

Secret Correlation in Repeated Games with Imperfect Monitoring

Olivier Gossner

Paris-Jourdan Sciences Economiques, UMR CNRS-EHESS-ENPC-ENS 8545, 48 Boulevard Jourdan, 75014 Paris, France and
MEDS, Kellogg School of Management, Northwestern University, 2001 Sheridan Road, Evanston, Illinois 60208-2001,
gossner@ens.fr

Tristan Tomala

CEREMADE, UMR CNRS 7534, Université Paris-Dauphine, Place du Maréchal de Lattre de Tassigny,
75775 Paris, Cedex 16, France, tomala@ceremade.dauphine.fr

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 tradeoff for the team between optimization of stage payoffs and generation of signals for future correlation.

Key words: stochastic process; min–max values; signals; entropy; repeated games; secret correlation

MSC2000 subject classification: Primary: 60G35, 94A17; secondary: 91A20

OR/MS subject classification: Primary: probability, entropy; secondary: games

History: Received June 15, 2005; revised February 24, 2006.

1. Introduction. In 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 devises coordinated plans for the team members, and generals of allied armies need to keep their coordinated plans secret from enemies. On the Internet, coordinated attacks of systems (e.g., by viruses) are known to be much more dangerous than uncoordinated attacks. 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 [2]), or be the result of communication between the players in the group (Forges [5]). In the model of repeated games with imperfect monitoring, each player observes a signal that depends stochastically on chosen actions (deterministic signals is a particular case); the signals may be correlated. These games feature both correlated signals and communication possibilities since actions may be used as messages.

This article explores the possibilities of secret correlation between team members in a repeated game with imperfect monitoring. In our model, two teams are matched against each other. Each member i of team I has an action set A^i . Team II is viewed as single player with action set 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 us to study the optimal tradeoffs 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, or the discount factor goes to one. We emphasize three reasons why we believe characterizing the max min value is important:

First, such characterizations are important for the general study of repeated games with imperfect monitoring because they provide the individually rational levels. Some generalizations of the Folk Theorem from the perfect monitoring case to imperfect monitoring, such as Renault and Tomala [20] and Hörner and Olszewski [12] show that the set of equilibrium payoffs of repeated games is the set of feasible and individually rational payoffs, but do not characterize the individually rational levels. Such a characterization completes these studies, thus providing full descriptions of the sets of equilibrium payoffs of the repeated games.

Second, von Stengel and Koller [22] proved that, in finite games where a team of players is matched against 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 repeated game with imperfect monitoring is the max min we characterize.

Finally, 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, only partial correlation may be achievable; the max min of the repeated game may lie between these two values. The study of the endogenous emergence of secret correlation of a group of players is interesting in itself. This article studies secret correlation as arising from monitoring structures. Gossner [8, 9] and Bavly and Neyman [3] studied its emergence through limitations of computational capacities of the players.

The problem faced by the team consists in finding the optimal tradeoff 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 the signals' 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 §2, present the model and definitions in §3, and the main result in §4. We discuss examples in §5, and computational aspects in §6. The proof of the main results is given in §7. For simplicity, the model and the main result are first stated for a simple class of signalling structures; we extend our main result to more general signalling structures in §8. Finally, we show consequences for the Folk Theorem in §9.

2. Examples. We consider a three-player game where the teams are $I = \{1, 2\}$ and $II = \{3\}$. Player 1 chooses rows, Player 2 chooses columns, and Player 3 chooses matrices. The payoffs to the team are given by

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

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 $\frac{1}{4}$.

Now assume 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. The team can then use the first move of Player 1 as a correlation device, and thus can guarantee the max min of the one-shot game in long repetitions, where the max runs over the set of all probability distributions on $A^1 \times A^2$. That is, from the second stage on, I guarantees $\frac{1}{2}$.

Now consider the case where team members observe each other's actions and the signal of Player 3 is given by the following matrix:

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

Player 3 thus learns at each stage whether Players 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 $\mathbf{a}_1^1 = b$. 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 . Actions of 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:

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

As in the latter two cases, 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 Stage 1 and at Stage 2, and play (a, a) if $\mathbf{a}_1^1 = b$ and (b, b) if $\mathbf{a}_1^1 = a$. Unlike in the previous examples, 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 tradeoff between generating signals for correlation and using this correlation appears here, Stage 1 generates a correlation device, and the Stage 2 uses it. Playing this two-stage strategy cyclically, the team guarantees $\frac{3}{8}$ and we will see that this is not optimal.

This game with the latter signaling structure serves in the sequel of the paper for further illustrations. We shall therefore refer to it as *the main example*.

3. Model and definitions.

3.1. The repeated game. Let $I = \{1, \dots, |I|\}$ 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 $t = 1, 2, \dots$, each player chooses an action in his own set of actions; 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 $g: A \times B \rightarrow \mathbb{R}$. 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 . Note that the model is designed to preserve transparency; §8 presents extensions of the model and results to a larger class of signalling structures.

For each finite set E , we let $\Delta(E)$ be the set of probabilities on E . We write an element $x \in \Delta(E)$ as a vector $x = (x(e))_{e \in E}$ with $x(e) \geq 0$ and $\sum_e x(e) = 1$. 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 for player II is an element h_n^{II} of $H_n^{II} = (B \times S)^n$; by convention, H_0 and H_0^{II} are arbitrary singletons. A behavioral strategy σ^i for a team player i is a mapping $\sigma^i: \bigcup_{n \geq 0} H_n \rightarrow \Delta(A^i)$; a behavioral strategy τ for player II is a mapping $\tau: \bigcup_{n \geq 0} H_n^{II} \rightarrow \Delta(B)$. 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 $\gamma_n(\sigma, \tau) = \mathbf{E}_{\sigma, \tau}[(1/n) \sum_{m=1}^n g(\mathbf{a}_m, \mathbf{b}_m)]$. The n -stage max min payoff of team I denoted v_n is

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

The uniform max min payoff of player II denoted v_∞ is defined as follows:

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

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

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

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

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

The max min v_1 of the one-shot game is simply $\max_{x \in \otimes_{j \neq I} \Delta(A^j)} \min_b g(x, b)$, where g is extended to mixed action in the usual way. We call this the *independent* max min. This is the best that the team can guarantee in the one-shot game with independent mixed strategies. This quantity can also be guaranteed in every version of the repeated game by playing independent and identically distributed (i.i.d.) a mixed strategy profile x that achieves the maximum in v_1 . Therefore,

$$\forall n, \forall \lambda, \quad v_n \geq v_1, \quad v_\lambda \geq v_1 \quad \text{and} \quad v_\infty \geq v_1.$$

On the other hand, in any version of the repeated game, the team cannot guarantee more than the value of the two-person zero-sum game defined by (I, II, A, B, g) . Let us denote $\text{val } g = \max_{x \in \Delta(A)} \min_b g(x, b)$ this value, and call it the *correlated* max min. One has

$$\forall n, \forall \lambda, \quad v_n \leq \text{val } g, \quad v_\lambda \leq \text{val } g \quad \text{and} \quad v_\infty \leq \text{val } g.$$

3.2. Information theory tools. The entropy of a finite random variable \mathbf{x} with law P is, by definition,

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

where \log denotes the logarithm with Base 2, and $0 \log 0 = 0$. 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 $P(\mathbf{x} | y)$ when this conditional distribution is well defined:

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

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

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

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

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

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 and of v_∞ .

4.1. Correlations systems. Let σ be a strategy. Suppose that at stage n , the history for player II is $h_n^{\text{II}} = (b_1, s_1, \dots, b_n, s_n)$. Let $h_n = (a_1, b_1, s_1, \dots, a_n, b_n, s_n)$ 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—he believes that player I plays $\sigma(h_n)$ with probability $\mathbf{P}_\sigma(h_n | h_n^{\text{II}})$. The distribution of the action profile \mathbf{a}_{n+1} given the information h_n^{II} of player II is $\sum_{h_n} \mathbf{P}_\sigma(h_n | h_n^{\text{II}}) \sigma(h_n)$, an element of $\Delta(A)$ the set of correlated distributions on A .

DEFINITION 1. 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. 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 2. 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

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

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

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

and therefore,

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

The entropy variation is thus the new entropy of the information possessed by I and not by II minus the initial entropy.

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

DEFINITION 3. Given a correlation system c , the distribution of the action profile for the team is $x_c \in \Delta(A)$ such that for each $a \in A$, $x_c(a) = \int_X \prod_i x_i(a_i) dc(x)$. The optimal payoff yielded by c is $\pi(c) = \min_{b \in B} g(x_c, b)$.

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

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

LEMMA 4. V is compact.

PROOF. Since \mathbf{s} is independent of \mathbf{x} conditionally on \mathbf{a} , 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

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

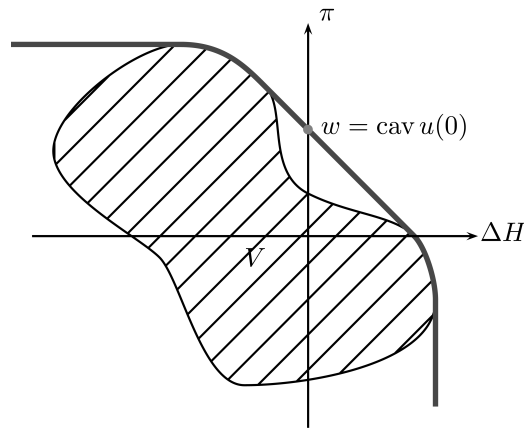


FIGURE 1. The set V , the graph of $\text{cav } u$, and $w = \text{cav } u(0)$.

From the definitions of entropy and conditional entropy, recalling that the law of \mathbf{a} given $\{\mathbf{x} = x\}$ is x ,

$$\Delta H(c) = \int_X H(x) dc(x) + \sum_a x_c(a) H(q(\cdot | a)) - H\left(\sum_a x_c(a) q(\cdot | a)\right),$$

which is clearly a continuous function of c . ΔH and π are thus continuous on the compact set C so V is compact. \square

We introduce the following notation:

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

This is the highest payoff associated with a convex combination of correlation systems under the constraint that the average entropy variation is nonnegative. For every correlation system c such that \mathbf{x} is almost surely constant, $\Delta H(c) \geq 0$ thus V intersects the half-plane $\{x_1 \geq 0\}$. Since V is compact, so is its convex hull and the supremum is indeed a maximum. The set V need not be convex as shown in Goldberg [7]; the supremum in the definition of w above might not be achieved by a point in V , but might be achieved by a convex combination involving two points of V with nonzero weights.

For computations, it is convenient to express the number w through the boundary of $\text{co } V$. 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 $\text{cav } u$ be the least concave function pointwise greater than u . Then

$$w = \text{cav } u(0).$$

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

Figure 1 illustrates how the map $\text{cav } u$ and the value w are derived from the set V .

4.2. A characterization of asymptotic max min values.

THEOREM 5. *The max min 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 as n goes to infinity. This limit coincides with the uniform max min which is*

$$\lim_{\lambda} v_{\lambda} = \lim_n v_n = v_{\infty} = w.$$

5. Examples.

5.1. Perfect observation. We say that the observation is *perfect* when the signal s reveals the action profile a , i.e., $a \neq a' \Rightarrow \text{supp } q(\cdot | a) \cap \text{supp } q(\cdot | a') = \emptyset$. 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., $w = v_1 = \max_{x \in X} \min_b g(x, b)$. Now we verify 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}) is independent of \mathbf{x} . This implies that $\Delta H(c) = 0$ if and only if c is a Dirac measure on some $x \in X$.

Now let $(x_1, x_2) \in \text{co } V$ such that $x_1 \geq 0$. We can write (x_1, x_2) as a convex combination:

$$(x_1, x_2) = \sum_k \lambda_k (\Delta H(c_k), \pi(c_k)).$$

From the above discussion, for each k such that $\lambda_k > 0$, c_k is a Dirac measure on some $x_k \in X$; thus $\pi(c_k) = \min_b g(x_k, b) \leq v_1$. Therefore, $x_2 \leq v_1$ and also $w \leq v_1$, hence the equality.

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, there is no limitation on the correlation the team may achieve by exchanging some messages; thus $w = \text{val } g = \max_{x \in \Delta(A)} \min_b g(x, b)$, which is the correlated max min of player II. Applying our main theorem, we remark that if observation is trivial, $\Delta H(c) \geq 0$ for each c . Let $x \in \Delta(A)$ that achieves the max in $\text{val } g$ and let c be such that the distribution induced on actions is $x_c = x$ (e.g., decompose x as a convex combination of pure action profiles). One has $\Delta H(c) \geq 0$ and $\pi(c) = \text{val } g$; thus $w = \text{val } g$.

5.3. $\frac{3}{8}$ is not optimal in the main example. We revisit our main example, i.e., the following three-player game where Player 1 chooses rows, Player 2 chooses columns, and Player 3 chooses matrices:

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

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 $2n + 1$. At stage $2n + 2$, the team plays (a, a) if $\mathbf{a}_{2n+1}^1 = a$ and (b, b) if $\mathbf{a}_{2n+1}^1 = b$. This strategy consists of playing alternately 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 variation of 0. Now we deduce from Theorem 5 the existence of strategies for Players 1 and 2 that guarantee more than $\frac{3}{8}$. By Theorem 5, 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 nonnegative average entropy variation.

Define the correlation system c_ε which puts equal weights on $(1 - \varepsilon, \varepsilon) \otimes (1, 0)$ and $(\varepsilon, 1 - \varepsilon) \otimes (0, 1)$: $c_\varepsilon(\{(1 - \varepsilon, \varepsilon) \otimes (1, 0)\}) = c_\varepsilon(\{(\varepsilon, 1 - \varepsilon) \otimes (0, 1)\}) = \frac{1}{2}$. We have $\pi(c_\varepsilon) = (1 - \varepsilon)/2$ 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 $(\Delta H(c_\varepsilon), \pi(c_\varepsilon))$ lies above the line

$$\left\{ \lambda \left(-1, \frac{1}{2} \right) + (1 - \lambda) \left(1, \frac{1}{4} \right), \lambda \in [0, 1] \right\}.$$

For this ε , there exists $0 \leq \lambda \leq 1$ 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 more than $\frac{3}{8}$.

Figure 2 gives a geometric illustration of the fact that playing c_ε and c_{+1} with frequencies λ and $1 - \lambda$ yields a payoff above $\frac{3}{8}$.

6. Computing w . In §4, the maxmin w is characterized as $\text{cav } u(0)$ with $u(h) = \max\{\pi(c) \mid c \in C, \Delta H(c) \geq h\}$ so the numerical computation of w consists of computing the function $u(h)$, i.e., in solving the associated optimization problem. This task proves to be difficult. In the paper, Gossner et al. [10], we develop

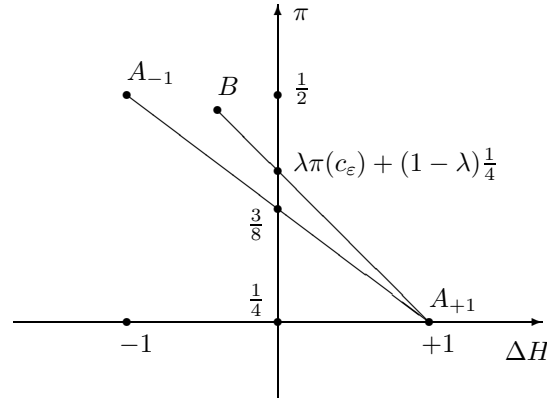


FIGURE 2. $A_{+1} = (\Delta H(c_{+1}), \pi(c_{+1}))$, $A_{-1} = (\Delta H(c_{-1}), \pi(c_{-1}))$, $B = (\Delta H(c_\epsilon), \pi(c_\epsilon))$.

tools to solve it starting from the following observations. In this maximization problem, the objective function $\pi(c) = \min_b g(x_c, b)$ depends on the correlation system c through its barycenter x_c only. Also, if we look at the constraint in the expression

$$\Delta H(c) = \int_X H(x) dc(x) + \sum_a x_c(a)H(q(\cdot | a)) - H\left(\sum_a x_c(a)q(\cdot | a)\right),$$

the second and third terms depend only on x_c . Only the first term $H(\mathbf{a} | \mathbf{x}) = \int H(x) dc(x)$ depends on the way the distribution c averages on x_c . We argue that fixing the barycenter x_c , we may choose any other c' that also averages on x_c provided that $\int H(x) dc'(x) \geq \int H(x) dc(x)$. In Gossner et al. [10], we study the problem of how to generate a correlated distribution of actions x^* through a correlation system c , while maximizing the expected entropy: $\max_{c: x_c = x^*} \int H(x) dc(x)$. Note that this latter problem is independent both of the game and of the signalling structure; thus its solution is helpful in solving all the instances covered by Theorems 5 and 14. The paper Gossner et al. [10] studied this auxiliary problem and solved the case where the team consists of two players, each of them having two actions. The solution and its proof are rather involved and the reader is referred to Gossner et al. [10] for the statement of the solution. Building on this result, two examples of games and signalling structures have been completely resolved so far: one in Gossner et al. [10] and one in Goldberg [7].

Note that for each $h \in \mathbb{R}$, either $\text{cav } u(h) = u(h)$ or $\text{cav } u$ is linear on some interval containing h . Thus, either $\text{cav } u(0) = \pi(c)$ for some c s.t. $\Delta H(c) \geq 0$ or there exists c_1, c_2 , and $\lambda \in (0, 1)$ s.t. $\text{cav } u(0) = \lambda\pi(c_1) + (1 - \lambda)\pi(c_2)$ and $\lambda\Delta H(c_1) + (1 - \lambda)\Delta H(c_2) \geq 0$. In the first case, the optimal strategy can be thought of 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 repeatedly plays two phases. Assume without loss of generality $\Delta H(c_1) > 0$. In a first phase, the optimal strategy plays c_1 to accumulate entropy; in a second phase, the optimal strategy plays c_2 , spending entropy to yield a good payoff. The relative lengths of these phases are $(\lambda, 1 - \lambda)$. Gossner et al. [10] showed that our main example is of the first kind and Goldberg [7] exhibited an example of the second.

We consider once more the main example. In this case, Gossner et al. [10] proved that the only points v in the set V that are undominated (i.e., $(v + \mathbb{R}_+^2) \cap V = \{v\}$) are of the form $v = (\Delta H(c), \pi(c))$ where c is a correlation system of the form

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

for some $x \in [0, 1]$. Such a correlation system has the following properties: for each x , the marginal distribution of actions under c is $(\frac{1}{2}, \frac{1}{2})$ for each player and the respective probabilities of (a, a) and (b, b) are equal. It follows that the associated payoff is $\pi(c) = \frac{1}{2}x^2 + \frac{1}{2}(1 - x)^2$; the entropy variation is $\Delta H(c) = 2H(x, 1 - x) - 1$.

The graph of $h \mapsto u(h)$ is then the parametric curve:

$$\left\{ \left(2H(x, 1 - x) - 1, \frac{1}{2}x^2 + \frac{1}{2}(1 - x)^2 \right) : x \in [0, 1] \right\}.$$

This curve is concave (this is easily checked by computing the slope of this curve at x) thus $w = \text{cav } u(0) = u(0)$; i.e., $w = \frac{1}{2}x^2 + \frac{1}{2}(1 - x)^2$ where $0 < x < \frac{1}{2}$ is such that $\Delta H(c) = 2H(x, 1 - x) - 1 = 0$. Numerically, this gives $w \cong 0.401$.

The graph of u can be seen in Figure 3.

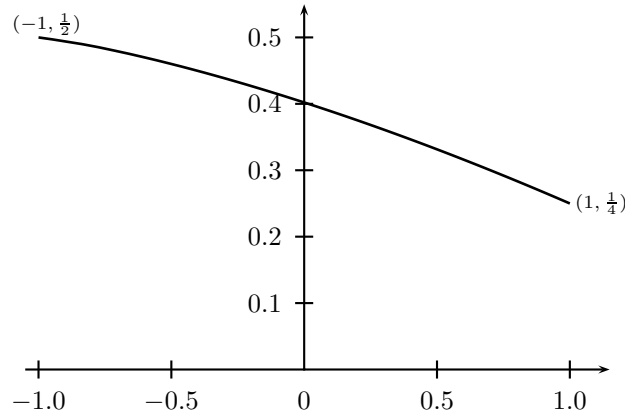


FIGURE 3. The graph of u for the main example.

7. Proof of the main results.

7.1. Player II defends w . Here we prove that for every strategy of the team, if player II plays stage-best replies, the average vector of (payoffs, entropy variation) generated belongs to V . This latter implies that no strategy for the team can guarantee a better payoff than w . The proof follows the same lines as some previous papers using entropy methods (see, e.g., Neyman and Okada [18, 19] and Gossner and Vieille [?]).

DEFINITION 6. 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, $\tau_\sigma(\emptyset) \in \arg \min_b g(\sigma(\emptyset), b)$ where \emptyset is the null history that starts the game. Assume that τ_σ is defined on histories of length less than $n + 1$. For every history h_n^{II} of player II, let $x_{n+1}(h_n^{\text{II}}) \in \Delta(A)$ be the distribution of the action profile of the team at stage $n + 1$ given h_n^{II} and let $\tau_\sigma(h_n^{\text{II}})$ be in $\arg \min_b g(x_{n+1}(h_n^{\text{II}}), b)$.

LEMMA 7. *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) ; let $\mathbf{h}_m^{\text{II}} = (\mathbf{b}_1, \mathbf{s}_1, \dots, \mathbf{b}_{m-1}, \mathbf{s}_{m-1})$ be the history of player II before stage m , and let $\mathbf{h}_m = (\mathbf{a}_1, \mathbf{b}_1, \mathbf{s}_1, \dots, \mathbf{a}_{m-1}, \mathbf{b}_{m-1}, \mathbf{s}_{m-1})$ be the total history. Let $\mathbf{x}_m = \sigma(\mathbf{h}_m)$ and $c_m(\mathbf{h}_m^{\text{II}})$ be the distribution of \mathbf{x}_m conditional on \mathbf{h}_m^{II} ; i.e., $c_m(\mathbf{h}_m^{\text{II}})$ 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 $\min_b g(\mathbf{E}_{\sigma, \tau}[\mathbf{x}_m | \mathbf{h}_m^{\text{II}}], b) = \pi(\mathbf{c}_m)$ from the definition of τ and thus $\gamma_n(\sigma, \tau) = \mathbf{E}_{\sigma, \tau}[(1/n) \sum_{m=1}^n \pi(\mathbf{c}_m)]$.

We set $H_m = H(\mathbf{a}_1, \dots, \mathbf{a}_m | \mathbf{h}_{m+1}^{\text{II}})$. Using the additivity of entropy, we have

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

Thus,

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

where the second equality holds since \mathbf{a}_m and \mathbf{b}_m are independent conditional on \mathbf{h}_m^{II} , the third equality holds since b_m is h_m^{II} -measurable, and the fourth equality holds since (a_m, s_m) depends on h_m only through x_m . We deduce

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

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

COROLLARY 8. *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, e.g., Lehrer and Sorin [17]):

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

From Lemma 7, we get $\gamma_\lambda(\sigma, \tau_\sigma) \leq w$ and thus $v_\lambda \leq w$. \square

7.2. v_n converges to w . We introduce 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 $i \in I$, $\sigma^i: \bigcup_n (A \times S)^n \rightarrow \Delta(A^i)$. Against a profile of autonomous strategies, the myopic best reply is a true best reply.

LEMMA 9. *Let σ be a profile of autonomous strategies, for each stage n and strategy τ for player II, $\mathbf{E}_{\sigma, \tau_\sigma} g(\mathbf{a}_n, \mathbf{b}_n) \leq \mathbf{E}_{\sigma, \tau} g(\mathbf{a}_n, \mathbf{b}_n)$. Thus τ_σ is player II's best reply in any version of the repeated game.*

PROOF. Consider the optimization problem of player II,

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

Since player II's moves do not influence the play of the team, this optimization problem is equivalent to solving $\min_b \mathbf{E}_{\sigma} [g(\mathbf{a}_n, b) | h_n^I]$ for n and each history h_n^I . The same argument applies in the n -stage game. \square

Now we associate autonomous strategies to distributions on strings of actions and signals. Note that for every autonomous strategy σ , the induced distribution \mathbf{P}_σ on $(A \times S)^\infty$ is such that for every history $h_n^I = (a_1, s_1, \dots, a_n, s_n)$, $\mathbf{P}_\sigma(a_{n+1}, s_{n+1} | h_n^I) = \prod_i \sigma^i(h_n^I)(a^i) \cdot q(s | a)$. We let Y be the set of probability distributions y on $A \times S$ for which there exists $x \in X$ such that for each (a, s) , $y(a, s) = \prod_i x^i(a^i) \cdot q(s | a)$.

We call a distribution \mathbf{P} on $(A \times S)^\infty$ a Y -distribution if at each stage n , after \mathbf{P} -almost every history $h_n^I = (a_1, s_1, \dots, a_n, s_n) \in H_n^I$, the distribution of $(\mathbf{a}_{n+1}, \mathbf{s}_{n+1})$ conditional on h_n^I , $\mathbf{P}(\mathbf{a}_{n+1}, \mathbf{s}_{n+1} | h_n^I)$ belongs to Y . Every autonomous strategy profile induces a Y -distribution; conversely, a Y -distribution defines an autonomous strategy profile (up to histories with probability 0: for these histories, the strategy is arbitrarily defined).

Given an autonomous strategy profile σ or equivalently a Y -distribution, consider the random correlation system at stage n : given \mathbf{h}_n^I , \mathbf{c}_n is the distribution of $\sigma(\mathbf{h}_n^I)$. The random variable \mathbf{c}_n is \mathbf{h}_n^I -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}_n^I . We define the random variable

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

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

$$\gamma_n(\sigma, \tau_\sigma) = \mathbf{E}_{\sigma, \tau_\sigma} \left[\frac{1}{n} \sum_{m=1}^n \pi(c_m) \right] = \mathbf{E}_{\sigma, \tau_\sigma} [\mathbf{E}_{\mathbf{d}_n} \pi] = \mathbf{E}_{\delta_n(\sigma)} \pi.$$

From Gossner and Tomala [11, Theorem 2.2], we deduce:

LEMMA 10. *For every $\delta \in \Delta(C)$ such that $\mathbf{E}_\delta \Delta H \geq 0$, there exists a Y -distribution \mathbf{P} on $(A \times S)^\infty$ such that $\mathbf{E}_\mathbf{P} \mathbf{d}_n$ weak-* converges to δ .*

Since a Y -distribution \mathbf{P} corresponds to an autonomous strategy, there exists σ autonomous such that $\delta_n(\sigma)$ weak- converges to δ .*

Note that Theorem 2.2 (Gossner and Tomala [11]) applies to an observer (here player II) who gets deterministic signals on a stochastic process. The process may be constrained in such a way that transitions belong to a fixed closed subset of probability distributions. When applying Theorem 2.2 to prove Lemma 10, we assume that the team chooses the pair (a, s) at each stage and is constrained in that the law of (a, s) conditional on the past history belongs to $y \in Y$. Since the transition q is fixed, choosing $y = x \otimes q \in Y$ is equivalent to choosing $x \in X$, so the construction is legitimate.

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 lemma yields the existence of an autonomous strategy σ such that $\lim_n \gamma_n(\sigma, \tau_\sigma) = \mathbf{E}_\delta \pi$. From Lemma 9, this gives $\liminf_n v_n \geq \mathbf{E}_\delta \pi$. \square

We may now conclude the proof. The set of vectors $(\mathbf{E}_\delta \Delta H, \mathbf{E}_\delta \pi)$ as δ varies in $\Delta(C)$ is $\text{co } V$; thus $\sup\{\mathbf{E}_\delta \pi \mid \delta \in \Delta(C), \mathbf{E}_\delta \Delta H \geq 0\} = w$. From Lemmas 7 and 11 we get $\lim_n v_n = w$.

7.3. 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 $\gamma_n(\sigma, \tau_\sigma) \geq w - \varepsilon/2$. 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

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

So,

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

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

7.4. Proof of the existence and value of v_∞ . From Lemma 7, by playing stage-best replies player II defends w . On the other hand, team I guarantees v_n by playing cyclically an optimal strategy in the n -stage game, thus I guarantees $\lim v_n = w$.

8. More general signalling structures. The method developed in this paper, and thus Theorem 5, extends to a larger class of signals than those presented in §3. Note, indeed, that our proof relies on only the following three conditions: (i) the signal of player II does not depend on his own action; (ii) the information regarding actions and signals that are unobserved by player II is symmetric within the team; and (iii) each team player knows the information of player II regarding those actions and signals. Condition (i) means that the entropy variation is not controlled by player II, which ensures that player II best-responds in the repeated game by optimizing myopically. Otherwise, player II faces a tradeoff between best-responding (potentially allowing team players to get a large amount of entropy) and minimizing the entropy produced by his choice. The case where the entropy variation depends on the action of player II is under investigation. Conditions (ii) and (iii) mean that each team player is able to compute the entropy variation. They thus agree on how to use the available entropy for correlation.

Consider the following signalling structure: if the team plays the action profile a and player II plays action b , then

- a pair of signals $(s, t) \in S \times T$ is drawn from a pair of finite sets S, T according to $q(\cdot \mid a)$ with $q: A \rightarrow \Delta(S \times T)$. The tuple (a, s, t) is observed by each team player. Player II observes (b, s) ;
- each player $i \in I$ observes a private signal $r^i = f^i(a, b)$ which is a deterministic function of the action profile.

These signalling structures generalize those of §3 in two respects. First, team players do not fully observe the move of player II. Second, they get to observe a random signal t that depends on the action profile. For instance, the generalization includes the case where all actions are perfectly observed and the team gets to privately observe at each stage the realization of a random variable. Note that the more-general signaling structure satisfies the requirements (i), (ii), and (iii) above: the only information asymmetry within the team is about the move of player II, which cannot be used for correlation. Theorem 5 extends naturally to these signalling structures.

The definition of the optimal payoff $\pi(c)$ associated with a correlation system c is unchanged. The definition of the entropy variation generalizes as follows:

DEFINITION 13. Let c be a correlation system and $(\mathbf{x}, \mathbf{a}, \mathbf{s}, \mathbf{t})$ be a random variable in $X \times A \times S \times T$ 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}, \mathbf{t}) given $\{\mathbf{a} = a\}$ is $q(\cdot \mid a)$. The entropy variation of c is

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

With this adaptation, we still consider the set of feasible vectors $(\Delta H(c), \pi(c))$ in the (entropy variation, payoff) plane:

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

and we derive the quantity

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

THEOREM 14. *Under the generalized signaling structure, the max min 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. This limit is*

$$\lim_{\lambda} v_{\lambda} = \lim_n v_n = w.$$

Furthermore, the uniform max min exists and takes the value w .

PROOF. We begin by extending the proof of Lemma 7. We first modify the signalling structure by assuming that the actions of player II are publicly observable. In this modified game, since the signals r^i are deterministic functions of b , the set of strategies of the team is larger, so if player II defends w in the modified game, it is also true in the original game.

The crux is to prove that

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

Set $H_m = H(\mathbf{a}_1, \mathbf{t}_1, \dots, \mathbf{a}_m, \mathbf{t}_m \mid \mathbf{h}_{m+1}^{\text{II}})$. Using the additivity of entropy,

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

Thus,

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

and the rest of the proof carries out.

Second, to prove that the team guarantees w , define an autonomous strategy as a strategy that does not depend on the signals r^i ; i.e., it depends solely on the action profiles a and on the signals (s, t) . We let $Y \subset \Delta(A \times S \times T)$ be the set of probability distributions y such that $\forall(a, s, t), y(a, s, t) = x(a)q(s, t \mid a)$ for some $x \in X$. This is the set of distributions on $A \times S \times T$ that can be obtained by a profile of mixed strategies x of the team and the transition q . An autonomous strategy can be identified with a probability distribution \mathbf{P} on $(A \times S \times T)^\infty$ such that at each stage n , after \mathbf{P} -almost every history $h_n^1 = (a_1, s_1, t_1, \dots, a_n, s_n, t_n) \in H_n^1$, the distribution $\mathbf{P}(\mathbf{a}_{n+1}, \mathbf{s}_{n+1}, \mathbf{t}_{n+1} \mid h_n^1)$ belongs to Y . We may thus apply Lemma 10 to Y -distributions and conclude as in Theorem 5. \square

9. Consequences for 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 et al. [1], Lehrer [14], Fudenberg et al. [6]), equilibria in which a communication mechanism serves to resolve information asymmetries (see Compte [4], Kandori and Matsushima [13], Renault and Tomala [21]), or two-player games (Lehrer [15, 16]). In our approach, we tackle information asymmetries by measuring them with the entropy function.

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

Consider a game in which the set of players is $\{1, \dots, n\}$, $n \geq 4$, and in which i 's finite action set is A^i . Players $i = 2, \dots, n - 1$, have perfect observation: they observe $s^i = (a^1, \dots, a^n)$. Player 1 observes every

player but player n 's signal is $s^1 = (a^1, \dots, a^{n-1})$. Player n observes every player but Player 1's signal is $s^n = (a^2, \dots, a^n)$. This structure of signals is represented in Renault and Tomala [20] 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 two-connected: there are at least two distinct paths from i to j for each pair (i, j) .

Let $\text{cog}(A)$ be the set of feasible payoffs. To define the individually rational level of player i in the repeated game, we consider the game played by the team $-i$ —i.e., all players but i —against player i (thus, with payoff $-g^i$), and we let v^i be the associated uniform value. We set then $IR = \{x \in \mathbb{R}^n, x^i \geq v^i\}$, the set of individually rational payoffs with respect to the min max values of the repeated game. Renault and Tomala [20] proved that in a repeated game where each player monitors the actions of his neighbors in a fixed graph the set of uniform equilibrium payoffs equals $\text{cog}(A) \cap IR$ when the graph is two-connected. However, Renault and Tomala [20] left open the characterization of min max values of the repeated game.

Since each Player $1 < i < n$ has perfect observation, his individually rational level v^i in the repeated game equals his independent min max v^i . Regarding player n (respectively, Player 1), we may apply Theorem 14. Signals are deterministic: team players in $\{1, \dots, n-1\}$ fully observe the team action profile, and each of them gets to observe a signal on the move of player n . Player 1 observes a constant signal and the other players observe this move. We thus get a complete characterization of the set of uniform equilibrium payoffs.

Lehrer [14] characterized Nash equilibrium payoffs for all repeated games having a semistandard 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.

References

- [1] Abreu, D., D. Pearce, E. Stacchetti. 1990. Toward a theory of discounted repeated games with imperfect monitoring. *Econometrica* **58** 1041–1063.
- [2] Aumann, R. J. 1974. Subjectivity and correlation in randomized strategies. *J. Math. Econom.* **1** 67–95.
- [3] Bavly, G., A. Neyman. 2003. Online concealed correlation by boundedly rational players. Discussion Paper Series 336, Center for Rationality and Interactive Decision Theory, Hebrew University, Jerusalem, Israel.
- [4] Compte, O. 1998. Communication in repeated games with imperfect private monitoring. *Econometrica* **66** 597–626.
- [5] Forges, F. 1986. An approach to communication equilibria. *Econometrica* **54** 1375–1385.
- [6] Fudenberg, D., D. K. Levine, E. Maskin. 1994. The Folk theorem with imperfect public information. *Econometrica* **62** 997–1039.
- [7] Goldberg, Y. 2007. Secret correlation in repeated games with imperfect monitoring: The need for nonstationary strategies. *Math. Oper. Res.* **32**(2) 425–435.
- [8] Gossner, O. 1998. Repeated games played by cryptographically sophisticated players. DP 9835, CORE, Université Catholique de Louvain, Louvain, Belgium.
- [9] Gossner, O. 2000. Sharing a long secret in a few public words. DP 2000-15, THEMA, Nanterre, France.
- [10] Gossner, O., R. Laraki, T. Tomala. 2006. On the optimal use of coordination. *Math. Programming B.* Forthcoming.
- [11] Gossner, O., T. Tomala. 2006. Empirical distributions of beliefs under imperfect observation. *Math. Oper. Res.* **31**(1) 13–30.
- [12] Hörner, J., W. Olszewski. 2006. The Folk theorem with private almost-perfect monitoring. *Econometrica* **74**(6) 1499–1544.
- [13] Kandori, M., H. Matsushima. 1998. Private observation, communication and collusion. *Rev. Econom. Stud.* **66** 627–652.
- [14] Lehrer, E. 1990. Nash equilibria of n player repeated games with semi-standard information. *Internat. J. Game Theory* **19** 191–217.
- [15] Lehrer, E. 1991. Internal correlation in repeated games. *Internat. J. Game Theory* **19** 431–456.
- [16] Lehrer, E. 1992. Correlated equilibria in two-player repeated games with nonobservable actions. *Math. Oper. Res.* **17** 175–199.
- [17] Lehrer, E., S. Sorin. 1992. A uniform Tauberian theorem in dynamic programming. *Math. Oper. Res.* **17** 303–307.
- [18] Neyman, A., D. Okada. 1999. Strategic entropy and complexity in repeated games. *Games and Econom. Behav.* **29** 191–223.
- [19] Neyman, A., D. Okada. 2000. Repeated games with bounded entropy. *Games and Econom. Behav.* **30** 228–247.
- [20] Renault, J., T. Tomala. 1998. Repeated proximity games. *Internat. J. Game Theory* **27** 539–559.
- [21] Renault, J., T. Tomala. 2004. Communication equilibrium payoffs in repeated games with complete information and imperfect monitoring. *Games Econom. Behavior* **49**(2) 313–344.
- [22] von Stengel, B., D. Koller. 1997. Team max min equilibria. *Games and Econom. Behav.* **21** 309–321.