Correlated Equilibria in Two-Player Repeated Games with Nonobservable Actions
Author(s): Ehud Lehrer
Source: *Mathematics of Operations Research*, Vol. 17, No. 1 (Feb., 1992), pp. 175-199
Published by: INFORMS
Stable URL: http://www.jstor.org/stable/3689900
Accessed: 22/08/2011 07:45

MATHEMATICS OF OPERATIONS RESEARCH
Vol. 17, No. 1, February 1992
Printed in U.S.A.

CORRELATED EQUILIBRIA IN TWO-PLAYER REPEATED GAMES WITH NONOBSERVABLE ACTIONS*

EHUD LEHRER

Four kinds of correlated equilibrium payoff sets in undiscounted repeated games with nonobservable actions are studied. Three of them, the upper, the uniform, and Banach lead to the same payoff set, whereas the lower one in general is associated with a larger set. The extensive form correlated equilibrium is also explored. It turns out that both the regular and extensive form correlated equilibria yield the same sets of payoffs.

1. Introduction. In repeated games with nonobservable actions a player gets, after each stage, a signal that depends on the joint action played. This signal does not reveal necessarily the opponents' actions nor does it reveal their payoffs. The question naturally arises: What are the possible equilibrium outcomes and how do players use the information they collected during the game?

We confine ourselves to undiscounted repeated games, where the payoffs are determined by the limit of partial average of the stage payoffs. This model enables one to examine the long-run impact of imperfect monitoring. The paper characterizes several types of long-term correlated equilibrium payoffs in two-player repeated games with nonobservable actions.

Correlated equilibrium (introduced by Aumann in [A1]) allows the players to utilize an exogenous mediator who provides each one with private information. The players, based on this private information, adopt a pure strategy to be played in the repeated game. Such coordination between players may, in general, sustain equilibrium payoffs that were not supportable by an equilibrium without it (namely, by regular Nash equilibrium). The correlated equilibrium can be thought of also as a Nash equilibrium of an extended game in which a mediator sends messages to the players and then they choose strategies.

Correlated equilibrium is a more attractive solution concept than Nash equilibrium for several reasons: (1) it better reflects real-life phenomena in which players may condition their behavior on their private information; (2) it allows for coordination excluded by Nash equilibrium; and (3) it is simpler to compute (see [GZ] and [HS]). In repeated games with imperfect monitoring the introduction of a mediator facilitates the characterization of the equilibrium payoffs set and simplifies the supporting equilibrium strategies.

In addition to the regular correlated equilibrium in which a mediator coordinates between the players before starting the game and then disappears, we present the extensive form correlated equilibrium (introduced by Forges [F1]). In this type of correlated equilibrium the mediator remains active all over the game. He sends messages to each participating player before each stage. In general, the extensive form type sustains a larger set of equilibrium payoffs than the regular one. However,

*Received November 11, 1987; revised July 22, 1990.
AMS 1980 subject classification. Primary: 90D05. Secondary: 90D20.
IAOR 1973 subject classification. Main: Games.
OR/MS Index 1978 subject classification. Primary: 231 Games/group decisions.
Key words. 2-person repeated games, equilibrium payoff sets, correlated equilibrium.

0364-765X/92/1701/0175/$01.25

so it turns out, in the model investigated here both yield the same set of equilibrium payoffs.

Four types of long-run equilibria are defined: the upper, the uniform, Banach, and lower. The payoffs sets associated with the first three coincide, whereas the one associated with the lower equilibrium is usually greater. The various types of equilibria differ in the ways players evaluate possible deviations. The upper equilibrium corresponds to "optimistic" players for whom the best periods matter most. The uniform equilibrium (see [S]) concept views the infinite game as an "approximation" to large finitely repeated games. Thus, a joint strategy is a uniform equilibrium if it induces an ϵ-equilibrium in a sufficiently long, finitely repeated game. The Banach equilibrium concept incorporates a Banach limit in order to evaluate profitability of possible deviation. The lower equilibrium relates to "pessimistic" players, taking into account the worst averages they are about to experience.

We find, unexpectedly, that the set of lower correlated equilibrium payoffs coincide with the set of Nash lower equilibrium payoffs. In other words, the correlation device does not enlarge the players' possibilities (in terms of payoffs). However, the payoff sets corresponding to other correlated equilibria types are, in general, larger than the respective Nash equilibrium payoffs sets.

In order to describe the main results of the paper, two relations between a player's actions must be introduced. Two actions of a player are *indistinguishable* if they yield the same signal for the opponent, no matter what the latter is playing. In other words, the opponent cannot distinguish between two indistinguishable actions of a player. One might think that if a player who is assigned to play a certain action decides to play another action, indistinguishable from the assigned one, the opponent will not be able to detect the deviation. As was pointed out by [L2] and [L4], this is not the case.

A player can deviate to an indistinguishable action but a less informative one, that is, to an action by which he is able to collect less information. By playing a less informative action a player will know less about previous actions of his opponent. In a communication phase of the repeated game strategies, to be described in detail below, the player can discern that his opponent knows less than what he should know had he adhered to the prescribed action. Thereby, players can detect a deviation to an action which is indistinguishable from the prescribed action but less informative than it. Thus, in order to define an undetectable deviation one should introduce another relation. An action a' is *more informative* than a, if by playing a', a player can distinguish between two actions of his opponent better than by playing a.

It is shown that any deviation from the prescribed action to another, either distinguishable from it or less informative than it, is detectable. Moreover, any other deviation is not detectable. In equilibrium, a player will not have an incentive to deviate because all possible deviations are either detectable (and the player is threatened by punishment) or undetectable but also unprofitable.

The set of upper, uniform, or Banach correlated equilibrium payoffs is characterized as the set of all the individually rational payoffs of the following form. They should be associated with correlated actions (probability distribution over the joint pure actions) in which any action assigned a positive probability is a best response among the class of actions, which are indistinguishable from and more informative than itself.

On the other hand, the set of lower correlated equilibrium payoffs set is characterized by the individually rational payoffs associated with two (possibly different) correlated actions. In the first one, actions of player 1 are best responses (among the class of actions, etc., as above) and in the second, actions of player 2 are best responses (among the class of actions, etc.). Obviously, this set is larger than the one corresponding to the upper equilibrium.

The paper contains six sections. The model and various equilibria types are presented in §2. §3 is devoted to the definition of the relation just described, and to the formulation of the main theorems. §4 and §5 provide the proofs of the theorems. §6 contains comments on some alternative approaches and on some possible extensions.

2. The model.

2.1. *The components of the game.* The two-player repeated game with nonobservable actions consists of:

(i) Two finite sets of actions Σ_1 and Σ_2. Set $\Sigma = \Sigma_1 \times \Sigma_2$.

(ii) Two information functions l_1 and l_2 and two signals sets L_1 and L_2, s.t. l_1: $\Sigma \to L_i$. Elements in L_i are called *signals*.

(iii) Two payoff functions h_1 and h_2, where h_i: $\Sigma \to \mathbb{R}$.

2.2. *Pure strategies.* A *pure strategy* of player i is a sequence of functions $(f^1, f^2, \dots)$ s.t. f^t: $L_i^{t-1} \to \Sigma_i$, where L_i^{t-1} is the Cartesian product of L_i with itself $t-1$ times. Denote by Σ_i^* the set of all pure strategies of player i in the repeated game. A pair of pure strategies $(f, g) \in \Sigma_1^* \times \Sigma_2^*$ is called a joint strategy.

A joint strategy (f, g) induces two sequences $\{x_i^t\}_{t=1}^{\infty}$, $i = 1, 2$, of numbers, where x_i^t is the payoff of player i at stage t.

2.3. *Upper correlated equilibrium.* An *upper correlated equilibrium* is a tuple $(A \times B, \mathscr{A} \times \mathscr{B}, P, \sigma, \tau)$, where

(i) $A \times B$ is a product set of points;

(ii) $\mathscr{A} \times \mathscr{B}$ is a σ-algebra of $A \times B$;

(iii) P is a probability measure defined on $\mathscr{A} \times \mathscr{B}$;

(iv) σ is a measurable function from $(A, \mathscr{A})$ to Σ_1^*;

(v) τ is a measurable function from $(B, \mathscr{B})$ to Σ_2^*, satisfying:

$$\text{(1a)} \qquad \lim_T E_{\sigma,\tau,P}\left[(1/T) \sum_{t=1}^{T} x_i^t\right] \quad \text{exists for } i = 1, 2.$$

Denote it by $H_i^*(\sigma, \tau)$.

$$\text{(1b)} \qquad \limsup_T E_{\bar{\sigma},\tau,P}\left[(1/T) \sum_{t=1}^{T} x_1^t\right] \leqslant H_1^*(\sigma, \tau) \quad \text{for all } \bar{\sigma}.$$

$$\text{(1c)} \qquad \limsup_T E_{\sigma,\bar{\tau},P}\left[(1/T) \sum_{t=1}^{T} x_2^t\right] \leqslant H_2^*(\sigma, \tau) \quad \text{for all } \bar{\tau}.$$

Denote by UCEP the set of all the upper correlated equilibrium payoffs $(H_1^*(\sigma, \tau), H_2^*(\sigma, \tau))$.

2.4. *The lower correlated equilibrium.* The *lower correlated equilibrium* is defined as the upper one with the change that liminf replaces limsup in (1b) and (1c).

Denote by LCEP the set of all lower correlated equilibrium payoffs. Obviously UCEP $\subseteq$ LCEP. The lower and the upper equilibria differ in the way an infinite stream of payoffs is evaluated by the players. The former corresponds to "pessimistic" players taking into account the worst averages they are about to experience, while the latter assumes "optimistic" players for which the best periods matter most.

2.5. *The uniform correlated equilibrium.* A *uniform correlated equilibrium* is a tuple $U = (A \times B, \mathscr{A} \times \mathscr{B}, P, \sigma, \tau)$ for which and for every $\epsilon > 0$ there is T_0 s.t. if $T \geqslant T_0$ then U induces an ϵ-Nash equilibrium in the extended (including the messages of the mediator) T-truncated game. In other words, (1a) is satisfied and for every $\epsilon > 0$ there is a T_0 s.t. $T \geqslant T_0$ implies

$$(1b') \qquad E_{\bar{\sigma},\tau,P}\left[(1/T)\sum_{t=1}^{T} x_1^t\right] \leqslant H_1^*(\sigma,\tau) + \epsilon \quad \text{for all } T \geqslant T_0,$$

and a similar condition for player 2.

Denote the set of uniform correlated equilibrium payoffs by UNIC.

It is clear that UNIC $\subseteq$ UCEP.

2.6. *The Banach correlated equilibrium.* Let L be a Banach limit. A *Banach correlated equilibrium* is a tuple $(A \times B, \mathscr{A} \times \mathscr{B}, P, \sigma, \tau)$ which satisfies

$$L\left\{E_{\bar{\sigma},\tau,P}(1/T)\sum_{t=1}^{T} x_1^t\right\}_T \leqslant L\left\{E_{\sigma,\tau,P}(1/T)\sum_{t=1}^{T} x_1^t\right\}_T$$

for all $\bar{\sigma}$, and a similar inequality for $\bar{\tau}$, replacing x_1^t by x_2^t.

Denote by CEP_L the set of all L-Banach equilibrium payoffs.

2.7. *Description of the game in words.* Before the game starts a mediator chooses a point $(\alpha, \beta) \in A \times B$ according to P. He informs player 1 (hereafter PI) of α and player 2 (PII) of β. α and β are called messages. PI then plays in the repeated game according to the pure strategy $\sigma_\alpha = \sigma(\alpha)$ and PII plays according to $\tau_\beta = \tau(\beta)$, i.e., at the first stage PI plays σ_α^1 and PII plays τ_β^1. Denoting $z^1 = (\sigma_\alpha^1, \tau_\beta^1)$, player i receives the signal $s_i^1 = l_i(z^1)$ and the payoff $x_i^1 = h_i(z^1)$. At the second stage PI plays $\sigma_\alpha^2(s_1^1)$ and PII plays $\tau_\beta^2(s_2^1)$. Denoting $z^2 = (\sigma_\alpha^2(s_1^1), \tau_\beta^2(s_2^1))$, player i gets the signal $s_i^2 = l_2(z^2)$ and the payoff $x_i^2 = h_i(z^2)$, and so forth.

The choice of the particular pure strategies is done by functions σ and τ. These choice functions are in equilibrium if any other player's choice function would not increase his expected payoff in the repeated game, evaluated with either the upper, lower, Banach limit, or sufficiently large partial averages (which correspond to uniform equilibrium).

EXAMPLE 1. The repeated game of:

	b_1	b_2	b_3	b_4
a_1	6, 6	2, 7	6, 6	0, 0
a_2	7, 2	0, 0	0, 0	0, 0
a_3	6, 6	0, 0	0, 0	0, 0
a_4	0, 0	0, 0	0, 0	0, 0

payoffs

	b_1	b_2	b_3	b_4
a_1	λ, λ	λ, η	λ, γ	λ', δ
a_2	η, λ	η, η	η, γ	η', δ
a_3	γ, λ	γ, η	γ, γ	γ, δ
a_4	δ, λ'	δ, η'	δ, γ	ϵ, ϵ

signals

In this example, $\Sigma_1 = \{a_1, a_2, a_3, a_4\}$, $\Sigma_2 = \{b_1, b_2, b_3, b_4\}$ and $L_i = \{\lambda, \eta, \gamma, \delta, \lambda', \eta', \epsilon\}$, $i = 1, 2$. If, for instance, PI played a_2 and PII played b_1, the payoffs are 7 and 2 for PI and PII, respectively, and the signals are η and λ for PI and PII, respectively.

REMARK 1. In the framework described here the players are not allowed to randomize. Any randomization, if and when it takes place, should be provided by the mediator. However, all the messages are given to the players before starting the game. Therefore, the message should contain a random signal on which the players base their actions when the need of randomization arises (e.g., in case of punishment or in a case where a random stage is chosen). For our purposes it will be enough if player i will get in addition to previously mentioned messages also a string $(s_i^1, s_i^2, s_i^3, \dots)$, where s_i^t is drawn randomly from $[0, 1]$ according to the uniform distribution *independently of all other messages' components*. (Actually, it would suffice to get a message consisting of one number which is independently drawn from $[0, 1]$ according to the uniform distribution.)

In the sequel, when it is said that a player randomizes, it should be understood as a player bases his action on the random message he got from the mediator.

2.8. *An extensive form correlated equilibrium*. As opposed to the correlated equilibrium, where the mediator correlates between the players only before the game starts, we consider here a mediator who is active at all stages. Before stage t the mediator selects a message $(\alpha_t, \beta_t) \in (A_t, B_t)$ according to a probability distribution P_t, which may depend on his previous selected messages $\{(\alpha_s, \beta_s)\}_{s<t}$. Relying on all the signals and messages previously received, a player chooses an action to be played at stage t. For a more elaborate study see [F1] and [F2].

Precisely, an *upper extensive form correlated equilibrium* is a tuple $((\times_{t=1}^{\infty} A_t) \times (\times_{t=1}^{\infty} B_t), \mathscr{A} \times \mathscr{B}, P, f, g)$, where:

(i) $((\times_{t=1}^{\infty} A_t) \times (\times_{t=1}^{\infty} B_t), \mathscr{A} \times \mathscr{B}, P)$ is a product sample space;

(ii) $f = (f^1, f^2, \dots)$, where f^t is a measurable function from $L_1^{t-1} \times A_1 \times \cdots \times A_t$ to Σ_1; and

(iii) $g = (g^1, g^2, \dots)$, where g^t is a measurable function from $L_2^{t-1} \times B_1 \times \cdots \times B_t$ to Σ_2, satisfying convergence and incentive compatibility conditions. Namely, having the properties (1a)–(1c), replacing σ with f, τ with g, $\bar{\sigma}$ with $\bar{f}$, and $\bar{\tau}$ with $\bar{g}$.

Similarly, lower, uniform and Banach extensive form correlated equilibria can be defined. In the sequel, an asterisk attached to a correlated equilibrium payoffs set will indicate that the set corresponds to extensive form correlated equilibrium.

Notice that any correlated equilibrium payoff is an extensive form correlated equilibrium payoff (UECP $\subseteq$ UECP*, LECP $\subseteq$ LECP*, UNIC $\subseteq$ UNIC*, and $\text{CEP}_L \subseteq \text{CEP}_L^*$), but the opposite, typically, is incorrect. The following game, quoted from [M], is an example in which there is an extensive form correlated equilibrium payoff that is not a correlated equilibrium payoff.

EXAMPLE 2. PI chooses either t or $-t$. If he chooses t, both players get 2; otherwise, he should choose between m and b. PII is informed of the first choice of PI (i.e., either t or $-t$), but not of the second one. After PI takes his actions, PII should choose between l and r. The various combinations of actions result in payoffs depicted in Figure 1.

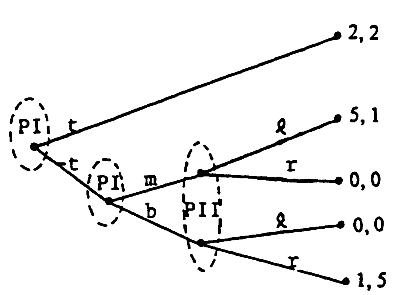

FIGURE 1

A mediator recommends to PI at the first node to choose $-t$. After PI has made his choice, the mediator picks either (m, l) or (b, r), with probability $\frac{1}{2}$ each. If the former was the outcome, he recommends PI to play m and PII to play l, and if the latter was the outcome he recommends b to PI and r to PII.

Notice that after the choice of $-t$ has been made (and then t is no longer available), neither player can gain by disobeying the mediator's recommendations, provided that the other player obeys them. The expected payoff of this correlation mechanism is 3 for both players.

However, $(3,3)$ cannot be achieved as a correlated equilibrium payoff, because PI would not be willing to play consecutively $-t$ and then b (and get the payoff 1) if he can play the dominating action t (and get the payoff 2).

3. The main theorems. In order to state the main results, a few notations are needed.

3.1. *Comparison of actions.* We first define two relations on the set of actions Σ_i. The first one is an equivalence relation and the other is a partial order.

Two actions, $a', a \in \Sigma_i$ are *indistinguishable* (denoted $a' \sim a$) if $l_{3-i}(a, b) = l_{3-i}(a', b)$ for any $b \in \Sigma_{3-i}$. In words, a and a' are indistinguishable if by playing any b, player $3 - i$ cannot distinguish between them.

The action a' is *more informative* than the action a if $l_i(a, b) \neq l_i(a, b')$ implies $l_i(a', b) \neq l_i(a', b')$, for any $b, b' \in \Sigma_{3-i}$. In other words, a' is more informative than a if, whenever by playing a, player i can distinguish between b and b' he can do so also by playing a'. Namely, by playing a', player i can collect more information about his opponent's action than by playing a.

An action a is *strictly less informative* than a' if a' is more informative than a and a is not more informative than a'.

EXAMPLE 3. In Example 1, a_1 and a_2 are indistinguishable, while a_1 and a_4 are distinguishable. In the same example a_2 is more informative than a_3, while a_3 is not more informative than a_2, because by playing a_2 PI can distinguish between b_3 and b_4, while he cannot do so by playing a_3.

3.2. *Correlated actions and the sets* $B_i, \tilde{B}_i$. Denote by Δ the set of probability distributions on Σ. Elements of Δ are referred to as *correlated actions*. We can extend the domain of the payoff function h_i to correlated actions. For any $Q \in \Delta$, define $h_i(Q) = \sum_{(a,b) \in \Sigma} Q(a,b) h_i(a,b)$. Let $h(Q) = (h_1(Q), h_2(Q))$.

For any joint pure action, $(a_0, b_0) \in \Sigma$, define

$$h_1(Q|a_0) = \sum_{b \in \Sigma_2} Q(a_0, b) h_1(a_0, b) \quad \text{and}$$

$$h_2(Q|b_0) = \sum_{a \in \Sigma_1} Q(a, b_0) h_2(a, b_0),$$

i.e., $h_1(Q|a_0)$ and $h_2(Q|b_0)$ are the unnormalized (the probabilities do not sum up to 1) expected payoffs of PI and PII, given the actions a_0 and b_0, respectively.

Now we are ready to define two sets, B_1 and B_2, consisting of correlated actions. A distribution Q in Δ is in B_1 if any action a_0, assigned a positive probability by Q, is a best response, versus the expected mixed action of PII, among the class of actions that are both indistinguishable from and more informative than itself. Precisely,

DEFINITION 1

$$B_1 = \Big\{ Q \in \Delta | h_1(Q|a_0) \geqslant \sum_{b \in \Sigma_2} Q(a_0, b) h_1(a, b)$$

$$\text{for all } \ a_0, a \in \Sigma_1 \ \text{ satisfying } \ a \sim a_0 \ \text{ and } a \text{ is more informative than } a_0 \Big\}.$$

Notice that if instead of playing a_0 PI plays a, his unnormalized payoff is $\sum_{b\in\Sigma_2} Q(a_0, b)h_1(a, b)$. B_2 is defined in a similar way.

In Definition 1 we required that a_0 is a best response among all actions a that satisfy $a \sim a_0$ and that a is more informative than a_0. Had we required in Definition 1 only $a \sim a_0$, namely, that a_0 is a best response among a greater class of actions (of those indistinguishable from a_0), we would have obtained a smaller set of distributions. Denote the set of all these distributions by $\tilde{B}_1$. Formally,

$$\tilde{B}_1 = \left\{ Q \in \Delta | h_1(Q|a_0) \geqslant \sum_{b\in\Sigma_2} Q(a_0, b)h_1(a, b) \right.$$
$$\left. \text{for all } a_0, a \in \Sigma_1 \text{ satisfying } a \sim a_0 \right\}.$$

$\tilde{B}_2$ is defined similarly. The set C_i (resp., D_i), defined initially in [L2], is defined like B_i (resp., $\tilde{B}_i$) with the additional qualification that the distribution Q is a product of its marginal distributions. Precisely, let Δ' be the set of all $Q \in \Delta$ s.t. Q is a product of its marginal distributions. Define $C_i = B_i \cap \Delta'$ and $D_i = \tilde{B}_i \cap \Delta'$. Thus, $C_i \subseteq B_i$ and $D_i \subseteq \tilde{B}_i$, $i = 1, 2$.

REMARK 2. (i) Any distribution induced by a Nash equilibrium (of the one-shot game) is included in $B_1 \cap B_2$. Thus, $B_1 \cap B_2$ is nonempty.

(ii) It can be easily verified that B_i and $\tilde{B}_i$ are convex sets.

EXAMPLE 4 (inspired by [A2]). In the following game, the information is either standard (the players are informed of the joint action played) or trivial (each player is informed solely of his own action). Standard information is indicated by an asterisk.

	b_1	b_2	b_3
a_1	0, 0	8, 0	0, 0
a_2	0, 8	6, 6	2, 7
a_3	0, 0	7, 2	0, 0*

Notice that under a standard-trivial information structure (like the one just described) two actions are indistinguishable if and only if both yield trivial information regardless of the opponent's action. Moreover, if an action a' is strictly more informative than a, then there is no action indistinguishable from the first.

Let Q be the distribution assigning $\frac{1}{3}$ to each of the pairs (a_3, b_2), (a_2, b_2) and (a_2, b_3). Notice that Q is not a correlated equilibrium in the one-shot game, since PI (resp., PII) can gain by deviating from a_3 (resp., b_3) to a_1 (resp., b_1). However, $a_1 \nsim a_3$ and $b_1 \nsim b_3$. Thus, $Q \in B_1 \cap B_2$.

3.3. *Some properties of* B_i, $\tilde{B}_i$. The following propositions are interesting in their own right and they will be used in Theorem 1's proof.

PROPOSITION 1. $\operatorname{conv} h(C_i) = h(B_i)$, *for* $i = 1, 2$.

PROOF. We will show this for $i = 1$. By Definition 1, $h(C_1) \subseteq h(B_1)$. By Remark 1(ii), $\operatorname{conv} h(C_1) \subseteq h(B_1)$. It remains to show the inverse inclusion.

Let $Q \in B_1$. Denote by Q_1 the marginal distribution over Σ_1. For any $a \in \Sigma_1$ s.t. $Q_1(a) > 0$, we have

$$\left(a, \sum_{b\in\Sigma_2} (Q(a, b)/Q_1(a))\delta_b \right) \in C_1,$$

where δ_b is the probability measure assigning a mass 1 to b. Thus,

$$h(Q) = \sum_{a \in \Sigma_1, Q_1(a) > 0} Q_1(a) h\left(a, \sum_{b \in \Sigma_2} (Q(a,b)/Q_1(a))\delta_b\right)$$

is included in conv $h(C_1)$. Therefore, $h(B_1) \subseteq \text{conv } h(C_1)$, and the proposition follows. //

Similarly, one can obtain:

PROPOSITION 2. conv $h(D_i) = h(\tilde{B}_i)$, *for* $i = 1, 2$.

Let Q be a correlated action. Denote by $UD_i(Q)$ (for undetectable) the set of all the correlated actions that are the results of undetectable deviations of player i. Namely, $Q' \in UD_1(Q)$ if (i) there is a map γ from Σ_1 to Σ_1 such that $\gamma(a) \sim a$ and $\gamma(a)$ is more informative than a and (ii) $Q'(a', b) = \Sigma_{a \in \gamma^{-1}(a')} Q(a, b)$ for any $(a', b) \in \Sigma$. $UD_2(Q)$ is defined similarly. In words, γ attaches to any action a an action which is indistinguishable from and more informative than itself. Moreover, the probability of the pair (a', b) w.r.t.[1] Q' is the sum of all the probabilities w.r.t. Q of all the pairs (a, b), where a is attached to a'.

The class $BR_i(Q)$ (for best response) is the set of the distributions in $UD_i(Q)$ which yield the maximum payoff to player i, i.e., $BR_i(Q) = \{Q' | h_i(Q') \geqslant h_i(Q'') \text{ for all } Q'' \in UD_i(Q)\}$. Clearly, $BR_i(Q) \subseteq B_i$.

The following lemma will be useful in §4.

LEMMA 1. *Suppose that* K *is a straight line that divides* $\mathbb{R}^2$ *into two parts,* K^- *and* K^+. *Furthermore, suppose that* $h(B_i) \subseteq K^-$ *and*[2] $\text{dist}(h(B_i), K) = d > 0$. *Then there exists* $\epsilon > 0$ *s.t.* $h(Q) \in K^+$ *and* $Q' \in BR_i(Q)$ *imply*

$$h_i(Q') \geqslant h_i(Q) + \epsilon, \qquad i = 1, 2.$$

PROOF. Assume to the contrary that there exists a sequence $Q^n \in \Delta$ which satisfies (i) $h(Q^n) \in K^+$ and (ii) for every $\bar{Q}^n \in BR_i(Q^n)$ the following holds: $h_i(\bar{Q}^n) < h_i(Q^n) + \epsilon_n$, where $\epsilon_n \to 0$. We can assume that $Q^n \to Q$. Since h is continuous, $h(\bar{Q}) \leqslant h(Q)$ for all $\bar{Q} \in UD_i(Q)$. Thus, $Q \in B_i$.

On the other hand, $\text{dist}(h(Q^n), h(B_i)) \geqslant d$ and therefore $\text{dist}(h(Q), h(B_i)) \geqslant d$, a contradiction. //

A similar statement holds for $h(\tilde{B}_i)$ as well. The following lemma will be used in §5.

LEMMA 2. *Let* K *be a straight line satisfying* $\text{dist}(h(B_1 \cap B_2), K) = d > 0$. *Then there exists an* $\epsilon > 0$ *s.t. for all correlated actions* Q, *if* K *separates between* $h(Q)$ *and* $h(B_1 \cap B_2)$ *then there is an* i *satisfying*

$$h_i(\bar{Q}) \geqslant h_i(Q) + \epsilon \quad \text{for all } \bar{Q} \in BR_i(Q).$$

PROOF. Assume to the contrary that there are sequences of correlated actions $\{Q^n\}$, $\{\bar{Q}_1^n\}$ and $\{\bar{Q}_2^n\}$ satisfying: (i) $h_i(\bar{Q}^n) < h_i(Q^n) + \epsilon_n$ for $i = 1, 2$, where $\epsilon_n \to 0$, and (ii) $\bar{Q}_i^n \in BR_i(Q^n)$, $i = 1, 2$.

W.l.o.g.[3] we can assume that $Q^n \to Q$. Thus, $Q \in B_1 \cap B_2$. On the other hand, since h is continuous, $\text{dist}(h(Q), h(B_1 \cap B_2)) \geqslant d$, a contradiction. //

[1] With respect to.
[2] $\text{dist}(\cdot, \cdot)$ is the distance induced by the Euclidean metric.
[3] Without loss of generality.

3.4. *The lower equilibrium.* The lower Nash equilibrium (not the correlated one) is defined like the correlated equilibrium with the further qualification that the probability measure P on $A \times B$ is the product of its marginal distributions. In other words, the distribution according to which a player picks his pure strategy, before playing the game, is fixed across messages his opponent gets. In [L4] the set of all the lower equilibrium payoffs, LEP, is characterized. This characterization is done by the sets C_i, D_i. For the sake of completeness we present it here.

A player has *trivial information* if all his opponent's actions are indistinguishable from one another. The main result of [L4] is:

$$\text{LEP} = \operatorname{conv} h(C_1) \cap \operatorname{conv} h(C_2) \cap IR$$

if both players do not have trivial information, and

$$\text{LEP} = \operatorname{conv} h(D_1) \cap \operatorname{conv} h(D_2) \cap IR$$

otherwise, where IR is the set of the individually rational payoffs.

3.5. *Characterization of LCEP.* In the case where the information of both players is not trivial the characterization will be done by using B_i, and in the trivial case by using $\tilde{B}_i$.

THEOREM 1. *In two-player games*[4] *the following hold*:

(i) *if both players have nontrivial information, then* LCEP = LCEP* = $\operatorname{conv} h(C_1) \cap \operatorname{conv} h(C_2) \cap IR = h(B_1) \cap h(B_2) \cap IR$; *and*

(ii) *if at least one of the players has trivial information, then* LCEP = LCEP* = $\operatorname{conv} h(D_1) \cap \operatorname{conv} h(D_2) \cap IR = h(\tilde{B}_1) \cap h(\tilde{B}_2) \cap IR$.

In words, the lower correlated equilibrium payoffs set and the extensive form correlated equilibrium payoffs set coincide. Moreover, in the nontrivial case, they are equal to the set of payoffs associated with a correlated action in B_1 and (possibly different) correlated action in B_2.

One of the implications of Theorem 1 is:

COROLLARY 1. LCEP = LEP. //

In other words, the introduction of a mediator to the game does not enlarge the set of lower equilibrium payoffs.

REMARK 3. In a case in which both players have trivial information, LCEP equals the set of correlated equilibrium payoffs of the one-shot game.

EXAMPLE 5. One can compute $h(B_i)$ of Example 1 and find

$$IR \cap h(B_i) = \operatorname{conv}\{(0,0),(7,2),(2,7),(6,6)\}, \qquad i = 1,2.$$

Thus, LCEP = $\operatorname{conv}\{(0,0),(7,2),(2,7),(6,6)\}$, which coincides with the feasible and individually rational payoffs.

EXAMPLE 6. In Example 4 the payoff $(6,6)$ is not in $h(B_i)$, $i = 1,2$, and thus $(6,6) \notin$ LCEP. Thus, not all the feasible payoffs are necessarily associated with lower correlated equilibrium.

3.6. *The characterization of UCEP.* The upper equilibrium is more restrictive. This fact is reflected in the characterization of the corresponding payoffs set. While a typical payoff in LCEP is associated with *two* correlated actions (one in B_1 and one

[4]Here and in the sequel, "games" refers to repeated games with nonobservable actions.

in B_2), a payoff in UCEP is associated with *one* correlated action which is in both B_1 and B_2.

THEOREM 2. *In two-player games,* (i) *if both players have nontrivial information, then*

$$\text{UCEP} = \text{UCEP}^* = \text{UNIC} = \text{UNIC}^* = \text{CEP}_L = \text{CEP}_L^* = h(B_1 \cap B_2) \cap IR,$$

for all Banach limit L; and (ii) *if at least one player has trivial information, then*

$$\text{UCEP} = \text{UCEP}^* = \text{UNIC} = \text{UNIC}^* = \text{CEP}_L = \text{CEP}_L^* = h\left(\tilde{B}_1 \cap \tilde{B}_2\right) \cap IR,$$

for all Banach limit L.

EXAMPLE 7. In Example 4, since $IR = \mathbb{R}^2_+$, one obtains

$$h(B_1 \cap B_2) = \text{UCEP} = \text{conv}[\{\alpha(7,2) + \beta(2,7) + \gamma(6,6) | \\ \alpha + \beta + \gamma = 1; \alpha, \beta, \gamma \geqslant 0; \gamma \leqslant \alpha; \gamma \leqslant \beta\} \cup \{(0,8),(8,0),(0,0)\}].$$

4. The proof of Theorem 1. From here on it is assumed that h is bounded between 0 and 1. We will use the result quoted in §3.4 above.

The first step in the proof is to show that

$$h(B_1) \cap h(B_2) \cap IR \subseteq \text{LCEP}. \tag{4.1}$$

It is clear that any lower equilibrium payoff is also a correlated equilibrium. Thus, LEP $\subseteq$ LCEP. By Proposition 1 and by §3.4:

$$\text{conv}\, h(C_1) \cap \text{conv}\, h(C_2) = h(B_1) \cap h(B_2).$$

Therefore (4.1) is established. It remains to show the converse inclusion. We will show that LCEP* $\subseteq h(B_1) \cap h(B_2) \cap IR$.

Assume to the contrary that $U = ((\times_{t=1}^{\infty} A_t) \times (\times_{t=1}^{\infty} B_t), P, f, g)$ is an extensive form correlated equilibrium and that the payoff associated with it, (w_1, w_2), lies outside of $h(B_1) \cap h(B_2)$. W.l.o.g. we may assume that $(w_1, w_2) \notin h(B_2)$. We will define a function $\bar{g}$ (a deviation, according to which PII chooses his pure strategy), which results in a higher payoff for PII. Precisely,

$$\liminf_T E_{f,\bar{g},P}\left((1/T)\sum_{t=1}^{T} x_2^t\right) > w_2.$$

Thereby, we will prove that U is not an equilibrium. The deviation $\bar{g}$ is described as follows. Instead of playing the prescribed action (defined by g) PII plays the best undetectable deviation. However, the play of PII should be continued in a consistent way, so as not to affect the distribution of PI's signals. Lemma 4 ensures that there exists such a continuation. In order to verify that, indeed, $\bar{g}$ is a profitable deviation, we show that on a large set of states (Lemma 3), PII increases his expected payoff by at least $\epsilon > 0$ (Lemma 1).

Let K be a straight line that divides the plan into two disjoint parts: K^-, the open one, and K^+, the closed one. Moreover, assume that $(w_1, w_2) \in K^-$ and $h(B_2) \subseteq K^+$, and that $\operatorname{dist}((w_1, w_2), K) = \operatorname{dist}(h(B_2), K) = d > 0$. There exists such a separating line because $h(B_2)$ is closed and convex.

Denote by R^t the distribution over histories (consisting of messages and joint actions) of length $t - 1$. Recall that together with R^t the functions f and g induce a correlated action to be played at stage t. In other words, the histories and f, g induce a distribution over joint actions. This distribution, denoted by Q^t, indicates the probability for any joint action to be played at stage t had the players adhered to (f, g).

The following lemma states that the set of stages t on which Q^t is associated with a payoff in K^- (far away from $h(B_2)$) is relatively a large set.

LEMMA 3. *The set of stages $M = \{t | h(Q^t) \in K^-\}$ has a positive lower density, η, i.e.,*

$$\liminf_T |M \cap \{1, \dots, T\}|/T = \eta > 0.$$

PROOF. Notice that by the definition of (w_1, w_2) one obtains

$$(4.2) \qquad (w_1, w_2) = \lim_T \frac{1}{T} \sum_{t=1}^{T} h(Q^t).$$

Suppose to the contrary that $\eta = 0$. Thus, there exists a sequence $\{T_n\}$ satisfying

$$|M \cap \{1, \dots, T_n\}|/T_n = \eta_n \to 0.$$

For every n one gets

$$(4.3) \quad \frac{1}{T_n} \sum_{t=1}^{T_n} h(Q^t) = \frac{1}{T_n} \sum_{t \in M, t \leqslant T_n} h(Q^t) + \frac{1}{T_n} \sum_{t \notin M, t \leqslant T_n} h(Q^t)$$

$$\leqslant \frac{1}{T} \sum_{t \notin M, t \leqslant T_n} h(Q^t) + \eta_n(1, 1)$$

$$= (1 - \eta_n) \left[(1/T_n)(1 - \eta_n) \sum_{t \notin M, t \leqslant T_n} h(Q^t) \right] + \eta_n(1, 1).$$

The term in brackets is a convex combination of payoffs in K^+, which is in K^+ (recall that K^+ is convex and closed). Thus, the right side of (4.3) converges to a point in K^+. This contradicts (4.2), and the lemma follows. //

The following lemma mimics the functions g and $\bar{g}$ after a certain stage, say, $t - 1$. I is the set of all PI's histories of length $t - 1$ and $J, \bar{J}$ stand for the set of all PII's histories of the same length. The sample space $(I \times J, \mu)$, where μ is the probability, defined on $I \times J$, is the distribution over the joint histories induced by the original extensive form correlated equilibrium, U. The sample space $(I \times \bar{J}, \bar{\mu})$ is the one induced by the deviation $\bar{g}$. By playing undetectable deviations (in particular, more informative action) in the previous stages, PII did not lose the ability to distinguish between actions of PI. This fact is represented in the lemma by a map, ψ, between possible histories that correspond to g and actual histories that correspond to $\bar{g}$. The conclusion of the lemma is that PII can pretend as if he abides by the prescribed action e (to be played at stage t), while actually he plays $\bar{e}$. $\bar{e}$ will be utilized later in the construction of $\bar{g}$.

LEMMA 4. *Let there be given two finite sample spaces* $(I \times J, \mu)$ *and* $(I \times \bar{J}, \bar{\mu})$, *and let* ψ *be a one-to-one function* ψ: $I \times J \to I \times \bar{J}$, *which satisfies the following*:

(i) $\psi_1(\alpha, \beta) = \alpha$, *where* $\psi = (\psi_1, \psi_2)$,

(ii) ψ *is measure preserving, i.e.,* $\mu(\psi^{-1}(\alpha, \bar{\beta})) = \bar{\mu}(\alpha, \bar{\beta})$,

(iii) *if* $(\alpha', \bar{\beta}), (\alpha, \bar{\beta}) \in$ *support*$(\bar{\mu})$, $\psi(\alpha', \beta) = (\alpha', \bar{\beta})$ *and* $\psi(\alpha, \gamma) = (\alpha, \bar{\beta})$, *then* $\beta = \gamma$.

Then, for any function e: $B \to \Delta^n$ (*the unit simplex in* $\mathbb{R}^n$), *there is a function* $\bar{e}$: $\bar{B} \to \Delta^n$ *s.t., for all* $\alpha \in I$, $E_\mu(e|\alpha) = E_{\bar{\mu}}(\bar{e}|\alpha)$.

PROOF. Denote by μ_1, μ_2 (resp., $\bar{\mu}_1, \bar{\mu}_2$) the marginal distributions of μ (resp., $\bar{\mu}$) over I, J (resp., $I, \bar{J}$). W.l.o.g. we can assume that $\bar{\mu}_2(\bar{\beta}) > 0$ for all $\bar{\beta} \in \bar{J}$, and $\mu_1(\alpha) > 0$ for all $\alpha \in I$. For every $\bar{\beta} \in \bar{B}$, in order to define $b(\bar{\beta})$, take any $(\alpha, \bar{\beta}) \in$ support$(\bar{\mu})$ and any (α, β) satisfying $\psi(\alpha, \beta) = (\alpha, \bar{\beta})$ and set $b(\bar{\beta}) = \beta$. By (iii), $b(\bar{\beta})$ is well defined. $\psi(\cdot)$ and $b(\cdot)$ are one to one. Define $\bar{e}(\bar{\beta}) = e(b(\bar{\beta}))$. By (ii), we obtain

$$E_{\bar{\mu}}(\bar{e}|\alpha) = \sum_{\bar{\beta} \in \bar{J}} \bar{e}(\bar{\beta})\bar{\mu}(\alpha, \bar{\beta})/\bar{\mu}_1(\alpha) = \sum_{\bar{\beta} \in \bar{J}} e\big(b(\bar{\beta})\big)\mu\big(\psi^{-1}(\alpha, \bar{\beta})\big)/\bar{\mu}_1(\alpha)$$

(by (i) and (ii))

$$= \sum_{\bar{\beta} \in \bar{J}} e\big(b(\bar{\beta})\big)\mu\big(\psi^{-1}(\alpha, \bar{\beta})\big)/\mu_1(\alpha) = \sum_{\beta \in J} e(\beta)\mu(\alpha, \beta)/\mu_1(\alpha)$$

$$= E_\mu(e|\alpha). \quad \backslash\backslash$$

Now we are ready to define $\bar{g} = (\bar{g}^1, \bar{g}^2, \dots)$. It will be done by defining first a sequence of functions $\bar{g}_n = (\bar{g}_n^1, \bar{g}_n^2, \dots)$, and second by defining $\bar{g}$ as the diagonal, i.e., $\bar{g}^n = g_n^n$ for all n.

The function $\bar{g}_n$ is an improvement of $\bar{g}_{n-1}$ in the sense that $\bar{g}_n$ agrees with $\bar{g}_{n-1}$ on the first $n - 1$ stages, and it increases PII's payoff without being detectable. Furthermore, at the rest of the stages, $\bar{g}_n$ is a continuation of the play without giving a chance to PI to detect the previous deviation.

$(\bar{g}_n)_n$ is defined inductively. Set $\bar{g}_1 = g_1$, the original function. Suppose that $\bar{g}_j$ is defined for all $j < n$. Define $\bar{g}_n^t = \bar{g}_{t-1}^t$ for all $t < n$.

Recall that $\bar{g}_n^n$ (the nth function of the strategy $\bar{g}_n$) maps elements consisting of $v \in L_2^{n-1}$ and a string of messages, $\beta_1, \dots, \beta_n$, to actions in Σ_2. Denote for such v and $\beta_1, \dots, \beta_n$

$$k^n(v, \beta_1, \dots, \beta_n)$$

$$= \sum_{u \in L_1^{n-1}} \sum_{\alpha_1, \dots, \alpha_n} \mathrm{pr}(\alpha_1, \dots, \alpha_n, \beta_1, \dots, \beta_n, u, v) \cdot f^n(u, \alpha_1, \dots, \alpha_n),$$

where the probability pr$(\cdot)$ is the probability induced by f, $\bar{g}_{n-1}$ and $\{P_t\}_{t \leqslant n}$ and α_t is the message PI got at stage t. Thus, $k^n(v, \beta_1, \dots, \beta_n)$ is the expected mixed action PI is supposed to play, given that the history of PII is $(v, \beta_1, \dots, \beta_n)$. $\bar{g}_n^n(v, \beta_1, \dots, \beta_n)$ will be defined as a best response versus $k^n(v, \beta_1, \dots, \beta_n)$, among all the actions that are indistinguishable from, and more informative than $\bar{g}_{n-1}^n(v, \beta_1, \dots, \beta_n)$.

We will define $\bar{g}_n^t$ for $t > n$, using Lemma 4. Let $t = n + 1$. I is the set of all the $(u, \alpha_1, \dots, \alpha_{n+1})$ and J is the set of all the $(v, \beta_1, \dots, \beta_{n+1})$, where $u \in L_1^n$ and $v \in L_2^n$, μ is the probability distribution induced by f and $\bar{g}_{n-1}^1, \bar{g}_{n-1}^2, \dots, \bar{g}_{n-1}^n$, and where $\bar{\mu}$ is the one induced by f and $\bar{g}_n^1, \dots, \bar{g}_n^n$. ψ on $I \times J$ is defined as follows:

$$\psi_1((u, \alpha_1, \dots, \alpha_{n+1}), (v, \beta_1, \dots, \beta_{n+1})) = (u, \alpha_1, \dots, \alpha_{n+1}) \quad \text{and}$$

$$\psi_2((u, \alpha_1, \dots, \alpha_{n+1}), (v, \beta_1, \dots, \beta_{n+1})) = (\bar{v}, \beta_1, \dots, \beta_{n+1}),$$

where $\bar{v} = (\bar{v}_1, \dots, \bar{v}_n) \in L_2^n$ coincides with v on the first $n - 1$ coordinates and $\bar{v}_n = l_2(f_1^n(u, \alpha_1, \dots, \alpha_n), \bar{g}_n^n(v, \beta_1, \dots, \beta_n))$.

In order to use Lemma 4 we have to show that ψ satisfies the hypothesis of the lemma. Obviously, ψ is a one-to-one function. By the definition of ψ, (i) holds. Since $\bar{g}_n^n(v, \beta_1, \dots, \beta_n)$ is indistinguishable from $\bar{g}_{n-1}^n(v, \beta_1, \dots, \beta_n)$, (ii) is satisfied and because the former is more informative than the latter, (iii) is implied.

Apply Lemma 4 for $e = \bar{g}_{n-1}^{n+1}$, which is defined on histories of length n, to obtain the function $\bar{e}$. $\bar{e}$ satisfies $E_\mu(e|\alpha) = E_{\bar{\mu}}(\bar{e}|\alpha)$ for all $\alpha \in A$. Define $\bar{g}_n^{n+1} = \bar{e}$.

In words, PII adjusts his behavior. Instead of playing according to $\bar{g}_{n-1}^{n+1}$ he plays according to $\bar{g}_n^{n+1}$. However, PI cannot differentiate between the two since both induce the same mixed action, no matter what the history of PI is.

So far we defined $\bar{g}_n$ up to stage $n + 1$. In order to continue defining $\bar{g}_n^{n+2}, \bar{g}_n^{n+3}, \dots,$ we should repeatedly use Lemma 4. $\bar{g}_n^{n+1}$, just defined, induces, together with f, a distribution over the joint histories. By playing according to $\bar{g}_n^{n+1}$, PII does not lose information, in the sense that a function ψ, applied to histories of length $n + 1$, can be found so as to satisfy hypotheses (i)–(iii) of Lemma 4. Thus, $\bar{g}_n^{n+2}$ can be defined without affecting the distribution PI is expecting (from $\bar{g}_{n-1}^{n+2}$). In the same way, all the strategy $\bar{g}_n$ is defined, thereby ensuring that

$$(4.4) \qquad E_{f, \bar{g}_n}(x^t) = E_{f, \bar{g}_{n-1}}(x^t) \quad \text{for all } t > n.$$

Namely, the expected payoffs after stage n are not changed by $\bar{g}_n$. Moreover, letting Q^n (resp., $\bar{Q}^n$) denote the probability distribution of the set of joint actions (to be played at stage n) induced by f and g (resp., $\bar{g}_n$), one obtains

$$(4.5) \qquad \bar{Q}^n \in BR_2(Q^n).$$

This is because $\bar{g}_n^n$ was defined as a best response among all the actions indistinguishable from and more informative than the prescribed one. In other words,

$$(4.6) \qquad E_{f, \bar{g}_n}(x_2^n) \geqslant E_{f, \bar{g}_{n-1}}(x_2^n)$$

and $E_{f, \bar{g}_n}(x^n) \in h(B_2)$.

Define $\bar{g}^n = \bar{g}_n^n$. (4.6) and (4.4) imply that

$$(4.7) \qquad E_{f, \bar{g}}(x_2^t) = E_{f, \bar{g}_t}(x_2^t) \geqslant E_{f, \bar{g}_{t-1}}(x_2^t)$$

$$= E_{f, \bar{g}_{t-2}}(x_2^t) = \cdots = E_{f, g}(x_2^t) \quad \text{for all } t.$$

(4.5) and (4.7) and Lemma 1 imply that there is an $\epsilon > 0$ satisfying

$$(4.8) \qquad E_{f, \bar{g}}(x_2^n) \geqslant E_{f, g}(x_2^n) + \epsilon \quad \text{for all } n \in M.$$

From Lemma 3 and (4.8) it follows that

$$\liminf_T E_{f,\bar{g}}\left(\frac{1}{T}\sum_{t=1}^{T} x_2^t\right) \geqslant \lim_T E_{f,g}\left(\frac{1}{T}\sum_{t=1}^{T} x_2^t\right) + \epsilon\eta .$$

It shows that PII has a profitable deviation, $\bar{g}$, which establishes the fact that U is not an extensive form correlated equilibrium. Recall that it derives from the assumption that the payoff associated with U is not in $h(B_2)$. Thus, we have shown that LCEP* $\subseteq h(B_1) \cap h(B_2) \cap IR$ in the nontrivial case and the proof of Theorem 1 is concluded. //

5. **Proof of Theorem 2.** We consider here only the nontrivial case; the other case is left to the reader.

The proof will be divided into *three* steps. In the first one, it is shown that UCEP* $\subseteq h(B_1 \cap B_2) \cap IR$. Since UNIC* $\subseteq$ UCEP*, it will provide also a proof to UNIC* $\subseteq h(B_1 \cap B_2) \cap IR$. In the second step it will be shown that $\text{CEP}_L^* \subseteq h(B_1 \cap B_2) \cap IR$ for every Banach limit L.

The first two steps are proven by the same method. It is assumed, to the contrary, that there is an equilibrium (the one in question) payoff not in $h(B_1 \cap B_2) \cap IR$. Since any equilibrium payoff should be in IR it can be assumed that the payoff is not in $h(B_1 \cap B_2)$. Based on this assumption, a profitable deviation is constructed in the way it has been built in the previous section. The existence of profitable deviation contradicts the fact that the payoff is associated with an equilibrium.

The third step is devoted to the converse direction. It is shown that $h(B_1 \cap B_2) \cap IR \subseteq$ UNIC. Since UNIC is the smallest set of correlated equilibrium payoffs mentioned in this paper, this step concludes the proof of the theorem.

Step 1. UCEP* $\subseteq h(B_1 \cap B_2) \cap IR$. It is obvious that UCEP* $\subseteq IR$. Assume that $(w_1, w_2) \notin IR \setminus h(B_1 \cap B_2)$ and that $U = ((\times A_t) \times (\times B_t), \mathscr{A} \times \mathscr{B}, P, f, g)$ is an extensive form correlated equilibrium associated with (w_1, w_2).

Let K be a separating straight line between (w_1, w_2) and $h(B_1 \cap B_2)$ so that $\text{dist}((w_1, w_2), K) = \text{dist}(h(B_1 \cap B_2), K) = d > 0$. Denote the half-plane that contains (w_1, w_2) by K^-.

Denote by Q^n the distribution on Σ induced by U. Set $M = \{t | h(Q^t) \in K^-\}$. In words, t is the set of the stages on which the expected payoff is far away from $h(B_1 \cap B_2)$. On this set of stages the deviating player will benefit at least by $\epsilon > 0$.

By Lemma 2 there is an $\epsilon > 0$ s.t. if Q satisfies $h(Q) \in K^-$, then there is i s.t. $h_i(\bar{Q}) \geqslant h_i(Q) + \epsilon$ for all $\bar{Q} \in BR_i(Q)$. Thus, M can be written as a union of M_i, $i = 1, 2$, where $M_i = \{t \in M | h_i(\bar{Q}) \geqslant h_i(Q) + \epsilon$ for all $\bar{Q} \in BR_i(Q)\}$.

LEMMA 5. *There is i s.t. M_i has a positive upper density, i.e.,*

$$\limsup_T |M \cap \{1, \dots, T\}| / T = \eta > 0.$$

PROOF. By Lemma 3,

$$\begin{aligned} 0 &< \liminf_T |M \cap \{1, \dots, T\}| / T \\ &= \liminf_T |(M_1 \cup M_2) \cap \{1, \dots, T\}| / T \\ &\leqslant \limsup_T |M_1 \cap \{1, \dots, T\}| / T + \limsup_T |M_2 \cup \{1, \dots, T\}| / T. \end{aligned}$$

Therefore, one of the terms on the right side should be positive. //

W.l.o.g., i of Lemma 5 equals 2. Define the deviation of PII, $\bar{g}$, as it was defined in the previous section. The deviation $\bar{g}$ results in (similar to (4.8)):

$$E_{f,\bar{g}}(x_2^n) \geqslant E_{f,g}(x_2^n) + \epsilon \quad \text{for all } n \in M_2,$$

where ϵ is the one obtained by Lemma 2 and employed in the definition of M_2. Moreover, as in (4.7),

$$E_{f,\bar{g}}(x_2^t) \geqslant E_{f,g}(x_2^t) \quad \text{for all } t.$$

Thus,

$$\limsup_T (1/T) \sum_{t=1}^{T} E_{f,\bar{g}}(x_2^t)$$

$$= \limsup_T (1/T) \left[\sum_{t \leqslant T, t \notin M_2} E_{f,\bar{g}}(x_2^t) + \sum_{t \leqslant T, t \in M_2} E_{f,\bar{g}}(x_2^t) \right]$$

$$\geqslant \limsup(1/T) \left[\sum_{t \leqslant T, t \notin M_2} E_{f,g}(x_2^t) + \sum_{t \leqslant T, t \in M_2} E_{f,g}(x_2^t) + \epsilon \right]$$

$$= \limsup(1/T) \left[\sum_{t=1}^{T} E_{f,g}(x_2^t) + \sum_{t \leqslant T, t \in M_2} \epsilon \right]$$

$$= \lim(1/T) \sum_{t=1}^{T} E_{f,g}(x_2^t) + \limsup_T (1/T) \sum_{t \leqslant T, t \in M_2} \epsilon \geqslant w_2 + \eta\epsilon.$$

Hence, $\bar{g}$ is a profitable deviation which contradicts the assumption saying that U is an equilibrium. We therefore conclude that

$$\text{UCEP}^* \subseteq h(B_1 \cap B_2) \cap IR.$$

Step 2. $\text{CEP}_L^* \subseteq h(B_1 \cap B_2) \cap IR$. The key to the proof of the previous step was to show that there is a significantly large set of stages (M_2 had a positive upper density) on which PII gains by at least ϵ, while on the other stages he guarantees at least what he would get had he adhered to the prescribed strategy.

Fix a Banach limit L. The objective of the proof is to show that the deviator may profit on a big set (w.r.t. L) at least by ϵ without losing on the other stages. The next lemma deals with the size of the set on which the deviator gains.

LEMMA 6. *Suppose that Q^n is a sequence of correlated actions satisfying*

$$(w_1, w_2) = L(h(Q^n)) \notin h(B_1 \cap B_2).$$

Suppose, furthermore, that $Q_1^n \in BR_i(Q^n)$ for all n and i. Then there is i s.t.

$$L\left((1/N) \sum_{n=1}^{N} h_i(Q_i^n) \right) > w_i. \tag{5.1}$$

In words, there is a player who gets, in the long run, more than his prescribed payoff.

PROOF. Recall that $Q' \in UD_i(Q)$ is characterized by an admissible function $\gamma\colon \Sigma_i \to \Sigma_i$ (i.e., $\gamma(a)$ is indistinguishable from and more informative than a for all $a \in \Sigma_i$). Thus, one can divide $\mathbb{N}$ into a finite number of sets M_i^γ (M_i^γ is the set of all stages n for which Q_i^n is characterized by the function γ).

For $A \subseteq \mathbb{N}$ denote by F_A the infinite sequence whose nth coordinate equals $|A \cap \{1, \ldots, n\}|/n$. Thus, for all i, $\Sigma_\gamma L(F_{M_i^\gamma}) = 1$, where the summation is taken over all admissible functions γ. Fix γ. Denote $P_\gamma^n = Q^n$ and $\bar{P}_\gamma^n = Q_i^n$ if $n \in M_i^\gamma$ and 0 (the matrix 0) otherwise. Since all Q_i^n ($n \in M$) retain the same system of linear weak inequalities and since Banach limit preserves such inequalities one obtains:

(1) $L((1/N)\Sigma_{n=1}^N \bar{P}_\gamma^n)$ is a matrix whose entries sum up to $L(F_{M_i^\gamma})$.

(2) If $L(F_{M_i^\gamma}) > 0$, then the matrix

$$L\left((1/N) \sum_{n=1}^{N} \bar{P}_\gamma^n\right) / L\left(F_{M_i^\gamma}\right) \in BR_i\left(L\left((1/N) \sum_{n=1}^{N} P_\gamma^n\right) / L\left(F_{M_i^\gamma}\right)\right).$$

Since $L((1/N)\Sigma_{n=1}^N h(Q^n)) \notin h(B_1 \cap B_2)$ and since there are a finite number of admissible functions, there exist i and an admissible function γ which satisfy $L(F_{M_i^\gamma}) > 0$ and

$$h_i\left(L\left((1/N) \sum_{n=1}^{N} \bar{P}_\gamma^n\right) / L\left(F_{M_i^\gamma}\right)\right) > h_i\left(L\left((1/N) \sum_{n=1}^{N} P_\gamma^n\right) / L\left(F_{M_i^\gamma}\right)\right).$$

Since for all other γ's (A.1) is satisfied with weak inequality, providing that $L(F_{M_i^\gamma}) > 0$, the proposition follows. //

Suppose that the extensive form correlated equilibrium $U = ((\times A_t) \times (\times B_t), \mathscr{A} \cap \mathscr{B}, P, f, g)$ is given and, moreover, the payoff associated with U (w.r.t. L) is $(w_1, w_2) \notin h(B_1 \cap B_2) \cap IR$. Since (w_1, w_2) should be in IR we may assume that $(w_1, w_2) \notin h(B_1 \cap B_2)$.

Denote by Q^n the correlated action at stage n induced by f and g. Thus, $L((1/N)\Sigma_{n=1}^N h(Q^n)) = (w_1, w_2)$. Let Q_i^n be an element in $BR_i(Q^n)$. By Lemma 6 there is a player i for which (5.1) holds. W.l.o.g., $i = 2$.

Define the deviation of PII, $\bar{g}$, as it was defined in §4. Denote by $\bar{Q}^n$ the correlated action at time n induced by $f, \bar{g}$. By (4.5) and (5.1), replacing Q_i^n by $\bar{Q}^n$, one obtains

$$L\left((1/N) \sum_{n=1}^{N} h_2(\bar{Q}^n)\right) > w_2. \tag{5.2}$$

The parenthetical term of (5.2) is PII's payoff up to stage N associated with $f, \bar{g}$. Thus, $\bar{g}$ is a profitable deviation of PII. This concludes the proof of Step 2.

Step 3. $h(B_1 \cap B_2) \cap IR \subseteq$ UNIC. The objective is to construct for any $Q \in B_1 \cap B_2$ satisfying $h(Q) \in IR$ a uniform correlated equilibrium associated with the payoff $h(Q)$. Before going into technical details, I would like to give an informal, textual description of the correlated strategy, defined later. Suppose that before starting the game a mediator draws a joint action in Σ according to Q, infinitely many times. Each draw is independent of the other. Let the outcomes at the tth draw be denoted by (a_t, b_t). The message for player 1 is $(a_1, a_2, \ldots)$ and the message for player 2 is $(b_1, b_2, \ldots)$. Suppose, furthermore, that PI plays a_t at stage t and PII plays b_t. Does this generate an equilibrium? Certainly not. The player might have incentives to deviate and disobey the recommendation of the mediator. However, we assumed $Q \in B_1 \cap B_2$. Thus, any profitable deviation is either distinguishable from or less informative than the recommended action.

With a slight modification the former correlated strategy would be able to cope with deviations from the assigned actions to distinguishable ones. Such deviations change the distribution of the opponent's signals. Hence, by making tests on the previous signals, a player can detect, with high precision, his opponent's deviations. But for this test, sometimes a player should play some actions outside the support of Q. In other words, in order to detect deviation to an action, distinguishable from the recommended one, a player may need to play actions that are assigned zero probability by Q. For this purpose we modify the way the mediator chooses the messages $(a_1, a_2, \dots), (b_1, b_2, \dots)$.

Let Q^t be a perfectly mixed correlated action, assigning each joint action a positive probability. Furthermore, Q^t tends to Q. The pair (a_t, b_t) is drawn according to Q^t independently of the previous draws. PI plays a_t and PII plays b_t at stage t. The players are supposed to check the signals they got and compare it with the expected signals. If a player finds a discrepancy between the two, he should start punishing his opponent for a long period of time and then resume the game from the beginning. Does this form an equilibrium? Still not. The reason is that a player can deviate to an action indistinguishable from, yet less informative than the recommended one.

Say, for instance, that a and a' are two actions of PI, indistinguishable one from the other. Moreover, a is strictly less informative than a'. In other words, PI, by playing a', can distinguish between the actions b and b' of PII, while by playing a he cannot. Suppose that in a certain stage PII knows when PI is supposed to play a', while PI does not know that PII knows it. In this case, PI may deviate to a (he does not suspect that PII will detect it) and thereby lose his ability to distinguish between b and b'. However, if PII plays with probability $\frac{1}{2}$ either b or b', and PI reports to PII (by a method that will be described in the sequel) whether or not he observed the signal corresponding to b or b', PI is going to be mistaken with probability $\frac{1}{2}$. This is so because by playing a, PI observes the same signal, no matter if PII plays b or b'. PII knows that PI was supposed to play a' and to be able to report it correctly. Thus, PII, with probability $\frac{1}{2}$, infers that PI has deviated to a less informative action, and he can start punishing PI.

The previous explanation suggests that in a correlated strategy we are about to define, there will be stages in which PII (resp., PI) knows what PI (resp., PII) is supposed to play while PI (resp., PII) does not know that PII (resp., PI) knows it. Moreover, there should be a way to communicate, so that one player will be able to elicit information from the other.

The mediator, instead of drawing (a_t, b_t) from Σ, draws (a_t, b_t, i_t), where $i_t \in \{0, 1, 2\}$. In a case where $i_t = 0$, the mediator sends PI and PII the messages a_t and b_t, respectively. However, if $i_t = 1$, he informs PI of (a_t, b_t) and PII of b_t. If $i_t = 2$, he informs PII of (a_t, b_t) and PI of a_t. In other words, if $i_t \neq 0$, the player i_t knows what his opponent is recommended to act while the latter does not know that player i_t knows it. Obviously, the majority of the weight (w.r.t. which (a_t, b_t, i_t) is drawn) should be put on $(a_t, b_t, 0)$. Thus, the additional possible selections of the mediator do not distort the payoffs by much.

How do players communicate? We assume that each player has nontrivial information. Therefore, each player can distinguish between two of his opponent's actions at least. Different combinations of these two actions can encode different reports. The correlated strategy hereby defined will specify who and when should report on what. This should be designed carefully, because the mediator should not reveal to PI on which actions he will have to report. This information will be disclosed only to PII. When the time comes, PII will announce (by a special combination of actions) what stage PI should report on. The uncertainty about the stage on which PI will have to report will prevent him from deviating at *all* stages.

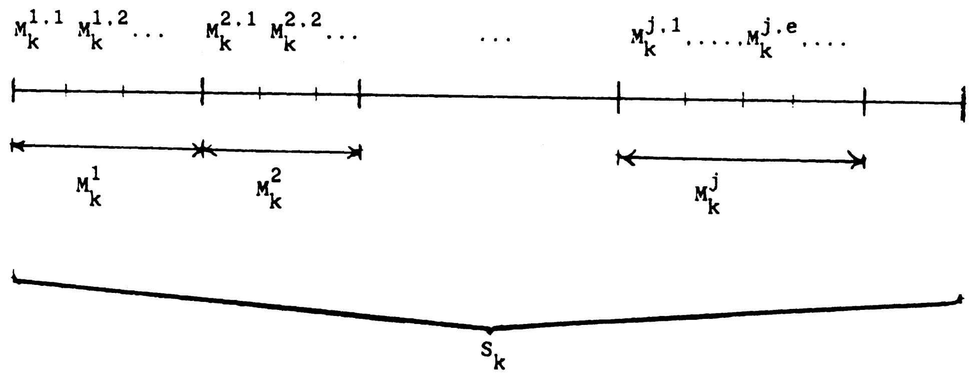

FIGURE 2

Before coming to the rigorous formal definition of the strategy, recall Remark 1.

Divide the set of stages $\mathbb{N}$, into consecutive superblocks, $S_1, S_2, \ldots$. Each block S_k will be divided into consecutive blocks, $M_k^1, M_k^2, \ldots$, and each M_k^j will be divided into subblocks $M_k^{j,1}, M_k^{j,2}, \ldots$.

Let $|M_k^{j,e}| = k + 4[\lg k + 1]$, $|M_k^j| = k^5|M_k^{j,e}|$ and $|S_k| = k|M_k^j|$. In words, the size of the subblock $M_k^{j,e}$ is $k + 4[\lg k + 1]$. There are k^5 subblocks of the same size in any block, and there are k blocks of the same size in any superblock.

Any subblock is divided into three phases. The first one, called the master phase and denoted $M_k^{j,e}(1)$, the players are supposed to play according to the message of the mediator. In the second phase, denoted $M_k^{j,e}(2)$, the players will choose randomly a stage on which they will have to report to their opponents on the third phase, $M_k^{j,e}(3)$. The last two phases will last $2[\lg k + 1]$ stages each. Since the length of the first phase is k and since $4[\lg k + 1]/k$ tends to zero, the master phase is dominant in the sense of determining the average payoff.

Denote $\overline{\Sigma} = \Sigma \cup (\Sigma \cup \{1\}) \cup (\Sigma \times \{2\})$. Let Q_k be a probability measure assigning each point $(a, b) \in \overline{\Sigma}$ the probability $Q(a, b)(1 - 2|\Sigma|/k)$ and other points probability $1/k$ each. Let $Z_1, Z_2, \ldots$ be a sequence of i.i.d. random variables attaining values in $\overline{\Sigma}$ and Z_t is distributed like Q_k whenever $t \in S_k$. The message $(\alpha_1, \alpha_2, \ldots)$ the mediator sends to PI is defined by:

$$\alpha_t = \begin{cases} a & \text{if } Z_t = (a,b) \quad \text{or} \quad (a,b,2), \\ (a,b) & \text{if } Z_t = (a,b,1). \end{cases}$$

In words, the mediator chooses randomly an element of $\overline{\Sigma}$ and informs PI of the action a if (a, b) or $(a, b, 2)$ was chosen and of the joint action (a, b) if $(a, b, 1)$ was chosen. Thus, when PI gets (a, b) as a message, he knows with precision what PII is supposed to play.

The message for PII is defined similarly.

Denote by (a_t, b_t) the joint action corresponding to (α_t, β_t). We can now describe the strategies in $t \in M_k^{j,e}(1)$, i.e., in the master phase of the subblock $M_k^{j,e}$. PI should play a_t and PII b_t, unless one of them ascribes a deviation to his opponent in one of the previous blocks of the same superblock.

When does a player ascribe a deviation to his opponent? Each player checks his opponent as to whether the latter deviated to actions that are indistinguishable from the prescribed one or to less informative actions. Either checking requires a different method. A player checks possible deviation to actions indistinguishable from the recommended ones by comparing the expected signal to the actual one at the stages when a player knows exactly what his opponent is supposed to play.

For instance, in a case where $\alpha_t = (a, b)$, PI knows that PII is supposed to play b. Therefore, he expects to observe the signal $l_1(a, b)$. If at that stage the signal was different, he knows that his opponent deviated.

Precisely, PI ascribes a deviation to PII in M_k^j if:

(5.3) there are $t \in M_k^j$ and $(a, b) \in \Sigma$ satisfying (1) $\alpha_t = (a, b)$ and (2) the signal PI observed at t, differs from $l_1(a, b)$.

In order to detect deviations to strictly less informative actions, a different procedure must be introduced. The parts $M_k^{j,e}(2)$ and $M_k^{j,e}(3)$ are devoted to it. The idea of the following procedure was introduced first by S. Sorin [S].

Each player has nontrivial information. Thus, there are $v_1, v_1', u_1 \in \Sigma_1$ and $v_2, v_2', u_2 \in \Sigma_2$ so that

$$l_1(u_1, v_2) \neq l_1(u_1, v_2') \quad \text{and} \quad l_2(v_1, u_2) \neq l_2(v_1', u_2).$$

By playing sequentially either v_2 or v_2', PII can send to PI a string of signals consisting of $l_1(u_1, v_2)$ and $l_1(u_1, v_2')$, providing that at the same time PI plays u_1. In the first half of the second phase $M_k^{j,e}(2)$, PI plays u_1 and PII plays with probability $\frac{1}{2}$ each v_2 and v_2'. They repeat this procedure $m_k^j = [\lg k + 1]$ times. Thus, after m_k^j stages PI observes a random string of length m_k^j consisting of $l_1(u_1, v_2)$ and $l_1(u_1, v_2')$. Such strings can encode stages in $M_k^{j,e}(1)$. Any encoding induces a distribution on the states in $M_k^{j,e}(1)$. If the encoding is appropriately designed, each stage in $M_k^{j,e}(1)$ is assigned a probability of at least $1/2k$ by the induced distribution (recall $|M_k^{j,e}(1)| = k$). To sum up, PII chooses and reports to PI a random stage, denoted by $t_k^{j,e}(\mathrm{I})$.

In the second half of the second phase, PI chooses a random stage from $M_k^{j,e}(1)$, denoted by $t_k^{j,e}(\mathrm{II})$, and reports on it to PII. This is done by playing m_k^j times either v_1 or v_1' with probability $\frac{1}{2}$ each, while PII plays u_2.

In the third phase, PI reports on the signal he received at stage $t_k^{j,e}(\mathrm{I})$ and PII reports on the signal of the stage $t_k^{j,e}(\mathrm{II})$.

How to report on a signal? There are finitely many possible signals. Each of them can be encoded by a finite string of two different symbols. In any stage of the first part of $M_k^{j,e}(3)$, PII plays u_2 while PI plays either v_1 or v_1', so as to transmit the string consisting of $l_2(v_1, u_2)$ and $l_2(v_1', u_2)$ which encodes the symbol he (PI) has received in the stage $t_k^{j,e}(\mathrm{I})$. In other words, at the end of the first portion of $M_k^{j,e}(3)$, PII can look at the string of the signals he received in $M_k^{j,e}(3)$ and infer about the identity of the signal on which PI reported.

Similarly, PII reports to PI on his signal. Namely, while PI plays u_1, PII plays sequentially either v_2 or v_2' according to the string encoding the signal PII received at the stage $t_k^{j,e}(\mathrm{II})$.

Denote the signal reported to PI (resp., PII) by PII (resp., PI) as $s_k^{j,e}(\mathrm{II})$ (resp., $s_k^{j,e}(\mathrm{I})$). To sum up, after the last stage of $M_k^{j,e}(3)$, PI knows that PII reported on $s_k^{j,e}(\mathrm{II})$ as the signal he received at stage $t_k^{j,e}(\mathrm{II})$, and PII knows that PI reported on $s_k^{j,e}(\mathrm{I})$ as the signal of the stage $t_k^{j,e}(\mathrm{I})$. Both players can check whether these reports are consistent with the strategies and with the actions they played in $t_k^{j,e}(\mathrm{I})$ and $t_k^{j,e}(\mathrm{II})$.

When no ambiguity arises, we denote $t_k^{j,e}(\mathrm{I})$ by $t(\mathrm{I})$ and $t_k^{j,e}(\mathrm{II})$ by $t(\mathrm{II})$.

The report of PII is *inconsistent* if $\alpha_{t(\mathrm{II})} = (a_{t(\mathrm{II})}, b_{t(\mathrm{II})})$, i.e., PI knows that PII is supposed to play $b_{t(\mathrm{II})}$ at stage $t(\mathrm{II})$ and

$$(5.4a) \qquad s_k^{j,e}(\mathrm{II}) \neq l_2(a_{t(\mathrm{II})}, b_{t(\mathrm{II})}).$$

In that case PI ascribes a deviation in $M_k^{j,e}$ to PII. The report of PI is inconsistent if $\beta_{t(\mathrm{I})} = (a_{t(\mathrm{I})}, b_{t(\mathrm{I})})$. Namely, PII knows what PI is supposed to play at $t(\mathrm{I})$ and

$$(5.4b) \qquad s_k^{j,e}(\mathrm{I}) \neq l_1(a_{t(\mathrm{I})}, b_{t(\mathrm{I})}).$$

In such a case, PII ascribes a deviation in $M_k^{j,e}$ to PI.

To recapitulate, PI attributes a deviation in $M_k^{j,e}$ to PII if either (5.3) or (5.4a) holds. In this case, PI should punish PII by playing the mixed action that "minmaxes" PII. Notice that in our model of deterministic signalling, a player ascribes a deviation to his opponent only when the latter had, indeed, deviated.

We will denote the strategies defined above by $U = (A \times B, P, \sigma, \tau)$, where A and B are the sets of the strings $(\alpha_1, \alpha_2, \dots)$ and $(\beta_1, \beta_2, \dots)$, respectively; P is the probability measure induced by the random selection of the mediator described above and σ, τ are the strategies of PI and PII, respectively.

In what follows we show that the payoff associated with U is $h(Q)$ and that U is a uniform equilibrium. The next proposition shows the first assertion.

PROPOSITION 3. *$\lim_T (1/T)\Sigma_{t=1}^T E_{\sigma,\tau,P}(x^t)$ exists and equals $h(Q)$.*

PROOF. Suppose that both players adhere to the strategies described above. In that case a deviation is detected with probability 0. Thus the expected payoff at time t is close to $h(Q)$ up to $2|\Sigma|/k$ (recall that Q_k assigns a total probability of $2|\Sigma|/k$ to points out of Σ). This concludes the proof. //

In the next proposition it is proven that U is a uniform correlated equilibrium. It cannot be assumed that an action of a player at one stage is independent of previous ones. Therefore, we need the following generalization of Chebyshev inequality, which is quoted from [L1].

LEMMA 7. *Let $R_1, \dots, R_n$ be a sequence of identically distributed Bernoulli random variables, with parameter p (i.e., $\mathrm{pr}(R_1 = 1) = p = 1 - \mathrm{pr}(R_1 = 0)$). Let $Y_1, \dots, Y_n$ be a sequence of Bernoulli random variables such that for each $i \leqslant m \leqslant n$, R_m is independent of $R_1, \dots, R_{m-1}, Y_1, \dots, Y_m$. Then, for every $\epsilon > 0$,*

$$\mathrm{pr}\left\{\left|\frac{R_1Y_1 + \cdots + R_nY_n}{n} - p\frac{Y_1 + \cdots + Y_n}{n}\right| \geqslant \epsilon\right\} \leqslant 1/n\epsilon^2. \quad //$$

We are now in a position to finish the proof of Theorem 2 by showing:

PROPOSITION 4. *U is a uniform correlated equilibrium.*

PROOF. Assume that PI plays σ.

We show that PII can gain at the block M_k^j by more than $h(Q) + 1/\sqrt{k}$ without being detected only with probability $O(k)$. Therefore, a profitable deviation will lead with probability $1 - O(k)$ to a punishment.

Denote by $\overline{M}_k^j = \bigcup M_k^{j,e}(1)$, where the union is taken over all the subblocks $M_k^{j,e}$ contained in M_k^j. $\overline{M}_k^j$ is the union of all the master phases. Recall that x_2^t is the payoff of PII at stage t.

Suppose that

$$(5.5) \qquad \left(1/|\overline{M}_k^j|\right) \sum_{t \in \overline{M}_k^j} x_2^t > h_2(Q) + 1/\sqrt{k}.$$

Fix $b \in \Sigma_2$. Denote by F_b the set of all stages in $\overline{M}_k^j$ in which PII was informed of the action b. Precisely, $F_b = \{t \in \overline{M}_k^j | \beta_t = b\}$. Denote by $Q_k^1(a)$, $Q_k^2(b)$ the probabilities assigned by Q_k to the actions $a \in \Sigma_1$ and $b \in \Sigma_2$, respectively, i.e., $Q_k^1(a) = Q_k(\{(a,b) | b \in \Sigma_2\})$, $Q_k^2 = Q_k(\{(a,b) | a \in \Sigma_1\})$.

By Chebyshev inequality with probability of at least

$$1 - |\Sigma_2| k^2 / |\overline{M}_k^j| = 1 - |\Sigma_2| / k^4 = 1 - c_1(k)$$

the following holds:

$$|F_b| > |\overline{M}_k^j| \left(Q_k^2(b) - 1/k \right) \quad \text{for all } b \in \Sigma_2. \tag{5.6}$$

Now we evaluate the probability that (5.4a) does not hold given that (5.3) does not hold and that (5.5) and (5.6) hold. In words, the probability that PI discovers a deviation of PII in $\overline{M}_k^j$ (by means of PI's reports), given that (1) PI did not detect a deviation by an inconsistent signal; (2) PII's average payoff in $\overline{M}_k^j$ is greater than $h_2(Q) + 1/\sqrt{k}$; and (3) the relative frequency of any action b (of PII) is close to its expectation.

If (5.5) holds, then there is an action b_0 satisfying

$$Q_k^2(b_0) > 0 \quad \text{and} \tag{5.7a}$$

$$h_2(p_{b_0}, b_0) + 1/2\sqrt{k} < \left(1/|F_{b_0}|\right) \sum_{t \in F_{b_0}} x_2^t, \tag{5.7b}$$

where p_{b_0} is PI's mixed action induced by Q_k given b_0 (i.e., $p_{b_0}(a) = Q_k(a, b_0)/Q_k^2(b_0)$ for all $a \in \Sigma_1$). However, according to (5.6), F_{b_0} contains a large number of states. Therefore, the relative frequency of the times PI plays a is close to $P_{b_0}(a)$, with high probability.

Formally, for any $t \in F_{b_0}$, denote $R_t(a) = 1$ if PI plays a at stage t, and 0 otherwise. Denote $Y_t(b) = 1$ if PII played b at stage t and 0 otherwise. By Lemma 7, with probability of at most $|\Sigma| k^2 / |F_{b_0}| = c_2(k)$, there holds

$$\left(1/|F_{b_0}|\right) \left| \sum_{t \in F_{b_0}} \left(Y_t(b) R_t(a) - p_b(a) Y_t(b) \right) \right| > 1/k \quad \text{for all } (a,b) \in \Sigma. \tag{5.8}$$

Notice that in view of (5.6), $c_2(k)$ tends to zero as k goes to infinity. From (5.7b) and (5.8) one obtains that, given (5.5) and (5.6), with probability of at least $1 - c_2(k)$, the following holds:

$$\begin{aligned} h_2(p_{b_0}, b_0) + 1/2\sqrt{k} &< \left(1/|F_{b_0}|\right) \sum_{t \in F_{b_0}} x_2^t \\ &= \left(1/|F_{b_0}|\right) \sum_{t \in F_{b_0}} \sum_{a,b} Y_t(b) R_t(a) h_2(a,b) \\ &\leqslant \sum_{a,b} \left(\left(1/|F_{b_0}|\right) \sum_{t \in F_{b_0}} p_{b_0}(a) Y_t(b) h_2(a,b) + 1/k \right) \\ &= \sum_b \left(1/|F_{b_0}|\right) \sum_{t \in F_{b_0}} Y_t(b) h_2(p_{b_0}, b) + |\Sigma_1|/k. \end{aligned} \tag{5.9}$$

Denote

$$q(b) = \left(1/|F_{b_0}|\right) \sum_{t \in F_{b_0}} Y_t(b) \quad \text{for all } b \in \Sigma_2. \tag{5.10}$$

Notice that q is a random variable attaining values in the set of PII's mixed actions. The right side of (5.9) equals

$$\sum_b q(b) h_2\left(p_{b_0}, b\right) + |\Sigma_1|/k \leqslant h_2\left(p_{b_0}, q\right) + |\Sigma|/k. \tag{5.11}$$

From (5.9) and (5.11) we get

$$h_2\left(p_{b_0}, b_0\right) + 1/2\sqrt{k} < h_2\left(p_{b_0}, q\right) + |\Sigma|/k. \tag{5.12}$$

To sum up, if (5.5) and (5.6) hold, then (5.12) holds with probability of at least $1 - c_2(k)$.

The next step is to show that if (5.3) does not hold, then with high probability those actions assigned a relatively high probability by q (recall (5.10)) are indistinguishable from b_0.

Fix $a \in \Sigma_1$. Define $R'_t = 1$ if $\alpha_t = (a, b)$, namely, PI knows that PII is supposed to play b, and 0 otherwise. Define $Y'_t = 1$ if PII plays at t an action b' that satisfies $l_1(a, b) = l_1(a, b')$. Again, by Lemma 7:

$$\operatorname{pr}\left\{ \left| \left(1/|F_{b_0}|\right) \sum_{t \in F_{b_0}} \left[(1 - Y'_t) R'_t - \operatorname{pr}\{R'_t = 1\}(1 - Y'_t)\right] \right| > 1/k^2 \right\} < k^4/|F_{b_0}|.$$

Assuming that (5.3) does not hold (namely, that $(1 - Y'_1)R'_t = 0$ for all $t \in F_{b_0}$), we obtain that with probability of at least $1 - k^4/|F_{b_0}|$ the following occurs: $(1/|F_{b_0}|)2|\Sigma|k\Sigma_{t \in F_{b_0}}(1 - Y'_t) \leqslant 1/k^2$. In other words:

$$\left(1/|F_{b_0}|\right) \sum_{t \in F_{b_0}} Y'_t > 1 - 2|\Sigma|/k. \tag{5.13}$$

Recall that (5.13) was obtained for a fixed a. Denote by $Y''_t = 1$ if PII plays an action indistinguishable from b_0, and 0 otherwise. By applying (5.13) to every a, one obtains

$$\left(1/|F_{b_0}|\right) \sum_{t \in F_{b_0}} Y''_t > 1 - 2|\Sigma|^2/k \tag{5.14}$$

with probability of at least $1 - |\Sigma_1|k^4/|F_{b_0}| = 1 - c_3(k)$. From (5.14) we deduce that with probability of at least $1 - c_3(k)$, the mixed action q assigns a probability of at least $1 - 2|\Sigma|^2/k$ to actions that are indistinguishable from b_0. In view of (5.12), recalling that $Q \in B_2$, q should assign a probability of at least $1/4\sqrt{k}$ to actions that are strictly less informative than b_0. Precisely, with probability of at least $1 - c_2(k) - c_3(b)$ the following holds:

$$\sum_b \left(1/|F_{b_0}|\right) \sum_{t \in F_{b_0}} Y_t(b) > 1/4\sqrt{k}, \tag{5.15}$$

where the summation is taken over all the actions b that are strictly less informative than b_0.

Denote $F^e_{b_0} = M^{j,e}_k \cap F_{b_0}$. $F^e_{b_0}$ is the set of all the states in $M^{j,e}_k$ in which b_0 was the message of PII. Given (5.6), inequality (5.15) implies that the fraction of those "good" subblocks $M^{j,e}_k$ for which

$$\sum_b \left(1/|F^e_{b_0}|\right) \sum_{t\in F^e_{b_0}} Y_t(b) > 1/k, \tag{5.16}$$

where the summation is like that of (5.15) and

$$|F^e_{b_0}|/|M^{j,e}_k| > 1/k \tag{5.17}$$

is greater than $1/k$. Thus, there are at least k^4 "good" subblocks satisfying (5.16) and (5.17). The probability of detecting a deviation in such a subblock is the probability to choose t satisfying $Y_t(b) = 1$ for some b which is strictly less informative than b_0, times the probability to play an action a about which PII would have known more had he played b_0 (and not b).

In other words, the probability of detecting a deviation in a "good" subblock is $|F^e_{b_0}|/2|M^{j,e}_k|k^2$ (the number 2 appears because PII has probability $\frac{1}{2}$ of guessing the signal correctly. One k stands for $1/k$ in (5.16). The other k in the denominator is for the probability of PI to get a message of the type (a, b_0), where a is an action by which PI can detect a deviation to a less informative action than b_0). By (5.17) it is greater than $1/k^3|\Sigma_1|$. Since there are at least k^4 "good" subblocks, the probability of evading PI's detection is at most $(1 - (1/k^3)|\Sigma_1|)^{k^4} = e^{-O(k)}$.

To recapitulate, given that (5.3) does not hold and that (5.5) and (5.6) do hold, with probability of at least $1 - c_2(k) - c_3(k) - e^{-O(k)}$ PI will discover a deviation by PII and the latter will be punished thereafter. However, (5.6) holds with probability of at least $1 - c_1(k)$. Thus, given that (5.3) does not hold and that (5.5) does hold, PI discovers a deviation with probability of at least $1 - c_1(k) - c_2(k) - c_3(k) - e^{-O(k)}$.

Notice that (5.5) deals with stages in $\overline{M}^j_k$. The remaining stages in M^j_k are negligible in the sense that $|M^j_k \setminus \overline{M}^j_k|/|M^j_k|$ is of the order of $(\lg k)/k$—in particular, smaller than $1/\sqrt{k}$. Thus, the event defined in (5.5) is included in the one defined by

$$\left(1/|M^j_k|\right) \sum_{t\in M^j_k} x^t_2 > h_2(Q) + 2/\sqrt{k} = h_2(Q) + d_1(k). \tag{5.5$'$}$$

To conclude the proof, notice that the length of any block M^j_k compared to the total length of its precedents goes to zero. Namely, $|M^j_k|/\min M^j_k \leqslant 1/2k = d_2(k)$. Thus, for any $t \geqslant \max S_{k-1}$, the correlated equilibrium U induces a $[c_1(k) + c_2(k) + c_3(k) + e^{-O(k^2)} + d_1(k) + d_2(k)]$-correlated equilibrium in the t-fold repeated game. Since the term in brackets goes to zero as k tends to infinity, it follows that U is a uniform equilibrium. //

6. Related topics and concluding remarks.

6.1. *Extensive form correlated equilibrium simplifies strategies.* In the presence of an active mediator, namely, a mediator who gives messages before any stage, the proof given in the last step of the previous section could have been simpler. Instead of choosing randomly a stage in the second phase of each subblock, the mediator can provide that information, i.e., the mediator chooses randomly two stages, t_1 and t_2. On the first, PI has to report, whereas on the second one PII has to report. Moreover, when t_1 is informed to PI, PII is provided with additional information: the action that was recommended to PI at stage t_1. Thus, PII knows exactly what should be the signal observed by PI, and therefore what signal PI should report on.

6.2. *Pointwise version of convergence.* It can be proven that the strategy U, constructed at the last stage of the previous section, yields that the partial averages converge *almost surely* to $h(Q)$. Moreover, it can be shown that the upper limit of the partial averages is less than the prescribed payoff with *probability one* under any deviation.

6.3. *Getting rid of the mediator.* The mediator provides private information to each one of the participating players. However, in repeated games with imperfect monitoring, even when players start with common knowledge and even if no exogenous correlation device exists, players can acquire private information during the course of the game. What kind of correlations can emerge from histories (which are private knowledge), and what is an efficient way to utilize the internal coordination (perhaps to achieve a higher level of cooperation) is still unsolved.

One simple case in which the phenomenon of internal correlation is demonstrated is given in [L3]. This is the case of standard-trivial information (recall Example 4).

6.4. *Correlated equilibrium in games with more than two players.* In a case of two players, a deviation can be attributed to one player—the opponent. However, if there are more than two players, who should be blamed for the deviation and who should be punished? In some cases a punishment of one player may benefit another. The latter has an incentive to pretend as if the former had deviated and to gain by the resulting punishment. The question is then to describe those outcomes that are supported by a "steady" behavioral pattern.

Another difficulty involved is to describe the most efficient way to punish a player. Typically, cooperation is needed to effectively punish a deviator. However, it may be the case that not all the players noticed the deviation, and the information about the alleged deviation should be spread among the players. How to do that in an "optimal" way without violating incentive compatibility is a subject for future study.

6.5. *Games with stochastic signaling.* In cases where the signaling is stochastic, even when a player knows what his opponent is supposed to play, he does not know exactly what was the resulting signal. It takes the right definition of indistinguishability and of being more informative to extend the deterministic results to stochastic ones. It seems that checking deviation would require statistical tests both on the signals received during the master phase and on the reports transmitted during the communication phases.

Acknowledgements. I would like to thank Robert J. Aumann, Francoise Forges, Abraham Neyman, Sylvain Sorin, and an anonymous referee of *Mathematics of Operations Research* for stimulating discussions and helpful comments.

References

[A1] Aumann, R. J. (1974). Subjectivity and Correlation in Randomized Strategies. *J. Math. Economics* **1** 67–95.

[A2] ______. (1981). Survey of Repeated Games. In *Essays in Game Theory and Mathematical Economics in Honor of Oskar Morgenstern*. Bibliographisches Institute, Mannheim/Wein/Zurich, 11–42.

[F1] Forges, F. (1986). An Approach to Communication Equilibria. *Econometrica* **54**, **6** 1375–1385.

[F2] ______. (1988). Communication Equilibria in Repeated Games with Incomplete Information, *Math. Oper. Res.* **13** 191–231.

[GZ] Gilboa, I. and Zemel, E. (1989). Nash and Correlated Equilibria: Some Complexity Considerations. *Games and Economic Behavior*, 80–93.

[HS] Hart, S. and Schmeidler, D. (1989). Existence of Correlated Equilibria. *Math. Oper. Res.* **14** 18–25.

[L1] Lehrer, E. (1990). Nash Equilibria of n-Player Repeated Games with Semi-Standard Information. *Internat. J. Game Theory* **19** 191–217.

[L2] ______. (1992). Two Player Repeated Games with Non observable Actions and Observable Payoffs. *Math. Oper. Res.* **17** 200–224.

[L3] ______. (1991). Internal Correlation in Repeated Games. *Internat. J. Game Theory* **19** 431–456.

[L4] ______. (1989). Lower Equilibrium Payoffs in Repeated Games with Non-Observable Actions. *Internat. J. Game Theory* **18** 57–289.

[M] Myerson, R. (1986). Multistage Games with Communication. *Econometrica* **54** 323–358.

[S] Sorin, S. (1990). Supergames. *Game Theory and Applications*, T. Ichiichi et al. (Eds.), Academic Press, New York, 46–63.

DEPARTMENT OF MANAGERIAL ECONOMICS AND DECISION SCIENCES, J. L. KELLOGG GRADUATE SCHOOL OF MANAGEMENT, AND DEPARTMENT OF MATHEMATICS, NORTHWESTERN UNIVERSITY, 2001 SHERIDAN ROAD, EVANSTON, ILLINOIS 60208