of $\mathbf{s}$ which is observed by player II. If $\mathbf{x}$ is finite, from the additivity formula we obtain:

$$H(\mathbf{x}, \mathbf{a}, \mathbf{s}) = H(\mathbf{x}) + H(\mathbf{a}, \mathbf{s} \mid \mathbf{x}) = H(\mathbf{s}) + H(\mathbf{x}, \mathbf{a} \mid \mathbf{s})$$

and therefore,

$$\Delta H(c) = H(\mathbf{x}, \mathbf{a} \mid \mathbf{s}) - H(\mathbf{x})$$

The entropy variation is then written as the difference between the entropy of the secret information of the team after stage n and before stage n.

We define now, given a correlation system c, the payoff obtained when player II plays a best reply to the expected distribution on A.

Definition 5. *Given a correlation system c, the distribution of the action profile for the team is $x_c \in \Delta(A)$ such that for each $a \in A$, $x_c(a) = \int_X (\Pi_i x_i(a_i)) dc(x)$. The optimal payoff yielded by c is $\pi(c) = \min_{b \in B} g(x_c, b)$, where g is extended to mixed actions in the usual way.*

We consider the set of feasible vectors $(\Delta H(c), \pi(c))$ in the (entropy variation, payoff) plane:

$$V = \{(\Delta H(c), \pi(c)) \mid c \in C\}$$

Lemma 6. *V is compact.*

Proof. Since the signal $\mathbf{s}$ depends on $\mathbf{a}$ only, the additivity formula gives $H(\mathbf{a}, \mathbf{s}|\mathbf{x}) = H(\mathbf{a}|\mathbf{x}) + H(\mathbf{s}|\mathbf{a})$ and the entropy variation is:

$$\Delta H(c) = H(\mathbf{a}|\mathbf{x}) + H(\mathbf{s}|\mathbf{a}) - H(\mathbf{s})$$

From the definitions of entropy and conditional entropy:

$$\Delta H(c) = \int_X H(x) dc(x) + \sum\nolimits_a x_c(a) H(q(\cdot|a)) - H(\sum\nolimits_a x_c(a) q(\cdot|a))$$

which is clearly a continuous function of c. Both ΔH and π are thus continuous on the compact set C so that the image set V is compact. □

We introduce the following notation:

$$w = \sup\{x_2 \in \mathbb{R} \mid (x_1, x_2) \in \operatorname{co} V, x_1 \geq 0\}$$

This is the highest payoff associated to a convex combination of correlations systems under the constraint that the average entropy variation is non-negative. For every correlation system c such that $\mathbf{x}$ is a.s. constant, $\Delta H(c) \geq 0$ thus V intersects the half-plane $\{x_1 \geq 0\}$ and since V is compact the supremum is indeed a maximum. The set V is not convex in general, and an example of non-convex V is provided by Goldberg [Gol03]. In this case, the supremum is not be achieved by a single point but by a convex combination of two points .

For computations, it is convenient to express the number w through

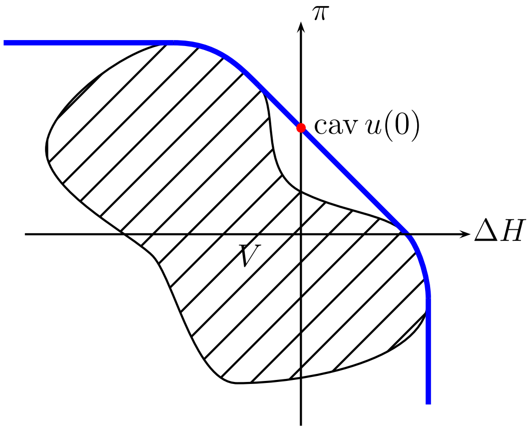

Figure 1. The set V, the graph of $\operatorname{cav} u$, and $\operatorname{cav} u(0)$

the boundary of $\operatorname{co} V$. Define for each real number h:

$$u(h) = \max\{\pi(c) \mid c \in C, \Delta H(c) \geq h\}$$

From the definition of V we have for each h:

$$u(h) = \max\{x_2 \mid (x_1, x_2) \in V, x_1 \geq h\}$$

Since V is compact, $u(h)$ is well defined. Let $\operatorname{cav} u$ be the least concave function pointwise greater than u. Then:

$$\sup\{x_2 \in \mathbb{R} \mid (x_1, x_2) \in \operatorname{co} V, x_1 \geq 0\} = \operatorname{cav} u(0)$$

Indeed, u is upper-semi-continuous, non-increasing and the hypograph of u is the comprehensive set $V^* = V - \mathbb{R}^2_+$ associated to V. This implies that $\operatorname{cav} u$ is also non-increasing, l.s.c. and its hypograph is $\operatorname{co} V^*$.

4.2. **The main result.**

Theorem 7. *The maxmin value of the λ-discounted game and of the n-stage game both converge to the same limit respectively as λ goes to 1 and n goes to infinity and this limit is:*

$$\lim_{\lambda} v_\lambda = \lim_{n} v_n = w$$

4.3. **Example.** We take back the last example of section 2 i.e. the following 3-player game where player 1 chooses rows, player 2 chooses columns and player 3 chooses matrices.

$$\begin{array}{ccc} & \begin{array}{cc} a & b \end{array} & \begin{array}{cc} a & b \end{array} \\ \begin{array}{c} a \\ b \end{array} & \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} & \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \\ & L & R \end{array}$$

The signals are given by the moves of player 2 i.e.:

$$\begin{array}{cc} & \begin{array}{cc} a & b \end{array} \\ \begin{array}{c} a \\ b \end{array} & \begin{pmatrix} s & s' \\ s & s' \end{pmatrix} \end{array}$$

Consider the following cyclic strategy: the team plays the mixed action profile $(\frac{1}{2}, \frac{1}{2}) \otimes (\frac{1}{2}, \frac{1}{2})$ at stage $2n+1$ and at stage $2n+2$, the team plays (a, a) if $\mathbf{a}^1_{2n+1} = a$ and (b, b) if $\mathbf{a}^1_{2n+1} = a$. This strategy consists in playing alternatively two correlation systems. Let c_{+1} be the Dirac measure on $(\frac{1}{2}, \frac{1}{2}) \otimes (\frac{1}{2}, \frac{1}{2})$ and c_{-1} which puts equal weights on $(1, 0) \otimes (1, 0)$ and on $(0, 1) \otimes (0, 1)$ i.e. $c_{-1} \in \Delta(X)$ and $c_{-1}(\{(1, 0) \otimes (1, 0)\}) = c_{-1}(\{(0, 1) \otimes (0, 1)\}) = \frac{1}{2}$. We have $\pi(c_{+1}) = \frac{1}{4}$, $\Delta H(c_{+1}) = +1$, $\pi(c_{-1}) = \frac{1}{2}$ and $\Delta H(c_{-1}) = -1$ since the move of player 2 at an even stage reveals the action of player 1 at the previous stage. The so-defined strategy, playing c_{+1} at odd stages and c_{-1} at even stages gives an average payoff of $\frac{3}{8}$ and an average entropy

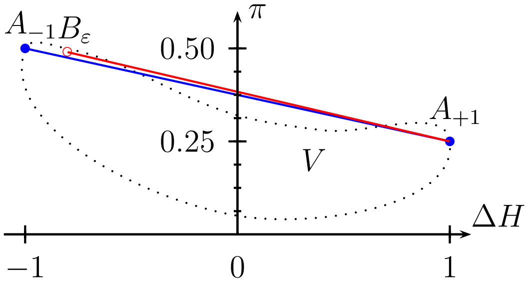

FIGURE 2. The existence of a point B_ε in V above the line between A_{-1} and A_{+1} implies the team can guarantee more than $\frac{3}{8}$.

variation of 0.

We now prove the existence of strategies for players 1 and 2 that guarantee more than $\frac{3}{8}$. By theorem 7, it is enough to show the existence of a convex combination of two correlation systems yielding an average payoff larger than $\frac{3}{8}$ and a non-negative average entropy variation.

Define the correlation system c_ϵ which puts equal weights on $(1-\varepsilon,\varepsilon)\otimes(1,0)$ and $(\varepsilon,1-\varepsilon)\otimes(0,1)$: $c_\epsilon(\{(1-\varepsilon,\varepsilon)\otimes(1,0)\}) = c_\epsilon(\{(\varepsilon,1-\varepsilon)\otimes(0,1)\}) = \frac{1}{2}$. Let A_{+1}, A_{-1}, and B_ε be the points with coordinates $(\Delta H(c_{+1}),\pi(c_{+1}))$, $(\Delta H(c_{-1}),\pi(c_{-1}))$, and $(\Delta H(c_\varepsilon),\pi(c_\varepsilon))$ respectively. We have $\pi(c_\varepsilon)=\frac{1-\varepsilon}{2}$ and $\Delta H(c_\varepsilon)=h(\varepsilon)-1$ where for $x\in\,]0,1[$, $h(x)=-x\log(x)-(1-x)\log(1-x)$, $h(0)=h(1)=0$. Using that $h'(0)=+\infty$, we deduce the existence of $\varepsilon>0$ such that B_ε lies above the line joining A_{-1} and A_{+1}.

For this ε, there exists $0\le\lambda\le1$ such that $\lambda\Delta H(c_\varepsilon)+(1-\lambda)\Delta H(c_{+1})=0$ and $\lambda\pi(c_\varepsilon)+(1-\lambda)\pi(c_{+1})>\frac{3}{8}$, which implies that the team can guarantee (strictly) more than $\frac{3}{8}$. This is illustrated in figure 2.

4.4. **Proof of the main theorem.**

4.4.1. *Player II defends w.* We prove here that for every strategy of the team, if player II plays stage-best replies, the average vector of (payoffs, entropy variation) after any number of stages belongs to V. This will imply that no strategy for the team can guarantee a better payoff than w. The proof follows similar lines as for instance the previous papers [NO99],[NO00], [GV02].

Lemma 8. *Player II defends w in every n-stage game, i.e. for each integer n and strategy profile for the team σ:*

$$\gamma_n(\sigma, \tau_\sigma) \leq w$$

Therefore, for each n, $v_n \leq w$.

Proof. Let σ be a strategy for the team and set $\tau = \tau_\sigma$. Let $\mathbf{a}_m, \mathbf{b}_m, \mathbf{s}_m$ be the sequences of random action profiles and signals associated to (σ, τ), $\mathbf{h}_m^{II} = (\mathbf{b}_1, \mathbf{s}_1, \ldots, \mathbf{b}_{m-1}, \mathbf{s}_{m-1})$ be the history of player II before stage m and $\mathbf{h}_m = (\mathbf{a}_1, \mathbf{b}_1, \mathbf{s}_1, \ldots, \mathbf{a}_{m-1}, \mathbf{b}_{m-1}, \mathbf{s}_{m-1})$ be the total history. Let $\mathbf{x}_m = \sigma(\mathbf{h}_m)$ and $c_m(\mathbf{h}_m^{II})$ be the distribution of $\mathbf{x}_m$ conditional on $\mathbf{h}_m^{II}$ i.e. $c_m(\mathbf{h}_m^{II})$ is the correlation system at stage m after history $\mathbf{h}_m^{II}$. Under (σ, τ), the payoff at stage m after $\mathbf{h}_m^{II}$ is $\min_b g(\mathbf{E}_{\sigma,\tau}[\mathbf{x}_m | \mathbf{h}_m^{II}], b) = \pi(\mathbf{c}_m)$ from the definition of τ and thus $\gamma_n(\sigma, \tau) = \mathbf{E}_{\sigma,\tau}[\frac{1}{n} \sum_{m=1}^{n} \pi(\mathbf{c}_m)]$.

We set $H_m = H(\mathbf{a}_1, \ldots, \mathbf{a}_m \mid \mathbf{h}_{m+1}^{II})$ and using the additivity of entropy, we have:

$$\begin{aligned} H(\mathbf{a}_1, \dots, \mathbf{a}_m, \mathbf{b}_m, \mathbf{s}_m | \mathbf{h}^{II}_m) &= H(\mathbf{b}_m, \mathbf{s}_m | \mathbf{h}^{II}_m) + H_m \\ &= H_{m-1} + H(\mathbf{a}_m, \mathbf{b}_m, \mathbf{s}_m | \mathbf{h}_m) \end{aligned}$$

Thus,

$$\begin{aligned} H_m - H_{m-1} &= H(\mathbf{a}_m, \mathbf{b}_m, \mathbf{s}_m | \mathbf{h}_m) - H(\mathbf{b}_m, \mathbf{s}_m | \mathbf{h}^{II}_m) \\ &= H(\mathbf{a}_m, \mathbf{s}_m | \mathbf{h}_m) - H(\mathbf{s}_m | \mathbf{h}^{II}_m) + H(\mathbf{b}_m | \mathbf{h}_m) - H(\mathbf{b}_m | \mathbf{h}^{II}_m) \\ &= H(\mathbf{a}_m, \mathbf{s}_m | \mathbf{h}_m) - H(\mathbf{s}_m | \mathbf{h}^{II}_m) \\ &= H(\mathbf{a}_m, \mathbf{s}_m | \mathbf{x}_m, \mathbf{h}^{II}_m) - H(\mathbf{s}_m | \mathbf{h}^{II}_m) \\ &= \mathbf{E}_{\sigma,\tau} \Delta H(c_m(\mathbf{h}^{II}_m)) \end{aligned}$$

where the second equality holds since $\mathbf{a}_m$ and $\mathbf{b}_m$ are independent conditional on $\mathbf{h}^{II}_m$, the third uses that $\mathbf{b}_m$ is $\mathbf{h}^{II}_m$-measurable and the fourth that $(\mathbf{a}_m, \mathbf{s}_m)$ depends on $\mathbf{h}_m$ only through $\mathbf{x}_m$. We deduce:

$$\sum\nolimits_{m=1}^{n} \mathbf{E}_{\sigma,\tau} \Delta H(c_m(\mathbf{h}^{II}_m)) = H(\mathbf{a}_1, \dots, \mathbf{a}_n \mid \mathbf{b}_1, \mathbf{s}_1 \dots, \mathbf{b}_n, \mathbf{s}_n) \geq 0.$$

Therefore the vector $(\frac{1}{n} \sum_{m=1}^{n} \mathbf{E}_{\sigma,\tau} \Delta H(c_m(\mathbf{h}^{II}_m)), \gamma_n(\sigma, \tau))$ is in $\operatorname{co} V \cap \{x_1 \geq 0\}$. $\square$

Corollary 9. *Player II defends w in every λ-discounted game, i.e. for each $\lambda \in (0, 1)$ and strategy profile for the team σ:*

$$\gamma_\lambda(\sigma, \tau_\sigma) \leq w$$

Therefore, for each λ, $v_\lambda \leq w$.

Proof. The discounted payoff is a convex combination of the average payoffs (see Lehrer and Sorin [LS92]):

$$\gamma_\lambda(\sigma,\tau) = \sum\nolimits_{n\geq 1}(1-\lambda)^2 n\lambda^{n-1}\gamma_n(\sigma,\tau)$$

From lemma 8 we get $\gamma_\lambda(\sigma,\tau_\sigma) \leq w$ and thus $v_\lambda \leq w$.

$\square$

4.4.2. *v_n converges to w.* We call a distribution $\mathbf{P}$ on $(A \times S)^\infty$ an X-distribution if at each stage n, after $\mathbf{P}$-almost every history $h^I_n = (a_1, s_1, \ldots, a_n, s_n) \in H^I_n$, the distribution of $\mathbf{a}_{n+1}$ conditional on h^I_n, $\mathbf{P}(\mathbf{a}_{n+1}|h^I_n)$ belongs to X. Every autonomous strategy profile induces an X-distribution and conversely an X-distribution defines an autonomous strategy profile.

Given an autonomous strategy profile σ or equivalently an X-distribution, consider the random correlation system at stage n: given $\mathbf{h}^{II}_n$, $\mathbf{c}_n$ is the distribution of $\sigma(\mathbf{h}^I_n)$ given $\mathbf{h}^{II}_n$. The random variable $\mathbf{c}_n$ is $\mathbf{h}^{II}_n$-measurable with values in $C = \Delta(X)$. We consider the empirical distribution of correlation systems up to stage n, i.e. the time frequencies of correlation systems appearing along the history $\mathbf{h}^{II}_n$. We define it as the random variable:

$$\mathbf{d}_n = \frac{1}{n}\sum\nolimits_{m\leq n}\epsilon_{\mathbf{c}_m}$$

where ϵ_c denotes the Dirac measure on c. The random variable $\mathbf{d}_n$ has values in $D = \Delta(C)$. If we let $\delta = \mathbf{E}_\sigma \mathbf{d}_n$ be the barycenter of $\mathbf{d}_n$ i.e. the element of D such that for any real-valued continuous function f on C, $\mathbf{E}_\sigma[\int f(b)d\mathbf{d}_n(b)] = \int f(b)d\delta(b)$, the average payoff under (σ,τ_σ) writes:

$$\gamma_n(\sigma,\tau_\sigma) = \mathbf{E}_{\sigma,\tau_\sigma}[\frac{1}{n}\sum\nolimits_{m=1}^{n}\pi(c_m)] = \mathbf{E}_{\sigma,\tau_\sigma}[\mathbf{E}_{\mathbf{d}_n}\pi] = \mathbf{E}_\delta \pi$$

We use the following result from Gossner and Tomala [GT04]:

Theorem 10 ([GT04], thm. 9)**.** *For every* $\delta \in \Delta(C)$ *such that* $\mathbf{E}_\delta \Delta H \geq 0$*, there exists an* X*-distribution* $\mathbf{P}$ *on* $(A \times S)^\infty$ *such that* $\mathbf{E}_\mathbf{P}\mathbf{d}_n$ *weak-*$*$ *converges to* δ*.*

Since any X-distribution $\mathbf{P}$ corresponds to an autonomous strategy, we get:

Lemma 11. $\liminf_n v_n \geq \sup\{\mathbf{E}_\delta \pi \mid \delta \in \Delta(C), \mathbf{E}_\delta \Delta H \geq 0\}$.

Proof. For each δ such that $\mathbf{E}_\delta \Delta H \geq 0$, the previous theorem yields the existence of an autonomous strategy σ such that $\lim_n \gamma_n(\sigma, \tau_\sigma) = \mathbf{E}_\delta \pi$. From lemma 2 this gives $\liminf_n v_n \geq \mathbf{E}_\delta \pi$. □

We may now conlude the proof. The set of vectors $(\mathbf{E}_\delta \Delta H, \mathbf{E}_\delta \pi)$ as δ varies in $\Delta(C)$ is $\operatorname{co} V$ and thus $\sup\{\mathbf{E}_\delta \pi \mid \delta \in \Delta(C), \mathbf{E}_\delta \Delta H \geq 0\} = w$. From lemmata 8 and 11 we get $\lim_n v_n = w$.

4.4.3. *v_λ converges to w.* Since $v_\lambda \leq w$ it is enough to prove the following lemma.

Lemma 12. $\forall \varepsilon > 0$*,* $\exists \sigma, \exists \lambda_0$*, such that* $\forall \lambda \geq \lambda_0$*,* $\gamma_\lambda(\sigma, \tau_\sigma) \geq w - \varepsilon$*.*

Proof. For $\varepsilon > 0$, choose σ autonomous such that $\gamma_n(\sigma, \tau_\sigma) \geq w - \frac{\varepsilon}{2}$. Define a cyclic strategy σ^* as follows: play σ until stage n and restart this strategy every n stages. Set y_m as the expected payoff under $(\sigma^*, \tau_{\sigma^*})$ at stage m. Since σ^* is cyclic, τ_{σ^*} is also cyclic and:

$$\gamma_\lambda(\sigma^*, \tau_{\sigma^*}) = \sum\nolimits_{m=1}^{n} (1-\lambda)\lambda^{m-1} y_m + \lambda^n \gamma_\lambda(\sigma^*, \tau_{\sigma^*}))$$

So,

$$\gamma_\lambda(\sigma^*, \tau_{\sigma^*}) = \sum\nolimits_{m=1}^{n} (1-\lambda)\frac{\lambda^{m-1}}{1-\lambda^n} y_m$$

Then, $\lim_{\lambda\to 1}\gamma_\lambda(\sigma^*, \tau_{\sigma^*}) = \frac{1}{n}\sum_{m=1}^n y_m \geq w - \frac{\varepsilon}{2}$ which ends the proof. □

5. Perfect and trivial observation

5.1. **Perfect observation.** We say that the observation is perfect when the signal s reveals the action profile a i.e. $a \neq a' \Rightarrow \operatorname{supp} q(\cdot|a) \cap \operatorname{supp} q(\cdot|a') = \emptyset$. It is well known that in this case, the max min of the repeated game is the independent max min of player II, i.e. $w = \max_{x\in X}\min_b g(x, b)$. We verify now that our main theorem gives the same value.

Since the observation is perfect, $H(\mathbf{a}|\mathbf{s}) = 0$ and $\Delta H(c) = H(\mathbf{s}|\mathbf{x}) - H(\mathbf{s}) \leq 0$ for each correlation system c and $\Delta H(c) = 0$ if and only if $\mathbf{s}$ (and thus $\mathbf{a}$) in independent of $\mathbf{x}$. This implies that $\Delta H(c) = 0$ if and only if c is a Dirac measure on some $x \in X$. We let C_d be the set of correlation systems whose support is a subset of $\{\epsilon_x, x \in X\}$ where ϵ_x denotes the Dirac measure on x. From the above discussion it follows that for every distribution δ, $\mathbf{E}_\delta \Delta H \geq 0$ if and only if the support of δ is a subset of C_d. This has a clear interpretation: if the observation is perfect, at each stage, the next moves of the team are independent conditional on the signals of player II. Thus, $w = \sup\{\pi(\epsilon_x), x \in X\}$ that is $w = \max_{x\in X}\min_b g(x, b)$.

5.2. **Trivial observation.** We say that the observation is trivial when the signal s does not depend on the action profile a. In this case, the team can randomize actions for a number of stages, and use these first actions as a correlating device in all subsequent stages. This way, the team can guarantee $w = \max_{x\in\Delta(A)}\min_b g(x, b)$ which is the correlated minmax of player II. Applying our main theorem, we remark that if observation is trivial $\Delta H(c) \geq 0$ for each c and thus every distribution δ verifies $\mathbf{E}_\delta \Delta H \geq 0$, therefore $w = \sup\{\pi(c), c \in C\} = \max_{x\in\Delta(A)} \min_b g(x,b)$.

6. Numerical examples

In section 4, the max min w is characterized as $\operatorname{cav} u(0)$ with $u(h) = \max\{\pi(c) \mid c \in C, \Delta H(c) \geq h\}$ so the numerical computation of w consists in computing the function $u(h)$. Note that for each $h \in \mathbb{R}$, either $\operatorname{cav} u(h) = u(h)$ or $\operatorname{cav} u$ is linear on some interval containing h. Thus, either $\operatorname{cav} u(0) = \pi(c)$ for some c s.t. $\Delta H(c) \geq 0$ or there exists c_1, c_2 and $\lambda \in (0,1)$ s.t. $\operatorname{cav} u(0) = \lambda\pi(c_1) + (1-\lambda)\pi(c_2)$ and $\lambda\Delta H(c_1) + (1-\lambda)\Delta H(c_2) \geq 0$. In the first case, the optimal strategy can be regarded as "stationary" (in the space of correlation systems) since only one correlation system is used at almost all stages. In the second case, the strategy has two phases. Assume w.l.o.g. $\Delta H(c_1) > 0$, in a first phase the optimal strategy plays c_1 to accumulate entropy and in a second phase, the optimal strategy plays c_2 that spends entropy and yields a good payoff. The relative lengths of these phases are $(\lambda, 1-\lambda)$. We give now examples illustrating both cases. The computation of $u(h)$ is studied in Gossner et al. [GLT03] and Goldberg [Gol03].

6.1. **An example of optimal "stationary" correlation.** The paper by Gossner et al. [GLT03] is devoted to the computation of the max min for the game given last example of section 2:

$$\begin{array}{c} \\ a \\ b \end{array} \begin{array}{c} \begin{array}{cc} a & b \end{array} \\ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \\ L \end{array} \begin{array}{c} \begin{array}{cc} a & b \end{array} \\ \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \\ R \end{array}$$

where the signal is given by the matrix:

$$\begin{array}{c|cc} & a & b \\ \hline a & s & s' \\ b & s & s' \end{array}$$

In this case, Gossner et al. [GLT03] prove that there exists a correlation system c such that $\Delta H(c) = 0$ and $w = \pi(c)$. This system can be written:

$$c = \frac{1}{2}\epsilon_{(x,1-x)\otimes(x,1-x)} + \frac{1}{2}\epsilon_{(1-x,x)\otimes(1-x,x)}$$

where $0 < x < \frac{1}{2}$ is such that $\Delta H(c) = 2H(x, 1-x) - 1 = 0$. Numerically, this gives $w \cong 0.401$.

6.2. **An example of optimal convexification.** Another example is studied by Goldberg [Gol03]. Consider the game where payoffs for players 1 and 2 are given by:

$$\begin{array}{c} \\ a \\ b \end{array}\begin{array}{c} a\quad b \\ \begin{pmatrix} 1 & 0 \\ 3 & 1 \end{pmatrix} \\ L \end{array}\quad\begin{array}{c} a\quad b \\ \begin{pmatrix} 1 & 3 \\ 0 & 1 \end{pmatrix} \\ R \end{array}$$

Players 1 and 2 observe the action profiled played, and player 3 observes a signal on player 1 and 2's action given by the matrix:

$$\begin{array}{c|cc} & a & b \\ \hline a & s & s'' \\ b & s' & s \end{array}$$

The independent max min is $\frac{5}{4}$ and is obtained by the mixed action $(\frac{1}{2}, \frac{1}{2}) \otimes (\frac{1}{2}, \frac{1}{2})$. The correlated max min is $\frac{3}{2}$ and is obtained by the distribution of actions $\frac{1}{2}(1,0) \otimes (1,0) + \frac{1}{2}(0,1) \otimes (0,1)$.

Using theorem 7, Goldberg [Gol03] proves that the limiting maxmin w is $\frac{4}{3}$ and is achieved by the following strategy: play $(\frac{1}{2}, \frac{1}{2}) \otimes (\frac{1}{2}, \frac{1}{2})$ until the signal s appears. If s appears at stage n, play (a, b) at stage $n+1$ if (a, a) was played at stage n, and play (b, a) if (b, b) was played. Resume playing $(\frac{1}{2}, \frac{1}{2}) \otimes (\frac{1}{2}, \frac{1}{2})$ at stage $n + 2$.

Under this strategy, two correlations systems are used: c_1 which is the Dirac measure on $(\frac{1}{2}, \frac{1}{2}) \otimes (\frac{1}{2}, \frac{1}{2})$ and $c_2 = \frac{1}{2}\epsilon_{(0,1)\otimes(1,0)} + \frac{1}{2}\epsilon_{(1,0)\otimes(0,1)}$. The strategy induces an irreducible Markov chain on the finite state space $\{c_1, c_2\}$ and in the long run it spends two thirds of the time in c_1. Since $\pi(c_1) = \frac{5}{4}$ and $\pi(c_2) = \frac{3}{2}$ this strategy yields an average payoff of $\frac{4}{3}$. We check now that the average entropy variation is non-negative.

Under c_1, with probability $\frac{1}{2}$, players 1 and 2 play different actions and thus the action profile is revealed by the signal: no entropy is gained. With probability $\frac{1}{2}$ players 1 and 2 play the same action inducing the signal s. The conditional distribution puts then equal probabilities on (a, a) and (b, b): 1 bit of entropy is gained. Thus $\Delta H(c_1) = \frac{1}{2} \times 0 + \frac{1}{2} \times 1 = \frac{1}{2}$. Under c_2, players 1 and 2 play different actions with probability 1 and their action profile is revealed by the signal. All entropy is lost and $\Delta H(c_1) = -1$. One has $\frac{2}{3}\Delta H(c_1) + \frac{1}{3}\Delta H(c_2) = 0$.

7. Relation with the Folk Theorem

In repeated games with imperfect monitoring, information asymmetries raise a number of difficulties that cause the set of equilibrium payoffs to be hard to characterize in general. For this reason, the central results consider public equilibria (Abreu, Pierce and Stachetti [APS90], Lehrer [Leh90], and Fudenberg, Levine and Maskin [FLM94]), equilibria in which a communication mechanism serves to resolve information asymmetries (see Compte [Com02], Renault and Tomala [RT00]), or 2-player games ([Leh91], [Leh92]). In our approach, we tackle information asymmetries by measuring them with the entropy function.

The previous examples show 3-player games in which our main theorem allows to characterize the individually rational payoff of one player in the repeated game. We now present a signalling structure for which our theorem allows for a characterization of all individually rational payoffs.

Consider a game in which player the set of players is $I = \{1, \ldots, n\}$, $n \geq 4$, and i's finite action set is A^i. Players i ,$1 < i < n$ have perfect observation: they observe $s^i = (a^1, \ldots, a^n)$. Player 1 observes every player but player n: his signal is $s^1 = (a^1, \ldots, a^{n-1})$. Player n observes every player but player 1: his signal is $s^n = (a^2, \ldots, a^n)$. This structure of signals is represented in Renault and Tomala [RT98] by a graph whose nodes are the players and where there is an edge between i and j whenever i and j monitor each other. The graph described here is 2-connected: there are at least two distinct paths from i to j for each pair (i, j).

Since each player $1 < i < n$ has perfect observation, his individually rational levels v^i in the repeated game equals his minmax in mixed strategies of the one-shot game.

Consider a modified version of the repeated game in which player 1 also observes the moves of player n. This modified signalling structure between the team of players $\{1, \ldots, n-1\}$ and player n fulfill the conditions of our theorem. Let then v^n be the uniform $\min\max$ of player n in the modified repeated game. Since players $\{1, \ldots, n-1\}$ can push player n's payoff down to v^n with autonomous strategies, they can also push player n down v^n in the original game. On the other hand, since player n can defend v^n in the modified repeated game, he can also defend v^n in the original one. Hence the individually rational level of player n in the original game is v^n. The same reasoning shows that our theorem applies to a characterization of the min max for player 1.

Let $\operatorname{co} g(A)$ be the set of feasible payoffs, and $IR = \{x \in \mathbb{R}^n, x^i \geq v^i\}$ be the set of individually rational payoffs. Renault and Tomala [RT98] prove that when the signals are represented by a 2-connected graph, the set of uniform equilibrium payoffs is $\operatorname{co} g(A) \cap IR$.

Lehrer [Leh90] characterizes Nash equilibrium payoffs for all repeated games having a semi-standard signalling structure. Our example constitutes –as far as we know– the only other n-player signalling structure for which a characterization of Nash equilibrium payoffs is known for all payoff functions.

8. Extensions

8.1. **More general signals.** Theorem 7 remains true if team players do not fully observe the move b of player II, i.e. if whenever player II plays b, a signal r is drawn according to some distribution $p(\cdot|b)$ that depends on b only and is publicly announced. Indeed, w can be guaranteed with autonomous strategies by the team. On another hand, the proof of lemma 8 extends to show that the team cannot guarantee more than w.

8.2. **The uniform approach.** Our result characterize the limit of the n-stage and λ-discounted values of long games. In fact, a stronger result obtains if one observes that the same strategies perform uniformly good in all long enough games. The corresponding notion is of that uniform $\max\min$ payoff of player II (see Aumann Maschler [AM95]):

(1) The team I guarantees $v \in \mathbb{R}$ if:

$$\forall \varepsilon > 0, \exists \sigma = (\sigma^i)_{i\in I}, \exists N \text{ s.t. } \forall \tau, \forall n \geq N, \gamma_n(\sigma,\tau) \geq v - \varepsilon.$$

(2) Player II defends $v \in \mathbb{R}$ if:

$$\forall \varepsilon > 0, \forall \sigma, \exists \tau, \exists N \text{ s.t. } \forall n \geq N, \gamma_n(\sigma,\tau) \leq v + \varepsilon.$$

(3) The uniform $\max\min$, if it exists, is $v_\infty \in \mathbb{R}$ such that I guarantees v_∞ and II defends v_∞.

We get here:

Theorem 13. *The uniform* $\max\min$ *exists and* $v_\infty = w$.

Proof. From lemma 8, player II defends w. For each δ such that $\mathbf{E}_\delta \Delta H \geq 0$, theorem 10 yields the existence of an autonomous strategy σ such that $\lim_n \gamma_n(\sigma, \tau_\sigma) = \mathbf{E}_\delta \pi$ and thus the team guarantees w.

□

References

[AM95] R. J. Aumann and M. B. Maschler. *Repeated games with incomplete information.* with the collaboration of R. E. Stearns. MIT Press, Cambridge, 1995.

[APS90] D. Abreu, D. Pearce, and E. Stacchetti. Toward a theory of discounted repeated games with imperfect monitoring. *Econometrica*, 58:1041–1063, 1990.

[Aum74] R.J. Aumann. Subjectivity and correlation in randomized strategies. *Journal of Mathematical Economics*, 1:67–95, 1974.

[Com02] O. Compte. Communication in repeated games with imperfect private monitoring. *Econometrica*, 66:597–626, 2002.

[FLM94] D. Fudenberg, D. K. Levine, and E. Maskin. The folk theorem with imperfect public information. *Econometrica*, 62:997–1039, 1994.

[FM86] D. Fudenberg and E. Maskin. The folk theorem in repeated games with discounting or with incomplette information. *Econometrica*, 54:533–554, 1986.

[For86] F. Forges. An approach to communication equilibria. *Econometrica*, 54:1375–1385, 1986.

[GLT03] O. Gossner, R. Laraki, and T. Tomala. On the optimal use of coordination. mimeo, 2003.

[Gol03] Y. Goldberg. On the minmax of repeated games with imperfect monitoring: A computational example. Discussion Paper Series 345, Center the Study of Rationality, Hebrew University, Jerusalem, 2003.

[GT04] O. Gossner and T. Tomala. Empirical distributions of beliefs under imperfect observation. Cahiers du CEREMADE 410, Université Paris Dauphine, Paris, 2004.

[GV02] O. Gossner and N. Vieille. How to play with a biased coin? *Games and Economic Behavior*, 41:206–226, 2002.

[Leh90] E. Lehrer. Nash equilibria of n player repeated games with semi-standard information. *International Journal of Game Theory*, 19:191–217, 1990.

[Leh91] E. Lehrer. Internal correlation in repeated games. *International Journal of Game Theory*, 19:431–456, 1991.

[Leh92] E. Lehrer. Correlated equilibria in two-player repeated games with nonobservable actions. *Mathematics of Operations Research*, 17:175–199, 1992.

[LS92] E. Lehrer and S. Sorin. A uniform tauberian theorem in dynamic programming. *Mathematics of Operations Research*, 17:303–307, 1992.

[NO99] A. Neyman and D. Okada. Strategic entropy and complexity in repeated games. *Games and Economic Behavior*, 29:191–223, 1999.

[NO00] A. Neyman and D. Okada. Repeated games with bounded entropy. *Games and Economic Behavior*, 30:228–247, 2000.

[RT98] J. Renault and T. Tomala. Repeated proximity games. *International Journal of Game Theory*, 27:539–559, 1998.

[RT00] J. Renault and T. Tomala. Communication equilibrium payoffs of repeated games with imperfect monitoring. Cahiers du CEREMADE 0034, Université Paris Dauphine, Paris, 2000.

[vSK97] B. von Stengel and D. Koller. Team max min equilibria. *Games and Economics Behavior*, 21:309–321, 1997.

CERAS, URA CNRS 2036
E-mail address: `Olivier.Gossner@mail.enpc.fr`

CEREMADE, UMR CNRS 7534 Université Paris 9 – Dauphine
E-mail address: `tomala@ceremade.dauphine.fr`